University of Massachusetts Medical School
eScholarship@UMMS

GSBS Dissertations and Theses | Graduate School of Biomedical Sciences

2003-01-10

HIV-1 Gene Expression: Transcriptional Regulation and RNA Interference Studies: a Dissertation

Ya-Lin Chiu
University of Massachusetts Medical School

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss

 Part of the Amino Acids, Peptides, and Proteins Commons, Genetic Phenomena Commons, Immune System Diseases Commons, Nucleic Acids, Nucleotides, and Nucleosides Commons, Virus Diseases Commons, and the Viruses Commons

Repository Citation
Chiu Y. (2003). HIV-1 Gene Expression: Transcriptional Regulation and RNA Interference Studies: a Dissertation. GSBS Dissertations and Theses. https://doi.org/10.13028/by53-2r17. Retrieved from https://escholarship.umassmed.edu/gsbs_diss/118

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

HIV-1 GENE EXPRESSION: TRANSCRIPTIONAL REGULATION AND RNA INTERFERENCE STUDIES

A Dissertation Presented

By

YA-LIN CHIU

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

JANUARY 10, 2003

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY

HIV-1 GENE EXPRESSION: TRANSCRIPTIONAL REGULATION AND RNA INTERFERENCE STUDIES

A Dissertation Presented

By

YA-LIN CHIU

Approved as to style and content by:

Michael R. Green, Chair of Committee

Craig L. Peterson, Member of Committee

Joel D. Richter, Member of Committee

Stewart Shuman, Member of Committee

Mario Stevenson, Member of Committee

Tariq M. Rana, Dissertation Mentor

Anthony Carruthers, Dean of the
Graduate School of Biomedical Sciences

Department of Biochemistry and Molecular Pharmacology

January 10, 2003

DEDICATION

This dissertation is dedicated to my parents,
Wen-Shin Lai and Wen-Ling Chiu

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere thanks to my mentor, Dr. Tariq M. Rana for his guidance and continuous support. I am especially grateful for his patience, allowing me to explore and develop my scientific abilities at my own pace. From the bottom of my heart, I appreciated all his time and effort.

I would like to thank my research advisory committee members, Drs. Michael R. Green, Craig L. Peterson, Joel D. Richter, Stewart Shuman and Mario Stevenson for helping to clear my thoughts and providing constructive comments.

I remain deeply grateful to my graduate (MS. degree) advisor, Dr. Guang-Hsiung Kou and Chu-Fang Lo at National Taiwan University. Their encouragement and support were my inspiration to continue my education and begin PhD. studies in the United States.

I would like to thank the members of the Dr. Tariq M. Rana's Laboratory, both past and present, for supporting me like a family.

I would also like to thank the members of Drs. Leroy F. Liu, Marc R. Gartenberg, Michael Hamsey and Nancy Walworth's Laboratories at UMDNJ/Rutgers for all their kindly help and friendship.

Lastly, I would like to thank my family and friends in Taiwan, especially my parents, my finance and his family for their love, understanding and continuous support.

CONTRIBUTIONS

In vitro GST pull-down assay

Dr. Elizabeth Coronel (Member of Dr. Tariq M. Rana's Laboratory)

Purification of Mammalian capping enzyme (Mce1)

Dr. Stewart Shuman (Sloan-kettering Institute, New York)

Dr. C. Kiong Ho (Sloan-kettering Institute, New York)

Purification of Human mRNA cap methyltransferase (Hcm1)

Dr. Beate Schwer (Weill Medical College of Cornell University, New York)

Dr. Nayanendu Saha (Weill Medical College of Cornell University, New York)

HIV infectivity assay

Dr. Mario Stevenson (UMass Center For AIDS Research)

Dr. Jean-Marc Jacque (UMass Center For AIDS Research)

Genechip hybridization, Staining, Washing, Scanning and Detection

Phyllis Spatrick (UMass Genomic Core Facility)

Genechip data analysis by using Affymetrix software

Dr. Hong Cao (Member of Dr. Tariq M. Rana's Laboratory)

ABSTRACT

Gene expression of human immunodeficiency virus type-1 (HIV-1), which causes Acquired Immunodeficiency Syndrome (AIDS), is regulated at the transcriptional level, where negative factors can block elongation that is overcome by HIV Tat protein and P-TEFb. P-TEFb, a positive elongation transcription factor with two subunits, CDK9 and Cyclin T1 (CycT1), catalyzes Tat-dependent phosphorylation of Ser-5 in the Pol II C-terminal domain (CTD), allowing production of longer mRNAs. Ser-5 phosphorylation enables the CTD to recruit mammalian mRNA capping enzyme (Mce1) and stimulate its guanylyltransferase activity. This dissertation demonstrates that stable binding of Mce1 and cap methyltransferase to template-engaged Pol II depends on CTD phosphorylation, but not on nascent RNA. Capping and methylation doesn't occur until nascent pre-mRNA become 19–22 nucleotides long. A second and novel pathway for recruiting and activating Mce1 involved direct physical interaction between the CTD, Tat and Mce1. Tat stimulated the guanylyltransferase and triphosphatase activities of Mce1, thereby enhancing the otherwise low efficiency of cotranscriptional capping of HIV mRNA. These findings imply that multiple mechanisms exist for coupling transcription elongation and mRNA processing at a checkpoint critical to HIV gene expression.

To elucidate P-TEFb's function in human (HeLa) cells, RNA interference (RNAi) was used to degrade mRNA for hCycT1 or CDK9. Down-regulation of P-TEFb expression by RNAi can be achieved without causing major toxic or lethal

effects and can control Tat transactivation and HIV replication in host cells. High-density oligonucleotide arrays were used to determine the effect of P-TEFb knockdown on global gene expression. Of 44,928 human genes analyzed, 25 were down-regulated and known or likely to be involved in cell proliferation and differentiation. These results provide new insight into P-TEFb function, its potent role in early embryonic development and strong evidence that P-TEFb is a new target for developing AIDS and cancer therapies.

To fulfill the promise of RNAi for treating infectious and human genetic diseases, structural and functional mechanisms underlying RNAi in human cells were studied. The status of the 5' hydroxyl terminus of the antisense strand of short interfering RNA (siRNA) duplexes determined RNAi activity, while a 3' terminus block was tolerated *in vivo*. A perfect A-form helix in siRNA was not required for RNAi, but was required for antisense-target RNA duplexes. Strikingly, crosslinking siRNA duplexes with psoralen did not completely block RNAi, indicating that complete unwinding of the siRNA helix is not necessary for RNAi *in vivo*. These results suggest that RNA amplification by RNA-dependent RNA polymerase is not essential for RNAi in human cells.

TABLE OF CONTENTS

Page No.

Title page.....i

Approval page.....ii

Dedication.....iii

Acknowledgement.....iv

Contributions.....v

Abstract.....vi

Table of contents.....viii

List of figures.....ix

List of abbreviations.....xii

Chapter I: Introduction.....1

Chapter II: HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA.....36

Chapter III: Tat stimulates cotranscriptional capping of HIV mRNA.....76

Chapter IV: Human transcription elongation factor (CDK9/CyclinT1): A new therapeutic target for AIDS and cancer?.....122

Chapter V: RNA interference in human cells: Basic structural and functional features of small interfering RNA.....170

Reference.....219

Appendix.....244

Curriculum vitae.....265

LIST OF FIGURES

Page No.

Chapter I: Introduction

Figure 1 Functional domains of HIV-1 Tat protein and secondary structure of wild-type TAR RNA.....22

Figure 2 Acquisition of the 5'-terminal m7GpppN cap through a series of three enzymatic reactions.....24

Chapter II: HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA

Figure 1 Tat interacts directly with Mce1 *in vitro*.....60

Figure 2 The C-terminal segment of Tat containing the RNA binding domain is sufficient for interaction with Mce1.....62

Figure 3 Tat enhances guanylylation of Mce1.....64

Figure 4 Tat enhances the RNA triphosphatase activity of full-length Mce1 but not the isolated triphosphatase domain.....66

Figure 5 The RNA triphosphatase domain of mammalian capping enzyme can bind to Tat.....68

Figure 6 Tat enhances RNA cap formation.....70

Figure 7 Tat enhances capping of TAR RNA.....72

Figure 8 Emerging connections between CTD phosphorylation, capping, and transcription elongation.....74

Chapter III: Tat stimulates cotranscriptional capping of HIV mRNA

Figure 1 Mammalian capping enzyme is not a component of PICs but can associate with TECs on the HIV-1 transcription unit.....108

Figure 2 HIV-1 mRNA capping during transcription elongation.....110

Figure 3 Mce1 binding to Pol II transcription complexes requires CTD phosphorylation but does not require nascent RNA.....112

Figure 4 Dissociation of Mce1 from the transcription complex during elongation is stimulated by a factor in nuclear extract.....114

Figure 5 Hcm1 associates with transcription elongation complexes and methylates the 5' Cap of nascent RNAs.....116

Figure 6 Tat stimulates cotranscriptional capping by Mce1.....118

Figure 7 Interaction of Tat and Mce1 *in vivo*.....120

Chapter IV: Human transcription elongation factor (CDK9/CyclinT1): A new therapeutic target for AIDS and cancer?

Figure 1 Specific down-regulation of P-TEFb expression by RNAi.....156

Figure 2 Kinetics of hCycT1 and CDK9 knockdown by siRNA.....159

Figure 3 P-TEFb down-regulation can be achieved without causing major toxic and lethal effect in HeLa cells.....161

Figure 4 hCycT1 and CDK9 duplex siRNAs inhibit HIV-1 Tat transactivation in Magi cells.....163

Figure 5 Small interfering RNAs targeting Cyclin T1 or CDK9 modulate HIV-1 infectivity.....166

Figure 6 Genomewide analysis of gene expression in P-TEFb

knockdown HeLa cells.....168

Chapter V: RNA interference in human cells: Basic structural and functional features of small interfering RNA

Figure 1 Dual fluorescence reporter assay system for RNAi analysis in HeLa cells.....200

Figure 2 Analysis of specific RNAi activities by western blotting.....202

Figure 3 Expression of GFP in HeLa cells treated with antisense or double-stranded siRNA targeting GFP.....204

Figure 4 Modification of GFP siRNA duplexes.....206

Figure 5 Fluorescence images showing RNA interference effects in living HeLa cells transfected with modified siRNA duplexes.....208

Figure 6 Quantitative analysis of RNAi effects in HeLa cells transfected with modified siRNAs.....210

Figure 7 Isolation of 5' end phosphorylated and 3' end biotinylated siRNA from HeLa cells.....212

Figure 8 RNA interference activities of covalently photocrosslinked duplex RNA in HeLa cells.....214

Figure 9 Isolation of psoralen-crosslinked siRNA from human cells.....217

LIST OF ABBREVIATIONS

AIDS	Acquired Immunodeficiency Syndrome
as	antisense strand
β-gal	β-galactosidase
CDK7	cyclin dependent kinase 7
CDK9	cyclin dependent kinase 9
CTD	carboxy-terminal domain
CTR	C-terminal repeat
CycT1	Cyclin T1
DRB	5,6-dichloro-1-β-D-ribofuranosylbenzimidazole
ds	double-stranded
DSIF	DRB sensitivity-inducting factor
GFP	green fluorescent protein
GTF	general transcription factor
Hcm1	human mRNA cap methyltransferase
HIV-1	Human immunodeficiency virus type-1
HMT	4'-hydroxymethyl-4,5',8-trimethylpsoralen
LTR	Long terminal repeat
Magi	multinucleate activation of galactosidase indicator cells
Mce1	mammalian capping enzyme 1
mm	mismatch
N3	amino modification
NELF	negative elongation factor
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis

PIC	pre-initiation complex
Pmn	Puromycin
P-TEFb	positive transcriptional elongation factor b
PTGS	post transcriptional gene silencing
RDRP	RNA-dependent RNA polymerase
RFP	red fluorescent protein
RISC	RNA-induced silencing complex
RNAi	RNA interference
RNA Pol II	RNA Polymerase II
RNA Pol IIA	unphosphorylated RNA Polymerase II
RNA Pol IIO	phosphorylated RNA Polymerase II
RT-PCR	reverse transcriptase-polymerase chain reaction
SDS	sodium-dodecyl sulfate
siRNA	small interfering RNA
ss	sense strand
TAK	Tat-associated kinase
TAR	*trans*-activation responsive
Tat	*trans*-activator of transcription
TEC	transcription elongation complex
TFIIH	transcription factor IIH
TRM	Tat-TAR RNA recognition motif

CHAPTER I

INTRODUCTION

Part I. Transcriptional Regulation of HIV-1 Gene Expression

1. Human immunodeficiency virus type-1

Human immunodeficiency virus type-1 (HIV-1), the etiological agent for Acquired Immunodeficiency Syndrome (AIDS) (7, 69, 132), is a member of the primate lineages of the lentivirus family of retroviruses. HIV-1 contains two identical copies of a linear single-stranded RNA genome of ~9.2 kb (63, 109). The virion RNA genome is converted into a linear double-strand (ds) DNA by viral reverse transcriptase in infected cells and then integrated into the host genome as proviral DNA. This ds DNA genome encodes nine open reading frames (63, 236). Three of them are polyproteins named gag, pol and env, which are subsequently proteolyzed into individual proteins. *Gag* and *env* encode the structural proteins that build up the core and the envelope of the virion. Gag (group specific antigen) precursor protein is cleaved into MA (matrix), CA (capsid), NC (nucleocapsid) and p6. Env (envelope) precursor protein is cleaved into SU (surface or gp120) and TM (transmembrane or gp41). Pol (Polymerase) comprises three proteins, PR (protease), RT (reverse transcriptase), and IN (integrase) and possesses essential enzymatic activities. *Nef*, *vpu*, *vif*, and *vpr* genes encode accessory proteins that are dispensable for viral growth in tissue culture cells but appear to help virus infectivity and replication *in vivo* (63). *Tat* and *rev* are two regulatory genes that can act *in trans* to control HIV-1 gene

expression and up-regulate virus replication through transactivating proteins (43, 63, 114).

2. Tat (*Trans*-Activator of Transcription) and TAR (*Trans*-Acting Responsive) RNA

Tat, the HIV-1 *trans*-activator of transcription, is a small nuclear protein encoded by two separate exons. Tat is essential for virus replication and is conserved in all primate lentiviruses (107). Naturally occurring HIV-1 Tat comprises 101 amino acids (107). However, an 86-amino acids version of Tat, which arose from the 101-amino acid form as a consequence of laboratory tissue culture passage, is commonly used for research and completely reflects biological functions for full-length Tat. (163). The extreme C-terminus of native Tat is apparently not needed for *ex vivo* propagation of HIV-1, although these amino acids remain conserved by viruses that replicate *in vivo* (183).

Tat protein is composed of several functional regions (Figure 1A)(183). The first 72 amino acids encoded by the first exon are both necessary and sufficient for TAR RNA binding (see below) and transcriptional activation. The acidic N-terminal region has a periodic arrangement of acidic, polar, and hydrophobic residues consistent with an amphipathic α-helix. A cysteine rich region (amino acids 22 to 37) contains seven cysteine residues. Individual mutations in six of the seven cysteines abolish Tat function. The "core" sequence of Tat (amino acids 37 to 48) contains hydrophobic amino acids and is essential

for *trans*-activation and specific high affinity binding of TAR RNA. The basic RNA-binding region (amino acids 48 to 59), adjacent to the core domain, contains six arginines and two lysines. This domain has TAR RNA-binding properties and mediates nuclear localization of Tat protein. A glutamine-rich region at the carboxyl-terminus of the first exon contains several regularly spaced glutamines. The second exon encodes a 14–amino-acid C-terminal sequence that is not required for *trans*-activation but does contain an RGD (arginine-glycine-aspartate) motif (amino aicds 78-80), which is used as a cell adhesion signal for binding to cellular integrins (18). Based on mutational analysis, Tat also can be divided into two functional domains. The first domain is the activation domain (amino acid 1 to 47) that mediates interaction with other cellular proteins required for Tat transactivation (204). The second functional domain, containing a basic region, is required for RNA binding as well as nuclear localization activities of Tat (52).

As a potent activator of HIV-1 long terminal repeat (LTR) region-specific transcription, Tat protein acts by binding to a nascent RNA element named TAR (*trans*-acting responsive element) (14, 203). TAR RNA is a 59-base stem-bulge-loop structure located at the 5' ends of all nascent HIV-1 transcripts (14). TAR RNA was originally localized to nucleotide +1 to +80 within the viral LTR (193). Subsequent deletion studies have established that the region from +19 to +42 incorporates the minimal domain that is both necessary and sufficient to support Tat transactivation *in vivo* (72, 106, 183, 203). The functional domain of TAR

RNA contains a six-nucleotide loop and a three-nucleotide pyrimidine bulge that separate two helical stem regions (Figure 1B). Key elements required for TAR RNA recognition by Tat have been defined by various approaches, such as mutagenesis, chemical probing, and peptide binding studies (183). It has been demonstrated that Tat interacts with U23 in the trinucleotide bulge domain. C24 and U25 function as spacers because they can be substituted by other nucleotides or linkers (33, 222). In addition to the trinucleotide bulge region, two base-pairs (bp) above and below the bulge also contribute significantly to Tat binding (33, 251). By using a site-specific photocrosslinking, Tat has been shown to interact with U23, U38, and U40 in the major groove of TAR RNA (246, 247). The loop region of TAR RNA is necessary for efficient transactivation, but not required for Tat binding. Recently, a cellular loop-binding protein has been identified and is required for Tat transactivation (252).

3. RNA polymerase II transcription complex

Transcription, copying the sequence of DNA template into a complementary RNA transcript by the enzyme RNA polymerase, is the first step of the central dogma. RNA polymerase, the enzyme that catalyzes DNA transcription, is a complex molecule containing many polypeptide chains. In eukaryotic cells there are three distinct RNA polymerase enzymes, designated RNA polymerases I, II and III, The three eukaryotic RNA polymerase were initially distinguished by their chromatographic property differences during

purification and by their sensitivity to α-amanitin. Sensitivity of RNA polymerase to α-amanitin indicates that each RNA polymerase is responsible for the transcription of a subclass of nuclear genes. Among them, RNA polymerase II (RNA Pol II) mediates the synthesis of the mature and functional messenger RNA (mRNA) that serves as a message for translation into protein.

Synthesis of mRNA by eukaryotic RNA Pol II can be divided into at least four major steps, i.e. preinitiation or the assembly of transcription complexes on a promoter, initiation, elongation, and termination and dissociation of Pol II from the DNA template (39, 60). In addition to RNA Pol II, a universal set of proteins called general transcription factors (GTFs), including, TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH, assemble at the core promoter of a gene and form the pre-initiation complex (PIC)(168). Following PIC formation, DNA melting, initiation, and promoter clearance take place rapidly in an energy-dependent manner (192). TFIIH plays a major role and may be the source of ATP dependence to facilitate these events. The helicase activity of TFIIH melts the ds DNA template through hydrolysis of the β-γ bond of ATP (or dATP) at about 10 bp region, just upstream of the transcription start site, to form an activated PIC. The kinase activity of TFIIH can phosphorylate the RNA polymerase II to facilitate promoter clearance and leads transcription complex into the elongation stage (142, 223). RNA pol II and the GTFs are sufficient to mediate basal level transcription. However, highly ordered, gene-specific transcription requires the cooperative

action of many cellular proteins and is tightly regulated by transcription regulators at many levels (240).

4. CTD (Carboxy-Terminal Domain) of RNA polymerase II

RNA Pol II is a large multisubunit enzyme complex. The largest subunit of RNA Pol II from eukaryoties contains a unique feature at its C-terminus, the so called CTD (carboxy-terminal domain) (45). Neither eukaryotic RNA polymerase I, III nor prokaryotic RNA polymerase contains a region resembling the CTD. In eukaryotic cells, the CTD of RNA Pol II is composed of tandem repeats (52 in vertebrates and 26 in budding yeast) of a heptapeptide with the conserved consensus motif Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS)(45). Each cycle of transcription involves the dynamic phosphorylation and dephosphorylation of the CTD on the serine residues 2 and 5 (Ser-2 and Ser-5) of each repeated unit. A large number of CTD kinases including TFIIH, cdc2, CTK1 and CDK9, have been implicated in the phosphorylation of the CTD at multiple sites *in vitro* and cause a characteristic mobility shift in SDS-polyacrylamide gel electrophoresis (PAGE) (12). A phosphatase responsible for dephosphorylation of the CTD has been discovered (12, 23, 24). RNA Pol II with an unphosphorylated CTD is designated RNA Pol IIA and polymerase with a highly phosphorylated CTD is designated RNA Pol IIO. RNA Pol IIA is efficiently recruited to promoters during assembly of a preinitiation complex (110, 141) whereas RNA Pol IIO is associated with the elongation complex (46). The CTD of RNA Pol II is required for transcriptional

activation or repression and for controlling mRNA synthesis and maturation events, such as capping, splicing and cleavage/polyadenylation (53, 92, 93, 152, 153, 270). The CTD functions as a platform for the recruitment and assembly of factors involved in pre-mRNA processing (13, 182).

5. Transcriptional elongation regulators

After successful initiation of RNA synthesis, the RNA pol II transcription complex enters the elongation stage. Transcriptional elongation by RNA Pol II is controlled by a number of *trans*-acting factors called transcription elongation factors (39, 40). During mRNA synthesis, RNA Pol II frequently encounters blocks to elongation; a primary cause of pause or arrest is thought to be the secondary structure formed by nascent transcripts. Transcription factors such as TFIIF, elongatin, ELL and TFIIS interact with RNA Pol II and thereby prevent it from pausing or reactivating from arrest (185, 207).

Recently, proteins involved in regulation of transcriptional elongation were discovered during studies aimed at understanding the mechanism of transcription inhibition by a nucleoside analog, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). DRB was originally discovered as an inhibitor of hnRNA and mRNA synthesis in HeLa cells (103, 202, 231, 232). Treatment of mammalian cells with DRB is lethal, but its initial effect is a dramatic reduction of mRNA (202). DRB treatment has caused the production of shortened transcripts from a variety of genes, suggesting that RNA pol II elongation was affected (232). However, DRB

had no effect on promoter-independent RNA pol II transcription and on transcription reconstituted by purified general transcription factors and RNA pol II (32, 119). These studies suggested that cellular proteins other than RNA Pol II and general transcription factors confer DRB sensitivity on elongation. The recent discovery of this class of transcription elongation factors, which include DRB-sensitivity inducing factor (DSIF), negative elongation factor (NELF), and positive transcription elongation factor (P-TEFb), has shed new light on the control of RNA Pol II elongation (40, 181, 243, 262).

5.1. P-TEFb (positive-transcriptional elongation factor b)

P-TEFb, the first and only known component of positive transcription factor, was originally identified and purified from *Drosophila melanogaster* Kc cell nuclear extract (147). P-TEFb is a factor required for the production of long transcripts by Pol II *in vitro* and is required for the bypass of an early block to elongation (147). P-TEFb consists of two subunits: the catalytic subunit cyclic-dependent kinase CDK9 (previously named PITALRE) (76) and the regulatory subunit Cyclin T1 (76, 175, 176, 252). CDK9 is a cdc2-related serine/threonine kinase (48, 73) and can form a complex with several cyclin partners in human cells: CyclinT1, Cyclin T2a, Cyclin T2b, and Cyclin K (66, 176). In HeLa cells, roughly 80% of CDK9 binds with Cyclin T1 and 10% with Cyclin T2a and Cyclin T2b (176). The assembly of P-TEFb involves the sequential interactions of Hsp70 and Hsp90/Cdc37 with CDK9. These two chaperone complexes,

CDK9/Hsp70 and CDK9/Hsp90/Cdc37, are important and essential precursors for the generation of the mature and active CDK9/CyclinT1 complex (167).

CDK9 and its Cyclin T partners are expressed in a wide variety of tissues (6, 49). *In situ* immunohistology and immunofluorescence studies of mouse embryos have indicated that murine CDK9/PITALRE protein appears to be expressed predominantly in tissues that are terminally differentiated, such as developing brain and dorsal root ganglia, areas of skeletal muscle, cardic muscle, and the lining of the developing intestinal epithelium (6). Analysis of the expression pattern of murine CDK9/PITALRE protein in adult mouse tissues by immunoblotting indicates that murine CDK9/PITALRE expression is ubiquitous, although steady-state protein levels are markedly higher in the brain, liver, lung, spleen and kidney. Kinase activity of CDK9/PITALRE kinase has also been detected in the same adult tissues and was highest in mouse brain, liver, spleen and lung. Only minimal kinase activity was found in the heart, muscle, and kidney. The expression pattern of CycT1 in adult human tissues was investigated by immunohistochemistry (49). Tissues of mesenchymal origin, such as cardiovascular and connective tissues, skeletal muscle cells, myocardial cells, adipocytes, chondrocytes and endothial cells, blood and lymphoid tissues showed high immunoreactivity. Astrocytes, oligodendroglial and microglial cells of brain tissue also showed high expression levels of Cyclin T1, while endocrine and reproductive systems showed low Cyclin T1 expression.

Although P-TEFb is not involved in the cell cycle, it has been suggested that P-TEFb may serve as a multifunctional cyclin-dependent kinase, which regulates various cellular events (216). It has also been proposed that the P-TEFb complex is required for global gene expression during embryonic development of *Caenorhabditis elegans* (208). Knockdown of CDK9 or CycT1 genes by siRNA in *C. elegans* embryos inhibits transcription of embryonic genes.

It has been demonstrated that P-TEFb is a Tat-associated-kinase-complex (TAK) and that the effect of Tat on transcription elongation depends on P-TEFb (265, 277). Studies of the general transcription elongation inhibitor, DRB, have shown that it inhibits P-TEFb protein kinase activity, thereby blocking a step in transcription elongation where P-TEFb is required. Pol II CTD and the Spt5 subunit of the DSIF complex have been identified as substrates for P-TEFb activity and will be discussed later. Depletion of P-TEFb from HeLa nuclear extracts greatly reduces the ability of RNA pol II to produce full-length transcripts and eliminates the DRB sensitivity of nuclear extracts (144, 146, 147, 175, 277). Moreover, in addition to DRB, P-TEFb is strongly inhibited by flavopiridol (L86-8275, HMR1275), a cyclin-dependent kinase (CDK) inhibitor being tested in clinical trials as a potential anti-cancer drug (117, 205). However, flavopiridol's target and the mechanism for its antiproliferative effects are unknown. Recent evidence indicates that flavopiridol inactivates P-TEFb and inhibits HIV-1 transcription (25, 26).

5.2. DSIF (DRB sensitivity-inducing factor) and NELF (negative elongation factor)

DSIF and NELF are two negative transcription elongation factors (N-TEF) that can impede RNA Pol II elongation (242, 261). DSIF was primarily identified for its ability to repress transcription in the presence of DRB (242). DSIF is a heterodimer comprising mammalian homologues of *Saccharomyces cerevisiae* Spt5 and Spt4 transcription factors. Spt4 possesses a putative zinc finger domain and most likely interacts with Spt5 through mainly hydrophobic interactions (261). Spt5 has several distinct domains including four KOW repeats and two C-terminal elements, CTR1 (C-terminal repeat 1) and CTR2 (103, 261). KOW domains, which are also found in *Escherichia coli* RNA polymerase-binding protein NusG, play a role in binding to RNA Pol II and affecting transcriptional termination and anti-termination (134). The CTR domains possess multiple amino acid repeats that are rich in serine and threonine residues, a similar composition to the C-terminal domain of RNA Pol II, which may provide phosphorylation sites for cellular kinases (103, 218).

Immunodepletion of DSIF p160 (hSpt5) from HeLa cell nuclear extract results in DRB-insensitive transcription and adding back DSIF restores DRB-sensitivity (242). Recent studies have shown that Spt4 and Spt5 function during early transcription elongation process, which is regulated by P-TEFb (242, 243). Immunodepletion of DSIF and P-TEFb restored transcription to normal levels, and addition of recombinant DSIF was able to repress transcription in a dose-

dependent manner (243). These results indicated that DSIF plays a role as a negative regulator in transcription in the absence of P-TEFb (243). However, In yeast, Spt4, Spt5, and Spt6 appear to have a positive effect on transcription elongation through the modulation of chromatin structure (80, 224, 225). Results from studies of *Drosophila* polytene chromosomes have shown that Spt5 and Spt6 co-localize with phosphorylated RNA Pol II and Cyclin T1 at a large number of transcriptionally active sites (3, 112). These findings provide support for the closely related role of Spt5 and Spt6 in active transcription elongation *in vivo*.

Purified NELF is composed of five polypeptides, A to E (66, 61, 59, 58, and 46 kDa, respectively) (261). NELF-A, the largest subunit of NELF, is encoded by Wolf-Hirschhorn syndrome (WHSC) 2 gene. A sequence analysis showed limit homology (27% identity) between the N-terminal half of NELF-A/WHSC2 and hepatitis delta antigen (HDAg), the only protein encoded by hepatitis delta virus (HDV) (260). NELF-E is the smallest subunit containing repeats of the dipeptide arg-asp (RD repeats), an RNA recognition motif (RRM), and a leucine zipper at the C-terminal end. RD is a putative RNA-binding protein with unknown function (261). NELF and DSIF function to cooperatively repress transcriptional elongation by RNA Pol II (261).

Unlike P-TEFb, neither DSIF nor NELF has protein kinase activity, suggesting that they are not direct targets of DRB. The connection between DRB and these three elongation factors was revealed by the finding that only active P-TEFb can counteract the negative effects of DSIF and NELF in a transcription

extract (278). Immunodepletion and *in vitro* transcription studies suggest that the main mechanism for P-TEFb control of RNA pol II transcription elongation *in vitro* is to alleviate the inhibition effect of DSIF and NELF by phosphorylating the CTD of RNA Pol II (263) and also the Spt5 subunit of DSIF (178, 243). Furthermore, recent studies have demonstrated a chromatin-specific elongation factor, FACT, which removes the negative regulation of transcription imposed by DSIF and NELF in cooperation with P-TEFb (241).

6. Coupling of transcription elongation and mRNA capping

mRNA processing plays an important role in the expression of eukaryotic genes and is specifically targeted to transcripts made by RNA Pol II. Placing a mammalian Pol II transcription unit under the control of a Pol III promoter results in a failure to cap, splice, or polyadenylate transcripts (211). How is pre-mRNA "identity" established? Accumulated evidence indicates that the unique Pol II CTD appears to occupy a central position as a platform for interacting with macromolecular assemblies that regulate mRNA synthesis and processing (53, 92, 93, 152, 153, 270). However, the CTD is more than a passive scaffold; indeed, binding to the CTD can regulate the activity of protein partners (13).

Acquisition of the 5'-terminal m7GpppN cap is the first modification event in mRNA biogenesis and capping facilitates downstream transactions such as splicing, polyadenylation, transport and translation. RNA capping is essential for cell growth, *i.e.* mutations of the triphosphatase, guanylyltransferase, or

methyltransferase components of the capping apparatus that abrogate their catalytic activity are lethal *in vivo* (65, 94, 145, 161, 200, 201, 214, 245).

Capping entails a series of three enzymatic reactions: (i) RNA triphosphatase (RTP) removes the phosphate from the 5'-triphosphate end of the nascent mRNA to form a diphosphate terminus; (ii) RNA guanylyltransferase (RGT) transfers GMP from GTP to the diphosphate RNA terminus to form GpppRNA; and (iii) RNA (guanine-7)-methyltransferase (MT) adds a methyl group to the N7 position of the cap guanine (213). The yeast *S. cerevisiae* encodes three proteins, each possessing one activity (213) while the mammalian capping apparatus consists of two components: a bifunctional triphosphatase-guanylyltransferase (mammalian capping enzyme, Mce1) and a separate cap methyltransferase (Hcm1) (94, 177, 195, 269).

Capping occurs shortly after transcription initiation when the 5'-end of the nascent RNA chain is extruded from the RNA-binding pocket of the elongating RNA polymerase (41, 81, 184). Capping of cellular RNAs *in vivo* is specific to transcripts synthesized by RNA polymerase II (Pol II) (79, 139). It has been suggested that this specific targeting is achieved through direct physical interaction of one or more components of the capping apparatus with the phosphorylated CTD of the largest subunit of Pol II (30, 31, 93, 94, 152, 177, 269). In the budding yeast *S. cerevisiae*, the guanylyltransferase (Ceg1) and methyltransferase (Abd1) bind directly to the phosphorylated CTD (31, 152). The mammalian capping enzyme, Mce1 binds to the phosphorylated CTD but not to

an unphosphorylated CTD (93, 94, 152, 269). Binding to CTD phosphorylated at Ser-5 of the YSPTSPS heptad stimulates the guanylyltransferase activity of Mce1 (93, 253). Recruitment of the capping apparatus to the yeast Pol II elongation complex *in vivo* requires the action of the TFIIH-associated CTD kinase Kin28 (equivalent to CDK7 in mammals), which phosphorylates Ser-5 of the CTD (128, 191, 199).

There is a fascinating link between capping and transcription elongation, which may serve as a checkpoint to ensure that Pol II commits to productive elongation only after transcript has been capped. For example, it was reported that hSpt5, the human homolog of yeast elongation factor Spt5, which also plays a role in Tat transactivation of HIV-1 gene expression at the level of transcription elongation (123, 256) interacts directly with the mammalian capping enzyme and stimulates its guanylyltransferase activity (253).

What happens after the cap has been added? *In vivo* crosslinking experiments In budding yeast showed that the complex of RNA triphosphatase and guanylyltransferase is released early, in a step coupled with removal of Ser-5 phosphorylation by the CTD phosphatase Fcp1 (199) or another phosphatase (29). Methyltransferase remains associated with Pol II even at the 3' end of the gene, indicating that the methyltransferase may travel with Pol II for the whole length of the gene or it could be bound and released in a dynamic equilibrium throughout elongation of the RNA chain (199). At present, no evidences in mammalian systems has shown whether cellular capping enzymes remain

associated with the elongation RNA polymerase after the cap structure has been formed or whether they are jettisoned to make room for other processing assemblies.

7. Regulation of HIV-1 gene expression

In theory, each of transcription steps, i.e, preinitiation, initiation, elongation, and termination (39, 60) can be regulated, resulting in either overall stimulation or repression of transcription. While it is widely accepted that the preinitiation step plays a criticial role in transcriptional regulation, over the past several years increasing attention has focused on the role of the elongation step. A number of genes including *c-myc*, *c-fos*, *c-myb*, *hsp70*, and HIV long terminal repeat (LTR) region, are potentially regulated at the elongation stage of transcription (109, 129, 194, 268). Among them, the mechanism of HIV-1 gene expression is one elegant example of transcription elongation control.

As described above, RNA synthesis from HIV proviral DNA is mediated by host RNA polymerase II (RNA Pol II). Shortly after the initation of transcription, RNA Pol II is subjected to both negative and positive control by the transcription regulators listed above. In the absence of HIV-1 Tat, the initiated transcripts prematurely terminate within 100 nucleotides (nt) of the start site. In the presence of Tat, elongation is very effective, and hyperphosphorylation of the CTD of RNA Pol II has been suggested as the molecular event underlying Tat transactivation (108, 113). In addition to the binding of Tat with TAR RNA, other cellular factors

are also involved in the stimulation of RNA pol II elongation processivity by Tat. A gene located at human chromosome 12 encodes a species-specific cofactor for Tat (1). A cellular protein kinase complex called TAK (Tat-associated kinase) was identified that specifically binds to the activation domain of Tat and can phosphorylate the CTD of RNA Pol II (90, 91). The kinase component of TAK was then shown to be identical to a previously identified cdc2-related kinase named PITALRE. PITALRE has also been identified as the kinase component of an essential multiprotein complex, P-TEFb, which is involved in transcription elongation in *Drosophila* and mammals (175, 277). Furthermore, PITALRE kinase was identified during an *in vitro* screening of drug inhibitors of Tat activity (144). Moreover, Cyclin T was independently cloned as a Tat-associated protein (252).

The detailed interactions of Tat, TAR, and P-TEFb have been studied very well (181, 183). Although CDK9 can form a complex with various cyclin partners, it is clear that CDK9/Cyclin T1 is the only target of Tat. Human Cyclin T2a, Cyclin T2b as well as Cyclin K do not support Tat transactivation (255). Rodent cells can be made permissive for Tat transactivation by making them express human Cyclin T1 or by making a single amino acid (A261C) change in mouse Cyclin T1. The cysteine residue at position 261 of Cyclin T1 is required for a zinc-dependent interaction between Cyclin T1 and Tat (16, 71, 255).

Tat appears to contact residues in the carboxy-terminal boundary of human Cyclin T1's (hCycT1's) cyclin domain that are not critical for binding of

hCycT1 to CDK9 (16, 68, 71, 104, 255, 276). The hCycT1 sequence containing amino acids 1-272 is sufficient to form complexes with Tat-TAR, CDK9, and to activate transcription by Tat (16, 68, 71, 104, 255, 276). The sequence and structural determinants for high affinity hCycT1-Tat-TAR RNA ternary complex formation have been identified (188). Human Cyclin T1 residues 250-262 represent the Tat-TAR RNA recognition motif (TRM)(71) and have been shown to interact with one side of the TAR RNA loop and enhance interaction of Tat residue K50 with the other side of the loop (189). The binding of Cyclin T1 with Tat increases the binding affinity of Tat for TAR (71, 252, 273). hCycT1 and Tat binding to TAR RNA is highly cooperative, with a capacity of 85%, Hill coefficient of 2.7 and a dissociation constant (K_D) of 2.45 nM (188). All these results indicate that TAR RNA provides a scaffold for two protein partners to bind and assemble a regulatory switch in HIV replication.

Besides P-TEFb-Tat-TAR complex, Pol II CTD, Spt5, and Tat are also intimately connected to regulation of HIV gene expression (reviewed in 181). Human Spt5 and its binding partner hSpt4 comprise the transcription elongation regulatory factor DSIF (DRB sensitivity-inducing factor). DSIF binds to Pol II and, in concert with NELF (negative elongation factor), represses elongation at promoter-proximal positions in the transcription unit. Escape from the repressive effect of DSIF/NELF requires the action of P-TEFb, a DRB-sensitive protein kinase that phosphorylates both the Pol II CTD and the Spt5 subunit of DSIF. The interactions of these regulatory factors with Pol II at discrete functional

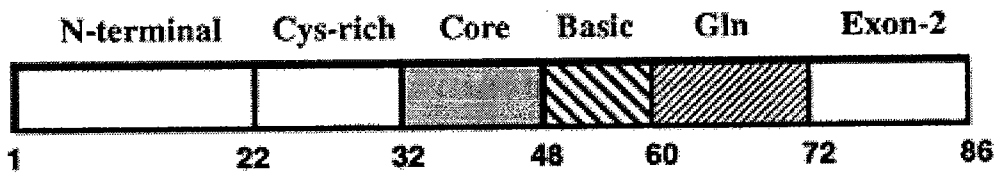
stages of transcription from the HIV-1 LTR promoter *in vitro* have been analyzed (178, 179). P-TEFb is a component of the Pol II preinitiation complex (PIC) and travels with the transcription elongation complex (TEC) along the HIV transcription unit (179). In contrast, DSIF and NELF are not present in the PIC, but associate with the TEC at promoter proximal positions and then travel with the TECs down the template (178). It has been demonstrated that the kinase activity of P-TEFb is able to phosphorylate a number of cellular proteins, including the CTD of RNA Pol II and Spt5 *in vitro* (101, 103, 179, 256, 275, 276). Furthermore, P-TEFb functions to counteract negative elongation factors DSIF as well as NELF (261). Although DSIF was discovered as a negative elongation factor, Spt5, the subunit of DSIF, was identified independently as a factor involved in HIV-1 Tat transactivation and functions as a positive elongation factor (123, 256). Taken together, the regulation of HIV-1 gene expression requires not only Tat-TAR-P-TEFb interaction, but also the interactions between RNA Pol II transcription complexes and the positive as well as negative transcriptional elongation factors.

How does capping fit into the scheme of HIV transcription regulation? At present, little is known about the dynamics of the interactions of mammalian capping and methylating enzymes with mammalian Pol II during the transcription cycle or about the efficiency and timing of the cap guanylylation and methylation steps. It has been proposed that Tat expression would result in a marked increase in the steady state level of the mRNA and stimulated the translation of

mRNAs synthesized from the HIV transcription unit (44). Since Tat transactivation and capping are both correlated with CTD phosphorylation at an early stage of transcription elongation, it is conceivable that Tat may interact physically or functionally with mammalian capping enzyme. Spt5-induced arrest at promoter-proximal sites in the HIV transcription unit might ensure a temporal window for recruitment of the capping enzymes and stimulate modification of the 5' end of the HIV pre-mRNA.

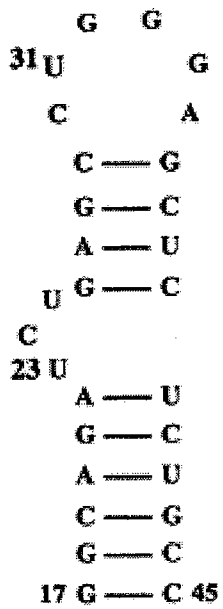
Figure 1. (A) Functional domains of HIV-1 Tat protein. (B) Secondary structure of wild-type TAR RNA. Wild-type TAR-RNA spans the minimal sequences that are required for Tat *trans*-activation.

(A)



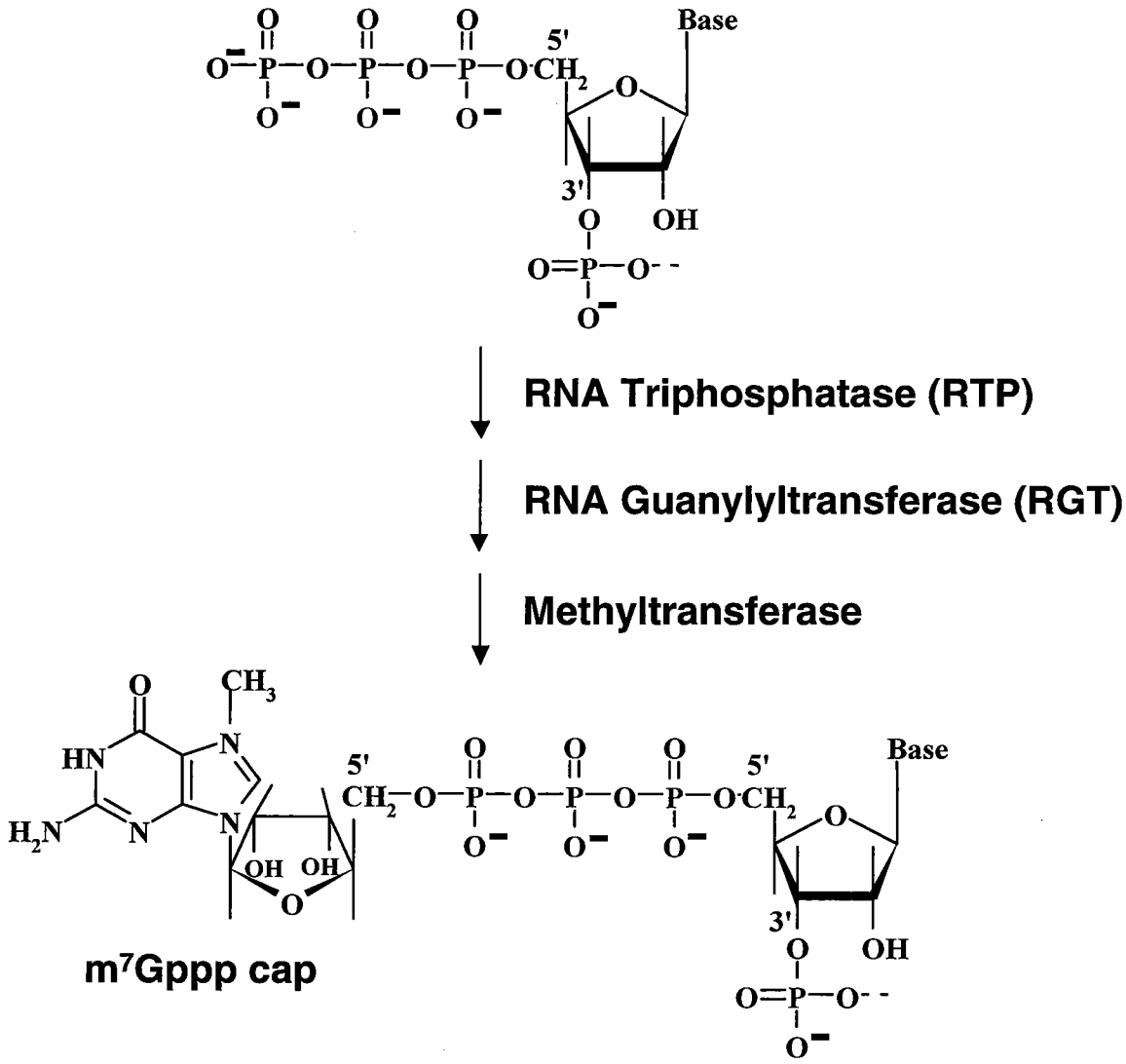
HIV-1 Tat

(B)



TAR RNA

Figure 2. Acquisition of the 5'-terminal m7GpppN cap through a series of three enzymatic reactions. Capping entails a series of three enzymatic reactions: (i) RNA triphosphatase (RTP) removes the phosphate from the 5'-triphosphate end of the nascent mRNA to form a diphosphate terminus; (ii) RNA guanylyltransferase (RGT) transfers GMP from GTP to the diphosphate RNA terminus to form GpppRNA; and (iii) RNA (guanine-7)-methyltransferase (MT) adds a methyl group to the N7 position of the cap guanine.


Base
5'
3'
OH
RNA Triphosphatase (RTP)
RNA Guanylyltransferase (RGT)
Methyltransferase
m7Gppp cap

Part II. Structural and Functional Study of RNA Interference

1. RNA interference is broadly distributed in nature

RNAi interference (RNAi) is the process whereby ds RNA induces the sequence-specific degradation of homologous mRNA. Fire and associates first reported on RNA interference when they demonstrated potent and specific genetic interference upon injecting dsRNA into *C. elegans* (62). However, the underlying phenomenon, known as posttranscriptional gene silencing phenomena (PTGS), had initially been described in transgenic plants (as PTGS or cosuppression) and *Neurospora crassa* (as quelling) (reviewed in 84, 206). PTGS was observed when the introduction of all or a portion of a transgene occasionally resulted in the loss of expression from the corresponding endogenous gene with no disruption in gene transcription. It has become clear that dsRNA-induced silencing phenomena are present in evolutionarily diverse organisms, e.g., nematodes, plants, fungi, and trypanosomes (8, 37, 62, 84, 122, 150, 206, 215, 238, 248). Although PTGS and RNAi were identified independently, genetic and biochemical analyses suggest a strong evolutionary link between these multi-step pathways (8-10).

2. RNA interference pathway

The discovery of RNAi was followed by studies of its mechanism. Work in *C. elegans* indicated that RNAi had at least two important steps, the first of which probably involved the generation of a sequence-specific silencing agent (78). A

strong candidate for this agent was a class of short RNAs originally reported by Andrew Hamilton and David Baulcombe (82). They found that *Arabidopsis* plants undergoing transgene- or virus-induced PTGS contained 21-25 nt RNAs that were complementary to both strands of the silenced gene and that had been processed from a long dsRNA precursor. An ATP-dependent, processive cleavage of dsRNA into 21- to 23-nt short interfering RNAs (siRNAs) was recapitulated *in vitro* using *Drosophila* embryo lysates and S2 cell extracts (15, 271). Accumulated evidence revealed that 21- to 23-nt RNAs were associated with RNAi silencing in *C. elegans* and *Drosophila* (83, 171, 264). Cloning and sequencing of these native siRNA duplexes revealed a very specific structure: 21-to 23-nt dsRNAs with 2-nt 3' end overhangs (57). Cleavage of the dsRNA *in vitro* and in cultured *Drosophila* cells requires the multidomain RNase III protein, Dicer, a member of the RNase III family of dsRNA-specific endonucleases (15). Cleavage products containing 5' phosphate and 3' hydroxyl termini are then incorporated into a protein complex called RNA-induced silencing complex (RISC) (83). ATP-dependent unwinding of the siRNA duplex generates an active complex, RISC* (the asterisk indicates the active conformation of the complex) (166). Guided by the antisense strand of siRNA, RISC* recognizes and cleaves the corresponding mRNA (57, 83, 166).

Functional anatomy studies of synthetic siRNA in *Drosophila* cell lysates have demonstrated that each siRNA duplex cleaves its target RNA at a single site (58). The 5' end of the guide siRNA sets the ruler for defining the position of

target RNA cleavage (58). 5' phosphorylation of the antisense strand is required for effective RNA interference *in vitro* (166). Mutation studies have shown that a single mutation within the center of a siRNA duplex discriminates between mismatched targets (58). However, none of these phenomena have been demonstrated *in vivo*, especially in mammalian systems.

A particularly fascinating aspect of RNAi is its extraordinary efficiency. Conversion of the long trigger dsRNA into many 21- to 23-nt siRNA fragments would, itself, provide some degree of amplification. Another plausible explanation for the potency of interference is that the RISC* is a multiple-turnover enzyme (98), which can catalytically perform the targeting and cleavage activity.

3. Enzymes of RNA interference

Several groups are actively pursuing the identification and characterization of enzymes in RNAi and cosuppression. In *C. elegans*, initial mutant screens have generated ~80 candidates, of which five have been specifically identified: RDE-1, RDE-2, RDE-3, RDE-4 and Mut-7 (78, 121, 122, 226). Selection of mutations after cosuppression in *Arabidopsis* have identified homologs of the genes coding for the enzymes in *C. elegans* (47, 59, 160).

DCR-1 (the gene coding for Dicer) comprises an ATP-dependent RNA helicase domain, a Piwi/Argonaute/Zwille (PAZ) domain, a tandem RNase III domain, and a dsRNA-binding-domain. Deletion of the *Dcr-1* in worms abrogates RNAi and leads to misregulation of developmental timing (77, 120). Dicer is

required for cleavage of the dsRNA *in vitro* and in cultured *Drosophila* cells to produce siRNAs with characteristic structure of 5' phosphate and 3' OH termini with 2-nt overhangs. The presence of two RNase III domains suggests that Dicer might cleave dsRNA as a monomer. The dsRNA-binding domain could position the enzyme on the substrate, and the two catalytic domains could hydrolyze bonds in both strands of the dsRNAs.

The involvement of RNA-dependent RNA polymerase (RdRP) in the amplification process has recently been postulated since genetic screening has identified the gene for RdRP as a requirement for gene silencing in plants, fungi, and worms (36, 47, 160, 215). A random degradative PCR model has been suggested (136, 164, 215), in which siRNA serves as the primer for the RdRP reaction. The siRNA-primed RdRP converts target mRNA into dsRNA, which can serve as Dicer substrates, initiating the RdRP chain reaction. Certain structural features of siRNA, including the 3' hydroxyl group and 5' phosphate group, are critical for the RdRP reaction and RNA ligation (136, 215). Although RdRP activity has been reported in *Drosophila* embryo lysates (136), a homolog of RdRP has not been identified in available mammalian genomic sequences.

Rde-1, which is important for RNAi in *C. elegans*, is a member of a large *Piwi/STING/Argonaute/Zwille/eIF2C* family. There are 23 related genes in *C. elegans* (226), four family members in *Drosophila* and several in humans. *Arabidopsis* encodes eight genes related to *rde-1*. Mutation in two of these genes, *ago-1* and *zll/pnh*, results in developmental defects. Recently, EIF2C has

been identified as the major component of RISC complex isolated from HeLa cytolasma extract by biotinylated double stranded RNA pull down assay (148).

Rde-4 encodes a dsRNA binding protein required for RNAi (227). The stable interaction between RDE-4 and dsRNA requires *rde-1* (+) activity *in vivo*. RDE-4 protein also interacts *in vivo* with DCR-1 proteins and with a conserved DexH-box helicase that is required for RNAi in *C. elegans* (227). RNA sequences bound to RDE-4 were restricted to regions found within the trigger dsRNA, suggesting that RDE-4 functions during the initial steps of RNAi to recognize foreign dsRNA and to present this dsRNA to DCR-1 for processing (227).

Mut-7, the putative *C. elegans* exonuclease, acts downstream in the RNAi pathway, as well as in repressing transposition of mobile genetic elements (59). Its homolog in *Neurospora*, QDE-3, which belongs to the RecQ DNA/RNA helicase family (38), is essential for the post transcriptional gene silencing phenomenon, quelling. However, WRN, BLM and RecQ1 are not involoved in dsRNA-induced sequence-specific mRNA degradation in mammals, suggesting potential differences with the mammalian RNAi pathway (220).

4. dsRNA-induced gene silencing in somatic mammalian cells

Tuschl and colleagues (55) have demonstrated that RNAi can be induced in numerous mammalian cell lines by introducing synthetic 21-nt siRNAs. By virtue of their small size, these siRNAs avoid provoking an interferon response that activates the protein kinase PKR (219). The principal method used to

introduce the small siRNAs was by co-transfection (using lipofectamine) with plasmids expressing marker genes or control sequences. The 21- to 23-nt dsRNAs consisted of chemically synthesized and annealed single-stranded sense and antisense RNAs with 2-nt 3' overhangs. Several somatic mammalian cell lines were tested, including murine, non-human primate and human-derived cells. All theses cell lines showed evidence for sequence-specific inhibition of the target gene when the appropriate small dsRNA was transfected. Ability of dsRNA to down-regulate the expression of an endogenously expressed protein has been investigated as well. dsRNAs corresponding to the nuclear envelope protein lamin A/C induced ~90% reduction in protein levels 40-45 hous after initiation of transfection (55). More and more researchers are seizing siRNA technology to elucidate mammalian gene function, and the types of genes that are being knocked down vary widely, from structural to catalytic proteins (155).

5. Delivery vectors for siRNA

Synthetic siRNAs can be delivered into cells by cationic liposome (lipofectamine) transfection as described above. However, these exogenously added siRNAs have only a short-term silencing effect (4~5 days) (55). Several strategies using recombinant DNA constructs to express siRNA duplexes within cells have been explored to see if they might allow long-term target gene suppression. Mammalian Pol III promoter system (U6 small nuclear RNA [snRNA] promoter) has been reported to be capable of expressing functional ds

siRNAs (131, 157, 172, 221, 267). Transcriptional termination by RNA Pol III occurs at runs of four consecutive T residues in the DNA template, providing a mechanism to end a siRNA transcript at a sepecific sequence. The siRNA corresponds to the sequence of the target gene (21 mer), and 5'-3' and 3'-5' orientations can be expressed in the same or separate constructs. Hairpin siRNAs, driven by U6 snRNA promoter and expressed in cultured cells, can inhibit expression of their target genes (131, 157, 172, 221, 267). Constructs containing siRNA sequences under the control of the T7 promoter can also be made into functional siRNAs when cotransfected into cells with a vector expressing T7 RNA polymerase (105). It has been shown that animal cells express a range of ~22-nt noncoding RNAs termed micro RNA (miRNAs), which can regulate gene expression at posttranscriptional or translational levels during embryonic development. One common feature of miRNAs is that they are all excised from a ~70-nt precursor RNA stem-loop, probably by Dicer. By substituting the stem sequences of the miRNA precursor with a designed miRNA sequence, a vector under control of the RNA Pol II promoter can be constructed, and can be used to produce siRNAs that initiate RNAi against specific mRNA targets in mammalian cells (272). When expressed by DNA vectors containing polymerase III promoters, micro-RNA designed hairpin are also active on silencing gene expression (154). Viral-mediated delivery mechanism can result in specific silencing of targeted genes through expression of siRNA. Recombinant adenoviruses harboring siRNA under RNA Pol II promoter transcription control

have been generated. Infection of HeLa cells by these recombinant adenoviruses diminishes endogenous target gene expression. This recombinant adenovirus vector delivery strategy holds promise in therapeutic applications, since such vectors reduce the expression of genes targeted by the siRNAs injected into transgenic mice. (258).

6. RNAi in reverse genetics and its potential as a therapeutic approach

Reverse genetics is a genetic analysis that proceeds from genotype to phenotype by gene-manipulation techniques. The ability of siRNA technology to knock down the expression of any given gene promises to revolutionize reverse genetic approaches. In *C. elegans*, the direct injection of dsRNA has been used to establish or confirm the function of numerous genes and to assign phenotypic effects to many genes that had previously been identified only as computer-predicted open reading frames (64, 75). The direct injection of dsRNA into *Drosophila* embryos has been used to study the function of several genes expressed during development (198) Reports using small dsRNA-triggered RNAi in human cells to determine gene function have recently been published (56, 155).

PTGS has been suggested to protect plant tissues from viral infection (249). Initial studies to test the potential application of synthetic siRNAs to inhibit virus infection in vertebrates have been promising (155). Decreased mRNA expression by the respiratory syncytial virus (RSV), a negative-strand virus and

the causative agent of a severe respiratory disease, has been successfully demonstrated by using RNAi technology (17). siRNA-mediated silencing of HIV has also been seen in cultured cell lines and in human primary T cells (22, 35, 95, 105, 131, 165). These studies show that siRNAs can inhibit viral replication at several stages of infection. Infection can also be blocked by targeting either viral genes (*tat*, *rev*, *nef*, *vif*, *p24*) or host genes encoding CD4 and CCR5 receptors that are involved in the viral life cycle. siRNA can also confer resistance to poliovirus (74).

Recent data indicate that the *in vivo* delivery of siRNAs in mice is possible. RNase III prepared siRNA, delivered by electroporation into postimplantation mouse embryos, efficiently silenced reporter gene expression in different regions of the neural tube or other cavities of the mouse embryo (21). Rapid injection of a siRNA-containing physiological solution into the tail vein of a postnatal mice inhibited reporter transgene expression (133, 151). This RNAi effect depended on the siRNA dose and persisted for ~4 days after siRNA delivery (133). The therapeutic potential of this technique has been demonstrated by effective targeting of a sequence from hepatitis C virus (151). siRNA expressed from viral vectors in transgenic mice specifically reduce neuronal levels of polyglutamine protein with a corresponding decrease in disease protein aggregation, which is the cause of at least nine inherited neurodegenerative diseases, providing therapy for dominantly human diseases (258).

It is certain that the ability of siRNAs to silence specific genes, either when transfected directly as siRNAs or generated from DNA vectors, will make RNAi technology a powerful method in functional studies of human genes. However, to fulfil the promise of using RNAi technology as a therapeutic approach to treating infectious disease and human genetic diseases, structural and functional studies are needed to elucidate the basic mechanism underlying RNA interference in human cells.

CHAPTER II

HIV-1 TAT PROTEIN INTERACTS WITH MAMMALIAN CAPPING ENZYME AND STIMULATES CAPPING OF TAR RNA

ABSTRACT

HIV gene expression is subject to a transcriptional checkpoint, whereby negative transcription elongation factors induce an elongation block that is overcome by HIV Tat protein in conjunction with P-TEFb. P-TEFb is a cyclin-dependent kinase that catalyzes Tat-dependent phosphorylation of Ser-5 of the Pol II C-terminal domain (CTD). Ser-5 phosphorylation confers on the CTD the ability to recruit the mammalian mRNA capping enzyme (Mce1) and stimulate its guanylyltransferase activity. Here we show that Tat spearheads a second and novel pathway of capping enzyme recruitment and activation via a direct physical interaction between the C-terminal domain of Tat and Mce1. Tat stimulates the guanylyltransferase and triphosphatase activities of Mce1 and thereby enhances the otherwise low efficiency of cap formation on a TAR stem-loop RNA. Our findings suggest that multiple mechanisms exist for coupling transcription elongation and mRNA processing.

INTRODUCTION

mRNA processing plays an important role in the expression of eukaryotic genes, and the earliest modification event is the formation of the 5'-terminal m7GpppN cap. Capping entails a series of three enzymatic reactions: (i) RNA triphosphatase removes the phosphate from the 5'-triphosphate end of the nascent mRNA to form a diphosphate terminus; (ii) RNA guanylyltransferase transfers GMP from GTP to the diphosphate RNA terminus to form GpppRNA; and (iii) RNA (guanine-7)-methyltransferase adds a methyl group to the N7 position of the cap guanine (213). RNA capping is essential for cell growth, *i.e.* mutations of the triphosphatase, guanylyltransferase, or methyltransferase components of the capping apparatus that abrogate their catalytic activity are lethal *in vivo* (65, 94, 145, 201, 214, 245). The m7G cap facilitates translation initiation (161). Failure to cap pre-mRNAs results in their accelerated decay through the agency of a 5' exoribonuclease (200).

Capping occurs shortly after transcription initiation when the 5'-end of the nascent RNA chain is extruded from the RNA-binding pocket of the elongating RNA polymerase (41, 81, 184). Capping of cellular RNAs *in vivo* is specific to transcripts synthesized by RNA polymerase II (Pol II) (79, 139). It has been suggested that this specific targeting is achieved through direct physical interaction of one or more components of the capping apparatus with the phosphorylated CTD of the largest subunit of Pol II (30, 31, 93, 94, 152, 177, 269).

The CTD is unique to Pol II and consists of a tandemly repeated heptapeptide motif with the consensus sequence YSPTSPS that is differentially phosphorylated during the transcription cycle (46). Phosphorylation of the CTD correlates with the release of preinitiation complexes from the promoter and recruitment of the capping enzyme to the transcription elongation complex. In the budding yeast *Saccharomyces cerevisiae*, the guanylyltransferase (Ceg1) and methyltransferase (Abd1) bind directly to the phosphorylated CTD (31, 152). The mammalian capping enzyme, Mce1, a bifunctional 597-amino acid polypeptide with both RNA triphosphatase and guanylyltransferase activities, binds to the phosphorylated CTD but not to an unphosphorylated CTD (93, 94, 152, 269). Binding to CTD phosphorylated at Ser-5 of the YSPTSPS heptad stimulates the guanylyltransferase activity of Mce1 (93, 253). Although interaction between Pol II and capping enzymes offers an elegant explanation of the specific targeting of capping enzyme to nascent pre-mRNAs, it is conceivable that other factors are also involved in linking capping to transcription. For example, it was reported that hSpt5, the human homolog of yeast elongation factor Spt5, interacts directly with the mammalian capping enzyme and stimulates its guanylyltransferase activity (253). hSpt5 also plays a role in Tat transactivation of HIV-1 gene expression at the level of transcription elongation (123, 256).

HIV-1 Tat is a small RNA-binding protein required for efficient transcription of HIV genes. Tat binds specifically to a structured RNA element, TAR, located at the 5'-end of the nascent HIV transcript. Tat contains two important functional

domains: an arginine-rich region that mediates the binding of Tat to TAR RNA and an activation domain that mediates interactions with cellular factors. Tat functions through TAR to control an early step in transcription elongation that is sensitive to protein kinase inhibitors and requires the Pol II CTD (108). Tat increases the processivity of RNA polymerase complexes that would otherwise prematurely terminate. This function of Tat is predicated on its ability to enhance the activity of a positive transcription elongation factor, P-TEFb (181).

Components of the P-TEFb complex required for its activity include a catalytic protein kinase subunit Cdk9 (previously known as PITALRE) (144, 277) and regulatory subunits cyclin T1, cyclin T2a, or cyclin T2b, which associate with Cdk9 and increase its kinase activity (176, 252, 277). Cyclin T1 interacts directly with the activation domain of Tat. When the proteins are bound to TAR RNA, Tat interacts with the bulge region, whereas cyclin T1 binds to the loop segment (252). Phosphorylation of the Pol II CTD by P-TEFb kinase is stimulated by Tat and leads to the formation of processive transcription elongation complexes.

Because Tat transactivation and capping are both correlated with CTD phosphorylation at an early stage of transcription elongation, it is conceivable that Tat may interact physically or functionally with mammalian capping enzyme. Here, we show a direct association between Tat and Mce1 *in vitro*. We find that Tat stimulates mRNA capping *in vitro* by enhancing the triphosphatase and guanylyltransferase activities of Mce1. Moreover, Tat stimulates the capping of TAR mRNA, which is not guanylylated efficiently by Mce1, presumably because

the 5'-terminus is encompassed within a stable RNA hairpin. We suggest a model whereby Tat stimulation of TAR mRNA capping contributes to the activation of HIV gene expression.

MATERIALS AND METHODS

Expression and purification of recombinant wild-type and mutant Tat. Recombinant wild-type HIV-1 Tat, hemagglutinin (HA)-tagged Tat, and Tat deletion mutants were expressed in *Escherichia coli* as glutathione *S*-transferase (GST) fusion proteins. These fusion proteins consisted of an N-terminal GST moiety followed by a thrombin cleavage site and variable C-terminal polypeptides segments comprising wild-type Tat-(1-86) with or without HA tag; Tat2/36 (a deletion from amino acids 2-36 in the transactivation domain); Tat48 (a deletion of amino acids 49-86 including the RNA binding domain). Recombinant fusion proteins were purified to apparent homogeneity from bacterial lysates by glutathione-Sepharose affinity chromatography. Briefly, lysates were mixed with a 1 ml slurry of glutathione-Sepharose beads (Amersham Pharmacia Biotech) for 1 h at 4 °C. The beads were then poured into a column and washed with 20 ml of PBS (136 mM NaCl, 2.6 mM KCl, 10 mM Na_2HPO4, 1.76 mM KH_2PO4, pH 7.4) containing 1% Triton X-100, 1 mM EDTA, and 50 μg/ml phenylmethylsulfonyl fluoride. The immobilized GST·Tat beads were used in protein binding assays described below. Alternatively, the Tat proteins were recovered from glutathione-Sepharose beads by thrombin cleavage according to previously described procedures (187). The eluted Tat proteins were stored in buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM DTT, and 2.5 mM $CaCl_2$ at 80 °C.

Recombinant capping enzymes. Full-length mouse capping enzyme Mce1, the N-terminal RNA triphosphatase domain Mce1-(1-210), and the C-

terminal RNA guanylyltransferase domain Mce1-(211-597) were produced in bacteria as N-terminal His-tagged fusions and purified as described (94).

***In vitro* assay of the binding of GST·Tat to mammalian capping enzyme.** Reaction mixtures containing 3-5 μg of the wild-type or mutant (2/36, 48) GST·Tat proteins bound to glutathione-Sepharose beads and 1 μg of the wild-type or mutant Mce1 proteins in 300 μl of binding buffer (20 mM Tris-HCl, pH 7.9, 1% Triton X-100, 0.5% Nonidet P-40, 5 mM DTT, 0.2 mM ZnCl2, 0.1% bovine serum albumin) supplemented with a protease inhibitor mixture (Roche Molecular Biochemicals) were incubated at 4°C for 2 h. The beads were washed four times in 600 μl of washing buffer (30 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM $CaCl_2$, 5 mM DTT) and then resuspended in 200 μl of washing buffer and split into two equal aliquots. One of the aliquots was treated with 10 units of thrombin (Amersham Pharmacia Biotech) for 10 min at room temperature. Thrombin digestion was then quenched by adding 20 μg of phenylmethylsulfonyl fluoride. The second aliquot was not treated with thrombin. The beads were then separated from the supernatant by centrifugation. The supernatant fractions were resolved by SDS-PAGE. The protein contents of the gel were transferred to a polyvinylidene difluoride membrane that was then immunoblotted with rabbit antiserum raised against the guanylyltransferase domain Mce1-(211-597). Immune complexes were detected by using an ECL Western blotting kit according to the instructions of the vendor (Amersham Pharmacia Biotech).

***In vitro* assay of the binding of HA·Tat to the triphosphatase and guanylyltransferase domains of mammalian capping enzyme.** 40 μg of His-tagged Mce1-(1-210) or Mce1-(211-597)was adsorbed to 100 μl of Ni2+-agarose beads (Qiagen) during a 1-h incubation at 4°C in 500 μl of Ni2+-binding buffer A (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10% glycerol) supplemented with protease inhibitor mixture. The beads were washed four times with 1 ml of binding buffer A and resuspended in 250 μl of binding buffer A. The beads were then incubated for 1 h at 4°C with 2 μg of purified HA-tagged Tat protein in binding buffer A. The beads were washed six times with 1 ml of binding buffer A, and the bound proteins were then stripped from the beads by boiling in 25 μl of SDS-PAGE loading buffer (50 mM Tris-HCl, pH 6.8,12% glycerol, 4% SDS, 100 mM DTT, and 0.01% Coomassie Blue G-250). Polypeptides were resolved on 8% polyacrylamide gels. The gel contents were transferred to a polyvinylidene difluoride membrane, and Tat protein was detected by immunoblotting with a biotin-conjugated antibody against the HA tag. The blot was developed using a chemiluminescence kit.

Guanylyltransferase assay. Guanylyltransferase activity of capping enzyme was assayed by the formation of the covalent enzyme-GMP intermediate. Reaction mixtures (20 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM DTT, 5 mM $MgCl_2$, 1.25 μM [^{32}P]GTP, and capping enzyme and Tat proteins as specified were incubated for 15 min at 37 °C. The reaction was halted by addition of SDS to 1% final concentration. The reaction mixtures were analyzed by SDS-

PAGE. Capping enzyme-[^{32}P]GMP complexes were visualized by autoradiography and quantified by scanning the gel with a phosphorimager.

RNA triphosphatase assay. RNA triphosphatase activity was assayed by liberation of ^{32}Pi from [^{32}P]GTP-labeled poly(A) RNA synthesized by T7 RNA polymerase transcription. The template strand encoded the sequence for poly(A) RNA starting with G at position +1. *In vitro* transcription reactions were carried out as described previously (97) in the presence of [^{32}P]GTP. 5'-GTP-terminated 29-mer poly(A) was purified on a 15% polyacrylamide, 7 M urea denaturing gel. RNA triphosphatase reaction mixtures (10 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM DTT, 10 pmol of 5'-GTP-terminated poly(A), and capping enzyme and Tat proteins as specified were incubated for 15 min at 37 °C. The reaction was halted by addition of 1 μl of 88% formic acid. Aliquots of the mixtures were applied to a polyethyleneimine-cellulose TLC plate, which was developed with 0.75 M potassium phosphate (pH 4.3). The release of ^{32}Pi was quantified by scanning the TLC plate with a phosphorimager.

RNA capping assay. Triphosphate-terminated 17-mer RNA with no apparent secondary structure and TAR RNA containing a bulge and loop region (see Figure 7A) were prepared by *in vitro* T7 polymerase transcription and then purified by electrophoresis through a 20% polyacrylamide, 7 M urea gel as described previously (97). [^{32}P]GMP was incorporated at internal positions in the RNAs by including [^{32}P]GTP in the transcription reactions. Capping reaction mixtures (20 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM DTT, 50 μM GTP, 2.5

mM $MgCl_2$ 10 pmol of RNA, 20 units of RNase inhibitor (Promega), and capping enzyme and Tat proteins as specified were incubated for 15 min at 37 °C. The reaction was quenched by adding 200 µl of stop solution (0.3 M Tris-HCl, pH 7.5, 0.3 M sodium acetate, 0.5% SDS, 2 mM EDTA). The mixtures were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and then with chloroform. RNAs were recovered by ethanol precipitation and then analyzed by electrophoresis through a 15% polyacrylamide gel containing 7 M urea in Tris-Borate-EDTA. Labeled RNA products were visualized by autoradiography. The internally labeled capped RNA product migrated more slowly than the uncapped substrate RNA. The extent of capping was quantitated by scanning the gel with a phosphorimager.

RESULTS

Tat directly interacts with mammalian capping enzyme *in vitro*.

To test for interaction between Tat and mammalian capping enzyme, a purified GST·Tat fusion protein was linked to glutathione-Sepharose beads, and the beads were incubated with purified recombinant full-length Mce1. After washing the beads with buffer to remove unbound protein, the bound Mce1 was recovered from the beads by treatment with thrombin, which cleaved the GST·Tat fusion protein between the GST and Tat domains (Figure 1A). Released Mce1 was detected in the supernatant fraction by immunoblotting with antiserum raised against the C-terminal guanylyltransferase domain (Figure 1B, lane 3). Mce1 was not detected in the supernatant when the thrombin cleavage step was omitted (Figure 1B, lane 2). Alanine mutations of the active site cysteine of the RNA triphosphatase domain of Mce1 (C126A) or the active site lysine of the guanylyltransferase domain of Mce1 (K294A) that abrogate the triphosphatase and guanylyltransferase activities, respectively, did not interfere with the binding of Mce1 to immobilized Tat (Figure 1B, lanes *6* and 9). These results indicate that mammalian capping enzyme can interact directly with Tat *in vitro* independent of the competence of Mce1 to catalyze phosphoryl or nucleotidyl transfer.

The C-terminal segment of Tat containing the RNA binding domain suffices for binding to mammalian capping enzyme.

Tat protein can be divided into two major functional domains (Figure 2A). The transactivation domain (amino acids 1-48) is required for recruitment of cyclin T1 by Tat to the HIV-1 long terminal repeat (LTR) promoter (252). The C-terminal domain (amino acids 49-86) includes a basic region and is required for both RNA binding and nuclear localization of Tat (52). Two truncated versions of Tat, Tat2/36 and Tat48, which are deleted in the transactivation domain and RNA binding domain, respectively, were expressed as GST fusion proteins and tested for binding to the guanylyltransferase domain of mammalian capping enzyme, Mce1-211-597). The guanylyltransferase bound to beads containing immobilized wild-type Tat and Tat2/36 (Figure 2B, lanes 3 and 5), but not to Tat48 (Figure 2B, lane7). Similar results were obtained for binding of full-length Mce1 to the truncated Tat proteins (data not shown). We conclude that the Tat segment from amino acids 37-86 suffices for the binding of capping enzyme and that the transactivation domain *per se* does not interact with capping enzyme or its guanylyltransferase component.

Tat binding stimulates the activity of mammalian guanylyltransferase.

Are there functional consequences for the interaction of mammalian capping enzyme with Tat? To address this question, we tested the effects of full-length Tat and truncated Tat derivatives on the guanylyltransferase and triphosphatase activities of Mce1. The 597-amino acid mammalian capping enzyme consists of an N-terminal triphosphatase domain (amino acids 1-210)

and a C-terminal guanylyltransferase domain (amino acids 211-597). The guanylyltransferase component of the enzyme catalyzes two sequential nucleotidyl transfer reactions involving a covalent enzyme-guanylate intermediate (212). In the first partial reaction, nucleophilic attack on the phosphate of GTP by enzyme results in liberation of pyrophosphate and formation of a covalent adduct in which GMP is linked via a phosphoamide bond to the -amino group of a Lys-294 (94, 269). The nucleotide is then transferred to the 5'-end of the RNA acceptor to form an inverted (5')-(5') triphosphate bridge structure, GpppN.

The extent of Mce1·[^{32}P]GMP complex formation during reaction with 1.25 μM GTP was proportional to input protein up to 0.15 μM Mce1 and leveled off as Mce1 was increased to 0.5 μM (Figure 3A). In the linear range of Mce1-dependence, ~9% of the input protein molecules were labeled with [^{32}P]GMP. Enzyme-[32P]GMP complex formation by 0.05 μM Mce1 was stimulated by Tat and Tat2/36, but Tat48 had no effect (Figure 3B). Optimal stimulation (4-5-fold) was attained at a 2:1 molar ratio of Tat to Mce1 (Figure 3C). The activity of the autonomous guanylyltransferase domain Mce1-(211-597) was similarly stimulated by Tat and Tat2/36, but not Tat48 (data not shown). Tat had no stimulatory effect on enzyme-GMP formation by purified recombinant yeast guanylyltransferase Ceg1 (data not shown). These results indicate that Tat interaction with Mce1 stimulates the guanylyltransferase activity of mammalian capping enzyme and the Tat domain from amino acids 37-86 suffices for this function.

Tat stimulates the RNA triphosphatase activity of mammalian capping enzyme.

The RNA triphosphatase domain of mammalian capping enzyme displays extensive amino acid sequence similarity to protein tyrosine phosphatases and dual-specificity protein phosphatases. By analogy to the protein phosphatases, it is proposed that mammalian RNA triphosphatase executes a two-step phosphoryl transfer reaction involving a covalent enzyme-(cysteinyl-*S*)-phosphate intermediate (94, 228, 253). In the first partial reaction, nucleophilic attack by a cysteine thiolate (Cys-126 in Mce1) on the -phosphate of RNA results in release of diphosphate-terminated RNA and formation of a phosphoenzyme. The phosphate is then transferred from Cys-126 to water to release Pi. RNA triphosphatase activity is assayed by the release of ^{32}Pi from ^{32}P-labeled triphosphate-terminated RNA.

Hydrolysis of 1 μM [^{32}P]GTP-labeled poly(A) by Mce1 increased linearly from 2-8 nM enzyme and was nearly quantitative at 10 nM (Figure 4A). The titration profile of the isolated triphosphatase domain Mce1-(1-210) was more sigmoidal, but the slope of the curve in the linear range (4-12 nM enzyme) was similar to that of Mce1 (Figure 4A). RNA -phosphate hydrolysis by 3 nM Mce1 was stimulated 6-8-fold by the inclusion of 6-18 nM Tat (Figure 4B). Note that Tat by itself had no detectable RNA triphosphatase activity at the highest level of input Tat used in this experiment (data not shown). The remarkable finding was

that Tat did not stimulate RNA -phosphate hydrolysis by 5 nM of the N-terminal RNA triphosphatase domain Mce1(1-210), which catalyzed a similar level of basal RNA hydrolysis as 3 nM Mce1 (Figure 4B).

To investigate whether Tat interacts directly with the RNA triphosphatase domain, we established an affinity chromatography assay using His-tagged capping enzyme domains immobilized on Ni2+-agarose beads (Figure 5A). The beads were incubated with HA-tagged Tat. After washing the beads with buffer to remove unbound protein, the bead-bound material was stripped from the beads with SDS, and the presence of HA·Tat in the SDS-eluate was detected by immunoblotting with anti-HA antibody. As shown in Figure 5B, HA·Tat was adsorbed to beads containing the immobilized RNA triphosphatase domain Mce1-(1-210) (lane 3), whereas only a scant amount of HA·Tat was retained on the control Ni2+-agarose beads (lane 1). HA·Tat also bound to beads containing immobilized guanylyltransferase domain Mce1-(211-597) (lane 2). This experiment reciprocates the finding that soluble guanylyltransferase domain was bound to immobilized Tat (Figure 2).

The results presented thus far demonstrate that Tat interacts with the isolated RNA triphosphatase and guanylyltransferase domains of mammalian capping enzyme. Whereas Tat-binding stimulates the guanylyltransferase activity of Mce1 and the C-terminal domain, Tat-stimulation of the RNA triphosphatase reaction appears to occurs only in the context of the full-length Mce1.

Tat stimulates RNA cap formation by Mce1.

The enzymatic addition of an unlabeled cap guanylate to the 5'-end of an internally labeled RNA molecule results in a characteristic slowing of the electrophoretic mobility of the RNA, equivalent to about a 2-nucleotide increase in apparent chain length (51). The change in mobility upon addition of capping enzyme has been used to detect cap formation on RNAs as long as 78 nucleotides (50). Here we studied the complete capping reaction of Mce1 using a synthetic 17-mer triphosphate-terminated RNA substrate that was labeled internally with [^{32}P]GMP (Figure 6A). Incubation of 0.5 μM RNA substrate with Mce1 in the presence of 50 μM cold GTP and magnesium chloride resulted in transfer of GMP to the 5'-end, yielding a capped species that migrated more slowly than the input substrate RNA (Figure 6B). Formation of the capped species was dependent on inclusion of GTP in the reaction mixture (not shown). The extent of capping was proportional to input Mce1 and saturated at 0.4 μM Mce1 with about 70% of the input RNA being capped (Figure 6C). In the linear range of enzyme-dependence, 1 pmol of RNA was capped per pmol of the Mce1.

The level of capping catalyzed by 50 nM Mce1 (1 pmol of input Mce1) was increased 6-fold by Tat (Figure 6D). The stimulation was Tat concentration-dependent and saturation was attained at a 4:1 molar ratio of Tat to Mce1. An identical capping stimulation profile was observed for the deletion mutant Tat2/36 containing an intact RNA binding domain. Tat48 containing the activation domain had no salutary effect on RNA capping (Figure 6D). The ability of Tat or its

component domains to stimulate cap formation by Mce1 correlated perfectly with the capacity to bind to Mce1 (Figure 3).

Tat enhances TAR RNA capping.

Tat functions through TAR RNA to control an early step in transcription elongation that depends on the Pol II CTD (43, 108). In light of our findings that Tat stimulates mammalian capping enzyme, we envisioned that Tat might enhance the capping of its target TAR mRNA, which might otherwise be inefficiently capped because the 5'-end is encompassed within a stable RNA duplex (67). To test this hypothesis, we prepared internally labeled Non-TAR RNA (17-mer) and TAR RNA (29-mer) substrates (Figure 7A) and tested them for capping in the same reaction mixtures with limiting Mce1. Capping reactions containing 75 nM Mce1 and 250 nM each of the Non-TAR and TAR substrates were supplemented with increasing amounts of Tat protein and the labeled products were resolved by PAGE (Fig. 7*B*). In a competitive situation in the absence of Tat (Figure 7B, lane 2), Mce1 favored the Non-TAR substrate (1 pmol of cap formed) over the TAR substrate (0.3 pmol of cap formed) (Figure 7C). Tat enhanced capping of both Non-TAR and TAR RNA substrates at a 2:1 ratio of Tat to Mce1 (Figure 7C). Yet, the relative Tat stimulation of capping of the TAR RNA (5-fold) was greater than that of the Non-TAR substrate (2-fold).

DISCUSSION

Selective targeting of caps to Pol II transcripts *in vivo* is achieved, at least in part, through direct physical interaction of the capping apparatus with the phosphorylated CTD of Pol II. In addition to recruiting capping enzyme to the Pol II elongation complex, the phosphorylated CTD stimulates the guanylyltransferase activity of the mammalian capping enzyme (93). This simple and appealing model for targeting and regulation of capping is belied by the underlying complexity of CTD phosphorylation (and dephosphorylation) and by mounting evidence that regulation of transcription elongation is a key facet of cotranscriptional mRNA processing. The data presented here illuminate a new pathway of capping enzyme recruitment and activation by the HIV Tat protein. Our findings contribute to an emerging picture of how elongation and processing are coupled, especially during HIV gene expression.

In mammalian cells, the timely acquisition of the cap may promote subsequent mRNA-specific processing steps (splicing and polyadenylation) and protect the nascent mRNA from exonucleolytic decay. A clear advantage would accrue from a mechanism whereby capping is restricted to Pol II complexes that are committed to productive elongation, insofar as the capping of short transcripts that are subsequently aborted and released would generate a population of non-coding capped RNAs that could compete with *bona fide* mRNAs in cap-dependent transactions. CTD hyperphosphorylation has often been correlated with the establishment of a stable elongation complex, but there

is little information as to the exact nature of the phosphorylation array at any point in the transcription cycle and there is still uncertainty concerning the relative contributions of different CTD kinases (TFIIH and P-TEFb) to the establishment and remodeling of the CTD phosphorylation array. The timing during early elongation of the critical CTD phosphorylation steps that permit capping enzyme recruitment are not well defined and may even vary for different transcription units. Although it is clear that short nascent transcripts (on the order of 30 nucleotides) *can* be capped, such results reflect the action of the capping enzymes on arrested polymerase elongation complexes, where there is no kinetic competition between ongoing elongation and capping. The RNA size threshold for capping in this experimental setting simply reflects the steric constraints on capping of RNA chains held within the RNA binding pocket of the polymerase. The unimpeded rate of elongation by RNA polymerase (~20 nucleotides/s) is faster than estimates of the rates of the RNA triphosphatase andguanylyltransferase reactions (reviewed in ref. (210)), which sets up a situation in which RNA polymerase might "outrun" the capping enzyme. It is not known whether capping enzyme has a narrow or wide window for action on nascent 5' ends *in vivo*, *i.e.* whether capping enzyme would dissociate from the elongation complex after polymerase has proceeded a certain distance down the transcription unit.

Thus, there is advantage in imposing an elongation checkpoint to maximize the opportunity for the capping apparatus to bind the elongation

complex. This scenario has been studied most thoroughly for HIV transcription, where the negative elongation factor DSIF induces an elongation block that is overcome by P-TEFb, a cyclin-dependent protein kinase that phosphorylates the CTD (181) (Figure 8). One key function of the Tat-TAR RNA complex is to recruit P-TEFb to the nascent HIV mRNA through an interaction of the cyclin T1 subunit of P-TEFb with the activation domain of Tat. The CTD kinase function of P-TEFb is essential for Tat-TAR stimulation of HIV transcription, whereas the CTD kinase activity of TFIIH is apparently not critical for overriding the effects of DSIF (241, 275). Indeed, although TFIIH and P-TEFb are both associated with very early HIV elongation complexes halted at position +14, TFIIH dissociates by the time the polymerase moves to position +30 or +36 (prior to elaboration of the TAR site in the nascent RNA), whereas P-TEFb remains associated with the elongation complex at least up to position +79 when TAR is formed (179, 275). The presence of Tat triggers a new round of P-TEFb-catalyzed CTD phosphorylation on elongation complexes at position +79 (275).

An exciting connection between Tat, P-TEFb, and cap formation emerges from the recent report that Tat alters the phosphorylation site-specificity of P-TEFb in the context of the HIV transcription complex. P-TEFb phosphorylates Ser-2 of the CTD heptad in the absence of Tat, but it phosphorylates both Ser-2 and Ser-5 when Tat is present (275). CTD phosphorylation on either Ser-2 or Ser-5 suffices to bind mammalian capping enzyme, but CTD-stimulation of capping activity is specific for the Ser-5-PO4 CTD array (93). Thus, the

Tat/TAR/P-TEFb complex helps craft a CTD array that both recruits and activates Mce1.

Here we have shown that Tat spearheads a second and novel pathway of capping enzyme recruitment and activation via a direct physical interaction between Tat and Mce1. Unlike, the Tat-P-TEFb interaction, which requires the N-terminal transactivation domain of Tat, the binding of Tat to capping enzyme is via the C-terminal domain that includes the TAR RNA binding site. The Tat-Mce1 interaction results in a significant stimulation of the guanylyltransferase and triphosphatase activities of Mce1 and it thereby enhances the efficiency of cap formation on defined RNA substrates. Tat is unique among the recently discovered regulators of mammalian capping enzyme because as it is the only factor that up-regulates the triphosphatase component of the bifunctional enzyme. The effects of Ser-5-PO4 CTD and hSpt5 on Mce1 are limited to stimulation of the guanylyltransferase (93, 253). Tat-stimulation of Mce1 triphosphatase activity occurs in the context of the native two-domain enzyme, but not with the isolated N-terminal triphosphatase component, even though Tat can bind to the isolated triphosphatase domain. Thus, it appears that contacts of Tat to both domains of Mce1 are needed to elicit the increase in triphosphatase activity. In contrast, Tat is equally capable of stimulating theguanylyltransferase activity of the full-length Mce1 and its isolated C-terminal domain. Ser-5-PO4 CTD can also up-regulate either Mce1 or the guanylyltransferase domain, but hSpt5 (which is a subunit of DSIF) only affects the activity of full-length Mce1.

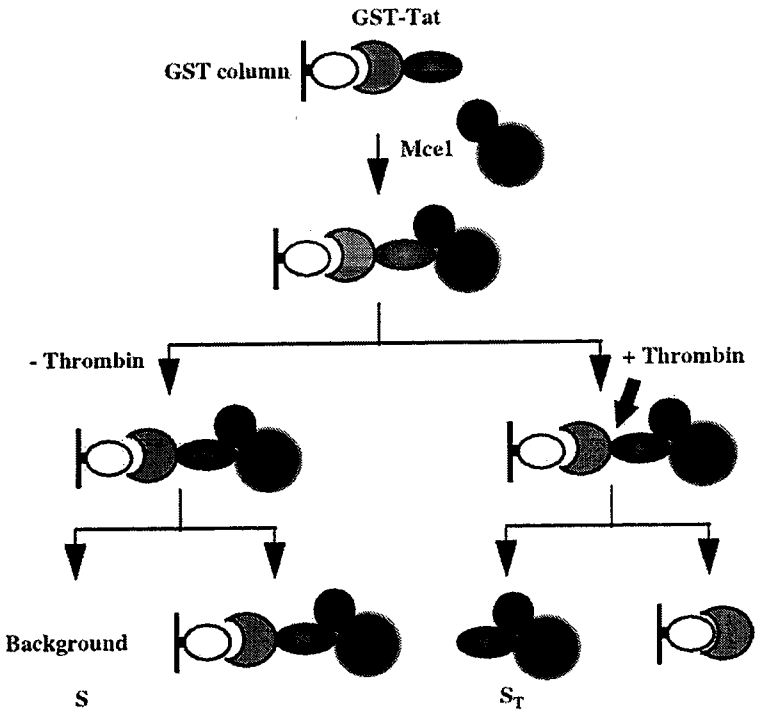
There are now 3 possible pathways of up-regulating HIV mRNA capping involving factors that are known to be associated with the HIV transcription complex: (i) direct Tat activation of Mce1; (ii) CTD-PO4 activation of Mce1, via Tat-stimulation of the CTD kinase activity of P-TEFb; (iii) DSIF activation of Mce1 via its hSpt5 subunit (Figure 8). There is potential for cross-talk between these pathways, insofar as P-TEFb phosphorylation of CTD clearly affects DSIF function and P-TEFb is also capable of phosphorylating the hSpt5 subunit of DSIF (103). Phosphorylation of hSpt5 by P-TEFb has been reported to block its ability to elicit a transcription elongation block (103). The effects of hSpt5 phosphorylation on its interaction with the capping enzyme are unclear. The attractive feature of the direct Tat-Mce1 activation pathway in HIV gene expression is that it is independent of CTD phosphorylation status and provides a means for direct and specific recruitment of Mce1 to the nascent HIV transcript containing bound Tat.

Why might HIV go to such lengths to employ Tat to recruit and activate Mce1? The formation of stable RNA secondary structures in which the 5' end of the HIV transcript is encompassed within a duplex stem may limit access of the 5' terminus to the active sites of the capping enzyme (67). Such a stem structure is formed by nascent HIV mRNA even prior to the synthesis of the TAR sequence and, although the secondary structure changes after TAR is formed, the 5' end remains held within a duplex stem (170). We observed a structured TAR substrate was capped less effectively by purified Mce1 than an RNA with no

apparent secondary structure. Although Tat stimulated the capping of both Non-TAR and TAR RNAs, the-fold-stimulation of the structured TAR substrate was greater than that of the unstructured transcript. Thus regulation of capping by Tat would have the most impact where capping is inherently weak. Finally, a Tat-dependent enhancement of mRNA cap formation may account for the finding that Tat stimulates the translation of mRNAs synthesized from the HIV transcription unit (44).

Figure 1. Tat interacts directly with Mce1 *in vitro.* A, experimental design to test the interaction between Mce1 and Tat protein. Purified GST·Tat fusion protein was linked to glutathione-Sepharose beads and incubated with purified Mce1. After loading to the column and washing with buffer (30mM Tris-HCl, pH 8.0, 150mM NaCl, 2mM $CaCl_2$, 5mM DTT), the sample was divided into two aliquots. Bound Mce1 was released by thrombin treatment (supernatant phase, *ST*). Supernatant without thrombin treatment (*S*) shows the background. B, full-length Mce1 (lane 1) and two active site mutants, Mce1(C126A) (lane 4) and Mce1(K294A) (lane 7), were assayed for binding to GST·Tat. The input and supernatant fractions were analyzed by SDS-PAGE; the wild-type and mutant Mce1 polypeptides were detected by immunoblotting.

A


GST-Tat
GST column
Mce1
- Thrombin
+ Thrombin
Background
S
S_T

B

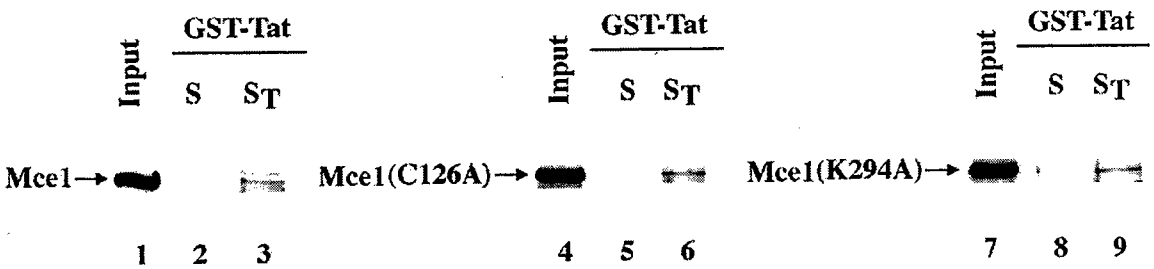
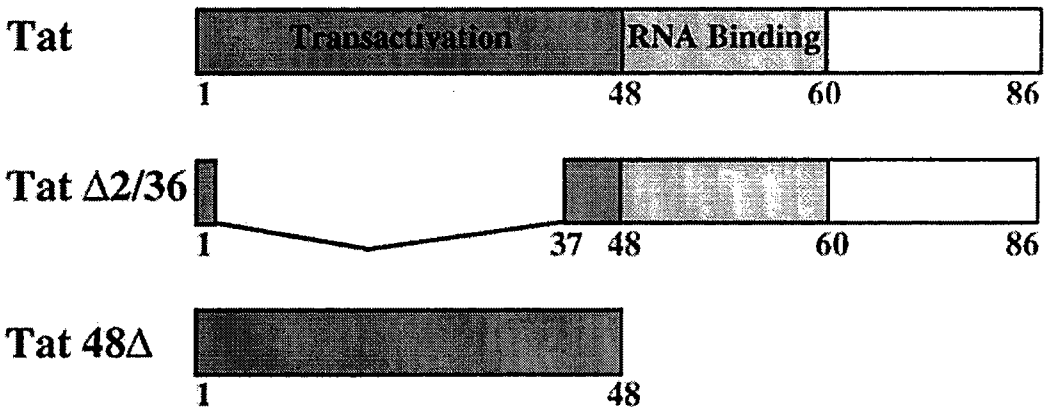

Input
GST-Tat
S
S_T
Mce1→
Mce1(C126A)→
Mce1(K294A)→
1 2 3
4 5 6
7 8 9

Figure 2. The C-terminal segment of Tat containing the RNA binding domain is sufficient for interaction with Mce1. A, schematic representation of the domains of HIV-1 Tat protein and the Tat deletion mutants that were used in this binding assay. *B*, purified GST-Tat, GST-Tat2/36 (transactivation domain deletion) and GST-Tat48 (RNA binding domain deletion) were incubated with purified Mce1-(211-597). Binding was assayed as shown in Figure 1. An immunoblot of the input and supernatant fractions is shown.

A


Tat
Transactivation
RNA Binding
1
48
60
86
Tat Δ2/36
1
37
48
60
86
Tat 48Δ
1
48

B

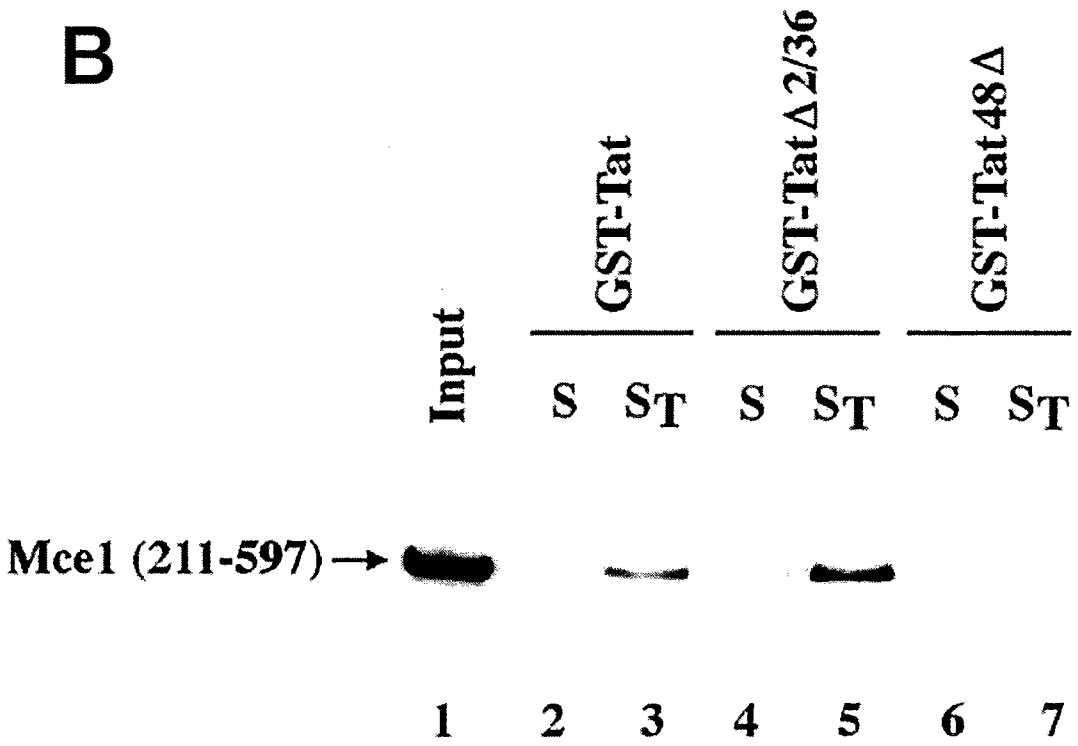

Input
GST-Tat
GST-Tat Δ2/36
GST-Tat 48Δ
S
ST
S
ST
S
ST
Mce1 (211-597) →
1
2
3
4
5
6
7

Figure 3. Tat enhances guanylylation of Mce1. A, guanylyltransferase activity of purified Mce1 was assayed as described under "Experimental Procedures". Mce1·[^{32}P]GMP complex formation is plotted as a function of input Mce1. B, effect of Tat. Mce1 (1 pmol) was assayed for guanylyltransferase activity in the presence of increasing amounts (0, 1, 2, 4, and 6 pmol) of wild-type Tat (lanes 1-5), Tat2/36 (lanes 6-10), and Tat48 (lanes 11-15). The reaction products were analyzed by SDS-PAGE, and Mce1·[^{32}P]GMP complexes were detected by autoradiography. C, Mce1·[^{32}P]GMP complex formation is plotted as a function of the Tat/Mce1 ratio.

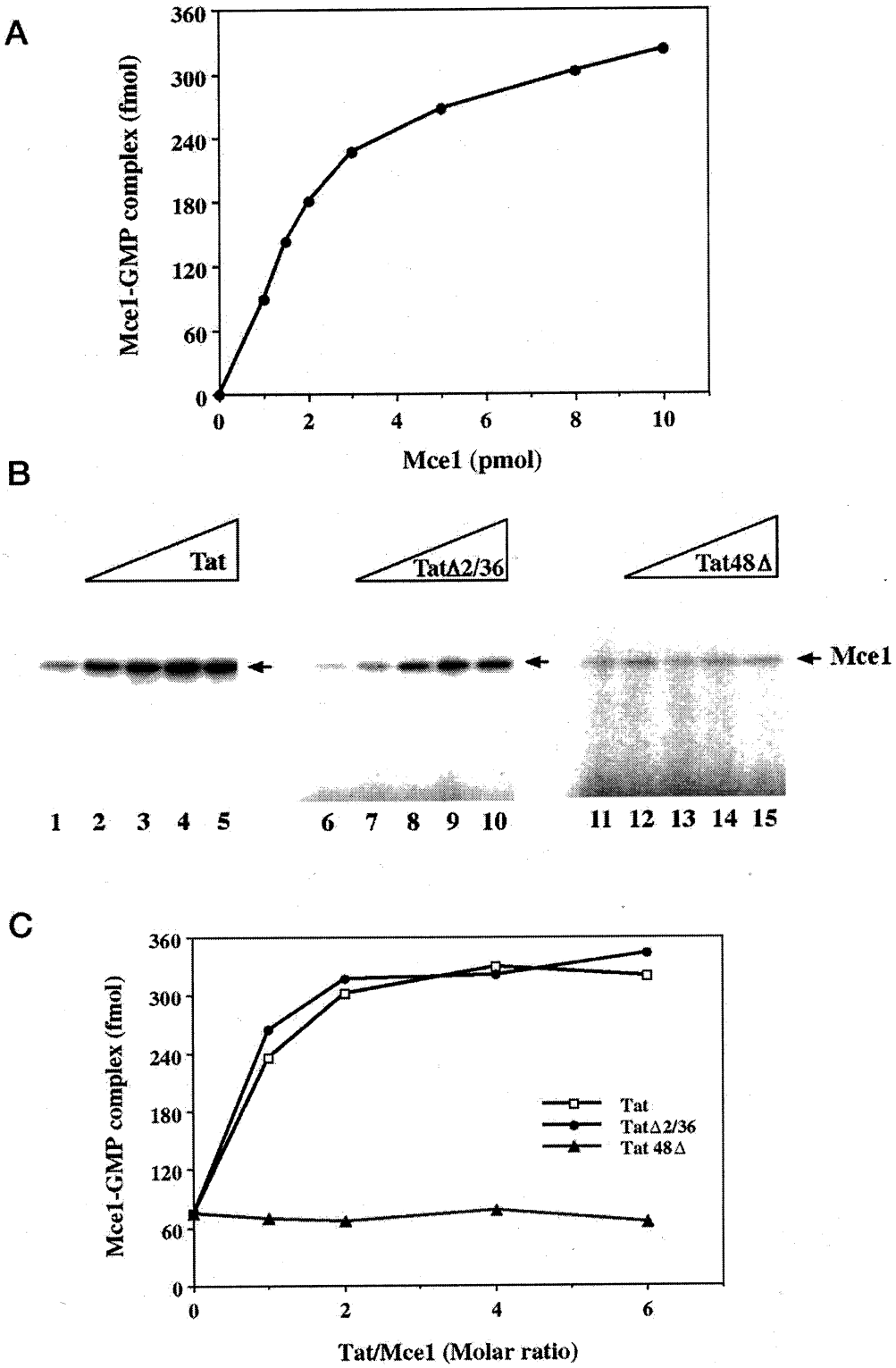
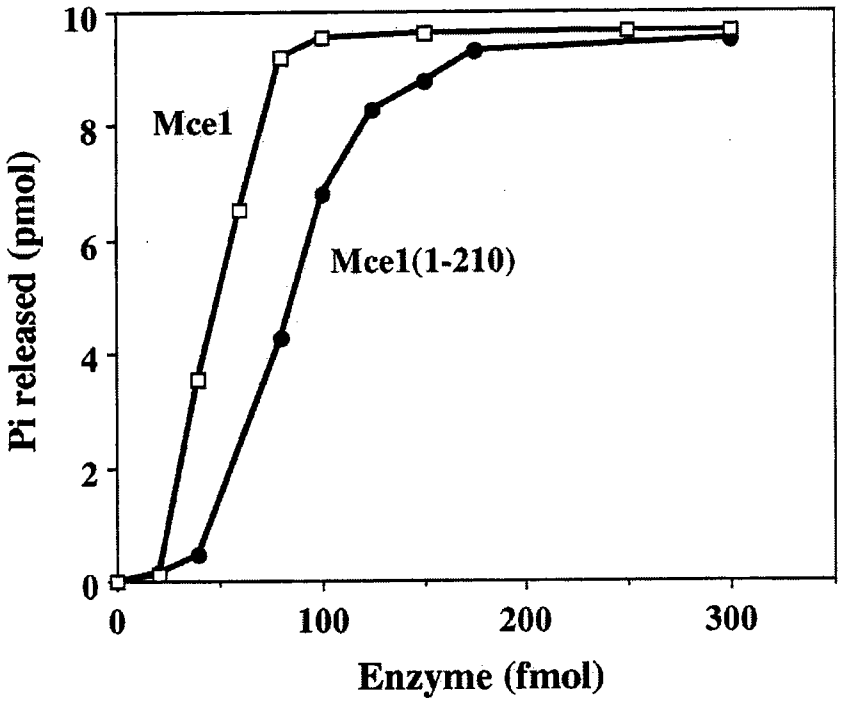

A
Mce1-GMP complex (fmol)
360
300
240
180
120
60
0
0
2
4
6
8
10
Mce1 (pmol)
B
Tat
TatΔ2/36
Tat48Δ
Mce1
1 2 3 4 5
6 7 8 9 10
11 12 13 14 15
C
Mce1-GMP complex (fmol)
360
300
240
180
120
60
0
0
2
4
6
Tat
TatΔ2/36
Tat 48Δ
Tat/Mce1 (Molar ratio)

Figure 4. Tat enhances the RNA triphosphatase activity of full-length Mce1 but not the isolated triphosphatase domain. A, RNA triphosphatase activity was assayed as described under "Experimental Procedures." ^{32}Pi release is plotted as function of input Mce1 or Mce1-(1-210). B, effect of Tat. Triphosphatase reaction mixtures contained 30 fmol of Mce1 or 50 fmol of Mce1-(1-210) plus increasing concentrations of Tat. The extent of ^{32}Pi release from [^{32}P]GTP-labeled RNA is plotted as a function of the molar ratio of Tat to capping enzyme.

A


Mce1
Mce1(1-210)
Pi released (pmol)
Enzyme (fmol)
0
2
4
6
8
10
0
100
200
300

B

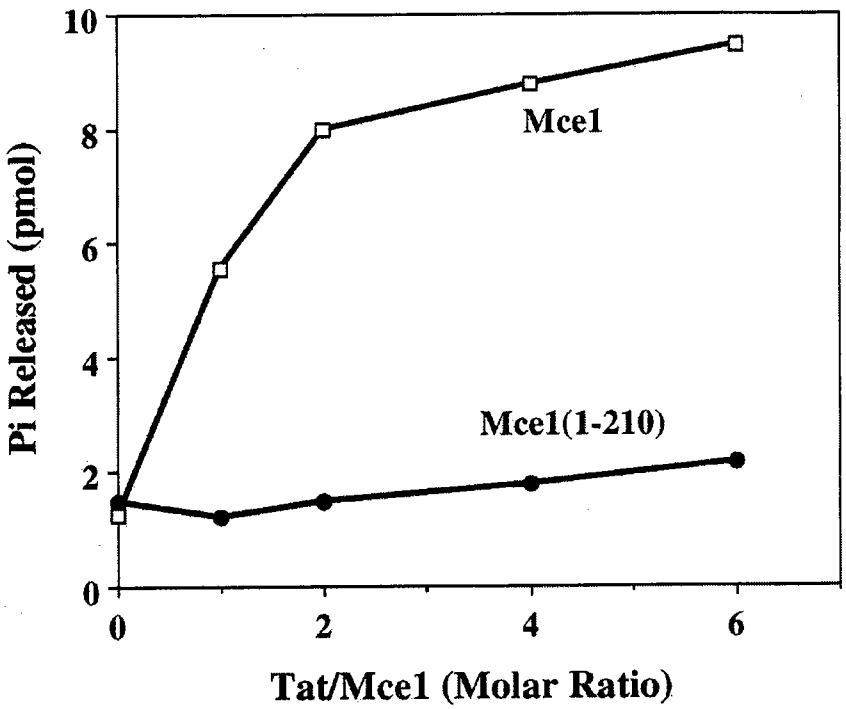

Mce1
Mce1(1-210)
Pi Released (pmol)
Tat/Mce1 (Molar Ratio)
0
2
4
6
8
10
0
2
4
6

Figure 5. The RNA triphosphatase domain of mammalian capping enzyme can bind to Tat. A, HA·Tat was reacted with Ni2+-agarose beads alone (1), with beads containing His-tagged Mce1-(211-597) (2), and beads containing His-tagged Mce1-(1-210) (3). The bound material was eluted with SDS and resolved by SDS-PAGE. *B,* Tat protein in the input sample and the SDS eluates was detected by immunoblotting using antibody directed against the HA tag.

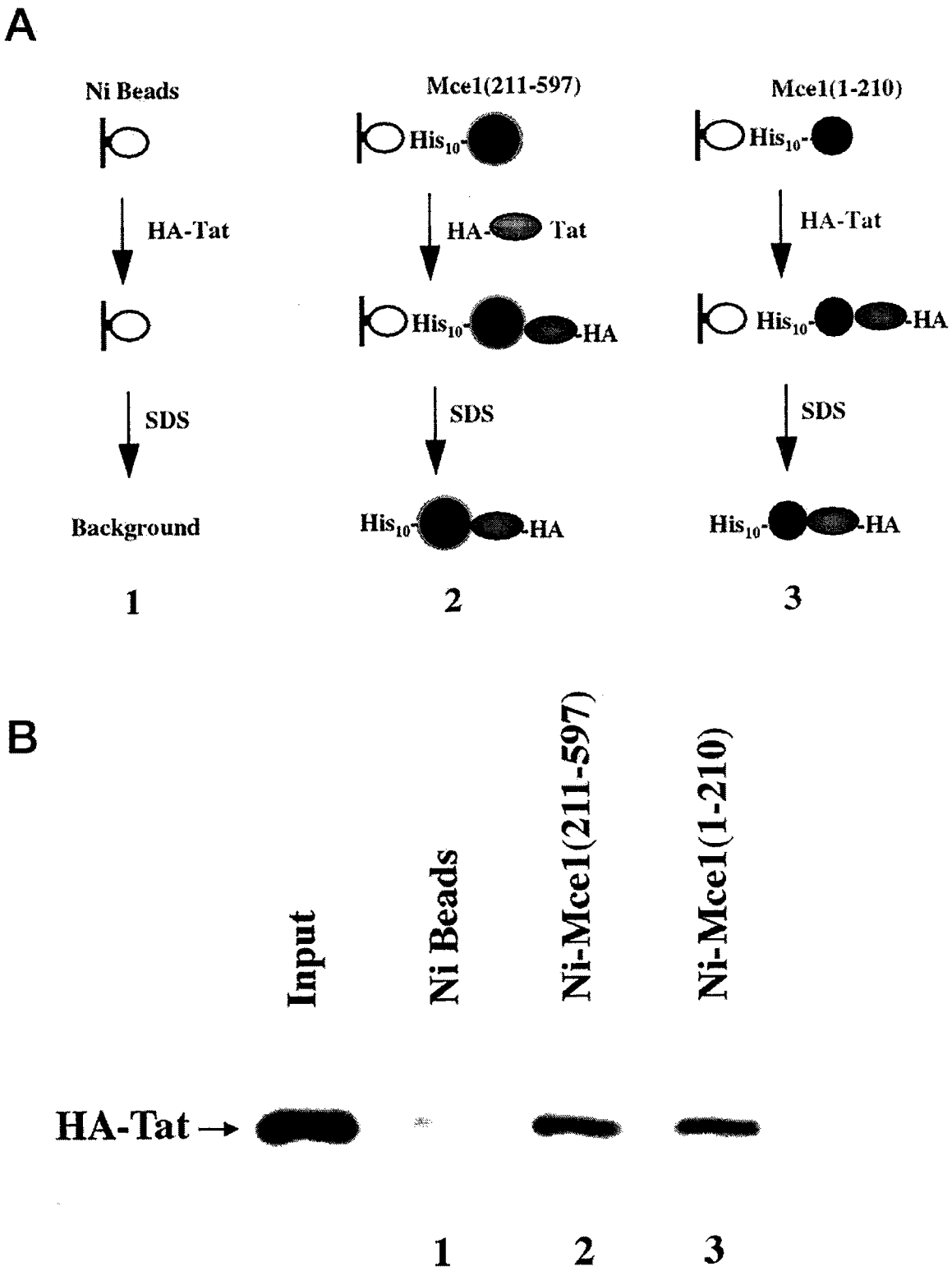
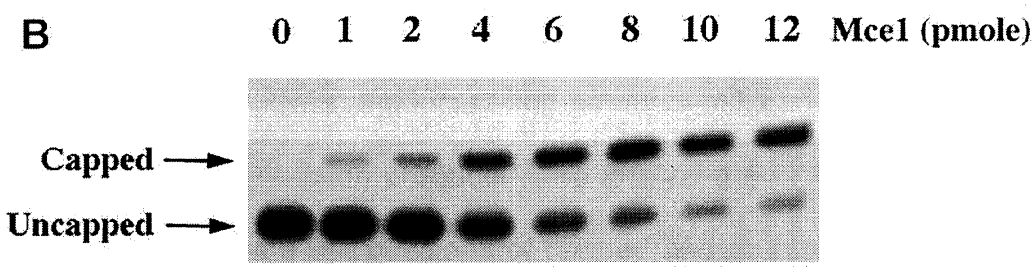

A
Ni Beads
Mce1(211-597)
Mce1(1-210)
HA-Tat
HA-Tat
HA-Tat
SDS
SDS
SDS
Background
1
2
3
B
Input
Ni Beads
Ni-Mce1(211-597)
Ni-Mce1(1-210)
HA-Tat
1
2
3

Figure 6. Tat enhances RNA cap formation. A, sequence of the 17-mer substrate RNA used in capping reactions. B, capping reactions contained 10 pmol of internally labeled 17-mer RNA and Mce1 as specified. The radiolabeled reaction products were analyzed by PAGE. An autoradiogram of the gel is shown. The positions of capped and uncapped RNAs are indicated on the left. C, the extent of cap formation in (B) is plotted as a function of input enzyme. D, effect of Tat on RNA capping. In these reactions, we used 10 pmol of RNA, 1 pmol of Mce1, and increasing concentration of various Tat proteins. Quantitative analysis of Mce1 capping activity in the presence of various Tat sequences.

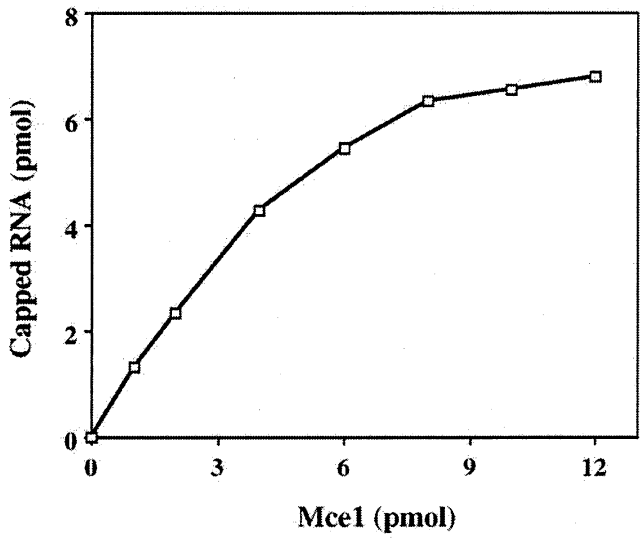
A

pppGCUGGCUCUCUGGCGCC

B



C



D

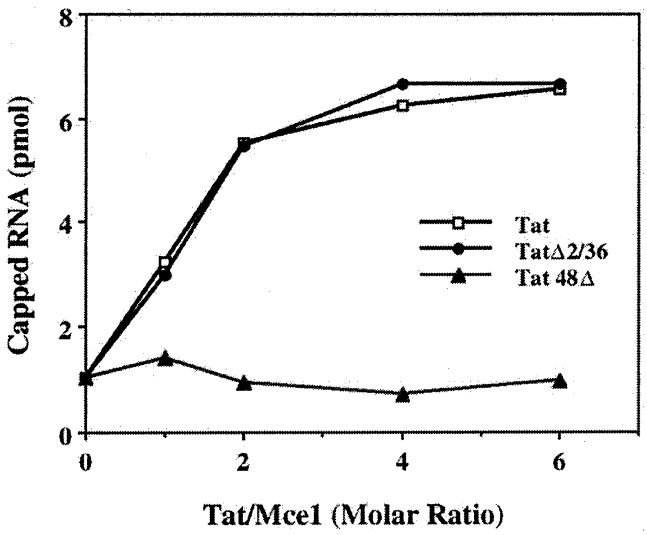


Figure 7. Tat enhances capping of TAR RNA. A, sequence and secondary structure of TAR RNA (29-mer) used in the RNA cap formation assay. B, capping reaction mixtures (20 μl) contained 50 μM GTP, 75 nM Mce1, 250 nM Non-TAR and 250 nM TAR RNAs and increasing amounts of Tat (0, 1.5, 3, 6, and 9 pmol). The radiolabeled reaction products were analyzed by PAGE. An autoradiogram of the gel is shown. The positions of capped and uncapped TAR and Non-TAR RNAs are indicated by *arrows. C,* the extent of cap formation is plotted as a function of the Tat/Mce1 molar ratio for Non-TAR RNA (right y axis) and TAR RNA (left y axis).

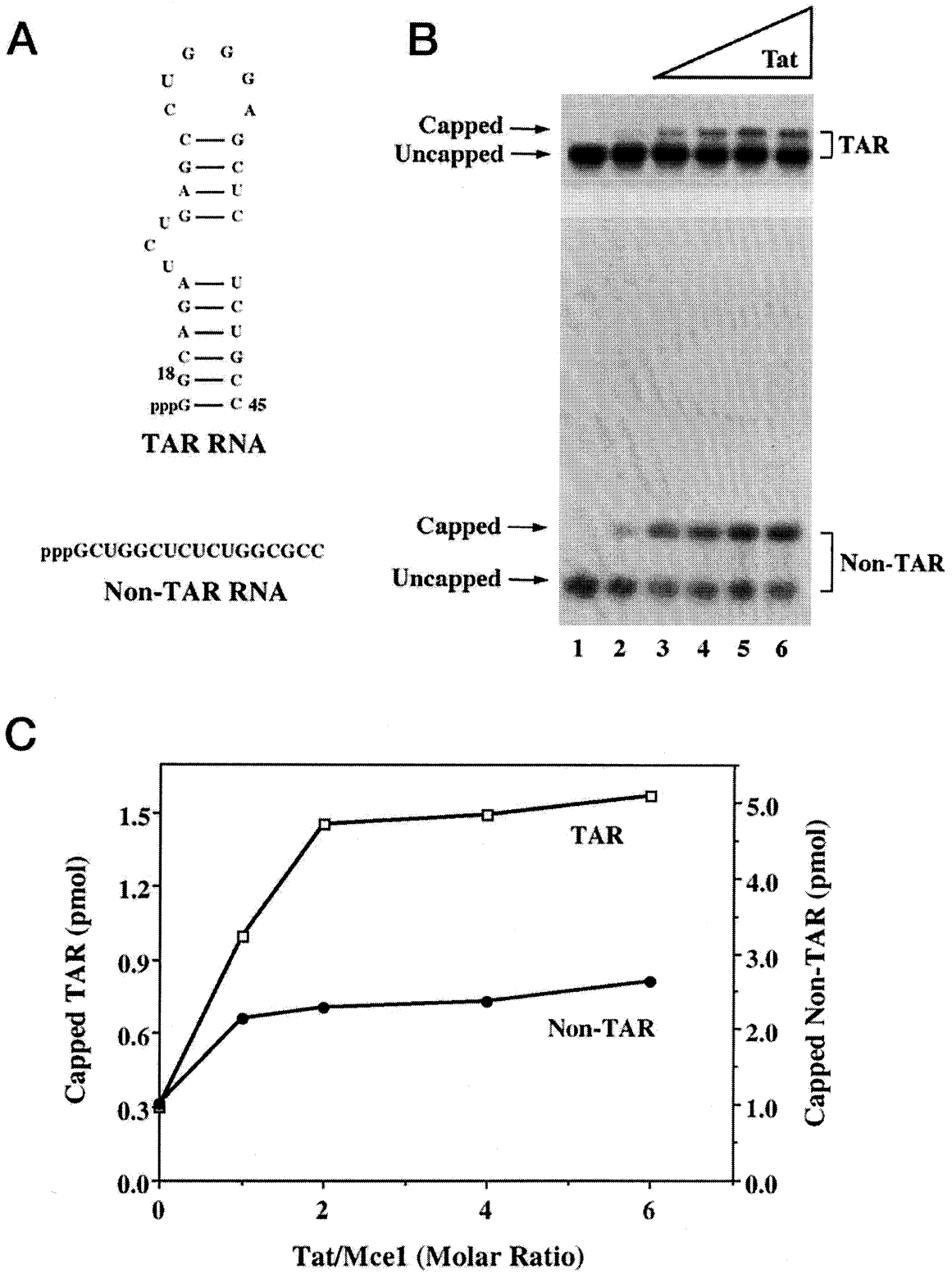
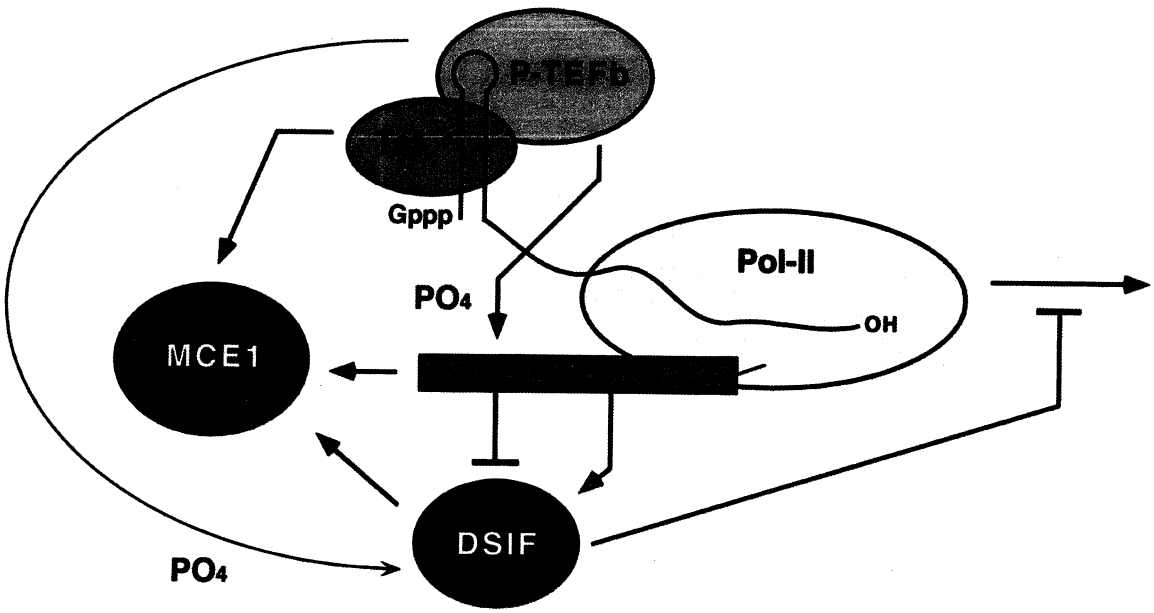

A
TAR RNA
pppGCUGGCUCUCUGGCGCC
Non-TAR RNA
B
Tat
Capped
Uncapped
TAR
Capped
Uncapped
Non-TAR
1 2 3 4 5 6
C
Capped TAR (pmol)
Capped Non-TAR (pmol)
TAR
Non-TAR
Tat/Mce1 (Molar Ratio)

Figure 8. Emerging connections between CTD phosphorylation, capping, and transcription elongation. See text for details.


P-TEFb
Gppp
PO4
Pol-II
OH
MCE1
DSIF
PO4

CHAPTER III

TAT STIMULATES COTRANSCRIPTIONAL CAPPING OF HIV MESSENGER RNA

ABSTRACT

Here we investigated how capping and methylation of HIV pre-mRNAs are coupled to Pol II elongation. Stable binding of the capping enzyme (Mce1) and cap methyltransferase (Hcm1) to template-engaged Pol II depends on CTD phosphorylation, but not on nascent RNA. Both Mce1 and Hcm1 travel with Pol II during elongation. The capping and methylation reactions cannot occur until the nascent pre-mRNA has attained a chain length of 19–22 nucleotides. HIV pre-mRNAs are capped quantitatively when elongation complexes are halted at promoter-proximal positions, but capping is much less efficient during unimpeded Pol II elongation. Cotranscriptional capping of HIV mRNA is strongly stimulated by Tat, and this stimulation requires the C-terminal segment of Tat that mediates its direct binding to Mce1. Our findings implicate capping in an elongation checkpoint critical to HIV gene expression.

INTRODUCTION

mRNA capping occurs cotranscriptionally by a series of three enzymatic reactions in which the 5' triphosphate terminus of the pre-mRNA is cleaved to a diphosphate by RNA triphosphatase, then capped with GMP by RNA guanylyltransferase, and methylated by RNA (guanine-N7) methyltransferase. Targeting of cap formation to cellular transcripts made by RNA polymerase II (Pol II) is achieved through physical interactions of the capping enzymes with the phosphorylated carboxyl-terminal domain (CTD) of the largest subunit of Pol II (31, 152, 173, 269). The CTD, which is composed of a tandemly repeated heptad motif YSPTSPS, undergoes extensive serine phosphorylation and dephosphorylation during the transcription cycle. Recruitment of the capping apparatus to the yeast Pol II elongation complex *in vivo* requires the action of the TFIIH-associated CTD kinase Kin28 (equivalent to Cdk7 in mammals), which phosphorylates Ser5 of the CTD (128, 191, 199). Other factors may also be involved in coupling capping to Pol II transcription. For example, the Pol II transcription elongation factor Spt5 interacts directly with the triphosphatase and guanylyltransferase components of the capping apparatus in mammals and in the fission yeast *Schizosaccharomyces pombe* (174, 253). Many eukaryotic viruses that replicate in the nucleus exploit the host's enzymes to transcribe and cap their mRNAs. Human immunodeficiency virus (HIV) is of particular interest because it appears to impose an additional mechanism to target the capping enzyme to the viral transcription unit, whereby the HIV-encoded Tat protein binds

to the triphosphatase and guanylyltransferase domains of the mammalian capping enzyme and upregulates both catalytic activities (27).

The Pol II CTD, Spt5, and Tat are intimately connected to the regulation of HIV gene expression (reviewed in 181). Human Spt5 and its binding partner hSpt4 comprise the transcription elongation regulatory factor DSIF (DRB sensitivity-inducing factor). DSIF binds to Pol II and, in concert with NELF (negative elongation factor), represses elongation at promoter-proximal positions in the transcription unit. Escape from the repressive effect of DSIF/NELF requires the action of P-TEFb (positive transcription elongation factor b), a DRB-sensitive protein kinase that phosphorylates both the Pol II CTD and the Spt5 subunit of DSIF. P-TEFb is composed of two subunits, Cdk9 and cyclin T1, and it binds to the HIV Tat protein via the cyclin T1 subunit (181, 252). In previous studies we have analyzed the interactions of these regulatory factors with Pol II at discrete functional stages of transcription from the HIV-1 LTR promoter in vitro. We found that P-TEFb is a component of the Pol II preinitiation complex (PIC) and that it travels with the transcription elongation complex (TEC) as in moves along the HIV transcription unit (179). In contrast, DSIF and NELF are not present in the PIC, but associate with the TEC at promoter proximal positions and then travel with the TECs down the template (178).

How does capping fit into the scheme of HIV transcription regulation? The mammalian capping apparatus consists of two components: a bifunctional triphosphatase-guanylyltransferase (Mce1) and a separate cap

methyltransferase (Hcm1) (94, 177, 195, 269). Our finding that Tat enhances the efficiency of capping of isolated HIV mRNA by Mce1 in vitro led to the suggestion that Spt5-induced arrest at promoter-proximal sites in the HIV transcription unit might ensure a temporal window for recruitment of the capping enzymes and stimulate modification of the 5' end of the HIV pre-mRNA (27).

At present, little is known about the dynamics of the interactions of mammalian capping and methylating enzymes with mammalian Pol II during the transcription cycle or about the efficiency and timing of the cap guanylylation and methylation steps. To rectify these gaps in our knowledge, we monitored the binding of Mce1 and Hcm1 to Pol II complexes arrested at discrete steps of the initiation and elongation reactions on the HIV transcription unit. We report that the capping enzymes are associated with TECs, but not PICs. Whereas Mce1 and Hcm1 are bound stably to TECs containing nascent RNAs as short as 14 nucleotides, the capping and methylation reactions do not occur until the nascent pre-mRNA has attained a chain length of 19–22 nucleotides, at which point an ~4–7 nucleotide 5' segment is extruded from Pol II. Mce1 and Hcm1 binding to template-engaged Pol II does not require RNA synthesis, but does require CTD phosphorylation. Both Mce1 and Hcm1 travel with Pol II during elongation down the template and remain poised to quantitatively cap the nascent pre-mRNA when elongation is halted. In contrast, capping of nascent HIV mRNA is much less efficient during unimpeded Pol II elongation. Under such conditions, cotranscriptional capping of HIV mRNA is strongly stimulated by Tat, and this

stimulation of capping requires the C-terminal segment of Tat that mediates its direct binding to Mce1. Our findings implicate cap formation as a component of a transcription elongation checkpoint critical to HIV gene expression.

MATERIALS AND METHODS

Recombinant capping enzymes and Tat. The bifunctional mammalian triphosphatase-guanylyltransferase Mce1 and the cap guanine N7-methyltransferase Hcm1 were produced in *E. coli* as N-terminal His-tagged fusions and purified by Ni-agarose affinity chromatography as described previously (94, 195). A new version of Hcm1 with an N-terminal hemagglutinin (HA) tag and a His tag was produced and purified by the same procedure. HIV-1 Tat was produced in *E. coli* as a glutathione S-transferase fusion and purified by glutathione-Sepharose affinity chromatography (187).

DNA templates. The DNA templates used for preparing transcription ternary complexes were generated by PCR amplification using the HIV-1 LTR promoter-containing plasmid p10SLT as a template (116) and primer pairs corresponding to p10SLT plasmid sequences 5'-ACCAGTTGAACCAGAGC and 5'-CACACTGACTAAAAGGGT. The first primer contained a biotin at the 5' end which was used to immobilize the DNA on beads. Gel-purified duplex PCR products (0.25 μg of DNA) were adsorbed to 25 μl of streptavidin-coated magnetic beads (Dynal Inc.) by overnight incubation at 22°C in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 1 M NaCl. The beads were washed with the same buffer to remove any free DNA and then stored in 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA.

Stepwise walking of Pol II. Stepwise transcriptions were performed as described previously (178, 179). Preinitiation complexes (PICs) were assembled

by incubating the immobilized DNA template (200 ng) in a volume of 25 μl containing 12 μl of HeLa cell nuclear extract, 6 mM $MgCl_2$, and 0.5 μg of poly(dA-dT) for 15 min at 30°C. Pol IIO complexes at the PIC stage were formed by including 200 μM dATP during the incubation. To remove unbound materials, the bead-bound PICs were pelleted and then washed with 25 μl of buffer A (20 mM HEPES, pH 7.9, 100 mM KCl, 20% (v/v) glycerol, 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethyl-sulfonyl fluoride, 6 mM $MgCl_2$). PICs were walked to template position +14U by incubation with 12.5 mM phosphocreatine, 20 μM CTP, GTP, UTP, and dATP for 5 min at room temperature. TECs stalled at +14U were washed with 25 μl of buffer B (buffer A containing 0.05% Nonidet P-40 and 0.015% Sarkosyl) and twice with 25 μl of buffer C (buffer A containing 0.05% Nonidet P-40). The TECs were walked stepwise along the DNA template by serial incubations with different combinations of NTPs (20 μM each). For all experiments that required additional purified proteins (or cellular factors supplied by nuclear extract), the TECs were incubated with the purified proteins (or nuclear extract) at 30°C for 10 min and then washed three times with buffer C before the next manipulation. For chase experiments, the stalled TECs were incubated with 2 μM ATP, CTP, GTP, and UTP for 10 min at 30°C.

Western blotting. For the isolation of ternary complexes on immobilized DNA template, the PICs and TECs stalled at different steps were released by restriction enzyme digestion at 30°C. The supernatant phase containing released ternary complexes was resolved by 8% SDS-PAGE, and the gel contents were

transferred to a polyvinylidene difluoride membrane (Bio-Rad), which was then immunoblotted with antibodies raised against either the largest subunit of Pol II or Mce1 or Hcm1. Immunoreactive polypeptides were visualized by chemiluminescence using either an ECL kit from Amersham Pharmacia Biotech or a BM chemiluminescence Blotting Kit from Roche Molecular Biochemicals. Pol II antibodies were a gift of Dr. Michael Dahmus. Rabbit antibodies were raised against the C-terminal domain of Mce1. Anti-GFP antibodies were purchased from Clontech.

Capping of nascent RNAs in stalled transcription elongation complexes. TECs stalled at different positions were formed by stepwise transcription, and nascent RNAs were labeled internally by including [-32P]CTP in the reaction mixture. TECs were incubated with 20 pmol of Mce1 at 30°C for 10 min and washed three times with 200 μl of transcription buffer C (as described in "Stepwise Walking of Pol II") to remove the unbound Mce1. Mce1-treated TECs were incubated in capping reaction mixtures (20 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM DTT, 50 μM GTP, and 2.5 mM $MgCl_2$ for 10 min at 30°C. The reaction was quenched by adding 200 μl of stop solution (0.3 M Tris-HCl, pH 7.5, 0.3 M sodium acetate, 0.5% SDS, 2 mM EDTA). The mixtures were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and then with chloroform. RNAs were recovered by ethanol precipitation and then analyzed by electrophoresis through a 15% polyacrylamide gel containing 7 M urea in Tris-borate, EDTA buffer. Labeled RNA products were visualized with a phosphorimager.

Cap methylation of nascent RNAs in stalled transcription complexes. Nascent unlabeled RNAs in TECs stalled at A22 were 5' 32P cap-labeled by Mce1 in capping reaction mixtures as described above containing 1.25 μM [^{32}P]GTP. The ^{32}P cap-labeled A22 TECs were then incubated with 20 pmol of Hcm1 for 10 min at 30°C. After removing the unbound Hcm1, TECs were walked to different positions, released from the beads by restriction digestion, and incubated in a methyltransferase reaction mixture (20 μl) containing 50 mM Tris-HCl (pH 8.0), 5 mM DTT, and 50 μM AdoMet (Aldrich) for 10 min at 30°C. The reaction was quenched by adding 200 μl of stop solution. RNAs were recovered by phenol-extraction and ethanol precipitation and dissolved in nuclease S1 buffer (pH 5.5) (GIBCO-BRL). The samples were incubated with 20 U of nuclease S1 (GIBCO-BRL) for 2 hr at 37°C. The digests were than spotted on polyethyleneimine-cellulose TLC plates that were developed with 0.2 M $(NH4)_2SO_4$. Cap dinucleotides were visualized by scanning the TLC plate with a phosphorimager.

RNA capture assay. TECs containing ^{32}P-CMP labeled nascent RNA were stalled at +14U. The U14 TECs were incubated with 20 pmol of purified Mce1 for 10 min at 30°C. The unbound proteins were removed by washing three times with 200 μl of transcription buffer C. TECs were chased by the addition of 2 μM NTPs for 90 s in the presence of 10 pmol of wild-type or mutant Tat proteins. RNAs were recovered by phenol extraction and ethanol precipitation, then treated with 50 pmol of purified Hcm1 for 1 hr at 37°C. Hcm1 will only methylate

the 5' termini of RNAs that had been guanylylated during the transcription elongation reaction. m^7G-capped RNAs were recovered by "Capture" affinity chromatography, a method based on the selective binding of m^7GpppRNA to immobilized cap binding protein eIF4E. The recombinant GST-eIF4E fusion protein was expressed in bacteria (with an expression plasmid kindly provided by Dr. J. Pelletier, McGill University), then purified and immobilized on glutathione-Sepharose beads according to previously described procedures with minor modifications (54, 152). RNAs were preheated for 2 min at 95°C in 100 µl of cap binding buffer (17 mM Hepes, pH 7.9, 85 mM NaCl, 80 µM EDTA, 1 mM DTT, 17% glycerol, 0.08% NP-40, 1% polyvinyl alcohol, 200 µg/ml calf liver tRNA) and then cooled on ice. 6 mM DTT was added, followed by 2 U of RNase inhibitor (Promega). The RNAs were mixed with immobilized GST-eIF4E in 300 µl of cap binding buffer for 1 hr at 4°C. The beads were washed six times with 600 µl of washing buffer (20 mM Hepes, pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 20% glycerol, 0.1% NP40). The bound RNAs were eluted from the beads by phenol extraction and recovered by ethanol precipitation and then analyzed by 12% polyacrylamide/7 M urea PAGE. Labeled RNAs were visualized by scanning the gel with a phosphorimager.

Coimmunoprecipitation of Tat with Mce1. pEGFP-Mce1 expression plasmids were constructed by fusing the coding sequences for EGFP and Mce1 according to protocols provided by the pEGFP vendor (Clontech). HA-Tat-tagged plasmids were kindly provided by Dr. Katherine Jones. In these plasmids,

expression of the EGFP-Mce1 and HA-Tat proteins is driven by a CMV promoter. The plasmids were cotransfected into HL3T1 cells with Lipofectamine (GIBCO-BRL). After 48 hr, the cells were lysed with RIPA buffer (20 mM Tris-HCl, pH 8.0, 0.5% Nonidet P-40, 1% Triton X-100, and 150 mM KCl, 5 mM DTT). Cell lysates (120 μg of protein) were immunoprecipitated with anti-GFP or anti-Mce1 antibodies which had been adsorbed to protein G-Sepharose beads (Amersham Pharmacia Biotech) in RIPA buffer during an overnight incubation at 4°C. The beads were washed three times with 300 μl of RIPA buffer containing 0.015% Sarkosyl and 1 M KCl and once with 300 μl of RIPA buffer. The bead-bound proteins were eluted with SDS, resolved by 8% SDS-PAGE, and probed by Western blotting with anti-HA antibodies.

RESULTS

Capping enzyme is not a component of PICs formed at the HIV-1 LTR Promoter but is present in TECs

Pol II PICs were assembled on immobilized DNA templates containing the HIV-1 LTR promoter and then probed by Western blotting for the presence of Pol II and Mce1. PICs were prepared by incubating the bead-bound DNA with NTP-depleted HeLa cell nuclear extract and then released from the beads by cleaving the template upstream of the promoter with the restriction enzyme *Bsp*E1 (Figure 1B). Whereas Pol II was present in PICs, Mce1 was not, although Mce1 was readily detected in the nuclear extract (Figure 1C, lanes 1 and 2). Moreover, supplementation of the PICs with purified recombinant Mce1 resulted in only trace levels of binding of the capping enzyme to the PICs (Figure 1C, lane 3).

To determine if and when Mce1 associates stably with the Pol II transcription complex, we converted the PICs into elongation complexes arrested at discrete promoter-proximal sites by provision of only three NTPs for initiation and elongation. The 5' sequence of the HIV-1 pre-mRNA is shown in Figure 1A. Elongation complexes starved for ATP were stalled at template position +14U. After stripping the bead-bound TECs with sarkosyl to prevent further initiation events, the U14 TECs were walked to +30C. The arrested TECs were supplemented with recombinant Mce1 and then washed with transcription buffer to remove unbound proteins. The TECs were released from the beads by cleavage with *Pvu*II and then probed by Western blotting for Pol II and Mce1.

Capping enzyme was detected in the U14 and C30 TECs after incubation with purified Mce1 (Figure 1C, lanes 5 and 7). There was no Mce1 present in TECs that had not been supplemented (Figure 1C, lanes 4 and 6). The PICs and TECs contained similar amounts of the Pol II large subunit.

Nascent HIV-1 Ttranscripts as short as 22 nucleotides are capped *in vitro* by Mce1

To determine when nascent HIV transcripts acquire their caps, we arrested TECs at templates positions +14U, +22A, +30C, and +46U, washed them to remove nucleotides, supplemented them with Mce1, and washed again to remove unbound Mce1. These TECs were then incubated with GTP, the substrate for the guanylyltransferase component of Mce1. The addition of an unlabeled cap guanylate to the 5' end of an internally labeled RNA molecule results in a characteristic slowing of the electrophoretic mobility of the RNA, equivalent to about a 1 nucleotide increase in apparent chain length (50). Here we studied the capping reaction of Mce1 using stalled transcription complexes containing RNA transcripts labeled internally with 32P-CMP during a stepwise transcription-walk down the template (Figure 2).

We found that A22, C30, and U46 TECs catalyzed near-quantitative addition of a cap guanylate to their nascent RNAs in a reaction that depended on prior incubation with Mce1 and the provision of GTP (Figure 2B). However, the nascent 14-mer RNA in the U14 TECs was refractory to guanylylation (Figure

2B), even though Mce1 is present in the U14 transcription complexes (Figure 1C). These data show that capping of HIV mRNA within paused elongation complexes can occur during a defined interval between the synthesis of 14-mer and 22-mer nascent transcripts. We infer that the 5' end of the U14 transcript is inaccessible to the active site(s) of Mce1 because it is sequestered within the nascent RNA binding pocket of Pol II. Thus, the window for capping on paused TECs reflects a steric constraint on the minimum substrate size.

HIV-1 mRNA transcripts can be capped by the mammalian capping enzyme traveling with the early transcription elongation complexes

To probe the window for cap formation, we asked whether Mce1 that was recruited to stalled U14 TECs is available to cap the nascent RNA at a later stage during a stepwise transcription walk (Figure 2C). The size heterogeneity of the RNAs contained in U14 TECs, being either 12, 13, or 14 nucleotides (Figure 2D, lanes 1 and 2), reflects the capacity of Pol II to initiate at any of the three G residues at +1 to +3 of the transcription unit (the degree of heterogeneity varies from one experiment to another). We incubated stalled U14 TECs with or without purified Mce1, removed the unbound Mce1 by washing, and allowed Pol II to elongate to +30C. Whereas the original RNAs in the U14 complex were not capped in response to Mce1, they were capped nearly quantitatively by the time Pol II had extended the chain to 30 nucleotides. Thus, the capping enzyme binds to the TEC before the 5' end is extruded from the polymerase and remains

poised during subsequent elongation steps to cap the nascent chain after it is extruded from the polymerase.

Capping enzyme association with Pol II elongation complexes does not require RNA

Our finding that Mce1 can load onto the U14 TECs, at a stage when the nascent mRNA is not exposed to cap addition, suggested that the recruitment of Mce1 to the Pol II elongation complexes may be RNA independent. To address this question directly, we first treated A22 TECs with RNase A to remove the segment of the nascent RNA that had been extruded from Pol II and then examined the binding of Mce1 to the digested A22 TECs (designated A22' in Figure 3). RNase A converted the 32P-labeled A22 transcript into a discrete product of ~15 nucleotides (Figure 3B). After washing to remove RNase A, the A22' TECs remained competent to elongate the digested RNA during a subsequent walk to +26G (Figure 3B). This result verifies that the RNase digestion removed only the 5' end of the A22 RNA and left the 3' OH end of the RNA intact at the polymerase active site. Neither the digested 15-mer RNA in the A22' TEC nor the 19-mer RNA in the G26' TEC could be capped in vitro by Mce1, because the 5' OH generated by RNase A cleavage cannot serve as a substrate for RNA guanylyltransferase (Figure 3B). Nonetheless, Mce1 bound to the RNase-digested A22' TEC as well as it did to the intact A22 TEC, as gauged by Western blotting of the TECs released from the beads by *Pvu*II digestion of the template (Figure 3C). Based on these results, we conclude that Mce1

association with Pol II elongation complexes does not require a direct interaction between Mce1 and nascent RNA.

CTD phosphorylation is required for capping enzyme association with Pol II at the HIV-1 LTR promoter

Mce1 is likely to interact with the Pol II elongation complex through direct protein-protein contacts. Mce1 binds in vitro to the phosphorylated Pol II CTD, but not to the unphosphorylated CTD; the CTD interaction is mediated by the C-terminal guanylyltransferase domain of Mce1. Binding of Mce1 to a CTD peptide phosphorylated at position Ser5 stimulates Mce1 guanylyltransferase activity (93). To determine whether CTD phosphorylation is required for recruitment of Mce1 to the HIV LTR promoter, we prepared PICs containing either the hypophosphorylated IIA or the phosphorylated IIO forms of Pol II by incubating the PICs with or without dATP (which serves as a substrate for the TFIIH-associated CTD kinase, but not for RNA synthesis by Pol II) (Figure 3D). The PICs were then incubated with purified Mce1 and washed to remove unbound protein. The PICs were released from the beads by digestion with *Bsp*EI, and their associated polypeptides were resolved by electrophoresis through a 5% polyacrylamide gel containing SDS. Western blotting with Pol II antibody showed that the PICs not exposed to dATP contained only the IIA form, whereas inclusion of dATP elicited its conversion to the IIO form (Figure 3E). Mce1 was associated exclusively with PICs that had undergone CTD phosphorylation in the presence of dATP (Figure 3E). We conclude that Pol II CTD phosphorylation is

necessary for recruitment of mammalian capping enzyme to the transcription machinery at the HIV LTR promoter.

Mce1 release from the Pol II complex during transcription elongation requires a factor present in nuclear extract

To determine the fate of Mce1 after cap guanylylation, TECs stalled at +22A with capped nascent transcripts were prepared and then incubated with or without nuclear extract (Figure 4A). After washing to remove unbound proteins, the NE-treated and control TECs were chased for 10 min to the end of the template. TECs were released by restriction enzyme digestion, and the associated proteins were detected by immunoblotting. RNA transcript analysis showed that the stalled A22 TECs with capped nascent RNAs had elongated to produce an ~168 nucleotide transcript (Figure 4B). Most of the Mce1 that had been associated with the arrested A22 TECs remained bound to the transcription complexes after elongation to the 3' end of the template (Figure 4C). However, the transient exposure of the A22 TECs to nuclear extract resulted in the release of the majority of the Mce1 from the transcription complex (Figure 4C). We conclude that Mce1 can dissociate from Pol II during later stages of transcription elongation and that nuclear extract contains a factor (or factors) that stimulate this dissociation step.

Cap methyltransferase Hcm1 associates stably with the transcription complex and methylates the caps of nascent mRNAs

The enzyme RNA (guanine-N7) methyltransferase (referred to hereafter as cap methyltransferase) catalyzes the transfer of a methyl group from AdoMet to the GpppRNA terminus to produce m7GpppRNA and AdoHcy. To determine when the cap structure of nascent HIV-1 transcripts becomes methylated, we incubated stalled A22 TECs with purified Mce1 and [α-^{32}P]GTP so that the nascent A22 RNA will be 5' cap labeled with 32P-GMP. After removing the unbound protein and free [α-^{32}P]GTP by washing with transcription buffer, the cap-labeled A22 TECs were incubated with purified human cap methyltransferase (Hcm1) for short periods of time. After washing away the unbound Hcm1, the A22 TECs were walked stepwise down the template by incubation with different sets of three NTPs. Stalled TECs were released from the beads by restriction enzyme digestion. At this point, the released transcription complexes were incubated in the presence of AdoMet, and the labeled RNAs were recovered by phenol extraction and ethanol precipitation. An electrophoretic analysis of the labeled RNAs recovered from A22, U46, and C61 TECs and complexes chased to near the 3' end of the template is shown in Figure 5B. To analyze the cap structures present on the nascent transcripts, the isolated RNAs were digested to cap dinucleotides with nuclease S1 and then analyzed by polyethyleneimine-cellulose thin layer chromatography, which resolves the GpppG cap from the methylated cap m7GpppG (Figure 5C). The TLC analysis shows that the caps found in A22 complexes that had not been exposed to exogenous Hcm1 were exclusively unmethylated, but that transient exposure to

Hcm1 resulted in stable binding of the cap methyltransferase to the A22 TEC, such that 25% of the caps were methylated during a subsequent incubation of the A22 TECs with AdoMet (Figures 5C and 5D). The efficiency of cap methylation increased steadily as the elongation complex progressed down the template, implying that (i) access of Hcm1 to the 5' GpppRNA end may increase as more RNA is extruded from the polymerase or (ii) other protein factors that impede such access may dissociate from the 5' end as the polymerase elongates. By the time Pol II had approached the 3' margin of the 168 nucleotide transcription unit, nearly 90% of the caps could be methylated by Hcm1 associated with the TEC (Figures 5C and 5D). Cap methylation by Hcm1 depended on the inclusion of AdoMet in the reaction mixture (data not shown).

The binding of Hcm1 to template-engaged Pol II was gauged by incubating pol IIA PICs, pol IIO PICs, and TECs with HA-tagged Hcm1, followed by washing to remove unbound protein, release of the Pol II complexes by restriction endonuclease digestion, and probing their protein content by Western blotting with antibodies against Pol II and the HA-tag of HA-Hcm1 (Figure 5E). We found that the cap methyltransferase is excluded from the PIC containing Pol IIA, but present at similar levels (relative to Pol II) in PICs containing pol IIO and in TECs paused at promoter proximal (U14, U22, U46) and promoter distal (+168) positions on the template. That the methyltransferase remains physically associated with the elongating polymerase complex confirms the functional analysis of cap methylation by TECs paused at proximal and distal template

sites. An instructive finding was that transient exposure of the U46 TECs to nuclear extract after they had been incubated with Hcm1 did not diminish the amount of Hcm1 present in TECs that were subsequently chased to the end of the template (Figure 5E, lane 6). Thus, Hcm1 is not susceptible to dissociation by the factor(s) in the nuclear extract that triggers release of Mce1 during elongation (compare Figures 4C and 5E).

Tat stimulates the capping of nascent HIV-1 RNA

Although our experiments show that short nascent Pol II transcripts (on the order of 22 nucleotides) can be capped and methylated, such results reflect the action of the mammalian capping enzymes on arrested polymerase elongation complexes, where there is no kinetic competition between ongoing elongation and capping. The unimpeded rate of elongation by RNA polymerases (~20 nucleotides/s) is faster than the rate of the overall RNA capping reaction (94). Thus, it is not clear when (or how efficiently) capping would occur during reactions in which elongation is ongoing. To address this issue for transcription directed by the HIV-1 LTR promoter, we analyzed the abundance of capped RNAs via "Capture" affinity chromatography with immobilized eIF4E (Figure 6A) (54).

TECs containing internally ^{32}P-CMP-labeled nascent RNAs were stalled at position +14U and then incubated with purified Mce1. As shown above, Mce1 can associate with TECs at this stage, but cannot cap the nascent transcript. The

U14 TECs were then chased by the addition of 2 μM concentration of all four NTPs for 90 s. The labeled nascent RNAs were isolated and then reacted with Hcm1 and AdoMet in order to methylate any RNAs that had been guanylylated by Mce1 during the chase phase. The m7G-capped nascent transcripts were isolated by eIF4E affinity chromatography, and the bound material ("B") was analyzed by denaturing PAGE in parallel with a sample of the unfractionated input nascent RNA population ("I") and the supernatant phase containing uncapped RNA ("S"). The specificity and efficiency of this assay is demonstrated in Figure 6B (lanes 2–6), whereby ^{32}P-CMP labeled RNAs in A22 TECs that were guanylylated by Mce1 (compare lanes 2 and 3 and note the mobility shift resulting from cap guanylylation) and split into two portions, one of which was treated with Hcm1 and AdoMet. Whereas <5% of the unmethylated GpppRNA was retained on the eIF4E affinity column, the input Hcm1-treated RNA was retained nearly quantitatively (95% bound) (lane 4 versus lane 6).

In contrast to the high efficiency of cap guanylylation observed for arrested TECs (nearly 100%), we found that when U14 TECs were chased for 90 s (yielding a heterogeneous population of transcripts from ~19 to 70 nucleotides in length), only 25% of the nascent chains had 5' GpppN caps that could be converted into m7G caps by Hcm1 and thus bind to eIF4E (Figures 6B and 6C). This result raised an important question of whether additional factors might be required to more tightly couple capping of HIV mRNAs to early transcription elongation.

Previously, we showed that the HIV Tat protein, which activates gene expression from the HIV-1 LTR at the early elongation phase of the transcription cycle, interacts directly with Mce1 in vitro and stimulates the activity of both the RNA triphosphatase and RNA guanylyltransferase components (27). We also found that Tat stimulated the capping of isolated HIV TAR mRNA, which is inefficiently guanylylated by Mce1. To analyze the effect of Tat on mRNA capping during HIV-1 transcription elongation, we loaded Mce1 on stalled U14 TECs and included recombinant Tat during the chase phase. More than 80% of the nascent RNAs chased in the presence of Tat were bound to the eIF4E beads after reaction with Hcm1 (Figures 6B, lane 11, and 6C). The size distribution of the capped RNA population was indistinguishable from that of the input RNA. Thus, Tat significantly increased the efficiency of cotranscriptional capping.

Tat is composed of two major functional domains: an N-terminal transactivation domain (amino acids 1–48) that interacts with the cyclin T1 subunit of P-TEFb, and a C-terminal domain (amino acids 49–86) that mediates RNA binding and nuclear localization. Two truncated versions of Tat—Tat2/36 and Tat48—which are deleted for the transactivation domain and RNA binding domain, respectively, were tested for their effects on cotranscriptional cap formation. Tat48 did not significantly enhance cap guanylylation (29% of nascent RNAs "captured") compared to the basal level of 25% in the absence of Tat. However, Tat2/36 did stimulate capping to an extent of 52% (Figures 6B and 6C). We showed previously that Mce1 binds in vitro to either wild-type Tat or

Tat2/36 (but not to Tat48) and that the interactions with Tat or Tat2/36 stimulate the Mce1 guanylyltransferase activity (27). Thus, the ability of Tat to stimulate capping of nascent RNA during transcription elongation correlates with its capacity to bind and activate the isolated mammalian capping enzyme.

Interaction of Tat and mammalian capping enzyme *in vivo*

Direct association of Tat and Mce1 has been demonstrated by protein-affinity chromatography in vitro (27). To analyze the interaction of Tat and Mce1 in vivo, we contransfected human HL3T1 cells with plasmids expressing GFP-Mce1 and HA-Tat fusion proteins. Immune precipitates of cell lysates were prepared using anti-GFP antibodies and then probed by Western blotting as shown in Figure 7. IP with anti-GFP antibody coprecipitated HA-Tat from cells that had been transfected to produce GFP fused to either full-length Mce1 or to the guanylyltransferase domain Mce1(211–597), but not from cells producing GFP alone. Similar results were obtained when the IP step was performed using anti-Mce1 antibody (data not shown). Control immunoblots of the unfractionated lysates confirmed that each of the expected proteins was produced in the transfected cells (Figure 7). We conclude that Tat can interact with mammalian capping enzyme in vivo in mammalian cells.

DISCUSSION

Freeze-frame analysis of cap formation and capping enzyme recruitment during transcription by human Pol II

Here we investigated how capping and methylation of HIV pre-mRNAs is coupled to Pol II elongation under both static and dynamic conditions. The static or "freeze-frame" approach entails arresting RNA polymerase stepwise at discrete template positions and then querying whether nascent chains within the arrested TECs can be modified when purified capping enzymes are added (81). Because the 5' end of the substrate RNA is a "sitting target" within the arrested TECs, the freeze-frame analysis focuses on the steric constraints on access of the 5' RNA end to the active site(s) of the various cap-forming enzymes. We found that cotranscriptional cap guanylylation of HIV pre-mRNA occurred with virtually 100% efficiency on nascent RNAs 22 nucleotides long, whereas RNAs 14 nucleotides long could not be capped, implying that the 5' end of the growing RNA is accessible to the active site of Mce1 only after a critical chain length is extruded from Pol II. RNase footprinting showed that Pol II protects a 15 nucleotide RNA segment extending back from the 3' growing point of the chain. Our findings about the timing of cap formation under static conditions agree with earlier studies showing that: (i) 6- to 17-mer nascent RNAs synthesized in vitro by Pol II from the adenovirus major late promoter were uncapped (41) and (ii) nascent RNA within TECs stalled in vivo at promoter-proximal positions on the

Drosophila heat shock genes were capped only after attaining a minimal chain length of 20 nucleotides (184). The efficiency of capping of Hsp70 nascent transcripts increased within a narrow window from 50% for 20-mer RNAs to nearly 100% for 30-mer RNAs.

To our knowledge, there have been no prior studies comparing the timing of the cap guanylylation and guanine-N7 methylation steps during transcription elongation. The analysis of nascent heat shock transcripts within arrested TECs relied on removal of the cap with tobacco acid pyrophosphatase and did not discriminate GpppRNA and m7GpppRNA caps (184). Whereas it is obvious that guanylylation is a prerequisite for methylation, it was not clear if the two reactions occur in rapid succession during elongation. Here we demonstrated a distinct lag between guanylylation and methylation. Mce1 readily guanylylated the 22-nucleotide nascent HIV RNA, but only 25% of the caps were methylated by Hcm1. The extent of cap methylation increased to 65% as the TEC was walked stepwise to +46 and to 85% by the time Pol II reached ~+168. Conceivably, Hcm1 may require the extrusion of a longer 5' segment of RNA from the polymerase than does Mce1, or the Hcm1 may be impeded by (or compete with) another factor that interacts with the 5' cap guanylate (Mce1 itself may be such a competing factor).

Mce1 and Hcm1 bind stably to the TEC prior to the extrusion of the 5' end of the nascent RNA out of the RNA polymerase. Indeed, Mce1 binding to the TEC is unaffected by RNase digestion of any 5' RNA segment that is extruded.

Mce1 and Hcm1 can be recruited to the Pol II complex even prior to the initiation of chain elongation, provided that CTD phosphorylation has already occurred. Neither Mce1 nor Hcm1 binds to PICs containing Pol IIA. Thus, CTD phosphorylation is required (and perhaps sufficient) for recruitment of both components of the mammalian capping apparatus.

A key finding is that both Mce1 and Hcm1 remain bound to the Pol II TEC even after the cap guanylylation and N7 methylation reactions have been executed. Transient exposure of TECs to nuclear extract stimulated the dissociation of Mce1 from the TEC during a subsequent chase down the template, but Hcm1 was refractory to dissociation by nuclear extract and traveled with the TEC to the end of the template (near +168). We infer the existence of factor(s) in the extract that trigger the release and recycling of Mce1 after it modifies the 5' RNA end. Our findings concerning the dynamic association of mammalian capping enzymes with the TEC in a reconstituted in vitro elongation/processing system echo recent *in vivo* findings in *S. cerevisiae*, where physical association of the triphosphatase and guanylyltransferase enzymes with actively transcribed genes is maximal at the 5' end of the transcription unit and decays incrementally toward the 3' end of the gene, whereas the cap methyltransferase remains on the template DNA across the full length of the gene (128, 199). Our *in vitro* experiments show that dissociation of the guanylyltransferase is an active process that depends on a factor extrinsic to the TEC. Candidates for such a role include: (i) phosphatases and kinases that

remodel the CTD phosphorylation array and thereby affect CTD interaction with Mce1 (29, 88, 93) or (ii) factors that interact with Mce1 and regulate its affinity for the CTD. Although it remains unclear why the cap methyltransferase would travel with the yeast and human Pol II TECs after the cap is methylated, it is worth considering that this enzyme might play additional roles in transcription or mRNA maturation.

Cap formation during continuous Pol II elongation

The freeze-frame view of capping enlightens us with respect to the earliest step at which capping enzymes can associate with template-engaged Pol II and the window in elongation during which the 5' modifications can first be executed by mammalian capping enzymes. But this analysis does not address whether capping actually occurs at these early stages when Pol II elongation is ongoing or whether cotranscriptional capping is subject to regulation by *trans*-acting factors. The present study provides answers to both of these key questions.

The timing of cap guanylylation under dynamic transcription conditions is determined by a kinetic balance between the rate of Mce1 binding to the TEC, the rates of the serial triphosphatase and guanylyltransferase reactions, the rate of polymerase movement, and the rate of premature dissociation of Mce1 from the TEC. Our experimental strategy entailed preloading Mce1 on TECs prior to extrusion of the 5' end and then gauging the overall efficiency of guanylylation during a brief chase in the presence of all four NTPs. Because we showed that

Mce1 remains bound and travels with the TEC under these reaction conditions, the experiment focuses exclusively on the kinetic balance between cap formation and polymerase movement. The instructive finding was that capping of nascent HIV transcripts under dynamic conditions was much less efficient (25% overall) than capping in the context of an arrested Pol II TEC (nearly 100%).

The inherently low efficiency of cotranscriptional capping of nascent HIV mRNA under dynamic conditions provides a rationale for the elongation checkpoint that regulates HIV gene expression. We hypothesize that DSIF/NELF-induced arrest of Pol II at promoter-proximal sites on the HIV transcription unit expands the kinetic window during which Mce1 can bind to the TEC and catalyze cap guanylylation (in essence providing a "sitting target" for the capping enzyme so that short nascent chains can be more efficiently capped). The cap promotes downstream mRNA processing steps, especially the splicing of the 5' proximal intron, and it protects mRNA from exonucleolytic decay. If Pol II commits prematurely to traversing the entire transcription unit without the benefit of a cap, it runs a risk of failing to excise the first intron in a timely fashion, or perhaps at all, a process that would yield a nonfunctional or misfunctional transcript. Accelerated 5' to 3' decay of unguanylated nascent RNAs would also result in a wasteful round of transcription.

Tat stimulates cotranscriptional capping of HIV mRNA

Our earlier findings that Tat binds directly to Mce1 *in vitro* and stimulates its triphosphatase and guanylyltransferase activity raised the prospect that Tat may regulate cap formation during HIV transcription. Here we showed that recombinant Tat protein stimulates cotranscriptional cap guanylylation during ongoing Pol II elongation. Whereas only 25% of the nascent HIV mRNAs were capped during a chase in the absence of Tat, more than 80% were guanylylated in the presence of Tat. Thus, our results provide an instance in which the efficiency of cotranscriptional capping is demonstrated to be regulated by a *trans*-acting factor.

A highly instructive point about the Tat-mediated stimulation of HIV RNA guanylylation is that it applies to nascent RNAs ranging from 19 to 70 nucleotides in length (Figure 6). The majority of the transcripts whose capping is enhanced by Tat are shorter than the 45 nucleotides of nascent 5' RNA that would be required to form the minimal TAR secondary structure. Thus, we conclude that the salutary effect of Tat on cotranscriptional capping precedes, and is independent of, its interaction with TAR. The C-terminal domain of Tat that suffices for binding and stimulation of isolated Mce1 and for stimulation of cotranscriptional capping includes the segment required for Tat binding to TAR. Thus, it is sensible that Tat stimulation of capping is temporally unlinked from TAR binding, given the likelihood that the two binding events (to Mce1 or TAR) are mutually exclusive for a given molecule of Tat protein. A second key point about the Tat effect on capping is that it occurs on TECs that have already been

loaded with Mce1, i.e., the effect is directly on cap formation rather than on the recruitment of Mce1 to the TECs.

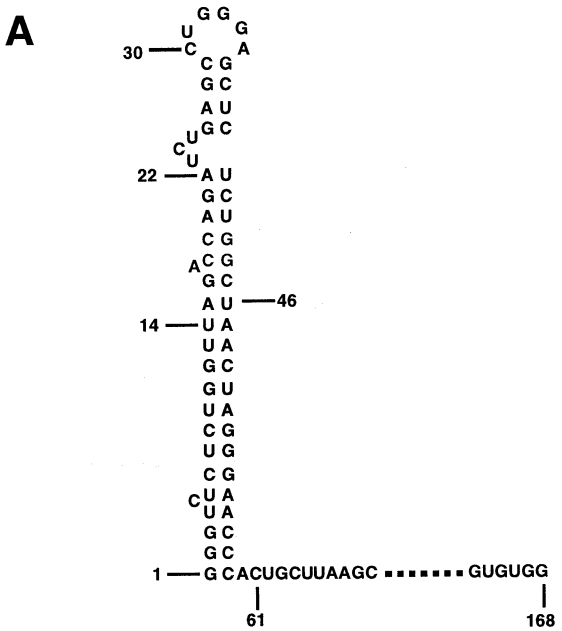
Because Tat binds directly to Mce1 in solution, it is conceivable that Tat joins the TEC at a stage prior to TAR formation, using Mce1 as a bridge. Alternatively, Tat may interact with the TEC via a parallel pathway that is independent of Mce1. By applying the freeze-frame approach to gauge the association of Tat with Pol II TECs on the HIV transcription unit, we have found that Tat is not present in PICs, but is present in U14 TECs formed in the absence of Mce1 (274). Thus, Tat requires neither TAR nor capping enzyme to bind to the Pol II complex. Nonetheless, it is likely that the functional status of Tat and its macromolecular interactions are subject to change as Mce1 and TAR are elaborated as components of the TEC.

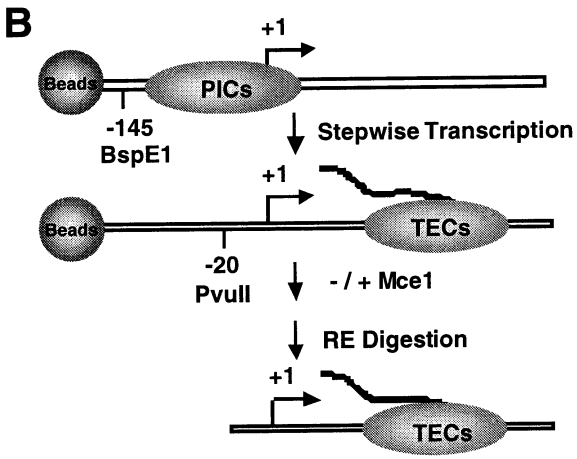
A revised view of Tat function in HIV transcription

A dominant paradigm for Tat upregulation of HIV gene expression at the level of transcription elongation revolves around the ability of the Tat-TAR complex to bind P-TEFb, stimulate its phosphorylation of the CTD and Spt5, and thereby override the elongation arrest elicited by DSIF and NELF. Here we show that Tat plays a TAR-independent role in stimulating cotranscriptional capping of nascent HIV mRNAs. This function rationalizes elongation arrest and restart as a means to ensure a temporal window for efficient cap formation. In this revised regulatory axis, Tat elicits a stimulation of capping (in concert perhaps with Spt5

and CTD-PO4) and triggers the release from elongation arrest in concert with P-TEFb.

Figure 1. Mammalian capping enzyme is not a component of PICs but can associate with TECs on the HIV-1 transcription unit. (A) RNA sequence of the early transcribed region (+1 to +70) in the HIV-1 promoter. TAR RNA secondary structure and various positions of transcription arrest are shown. (B) PICs were prepared by incubating an immobilized DNA template containing the HIV-1 LTR promoter with HeLa nuclear extract. TECs were formed by stepwise transcription. PICs and stalled TECs were incubated with purified Mce1; then, unbound proteins were removed by washing with buffer. To isolate PICs, DNA was cleaved at the indicated *Bsp*E1 restriction site. For isolating TECs, RNA Pol II complexes were released from the beads by cleavage at the indicated *Pvu*II restriction site. (C) Protein contents of PICs (lanes 2 and 3) or TECs stalled at U14 (lanes 4 and 5) and C30 (lanes 6 and 7) were resolved by SDS-PAGE followed by immunoblotting with antibodies against RNA Pol II and Mce1. Complexes were isolated without (lanes 2, 4, and 6) and with Mce1 supplementation (lanes 3, 5 and 7). Lane 1 contains 10% of the nuclear extract used in the PIC assembly reaction.


A
30
22
14
1
46
GCACUGCUUAAGC
GUGUGG
61
168


B
+1
Beads
PICs
-145
BspE1
Stepwise Transcription
+1
Beads
TECs
-20
PvuII
- / + Mce1
RE Digestion
+1
TECs

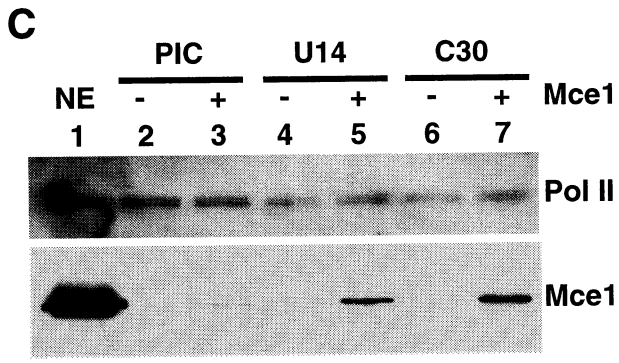

C
PIC
U14
C30
NE
-
+
-
+
-
+
Mce1
1
2
3
4
5
6
7
Pol II
Mce1

Figure 2. HIV-1 mRNA capping during transcription elongation. (A) ^{32}P-CMP-labeled TECs stalled at different positions on the template were prepared by stepwise transcription and then incubated with or without Mce1. After removing the unbound Mce1 by washing with transcription buffer, the TECs were incubated with or without 50 μM GTP in reaction mixtures (20 μl) containing 50 mM Tris-HCl, pH 8.0, 5 mM DTT, and 2.5 mM $MgCl_2$. (B) PAGE analysis of radiolabeled RNAs isolated from U14 (lanes 1–4), A22 (lanes 5–8), C30 (lanes 9–12), or U46 (lanes 13–16) TECs after reaction with Mce1 and GTP as indicated above the lanes. Capped and uncapped species are indicated by arrows. Lane M contains radiolabeled 18-, 33-, and 43-mer DNA oligonucleotide size markers. (C and D) Stalled 32P-CMP-labeled U14 TECs were incubated with (lane 2) or without Mce1 (lane 1), washed, and then walked to +30C (lanes 3 and 4). Labeled RNAs recovered from the U14, and C30 TECs were analyzed by PAGE.

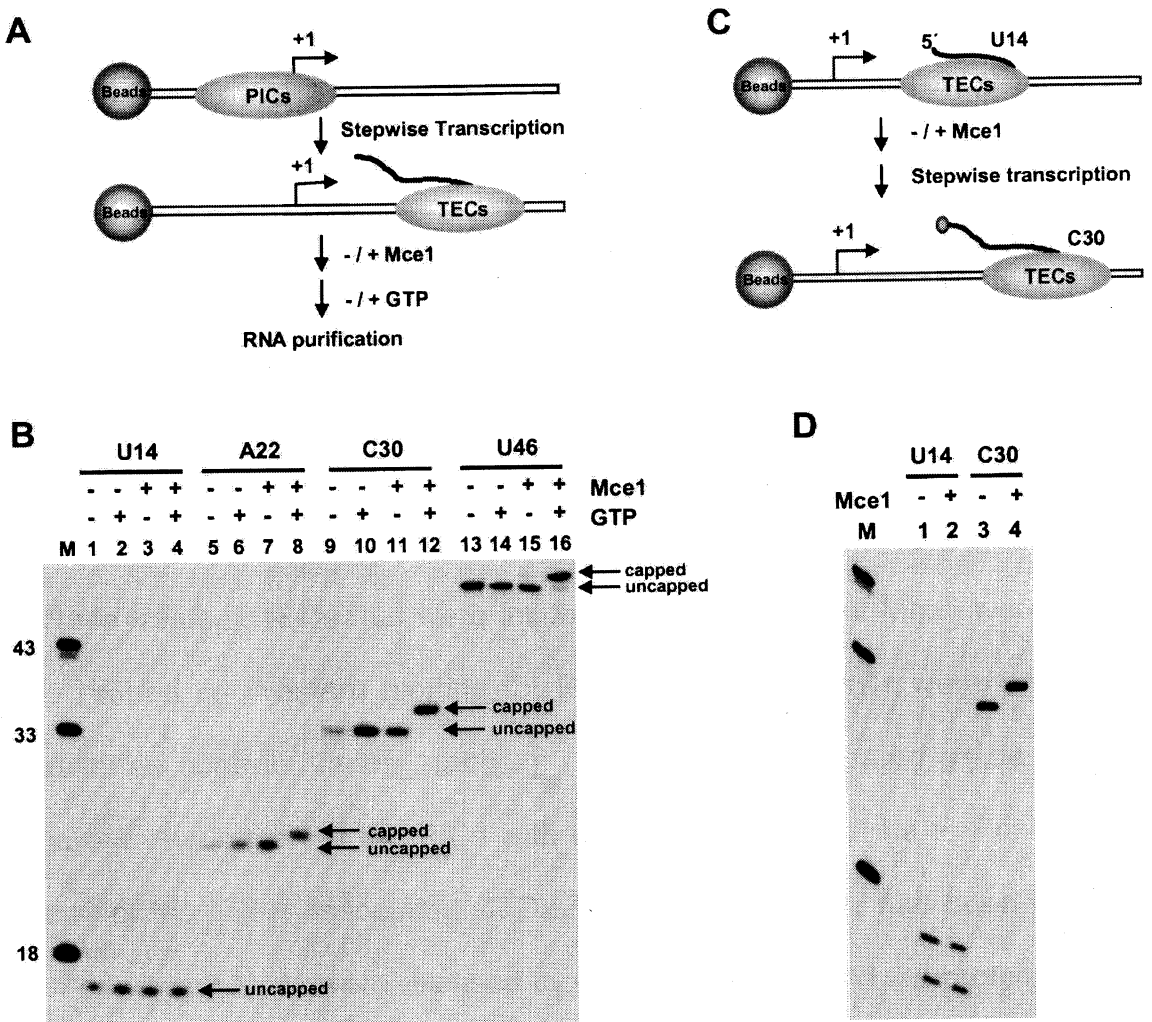
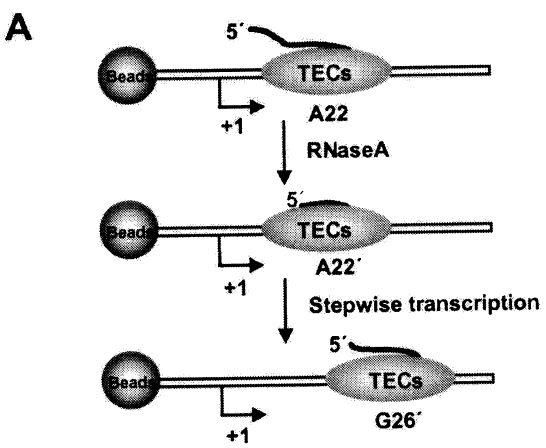
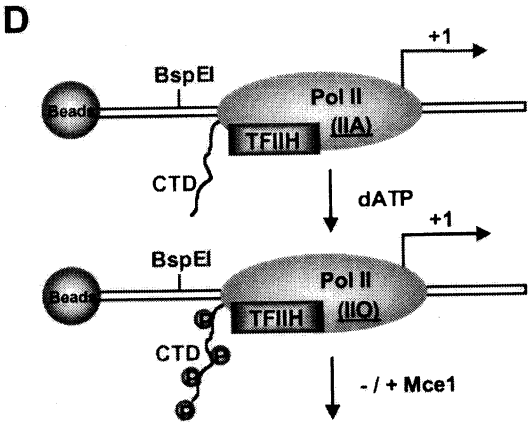
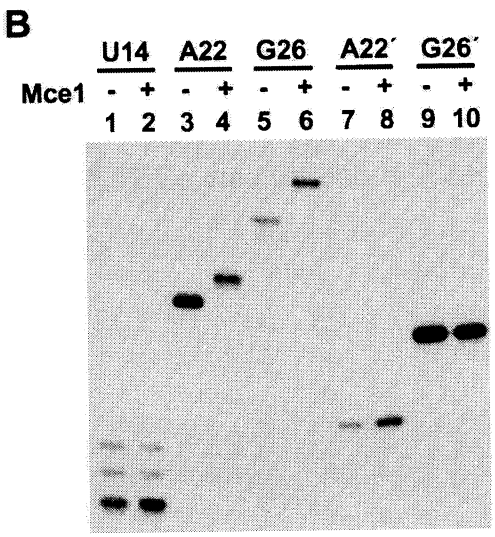
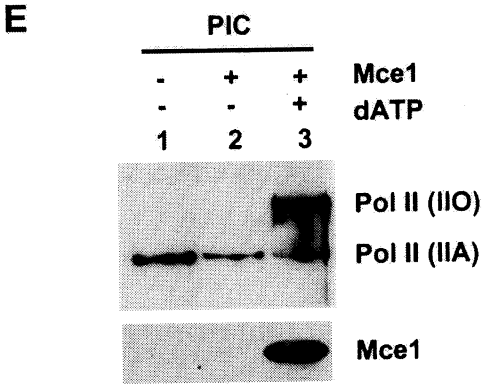

A
+1
Beads
PICs
Stepwise Transcription
+1
Beads
TECs
- / + Mce1
- / + GTP
RNA purification
C
+1
5'
U14
Beads
TECs
- / + Mce1
Stepwise transcription
+1
C30
Beads
TECs
B
U14
A22
C30
U46
Mce1
GTP
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
capped
uncapped
43
33
18
capped
uncapped
capped
uncapped
uncapped
D
U14
C30
Mce1
M
1 2 3 4

Figure 3. Mce1 binding to Pol II transcription complexes requires CTD phosphorylation but does not require nascent RNA. (A) Stalled A22 TECs were treated with 10 μg of RNase A for 5 min at 30°C to digest the extruded 5' end of the RNA (A22') and then walked to position +26G (G26'). (B) PAGE analysis of radiolabeled RNAs isolated from TECs at U14 (lanes 1 and 2), at A22 before (lanes 3 and 4) and after (lanes 7 and 8) RNase A treatment, and after walking from +22A to +26G without (lanes 5 and 6) or after (lanes 9 and 10) RNase treatment. The A22 and A22' TECs had been incubated with or without Mce1 prior to walking to +26G as indicated above the lanes. (C) A22 and A22' (RNase treated) TECs that were incubated with or without Mce1 were isolated and probed by Western blotting for Pol II and Mce1. PICs were analyzed in parallel as controls (lanes 1 and 2). (D) PICs were prepared as described in Figure 1. Incubation of the PICs with dATP converted Pol IIa to Pol IIo by the action of the TFIIH-associated CTD kinase. (E) PICs that had been incubated with or without Mce1 were released by *Bsp*E1 digestion and the protein contents were resolved by 5% SDS-PAGE and probed by Western blotting for the Pol II largest subunit and Mce1.


A
5′
Beads
TECs
A22
+1
RNaseA
5′
Beads
TECs
A22′
+1
Stepwise transcription
5′
Beads
TECs
G26′
+1


D
+1
BspEI
Beads
Pol II
(IIA)
TFIIH
CTD
dATP
+1
BspEI
Beads
Pol II
(IIO)
TFIIH
CTD
- / + Mce1


B
U14
A22
G26
A22′
G26′
Mce1 - + - + - + - + - +
1 2 3 4 5 6 7 8 9 10


E
PIC
- + + Mce1
- - + dATP
1 2 3
Pol II (IIO)
Pol II (IIA)
Mce1

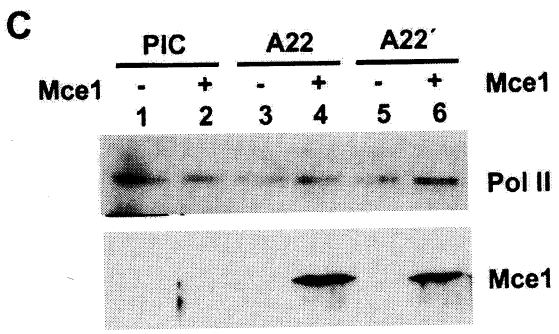
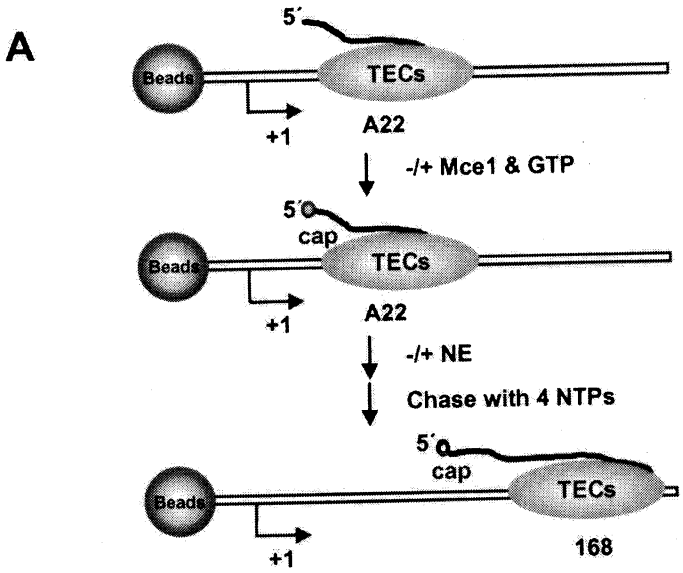
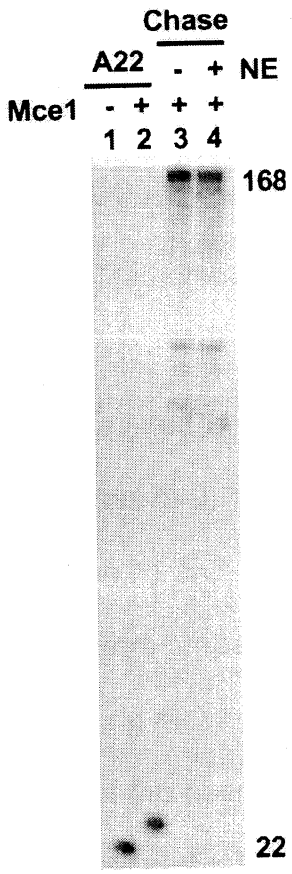

C
PIC
A22
A22′
Mce1 - + - + - + Mce1
1 2 3 4 5 6
Pol II
Mce1

Figure 4. Dissociation of Mce1 from the transcription complex during elongation is stimulated by a factor in nuclear extract. (A) Stalled A22 TECs were incubated for 10 min in a capping reaction mixture containing 20 pmol of Mce1, 50 μM GTP, 50 mM Tris-HCl, pH 8.0, 5 mM DTT, and 2.5 mM MgCl2. After washing to remove Mce1 and GTP, the TECs were incubated with or without nuclear extract (NE; 20 μl). After removing the unbound proteins by washing with transcription buffer, all four NTPs were added to chase the TECs to the +168 position. (B) PAGE analysis of radiolabeled RNAs isolated from the A22 TECs and TECs chased to the end of the template after exposure to Mce1 and nuclear extract as indicated. (C) The presence of Pol II and Mce1 in the A22 and chased TECs was probed by Western blotting. PICs were analyzed in parallel as controls.


A
5´
Beads
TECs
+1
A22
-/+ Mce1 & GTP
5´
cap
Beads
TECs
+1
A22
-/+ NE
Chase with 4 NTPs
5´
cap
Beads
TECs
+1
168


B
Chase
A22
- + NE
Mce1 - + + +
1 2 3 4
168
22

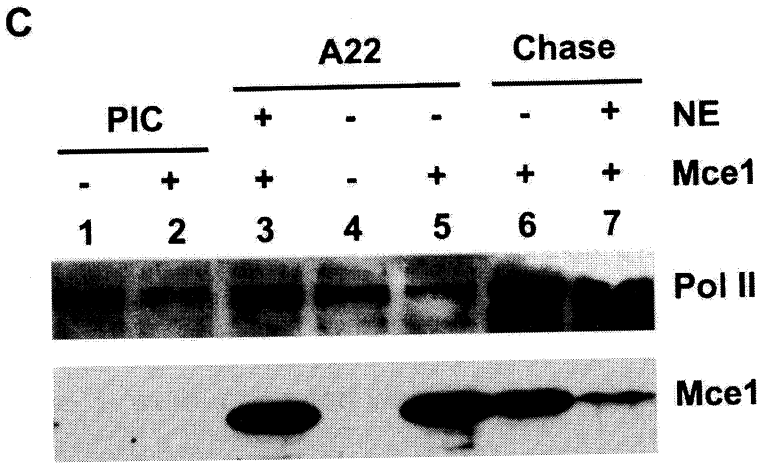
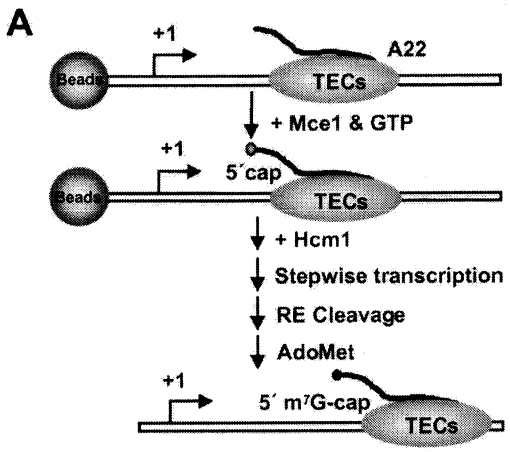
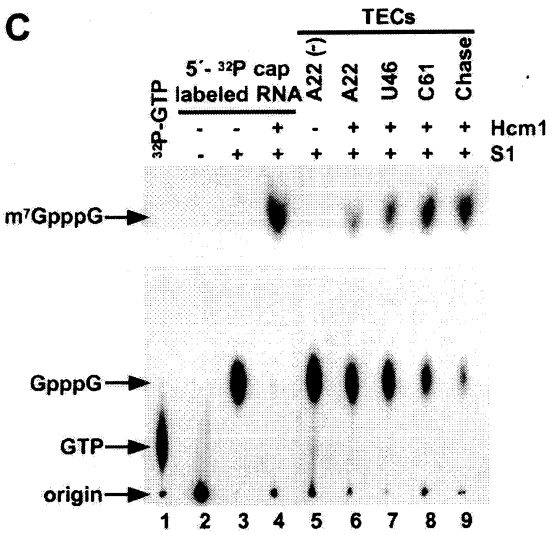
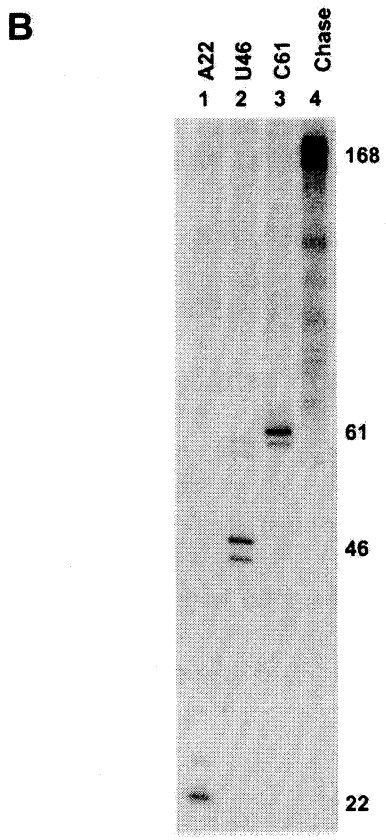
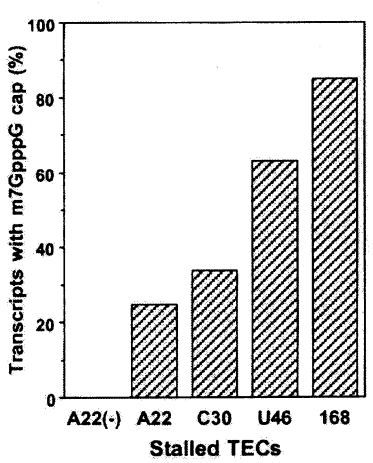

C
A22
Chase
PIC
+ - - - + NE
- + + - + + + Mce1
1 2 3 4 5 6 7
Pol II
Mce1

Figure 5. Hcm1 associates with transcription elongation complexes and methylates the 5' Cap of nascent RNAs. (A) Stalled A22 TECs were 5' ^{32}P-cap labeled by Mce1, washed, and then incubated with Hcm1. After washing to remove unbound Hcm1, TECs were walked to distal template positions, released from the beads, and then incubated in methylation reaction mixtures containing 50 mM Tris-HCl (pH 8.0), 5 mM DTT, and 50 μM AdoMet. (B) Radiolabeled RNAs isolated from the stalled TECs were analyzed by PAGE. (C) The cap structures of the RNAs isolated from the indicated TECs were analyzed by S1 nuclease digestion followed by polyethyleneimine-cellulose TLC (lanes 5–9). The chromatographic origin (lane 2) and the positions of [^{32}P]GTP (lane 1) and ^{32}P-labeled cap dinucleotide GpppG (lane 3) or m7GpppG (lane 4) markers prepared by in vitro guanylylation and methylation of isolated 17-mer RNA using Mce1 and Hcm1 are indicated on the left. (D) The extent of cap methylation of the nascent cap-labeled RNAs by the indicated TECs was quantitated by scanning the TLC plate with a phosphorimager. (E) PICs (lanes 1 and 2) and stalled U14 TECs (lane 3) were incubated with 20 pmol of HA-tagged Hcm1 for 10 min at 30°C. The U14 TECs were walked to A22 (lane 4) or U46 (lane 5). Nuclear extract was added to TECs stalled at U46, washed to remove unbound proteins, and the TECs were chased to +168 (lane 6). The presence of Pol II and HA-Hcm1 in the PICs and TECs was probed by Western blotting.


A
+1
A22
Beads
TECs
+ Mce1 & GTP
+1
5´cap
Beads
TECs
+ Hcm1
Stepwise transcription
RE Cleavage
AdoMet
+1
5´ m7G-cap
TECs


C
TECs
5´- 32P cap labeled RNA
32P-GTP
A22 (-)
A22
U46
C61
Chase
Hcm1
S1
m7GpppG
GpppG
GTP
origin
1 2 3 4 5 6 7 8 9


B
A22
U46
C61
Chase
1 2 3 4
168
61
46
22


D
Transcripts with m7GpppG cap (%)
100
80
60
40
20
0
A22(-) A22 C30 U46 168
Stalled TECs

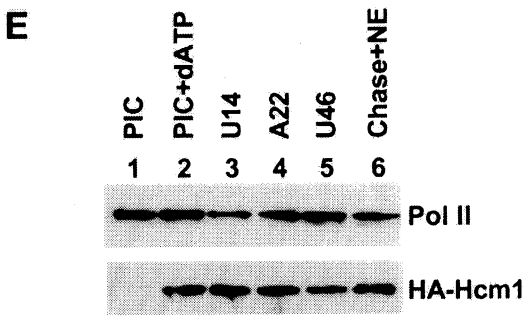
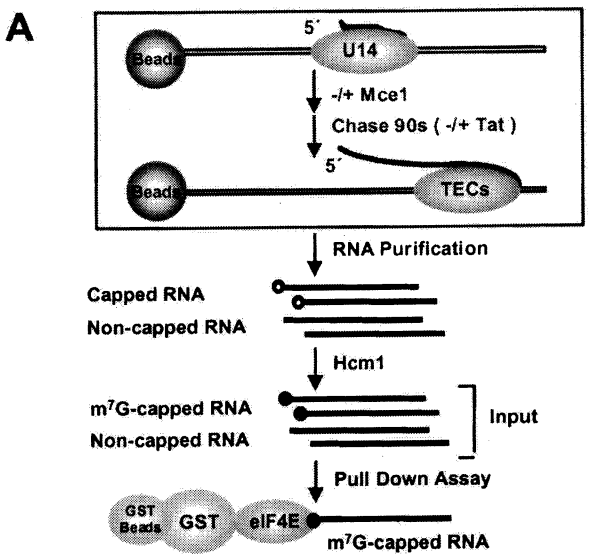
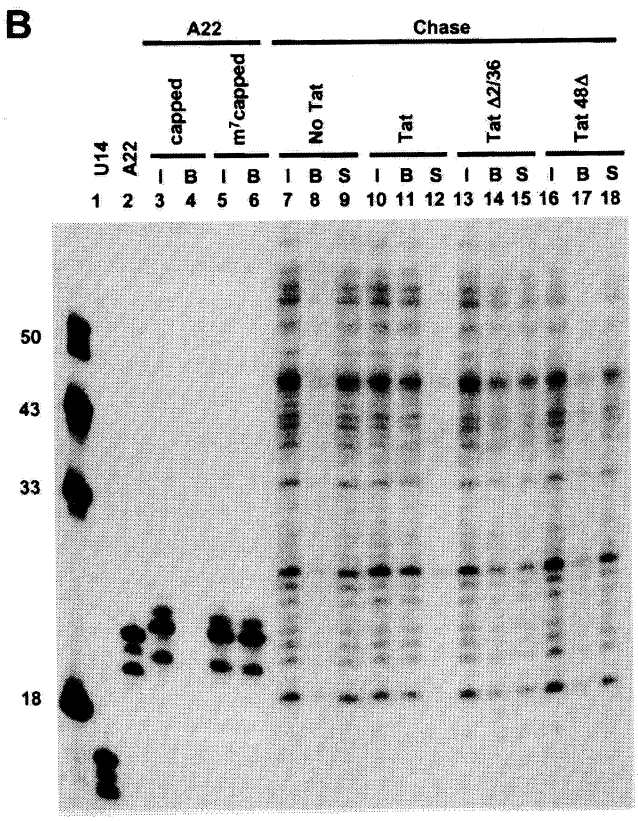

E
PIC
PIC+dATP
U14
A22
U46
Chase+NE
1 2 3 4 5 6
Pol II
HA-Hcm1

Figure 6. Tat stimulates cotranscriptional capping by Mce1. (A) Nascent RNAs were labeled internally with ^{32}P-CMP during stepwise transcription. TECs stalled at U14 were incubated with Mce1. After washing to removing unbound Mce1, TECs were chased for 90s in the absence or presence of the Tat proteins (10 pmol) as indicated above the lanes in (B). RNAs isolated from TECs were incubated with Hcm1 as described in Figure 5. m7G-capped RNAs were isolated by eIF4E affinity chromatography. (B) PAGE analysis of the input RNA (I), the RNAs bound by the GST-eIF4E fusion proteins (B), and the uncapped RNAs in the supernatant phase (S). To evaluate the specificity and efficiency of the capture assay, G-capped transcripts were isolated from stalled A22 TECs treated with Mce1 and analyzed by eIF4E affinity chromatography (lanes 3 and 4) in parallel with m7G-capped transcripts isolated from stalled A22 TECs treated with Mce1 and Hcm1 plus AdoMet (lanes 5 and 6). Other samples are as follows: transcripts isolated from stalled U14 and A22 TECs (lanes 1 and 2); transcripts isolated from U14 TECs after a chase with 4 NTPs in the absence (lanes 7–9) or presence of Tat (lanes 10–12), Tat2/36 (lanes 13–15), or Tat48 (lanes 16–18). (C) The percent of the input RNA sample that bound to eIF4E was quantitated by scanning the gel in (B) with a phosphorimager.


A
5´
U14
Beads
-/+ Mce1
Chase 90s (-/+ Tat)
5´
Beads
TECs
RNA Purification
Capped RNA
Non-capped RNA
Hcm1
m7G-capped RNA
Non-capped RNA
Input
Pull Down Assay
GST Beads
GST
eIF4E
m7G-capped RNA


B
A22
Chase
U14
A22
capped
m7capped
No Tat
Tat
Tat Δ2/36
Tat 48Δ
I B I B I B S I B S I B S I B S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
50
43
33
18

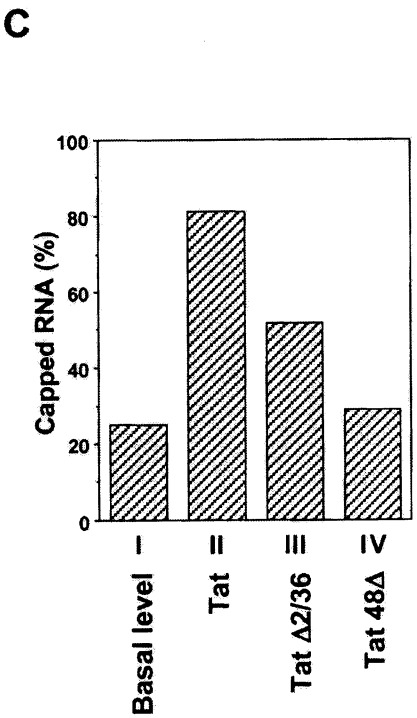
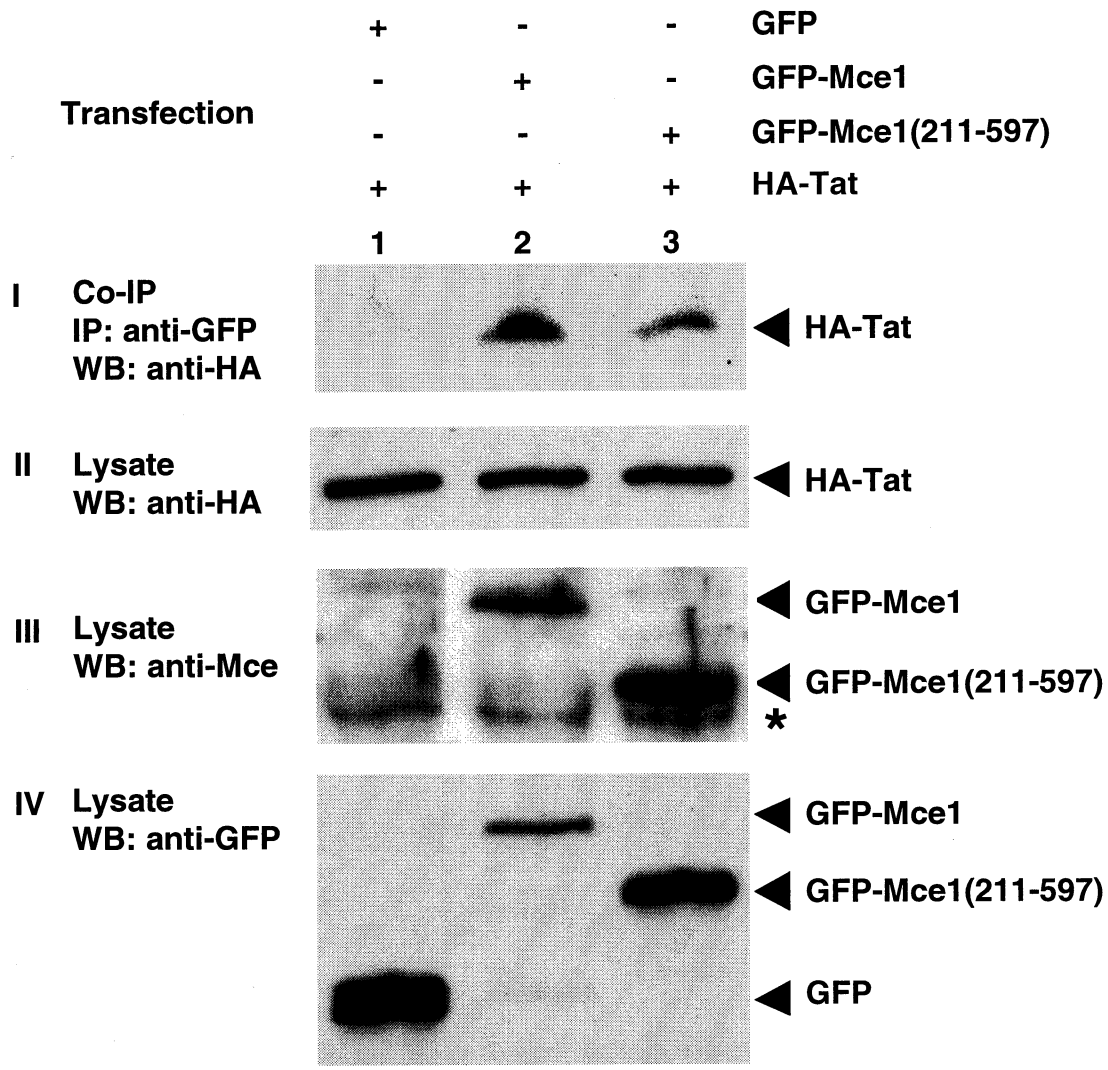

C
Capped RNA (%)
100
80
60
40
20
0
I
II
III
IV
Basal level
Tat
Tat Δ2/36
Tat 48Δ

Figure 7. Interaction of Tat and Mce1 *in vivo*. HL3T1 cells were transfected with plasmids expressing HA-Tat and either GFP or the indicated GFP-Mce1 fusion proteins. Cells were lysed with RIPA buffer at 48 hr posttransfection. (I) The cell lysates were immunoprecipitated (IP) with anti-GFP antibody, the precipitates were analyzed by SDS-PAGE, and the gel contents were probed by Western blotting (WB) with anti-HA antibody. (II–IV) Unfractionated lysate (60 _g of protein) was resolved by SDS-PAGE and probed by Western blotting using antibodies against HA (II), Mce1 (III), and GFP (IV). The identities of the immunoreactive polypeptides are indicated by arrowheads on the right. The asterisk (*) in (III) indicates the endogenous untagged Mce1 polypeptide.


Transfection
+ - - GFP
- + - GFP-Mce1
- - + GFP-Mce1(211-597)
+ + + HA-Tat
1 2 3
I Co-IP
IP: anti-GFP
WB: anti-HA
HA-Tat
II Lysate
WB: anti-HA
HA-Tat
III Lysate
WB: anti-Mce
GFP-Mce1
GFP-Mce1(211-597)
*
IV Lysate
WB: anti-GFP
GFP-Mce1
GFP-Mce1(211-597)
GFP

CHAPTER IV

HUMAN TRANSCRIPTION ELONGATION FACTOR (CDK9/CYCLINT1): A NEW THERAPEUTIC TARGET FOR AIDS AND CANCER?

ABSTRACT

Regulation of mRNA transcription plays a central role in mammalian cell growth and development. A large number of genes in human and other eukaryotic cells, as well as their viruses, are specifically regulated at the level of transcription elongation. P-TEFb, a positive elongation factor composed of two subunits, CDK9 and Cyclin T1 (CycT1), allows the transition to productive elongation, producing longer mRNAs. Replication of human immunodeficiency virus type 1 (HIV-1) requires Tat protein, which activates elongation of RNA polymerase II transcription at the HIV-1 promoter by interacting with human CycT1 (hCycT1). The trans-activation domain of Tat binds directly to the hCycT1 subunit of P-TEFb and induces loop sequence-specific binding of P-TEFb onto nascent HIV-1 trans-activation responsive (TAR) RNA. Using RNA interference to specifically degrade mRNA for hCycT1 or CDK9, we show here that down regulation of P-TEFb expression in HeLa cells achieved without causing major toxic or lethal effects and can control Tat transactivation and HIV replication in host cells. In addition, we have used high-density oligonucleotide arrays to determine the effect of P-TEFb knockdown on global gene expression. Of 44,928 human genes analyzed, 25 were down-regulated and known or likely to be involved in controlling and mediating cell proliferation and differentiation. Our results provide new insight into P-TEFb function, its potent role in early embryonic development and strong evidence that P-TEFb is a new target for developing AIDS and cancer therapies.

INTRODUCTION

Regulation of mRNA transcription plays a central role in mammalian cell growth and development. A large number of genes in human and other eukaryotic cells, as well as their viruses, are specifically regulated at the level of transcription elongation. Among these genes are several protooncogenes (*c-myc*, *c-myb, c-fos*), the gene for adenosine deaminase, a collection of stress response genes, and HIV-1 and HIV-2. During elongation, RNA polymerase II (RNA pol II) can pause, get arrested, pass through terminator sequences, or terminate transcription. Recent studies have uncovered that shortly after initiation, RNA pol II faces a barrier of negative transcription elongation factors (N-TEF) and enters abortive elongation (181). Positive transcription elongation factors (P-TEF) lower the barrier of N-TEF and help RNA pol II escape from this transition phase, which could lead to premature termination of transcription (181). A positive elongation factor, P-TEFb, which is composed of two subunits (CDK9 and Cyclin T1 (71)) allows the transition to productive elongation, producing longer mRNA transcripts (181).

One elegant example of transcription elongation control is the mechanism of HIV-1 gene expression (reviewed in 43, 108, 233). HIV-1 encodes a small regulatory protein, Tat, which is required for efficient transcription of viral genes. Tat enhances the processivity of RNA pol II elongation complexes that initiate in the HIV long terminal repeat (LTR) region. Tat activates transcription by binding to a highly structured RNA element, trans-activation responsive (TAR) RNA,

which is located at the 5'-end of nascent viral transcripts (183). Tat functions through TAR RNA to control an early transcription elongation step that is sensitive to protein kinase inhibitors and requires the carboxyl-terminal domain (CTD) of the large subunit of RNA pol II (108). Tat interacts with the human Cyclin T1 (hCycT1) subunit of P-TEFb and recruits the kinase complex to the pol II elongation complex (43, 181, 233, 252).

Recruitment of P-TEFb to TAR RNA has been proposed to be both necessary and sufficient to activate transcription elongation from the HIV-1 LTR promoter (16). Neither hCycT1 nor the P-TEFb complex bind TAR RNA in the absence of Tat, indicating that binding to RNA is highly co-operative for both Tat and P-TEFb (70, 252). Tat appears to contact residues in the carboxy-terminal boundary of hCycT1's cyclin domain that are not critical for binding of hCycT1 to CDK9 (16, 68, 71, 104, 255, 276). Mutagenesis studies have shown that the hCycT1 sequence containing amino acids 1-272 is sufficient to form complexes with Tat-TAR, CDK9, and to activate transcription by Tat (16, 68, 71, 104, 255, 276). Human Cyclin T1 residues 250-262 represent the Tat-TAR RNA recognition motif (TRM)(71). TRM is required but not sufficient to form the hCycT1-Tat-TAR RNA ternary complex. N-terminal residues in the cyclin box are also necessary. We have recently identified the sequence and structural determinants for high affinity hCycT1-Tat-TAR RNA ternary complex formation (188). Our results show that hCycT1 and Tat binding to TAR RNA is highly cooperative, with a capacity of 85%, Hill coefficient of 2.7 and a dissociation

constant (K_D) of 2.45 nM (188). In another study, we showed that residues 252-260 of hCycT1 interact with one side of the TAR RNA loop and enhance interaction of Tat residue K50 with the other side of the loop (189). Our results show that TAR RNA provides a scaffold for two protein partners to bind and assemble a regulatory switch in HIV replication.

The pol II CTD, Spt5 and Tat are intimately connected to the regulation of HIV gene expression. Human Spt5 and its binding partner hSpt4 comprise the transcription elongation regulatory factor DSIF (DRB sensitivity inducing factor)(242). DSIF binds to pol II and, in concert with NELF (negative elongation factor), represses elongation at promoter-proximal positions in the transcription unit (186, 261). Escape from the repressive effect of DSIF/NELF requires the action of P-TEFb, a DRB-sensitive protein kinase that phosphorylates both the pol II CTD and the Spt5 subunit of DSIF (103, 123, 178, 243). In previous studies we have analyzed the *in vitro* interactions of the CDK9 and hCycT1 subunits of P-TEFb and the HIV Tat protein with pol II at discrete functional stages of transcription from the HIV-1 LTR promoter. We found that P-TEFb is a component of the pol II preinitiation complex (PIC) and that it travels with the transcription elongation complex (TEC) as it moves along the HIV transcription unit unit (179). In contrast, DSIF and NELF are not present in the PIC, but associate with the TEC at promoter proximal positions and then travel with the TECs down the template (178). Our current model for understanding HIV-1 gene regulation is depicted in Figure 1A.

To understand the cellular functions of P-TEFb and to test the notion that HIV infectivity can be modulated by targeting host proteins such as P-TEFb without causing major toxic or lethal effects, we used RNA interference (RNAi) to specifically degrade mRNA for hCycT1 or CDK9. RNAi is a remarkably efficient process whereby double-stranded RNA (dsRNA) induces the sequence-specific degradation of homologous mRNA in animals and plant cells (99, 206). In mammalian cells, RNAi can be triggered by 21-nucleotide (nt) duplexes of small interfering RNA (siRNA)(28, 55). Our results show that specific down-regulation of P-TEFb expression in human cells can be achieved without causing major toxic or lethal effects and can be used to control Tat transactivation and HIV replication in host cells. To determine the effect of P-TEFb knockdown on global gene expression, we used high-density oligonucleotide arrays to analyze the expression of 44,928 human genes. We observed down-regulation of genes that are known or likely to be involved in controlling and mediating cell proliferation and differentiation. The data presented in this chapter provide new insight into P-TEFb function and its role in early embryonic development and offer strong evidence that P-TEFb is a new target for developing AIDS and cancer therapies.

MATERIALS AND METHODS

siRNA preparation. 21-nt ds RNAs were chemically synthesized as 2' bis(acetoxyethoxy)-methyl ether-protected oligos by Dharmacon (Lafayette, CO). Synthetic oligonucleotides were deprotected, annealed and purified according to the manufacturer's recommendation. Successful duplex formation was confirmed by 20% non-denaturing polyacrylamide gel electrophoresis (PAGE). All siRNAs were stored in DEPC (0.1% diethyl pyrocarbonate)-treated water at –80°C. Sequences of siRNA duplexes were designed according to the manufacturer's recommendation and subjected to a BLAST search against the NCBI EST library to ensure that they targeted only the desired genes. siRNA sequences used in our experiments were: hCycT1 ds (5'-UCCCUUCCUGAUACUAGAAdTdT-3');

hCycT1 mm (5'-UCCCUUCCGUAUACUAGAAdTdT-3');

CDK9 ds (5'-CCAAAGCUUCCCCCUAUAAdTdT-3');

CDK9 mm (5'-CCAAAGCUCUCCCCUAUAAdTdT-3');

CDK7 ds (5'-UUGGUCUCCUUGAUGCUUUdTdT-3');

Tat ds (5'-GAAACGUAGACAGCGCAGAdTdT-3');

GFP ds (5'-GCAGCACGACUUCUUCAAGdTdT-3'); and

RFP ds (5'-GUGGGAGCGCGUGAUGAACdTdT-3').

Underlined residues represent the mismatched sequence to their targets.

Culture and transfection of cells. HeLa cells were maintained at 37°C in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS), 100 unit/ml penicillin and 100 μg/ml streptomycin

(Invitrogen). Magi (multinucleate activation of galactosidase indicator) cells harboring the endogenous HIV LTR-β-gal gene were maintained at 37°C in DMEM, supplemented with 10% FBS, 0.2 mg/ml Geneticin (G418) and 0.1 mg/ml hygromycin B (Roche Molecular Biochemicals). Cells were regularly passaged at sub-confluence and plated at 70% confluency 16 h before transfection. Lipofectamine (Invitrogen)-mediated transient cotransfections of reporter plasmids and siRNAs were performed in duplicate 6-well plates (Falcon) as described by the manufacturer for adherent cell lines. A standard transfection mixture containing 100 nM siRNA and 10 μl lipofectamine in 1 ml serum-reduced OPTI-MEM (Invitrogen) was added to each well. Cells were incubated in transfection mixture for 6 h and further cultured in antibiotic-free DMEM. For Western blot analysis at various time intervals, the transfected cells were washed twice with phosphate buffered saline (PBS, Invitrogen), flash frozen in liquid nitrogen, and stored at -80°C for analysis. For *in vivo* assays of Tat-mediated transactivation at 48 h post transfection, Magi cells were subjected to β-gal staining directly or flash frozen in liquid nitrogen and stored at -80°C for β-galactosidase assay as described below.

Western blotting. Cells treated with siRNA were harvested as described above and lysed in ice-cold reporter lysis buffer (Promega) containing protease inhibitor (complete, EDTA-free, 1 tablet/10 ml buffer, Roche Molecular Biochemicals). After clearing the resulting lysates by centrifugation, protein in clear lysates was quantified by Dc protein assay kit (Bio-Rad). Proteins in 60 μg

of total cell lysate were resolved by 10% SDS-PAGE, transferred onto a polyvinylidene difluoride membrane (PVDF, Bio-Rad), and immuno-blotted with antibodies against hCycT1 and CDK9 (Santa Cruz). Protein content was visualized with a BM Chemiluminescence Blotting Kit (Roche Molecular Biochemicals). The blots were exposed to x-ray film (Kodak MR-1) for various times (30 s to 5 min).

RT-PCR for amplification of hCycT1 and CDK9 mRNA. Total cellular mRNA was prepared from HeLa cells with or without hCycT1/CDK9 siRNA treatment using a Qiagen RNA mini kit, followed by an oligotex mRNA mini kit (Qiagen). RT-PCR was performed using a SuperScript One-Step RT-PCR kit with platinum *Taq* (Invitrogen) and 40 cycles of amplification. Each RT-PCR reaction included 100 ng total cellular mRNA, gene-specific primer sets for hCycT1 and CDK9 amplification (0.5 μM for each primer), 200 μM dNTP, 1.2 mM $MgSO_4$ and 1U of RT/platinum *Taq* mix. Primer sets for hCycT1 produced 2178 bp products, while CDK9 primer sets produced 1116 bp products. RT-PCR products were resolved in 1% agarose gel and viewed by ethidium bromide staining.

Plasmids harboring the HIV-1 Tat sequence . pTat-RFP plasmids were constructed by fusing the DNA sequence of HIV-1 Tat with DNA sequences of DsRed1-N1, harboring coral (*Discosoma spp.*)-derived red fluorescent protein (RFP), per the manufacturer's recommendation (Clontech). Cytomegalovirus promoter control drove the expression of Tat-RFP fusion proteins, which were

easily visualized in living cells by fluorescence microscopy (Zeiss). Expression of Tat-RFP fusion proteins was also quantified by directly exciting the RPF fluorophore in clear cell lysates and measuring fluorescence, as described below.

β-galactosidase staining of cells. Magi cells were transfected with Tat-containing plasmids in the absence or presence of siRNAs. At 48h post transfection, cells were washed twice with PBS and fixed 5 min in fixative (1% formaldehyde and 0.2% glutaraldehyde in PBS) at room temperature. After washing twice with PBS, cells were covered with staining solution (PBS containing 4 mM potassium ferrocyanide, 4mM potassium ferricyanide, 2 mM $MgCl_2$ and 0.4 mg/ml X-gal [Promega]) and incubated at 37°C for exactly 50 min. Plates were washed twice with PBS. Cell counts represent number of β-gal positive (blue) cells per 100-power field.

β-Galactosidase enzyme assay. Magi cells were transfected with Tat-containing plasmids in the absence or presence of siRNAs. At 48 h post transfection, cells were harvested and clear cell lysates were prepared and quantified as described above. Total cell lysate (120 μg) in reporter lysis buffer (150 μl) was subjected to standard β-galactosidase assay by adding 150 μl 2X β-galactosidase assay buffer (Promega) and incubating at 37°C for 30 min. The reactions were stopped by adding 500 μl 1M sodium carbonate and briefly vortexing. Absorbance was read immediately at 420nm. The same amount of cell lysate was subjected to fluorescence measurements on a PTI (Photon Technology International) fluorescence spectrophotometer, with slit widths set at

4 nm for both excitation and emission wavelengths. All experiments were carried out at room temperature. Fluorescence of Tat-RFP in the cell lysate was detected by exciting at 568 nm and recording the emission spectrum from 588 nm to 650 nm; the spectrum peak at 583 nm represents the maximum fluorescence intensity of Tat-RFP. Tat transactivation was determined by calculating the ratio of β-galactosidase activity (absorbance at 420 nm) of the pTat-RFP transfected cells to that of cells without pTat-RFP plasmid treatment. The inhibitory effect of siRNA treatment was determined by normalizing Tat-transactivation activity to the amount of Tat-RFP protein (represented by RFP fluorescence intensity) in the presence and absence of siRNA.

***In vivo* fluorescence analysis.** pEGFP-C1 reporter plasmids (1 μg) and siRNA (100 nM) were cotransfected into HeLa cells by lipofectamine as described above, except that cells were cultured on 35 mm plates with glass bottoms (MatTek Corporation, Ashland MA) instead of standard 6-well plates. Fluorescence in living cells was visualized 50 h post transfection by conventional fluorescence microscopy (Zeiss). For GFP fluorescence detection, FITC filter was used.

GeneChip® experiments. HeLa cells were transfected with hCycT1/CDK9 ds siRNA. For comparison, cells treated with nonrelated ds siRNA (RFP ds) were used as control. At 48 h post transcription, total cellular mRNA was prepared from HeLa cells with or without hCycT1/CDK9 ds siRNA treatment using a Qiagen RNA mini kit followed by oligotex mRNA mini kit. Double-

stranded cDNAs were synthesized from 2 μg total mRNA using the Superscript Choice System for cDNA synthesis (Invitrogen) with the T7-(dT)24 primer following the manufacturer's recommendations. cDNAs were cleaned up by phase lock gel (PLG) (Brinkman Instrument)-phenol/chloroform extraction and concentrated by ethanol precipitation. Biotin-labeled cRNA was synthesized from cDNA by *in vitro* transcription using the Bioarray HighYield RNA transcript Labeling Kit (Affymetrix) following vendor's recommendation. *In vitro* transcription products were cleaned up using RNeasy spin columns (Qiagen) and fragmented into 35-200 base units by metal-induced hydrolysis in fragmentation buffer (40mM Tris-acetate, pH 8.1, 100mM KOAc, 30mM MgOAc). Fragmented cRNA was then subjected to Affymetrix Human Genome U133A and U133B GeneChip® sets in hybridization buffer (100mM MES, 1M NaCl, 20mM EDTA, 0.01% Tween-20). All hybridization, washing, staining and scanning procedures were performed in the UMASS Medical School Genomics Core Facility following the manufacturer's protocols. GeneChip® images were analyzed with Affymetrix Microarray Suite V5.0 and Affymetrix Data Mining Tool V3.0. The HG-U133 GeneChip® includes HG-133A and HG-133B sets of arrays containing 22,283 and 22,645 genes, respectively. Each gene was represented by a probe set containing 11 probe pairs. 100 human maintenance genes on both arrays serve as a tool to normalize and scale the data prior to making comparisons. Signal intensities of all probe sets were scaled to a target value of 150. Results of Detection Call, Change Call and Signal Log Ratio were obtained by applying the

default parameters to statistical algorithms for both absolute and comparison analyses. For comparison analysis, data from cells treated with nonrelated ds siRNA (RFP ds) were used as baseline array. Each probe pair in a probe set is considered as having a potential vote in determining whether the measured transcript is detected (Present) or not detected (Absent). The Detection algorithm uses probe pair intensities to generate a Detection *p*-value and assign a Present, Marginal or Absent call. Only probe pairs in the baseline array assigned a Present call were considered for further analysis. The comparison analysis compares the difference values of each probe pair in the baseline array to its matching probe pair on the experiment array. A Change *p*-value is calculated indicating an increase, decrease or no change in gene expression. The Signal Log Ratio (base on 2) estimates the magnitude and direction of change of a transcript when two arrays are compared. Log ratio larger than 1 indicating more than 2 fold up-regulation while log ratio smaller than 1 indicating more than 2 fold down-regulation.

Magi infectivity assay. HeLa-CD4-LTR/β-gal indicator (Magi) cells (124) were plated in 24-well plates ($7.5x10^5$ cells per well) and transfected with siRNAs as previously described (105). siRNA (60 pmol) was transfected into cells using oligofectamine (2μl, Invitrogen) for 3 h in serum-free DMEM (GIBCO). Cells were rinsed twice and top-layered in 500 μl of DMEM-10% FBS. Sixteen hours after transfection, cells were trypsinized and seeded in 96-well microtiter plates ($4x10^4$

cells per well), incubated 3 h and infected. HIV-1 virions (normalized to RT activity in cpm) were added in doubling dilutions to duplicate wells. Thirty-six hours post infection, cells were harvested to quantify β-galactosidase activity.

RESULTS AND DISCUSSION

Specific knockdown of P-TEFb expression by siRNA in HeLa cells

We used RNAi to inhibit hCycT1 and CDK9 expression in cultured human (HeLa) cell lines. The short interfering RNA (siRNA) sequence targeting hCycT1 was from position 351 to 371 relative to the start codon, and the CDK9 siRNA sequence was from position 258 to 278 relative to the start codon. Using lipofectamine, we transfected HeLa cells with hCycT1 or CDK9 siRNA duplex, targeting either hCycT1 or CDK9. To analyze RNAi effects, we prepared lysates from siRNA duplex-treated cells at various times after transfection. Western blot experiments were carried out (Figure 1B) using anti-hCycT1 and anti-CDK9 antibodies. Analysis of immunoblotting experiments reveals that the siRNA targeting hCycT1 inhibited hCycT1 expression (Figure 1B, lanes 8-14, upper panel). siRNA targeting CDK9 was similarly specific against CDK9 expression (Figure 1B, lanes 15-21, lower panel). This RNAi effect depended on the presence of a 21-nt duplex siRNA harboring a sequence complementary to the target mRNA, but not on the antisense strand siRNA (data not shown) nor on an irrelevant control siRNA, which targeted a coral (*Discosoma spp.*)-derived red fluorescent protein (RFP) (Figure 1B, lanes 1-7). As a specificity control, cells were also transfected with mutant siRNAs (mismatched siRNA) of hCycT1 or CDK9, which have two nucleotide mismatches between the target mRNA and the antisense strand of siRNA at the putative cleavage site of the mRNA. Mutant

siRNAs showed no interference activity (Figure 1B, lanes 22-28 and 29-35), indicating the specificity of the RNAi effect.

Kinetics of P-TEFb mRNA interference in HeLa cells

RNA interference is a highly efficient process because a few dsRNA molecules are sufficient to inactivate a continuously transcribed target mRNA for long periods of time. Experiments have shown in plants and worms (38, 47, 78) that this inactivation can spread throughout the organism and is often heritable to the next generation. Mutations in genes encoding proteins related to RNA-dependent RNA polymerase (RdRP) affect RNAi-type processes in *Neurospora, Caenorhabditis elegans* and plants (38, 47, 136, 160, 217), and the involvement of RdRP in amplifying RNAi has been postulated (136).

Western blot experiments (Figure 1B) using anti-hCycT1 and anti-CDK9 antibodies also revealed the kinetics of gene suppression and persistence of RNAi in HeLa cells. Although RNAi can suppress expression of hCycT1 and CDK9 proteins up to 66 h post transfection, maximum activities were observed at 42-54 h, and inhibition by siRNAs did not persist. After reaching maximal activity at 42-54 h post transfection, RNA interference started to decrease at 54 h, with protein amount showing gradual recovery to normal levels between 66 to 90 h (3 to 4 days) post transfection (Figure 2). Similar phenomena were demonstrated in 293T cells, a human T lymphocyte cell line (data not shown). The recovery of target gene expression indicates that RNAi by exogenous siRNA duplexes does not last forever in mammalian cells. These findings suggest that the amplification

system driven by RdRP in plants and nematodes may not exist or has very little effect on siRNA-mediated gene silencing in mammalian cells.

Another intriguing finding revealed by this Western analysis is that treating cells with hCycT1 siRNA also down-regulated the level of CDK9 protein. The kinetics of CDK9 knockdown by hCycT1 siRNA shows a pattern similar to hCycT1 knockdown (Figure 1B, lanes 8-14, lower panel). No homologous sequence has been found between the CDK9 mRNA and hCycT1 siRNA used in this assay, suggesting that CDK9 knockdown by hCycT1 siRNA was not due to cleavage of CDK9 mRNA by the RNAi pathway. Two possibilities could explain this unexpected down-regulation. First, hCycT1 knockdown may affect the stability of CDK9, the kinase component of the P-TEFb complex. Second, since P-TEFb is a positive transcription factor, it is possible that the P-TEFb complex is required for transcription of the CDK9 gene. Thus, the down-regulation of hCycT1 by hCycT1 siRNA would also down-regulate intracellular CDK9 levels.

Specific down-regulation of P-TEFb by siRNA at the mRNA level and stability of CDK9

To determine the specificity of P-TEFb knockdown by siRNA at the mRNA level and to distinguish the two hypotheses proposed above, we performed RT-PCR to reveal the effect of siRNA on the level of mRNA involved in P-TEFb expression. As shown in Figure 1C, transfection of cells with siRNA duplex targeting hCycT1 (hCycT1 ds) significantly reduced hCycT1 expression (Figure

1C, lanes 1-7, upper panel), but had no effect on CDK9 mRNA (Figure 1C, lanes 1-7, lower panel). On the other hand, transfection of cells with siRNA duplex targeted to CDK9 (CDK9 ds) significantly interfered with the expression of CDK9, but not hCycT1 (Figure 1C, lanes 8-14). The siRNA duplex started to cause an RNAi effect as early as 6-18 h post transfection and gradually increased with time, peaking at 30 h, and decreased between 54-66 h. The time-dependent effect of siRNA indicates that siRNAs need to be processed or assembled into an active complex with cellular factors for effective RNA interference. A time lag was also seen between the degradation of target mRNA (starting at 6 h post siRNA transfection, as shown by semi-quantitative RT-PCR in Figure 1C) and the half-life of the existing protein expressed by the target gene, because protein levels didn't show any down-regulation until 18-30 h post siRNA transfection (Figure 1B). Combined with Western blot analysis (Figure 1B), semi-quantitative RT-PCR (Figure 1C) not only confirms the specific knockdown of P-TEFb by siRNA at the mRNA level, but also suggests that forming a complex with hCycT1 is a prerequisite for maintaining the stability of CDK9 proteins in living cells. Thus, hCycT1 siRNA down-regulated hCycT1 levels by the RNAi pathway, while down-regulating CDK9 levels by promoting its degradation without affecting its gene expression at the mRNA level.

hCycT1 and CDK9 knockdown can be achieved without causing major toxic or lethal effects

To analyze cell viability *in vivo* under siRNA treatment, pEGFP-C1 reporter plasmid (harboring enhanced green fluorescent protein [GFP]) and siRNAs were cotransfected into HeLa cells using lipofectamine. Reporter gene (GFP) expression, driven by cytomegalovirus (CMV) immediate early promoter, was monitored in living cells (Figure 3A, upper panels). Cellular morphology and density were monitored by phase contrast microscopy (Figure 3A, lower panels). GFP expression was not affected by hCycT1 or CDK9 knockdown (Figure 3A, compare panels a-c). Cells with P-TEFb knockdown had normal shape and growth rate. At 50 h post transfection, cell density reached ~90% to 100% confluency (Figure 3A, compare panels e-g).

For comparison, we transfected cells with siRNA targeting CDK7, a well-characterized kinase required for TFIIH, an essential transcription factor, to phosphorylate the CTD of RNA pol II at the step of promoter clearance during initiation of transcription. Kin28, a protein in *Saccharomyces cerevisiae* that is equivalent to CDK7 in mammals, is an essential gene product that phosphorylates Ser5 of the CTD YSPTSPS repeat region (128, 191, 199) and is required to recruit the mRNA capping enzyme to the transcription machinery (31, 152, 153, 269). CDK7 is a bifunctional enzyme in larger eukaryotes, promoting both CDK activation and transcription (86). Reduction of CDK7 levels by RNAi led to a lower reporter (GFP) expression and an arrest in cellular growth (Figure 3A, panel d). CDK7 knockdown cells were smaller than control cells and showed blebbing (Figure 3A, panel h).

We also analyzed cellular viability under various siRNA treatments. At various times after transfection, cell viability was assessed by trypan blue exclusion (see Experimental Procedures). Over a 66 h time course experiment, the rate of cell death in P-TEFb (hCycT1 or CDK9) knockdown cells was comparable to that in control cells with unrelated siRNA treatment, while CDK7 knockdown cells showed a significant increase in cell death (Figure 3B). These results indicate that transient P-TEFb knockdown can be achieved without causing major toxic or lethal effects to human cells, while a much more stringent threshold for CDK7 is required to maintain cell viability and growth.

hCycT1 and CDK9 RNAi inhibit HIV-1 Tat transactivation in human cells

A dominant paradigm for Tat up-regulation of HIV gene expression at the level of transcription elongation revolves around the ability of the Tat-TAR RNA complex to bind P-TEFb and stimulate phosphorylation of the CTD and Spt5, thereby overriding the elongation arrest elicited by DSIF and NELF (178, 181). To test whether siRNAs that targeted sequence elements of P-TEFb would specifically block Tat transactivation, we cotransfected Magi cells with the Tat expression construct pTat-RFP and hCycT1 or CDK9 ds siRNA or as controls, antisense RNA strands or mutant siRNAs. Magi, a HeLa cell line harboring a single copy of persistently transfected HIV-1 LTR-β-galactosidase gene, is programmed to express the CD4 receptor and the CCR5 coreceptor for HIV-1 (124). We confirmed that the HIV-1 Tat-RFP fusion protein was expressed under

control of the CMV early promoter by Western blot, using anti-RFP antibody (data not shown).

Tat-RFP strongly enhanced gene expression, which is under control of the HIV-1 LTR promoter in transfected Magi cells. Tat transactivation was determined by calculating the ratio of β-gal activity in pTat-RFP transfected cells to the activity in cells without pTat-RFP treatment. Inhibitory activity was determined by normalizing Tat-transactivation activity to the amount of Tat-RFP protein (represented by RFP fluorescence intensity as described in Experimental Procedures) in the presence and absence of siRNA. Under standard experimental conditions, Tat-RFP enhanced gene transactivation 20- to 25-fold (Figure 4C, lane 1). This activation was strongly inhibited by cotransfecting host Magi cells with the specific ds siRNAs targeting hCycT1 and CDK9 (Figure 4C, lanes 4 and 5), but not with antisense (as) RNA strands (Figure 4C, lanes 2 and 3), mutant (mm) siRNAs (Figure 4C, lanes 6 and 7), or an unrelated control siRNA (Figure 4C, lane 8). Specific RNA interference with hCycT1 and CDK9 expression in Magi cells was demonstrated by Western blot analysis (Figure 4A, lanes 3 and 4), and the inhibition of Tat transactivation correlated well with the knockdown of hCycT1 and CDK9 protein levels by the hCycT1 and CDK9 siRNAs (compare Figure 4C, lanes 4 and 5 to Figure 4A, lanes 3 and 4). As shown in Figure 4B, syncytia formation and LTR activation (represented by cells with blue β-gal staining) were reduced in hCycT1 ds siRNA-treated cells (Figure 4B, compare panels d to h and b to f). From these results, we conclude that

siRNA targeting P-TEFb can inhibit Tat-transactivation in human cells without causing major toxic or lethal effects.

hCycT1 and CDK9 RNAi inhibit HIV-1 infectivity

Can we inhibit HIV replication by targeting the human P-TEFb complex? To investigate this question, we transfected HeLa-CD4-LTR/β-gal (Magi) cells with homologous and mismatched siRNAs directed against hCycT1 or CDK9 and 16 h later infected the Magi cells with NL-GFP, an infectious molecular clone of HIV-1. HIV-1 Tat-mediated transactivation of the LTR led to β-gal production that was quantified 36 h post-infection. As shown in Figure 5, ds siRNA directed against hCycT1 or CDK9 inhibited viral infectivity. Doubling dilutions of the inoculums are consistent with an 8-fold decrease in viral infectivity. Control experiments using siRNA duplexes containing mismatched sequences (see Experimental Procedures) and an unrelated ds siRNA against the RFP sequence showed no antiviral activities. Consistent with publised studies (105), siRNA targeting GFP-Nef and Tat led to an 8-fold decrease in viral infectivity. No significant toxicity or cell death was observed during these experiments. These results demonstrate that HIV infectivity can be modulated by siRNAs targeting CycT1 or CDK9, both components of P-TEFb.

Genomewide analysis of gene expression in P-TEFb knockdown HeLa cells

To investigate the effects of P-TEFb knockdown on global gene expression, we isolated total mRNA from HeLa cells treated with and without siRNA directed against hCycT1 or CDK9 at 48h post transfection and analyzed the expression of various genes using the GeneChip® Human Genome U133 (HG-U133) from Affymetrix. The HG-U133 includes HG-133A and HG-133B sets of arrays containing 22,283 and 22,645 genes, respectively. Each gene was represented by a probe set containing 11 probe pairs. Another 100 human maintenance genes on both arrays serve as a tool to normalize and scale the data prior to making comparisons. For comparison analysis, data from cells treated with nonrelated ds siRNA (RFP ds) were used as baseline. The comparison analysis compares the difference values of each probe pair in the baseline array to its matching probe pair on the experiment array. A Change *p*-value is calculated indicating an increase, decrease or no change in gene expression. The Signal Log Ratio (base on 2) estimates the magnitude and direction of change of a transcript when two arrays are compared. Log ratio larger than 1 indicating more than 2 fold up-regulation while log ratio smaller than 1 indicating more than 2 fold down regulation (see Material and Methods for detail).

Genomewide analysis of the gene expression profile of P-TEFb knockdown cells revealed that 202 genes out of 44,928 were down-regulated. A complete list of down- and up-regulated genes with names and accession numbers is listed in Appendix (see Appendix Table 1). Of those 202 genes, 53 are known and were classified according to their function or protein product

activity (Figure 6A). It is widely accepted that P-TEFb is a positive transcription factor during RNA pol II elongation. Down-regulated genes in the P-TEFb knockdown background are presumably those required by P-TEFb for normal levels of expression, especially at the transcription elongation level.

P-TEFb regulates gene involved in controlling and mediating cell proliferation and differentiation

Broadly speaking, about 50% (25 out of 53) of the down-regulated genes are known or likely to be involved in controlling and mediating cell proliferation and differentiation. These genes can be further divided into three classes (Figure 6A, I – III). The first class (Figure 6A, I) includes genes directly linked to cellular proliferation and differentiation. Most of these genes belong to the protein tyrosine kinase (PTK) superfamily. PTKs catalyze phosphate transfer from ATP to tyrosine residues on protein substrates, activating numerous signaling pathways leading to cell proliferation, differentiation, migration, or metabolic changes and playing a prominent role in the control of a variety of cellular processes during embryonic development (96). Two classes of PTKs are affected by P-TEFb knockdown: transmembrane receptor protein tyrosine kinases (RTKs) and non-receptor tyrosine kinases (NRTKs). AXL receptor tyrosine kinase (AXL), discoidin domain receptor 1 (DDR1), epidermal growth factor receptor (EGFR) and fibroblast growth factor receptor (FGFR) belong to the RTKs, while cell adhesion kinase (CAK) belongs to the NRTKs. AXL, DDR1

and EGFR are required for the epithelial-mesenchymal transition during embryonic development, which allows multicellular organisms to get past the blastula stage (234).

Transforming growth factor beta (TGF-beta), also down-regulated by P-TEFb knockdown (Figure 6A, I), binds to another membrane receptor family with diverse functions during embryonic development and adult tissue homeostasis (5, 149). The genes for pre-T/NK cell-associated proteins (fasciculation and elongation protein zeta 2) are preferentially expressed in early stages of human T/NK cells and brain, suggesting that they play a role in early development (102, 254). Brain-derived neurotrophic factor (BDNF) has been implicated in activity-dependent plasticity of neuronal function and network arrangement (259).

A second class of genes affected by P-TEFb knockdown is functionally linked to the cell membrane and extracellular matrix (Figure 6A, II). Junction plakoglobulin (JUP) cooperates with beta-catenin and promotes epithelial growth and morphogenesis (87). Moesin (MSN), a plasma membrane protein associated with the underlying cytoskeleton, determines cell shape and participates in adhesion, motility and signal transduction pathways (19). This class of genes is also required for embryonic development, especially for angiogenesis and neuronal development. For example, integrin functions in epithelial cell organization and synaptogenesis during development of the CNS (100, 156). Epithelial membrane protein 1 (EMP1), which is highly expressed by immature

neurons in the embryonic stage, functions in neuronal differentiation and neurite outgrowth (257).

A third class of genes is involved in signal transduction (Figure 6A, III), the events downstream of membrane receptor activation. The inositol 1,4,5-triphosphate receptor (ITPR1) responds to G-protein-coupled receptor activation (115), while mitogen-activated protein kinase 6 (MAPK6) propagates cell proliferation/differentiation signals from receptor tyrosine kinase activation (96, 115). Non-receptor type protein tyrosine phosphatase regulates phosphotyrosine signalling events during complex ectodermal-mesenchymal interactions that regulate mammalian limb development (4, 196). Protein phosphatase 1 (PPP1CB) regulates the phosphorylation status of anti-apoptotic and pro-apoptotic proteins and their cellular activity in the apoptosis cascade (127). Dual specificity phosphatase 2 (DUSP2) participates in the regulation of intracellular signal transduction mediated by MAP kinases (266).

There is cross talk among these three gene classes (Figure 6A, I, II, and III). For example, collagen (type III, alpha 1), an extracellular matrix membrane protein can be a direct ligand for tyrosine kinase receptors (class I) and integrin (class II)(209). Coordinated regulation of PTK (class I) and protein phosphatase (class II) also acts at different steps in a common signal transduction process (11). Non-receptor type protein tyrosine phosphatase (class III) positively regulates BDNF-promoted (class I) survival of ventral mesencephalic dopaminergic neurons (229). FGFR (class I), EMP1 (class II), pre-T/NK cell-

associated protein (fasciculation and elongation protein zeta 2; class I) and integrin (class II) all participate in neuronal development. Nedasin (S form, class X), which is predominantly expressed in the brain, plays a role in the formation and structural changes of synapses during neuronal development by modifying clustering of neurotransmitter receptors at synaptic sites (130). Down-regulation of the genes coding for these last five proteins by P-TEFb knockdown indicates an important role for P-TEFb during neuronal development.

From the expression profile of these three classes of genes, we conclude that P-TEFb regulates expression of genes involved in controlling and mediating cell proliferation and differentiation. We suggested that P-TEFb is essential for embryonic gene expression and development, while knockdown of its subunits (hCycT1 and CDK9) may not cause major toxic or lethal effects at the adult stage. Although P-TEFb knockdown or knockout has never been explored in mammalian embryonic tissue, it has been proposed that the P-TEFb complex is required for global gene expression during embryonic development of *C. elegans* (208). Knockdown of CDK9 or CycT1 siRNA in *C. elegans* embryos inhibits transcription of embryonic genes, including the MAP kinase pathway and cell cycle-related genes (208).

The role of P-TEFb in human cancers

An intriguing finding is that genes linked to embryonic development and showing down-regulation in P-TEFb knockdown cells (as described above) also

participate in tumorogenesis and metastasis. Dysfunction of protein tyrosine kinases (Figure 6A, I) or aberrations in key components of the signaling pathways they activate can lead to severe pathologies such as cancer, diabetes and cardiovascular disease. For example, overexpression of EGFR (Figure 6A, I) has been implicated in mammary carcinomas, squamous carcinomas and glioblastomas (197). AXL, another receptor tyrosine kinase (Figure 6A, I), was originally identified with oncogenic potential and transforming activity in myeloid leukemia cells (20). Elevated TGF-beta levels can contribute to tumor progression and metastasis (149). Lysyl oxidase (LOX class II), an extracellular matrix remodeling enzyme, is up-regulated in prostatic tumor, cutaneous and uveal cell lines (126). Down-regulating these genes by P-TEFb knockdown could become a new therapeutic strategy for inhibiting tumorogenesis and metastasis.

Genes involved in mediating progression through the cell cycle and as checkpoints in cancer were regulated by P-TEFb (Figure 6A, IV). Cyclin G1 is the downstream target of the P53 pathway and plays a role in G2/M arrest, damage recovery and growth promotion after cellular stress (125). Cyclin D, a cell-cycle requlatory protein essential for G1/S transition, has been identified as a potential transforming gene in lymphoma (159). Misregulation of the activity of its partner, CDK4/6, by overexpression of Cyclin D leads to hyperproliferative defects and tumor progression (169). Several marker genes in cancer cells (class V) are also regulated by P-TEFb. For example, breast cancer-specific protein 1 (BCSG1) is overexpressed in advanced, infiltrating breast cancer and colorectal tumors

(140). Another example is soluble urokinase plasminogen activator receptor (SUPAR), which is present in high concentrations in cystic fluid form ovarian cancer, tumor tissue of primary breast cancer, and gynecological cancer (190, 244). Although the functions of these marker genes are still unknown, their high correlation with cancer has been used for prognosis in cancer therapy. The down-regulation of cyclin D and cancer marker genes by P-TEFb knockdown offer more encouraging evidence to support our proposal that P-TEFb could become a new and potent target for cancer therapy.

Another strong line of evidence in support of P-TEFb's role in cellular proliferation comes from clinical studies. Flavopiridol (L86-8275, HMR1275) is a cyclin-dependent kinase (CDK) inhibitor being tested in clinical trials as a potential anti-cancer drug (117, 205) but its target and the mechanism for its antiproliferative effects are unknown. Recent evidence indicates that flavopiridol inactivates P-TEFb and inhibits HIV-1 transcription (25, 26). Based on our results showing that P-TEFb knockdown in human cells down-regulates genes involved in controlling proliferation and differentiation, we propose that the antiproliferative effects of flavopiridol are due to P-TEFb inhibition.

Links between P-TEFb and stress responses

Another interesting group of genes down-regulated by P-TEFb are those involved in responding to stress or oxidant-mediated regulation (Figure 6A, VII). There are three known major redox signaling systems in eukaryotic cells, namely

glutathione/glutathione reductase, thioredoxin/thioredoxin reductase and glutaredoxin (GLRX, thioltransferase)(180). SH3 domain-binding glutamine-rich 3-like protein (SH3BGRL3) also belongs to the thioredoxin family. Glutathione S-transferase M4 (GSTM4) is involved in detoxifying reactive electrophiles, such as drug or foreign compounds, by catalyzing their reaction with glutathione (GSH)(42). Oxidant-mediated regulation by GSH systems plays a direct role in cellular signaling through thiol-disulfide exchange reactions with membrane-bound receptor proteins, transcription factors, and regulatory proteins in the cell (42). During stress responses, redox regulation has an important function in biological events such as DNA synthesis, enzyme activation, gene expression and cell cycle regulation. Down-regulation of genes involved in these events by P-TEFb knockdown in cells indicates an important role for P-TEFb in stress responses, especially oxidant-mediated regulation of cell proliferation. Moreover, glutathione (GSH) has been linked to multi-drug resistance (42), promising evidence for P-TEFb as a therapeutic target in cancer research.

Unlike the covalent modifications in redox regulation described above, the BAG family of modulating proteins (Figure 6A, VII) functions through alterations in conformation and influences signal transduction through non-covalent post-translational modifications (230). The BAG family molecular chaperone regulator-2 (BAG-2) belongs to this family, which contains an evolutionarily conserved "BAG domain" that allows its members to interact with and regulate the Hsp 70 (heat shock protein 70) family of molecular chaperones (230). Like Hsp 70, BAG-

family proteins have been reported to mediate the physiological stress signaling pathway that regulates cell division, death, migration and differentiation (230). P-TEFb has been shown to be recruited to heat shock loci in *Drosophila melongaster* and to co-localize with Hsp 70 and Hsp 90 upon heat shock stress (138). Down-regulation of the BAG-2 gene in P-TEFb knockdown cells indicates an important role for P-TEFb in regulating Hsp70 molecular chaperones in human cells. CDK9 itself has been proposed to form complexes with Hsp 70 and Hsp90/cdc37, thereby involving this chaperone-dependent pathway in the stablization/folding of CDK9 as well as the assembly of an active CDK9/CycT1 complex (167).

Role of P-TEFb in Metabolism and Biosynthesis

Several genes involved in metabolism and biosynthesis are also regulated by P-TEFb (Figure 6A, IX). Human iron-sulfur protein or ferredoxin (FDX1) serves as an electron transport intermediary for mitochondrial cytochrome P450 involved in steroid, vitamin D, and bile acid metabolism (135). Low-density lipoprotein receptor-related protein 5 (LRP5) contains conserved modules characteristic of the low-density lipoprotein (LDL) receptor family, genetically associated with Type 1 diabetes (61). Because alterations in LRP5 expression may be responsible for susceptibility to diabetes, LRP5 may therefore be a potential target for therapeutic intervention. The vacuolar (H+)-ATPases (or V-ATPases) function in the acidification of intracellular compartments in eukaryotic

cells. Eukaryotic translation elongation factor 1, alpha 2 isoform (EEF1A2), a key factor in protein synthesis, has been shown to have oncogenic properties: it enhances focus formation, allows anchorage-independent growth and decreases doubling time of fibroblasts (2). EEF1A2 is amplified in 25% of primary ovarian tumors, its expression makes NIH3T3 fibroblasts tumorogenic and it increases the growth rate of ovarian carcinoma cells (2).

Role of P-TEFb in Cell Cycle Regulation

Also affected by P-TEFb knockdown were genes involved in cell cycle regulation (Figure 6A, VI), which in turn controls proliferation and differentiation. The retinoblastoma (RB) protein regulates both the cell cycle and tissue-specific transcription by modulating the activity of its associated factors (143). Efforts to identify such cellular targets have led to the isolation of two novel proteins, RB-associated protein (RBP21) and RB- and p300-binding protein EID-1 (an E1A-like inhibitor of differentiation)(143). Although its cellular function is still unclear, RBP21 is widely expressed in various human tissues and cancer cell lines. EID-1 is a potent inhibitor of differentiation, an activity that has been linked to its ability to inhibit p300 (and the highly related molecule, CREB-binding protein, or CBP) acetylation of histones (158). EID-1, which is rapidly degraded by the proteasome as cells exit the cell cycle, may act at a nodal point that couples exit from the cell cycle to transcriptional activation of genes required for differentiation (158). Regulation of EID-1 expression by P-TEFb knockdown provides evidence

that P-TEFb is also involved in cell cycle regulation, especially RB-linked regulation of proliferation and differentiation.

Up-regulated genes

Four major classes of up-regulated genes were observed, including those involved in signal transduction (Figure 6B, I), transcription regulation (Figure 6B, II), cell cycle regulation (Figure 6B, IV) and metabolism and biosynthesis (Figure 6B, VI). This up-regulation may not be a direct effect of P-TEFb knockdown but rather a secondary or correlated effect, to which the cell responds by overexpressing certain genes to compensate the loss of function of genes modulated by P-TEFb. The up-regulation of genes involved in signal transduction, transcription and cell cycle regulation (Figure 6B, I, II, and IV) suggests that these genes could complement cellular functions in P-TEFb knockdown cells or play a role in overcoming effects of down-regulated genes. An interesting correlation was observed between up-regulation and the functions of two classes of genes, i.e., transcription regulation and metabolism/biosynthesis (Figure 6B, III and VI). Translation regulator gene *eIF-2*, which controls a signaling pathway to activate genes involved in amino acid biosynthesis (85) was up-regulated.

CONCLUSIONS

Our results raise new and intriguing questions for P-TEFb function and its role in regulating gene expression. First, is P-TEFb required for cellular viability? Recent evidence indicates that P-TEFb is essential for transcription in *C. elegans* embryos (208). P-TEFb may also be essential during mammalian embryonic development to control transcription elongation and ensure fidelity of mRNA production. This requirement for P-TEFb could explain why it has been difficult to create P-TEFb knockout mice. At later stages in the life cycle, knockdown of P-TEFb may be achieved without causing major toxic or lethal effects. Second, does P-TEFb regulate expression of genes involved in proliferation and cancer? Our results support this notion because ds siRNA directed against hCycT1 or CDK9 lowered levels of proliferation and cancer genes, indicating that expression of these genes is regulated by P-TEFb and negative elongation factors. It is tempting to postulate that HIV has evolved a mechanism of gene expression similar to the expression of genes involved in human cancers. Identifying how these genes and signaling pathways are regulated by P-TEFb function presents a significant and exciting future challenge.

Figure 1. Specific down-regulation of P-TEFb expression by RNAi. (A) Model for HIV Tat transactivation involving the human P-TEFb (CyclinT1/CDK9) complex. Human RNA pol II initiates transcription from HIV promoter DNA, and TFIIH kinase assists in promoter clearance steps. TFIIH leaves the elongation complex when 30-36 nt mRNA is transcribed. Elongation is inefficient in the absence of Tat protein. Tat recruits the P-TEFb kinase complex to TAR RNA by contacting the RNA and CycT1 components of the complex. CDK9 kinase then phosphorylates RNA pol II and Spt5, a component of negative elongation factor (N-TEF) in the elongation complex, which leads to processive transcription elongation. **(B) Analysis of specific hCycT1 and CDK9 RNAi activities by Western blotting.** HeLa cells were transfected with double-stranded (ds) siRNAs targeting RFP (control, lanes 1-7), hCycT1 (lanes 8-14), or CDK9 (lanes 15-21). Cells were also transfected with mutant siRNAs (hCycT1 mismatch [lanes 22-28] or CDK9 mismatch [lanes 29-35]) having 2 nucleotide mismatches between the target mRNA and the antisense strand of siRNA at the hypothetical cleavage site of the mRNA. Cells were harvested at various times post transfection, their protein content resolved on 10% SDS-PAGE, transferred onto PVDF membranes, and immunoblotted with antibodies against hCycT1 and CDK9. **(C) Analysis of specific hCycT1 and CDK9 RNAi activities by RT-PCR.** HeLa cells were transfected with hCycT1 ds siRNA (lanes 1-7) and CDK9 ds siRNA (lanes 8-14), harvested at various times after transfection and mRNAs extracted. One-step RT-PCR was performed, setting the specific primer for

hCycT1 and CDK9 amplification (see Experimental Procedure for details). RT-PCR products were resolved in 1% agarose gel and viewed by ethidium bromide staining.

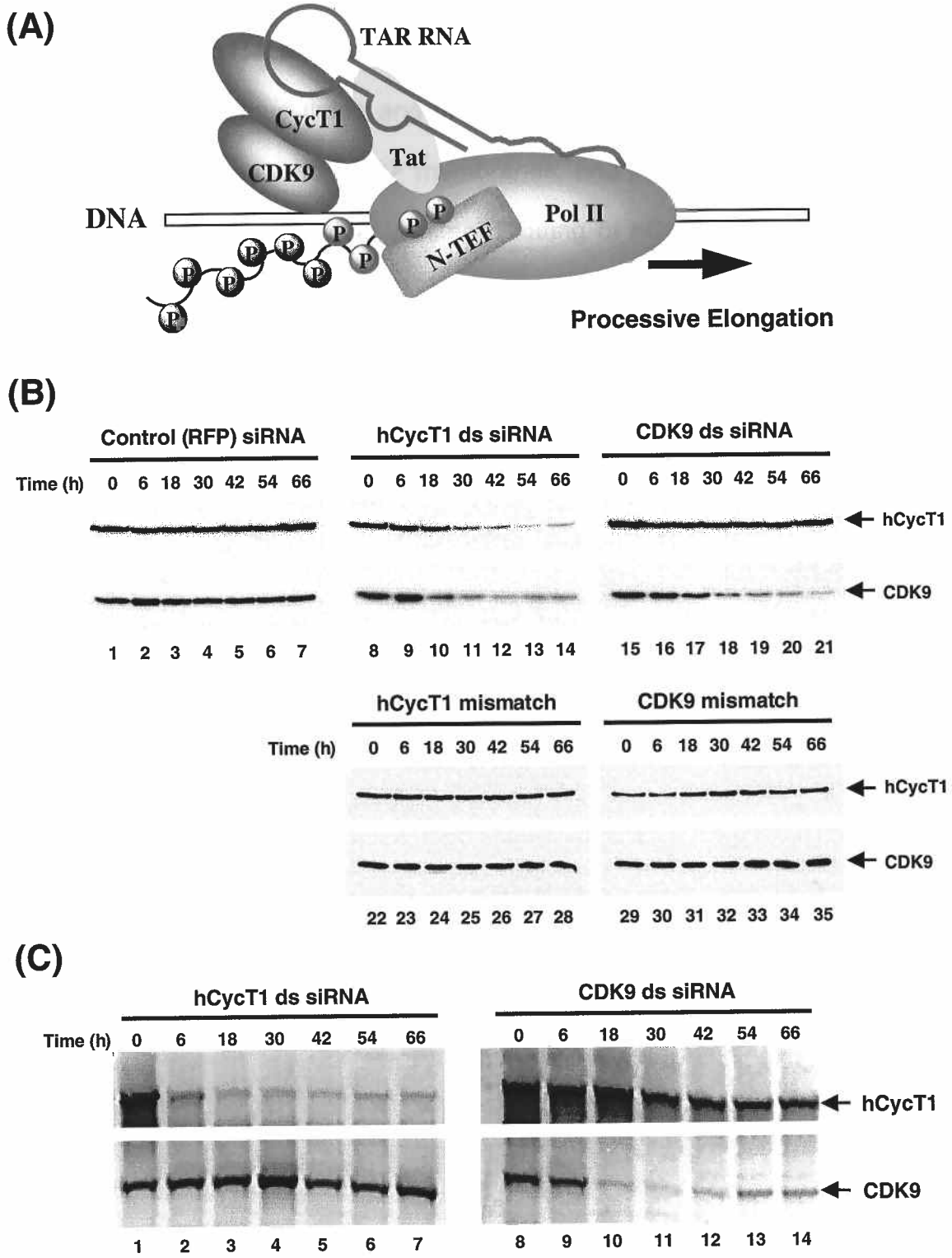

(A)
TAR RNA
CycT1
CDK9
Tat
DNA
Pol II
N-TEF
Processive Elongation
(B)
Control (RFP) siRNA
hCycT1 ds siRNA
CDK9 ds siRNA
Time (h) 0 6 18 30 42 54 66
hCycT1
CDK9
1 2 3 4 5 6 7
8 9 10 11 12 13 14
15 16 17 18 19 20 21
hCycT1 mismatch
CDK9 mismatch
22 23 24 25 26 27 28
29 30 31 32 33 34 35
(C)
hCycT1 ds siRNA
CDK9 ds siRNA
Time (h) 0 6 18 30 42 54 66
hCycT1
CDK9
1 2 3 4 5 6 7
8 9 10 11 12 13 14

Figure 2. Kinetics of hCycT1 and CDK9 silencing by siRNA. ds siRNAs were transfected into HeLa cells as described in Experimental Procedures. Cells were harvested at various times and protein contents were resolved on 10% SDS-PAGE, transferred onto PVDF membranes, and immunoblotted with antibodies against hCycT1 and CDK9. Western Blotting of hCycT1 and CDK9 over a 90-hour time course are shown.

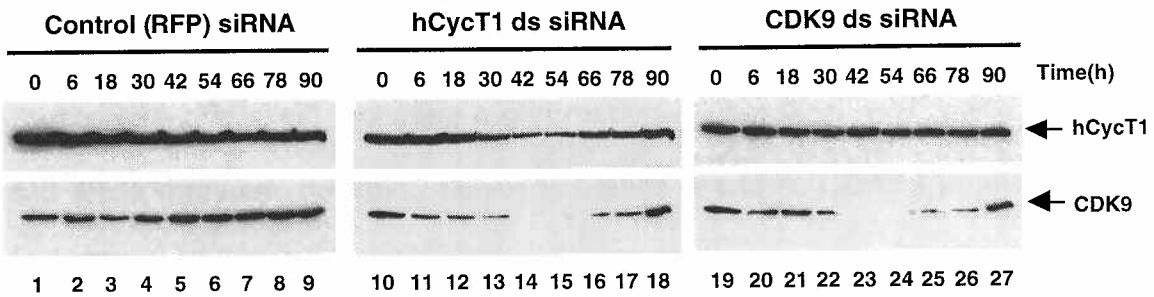
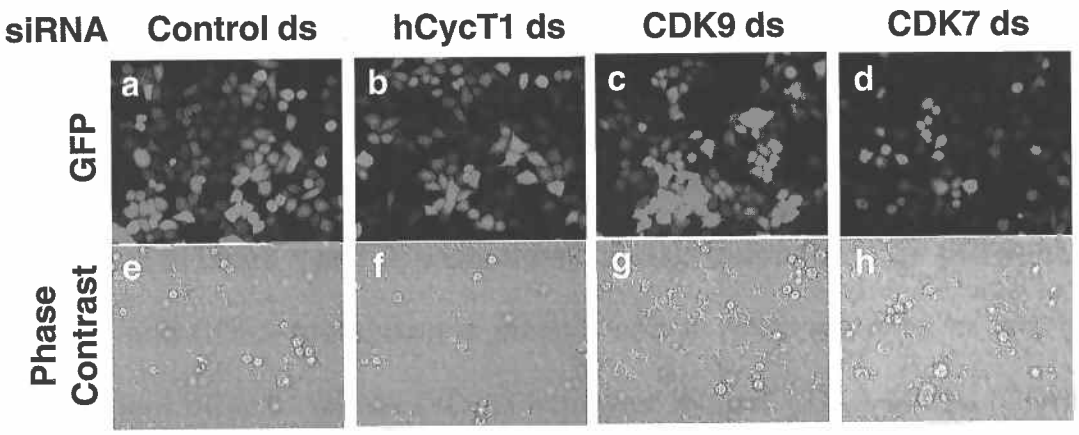

Control (RFP) siRNA
hCycT1 ds siRNA
CDK9 ds siRNA
0 6 18 30 42 54 66 78 90
0 6 18 30 42 54 66 78 90
0 6 18 30 42 54 66 78 90
Time(h)
hCycT1
CDK9
1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18
19 20 21 22 23 24 25 26 27

Figure 3. P-TEFb down-regulation can be achieved without causing major toxic or lethal effects to HeLa cells. (A) Analysis of cell viability by *in vivo* fluorescence analysis. HeLa cells were cotransfected by lipofectamine with pEGFP-C1 reporter (GFP) plasmid and siRNAs (see Experimental Procedures). Four siRNA duplexes, including a control duplex targeting RFP (panels a and e) and three duplexes targeting hCycT1 (panels b and f), CDK9 (panels c and g), and CDK7 (panels d and h), were used in these experiments. Reporter gene expression was monitored at 50 h post transfection by fluorescence imaging in living cells (upper panels). Cellular shape and density were recorded by phase contrast microscopy (lower panels). **(B) Analysis of cell viability by counting trypan blue-stained cells.** HeLa cells were cotransfected by lipofectamine with pEGFP-C1 reporter (GFP) plasmid and siRNAs (see Experimental Procedures). Four siRNA duplexes, including a control unrelated duplex (light blue) and three duplexes targeting hCycT1 (green), CDK9 (dark blue), and CDK7 (red), were used in these experiments. At various times after transfection, cells floating in the medium were collected and counted in the presence of 0.2% trypan blue (see Experimental Procedures). Cells that took up dye (stained blue) were not viable.

(A)


siRNA
Control ds
hCycT1 ds
CDK9 ds
CDK7 ds
GFP
Phase Contrast
a
b
c
d
e
f
g
h

(B)

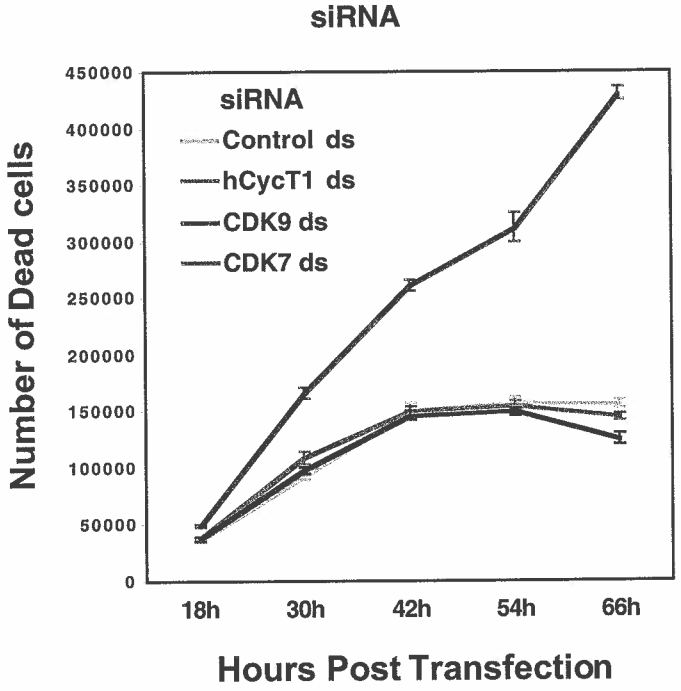
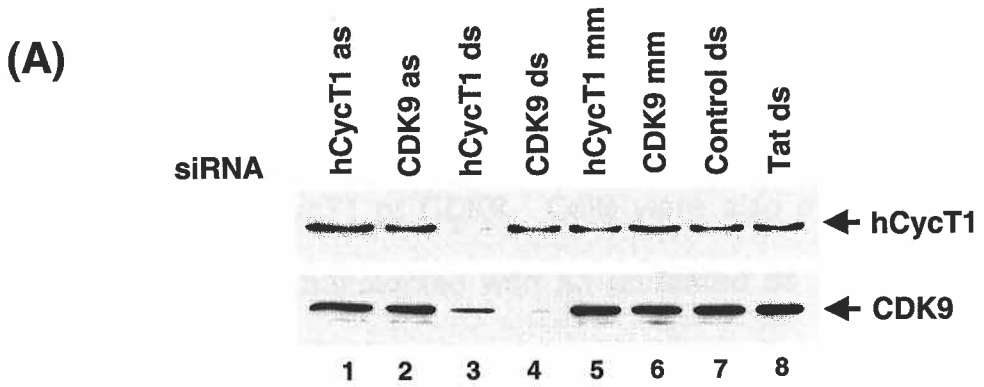
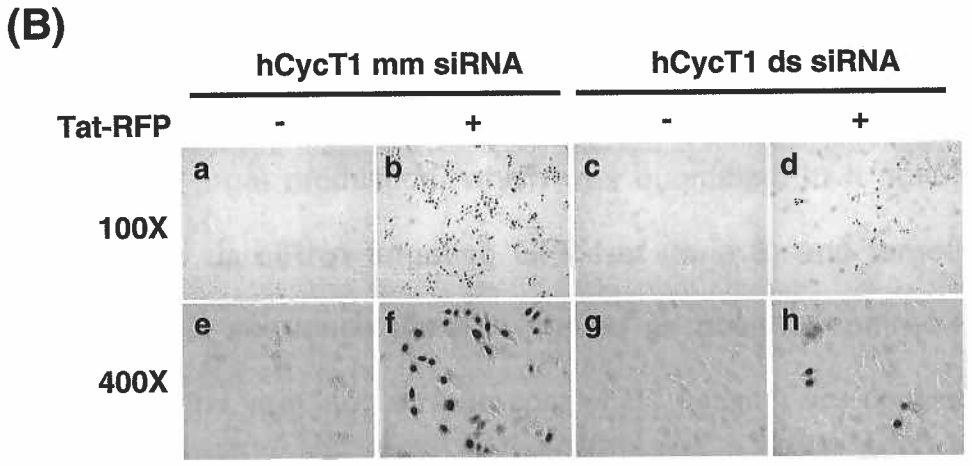

siRNA
siRNA
Control ds
hCycT1 ds
CDK9 ds
CDK7 ds
Number of Dead cells
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
18h
30h
42h
54h
66h
Hours Post Transfection

Figure 4. hCycT1 and CDK9 duplex siRNAs inhibit HIV-1 Tat transactivation in Magi cells. (A) Analysis of hCycT1 and CDK9 RNAi activities in Magi cells by Western blotting. Magi cells were co-transfected with pTat-RFP plasmid and various siRNAs. Cells were harvested at 48 h post transfection, resolved on 10% SDS-PAGE, transferred onto PVDF membranes, and immunoblotted with antibodies against hCycT1 (upper panel) and CDK9 (lower panel). RNAi activities in Magi cells treated with antisense (as) strands of hCycT1 and CDK9 siRNAs are shown in lanes 1 and 2, while those of cells treated with ds siRNA targeting hCycT1 and CDK9 are shown in lanes 3 and 4. RNAi activities in cells treated with mutant hCycT1 siRNA (hCycT1 mm) or mutant CDK9 siRNA (CDK9 mm) are shown in panels 5 and 6. GFP ds siRNA was used as an unrelated control (lane 7), while Tat ds RNAi was used to target mRNA encoding Tat (lane 8). **(B) Photomicograph of β-gal stained Magi cells.** Magi cells were either untransfected (panels a, c, e and g) or transfected (panels b, d, f and h) with pTat-RFP in the presence of mismatched hCycT1 siRNA (mm) (panels b and f) or hCycT1 ds siRNA (panels d and h). Syncytia formation and LTR activation (represented by blue β-gal-staining cells) were reduced in the hCycT1 ds siRNA-treated cells (panels d and h). **(C) Effect of P-TEFb knockdown by RNAi on Tat transactivation in Magi cells.** Twenty-four hours after pre-treating Magi cells with siRNA, they were cotransfected with pTat-RFP plasmid and various siRNAs. Cells were harvested 48h post pTat-RFP transfection, and activity of β-galactosidase in clear cell lysates was measured (see Experimental Procedures).

Tat transactivation was determined by the ratio of β-galactosidase activity in pTat-RFP-transfected cells to that of cells without pTat-RFP treatment. Inhibitory activity was determined by normalizing Tat transactivation activity to the amount of Tat RFP protein (see Experimental Procedures) in the presence or absence of siRNA treatment. Tat-RFP transfection (mock) is shown in lane 1. Magi cells were cotransfected with ds siRNAs targeting hCycT1 and CDK9 (lanes 4 and 5), with antisense (as) RNA strands (lanes 2 and 3), or mutant (mm) siRNAs (lanes 6 and 7). GFP ds siRNA was used as an unrelated control siRNA (lane 8), while Tat ds siRNA, targeting the mRNA encoding Tat sequence, was used as a positive control (lane 9). Means ± SD of two experiments are shown.


(A)
siRNA
hCycT1 as
CDK9 as
hCycT1 ds
CDK9 ds
hCycT1 mm
CDK9 mm
Control ds
Tat ds
hCycT1
CDK9
1 2 3 4 5 6 7 8


(B)
hCycT1 mm siRNA
hCycT1 ds siRNA
Tat-RFP
-
+
-
+
a
b
c
d
100X
e
f
g
h
400X

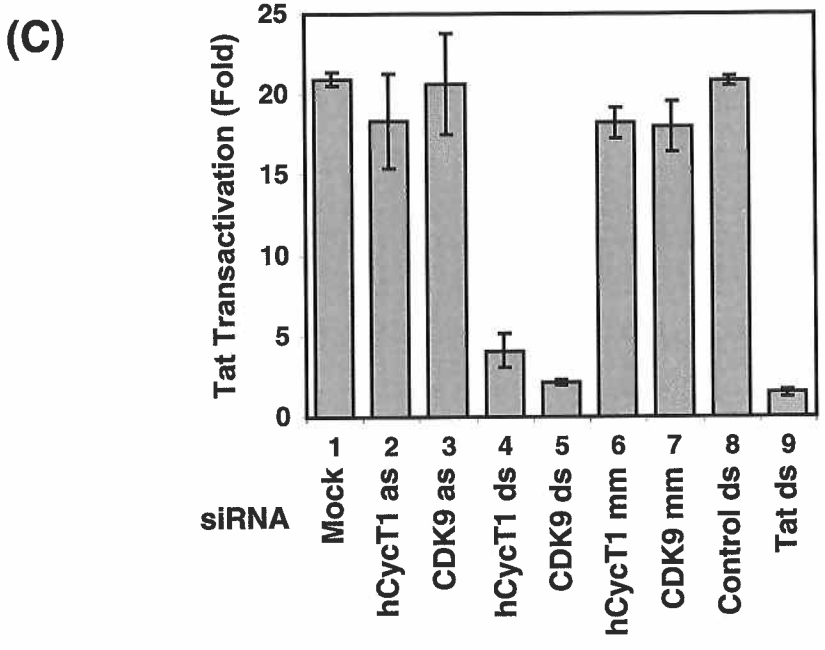

(C)
Tat Transactivation (Fold)
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9
siRNA
Mock
hCycT1 as
CDK9 as
hCycT1 ds
CDK9 ds
hCycT1 mm
CDK9 mm
Control ds
Tat ds

Figure 5. Small interfering RNAs targeting Cyclin T1 or CDK9 modulate HIV-1 infectivity. HeLa-CD4-LTR/β-galactosidase (Magi) cells were transfected with homologous (ds, lanes 3 and 4) and mismatched (mm, lanes 5 and 6) siRNAs directed against CycT1 or CDK9. Cells were also mock transfected without siRNA (lane 2) or transfected with an unrelated ds siRNA against the RFP sequence (lane 7). Sixteen hours later, cells were infected with NL-GFP, an infectious molecular clone of HIV-1. Cells infected with virus and not treated with oligofectamine are shown in lane 1. HIV-1 Tat-mediated transactivation of the LTR led to β-gal production, which was quantified 36 h post-infection. Cells treated with ds siRNA targeting GFP-Nef (lane 8) and targeting the mRNA encoding Tat sequence (lane 9) served as positive controls. Serial double dilutions of the viral inoculum (in cpm of RT activity) are consistent with 8-fold decreases in viral infectivity.

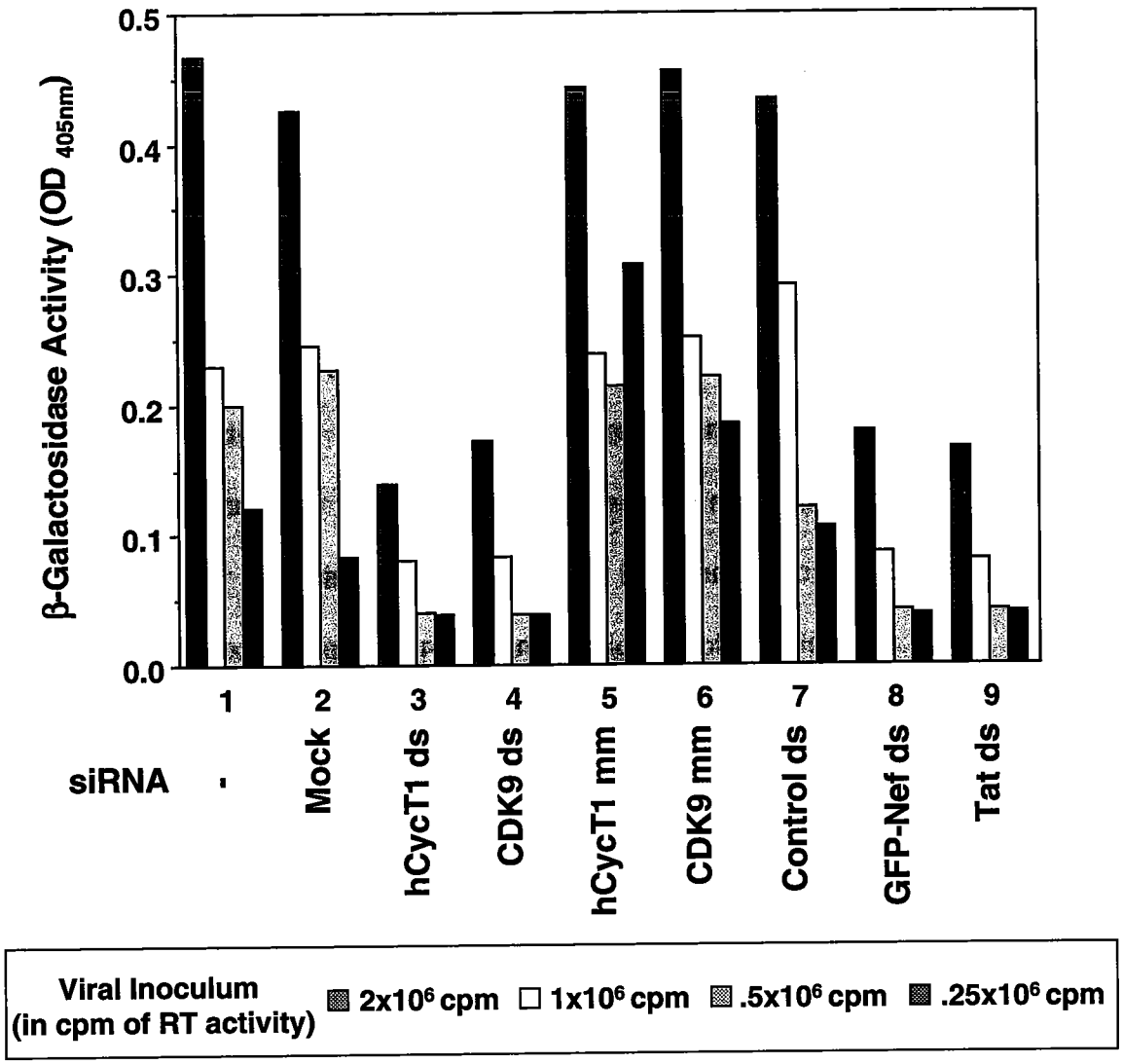
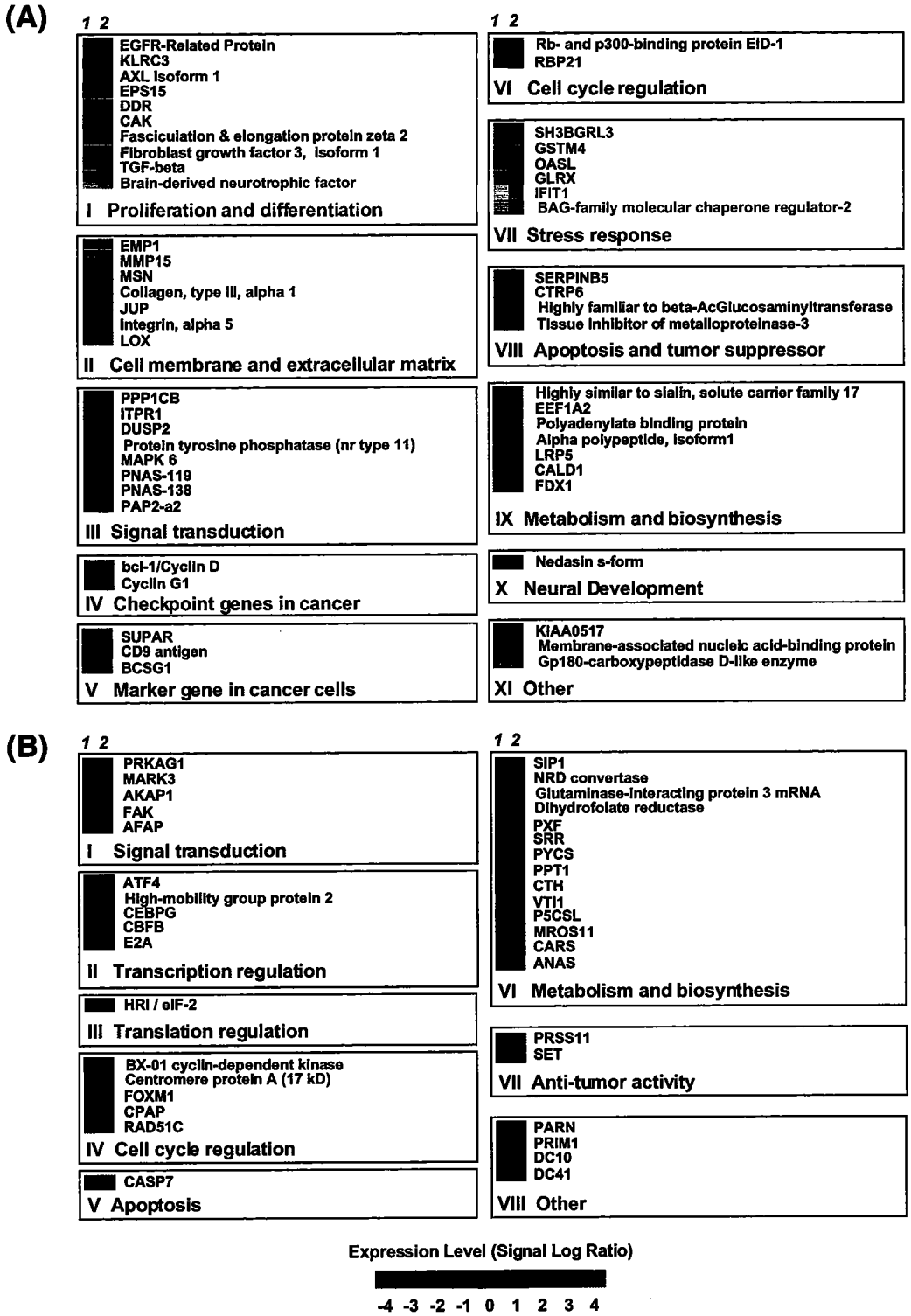

β-Galactosidase Activity (OD 405nm)
0.5
0.4
0.3
0.2
0.1
0.0
1
2
3
4
5
6
7
8
9
siRNA
-
Mock
hCycT1 ds
CDK9 ds
hCycT1 mm
CDK9 mm
Control ds
GFP-Nef ds
Tat ds
Viral Inoculum
(in cpm of RT activity)
2x10⁶ cpm
1x10⁶ cpm
.5x10⁶ cpm
.25x10⁶ cpm

Figure 6. Genomewide analysis of gene expression in P-TEFb knockdown HeLa cells. HeLa cells were treated with and without ds siRNA directed against hCycT1 or CDK9, and total mRNA was isolated. Total mRNA was used to synthesize ds cDNAs, from which biotin-labeled cRNA was synthesized and fragmented. Fragmented cRNA was then subjected to high-density oligonucleotide microarray hybridization (GeneChip®) using Human Genome U133 from Affymetrix (see Experimental Procedures). Of the 44,928 genes expressed, 90 are displayed by class, based on their putative functions. Each row represents one gene. Column 1 indicates hCycT1 ds siRNA treatment and column 2 indicates CDK9 ds siRNA treatment. The down- (A) and up- (B) regulated genes are represented by green and red, respectively. The brightness of each color reflects the magnitude of the gene expression level (Signal Log Ratio). See Experimental Procedure for details of analysis.


(A)
1 2
EGFR-Related Protein
KLRC3
AXL Isoform 1
EPS15
DDR
CAK
Fasciculation & elongation protein zeta 2
Fibroblast growth factor 3, isoform 1
TGF-beta
Brain-derived neurotrophic factor
I Proliferation and differentiation
EMP1
MMP15
MSN
Collagen, type III, alpha 1
JUP
Integrin, alpha 5
LOX
II Cell membrane and extracellular matrix
PPP1CB
ITPR1
DUSP2
Protein tyrosine phosphatase (nr type 11)
MAPK 6
PNAS-119
PNAS-138
PAP2-a2
III Signal transduction
bcl-1/Cyclin D
Cyclin G1
IV Checkpoint genes in cancer
SUPAR
CD9 antigen
BCSG1
V Marker gene in cancer cells
1 2
Rb- and p300-binding protein EID-1
RBP21
VI Cell cycle regulation
SH3BGRL3
GSTM4
OASL
GLRX
IFIT1
BAG-family molecular chaperone regulator-2
VII Stress response
SERPINB5
CTRP6
Highly familiar to beta-AcGlucosaminyltransferase
Tissue inhibitor of metalloproteinase-3
VIII Apoptosis and tumor suppressor
Highly similar to sialin, solute carrier family 17
EEF1A2
Polyadenylate binding protein
Alpha polypeptide, isoform1
LRP5
CALD1
FDX1
IX Metabolism and biosynthesis
Nedasin s-form
X Neural Development
KIAA0517
Membrane-associated nucleic acid-binding protein
Gp180-carboxypeptidase D-like enzyme
XI Other
(B)
1 2
PRKAG1
MARK3
AKAP1
FAK
AFAP
I Signal transduction
ATF4
High-mobility group protein 2
CEBPG
CBFB
E2A
II Transcription regulation
HRI / eIF-2
III Translation regulation
BX-01 cyclin-dependent kinase
Centromere protein A (17 kD)
FOXM1
CPAP
RAD51C
IV Cell cycle regulation
CASP7
V Apoptosis
1 2
SIP1
NRD convertase
Glutaminase-interacting protein 3 mRNA
Dihydrofolate reductase
PXF
SRR
PYCS
PPT1
CTH
VTI1
P5CSL
MROS11
CARS
ANAS
VI Metabolism and biosynthesis
PRSS11
SET
VII Anti-tumor activity
PARN
PRIM1
DC10
DC41
VIII Other
Expression Level (Signal Log Ratio)
-4 -3 -2 -1 0 1 2 3 4

CHAPTER V

RNA INTERFERENCE IN HUMAN CELLS: BASIC STRUCTURAL AND FUNCTIONAL FEATURES OF SMALL INTERFERING RNA

ABSTRACT

We investigated the mechanism of RNA interference (RNAi) in human cells. Here we demonstrate that the status of the 5' hydroxyl terminus of the antisense strand of a siRNA determines RNAi activity, while a 3' terminus block is tolerated in vivo. 5' hydroxyl termini of antisense strands isolated from human cells were phosphorylated, and 3' end biotin groups were not efficiently removed. We found no requirement for a perfect A-form helix in siRNA for interference effects, but an A-form structure was required for antisense-target RNA duplexes. Strikingly, crosslinking of the siRNA duplex by psoralen did not completely block RNA interference, indicating that complete unwinding of the siRNA helix is not necessary for RNAi activity in vivo. These results suggest that RNA amplification by RNA-dependent RNA polymerase is not essential for RNAi in human cells.

INTRODUCTION

RNAi interference (RNAi) is the process whereby double-stranded RNA (dsRNA) induces the sequence-specific degradation of homologous mRNA. Although RNAi was first discovered in *Caenorhabditis elegans* (62), similar phenomena had been reported in plants (posttranscriptional gene silencing [PTGS]) and in *Neurospora crassa* (quelling) (reviewed in 84, 206). It has become clear that dsRNA-induced silencing phenomena are present in evolutionarily diverse organisms, e.g., nematodes, plants, fungi, and trypanosomes (8, 37, 62, 84, 122, 150, 206, 215, 238, 248). Biochemical studies in *Drosophila* embryo lysates and S2 cell extracts have begun to unravel the mechanisms by which RNAi works (15, 239, 271).

RNAi is initiated by an ATP-dependent, processive cleavage of dsRNA into 21 to 23 nucleotide (nt) short interfering RNAs (siRNAs) (82, 83, 271) by the enzyme Dicer, a member of the RNase III family of dsRNA-specific endonucleases (15). These native siRNA duplexes containing 5' phosphate and 3' hydroxyl termini are then incorporated into a protein complex called RNA-induced silencing complex (RISC) (83). ATP-dependent unwinding of the siRNA duplex generates an active complex, RISC* (the asterisk indicates the active conformation of the complex) (166). Guided by the antisense strand of siRNA, RISC* recognizes and cleaves the corresponding mRNA (57, 83, 166).

Recently, Tuschl and colleagues (55) have demonstrated that RNAi can be induced in numerous mammalian cell lines by introducing synthetic 21 nt

siRNAs. By virtue of their small size, these siRNAs avoid provoking an interferon response that activates the protein kinase PKR (219). Functional anatomy studies of synthetic siRNA in *Drosophila* cell lysates have demonstrated that each siRNA duplex cleaves its target RNA at a single site (58). The 5' end of the guide siRNA sets the ruler for defining the position of target RNA cleavage (58). 5' phosphorylation of the antisense strand is required for effective RNA interference in vitro (166). Mutation studies have shown that a single mutation within the center of a siRNA duplex discriminates between mismatched targets (58). These experiments showed a more stringent requirement for the antisense strand of the trigger dsRNA as compared to the sense strand (78, 171). However, none of these phenomena have been demonstrated *in vivo*, especially in mammalian systems.

A particularly fascinating aspect of RNAi is its extraordinary efficiency. It has been estimated that in *Drosophila* embryos, ~35 molecules of dsRNA can silence a target mRNA thought to be present at >1000 copies per cell (118). Conversion of the long trigger dsRNA into many 21 to 23 nt siRNA fragments would, itself, provide some degree of amplification. Another plausible explanation for the potency of interference is that the RISC* is a multiple-turnover enzyme, which can catalytically perform the targeting and cleavage activity. The involvement of RNA-dependent RNA polymerase (RdRP) in the amplification process has recently been postulated since genetic screening has identified the gene for RdRP as a requirement for gene silencing in plants, fungi, and worms

(36, 47, 160, 215). A random degradative PCR model has been suggested (136, 164, 215), in which siRNA serves as the primer for the RdRP reaction. The siRNA-primed RdRP converts target mRNA into dsRNA, which can serve as Dicer substrates, initiating the RdRP chain reaction. The polarity of the RdRP reaction limits the synthesis of secondary siRNAs to the region upstream of the trigger sequence. Certain structural features of siRNA, including the 3' hydroxyl group and 5' phosphate group, are critical for the RdRP reaction and RNA ligation (136, 215). A number of apparent constraints on the role of RdRP activity in RNAi, however, severely limit models in which RdRP carries out a multiround replication of a double-stranded trigger. Experiments using siRNA with two differentially modified strands (171, 264) have shown a more stringent requirement for the antisense strand of the original trigger as compared to the sense strand. Multiround replication of a double-stranded trigger by RdRP would result in loss of memory of the difference between the original two strands and thus be incompatible with the observed effects of strand-specific modification. Although RdRP activity has been reported in *Drosophila* embryo lysates (136), a homolog of RdRP has not been identified in available mammalian genomic sequences.

A number of basic questions remain to be answered regarding RNA interference in human cells. (1) What is the kinetics of RNAi? (2) Is 5' OH required for kinase activity and is it phosphorylated in vivo? (3) Is 3' OH required for RNA-dependent RNA polymerase-like activity? (4) Is an A-form helix of

dsRNA required for RNAi? If yes, at which stage during the pathway? (5) Is complete unwinding of the dsRNA helix necessary to cause RNAi effects? This chapter addresses these issues.

MATERIALS AND METHODS

siRNA preparation. Twenty-one nucleotide RNAs were chemically synthesized as 2' bis(acetoxyethoxy)-methyl ether-protected oligos by Dharmacon (Lafayette, CO). Synthetic oligonucleotides were deprotected, annealed, and purified as described by the manufacturer. Successful duplex formation was confirmed by 20% nondenaturing polyacrylamide gel electrophoresis (PAGE). All siRNAs were stored in DEPC (0.1% diethyl pyrocarbonate)-treated water at -80°C. The sequences of GFP or RFP target-specific siRNA duplexes were designed according to the manufacturer's recommendation and subjected to a BLAST search against the human genome sequence to ensure that no endogenous genes of the genome were targeted.

Culture and transfection of cells. HeLa cells were maintained at 37°C in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS), 100 units/ml penicillin, and 100 μg/ml streptomycin (Invitrogen). Cells were regularly passaged at subconfluence and plated 16 h before transfection at 70% confluency. Lipofectamine (Invitrogen)-mediated transient cotransfections of reporter plasmids and siRNAs were performed in duplicate 6-well plates as described by the manufacturer for adherent cell lines. A transfection mixture containing 0.66 μg pEGFP-C1 and 1.33 μg pDsRed1-N1 reporter plasmids (Clontech), 50 nM siRNA, and 10 μl lipofectamine in 1 ml serum-reduced OPTI-MEM (Invitrogen) was added to each well. Cells were incubated in transfection mixture for 6 h and further cultured in

antibiotic-free DMEM. Cells were treated under the same conditions without siRNA for mock experiments. At various time intervals, the transfected cells were washed twice with phosphate-buffered saline (PBS, Invitrogen), flash frozen in liquid nitrogen, and stored at -80°C for reporter gene assays.

***In vivo* fluorescence analysis**. pEGFP-C1, pDsRed1-N1 reporter plasmids and 50 nM siRNA were cotransfected into HeLa cells by lipofectamine as described above except that cells were cultured on 35 mm plates with glass bottoms (MatTek Corporation, Ashland, MA) instead of standard 6-well plates. Fluorescence in living cells was visualized 48 hr posttransfection by conventional fluorescence microscopy (Zeiss). For GFP and RFP fluorescence detection, FITC and CY3 filters were used, respectively.

Dual fluorescence reporter gene assays. pEGFP-C1, pDsRed1-N1 reporter plasmids and 50 nM siRNA were cotransfected into HeLa cells. EGFP-C1 encoded enhanced green fluorescence protein (GFP), while DsRed1-N1 encoded red fluorescence protein (RFP). Cells were harvested as described above and lysed in ice-cold reporter lysis buffer (Promega) containing protease inhibitor (complete, EDTA-free, 1 tablet/10 ml buffer, Roche Molecular Biochemicals). After clearing the resulting lysates by centrifugation, protein in the clear lysate was quantified by Dc protein assay kit (Bio-Rad). One hundred twenty micrograms of total cell lysate in 160 μl reporter lysis buffer was measured by fluorescence spectroscopy (Photon Technology International). The slit widths were set at 4 nm for both excitation and emission. All experiments

were carried out at room temperature. Fluorescence of GFP in cell lysates was detected by exciting at 488 nm and recording from 498–650 nm. The spectrum peak at 507 nm represents the fluorescence intensity of GFP. Fluorescence of RFP in the same cell lysates was detected by exciting at 568 nm and recording from 588–650 nm; the spectrum peak at 583 nm represents the fluorescence intensity of RFP. The fluorescence intensity ratio of target (GFP) to control (RFP) fluorophore was determined in the presence of siRNA duplex and normalized to that observed in the presence of antisense strand siRNA. Normalized ratios less than 1.0 indicate specific interference.

Western blotting. Cell lysates were prepared from siRNA-treated cells and analyzed as described above. Proteins in 30 μg of total cell lysate were resolved by 10% SDS-PAGE, transferred onto a polyvinylidene difluoride membrane (PVDF membrane, Bio-Rad), and immunoblotted with antibodies against EGFP and DsRed1-N1 (Clontech). For loading control, the same membrane was also blotted with anti-actin actibody (Santa Cruz). Protein content was visualized with a BM Chemiluminescence Blotting Kit (Roche Molecular Biochemicals). The blots were exposed to X-ray film (Kodak MR-1) for various times (between 30 s and 5 min).

Psoralen photocrosslink of siRNA duplex. Forty micrograms of siRNA duplex was incubated with 132 μM of a psoralen derivative, 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) in 200 μl DEPC-treated water at 30°C for 30 min. Mixtures of siRNA duplex and HMT were exposed to UV 360 nm at 4°C for 20

min, then denatured by mixing with 400 µl formamide/formaldehyde (12.5:4.5) RNA loading buffer and heating at 95°C for 15 min. Crosslinked siRNA duplex and noncrosslinked siRNA were resolved by 20% PAGE containing 7 M urea in Tris-borate-EDTA. Crosslinked siRNA duplexes appeared as a population with retarded electrophoretic mobility compared to the noncrosslinked species. RNAs were cut from the gel and purified by C18 reverse-phase column chromatography (Waters). Purified crosslinked dsRNA and noncrosslinked dsRNA were used in dual fluorescence reporter assays as described above, except that all procedures were performed in the dark to avoid light effects on psoralen. To ensure that the crosslink depended on the presence of psoralen, part of the UV 360 nm treated mixture was also subjected to UV 254 nm at 4°C for 20 min. Photoreverse-crosslinked siRNA migrated in 20% polyacrylamide-7 M urea gels with similar mobility to the siRNA duplex without HMT treatment.

Biotin pull-out assay for siRNA isolation from human cells. Antisense strands of the siRNA duplex were chemically synthesized and biotin-conjugated at the 3' end (Dharmacon, Lafayette, CO). Synthetic oligonucleotides were deprotected and annealed with the unmodified sense strand RNA to form duplex siRNA (ss/as3'-Biotin). HeLa cells, which had been plated at 70% confluency in 100 mm dishes, were cotransfected with duplex siRNA (~600 pmole) and EGFP-C1 plasmid (1 µg) by a lipofectamine-mediated method as described above. At various times, the transfected cells were washed twice with PBS (Invitrogen) and flash frozen in liquid nitrogen. Low molecular weight RNA was isolated from the

cells using a Qiagen RNA/DNA Mini Kit. Biotinylated siRNA was pulled out by incubating purified RNA with streptavidin-magnetic beads (60 μl) in TE buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA) containing 1 M NaCl at room temperature for 3 hr. The beads were washed four times with 200 μl TE buffer, resuspended in 100 μl TE buffer, and split into two equal aliquots. To one aliquot (50 μl), we added 50 units of shrimp alkaline phosphatase (SAP, Roche Molecular Biochemicals) in 1X SAP buffer and incubated at 37°C for 1 hr. The SAP reaction was then stopped by heating at 65°C for 15 min and washed four times with 200 μl TE buffer. The other aliquot was not treated with SAP. Aliquots of beads with or without SAP treatment were incubated with 30 units T4 polynucleotide kinase (T4 PNK, Roche Molecular Biochemicals) in 30 μl 1X PNK buffer containing 0.2 mCi -32P ATP at 37°C for 1 hr. RNA products were resolved on 20% polyacrylamide-7 M urea gels and 32P-labeled RNAs were detected by phosphorimaging.

RESULTS

Dual fluorescence reporter system for RNAi analysis in mammalian cells

To explore the functional anatomy of siRNA in mammalian cells, we established a dual fluorescence reporter system, using HeLa cells as a model system. Two reporter plasmids were used: pEGFP-C1 and pDsRed1-N1, harboring enhanced green fluorescent protein (GFP) or coral (*Discosoma spp.*)-derived red fluorescent protein (RFP), respectively. The expression of these reporter genes was under cytomegalovirus promoter control and could be easily visualized by fluorescence microscopy in living cells. The siRNA sequence targeting GFP was from position 238–258 relative to the start codon, and the RFP siRNA sequence was from position 277–297 relative to the start codon (Figure 1A). Using lipofectamine, we cotransfected HeLa cells with pEGFP-C1 and pDsRed1-N1 expression plasmids and siRNA duplex, targeting either GFP or RFP. Fluorescence imaging was used to monitor GFP and RFP expression levels. As shown in Figures 1Ba and 1Bb, mock treatment (without siRNA) allowed efficient expression of both GFP and RFP in living cells. Transfection of cells with siRNA duplex targeting GFP (GFP ds) significantly reduced GFP expression (Figure 1Bc) but had no effect on RFP expression (Figure 1Bd) compared with mock-treated cells (Figures 1Ba and 1Bb). On the other hand, transfection of cells with siRNA duplex targeted to RFP (RFP ds) significantly interfered with the expression of RFP but not GFP (Figures 1Be and 1Bf).

To quantify RNAi effects, we prepared lysates from siRNA duplex-treated cells at 42 hr posttransfection. GFP and RFP fluorescence in clear lysates was measured on a fluorescence spectrophotometer. The peak at 507 nm (Figure 1C, left panel) represents the fluorescence intensity of GFP, and the peak at 583 nm (Figure 1C, right panel) represents the fluorescence intensity of RFP. GFP fluorescence intensity of GFP ds-treated cells (Figure 1C, left panel, green line) was only 5% of mock-treated (black line) or RFP ds-treated cells (cyan line). In contrast to GFP fluorescence, RFP fluorescence intensity (Figure 1C, right panel) significantly decreased only in cells treated with RFP ds (red line), indicating the specificity of the RNAi effect.

To confirm these findings on RNAi effects in living mammalian cells, we carried out Western blotting experiments (Figure 2) using anti-GFP and anti-RFP antibodies. Analysis of immunoblots revealed that the siRNA targeting GFP inhibited only GFP expression without affecting RFP levels (Figure 2A, lanes 9–14); siRNA targeting RFP was similarly specific against RFP expression (Figure 2B, lanes 9–14). This RNA interference effect depended on the presence of 21 nucleotide duplex siRNA but not of the antisense strand siRNA (Figures 2A and 2B, compare right and left panels). These results demonstrate that we have established a reliable and quantitative system for studying specific RNA interference in HeLa cells.

Kinetics of RNA interference in HeLa cells

One of the many intriguing features of gene silencing by RNA interference is its unusually high efficiency—a few dsRNA molecules suffice to inactivate a continuously transcribed target mRNA for long periods of time. It has been demonstrated in plants (36, 47) and worms (78) that this inactivation can spread throughout the organism and is often heritable to the next generation. Mutations in genes encoding a protein related to RdRP affect RNAi-type processes in *Neurospora* (QDE-1; 36), *C. elegans* (EGO-1; 217), and plants ([SGS2; 160] and [SDE-1; 47]). The involvement of RdRP in amplifying RNAi has been postulated (136).

To understand the kinetics of gene suppression and persistence of RNA interference in HeLa cells, we prepared lysates from cells cotransfected with GFP siRNA and dual fluorescence reporter plasmids, pEGFP-C1 and pDsRed1-N1. In this experiment, GFP was the target of the duplex siRNA, while RFP was used as a control for transfection efficiency and specificity of RNA interference. Emission spectra of GFP in cell lysates at various times after transfection (Figure 3) show that siRNA duplex caused an RNA interference effect as early as 6 hr posttransfection. This effect gradually increased with time, peaking at 42 hr, then started to decrease at 66 hr (Figure 3, green lines). As a control experiment, GFP expression in the presence of antisense strand was also monitored and showed no RNAi effects (Figure 3, blue lines). Thus, RNA interference can last for at least 66 hr in HeLa cells (Figure 3, green lines).

To quantify the kinetics of RNA interference, we measured the fluorescence intensity ratio of target (GFP) to control (RFP) fluorophore in the presence of siRNA duplex (ds) and normalized it to the ratio observed in the presence of antisense strand siRNA (as). Normalized ratios less than 1.0 indicate specific interference. As shown in Figure 1D, at 6 hr posttransfection GFP duplex siRNA (green bars) inhibits 60% of GFP expression compared to antisense strand siRNA (blue bars). RNA interference reached its maximum (92% inhibition) at 42 hr posttransfection; only 8% of normal GFP expression was observed in duplex siRNA-treated cells. These results show that RNA interference can suppress target protein expression up to 66 hr, although maximum activities were observed at 42–54 hr posttransfection.

Free 5' OH groups on the antisense strand of the siRNA duplex are required for RNA interference in vivo

Synthetic 21 nucleotide siRNA duplexes with 5' hydroxyl termini and 3' overhang have been shown to specifically suppress expression of endogenous and heterologous genes in *Drosophila* extracts (57) and mammalian cell lines (55). Nonetheless, native siRNA, processed by Dicer cleavage of dsRNA, contains 5' phosphate ends (57). It has been demonstrated *in vitro* that *Drosophila* embryo lysates contain a potent kinase activity that phosphorylates the 5' hydroxyl termini of synthetic siRNAs (166). The 5' phosphate is required on the siRNA strand that guides target cleavage in RNA interference (166).

To examine the importance of 5' termini of siRNA in RNA interference in human cells, we modified synthetic siRNAs targeting GFP by using an amino group with a 3-carbon linker (5' N3, Figure 4A) to block their 5' termini. Synthetic siRNAs with this modification lacked a hydroxyl group to be phosphorylated by kinases in vivo. This modification could also block access to siRNA by cellular factors that might require recognizing the 5' OH termini. We annealed unmodified siRNA strands with 5' modified strands, producing siRNA duplexes with 5' modification at only the sense strand (5'-N3ss/as), at only the antisense strand (ss/5'-N3as), or at both strands (5'-N3ss/5'-N3as) (Figure 4B). RNAi effects of these siRNA duplexes were analyzed in our dual fluorescence reporter system as described in Figure 1. 5' modification of the sense strand had no effect on RNAi activity (Figures 5b and 5c), whereas 5' modification of the antisense strand completely abolished the RNAi effect (Figures 5d and 5e; Figures 6A and 6B, upper panels). HeLa cells transfected with antisense strand (as) siRNA as control showed no RNAi activity (Figure 5a). These results demonstrate that the 5' OH in the antisense strand of the siRNA duplex is an important determinant of RNAi activity in human cells.

Blocking the 3' nd oef siRNAs has little effect on RNA interference *in vivo*

To determine the effect of 3' OH groups on RNAi activity, we synthesized siRNA duplexes containing a 3' end blocked with 3' puromycin (3'-Pmn, Figure 4A) or biotin instead of 3' OH groups on the overhang deoxythymidine (Figure

4B). These 3' end modifications would block any processing of the siRNA duplex that required a free 3' hydroxyl group. We prepared three combinations of siRNA duplexes containing 3' puromycin: 3' blocked at only the sense strand (ss3'-Pmn/as), at only the antisense strand (ss/as3'-Pmn), or at both strands (ss3'-Pmn/as3'-Pmn) (Figure 4B). We also prepared a siRNA duplex containing biotin at the 3' end of antisense strand (ss/as3'-Biotin). The RNAi activities of these siRNA duplexes were analyzed in our dual fluorescence reporter system. Results of these experiments indicate that a 3' block at either the sense or antisense strand of siRNA duplex had little effect on its RNA interference activity (Figures 5f–5i; Figures 6A and 6B, middle panels). Furthermore, biotin pull-out experiments showed that the 3' end biotin groups on the antisense strand were not efficiently removed during RNAi activities in HeLa cells (Figure 7, see below). Modifications could be introduced in the 3' overhangs without affecting siRNA efficacy, suggesting that RNA interference in mammalian cells does not occur through the recently reported RdRP-dependent degradative PCR mechanism (Lipardi et al., 2001; Sijen et al., 2001), which requires a free 3' hydroxyl group.

A-Form helix of siRNA is absolutely required for effective RNA interference *in vivo*

Synthetic and native siRNAs, generated from ATP-dependent cleavage of double-strand RNA, have been proposed to act as "guide RNAs" that target an associated nuclease complex, the RISC (RNA-induced silencing complex), to the

corresponding mRNA through strand complementarity (83, 166). How are these siRNA duplexes recognized and incorporated into the RISC protein complex? siRNA duplexes are readily characterized by their A-form helix, which can be distinguished from the structures of B-form helix DNA and single-stranded RNA in the cell. A single mismatch between a target mRNA and its guide strand siRNA completely prevents target RNA cleavage in *Drosophila* embryo lysates (58). Although the mechanism of target recognition has not been experimentally demonstrated, this finding indicates that recognition requires exact complementarity between the guide strand and target mRNA.

These observations raise two fundamental questions regarding RNAi effects in vivo. (1) Is an A-form RNA helix required in the siRNA structure? (2) Is an A-form helix recognized by proteins after the antisense strand of siRNA duplex is hybridized with the target mRNA? To address these questions, we designed three siRNA duplexes containing internal bulge structures in the RNA helices (Figure 4B). The A-form RNA helix has a deep, narrow major groove and a shallow, wide minor groove. More than one nucleotide bulge has been shown to distort RNA helical structures, widening the major groove and enhancing accessibility to its functional groups (162, 250, 251) . We decided to use 2 nt bulges to generate distorted A-form helices in siRNAs. We synthesized mutant siRNA by introducing two extra nucleotides into the sense or antisense strand of siRNA duplexes. Combining these mutant siRNA strands with original siRNA sequences produced three siRNA duplexes with an internal bulge at only the

sense strand (ss-bulge/as), at only the antisense strand (ss/as-bulge), or at both strands (ss-bulge/as-bulge) (Figure 4B). This design of bulge-containing siRNAs could dissect the requirement for the A-form helix at two different steps of RNA interference: (1) siRNA recognition by RISC and (2) RISC targeting of mRNA via the guiding siRNA. siRNA duplexes with an internal bulge at only the sense strand (ss-bulge/as) caused a structural change in the siRNA duplex (an imperfect A-form) without affecting the complementarity between target mRNA and the antisense strand, which acts as the guiding strand in the RNA interference pathway. RNA interference by these siRNA duplexes was analyzed and quantified in our dual fluorescence reporter system as described above.

Surprisingly, the siRNA duplex containing a bulge in its sense strand retained most of its RNA interference activity (compare Figures 5b and 5j; Figures 6A and 6B , lower panels, green line and bars), indicating that an A-form siRNA helix is not essential for effective RNA interference in vivo. However, bulges in the antisense strand or both strands of duplex siRNA completely abolished RNA interference ability (Figures 5k and 5l; Figures 6A and 6B, lower panels, dark and light blue line and bars), indicating that effective RNA interference in vivo absolutely requires A-form helix formation between target mRNA and its guiding antisense strand.

5' OH groups on the antisense strand of the siRNA duplex are phosphorylated *in vivo*

To analyze the phosphorylation status of the 5' termini of siRNA and to probe the participation of siRNA 3' termini in the RNA interference pathway in vivo, we transfected HeLa cells with 21 nt RNAs containing biotin at the 3' terminal of the antisense strand (ss/as3'-Biotin) and isolated the biotinylated siRNA at various times after transfection (see Experimental Procedures). In brief, streptavidin magnetic beads were used to pull out biotinylated siRNAs from transfected cells, washed to remove unbound RNA, and split into two aliquots. One aliquot was dephosphorylated with shrimp alkaline phosphatase (SAP), and the RNA 5' ends were labeled with 32P by T4 polynucleotide kinase (PNK) reaction. The other aliquot was subjected to 5' end radiolabeling with polynucleotide kinase without prior dephosphorylation reaction with SAP. RNA was resolved on 20% polyacrylamide-7 M urea gels and visualized by phosphorimager analysis. Cells without siRNA treatment showed no detectable signal after biotin pull-out assay (Figure 7, lane 4), indicating the absence of nonspecific RNA-bead interactions. Efficient 5' end radiolabeling was observed only when RNA was pretreated with phosphatase (compare lanes 5–9 and 10–14), indicating that the 5' termini of siRNA did not contain free OH groups in vivo. Although phosphorylating with SAP and quenching the phosphatase reaction by heating resulted in some RNA degradation, the efficiency of the kinase reaction after SAP treatment is obvious. These results indicate that 5' OH groups are phosphorylated in vivo for RNAi activities.

These experiments demonstrate three key findings. First, biotinylated-siRNA can be isolated from HeLa cells at 6 to 54 hr posttransfection (Figure 7, lanes 5–9). The amount of isolated siRNA decreased in a time-dependent manner, indicating the degradation of siRNA in vivo. Our dual fluorescence assays showed that RNA interference mediated by 3' end biotinylated siRNA was as effective as unmodified siRNA (Figures 5f and 5b; Figures 6A and 6B , middle panel). RNA interference is seen as early as 6 hr post-siRNA transfection and can be maintained for 42 hr posttransfection. Our ability to isolate biotin-RNA from cells after RNA interference had been initiated indicates that biotin was not removed from the RNA and rules out the possibility of siRNA 3' OH termini involvement in the RNA interference pathway in human cells.

Second, in this biotin pull-out assay, only siRNA with 5' OH ends can be 32P-labeled by T4 PNK. As shown in Figure 7, the siRNA without SAP treatment was not efficiently labeled by T4 PNK (e.g., compare lane 10 to lane 5 and lane 11 to lane 6), indicating that the 5' termini of siRNA did not contain free OH groups in vivo. These 5' terminal groups can be removed by alkaline phosphatase treatment for subsequent radiolabeling (Figure 7, lanes 5–9), indicating that the 5' termini of the siRNA had been phosphorylated in vivo.

Third, only the antisense strand is recovered by biotin pull-out assays. siRNA duplexes were 5' end labeled with 32P by T4 PNK, heat denatured (10 min at 95°C), and analyzed on a polyacrylamide-7 M urea denaturing gel. As shown in Figure 7 (lane 3), two single-stranded RNA species corresponding to

the sense and biotinylated-antisense strands were observed, indicating that the siRNA duplexes were fully denatured under these conditions. Denatured siRNA duplexes contained approximately equal molar amounts of the sense and the antisense strands of RNA (Figure 7, lane 3). The cells were transfected with duplex siRNA, but the major products of the isolated siRNA (Figure 7, lanes 5–9) by biotin pull-out assay exhibited electrophoretic moblities identical to the antisense strand (lane 3), indicating that only biotinylated antisense strands were being recovered. These results suggest that RISC melts the duplex siRNA and separates the antisense from the sense strand during RNA interference in vivo.

Complete unwinding of siRNA duplex is not necessary for RNA interference pathway *in vivo*

ATP-dependent unwinding of the siRNA duplex in the RISC has been proposed to activate the complex to generate RISC*, which is competent to mediate RNAi (166). Although unwinding of siRNA in *Drosophila* embryo lysates has been demonstrated in the presence of ATP, the efficiency of unwinding seems low since only 5% of unwound siRNA was detected (166).

To examine whether or not the siRNA duplex in human cells is completely unwound, we performed RNA interference experiments with siRNA duplexes covalently crosslinked by psoralen photochemistry. Psoralens are bifunctional furocoumarins that intercalate between the base pairs of double-stranded nucleic acids and can photoreact with pyrimidine bases to form monoadducts and

crosslinks (for review see 34). The structure of the psoralen derivative, 4'-(hydroxymethyl)-4,5',8-trimethylpsoralen (HMT) used in this study is shown in Figure 8A. Psoralen crosslinking involves two successive photochemical reactions that take place at the 3,4 or 4',5' double bonds of psoralen (34). Upon long-wave UV irradiation (320–400 nm), the intercalated psoralen can photoreact with adjacent pyrimidine bases to form either furan-side or pyrone-side monoadducts, which are linked to only one strand of the helix (34). By absorbing a second photon, the furan-side monoadducts can be driven into diadducts, which are covalently linked to both strands of the helix (89, 111). Psoralen crosslink formation occurs only when psoralen adds to adjacent and opposite pyrimidine bases in the double helix. The reaction is primarily with uracil in native RNAs, but reactions with cytidine have also been reported (137, 235, 237). Based on psoralen photoreactivity, three possible psoralen crosslink sites in the GFP siRNA duplex are shown in Figure 8B. Note that there is no chance for all three sites to be crosslinked in one RNA.

Unlike the noncrosslinked ds siRNA, the two strands of the crosslinked siRNA duplex could not separate from each other under denaturing conditions so that the crosslinked siRNA duplex showed characteristically retarded mobility in polyacryalmide gel electrophoresis (PAGE) containing 7 M urea (Figure 8C). Crosslinking efficiency depended on the psoralen concentration (Figure 8C, lanes 2 and 3). To further verify the presence of crosslinks in the RNA helix and rule out the possibility of only monoadduct formation, the psoralen crosslinks

were irradiated with short wave UV (254 nm), which showed photoreversal of the crosslinked bonds (Figure 8C, lane 4). The crosslinked siRNA duplex (Figure 8C, lane 3, upper band) was excised from the gel and purified. As control, the noncrosslinked siRNA that was irradiated with long wave UV (360 nm) (Figure 8C, lane 3, lower band) was also purified by the same method. The structures of the purified noncrosslinked and psoralen crosslinked siRNA duplexes were confirmed by PAGE containing 7 M urea (Figure 8C, lanes 5 and 6). Fluorescence imaging of living cells treated with crosslinked siRNA duplex showed that the siRNA duplex's inability to separate on PAGE did not completely abolish its RNA interference activity (Figure 8D, ds-XL). Quantitative analysis of GFP fluorescence intensity indicated that crosslinked siRNA retained 30% of its RNAi activity (Figure 8E, blue line). These results demonstrate that a complete unwinding of the siRNA duplex is not required for gene silencing in vivo (see Discussion).

There is a possibility that the psoralen crosslink of RNA can be photoreversed during transfection, repaired or removed by some unknown mechanism inside the cells, which might cause the partial RNA interference effect in vivo observed in Figures 8D and 8E. To rule out this possibility, we carried out a psoralen crosslinking experiment with siRNA duplex containing biotin at the 3' end of the antisense strand. The crosslinked duplex (ss/as3'-Biotin-XL) was isolated and purified as described above and transfected into HeLa cells by lipofectamine. Biotinylated siRNA was isolated from the cells 30 hr

posttransfection by biotin pull-out assay, SAP treated, and 32P-labeled by T4 PNK as described above. The biotinylated siRNA was still crosslinked (Figure 9, lane 7) at 30 hr posttransfection. When UV irradiated (254 nm), this higher molecular weight siRNA species was converted into two RNA species corresponding to sense and antisense strands (Figure 9, lane 8), indicating the reversibility of the psoralen crosslink. These results show that crosslinked siRNA duplexes can enter the RNAi pathway.

DISCUSSION

By using a quantitative dual fluorescence-based system, we have dissected the kinetics and a number of important parameters involved in the RNAi pathway in cultured human cells. Our results highlight the role of free 5' end hydroxyl groups and the requirement of an A-form helical structure between the antisense strand and the target mRNA. We also found that a complete unwinding of the siRNA helix is not necessary to cause RNAi effects in vivo.

The time-dependent effect of siRNA may reflect a time lag between target mRNA degradation and the half-life of the existing protein expressed from the target gene. This time dependence may also indicate that the siRNAs need to be processed or assembled into an active complex with cellular factors for effective RNA interference.

Although RNA interference lasted at least 66 hr in HeLa cells, quantitative analysis indicated that inhibition by siRNAs did not persist. After reaching maximal activity at 42 hr posttransfection, RNA interference started to decrease at 54 hr, with only 70% inhibition activity at 66 hr. We also found that 5%–10% protein expressed from the genes targeted by siRNA remained at 42 hr posttransfection, but protein amount showed gradual recovery to normal levels between 66 to 90 hr (3 to 4 days) posttransfection (Y.-L.C. and T.M.R., unpublished data). The recovery of target gene expression also indicates that RNA interference by exogenous siRNA duplex does not exist forever in mammalian cells. These findings suggest that the proposed amplification system

driven by RdRP and present in plants and nematodes may not exist or has very little effect on siRNA-mediated gene silencing in mammalian cells.

Recent studies have shown that synthetic siRNAs containing 5'-OH termini can successfully induce RNAi effects in *Drosophila* embryo lysates (58, 166) and cultured mammalian cells (55). A model involving a 5' end kinase activity necessary for RNA interference has been proposed (166). To our knowledge, however, there is no evidence that the 5' end hydroxyl is required for in vivo interference activity. Our results show that replacing the 5' OH, a kinase target site, with amino groups inhibited RNAi activity. Further isolation of siRNA by biotin pull-out experiments revealed that prior phosphatase activity was required for in vitro 5' end radiolabeling by a polynucleotide kinase. Taken together, these results provide strong evidence for the requirement of 5' end kinase activity for RNA interference effects in vivo.

What about a free 3' end for RNAi effects in vivo? An RNA-directed RNA polymerase (RdRP) chain reaction, primed by siRNA, has recently been proposed to amplify the interference effects of a small amount of trigger RNA (reviewed in 164). Lipardi et al. (2001) have shown siRNA-primed RNA synthesis in *Drosophila* embryo lysates and suggested that RNAi in *Drosophila* involves an RdRP where siRNA primes the conversion of target RNA to dsRNA. Further evidence of RdRP involvement in the RNAi pathway in *C. elegans* has been provided in studies (215) showing target RNA-templated synthesis of new dsRNA. These studies highlight the importance of a 3' hydroxyl in priming

subsequent RdRP reactions. An RdRP homolog has not yet been identified in the human genome, suggesting the presence of a separate enzyme that can carry out primer-dependent replication of an RNA template. Our results demonstrate that blocking the 3' position did not significantly affect RNAi activity of siRNA in human cells. Results of our kinetic experiments show that the interference effect lasted only ~4 days, indicating the absence of an amplification mechanism in human cells. In addition, our biotin pull-out experiments show that the 3' end biotin groups on the antisense strand were not efficiently removed during RNAi activities in HeLa cells. Based on these studies, we suggest a model where RNA amplification by RNA-dependent RNA polymerase is not essential for RNA interference in mammalian cell lines.

It is interesting to note that we found no requirement for a perfect A-form helix in siRNA for interference effects in HeLa cells, but an A-form structure was required for antisense-target RNA duplexes. These results suggest an RNAi mechanism where RISC formation does not involve perfect RNA helix recognition, but RISC* assembly requires an A-form helical structure.

The most intriguing results were obtained by crosslinking siRNAs and testing their interference activities in HeLa cells. Psoralen crosslinked siRNA duplexes retained 30% of RNA interference activity. This result can be explained by psoralen photocrosslinking chemistry. There are three possible sites in the GFP siRNA duplex where psoralen can crosslink, yet the crosslinking reaction is not efficient enough to create multiple crosslinks in a single given siRNA duplex

(34, 235). Thus, in the purified crosslinked siRNA duplex population, about one-third had crosslinking at the site near the 5' end of the antisense strand, about one-third had crosslinking in the middle region, and the rest had crosslinking near the 3' end of the antisense strand.

We already showed that accessibility to the 5' termini of the antisense strand is required for efficient RNA interference *in vivo*. 5' phosphorylation of the antisense strand is also required for RNA interference in vitro (166). The cleavage site on target mRNA has been shown to be determined by the 5' end position of the target-recognizing siRNA (58). Based on these findings, we suggest that unwinding of the siRNA duplex would start from the 5' end of the antisense strand, which sets the ruler for target mRNA cleavage. If crosslinking occurred near the 5' end of the antisense strand, it would completely prohibit the unwinding of the siRNA duplex and block access to the 5' termini of the antisense strand, which would completely abolish the RNAi effect. If crosslinking occurred in the middle of the siRNA duplex, near the cleavage site of mRNA, we suggest that, although the siRNA duplex could still undergo some unwinding, this crosslink might interfere with the pairing between target mRNA and the guiding siRNA, thus also blocking the RNAi effect. If crosslinking occurred near the 3' end of the antisense strand, the duplex RNA could unwind, not completely but sufficient for the antisense strand to hybridize to the target mRNA. We have already shown that blocking either the 3' end of the antisense strand or the 5' end of the sense strand has no significant effect on its RNAi activity. It would thus be

reasonable to believe that a siRNA duplex with crosslinking near the 3' end of the antisense strand may still be competent in RNA interference. This hypothesis also explains the remaining 30% RNAi activity in the psoralen-crosslinked siRNA duplex.

These results suggest a possible model for the RNAi pathway in human cells. An RNA-protein complex containing siRNA (RISC) is assembled without the requirement for an A-form RNA helix and/or a free 3'-OH. The 5'-OH of the siRNA duplex is phophorylated by a kinase. During activation of RISC to RISC*, a 5'3' helicase unwinds the RNA duplex to allow hybridization between the antisense strand of siRNA and the target RNA. The requirement of a perfect A-form helix at this stage strongly suggests that another protein (or protein complex) binds this RNA duplex, either in a structural role and/or assisting in the cleavage of mRNA. A complete unwinding of the siRNA duplex is not required for this process, nor can this interference activity be amplified via the 3' end. However, unwinding of the duplex up to the cleavage site may be necessary so that the antisense strand can form an A-form helix with the target strand for further protein interactions. These results also argue against the involvement of RNA amplification mechanism(s) for RNA interference in human cells.

In summary, our results provide new insight into the mechanism of RNAi in mammalian cells, and these results could guide the design of siRNA structures useful in probing biological questions and in functional genomic studies.

Figure 1. Dual fluorescence reporter assay system for RNAi analysis in HeLa cells. (A) Graphical representation of dsRNAs used for targeting GFP mRNA and RFP mRNA. GFP and RFP were encoded by the pEGFP-C1 and pDsRed1-N1 reporter plasmid, respectively. siRNAs were synthesized with 2 nt deoxythymidine overhangs at the 3' end. The position of the first nucleotide of the mRNA target site is indicated relative to the start codon of GFP mRNA or RFP mRNA. The sequence of the antisense strand of siRNA is exactly complementary to the mRNA target site. (B) Fluorescence images showing specific RNA interference effects in living HeLa cells. Fluorescence in living cells was visualized by fluorescence microscopy at 48 hr posttransfection. (a) and (b), images of mock-treated cells (no siRNA added); (c) and (d), images of GFP siRNA-treated cells; (e) and (f), images of RFP siRNA-treated cells. (C) Quantitative analysis of RNAi effects in HeLa cells. Fluorescence emission spectra of GFP and RFP in total cell lysates were detected by exciting at 488 and 568 nm, respectively. (D) Kinetics of RNAi effects in HeLa cells. Ratios of normalized GFP to RFP fluorescence intensity over a 66 hr time course. The fluorescence intensity ratio of target (GFP) to control (RFP) protein was determined in the presence of double-strand (ds) RNA (green bars) and normalized to the ratio observed in the presence of antisense strand (as) RNA (blue bars). Normalized ratios less than 1.0 indicate specific RNA interference. Maximal RNAi effect occurred at 42 hr posttransfection.

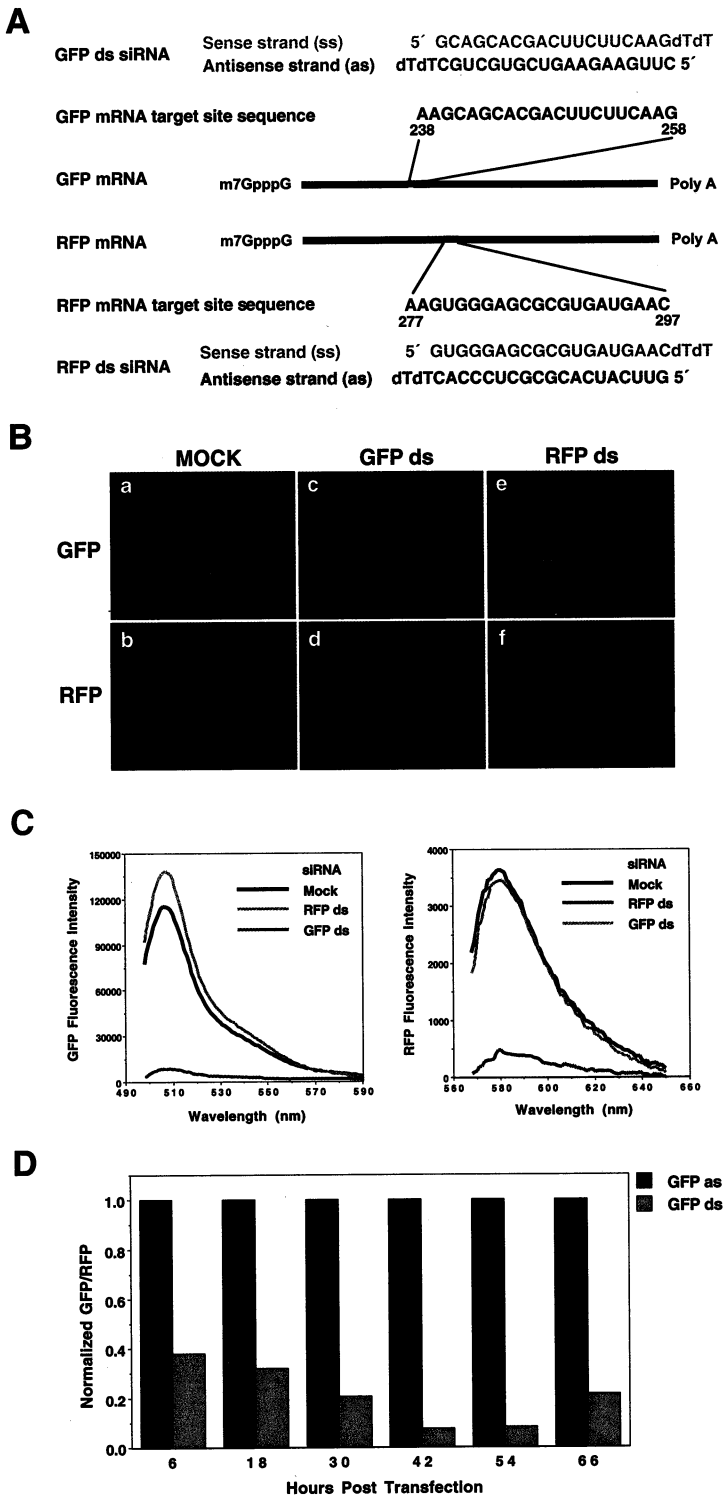

A
GFP ds siRNA
Sense strand (ss) 5′ GCAGCACGACUUCUUCAAGdTdT
Antisense strand (as) dTdTCGUCGUGCUGAAGAAGUUC 5′
GFP mRNA target site sequence
AAGCAGCACGACUUCUUCAAG
238
258
GFP mRNA
m7GpppG
Poly A
RFP mRNA
m7GpppG
Poly A
RFP mRNA target site sequence
AAGUGGGAGCGCGUGAUGAAC
277
297
RFP ds siRNA
Sense strand (ss) 5′ GUGGGAGCGCGUGAUGAACdTdT
Antisense strand (as) dTdTCACCCUCGCGCACUACUUG 5′
B
MOCK
GFP ds
RFP ds
GFP
RFP
a
b
c
d
e
f
C
GFP Fluorescence Intensity
150000
120000
90000
60000
30000
0
490
510
530
550
570
590
Wavelength (nm)
siRNA
Mock
RFP ds
GFP ds
RFP Fluorescence Intensity
4000
3000
2000
1000
0
560
580
600
620
640
660
Wavelength (nm)
siRNA
Mock
RFP ds
GFP ds
D
Normalized GFP/RFP
1.0
0.8
0.6
0.4
0.2
0.0
6
18
30
42
54
66
Hours Post Transfection
GFP as
GFP ds

Figure 2. Analysis of specific RNAi activities by western blotting. Antisense and double-strand RNA are indicated as as and ds, respectively. GFP as ([A], left panel), GFP ds ([A], right panel), RFP as ([B], left panel), or RFP ds ([B], right panel) were cotransfected with pEGFP-C1 and pDsRed1-N1 reporter plasmids into HeLa cells. Cells were harvested at various times, resolved on 10% SDS-PAGE, transferred onto PVDF membranes, and immunoblotted with antibodies against EGFP and DsRed1-N1. The membrane was stripped and reprobed with anti-actin antibody to check for equal loading of total proteins.

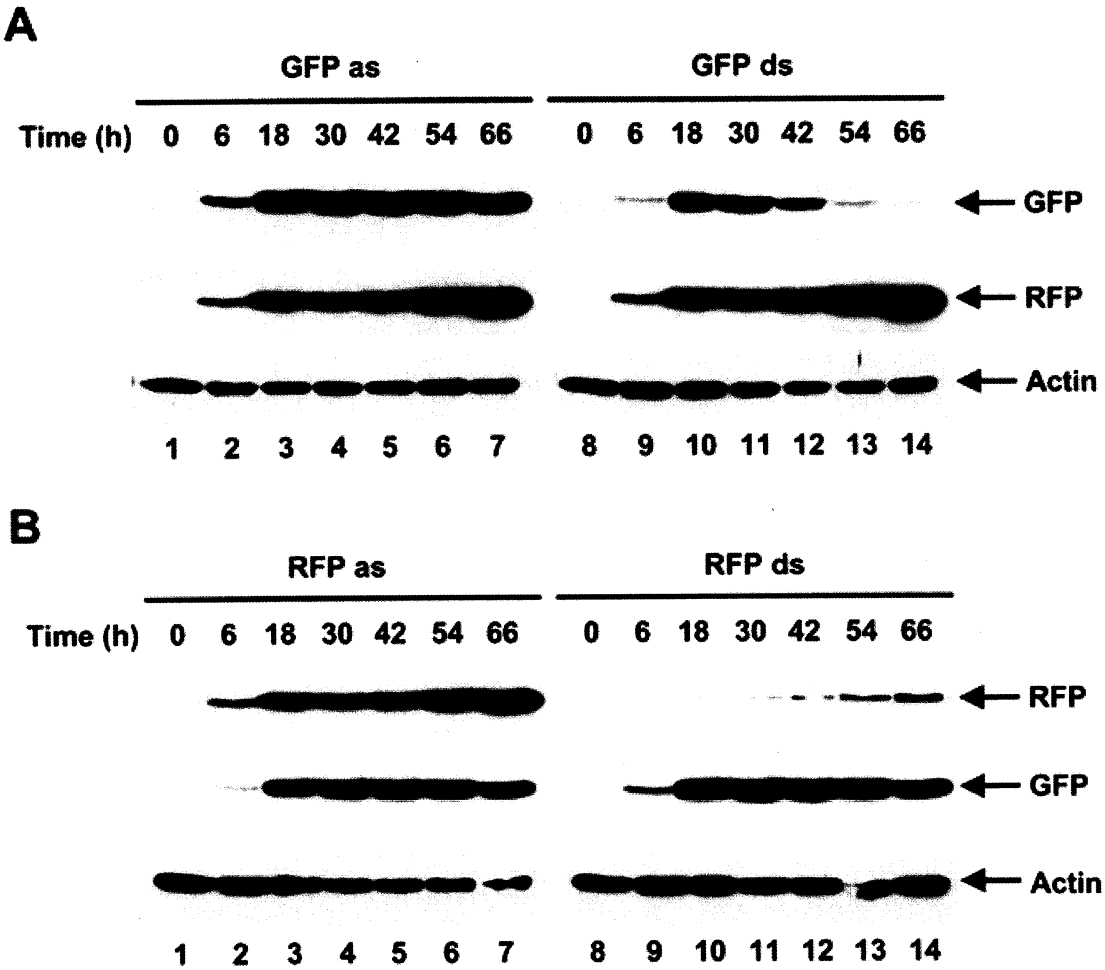

A
GFP as
GFP ds
Time (h) 0 6 18 30 42 54 66 0 6 18 30 42 54 66
GFP
RFP
Actin
1 2 3 4 5 6 7 8 9 10 11 12 13 14
B
RFP as
RFP ds
Time (h) 0 6 18 30 42 54 66 0 6 18 30 42 54 66
RFP
GFP
Actin
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 3. Expression of GFP in HeLa cells treated with antisense or double-stranded siRNA targeting GFP. Transfected cells were harvested at various times after transfection and total cell lysates were analyzed by fluorescence spectroscopy. Fluorescence emission spectra of GFP and RFP were detected by exciting at 488 and 568 nm, respectively.

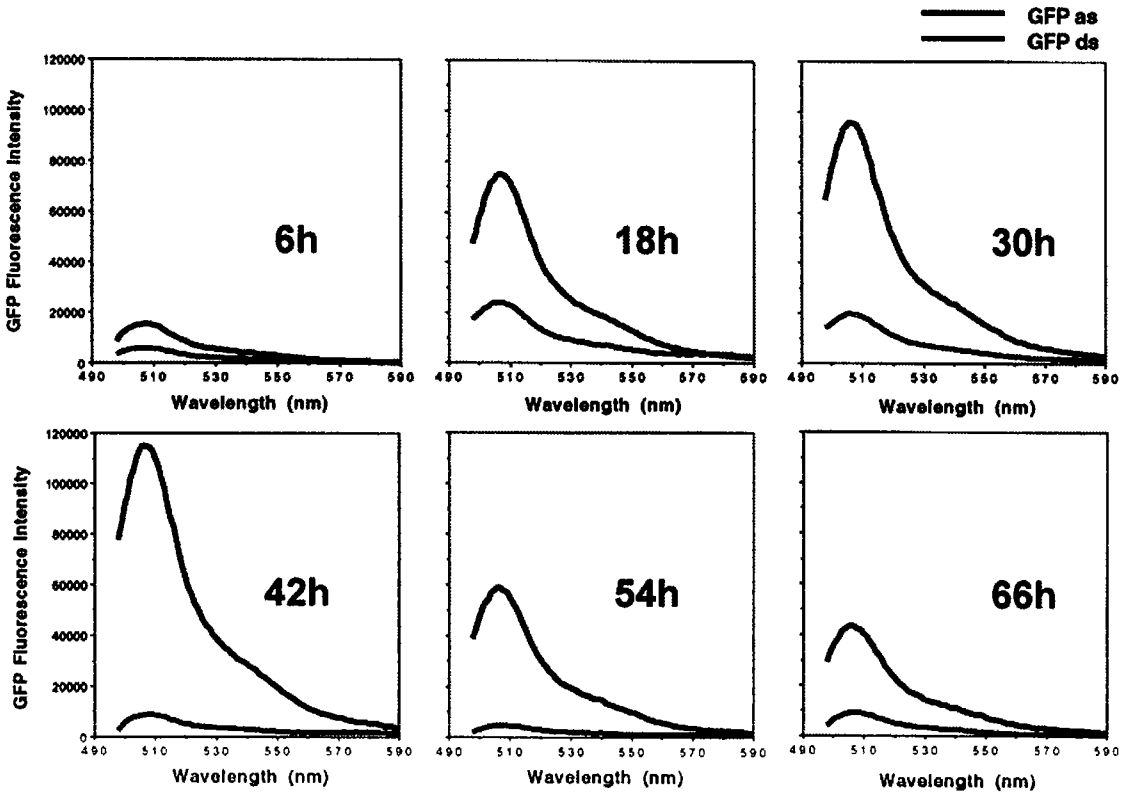
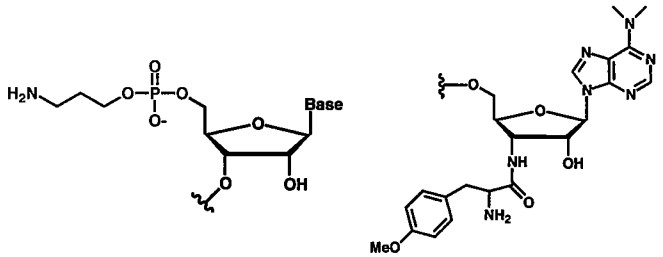

GFP as
GFP ds
GFP Fluorescence Intensity
120000
100000
80000
60000
40000
20000
0
490
510
530
550
570
590
Wavelength (nm)
6h
18h
30h
42h
54h
66h

Figure 4. Modification of GFP siRNA duplexes. (A) Structure of 5'-N3 (amino group with 3-carbon linker, red) and 3'-Pmn (puromycin, blue) modifications. (B) Classification and nomenclature of the modified siRNAs. Sense (top row, purple) and antisense (bottom row, black) strands of siRNA species are shown with their 5'-N3 (red) and 3'-Pmn or biotin (blue) modifications. A dinucleotide internal bulge structure (green) was introduced in sense, antisense, or duplex RNAs.

A



5´- Amino modification = N3 **3´- Puromycin = Pmn**

B

GFP siRNA

	ds (WT)	GCAGCACGACUUCUUCAAGdTdT dTdTCGUCGUGCUGAAGAAGUUC
5´ Modified siRNA	5´-N3ss/as	N3 - GCAGCACGACUUCUUCAAGdTdT dTdTCGUCGUGCUGAAGAAGUUC
	ss/5´-N3as	GCAGCACGACUUCUUCAAGdTdT dTdTCGUCGUGCUGAAGAAGUUC - N3
	5´-N3ss/5´-N3as	N3 - GCAGCACGACUUCUUCAAGdTdT dTdTCGUCGUGCUGAAGAAGUUC - N3
3´ Modified siRNA	ss3´-Pmn/as	GCAGCACGACUUCUUCAAGdTdT - Pmn dTdTCGUCGUGCUGAAGAAGUUC
	ss/as3´-Pmn	GCAGCACGACUUCUUCAAGdTdT Pmn - dTdTCGUCGUGCUGAAGAAGUUC
	ss3´-Pmn/as3´-Pmn	GCAGCACGACUUCUUCAAGdTdT - Pmn Pmn - dTdTCGUCGUGCUGAAGAAGUUC
	ss/as3´-Biotin	GCAGCACGACUUCUUCAAGdTdT Biotin - dTdTCGUCGUGCUGAAGAAGUUC
Bulge-containing siRNA	ss-bulge/as	UG GCAGCACGA CUUCUUCAAGdTdT dTdTCGUCGUGCU GAAGAAGUUC
	ss/as-bulge	GCAGCACGA CUUCUUCAAGdTdT dTdTCGUCGUGCU GAAGAAGUUC CA
	ss-bulge/as-bulge	UG GCAGCACGA CUUCUUCAAGdTdT dTdTCGUCGUGCU GAAGAAGUUC CA

Figure 5. Fluorescence images showing RNA interference effects in living HeLa cells transfected with modified siRNA duplexes. HeLa cells were cotransfected by lipofectamine with pEGFP-C1, pDsRed1-N1 reporter plasmids and siRNA with a 5' modification (c–e), 3' modification (f–i), or internal bulge (j–l). Fluorescence in living cells was visualized at 48 hr posttransfection. GFP fluorescence (left panels) and phase contrast images (right panels) are shown. RNA used in each experiment is indicated on the left of each pair of panels.

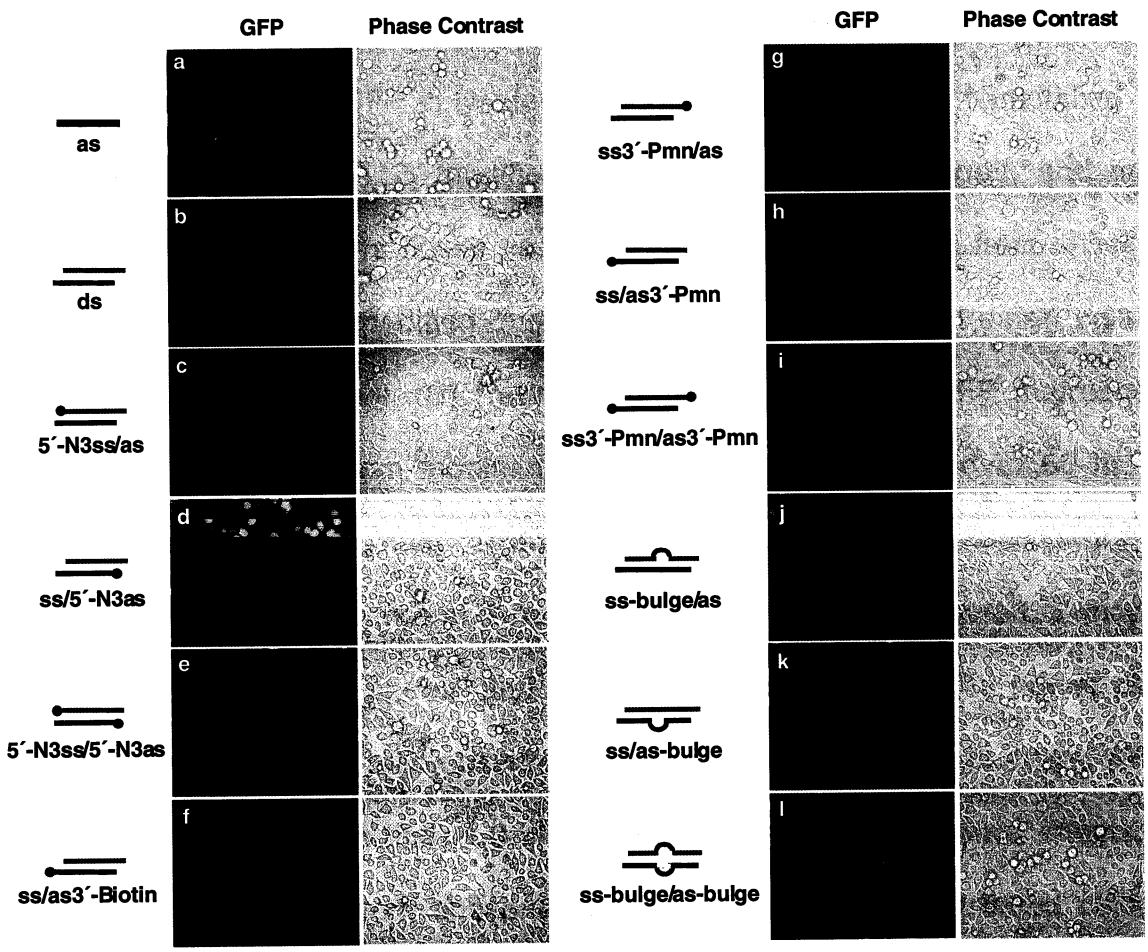

GFP
Phase Contrast
GFP
Phase Contrast
a
b
c
d
e
f
g
h
i
j
k
l
as
ds
5´-N3ss/as
ss/5´-N3as
5´-N3ss/5´-N3as
ss/as3´-Biotin
ss3´-Pmn/as
ss/as3´-Pmn
ss3´-Pmn/as3´-Pmn
ss-bulge/as
ss/as-bulge
ss-bulge/as-bulge

Figure 6. Quantitative analysis of RNAi effects in HeLa cells transfected with modified siRNAs. pEGFP-C1 (as reporter), pDsRed1-N1 (as control) plasmids and 50 nM siRNA were cotransfected into HeLa cells by lipofectamine. Cells were harvested at various times after transfection. Fluorescence emission spectra of GFP and RFP in total cell lysates were detected by exciting at 488 and 568 nm, respectively. (A) GFP emission spectra of modified siRNA-treated cells. Emission spectra of GFP in lysates from cells transfected with 5'-modified GFP siRNAs (upper panel), 3'-modified GFP siRNAs (middle panel), and bulge-containing GFP siRNAs (lower panel). For comparison, results from antisense- (as, red line) and unmodified duplex siRNA- (ds, black line) treated cells are included in each panel. (B) Ratios of normalized GFP to RFP fluorescence intensity in lysates from modified siRNA-treated HeLa cells over 66 hr. The fluorescence intensity ratio of target (GFP) to control (RFP) fluorophore was determined in the presence of 5'-modified GFP siRNAs (upper panel), 3'-modified GFP siRNAs (middle panel), and bulge-containing GFP siRNAs (lower panel) and normalized to the ratio observed in the presence of antisense strand siRNA. Normalized ratios less than 1.0 indicate specific RNA interference effects. For comparison, results from antisense RNA and duplex siRNA-treated cells are included in each panel (as, orange bars; ds, yellow bars).

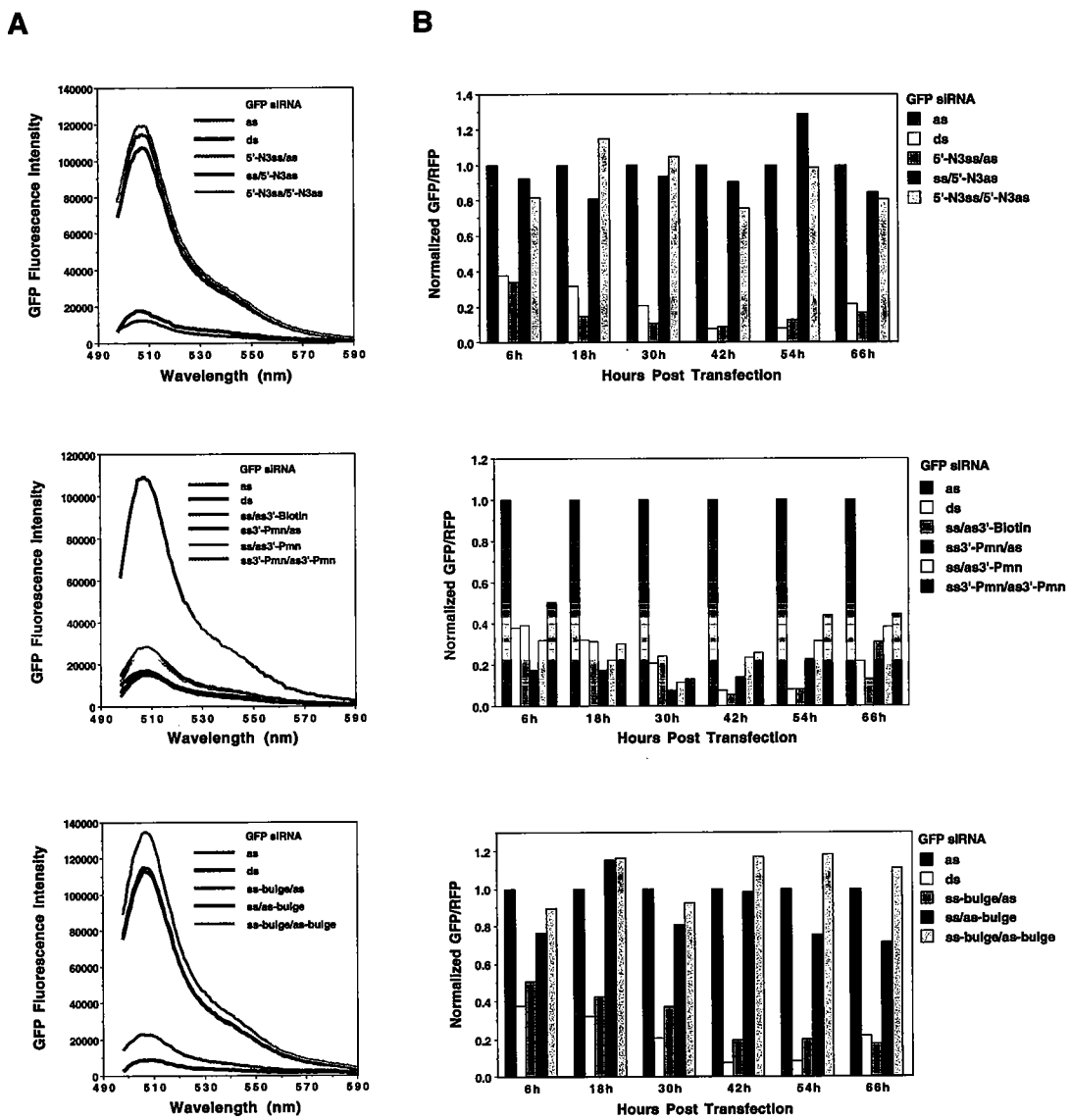

A
B
GFP Fluorescence Intensity
Wavelength (nm)
GFP siRNA
as
ds
5'-N3ss/as
ss/5'-N3as
5'-N3ss/5'-N3as
Normalized GFP/RFP
Hours Post Transfection
6h
18h
30h
42h
54h
66h
ss/as3'-Biotin
ss3'-Pmn/as
as/as3'-Pmn
ss3'-Pmn/as3'-Pmn
ss-bulge/as
as/as-bulge
ss-bulge/as-bulge

Figure 7. Isolation of 5' end phosphorylated and 3' end biotinylated siRNA from HeLa cells. HeLa cells were cotransfected with biotinylated GFP duplex siRNA (ss/as3'-Biotin) and pEGFP-C1 plasmid as described in the Experimental Procedures. The siRNA was isolated by pull-out assay and subjected to phosphatase and kinase reactions (see Experimental Procedures). In brief, streptavidin magnetic beads were used to pull out biotinylated siRNAs from transfected cells, washed to remove unbound RNA, and split into two aliquots. One aliquot was dephosphorylated with shrimp alkaline phosphatase (SAP), and the RNA 5' ends were labeled with 32P by T4 polynucleotide kinase (PNK) reaction. The other aliquot was not dephosphorylated. RNA was resolved on 20% polyacrylamide-7 M Urea gels and visualized by phosphorimager analysis. Lanes 1–3 (marker lanes) contain 5' end-labeled RNA: lane 1, sense strand (ss); lane 2, 3' biotinylated antisense strand (as3'-Biotin); lane 3, heat denatured (10 min at 95°C) siRNA duplex (ss/as3'-Biotin). Lanes 5–14, isolated biotinylated siRNA with SAP treatment (lanes 5–9) or without (lanes 10–14). Lane 4, RNA isolated as above from HeLa cells without siRNA transfection.

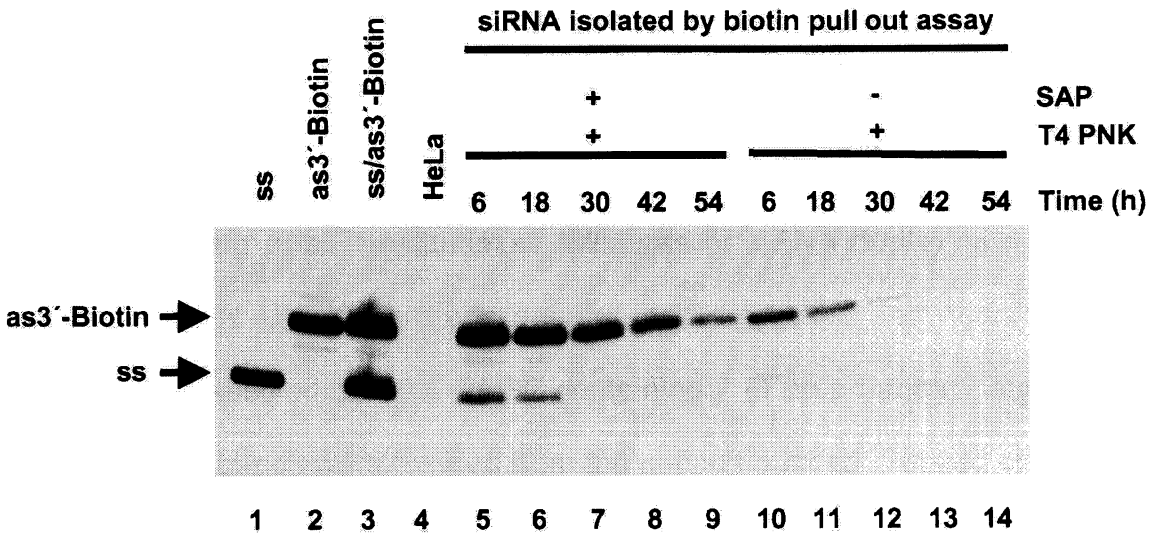

siRNA isolated by biotin pull out assay
ss
as3´-Biotin
ss/as3´-Biotin
HeLa
+
+
-
+
SAP
T4 PNK
6 18 30 42 54 6 18 30 42 54 Time (h)
as3´-Biotin
ss
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 8. RNA interference activities of covalently photocrosslinked duplex RNA in HeLa cells. (A) Structure of a psoralen derivative, 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT), used to crosslink the duplex RNA. (B) Photocrosslinking sites in GFP siRNA. Three preferred sites for psoralen addition to a duplex RNA are shown by cyan letters with red bars indicating the C-U crosslinks formed by UV irradiation in the presence of HMT. (C) Psoralen photocrosslinking of siRNA duplexes. Mixtures of siRNA duplex and psoralen were exposed to UV 360 nm and denatured. Crosslinked and noncrosslinked siRNAs were resolved on 20% PAGE containing 7 M urea (lanes 2 and 3). UV-irradiated RNA bands were excised from the gel and purified. Purified crosslinked dsRNA (ds-XL) and noncrosslinked dsRNA (ds*) are shown in lanes 6 and 5, respectively. To confirm the nature and purity of the crosslink, a portion of the 360 nm UV-irradiated sample (lane 3) was UV irradiated at 254 nm. Photoreversal of psoralen crosslinked siRNA resulted in products with similar electrophoretic mobility to the siRNA duplex without HMT treatment (lane 4). (D) Fluorescence images showing RNA interference effects of psoralen photocrosslinked siRNAs in living HeLa cells. Purified crosslinked ds siRNA (ds-XL, bottom panels) was cotransfected with reporter pEGFP-C1 and control pDsRed1-N1 plasmids into HeLa cells for dual fluorescence reporter assays. Fluorescence (left panels) and phase contrast (right panels) images of living cells were taken 48 hr posttransfection. For comparison, images from noncrosslinked ds siRNA (ds*, middle panels) and antisense siRNA (as, top panels) are also

shown. (E) GFP emission spectra of psoralen photocrosslinked siRNA duplex-treated cells. Cell lysates were prepared from HeLa cells treated with antisense siRNA (as), unmodified UV-irradiated duplex siRNA (ds*), and crosslinked ds siRNA (ds-XL) and analyzed by fluorescence spectroscopy. Fluorescence emission spectra of GFP and RFP were detected by exciting at 488 and 568 nm, respectively. GFP emission spectra are shown normalized to RFP expression.

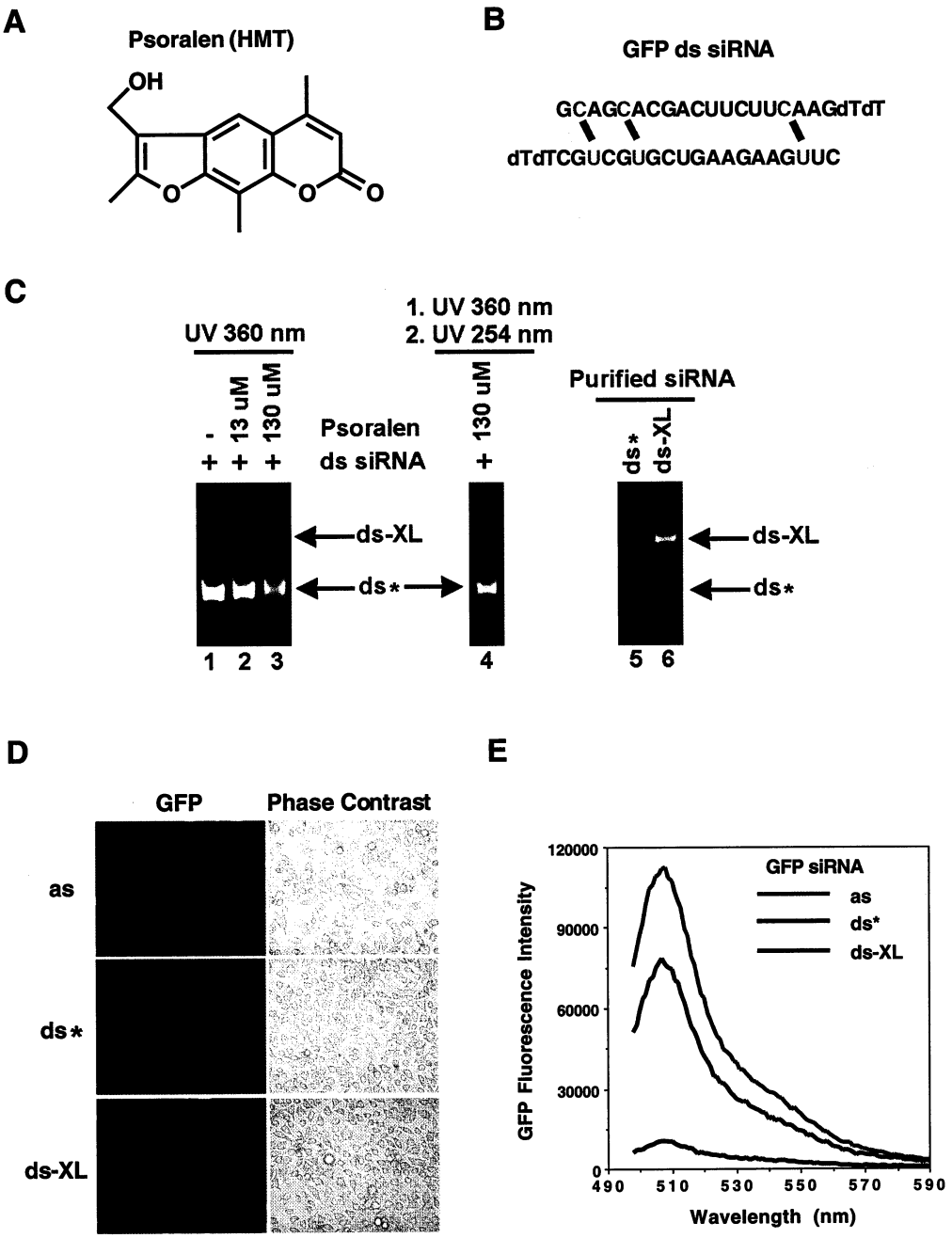
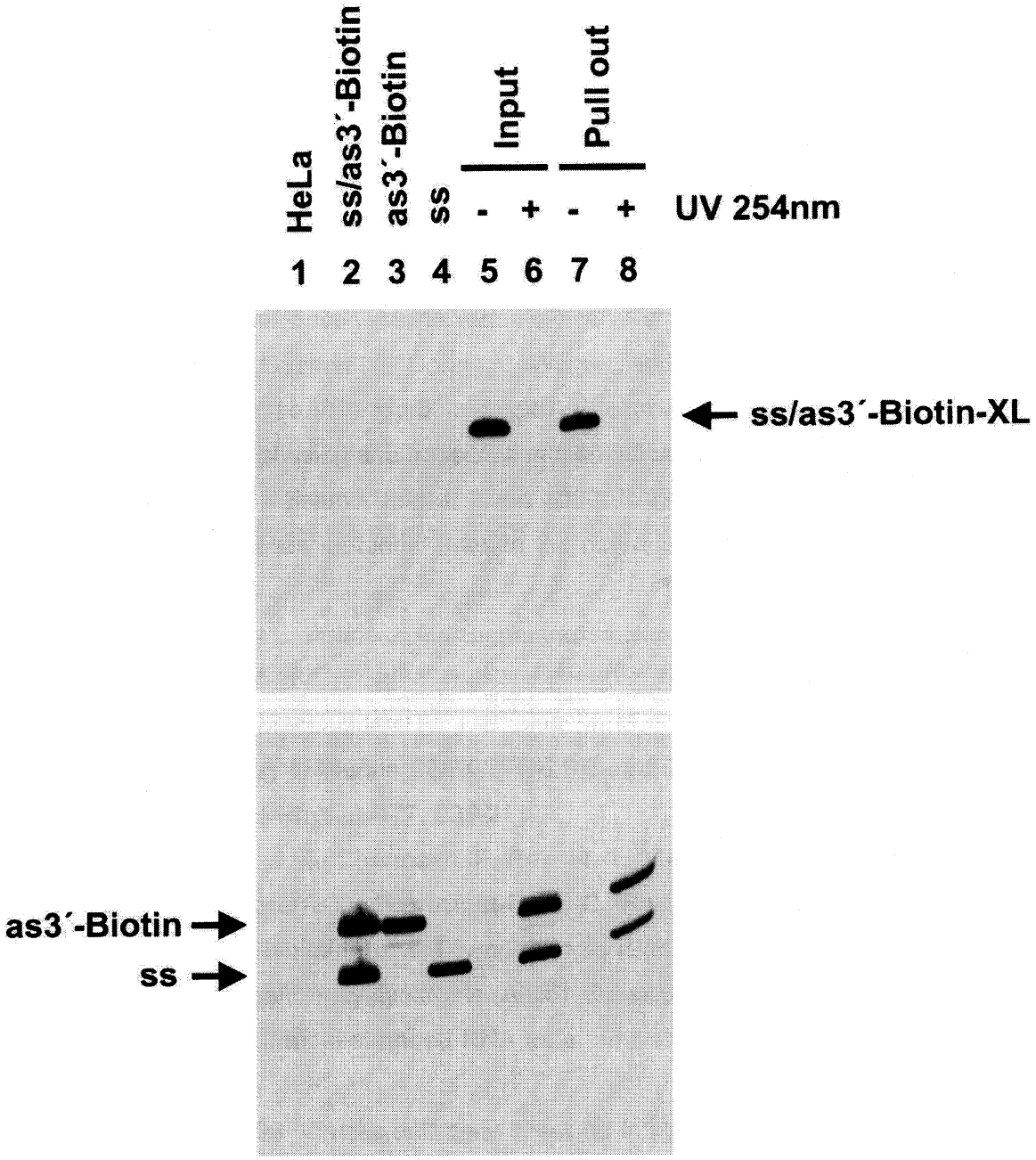

A
Psoralen (HMT)
OH
O
O
O
B
GFP ds siRNA
GCAGCACGACUUCUUCAAGdTdT
dTdTCGUCGUGCUGAAGAAGUUC
C
UV 360 nm
1. UV 360 nm
2. UV 254 nm
- + + +
13 uM
130 uM
Psoralen
ds siRNA
130 uM
+
Purified siRNA
ds*
ds-XL
ds-XL
ds*
ds-XL
ds*
1 2 3
4
5 6
D
GFP
Phase Contrast
as
ds*
ds-XL
E
GFP siRNA
as
ds*
ds-XL
GFP Fluorescence Intensity
120000
90000
60000
30000
0
490 510 530 550 570 590
Wavelength (nm)

Figure 9. Isolation of psoralen-crosslinked siRNA from human cells. siRNA duplexes were conjugated with 3' biotin (ss/as3'-Biotin), psoralen crosslinked, and purified as described in Figure 6 and in the Experimental Procedures. HeLa cells were cotransfected by lipofectamine with crosslinked siRNA (ss/as3'-Biotin-XL) and pEGFP-C1 plasmid, and siRNAs were isolated by biotin pull-out assay at 30 hr posttransfection as described in the Experimental Procedures. In brief, streptavidin-magnetic beads with biotinylated siRNA were subjected to phosphatase treatment and 5' end labeled with 32P. RNA was resolved on 20% polyacrylamide-7 M urea gels and visualized by phosphorimager analysis. Lane 1, RNA from HeLa cells without siRNA transfection. Lane 2, 32P-labeled noncrosslinked siRNA duplex (ss/as3'-Biotin). Lane 3, 32P-labeled 3' biotinylated antisense strand siRNA (as3'-Biotin). Lane 4, 32P-labeled sense strand RNA (ss). Lane 5, 32P-labeled crosslinked siRNA duplex (ss/as3'-Biotin-XL). Lanes 7 and 8, siRNA isolated from HeLa cells treated with crosslinked siRNA duplex (ss/as3'-Biotin-XL). Lanes 6 and 8, UV irradiation (254 nm) of crosslinked siRNA to photoreverse the psoralen crosslinks.


HeLa
ss/as3´-Biotin
as3´-Biotin
ss
Input
Pull out
- + - +
UV 254nm
1 2 3 4 5 6 7 8
ss/as3´-Biotin-XL
as3´-Biotin
ss

REFERENCES

1. Alonso, A., D. Derse, and B. M. Peterlin 1992. Human chromosome 12 is required for optimal interactions between Tat and TAR of human immunodeficiency virus type 1 in rodent cells. J Virol. 66: 4617-4621.
2. Anand, N., S. Murthy, G. Amann, M. Wernick, L. A. Porter, I. H. Cukier, C. Collins, J. W. Gray, J. Diebold, D. J. Demetrick, and J. M. Lee 2002. Protein elongation factor EEF1A2 is a putative oncogene in ovarian cancer. Nat Genet. 31: 301-305.
3. Andrulis, E. D., E. Guzman, P. Doring, J. Werner, and J. T. Lis 2000. High-resolution localization of *Drosophila* spt5 and spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev. 14: 2635-2649.
4. Arregui, C. O., J. Balsamo, and J. Lilien 2000. Regulation of signaling by protein-tyrosine phosphatases: potential roles in the nervous system. Neurochem Res. 25: 95-105.
5. Attisano, L., and J. L. Wrana 2002. Signal transduction by the TGF-beta superfamily. Science. 296: 1646-1647.
6. Bagella, L., T. K. MacLachlan, R. J. Buono, M. M. Pisano, A. Giordano, and A. De Luca 1998. Cloning of murine CDK9/PITALRE and its tissue-specific expression in development. J Cell Physiol. 177: 206-213.
7. Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. 220: 868-871.
8. Bass, B. L. 2000. Double-stranded RNA as a template for gene silencing. Cell. 101: 235-238.
9. Baulcombe, D. 1999. Viruses and gene silencing in plants. Arch Virol Suppl. 15: 189-201.
10. Baulcombe, D. C. 1999. Gene silencing: RNA makes RNA makes no protein. Curr Biol. 9: R599-601.
11. Bauman, A. L., and J. D. Scott 2002. Kinase- and phosphatase-anchoring proteins: harnessing the dynamic duo. Nat Cell Biol. 4: E203-206.

12. Bensaude, O., F. Bonnet, C. Casse, M. F. Dubois, V. T. Nguyen, and B. Palancade 1999. Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD). Biochem Cell Biol. 77: 249-255.
13. Bentley, D. 2002. The mRNA assembly line: transcription and processing machines in the same factory. Curr Opin Cell Biol. 14: 336-342.
14. Berkhout, B., R. H. Silverman, and K. T. Jeang 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. Cell. 59: 273-282.
15. Bernstein, E., A. A. Caudy, S. M. Hammond, and G. J. Hannon 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 409: 363-366.
16. Bieniasz, P. D., T. A. Grdina, H. P. Bogerd, and B. R. Cullen 1999. Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. Proc Natl Acad Sci USA. 96: 7791-7796.
17. Bitko, V., and S. Barik 2001. Phenotypic silencing of cytoplasmic genes using sequence-specific double-stranded short interfering RNA and its application in the reverse genetics of wild type negative-strand RNA viruses. BMC Microbiol. 1: 34-44.
18. Brake, D. A., C. Debouck, and G. Biesecker 1990. Identification of an Arg-Gly-Asp (RGD) cell adhesion site in human immunodeficiency virus type 1 transactivation protein, tat. J Cell Biol. 111: 1275-1281.
19. Bretscher, A., K. Edwards, and R. G. Fehon 2002. ERM proteins and merlin: integrators at the cell cortex. Nat Rev Mol Cell Biol. 3: 586-599.
20. Burchert, A., E. C. Attar, P. McCloskey, Y. W. Fridell, and E. T. Liu 1998. Determinants for transformation induced by the Axl receptor tyrosine kinase. Oncogene. 16: 3177-87.
21. Calegari, F., W. Haubensak, D. Yang, W. B. Huttner, and F. Buchholz 2002. Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. Proc Natl Acad Sci USA. 99: 14236-14240.
22. Capodici, J., K. Kariko, and D. Weissman 2002. Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference. J Immunol. 169: 5196-5201.

23. Chambers, R. S., and M. E. Dahmus 1994. Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. J Biol Chem. 269: 26243-26248.

24. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J Biol Chem. 270: 14962-14969.

25. Chao, S. H., K. Fujinaga, J. E. Marion, R. Taube, E. A. Sausville, A. M. Senderowicz, B. M. Peterlin, and D. H. Price 2000. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. J Biol Chem. 275: 28345-28348.

26. Chao, S. H., and D. H. Price 2001. Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription *in vivo*. J Biol Chem. 28: 28345-28348.

27. Chiu, Y. L., E. Coronel, C. K. Ho, S. Shuman, and T. M. Rana 2001. HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA. J Biol Chem. 276: 12959-12966.

28. Chiu, Y. L., and T. M. Rana 2002. RNAi in human cells: basic structural and functional features of small interfering RNA. Mol Cell. 10: 549-561.

29. Cho, E. J., M. S. Kobor, M. Kim, J. Greenblatt, and S. Buratowski 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev. 15: 3319-3329.

30. Cho, E. J., C. R. Rodriguez, T. Takagi, and S. Buratowski 1998. Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. Genes Dev. 12: 3482-3487.

31. Cho, E. J., T. Takagi, C. R. Moore, and S. Buratowski 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain [see comments]. Genes Dev. 11: 3319-3326.

32. Chodosh, L. A., A. Fire, M. Samuels, and P. A. Sharp 1989. 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription elongation by RNA polymerase II in vitro. J Biol Chem. 264: 2250-2257.

33. Churcher, M. J., C. Lamont, F. Hamy, C. Dingwall, S. M. Green, A. D. Lowe, J. G. Butler, M. J. Gait, and J. Karn 1993. High affinity binding of TAR RNA by the human

immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. J Mol Biol. 230: 90-110.

34. Cimino, G. D., H. B. Gamper, S. T. Isaacs, and J. E. Hearst 1985. Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. Annu Rev Biochem. 54: 1151-1193.
35. Coburn, G. A., and B. R. Cullen 2002. Potent and specific inhibition of human immunodeficiency virus type 1 replication by RNA interference. J Virol. 76: 9225-9231.
36. Cogoni, C., and G. Macino 1999. Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature. 399: 166-169.
37. Cogoni, C., and G. Macino 2000. Post-transcriptional gene silencing across kingdoms. Curr Opin Genet Dev. 10: 638-643.
38. Cogoni, C., and G. Macino 1999. Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. Science. 286: 2342-2344.
39. Conaway, J. W., and R. C. Conaway 1999. Transcription elongation and human disease. Annu Rev Biochem. 68: 301-319.
40. Conaway, J. W., A. Shilatifard, A. Dvir, and R. C. Conaway 2000. Control of elongation by RNA polymerase II. Trends Biochem Sci. 25: 375-380.
41. Coppola, J. A., A. S. Field, and D. S. Luse 1983. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. Proc Natl Acad Sci USA. 80: 1251-1255.
42. Cotgreave, I. A., and R. G. Gerdes 1998. Recent trends in glutathione biochemistry--glutathione-protein interactions: a molecular link between oxidative stress and cell proliferation? Biochem Biophys Res Commun. 242: 1-9.
43. Cullen, B. R. 1998. HIV-1 auxiliary proteins: making connections in a dying cell. [Review] [50 refs]. Cell. 93: 685-692.
44. Cullen, B. R. 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell. 46: 973-982.
45. Dahmus, M. E. 1995. Phosphorylation of the C-terminal domain of RNA polymerase II. Biochim Biophys Acta. 1261: 171-182.
46. Dahmus, M. E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. J Biol Chem. 271: 19009-19012.

47. Dalmay, T., A. Hamilton, S. Rudd, S. Angell, and D. C. Baulcombe 2000. An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell. 101: 543-553.
48. De Falco, G., and A. Giordano 1998. CDK9 (PITALRE): a multifunctional cdc2-related kinase. J Cell Physiol. 177: 501-506.
49. De Luca, A., P. Russo, A. Severino, A. Baldi, T. Battista, I. Cavallotti, L. De Luca, F. Baldi, A. Giordano, and M. G. Paggi 2001. Pattern of expression of cyclin T1 in human tissues. J Histochem Cytochem. 49: 685-692.
50. Deng, L., J. Hagler, and S. Shuman 1996. Factor-dependent release of nascent RNA by ternary complexes of vaccinia RNA polymerase. J Biol Chem. 271: 19556-19562.
51. Deng, L., and S. Shuman 1996. An ATPase component of the transcription elongation complex is required for factor-dependent transcription termination by vaccinia RNA polymerase. J Biol Chem. 271: 29386-29392.
52. Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner 1990. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. Embo J. 9: 4145-4153.
53. Du, L., and S. L. Warren 1997. A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. J Cell Biol. 136: 5-18.
54. Edery, I., L. L. Chu, N. Sonenberg, and J. Pelletier 1995. An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). Mol Cell Biol. 15: 3363-3371.
55. Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 411: 494-498.
56. Elbashir, S. M., J. Harborth, K. Weber, and T. Tuschl 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods. 26: 199-213.
57. Elbashir, S. M., W. Lendeckel, and T. Tuschl 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. 15: 188-200.
58. Elbashir, S. M., J. Martinez, A. Patkaniowska, W. Lendeckel, and T. Tuschl 2001. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. Embo J. 20: 6877-6888.

59. Fagard, M., S. Boutet, J. B. Morel, C. Bellini, and H. Vaucheret 2000. AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci USA. 97: 11650-11654.

60. Fiedler, U., and H. T. Timmers 2001. Analysis of the open region of RNA polymerase II transcription complexes in the early phase of elongation. Nucleic Acids Res. 29: 2706-2714.

61. Figueroa, D. J., J. F. Hess, B. Ky, S. D. Brown, V. Sandig, A. Hermanowski-Vosatka, R. C. Twells, J. A. Todd, and C. P. Austin 2000. Expression of the type I diabetes-associated gene LRP5 in macrophages, vitamin A system cells, and the Islets of Langerhans suggests multiple potential roles in diabetes. J Histochem Cytochem. 48: 1357-1368.

62. Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans [see comments]. Nature. 391: 806-811.

63. Frankel, A. D., and J. A. Young 1998. HIV-1: fifteen proteins and an RNA. Annu Rev Biochem. 67: 1-25.

64. Fraser, A. G., R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, and J. Ahringer 2000. Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature. 408: 325-330.

65. Fresco, L. D., and S. Buratowski 1994. Active site of the mRNA-capping enzyme guanylyltransferase from Saccharomyces cerevisiae: similarity to the nucleotidyl attachment motif of DNA and RNA ligases. Proc Natl Acad Sci USA. 91: 6624-6628.

66. Fu, T. J., J. Peng, G. Lee, D. H. Price, and O. Flores 1999. Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. J Biol Chem. 274: 34527-34530.

67. Fuerst, T. R., and B. Moss 1989. Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells. Importance of the 5' untranslated leader. J Mol Biol. 206: 333-348.

68. Fujinaga, K., T. P. Cujec, J. Peng, J. Garriga, D. H. Price, X. Grana, and B. M. Peterlin 1998. The ability of positive transcription elongation factor B to transactivate

human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. J Virol. 72: 7154-7159.

69. Gallo, R. C., P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch, and M. Popovic 1983. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). Science. 220: 865-867.
70. Garber, M. E., P. Wei, and K. A. Jones 1998. HIV-1 Tat interacts with cyclin T1 to direct the P-TEFb CTD kinase complex to TAR RNA. Cold Spring Harb Symp Quant Biol. 63: 371-380.
71. Garber, M. E., P. Wei, V. N. KewalRamani, T. P. Mayall, C. H. Herrmann, A. P. Rice, D. R. Littman, and K. A. Jones 1998. The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. Genes Dev. 12: 3512-3527.
72. Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. EMBO J. 8: 765-778.
73. Garriga, J., E. Segura, X. Mayol, C. Grubmeyer, and X. Grana 1996. Phosphorylation site specificity of the CDC2-related kinase PITALRE. Biochem J. 320: 983-989.
74. Gitlin, L., S. Karelsky, and R. Andino 2002. Short interfering RNA confers intracellular antiviral immunity in human cells. Nature. 418: 430-434.
75. Gonczy, P., C. Echeverri, K. Oegema, A. Coulson, S. J. Jones, R. R. Copley, J. Duperon, J. Oegema, M. Brehm, E. Cassin, E. Hannak, M. Kirkham, S. Pichler, K. Flohrs, A. Goessen, S. Leidel, A. M. Alleaume, C. Martin, N. Ozlu, P. Bork, and A. A. Hyman 2000. Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. Nature. 408: 331-336.
76. Grana, X., A. De Luca, N. Sang, Y. Fu, P. P. Claudio, J. Rosenblatt, D. O. Morgan, and A. Giordano 1994. PITALRE, a nuclear CDC2-related protein kinase that phosphorylates the retinoblastoma protein in vitro. Proc Natl Acad Sci USA. 91: 3834-3838.
77. Grishok, A., A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D. L. Baillie, A. Fire, G. Ruvkun, and C. C. Mello 2001. Genes and mechanisms related to RNA

interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell. 106: 23-34.

78. Grishok, A., H. Tabara, and C. C. Mello 2000. Genetic requirements for inheritance of RNAi in C. elegans. Science. 287: 2494-2497.
79. Gunnery, S., and M. B. Mathews 1995. Functional mRNA can be generated by RNA polymerase III. Mol Cell Biol. 15: 3597-3607.
80. Guo, S., Y. Yamaguchi, S. Schilbach, T. Wada, J. Lee, A. Goddard, D. French, H. Handa, and A. Rosenthal 2000. A regulator of transcriptional elongation controls vertebrate neuronal development. Nature. 408: 366-369.
81. Hagler, J., and S. Shuman 1992. A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. Science. 255: 983-986.
82. Hamilton, A. J., and D. C. Baulcombe 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. 286: 950-952.
83. Hammond, S. M., E. Bernstein, D. Beach, and G. J. Hannon 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature. 404: 293-296.
84. Hammond, S. M., A. A. Caudy, and G. J. Hannon 2001. Post-transcriptional gene silencing by double-stranded RNA. Nat Rev Genet. 2: 110-119.
85. Harding, H. P., I. Novoa, Y. Zhang, H. Zeng, R. Wek, M. Schapira, and D. Ron 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell. 6: 1099-1108.
86. Harper, J. W., and S. J. Elledge 1998. The role of Cdk7 in CAK function, a retro-retrospective. Genes Dev. 12: 285-289.
87. Hatsell, S., and P. Cowin 2001. Deconstructing desmoplakin. Nat Cell Biol. 3: E270-272.
88. Hausmann, S., and S. Shuman 2002. Characterization of the CTD phosphatase Fcp1 from fission yeast. Preferential dephosphorylation of serine 2 versus serine 5. J Biol Chem. 277: 21213-21220.
89. Hearst, J. E., S. T. Isaacs, D. Kanne, H. Rapoport, and K. Straub 1984. The reaction of the psoralens with deoxyribonucleic acid. Rev Biophys. 17: 1-44.
90. Herrmann, C. H., and A. P. Rice 1995. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal

domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. J Virol. 69: 1612-1620.

91. Herrmann, C. H., and A. P. Rice 1993. Specific interaction of the human immunodeficiency virus Tat proteins with a cellular protein kinase. Virology. 197: 601-608.
92. Hirose, Y., and J. L. Manley 1998. RNA polymerase II is an essential mRNA polyadenylation factor. Nature. 395: 93-96.
93. Ho, C. K., and S. Shuman 1999. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. Mol Cell. 3: 405-411.
94. Ho, C. K., V. Sriskanda, S. McCracken, D. Bentley, B. Schwer, and S. Shuman 1998. The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem. 273: 9577-9585.
95. Hu, W. Y., C. P. Myers, J. M. Kilzer, S. L. Pfaff, and F. D. Bushman 2002. Inhibition of retroviral pathogenesis by RNA interference. Curr Biol. 12: 1301-1311.
96. Hubbard, S. R., and J. H. Till 2000. Protein tyrosine kinase structure and function. Annu Rev of Biochem. 69: 373-398.
97. Huq, I., Y. H. Ping, N. Tamilarasu, and T. M. Rana 1999. Controlling human immunodeficiency virus type 1 gene expression by unnatural peptides. Biochemistry. 38: 5172-5177.
98. Hutvagner, G., and P. D. Zamore 2002. A microRNA in a multiple-turnover RNAi enzyme complex. Science. 297: 2056-2060.
99. Hutvagner, G., and P. D. Zamore 2002. RNAi: nature abhors a double-strand. Curr Opin Genet Dev. 12: 225-232.
100. Hynes, R. O. 2002. A reevaluation of integrins as regulators of angiogenesis. Nat Med. 8: 918-921.
101. Isel, C., and J. Karn 1999. Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. J Mol Biol. 290: 929-941.
102. Ishii, H., R. Baffa, S. I. Numata, Y. Murakumo, S. Rattan, H. Inoue, M. Mori, V. Fidanza, H. Alder, and C. M. Croce 1999. The FEZ1 gene at chromosome 8p22

encodes a leucine-zipper protein, and its expression is altered in multiple human tumors. Proc Natl Acad Sci USA. 96: 3928-3933.

103.Ivanov, D., Y. T. Kwak, J. Guo, and R. B. Gaynor 2000. Domains in the SPT5 Protein That Modulate Its Transcriptional Regulatory Properties. Mol Cell Biol. 20: 2970-2983.

104.Ivanov, D., Y. T. Kwak, E. Nee, J. Guo, L. F. Garcia-Martinez, and R. B. Gaynor 1999. Cyclin T1 domains involved in complex formation with Tat and TAR RNA are critical for tat-activation. J Mol Biol. 288: 41-56.

105.Jacque, J. M., K. Triques, and M. Stevenson 2002. Modulation of HIV-1 replication by RNA interference. Nature. 418: 435-438.

106.Jakobovits, A., D. H. Smith, E. B. Jakobovits, and D. J. Capon 1988. A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. Mol Cell Biol. 8: 2555-2561.

107.Jeang, K. T., and A. Gatignol 1994. Comparison of regulatory features among primate lentiviruses. Curr Top Microbiol Immunol. 188: 123-144.

108.Jones, K. A. 1997. Taking a new TAK on tat transactivation [comment]. Genes Dev. 11: 2593-2599.

109.Jones, K. A., and B. M. Peterlin 1994. Control of RNA initiation and elongation at the HIV-1 promoter. Annu Rev Biochem. 63: 717-743.

110.Kang, M. E., and M. E. Dahmus 1993. RNA polymerases IIA and IIO have distinct roles during transcription from the TATA-less murine dihydrofolate reductase promoter. J Biol Chem. 268: 25033-26040.

111.Kanne, D., K. Straub, H. Rapoport, and J. E. Hearst 1982. Psoralen-deoxyribonucleic acid photoreaction. Characterization of the monoaddition products from 8-methoxypsoralen and 4,5'8-trimethylpsoralen. Biochemistry. 21: 861-871.

112.Kaplan, C. D., J. R. Morris, C. Wu, and F. Winston 2000. Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in *D. melanogaster*. Genes Dev. 14: 2623-2634.

113.Karn, J. 1999. Tackling Tat. J Mol Biol. 293: 235-254.

114.Karn, J., C. Dingwall, M. J. Gait, S. Heaphy, and M. A. Skinner 1991. Regulation of HIV-1 gene expression by the RNA-binding proteins tat and rev. Nucleic Acids Mol Biol. 194-218.

115.Katso, R., K. Okkenhaug, K. Ahmadi, S. White, J. Timms, and M. D. Waterfield 2001. Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. Annu Rev of Cell Dev Biol. 17: 615-75.

116.Keen, N. J., M. J. Gait, and J. Karn 1996. Human immunodeficiency virus type-1 Tat is an integral component of the activated transcription-elongation complex. Proc Natl Acad Sci USA. 93: 2505-2510.

117.Kelland, L. R. 2000. Flavopiridol, the first cyclin-dependent kinase inhibitor to enter the clinic: current status. Expert Opin Investig Drugs. 9: 2903-2911.

118.Kennerdell, J. R., and R. W. Carthew 1998. Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell. 95: 1017-1026.

119.Kephart, D. D., N. F. Marshall, and D. H. Price 1992. Stability of Drosophila RNA polymerase II elongation complexes in vitro. Mol Cell Biol. 12: 2067-2077.

120.Ketting, R. F., S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, and R. H. Plasterk 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. 15: 2654-2659.

121.Ketting, R. F., T. H. Haverkamp, H. G. van Luenen, and R. H. Plasterk 1999. Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell. 99: 133-141.

122.Ketting, R. F., and R. H. Plasterk 2000. A genetic link between co-suppression and RNA interference in C. elegans. Nature. 404: 296-298.

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

124.Kimpton, J., and M. Emerman 1992. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. J Virol. 66: 2232-2239.

125.Kimura, S. H., M. Ikawa, A. Ito, M. Okabe, and H. Nojima 2001. Cyclin G1 is involved in G2/M arrest in response to DNA damage and in growth control after damage recovery. Oncogene. 20: 3290-3300.

126.Kirschmann, D. A., E. A. Seftor, S. F. Fong, D. R. Nieva, C. M. Sullivan, E. M. Edwards, P. Sommer, K. Csiszar, and M. J. Hendrix 2002. A molecular role for lysyl oxidase in breast cancer invasion. Cancer Res. 62: 4478-4483.

127.Klumpp, S., and J. Krieglstein 2002. Serine/threonine protein phosphatases in apoptosis. Curr Opin Pharmacol. 2: 458-462.

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

129.Krumm, A., T. Meulia, M. Brunvand, and M. Groudine 1992. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. Genes Dev. 6: 2201-2213.

130.Kuwahara, H., N. Araki, K. Makino, N. Masuko, S. Honda, K. Kaibuchi, K. Fukunaga, E. Miyamoto, M. Ogawa, and H. Saya 1999. A novel NE-dlg/SAP102-associated protein, p51-nedasin, related to the amidohydrolase superfamily, interferes with the association between NE-dlg/SAP102 and N-methyl-D-aspartate receptor. J Biol Chem. 274: 32204-32214.

131.Lee, N. S., T. Dohjima, G. Bauer, H. Li, M. J. Li, A. Ehsani, P. Salvaterra, and J. Rossi 2002. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. Nat Biotechnol. 20: 500-505.

132.Levy, J. A., A. D. Hoffman, S. M. Kramer, J. A. Landis, J. M. Shimabukuro, and L. S. Oshiro 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science. 225: 840-842.

133.Lewis, D. L., J. E. Hagstrom, A. G. Loomis, J. A. Wolff, and H. Herweijer 2002. Efficient delivery of siRNA for inhibition of gene expression in postnatal mice. Nat Genet. 32: 107-108.

134.Li, J., R. Horwitz, S. McCracken, and J. Greenblatt 1992. NusG, a new *Escherichia coli* elongation factor involved in transcriptional antitermination by the N protein of phage lambda. J Biol Chem. 267: 6012-6019.

135.Lill, R., K. Diekert, A. Kaut, H. Lange, W. Pelzer, C. Prohl, and G. Kispal 1999. The essential role of mitochondria in the biogenesis of cellular iron-sulfur proteins. Biol Chem. 380: 1157-1166.

136.Lipardi, C., Q. Wei, and B. M. Paterson 2001. RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. Cell. 107: 297-307.

137.Lipson, S. E., G. D. Cimino, and J. E. Hearst 1988. Structure of M1 RNA as determined by psoralen cross-linking. Biochemistry. 27: 570-575.

138.Lis, J. T., P. Mason, J. Peng, D. H. Price, and J. Werner 2000. P-TEFb kinase recruitment and function at heat shock loci. Genes Dev. 14: 792-803.

139.Lo, H. J., H. K. Huang, and T. F. Donahue 1998. RNA polymerase I-promoted HIS4 expression yields uncapped, polyadenylated mRNA that is unstable and inefficiently translated in Saccharomyces cerevisiae. Mol Cell Biol. 18: 665-675.

140.Lu, A., A. Gupta, C. Li, T. E. Ahlborn, Y. Ma, E. Y. Shi, and J. Liu 2001. Molecular mechanisms for aberrant expression of the human breast cancer specific gene 1 in breast cancer cells: control of transcription by DNA methylation and intronic sequences. Oncogene. 20: 5173-5185.

141.Lu, H., O. Flores, R. Weinmann, and D. Reinberg 1991. The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc Natl Acad Sci USA. 88: 10004-10008.

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

143.MacLellan, W. R., G. Xiao, M. Abdellatif, and M. D. Schneider 2000. A novel Rb- and p300-binding protein inhibits transactivation by MyoD. Mol Cell Biol. 20: 8903-8915.

144.Mancebo, H. S., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price, and O. Flores 1997. P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev. 11: 2633-2644.

145.Mao, X., B. Schwer, and S. Shuman 1996. Mutational analysis of the Saccharomyces cerevisiae ABD1 gene: cap methyltransferase activity is essential for cell growth. Mol Cell Biol. 16: 475-480.

146.Marshall, N. F., J. Peng, Z. Xie, and D. H. Price 1996. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem. 271: 27176-27183.

147.Marshall, N. F., and D. H. Price 1995. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J Biol Chem. 270: 12335-12338.

148.Martinez, J., A. Patkaniowska, H. Urlaub, R. Luhrmann, and T. Tuschl 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell. 110: 563-574.

149.Massague, J. 2000. How cells read TGF-beta signals. Nat Rev Mol Cell Biol. 1: 169-178.

150.Matzke, M., A. J. Matzke, and J. M. Kooter 2001. RNA: guiding gene silencing. Science. 293: 1080-1083.

151.McCaffrey, A. P., L. Meuse, T. T. Pham, D. S. Conklin, G. J. Hannon, and M. A. Kay 2002. RNA interference in adult mice. Nature. 418: 38-39.

152.McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, S. Foster, S. Shuman, and D. L. Bentley 1997. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II [see comments]. Genes Dev. 11: 3306-3318.

153.McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. D. Patterson, M. Wickens, and D. L. Bentley 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature. 385: 357-361.

154.McManus, M. T., C. P. Petersen, B. B. Haines, J. Chen, and P. A. Sharp 2002. Gene silencing using micro-RNA designed hairpins. RNA. 8: 842-850.

155.McManus, M. T., and P. A. Sharp 2002. Gene silencing in mammals by small interfering RNAs. Nat Rev Genet. 3: 737-747.

156.Milner, R., and I. L. Campbell 2002. The integrin family of cell adhesion molecules has multiple functions within the CNS. J Neurosci Res. 69: 286-291.

157.Miyagishi, M., and K. Taira 2002. U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. Nat Biotechnol. 20: 497-500.

158.Miyake, S., W. R. Sellers, M. Safran, X. Li, W. Zhao, S. R. Grossman, J. Gan, J. A. DeCaprio, P. D. Adams, and W. G. Kaelin, Jr. 2000. Cells degrade a novel inhibitor

of differentiation with E1A-like properties upon exiting the cell cycle. Mol Cell Biol. 20: 8889-8902.

159.Motokura, T., and A. Arnold 1993. Cyclin D and oncogenesis. Curr Opin Genet Dev. 3: 5-10.

160.Mourrain, P., C. Beclin, T. Elmayan, F. Feuerbach, C. Godon, J. B. Morel, D. Jouette, A. M. Lacombe, S. Nikic, N. Picault, K. Remoue, M. Sanial, T. A. Vo, and H. Vaucheret 2000. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell. 101: 533-542.

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

162.Neenhold, H. R., and T. M. Rana 1995. Major groove opening at the HIV-1 Tat binding site of TAR RNA evidenced by a rhodium probe. Biochemistry. 34: 6303-6309.

163.Neuveut, C., and K. T. Jeang 1996. Recombinant human immunodeficiency virus type 1 genomes with tat unconstrained by overlapping reading frames reveal residues in Tat important for replication in tissue culture. J Virol. 70: 5572-5581.

164.Nishikura, K. 2001. A short primer on RNAi: RNA-directed RNA polymerase acts as a key catalyst. Cell. 107: 415-418.

165.Novina, C. D., M. F. Murray, D. M. Dykxhoorn, P. J. Beresford, J. Riess, S. K. Lee, R. G. Collman, J. Lieberman, P. Shankar, and P. A. Sharp 2002. siRNA-directed inhibition of HIV-1 infection. Nat Med. 8: 681-686.

166.Nykanen, A., B. Haley, and P. D. Zamore 2001. ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell. 107: 309-321.

167.O'Keeffe, B., Y. Fong, D. Chen, S. Zhou, and Q. Zhou 2000. Requirement for a kinase-specific chaperone pathway in the production of a Cdk9/cyclin T1 heterodimer responsible for P-TEFb-mediated tat stimulation of HIV-1 transcription. J Biol Chem. 275: 279-287.

168.Orphanides, G., T. Lagrange, and D. Reinberg 1996. The general transcription factors of RNA polymerase II. Genes Dev. 10: 2657-2683.

169.Ortega, S., M. Malumbres, and M. Barbacid 2002. Cyclin D-dependent kinases, INK4 inhibitors and cancer. Biochim Biophys Acta. 1602: 73-87.

170.Palangat, M., T. I. Meier, R. G. Keene, and R. Landick 1998. Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure. Mol Cell. 1: 1033-1042.

171.Parrish, S., J. Fleenor, S. Xu, C. Mello, and A. Fire 2000. Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. Mol Cell. 6: 1077-1087.

172.Paul, C. P., P. D. Good, I. Winer, and D. R. Engelke 2002. Effective expression of small interfering RNA in human cells. Nat Biotechnol. 20: 505-508.

173.Pei, Y., S. Hausmann, C. K. Ho, B. Schwer, and S. Shuman 2001. The length, phosphorylation state, and primary structure of the RNA polymerase II carboxyl-terminal domain dictate interactions with mRNA capping enzymes. J Biol Chem. 276: 28075-28082.

174.Pei, Y., and S. Shuman 2002. Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. J Biol Chem. 277: 19639-19648

175.Peng, J., N. F. Marshall, and D. H. Price 1998. Identification of a cyclin subunit required for the function of Drosophila P-TEFb. J Biol Chem. 273: 13855-13860.

176.Peng, J., Y. Zhu, J. T. Milton, and D. H. Price 1998. Identification of multiple cyclin subunits of human P-TEFb. Genes Dev. 12: 755-762.

177.Pillutla, R. C., Z. Yue, E. Maldonado, and A. J. Shatkin 1998. Recombinant human mRNA cap methyltransferase binds capping enzyme/RNA polymerase IIo complexes. J Biol Chem. 273: 21443-21446.

178.Ping, Y. H., and T. M. Rana 2001. DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. J Biol Chem. 276: 12951-12958.

179.Ping, Y. H., and T. M. Rana 1999. Tat-associated kinase (P-TEFb): a component of transcription preinitiation and elongation complexes. J Biol Chem. 274: 7399-7404.

180.Powis, G., M. Briehl, and J. Oblong 1995. Redox signalling and the control of cell growth and death. Pharmac. Ther. 149-173.

181.Price, D. H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Mol Cell Biol. 20: 2629-2634.

182.Proudfoot, N. 2000. Connecting transcription to messenger RNA processing. Trends Biochem Sci. 25: 290-293.

183.Rana, T. M., and K. T. Jeang 1999. Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. Arch Biochem Biophys. 365: 175-185.

184.Rasmussen, E. B., and J. T. Lis 1993. In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proc Natl Acad Sci USA. 90: 7923-7927.

185.Reines, D., J. W. Conaway, and R. C. Conaway 1996. The RNA polymerase II general elongation factors. Trends Biochem Sci. 21: 351-355.

186.Renner, D. B., Y. Yamaguchi, T. Wada, H. Handa, and D. H. Price 2001. A highly purified RNA polymerase II elongation control system. J Biol Chem. 276: 42601-42609.

187.Rhim, H., C. O. Echetebu, C. H. Herrmann, and A. P. Rice 1994. Wild-type and mutant HIV-1 and HIV-2 Tat proteins expressed in *Escherichia coli* as fusions with glutathione S-transferase. J AIDS. 7: 1116-1121.

188.Richter, S., H. Cao, and T. M. Rana 2002. Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation. Biochemistry. 41: 6391-6397.

189.Richter, S., Y. H. Ping, and T. M. Rana 2002. TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication. Proc Natl Acad Sci USA. 99: 7928-7933.

190.Riisbro, R., I. J. Christensen, T. Piironen, M. Greenall, B. Larsen, R. W. Stephens, C. Han, G. Hoyer-Hansen, K. Smith, N. Brunner, and A. L. Harris 2002. Prognostic significance of soluble urokinase plasminogen activator receptor in serum and cytosol of tumor tissue from patients with primary breast cancer. Clin Cancer Res. 8: 1132-1141.

191.Rodriguez, C. R., E. J. Cho, M. C. Keogh, C. L. Moore, A. L. Greenleaf, and S. Buratowski 2000. Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. Mol Cell Biol. 20: 104-112.

192.Roeder, R. G. 1996. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem Sci. 21: 327-335.

193.Rosen, C. A., J. G. Sodroski, and W. A. Haseltine 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell. 41: 813-823.

194.Rougvie, A. E., and J. T. Lis 1988. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. Cell. 54: 795-804.

195.Saha, N., B. Schwer, and S. Shuman 1999. Characterization of human, Schizosaccharomyces pombe, and Candida albicans mRNA cap methyltransferases and complete replacement of the yeast capping apparatus by mammalian enzymes. J Biol Chem. 274: 16553-16562.

196.Saxton, T. M., B. G. Ciruna, D. Holmyard, S. Kulkarni, K. Harpal, J. Rossant, and T. Pawson 2000. The SH2 tyrosine phosphatase shp2 is required for mammalian limb development. Nat Genet. 24: 420-423.

197.Schlessinger, J. 2002. Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. Cell. 110: 669-672.

198.Schmid, A., B. Schindelholz, and K. Zinn 2002. Combinatorial RNAi: a method for evaluating the functions of gene families in Drosophila. Trends Neurosci. 25: 71-74.

199.Schroeder, S. C., B. Schwer, S. Shuman, and D. Bentley 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev. 14: 2435-2440.

200.Schwer, B., X. Mao, and S. Shuman 1998. Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. Nucleic Acids Res. 26: 2050-2057.

201.Schwer, B., and S. Shuman 1994. Mutational analysis of yeast mRNA capping enzyme. Proc Natl Acad Sci USA. 91: 4328-4332.

202.Sehgal, P. B., J. E. Darnell, Jr., and I. Tamm 1976. The inhibition by DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) of hnRNA and mRNA production in HeLa cells. Cell. 9: 473-480.

203.Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev. 3: 547-558.

204.Selby, M. J., and B. M. Peterlin 1990. *Trans*-activation by HIV-1 Tat via a heterologous RNA binding protein. Cell. 62: 769-776.

205.Senderowicz, A. M. 2000. Small molecule modulators of cyclin-dependent kinases for cancer therapy. Oncogene. 19: 6600-6606.

206.Sharp, P. A. 2001. RNA interference--2001. Genes Dev. 15: 485-490.

207.Shilatifard, A. 1998. Factors regulating the transcriptional elongation activity of RNA polymerase II. FASEB J. 12: 1437-1446.

208.Shim, E. Y., A. K. Walker, Y. Shi, and T. K. Blackwell 2002. CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the C. elegans embryo. Genes Dev. 16: 2135-2146.

209.Shrivastava, A., C. Radziejewski, E. Campbell, L. Kovac, M. McGlynn, T. E. Ryan, S. Davis, M. P. Goldfarb, D. J. Glass, G. Lemke, and G. D. Yancopoulos 1997. An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors. Mol Cell. 1: 25-34.

210.Shuman, S. 1995. Capping enzyme in eukaryotic mRNA synthesis. Prog. Nucleic Acid Res. Mol. Biol. 50: 101-29.

211.Shuman, S. 1997. Origins of mRNA identity: capping enzymes bind to the phosphorylated C- terminal domain of RNA polymerase II. Proc Natl Acad Sci USA. 94: 12758-12760.

212.Shuman, S. 1982. RNA capping by HeLa cell RNA guanylyltransferase. Characterization of a covalent protein-guanylate intermediate. J Biol Chem. 257: 7237-7245.

213.Shuman, S. 2001. Structure, mechanism, and evolution of the mRNA capping apparatus. Prog. Nucleic Acid Res. Mol. Biol. 66: 1-40.

214.Shuman, S., Y. Liu, and B. Schwer 1994. Covalent catalysis in nucleotidyl transfer reactions: essential motifs in Saccharomyces cerevisiae RNA capping enzyme are conserved in Schizosaccharomyces pombe and viral capping enzymes and among polynucleotide ligases. Proc Natl Acad Sci USA. 91: 12046-12050.

215.Sijen, T., and J. M. Kooter 2000. Post-transcriptional gene-silencing: RNAs on the attack or on the defense? Bioessays. 22: 520-531.

216.Simone, C., and A. Giordano 2001. New insight in cdk9 function: from Tat to MyoD. Front Biosci. 6: D1073-1082.

217.Smardon, A., J. M. Spoerke, S. C. Stacey, M. E. Klein, N. Mackin, and E. M. Maine 2000. EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol. 10: 169-178.

218.Stachora, A. A., R. E. Schafer, M. Pohlmeier, G. Maier, and H. Ponstingl 1997. Human Spt5 protein, a putative modulator of chromatin structure, is reversibly phosphorylated in mitosis. FEBS Lett. 409: 74-78.

219.Stark, G. R., I. M. Kerr, B. R. Williams, R. H. Silverman, and R. D. Schreiber 1998. How cells respond to interferons. Annu Rev Biochem. 67: 227-264.

220.Stein, P., P. Svoboda, D. J. Stumpo, P. J. Blackshear, D. B. Lombard, B. Johnson, and R. M. Schultz 2002. Analysis of the role of RecQ helicases in RNAi in mammals. Biochem Biophys Res Commun. 291: 1119-1122.

221.Sui, G., C. Soohoo, B. Affar el, F. Gay, Y. Shi, and W. C. Forrester 2002. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proc Natl Acad Sci USA. 99: 5515-5520.

222.Sumner-Smith, M., S. Roy, R. Barnett, L. S. Reid, R. Kuperman, U. Delling, and N. Sonenberg 1991. Critical chemical features in trans-acting-responsive RNA are required for interaction with human immunodeficiency virus type 1 Tat protein. J Virol. 65: 5196-5202.

223.Svejstrup, J. Q., P. Vichi, and J. M. Egly 1996. The multiple roles of transcription/repair factor TFIIH. Trends Biochem Sci. 21: 346-350.

224.Swanson, M. S., E. A. Malone, and F. Winston 1991. SPT5, an essential gene important for normal transcription in *Saccharomyces cerevisiae*, encodes an acidic nuclear protein with a carboxy-terminal repeat [published erratum appears in Mol Cell Biol 1991 Aug;11(8):4286]. Mol Cell Biol. 11: 3009-3019.

225.Swanson, M. S., and F. Winston 1992. SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in *Saccharomyces cerevisiae*. Genetics. 132: 325-336.

226.Tabara, H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, L. Timmons, A. Fire, and C. C. Mello 1999. The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell. 99: 123-132.

227.Tabara, H., E. Yigit, H. Siomi, and C. C. Mello 2002. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH- box helicase to direct RNAi in C. elegans. Cell. 109: 861-871.

228.Takagi, T., C. R. Moore, F. Diehn, and S. Buratowski 1997. An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. Cell. 89: 867-873

229.Takai, S., M. Yamada, T. Araki, H. Koshimizu, H. Nawa, and H. Hatanaka 2002. Shp-2 positively regulates brain-derived neurotrophic factor-promoted survival of cultured ventral mesencephalic dopaminergic neurons through a brain immunoglobulin-like molecule with tyrosine-based activation motifs/Shp substrate-1. J Neurochem. 82: 353-364.

230.Takayama, S., and J. C. Reed 2001. Molecular chaperone targeting and regulation by BAG family proteins. Nat Cell Biol. 3: E237-241.

231.Tamm, I., and T. Kikuchi 1979. Early termination of heterogeneous nuclear RNA transcripts in mammalian cells: accentuation by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole. Proc Natl Acad Sci USA. 76: 5750-5754.

232.Tamm, I., T. Kikuchi, J. E. Darnell, Jr., and M. Salditt-Georgieff 1980. Short capped hnRNA precursor chains in HeLa cells: continued synthesis in the presence of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole. Biochemistry. 19: 2743-2748.

233.Taube, R., K. Fujinaga, J. Wimmer, M. Barboric, and B. M. Peterlin 1999. Tat transactivation: a model for the regulation of eukaryotic transcriptional elongation. Virology. 264: 245-253.

234.Thiery, J. P. 2002. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer. 2: 442-454.

235.Thompson, J. F., and J. E. Hearst 1983. Structure of E. coli 16S RNA elucidated by psoralen crosslinking. Cell. 32: 1355-1365.

236.Turner, B. G., and M. F. Summers 1999. Structural biology of HIV. J Mol Biol. 285: 1-32.

237.Turner, S., and H. F. Noller 1983. Identification of sites of 4'-(hydroxymethyl)-4,5',8-trimethylpsoralen cross-linking in Escherichia coli 23S ribosomal ribonucleic acid. Biochemistry. 22: 4159-4164.

238.Tuschl, T. 2001. RNA Interference and Small Interfering RNAs. Chembiochem Europ J Chem Biol. 2: 239-245.

239.Tuschl, T., P. D. Zamore, R. Lehmann, D. P. Bartel, and P. A. Sharp 1999. Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev. 13: 3191-3197.

240.Uptain, S. M., C. M. Kane, and M. J. Chamberlin 1997. Basic mechanisms of transcript elongation and its regulation. Annu Rev Biochem. 66: 117-172.

241.Wada, T., G. Orphanides, J. Hasegawa, D. K. Kim, D. Shima, Y. Yamaguchi, A. Fukuda, K. Hisatake, S. Oh, D. Reinberg, and H. Handa 2000. FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH. Mol Cell. 5: 1067-1072.

242.Wada, T., T. Takagi, Y. Yamaguchi, A. Ferdous, T. Imai, S. Hirose, S. Sugimoto, K. Yano, G. A. Hartzog, F. Winston, S. Buratowski, and H. Handa 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. 12: 343-356.

243.Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa 1998. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. Embo J. 17: 7395-7403.

244.Wahlberg, K., G. Hoyer-Hansen, and B. Casslen 1998. Soluble receptor for urokinase plasminogen activator in both full-length and a cleaved form is present in high concentration in cystic fluid from ovarian cancer. Cancer Res. 58: 3294-3298.

245.Wang, S. P., and S. Shuman 1997. Structure-function analysis of the mRNA cap methyltransferase of Saccharomyces cerevisiae. J Biol Chem. 272: 14683-14689.

246.Wang, Z., and T. M. Rana 1996. RNA conformation in the Tat-TAR complex determined by site-specific photo-cross-linking. Biochemistry. 35: 6491-6499.

247.Wang, Z., and T. M. Rana 1998. RNA-protein interactions in the Tat-trans-activation response element complex determined by site-specific photo-cross-linking. Biochemistry. 37: 4235-4243.

248.Waterhouse, P. M., M. B. Wang, and E. J. Finnegan 2001. Role of short RNAs in gene silencing. Trends Plant Sci. 6: 297-301.

249.Waterhouse, P. M., M. B. Wang, and T. Lough 2001. Gene silencing as an adaptive defence against viruses. Nature. 411: 834-842.

250.Weeks, K. M., and D. M. Crothers 1993. Major groove accessibility of RNA. Science. 261: 1574-1577.

251.Weeks, K. M., and D. M. Crothers 1991. RNA recognition by Tat-derived peptides: interaction in the major groove? Cell. 66: 577-588.

252.Wei, P., M. E. Garber, S. M. Fang, W. H. Fischer, and K. A. Jones 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell. 92: 451-462.

253.Wen, Y., and A. J. Shatkin 1999. Transcription elongation factor hSPT5 stimulates mRNA capping. Genes Dev. 13: 1774-1779.

254.Whitehouse, C., J. Chambers, K. Howe, M. Cobourne, P. Sharpe, E. Solomon, H. Ishii, R. Baffa, S. I. Numata, Y. Murakumo, S. Rattan, H. Inoue, M. Mori, V. Fidanza, H. Alder, and C. M. Croce 2002. NBR1 interacts with fasciculation and elongation protein zeta-1 (FEZ1) and calcium and integrin binding protein (CIB) and shows developmentally restricted expression in the neural tube. Eur J Biochem. 269: 538-545.

255.Wimmer, J., K. Fujinaga, R. Taube, T. P. Cujec, Y. Zhu, J. Peng, D. H. Price, and B. M. Peterlin 1999. Interactions between Tat and TAR and human immunodeficiency virus replication are facilitated by human cyclin T1 but not cyclins T2a or T2b. Virology. 255: 182-189.

256.Wu-Baer, F., W. S. Lane, and R. B. Gaynor 1998. Role of the human homolog of the yeast transcription factor SPT5 in HIV-1 Tat-activation. J Mol Biol. 277: 179-197.

257.Wulf, P., and U. Suter 1999. Embryonic expression of epithelial membrane protein 1 in early neurons. Brain Res Dev Brain Res. 116: 169-180.

258.Xia, H., Q. Mao, H. L. Paulson, and B. L. Davidson 2002. siRNA-mediated gene silencing in vitro and in vivo. Nat Biotechnol. 20: 1006-1010.

259.Yamada, M. K., K. Nakanishi, S. Ohba, T. Nakamura, Y. Ikegaya, N. Nishiyama, and N. Matsuki 2002. Brain-derived neurotrophic factor promotes the maturation of GABAergic mechanisms in cultured hippocampal neurons. J Neurosci. 22: 7580-7585.

260.Yamaguchi, Y., J. Filipovska, K. Yano, A. Furuya, N. Inukai, T. Narita, T. Wada, S. Sugimoto, M. M. Konarska, and H. Handa 2001. Stimulation of RNA polymerase II elongation by hepatitis delta antigen. Science. 293: 124-127.

261.Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, J. Hasegawa, and H. Handa 1999. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell. 97: 41-51.

262.Yamaguchi, Y., T. Wada, and H. Handa 1998. Interplay between positive and negative elongation factors: drawing a new view of DRB. Genes Cells. 3: 9-15.

263.Yamaguchi, Y., T. Wada, D. Watanabe, T. Takagi, J. Hasegawa, and H. Handa 1999. Structure and function of the human transcription elongation factor DSIF. J Biol Chem. 274: 8085-8092.

264.Yang, D., H. Lu, and J. W. Erickson 2000. Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in Drosophila embryos. Curr Biol. 10: 1191-1200.

265.Yang, X., M. O. Gold, D. N. Tang, D. E. Lewis, C. E. Aguilar, A. P. Rice, and C. H. Herrmann 1997. TAK, an HIV Tat-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. Proc Natl Acad Sci USA. 94: 12331-12336.

266.Yi, H., C. C. Morton, S. Weremowicz, O. W. McBride, and K. Kelly 1995. Genomic organization and chromosomal localization of the DUSP2 gene, encoding a MAP kinase phosphatase, to human 2p11.2-q11. Genomics. 28: 92-96.

267.Yu, J. Y., S. L. DeRuiter, and D. L. Turner 2002. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. Proc Natl Acad Sci USA. 99: 6047-6052.

268.Yue, X., P. Favot, T. L. Dunn, A. I. Cassady, and D. A. Hume 1993. Expression of mRNA encoding the macrophage colony-stimulating factor receptor (c-fms) is controlled by a constitutive promoter and tissue-specific transcription elongation. Mol Cell Biol. 13: 3191-3201.

269.Yue, Z., E. Maldonado, R. Pillutla, H. Cho, D. Reinberg, and A. J. Shatkin 1997. Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II [see comments]. Proc Natl Acad Sci USA. 94: 12898-12903.

270.Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc Natl Acad Sci USA. 93: 6975-6980.

271.Zamore, P. D., T. Tuschl, P. A. Sharp, and D. P. Bartel 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell. 101: 25-33.

272.Zeng, Y., E. J. Wagner, and B. R. Cullen 2002. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol Cell. 9: 1327-1333.

273.Zhang, J., N. Tamilarasu, S. Hwang, M. E. Garber, I. Huq, K. A. Jones, and T. M. Rana 2000. HIV-1 TAR RNA enhances the interaction between tat and cyclin T1. J Biol Chem. 275: 34314-34319.

274.Zhou, C., and T. M. Rana 2002. A bimolecular mechanism of HIV-1 Tat protein interaction with RNA polymerase II transcription elongation complexes. J Mol Biol. 320: 925-942.

275.Zhou, M., M. A. Halanski, M. F. Radonovich, F. Kashanchi, J. Peng, D. H. Price, and J. N. Brady 2000. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Mol Cell Biol. 20: 5077-5086.

276.Zhou, Q., D. Chen, E. Pierstorff, and K. Luo 1998. Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. EMBO J. 17: 3681-3691.

277.Zhu, Y., T. Pe'ery, J. Peng, Y. Ramanathan, N. Marshall, T. Marshall, B. Amendt, M. B. Mathews, and D. H. Price 1997. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev. 11: 2622-2632.

278.Zorio, D. A., and D. L. Bentley 2001. Transcription elongation: the 'Foggy' is lifting... Curr Biol. 11: R144-146.

APPENDIX

Table 1. Genomewide analysis of gene expression in P-TEFb knockdown HeLa cells. Total mRNA from HeLa cells treated with and without siRNA directed against hCycT1 (Cyclin T Δ) or CDK9 (CDK9 Δ) at 48h post transfection and the expression of various genes were analyzed by using the GeneChip® Human Genome U133 (HG-U133) from Affymetrix. For comparison analysis, data from cells treated with nonrelated ds siRNA (RFP ds) were used as baseline. A Change *p*-value is calculated indicating an increase (I), decrease (D) or no change in gene expression. The Signal Log Ratio (base on 2) estimates the magnitude and direction of change of a transcript when two arrays are compared. A complete list of down- and up-regulated genes with names and accession numbers is shown here.

Table 1.

ID #	CyclinT Δ		CDK9 Δ		Description
	Signal log ratio	Change	Signal log ratio	Change	
230465_at	-4.5	D	-1.02	D	Consensus includes gb:AI806674 /FEA=EST /DB_XREF=gi:5393240 /DB_XREF=est:wf35d02.x1 /CLONE=IMAGE:2357571 /UG=Hs.121898 ESTs
204379_s_at	-3.18	D	-3.36	D	gb:NM_000142.2 /DEF=Homo sapiens fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism) (FGFR3), transcript variant 1, mRNA. /FEA=mRNA /GEN=FGFR3 /PROD=fibroblast growth factor receptor 3, isoform 1precursor /DB_XREF=gi:13112046 /UG=Hs.1420 fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism) /FL=gb:NM_000142.2 gb:M58051.1
234986_at	-2.96	D	-1.57	D	
229256_at	-2.88	D	-2.52	D	
225710_at	-2.75	D	-2.68	D	
203650_at	-2.56	D	-1.85	D	
226625_at	-2.5	D	-1.17	D	
203153_at	-2.49	D	-1.23	D	gb:NM_001548.1 /DEF=Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), mRNA. /FEA=mRNA /GEN=IFIT1 /PROD=interferon-induced protein withtetratricopeptide repeats 1 /DB_XREF=gi:4504584 /UG=Hs.20315 interferon-induced protein with tetratricopeptide repeats 1 /FL=gb:M24594.1 gb:NM_001548.1
201940_at	-2.36	D	-1.4	D	
201324_at	-2.35	D	-1.37	D	gb:NM_001423.1 /DEF=Homo sapiens epithelial membrane protein 1 (EMP1), mRNA. /FEA=mRNA /GEN=EMP1 /PROD=epithelial membrane protein 1 /DB_XREF=gi:4503558 /UG=Hs.79368 epithelial membrane protein 1 /FL=gb:U77085.1 gb:U43916.1 gb:NM_001423.1
224893_at	-2.31	D	-1.33	D	
204731_at	-2.29	D	-1.04	D	
204396_s_at	-2.27	D	-1.09	D	
203646_at	-2.24	D	-1.21	D	gb:NM_004109.2 /DEF=Homo sapiens ferredoxin 1 (FDX1), nuclear gene encoding mitochondrial protein, mRNA. /FEA=mRNA /GEN=FDX1 /PROD=ferredoxin 1 precursor /DB_XREF=gi:13677224 /UG=Hs.744 ferredoxin 1 /FL=gb:NM_004109.2 gb:J03548.1 gb:M18003.1 gb:M34788.1
35148_at	-2.21	D	-2.71	D	Cluster Incl. AC005954:Homo sapiens chromosome 19, cosmid R28784 /cds=(0,2858) /gb=AC005954 /gi=3851201 /ug=Hs.25527 /len=2859
222692_s_at	-2.19	D	-1.85	D	
225847_at	-2.16	D	-1.89	D	

200600_at	-2.12	D	-1.07	D	gb:NM_002444.1 /DEF=Homo sapiens moesin (MSN), mRNA. /FEA=mRNA /GEN=MSN /PROD=moesin /DB_XREF=gi:4505256 /UG=Hs.170328 moesin /FL=gb:M69066.1 gb:NM_002444.1
238417_at	-2.12	D	-1.79	D	
214121_x_at	-2.05	D	-0.97	D	
225479_at	-2.05	D	-1.45	D	
201941_at	-2.04	D	-0.99	D	
225032_at	-2.04	D	-1.7	D	
202686_s_at	-2.03	D	-2.81	D	gb:NM_021913.1 /DEF=Homo sapiens AXL receptor tyrosine kinase (AXL), transcript variant 1, mRNA. /FEA=mRNA /GEN=AXL /PROD=AXL receptor tyrosine kinase isoform 1precursor /DB_XREF=gi:11863122 /UG=Hs.83341 AXL receptor tyrosine kinase /FL=gb:NM_021913.1
212993_at	-2.01	D	-1.4	D	
225179_at	-2.01	D	-1.29	D	
218618_s_at	-1.99	D	-1.79	D	gb:NM_022763.1 /DEF=Homo sapiens hypothetical protein FLJ23399 (FLJ23399), mRNA. /FEA=mRNA /GEN=FLJ23399 /PROD=hypothetical protein FLJ23399 /DB_XREF=gi:12232434 /UG=Hs.299883 hypothetical protein FLJ23399 /FL=gb:NM_022763.1
224989_at	-1.97	D	-3.09	D	Consensus includes gb:AI824013 /FEA=EST /DB_XREF=gi:5444684 /DB_XREF=est:wj29e07.x1 /CLONE=IMAGE:2404260 /UG=Hs.122489 Homo sapiens cDNA FLJ13289 fis, clone OVARC1001170
209877_at	-1.95	D	-1.11	D	gb:AF010126.1 /DEF=Homo sapiens breast cancer-specific protein 1 (BCSG1) mRNA, complete cds. /FEA=mRNA /GEN=BCSG1 /PROD=BCSG1 protein /DB_XREF=gi:2281473 /UG=Hs.63236 synuclein, gamma (breast cancer-specific protein 1) /FL=gb:AF010126.1 gb:AF017256.1 gb:NM_003087.1
228745_at	-1.92	D	-1.14	D	
200831_s_at	-1.91	D	-1.33	D	
229553_at	-1.89	D	-1.07	D	
205619_s_at	-1.88	D	-1.4	D	
212731_at	-1.86	D	-1.73	D	
213372_at	-1.86	D	-1	D	
201490_s_at	-1.85	D	-0.97	D	
207357_s_at	-1.84	D	-1	D	
238320_at	-1.83	D	-1.23	D	Consensus includes gb:AV659198 /FEA=EST /DB_XREF=gi:9880212 /DB_XREF=est:AV659198 /CLONE=GLCFUE07 /UG=Hs.11367 ESTs
211162_x_at	-1.8	D	-2.16	D	
226685_at	-1.8	D	-1.41	D	

202342_s_at	-1.79	D	-1.35	D	gb:NM_015271.1 /DEF=Homo sapiens tripartite motif protein TRIM2 (KIAA0517), mRNA. /FEA=mRNA /GEN=KIAA0517 /PROD=tripartite motif protein TRIM2 /DB_XREF=gi:13446226 /UG=Hs.12372 tripartite motif protein TRIM2 /FL=gb:AF220018.1 gb:NM_015271.1
236262_at	-1.79	D	-1.38	D	
226785_at	-1.79	D	-1.65	D	
213012_at	-1.78	D	-1.36	D	
225319_s_at	-1.76	D	-0.98	D	
201943_s_at	-1.75	D	-1.07	D	
232125_at	-1.74	D	-1.92	D	
208867_s_at	-1.73	D	-0.96	D	
225266_at	-1.73	D	-1.82	D	
208051_s_at	-1.71	D	-1.66	D	
227052_at	-1.71	D	-2.72	D	
226982_at	-1.71	D	-1.77	D	
212919_at	-1.7	D	-0.96	D	
201389_at	-1.69	D	-1.38	D	gb:NM_002205.1 /DEF=Homo sapiens integrin, alpha 5 (fibronectin receptor, alpha polypeptide) (ITGA5), mRNA. /FEA=mRNA /GEN=ITGA5 /PROD=integrin alpha 5 precursor /DB_XREF=gi:4504750 /UG=Hs.149609 integrin, alpha 5 (fibronectin receptor, alpha polypeptide) /FL=gb:NM_002205.1
203231_s_at	-1.68	D	-1.02	D	
209406_at	-1.67	D	-0.98	D	gb:AF095192.1 /DEF=Homo sapiens BAG-family molecular chaperone regulator-2 mRNA, complete cds. /FEA=mRNA /PROD=BAG-family molecular chaperone regulator-2 /DB_XREF=gi:4322819 /UG=Hs.55220 BCL2-associated athanogene 2 /FL=gb:AF095192.1 gb:AL050287.1 gb:NM_004282.2
201681_s_at	-1.64	D	-1	D	
227112_at	-1.64	D	-1.56	D	
201971_s_at	-1.61	D	-1.1	D	gb:NM_001690.1 /DEF=Homo sapiens ATPase, H+ transporting, lysosomal (vacuolar proton pump), alpha polypeptide, 70kD, isoform 1 (ATP6A1), mRNA. /FEA=mRNA /GEN=ATP6A1 /PROD=ATPase, H+ transporting, lysosomal (vacuolarproton pump), alpha polypeptide, 70kD, isoform 1 /DB_XREF=gi:4502304 /UG=Hs.281866 ATPase, H+ transporting, lysosomal (vacuolar proton pump), alpha polypeptide, 70kD, isoform 1 /FL=gb:L09235.1 gb:NM_001690.1 gb:AF113129.1
218196_at	-1.61	D	-1.89	D	
225750_at	-1.59	D	-1.28	D	Consensus includes gb:BE966748 /FEA=EST /DB_XREF=gi:11772486 /DB_XREF=est:601661247R1 /CLONE=IMAGE:3916235 /UG=Hs.10949 Homo sapiens cDNA FLJ14162 fis, clone NT2RM4002504

225242_s_at	-1.59	D	-3.1	D	
241879_at	-1.58	D	-1.24	D	
201149_s_at	-1.57	D	-1.26	D	gb:U67195.1 /DEF=Human tissue inhibitor of metalloproteinase-3 mRNA, complete cds. /FEA=mRNA /PROD=tissue inhibitor of metalloproteinase-3 /DB_XREF=gi:1519557 /UG=Hs.245188 tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory) /FL=gb:U67195.1 gb:U02571.1 gb:U14394.1 gb:NM_000362.2
214895_s_at	-1.56	D	-1.11	D	
226962_at	-1.55	D	-1.37	D	
204298_s_at	-1.52	D	-1.16	D	gb:NM_002317.1 /DEF=Homo sapiens lysyl oxidase (LOX), mRNA. /FEA=mRNA /GEN=LOX /PROD=lysyl oxidase /DB_XREF=gi:4505008 /UG=Hs.102267 lysyl oxidase /FL=gb:M94054.1 gb:AF039291.1 gb:NM_002317.1
215177_s_at	-1.52	D	-1.68	D	
232079_s_at	-1.52	D	-2.59	D	
231550_at	-1.51	D	-1.15	D	
201942_s_at	-1.5	D	-1.56	D	gb:D85390.1 /DEF=Homo sapiens mRNA for gp180-carboxypeptidase D-like enzyme, complete cds. /FEA=mRNA /PROD=gp180-carboxypeptidase D-like enzyme /DB_XREF=gi:3641620 /UG=Hs.5057 carboxypeptidase D /FL=gb:U65090.1 gb:D85390.1 gb:NM_001304.2
208712_at	-1.49	D	-1.01	D	gb:M73554.1 /DEF=Human bcl-1 mRNA, complete CDS. /FEA=mRNA /GEN=bcl-1 /PROD=bcl-1 /DB_XREF=gi:179364 /UG=Hs.82932 cyclin D1 (PRAD1: parathyroid adenomatosis 1) /FL=gb:BC000076.1 gb:M73554.1
227062_at	-1.49	D	-1	D	
201409_s_at	-1.48	D	1	I	gb:NM_002709.1 /DEF=Homo sapiens protein phosphatase 1, catalytic subunit, beta isoform (PPP1CB), mRNA. /FEA=mRNA /GEN=PPP1CB /PROD=protein phosphatase 1, catalytic subunit, betaisoform /DB_XREF=gi:4506004 /UG=Hs.21537 protein phosphatase 1, catalytic subunit, beta isoform /FL=gb:NM_002709.1 gb:AF092905.1
201015_s_at	-1.47	D	-1.09	D	gb:NM_021991.1 /DEF=Homo sapiens junction plakoglobin (JUP), transcript variant 2, mRNA. /FEA=mRNA /GEN=JUP /PROD=junction plakoglobin, isoform 1 /DB_XREF=gi:12056467 /UG=Hs.2340 junction plakoglobin /FL=gb:NM_021991.1 gb:BC000441.1
224874_at	-1.47	D	-0.98	D	
201617_x_at	-1.45	D	-1.03	D	gb:NM_004342.2 /DEF=Homo sapiens caldesmon 1 (CALD1), mRNA. /FEA=mRNA /GEN=CALD1 /PROD=caldesmon 1 /DB_XREF=gi:11091984 /UG=Hs.325474 caldesmon 1 /FL=gb:NM_004342.2 gb:M64110.1
224691_at	-1.45	D	-1.17	D	

228418_at	-1.44	D	1.1	I	
224597_at	-1.44	D	-1.02	D	
206584_at	-1.43	D	-1.87	D	
235198_at	-1.43	D	-2.61	D	
208779_x_at	-1.42	D	-1.54	D	gb:L20817.1 /DEF=Homo sapiens tyrosine protein kinase (CAK) gene, complete cds. /FEA=mRNA /GEN=CAK /PROD=tyrosine protein kinase /DB_XREF=gi:306474 /UG=Hs.75562 discoidin domain receptor family, member 1 /FL=gb:L20817.1
228220_at	-1.42	D	-1.6	D	
209748_at	-1.41	D	1.07	I	
211985_s_at	-1.41	D	-1.11	D	
239503_at	-1.41	D	-1.42	D	Consensus includes gb:AI803010 /FEA=EST /DB_XREF=gi:5368482 /DB_XREF=est:tj60c11.x1 /CLONE=IMAGE:2145908 /UG=Hs.126877 ESTs
224702_at	-1.4	D	-1.59	D	Consensus includes gb:BG501219 /FEA=EST /DB_XREF=gi:13462747 /DB_XREF=est:602546289F1 /CLONE=IMAGE:4668574 /UG=Hs.10248 Homo sapiens cDNA FLJ20167 fis, clone COL09512
228708_at	-1.39	D	-1.71	D	
32502_at	-1.38	D	-0.96	D	Cluster Incl. AL041124:DKFZp434D0316_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434D0316 /clone_end=3 /gb=AL041124 /gi=5410060 /ug=Hs.6748 /len=719
223571_at	-1.38	D	-1.02	D	gb:AF329842.1 /DEF=Homo sapiens complement-c1q tumor necrosis factor-related protein (CTRP6) mRNA, complete cds. /FEA=mRNA /GEN=CTRP6 /PROD=complement-c1q tumor necrosis factor-relatedprotein /DB_XREF=gi:13274530 /UG=Hs.22011 Homo sapiens complement-c1q tumor necrosis factor-related protein (CTRP6) mRNA, complete cds /FL=gb:AF329842.1
202351_at	-1.37	D	-1.04	D	
213552_at	-1.37	D	-1.44	D	
211708_s_at	-1.36	D	-2.08	D	gb:BC005807.1 /DEF=Homo sapiens, clone MGC:10264, mRNA, complete cds. /FEA=mRNA /PROD=Unknown (protein for MGC:10264) /DB_XREF=gi:13543283 /FL=gb:BC005807.1
217887_s_at	-1.36	D	-1.33	D	gb:NM_001981.1 /DEF=Homo sapiens epidermal growth factor receptor pathway substrate 15 (EPS15), mRNA. /FEA=mRNA /GEN=EPS15 /PROD=epidermal growth factor receptor pathwaysubstrate 15 /DB_XREF=gi:4503592 /UG=Hs.79095 epidermal growth factor receptor pathway substrate 15 /FL=gb:NM_001981.1 gb:U07707.1
219023_at	-1.36	D	1.01	I	
243495_s_at	-1.35	D	-1.72	D	Consensus includes gb:AL036450 /FEA=EST /DB_XREF=gi:5406002 /DB_XREF=est:DKFZp564D1062_r1

					/CLONE=DKFZp564D1062 /UG=Hs.103238 ESTs
209064_x_at	-1.34	D	-1.71	D	gb:AL136920.1 /DEF=Homo sapiens mRNA; cDNA DKFZp586C051 (from clone DKFZp586C051); complete cds. /FEA=mRNA /GEN=DKFZp586C051 /PROD=hypothetical protein /DB_XREF=gi:12053334 /UG=Hs.109643 polyadenylate binding protein-interacting protein 1 /FL=gb:AL136920.1
214975_s_at	-1.34	D	-1.9	D	
229461_x_at	-1.34	D	-2.16	D	
227312_at	-1.34	D	-1.12	D	
225522_at	-1.34	D	-1.22	D	
226390_at	-1.33	D	-1.07	D	
226254_s_at	-1.33	D	-1.83	D	Consensus includes gb:AI912523 /FEA=EST /DB_XREF=gi:5632378 /DB_XREF=est:tz25f09.x1 /CLONE=IMAGE:2289641 /UG=Hs.126914 KIAA1430 protein
224618_at	-1.33	D	-1.43	D	
226633_at	-1.32	D	-1.55	D	
204855_at	-1.3	D	-1.61	D	gb:NM_002639.1 /DEF=Homo sapiens serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5 (SERPINB5), mRNA. /FEA=mRNA /GEN=SERPINB5 /PROD=serine (or cysteine) proteinase inhibitor, cladeB (ovalbumin), member 5 /DB_XREF=gi:4505788 /UG=Hs.55279 serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5 /FL=gb:NM_002639.1 gb:U04313.1
221568_s_at	-1.3	D	-1.07	D	
208796_s_at	-1.29	D	-1.03	D	gb:BC000196.1 /DEF=Homo sapiens, cyclin G1, clone MGC:643, mRNA, complete cds. /FEA=mRNA /PROD=cyclin G1 /DB_XREF=gi:12652880 /UG=Hs.79101 cyclin G1 /FL=gb:L49504.1 gb:U47413.1 gb:BC000196.1 gb:D78341.1 gb:NM_004060.2
225366_at	-1.29	D	-1.74	D	
226817_at	-1.27	D	-1.11	D	
219594_at	-1.26	D	-1.35	MD	
230788_at	-1.26	D	-1.3	D	Consensus includes gb:BF059748 /FEA=EST /DB_XREF=gi:10813644 /DB_XREF=est:7k65h01.x1 /CLONE=IMAGE:3480432 /UG=Hs.116346 ESTs, Highly similar to A46297 beta-1,6-N-acetylglucosaminyltransferase H.sapiens
226310_at	-1.26	D	-1.16	D	
224800_at	-1.26	D	-1.79	D	
1007_s_at	-1.25	D	-1.19	D	U48705 /FEATURE=mRNA /DEFINITION=HSU48705 Human receptor tyrosine kinase DDR gene, complete cds
206315_at	-1.25	D	-1.82	D	
212690_at	-1.25	D	-1.38	D	

218706_s_at	-1.25	D	-1.15	D	
201469_s_at	-1.24	D	-1.62	D	
238587_at	-1.24	D	-2.66	D	
226051_at	-1.24	D	-0.96	D	
225405_at	-1.24	D	-1.03	MD	
201602_s_at	-1.23	D	-1.02	D	
209380_s_at	-1.23	D	-1.58	D	
239269_at	-1.23	D	-0.97	D	
227740_at	-1.23	D	-1.14	D	Consensus includes gb:AW173222 /FEA=EST /DB_XREF=gi:6439170 /DB_XREF=est:xj84g11.x1 /CLONE=IMAGE:2663972 /UG=Hs.127310 ESTs
225575_at	-1.23	D	-1.39	D	
212488_at	-1.22	D	-1.4	D	
227853_at	-1.22	D	-1.05	D	
224617_at	-1.22	D	-1.47	D	
205905_s_at	-1.21	D	-0.96	D	
223441_at	-1.21	D	-1.68	D	Consensus includes gb:AK026921.1 /DEF=Homo sapiens cDNA: FLJ23268 fis, clone COL08932, highly similar to HSA387747 Homo sapiens mRNA for sialin. /FEA=mRNA /DB_XREF=gi:10439893 /UG=Hs.117865 solute carrier family 17 (anionsugar transporter), member 5 /FL=gb:AF244577.1
205552_s_at	-1.2	D	-1.29	D	
210749_x_at	-1.2	D	-1.55	D	gb:L11315.1 /DEF=Homo sapiens receptor tyrosine kinase mRNA, complete cds. /FEA=mRNA /PROD=receptor tyrosine kinase /DB_XREF=gi:403386 /UG=Hs.75562 discoidin domain receptor family, member 1 /FL=gb:L11315.1
235258_at	-1.2	D	-1.03	D	
235125_x_at	-1.2	D	-1.06	D	
228531_at	-1.2	D	-1.18	D	
204794_at	-1.19	D	-1.34	D	gb:NM_004418.2 /DEF=Homo sapiens dual specificity phosphatase 2 (DUSP2), mRNA. /FEA=mRNA /GEN=DUSP2 /PROD=dual specificity phosphatase 2 /DB_XREF=gi:12707563 /UG=Hs.1183 dual specificity phosphatase 2 /FL=gb:NM_004418.2 gb:L11329.1
207169_x_at	-1.19	D	-1.69	D	
202893_at	-1.17	D	-1.02	D	
209468_at	-1.17	D	-0.98	D	gb:AB017498.1 /DEF=Homo sapiens LRP5 mRNA for Lipoprotein Receptor Related Protein 5, complete cds. /FEA=mRNA /GEN=LRP5 /PROD=Lipoprotein Receptor Related Protein 5 /DB_XREF=gi:3582144 /UG=Hs.6347 low density lipoprotein receptor-related protein 5 /FL=gb:AB017498.1 gb:AF064548.1 gb:AF077820.1 gb:NM_002335.1

227685_at	-1.17	D	-1.48	D	
201506_at	-1.15	D	-1.13	D	
203679_at	-1.15	D	-1.15	D	
206382_s_at	-1.15	D	-0.98	D	gb:NM_001709.1 /DEF=Homo sapiens brain-derived neurotrophic factor (BDNF), mRNA. /FEA=mRNA /GEN=BDNF /PROD=brain-derived neurotrophic factor /DB_XREF=gi:4502392 /UG=Hs.56023 brain-derived neurotrophic factor /FL=gb:NM_001709.1
213623_at	-1.15	D	-1.46	D	
228611_s_at	-1.15	D	-1.25	D	
213151_s_at	-1.14	D	-1.84	D	
213805_at	-1.14	D	-1.44	D	
225571_at	-1.14	D	-1.22	D	
202502_at	-1.13	D	1.07	I	
214443_at	-1.13	D	-1.51	D	
223006_s_at	-1.13	D	-1.02	D	
202067_s_at	-1.12	D	-2.56	D	
205808_at	-1.11	D	-2.05	D	
212589_at	-1.11	D	1.18	I	
64883_at	-1.11	D	-0.97	D	Cluster Incl. AI744083:wc36b04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2317231 /clone_end=3 /gb=AI744083 /gi=5112371 /ug=Hs.65406 /len=608
226671_at	-1.11	D	-1.02	D	
225813_at	-1.11	D	-1.15	D	
212340_at	-1.09	D	-1.08	D	
229450_at	-1.09	D	-1.24	D	
227771_at	-1.08	D	-1.53	D	
204364_s_at	-1.07	D	-1.39	D	
207147_at	-1.07	D	-1.02	D	
212489_at	-1.07	D	-1.69	D	
226045_at	-1.07	D	-1.34	D	
221269_s_at	-1.06	D	-1.12	D	gb:NM_031286.1 /DEF=Homo sapiens SH3BGRL3-like protein (SH3BGRL3), mRNA. /FEA=mRNA /GEN=SH3BGRL3 /PROD=SH3BGRL3-like protein /DB_XREF=gi:13775197 /FL=gb:NM_031286.1
228479_at	-1.06	D	-1.06	D	
226115_at	-1.06	D	-1.23	D	
222996_s_at	-1.06	D	-1.06	D	gb:BC002490.1 /DEF=Homo sapiens, hypothetical protein, clone MGC:915, mRNA, complete cds. /FEA=mRNA /PROD=hypothetical protein /DB_XREF=gi:12803342 /UG=Hs.15093 hypothetical protein /FL=gb:BC002490.1 gb:AF151029.1 gb:NM_016463.1
203071_at	-1.05	D	-1.14	D	
231716_at	-1.05	D	-0.95	D	Consensus includes gb:AF255304.1 /DEF=Homo sapiens membrane-associated nucleic acid

					binding protein mRNA, complete cds. /FEA=mRNA /PROD=membrane-associated nucleic acid bindingprotein /DB_XREF=gi:9837126 /UG=Hs.112227 membrane-associated nucleic acid binding protein /FL=gb:AF255304.1
225607_at	-1.05	D	-1.03	D	
225912_at	-1.04	D	-1.26	D	
205594_at	-1.03	D	-1.52	D	
211161_s_at	-1.03	D	-1.28	D	gb:AF130082.1 /DEF=Homo sapiens clone FLC1492 PRO3121 mRNA, complete cds. /FEA=mRNA /PROD=PRO3121 /DB_XREF=gi:11493468 /UG=Hs.119571 collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant) /FL=gb:AF130082.1
203203_s_at	-1.01	D	0.98	I	
210797_s_at	-1.01	D	-1.57	D	
215489_x_at	-1.01	D	-1.29	D	
201852_x_at	-1	D	-1.21	D	
202667_s_at	-1	D	-1.14	D	
203325_s_at	-1	D	-1.22	D	
219297_at	-1	D	0.96	I	
243357_at	-1	D	-1.14	D	
225725_at	-1	D	-1.8	D	
224209_s_at	-1	D	-1.19	D	gb:AF019638.1 /DEF=Homo sapiens nedasin s-form mRNA, complete cds. /FEA=mRNA /PROD=nedasin s-form /DB_XREF=gi:6469319 /UG=Hs.239147 guanine deaminase /FL=gb:AF019638.1
213508_at	-0.99	D	-1.32	D	
229399_at	-0.99	D	-2.19	D	
225192_at	-0.99	D	-1.59	D	
242931_at	-0.98	D	-1.41	D	
203743_s_at	-0.97	D	-1.06	D	
205383_s_at	-0.97	D	-1.24	D	
226711_at	-0.97	D	-1.26	D	Consensus includes gb:BF590117 /FEA=EST /DB_XREF=gi:11682441 /DB_XREF=est:nab19b11.x1 /CLONE=IMAGE:3266253 /UG=Hs.106131 ESTs
217738_at	-0.96	D	-1.09	D	
217765_at	-0.96	D	-1.02	D	
229010_at	-0.96	D	-1.17	D	
227082_at	-0.96	D	-1.56	D	
226312_at	-0.96	D	-0.95	D	
226237_at	-0.96	D	-2.07	D	
225950_at	-0.96	D	1.38	I	
204540_at	-0.95	D	-1.22	D	gb:NM_001958.1 /DEF=Homo sapiens eukaryotic translation elongation factor 1 alpha 2 (EEF1A2), mRNA. /FEA=mRNA /GEN=EEF1A2

					/PROD=eukaryotic translation elongation factor 1 alpha2 /DB_XREF=gi:4503474 /UG=Hs.2642 eukaryotic translation elongation factor 1 alpha 2 /FL=gb:BC000432.1 gb:NM_001958.1
220178_at	0.95	I	1.43	I	gb:NM_021731.1 /DEF=Homo sapiens hypothetical protein PP3501 (PP3501), mRNA. /FEA=mRNA /GEN=PP3501 /PROD=hypothetical protein PP3501 /DB_XREF=gi:11119425 /UG=Hs.301406 hypothetical protein PP3501 /FL=gb:NM_021731.1
229075_at	0.95	I	1.07	I	Consensus includes gb:AI754871 /FEA=EST /DB_XREF=gi:5133135 /DB_XREF=est:cr31g04.x1 /CLONE=HBMSC_cr31g04 /UG=Hs.111207 ESTs
209849_s_at	0.96	I	0.98	I	gb:AF029669.1 /DEF=Homo sapiens Rad51C (RAD51C) mRNA, complete cds. /FEA=mRNA /GEN=RAD51C /PROD=Rad51C /DB_XREF=gi:2909800 /UG=Hs.11393 RAD51 (S. cerevisiae) homolog C /FL=gb:AF029669.1
212483_at	0.96	I	1	I	
218294_s_at	0.96	I	2.14	I	gb:AF267865.1 /DEF=Homo sapiens DC41 mRNA, complete cds. /FEA=mRNA /PROD=DC41 /DB_XREF=gi:12006056 /UG=Hs.271623 nucleoporin 50kD /FL=gb:AF267865.1 gb:AF107840.1 gb:NM_007172.1 gb:AF116624.1
218542_at	0.96	I	1.02	I	
218585_s_at	0.96	I	1.15	I	
209154_at	0.97	I	1.28	I	gb:AF234997.1 /DEF=Homo sapiens glutaminase-interacting protein 3 mRNA, complete cds. /FEA=mRNA /PROD=glutaminase-interacting protein 3 /DB_XREF=gi:12005281 /UG=Hs.12956 Tax interaction protein 1 /FL=gb:AF028823.2 gb:NM_014604.1 gb:AF234997.1
231094_s_at	0.97	I	0.99	I	
223039_at	0.97	I	1.15	I	gb:BC004144.1 /DEF=Homo sapiens, Similar to PRO1992 protein, clone MGC:1842, mRNA, complete cds. /FEA=mRNA /PROD=Similar to PRO1992 protein /DB_XREF=gi:13278734 /UG=Hs.9850 Homo sapiens cDNA: FLJ21589 fis, clone COL06960 /FL=gb:BC004144.1
203563_at	0.98	I	1.5	I	gb:NM_021638.1 /DEF=Homo sapiens actin filament associated protein (AFAP), mRNA. /FEA=mRNA /GEN=AFAP /PROD=actin filament associated protein /DB_XREF=gi:11056013 /UG=Hs.80306 actin filament associated protein /FL=gb:AF188700.1 gb:NM_021638.1
212943_at	0.98	I	1.19	I	
218945_at	0.98	I	1.02	I	gb:NM_024109.1 /DEF=Homo sapiens hypothetical protein MGC2654 (MGC2654), mRNA. /FEA=mRNA /GEN=MGC2654 /PROD=hypothetical protein MGC2654 /DB_XREF=gi:13129121 /UG=Hs.165428 hypothetical protein MGC2654 /FL=gb:BC001908.1 gb:NM_024109.1
221802_s_at	0.99	I	1.62	I	
203150_at	1	I	1.04	I	

208709_s_at	1	I	1.39	I	gb:U64898.1 /DEF=Homo sapiens NRD convertase mRNA, complete cds. /FEA=mRNA /PROD=NRD convertase /DB_XREF=gi:2897866 /UG=Hs.4099 nardilysin (N-arginine dibasic convertase) /FL=gb:U64898.1
228867_at	1	I	1	I	Consensus includes gb:BE541548 /FEA=EST /DB_XREF=gi:9770193 /DB_XREF=est:601067822F1 /CLONE=IMAGE:3454205 /UG=Hs.123386 ESTs
225836_s_at	1	I	1.01	I	
201292_at	1.01	I	1.3	I	
204228_at	1.01	I	1.11	I	
211114_x_at	1.01	I	1.25	I	
201536_at	1.02	I	1.04	I	
219204_s_at	1.02	I	1.1	I	gb:NM_021947.1 /DEF=Homo sapiens serine racemase (SRR), mRNA. /FEA=mRNA /GEN=SRR /PROD=serine racemase /DB_XREF=gi:11345491 /UG=Hs.27335 serine racemase /FL=gb:AF169974.1 gb:NM_021947.1
224824_at	1.02	I	1.26	I	
206364_at	1.03	I	1.02	I	
201675_at	1.04	I	1.43	I	gb:NM_003488.1 /DEF=Homo sapiens A kinase (PRKA) anchor protein 1 (AKAP1), mRNA. /FEA=mRNA /GEN=AKAP1 /PROD=A kinase (PRKA) anchor protein 1 /DB_XREF=gi:4502014 /UG=Hs.78921 A kinase (PRKA) anchor protein 1 /FL=gb:BC000729.1 gb:NM_003488.1
241453_at	1.04	I	1.51	I	Consensus includes gb:AA912743 /FEA=EST /DB_XREF=gi:3052135 /DB_XREF=est:ol41d04.s1 /CLONE=IMAGE:1526023 /UG=Hs.126421 ESTs, Highly similar to FAK1_HUMAN FOCAL ADHESION KINASE 1 H.sapiens
223024_at	1.04	I	1.11	I	
202533_s_at	1.05	I	1.24	I	gb:BC003584.1 /DEF=Homo sapiens, dihydrofolate reductase, clone MGC:3153, mRNA, complete cds. /FEA=mRNA /PROD=dihydrofolate reductase /DB_XREF=gi:13097773 /UG=Hs.83765 dihydrofolate reductase /FL=gb:BC000192.1 gb:BC003584.1 gb:NM_000791.2
208653_s_at	1.05	I	1.29	I	
218979_at	1.05	I	1.62	I	gb:NM_024945.1 /DEF=Homo sapiens hypothetical protein FLJ12888 (FLJ12888), mRNA. /FEA=mRNA /GEN=FLJ12888 /PROD=hypothetical protein FLJ12888 /DB_XREF=gi:13376426 /UG=Hs.284137 hypothetical protein FLJ12888 /FL=gb:NM_024945.1
226181_at	1.05	I	1.18	I	
223513_at	1.06	I	1.06	I	gb:AF139625.1 /DEF=Homo sapiens centrosomal P4.1-associated protein (CPAP) mRNA, complete cds. /FEA=mRNA /GEN=CPAP /PROD=centrosomal P4.1-associated protein /DB_XREF=gi:10643591 /UG=Hs.283077

					centrosomal P4.1-associated protein; uncharacterized bone marrow protein BM032 /FL=gb:AF139625.1
203740_at	1.07	I	1.77	I	
219306_at	1.07	I	1.48	I	
238756_at	1.07	I	1.18	I	
228577_x_at	1.07	I	1.34	I	
218755_at	1.08	I	1.17	I	
219148_at	1.08	I	1.18	I	
225456_at	1.08	I	1.34	I	Consensus includes gb:AI708776 /FEA=EST /DB_XREF=gi:4998552 /DB_XREF=est:as35f01.x1 /CLONE=IMAGE:2319193 /UG=Hs.10130 ESTs
205063_at	1.09	I	1.65	I	gb:NM_003616.1 /DEF=Homo sapiens survival of motor neuron protein interacting protein 1 (SIP1), mRNA. /FEA=mRNA /GEN=SIP1 /PROD=survival of motor neuron protein interactingprotein 1 /DB_XREF=gi:4506960 /UG=Hs.102456 survival of motor neuron protein interacting protein 1 /FL=gb:AF027150.1 gb:NM_003616.1
225549_at	1.09	I	1.35	I	
225125_at	1.09	I	0.98	I	Consensus includes gb:BF978280 /FEA=EST /DB_XREF=gi:12345495 /DB_XREF=est:602148240F1 /CLONE=IMAGE:4307192 /UG=Hs.110702 Homo sapiens mRNA; cDNA DKFZp761E212 (from clone DKFZp761E212)
201185_at	1.1	I	1.07	I	gb:NM_002775.1 /DEF=Homo sapiens protease, serine, 11 (IGF binding) (PRSS11), mRNA. /FEA=mRNA /GEN=PRSS11 /PROD=protease, serine, 11 (IGF binding) /DB_XREF=gi:4506140 /UG=Hs.75111 protease, serine, 11 (IGF binding) /FL=gb:D87258.1 gb:NM_002775.1
211115_x_at	1.1	I	1.37	I	
213761_at	1.1	I	1.21	I	
219162_s_at	1.1	I	1.21	I	
236641_at	1.1	I	1.29	I	Consensus includes gb:AW183154 /FEA=EST /DB_XREF=gi:6451630 /DB_XREF=est:xj67b12.x1 /CLONE=IMAGE:2662271 /UG=Hs.116649 ESTs
227356_at	1.1	I	1.66	I	
225837_at	1.1	I	1.21	I	
202536_at	1.11	I	1.86	I	
225060_at	1.11	I	1.66	I	
202399_s_at	1.12	I	1.64	I	
211595_s_at	1.12	I	1.28	I	gb:AB049944.1 /DEF=Homo sapiens MRPS11 mRNA for mitochondrial ribosomal protein S11, complete cds. /FEA=mRNA /GEN=MRPS11 /PROD=mitochondrial ribosomal protein S11 /DB_XREF=gi:13620888 /FL=gb:AB049944.1
219045_at	1.12	I	1.36	I	

200975_at	1.13	I	1.03	I	gb:NM_000310.1 /DEF=Homo sapiens palmitoyl-protein thioesterase 1 (ceroid-lipofuscinosis, neuronal 1, infantile) (PPT1), mRNA. /FEA=mRNA /GEN=PPT1 /PROD=palmitoyl-protein thioesterase 1(ceroid-lipofuscinosis, neuronal 1, infantile) /DB_XREF=gi:4506030 /UG=Hs.3873 palmitoyl-protein thioesterase 1 (ceroid-lipofuscinosis, neuronal 1, infantile) /FL=gb:U44772.1 gb:NM_000310.1
208931_s_at	1.13	I	1.48	I	
229603_at	1.13	I	1.17	I	
201555_at	1.14	I	1.08	I	
202569_s_at	1.14	I	1.11	I	gb:NM_002376.1 /DEF=Homo sapiens MAPmicrotubule affinity-regulating kinase 3 (MARK3), mRNA. /FEA=mRNA /GEN=MARK3 /PROD=MAPmicrotubule affinity-regulating kinase 3 /DB_XREF=gi:4505102 /UG=Hs.172766 MAPmicrotubule affinity-regulating kinase 3 /FL=gb:M80359.1 gb:NM_002376.1
203905_at	1.14	I	1.18	I	gb:NM_002582.1 /DEF=Homo sapiens poly(A)-specific ribonuclease (deadenylation nuclease) (PARN), mRNA. /FEA=mRNA /GEN=PARN /PROD=poly(A)-specific ribonuclease (deadenylationnuclease) /DB_XREF=gi:4505610 /UG=Hs.43445 poly(A)-specific ribonuclease (deadenylation nuclease) /FL=gb:NM_002582.1
242617_at	1.14	I	0.96	I	
227249_at	1.14	I	1.45	I	
213007_at	1.15	I	0.99	I	
218460_at	1.15	I	0.97	I	
204622_x_at	1.16	I	1.01	I	
207165_at	1.16	I	1.01	I	
52164_at	1.16	I	1.06	I	Cluster Incl. AA065185:zm50e09.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-529096 /clone_end=3 /gb=AA065185 /gi=1559080 /ug=Hs.47008 /len=605
203790_s_at	1.17	I	1.25	I	
200639_s_at	1.18	I	1.35	I	
206085_s_at	1.18	I	1.07	I	gb:NM_001902.1 /DEF=Homo sapiens cystathionase (cystathionine gamma-lyase) (CTH), mRNA. /FEA=mRNA /GEN=CTH /PROD=cystathionase (cystathionine gamma-lyase) /DB_XREF=gi:4503124 /UG=Hs.19904 cystathionase (cystathionine gamma-lyase) /FL=gb:NM_001902.1
212651_at	1.18	I	1.67	I	
221521_s_at	1.18	I	1.08	I	gb:BC003186.1 /DEF=Homo sapiens, HSPC037 protein, clone MGC:673, mRNA, complete cds. /FEA=mRNA /PROD=HSPC037 protein /DB_XREF=gi:13112024 /UG=Hs.108196 HSPC037 protein /FL=gb:BC003186.1 gb:AF201939.1
221931_s_at	1.18	I	1.24	I	

201513_at	1.19	I	1.31	I	
222000_at	1.19	I	1.04	I	
201306_s_at	1.2	I	1.28	I	
222103_at	1.2	I	1.27	I	
225834_at	1.2	I	1.06	I	
218875_s_at	1.21	I	1.44	I	
207181_s_at	1.22	I	1.17	I	gb:NM_001227.1 /DEF=Homo sapiens caspase 7, apoptosis-related cysteine protease (CASP7), mRNA. /FEA=mRNA /GEN=CASP7 /PROD=caspase 7, apoptosis-related cysteine protease /DB_XREF=gi:4502580 /UG=Hs.9216 caspase 7, apoptosis-related cysteine protease /FL=gb:U37448.1 gb:U40281.1 gb:U67319.1 gb:U67320.1 gb:NM_001227.1
225195_at	1.22	I	0.97	I	
204203_at	1.23	I	0.99	I	gb:NM_001806.1 /DEF=Homo sapiens CCAATenhancer binding protein (CEBP), gamma (CEBPG), mRNA. /FEA=mRNA /GEN=CEBPG /PROD=CCAATenhancer binding protein gamma /DB_XREF=gi:4502768 /UG=Hs.2227 CCAATenhancer binding protein (CEBP), gamma /FL=gb:NM_001806.1 gb:U20240.1
212815_at	1.23	I	1.19	I	
222431_at	1.23	I	1.04	I	gb:AL136719.1 /DEF=Homo sapiens mRNA; cDNA DKFZp566G0346 (from clone DKFZp566G0346); complete cds. /FEA=mRNA /GEN=DKFZp566G0346 /PROD=hypothetical protein /DB_XREF=gi:12052956 /UG=Hs.289043 spindlin /FL=gb:AL136719.1 gb:AF087864.1 gb:AF317228.2 gb:AF106682.1 gb:NM_006717.1
219502_at	1.24	I	0.96	I	
223307_at	1.24	I	1.07	I	gb:BC002551.1 /DEF=Homo sapiens, Similar to gene rich cluster, C8 gene, clone MGC:2577, mRNA, complete cds. /FEA=mRNA /PROD=Similar to gene rich cluster, C8 gene /DB_XREF=gi:12803452 /UG=Hs.30114 Homo sapiens, Similar to gene rich cluster, C8 gene, clone MGC:2577, mRNA, complete cds /FL=gb:BC002551.1
217127_at	1.25	I	1.4	I	
222425_s_at	1.25	I	1.24	I	
224626_at	1.26	I	1.53	I	
201707_at	1.27	I	1.29	I	gb:NM_002857.1 /DEF=Homo sapiens peroxisomal farnesylated protein (PXF), mRNA. /FEA=mRNA /GEN=PXF /PROD=peroxisomal farnesylated protein /DB_XREF=gi:4506338 /UG=Hs.168670 peroxisomal farnesylated protein /FL=gb:BC000496.1 gb:NM_002857.1 gb:AB018541.1
203494_s_at	1.27	I	1.21	I	
208079_s_at	1.27	I	1.47	I	
208808_s_at	1.27	I	1.52	I	gb:BC000903.1 /DEF=Homo sapiens, high-mobility group (nonhistone chromosomal) protein

					2, clone MGC:5234, mRNA, complete cds. /FEA=mRNA /PROD=high-mobility group (nonhistone chromosomal)protein 2 /DB_XREF=gi:12654170 /UG=Hs.80684 high-mobility group (nonhistone chromosomal) protein 2 /FL=gb:BC000903.1 gb:BC001063.1
209714_s_at	1.27	I	1.23	I	gb:AF213033.1 /DEF=Homo sapiens isolate BX-01 cyclin-dependent kinase associated protein phosphatase mRNA, complete cds. /FEA=mRNA /PROD=cyclin-dependent kinase associated proteinphosphatase /DB_XREF=gi:12734643 /UG=Hs.84113 cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase) /FL=gb:AF213033.1 gb:AF213034.1 gb:AF213035.1 gb:AF213036.1 gb:AF213037.1 gb:AF213038.1 gb:AF213039.1 gb:AF213040.1 gb:AF213041.1 gb:AF213042.1 gb:AF213044.1 gb:AF213046.1 gb:AF213047.1 gb:AF213048.1 gb:AF213049.1 gb:AF213050.1 gb:AF213051.1 gb:AF213052.1 gb:AF213053.1 gb:U02681.1 gb:L25876.1 gb:NM_005192.1 gb:L27711.1
227157_at	1.28	I	1.08	I	
218886_at	1.29	I	2.09	I	
225649_s_at	1.29	I	1.49	I	Consensus includes gb:AA001414 /FEA=EST /DB_XREF=gi:1436899 /DB_XREF=est:ze45d08.s1 /CLONE=IMAGE:361935 /UG=Hs.100057 Homo sapiens cDNA: FLJ22902 fis, clone KAT05581
200631_s_at	1.3	I	1.15	I	gb:NM_003011.1 /DEF=Homo sapiens SET translocation (myeloid leukemia-associated) (SET), mRNA. /FEA=mRNA /GEN=SET /PROD=SET translocation (myeloid leukemia-associated) /DB_XREF=gi:4506890 /UG=Hs.145279 SET translocation (myeloid leukemia-associated) /FL=gb:U51924.1 gb:M93651.1 gb:NM_003011.1
201515_s_at	1.3	I	1.46	I	
212297_at	1.3	I	1.04	I	
213047_x_at	1.3	I	1.08	I	
218602_s_at	1.3	I	1.36	I	
231855_at	1.3	I	1.24	I	
223700_at	1.3	I	1.12	I	gb:AY028916.1 /DEF=Homo sapiens GAJ (GAJ) mRNA, complete cds. /FEA=mRNA /GEN=GAJ /PROD=GAJ /DB_XREF=gi:13488608 /UG=Hs.294088 Homo sapiens GAJ (GAJ) mRNA, complete cds /FL=gb:AY028916.1
222416_at	1.3	I	1.23	I	gb:U76542.1 /DEF=Human pyrroline-5-carboxylate synthase long form (P5CSL) mRNA, complete cds. /FEA=mRNA /GEN=P5CSL /PROD=pyrroline-5-carboxylate synthase long form /DB_XREF=gi:4335784 /UG=Hs.114366 pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase) /FL=gb:U68758.1 gb:U76542.1 gb:NM_002860.1
228812_at	1.31	I	1.33	I	

224601_at	1.31		1		
204616_at	1.32		1.57		
218883_s_at	1.32		1.04		gb:NM_024629.1 /DEF=Homo sapiens hypothetical protein FLJ23468 (FLJ23468), mRNA. /FEA=mRNA /GEN=FLJ23468 /PROD=hypothetical protein FLJ23468 /DB_XREF=gi:13375855 /UG=Hs.38178 hypothetical protein FLJ23468 /FL=gb:NM_024629.1
200629_at	1.33		0.97		
218545_at	1.33		1.3		
200640_at	1.34		1.38		
217791_s_at	1.34		1.18		gb:NM_002860.1 /DEF=Homo sapiens pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase) (PYCS), mRNA. /FEA=mRNA /GEN=PYCS /PROD=pyrroline-5-carboxylate synthetase (glutamategamma-semialdehyde synthetase) /DB_XREF=gi:4506348 /UG=Hs.114366 pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase) /FL=gb:U68758.1 gb:U76542.1 gb:NM_002860.1
201328_at	1.35		1.25		
232235_at	1.37		1.05		Consensus includes gb:AK021539.1 /DEF=Homo sapiens cDNA FLJ11477 fis, clone HEMBA1001746, weakly similar to Homo sapiens squamous cell carcinoma antigen recognized by T cell (SART-2) mRNA. /FEA=mRNA /DB_XREF=gi:10432739 /UG=Hs.124673 Homo sapiens cDNA FLJ11477 fis, clone HEMBA1001746, weakly similar to Homo sapiens squamous cell carcinoma antigen recognized by T cell (SART-2) mRNA
208654_s_at	1.38		1.58		
212367_at	1.38		1.34		
212444_at	1.38		1.1		
236300_at	1.38		1.2		
228507_at	1.38		1.26		
212688_at	1.39		1.08		
218373_at	1.39		1.1		gb:NM_022476.1 /DEF=Homo sapiens hypothetical protein FLJ13258 similar to fused toes (FLJ13258), mRNA. /FEA=mRNA /GEN=FLJ13258 /PROD=hypothetical protein FLJ13258 similar to fusedtoes /DB_XREF=gi:11968026 /UG=Hs.288929 hypothetical protein FLJ13258 similar to fused toes /FL=gb:NM_022476.1
204962_s_at	1.4		2		
213682_at	1.4		2.2		
210821_x_at	1.41		1.26		gb:BC002703.1 /DEF=Homo sapiens, centromere protein A (17kD), clone MGC:3892, mRNA, complete cds. /FEA=mRNA /PROD=centromere protein A (17kD) /DB_XREF=gi:12803732

					/UG=Hs.1594 centromere protein A (17kD) /FL=gb:BC002703.1
218458_at	1.42	I	1.27	I	gb:NM_022471.1 /DEF=Homo sapiens hypothetical protein FLJ13057 similar to germ cell-less (FLJ13057), mRNA. /FEA=mRNA /GEN=FLJ13057 /PROD=hypothetical protein FLJ13057 similar to germcell-less /DB_XREF=gi:11968020 /UG=Hs.243122 hypothetical protein FLJ13057 similar to germ cell-less /FL=gb:NM_022471.1
40189_at	1.42	I	1.18	I	Cluster Incl. M93651:Human set gene, complete cds /cds=(3,836) /gb=M93651 /gi=338038 /ug=Hs.145279 /len=2562
225006_x_at	1.42	I	1.73	I	
217958_at	1.43	I	1.5	I	
224560_at	1.43	I	1.68	I	
222889_at	1.43	I	1.37	I	Consensus includes gb:AI703304 /FEA=EST /DB_XREF=gi:4991204 /DB_XREF=est:wd82f11.x1 /CLONE=IMAGE:2338125 /UG=Hs.115660 hypothetical protein FLJ12810 /FL=gb:NM_022836.1
213359_at	1.44	I	1.38	I	
231579_s_at	1.46	I	1.19	I	
228597_at	1.46	I	1.7	I	
200896_x_at	1.48	I	1.45	I	
203341_at	1.49	I	1.66	I	
213523_at	1.49	I	1.37	I	
214949_at	1.49	I	1.32	I	
201001_s_at	1.5	I	1.49	I	
205053_at	1.5	I	1.11	I	gb:NM_000946.1 /DEF=Homo sapiens primase, polypeptide 1 (49kD) (PRIM1), mRNA. /FEA=mRNA /GEN=PRIM1 /PROD=primase, polypeptide 1 (49kD) /DB_XREF=gi:4506050 /UG=Hs.82741 primase, polypeptide 1 (49kD) /FL=gb:BC005266.1 gb:NM_000946.1
225865_x_at	1.51	I	1.81	I	
213322_at	1.52	I	1.09	I	
217736_s_at	1.54	I	1.64	I	gb:NM_014413.2 /DEF=Homo sapiens heme-regulated initiation factor 2-alpha kinase (HRI), mRNA. /FEA=mRNA /GEN=HRI /PROD=heme-regulated initiation factor 2-alpha kinase /DB_XREF=gi:11125767 /UG=Hs.258730 heme-regulated initiation factor 2-alpha kinase /FL=gb:NM_014413.2 gb:AL136563.1 gb:AF147094.1 gb:AF255050.1 gb:AF116634.1 gb:AF183414.1
48808_at	1.54	I	1.12	I	Cluster Incl. AI144299:qb59h06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1704443 /clone_end=3 /gb=AI144299 /gi=3666108 /ug=Hs.106843 /len=765
218131_s_at	1.55	I	1.05	I	
220607_x_at	1.55	I	1.43	I	

229700_at	1.58	I	1.37	I	
229299_at	1.59	I	1.21	I	
203745_at	1.6	I	1.25	I	
203126_at	1.61	I	1.73	I	
216484_x_at	1.62	I	1.62	I	
213959_s_at	1.63	I	1.02	I	
208405_s_at	1.65	I	1.81	I	
212021_s_at	1.65	I	1.27	I	
225343_at	1.65	I	1.72	I	
224819_at	1.66	I	1.5	I	
214948_s_at	1.67	I	1.37	I	
222844_s_at	1.67	I	1.34	I	gb:AF169974.1 /DEF=Homo sapiens serine racemase mRNA, complete cds. /FEA=mRNA /PROD=serine racemase /DB_XREF=gi:11034784 /UG=Hs.27335 serine racemase /FL=gb:AF169974.1 gb:NM_021947.1
225261_x_at	1.68	I	1.95	I	
205417_s_at	1.69	I	1.31	I	
209513_s_at	1.69	I	1.3	I	gb:BC004331.1 /DEF=Homo sapiens, Similar to RIKEN cDNA 2610207I16 gene, clone MGC:10940, mRNA, complete cds. /FEA=mRNA /PROD=Similar to RIKEN cDNA 2610207I16 gene /DB_XREF=gi:13279253 /UG=Hs.47986 Homo sapiens, Similar to RIKEN cDNA 2610207I16 gene, clone MGC:10940, mRNA, complete cds /FL=gb:BC004331.1
202402_s_at	1.71	I	1.02	I	gb:NM_001751.1 /DEF=Homo sapiens cysteinyl-tRNA synthetase (CARS), mRNA. /FEA=mRNA /GEN=CARS /PROD=cysteinyl-tRNA synthetase /DB_XREF=gi:10835050 /UG=Hs.159604 cysteinyl-tRNA synthetase /FL=gb:NM_001751.1 gb:BC002880.1 gb:AF288206.1 gb:AF288207.1
202534_x_at	1.71	I	1.15	I	
217959_s_at	1.73	I	1.35	I	
224715_at	1.76	I	1.23	I	
222843_at	1.77	I	1.77	I	
212474_at	1.79	I	1.81	I	
217794_at	1.79	I	1.36	I	
226546_at	1.79	I	1.2	I	
209512_at	1.85	I	1.12	I	gb:BC004331.1 /DEF=Homo sapiens, Similar to RIKEN cDNA 2610207I16 gene, clone MGC:10940, mRNA, complete cds. /FEA=mRNA /PROD=Similar to RIKEN cDNA 2610207I16 gene /DB_XREF=gi:13279253 /UG=Hs.47986 Homo sapiens, Similar to RIKEN cDNA 2610207I16 gene, clone MGC:10940, mRNA, complete cds /FL=gb:BC004331.1
212020_s_at	1.85	I	1.23	MI	
218563_at	1.86	I	2.38	I	
228992_at	1.87	I	1.96	I	

213310_at	1.9	I	1.37	I	
204634_at	1.91	I	1.71	I	
240983_s_at	1.96	I	1.09	I	
202580_x_at	1.97	I	1.55	I	gb:NM_021953.1 /DEF=Homo sapiens forkhead box M1 (FOXM1), mRNA. /FEA=mRNA /GEN=FOXM1 /PROD=forkhead box M1 /DB_XREF=gi:11386144 /UG=Hs.239 forkhead box M1 /FL=gb:NM_021953.1 gb:U83113.1 gb:L16783.1
203324_s_at	2.03	I	1.77	I	
202370_s_at	2.07	I	1.88	I	gb:NM_001755.1 /DEF=Homo sapiens core-binding factor, beta subunit (CBFB), transcript variant 2, mRNA. /FEA=mRNA /GEN=CBFB /PROD=core-binding factor, beta subunit, isoform 2 /DB_XREF=gi:13124872 /UG=Hs.179881 core-binding factor, beta subunit /FL=gb:NM_001755.1
221739_at	2.12	I	1.7	I	
230251_at	2.22	I	1.36	I	
200779_at	2.27	I	1	I	gb:NM_001675.1 /DEF=Homo sapiens activating transcription factor 4 (tax-responsive enhancer element B67) (ATF4), mRNA. /FEA=mRNA /GEN=ATF4 /PROD=activating transcription factor 4 /DB_XREF=gi:4502264 /UG=Hs.181243 activating transcription factor 4 (tax-responsive enhancer element B67) /FL=gb:M86842.1 gb:NM_001675.1
209096_at	2.29	I	2.29	I	
202209_at	2.34	I	2.48	I	
216483_s_at	2.38	I	1.67	I	
200749_at	2.56	I	2.52	I	
205047_s_at	3.28	I	1.78	I	gb:NM_001673.1 /DEF=Homo sapiens asparagine synthetase (ASNS), mRNA. /FEA=mRNA /GEN=ASNS /PROD=asparagine synthetase /DB_XREF=gi:4502258 /UG=Hs.75692 asparagine synthetase /FL=gb:M27396.1 gb:NM_001673.1
209452_s_at	3.58	I	3.31	I	gb:AF035824.1 /DEF=Homo sapiens vesicle soluble NSF attachment protein receptor (VTI1) mRNA, complete cds. /FEA=mRNA /GEN=VTI1 /PROD=vesicle soluble NSF attachment protein receptor /DB_XREF=gi:2687399 /UG=Hs.169206 vesicle-associated soluble NSF attachment protein receptor (v-SNARE; homolog of S. cerevisiae VTI1) /FL=gb:BC003142.1 gb:AF035824.1 gb:AF060902.1 gb:NM_006370.1

CURRICULUM VITAE

YA-LIN CHIU

Education and Academic Experience:

PhD, *Biochemistry and Molecular Pharmacology*
University of Massachusetts Medical School, MA, January 2003

PhD Candidate, *Molecular and Cellular Pharmacology*
Robert Wood Johnson Medical School-UMDNJ, NJ. Sep/1998-Dec/2001

Research Assistant 1997-1998
Department of Psychiatry, Cheng-Hsing Medical Center, Taipei, Taiwan

Master of Science in Zoology, June 1997
National Taiwan University, Taipei, Taiwan. 1995-1997

Bachelor of Science in Zoology, June 1995
National Taiwan University, Taipei, Taiwan. 1991-1995

Publications

Journal papers

1. **Chiu YL** and Rana TM. (2003). Chemical modification of small interfering RNAs imparts functional insight into the molecular mechanisms of RNA interference. *Submitted.*

2. **Chiu YL** and Rana TM. (2002). RNAi in human cells: Basic structural and functional features of small interfering RNA. *Mol. Cell* 10: 549-561

3. **Chiu YL**, Ho CK, Saha N, Schwer B, Shuman S and Rana TM (2002). Tat stimulates cotranscriptional capping of HIV mRNA. *Mol. Cell* 10: 585-597.

4. **Chiu YL**, Coronel E, Ho CK, Shuman S and Rana TM. (2001) HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA. *J. Biol. Chem.* 276 (16):12959-12966

5. Chen LL, Lo CF, **Chiu YL**, Chang CF, Kou GH. (2000) Natural and experimental infection of white spot syndrome virus (WSSV) in benthic larvae of mud crab Scylla serrata. *Dis. Aquat. Organ.* 40(2):157-161

6. Chen CH, **Chiu YL**, Shaw CK, Tsai MT, Hwang AL and Hsiao KJ. (1999) Detection of Borna disease virus RNA from peripheral blood cells in schizophrenic patients and mental health workers. *Mol. Psychiatry* 4(6):566-571.

7. Chen CH, **Chiu YL**, Wei FC, Koong FJ, Liu HC, Shaw CK, Hwu HG and Hsiao KJ. (1999) High seroprevalence of Borna virus infection in schizophrenic patients, family members and mental health workers in Taiwan. *Mol. Psychiatry* 4(1): 33-38

8. Kou, GH., Peng, SE., **Chiu, YL.**, and Lo, CF. (1998) Tissue distribution of white spot syndrome virus (WSSV) in shrimp and crabs. In Felgel TW (ed) Advances in shrimp biotechnology. National Center for Genetic Engineering and Biotechnology , Bangkok. pp267-271.

9. Lo CF, Ho CH, Chen CH, Liu KF, **Chiu YL**, Yeh PY, Peng SE, Hsu HC, Chang CF, Su MS, Wang CH and Kou GH (1997) Detection and tissue tropism of white spot syndrome baculovirus (WSBV) in captured brooders of *Penaeus mondon* with a special emphasis on reproductive organs. *Dis. Aquat. Org.* 30:53-72

10. Lo CF, Ho CH, Peng SE, Chen CH, Hsu HC, **Chiu YL**, Chang CF, Liu KF, Su MS, Wang CH and Kou GH (1996) Whit spot syndrome baculovirus (WSBV) detected in cultured and captured shrimp, crabs and other arthropods. *Dis. Aquat. Org.* 27: 215-225

11. Lo CF, Ho CH, Peng SE, Chen CH, Hsu HC, YT Chen, **Chiu YL**, Chang CF, Liu KF, Su MS, Wang CH and Kou GH (1996) Whit spot syndrome associated baculovirus (WSBV) diagnostic polymerase chain reaction. *COA. Fisheries Series* No. 56. 145-158.

Conference Poster

1. **Chiu YL** and Rana TM. HIV-1 Tat protein interacts with mammalian capping enzyme and stimulates capping of TAR RNA (2001). Cold Spring Harbor Laboratory Eukaryotic mRNA Processing meeting.

2. Wang CH, Lo CF, **Chiu YL** and Kou GH. Characterization of white spot syndrome baculovirus (WSBV) found in penaeid shrimp (1997). 28th Annual Meeting of the World Aquaculture Society. Seattle, Washington USA.

3. **Chiu YL**, Lo CF and Kou GH. THE infection of white spot syndrome associated baculovirus in portunid crabs in Taiwan (1996). 29th Annual

Meeting of the Society for Invertebrate Pathology and IIIth International Colloquium on Bacillus Thuringiensis. Cordoba, Spain.

4. Kou GH, Ho CH, Chen CH, **Chiu YL** and Lo CF. Detection of white spot syndrome associated baculovirus (WSBV) in cultured and wild-caught shrimps, crabs and other arthropods (1996). 29th Annual Meeting of the Society for Invertebrate Pathology and IIIth International Colloquium on Bacillus Thuringiensis. Cordoba, Spain.

Awards & Scholarships

2003 Dean's Award for Outstanding Academic Achievement (2003)
Graduate School of Biomedical Sciences
University of Massachusetts Medical School, Massachusetts, USA

Young Investigator Award (1999-2000 and 2000-2001)
Department of Molecular and Cellular Pharmacology,
RWJMS-UMDNJ, New Jersey, USA

Excellent Student Scholarship (1996-1997)
Ministry of Education, Taiwan

Outstanding Scholarship (Sep/1996)
National Sciences Council, Taiwan