2009

TRANSCRIPTIONAL, EPIGENETIC, AND SIGNAL EVENTS IN ANTIFOLATE THERAPEUTICS

Alexandra Racanelli
Virginia Commonwealth University

Virginia Commonwealth University
School of Medicine

This is to certify that the dissertation prepared by Alexandra C. Racanelli entitled TRANSCRIPTIONAL, EPIGENETIC, AND SIGNALING EVENTS IN ANTIFOLATE THERAPEUTICS has been approved by her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy.

Richard G. Moran, Ph.D., Director of Dissertation

Shirley M. Taylor, Ph.D., School of Medicine

Cynthia Nau Cornelissen, Ph.D., School of Medicine

Tomasz K. Kordula, Ph.D., School of Medicine

Stephen T. Sawyer, Ph.D., School of Medicine

Sumitra Deb, Ph.D., School of Medicine

William L. Dewey, Ph.D., Chair, Department of Pharmacology and Toxicology

Jerome F. Strauss, III, M.D., PhD., Dean, School of Medicine

Dr. F. Douglas Boudinot, Dean of the Graduate School

June 24, 2009

TRANSCRIPTIONAL, EPIGENETIC, AND SIGNALING EVENTS IN ANTIFOLATE THERAPEUTICS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

ALEXANDRA C. RACANELLI
Bachelor of Science, University of Mary Washington, 2001

Advisor: Richard G. Moran, PhD
Professor
Department of Pharmacology and Toxicology

Virginia Commonwealth University
Richmond, Virginia
June, 2009

Acknowledgement

I would like to take this opportunity to recognize some of the individuals who have guided and assisted me throughout my graduate education. I would like to express my sincere gratitude to my advisor, Dr. Richard Moran, for his mentorship, patience, and friendship. His approach to my training has included a delicate balance between fostering and promoting the exploration of my own scientific curiosities and providing the guidance and leadership to ensure my growth and productivity as a graduate student. For his willingness to be a "back-seat driver" even in times where I know he wanted to take the wheel I am truly grateful.

I would also like to thank the members of my graduate committee: Drs. Shirley Taylor, Cynthia Cornelissen, Tomasz Kordula, Steve Sawyer, and Sumitra Deb for their efforts and dedication towards my training. Their willingness to give time, advice, and guidance has been a tremendous help during the more challenging parts of my tenure as a graduate student. They have all taken the time to understand my goals, strengths, and weaknesses, which has greatly enhanced the benefit and productivity of our interactions on both an individual level and within the entire group.

I would also like to thank a few individuals who have been instrumental in my journey through graduate school. I am thankful for Dr. Linda Beck whose training prepared me for graduate school. I am also very grateful to Dr. Gordon Archer for giving me the opportunity to train in the physician-scientist program at VCU. I have benefited enormously from his dedication to the success of this program and all of the students. I would also like to thank Dr. Linda Costanzo who has always been the voice of reason and a listening friend.

Thank you to all my fellow lab members both past and present, Dr. Fiona Turner, Dr. Erin Perchiniak, Dr. Shane Kasten, Lin-Ying Xie, Cortney Heyer, Guoyan Gao, Scott Lawrence, and Scott Rothbart. I would also like to thank the members of the Taylor laboratory Erica Peterson, Dr. Dolores Arjona, Lisa Sale, and Prashant Thakkar. The personal and scientific support, as well as the comical relief, which often lightened difficult and frustrating situations, provided by these individuals has been a highlight of my the past four years. I am deeply grateful for the friendships I have developed with the members of these laboratories.

I would like to also thank my friends Emily White, Tracey Zola, Amanda George, and Amanda Derocco that have been by my side throughout my time at VCU, and their support and dedication to my success has been a true blessing. To my cousin, Vanessa Viola, who has provided endless technical and personal support throughout my graduate school education, thanks right brain!

A special thanks goes to my parents, Joyce and Eugene Racanelli, for their nurturing care and support that has always given me the strength to try even when I was unsure. Words cannot capture how much of my successes have been because of their love and parenting. To my sisters, Liz and Sam, I will be forever grateful for your patience, endless support, humor, and willingness to help wherever possible, LYCS. I would also

like to thank my grandmother (Gram), Lee Viola, for all of her love and support throughout the years. Finally, I would like to especially thank Paul, for standing by my side and giving me confidence and strength when I most needed them. His faith in my abilities and his genuine dedication and support towards my successes are unique gifts that I have had the good fortune to receive. Paul's selflessness and kindness are inspiring and I am truly blessed to have him in my life.

Table of Contents

List of Tables

Page

List of Figures

Page

Abbreviations

DDATHF	**5, 10-dideaza-5,6,7,8-tetrahydrofolic acid**
AICART	**aminoimidazaole ribunucleotide formyl transferase**
GART	**Glycinamide ribonucleotide formyl transferase**
TS	**thymidylate synthase**
DHFR	**dihydrofolate reductase**
LCA	**Lilly compound AICART inhibitor**
HAT	**Histone acetyltransferase**
HDAC	**Histone deacetylase**
ChIP	**chromatin immunoprecipitation**
PCR	**Polymerase Chain Reaction**
bp	**Base pair**
nt	**Nucleotide**
kD	**Kilodalton**
FBS	**Fetal Bovine Serum**
^{o}C	**Celsius**
µl	**microliter**
ml	**milliliter**
H3K4	**histone H3 lysine 4**
H3K27	**histone H3 lysine 27**
H3K36	**histone H3 lysine 36**

H3K9	**histone H3 lysine 9**
Me	**methylation**
Ac	**acetyation**
RNAPII	**RNA polymerase II**
P	**p53-/- mouse embryo fibroblasts**
PM	**p53-/-, DNMT1-/- mouse embryo fibroblasts**
MEF	**Mouse embryo fibroblast**
Q-PCR	**quantitative PCR**
Rt-PCR	**Reverse transcriptase PCR**
EtBR	**ethidium bromide**
mTOR	**mammalian target of rapamycin**
AMPK	**AMP-activated protein kinase**
TSC	**tuberous sclerosis complex**
TdR	**thymidine**
Hx	**Hypoxanthine**
IR	**Inosine**
AICA	**5-Amino-4-imidazolecarboxamide**
AICAR	**1-β-D-ribofuranosyl-5-aminoimidazole-4-carboxamide**
ZMP	**1-β-D-ribofuranosyl-5-Aminoimidazole-4-carboxamide-5′-phosphate**
DNMT	**DNA methyltransferase**
TI	**transcriptional interference**

Abstract

TRANSCRIPTIONAL, EPIGENETIC, AND SIGNAL EVENTS IN ANTIFOLATE THERAPEUTICS

By Alexandra Christin Racanelli, B.S

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2009

Major Advisor: Dr. Richard G. Moran, PhD
Professor, Department of Pharmacology and Toxicology

A targeted approach to the development of antifolate therapies has been sought for many years. Central to the success of such development is an understanding of the molecular mechanisms dictating the sensitivity of cells to antifolates and the fundamental differences of these processes between normal and neoplastic phenotypes. This dissertation addressed transcriptional mechanisms and cell-signaling events responsible for the efficacy of antifolate therapies. Transcriptional processes and cell signaling pathways are often aberrant in neoplastic tissues, providing a potential point of distinction between a normal and neoplastic cellular state.

Folylpolyglutamate synthetase (FPGS) catalyzes the formation of poly-γ-glutamate derivatives of folates and antifolates, which permits intracellular retention and accumulation of these compounds. The mouse *fpgs* gene uses two distant promoters to produce functionally distinct isozymes in a tissue-specific pattern. We questioned how the two promoters were differentially controlled. An analysis of DNA methylation and histone post-translational modifications across the length of the mouse *fpgs* gene showed that epigenetic mechanisms contributed to the tissue-specific control of the upstream (P1), but not the downstream (P2) *fpgs* promoter. RNAPII complexes and general transcription factors were present over P1 only when P1 was transcribed, but these components were present over P2 in most tissues, and promoter-proximal pausing was evident in brain. Clear promoter occlusion was found over P2 in liver. These studies concluded that tissue-specific coordination of dual promoters required multiple interacting controls.

The mammalian target of rapamycin (mTOR) controls protein translation initiation, and is central to a cell-signaling pathway rich in tumor suppressor and oncogenic proteins. mTOR dysregulation is a common feature of several human cancers and inhibition of this protein has been sought as an ideal cancer drug target. We have determined that antifolates inhibiting the two folate-dependent steps of purine synthesis (GART or AICART) activate AMP-dependent protein kinase (AMPK) and inhibit mTOR. The mechanism of AMPK stimulation appears to be mediated by either nucleotide depletion (GART inhibitors), or ZMP accumulation (AICART inhibitors). These studies discovered a new mechanism for antifolates that surprisingly defines them as molecularly targeted therapeutics.

Chapter 1: Introduction and Overview

Folates are forms of the essential water-soluble B vitamin, folic acid. Humans deficient in folates can suffer from a number of serious pathologies including birth defects, such as neural tube defects and cleft palate, cardiovascular disease, and megaloblastic anemia (229). The folic acid molecule consists of a pteridine moiety linked at carbon 6 by a methylene bridge to p-aminobenzoylglutamic acid. Within the cell, folic acid is reduced to 5,6,7,8-tetrahydrofolate (THF), and THF and N5 and N10-substituted reduced folates are the metabolically active forms of this vitamin (Figure 1-1) (9, 167, 199). Folate metabolism is compartmentalized in mammals between the cytosol and the

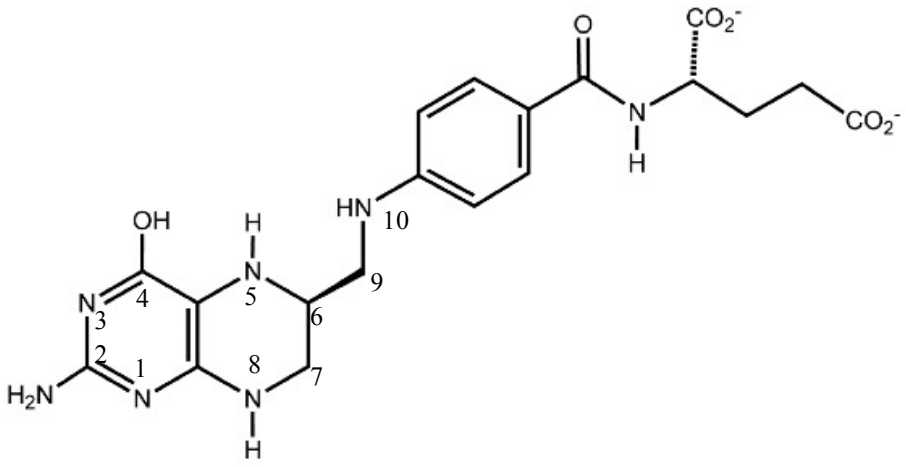


Figure 1-1. Structure of 5,6,7,8-Tetrahydrofolate (THF).

mitochondria, and early evidence showed that folic acid distributes equally between these two compartments in rat liver (47). Cellular folates represent a family of structurally related cofactors that play critical roles in one-carbon transfer reactions required for purine nucleotide and thymidylate synthesis, as well as the synthesis of methionine from homocysteine (9). Methionine is converted to S-adenosylmethionine, which serves as a methyl donor for several methylation reactions within the cell, including DNA and histone methylation (229). Folates are transported into the mitochondria via the mitochondrial folate carrier (158, 183, 241) and serve as cofactors for the mitochondrial serine hydroxymethyltransferase, for formation of formyl groups by the glycine cleavage system, and for the synthesis of the formylmethionine tRNA involved in the initiation of mitochondrial protein synthesis. Regulation of cellular folate levels is a key component of maintaining the fidelity of DNA and RNA synthesis, amino acid metabolism, and epigenetic processes within tissues (229). Several key folate-dependent steps in metabolism are illustrated in Fig. 1-2.

In mammals, circulating folates are monoglutamate forms, which are best viewed as the transport forms of this vitamin (24, 175). Dietary folates are absorbed primarily in the duodenum and upper jejunum of the small intestine through a high affinity proton-coupled folate transporter (PCFT) that functions optimally at low pH (190, 269). In humans, a loss-of-function mutation in the gene encoding the PCFT results in hereditary folate malabsorption disease (190). There are two folate transporters involved in moving folates across membranes at neutral pH: the reduced folate carrier (RFC) and the folate receptors (FR) (30, 79). The RFC moves folates in a facilitative manner using an anion exchange mechanism and the FRs transport folates intracellularly via endocytosis (269).

RFC null mice can survive embryonic development only if pregnant females are injected with folic acid, but pups die post-natally at 12 days due to hematological pathology; transport using this protein is critical throughout the life of the animal (269). On the other hand, mice where all FRs have been deleted die during development unless the pregnant females are given 5-formyl-THF; these mice survive gestation and develop into adulthood normally, highlighting the necessity of these receptors to maturing embryos (269).

Once inside the cell, folates serve as substrates for the enzyme folylpoly-γ-glutamate synthetase (FPGS), which catalyzes the conversion of monoglutamate folates to poly-γ-glutamate derivatives using ATP and glutamic acid (16, 167, 199) (Fig. 1-2). Mitochondrial and cytosolic forms of FPGS are encoded by the *fpgs* gene in dividing cells in mouse and humans using alternative start sites, which are distinguishable by the presence of a mitochondrial leader sequence at the N-terminus of the mitochondrial protein. The two isoforms allow the accumulation of folylpolyglutamates in the cytosol and mitochondria, the two cellular compartments involved in folate-dependent one-carbon metabolism (72). Folypolyglutamates are more efficient substrates for many folate-dependent enzymes and the cellular retention and accumulation of cofactors to high levels is necessary to meet the metabolic demands of rapidly dividing cells (167, 239). Mammalian cells that harbor inactivating FPGS mutations die from lack of end products of folate-dependent steps of metabolism in normal media (157, 239).

In the early 1940's, Dr. Sydney Farber at the Jimmy Fund in Boston made an observation that the administration of folic acid to children with acute lymphocytic leukemia (ALL) increased the progression and severity of the disease (65). Because of

this remarkable observation, he generated and tested the hypothesis that a folic acid antagonist might inhibit or arrest the proliferation of tumor cells and improve the mortality and morbidity of pediatric ALL, which at the time was an incurable disease. In 1948, he published the results of his study in the New England Journal of Medicine: administration of the folate antagonist 4-aminopteroyl-glutamic acid (aminopterin, AMT) to children with ALL frequently led to complete remissions of this disease, the first time such a remission was every observed (66). Farber's approach to target highly proliferative tumor cells by inhibiting steps critical to cellular division has been the basis for the development of chemotherapeutic agents and strategies for the past 60 years (162). This fundamental study marked the beginning of the use of folate antagonists, i.e. antifolates, as chemotherapeutic agents.

The use of AMT was replaced shortly after the original studies with a slightly less toxic compound, methotrexate (MTX) (64). Both MTX and AMT primarily inhibit dihydrofolate reductase (DHFR); the enzyme responsible for the reduction of folate and dihydrofolate to tetrahydrofolate. Because reduced folates are the active folate cofactors required for one carbon transfer metabolism, treatment with AMT or MTX leads to the inhibition of purine and thymidylate biosynthesis, with severe consequences on DNA and RNA production. MTX remains the first line therapy for ALL and is now even more frequently used in the treatment of several autoimmune diseases including psoriasis, rheumatoid arthritis, and sarcoidosis (189, 196, 197).

The basis for antifolate drug design efforts shifted toward the production of analogs that could be highly metabolized following the observation that MTX, like the naturally occurring folates, was metabolized to polyglutamate derivatives (17). MTX

polyglutamates accumulate to high intracellular levels with sufficient exposure of leukemic cells, and are retained for substantial periods of time following removal of extracellular drug (17, 73, 199). Long-term inhibition of DHFR is mediated by the poly-γ-glutamyl derivatives of MTX (199, 212) and the targets of this form of the drug are expanded to two other folate-dependent enzymes: thymidylate synthase (TS) (3) and 5-aminoimidazole carboxamide ribotide transformylase (AICART) (5). The intracellular formation of polyglutamate forms of MTX greatly enhance the cytotoxicity of this drug, and extend the length of time that intracellular metabolism is disrupted due to intracellular retention of the drug (41, 199). The importance of FPGS activity on the efficacy of MTX suggested that the rationale for selection of antifolate agents should target compounds that are efficient substrates of this protein.

In the late 1970's, the main target for the design of novel antifolate metabolites shifted from DHFR to the folate-dependent enzymes involved in thymidine and *de novo* purine synthesis. The first drug to emerge from these efforts was N^{10}-propargyl-5,8-dideazafolate (CB3717), a potent TS inhibitor (125). CB3717 was a moderately efficient substrate of FPGS; polyglutamated derivatives of CB3717 were 100-fold more potent as inhibitors of thymidylate synthase than the parent compound and were no longer efficient substrates for cellular efflux (117). Phase I clinical trials determined that the compound was active against breast, ovarian, and liver cancer, but CB3717 had remarkable liver and kidney toxicity due to the insolubility of the drug in acidic pH (35, 117). In spite of the fact that CB3717 proved to be too toxic for use in humans, the potent effect of this compound under clinical conditions was the first evidence suggesting that designing antifolates targeting thymidylate synthase or other folate-dependent biosynthetic enzymes

had great chemotherapeutic potential (35, 117). This led to the synthesis of a number of additional molecules that were more water-soluble and a second generation analogue of CB3717, N-(5-[N-(3,4-dihydro-2-methyl-4-oxoquinazolin-6-methyl)-N-methylamino]-2-thienoyl)-L-glutamic acid (DI694, ralitrexed, RTX), which was found to be more potent and had less toxic side-effects than the parent compound (118). Like CB3717, RTX is a substrate for FPGS, in fact, a far superior substrate, and polyglutamate deriatives are more potent TS inhibitors than the than unmodified form of this compound (117, 118). RTX is currently used in the treatment of advanced colorectal carcinoma disease in Europe and Asia, but was never approved for use in the United States (50).

A second line of drug development efforts for tetrahydrofolate antimetabolites resulted in the identification of 5,10-dideaza-5,6,7,8-tetrahydrofolate (lometrexol, DDATHF), a potent inhibitor of GART, an enzyme in *de novo* purine synthesis(18, 165, 237) (Figures 1-3 and 1-4). Cell culture experiments determined that DDATHF was a potent inhibitor of cell growth and a very efficient substrate for FPGS, allowing conversion of drug to polyglutamate forms at low μmolar concentrations in cells (18, 165). Further studies showed that the polyglutamated forms of this molecule were more potent inhibitors of GART than the parent drug (205). DDATHF was a promising antifolate; however, in phase I clinical trials DDATHF was found to induce severe unmanageable thrombocytopenia which prevented repetitive administration of the treatment (191). After nearly a decade of attempts to circumvent this toxicity, combination of DDATHF and folic acid tremendously improved the toxicity profile of lometrexol(195), but the use of this compound in the clinic was never successfully pursued.

The efficiency of both 6-R and 6-S diastereomers of DDATHF as a substrate for FPGS and the tight binding affinity of these compounds for GART initiated the synthesis of second-generation *de novo* purine synthesis inhibitors. Eliminating the C5 methylene position in DDATHF produced a potent cytotoxic agent, LY231514 (pemetrexed, PTX), which was originally tested in human CCRF-CEM cells(238)(Figure 1-4). Surprisingly, cell growth reversal studies determined that this agent was no longer an inhibitor of GART, but rather a potent inhibitor of TS (238)(Fig. 1-2). PTX is a very efficient substrate for FPGS, perhaps the most efficient FPGS substrate ever made, creating anabolites that are retained intracellularly after drug exposure to submicromolar concentrations of parent drug (224, 238). Very importantly, cell culture experiments determined that, unlike other traditional TS inhibitors, administration of exogenous thymidine, which prevents thymidylate synthase inhibition, does not completely reverse the cytotoxic effects observed from the drug; this suggested that PTX had at least one other target (224). A study of the activity of PTX against several recombinant rodent and human enzymes *in vitro* led to the conclusion that both GART and DHFR were potential secondary targets for the polyglutamate forms of PTX (224). PTX is currently used as first line therapy for the treatment of non-small cell lung cancer and mesothelioma and is being tested in phase II clinical trials for the treatment of renal cell cancer in combination with gemcitabine (42). In chapter 4 of this dissertation, we questioned the conclusion that GART and DHFR were the secondary targets of PTX, and found that in fact, the additional site of PTX-mediated inhibition was AICART, a site of action that had substantial impact for understanding the activity of this drug.

The rationale for the design of antifolates as chemotherapeutic agents has been based largely on the understanding that tumor cells are highly sensitive to alterations in the pathways involved in nucleotide biosynthesis and amino acid metabolism. This has proven very useful for the treatment of numerous types of cancers including ALL and colorectal carcinoma with MTX and Ralitrexed, respectively. However, as is the case for many classes of chemotherapeutic agents, the clinical use for many of these antimetabolites is limited by toxicity in normal dividing tissues, such as those found in the gastrointestinal tract. As such, a large effort has been made over the past two decades to design drugs against molecular targets that are unique to tumor cells. The development and use of gleevec for the treatment of chronic myelogenous leukemia is the prototypical paradigm for rational drug design (55). Gleevec specifically inhibits the kinase activity of the Bcr-Abl fusion protein made from the gene modified by the Philadelphia chromosome translocation, a protein that does not occur in humans without this translocation (55), making treatment based on Bcr-Abl similar to antimicrobial chemotherapy wherein drugs can be designed against proteins peculiar to prokaryotes.

Our laboratory has focused intensely on understanding the molecular basis for the effectiveness of antifolate agents in an effort to identify a potential mechanism that could be exploited to specifically target tumor cells with this class of antimetabolites. As discussed above, central to the therapeutic utility of a large number of antifolates is their ability to serve as an efficient substrate for polyglutamation, catalyzed by the enzyme FPGS. A few laboratories, including ours, have focused on understanding the control of this protein at the molecular level in normal tissues and tumor cells. FPGS protein levels in mammals are regulated in both a division-specific and tissue-specific fashion and both

of these controls occur at the transcriptional level (12, 71, 244). Abundant FPGS levels are detected in all normal and neoplastic dividing cells and are only found in a few select differentiated mammalian tissues (12, 166). In mouse, there are two isoforms of FPGS produced from two promoters, which include different first exons in the mRNA: one protein is found exclusively in dividing cells and the second is found only in differentiated tissue (244). The isoforms of FPGS identified have different substrate specificity attributed to the differences in the sequences of the N-terminal peptides (244). These studies suggested that the design of antifolates specific for polyglutamation by individual FPGS isoforms could be a potential mechanism of tissue-selective targeting of antifolate therapy for cancer. To date, appreciable levels of expression of only one isoform of FPGS has been detected in human tissues, even though the second *fpgs* promoter used in mice has been identified in humans (245). These observations initiated the studies presented in chapters 2 and 3 of this dissertation. We sought an understanding of the regulatory mechanisms involved in the synthesis of one FPGS isoform over a second in mouse tissues. We identified and explored the epigenetic and transcriptional interference mechanisms regulating the two promoters of the mouse *fpgs* gene, in an effort to understand how tandem promoters are coordinately controlled in mammals and, perhaps, to define the conditions where both promoters may be used in humans.

As discussed above, PTX inhibits TS primarily, but cell culture experiments suggest that this drug also has activity against additional folate-dependent enzymes (224). With the success of PTX in the clinical setting, understanding the additional targets of PTX became highly relevant to both current treatment regimens and to identifying additional cancer types that may be sensitive to PTX. We have recently performed a

series of cell culture experiments that have developed evidence that the target for PTX secondary to TS is the second folate-dependent enzyme in *de novo* purine synthesis AICART. These studies are discussed in chapter 4 of this dissertation. The significance of this is substantial, since the substrate for AICART is ZMP, a known activator of the AMP-activated protein kinase, AMPK (49), a key controlling element in the mTOR pathway (48). The mTOR pathway, responsible for balance of energy metabolism, protein and lipid synthesis, and growth, involves a series of upstream controlling proteins recognized as tumor suppressor proteins, including LKB1, PTEN, TSC1 and 2, and others recognized as cellular oncogenes, such as AKT and PI3 kinase (48). Aberrant mTOR regulation is a common phenotype to a variety of cancers due to loss of tumor suppressors or over expression of oncogenes within the pathway. Inhibition of this pathway by a clinically successful antifolate, PTX, suggests that its efficacy may involve mTOR inhibition and the development of additional antifolates specifically blocking AICART may be a promising molecular targeted approach towards neoplastic cells dependent on overactive mTOR.

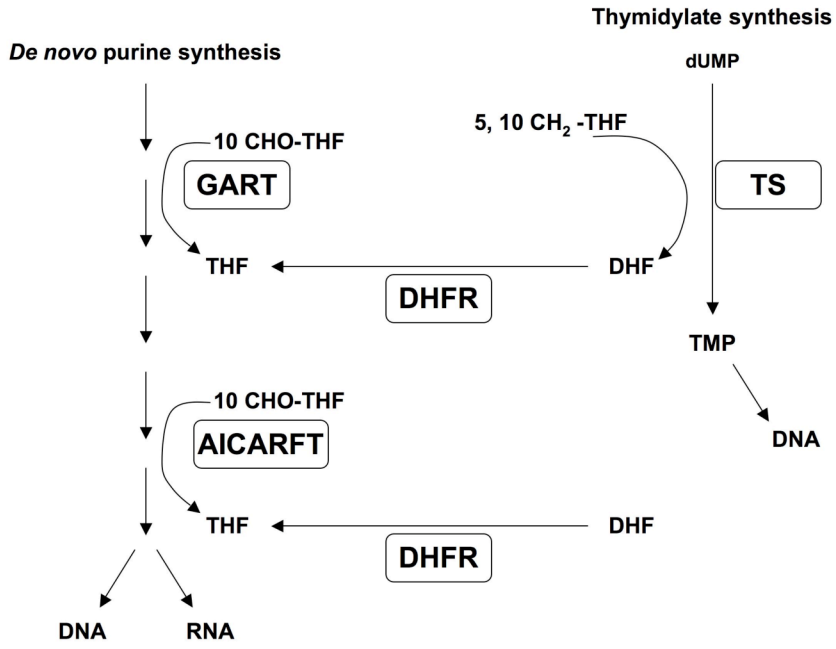


Figure 1-2. Pathways involved in folate metabolism. GART (glycinamide ribonucleotide formyltransferase), AICART (aminoimidazole carboxamide ribonucleotide formyltransferase), DHFR (Dihydrofolate reductase), TS (thymidylate synthase) THF (tetrahydrofolate), 10 CHO-THF (10-formyl tetrahydrofolate), 5, 10 CH_2 -THF (5, 10-methylene tetrahydrofolate). Adapted from Dr. Julie Bronder's dissertation (31).

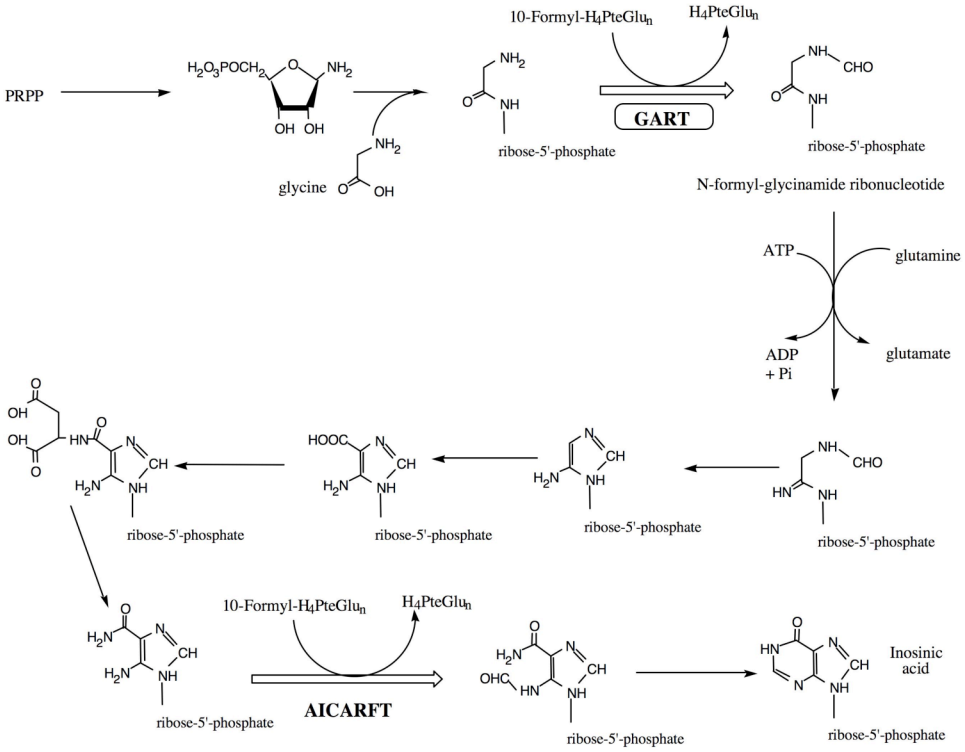


Figure 1-3. Steps in *de novo* purine synthesis. The two folate-dependent enzymes in this pathway are glycinamide ribonucleotide formyltransferase (GART) and aminoimidazole carboxamide formyltransferase (AICARFT).

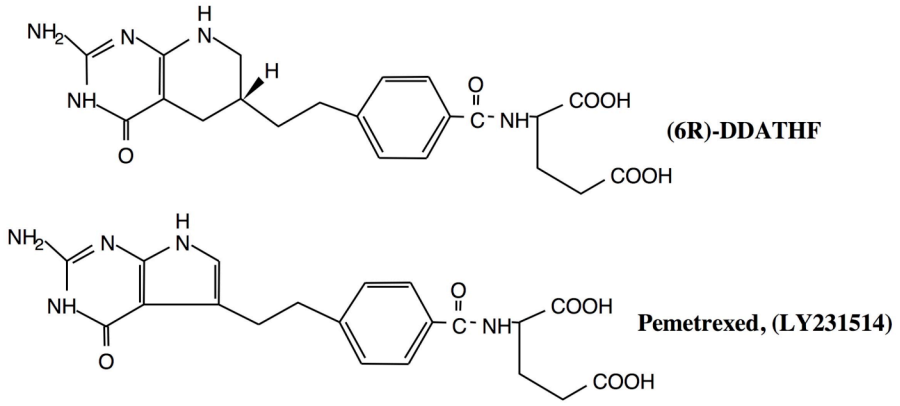


Figure 1-4. Chemical structures of the antifolate compounds (6R)-DDATHF and 231514.

Chapter 2: Multiple control mechanisms at the dual promoter mouse *fpgs* gene responsible for coordination of tissue-specific expression *in vivo*.

INTRODUCTION

I. Folylpoly-γ-glutamate synthetase (FPGS) and the tissue-specific expression pattern of the mouse *fpgs* gene.

In mammals, circulating folates are monoglutamate forms, which are best viewed as the transport forms of this vitamin (24, 175). Folates serve as cofactors in metabolic pathways that require one-carbon transfer; i.e., *de novo* purine synthesis and thymidylate synthesis (9). Dietary folates are transported into the cell by the folate receptor, the reduced folate carrier, and the proton-coupled folate transporter (30, 79, 190). After passage into peripheral cells, folates are converted to poly-g-glutamate derivatives by the enzyme folylpoly-g-glutamate synthetase (FPGS) (16, 167, 199). The addition of multiple glutamate residues to these molecules traps them within the intracellular compartment. Without this metabolic trapping mechanism, mammalian cells die for lack of the end products of folate metabolism (157).

FPGS is also necessary for the action of most antifolates, and point mutations in FPGS are a common mechanism for tumor cell resistance to these drugs (10). As is the case for endogenous folates, polyglutamated antifolate molecules are not substrates for

efflux and become concentrated within the intracellular compartment, maintaining the inhibitory effects of these molecules throughout the lifespan of the cell, a major therapeutic advantage for drugs that are efficient FPGS substrates. Since the cytotoxic effects of antifolate therapies depend upon the presence of FPGS, determining the levels of this protein in tumors and normal tissue was important to identify candidate tumor-types and normal tissues that would be highly sensitive to these therapies.

Early studies carried out by Moran and Colman used an *in vitro* enzyme assay to determine the FPGS activity in normal and neoplastic tissues of the mouse (166). The distribution of FPGS activity found in these mammalian tissues suggested that the expression of this protein was under strict regulation (12). High levels of enzyme activity were detected in all normal and neoplastic dividing cells, as well as certain tissues with a known dividing cell compartment, i.e. intestinal epithelium (12). In contrast, analysis of differentiated mouse tissues determined that only in mouse liver and kidney are measurable levels of FPGS expressed (12). Further analysis in developing rodent fetal tissues supported the idea that rapidly dividing cells require FPGS (12). In contrast, maturation of adult tissues and the induction of cellular growth arrest precipitated a rapid decline in FPGS expression (12, 59). From these studies, it was concluded that FPGS expression is controlled by mechanisms linked to both cellular proliferation and tissue specificity (12, 59).

The concept that transcriptional regulation is involved in establishing the FPGS activity patterns was first suggested through Northern blot analysis performed by Sarah Freemantle, a previous post-doctoral fellow in this laboratory, examining the levels of *fpgs*

mRNA in several tissues of the mouse (71) (Fig 2-1). The results of this experiment showed striking patterns of tissue-specific *fpgs* mRNA expression. Abundant levels of *fpgs* transcript were detected in mouse leukemic cells, L1210, and in two differentiated tissues: mouse liver and kidney (Fig 2-1). These expression profiles of *fpgs* mRNA determined for each tissue paralleled the FPGS pattern of activity found in previous studies (71). The Northern blot studies supported the original idea that the control of FPGS expression is based on mechanisms that are both division-specific and tissue-specific and suggested that regulation was occurring at the level of transcription.

The mouse *fpgs* gene has two promoters spaced 9.5 kilobases apart within a genomic locus spanning 20 kilobases (Fig. 2-2) (200, 201, 244). The downstream promoter (P2) is located in a CpG island and is activated when three Sp1 proteins bind to the consensus sites located immediately upstream of the transcriptional start site (43). The upstream promoter (P1) is classified as CpG-sparse, and while attempts have been made to understand the *cis* elements and *trans*-acting factors controlling P1, they remain poorly defined (43). This is most likely the case because studies performed to identify regulatory regions using reporter promoter constructs are based on transient transfections of dividing cells, but the P1 promoter is expressed only in differentiated tissue and the required factors may not be present in the transfected cell line. Transcripts generated at P1 include two upstream exons (A1a and A1b) linked to exons 2-15 (Figure 2-2) (202, 244). In contrast, when P2 is used, mature transcripts contain exons 1-15 (Figure 2-2). The alternate *fpgs* mRNAs produce two isozymes with regulatory consequences: the protein generated from

P2 transcript is subject to feedback inhibition by folate polyglutamates, and the enzyme encoded by P1 mRNA is less sensitive to this control (7).

Fiona Turner, a former student in this laboratory, used ribonuclease protection assays (RPAs) to define the conditions for the use of the two mouse *fpgs* promoters (244). These studies uncovered an additional layer of complexity at the mouse *fpgs* gene that greatly helped to explain earlier data. Dr. Turner's studies established the premise for the experiments discussed in this chapter: she established that the use of the P2 promoter is reserved for all neoplastic and normal dividing cells (244). This conclusion was stated on the basis of definitive RPAs demonstrating that *fpgs* mRNA from dividing cells, i.e. L1210, when hybridized to a probe from the P1-specific exon generated a protected fragment size supporting that the *fpgs* mRNA contained exons 1/4, but not exons A1a or A1b (Figure 2-3, left panel). However, when L1210 mRNA was hybridized to the P2 probe, the protected fragment size indicated the presence of exons 1-4 in the *fpgs* transcript (Figure 2-B, right panel). In contrast, P1 was found to be the promoter almost exclusively used in mouse liver and kidney, as evident by the fragment size detected when the P1 probe was hybridized to the *fpgs* mRNA (Figure 2-3, left panel) (243, 244). Hence, from these studies it appeared that transcriptional initiation at the mouse *fpgs* gene is tightly regulated to ensure that tissues expressing this gene are restricted to the use of only one promoter.

We were very interested in these observations and saw the mouse *fpgs* gene as an excellent model of a complex genetic locus employing the use of multiple mechanisms to achieve tissue-specific patterns of expression from two promoters. We designed a series of

studies aimed at understanding the transcriptional mechanisms dictating promoter choice at the mouse *fpgs* gene permitting the synthesis of one FPGS isoform over the other in a tissue-specific fashion. Our studies focused on mechanisms of epigenetic regulation and transcriptional interference at the mouse *fpgs* locus.

II. Basic mechanisms of transcription

A. Overview

The process of transcription can be divided into three general stages: initiation, elongation, and termination. Throughout the transcription cycle the transcriptional machinery is faced with substantial obstacles that must be managed in order to maintain high fidelity transcription. These challenges include but are not limited to: 1. steric hinderance by nucleosome placement along the DNA, 2. pre-mRNA processing involving splicing factors and machinery to ensure correct mature mRNA synthesis, and 3. 5' capping and 3' end polyadenylation. As a result, every phase of the transcription cycle requires a highly coordinated interplay between RNAPII, transcription factors, histone modifying enzymes, chromatin remodelers, and mRNA processing factors. We discuss the stages of transcription and some mechanisms involved in regulating transcription below as separate sections for the sake of clarity, keeping in mind that it is the interplay of all these events that coordinate the complex patterns of eukaryotic gene expression (Figure 2-4).

B. General features of transcriptional initiation:

The formation of a pre-initiation complex (PIC) occurs when the general transcription factors IID, IIB, IIE, IIF, IIH, TBP and RNAPII assemble at the promoter region of a gene (90). Activators bind *cis* elements located upstream of the core promoter and recruit coactivator complexes, such as SAGA and/or mediator, that mediate regional acetylation of histones and chromatin remodeling to allow PIC formation (90, 132). An open complex is formed when the helicase activity of TFIIH melts the DNA surrounding the transcriptional start site and the template is positioned in the active site of RNAPII (90, 132). Transcription initiation begins when the first phosphodiester bond of RNA is formed. Upon successful initiation, the complex undergoes a process referred to as promoter clearance, often involving a series of abortive initiation events (Fig 2-4). In cases where a gene is actively transcribed, RNAPII escapes from the factors bound to the promoter and proceeds down the template strand in the 3' to 5' direction.

C. General features of transcriptional elongation

During elongation, a set of regulatory proteins traverse the coding region through direct interactions with RNAPII, unlike the DNA binding proteins associated with initiation. The C-terminal domain (CTD) of the Rbp1 subunit of RNAPII provides the scaffold for secondary proteins to bind to the elongation complex (2, 90, 104, 133). This region is composed of 25-52 repeats of the heptad sequence YSPTSPS. The differential phosphorylation of the serines at positions 5 and 2 determine the proteins recruited to the enzyme along the length of the gene (90, 133) (Fig 2-4). At the 5' end of a gene, CDK7, a kinase in the TFIIH complex, phosphorylates serines at the 5 position of the heptad repeat

potentially generating a highly modified CTD (Figure 2-4). This modification recruits and activates the 5' mRNA capping machinery through binding guanyltransferase, a member of the capping complex (104). Additionally, phosphorylation of serine 5 of the CTD recruits the Polymerase Associated Factor (PAF), which facilitates the binding of histone H3 lysine 4 methyltransferase (HMTs), and FACT, a histone chaperone needed to mobilize nucleosomes off the DNA 3' to the elongation machinery and replace the nucleosomes behind the transit of the elongation complex (92, 141, 262). Towards the 3' end of a gene, phosphorylation of Serine 2 accumulates as a consequence of CDK9 kinase activity, a protein member of the P-TEFb elongation complex (2) (Figure 2-4). This modification recruits polyadenylation factors and histone H3 lysine 36 methyltransferase, proteins specific for the terminal phase of the transcription cycle (2, 130) (Figure 2-4). The distinguishing features of the proteins bound to RNAPII and the CTD PTMs during the transcription cycle provide methods for distinction between the different phases of transcription. We employed the use of these tools in our studies examining transcriptional processes occurring at the mouse *fpgs* gene.

D. Promoter-proximal pausing

Promoter-proximal pausing, also known as stalling or poising, was first described at *c-myc*, human-immunodeficiency virus (HIV), and heat shock response genes in *Drosophilia* (184). In the original studies at the *c-myc* gene, full-length mRNA levels of *c-myc* declined upon cellular differentiation, but nuclear run-on assays detected persistent RNA synthesis between exon 1 and intron 1 (19). The results from these experiments

suggested that the 5' end of the gene was engaged in the production of mRNA transcript, even when full-length *c-myc* mRNA was not being produced. As such, it appeared that a block at elongation, rather than initiation was controlling expression of *c-myc* mRNA (19). The components controlling promoter-proximal pausing have been identified largely through studies using transcription inhibitor 5,6-dichloro-1-b-Dribofuranosylbenzimidazole (DRB) (154). Increasing doses of DRB have been shown to greatly reduce the synthesis of 100 nt transcript *in vitro*, but short-length transcripts of 25-30 nt were able to accumulate even at high concentrations of drug, suggesting that drug-mediated inhibition of transcription occurred at a point in elongation (253). Two negative elongation factors, DRB sensitivity-inducing factor (DSIF) (252) and Negative Elongation Factor (Nelf) (264) were identified in studies examining the components responsible for DRB sensitivity (Figure 2-4). These proteins associate with RNAPII and have been shown to control promoter proximal pausing in both *Drosophila* and human cells (259). PTEFb is DRB-sensitive and responsible for the release of RNAPII complexes from a stalled mode to a productive elongation state (153, 154) through the phosphorylation of serine 2 of the CTD of RNAPII and DSIF and NELF (184) (Figure 2-4) (266). Current thought proposes that the phosphorylation of DSIF and NELF results in the release of NELF from the stalled complex, and the entrance of TFIIF, a factor that enhances productive elongation (46). While many of the major players involved in promoter-proximal pausing have been identified, the sequence of events leading to the establishment of a poised state and the release from pausing into a productive elongation mode are largely unknown.

E. Tissue-specific expression

Over two hundred different cell types comprise mammalian organisms and are created from varying transcription of an identical genomic sequence. A population of mRNA transcripts is common to all cells, and are generated from genes commonly referred to as housekeeping genes, however expression of unique subsets of genes are reserved for specific cell-types and the resulting expression patterns establish cellular identify. A major determinant of tissue-specific expression is the production of lineage-specific transcription factors in developing and differentiated tissues. Tissue-specific transcription factors bind *cis*-acting DNA elements and through direct interactions with basal transcriptional machinery and coactivators or corepressors impact gene expression. For example, there are six families of liver-specific factors characterized to date: HNF1, HNF3, HNF4, HNF6, C/EBP and D-binding protein (215), and their presence support the expression of liver-specific genes in developing and differentiated hepatocytes (266). The expression of liver-specific genes is hugely important in maintaining a differentiated state, and loss of HNF1 and HNF4 is a common feature of mammalian hepatoma cell lines. Re-expression of these proteins has been correlated with the reestablishment of liver-specific cellular processes found in mature hepatocytes (228).

The accessibility of consensus sequences in tissue-specific promoters to the DNA binding proteins is considered to play a major role in regulating patterns of tissue-specific gene expression (215). As such, patterns of chromatin histone post-translational modifications (PTMs) and DNA methylation should be considered regulators of tissue-specific expression. For instance, precedent literature argues that certain histone PTMs are

a prerequisite for the sequence-specific binding of the transcription factor Myc (86). The role of epigenetics in tissue-specific expression remains incompletely studied. Hence, we have considered the mouse *fpgs* gene as a model of tissue-specific expression and set out to examine the patterns of DNA methylation and histone PTMs across this locus in different tissues to address the potential roles of these epigenetic mechanisms in establishing the tissue-specific expression patterns observed.

III. Epigenetics mechanisms and transcription

A. General comments

Epigenetics is the "study of heritable changes in gene function that occur without changing the underlying DNA sequence". The field of epigenetics is extremely broad and covers topics of chromatin biology, DNA methylation, and non-coding RNAs. The importance of these processes in gene regulation is tremendous and genome-wide studies using advanced ChIP-Chip and ChIP-Seq technologies have drastically enhanced our awareness of the complexities involved in epigenetic regulation. However, we are far from understanding the entire story. In fact, a current challenge of the field is understanding the interplay between the different epigenetic mechanisms and defining the sequence of events leading to observable changes in gene regulation. In our studies, we have used the mouse *fpgs* gene as a model to consider the potential cooperation between chromatin structure and modifications and DNA methylation in the determination of tissue-specific patterns of expression in the adult animal.

B. Histone biology

A major step forward in understanding the chromatin structure was taken in 1974 by Robert Kornberg, when he proposed that the structure was based on a repeating unit of two molecules of the histones H2A, H2B, H3, and H4, and approximately 200 base pairs of DNA (Figure 2-5) (134). Interestingly, the concept that histones were not simply structural obstacles interfering with RNA synthesis was described ten years prior to Kornberg's chromatin structure theory, when a series of experiments revealed that nuclear histones were both acetylated and methylated (6). In this study, Allfrey proposed that such modifications, particularly acetylation, "may affect the capacity of histones to inhibit ribonucleic acid synthesis *in vivo*" (6).

The fields defining the roles of chromatin structure, modifications, and function in DNA replication and RNA transcription have greatly expanded the observations made in these initial studies. The positioning of nucleosomes within the genome appears to be highly regulated, supported by the fact that nucleosome free-regions exist at a large proportion of promoter regions across the genome (141). Assembly of Kornberg's histone-DNA units, now called the nucleosome, positions the amino-terminal tail of each histone such that it protrudes from the core structure and is available to serve as a substrate for modifying enzymes (135). Histone post-translational modifications (PTMs) include acetylation, methylation, phosphorylation, ubiquitination, and sumoylation (135) (Figure 2-5). The enzymes responsible for modifying histones are recruited to specific positions in chromatin often through protein-protein interactions with molecules involved in transcriptional elongation, DNA methylation, and variety of other cellular processes.

These interactions result in the deposition of histone PTMs across chromatin in a highly ordered fashion to create a pattern of histone PTMs, originally referred to as the histone code (119), which are now thought to be critically involved in the regulation of transcription.

The truly sweeping functions of histone PTMs are currently an area of intense investigation. Current thought proposes that the roles of histone PTMs in transcription can be classified into at least one of three general mechanisms: 1. Histone PTMs, such as histone acetylation, affect chromatin stability and structure leading to changes in DNA availability to factors in the cellular *milieu*, 2. Histone PTMs can alter the genomic processes occurring at a specific DNA location by either *preventing* or *recruiting* the binding of factors to modified areas of chromatin, and 3. Histone PTMs are dynamic and reversible and reflect the changing activities over a stretch of DNA sequence. Several histone PTMs known to be involved in transcription have been profiled across the locus of a highly active gene with a single promoter (247). In the studies discussed in this chapter, we extended these observations through the examination of several histone PTM profiles over the two promoters of the mouse *fpgs* gene in several tissues.

i. Chromatin modifications involved in transcriptional regulation

We elected to profile histone PTMs across the mouse *fpgs* gene that are known to play a role in the transcriptional activation or repression of genes in yeast and other systems (141). We studied histone H3 and H4 acetylation, as well as, histone methylation

at lysines 4, 9, 27, and 36. Composite profiles of these modifications generated from genome-wide studies in yeast are shown in Figure 2-6.

ii. General terms and concepts

Bromodomain: Proteins that contain this domain bind to acetyl-lysine residues in histones N-terminal peptide tails. Structural analysis of the p300/CBP coactivator identified three amino acid residues, Tyr 760, Tyr 802, and Asn 803, involved in recognition of the acetyl-lysine and these residues are conserved across family members (168). The aromatic residues form a hydrophobic binding pocket in which neutral acetylated lysines are bound preferentially to unmodified lysine residues (27). Additional proteins containing bromodomains have a range of functions and include HATs (i.e. GCN5), transcription factors (i.e. TAFII250), and chromatin remodelers (i.e. Brg1 in Swi/Snf complex) (168).

Chromodomain: This domain consists of 50 amino acids and is composed of α-helices and a three stranded anti-parallel β-strand sheets. Chromodomains are found in a variety of proteins and facilitate protein-protein and protein-DNA interactions (27). One important recognition site for chromodomains is methylated lysines within histones, which allows the docking of proteins involved in chromatin regulation directly to the nucleosome. Heterochromatin Protein 1 (HP1) was the first protein identified to bind histones through its chromodomain at methylated lysine 9 of histone H3 (172). Structural analysis of HP1 with a methylated histone H3 peptide revealed that three aromatic side chains Tyr 21, Trp 42, and Phe 45 form a hydrophobic binding pocket for the *N*-methyl groups on lysine 9

(172). Chromodomain containing proteins include chromatin remodeling complexes (e.g. CHD1 and Mi2) and Polycomb proteins (135).

SET domain containing methylases: This class of enzymes are responsible for the methylation of lysines at residues 4, 9, 27 and 36 of histone H3 (225). Structural studies have shown that the N-terminal and C-terminal regions of the domain orient to form a hydrophobic pore, referred to as the lysine-access channel, which positions lysines in close proximity to the methyl donor S-adenylhomocysteine (27).

Polycomb complexes: Polycomb group proteins (Pcg) mediate the stable repression of key developmental genes in embryonic stem cells to maintain a state of pluripotency (29). The PcG proteins form two complexes, Polycomb repressive complex (PRC) 1 and 2. PRC2 is composed of EE2, EZH2, and SUZ12 in mammals and is directed to target genes through recognition of the DNA *cis* element the Polycomb Response Element (Pre) (225). EZH2 catalyzes the addition of methyl groups to lysine 27 of histone H3(225). PRC1 directly binds to H3K27me3 and stabilizes compaction of chromatin and transcriptional repression. Genome-wide studies using mouse and human embryonic stem cells have located PRC2 and PRC1 components at genes predominately involved in development and differentiation (29, 139).

iii. Histone acetylation and deacetylation

Histone acetylation is the most comprehensively studied histone PTM to date and is a proven component of several aspects of histone biology, including nucleosome assembly, chromatin folding, and transcriptional control (218). The addition of an acetyl group to the

ε-amino of specific lysines within the histone tail neutralizes the positive charge of this region, which, in the context of the nucleosome, disrupts the charge-charge interaction between the positively charged histone tail and the negatively charged DNA (218). As a result, nucleosome stability is compromised and the availability of DNA sequence to the binding of cellular *trans* factors necessary for transcription is greatly enhanced (249). Additionally, acetylated histones are docking sites for proteins with bromodomains (135).

All four of the core histones serve as substrates for the family of histone acetyltransferases (HATs), which carry out the enzymatic acetylation of histone tails (218). Interestingly, several transcription regulatory proteins, i.e. TAFII250, possess acetyltransferase activity (230). In contrast, histone deacetylases (HDACs) are a class of proteins that remove acetyl groups from lysines. HDACs mediate the silencing of genes and are often associated in multi-subunit repressor complexes (218). HDACs detected at the coding regions of active genes have been shown to facilitate transcriptional elongation through preventing initiation at cryptic promoters located within the body of genes.

Recent studies have described that histone acetylation at active genes is restricted to the 5'end of the locus (Figure 2-6) (247). Genome-wide studies in yeast and higher eukaryotes have mapped the acetylation of lysines 9 and 14 of histone H3 and lysine 16 of histone H4 to the 5' end of active genes (186, 198, 216). Interestingly, abundant acetylation at these lysines has also been detected across promoter regions poised for transcriptional activation, such as the p21 promoter (81). Commonly detected profiles of histone H3 and H4 acetylation across an active genetic locus are illustrated in Fig 2-6.

iv. Histone lysine methylation

Genome-wide studies have shown that different histone methylation profiles can be used to distinguish genomic regions as transcriptionally active or silent. Both lysines and arginines within the histone tails are acceptor sites for methylation (225). The ε-amino group of histone lysine residues can accept up to three methyl groups and therefore can exist in either a mono-, di-, or tri-methylated form, adding an additional layer of complexity to this class of histone PTMs compared to histone acetylation (225). The family of enzymes responsible for methylating histone lysines is referred to as histone methyltransferases (HMTs), and the residues discussed below are substrates for the SET domain-containing class of HMTs (225). Originally, chromatin marks of histone methylation were thought to be permanent, but the first histone demethylase, LSD1, was identified a few years ago and several other classes of demethylases have since been identified (222). Because lysine residues 4, 9, 27, and 36 within histone H3 have a recognized involvement in the regulation of transcription in other systems, we questioned their role in the regulation of the tissue-specific expression patterns of the two mouse *fpgs* promoters.

a. Histone H3 lysine 4 methylation

Histone H3 lysine 4 methylation (H3K4me) is considered a mark of euchromatin, as such it is detected across genomic regions that either are currently generating transcript or are poised for activation (87, 207). The patterns of H3K4me1, 2, and 3 described in

yeast are shown in Fig 2-6. In yeast, H3K4me1 peaks at the 3'end, H3K4me2 is detected over the coding region, whereas H3K4me3 is restricted to the 5' ends of active genes (186) (Fig 2-6). Genome-wide maps of histone methylation patterns in mouse and human cells showed that all three methylation states correlate with transcriptional levels, but at highly active genes H3K4me3 was greatly increased and mono- and di- forms of this modification were substantially less abundant (15, 20). The link between H3K4me3 and the 5' end of genes has been greatly expanded upon over the past year (see discussion).

As discussed above, the phosphorylation of RNAPII within the CTD recruits the Paf1 complex, which binds directly to at the 5'-end of genes to the histone H3 lysine 4 methyltransferases, such as Set 1. This results in the deposition of methyl groups on lysine 4 of regionally positioned histone H3 (136). Interestingly, the recruitment of PAF is also thought to facilitate ubiquitination of histone H2B lysine 120 (261), a modification that has been shown to be important in the addition of the second and third methyl groups at lysine 4 of histone H3 (54).

The functions of H3K4me3 are still being uncovered but its importance in facilitating transcription is highlighted by the fact that the chromatin-remodeling complex, NURF, has been shown to bind to this PTM through a PHD finger domain (143). Interestingly, evidence suggests that the presence of methylation of histone H3 at lysine 4 may be integral in preventing transcriptional repression of marked promoters from *de novo* DNA methylation. Recently, LSD1, a demethylase specific for lysine 4 and 9 of histone H3, has been shown to be required for the maintenance of global DNA methylation in

mouse development (255).

b. Histone H3 lysine 36 methylation

In yeast and mammals, histone H3 lysine 36 methylation is also abundant across actively transcribed genes. While all three methylation states exist to some degree at active genes, most of the current work has focused on the tri-methylated state at lysine 36 of histone H3 and its role in transcriptional regulation. Genome-wide studies in yeast have determined that this histone PTM is most abundant towards the 3' end of active genes, in striking contrast to the pattern determined for H3K4me3 (Fig 2-6). Studies at highly active mammalian genes have also found H3K36me3 enrichment at the 3' end of these loci (11, 247). The contrast between the distribution of H3K4me3 and H3K36me3 across active genes is extremely interesting. Like the family of histone H3 lysine 4 methyltransferase, the localization of HMTs specific for lysine 36 (i.e. Set 2) are dependent upon the phosphorylation status of the CTD of RNAPII, but unlike H3K4me3, H3K36me3 is coincident with the phosphorylation of serine 2 residues within the CTD of RNAPII (130).

The function of H3K36me3 appears to mediate transcriptional repression in the coding region of active genes. The mechanism of H3K36me3-mediated transcriptional repression was uncovered in two studies that determined that H3K36me3 was bound by a Rpd3, a histone deacetylase complex, via the chromodomain of Eaf3, a subunit of the Rpd3 complex (38, 128). Thus, H3K36me3 appears to establish compact chromatin in the wake of elongating RNAPII, preventing initiation at cryptic promoters that have been

released from bound nucleosomes by the passage traveling RNAPII through the genomic region.

c. Histone H3 K27 methylation

The different methylation states of histone H3 lysine 27 are associated with both euchromatic and heterochromatic states of chromatin. H3K27me1 has been detected in human cells most abundantly at active promoters, whereas H3K27me2 and me3 levels were higher at silent promoters (15). H3K27me3 serves a docking site for the chromodomain of PRC1 (225), which results in polycomb-mediated heterochromatin formation (see above). The role of H3K27me1 at active promoters remains poorly understood.

d. Histone H3 K9 methylation

Histone H3 lysine 9 methylation is a histone PTM that has been detected at both euchromatic and heterochromatic regions of chromatin in higher eukaryotes. Lysine 9 methylation of histone H3 creates a binding site for the heterochromatin protein 1 (HP1) family of molecules. In euchromatin, both H3K9me2 and me3 are often found at promoters that are transcriptional silent and the coding regions of active genes (Fig 2-6). In contrast, H3K9me3 is the mark that is predominately found throughout the entire region of heterochromatin. Su(var)3 was the first HMT identified and it is responsible for the methylation of lysine 9 of histone H3 in *Drosophilia* (192). Several homologues of Su(var)*3* have been shown to methylate lysine 9 in mammals including Suv39h, G9a and Eu-HMTase I. Interestingly, studies in knockout ES cells determined that G9a and Eu-

HMTaseI are critical for *establishing* the patterns of H3K9me2 in euchromatin, whereas disruption of Suv39h results in the loss of H3K9me3 in heterochromatin (53).

A very interesting aspect of H3K9 methylation is the potential connection between this histone PTM and DNA methylation in mammals. Previous work has shown that H3K9 methylation directs DNA methylation in *Neurospora crassa* and *Arabidopsis thaliana*, but the causal relationship between these two epigenetic marks in mammals remains largely unknown (74). Evidence exists to supports each of these epigenetic modifications, i.e. DNA methylation and H3K9 methylation, as the primary mediator of transcriptional silencing (140, 210). For example, the methyl-binding protein -1 (MBD1) is known to interact with the lysine 9 specific HMT, SETDB1, suggesting that DNA methylation may be a prerequisite for H3K9 methylation (210). In contrast, studies in Suv39h -/- embryonic stem cells showed a substantial reduction in H3K9me3 and DNA methylation across silenced genomic regions, supporting the counter argument that H3K9 methylation mediates DNA methylation (140).

C. DNA methylation as an epigenetic mark and regulator of transcription:

DNA methylation occurs at CpG dinucleotides and is an epigenetic modification required by mammals to ensure proper embryonic development and survival (142). The DNA methyl transferases (Dnmts) are the enzymes responsible for establishing and maintaining genomic methylation patterns (131). The *de novo* Dnmts, Dnmt3a and Dnmt3b, modify unmethylated DNA, whereas Dnmt1 acts on a hemi-methylated substrate during replication. DNA methylation is thought to mediate transcriptional repression by

masking DNA binding regions from sequence specific transcription factors or by the recruitment of methyl binding proteins (MBPs) (Figure 2-7) (131). MBPs have been shown to complex with HDACS, chromatin remodeling complexes, and histone H3 lysine 9 methyltransferase to establish a repressive chromatin state (Figure 2-7) (123, 131, 210). Although the interplay of histone PTMs and DNA methylation has been shown to exist, the sequence of events leading to a transcriptionally inactive state remain open and unanswered.

Two types of promoters exist in the mammalian genome when classified by CpG dinucleotide content: regions with densely packed CpGs, called CpG islands, compose 60% of mammalian promoters and are usually found unmethylated (23). The second category includes promoters with a sparse number of CpGs that are often methylated (23). This profile of methylation is established during embryogenesis and is stable in normal somatic cells (193); however, certain tissue-specific genes undergo demethylation in their tissue of expression (32, 75, 227). Whether demethylation at these genes is a primary event or a consequence of transcriptional activation by other mechanisms is unknown.

While a requirement for DNA methylation has been well established in genomic imprinting, X-chromosome inactivation, and in the progression of certain cancers, its participation in the regulation of TSE in normal tissues remains unclear (121). Recently, genome-wide studies have provided evidence supporting a role for DNA methylation in TSE (58, 227). In those studies, a proportion of mammalian promoters were methylated in a tissue-specific manner, and this methylation was correlated with transcriptional silencing

(58, 227). In the experiments discussed in this chapter, we explored the role of DNA methylation in the regulation of the tissue-specific expression of the P1 promoter.

IV. Concepts of Transcriptional Interference

Transcriptional interference (TI) in broad terms describes a situation where transcriptional processes at one genomic region influence the potential for transcriptional activity at a second genomic region, which is usually located immediately upstream or downstream. TI mechanisms thought to be involved in gene regulation include: 1) competition for *cis* elements or *trans* factors required for transcriptional activation, which has been proposed to regulate the β-globin locus during mouse development (70); 2) histone PTMs positioned across a genomic region as a consequence of transcriptional activity at one locus preventing initiation at a second gene, i.e. H3K36me3 deposition in the body of a gene prevents initiation at intragenic promoters (128); and 3) promoter occlusion, where transcription complexes generated at one gene blocks initiation at a distal gene by preventing transcription initiation factors and machinery from properly interacting with the promoter region (28, 62, 83, 107, 155).

With the increasing number of genes found to have multiple promoters, it has become increasingly important to understand if and how TI may influence the conditions under which certain promoters within a genetic locus are used. The two promoters of the mouse *fpgs* gene are arranged in tandem and are regulated such that use of each promoter is controlled in a tissue-specific fashion. We questioned if and how processes of TI dictated promoter choice at the locus in different tissues of the mouse.

V. Objectives of this study

Previous studies in our laboratory and others determined that the mouse *fpgs* gene is expressed in a tissue-specific fashion and possess two promoters that are used under discrete conditions: the upstream P1 promoter is active exclusively in mouse liver and kidney, while the downstream P2 promoter is utilized by all dividing tissues and cells (244). In this study, we investigated how the two mouse *fpgs* promoters are controlled to accomplish this tissue-specific pattern of expression. We identified several mechanisms regulating expression from this locus including epigenetics, transcriptional interference, and factors pausing early stages of elongation. There were three critical pieces of data that lead us to explore the relevance of epigenetics and transcriptional interference in the control of the mouse *fpgs* gene: 1. P1 was determined to be within a tissue-specific differentially methylated region (TDMR), where methylation of P1 inversely correlates with promoter activity (Fig 2-9). 2. DNAse hypersensitivity analysis determined that the chromatin structure surrounding P2 is open and sensitive to DNAse I digestion in all tissues, but P1 is only digested in liver and kidney, when this promoter is active. It appeared that DNA methylation and chromatin state may determine the tissue-specific expression from P1. 3. Real-time analysis of *fpgs* mRNA levels generated from P2 in mouse liver determined that expression from P2 in this tissue is measurable, in contrast to the situation in mouse brain where *fpgs* mRNA levels are undetectable, but transcription is severely restricted when compared to the levels detected in L1210 cells (Table 2-2). These data suggested that the P2 promoter in mouse liver was capable of firing, but was being repressed, perhaps because of happenings at P1 in this tissue. Based on these studies, we

aimed to gain insight into the profiles of histone PTMs and transcriptional complexes across the entire *fpgs* locus in L1210 cells, mouse brain, and liver.

MATERIALS AND METHODS

Materials

C57BL/6 mice were used in all the mouse studies and were originally obtained from Charles River. Chemicals and reagents were purchased from Sigma (St. Louis, MO) and Fischer (Pittsburgh, PA). Antibodies were obtained from Upstate Biotechnology (Billerica, MA), Santa Cruz (Santa Cruz, CA), and Abcam (Cambridge, MA) and will be discussed individually in the sections to follow. Protein G Sepharose (cat #17-0618-01) beads were purchased from Amersham Biosciences (Uppsla, Sweden). 25 nmoles of DNA oligo primers were purchased for these studies from Intergrated DNA Technologies (IDT). Quanti-tect Sybr Green PCR Master Mix (product # 204143) was purchased from Qiagen (Valencia, CA) and used for all of the Real-time PCR analysis. A bath sonicator (Diagenode) was used to shear genomic DNA into 100-300 bp size fragments. The Real-time PCR machines used in these studies were BioRad DNA Engine Peltier Thermal Cyclers with a Chromo 4 Real-Time Detector attachment. Opticon Monitor Software was used to analyze the Real-Time data.

Cell culture

L1210 mouse leukemic suspension cells were maintained at a density between 10^5-10^6 cells/ml in RPMI-1640 medium (Gibco/Invitrogen) supplemented with 10% Fetal Bovine Serum (FBS) and stock cultures were fed every 2 days. Cells were grown at 37^0C in 5% CO_2.

Chromatin Immunopreciptation (ChIP)-histone modifications.

Harvesting and cross-linking mouse brain and liver

C57BL/6 were anesthetized using isoflurane and sacrificed by cervical dislocation. Blood was removed from the liver *in situ* by flushing cold phosphate buffered saline (PBS) into the hepatic vein using a 21-gauge needle and syringe. The organs were harvested and diced using a sterile razor blade and the tissue was weighed. Diced mouse liver or brain (300-500 mg) was placed in 30-40 mls of RPMI-1640 supplemented with 10% FBS and 1% formaldehyde (HCHO). The samples were rocked at 25^0C for 10 minutes (brain) or 15 minutes (liver). Fixation was stopped by the addition of glycine to the media to achieve a final concentration of 0.125M. The samples were rocked at 25^0C for an additional 5 min. Tissue samples were pelleted by centrifugation at 450 x *g* at 4^oC for 10 minutes. Pellets were resuspended in 5 mls of PBS with 1mM phenylmethanesulphonylfluoride (PMSF), swirled, and poured into a Dounce homogenizer. Samples were dounced 20-24x to achieve single cell suspension. PBS with 1mM PMSF (12 mls) was used to rinse out the dounce homogenizer and samples were passed through sterile gauze to remove tissue clumps. The suspensions were spun down at 1500 RPM at 4^oC for 10 minutes.

Cross-linking L1210 cells

L1210 cells were suspended at a density of 5 x 10^5/ml with 1% HCHO in RPMI-1640 medium in a T175 flask. The cell suspension was gently stirred with a small stir bar at 25^0C for 8 minutes, to ensure that the cells did not settle to the bottom of the flask during cross-linking. Glycine was then added to the media to achieve a final concentration of 0.125 M and the cells were mixed for an additional 5 minutes. Ten million cells were pelleted by centrifugation at 1500 RPM at 4^oC for 10 minutes, resuspended in 5 mls of PBS + 1mM PMSF, and again pelleted.

Cross-linked cells and mouse tissues were washed in buffer I (0.25% Triton X-100, 10mM EDTA pH 8.0, 0.5mM EGTA pH 7.5, 10mM Hepes, pH = 7.5), and then in buffer II (0.2M NaCl, 1mM EDTA pH 8.0, 0.5mM EGTA pH 7.5, 10mM Hepes, pH 7.5) All ChIP buffers contained 1 µg/ml pepstatin, 1 µg/ml leupeptin, 1 µg/ml aprotinin, and 1 mM fresh PMSF. Stock concentrations of 0.5M Tris, pH 7.5 and 8.0 were prepared by adding the recommended amounts of Trizma-HCL and Trizma-base using the Trizma chart to achieve the appropriate pH at 25^0C. The pH of the solution was then measured using at 25^0C using a Mettler Toledo pH meter. Hepes, buffer pH 7.5 (1M) was prepared and pH was adjusted using 8 N NaOH. Cells were pelleted and frozen at -80^0C or processed as described below.

Protein G Sepharose Bead preparation and Immunoprecipitation

The 300 mg of liver, 100 mg of brain, or 2.5 x10^6 L1210 cells were resuspended in 300 µL samples of lysis buffer 25 mM Tris buffer, pH 7.5 (containing 150 mM NaCl, 5 mM EDTA, 1%Triton X-100, 0.1 % SDS, and 0.5 % sodium deoxycholate), and sonication

was carried out in a bath sonicator (Diagenode) to achieve 100-300 bp DNA Fragmentation of the DNA was followed by reverse cross-linking and purifying DNA (for details see below) from 20 µl aliquots of sample followed by visualization of fragment sizes on a 1% agarose gel stained with Ethidium Bromide (Fig 2-8) . Sonication was performed with repeated cycles of 30 sec pulse and 30 sec off, followed by the addition of fresh cold ice water to the bath. The total sonication time for each sample was 25 minutes (liver and brain) and 20 minutes (L1210 cells). Following sonication, the samples were spun down at 21,000 x g, at 4^{o} C for 10 minutes. The supernatant was removed and all the samples were pooled for each tissue. The samples were then diluted to achieve the final tissue ratios used in the immunoprecipitation: 50 mg (liver), 30-40 mg (brain), and 1-2.0x10^{6} L1210 cells in 300 µl of lysis buffer. Three hundred microliters of sample lysates were aliquoted into 1.5 ml centrifuge tube. Thirty microliters of Protein G-Sepharose beads/sample were placed in a 1.5 ml centrifuge tube. The beads were aliquoted using a 20-200 µl range pipet tip with the end cut off. This assisted in pipetting the beads. One hundred microliters of beads were washed with 500 µl of freshly made lysis buffer and spun down at 240 x g at 4^{0}C for 1 min. This was repeated three times. One hundred microliters of lysis buffer was then added to the beads to obtain a 50% lysis buffer/Protein G Sepharose bead mix (final volume = 200 µl). Thirty µl of 50% slurry will be used for each reaction. Prior to use, the beads are blocked to help minimize non-specific IgG binding. To block 200 µl of a 50% slurry mix, 0.33 mg/ml of BSA and sonicated lambda are added to the 50% slurry (0.05 mg/ml per 30 µl or per reaction). Three hundred microliters of lambda were sonicated for 5 min set to 30 sec pulse and 30 sec off. The

50% slurry with BSA and sonicated lambda were rotated for 2-3 hours end-over-end at 4^0C. The beads were spun down at 240 x g at 4^0C for 1 min. The supernatant was removed and the beads were washed with 500 µl of lysis buffer and spun down at 240 *x g* at 4^0C for 1 min. This was repeated three times. The beads were then resuspended in 100 µl of lysis buffer to achieve a 50% slurry mix. The samples were incubated with 30 µl of 50% slurry for 1 hr at 4°C. This step is referred to as preclearing. Note: It is important to gently mix the 50% slurry with a pipet tip before pipetting. Following incubation, the beads were spun down at 240 x g at 4^0C for 1 min. The clarified lysates were placed in a fresh tube and the beads discarded. The lysates were rotated at 4°C overnight with 5 µg of antibodies against the following: Total Histone H3 (Abcam, ab1791), mono-methyl-H3K27 (Upstate, 07448), di-methyl-H3K27 (Upstate, 07452), tri-methyl-H3K27 (Upstate, 07449) tri-methyl-H3K9 (Upstate, 07442), tri-methyl H3K4 (Upstate, 07473), acetyl-H3 (Upsate, 06599), acetyl-H4 (Upstate, 06598), acetyl-H3K9 (Upstate, 07352), and IgG (Upstate Biotech, 12-371). Ten µg of antibody against tri-methyl H3K36 (Abcam, ab9050) was also used. Thirty microliters of 50% slurry of lysis buffer/Protein G-Sepharose beads were added to the antibody-lysate mix and rotated at 4°C for 1 hour. The beads were spun down (as above), the supernatant was removed and discarded. The supernatant for the sample with IgG antibody was placed in a fresh 1.5 ml centrifuge tube so that it could be used later as input DNA. The beads were washed two times with RIPA buffer (150mM NaCl, 50 MM Tris, pH 8, 0.1% SDS, 0.5% NaDoc, 1.0% NP-40), and once with high salt buffer (500 mM NaCl, 50 mM Tris, pH 8, 0.1% SDS, 1.0% NP-40), then LiCl buffer (250 mM LiCl, 50 mM Tris pH 8.0, 0.5% Na deoxycholate, 1.0% NP-40), and two times with

Tris EDTA, pH 8.0 (10mM Tris, 1mM EDTA, pH 8.0) (TE). For these washes, 500 µl of buffer was added to the beads and the suspension was rotated end-over-end for 10 minutes at 4°C in 1.5mL centrifuge tubes. Following these washes, 200 µl of elution buffer (2% SDS, 10 mM DTT, 0.1M $NaHCO_3$) was added to the beads and the samples were rotated for 15 minutes at 25^0C. This step caused the release of proteins and protein/DNA complexes bound to the antibodies by denaturing the antibodies and breaking the disulphide linkages between the large and small IgG subunits. The elution was performed twice. Concentrated (4M) NaCl was added to the eluted materials to a final concentration of 0.2M and they were placed at 65 ºC overnight to reverse the formaldehyde induced cross-links. The saved input material was also placed at 65 ºC.

DNA purification

Following these incubations, 1 ml of 100% ethanol was added to the 400 µl of eluted material and the mixtures were held at -80^0C for 30 minutes to precipitate DNA. Samples were spun down, washed with 70% ethanol, and resuspended in 180 µl of TE. The samples were incubated with 1 µl of 10 mg/ml of RNAse A, and incubated at 37^0C for 30 min. Twenty microliters of 5x Proteinase K digestion buffer (50 mM Tris, 25 mM EDTA, pH 8.0, 1.25% SDS) and 1 µl of 20 mg/ml Proteinase K to the samples and they were incubated at 42^0C for 1 hour. Incubation was followed by phenol chloroform extraction by adding equal volume of phenol:chloroform (1:1), pH 6.8. The samples were vortexed and spun at 16,000 x *g* for 5 min at 25^0C. The aqueous phase was recovered and precipitated in 2.5 x the volume of ice-cold 100% ethanol and $1/10^{th}$ the volume of 3M Na acetate, pH 5.2, and 1 µl of glycogen (20 µg) as a DNA carrier. The samples were held at -

80^0C for 30 min or overnight. The precipitated DNA was pelleted by centrifugation at 20,000 x *g* for 10 minutes at 4^0C. The samples were then washed with 1 ml of ice-cold 70% ethanol and again pelleted. The ethanol was decanted and the pellet was allowed to air dry for 5 min on the bench top. Samples were then resuspended in a final volume of 100 µl of TE.

The input material not precipitated before treatment was treatment with RNase A and Proteinase K. The steps following Proteinase K digestion were the same for the input material as were described for the samples. Real-time PCR was used to quantitate the content of DNA fragments using a 1 µl aliquot of each sample (see below for details). Real-time PCR reactions PCR primer pairs amplified regions of genomic DNA every 1.5-2.0 kb across the length of the *fpgs* gene (Fig 2-11). Prior to amplification of the immunoprecipitated material with primers specific for the mouse *fpgs* gene, real-time PCR was prepared using sample DNA and primers specific for regions in the *gapdh* and *globin* promoters and coding regions. These data served as controls for each immunoprecipitation with a given antibody. The GAPDH and Globin data is not presented in this chapter.

Chromatin immunoprecipitation-RNAPII and transcription factor studies.

The ChIP protocol described above was applied to study the distribution of RNAPII bound over smaller overlapping segments of the *fpgs* promoters and the distribution of transcription factors near P2. Cross-linked DNA fragments were sonicated down to 100-300 bp, as above. Lysates were rotated at 4ºC overnight with 8 µg of anti-RNAPII (Upstate Biotech, 05-623), 4 µg anti-IgG (Upstate Biotech, 12-371), or 5 µg of

antibodies against phosphoserine 5 peptide from RNAPII (Abcam, ab5131), phosphoserine 2 peptide from RNAPII (Abcam, ab5095), TBP (Santa Cruz, sc-273), TAFIIp250 (Santa Cruz, sc-17134), TFIIB (Santa Cruz, sc-225), NELF-A (Santa Cruz, sc-23599), or SP1 (Santa Cruz, sc-59). Real-time PCR was performed with 1 µl of final sample dissolved in 100 µl of TE. The primer sets used covered the entire upstream and downstream promoters of the *fpgs* gene. The P1 *fpgs* promoter region was divided into five segments, using primers that amplified nucleotides -379 to -119, -126 to 30, 7 to 153, 121 to 288, and 287 to 449 relative to the upstream transcriptional start site. The P2 *fpgs* promoter was divided into six fragments amplifying nucleotides -698 to -469, -475 to -333, -340 to -167, -194 to -52, -83 to relative to the transcriptional start site (Fig 2-, Appendix). In the studies using antibodies against general transcription factors the primers sets used were specific to P2; P1 was not amplified in these studies. The primer sets used across P2 for the general transcription factor studies amplified fragments -340 to -167, -194 to -52, -83 to relative to the transcriptional start site (Fig 2-18). The GAPDH and Globin primer sets discussed above were also used in these experiments as controls (data not shown).

Peptide competition study using total RNAPII, phospho-Serine 5, and phospho-Serine 2 RNAPII specific antibodies.

The specificities of the antibodies against RNAPII phosphorylated serine-5 CTD peptide and phosphorylated serine-2 CTD peptide were determined by incubating L1210 cell lysates and each antibody with increasing amounts of the serine 2 phosphopeptide (Abcam, ab 12793), serine 5 phosphopeptide (Abcam, 18488) and unphosphorylated CTD

peptide (Abcam, 12795) (See Table 2-1). Five µg of each antibody was rotated with either 0.1, or 1, or 10 µg of peptide in 100 µl of lysis buffer for three to four hours at 4^0C. In these studies, sonication of 2-3 million L1210 cells was carried out in 200 µl. The beads were prepared as discussed in the previous section. The lysates were prepared as discussed in the previous section. Following the pre-clearing step discussed above, the 100 µl antibody-peptide mix was added directly to the 200 µl of L1210 samples. The rest of the ChIP protocol was carried out as described above. Peptide inhibition for α-phospho-Serine 5 was determined using primers amplifying a region of the P2 *fpgs* promoter, and specificity for α-phospho-Serine 2 was tested across a fragment of the *gapdh* coding sequence.

Primer Design

Primers were designed with the assistance of the oligos deoxynucleotides analyzing tools on the IDT website (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/). The primers selected had a melting temperature ranging from 56^oC - 60^oC, a length between 18-24 bp, and a GC content between 50-60%. Primers that generated hairpins, homo-dimers, or hetero-dimers between sense and antisense primers with a delta G value larger than -6.0 kcal/mol were not selected. The delta G value was determined using the oligo analyzer (see above) on the IDT website. The mouse genome was searched for each primer sequence in order to identify primers that included sequences with high similarity to other genomic regions using the University of California Santa Cruz Genome Browser. If a high degree of similarity was detected primers were redesigned. Twenty-five nmoles of

each primer purified by standard desalting were ordered through the IDT website. The lyphollized primers were resuspended in High Performance Liquid Chromatography (HPLC) grade water (Fischer) to achieve a stock concentration of 150 μM. Primers were further diluted to 15 μM in HPLC grade water before they were added to any Polymerase Chain Reaction (PCR) mix. Primers designed following the parameters discussed above, usually allowed correct product formation in end-point or semi-quantitative PCR. However, even with the most stringent design it was sometimes the case that primer sets were not adaptable to Real-time PCR. In these cases, low-levels of non-specific products were formed in the PCR that were not detectable by visualization on agarose gel, but were detected as fluorescence in the Real-Time PCR machine. When multiple products are detected in a Real-Time PCR assay, the data is not interpretable. As a result, it proved useful to run test Real-Time PCR reactions prior to using a primer set in large-scale. For these preliminary experiments, 0.8 ng and 200 ng of genomic DNA and water served as the templates, and the reactions were performed using the real-time protocol described below. In this experiment, a single peak in the melting curve would indicate that the primer pair was suitable for real-time applications. Ten microliters of each reaction was also visualized on an agarose gel in order to identify primer dimers or non-specific products. The 0.8 ng and 200 ng of genomic DNA and water were used, because occasionally, they gave different results. There were several times when primer pairs would work well at 200 ng, but result in multiple products, or primer dimers, at low concentrations of template DNA or in the reactions where only water was added.

Semi-quantitative PCR

Semi-quantitative PCR was used in the initial studies examining RNAPII residence across P1 and P2 in L1210 cells and mouse liver. The HotStar Taq Plus Master Mix (Qiagen, Product # 203643) was used for all Semi-Quantitative experiments. A master mix was prepared using the HotStar Taq Mix, HPLC water, and 0.3 µM of sense and antisense primers. Twenty four µl of master mix was distributed into PCR strip tubes (Biorad). One µl of sample DNA or HPLC water or a fraction of input DNA (0.1%) was added to triplicate tubes. The strips were placed in the 96-well plate of the BioRad PCR machine and PCR was performed with the amplification conditions 95^{o}C, 45 sec; 57^{o}C, 45 sec; 72 ^{o}C, 45 sec, with a final extension at 72 ^{o}C for 5 minutes. All samples were first incubated at 95 ^{o}C for 15 minutes to activate the complex Taq polymerase. The number of PCR cycles used varied from 26 to 36. Initially, samples were removed at the cycles 26, 28, 30, and 32, following incubation at 72^{o}C, and transferred to a second 96-well format hot block set at 72^{o}C, to allow for the final extension incubation of 5 minutes for each sample. Ten microliters of each reaction mix were applied to a 1% agarose gel and separated DNA fragments were visualized following staining with ethidium bromide. The cycles selected varied and were chosen to capture the amplification reaction in the exponential range.

Real –time PCR

Real-time PCR was performed using Quantitect Sybr Green PCR Master Mix (Qiagen) and BioRad DNA Engine Peltier Thermal Cyclers with Chromo 4 Real-Time

Detector attachments. A master mix was prepared using the Sybr green mix, HPLC water, and 0.3 μM of each sense and antisense primers. Seventy-two microliters of the master mix was aliquoted into 0.6 ml PCR tubes. Three microliters of standard, sample, or input DNA or HPLC water were added to the appropriate PCR tubes (sub-master mix). As above, a fraction of the input DNA (0.1%) was used. The sub-master mix was vortexed and collected by brief centrifugation, and 25 μl of the mix was transferred to the Biorad Low-White Tube PCR strips in triplicate. Four standard reactions were prepared using 5-fold serial dilution of sonicated mouse genomic DNA (100 ng/μl, 20 ng/μl, 4 ng/μl, 0.8 ng/μl) (details on standard curve below). The reactions were pre-incubated at 95^{o}C for 15 min and the amplification condition for each primer pair were 95^{o}C for 45s; 58 ^{o}C for 45s, 72 ^{o}C for 45s followed by a final extension time at 72 ^{o}C for 5 min. Forty PCR cycles were performed for the real-time studies and a plate read step was programmed at the end of each cycle to capture the fluorescence in each tube. Melt curve analysis was programmed at the end of the run to measure fluorescence of each reaction in one degree increments between the temperatures of 50 ^{o}C- 100 ^{o}C. The melt curve analysis is an essential step in clarifying whether one specific product was amplified in the reaction (see below).

Analysis of Real-time PCR data

The standard curve is an essential component to analyzing real-time PCR data when using multiple primer sets. The standard curve is generated by plotting the log of the dilution factor against the Ct value (see below) obtained during amplification of each reaction. The equation of the linear regression line and the coefficient of determination

(R^2) were determined using the Opticon software and a R^2 value greater than 0.99 indicated that primer pairs were amplifying optimally. The data obtained from the samples will be assigned a value based on the Ct (see below) values determined for the standard reactions.

The threshold cycle number (C_T), or the cycle number at which amplified product accumulates to a level where fluorescent signal is detected, was used to quantify the amount of DNA in each reaction. Measured fluorescence for sample and input reactions were converted into values relative to mouse genomic DNA by comparing the C_T of the samples to the C_T values determined for the standard reactions. The values for the triplicate samples were averaged and divided by the averaged input values.

Melt curve analysis was used to identify if multiple products were formed during the different reactions. In this analysis, fluorescence in a tube is measured after the temperature is raised in 1^0C increments. As the temperature increased, double-stranded DNA products melt and fluorescence detected decreases. The negative first derivative was plotted as a function of temperature using the Opticon Monitor Software provided with the BioRad machines. Different products will melt at different temperatures and in this case the melt curve profile will have multiple peaks. A single peak indicated a specific product and data with multiple peaks was discarded as representing multiple products. Standard deviations were calculated using the formula: sample/input X (sqrt ((delta input/input)^2 + (delta sample/sample)^2)).

RESULTS

DNA cytosine methylation near the P1 and P2 promoters.

A previous student in this laboratory, Dr. Fiona Turner, initialized the studies on the epigenetic control mechanisms operative at control of the two mouse *fpgs* promoters by by performing bisulfite sequencing (Fig. 2-9) and DNAse hypersensitivity (Fig. 2-10) assays using genomic DNA surrounding the two *fpgs* promoters. The distribution of CpG dinucleotides in the regions of the two *fpgs* promoters differed significantly. The P1 promoter is in a region in which CpG dinucleotides are sparsely represented (19 dinucleotides distributed over the region from -650 to +900 nt, relative to the first transcriptional start site), whereas the P2 promoter is embedded in a CpG island (%GC=58.9, ObsCpG/ExpCpG= .71, CpGs distributed over the 1000 nt centered on the first transcriptional start site), as defined in (232).

We were very surprised by the results of the bisulfite sequencing. The methylation patterns defined over the two *fpgs* promoters were strikingly different (Figure 2-9). DNA from mouse liver showed an undermethylated state over P1, whereas L1210, bone marrow, and brain DNA had virtually quantitative methylation at each CpG over this same region. In striking contrast, the CpGs surrounding P2 were completely unmethylated in all tissues studied, including normal and neoplastic tissues in which the P2 promoter was transcriptionally active (L1210 and marrow). Liver, in which transcripts initiating at P2 were barely detectable, and brain, in which the entire *fpgs* gene was transcriptionally inactive (Figure 2-9) (Table 2-2). These data suggested that the two *fpgs* promoters represented two interesting cases, one where hypomethylation of a CpG-sparse promoter

predicted transcriptional activity, and the second a tissue-specific CpG-island promoter that was not methylated in any tissue, even those in which it was not expressed (Figure 2-9).

Mapping of chromatin structure near the two *fpgs* promoters.

The chromatin state of a region surrounding a promoter influences the availability of DNA *cis* elements to cellular *trans* factors, and when nucleosomes are tightly packed these interactions are restricted and transcriptional activation prevented. Dr. Turner asked if the chromatin state of P1 and/or P2 of the *fpgs* gene determined the tissue-specific activity of the two promoters. DNase I hypersensitivity assays were used to determine the accessibility of these regions to proteins by determining the cleavage patterns in genomic DNA after exposure of intact nuclei to DNase I (Figure 2-10). At low levels of DNase I digestion, a full-length 8.5 kb *Hin*DIII fragment hybridized to the probe for the P1 region but, with increasing amounts of DNase I, a smaller fragment (4.4 kb) was detected in DNA extracted from renal and hepatic nuclei (Fig. 2-10, left panels). This band was not seen in DNA extracted from mouse L1210 leukemia cells and mouse bone marrow, two dividing cell types that do not utilize the P1 promoter, nor was it observed in DNA from brain nuclei, which do not transcribe the *fpgs* gene from either promoter. When the P2 region was examined using a probe from the downstream *Hin*DIII fragment, a DNase hypersensitivity site was detected in DNA extracted from nuclei from all the tissues examined, i.e., L1210 cells, bone marrow, liver, kidney, and brain (Fig. 2-10, right panels). The DNase hypersensitivity sites found for each promoter mapped to positions about 130-

150 bp upstream of the transcriptional start sites defined as diagrammed in Fig. 2-10. This implied that the chromatin configuration centering on 130-150 bp upstream of P1 allowed access to this region only in those tissues that utilize this promoter (liver and kidney), but that the chromatin encompassing the P2 promoter was accessible in all tissues studied, whether or not this promoter was transcriptionally active.

Taken together, the DNA methylation and DNAse hypsersensitivity data indicated that epigenetic mechanisms working upstream of P1 may be causally involved in determining the activity of this promoter, but neither DNA methylation nor gross chromatin structure predicted the activity of P2 in the different tissues. We sought a more complete understanding of the epigenetic controls working at the *fpgs* gene suggested in these original experiments. We focused our efforts on profiling several histone PTMs across the entire 20 kb *fpgs* locus in tissues using P1 (mouse liver), using P2 (L1210 cells), and in a tissue where the gene is completely silenced (mouse brain). The questions we aimed to address in these studies included: 1. Are the patterns of histone PTMs across P1 different between tissues and is the repressive nature of DNA methylation reinforced by changes in these modifications? 2. Do the histone PTMs of nucleosomes spanning P2 in mouse liver explain why P2 expression is restricted in this tissue? 3. In brain, is the *fpgs* locus marked with histone PTMs commonly associated with a silenced/heterochromatic state? and 4. Are the patterns of histone PTMs across the dual promoter *fpgs* locus similar to those previously described for a highly expressed single promoter gene by Vakoc et. al (247)?

The histone marks studied were selected based on evidence that defined role(s) for these PTMS in transcriptional activation or silencing in yeast or mammalian cells or on a relationship with DNA methylation or proteins mediating DNA methylation. We chose to probe for antibodies against histone H3 acetylation (H3Ac), histone H4 acetylation (H4Ac), histone H3 lysine 36 tri-methylation (H3K36me3), histone H3 lysine 27 mono-, di-, and tri-methylation (H3K27me1, me2, and me3), histone H3 lysine 9 di- and tri methylation (H3K9me2 and me3), histone H3 lysine 4 di- and tri- methylation (H3K4me2 and me3), and total histone H3. Composite profiles of these histone PTMs obtained from work performed at single genes in yeast are reflected in Fig 2-6. We used these antibodies in our ChIP protocol for the individual tissues followed by Q-PCR to generate tissue-specific histone PTM profiles across the 20 kb *fpgs* gene. Prior to performing Q-PCR using the *fpgs* gene specific primers the enrichment of these antibodies in all three tissues where tested using primers in the GAPDH promoter and coding region and the globin promoter and coding region. Samples immunoprecipitated with an antibody raised against IgG were also analyzed as a non-specific background control in all of our experiments. Previous to our studies, Vakoc et. al mapped several histone PTMs (H3K4me3, H3K9me3, H3K27me1, 2, and 3, and H3K36me3) across the very highly expressed PABPC1 gene and these studies demonstrated the power of following histone marks along the body of a gene (247). We used this literature precedent as a guide to select regions of genomic DNA of the *fpgs* gene for amplification by Q-PCR. Primer pairs were designed to amplify regions every 1.5-2.0 kb down the entire length of the gene. The placement of these primer pairs is depicted in Figure 2-11.

ChIP walking across the *fpgs* gene describes the level of histone H3 in multiple mouse tissues.

In our initial studies we examined the density of nucleosomes over the length of the 20 kb *fpgs* locus in L1210 cells, mouse brain, and mouse liver using an antibody to total histone H3 levels (Figure 2-10). The data obtained from these experiments were critical to the interpretation of the histone PTM maps generated in the studies to follow (Figure 2-10). In order to understand the meaning of changes in levels of histone PTMs, we needed to ensure that observed shifts were not simply the result of changes in total nucleosome density. In theory, ChIP data for modified histone residues should be expressed as a ratio with total histone H3 to allow an estimation of the degree of modification per molecule of histone H3. For the most part, the data obtained in the histone PTM experiments were expressed as primary ChIP signals, to allow the reader to come closer to seeing the data, rather than heavily manipulated ratios. In a few cases, we used a ratio of histone PTM ChIP signal to total histone H3 signal to account for histone density changes (Fig 2-17).

In our studies, the density of histone H3 reached a local minimum over the P2 promoter in L1210 cells, mouse liver, and brain and co-localized with the DNase hypersensitivity site mapped to this P2 region in these tissues (Figure 2-10). Somewhat surprisingly, a similar histone H3 local minimum was not seen in liver over the P1 region, in spite of the DNase hypersensitivity site over that region (Figure 2-10). There was an observable increase in total H3 ChIP signal detected in L1210 cells and brain of the genomic regions starting 5 kb upstream to 12 kb downstream of the P2 promoter. (Figure 2-10) This pattern was not present in mouse liver. These data indicate that the position of

nucleosomes were altered by high levels of transcriptional elongation across the genomic regions surrounding the P2 promoter in mouse liver.

Histone acetylation concurs with transcriptional initiation over P1, but not P2.

Previous studies in yeast and human cells indicate that active promoters have substantial acetylation of lysines in the N-terminal peptides of histones H3 and H4 (144, 186). We probed the histones decorating the entire length of the *fpgs* gene, including segments containing P1 and P2 promoters, by ChIP using antibodies against acetylated histone H3 at lysines 9 and 14 and histone H4 at lysine 16 (Figure 2-12). As expected, abundant acetylation of both histone H3 and histone H4 was detected in mouse liver over P1 (Figure 2-12), high levels of acetylation started 1 kb upstream of the promoter and extended 3 kb into the *fpgs* gene. In L1210 cells, the nucleosomes over P1 were not enriched in acetylated histone H3 or histone H4 (Figure 2-12), but the levels increased substantially starting 2 kb upstream of P2 and spanned 2 to 3 kb for either histone H3 or histone H4 acetylation, respectively (Figure 2-12). The large regions of DNA over which histone H3 and histone H4 were acetylated were very surprising and were only captured because we followed the histone PTMs across several DNA segments that sampled the length of the entire gene. Interestingly, histone H3 and histone H4 acetylation was also found over P2 in mouse liver (Figure 2-12), although the span of genomic DNA over which histone acetylation was found over P2 in liver was much narrower than seen in L1210 cells (> 1.5 kb) (Figure 2-12). This difference was repeatable and was seen with an antibody to H3K9Ac. In mouse brain, a tissue in which both P1 and P2 are

transcriptionally inactive, the pattern of H3 acetylation was almost identical to that detected in L1210 cells (Figure 2-12), with a substantial level of histone H3 detected 2 kb upstream to 1 kb downstream of P2 (Figure 2-12). Interestingly, histone H4 acetylation detected in mouse brain was at undetectable levels upstream of P2 (Figure 2-12), unlike the levels detected in both liver and L1210 cells, and the span of nucleosomes acetylated at the region surrounding P2 was narrower than the signal measured over P2 in L1210 cells (Figure 2-12). These data suggest that deposition of histone H4 acetylation may be more closely linked to the transcriptional activity of a particular locus than histone H3 acetylation. Since histone acetylation is thought to mark an open promoter, finding H3 and H4 acetylation over P2 in mouse liver and brain raises the question of why P2 is not actively firing in these tissues.

The similarities detected between L1210 cells and mouse brain in the data on histone acetylation profiles in Figure 2-12 and the data to follow were, at the time, very surprising to us. We thought that the brain experiments would be an excellent control for our studies, which were originally meant to compare liver in which P1 is active and L1210 in which P2 is active. Our observations were put into perspective by studies in a hallmark genome-wide paper from the laboratory of Dr. Richard Young (87), which was published as we were preparing this study for resubmission to Molecular and Cellular Biology. Young's work concluded that, contrary to common belief, the promoter regions of most of the protein-coding genes in human embryonic stem cells were bound by proteins indicative of the process of transcriptional initiation (87). They also extended these observations to a

subset of differentiated cells. Since then other genome-wide studies have confirmed and extended these conclusions as will be presented in the discussion.

Histone H3 lysine 4 methylation over P1 and P2.

In an effort to understand the differential control of P1 and P2 in different tissues, we compared the patterns of histone H3 methylation over the length of the *fpgs* gene in liver and leukemic cells. These results were particularly meaningful when related to the pattern previously seen at a very highly expressed single promoter gene (247). Di- and tri-methylation of histone H3 at lysine 4 have been found to be marks of euchromatin and active transcription in yeast (186) and human cells (144). The distrubution of di- and tri-methylation on nucleosomes covering an active gene have been shown to be quite different, as shown in Figure 2-6. Studies performed in yeast at single genes have determined that di-methyl lysine 4 of histone H3 is present throughout euchromatic regions of chromatin, whereas tri-methylation of this residue is restricted to the 5' end of genes initiating transcription (186, 207). We examined the patterns of both di- and tri-methyl lysine 4 of histone H3 across the *fpgs* locus under different circumstances. In L1210 cells, the peak of di-methylation of histone H3 at lysine 4 encompassed the P2 promoter and spanned *ca* 4 kb of chromatin region (Figure 2-13). H3K4me2 was detected throughout the entire *fpgs* locus in mouse liver, with a peak spanning the 2 to 3 kb of chromatin downstream of P1, and a second narrower peak 1 kb downstream of P2 (Figure 2-13). Interestingly, these results paralleled almost exactly the patterns established using antibodies against acetylated histone H3 (Figure 2-13) in L1210 cells and mouse liver. In

striking contrast, trimethylation of histone H3 at lysine 4 formed a distinct peak over P1 in liver and P2 in L1210 cells; these peaks continued for 3-4 kb after the transcriptional initiation site (Figure 2-14). Surprisingly, a second smaller peak of H3K4me3 was found over P2 in mouse liver (Figure 2-14). We wondered if this low-level of signal reflected the processes producing the minimal amount of transcript generated from P2 in mouse liver (Table 2-2). Rather surprisingly, the pattern detected in brain was exactly the same as that seen in L1210 cells, in spite of the fact that transcription at P2 was high in L1210 and undetectable in mouse brain (Table 2-2). Clearly, the reasons for this histone PTM across P2 in either mouse liver or brain were more complicated than we originally thought. We concluded from this set of studies that the H3K4me3 mark over the P1 promoter exactly concurred with transcriptional activity, but there was no such relationship found over the CpG-island embedded P2.

One major limitation of the ChIP data obtained using mouse liver, brain, and L1210 cells was that the enrichment signals between tissues were not quantitatively comparable. The variations in tissue-type, treatment of starting material, and the efficiency of cross-linking procedures in different tissues prevented quantitative analysis across tissues. This limited the conclusions we were able to make from these data in a few cases and made interpretation challenging across samples. For example, the small peak of H3K4me3 detected over P2 in liver was similar in breadth to the peaks detected in L1210 cells and mouse brain (Figure 2-14). However, based on our experience with the antibodies generated against histone PTMs, the signal detected over P2 in mouse liver was unusually low and most likely substantially less than the levels detected in brain.

Histone H3 lysine 9 and lysine 27 methylation over P1 and P2.

Methylation of H3K9 and H3K27 has been linked to the formation of heterochromatic regions resulting in the repression of transcription in yeast and mammalian cells (225). More recently, H3K27me1 and H3K9me3 have been found in the coding regions of active mammalian genes (247). We examined the modifications of these lysines within histone H3 at the *fpgs* gene in order to determine if they were involved in the silencing of P2 in mouse liver and/or either promoter in mouse brain. We were particularly interested in the methylation status at histone H3K9 because of the well-studied relationship between DNA CpG methyl-binding proteins (i.e. MeCp2) and histone methyltransferases responsible for the histone H3K9me PTM (i.e. SETB1) (Figure 2-15). We expected to detect H3K9me2 and/or me3 across regions surrounding P1 in tissues where the surrounding DNA was found methylated, but not in mouse liver where lysine 9 of histone H3 was acetylated (Figure 2-15). In mouse liver, H3K9me3 was not detected anywhere along the *fpgs* gene (Figure 2-15). Surprisingly, H3K9me3 was also low across *fpgs* in L1210 cells. Dr. Turner had previously probed the P1 and P2 promoters in L1210 cells and mouse liver using an antibody against H3K9me2 in a ChIP assay. In those studies, H3K9me2 was not present at either promoter in liver or L1210 cells, but was found over the silenced globin promoter that was used as a positive control. We are still puzzled by the fact that H3K9 methylation was not detected over P1 in L1210 cells. It is interesting to note that few, if any, studies have examined the coexistence of H3K9me and DNA methylation at promoters that are as CpG sparse as P1, and it may be the case that this relationship only exists across regions of DNA that are CpG dense.

Mono-methylation of H3 at K27 was present across the *fpgs* gene in both liver and L1210 cells (Figure 2-16), in agreement with recent studies that indicate H3K27me1 as a marker of actively expressing genes in euchromatin (247). However, this mark was also abundant over the *fpgs* locus in brain (Figure 2-16). Within the accuracy of the DNAse hypersensitivity mapping estimates, the minimum of H3K27me1 found at P2 in all three tissues (L1210, liver, and brain) co-located with the DNase I hypersensitivity site in this region (Figure 2-16). Similar profiles of H3K27me3 were detected throughout the *fpgs* gene in all three tissues studied (Figure 2-16) and the levels measured were relatively constant across the body of the gene, particularly when the changes between fragments were compared to the major shifts observed in profiles of H3K4me3 (Figure 2-14) and H3K27me1 (Figure 2-16). We took this to suggest that trimethylation of lysine 27 of histone H3 was not causing the silencing of P2 in mouse liver or the complete silencing of the *fpgs* gene in mouse brain, in spite of the fact that several examples have been described where H3K27me3 causally silences mammalian genes through the recruitment of Polycomb complexes which results in the formation of heterochromatin (225). Interestingly, H3K27me2 reached a discernible peak between the two promoters in L1210 cells (Figure 2-16), with a maximum at 7 kb. The loss of this ChIP signal in L1210 after P2 may indicate either the absence of H3K27me2 at nucleosomes decorating the active downstream region in leukemic cells or it might indicate the presence of a second PTM across this region that masks the epitope recognized by the H3K27me2 antibody, e.g., phosphorylated H3 S28.

Histone H3 K36 methylation across the mouse *fpgs* gene

Methylation of histone H3 at K36 in yeast has been linked to regions of transcriptional elongation and was shown to participate in the recruitment of the histone deacetylase (HDAC), Rpd3, downstream of an active promoter in several cases (38, 128, 130). The recruitment of HDACs to H3K36me3 limited histone acetylation within transcribed regions, and yeast strains mutant in either Set2, the histone methyltransferase responsible for methylation of lysine 36, or in subunits of Rpd3, exhibited increased acetylation in open reading frames and an increase in aberrant transcripts generated from cryptic promoters within these regions (38, 128, 130). Hence, we thought H3K36me3 would be a probable candidate for the suppression of activation of P2 in mouse liver. We mapped the abundance of H3K36me3 across the *fpgs* gene to determine whether the spatial distribution of this PTM down the *fpgs* locus would suggest such a causal role (Figure 2-17). This did not seem to be the case, since H3K36me3 was present throughout the body of the *fpgs* gene, except for a distinct minimum at P2 in both liver and L1210 cells (Figure 2-17). The level of H3K36me3 near P2 was lower than could be explained by the decreased level of total H3 over this region, as best evidenced by the ratio of ChIP signal for H3K36me3 to total H3 (Figure 2-17). The density of H3K36me3 then built up towards the 3'-end of the *fpgs* gene in both tissues, similar to the higher distribution of this mark towards the 3'-end of other yeast and mammalian genes (186, 247). Interestingly, in both mouse liver and L1210 cells, the regions where H3K36me3 was found to be maximal were the regions in L1210 cells and mouse liver that histone acetylation reached a minimum. H3K36me3 was quite low across the body of the *fpgs* gene in brain from +3 to +15 kb,

consistent with the transcriptional inactivity of *fpgs* in this tissue (Figure 2-17). However, a marked increase in this modification occurred at the very 3' end of the *fpgs* locus in brain (Figure 2-17). We considered two possible explanations for this observation: 1. H3K36me3 detected in brain at the *fpgs* locus reflects the level of this PTM that accumulated during development of mouse brain. 2. The 3'end of the *fpgs* gene abuts the 3'end of the *eng* gene, and the distal H3K36me3 observed is "spill over" from the neighboring gene. Nevertheless, we could conclude that H3K36me3 was not directly involved in tissue-specific suppression of transcriptional initiation at P2, because it was low across this region in mouse liver.

Integration of information on histone PTMs and DNA methylation

The results from the experiments discussed above shed light on several aspects of the tissue-specific regulation of the two promoters of the *fpgs* gene. We determined that epigenetic mechanisms were critically involved in the control of the P1 promoter. Our studies implicated a link between DNA methylation and the tissue-specific silencing of a CpG-sparse promoter, a role for DNA methylation that had not yet been extensively studied. As we progressed through these studies, it became very clear that the two promoters were regulated by different mechanisms. Likewise, the possibility that several mechanisms were either independently or in combination regulating the P2 promoter was suggested by comparing the data obtained from mouse liver and brain. The histone PTM profiles described across the *fpgs* locus in mouse brain were almost identical to those found in L1210 cells, hence neither chromatin compaction nor histone modifications nor

DNA methylation were preventing this promoter from firing in mouse brain. In contrast, the histone acetylation and methylation of lysine 4 of histone H3 were present but interestingly, lower over P2 in mouse liver. We considered the possibility that the restricted histone PTM patterns detected over P2 in mouse liver reflected the small amount of transcription initiating at P2 in mouse liver (Table 2-2).

We hypothesized that the high level of transcriptional activity occurring at P1 in mouse liver may prevent successful transcriptional initiation at P2 in this tissue. We proposed that elongation complexes generated at P1 were limiting assembly of pre-initiation complexes at P2. Previous studies using *in vitro* model systems supported the notion that when two promoters were arranged in tandem the activity at the upstream promoter greatly influenced the activity measured from the downstream promoter, a process referred to as transcriptional interference.

RNAPII is detected across P1 and P2 in mouse liver.

In order to test this hypothesis, we designed a series of experiments to examine the presence of transcriptional complexes across P2 in mouse liver. In the initial studies, we fine-mapped the residence of RNAPII over the two promoters using six overlapping PCR fragments covering approximately 1 kilobase of P1 and P2. The location of the primers across both promoters is illustrated in Figure 2-18. ChIP was performed on L1210 cells and mouse liver using a commercial antibody generated against total RNAPII. In the initial studies, immunoprecipitated DNA was amplified using semi-quantitative PCR. Ten microliters of each reaction were visualized on a 1% agarose gel. As expected, RNAPII

was minimally detected over P1 in L1210 cells, when compared to the non-specific IgG control signal, but was found across most of the P2 promoter region (Fig 2-19). We detected RNAPII over P1 in mouse liver, but very interestingly, a substantial level of this protein was also found spread across most of the P2 promoter region (Figure 2-19).

The sensitivity of the semi-quantitative PCR was not sufficient to observe small changes between primer sets nor to visualize the differences quantitative aspects of the overall patterns of RNAPII residency between L1210 cells and mouse liver at P2. Hence, Q-PCR was used in the next set of studies. The profiles of RNAPII residency generated from these second generation studies were far more straightforward to interpret then the gel pictures from the semi-quantitative experiments. For these experiments, we adapted the primers described in Fig 2-18 to our Q-PCR protocol, with the exception of primer set #1 within the P1 region. We were not able to design primers for this particular fragment that did not give multiple products (identified by the melt curve, see methods) in the PCR reactions.

Fine-mapping of RNAPII across P1 and P2 in L1210 cells and mouse liver using Q-PCR.

The use of Q-PCR confirmed our original observations and greatly extended our understanding of the difference between the pattern of RNAPII over P2 in L1210 cells and mouse liver. RNAPII detected at P1 was very low in L1210 cells, but was still present at levels significantly higher than the IgG controls across this region (Figure 2-20 B, left data). In contrast, RNAPII occupancy was much greater in mouse liver near P1

fragments, peaking immediately 3' of the transcriptional start site and decreasing substantially between 200 and 400 bp downstream of the transcriptional start site (Figure 2-20 B). A similar pattern of RNAPII detection was seen for L1210 cells near the P2 promoter as seen in liver near P1: polymerase occupancy peaked at or slightly after the transcriptional initiation site (Fig. 2-20 B, right data) and rapidly fell off downstream (see below). However, in mouse liver, RNAPII was resident over the entire P2 region, increasing before the DNase I hypersensitivity site defined in Figure 2-10 and remaining high over the P2 transcriptional start site and substantially downstream (Figure 2-20 B). The pattern of RNAPII observed at P1, in mouse liver, and at P2, in L1210 cells, reflected the profile of RNAPII previously described at the active p21 gene in a human colorectal carcinoma cell line. Based on that study in Hct116 cells, we took the peaks of RNAPII around the TSS at P1 in mouse liver and P2 in L1210 cells to represent stalled molecules engaged in the assembly of initiation complexes (81). Likewise, the measurable decrease in RNAPII observed *ca* 400 bp into the *fpgs* gene at P1 in mouse liver (Figure 2-20 B), was evidence that elongating RNAPII complexes were moving rapidly through the coding region of the *fpgs* gene, and were much less likely to be captured then paused initiation molecules using ChIP technology. The profile of RNAPII detected at P2 in mouse liver was unique because substantial levels of protein were found across the entire promoter region covering the linear DNA fragments previously mapped as requisite transcription factor binding sites (43)(Figure 2-20 B). We wanted to determine if the accumulation of RNAPII detected at P2 in mouse liver was specific to this genomic region or was representative of processes occurring across the entire body of the *fpgs* gene in liver.

RNAPII has a bimodal distribution pattern along the *fpgs* gene in mouse liver.

We employed the primer sets used in the studies examining the histone PTMs to map the occupancy of RNAPII at intervals along the 20 kb *fpgs* gene. We found striking differences between the patterns seen in mouse liver and leukemic cells (Fig. 2-20 A). In L1210 cells, there was a single segment of DNA that co-precipitated with RNAPII, namely, which centered on the transcriptional start site at P2 (Figure 2-20 A). In liver, there were two peaks, one centered on P1 and a second higher peak beginning at P2. Given that transcripts in mouse liver that initiate from P2 were barely detectable (Table 2-2), it was most likely that the RNAPII over the P2 promoter represented enzyme incorporated into elongation complexes that had originated at P1. We designed several experiments to test this hypothesis.

The C-terminal domain (CTD) of RNAPII over P2 is phosphorylated at Serine 2 in liver.

It has been shown that phosphorylation of serine 5 residues in the repeated YSPTSPS peptide heptad in the CTD of RNAPII is present throughout transit of a gene, but is maximal on RNAPII found over the region of a gene associated with assembly of an initiation complex (81, 133). In contrast, phosphorylation of serine 2 in this motif accumulates only as the elongation complex travels through a transcribed sequence and is highest over the cleavage/polyadenylation signal (81, 133). We hypothesized that if the RNAPII complexes detected over P2 in mouse liver were predominately engaged in transcriptional elongation, a measurable increase in the level of protein using the phospho-

Serine 2 antibody would be located over P2, when directly compared to the amount measured over P1 in mouse liver. Because these studies were critically dependent on the specificity of the antibodies for the phospho-serine 2, phospho-serine 5, and unmodified peptides, we tested the specificities of these antibodies under our ChIP conditions.

Specificity of phospho-Serine 5 and phospho-Serine 2 antibodies

Phosphopeptides used to raise the Serine 5 and Serine 2 phosphospecific antibodies and an unphosphorylated peptide were purchased to test the specificity of the antibodies. Increasing concentrations of the peptides (0.1 μg, 1 μg, and 10 μg) were pre-incubated with the antibodies raised against phospho-Serine 5 and phospho-Serine 2 to determine how well these peptides would block the antibody-epitope recognition reaction (Figure 2-21). Following incubation, the antibody-peptide mix was added to sonicated L1210 cell lysates and the samples were incubated overnight at 4C and processed using our ChIP protocol (Figure 2-21). Q-PCR was performed on immunprecipitated DNA using primers against the P2 *fpgs* promoter and a fragment within the GAPDH coding region, locations were RNAPII complexes phosphorylated at Serine 5 or Serine 2 were predicted to be enriched, respectively. We chose the P2 *fpgs* promoter to assess the phospho-Serine 5 antibody because it represents the 5' end of a highly active gene, where serine 5 phosphorylation of the CTD should be maximal (Figure 2-21 A). Primers amplifying a portion of the coding region of the GAPDH gene, a region where serine 2 phosphorylation of the CTD of RNAPII should be substantial, were used to test the phospho-serine 2 antibodies (Figure 2-21 B). The enrichment of the P2 promoter in the immunoprecipitated

DNA using phospho-Serine 5 antibody was severely blocked when the antibody was pre-incubated with 1 or 10 μg of phospho-Serine 5 peptide (Figure 2-21), whereas incubation with either the unphosphorylated or phospho-Serine 2 peptide had no effect on the level of signal detected (Figure 2-21). The phospho-Serine 2 antibody was more sensitive to inhibition and detection of the GAPDH coding region within the samples was almost completely blocked by the addition of 0.1 μg of phospho-Serine 2 peptide (Figure 2-21), but neither the phospho-Serine 5 nor the unphosphorylated peptide affected the level of signal measured (Figure 2-21).

RNAPII CTD post-translational modifications over P1 and P2 of the mouse *fpgs* gene.

We followed phosphorylation of the C-terminal domain of RNAPII at serine 5 and serine 2 over the two *fpgs* promoters as an index of whether the RNAPII detected at P2 in mouse liver was binding in a futile initiation complex or was engaged in elongating a transcript initiated upstream. We used the signal measuring total RNAPII across the two promoters to normalize the levels obtained using the phospho-specific antibodies. As expected, minimal levels of RNAPII were detected over P1 in L1210 cells (Figure 2-22)). In contrast, a substantial amount of RNAPII was detected over P2 in L1210 cells, peaking near the +1 position corresponding to the TSS of P2 (Figure 2-22). Two experiments are presented for the RNAPII profiles over the two *fpgs* promoters in mouse liver. In the first experiment, serine 5 phosphorylated RNAPII over P1 was the most abundant signal, and total and phospho-serine 2 RNAPII signal were similar (Figure 2-23). In the second experiment, phospho-serine 5 and total RNAPII levels were similar and phospho-serine 2

levels were much lower (Figure 2-23). The patterns of enrichment signals detected for the three antibodies across either promoter in mouse liver or L1210 cells were not highly consistent between experiments. In fact, we found that changing the lot of antibody used could change the value of signal measured for a given antibody. These data highlight the importance of analyzing changes in levels within one antibody and comparing between experiments using levels detected at a control gene. In this case, we used the levels detected at P1 as the control to analyze the data collected for P2 in mouse liver. The levels of phospho-Serine 5 and phospho-Serine 2 RNAPII were normalized to the amount of total RNAPII in order to identify when changes in phospho-specific RNAPII levels were the result of an increase in overall RNAPII levels or due to the hyperphosphorylation of the RNAPII molecules present at a given position (Figure 2-23). Interestingly, in mouse liver, substantial levels of both phospho-Serine 5 and phospho-Serine 2 RNAPII molecules were detected over P1 (Figure 2-23). Ratios of the phospho-specific signals relative to the total level of RNAPII were generated using data collected on a fragment centered on the position 50 bp from the P2 promoter TSS position and the fragment centered on 75 bp downstream of the P1 promoter TSS (Figure 2-23). These data are shown in the bar graph in Figure 2-23. At P2 in L1210 cells, where this promoter is active, the ratios of phospho-Serine 5 and Serine 2 to total RNAPII were 2.0 and .66, respectively. Similarly, in mouse liver, when P1 is active, a ratio of 1.6 and 1.0 were determined for phospho-Serine 5 levels and .83 and .5 were calculated for the level of phospho-Serine 2 (Figure 2-23). The data collected at P2 in L1210 cells and P1 in mouse liver were taken as representative of the pattern in an initiation complex: serine 5 phosphorylation per molecule of RNAPII was

highest near nt +1, the transcriptional start site, while serine 2 phosphorylation was lower and proportional to the level of RNAPII throughout this region.

The profiles of RNAPII obtained at the P2 *fpgs* promoter in mouse liver were very interesting. The level of RNAPII phosphorylated at Serine 2 detected at P2 increased substantially compared to the amounts observed at P1 in mouse liver in both experiments. Accumulation of RNAPII complexes over P2 peaked at the position centered on 50 nt downstream of the P2 TSS (Figure 2-23). Phosphorylated Serine 2 RNAPII relative to the total RNAPII increased approximately 3-fold between the P1 promoter and the P2 promoter regions in liver (Figure 2-23). In contrast, changes in the level of phospho-Serine 5 between the two promoters were reflected by the changes in total RNAPII levels, indicating that the levels of serine 5 phosphorylation were not detectably different between P1 and P2 in mouse liver (Figure 2-23). We interpreted these data as direct evidence that the RNAPII detected over P2 in liver was a component of an actively transcribing elongation complex.

In what was initially a very surprising observation, the distribution of RNAPII across P2 was also easily detectable in mouse brain (Figure 2-24). However, this observation was in agreement with a human whole-genome tiling study that appeared just as we had done these studies that indicated that even inactive promoters are usually attempting to load initiation complexes (87)(see Discussion). Interestingly, the RNAPII was detected broadly across the P2 promoter in brain, similar to the data presented for the P2 promoter in mouse liver (Figure 2-23). We initially took this to suggest that the pattern of RNAPII occupancy observed at P2 in L1210 cells may be characteristic of promoters

engaged in high levels of active transcription. However, the phosphorylation pattern of RNAPII across P2 in mouse brain was, however, remarkably similar to that profile observered in L1210 cells: RNAPII present across the brain P2 was phosphorylated at serine 5 to a higher degree and serine 2 phosphorylation was minimal, indicating that these complexes were indeed futile initiation complexes.

Evidence that elongating RNAPII over P2 blocks assembly of transcriptional initiation complexes in liver.

Transcription from the mouse *fpgs* P2 promoter has been shown to depend on several Sp1 sites (43). A typical Sp1-driven promoter involves binding of TBP to Sp1 with subsequent recruitment of other components of the pre-initiation complex (PIC). We assessed the binding of components of the PIC to *fpgs* P2 in mouse liver, brain, and L1210 cells. The general transcription factors selected were Sp1, TBP, TFIIB, and TAFIIp250; the residence of these proteins at active promoters have previously been used as a mark of pre-initiation complex assembly at the p21 promoter (81). Since it appeared that the P2 promoter of the *fpgs* gene, at least in brain, may be controlled at the level of elongation, we also tested for the presence of NELF-A, a subunit of the Nelf complex, that has been suggested to play a critical role in RNAPII pausing at the early stages of transcriptional elongation in both *Drosophilia* and human experimental systems (169, 267). As shown in Figure 2-25, in L1210 cells and mouse brain, Sp1 and TBP were loaded onto the region of DNA immediately upstream of the transcriptional initiation site, as well as did TFIIB, TAFIIp250, and the NELF-A protein (Figure 2-25). These proteins persisted on the P2

locus past the point of initiation (+200 nt) (Figure 2-25). In striking contrast, Sp1 was very low at -200 nt in liver, and TBP was not detectable. TAF250 and NELF-A were also very low. At +200 nt in liver, Sp1, TFIIB, TAFIIp250, and Nelf A, were detected at low levels and the TBP was undetectable (Figure 2-25). Interestingly, high levels of Nelf-A, relative to the abundance of the general transcription factors detected at P2, were measured in mouse brain (Figure 2-25). This may suggest a role for Nelf in the silencing of P2 in mouse brain (see discussion). We concluded that the binding of several proteins needed for assembly of a successful PIC was being actively restricted at P2 in liver. This is consistent with a promoter interference mechanism whereby the elongating RNAPII complexes initiating at P1 were physically occluding the DNA over the P2 region in liver.

DISCUSSION

Genome-wide analysis of transcription start sites, in humans and mice, have revealed that over 50% of genes contain multiple promoters (37). Alternate promoter usage permits a single genetic locus to generate distinct protein products in response to different cellular backgrounds. How exactly is promoter choice dictated? Describing the mechanisms driving promoter selection is of great importance to our understanding of basic processes of transcription, but also has the potential to uncover the basis for new therapeutic strategies that could selectively target gene activity in one particular tissue-type. In our studies, we have used the mouse *fpgs* gene as a model to address the controls coordinating the use of two promoters in a tissue-specific fashion.

The mouse *fpgs* gene is transcribed from two alternative promoters, one of which (P2) drives transcription in normal dividing cells types and in tumors, the other (P1) controls transcription predominantly in two differentiated tissues, liver and kidney. We sought to explain how these two promoters are regulated in a virtually black and white manner; that is, tissues using P2 never initiate at P1, and tissues that initiate at P1 have a very limited, although detectable, usage of P2. Our analysis leads to the conclusion that activity at P1 correlates with the presence or absence of CpG methylation, making *fpgs* one of the few known cases in which methylation inversely correlates with expression in non-neoplastic tissues at a CpG-sparse promoter. In contrast, CpG methylation did not play a role in tissue-specific initiation from P2. An analysis of some of the most studied histone PTMs showed patterns across the span of this 20 kb gene only partly predictable from previous studies. Over the P2 promoter in mouse brain we detected RNAPII and general transcription factors, as well as, histone acetylation and lysine 4 of histone H3 methylation, suggesting that in mouse brain, RNAPII complexes were poised for transcriptional activation at P2. However, the real surprise in our results is that RNAPII does not seem to have access to P2 for transcriptional initiation in liver when P1 is active, despite open chromatin context and lack of CpG methylation in the P2 region. Hence, mouse *fpgs* is a clear case of transcriptional interference in an endogenous mammalian gene. The studies presented in this chapter have described several different mechanisms of transcription at the mouse *fpgs* gene, their implications and relevance to current literature is discussed below.

Tissue-specific chromatin state of the mouse fpgs gene

By profiling several histone PTMs across the mouse *fpgs* gene we gained an understanding for the chromatin context surrounding the *fpgs* gene in L1210 cells, mouse liver, and mouse brain. The marks of histone H3 and H4 acetlyation and tri-methylation of lysine 27 of histone H3 are distinguishing features of euchromatin and hetermochromatin, respectively (135). We were surprised to find detectable levels of histone acetylation (Figure 2-12) upstream of the P2 promoter in L1210 cells and mouse brain, since this region of the gene, in these tissues, was transcriptionally silent. Likewise, we expected to detect H3K27 tri-methylation across the genomic regions between P1 and P2 in L1210 cells, and likely across the entire locus in mouse brain. The levels of H3K27 tri-methylation detected followed similar patterns across the *fpgs* gene in all three tissues studied; thus, it was unlikely that this region was silenced by heterochromatic formation (Figure 2-16). Taken together, these data indicated that the *fpgs* locus was positioned in genomic regions of euchromatin in L1210 cells, mouse liver, and mouse brain independent of the transcriptional activity occurring at this locus.

Why is the mouse fpgs gene in brain not in a heterochromatic state? Genome-wide maps of histone PTMs in human embryonic stem cells and a few differentiated tissues have classified genes based on their "chromatin state" in pluripotent cells (161). They propose that the pattern of several histone PTMs, namely histone H3 lysine 4 and 27 tri-methylation, found at genes in embryonic stem cells is closely linked to cellular state and gene function. They found that almost all CpG-rich promoters in embryonic stem cells are associated with H3K4me3 (161). Additionally, upon differentiation, most of the CpG-rich

promoters retain this histone mark. On the other hand, CpG-poor promoters were typically not modified in embryonic stem cells and marks of activation or repression were only detected in differentiated states.

Interestingly, the two promoters of the mouse *fpgs* gene represent examples of the two classes of promoters identified based on chromatin state; the CpG poor P1 promoter was marked with H3K4me3 only in mouse liver, and the P2 CpG island promoter was marked with H3K4me3 in all tissues. Since such a large number of genes have been found to have multiple promoters, it will be important to investigate how different chromatin states at a single genetic locus will impact the overall chromatin context surrounding the entire transcriptional unit. With this in mind, we wondered if the persistent H3K4me3 mark at P2 in all tissues was sufficient to prevent heterochromatin formation across the *fpgs* gene in mouse brain.

Profiling of mono-methyl lysine 27 of histone H3 across the mouse fpgs gene in L1210 cells, mouse liver, and mouse brain reveals unexpected patterns from prior literature.

A very interesting observation was the presence of H3K27me1 throughout the entire coding region of the *fpgs* gene in all three tissues. At the time these studies were performed, little to no information was known about the presence of H3K27me1 at active genes. Since then, high-resolution genome-wide histone methylation analysis has shown H2K27me1, along with H3K9me1, in the coding regions of most active human genes (15). It has been proposed based on this global analysis that all mono-methylated forms of histone H3 lysines are associated with active transcription (15). However, we detected

abundant levels of mono-methylated H3K27 over *fpgs* in mouse brain, a tissue where this locus is completely silenced (Figure 2-16). Perhaps mono-methylation of H3K27 is not a mark of transcriptional activity, but rather a mark of euchromatin. The distinct patterns of observed of the mono-, di-, and tri- methylated forms of histone H3 at lysine 27 highlight the complexity of the mechanisms leading to a particular histone methylation pattern. How the distinct patterns are generated most likely depends on the nature of the surrounding histone PTMs, especially since most HMTs are able to methylate the ε-amino lysine at all three positions (225). The number of possible modified states is enormous and assessing each mark in the context of the surrounding histone PTMs, DNA methylation status, and transcriptional activity appears to be a necessary step in further understanding the biological consequences of a given histone PTM, or combinations of histone PTMs. It is highly likely that multiple PTMs in sum constitute a signal for molecular phenomena. Initial studies investigating this concept used mass spectrometry to map the combination of PTMs on histone H3 in *T. thermophila* and identified permutations of PTMs associated with transcriptional activation and repression (236). An example illustrating the interplay between histone PTMs has been described at active eukaryotic genes; histone H2B ubiquitination has been shown to enhance the conversion of di-methyl lysine 4 on focal histone H3 to the tri-methylated state (261). The postulate that specific combinations of histone PTMs on surrounding nucleosomes may dictate the methylation state of lysine 27 of histone H3 has yet to be tested.

Is tissue-specific methylation regulating the P1 CpG-sparse promoter?

The hypothesis that DNA methylation is involved in controlling tissue-specific expression has been tested in both genome-wide studies (227, 256) and in studies focusing on a few single genetic loci (75). However, these studies have largely addressed the methylation status of CpG islands across different tissues, in spite of the fact that most tissue-specific promoters are thought to be CpG-sparse. Similarly, the processes leading to transcriptional silencing at promoter regions via DNA methylation are understood largely in the context of aberrantly hypermethylated CpG islands in neoplastic cell lines (121) and the ability of these mechanisms to silence CpG-sparse promoters in normal tissues has only been tested in a few circumstances (75, 203, 254). Thus, the potential for epigenetic mechanisms to regulate tissue-specific expression in normal tissue remains unclear. In spite of this fact, current thought proposes that CpG-sparse promoters are methylated independent of activity state and less frequently rely on epigenetic mechanisms for transcriptional activation or repression (164). Our analysis of the DNA methylation across the P1 promoter of the mouse *fpgs* gene in tissues where this promoter is either transcriptionally active or silent argues a role for DNA methylation in the tissue-specific control of CpG-sparse promoters (Figure 2-9). Support for this concept has recently been provided. Bisulfite sequencing of the CpG dinucleotides in human chromosomes 6, 20, and 22 in several mouse tissues identified a population of differentially methylated CpG-sparse promoters (58). This population of promoters had not previously been detected because the limitiations of methods commonly used for genome-wide methylation anaylsis, e.g. MeDIP and RLGS, prevent the assessment of methylation at single CpG resolution (227,

256). Our analysis of epigenetic marks over the P1 promoter identified correlations between DNA methylation and histone PTMs previously predicted through studies at CpG-island promoters; i.e., three histone PTMs associated with active genes, H3Ac, H4Ac, and H3K4me3, were coincident with hypomethylation of regional CpG dinucleotides across P1 in mouse liver. However, we were surprised to find that in tissues where the CpGs of the P1 promoter were methylated, regional nucleosomes were not methylated at H3K9, a PTM commonly associated with DNA methylation at silenced promoters (Figure 2-15) (74). Lysine 9 methylation of H3 is tightly linked to DNA methylation and transcriptional repression in some systems, particularly *Arabidopsis thaliana* and *Neurospora* (74, 234), which has led to the proposal that DNA methylation is targeted to genes that have been silenced by other mechanisms. We took the absence of H3K9 methylation at P1 in the presence of DNA methylation to suggest that epigenetic control of the tissue-specific CpG-sparse P1 promoter may involve unique combinations of epigenetic marks to ensure proper patterns of expression.

While it is difficult to separate out cause and effect *in vivo*, studies using genetically modified mouse models with altered expression of proteins involved in establishing either DNA methylation or histone PTM patterns have tried to address the sequence of events controlling the patterns of epigenetic modifications leading to gene regulation (137). Recent studies have shown that the depletion of Dnmt1 in mouse embryonic fibroblasts resulted in a marked increase of H3K4 methylation and H3 acetylation (137). On the other hand, loss of either the histone demethylase, LSD1 (255) or the histone H3 lysine 9 methylase, G9a (53), in embryonic stem cells resulted in global

changes in DNA methylation patterns. These data suggest that the sequence of events leading to certain combinations of epigenetic marks is not universal. Further insight into the role of DNA methylation and histone PTMs in the regulation of gene expression may be gained from classifying promoters based on the primary epigenetic process that initiates mark the regulation of each individual promoter. In the absence of the expected repressive histone H3K9 modification at the chromatin surrounding the P1 *fpgs* promoter in tissues where this promoter is silent, our results suggest that DNA methylation is the prime determinant of tissue-specific chromatin condensation and repression of the P1 promoter. We concluded that, in adult dividing tissues of the mouse, the P1 *fpgs* promoter was locked in a transcriptionally inactive configuration by DNA methylation and that transcriptional activity at P1 in liver was reinforced by a series of histone PTMs commonly associated with transcriptional activity.

Potential mechanisms of regulation of P1 via DNA methylation.

We considered the mechanism of the proposed DNA methylation-induced transcriptional repression of the P1 promoter. DNA methyl-CpGs are binding sites for MBPs which serve as scaffolds for the recruitment of histone deacetylases (123, 170) and for chromatin remodeling complexes associated with transcriptional silencing. MBPs have been shown to preferentially bind promoter regions with more densely packed CpG dinucleotides, and the potential for this class of proteins to mediate transcriptional silencing at CpG sparse promoters needs to be further explored. Interestingly, in L1210 cells, where P1 is silent and methylated, low levels of RNAPII were reproducibly detected

across the promoter region, supporting the idea that the entire *fpgs* locus is in a dynamic state of euchromatin, where *trans* factors occasionally have access to DNA binding elements. In this context, the DNA methylation of the P1 promoter may be functioning to prevent the transcription factor and RNAPII binding events required to successfully assemble pre-initiation complexes at the P1 promoter. Support for this concept was recently furnished in a very important genome-wide study that correlated methylation status of promoters with histone PTMs in mouse embryonic stem cells and certain differentiated tissues (159). That study found that the DNA of CpG-sparse promoters were unmethylated in embryonic stem cells and regional nucleosomes were marked with H3K4 methylation but, in more differentiated cells, the methylation status of CpGs in these promoters distinguished a transcriptionally active gene from a silenced gene (159). In contrast, CpG island promoters were unmethylated across the different developmental stages. If methyl-CpGs block *cis* elements from transcription factor binding at P1 of the *fpgs* gene, high levels of tissue-specific transcription factors in the nucleus may relieve this occlusion, providing a discrete method of transcriptional activation for certain tissue-specific promoters. As was suggested by the case of *fpgs* P1, nucleosomes surrounding tissue-specific promoters controlled by DNA methylation may not be methylated at H3K9. This may have functional implications for these promoters by preventing highly stable repression and maintaining the ability for transcriptional activation to occur when the cellular context changes. Thus, the *fpgs* P1 promoter may represent a class of tissue-specific promoters identified by differential DNA methylation patterns and CpG sparsity, where histone PTMs associated with transcriptional activation are present because of the

presence of RNAPII initiation complexes, rather than being causative of transcriptional initiation and histone repressive marks are not used.

Are epigenetic mechanisms silencing P2 in mouse liver and brain?

The P2 region was devoid of DNA methylation in all tissues studied whether or not this promoter was used to initiate transcription (Figure 2-10). The histone PTMs near P2 show distinct patterns in different tissues: In either a tissue in which P2 was exclusively used (L1210 cells) or one in which the entire *fpgs* locus was silent (brain), there was a broad peak of H3Ac and H4Ac centered on the hypersensitivity site at this promoter, but there was a much more compact peak of these acetylated nucleosomes over P2 in liver (Figure 2-12). Likewise, a substantial peak of H3K4me3 centers on P2 in L1210 and brain, but a distinctly smaller peak was found over this same region in liver. It appeared from these data that the processes involved in regulating the silencing of P2 in mouse liver and brain were different (Figure 2-14).

Poised RNAPII at P2 in mouse brain: the role of negative elongation factors

In a very important recent paper, genome-wide analysis of human embryonic stem cells found that the promoters of most protein-encoding genes were associated with nucleosomes containing H3K4me3 and H3Ac and were also bound to RNAPII, while only a fraction of these transcriptionally-primed promoters produced mature full-length transcripts (87). Additional work in *Drosophilia* has confirmed this work and begun to address the functional significance of and mechanisms establishing the poised RNAPII

complexes at these promoters (169). Prior to these studies, promoters with assembled transcriptional complexes poised for activation were thought to be few (19), however it is now clear that there is a large number of genes regulated at the early stages of elongation. The active histone PTM (H3Ac, H4Ac, and H3K4me3) and the substantial level of RNAPII detected over P2 in mouse brain (Figure 2-24) placed this promoter in the class of promoters identified by Guenther *et. al* that are poised for transcriptional activation. This is very surprising in the case of *fpgs* P2, which will never again be activated in normal brain.

Current models of transcriptional poising propose that the phosphorylation of NELF by P-TEFb displaces NELF bound to the poised RNAPII, allowing TFIIF access to RNAPII; this release of NELF facilitates the transition from a stalled state to productive elongation (184). The role of NELF in establishing the poised state of RNAPII genome-wide has been supported through the use of siRNA strategies in fly systems (169). Two-thirds of the poised RNAPII complexes in *Drosophilia* were relieved when levels of NELF were reduced. In our studies, we probed for NELF at the P2 promoter of the mouse *fpgs* gene in L1210 cells, mouse liver, and brain (Figure 2-25). Since P2 in mouse brain appeared to be bound by stalled RNAPII elongation complexes we thought NELF would be detected at P2 in mouse brain, but not in L1210 cells. Surprisingly, NELF was detected at P2 in both L1210 cells and mouse brain. A similar observation was made at the p21 promoter in human colorectal carcinoma cells: NELF was detected at the p21 promoter when the promoter was bound by RNAPII in a stalled state and remained at the promoter upon transactivation of the p21 gene by p53 (81). While the presence of NELF may be an

essential component to poised RNAPII complexes, our data suggests that the recruitment of NELF to P2 is not a determinant of RNAPII poising at P2 in mouse brain. We took this to suggest that additional components, such as post-translational modifications or binding partners of NELF, are involved in the stalling of RNAPII at P2 in mouse brain. Recent genome-wide studies have detected NELF residence at promoters with poised RNAPII complexes and at promoters of highly active genes (138), suggesting that the observations we made at P2 in L1210 cells and mouse brain are universal. Interestingly, in *Drosophila* S2 cells NELF depletion up-regulated a number of rapidly responding genes, such as Hsp70, but a large proportion of genes previously shown to bind NELF showed a decreased level of expression (77). Both our studies and genome-wide studies suggest that the hypothesis that NELF recruitment and displacement coordinates promoter-proximal pausing is an oversimplification. It will be important to understand the different states of NELF at both active and stalled genes to understand the role of this protein in each condition.

Profiles of RNAPII at poised and active genes: insights into mechanisms of stalling

The profiles of RNAPII across the P2 promoter in L1210 cells and mouse brain were remarkably different. Molecules of RNAPII were reproducibly detected across the entire P2 region in mouse brain, whereas in L1210 cells a sharp peak of RNAPII was detected near the P2 TSS (Figure 2-22 and 2-24). One explanation for this observation is that differences in sonication between mouse brain and L1210 cells influenced the pattern of RNAPII detected across P2. However, data from a recent genome-wide study suggests

that our observations are not likely due to technical issues (214). RNAPII occupancy mapped across promoters bound by stalled and active molecules throughout the genome in $CD4^+$ T cells revealed similar profiles to those determined in our studies: sharp peaks of RNAPII occupancy at the TSS defined promoters actively engaged in transcription, and inactive promoters were bound by RNAPII across most of the promoter region (214). Overall these data suggest that on average the positioning of RNAPII at the promoter of active genes appears to be highly regulated, resulting in the observed narrow peak of occupancy near the TSS detected at the P2 promoter of the mouse *fpgs* gene, in L1210 cells (Fig 2-22), and at active promoters genome-wide. In contrast, the broad profile of RNAPII molecules bound to P2 in mouse brain (Figure 2-24) suggests that recruitment of enzymes to the P2 promoter region of the *fpgs* gene involves fundamentally different processes over the stalled vs. the productive chromosomal context. While it is difficult to separate out cause and effect, we expect that if the P2 promoter in mouse brain was activated, the pattern of RNAPII residency at P2 would reflect the profile observed in L1210 cells. The mechanisms causing the establishment of the two distinct patterns of RNAPII residency at active and stalled genes remains an open question. Interestingly, a high-resolution nucleosome-mapping study determined that the positioning of nucleosomes across stalled and active promoters were very similar, suggesting that nucleosome placement does not dictate the difference in RNAPII binding patterns at promoters of stalled and active genes (214). While the position of nucleosomes may be similar, the composition of individual nucleosomes may be drastically different between a stalled and active promoter regions. Conventional histones can be replaced within a

nucleosome with histone variants that differ in their primary sequence from canonical histones that have functional consequences for gene regulation (135). Two histone variants in particular, H2A.Z and H3.3, have been linked with transcriptional regulation. Genome-wide studies in yeast and humans have identified H2A.Z enrichment at promoter regions upstream and downstream of TSS (15). Likewise, a role for the histone variant H3.3 has been implicated through global analysis in *Drosophila*, where this variant was found localized to the promoter regions of active genes (163). It will be important to determine if nucleosomes surrounding promoter regions of stalled genes have a unique composition of histone variants.

We conclude that the silencing of P2 in mouse brain and mouse liver represents different phenomena. In brain, P2 and P1 are completely silenced whereas, in liver, initiations are blocked by elongation complexes. Accordingly, low levels of initiation can be detected from P2. The fact that transcriptional poised complexes assemble at P2 in brain makes it more likely that the absence of such complexes at P2 in liver is due to an active exclusion process.

Transcriptional interference at P2 in mouse liver: mechanism and consequences

In mouse liver, there were abundant transcripts from the *fpgs* gene, almost all initiating at the P1 promoter (Table 2-2). The profiles of histone PTMs commonly associated with the 5' end of active or poised genes (H3Ac, H4Ac, and H3K4me3) were detected over P2 in mouse liver, however the levels detected were substantially restricted compared to the amount observed at P2 in mouse brain or L1210 cells or at P1 in mouse

liver. We took this to suggest that something was limiting the assembly of RNAPII and initiation factors at the P2 promoter, in spite of active traverse of transcriptional elongation complexes through this region of genomic DNA. How was this happening? Literature precedent suggested a promoter interference mechanism whereby activity at the P1 promoter prevents transcriptional initiation at the P2 promoter (83, 155, 188). In higher eukaryotes, this method of transcriptional repression has been described in the tandemly arranged *unr* and N-*Ras* genes, where transcription of the *unr* gene interferes with initiation at the N-*Ras* promoter, which is spaced only 150 nt downstream of the *unr* termination signal (28). Similarly, deletion of an embryonic ß-like globin gene, Ey, caused the activation of a second ß-like globin gene, ßho, which is located immediately downstream of the deleted gene (107). However, the mouse *fpgs* gene represents an interesting case in which transcriptional interference is used to coordinate use of two potentially active promoters, spaced a fairly large distance (10 kb) apart. Our analysis of the P2 promoter in liver indicates that RNAPII elongation complexes are residing over an extended length of the *fpgs* genomic locus, covering the nucleosome-deficient DNase hypersensitive region before P2 and continuing downstream, physically limiting initiation at P2 (Figure 2-23). In support of this concept, we observed that the residence of several key general transcription factors at the P2 promoter in mouse liver was severely limited (Figure 2-25). Such a physical occlusion mechanism has been proposed to involve the decreased binding of transcription factors, such as Sp1, or through steric obstruction of DNA binding regions required for transcription factor interactions (83, 107, 155). We concluded that the mouse *fpgs* gene represents an example of this mechanism, one of the

few documented for an endogenous gene, and the only example in the literature to date demonstrating TI as a means of organizing transcription from a multi promoter gene.

What causes elongating RNAPII to stall at P2 in mouse liver?

Substantial levels of RNAPII molecules were detected at both the P1 and P2 promoters of the mouse *fpgs* gene in mouse liver. The accumulating complexes over P2 in mouse liver appeared to be hyperphosphorylated at serine 2 residues within the CTD of RNAPII (Figure 2-25), suggesting that these molecules of enzyme over P2 in mouse liver were engaged in elongation.

We considered several mechanisms as explanations for the accumulation of RNAPII observed at P2 in mouse liver. The presence of specific epigenetic marks and chromatin structure may play a role in establishing the profile of RNAPII at P2 in mouse liver and my analysis to date has focused on these factors. This analysis has not given us a concrete explanation of the RNAPII stalling. CpG methylation declines to almost undetectable levels 500 nt upstream of the P2 TSS, coincident with an increase in histone H3 and H4 acetylation and H3K4 tri-methylation in all tissues studied (Figure 2-9). The correlation of RNAPII accumulation with the hypomethylated CpG island in the coding region of the *fpgs* gene in mouse liver suggests that the observed change in DNA methylation and histone PTMs may impair the transit of elongating RNAPII across the P2 promoter region. Previous work using transgenic mice addressed the influence of intragenic DNA methylation on transcriptional elongation efficiency and determined that CpG methylation, coincident with loss of histone acetylation and H3K4 methylation,

impeded RNAPII processing across the methylated stretch of genomic DNA (146). Based on this evidence it seemed unlikely that loss DNA methylation or histone PTMs associated with the P2 promoter region in mouse liver were directly causing the accumulation observed. One potential epigenetic mechanism causing the increased RNAPII levels observed is the positioning of nucleosomes across the P2 promoter region in mouse liver. The DNAse Hypersensitivity site and Total histone H3 maps presented (Fig 2-10) assess gross chromatin structure, but recent high-resolution mapping studies highlight the functional significance of fine differences in nucleosome positioning not previously appreciated (214). It will be very interesting to determine the position of nucleosomes across the P2 region in L1210 cells, mouse liver, and mouse brain. These studies may find that the passage of RNAPII complexes across the P2 region in mouse liver is determined by chromatin structure.

Currently, we propose that elongating RNAPII complexes accumulate over P2 in mouse liver as a result of collisions with transiently bound initiation factors. The residence of general transcription factors and the accumulation of H3K4me3 at P2 in mouse liver, albeit at minimal levels, indicate that PICs are occasionally able to assemble at the P2 promoter in this tissue. We hypothesize that passage of RNAPII molecules across P2 is transiently blocked by the presence of initiation factors and the displacement of these factors is required for elongating RNAPII molecules to proceed down the rest of the gene (Figure 2-26). If this were the case, blocking PIC assembly at P2 in mouse liver should prevent the accumulation of RNAPII complexes over this region. It is interesting to note that mouse kidney also uses the P1 promoter to generate fpgs mRNA, but in this tissue the

P2 promoter also generates a substantial level of transcripts initiating at P2. The level of *fpgs* mRNA generated from P1 is higher in liver than in kidney, and we consider that the lower amount of initiation occurring at P1 in kidney may account for the higher level of P2 expression detected in mouse kidney. It will be interesting to explore the residence of RNAPII and histone PTMs across P2 in kidney and may provide further insight in to the mechanism of TI at the *fpgs* locus.

Are there consequences to the stalling of RNAPII at P2 in mouse liver?

The mouse *fpgs* gene generates two splice variants in a tissue-specific pattern from two different promoters (244). In dividing cells, such as L1210, transcription initiates at the P2 promoter and the mature mRNA includes exons 1-15 (Fig 2-3). In contrast, in mouse liver, a tissue where the P1 promoter is active, the 5' exons A1a and A1b are directly linked to exons 2-15, splicing out exon 1 (Fig 2-2). The distance between exons A1b and 2 is remarkably long and how the splicing machinery handles 10 kb of pre-mRNA in order to achieve the correct splicing pattern of exon A1b to exon 2 remains unclear. The generation of alternative splice variants is dependent upon the availability of splicing factors, splicing machinery, and regulatory sequences within the immature mRNA. However, work over the past ten years has shown that the rate of transcriptional elongation also directly impacts alternative splicing patterns (187). This concept was first addressed in experiments designed to test the hypothesis that mRNA processing occurred simultaneously with transcription. Proudfoot et al. altered the splicing pattern of the α-tropomyosin gene by artificially positioning RNAPII pause sites within the gene, thus

delaying the transcription of a splicing inhibitory element and causing the inclusion of an alternative exon in the mature mRNA (187). Additional studies used a mutant form of RNAPII, with a slower elongation rate, to show that the inclusion of alternate exons is enhanced when transcription is accomplished with the mutant form of RNAPII (51). We are very interested to address the postulate that the accumulation of RNAPII at the P2 promoter in mouse liver represents a stalled complex, which dictates the splicing pattern of *fpgs* mRNA detected in this tissue. The potential of alterations in the rate of RNAPII transit across the body of a gene to generate multiple splice variants from a single gene is a tremendously interesting concept, and the mouse *fpgs* gene may be an excellent model system to determine the sequence of events involved in this affect.

Two classes of promoters distinguished by methylation genome-wide?

Recent studies have examined RNAPII occupancy, H3Ac, and H3K4me3 at promoter regions in several differentiated mouse tissues. That study expanded on the previous observations made in genome-wide studies by classifying promoters based on their CpG status and describing patterns across several mouse tissues. In doing so, it became clear that CpG islands are most often bound by RNAPII and regional histones are acetylated and methylated at lysine 4 of histone H3 in a tissue-independent fashion (13). In contrast, CpG-sparse promoters were bound by RNAPII in a tissue-specific pattern and presumably in tissues only when they are expressed. The *fpgs* gene has one of each type of promoter identified in the abovementioned study (13). The difference between these two is that P1 appears to be controlled by DNA methylation, and P2 is not. It will be of

tremendously interesting to determine if DNA methylation is a distinguishing characteristic between the poised promoters and the tissue-specific promoters identified in the differentiated tissues of the mouse.

Overall, we conclude that the mouse *fpgs* gene uses at least two mechanisms, epigenetic marking of the P1 promoter and transcriptional interference at the P2 promoter, to ensure that each of two isoforms of FPGS are expressed as needed in dividing and select differentiated tissues.

Peptide name	Peptide sequence
RNAPII CTD repeat peptide	YSPTSPS
RNAPII CTD repeat peptide - phospho S2	YSPTSPS PO_4
RNAPII CTD repeat peptide - phospho S5	YSPTSPS PO_4

Table 2-1. Sequences of the C-terminal domain heptad repeat of RNAPII peptides used to test the specificity of the serines 2 and serines 5 phospho-specific antibodies. The position of the phosphate group in each synthetic peptide is illustrated**.**

	P1 (pg FPGS mRNA ± SD)	P2 (pg FPGS mRNA ± SD)
L1210	<0.0001	0.15 ± 0.014
Liver	0.21 ± 0.0054	0.001 ± 0.0001
Brain	<0.0001	<0.0001

Table 2-2. Quantitative RT-PCR determination of FPGS expression from P1 and P2 in mouse L1210, liver and brain. Total RNA (5µg) was reverse transcribed using Superscript III in a total volume of 20µl, and 1µl was used in real time pCR assays using primers specific for exons A1a/A1b to exon 3 (P1) or exon 1 to exon 3 (P2). Standard curves for quantitation were constructed using plasmids carrying cDNAs corresponding to the products of transcription from P1 (exons A1a through 3) and P2 (exons 1 through 3) , respectively.

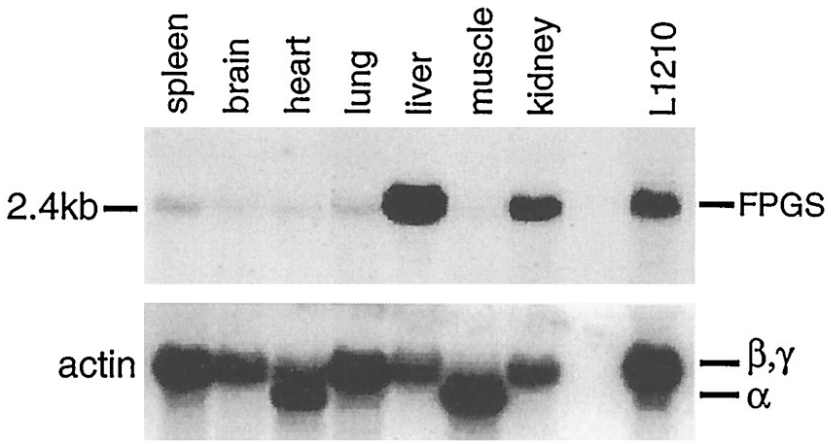


Figure 2-1. Northern analysis of FPGS gene expression in normal human and murine tissues. Total RNA (10 μg) from murine normal tissues from female DBA/2 mice was probed with a 1.7-kb downstream murine FPGS cDNA. **Freemantle, S. J. et al. J. Biol. Chem. 1997;272:25373-25379.**

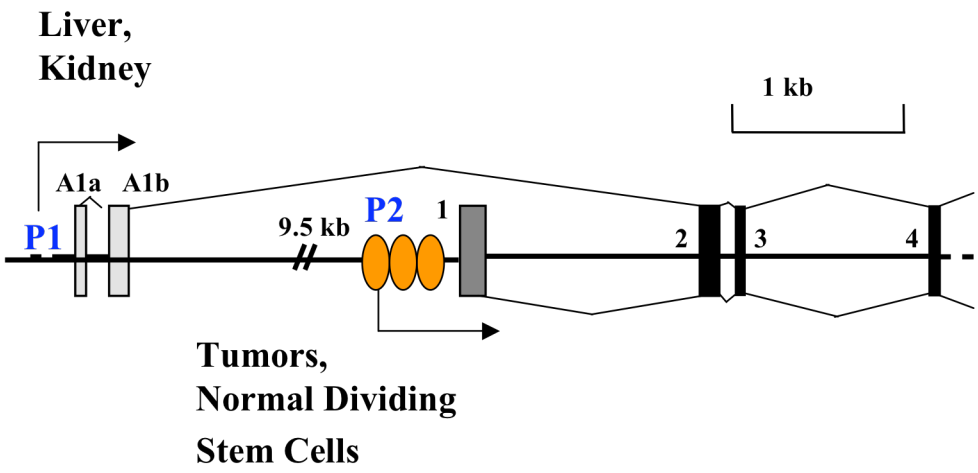
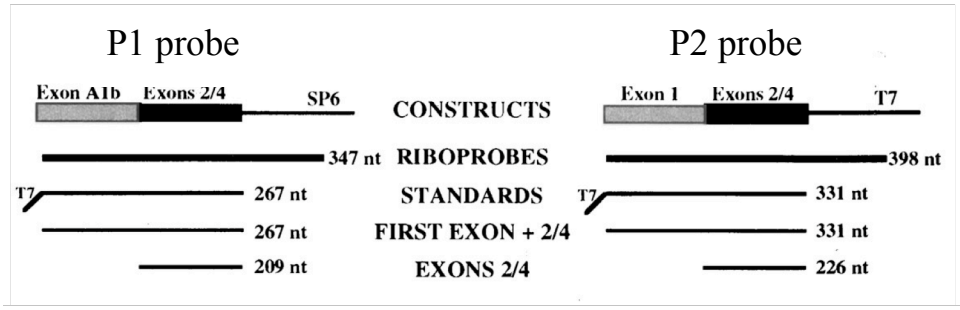


Figure 2-2. Schematic diagram of the tissue-specific expression patterns of the two promoters of the mouse *fpgs* gene. The two promoters of the *fpgs* gene are splaced 9.5 kb apart and are differentially activated; the P1 promoter is used exclusively in mouse liver and kidney and generates transcripts with exons A1a and A1b linked directly to exons 2-15, splicing out exon 1; the P2 promoter is used in all normal and neoplastic dividing cells and mature *fpgs* mRNA generated from P2 includes exon 1-15. The ovals at P2 mark the location of the three Sp1 sites regulating the P2 promoter.

A.



B.

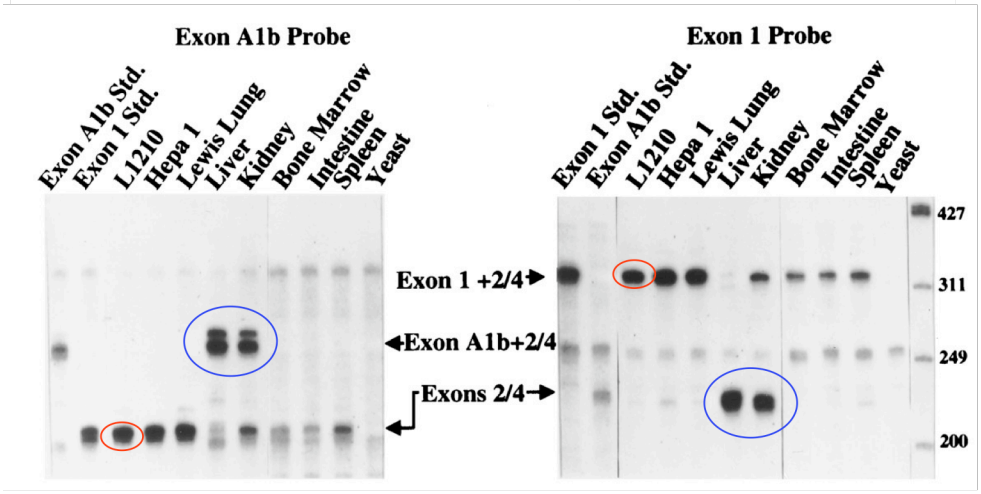


Figure 2-3. Tissue-specific use of P1 and P2 of the mouse *fpgs* gene. A. Schematic diagram of riboprobes, the PCR-generated RNA standards, and the expected protected fragment size. B. RPA using both probes on neoplastic and normal mouse tissues. Blue and red circles identify the fragments generated using mouse liver and L1210 cells, respectively. **Adapted from Turner et. al, Cancer Research, 59. 6074-6079, 1999.**

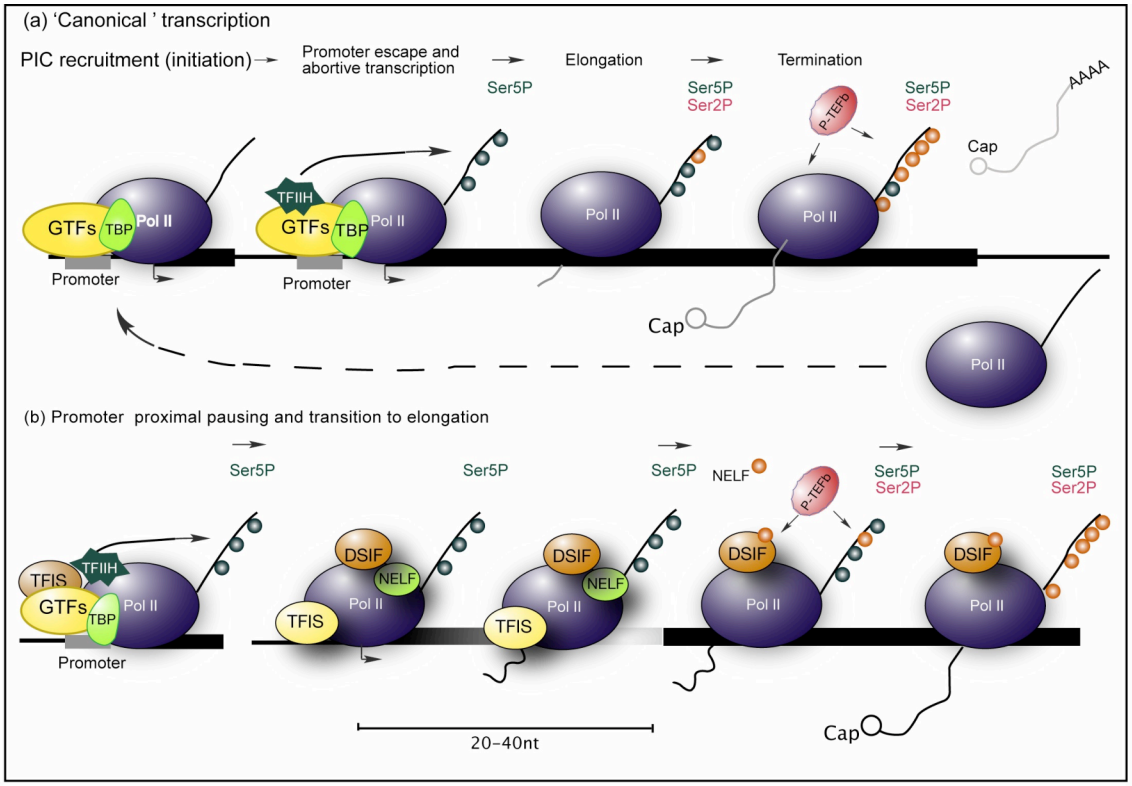


Figure 2-4. Stages of RNAPII-mediated transcription and promoter proximal pausing. **(a)** The phases of transcription: initiation, elongation, and termination. A pre-initiation complex is assembled with general transcription factors (GTFs) and RNAPII at the promoter region. Transcriptional elongation is marked by an the phosphorylation of the CTD of RNAPII. At the 3' end a termination signal is reaced and the mRNA molecule and RNAPII are released. **(b)** Steps of promoter-proximal pausing. In recent years, it has become apparent that a larger proportion of mammalian genes have a phase of promoter-proximal pausing during their transcription cycle. Adapted from Koch and Jourquin, Cell press, 2008.

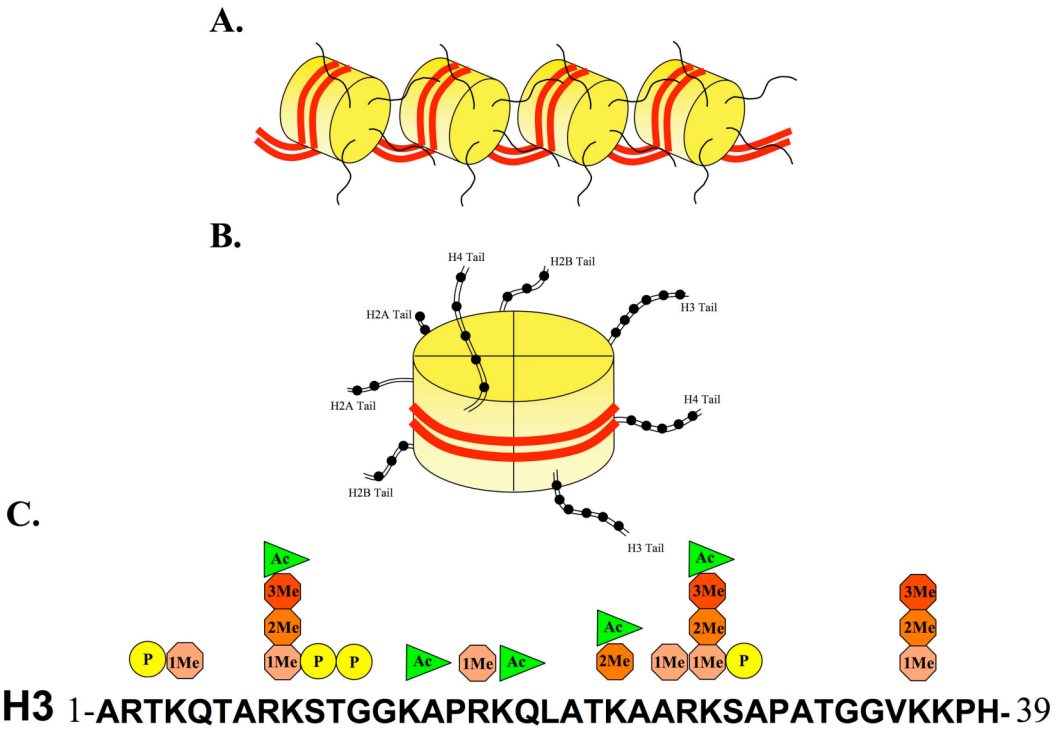


Figure 2-5. DNA wrapped around nucleosomes, histone tails, and histone tail post-translational modifications (PTMs). A. Nucleosomes (yellow cylinders) and DNA (red strands) interact to assemble chromatin. **B.** Two histone subunits of H2AB, H2B, H3, and H4 assemble into an octamer to form the basic unit of chromatin, the nucleosome. The N-terminal tail of each histone protrudes from the core of the nucleosome and key residues (black circles) serves as a substrate for histone modifying enzymes. **C.** A portion of the histone H3 tail with some of the potential PTMs. The potential modifications include: acetylation (Ac, green flag), mono-, di-, or tri- methylation (Me, octagon), phosphorylation (P, yellow circle)

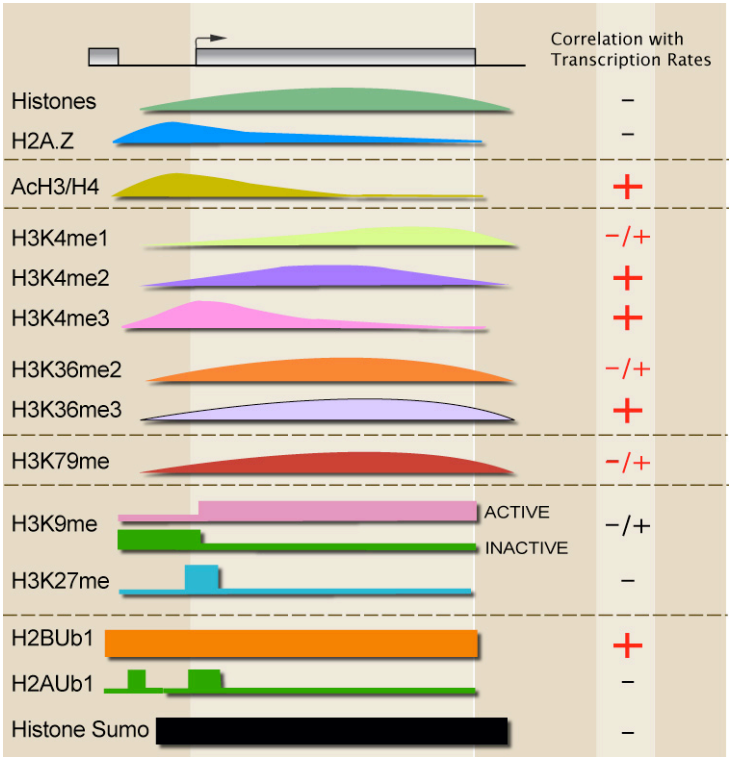


Figure 2-6. Composite profiles of the histone post-translational modifications (PTMs) studied across the mouse *fpgs* locus in L1210 cells, mouse liver, and mouse brain. The patterns are presented across a model gene and the correlation with active or repressed transcription rates is denoted by the + and – symbols. The presented profiles were obtained from genome-wide studies mainly in yeast, but H3K9 and H3K27 patterns were described in higher eukaryotes Adapted from Li and Workman, 2007 Cell.

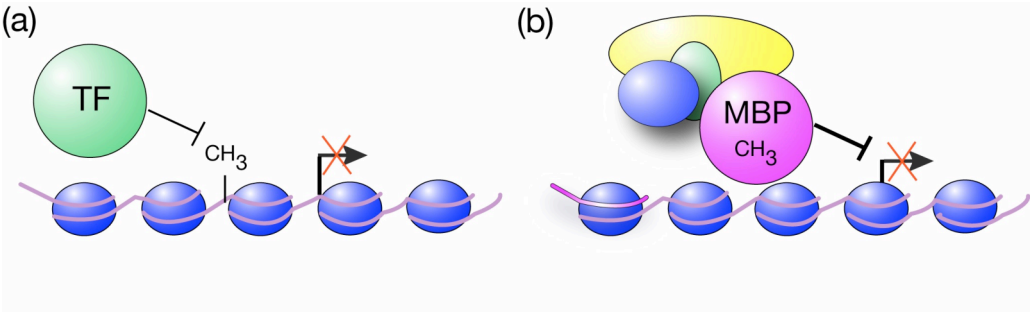


Figure 2-7. Mechanisms of DNA methylation-mediated repression. **A**. DNA methylation occludes *cis* elements from transcription factor binding. **B.** Methyl-binding proteins (MBPs) recognize methylated DNA and recruit histone deacetylases (HDACs) and histone H3 K9 methyltransferases. Adapted from Klose, Genomic Marks and Mediators 2006.

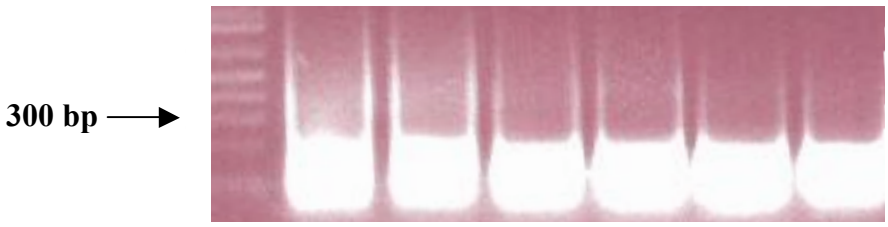


Figure 2-8. 1% agarose gel with sonicated material obtained for the ChIP studies presented in this chapter using L1210 cells, mouse liver, and mouse brain. Three hundered microliters of lysates were sonicated for 25 min (mouse liver and brain) and 20 min (L1210 cells). A 20 µl aliquot of sonicated material was incubated at 65^0C overnight followed by DNA purification and visualized on a 1% agarose gel stained with ethidium bromide. DNA fragments ranged from 100-300 bp on average.

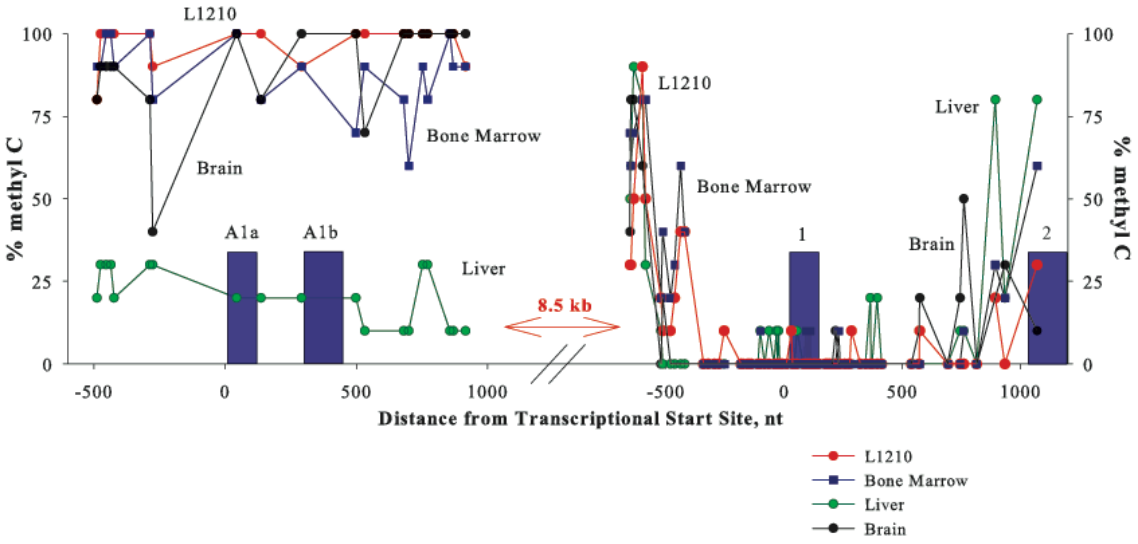
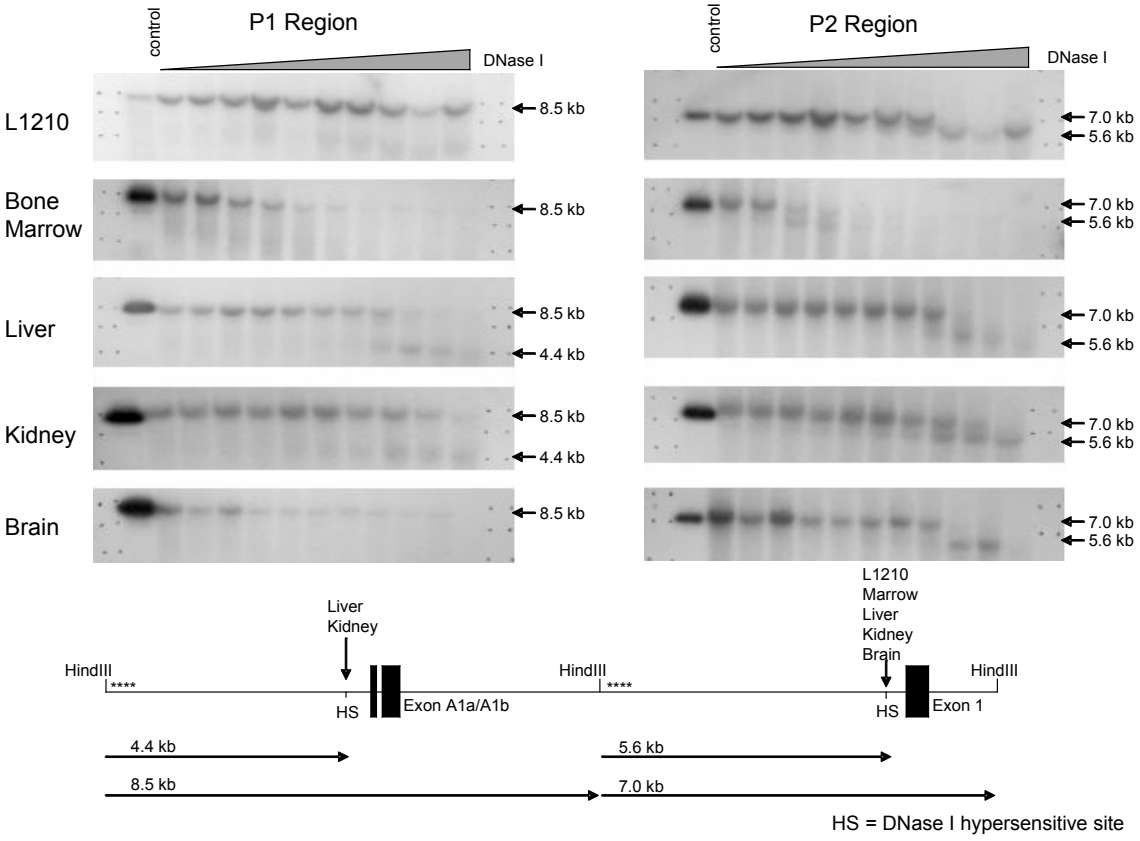
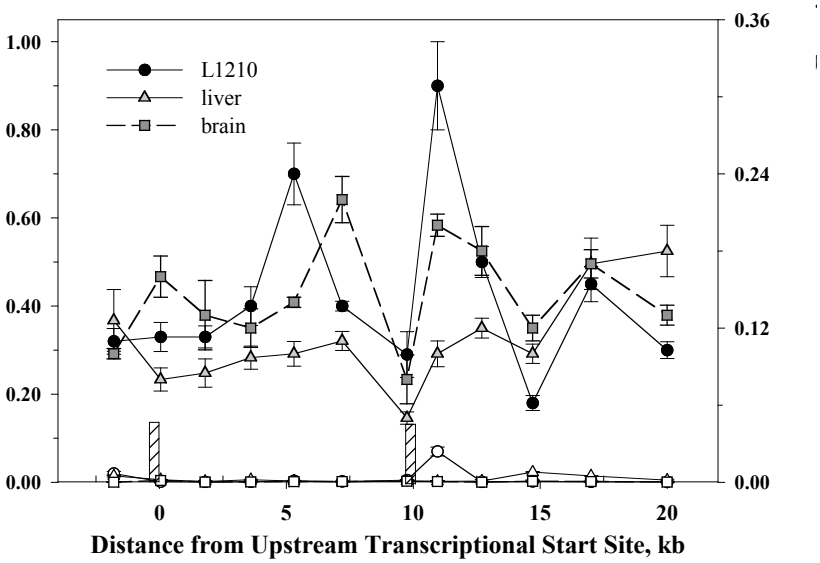


Figure 2-9. CpG methylation in the regions of the mouse *fpgs* promoters in various tissues. CpG methylation in the P1 promoter region correlates with transcriptional activity; the P2 promoter region is unmethylated regardless of activation. Genomic DNA was treated with 5M sodium bisulfite, amplified by PCR, cloned, and the sequences of 10 clones were determined for each region. At the position of each CpG dinucleotide in the genomic sequence, the percentage of clones containing a methylated cytosine is represented in the graph, plotted relative to the positions of exons A1a, A1b and 1 (filled boxes).

Figure 2-10. Chromatin density over the mouse *fpgs* locus. A. A prominent DNase I hypersensitivity site is present in the P1 region only in tissues in which the promoter is active; the P2 region is accessible to DNase I in expressing and non-expressing tissues. Nuclei were incubated with increasing concentrations of DNase I for 5 minutes at 25°C. Genomic DNA was extracted, digested with *Hin*DIII, and run on an agarose gel. Blots were probed with PCR-generated sequence at the 5' ends of the *Hin*DIII fragments (asterisks in lower diagram). The bands in the control lanes indicate the gel-mobility of the full-length *Hin*DIII fragments. The location of the DNase I hypersensitive sites (HSs) and the tissues containing such sites are indicated in the lower diagram. **B. ChIP determination of the level of histone H3 over the *fpgs* gene**. Chromatin from mouse liver (triangles), brain (squares), or L1210 leukemia cells (circles) was cross-linked, sonicated, and immunoprecipitated with either a pan-H3 antibody (filled symbols) or a non-specific IgG (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR using primers spaced 250-300 nt apart; symbols are placed at the middle of each PCR fragment. Input represents amplified product from 0.1% of the starting material. The positions of the two transcriptional start sites are shown as the cross-hatched bars. The experiments in this and all subsequent ChIP figures were performed at least twice, and the data are from a representative experiment.


control
P1 Region
DNase I
control
P2 Region
DNase I
L1210
Bone Marrow
Liver
Kidney
Brain
8.5 kb
4.4 kb
7.0 kb
5.6 kb
Liver
Kidney
L1210
Marrow
Liver
Kidney
Brain
HindIII
HindIII
HindIII
****
****
HS
HS
Exon A1a/A1b
Exon 1
4.4 kb
8.5 kb
5.6 kb
7.0 kb
HS = DNase I hypersensitive site

B.


L1210
liver
brain
1.00
0.80
0.60
0.40
0.20
0.00
0.36
0.24
0.12
0.00
0
5
10
15
20
Distance from Upstream Transcriptional Start Site, kb

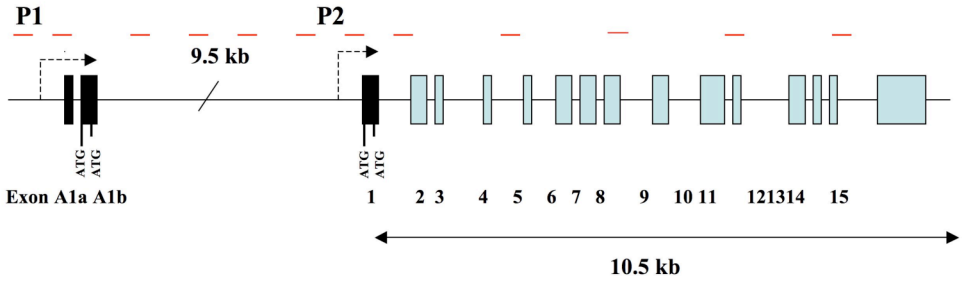
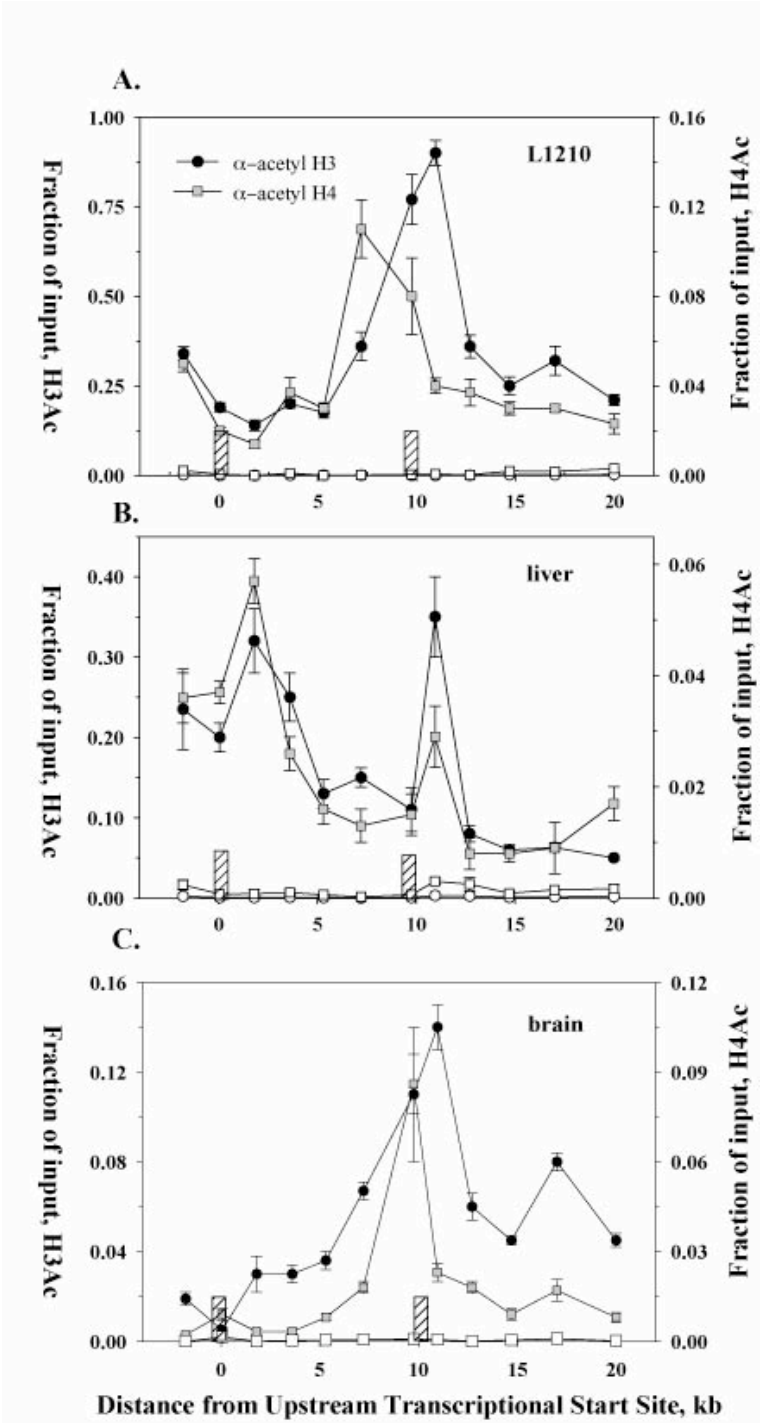


Figure 2-11. Schematic representation of amplified regions across the 20 kb of the mouse *fpgs* gene. The *fpgs* gene was divided into twelve fragments separated 1.5-2.0 kb apart**.** Each amplicon ranged between 150-300 bp in size.

Figure 2-12. Distribution of acetylated histone H3 and H4 along the *fpgs* locus in mouse tissues expressing transcript from P1 or P2. Chromatin from L1210 leukemia cells (A), mouse liver (B) and mouse brain (C) was cross-linked, sonicated, and immunoprecipitated with either an antibody against acetyl-H3 (circles), acetyl-H4 (squares), or a non-specific IgG (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR as in Figure 2B. The positions of the two transcriptional start sites are shown as the cross-hatched bars.


A.
L1210
α–acetyl H3
α–acetyl H4
Fraction of input, H3Ac
Fraction of input, H4Ac
B.
liver
C.
brain
Distance from Upstream Transcriptional Start Site, kb

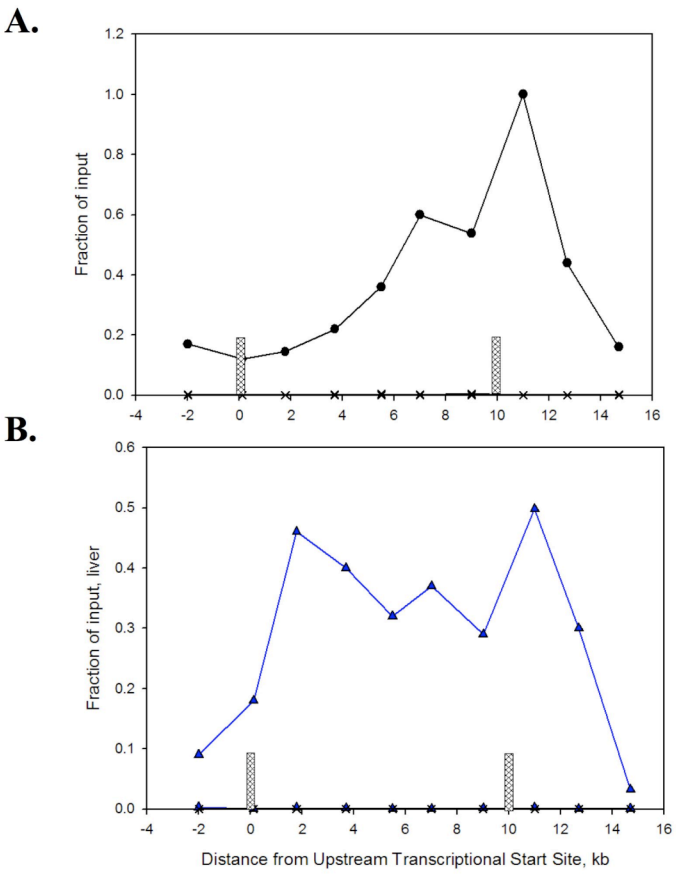
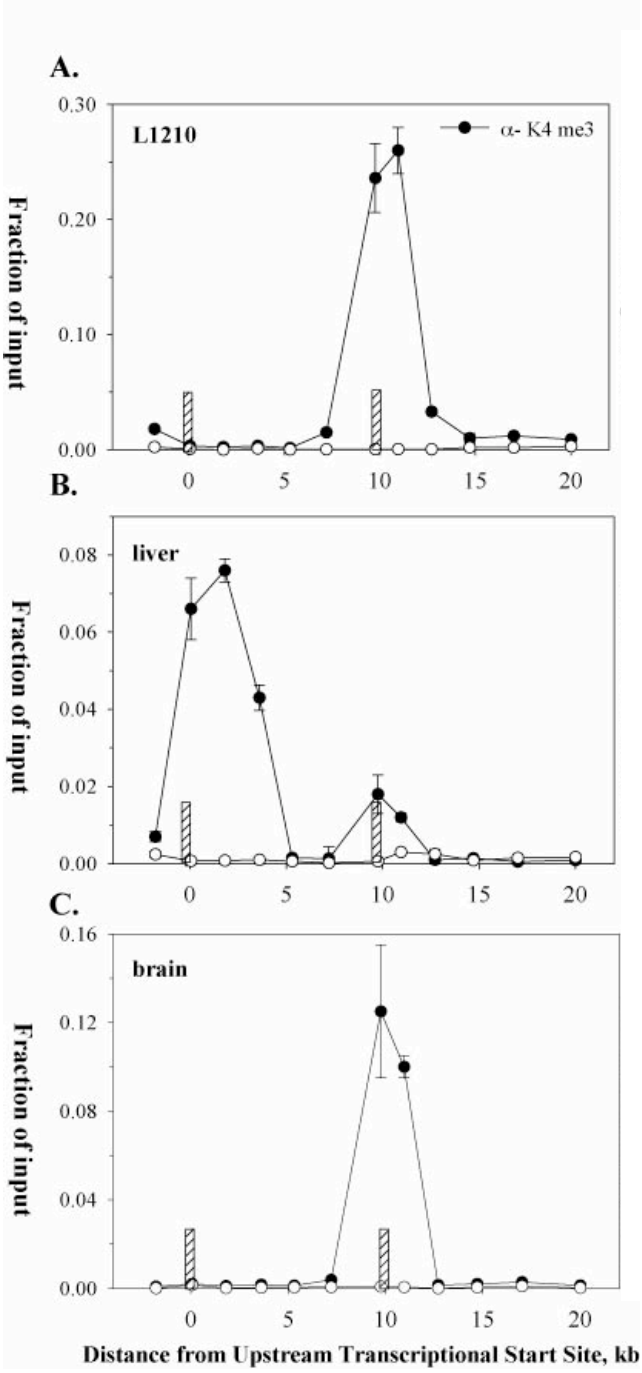


Figure 2-13. Analysis of the di-methyl histone H3 lysine modifications decorating the two *fpgs* promoter regions in mouse liver (B) and L1210 cells (A). ChIP analysis was performed using antibody raised against H3K4me2. Quantitation was performed using real-time PCR. Immunoprecipitation was also performed with an IgG antibody (X), as a non-specific control. The cross-hatched vertical bars represent the location of the P1 and P2 transcriptional start sites.

Figure 2-14. Distribution of histone H3 methylated at lysines 4 along the *fpgs* locus in mouse tissues expressing transcript from P1 (liver), P2 (L1210) or from neither promoter (brain). Chromatin from L1210 cells (**A**), mouse liver (**B**) and mouse brain (**C**) was cross-linked, sonicated, and immunoprecipitated with either an antibody against H3K4me3. For each immunoprecipitation, a separate non-specific IgG control was used (open symbols). The positions of the two transcriptional start sites are shown as the cross-hatched bars.


A.
L1210
α- K4 me3
Fraction of input
0.30
0.20
0.10
0.00
0
5
10
15
20
B.
liver
Fraction of input
0.08
0.06
0.04
0.02
0.00
C.
brain
Fraction of input
0.16
0.12
0.08
0.04
0.00
Distance from Upstream Transcriptional Start Site, kb

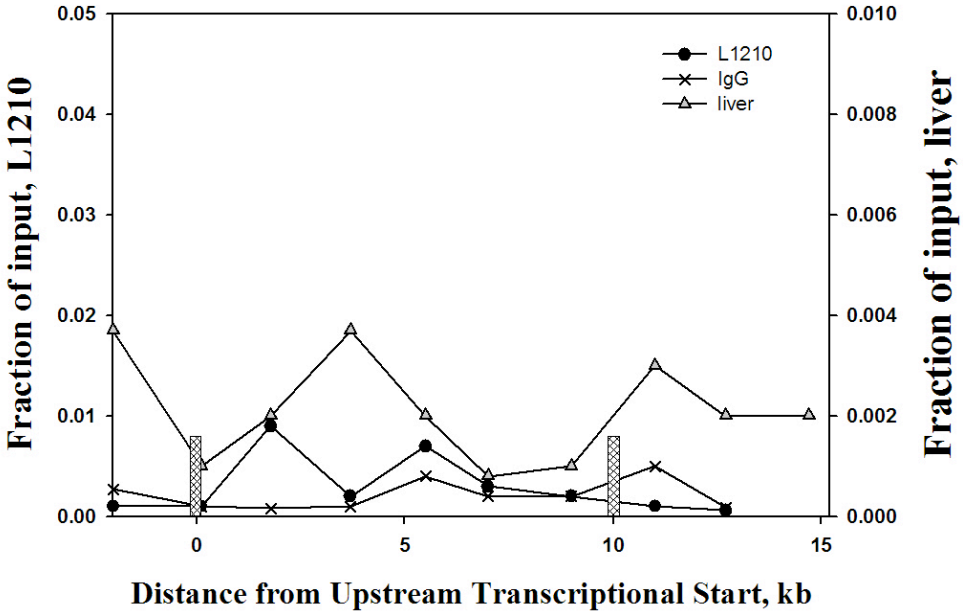


Figure 2-15. Analysis of the histone H3 K9 modifications decorating the two *fpgs* promoter regions in mouse liver (filled triangles) and L1210 cells (filled circles). The acetylation or tri-methylation of histone H3 K9 reflects the chromatin configuration and transcriptional activity of the P1 promoter. ChIP analysis was performed using antibody raised against H3K9me3. Quantitation was performed using real-time PCR. Immunoprecipitation was also performed with an IgG antibody (X), as a non-specific control. The cross-hatched vertical bars represent the location of the P1 and P2 transcriptional start sites.

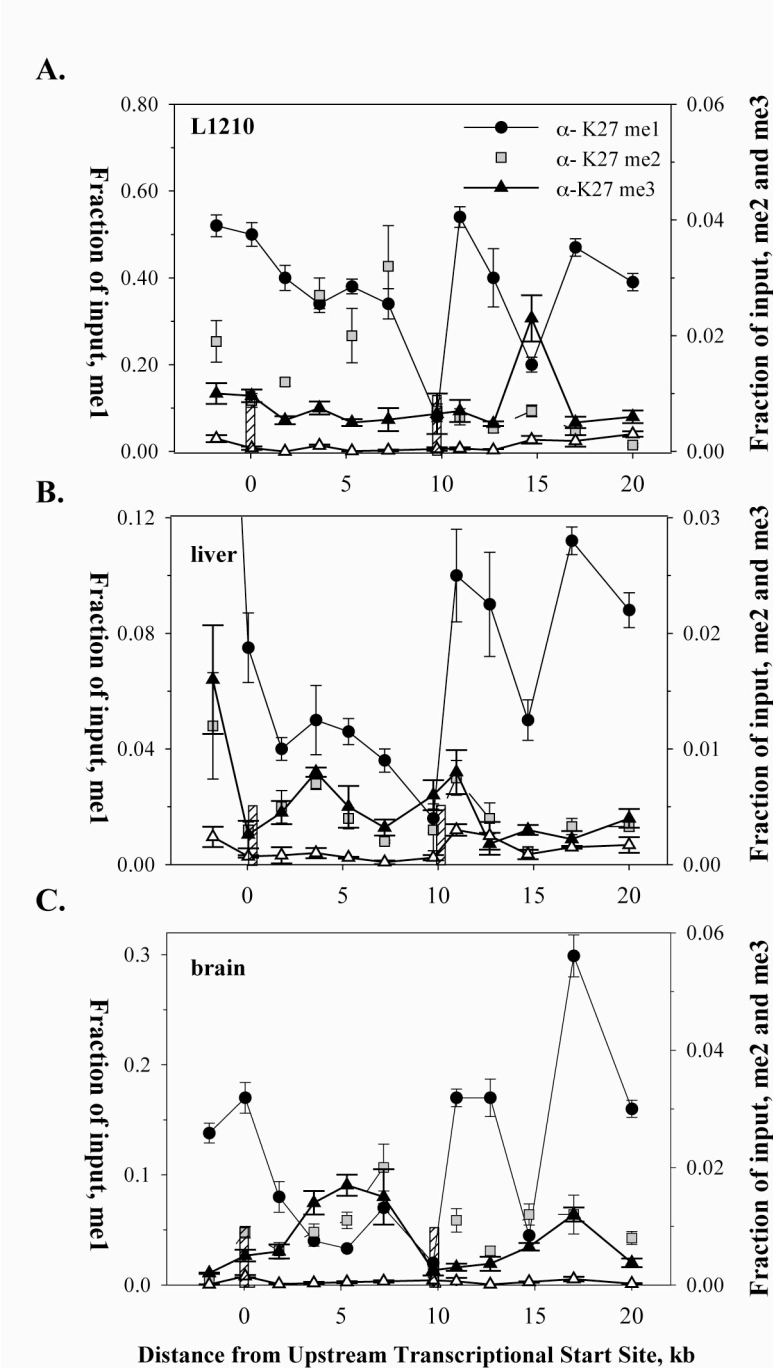
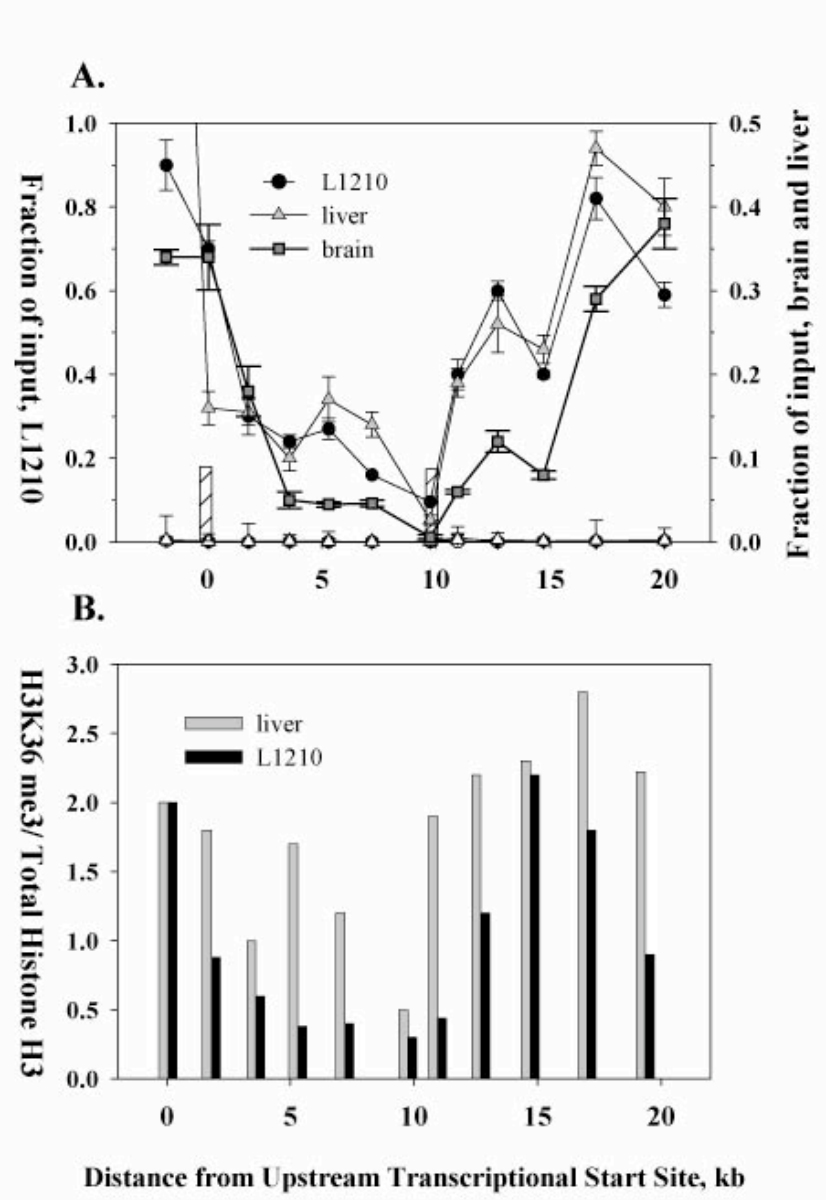


Figure 2-16. Distribution of histone H3 methylated at lysine 27 along the *fpgs* locus in mouse tissues expressing transcript from P1 (liver), P2 (L1210) or from neither promoter (brain). Chromatin from L1210 cells (**A**), mouse liver (**B**) and mouse brain (**C**) was cross-linked, sonicated, and immunoprecipitated with either an antibody H3K27me1, H3K27me2, or H3K27me3. For each immunoprecipitation, a separate non-specific IgG control was used (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR. The positions of the two transcriptional start sites are shown as the cross-hatched bars.

Figure 2-17. Distribution of histone H3 methylated at lysine 36 along the *fpgs* locus in mouse tissues. **A.** Chromatin from mouse liver (triangles), mouse brain (squares), and mouse L1210 leukemia cells (circles) was cross-linked, sonicated, and immunoprecipitated with either an antibody against H3K36me3 or IgG (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR as in **Figure 2B**. The positions of the two transcriptional start sites are shown as the cross-hatched bars. **B. The ratio of ChIP signal for H3K36me3 to that for total histone H3.** The amount of H3K36me3 is expressed as a ratio of the ChIP signal for total histone H3 to determine whether the minimum in the H3K36me3 signal seen in A reflects total H3 density or a lack of methylation at H3K36 on the regional nucleosomes.


A.
Fraction of input, L1210
Fraction of input, brain and liver
L1210
liver
brain
1.0
0.8
0.6
0.4
0.2
0.0
0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
B.
H3K36 me3/ Total Histone H3
liver
L1210
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
Distance from Upstream Transcriptional Start Site, kb

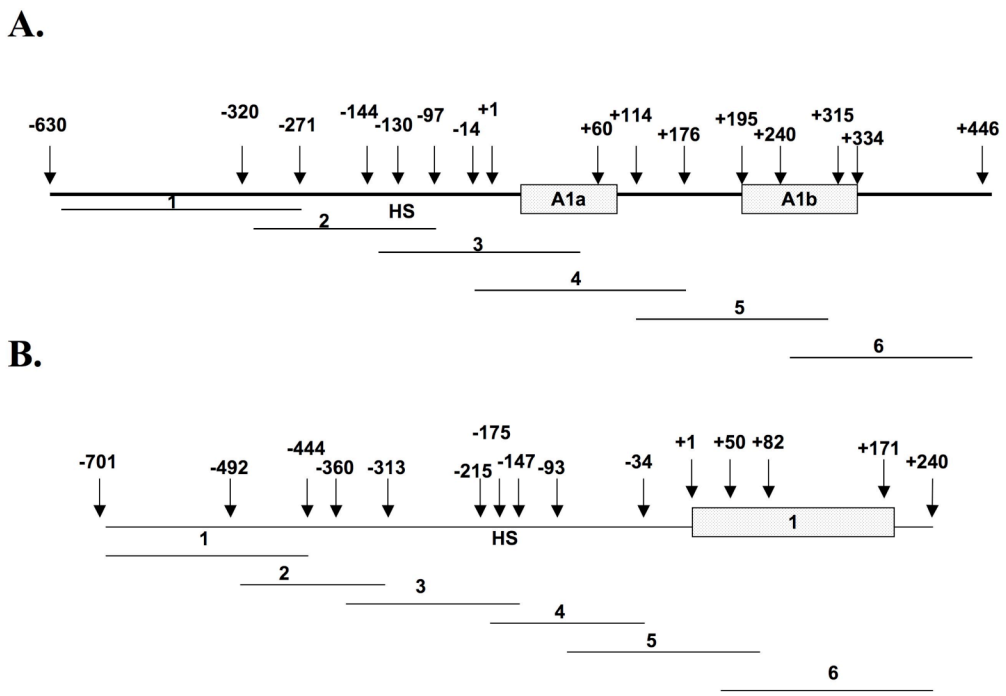


Figure 2-18. Schematic representation of overlapping primer sets used to amplify the P1 and P2 promoter in high-resolution ChIP experiments. The P1 promoter was divided in to six fragments covering the region between -630 and +315 nt, relative to the P1 promoter transcriptional start site. The P2 promoter was probed using six amplicons against the region between -701 and +240 nt, relative to the P2 promoter transcriptional start site. Adjacent products overlapped by *ca* 50 nt. The location of the DNase HS determined previously are labeled HS.

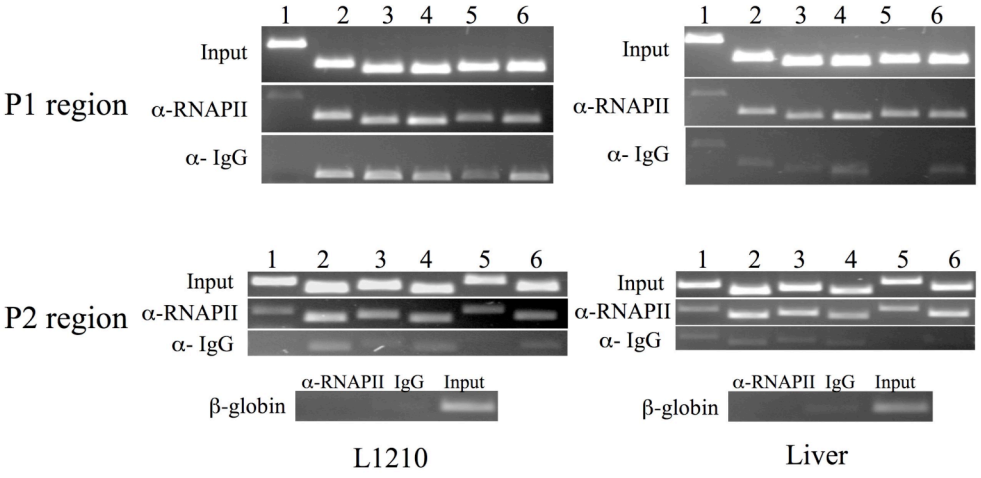


Figure 2-19. Semi-quantitative PCR identifies RNAPII across the P1 and P2 promoter regions in mouse liver and only at the P2 promoter in L1210 cells. High-resolution ChIP walking studies were performed using an antibody against RNAPII and IgG as a non-specific control. Six regions of genomic DNA, labeled 1-6, were amplified for both the P1 and P2 promoters. The location of each amplicon is illustrated in Fig. 2-18. 0.1% of total input DNA was added to PCR reactions using each primer pair. A portion of the β-globin promoter was amplified as a negative control for RNAPII binding. The reactions were run for 28 PCR cycles and 10 µl of product were run on a 1% agarose gel stained with ethidium bromide.

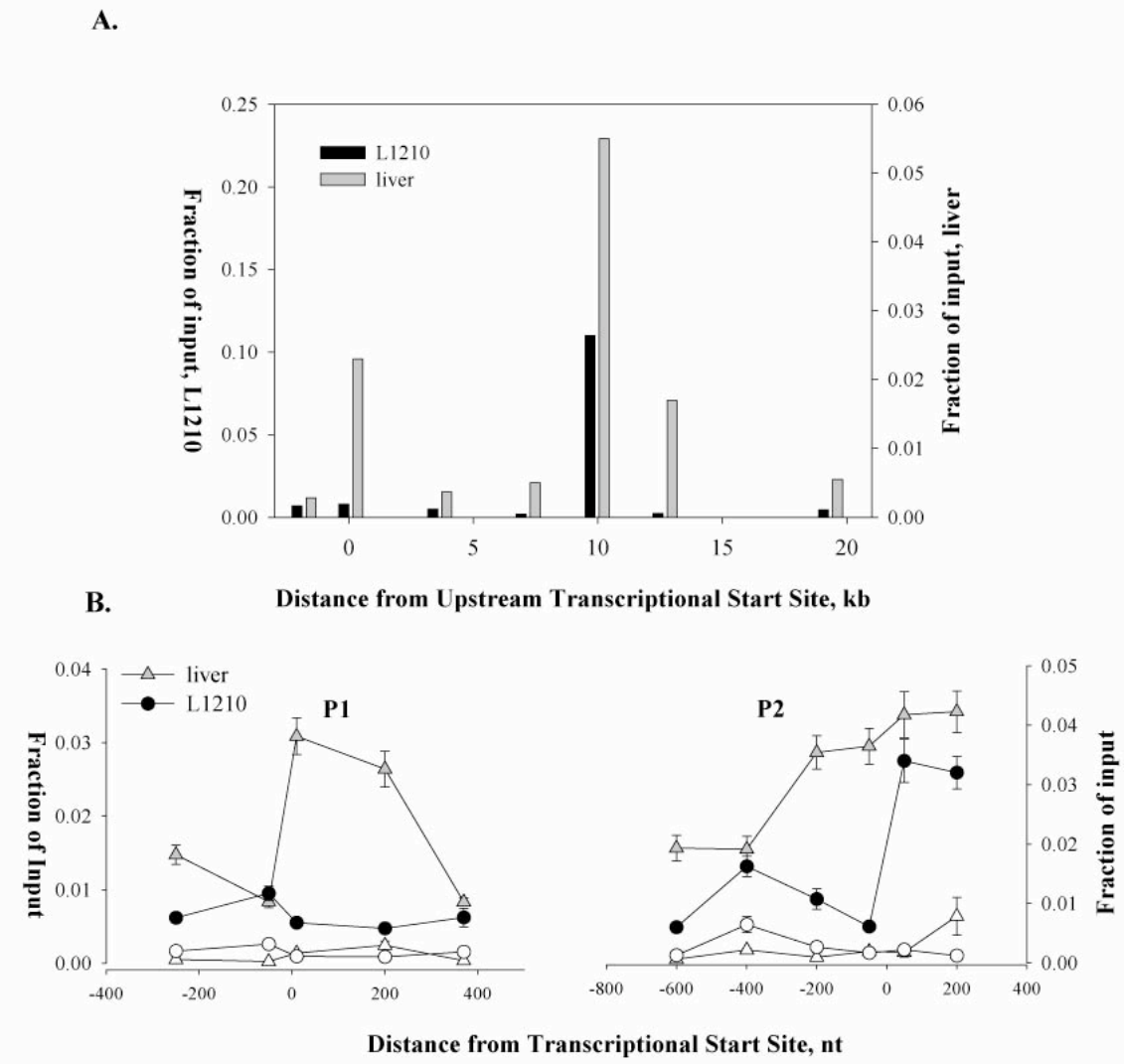
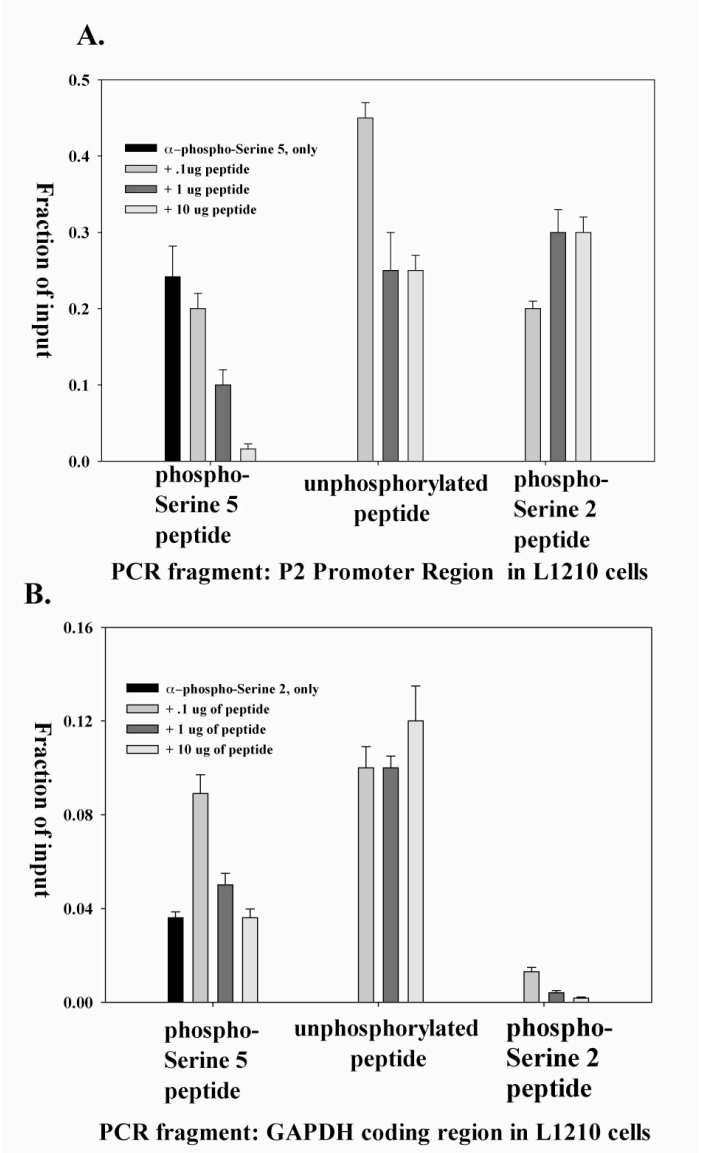


Figure 2-20. Residency of RNAPII over fragments of the two mouse *fpgs* promoters in mouse liver and L1210 cells. A. The amount DNA from regions spaced along the *fpgs* gene immunoprecipitated by an antibody to RNAPII was quantitated by real-time PCR as in **Figure 2B**. The position of P1 is plotted at 0 kb and P2 is at 10 kb on the abscissa. Data for L1210 cells is shown as black bars, for liver as grey bars. In **B**, the DNA regions immediately surrounding P1 (left panel) and P2 (right panel) were fine-mapped by ChIP using the same antibody. Non-specific binding was determined by immunoprecipitation with IgG, as shown by open triangles (liver) and circles (L1210). The primer set plotted at 0 in **A** is also plotted at 0 for the P1 blow-up in **B**. The primer set plotted at 10 kb in **A** is plotted at 0 for the P2 blow up in **B**. Data are plotted in **B** at the midpoint of each amplified region.

Figure 2-21. Phosphopeptides used to raise phosphospecific antibodies against the CTD of RNAPIII specifically block the ChIP signal under our conditions. Increasing concentrations of the phosphopeptides used to raise antibodies against RNAPII CTD phosphorylated serine 2 and RNAPII CTD phosphorylated serine 5 or the equivalent unphosphorylated peptide were incubated for 2 hours with antibodies prior to their use in ChIPs as described in the Materials and Methods. Subsequently, the antibodies were added to lysates of L1210 cells, and the DNA in precipitates was quantitated by real-time PCR using a set of primers directed against a region of the GAPDH promoter and against a region of the P2 promoter of L1210 cells, regions chosen for moderate signals against the phosphoserine 5 and phosphoserine 2 antibodies, respectively.


A.
0.5
0.4
0.3
0.2
0.1
0.0
Fraction of input
α–phospho-Serine 5, only
+ .1ug peptide
+ 1 ug peptide
+ 10 ug peptide
phospho-Serine 5 peptide
unphosphorylated peptide
phospho-Serine 2 peptide
PCR fragment: P2 Promoter Region in L1210 cells
B.
0.16
0.12
0.08
0.04
0.00
Fraction of input
α–phospho-Serine 2, only
+ .1 ug of peptide
+ 1 ug of peptide
+ 10 ug of peptide
phospho-Serine 5 peptide
unphosphorylated peptide
phospho-Serine 2 peptide
PCR fragment: GAPDH coding region in L1210 cells

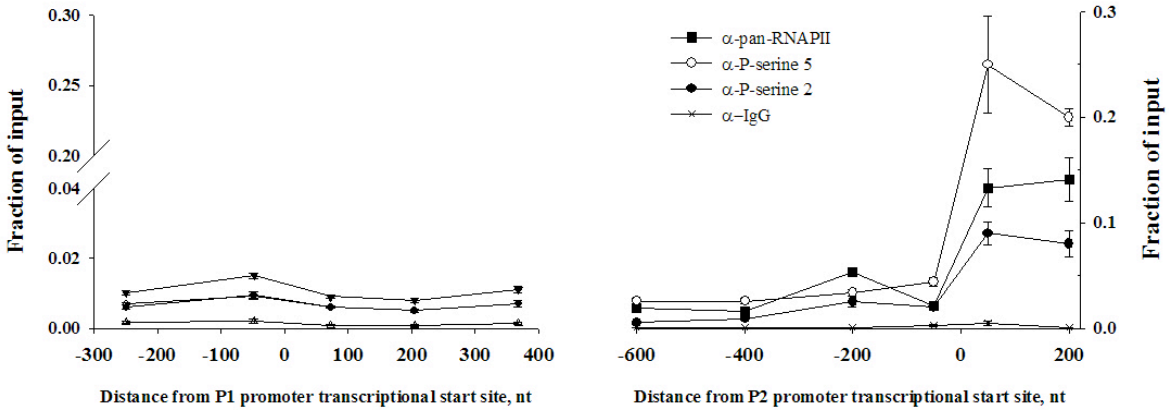
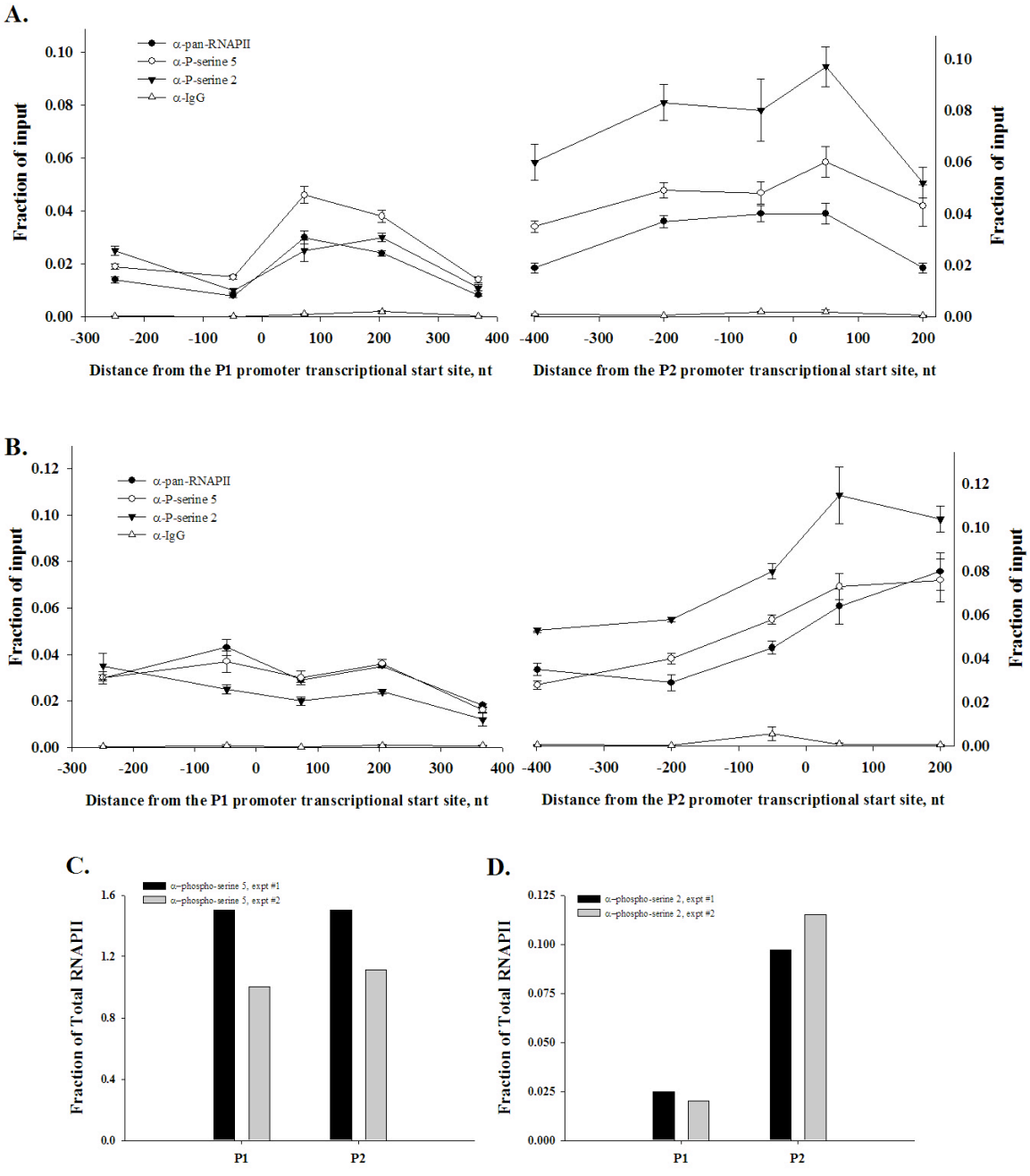


Figure 2-22. Phosphorylation of RNAPII at serine 2 and 5 detected over fragments of the P1 and P2 *fpgs* promoters L1210 cells. ChIP walking studies were performed using an antibody against RNAPII (filled circles), antibodies raised against the C-terminal domain heptad (YSPTSPS) repeat of RNAPII phosphorylated at serine 2 (inverted triangles) and serine 5 (open circles), as well as an antibody raised against IgG (open triangles) as a nonspecific control. Five or six fragments were amplified over the P1 and P2 *fpgs* promoters, respectively. Quantitation was performed using real-time PCR. The data points were plotted at the midpoint of each amplified region.

Figure 2-23. Phosphorylation of RNAPII at serine 2 and 5 detected over fragments of the P1 and P2 *fpgs* promoters in mouse liver. ChIP walking studies were performed using an antibody against RNAPII (filled circles), antibodies raised against the C-terminal domain heptad (YSPTSPS) repeat of RNAPII phosphorylated at serine 2 (inverted triangles) and serine 5 (open circles), as well as an antibody raised against IgG (open triangles) as a nonspecific control. Five or six fragments were amplified over the P1 and P2 *fpgs* promoters, respectively. Quantitation was performed using real-time PCR. The data points were plotted at the midpoint of each amplified region. The data from two separate experiments are presented in **A and B**. Ratios of enrichment signal of phospho-specific antibodies to total RNAPII signal were calculated for points at +75 nt and +50 nt for P1 and P2, respectively (**C and D)**.


A.
α-pan-RNAPII
α-P-serine 5
α-P-serine 2
α-IgG
Fraction of input
Distance from the P1 promoter transcriptional start site, nt
Distance from the P2 promoter transcriptional start site, nt
B.
α-pan-RNAPII
α-P-serine 5
α-P-serine 2
α-IgG
Fraction of input
Distance from the P1 promoter transcriptional start site, nt
Distance from the P2 promoter transcriptional start site, nt
C.
α–phospho-serine 5, expt #1
α–phospho-serine 5, expt #2
Fraction of Total RNAPII
P1
P2
D.
α–phospho-serine 2, expt #1
α–phospho-serine 2, expt #2
Fraction of Total RNAPII
P1
P2

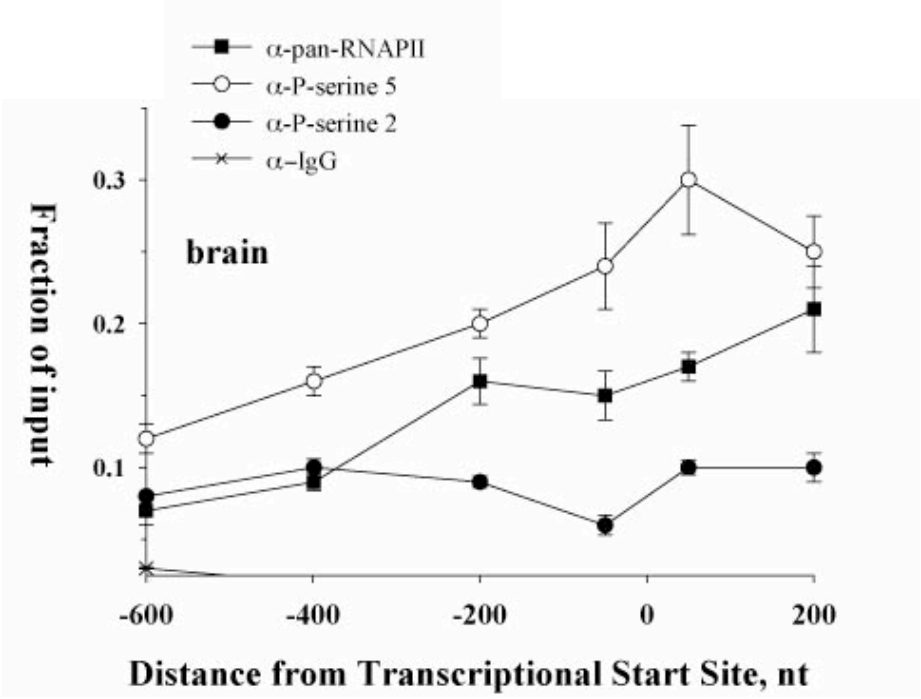


Figure 2-24. Phosphorylation of RNAPII at serine 2 and 5 detected over fragments of the P2 *fpgs* promoter in mouse brain. ChIP walking studies were performed using an antibody against RNAPII (filled squares), antibodies raised against the C-terminal domain heptad (YSPTSPS) repeat of RNAPII phosphorylated at serine 2 (filled circles) and serine 5 (open circles), as well as an antibody raised against IgG (X) as a nonspecific control. Quantitation was performed using real-time PCR. The data points were plotted at the midpoint of each amplified region.

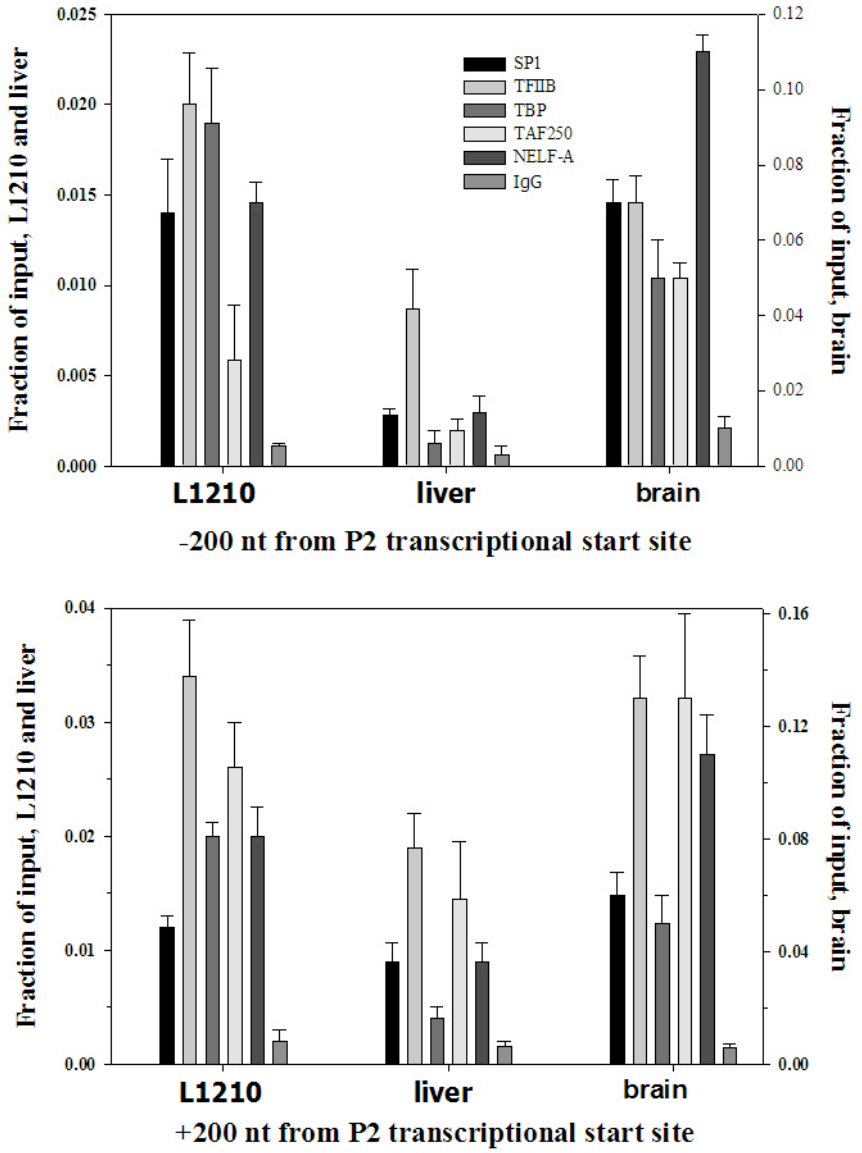


Figure 2-25. Presence of transcription factors over the mouse *fpgs* P2 promoter in L1210, liver and brain. ChIP analysis was performed using antibodies directed against the sequence-specific transcription factor Sp1, and general transcription factors TFIIB, TBP, TAF p250, and the NELF-A protein. Quantitation was performed by real time PCR using primers amplifying regions at -200 nt **(A)** and +200 nt **(B)** positions relative to the P2 transcriptional start site.

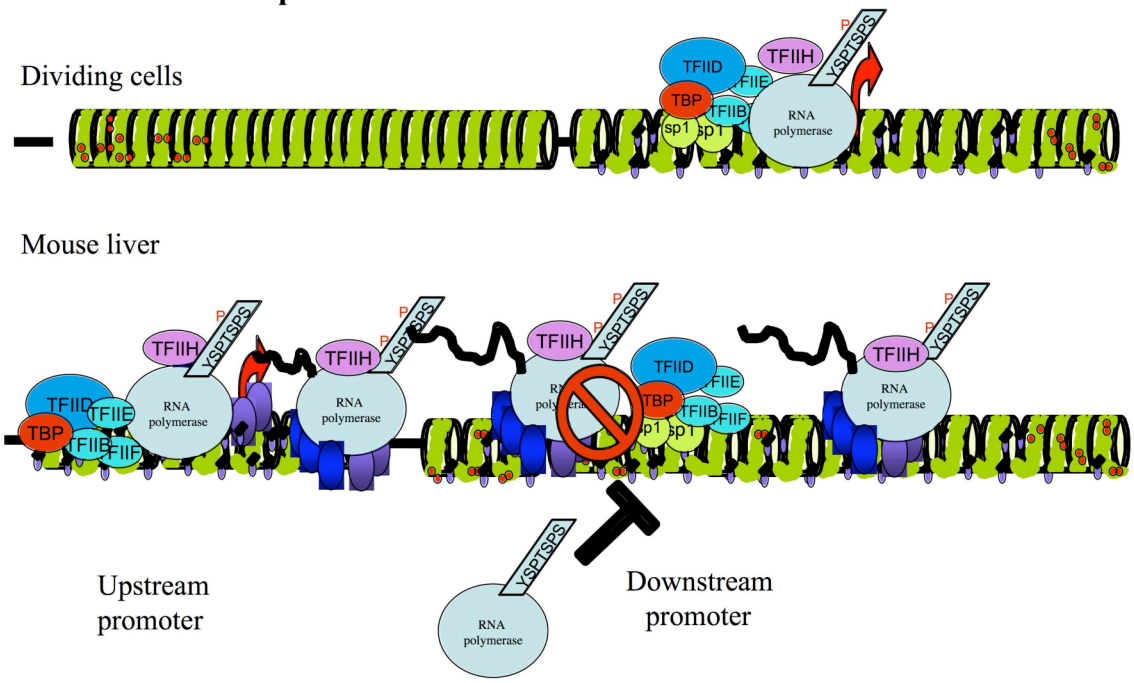


Figure 2-26. Schematic diagram of proposed mechanism of transcriptional interference occurring at the P2 promoter in mouse liver. In dividing cells, the initiation occurs at the P2 promoter and generates abundant *fpgs* transcript (Top). In mouse liver, transcriptional initiation occurs at the P1 promoter and elongation RNAPII complexes travel down the *fpgs* locus. We propose that RNAPII elongation complexes restricts assembly of a pre-initiation complex at the P2 promoter, thus preventing robust transcription from the P2 promoter in mouse liver.

Chapter 3: Disruption of individual steps in the control mechanisms of the mouse *fpgs* gene

INTRODUCTION

The studies presented in the previous chapter questioned how the two promoters of the mouse *fpgs* gene are regulated in a tissue-specific manner. Our data suggests that epigenetic mechanisms and transcriptional interference are involved in the control of the two mouse *fpgs* promoters.

A major challenge in further investigating these regulatory mechanisms is that the experiments needed require model systems where a primary change, i.e. loss of DNA methylation or loss of P1 activity, is established and the consequences on gene expression and epigenetic modifications at the mouse *fpgs* gene are measurable. As our initial studies were completed, two model systems became available that had single disturbances in the processes we were aiming to test: 1.) a knock-out mouse where the P1 promoter is completely removed was generated in our laboratory(263); and 2.) Cedar *et. al* engineered a mouse embryonic fibroblast immortal cell line that is severely depleted in DNA methylation (137).

Transcriptional Interference studies

In mouse liver, the upstream *fpgs* P1 promoter is active and the downstream P2 promoter is silenced. In previous studies, tracking the residence of RNAPII and general transcription factors across P2 in liver argued that the accumulation of elongation complexes restricted the assembly of pre-initiation complexes at this promoter. We took this to suggest that a mechanism of transcriptional interference at the mouse *fpgs* gene was involved in dictating the production of one isoform over a second in a tissue-specific fashion.

Transcriptional interference is a very broad term that includes any mechanism where the transcriptional activity of one gene impacts the activity at a second gene. TI mechanisms have been most extensively studied through engineered *in vitro* systems (1, 83, 188). Evidence from genome-wide studies supports the idea that TI may alter the expression of more genes than previously thought. Intergenic distances between genes have been mapped on human chromosomes 6, 20, and 22 and in many cases were found to be small, resulting in circumstances where regulatory regions frequently overlapped(233). Likewise, over 50% of mammalian genes have recently been identified to use multiple promoters(37). Whether and how TI is coordinating expression of either adjacent or multi-promoter mammalian genes are important questions that will require analysis on a case-by-case basis.

TI has been best illustrated *in vivo* at the *SRG1-SER3* locus in *S. cerevisiae.* The intergenic transcription of the *SRG1* gene represses the transcription of the adjacent *SER3* gene through direct TI (155). Transcriptional repression of the *SER3* gene was relieved in

SRG1 mutants or in mutants where a transcription termination signal was placed between the two genes (155). The TI mechanism was found to be promoter occlusion from studies showing that SRG1 activity severely limited transcription factor binding to the *SER3* promoter (155). Additional studies related this effect to amino-acid metabolism when it was found that *SRG1* expression was induced by high serine levels, which lead to the silencing of *SER3*, a serine biosynthetic gene (156). This was one of the first examples described *in vivo* where direct TI, or promoter occlusion, was proven to coordinate eukaryotic gene expression.

In our studies, we have used the mouse *fpgs* gene to study mechanisms of TI at an mammalian promoter *in vivo*. Lin-Ying Xie, a former student in the laboratory, generated a knock-out mouse line, where the P1 promoter and the two upstream exons A1a and A1b were deleted through homologous recombination (263). Homozygous deletion mice survived embryonic development, matured to adulthood, and reproduced normally (263). Our studies on the phenotype of these mice are not yet complete but, to date, the mice appear to be biologically normal.

In an effort to test our TI hypothesis, the address presented in this chapter compared expression levels and transcription factor binding at the P2 promoter in mouse liver harvested from wild-type and knock-out animals. We also used this mouse model to assess changes in the patterns of histone acetylation and histone H3 lysine 4 methylation across the mouse *fpgs* gene in the livers of genetically different mice. The P1 knock-out mouse model is a unique reagent allowing us to approach the mechanism of TI at an

endogenous locus by deletion of the proposed interfering gene; such studies have only previously been done in yeast models.

DNA methylation studies

DNA methylation analysis of the two mouse *fpgs* promoters showed that the CpG-sparse P1 promoter is differentially methylated between tissues, suggesting a role for DNA methylation in the regulation of the P1 promoter. Likewise, loss of DNA methylation was coincident with histone acetlyation and transcriptional initiation. These observations were truly a surprise since at the time we were unaware of any example described where loss of DNA methylation predicted the transcriptional activation of a tissue-specific CpG-sparse promoter. We were very interested to explore the possibility that DNA methylation was a primary determinant of transcriptional regulation at the P1 promoter.

Genome-wide studies have identified tissue-specific differentially methylated regions in both mouse and human (58, 227, 256). However, the primacy of DNA methylation in the regulation of tissue-specific transcriptional control remains controversial. The notion that CpG methylation may coordinate tissue-specific expression has been proposed for the past two decades (40, 235), but little evidence has supported this model (22, 206, 254). The methylation-mediated transcriptional repression of CpG-island promoters during the process of transformation has been well described (120, 131)but, tissue-specific promoters are now known to often be CpG-sparse and the ability for DNA methylation and methyl-binding proteins to promote silencing of this class of genes is poorly understood. Alterations of DNA methylation that parallel changes in expression at

tissue-specific and developmentally coordinated genes with non-CpG island promoters have been described in a few instances (75, 106, 112, 150). The maspin gene was one of the first cases identified *in vivo* where tissue-specific loss of DNA methylation resulted in production of mRNA. A case for the importance of this epigenetic mechanism in hematopoietic development has recently been made in several important studies at the mammalian globin locus. In this work, investigators tracked methylation and gene expression patterns in mice during development, which has lead to the proposal that DNA methylation and methyl-binding proteins may be involved in gene switching at this locus (106, 203).

Central to the interpretation of tissue-specific patterns of methylation is an understanding of how these genomic regions of differential methylation originate in developed tissues. The distribution of methyl groups at CpGs throughout the genome occurs at implantation with a phase of *de novo* methylation and is maintained throughout normal development and differentiation (193, 206). As such, loss of DNA methylation at a tissue-specific promoter must require a process of selective DNA demethylation. Whether loss of DNA methylation is a passive event, requiring multiple rounds of DNA replication, or an active process is a topic of controversy. However, there are an increasing number of examples found in the literature where rapid and cyclical changes in DNA methylation have been observed in mammalian cells in response to extracellular stimuli (126), leading to the proposal that this gene was regulated by an active process demethylation (32). Recent work illustrated an association between nucleotide-excision repair (NER) processes and genome-wide DNA demethylation (14, 213). However, the universality of NER

mechanisms in shaping the tissue-specific patterns of demethylation observed in mammals remains to be understood. Likewise, there is still a great deal to learn about how the timing and specifity of demethylation events are dictated, and whether or not DNA demethylation is a primary event in transcriptional activation. Our studies at the P1 promoter focus on the sequence of events leading to transcriptional activation, histone marks, and loss of DNA methylation at this region.

Methylation is the only known covalent modification of DNA in mammals and is catalyzed by a family of enzymes collectively known as the DNA methyltransferases (DNMTs) (131). These enzymes are classified based on a highly conserved catalytic domain at the C-terminus and the N-terminal domains of these proteins tend to be highly variable (131). To date, the identified mammalian DNMTs include: DNMT1, DNMT2, DNMT3a, and DNMT3b (131). Studies have determined that DNMT2 is not required for establishing or maintaining methylation patterns but instead, DNMT2 is a tRNA methyltransferase (80, 179). DNMT1 was the first enzyme purified in this family and was found to have activity towards nonmethylated and hemi-methylated DNA, with a preference towards a hemi-methylated substrate (21). Genetic deletion of the DNMT1 gene in developing mice leads to genome-wide loss of CpG methylation and embryonic lethality at E9.5 (142). Embryonic stem cells with DNMT1 mutant alleles retained detectable levels of DNA methylation, which lead to the proposal that additional methyltransferases were responsible for methylation in early stages of development (242). Subsequently, DNMT3a and 3b were cloned and characterized as the enzymes required for *de novo* methylation during development (177, 178). DNMT3L is a member of the *de*

novo methyltransferase family of proteins, which does not have any catalytic activity that cooperates with 3a and 3b to establish *de novo* methylation patterns (108). Current thought proposes that DNMT3a and 3b establish methylation patterns during development and DNMT1 uses a hemi-methylated substrate to maintain methylation in developed tissues (22, 131). How coordination of these enzymes affects tissue-specific patterns of methylation is not well understood.

A major challenge in the field of DNA methylation has been developing model systems to test the consequence of global loss of DNA methylation. It has been impossible to generate cell lines either from embryos of knock-out mouse models or in human carcinoma cells lines that are truly deficient in DNMT1. Previous literature predicts that if DNMT1 null cells survive in culture, they are not completely deficient in DNMT1, but instead are DNMT1 hypomorphs (44, 60). Additionally, in cells where the catalytic domain of DNMT1 has been conditionally deleted human carcinoma cells undergo mitotic catastrophe and substantial cell death within two days (44). While these issues highlight the importance of DNMT1 and DNA methylation for cellular survival, they have made it very difficult to ask fundamental questions regarding the consequences of primary changes in DNA methylation at specific CpGs.

An alternative approach to DNMT1 knockdown or knock-out animals or cell lines has been the use of nucleoside DNMT1 inhibitors 5-azacytidine (5aza) and 5aza-2-deoxycytidine (5daza) to pharmacologically induce loss of genome-wide patterns of DNA methylation (122). These compounds are cytidine analoges that are incorporated into DNA during replication. Binding of these analogues, once incorporated into DNA, to the

catalytic site of DNMT1 is covalent and pseudo-irreversible and traps the enzyme on the DNA, leading to depletion of cellular levels of DNMT1. As a result, these agents have been shown to cause demethylation and reactivation of many genes (60). While treatment with 5aza and 5daza has proven very useful, they have been shown to induce a DNA damage response, which leads to the expression of p53, p21, and numerous downstream targets (180). Distinguishing transcriptional effects caused by DNA damage vs. those resulting from a primary loss of methyl-CpGs is complicated by the fact that so many genes become activated during the DNA damage response.

In order to circumvent the problems generated by DNMT1 hypomorphs and inhibitors discussed above, Howard Cedar's group generated a fibroblast cell line that is homozygous for the DNMT1 mutant allele used in the original knock-out studies (142), and homozygous for a p53 mutant allele, resulting in a $p53^{-/-}Dnmt1^{-/-}$ genetic background. $p53^{-/-}Dnmt1^{-/+}$ animals were mated and the resulting embryos were harvested at E9.5, when the $p53^{-/-}Dnmt1^{-/-}$ embryos were identified by genotyping (137). In their report, they assayed global DNA methylation in these cells using nearest neighbor analysis and proposed that almost all CpG sites were unmethylated (137). This system was unique because the hypomethylated cells remained viable and proliferative in culture due to the loss of p53. Hence, these cells would support studies examining the primary effects of loss of DNA methylation on gene expression and chromatin structure, which had not previously been feasible. Likewise, a control cell line was easily available by generating mouse embryonic fibroblasts from $p53^{-/-}$ mice. With a primary change in DNA methyation, Cedar's group analyzed changes in histone acetylation and histone H3 lysine 4

methylation, as well as changes in gene expression. In these studies, they identified two classes of genes: those genes where loss of DNA methylation was coincident with an increase in gene expression, histone acetylation, and histone H3 lysine 4 methylation, and the second group of genes that required loss of DNA methylation and inhibition of histone deacteylation to show an increase in transcriptional activity. Some of the studies discussed in this chapter took advantage of this cell line to ask questions regarding the role of DNA methylation in the regulation of the P1 mouse *fpgs* promoter.

Objectives of study

Our previous work left us with a large number of unanswered questions regarding the histone PTMs, DNA methylation, and transcriptional interference involved in the tissue-specific regulation of the mouse *fpgs* gene. These questions include: 1. Is transcriptional interference the primary mechanism silencing P2 in mouse liver? 2. Is promoter occlusion the mechanism of transcriptional interference occurring at P2 in mouse liver? 3. Is DNA methylation necessary and sufficient for the silencing of P1 in dividing cells? and 4. Is DNA methylation across P1 necessary and sufficient to prevent histone acetylation and histone H3 lysine 4 methylation at nucleosomes surrounding P1 in dividing cell? The answers to these questions could not be pursued through observations of the wild-type mouse. In the studies presented in this chapter we used two model systems with primary disturbances in P1 activity and DNA methylation to explore the contribution of these factors in the use of the two mouse *fpgs* promoters.

MATERIALS AND METHODS

Materials

C57BL/6 mice were used in all the mouse studies and were originally obtained from Charles River Laboratories. P1 knock-out mice were back-crossed with wild-type C57BL/6 mice for six generations before they were used in the studies (263). Chemicals and reagents were purchased from Sigma (St. Louis, MO) and Fischer (Pittsburgh, PA). Antibodies were obtained from Upstate Biotechnology (Billerica, MA), Santa Cruz (Santa Cruz, CA), Abcam (Cambridge, MA), and Imigenex (San Diego, CA) and will be discussed individually in the sections to follow. 5-Aza-2'-deoxycytidine (A3656) and Trichostatin A (TSA) (T8552) were purchased from Sigma-Aldrich. Trizol reagent (Invitrogen, cat no. 15596-026) was used to isolate RNA. The puregene DNA purification system (Gentra, cat no. D-5000A) was used to isolate genomic DNA. EZ DNA Methylation-Direct Kit (Zymogen, cat. No D5020) was used for bisulfite conversion of genomic DNA. DNA oligo primers (25 nmoles) were purchased for these studies from Integrated DNA Technologies (IDT). Quanti-tect Sybr Green PCR Master Mix (product # 204143) was purchased from Qiagen (Valencia, CA) and used for all of the Real-time PCR analysis. Superscript III Reverse transcriptase First-Strand Synthesis System for RT-PCR (cat no. 1808-051) was used to generate cDNA. The Real-time PCR machines used in these studies were BioRad DNA Engine Peltier Thermal Cyclers with a Chromo 4 Real-Time Detector attachment. Opticon Monitor Software was used to analyze the Real-Time data.

Cell culture

P and PM MEFs were maintained in DMEM medium (Gibco/Invitrogen) supplemented with 15% Fetal Bovine Serum (FBS). Cells were grown at 37^0C in 5% CO_2.

Chromatin immunoprecipitation- P1 knock-out mouse liver studies.

The ChIP protocol described in chapter 2 for mouse liver was applied to profile histone H3 acetylation and histone H3 lysine 4 tri-methylation across the mouse *fpgs* gene and to assess Sp1, HNF4, and RNAPII binding to the P2 promoter in mouse liver from wild-type and P1 knock-out animals. Cross-linked DNA fragments were sonicated down to 100-300 bp and DNA size distribution was monitored on agarose gel. Lysates were rotated at 4ºC overnight with 8 µg of anti-RNAPII (Upstate Biotech, 05-623), 4 µg anti-IgG (Upstate Biotech, 12-371), or 5 µg of antibodies against phosphoserine 5 peptide from RNAPII (Abcam, ab5131), phosphoserine 2 peptide from RNAPII (Abcam, ab5095), HNF4 (Santa Cruz sc-8987), acetyl histone H3 lysine 9 (Upstate Biotech, 07-352), tri-methyl H3K4 (Upstate, 07473), or SP1 (Santa Cruz, sc-59). Real-time PCR was performed with 1 µl of final sample dissolved in 100 µl of TE. The primer sets used to cover the entire mouse *fpgs* gene and to amplify the P2 promoter at high-resolution were the same as those described in chapter 2.

RNA isolation from P and PM cells

Tissue-culture (100 mm) dishes were plated with P or PM cells and grown to 85% confluence. Pipettes and the workbench were wiped down with RNAse Zap solution

(Invitrogen) to limit the amount of RNAses in the environment. The protocol supplied with the trizol reagent was followed directly for the purification of RNA. The 100 mm dishes were placed on ice, medium was decanted, and 3 mls of trizol reagent were added to the plate. Cell lysates were taken up several times in a 1 ml filter pipette tip and transferred to a 14 ml polypropylene round-bottom Falcon tube (Falcon, cat no. 352059). The homogenized samples were incubated for 5 minutes at RT to allow complete dissociation of nucleoprotein complexes. Six hundred microliters of chloroform were added to each tube and the tubes were capped tightly and shaken by hand for 15 seconds, followed by 2-3 min incubation at RT. The samples were centrifuged at 12,000 x g for 15 min at 4^0C. The aqueous phase was transferred to a fresh Falcon tube and 1.5 mls of isopropyl alcohol was added to each tube to precipitate RNA. Samples were incubated for 10 minutes at RT followed by centrifugation at 12,000 xg for 10 minutes at 4^0C. A transluscent pellet was visible at the bottom of the tube and the supernatant was carefully removed by decanting to avoid disruption of the pellet. The RNA pellet was washed by adding 3 mls of 75% ethanol in diethylpyrocarbonate (DEPC)-treated water and the tubes were vortexed and spun down at 7500 *x g* for 5 minutes at 4^0C. The supernatant was again decanted and inverted on the workbench to remove excess ethanol. The pellets were air dried for no more than 2 minutes and resuspended in 100 µl of DEPC-treated water and transferred to a RNAse-free 1.5 ml microcentrifuge tube. The samples were passed through a pipette tip to ensure that the RNA was completely resuspended. If RNA was not completely in solution (flakes or pellet could still be visualized) the samples were incubated at 55^0C for 10 minutes. Total RNA was quantified using the NanoDrop ND-1000 spectrophotometer.

Five-hundred nanograms of each RNA sample were run on a 1% agarose gel stained with ethidium bromide at 100 volts for 30 min to assess the integrity of the RNA by visualizing the banding pattern of 28S RNA and 18S rRNA: if the 28S RNA band was approximately 2x more intense then the 18S rRNA, the RNA was considered to be intact. Purified RNA was stored at -80^{0}C in 3x volume of 100% ethanol or isopropanol and 1/10th the volume of NaAc at pH 5.2.

cDNA synthesis of RNA obtained from P and PM cells.

Total RNA (2.5 µg to 5 µg) from P and PM cells were used to synthesize cDNA with the Superscript III First-Strand Synthesis Reverse transcriptase Kit (Invitrogen). RNA was added to a RNAse-free 0.6 ml tube and combined with 1µl of 50µM oligo$(dT)_{20}$ primer, 1µl of 10mM dNTP mix, and DEPC-treated water up to 10µl. The mixture was incubated at 65^{0}C for 5 min to melt any RNA secondary structure. The oligo(dT) primer was used in these experiments to selectively convert polyadenylated full-length mature mRNA into cDNA. The samples were then placed on ice for at least 1 minute. In a separate 0.6 mL tube the cDNA synthesis mix was prepared by adding the following components, 20 µl of 10x RT buffer, 40 µl of 25mM $MgCl_2$, 20µl of 0.1M DTT, 10 µl of RNaseOUT (40U/µl), and 10µl of SuperScript III Reverse transcriptase (200U/µl) for 10 cDNA synthesis reactions. Ten microliters of cDNA synthesis mix was added to each RNA/primer mixture and mixed gently by flicking the tube. The samples were collected by brief centrifugation. The mixture was incubated at 50^{0}C for 50 minutes to allow first strand cDNA synthesis. The reaction was terminated through incubation at 85^{0}C for 5

minutes. Following termination, the samples were chilled on ice and were collected by brief centrifugation. One microliter of RNAse H was added to each tube and incubated for 20 min at 37^0C. This step removes the RNA template bound to the newly synthesized cDNA allowing for immediate use in PCR. In addition to the cDNA synthesis reactions carried out using total RNA, two control reactions were performed in parallel: 1) One reaction was prepared exactly as described above, except that only water was added to the primer mixture instead of RNA (no template control). This reaction controls for RNA or DNA contamination in the reagents of the kit. And 2) A second cDNA synthesis mixture was prepared without the addition of the SuperScript III RT (-RT cDNA synthesis mixture). Ten microliters of RNA/primer mixture was added to the –RT cDNA synthesis mixture and processed following the protocol outlined above. This reaction controls for the level of genomic DNA contaminating the RNA samples.

Primer design to amplify cDNA encoding mRNA generated at the P1 or the P2 mouse *fpgs* promoters.

The *fpgs* transcripts generated from the two promoters of the mouse *fpgs* gene are splice variants and we used this fact to design primers that would amplify one species over another. The sense primer was used to distinguish between the two types of transcripts: to select for P1-specific *fpgs* cDNA the sense primer was sequence within exon A1b (A1b forward), whereas the P2-specific *fpgs* cDNA was amplified using a sense primer positioned within exon 1 (Ex1 forward). The anti-sense primers used in these experiments were complimentary to sequences within exon A1b (A1b reverse), exon 3 (Ex3b reverse),

or exon 6 (Exon 6 reverse). The primers used in a given experiment are noted in the Figure Legends. The primers were designed, purchased, and resuspended as described in chapter 2.

Semi-quantitative PCR of cDNA generated from RNA of P and PM cells.

The semi-quantitative PCR experiments presented in this chapter were performed following the same procedure described in chapter 2 with the following exceptions: 1) 1µl of cDNA template was added to the 24 µl master mix. and 2) PCR reactions using primers specific to *fpgs* cDNA generated at the P1 promoter were removed at cycles 30, 32, 34, 36, and 38. GAPDH cDNA levels were used as a reference gene to control for variation in RNA extraction and cDNA synthesis between samples. PCR reactions amplifying GAPDH cDNA were removed at 8, 10, 12, and 14 cycles in order to visualize the production of products in the exponential range. The primers used to amplify the P1-specific *fpgs* cDNA in the semi-quantitative experiments were A1b forwarnd and Exon 3 reverse.

Real-time PCR of cDNA generated from RNA of P and PM cells.

Real-time PCR reaction mixtures were prepared as described in chapter 2 and 3µl of cDNA template was added to the 72 µl sub-master mix. Primers were designed to amplify *Slpi*, *cryaa*, and *GAPDH* cDNA in the real-time experiments. *Slpi* and *cryaa* were control genes for the P and PM cells and GAPDH was a reference gene used to compare between cDNA samples (discussed above). The primers used to amplify the *fpgs*-specific

cDNA species were A1b forward, A1b reverse, and Exon 6 reverse. The reactions were pre-incubated at 95^{o}C for 15 min and the amplification conditions for each primer pair were 95^{o}C for 45s; 58 ^{o}C for 45s, 72 ^{o}C for 45s followed by a final extension time at 72 ^{o}C for 5 min. Forty PCR cycles were performed for the real-time studies and a plate read step was programmed at the end of each cycle to capture the fluorescence in each tube. Melt curve analysis was programmed at the end of the run to measure melting curves of each PCR product in one degree increments between the temperatures of 50 ^{o}C- 100 ^{o}C. In PCR reactions where A1b forward and Exon 6 reverse were used the annealing temperature was set to 63 ^{o}C degrees. We found that amplification of P1-specific *fpgs* cDNA for forty cycles generated multiple sized products that encoded species of *fpgs* cDNA that would not generate a full-length protein due to the presence of pre-mature stop codons. These products were detected in melt curve analysis and by visualization on an agarose gel. By using an annealing temperature of 63 ^{o}C we were able to obtain single peak melt curves, which reflected the desired product as visualized on an agarose gel. In these experiments, a standard curve was not generated and the ΔCt method was used to analyze the real-time data (see below).

Analysis of Real-time PCR data

Real-time data was analyzed using the ΔCt method. This method allows relative quantification of a gene-specific cDNA relative to the level of the expression of a reference gene. In our experiments, GAPDH levels were used as the reference. As discussed in chapter 2, the threshold cycle number (C_T), is the cycle number at which amplified product

accumulates to a level where fluorescent signal is detected. The formula for the ΔCt method is

$$\text{Ratio} = 2^{\text{Ct(reference)-Ct(target)}}$$

The C_T value determined from the amplification of the cDNA of interest, e.g. P1-specific *fpgs* cDNA, was subtracted from the C_T value collected from reactions where GAPDH specific product was generated.

Chromatin immunoprecipitation-P and PM cells.

The ChIP protocol used in these studies was the same procedure described in chapter 2, with minor modifications as discussed below. P and PM were plated in 100 mm dishes and allowed to adhere to the plates overnight. The next day, cells were rocked for 8 minutes at RT with 1% HCHO in 10 mls of RPMI-1640 medium. Glycine was then added to the media to achieve a final concentration of 0.125 M and the cells were rocked for an additional 5 minutes. The media was aspirated from the plate and the cells were washed by the addition of 5 mls of PBS + 1mM PMSF to the plates. The PBS was removed and an additional 5 mls of PBS + 1mM PMSF was added to the plates. The cells were removed from the plate by scraping and cell suspension was then placed in a 15 ml centrifuge tube. The pelleted cells were washed and processed as described in chapter 2. Two million cells were resuspended in 300 µL samples of lysis buffer and sonication was carried out in a bath sonicator (Diagenode) to achieve 300-500 bp DNA. Additional sets of dishes were plated with P or PM cells when these experiments were initiated. The cells were trypsinized, pelleted by centrifugation, and were counted using the Beckman coulter

counter. The numbers obtained from these counts were used to estimate the number of cells that were cross-linked and lysed. Two million P and PM cells were sonicated using repeated cycles of 30 sec pulse and 30 sec off, followed by the addition of fresh cold ice water to the bath for a total sonication time of 25 minutes. Lysates were rotated at 4ºC overnight with 5 µg of antibodies against acetyl histone H3 lysine 9 (Upstate Biotech, 07-352), tri-methyl H3K4 (Upstate, 07473), or IgG (Upstate Biotech, 12-371). Real-time PCR was performed with 1 µl of final sample dissolved in 100 µl of TE. The primer sets used to cover the entire mouse *fpgs* gene and to amplify at high-resolution the P2 promoter were the same as those used in the studies described in chapter 2.

Treatment of P and PM cells with Trichostatin A (TSA).

P and PM MEFs were plated in 100 mm dishes and were treated when the plate was approximately 70% confluent. Ten milliliters of fresh media were added to the plates and half of the plates were treated with 0.06 µM of TSA. Three millimolar stock TSA solutions were prepared in ethanol. TSA was diluted in PBS and filter sterilized prior to use. The cells were exposed to drug for 12, 24, 48, and 72 hours and fresh drug was added every 24 hours. We found that, when MEFs were treated with TSA for more than 24 hours, the growth and viability of the cells were severely impacted. Hence, we focused our studies on 12 and 24-hour time points. At the indicated time points, RNA was isolated, cDNA synthesized, and analyzed by either semi-quantitative or real-time PCR.

Treatment of P and PM cells with 5-Aza-2'-deoxycytidine.

P and PM cells were plated in 100 mm dishes and were incubated overnight to allow cells to adhere to the plates. 5daza stock solutions were prepared in sodium phosphate buffer and resuspended to a final concentration of 1 mM, diluted in PBS and filter sterilized prior to use. Ten milliliters of fresh media were added to the plates and half of the plates were treated with 1μM 5daza for 24, 48, and 72 hours. Fresh media and drug were added every 24 hours. At the indicated time points, either RNA or DNA were isolated from the treated and untreated plates for future use in cDNA synthesis and bisulfite sequence analysis, respectively.

Bisulfite sequencing analysis

We used bisulfite conversion sequence analysis to quantify the DNA methylation present over the P1 mouse *fpgs* promoter in P and PM cells. Bisulfite treatment converts unmethylated cytidine (C) to uracil, but methyl cytidine is not susceptible to this reaction and remain unchanged. Amplification by PCR replaces the uracils with thymidines (T) and sequence analysis is used to determine which cytidine residues were methylated and unmethylated within a genomic region of interest.

Isolation of DNA

Genomic DNA was isolated from P and PM cells using the Puregene DNA purification kit. Cells were trypsinized and pelleted by centrifugation at 500 x g for 5 minutes. The supernatant was removed leaving 100-200 μl of residual liquid. The cell pellet was vortexed, resuspending cells to assist with lysis. Two to three million cells were

resuspended in 600 µl of lysis solution and pipetted up and down to generate a homogenous mixture. Cellular RNA was removed by the addition of RNAse A (3 µl of 4mg/ml stock) to the cell lysate. The sample was inverted 25 times and incubated at 37^{o}C for 30 min. The samples were placed on ice and cooled to room temperature and then were transferred to 14 ml Falcon tubes. Two-hundred microliters of protein precipitation solution was added to the RNAseA treated cell lysate. The samples were vortexed vigorously for 20 seconds and then incubated on ice for 5 minutes. The incubation on ice is not required, but I found that it helps the formation of a tight protein pellet after centrifugation. Following incubation, the samples were spun at 2,000 x g for 10 minutes. If the precipitated protein pellet was not tightly fixed to the bottom of the tube, the samples were vortexed and re-incubated on ice for 5 minutes followed by centrifugation. The supernatant was poured into a fresh Falcon tube containing 600 µl of isopropanol. The tubes were inverted gently fifty times followed by centrifugation at 2,000 x g for 15 minutes. The supernatant was poured off and the tubes were inverted on the bench top to drain additional liquid. The samples were washed in 600 µl of 70% ethanol, followed by centrifugation at 2,000 x g for 10 minutes. The ethanol was decanted leaving a small amount of residual liquid. The pellets were transferred to a 1.5 ml DNAse-free microcentrifuge tube. The samples were pelleted by centrifugation at 12,000 x g for 5 min. The remaining ethanol was removed and the pellet was allowed to dry on the bench top for 5 minutes. One-hundred microliters of DNA hydration solution (Tris-EDTA solution) was added to the tubes and the samples were placed at 65^{0}C for 1 hour and rotated end-over-

end overnight at RT. If genomic DNA was not completely in solution following overnight incubation, additional DNA hydration solution was added in 50 µl increments.

Bisulfite conversion

The EZ DNA methylation-Direct kit (Zymo research) was used for the bisulfite coversion of genomic DNA isolated from P and PM cells. Two-hundred to five-hundred nanograms of genomic DNA in a volume of 20 µl were added to 130 µl of CT conversion reagent solution to a 0.2 ml PCR tube. The tubes were placed in a thermal cycler (Biorad), which was programmed to incubate the samples at 98^{o}C for 8 minutes, 64^{o}C for 3.5 hours, and 4^{o}C for up to 20 hours. The zymo-spin IC column was placed in a collection tube and 600 µl of M-binding buffer was added to the column. The sample was loaded into the column containing the M-binding buffer. The column was capped and inverted several times. The sample was centrifuged at 10,000 x g for 30 seconds and the flow-through was discarded. At this point, the bisulfite converted genomic DNA was bound to the column. Desulphonation was carried out using 200 µl of M-desulphonation buffer and incubating the samples at RT for 20 minutes, followed by centrifugation at 10,000 x g for 30 seconds. The samples were washed with 200 µl of wash buffer to the column and spun down as described previously. The column was placed in a 1.5 ml microcentrifuge tube and the bisulfite converted DNA was eluted off the column using 10 µl of M-elution buffer. The tubes were centrifuged at 10,000 *x g* for 30 seconds to collect the converted DNA.

Bisulfite-Specific primer design

Bisulfite converted genomic DNA will be rich in adenine (A) and thymidine (T) bases, since all cytidine (C) residues will be converted to thymidines except for the few

that are in the context of a CpG dinucleotide and methylated. To design primers for PCR of bisulfite-treated DNA, I first converted the sequence of the genomic DNA of the P1 promoter *in silico* to reflect the bisulfite treated form. This made primer design and the subsequent sequence analysis very straightforward. The genomic sequence of the mouse *fpgs* gene was placed in a word document. The EDIT/REPLACE function was used to replace CG with XY. Then the EDIT/REPLACE function was used to replace C with T followed by replacement of the XY with CG. This generated a sequence where all cytidines were replaced by thymidines except for those in the context of a CpG dinucleotide. Primers were designed to amplify four 150 bp to 400 bp sized fragments from the -500 nt to +500 nt positions within the P1 promoter region relative to the transcriptional start site. The primers were 20 to 30 bases in length with Tm between 51^0C and 54^0C (the low Tm is a result of the AT-richness of the sequence) and did not include any CG residues (**this is an important point: if one includes CG residues in the primer design it is possible to preferentially amplify the methylated form of the desired product and overestimate the methylation status of a particular residue). We generated primers with at least 5-8 G residues to increase the Tm and the complexity of the primers in an effort to minimize non-specific binding. We positioned primers to generate small amplicons (< 400 bp) in order to increase the probability of generating an abundant PCR product, since fragmentation of genomic DNA does occur during bisulfite conversion. The fragments amplified were between -502 nt and -230 nt, -132 nt and +200, +264 nt and +560 nt, +609 nt and 1000 nt positions relative to the P1 promoter transcriptional start site. Primers were ordered and resuspended as described in chapter 2.

One microliters of bisulfite treated DNA were added to a 25 µl PCR reaction using the HotStarTaq (Qiagen) prepared as discussed in chapter 2. The AT-richness of the genomic DNA can decrease the specificity of PCR reactions. We used a touchdown PCR strategy to increase the specificity of the PCR amplifcation of the bisulfite converted DNA. Touchdown PCR enhances the specificity of the initial primer–template duplex formation by using initial annealing temperatures that are 5–10°C above the Tm of the primers. In the following cycles, the annealing temperature was decreased by 0.5^0C/cycle until the Tm of the primer is reached. This is followed by additional rounds of PCR using the annealing temperature of the Tm of the primer. The touchdown program used for all four primer pairs specific to bisulfite converted genomic DNA of the P1 promoter region was: 95^0C 15 min, 95^0C 30 sec, 62^0C->52^0C for 30 sec with a decrease of Tm 0.5^0C per cycle (20 cycles), 72^0C 30 sec, followed by 95^0C 30 sec, 52^0C 30 sec, 72^0C 30 sec, for 36 additional cycles, with a final extension step at 72^0C for 5 minutes. Ten microliters of reaction were visualized using 1% agarose gel stained with ethidium bromide. The touchdown PCR protocol generated single products at the correct size, which could easily be excised from the gel and purified.

Gel purification

The products of the bisulfite-specific PCR reaction were cut out of the agarose gel using a clean razor blade while visualizing the bands using a long-wavelength UV lamp. The gel slice was transferred to a pre-weighed 1.5 mL microcentrifuge tube and weighed again. The weight of the gel slice was obtained by subtracting the weight of the empty tube from the weight of the tube with the gel slice. The DNA was purified from the gel

slice using the Promega Wizard SV Gel and PCR Clean-up system (Promega). Membrane binding solution was added to the tube with the agarose gel slice (10µl/10 mg of gel slice). The samples were vortexed and incubated at 55^0C for 10 minutes, with frequent mixing, so that the gel slice was completely dissolved. The mixture was transferred to an assembled SV minicolumn and incubated for 1 minute at room temperature, to allow binding of the DNA to the column. This was followed by centrifugation of the column at 16,000 x g for 1 minute. The flow-through was discarded. Membrane wash solution (700 µl) was added to the column and spun down as above, followed by a second round of washing using 500 µl of wash solution. The flow-through was discarded and remaining ethanol was removed by a final spin at 16,000 x g for 1 minute. The column was placed in a fresh 1.5 mL microcentrifuge tube and 50 µl of Nuclease-free water was added to the center of the column and incubated for 1 min at room temperature. The DNA was collected by centrifugation for 1 minute at 16,000 x g in the microcentrifuge tube.

TOPO TA Cloning

The purified PCR products were cloned using the TOPO TA cloning system from Invitrogen. This system takes advantage of the fact that *Taq* polymerase adds a single deoxyadenosine to the 3' ends of PCR products. The vector has an overhanging 3' deoxythymidine residue, which allows efficient ligation of the PCR product into the vector. The plasmid vector pCRIV-TOPO was used for sequencing. Four microliters of PCR product were added to 1 µl of salt solution and Topo vector. The reaction was mixed gently and incubated for 5 minutes at RT and were then placed on ice. A reaction was also

prepared that only included vector, which served to control for random integration or re-ligation events that would result in false-positive colony formation after transformation.

Transformation of chemically competent TOP10 cells (Invitrogen)

The TOPO cloning reactions prepared in the last section were transformed into *E.coli*; uptake of the plasmid was detected through selection for antibiotic resistance. Blue-white screening was used to select colonies with plasmids that were successfully ligated. For each transformation one vial of TOP10 chemically competent cells (provided at a transformation efficiency of 1 x 10^9 cfu/ug of supercoiled DNA) and two Luna-Bertani (LB) agar plates formulated with 50 µg/ml of carbenicillin (Sigma) were used. SOC medium (Invitrogen) was thawed and warmed to RT. The LB plates were warmed at 37^0C for 30 minutes followed by the addition of 40 µl of X-gal (40 mg/ml) on each LB plate and re-incubation at 37^0C. The vials of one-shot TOP10 cells were thawed on ice. Two microliters of TOPO-cloning reaction for each PCR product was added to the vial, followed by gentle mixing. The TOP10 cells were incubated for 30 minutes on ice. The *E.coli* were heat shocked at 42^0C for 30 seconds to allow uptake of the recombinant plasmid. The tubes were transferred to ice and 250 µl of S.O.C. medium was added. The samples were shaken at 37^0C for 1 hour at 200 RPM to allow expression of the plasmid and 50 µl and 100 µl of each reaction (including the vector only control) were plated on LB-agar plates with carbenicillin and X-gal. The plates were incubated overnight at 37^0C. The transformation efficiency was controlled by for addition of 10 pg of puc19 plasmid to a vial of TOP10 cells and plating 10 µl of this mixture on an LB-agar plate.

Picking colonies

Blue-white screening increases the probability of selecting a colony that has plasmid with the desired insert: transformants with recombinant plasmid will have a disrupted lacZ-α gene and will not be able to metabolize X-gal, appearing as colorless colonies on the plate, whereas those that have been transformed with plasmid not containing an insert will metabolize X-gal and produce blue colonies. Ten white colonies were picked for each transformation reaction and placed in 4 mls of LB with 4 μl of carbenicillin (50 mg/ml stock). The bacteria were grown with shaking at 250 RPM at 37^0C overnight. Recombinant plasmid was isolated using the Wizard Plus SV Mini-Prep Kit according to the manufacturer's instructions (Promega, cat no A1460). Isolated plasmid from each clone was digested with *EcoRI* and the sizes of the digested products were visualized on a 1% agarose gel stained with ethidium bromide. Verified plasmid (100 ng/ul) was sequenced through the Nucleic Acids Research Facility (VCU) using the ABI 3700 Prism 96 Capillary sequencer. The sequence data was best when either vector specific M13 reverse primer or gene specific anti-sense primers were used.

Western blot analysis

Cellular lysate preparation

At the time of harvest, cells were washed with ice-cold PBS +1mM PMSF and removed from the plates using a cell scraper. Scraping was performed on ice. Cells were placed in a 15 ml conical and pelleted at 300 x g for 5 minutes at 4^oC, and washed again with 10 mls of ice-cold PBS + 1mM PMSF. Cells were again pelleted by centrifugation and resuspended in 1ml of ice-cold PBS + 1mM PMSF. The cells were transferred to a

pre-weighed 1.5 ml centrifuge tube and spun down at 12,000 x g for 5 minutes at 4^0C. The supernatant was aspirated and the cell pellets were weighed and lysed in 10x volumes of lysis buffer (62.5 mM Tris pH 6.8, 5% glycerol, 2% SDS, 5% 2-mercaptoethanol, and 1x protease inhibitor complete cocktail (Roche)). The samples were passed through a 21-gauge needle several times, and debris was pelleted by centrifugation at 10,000 x g for 5 minutes at 4^oC. This step was performed twice. We found that aspirating two times increased the amount of protein obtained from the lysates. The supernatant was aliquoted and stored at -80^0C. Protein concentrations were determined by the Biorad-Bradford protein assay according to the manufacturer's instructions against a standard curve generated by measuring the absorbance of known concentrations of bovine serum albumin at 595 nm.

SDS page and immunoblotting

Twenty micrograms of total cellular protein were diluted in 2x volume of Laemmli sample buffer (Biorad 161-0737). Lysates were boiled for 5 minutes followed by brief centrifugation to collect the samples. The pre-made SDS polyacrylamide gels (Biorad) were assembled in a electrophoretic gel box unit (100 cm x 6.5 cm), and running buffer (25 mM Tris base, 250 mM glycine, 0.1% SDS) was poured into the apparatus. The combs were removed and the wells were flushed with buffer to remove any debris. The samples and a pre-stained protein standard ladder (Biorad) were added to the wells. The total cellular proteins were resolved by electrophoresis on a 4-15% SDS polyacrylamide gel (Biorad) in running buffer. Gels were run for approximately 1 hr at 100 volts.

Proteins were transferred to an Immobilon polyvinylidene fluoride (PVDF) membrane (Millipore) in 1x Towbin buffer (25 mM Tris, pH 8.3, 192 mM glycine, 10% methanol, 0.04% SDS) using a semi-dry apparatus. Nine pieces of Whatman paper were soaked in Towbin buffer and placed into the bottom of the transfer unit. The PVDF membrane was soaked in 100% methanol and placed on top of the Whatman paper. The gel was placed on top of the PVDF membrane and covered with six additional pieces of Whatman paper. The transfer was run at 40 mAmps per gel for approximately 3 hours. Following transfer, the membranes were dipped in methanol and allowed to dry at room temperature for 15 minutes. Non-specific protein binding was blocked using StartingBlock Blocking Buffer (cat no. 37542, Pierce) for 1 hour at RT. Blots were washed with 0.05% TBS-T (0.5M Tris pH 7.5, 0.14M NaCl, 2.7 mM KCl, 0.05% Tween 20) three times for 5 minutes at RT. Primary antibodies were diluted in Startblocking buffer and incubated at 4^0C overnight. The primary antibodies used were DNMT3a (cat no. IMG-266A, Imgenex) at 1:250, DNMT3b (cat no. IMG-184A, Imgenex) at 1:250, and DNMT1 (ab19905, Abcam) 1:1000. Following incubation, the blots were washed three times in TBS-T for 5 minutes and membranes were incubated with goat-anti mouse (DNMT3 blots) or goat-anti-rabbit (DNMT1 blots) secondary antibody with horseradish peroxidase conjugate (1:10,000) for 1hr at room temperature (Pierce) in StartingBlock Buffer. The blots were again washed as described above. Chemiluminescence was detected using the SuperSignal West Pico and West Dura Chemiluminescent Substrate Kits (Pierce).

RESULTS

The studies presented in chapter 2 left us with a series of questions that could not be pursued through observations of the wild-type mouse. We have designed a series of experiments using two model systems with primary disturbances in P1 activity and DNA methylation to address these remaining questions.

Transcriptional interference studies:

The studies presented in this chapter addressing the contribution of transcriptional interference to the regulation of the P2 promoter of the *fpgs* gene in mouse liver have been pursued in collaboration with Lin-Ying Xie, a former student in the laboratory. She has generated the critical tool for these experiments: the P1 knock-out mouse (263).

Histone H3 acetylation is less abundant across the *fpgs* gene in knock-out liver.

We questioned if P1 activity was causative of histone H3 acetylation across P1 and P2 mouse liver by comparing the levels of histone H3 acetylation found in wild-type tissue with those detected in the P1 promoter knock-out mouse liver. ChIP analysis using an antibody raised against H3K9Ac was performed as and fragments of the *fpgs* gene were amplified using Q-PCR described in chapter 2. The amplicons were spaced 1.5-2.0 kb apart and the genomic region probed in these experiments spanned from -2 kb upstream to + 15 kb downstream of the P1 promoter transcriptional start site. Amplification of P1 in wild-type animals was not included in these studies, since these sequences were not present in the knock-out animals. The degree of H3K9 acetylation throughout the *fpgs*

gene was markedly lower in liver from the P1 knock-out mouse than in wild-type mouse liver (Figure 3-1). This observation is in agreement with the concept that histone H3 acetylation is commonly localized at nucleosomes surrounding active or poised genes (87). Surprisingly, the relative patterns of H3K9Ac between the two samples were almost super-imposable, and the profile determined in knock-out mouse liver was significantly different than the patterns previously described in L1210 cells or mouse brain (Fig 2-12). It should be noted that in knock-out liver, there was a measurable increase in H3K9Ac starting 2 kb upstream of the P2 transcriptional start site and peaking at promoter, whereas in wild-type liver H3K9Ac between these two amplicons remained constant (Figure 3-1). We took this change to reflect the increase in transcriptional activity observed at the P2 promoter in the liver of P1 knock-out animals. These data suggested that P1 transcriptional activation caused the high levels of histone H3 acetylation across the *fpgs* gene in mouse liver, but it appeared that additional tissue-specific mechanisms were involved in dictating the pattern of histone acetylation found in mouse liver.

Histone H3 lysine 4 tri-methylation across the *fpgs* gene in knock-out and wild-type mouse liver.

An enrichment of histone H3 lysine 4 tri-methylation has been associated with nucleosomes surrounding promoter regions bound by RNAPII complexes engaged in active transcription or in a poised state in both mammals and yeast (13, 186). We probed the *fpgs* gene in wild-type and knock-out mouse liver for this histone PTM and found patterns not entirely predicted by previous studies. H3K4me3 was found to be higher over

the P2 promoter region in knock-out mouse liver than in wild-type tissue (Figure 3-2); this supports the observation that P2 expression increases in the P1 knock-out mouse liver. Abundant levels of H3K4me3 between -2 kb and + 2 kb relative to the P1 transcriptional start site were detected in wild-type mouse liver, but were completely lost with the removal of P1 activity in knock-out mouse liver samples (Figure 3-2), suggesting that H3K4me3 of the nucleosomes surrounding this region was directly linked to the level of P1 transcriptional activity. The substantial level of H3K4me3 detected at + 3.5 kb relative to the P1 transcriptional start site in knock-out mouse liver was very surprising and implied that RNAPII complexes were binding to this region of the *fpgs* gene (Figure 3-2). We considered the possibility that loss of P1 activity permitted transcriptional initiation at a cryptic promoter in this tissue not previously identified through 5' RACE and RPAs (202, 244).

P2 expression and levels of the general transcription factor Sp1 and of HNF4, a tissue-specific factor, at the *fpgs* gene in wild-type and knock-out mouse liver.

In the previous chapter, we proposed that the activity of the P2 promoter in mouse liver was restricted by transcriptional interference through a mechanism of promoter occlusion. Using Q-Rt-PCR, Lin Xie determined that P2 expression in knock-out mouse liver increased between 4 and 6-fold when compared to the level of P2 expression in wild-type mouse liver (263). These data represent the influence of transcriptional interference on P2 expression in mouse liver. Interestingly, the level of P2 expression detected is at least two orders of magnitude lower than the amount of P2-specific transcript generated in

actively dividing cells, i.e. L1210 cells. In order to confirm that promoter occlusion was occurring at P2 in mouse liver, we probed for Sp1 at positions -200 and + 200 relative to the P2 transcriptional start site in wild type and knock-out mouse liver. Sp1 residence at P2 was negligible when compared to the non-specific IgG control in wild-type liver (Figure 3-3 A). In contrast, amplification of these fragments of P2 was enriched in knock-out liver (Figure 3-3 A). These data, coupled with Lin Xie's analysis of expression directly supported the idea that P1 activity repressed the P2 promoter by preventing the assembly of a PIC, through a mechanism of promoter occlusion.

In silico analysis of the mouse *fpgs* gene revealed consensus DNA binding sequences for hepatic nuclear factor 4 (HNF4) within the regions of the P1 and P2 promoters. HNF4 is a tissue-specific factor involved in directing the cell-type specific expression patterns detected in differentiated hepatocytes. HNF4 occupancy has been shown to recruit HATs, chromatin remodeling complexes, and general transcription factor components of the PIC (215). The role of HNF4 in the regulation of the mouse *fpgs* gene has not previously been studied. As discussed in the preceding chapter, the P2 promoter in wild-type mouse liver is not completely silenced, but the level of expression is extremely low when compared to L1210 cells (< 0.75%). As a starting point to understanding if HNF4 was involved in regulating P2, we assessed the residence of this factor at P2 (Figure 3-3 B). We detected HNF4 at positions -200 nt and +200 nt relative to the P2 transcriptional start site at similar levels in both wild-type and knock-out mouse liver, both substantially above the non-specific IgG control (Figure 3-3 B). Additionally, since the levels of HNF4 between the wild type and knock-out samples were comparable it appeared

that the residence of HNF4 at the P2 promoter was not affected by the transcriptional activity of the P1 promoter. This was quite a surprise, since Sp1 occupancy at the P2 promoter was found to inversely correlate with P1 activity in mouse liver (Figure 3-3 A). How molecules of HNF4, but not Sp1, were able to bind to the P2 promoter in wild-type mouse liver is a very interesting question that remains to be understood.

Occupancy of RNAPII across the P2 promoter in wild-type and knock-out mouse liver.

In our previous work, we mapped the residency of RNAPII across the P2 promoter region in mouse liver, L1210 cells, and mouse brain using high-resolution ChIP walking. The profiles determined for the three tissues were remarkably different (Figures 2-22 thru 24). We were interested in determining if the increase of P2 transcriptional activity observed in knock-out mouse liver would alter the pattern of RNAPII binding across the P2 promoter. We used ChIP walking to assess the residence of RNAPII across this region in knock-out mouse liver (Figure 3-4). P2 was divided into six overlapping fragments as described in Figure 2-18. ChIP analysis was performed on liver from wild-type and knock-out animals using antibodies generated against total, phospho-serine 5, and phospho-serine 2 RNAPII. As discussed in chapter 2, antibodies generated against the CTD of RNAPII phosphorylated at either serine 5 or serine 2 are often used as markers of early and late stages of elongation, respectively (81, 133). Total RNAPII was again detected across the P2 promoter region at similar levels in wild-type mouse liver (Figure 3-4 A) as in previous experiments (Figure 2-23). The occupancy of RNAPII complexes

phosphorylated at serines 5 and 2 reflected the levels of total RNAPII detected across P2 in wild type tissue (Figure 3-4 A). ChIP using knock-out liver showed total levels of RNAPII gradually increasing across P2 and peaking between the +1 and +200 positions relative to the P2 transcriptional start site (Figure 3-4 B). The level of RNAPII complexes phosphorylated at Serine 5 also reached a maximum between the +1 and +200 positions, reflecting an increase in total RNAPII levels (Figure 3-4 B). In contrast, the amount of RNAPII phosphorylated at Serine 2 did not increase with total RNAPII and remained fairly constant throughout the entire P2 promoter region in knock-out mouse liver (Figure 3-4 B). Since loss of P1 caused an increase in P2 transcriptional activity, we took the profiles of RNAPII detected in knock-out liver to represent complexes engaged in initiation and the early stages of elongation. The patterns of RNAPII complexes detected in knock-out mouse liver were different than those determined across the active P2 in L1210 cells (Figures 3-4 B and 2-22). In fact, the profile was more closely related to those previously described in mouse brain (Figures 3-4 B and 2-22), a tissue where RNAPII complexes are poised for activation. The studies using tissues from the knock-out animals are preliminary, but it is interesting to consider that the level of transcriptional activity generated at P2 may determine the distribution of RNAPII across the region. Also, since the levels of P2 expression in knock-out mouse liver were orders of magnitude lower than those detected in L1210 cells, it may be the case that a proportion of the RNAPII complexes detected across P2 in this tissue were actually poised but inactive as in mouse brain.

Overall, our studies confirm that transcriptional interference plays a role in silencing P2 in mouse liver, but other mechanisms are most likely involved in enhancing P2 activity to the levels observed in L1210 cells. Sp1 binding to the P2 promoter increased substantially in mouse liver when P1 activity was eliminated. This is one of the first pieces of data directly supporting a regulatory mechanism of promoter occlusion at a mammalian gene *in vivo*.

DNA methylation studies

We previously identified the P1 promoter of the mouse *fpgs* gene as a tissue-specific differentially methylated region: the CpGs within this promoter were methylated in tissues where P1 is silent, i.e. L1210 cells, and hypomethylated in tissues where P1 is active, i.e. mouse liver (Figure 2-9). We took this to suggest a role for DNA methylation in the regulation of the tissue-specific expression patterns generated from the P1 promoter. In order to test this hypothesis we designed a series of experiments using a model system generated by Cedar *et. al* (137), a mouse embryonic fibroblast (MEF) cell line reported to have low levels of global DNA methylation due to the homozygous disruption of the DNMT1 endogenous loci (PM cells). These cells are p53-deficient, which enables these cells to maintain viability after several passages in culture. The control for these experiments was a MEF cell line that is p53-deficient (P cells).

Expression of control genes and activity of the P2 mouse *fpgs* gene promoter in P and PM cells.

In the study published by Cedar *et. al* using the PM cell line, two groups of genes were defined that appeared to be regulated by DNA methylation. Group A included genes that were transcriptionally silent in P cells but active in PM cells, where global DNA methylation is greatly suppressed. In contrast, the genes within group B were only expressed in PM cells that were treated with a histone deacetylase inhibitor, trichostatin A (TSA) (137). The investigators concluded that genes in Group A were primarily regulated by DNA methylation, but additional mechanisms were involved in controlling the expression of genes within group B, since demethylation was not sufficient to activate transcription from the genes in this population (137). Dr. Cedar kindly sent us a flask of P and PM cells and we initiated our studies by measuring the expression of a gene within either group A (*Slpi*) or group B (*Cryaa*) to repeat the work Cedar published in the original paper. RNA was isolated from P and PM cells and was converted to cDNA and the expression of *Cryaa* and *Slpi* was assessed using Q-PCR. The ΔCt method was used to analyze *Cryaa* and *Slpi* expression relative to the expression of our chosen reference gene, glyceraldehydes-3-phosphate dehydrogenase (GAPDH). The results from this initial study were quite surprising (Figure 3-5). *Slpi* expression increased in PM cells, when compared to P cells by approximately 9-fold (Figure 3-5), as predicted by the fact that this gene was placed within group A. However, the increment of expression observed was much less than the 50-fold change reported previously for *Slpi* by Cedar's group. The *cryaa* gene was originally classified as a group B gene by Cedar et. al and thus we expected that little

to no change in expression levels would be detected between P and PM cells. In striking contrast, *cryaa* expression dramatically increased over 2000-fold in our experiments (Figure 3-5). These observations have been repeated, the primers have been checked, and the PCR products visualized on an agarose gel stained with ethidium bromide to verify that the product generated was the size predicted by the genomic sequence. Additionally, parallel experiments performed by Erica Peterson in Dr. Taylor's laboratory confirmed that the PM cells obtained from Cedar had markedly lower levels of DNMT1 than the P cells, confirming that we were working with the correct cell lines (data not shown). We concluded that the *cryaa* gene in our hands was clearly not a gene that should be classified within group B, but rather represented a gene within group A.

In dividing mouse cells, *fpgs* mRNA is generated from the sole use of the P2 promoter (244). We measured the level of *fpgs* cDNA complimentary to mRNA originating from transcriptional initiation at the P2 promoter in both P and PM cells. We applied real-time PCR and used the ΔCt method to analyze P2 expression relative to GAPDH (Figure 3-5). The level of P2 expression in the P and PM cell lines was found to be identical, suggesting that any transcriptional activity at P1 in PM cells was not interfering with initiation at P2.

Changes in P1 expression observed in the hypomethylated PM cell line.

To test our hypothesis that DNA methylation was involved in regulating the expression of the P1 promoter, we questioned if the global loss of DNA methylation observed in PM cells by Cedar *et. al* altered the expression of the P1 promoter. In the

experiments presented in Figure 3-6, cDNA encoding P1-specific *fpgs* mRNA was amplified using semi-quantitative PCR and real-time PCR analysis. Semi-quantitative PCR showed that P1-specific cDNA levels appeared at earlier cycles in PM cells than in P cells, suggesting that the amount of P1-specific cDNA in PM cells was slightly higher than in P cells (Figure 3-6 A). The amount of GAPDH cDNA determined by semi-quantitative PCR was very similar between P and PM cells, as judged by visualization of the products on an agarose gel stained with ethidium bromide (Figure 3-6 A). Real-time PCR analysis showed that in PM cells the P1-specific cDNA levels ranged from 2.5-5.0-fold higher than those detected in P cells (Figure 3-6 B). We used the two PCR approaches because we found that in P cells multiple products were generated at later PCR cycles. These products were clearly different sizes than the expected fragments, and sequencing of the products determined that they included varying segments of the intron between A1b and exon 1, and some of the products also included part of exon 1. All of the alternate species generated in P cells that were sequenced contained premature stop codons and were not predicted to generate full-length FPGS protein.

The alternate products posed a problem for analysis by real-time PCR, and therefore we designed primers within exon A1b to estimate of the difference between the levels of P1-specific *fpgs* cDNA in P and PM cells. This approach was flawed for a couple of reasons: 1) It is always a risk to amplify cDNA using primers within the same exon, since even small levels of contaminating genomic DNA can substantially influence the data. And 2) The multiple pseudo-products generated from P1 in P cells were still measured using this approach, and thus the levels calculated in P cells were most likely an

over estimate of the levels of correct *fpgs* cDNA transcript generated from P1 in P cells. However, we have verified the real-time data using semi-quantitative PCR and in later studies incorporating the use of an anti-sense primer complimentary to exon 6, which generated single products when the annealing temperature was set to 63^{o}C or above. As such, the combined data confirmed that, in PM cells, P1 expression increased on average between 2.5-5.0 fold compared to the levels determined in P cells.

Bisulfite sequencing of the P1 promoter region revealed a surprising pattern of DNA methylation in PM cells.

We performed bisulfite sequencing analysis of the 19 CpG dinucleotides within 1.5 kb of the P1 promoter spanning the region -500 nt to +1000 nt relative to the P1 transcriptional start site in P and PM cells. We sampled this region by designing primer pairs to amplify four 300-350 bp fragments. The CpGs within the P1 promoter in P cells were almost completely methylated, in agreement with the DNA methylation profile across P1 previously described for other murine dividing cells, i.e. L1210 cells (Figure 3-7, closed circles). However, the pattern of DNA methylation detected in PM cells across this region was very surprising: P1 remained abundantly methylated (Figure 3-7), in spite of the fact that genomic DNA from PM cells were found previously to be largely unmethylated (137). Remarkably, significant differences in the level of methylation at individual CpG residues throughout the region were apparent, and it appeared that loss of DNA methylation was reserved for the CpGs at positions -500, -250, and +50 nt relative to the P1 transcriptional start site. The methylation of the CpG at +50 nt approached zero, and its proximity to the

transcriptional start site was extremely interesting. Additionally, the methylation of the CpGs between +750 and +1000 nt was also decreased in PM cells compared to P cells by approximately 40%. However, the methylation levels across this 250 bp fragment remained substantially higher than the level of DNA methylation determined at these CpGs in mouse liver, which were found methylated in only 25% of the sequenced clones. How the DNA methylation is maintained across this region in PM cells remains an open question. It should be noted that work in Dr. Taylor's laboratory has shown that there are residual levels of DNMT1 in PM cells, supporting the idea that the PM cell line should be considered a DNMT1 hypomorph, rather than a complete knock-out cell line. These data suggested that the loss of CpG methylation at a few residues within the 1.5 kb region was sufficient to cause the increase observed in the transcriptional activity of P1 in PM cells.

Histone H3 acetylation and lysine 4 tri-methylation across the mouse *fpgs* promoters in P and PM cells.

The interplay between DNA methylation and several histone PTMs has been well described, but in many cases the causative epigenetic modification remains unknown (74). In chapter 2, we determined that loss of DNA methylation at P1 in mouse liver was coincident with an increase in P1 transcriptional activity and H3Ac and H3K4me3 of regional nucleosomes. In these studies, we used the P and PM cells to question if loss of DNA methylation was sufficient to cause an increase in H3Ac and/or H3K4me3. ChIP experiments were performed using P and PM cells with antibodies generated against H3K9Ac and H3K4me3. Real-time PCR was used to amplify regions of genomic DNA

within the P1 and P2 promoters and enrichment at the P2 promoter was used as a positive control. As expected, the histones surrounding the P2 promoter in both cell lines were decorated with abundant levels of H3Ac and H3K4me3 (Figure 3-8), reflecting the transcriptional activity of this promoter in these dividing cells. Histone H3 acetylation was also detected across the P1 promoter in both P and PM cells when compared to the non-specific IgG control (Figure 3-8). We were very surprised by the high levels of histone H3 acetylation detected over P1 in P cells. Histone acetylation at this promoter in L1210 cells was previously detected, but the levels found in P cells are more abundant. We are not yet clear on the cause of this increase in acetylation at P1 in P cells, but it may very well be a consequence of the p53-null status of the P cells. Interestingly, the amount of H3Ac found over P1 in PM cells was 1.5x higher than that observed in P cells, a factor that was calculated by normalizing the levels of H3Ac measured over P1 to those observed over the P2 promoter in the respective cell lines. We took this to suggest that the changes of DNA methylation patterns at the P1 promoter observed in PM cells were impacting the level of H3Ac across the promoter.

The presence of H3K4me3 across a promoter region in both yeast and mammals has been tightly linked to the presence of RNAPII complexes (13, 87, 186). The transcriptional activity of P1 was found to increase slightly in PM cells (Figure 3-6), however H3K4me3 was not detected across P1 in the hypomethylated PM cells (Figure 3-8). From these data, it appeared that the altered patterns of DNA methylation observed at positions -250 nt and +50 nt (Figure 3-5), and/or the slight increase in P1 activity detected in PM cells (Figure 3-6) were not sufficient to enhance H3K4me3 across this promoter; it

appeared that H3K4me3 at the P1 promoter was more tightly correlated with transcriptional activity, rather than DNA methylation patterns.

Inhibition of histone deacetylase activity enhances P1 activity in P and PM cells, but more substantially in PM cells.

Methyl-binding proteins are recruited to methylated CpG dinucleotides and often serve as a scaffold for large repressor complexes that include chromatin remodeling and HDAC enzymes (123, 170). As a result, histone deacetylation and DNA methylation have been shown to cooperate in mediating transcriptional silencing at several CpG-dense promoters. It was not clear if the link between DNA methylation and HDACs would translate to the CpG-sparse P1 promoter. As Cedar's group described, a large proportion of mammalian genes require both loss of methylation and inhibition of histone deacetylation to achieve transcriptional activation. We questioned if HDAC inhibition would enhance the P1 transcriptional activity in P and PM cells. Cells were treated with 0.06 μM TSA for 12, 24, 48, and 72 hours and P1-specific *fpgs* cDNA was measured using semi-quantitative and real-time PCR as described for Figure 3-6. *Slpi* expression was measured as a control. The growth of the two cell lines was significantly impacted by treatment with TSA for 48 and 72 hours, and to minimize the influence of secondary drug effects these cells were not processed. Treatment with TSA for 24 hours increased the expression of *Slpi* 7-fold and 22-fold in P and PM cells, respectively (data not shown). However, this treatment did not reproducibly alter the level of P1 transcriptional activity in either P or PM cells (data not shown). *Slpi* expression was also increased when cells were

treated for 12 hours with TSA but to a lesser extent than observed after 24 hours of treatment: expression increased 4-fold in P cells and 6-fold in PM cells (Figure 3-9). On the other hand, the level of P1 activity appeared to increase in both P and PM cells when cells were treated with TSA for 12 hours (Figure 3-9). P1-specific *fpgs* cDNA levels increased in P cells approximately 4-fold and increased in PM cells approximately 8-fold above the levels of P1 expression detected in the respective cell lines without TSA treatment (Figure 3-9). As such, HDAC inhibition coupled with the changes in DNA methylation appeared to further enhance P1 activity in PM cells. It is interesting to note that the level P1 expression detected in P cells treated with TSA approached the amount of P1-specific *fpgs* cDNA found in PM cells relative to GAPDH expression (Figure 3-9). Since DNA methylation was not substantially lost across P1 (Figure 3-7), we considered the possibility that the effect of TSA treatment on the expression of P1 in PM cells would be enhanced if the methyl groups of the CpGs across P1 were further depleted.

5-aza-deoxycytidine treatment of P and PM cells for 24 and 48 hours.

We could not fully understand the role of DNA methylation in the control of the tissue-specific expression pattern of the P1 promoter without further depleting the CpG dinucleotides in this region of methyl groups. We questioned if further inhibition of DNA methyltransferase activity would result in loss of DNA methylation and enhanced transcriptional activity at the P1 promoter in PM cells. To test this we treated P and PM cells with 1μM of 5-aza-deoxycytidine (5daza), an inhibitor of the DNA methyltransferases, for 24, 48, and 72 hours. The growth and viability of the P cells were

substantially impacted by treatment with 5daza for 48 and 72 hours, however PM cells appeared to grow normally and were unaffected by the presence of 5daza in the medium. The resistance of DNMT1 hypomorphs to 5daza treatment has been previously described in HCT116 cells (60). This observation is not surprising since the major target of this drug, DNMT1, is already substantially reduced in PM cells. In these studies, we isolated RNA and DNA from the treated P and PM cells for expression analysis and bisulfite sequencing.

Following 24 and 48-hour treatment with 5daza the expression of P1-specific *fpgs* transcript was measured in P and PM cells (Figures 3-10 and 3-11). P1 activity did not increase in PM cells treated for 24 hrs (Figure 3-10), however following 48 hours of treatment the level of P1 activity measured by *fpgs* cDNA increased approximately 5-fold compared to the amount of P1-specific *fpgs* cDNA detected in untreated PM cells (Figure 3-10). The doubling time of the PM cells is approximately 18 hours, and the effect of DNA methyltransferase inhibition on the transcriptional activity of the P1 promoter appeared to require two-rounds of DNA replication. It is possible that a substantial loss of DNA methylation across P1 in PM cells does not occur until 48 hours after treatment.

Bisulfite sequencing analysis was performed using genomic DNA isolated from PM cells after 24 hours of treatment with 5daza. We have also planned to perform this analysis using DNA harvested from PM cells treated for 48 hours with 5daza. The level of DNA methylation was in PM cells 24 hrs after treatment with 5daza (Figure 3-12 A, open circles). However, the change in DNA methylation was reserved to a cluster of CpGs between +100 nt and +750 nt positions relative to the P1 transcriptional start site, when

compared to the level of DNA methylation found in untreated PM cells (Figure 3-12 A, closed circles). The CpGs upstream of position -250 nt and downstream of +750 nt were methylated in a remarkably similar manner to the same residues in untreated PM cells (Figure 3-12 A). The distinctly localized change in DNA methylation observed in PM cells treated with 5daza reflected the punctate pattern of hypomethylation found at individual CpGs in the untreated PM cells (Figure 3-12 A, closed circles). These data identify four categories of CpGs within the P1 promoter, based on their sensitivity to loss of DNA methyltransferase activity: 1) CpGs located between -500 and -250 on average maintained methylation in PM and PM+5daza conditions, 2) CpGs within the region between -250 and +50 nt were the most sensitive to loss of DNMT1, 3) the fragment between +70 and + 700 included CpGs that were sensitive to 5daza treatment, and 4) CpGs between 750 and 1000 nt were moderately sensitive to the reduction of DNMT1 in PM cells, but were not affected by further DNA methyltransferase inhibition with 5daza treatment after 24 hrs. The factors determining the difference in sensitivity of individual clusters of CpGs within this promoter region remain unclear. However, since the level of P1 activity did not increase in PM cells after 24 hours of treatment with 5daza it appeared that the methylation status of the CpG residues between +70 and +750 nts was not involved in determining the transcriptional activity of this promoter. These data stress the importance of the CpG residues at the -250 nt and +50 nt positions in the transcriptional regulation of P1.

The effects of 5-daza treatment on P1 transcriptional activity and CpG methylation was also assessed in P cells (Figure 3-12 and 3-11). These data were very surprising. The

level of P1-specific *fpgs* cDNA increased approximately 20-fold and more than 100-fold following 24 and 48 hours of treatment, respectively (Figure 3-11). Relative to the P1 activity measured in PM cells under the same conditions, the level found in P cells was 3-fold and 10-fold greater than the amount of P1-specific *fpgs* cDNA detected in PM cells following 24 and 48 hours of treatment with 5daza, respectively (Figures 3-10 and 3-11).

Bisulfite sequencing of genomic DNA from P cells treated with 5daza for 24 hours revealed loss of methlaltion at individual CpGs (figure 3-12, open circles). Remarkably, the CpG residue at the +50 nt and between +300 nt and +550 nt positions relative to the P1 transcriptional start site appeared the most sensitive to DNA methyltransferase inhibition (Figure 3-12). In contrast, the genomic regions upstream of -250 nt and downstream of +750 nt remained almost 100% methylated (Figure 3-12). Because the increment in P1 activity is substantial in P cells after treatment with 5daza for 48 hours we are very interested in determining if this change in expression correlates with loss of DNA methylation throughout the P1 promoter under these conditions.

Measurment of DNMT3a and 3b levels in P and PM cells

The level of DNA methylation remaining in PM cells was substantial (Figure 3-7), and even after treatment with 5daza for 24 hours methylation levels remained (Figure 3-12). It is very likely that following 48 hours of treatment with 5daza, DNA methylation is reduced across the P1 promoter in PM cells, since transcription originating at this promoter increased at this time point (Figure 3-10). However, we were not convinced that the low levels of DNMT1 present in PM cells would be sufficient to maintain the abundant levels

of DNA methylation detected at P1 in these cells. We proposed that perhaps either DNMT3a or DNMT3b were coordinating the methylation of CpG dincucleotides within the P1 promoter. Cedar's group drew the conclusion that in these MEFs the levels of DNMT3a and 3b were likely to be quite low, since these enzymes are expressed in ES cells and their major function is *de novo* methylation during embryonic development (177). We designed primers to measure the expression levels of DNMT3a and DNMT3b in P and PM cells (Figure 3-13 A). The difference between the cDNA levels encoding DNMT3a and DNMT3b between P and PM cells was remarkable. DNMT3a expression was up regulated approximately 9-fold in the hypomethylated PM cell line (Figure 3-13). Similarly, DNMT3b levels were 4-fold higher in PM cells than the levels detected in P cells. Western blot analysis performed in collaboration with Erica Peterson, a member of Dr. Taylor's laboratory, confirmed that the amount of DNMT3a and DNMT3b protein levels in PM cells were substantially greater than the levels found in P cells (Figure 3-13 B). Interestingly, DNMT1, DNMT3a, and DNMT3b were all sensitive to 24-hour treatment with 1μM 5-daza, however DNMT3a appeared less sensitive than DNMT1 in PM cells (Figure 3-13 B and C). These data suggest that DNMTs 1, 3a, or 3b are potential candidates for the regulation of DNA methylation across the P1 promoter in PM cells. We know that the activity of DNMT 3a and 3b in PM cells is not able to compensate for the loss of DNMT1 globally, since genome-wide DNA methylation analysis has found these cells to be severely hypomethylated (137). However, it may be the case that DNMT3a or 3b activity sustains methylation of P1 in PM cells. It should also be noted that the level of DNMT3a in P cells is remarkably high, considering MEFs were not predicted to have any

detectable levels of *de novo* methyltransferases. Perhaps DNMT3a is responsible for the methylation of P1 in both P and PM cells, suggesting a new role for *de novo* methyltransferases in the methylation of tissue-specific promoters.

DISCUSSION

The results presented in chapter 2 suggested that epigenetic and transcriptional interference mechanisms were involved in coordinating the use of the two promoters. We sought to determine if these epigenetic factors are causative or consequential of the tissue-specific control of the mouse *fpgs* gene. In order to approach this question, we used a model system that caused a primary disturbance in DNA methylation at P1, and a P1-promoter knock, in which transcriptional interference at P2 was impossible and in which cross-talk between histone modifications and promoter activity could be tested. The patterns of DNA methylation at the P1 promoter are more complicated than we previously thought, but our data suggest that CpG methylation facilitates transcriptional silencing of the P1 promoter. Likewise, our studies supported transcriptional interference/occlusion at P2 in mouse liver, but also indicated other mechanisms enforcing tissue-specific restriction of P2 expression in liver.

Studies using the P1-knock-out mouse model

Tissue-specific patterns of histone marks are not simply a consequence of transcriptional activity.

Studies in yeast have best shown the functional consequences of individual histone PTMs in transcription (38, 128), but classifying histone PTMs as either causative or consequential of a particular transcriptional state at mammalian genes *in vivo* have been difficult. In the studies presented in this chapter, we used the P1 knock-out mouse model to determine if the distinct patterns of histone PTMs observed between mouse liver, brain, and L1210 cells were the consequence of P1 transcriptional activity in mouse liver. The levels of histone H3 acetylation across the *fpgs* gene in P1 knock-out liver were much lower than those detected in wild-type mouse liver, but the profiles described between the two genetically different tissues were remarkably similar (Figure 3-3). Furthermore, the pattern of histone H3 acetylation at the *fpgs* locus in knock-out liver was surprisingly different then those determined in L1210 cells and mouse brain, in spite of the fact that P1 activity was absent in all three of these tissues. Hence, it appeared that additional tissue-specific mechanisms were driving the positioning of histone acetylation across the coding region of the *fpgs* gene, and that transcriptional activity at P1 served to enhance the histone acetylation at those pre-determined positions. Our data suggests that histone marks across a tissue-specific gene are more than simply a consequence of transcriptional activity and support a functional role for these epigenetic modifications in the control of tissue-specific expression. Support for this concept has recently been furnished in studies mapping chromatin marks across enhancer and insulator elements at high resolution in several

human cell lines (102). These studies showed that insulator elements share similar epigenetic marks between different cell-types, but the chromatin signatures found across enhancer elements are often unique to a specific cellular background. Enhancers are *cis* acting elements that, when bound by sequence-specific DNA binding proteins can impact transcriptional activity from long range distances (>10 kb) and when positioned either upstream or downstream of the targeted promoter (26). We consider that enhancer elements present within the *fpgs* gene may be responsible for dictating the pattern of acetylated nucleosomes and perhaps tissue-specific expression at the *fpgs* gene in mouse liver.

The enrichment of histone H3 lysine 4 tri-methylation at regional nucleosomes has been clearly linked to the presence of RNAPII at the 5' end of genes in an active or poised state (13, 87). The abundant peak of H3K4me3 detected in wild-type mouse liver adjacent to P1 was absent in P1 knock-out mouse liver, presumably due to the lack of transcriptional initiation occurring at this region (Figure 3-2). However, we detected a substantial enrichment of H3K4me3 approximately 4 kb downstream of the +1 transcriptional start site of the P1 promoter in knock-out mouse liver. This peak was not detected in other tissues where P1 activity was also absent, i.e. L1210 cells and mouse brain. We took this to suggest that the genomic region 4 kb downstream of P1 in knock-out liver is permissive to RNAPII binding, and may represent a cryptic promoter that has previously not been identified. This occurrence may very well be related to the histone acetylation present at the *fpgs* gene in knock-out liver. Histone acetylation enhances the availability of *cis* DNA-binding elements to cellular *trans* factors (249). As such, in the

absence of P1 activity, the level of histone acetylation across the *fpgs* gene in mouse liver may be sufficient to permit binding of RNAPII at the +4 kb position.

Promoter occlusion at the P2 promoter of the mouse fpgs gene

We used the P1 knock-out mouse model to determine if transcriptional interference was involved in selecting the production of one FPGS isoform over another in mouse liver. Knock-out of P1 activity did indeed result in an increase in the transcriptional activation of the P2 promoter, confirming our previous conclusions in studies using wild-type mouse liver. In our laboratory, Lin-Ying Xie estimated that the increase in P2 activity in knock-out liver ranged between 4 to 6-fold when compared with levels of P2 expression found in wild-type liver (263). Removal of P1 activity was coincident with an increase in the binding of the transcription factor Sp1 to the P2 promoter in liver, supporting the idea that transcriptional initiation at the P2 promoter in wild-type tissue was blocked as a consequence of high levels of transcriptional activity originating at the P1 promoter (Figure 3-3). Interestingly, the expression of P2 in mouse kidney, a tissue that also uses P1, increased only approximately 1.5-fold in P1 knock-out tissue. We took this to suggest that the degree of transcriptional interference at P2 in mouse kidney was less than that occurring in mouse liver. This finding may be explained by previous data from RPAs, which showed that the amount of transcript generated from P1 was lower in kidney than in liver (Figure 2-2); this decrease in P1 activity in kidney was coincident with a higher level of detectable P2 expression in this tissue compared to liver (244). Taken together, it appears that the degree of transcriptional interference at P2 reflects the level of P1

transcriptional activity. It would be very interesting to determine if elongating RNAPII complexes accumulate over the P2 promoter in mouse kidney as they do in mouse liver. Overall, our studies using the P1-knock mouse liver concluded that a mechanism of promoter occlusion was limiting the P2 activity in wild-type mouse liver. This is one of the first proven cases where promoter occlusion was shown to control promoter usage of a mammalian gene *in vivo*.

Multiple control mechanisms of the P2 promoter

The P2 promoter represents a unique case where a CpG-island promoter is not broadly expressed, as is often the situation, since CpG-island promoters are often found at the 5'-ends of housekeeping genes. We proposed that P2 in mouse liver was controlled by transcriptional interference, however relief from transcriptional interference increased transcription at P2 only 4-6 fold, but was still 30-50 times lower than the level of P2 expression measured in dividing cells. Why is the P2 promoter not fully activated in knock-out mouse liver? One possibility is that the level of *trans* factors present in dividing cells causes robust transcriptional activation of the P2 promoter necessary to accommodate high levels of cellular proliferation. Perhaps enhancer elements along the length of the *fpgs* gene are differentially used to generate the dividing-cell specific levels of *fpgs* expression, as suggested by the genome-wide studies discussed previously (102). It may also be the case that repressive mechanisms, in addition to transcriptional interference, are restricting P2 expression in mouse liver. However, the fact that we detected HNF4 at P2 in mouse liver argues that transcriptional activation, rather than repression is the favored

transcriptional state of this promoter in mouse liver. HNF4 is also expressed in mouse kidney, and we believe that the redundant tissue-specific factors between liver and kidney may explain why measurable levels of P2 activity are detected in these tissues, compared to other differentiated tissues, such as mouse brain, where no *fpgs* transcripts are found (244).

In mouse brain, the P2 promoter was found to be poised for transcriptional activation (Figure 2-24), representing an additional mechanism of transcriptional regulation of this promoter. We are interested in determining if the poised state of the P2 promoter exists across multiple mouse tissues where P2 is silent. Genome-wide studies in human and mouse samples have shown that poised RNAPII complexes are common to a large number of genes and P2 appears representative of this class of promoters (13, 87).

A comprehension of the multitude of transcriptional mechanisms coordinating the expression of the P2 promoter of the mouse *fpgs* gene is critical to our understanding of the regulation of the human *fpgs* locus. In human tissue, *fpgs* mRNA is expressed in heart, liver, lung, and skeletal muscle (71, 244). Surprisingly, human *fpgs* mRNA is generated from the sole use of the P2 promoter (244). Since the effectiveness of antifolate-based chemotherapeutic regimens depend on the presence of FPGS protein in sensitive tumor cells, an awareness of the transcriptional mechanisms involved in coordinating the expression patterns of the P2 promoter in mouse tissues, might have important clinical implications.

DNA methylation studies

The studies described in this chapter using the P and PM cell lines were designed to ask the question: Is loss of DNA methylation sufficient to transcriptionally activate the P1 promoter in mouse dividing cells? Previous work demonstrated that methylated CpG dinucleotides were severely depleted in PM cells due to a substantial reduction of DNMT1 protein levels (137). Our interpretation of the results we obtained with these cells was complicated by the fact that the PM cells were not a simple case of a hypomethylated cell line existing in culture, as had previously been thought. However, we have made some surprising and interesting observations regarding the epigenetic control of the P1 promoter and the transcriptional regulation of the levels of DNMT3a and DNMT3b.

Mechanisms controlling the P1 promoter: epigenetics and tissue-specific factors

We tested the consequences of changes in epigenetic mechanisms on the activity of the P1 promoter using a PM MEF cell line and through the use of pharmacological agents, which lead to alterations of the cellular levels of DNA methylation and histone acteylation. The substantial increase in P1 transcriptional activity following 48 hours of treatment with 5daza in both P and PM cells suggest that loss of DNA methylation is sufficient for transcriptional activation of the P1 promoter. However, the levels of P1 activity detected in the treated P cells are at least three orders of magnitude lower than the levels found in mouse liver, suggesting that other mechanisms are involved in regulating this promoter. In support of this idea, recent experiments have determined that HNF4 is present at the P1 promoter in mouse liver. One possibility is that loss of DNA methylation permits basal

activation of the P1 promoter, but tissue-specific activators are required to achieve the levels of transcription measured in the liver. Whether loss of DNA methylation is required in order for *trans* factors to affect the transcriptional activity of P1 is a very interesting question and is currently being explored.

How are the methylation patterns across the P1 promoter established?

The data presented in this chapter stress that we are far from understanding the entire story linking DNA methylation and tissue-specific gene expression. The manner in which the pattern of DNA methylation across P1 is altered with loss of DNMT1 is remarkable: individual CpGs appear to lose methyl groups at dramatically different rates throughout the entire genomic region sampled in our bisulfite analysis. These data suggest that the mechanisms responsible for DNA methylation across the P1 promoter are multifaceted and methylation of individual CpGs within the same genomic region may be controlled differently. To the best of our knowledge, reduced levels of DNMT1 have been shown to cause loss of DNA methylation within CpG-rich promoter regions in a uniform fashion (60, 194), making the P1 promoter an interesting exception to this general rule. We do not yet understand how these patterns are established or their functional consequences, but our data argues that the rules dictating DNA methylation at CpG-rich promoters may be different than those of CpG-sparse promoters.

In PM cells, CpGs adjacent to the transcriptional start site of the P1 promoter are severely hypomethylated, yet the rest of the region retains high levels of methylation. One explanation for this pattern is that the remaining DNMT1 enzyme activity in PM cells

maintains the methylation of most of the P1 region, but is not able to methylate the sites adjacent to the transcriptional start site. However, DNMT1 has been shown to methylate a hemi-methylated substrate *in vitro* in a processive manner, arguing that partial methylation of CpG residues within a region of genomic DNA is unlikely mediated by this enzyme (250).

Alternatively, loss of methylation observed in PM cells at P1 may be a consequence of the hypomethylated state of these cells. The global levels of gene expression are altered in these cells and the pattern of methylation observed may be the result of a block in methylation caused by *trans*-acting factors binding. In this model, the transcription factor is expressed in PM cells but not P cells. Prevention of DNA methylation through transcription factor binding has previously been described (97, 235).

Treatment of PM cells with 5daza caused further loss of DNA methylation across the P1 promoter region. However, the effects of 5daza are not specific to DNMT1 and the levels of DNMT3a and DNMT3b were also affected by treatment with this agent. As such, the change in methylation observed in PM cells after 24-hour treatment with 5daza may be a consequence of the loss of DNMT 1, 3a, 3b, or a combination of the proteins.

When P cells were treated with 5daza for 24-hours, expression of P1 was increased and loss of methylation was observed in a pattern highly specific to individual CpG dinucleotides, similar to the profile found in PM cells. Subsequently, treatment for 48 hours resulted in a substantial increase in P1 expression that was higher than the levels observed in PM cells treated with 5daza. The difference in P1 expression between P and PM cells following 5daza treatment was suprising, since we expected that P1 might be

more sensitive to loss of DNMT1 levels in PM cells, where this protein had already been severely diminished. We have not yet explained this discrepancy, but we are strongly considering that alternate DNMT enzymes are involved in regulating the methylation across the P1 promoter. This is a divergent concept from the current thought that DNMT1 is the sole enzyme responsible for maintaining DNA methylation in differentiated cells or tissues.

Are DNMT3 protein levels regulated by DNA methylation or cellular levels of DNMT1?

The expression of the genes encoding the *de novo* methyltransferase enzymes, DNMT3a and DNMT3b, occurs most abundantly at early stages of embryonic development in the mouse (177). In the characterization of the PM cell line, Cedar et. *al* reported that the contribution of the *de novo* enzymes to genome-wide methylation in these cells were likely to be minimal. However, we detected higher levels of DNMT3a and DNMT3b at both the transcript and protein levels in PM cells when compared to P cells, suggesting that these genes may be regulated by DNA methylation or cellular levels of DNMT1. Additional reports have found that one of the two promoters within the *DNMT3a* locus is hypomethylated in HCT116 cells that are DNMT1 hypomorphs; however, the levels of *DNMT3a* transcript or protein were not measured (60). The coordination of DNMT expression through DNA methylation has previously been described at the DNMT3L locus, where methyl groups deposited by DNMT3a and DNMT3b activity led to the silencing of this gene (108). Whether the *de novo* enzymes DNMT3a and DNMT3b are regulated by mechanisms linked to cellular DNMT1 levels is not known. An

appreciation for how these enzymes are over-expressed in PM cells may shed light on this potentially important connection.

Are de novo methyltransferase enzymes responsible for methylation of P1?

The over expression of the *de novo* methyltransferase in PM cells may be responsible for maintaining the methylation levels observed across P1 in our studies. DNMT1 knock-out animal studies have clearly shown that the *de novo* enzymes are not compensatory for loss of DNMT1 on a genome-wide level (142), but a subset of genes may remain silenced in DNMT1 hypomorphs due to the activity of DNMT3a or DNMT3b that has not previously been identified. Interestingly, reasonable levels of DNMT3a but not DNMT3b were also found in P cells. This is in agreement with previous literature that showed the full-length form of DNMT3a at low levels in somatic tissue, but not DNMT3b (177). Is it possible that one of the *de novo* methyltransferases may be responsible for the methylation of P1 in PM cells and perhaps also P cells? Literature precedent suggests that DNMT3a may be a candidate enzyme regulating P1 methylation: 1) DNMT3a, but not DNMT3b, is detected in P cells and is also found at low levels in normal somatic tissues (177). 2) DNMT3b catalyzes the addition of methyl groups through a processive mechanism, similar to DNMT1, and loss of this protein is unlikely to cause the discrete hypomethylation of individual CpGs (82). And, 3) Dnmt3a has been shown to methylate DNA substrates *in vitro* in a distributive manner, resulting in products that are methylated in a non-uniform fashion. This has lead to the proposal that DNMT3a may require targeting to individual CpG sites (103). Perhaps a distributive enzyme would be more

easily used for the methylation of CpG-sparse promoters, but the processive DNMT1 for a CpG island promoter. As such, reduction of this protein may impact individual CpGs within a small genomic region differently. *In vitro* data also supports the idea that cooperation between the DNMTs may function to methylate the CpG-sparse promoter of the *fpgs* gene (67). An appreciation for how the DNA methylation patterns are established at the P1 promoter in P and PM cell lines is important to our understanding of the happenings at this promoter *in vivo*.

A class of promoters exists in both humans and mice that is methylated in a tissue-specific fashion, which inversely correlates with transcriptional activity. An awareness of how these differentially methylated regions are established is central to our ability to test their functional importance. We initiated these studies in the P and PM cells thinking that changes in DNMT1 would allow us to determine the contribution of DNA methylation to the control of the P1 promoter. Clearly, we underestimated the complexity of the patterns of methylation previously described at P1. We are eager to determine how DNA methylation is coordinated at P1 and if our observations are universal to tissue-specific CpG-sparse promoters genome-wide.

In this chapter, we used model systems with single genetic disturbances to study the impact of changes in P1 activity and DNA methylation on the use of the two mouse fpgs promoters. These studies are preliminary but, in spite of this fact, we have made some very interesting observations that require explanations. Future studies will be designed to address some of the remaining questions including: 1) Why is the P2 promoter not fully activated in P1 knock-out mouse liver? 2) What mechanism is determining the

tissue-specific patterns of histone marks found in mouse liver independent of the transcriptional activity at P1? 3) Why is P2 able to produce transcript in mouse liver and kidney, but not brain? 4) How are different methylation patterns of P1 established in adult mouse tissues, and why are certain CpGs sensitive to loss of cellular levels of DNMTs? And, 5) Are tissue-specific factors involved in the coordination of the P1 promoter and if so is loss of DNA methylation required for these *trans* factors to activate the P1 promoter?

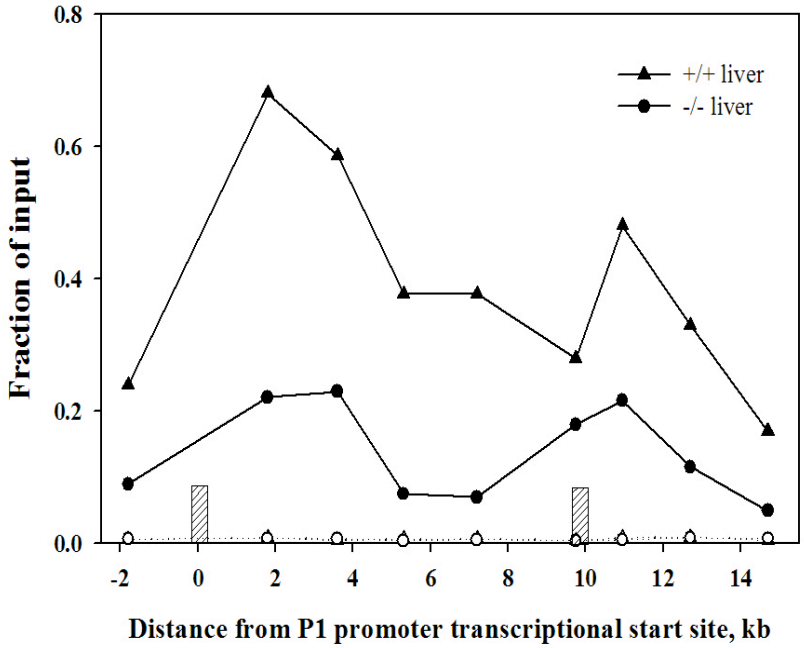


Figure 3-1. Histone H3 acetylation across the mouse *fpgs* gene in wild-type and P1 promoter knock out mouse liver. Chromatin from wild-type (closed triangles) and P1 knock out (closed circles) mouse liver was cross-linked, sonicated, and immunoprecipitated with an antibody against acetyl-H3K9 (closed symbols) or non-specific IgG (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR. The two transcriptional start sites are shown as the cross-hatched bars.

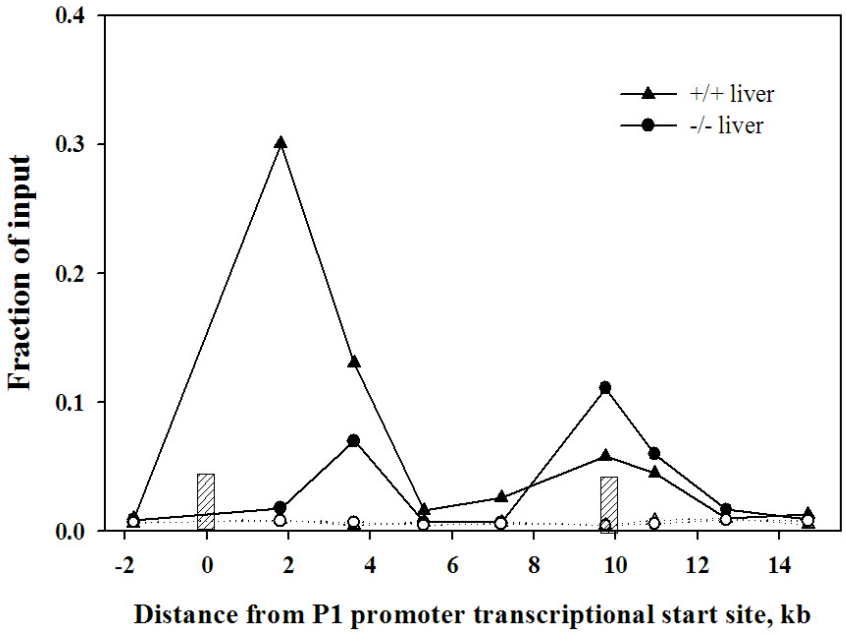


Figure 3-2. Histone H3 lysine 4 tri-methylation across the mouse *fpgs* gene in wild-type and P1 promoter knock out mouse liver. Chromatin from wild-type (closed triangles) and P1 knock out (closed circles) mouse liver was cross-linked, sonicated, and immunoprecipitated with an antibody against H3K4me3 (closed symbols) or non-specific IgG (open symbols). The content of DNA for various segments of the *fpgs* locus was determined by real-time PCR. The two transcriptional start sites are shown as the cross-hatched bars.

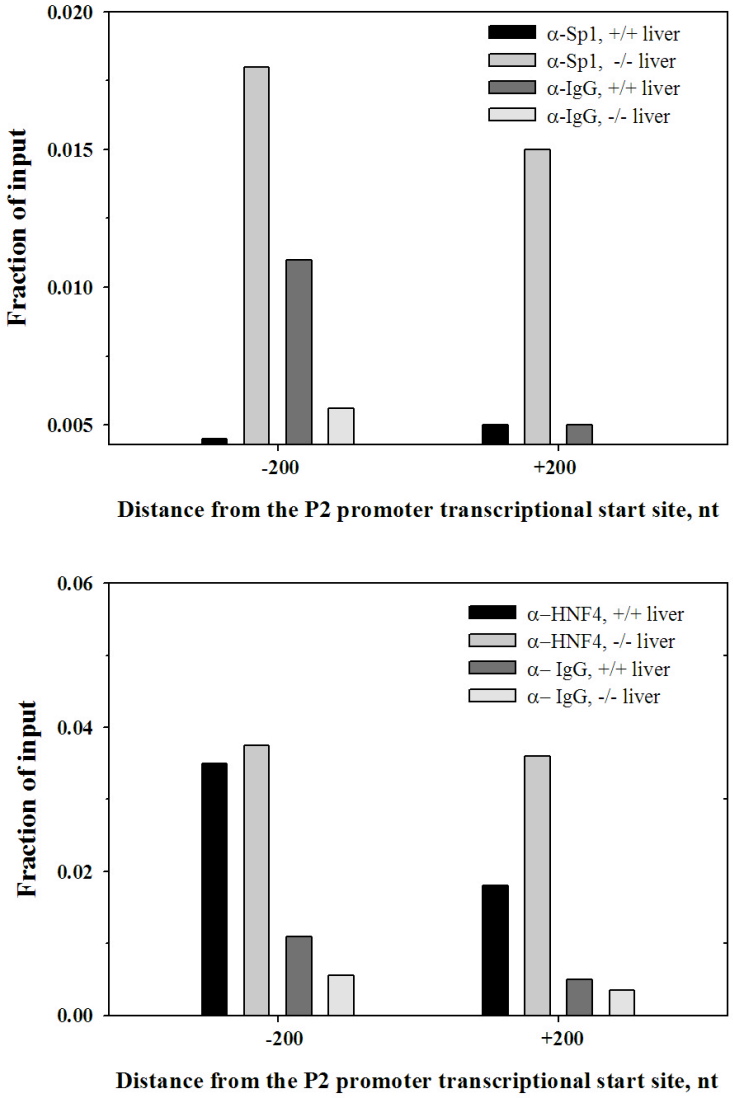


Figure 3-3. Residence of *trans* factors at positions the P2 promoter in wild-type and P1 knock out mouse liver. Chromatin from wild-type (black and dark grey vertical bars) and P1 knock out (light grey bars) mouse liver was cross-linked, sonicated, and immunoprecipitated with an antibodies against Sp1 (A), HNF4 (B) or non-specific IgG (A and B). The content of DNA at -200 nt and +200 nt relative to the P2 promoter transcriptional start site was determined by real-time PCR.

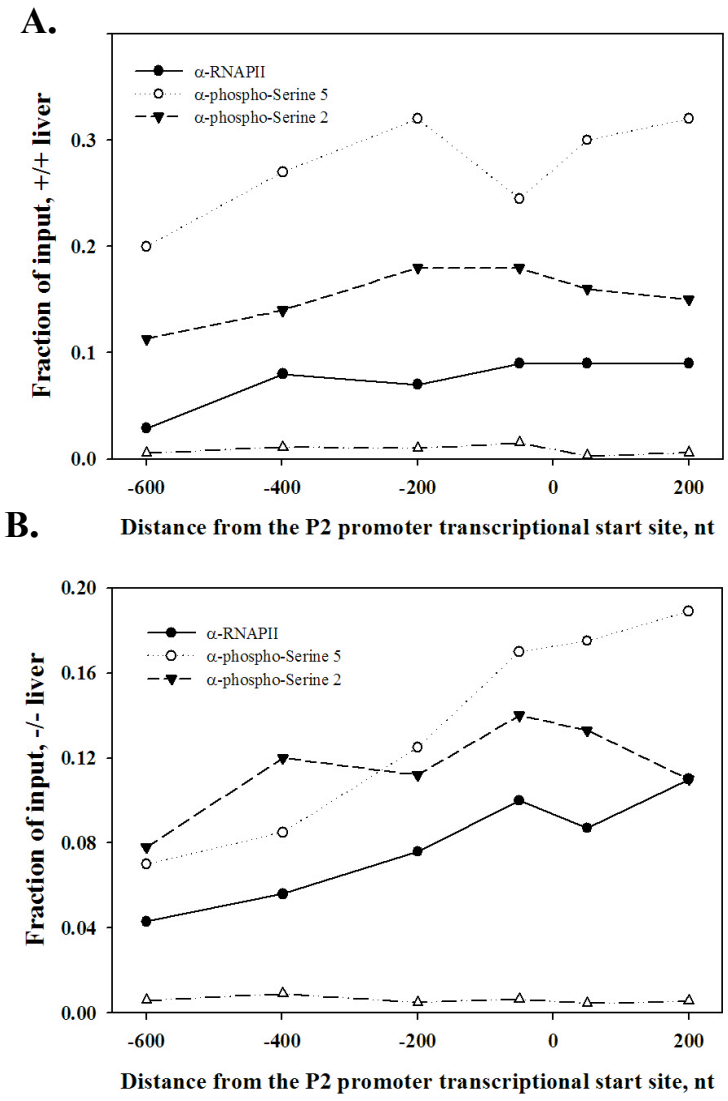


Figure 3-4. RNAPII occupancy across P2 in wild-type and P1 knock out mouse liver. Chromatin from wild-type (A) and P1 knock out (B) mouse liver was cross-linked, sonicated, and immunoprecipitated with an antibodies against total RNAPII (closed circles), phospho-serine 5 RNAPII (open circles), phospho-serine 2 (inverted triangles) or non-specific IgG (open triangles). The DNA content was determined by real-time PCR by amplifying six overlapping fragments. The points were plotted at the midpoint of each fragment.

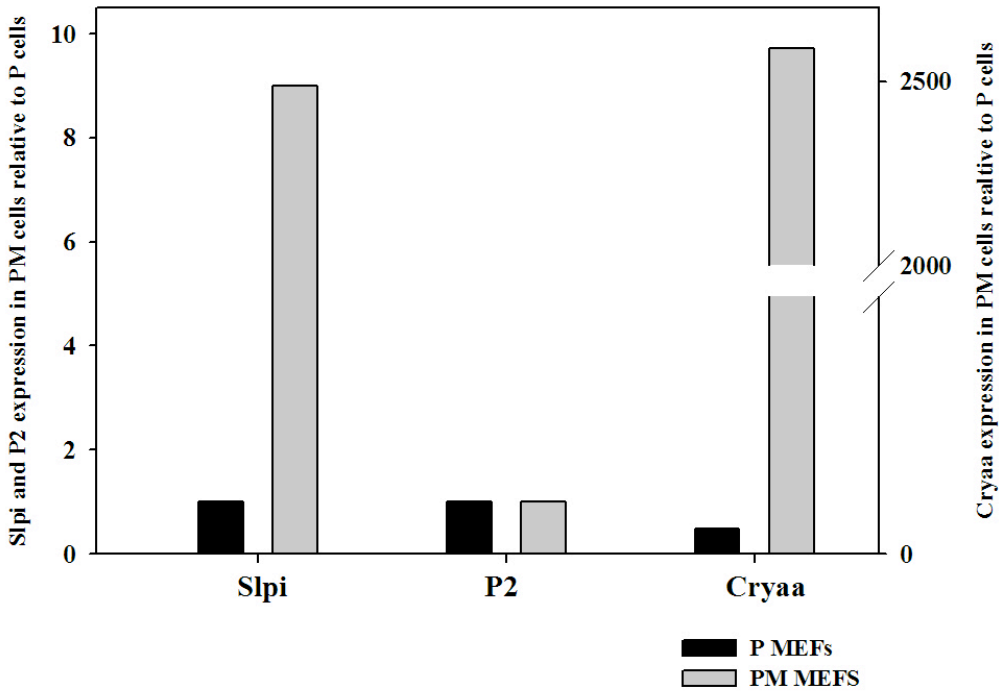


Figure 3-5. Expression of *Slpi*, *Cryaa*, and the P2-specific *fpgs* transcript in PM cells relative to levels detected in P cells. RNA was isolated from P (black vertical bars) and PM (grey bars) cells. 2.5 µg to 5 µg of RNA was converted to cDNA using Superscript III Reverse transcriptase kit. 1 µl of cDNA was added to a 25µl PCR reaction. cDNA content was measured using gene-specific primers in real-time PCR. The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each cell line relative to the amount of GAPDH. In these experiments, the level of Slpi, Cryaa, or P2-specific cDNA measured in P cells was normalized to 1.0.

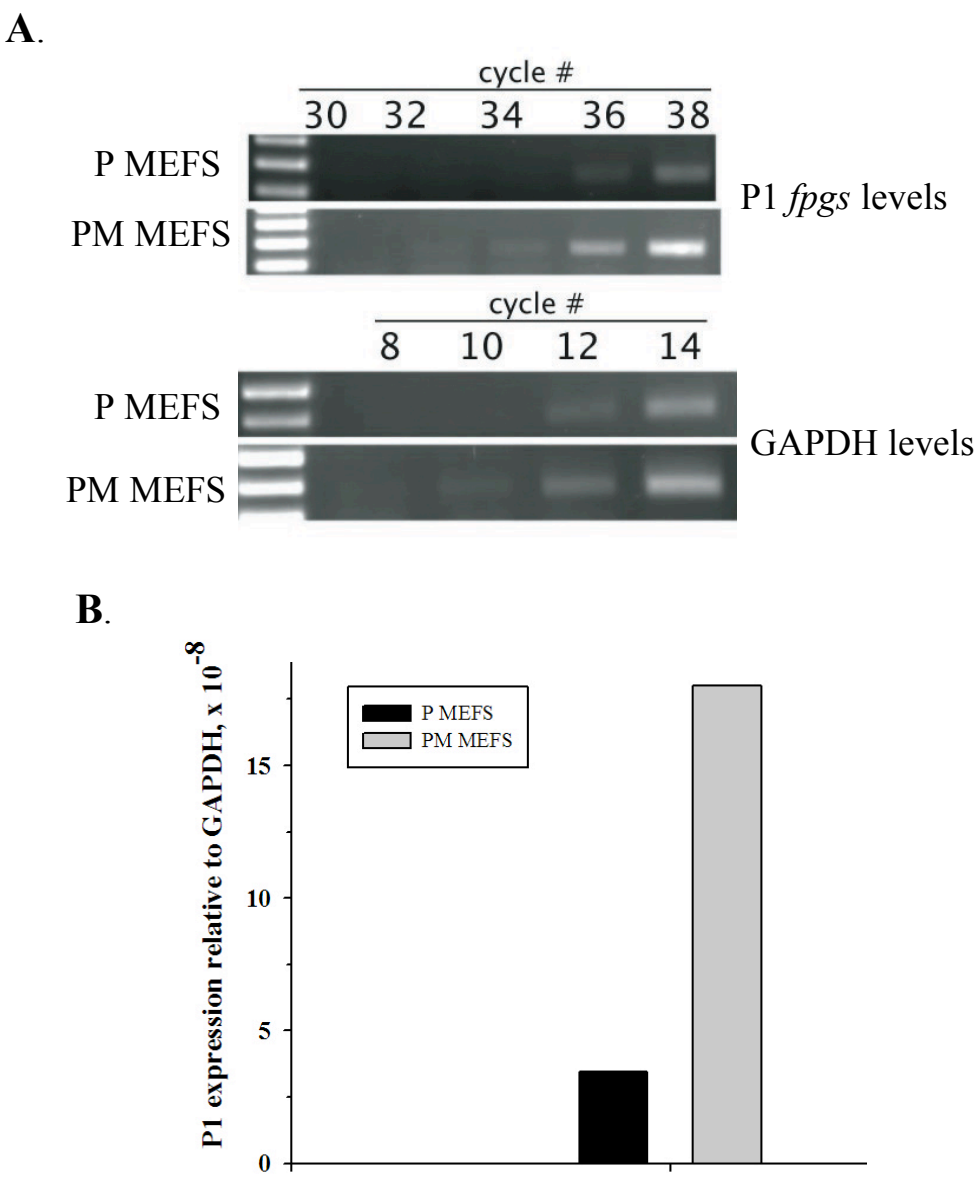


Figure 3-6. P1 transcriptional activity in P and PM cells. RNA was isolated from P (black vertical bars) and PM (grey bars) cells. 2.5 µg to 5 µg of RNA was converted to cDNA as in Figure 3-5. 1 µl of cDNA was added to a 25µl PCR reaction. (A) Semi-quantitative PCR was performed and reactions were terminated at 30, 32, 34, 36, and 38 cycles to quantify the level of P1-specific *fpgs* cDNA in P and PM cells. The primers used to detect this *fpgs* cDNA were specific for regions in A1b and exon 3. PCR reactions using primers specific for GAPDH cDNA were quenched at 8, 10, 12, and 14 cycles. (B) P1-specific *fpgs* cDNA was measured using real-time PCR with primers specific for regions within exon A1b. The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each cell line relative to the amount of GAPDH.

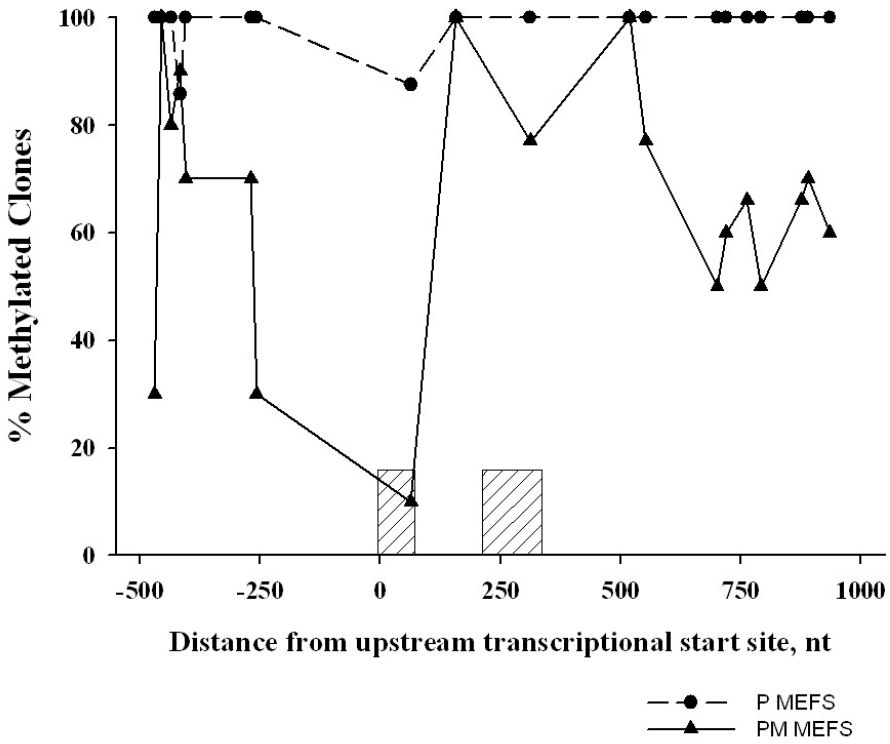


Figure 3-7. CpG methylation across the P1 promoter in P and PM cells. Genomic DNA was isolated from P (closed circles) and PM (closed triangles) cells. 200-500 ng of genomic DNA underwent bisulfite conversion. The P1 promoter was amplified in four fragments using bisulfite specific PCR primers. The region sampled spanned from -500 to +1000 nt relative to the P1 transcriptional start site. PCR products were gel-purified, cloned, and the sequences of 7-10 colonies were determined for each amplified region. At the position of each CpG dinucleotide in the genomic sequence, the percentage of clones containing a methylated cytosine is represented in the graph, plotted relative to the positions of exons A1a and A1b (cross-hatched vertical bars).

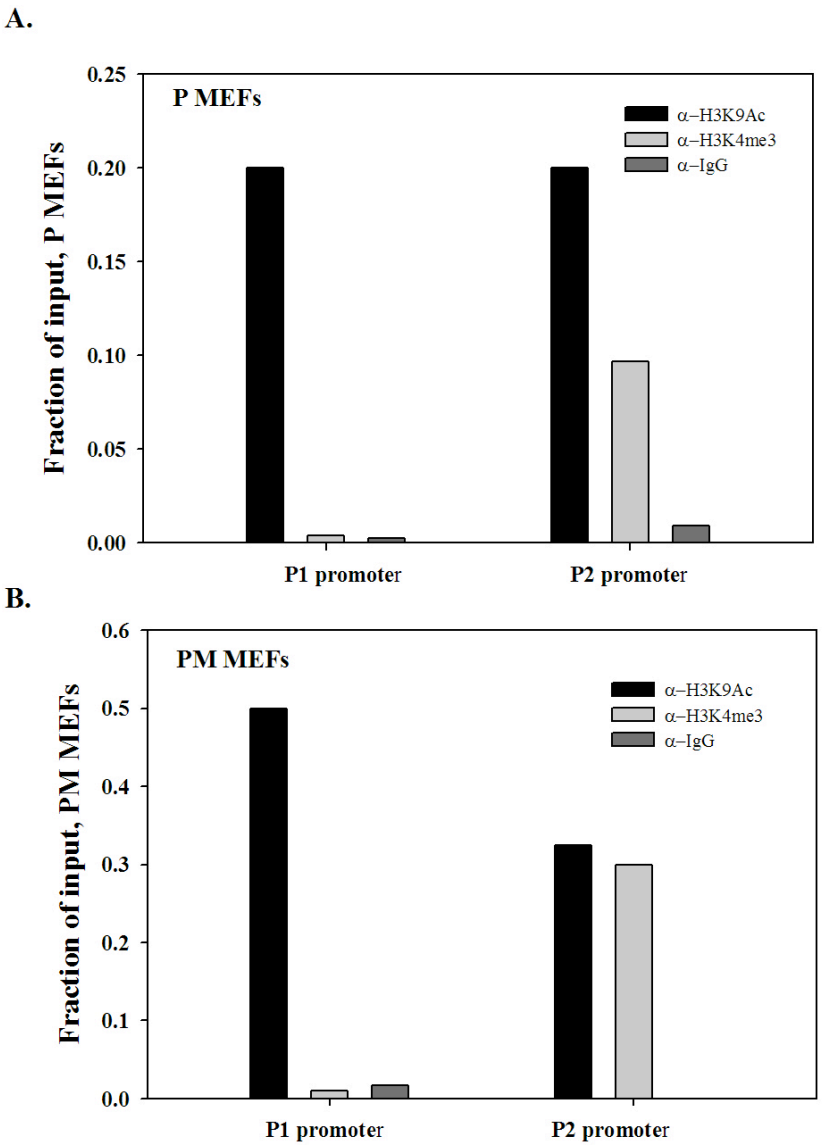


Figure 3-8. Histone H3 acetylation and lysine 4 tri-methylation at the two promoters of the mouse *fpgs* gene in P and PM cells. Chromatin from P (A) and PM (B) cells was cross-linked, sonicated, and immunoprecipitated with an antibodies against H3K9Ac (black bars), H3K4me3 (light grey bars) or non-specific IgG (dark grey bars). The midpoint of each amplified region was at +50 nt and -200 nt relative to the P1 and P2 promoter transcriptional start site, respectively. Real-time PCR determined the DNA content.

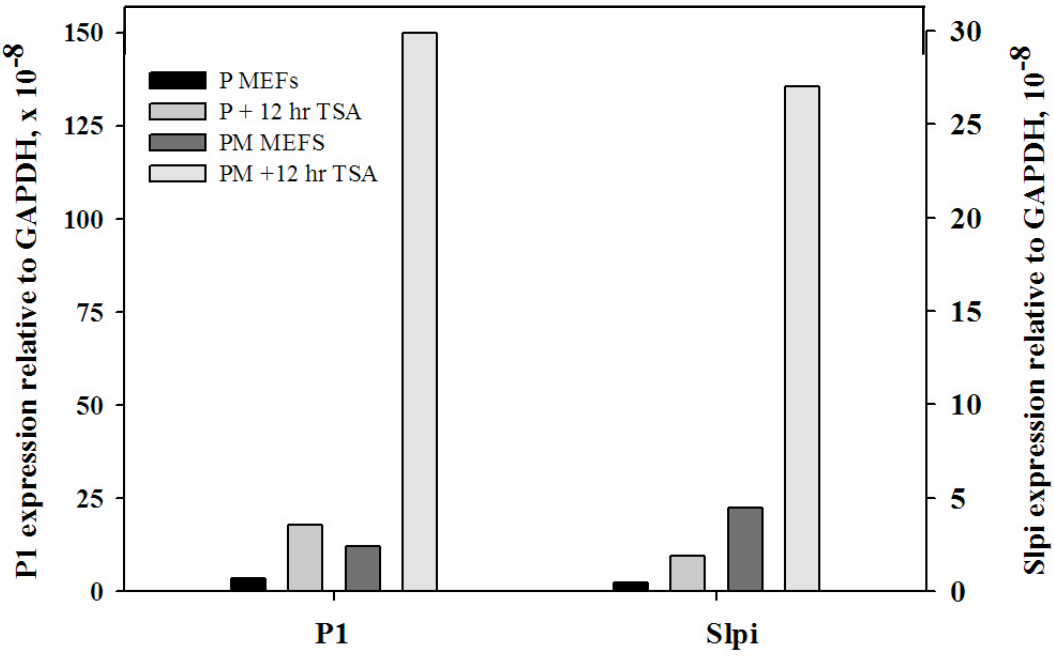


Figure 3-9. Inhibition of histone deacetylases activates transcription at P1 and *Slpi* in P and PM cells. Cells were treated with 0.06 µM TSA for 12 hours. RNA was isolated from P and PM cells and cDNA was synthesized as described in Figure 3-5. The primers used to detect P1-specific *fpgs* cDNA were specific for regions within exon A1b. Semi-quantitative PCR was also performed (data not shown). The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each cell line relative to the amount of GAPDH.

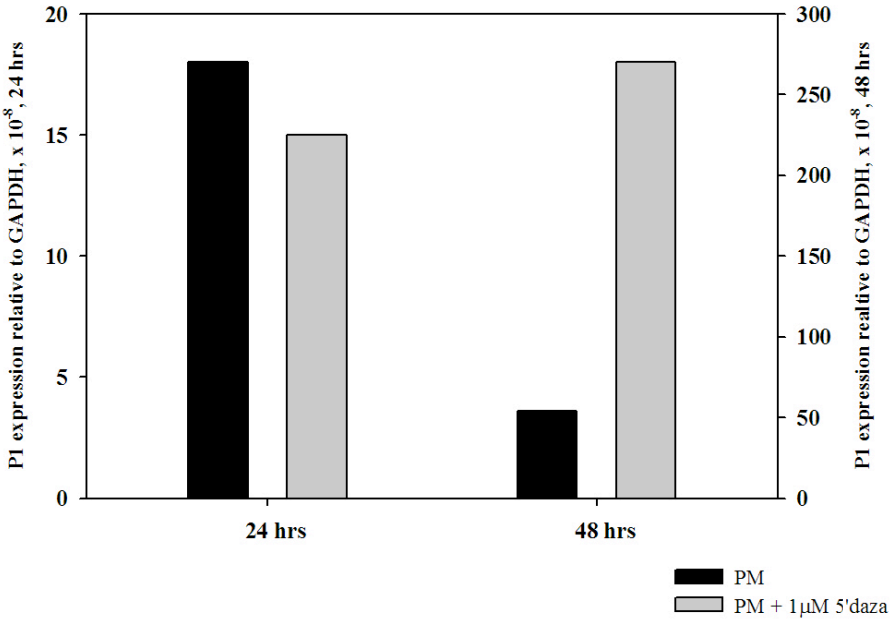


Figure 3-10. P1 transcriptional activity increases in PM cells after 24 and 48 hours of treatment with 5daza. RNA was isolated from PM cells (black bars) and PM cells treated with 1μM 5daza (grey bars) for 24 and 48 hrs. cDNA was generated as described in Figure 3-5. 1 μl of cDNA was added to a 25μl PCR reaction. The level of P1-specific cDNA at 24 hrs was measured using primers specific to regions within exon A1b. cDNA obtained from the 48 hr time point was amplified using a forward primer in exon A1b and a reverse primer in exon 6 of the mouse *fpgs* gene. Similar results were also collected using Semi-quantitative PCR. The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each condition relative to the amount of GAPDH.

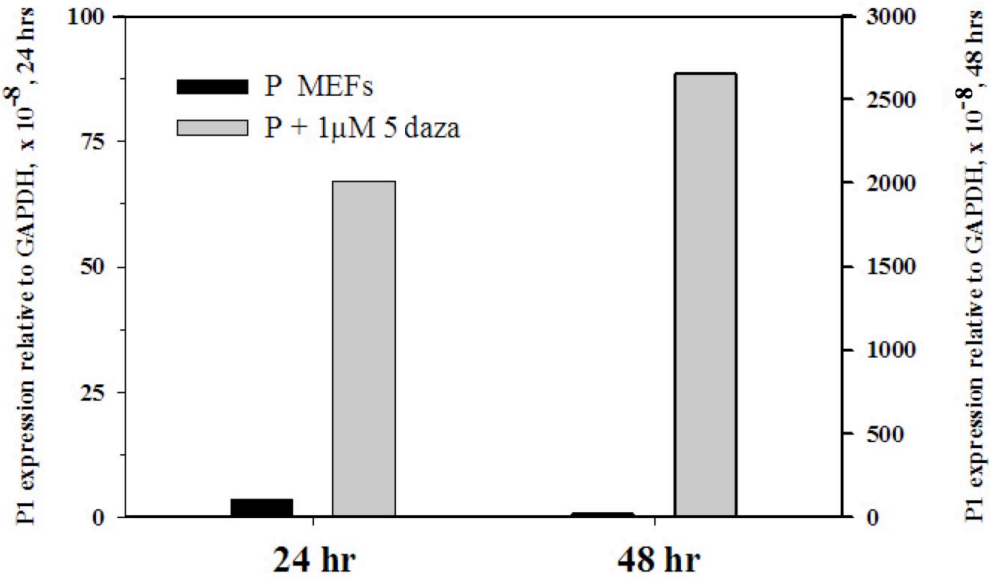
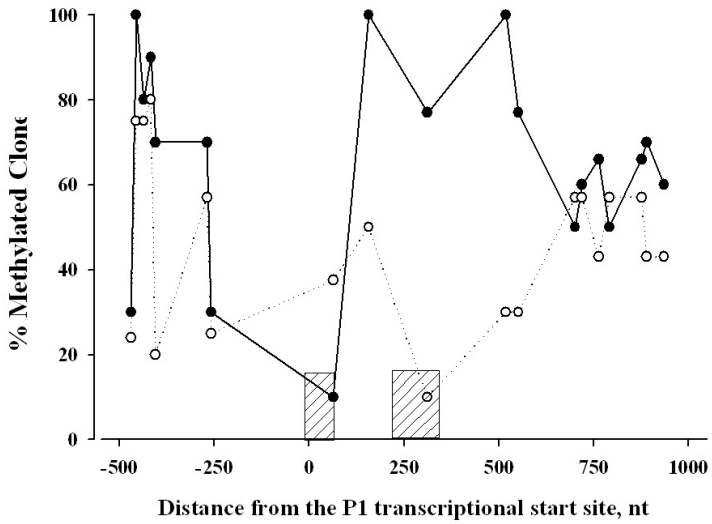


Figure 3-11. P1 transcriptional activity increases in P cells at after 24 and 48 hours of treatment with 5daza. RNA was isolated from P cells (black bars) and P cells treated with 1μM 5daza (grey bars) for 24 and 48 hrs. cDNA was generated as described in Figure 3-7. 1 μl of cDNA was added to a 25μl PCR reaction. The level of P1-specific cDNA at 24 hrs was measured using primers specific to regions within exon A1b. cDNA obtained from the 48 hr time point was amplified using a forward primer in exon A1b and a reverse primer in exon 6 of the mouse *fpgs* gene. Similar results were also collected using Semi-quantitative PCR. The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each condition relative to the amount of GAPDH.

Figure 3-12. CpG methylation across the P1 promoter in PM and P cells treated with 1μM of 5daza for 24 hours. Bisulfite sequencing was carried out as described in Figure 3-6 using genomic DNA isolated from P **(B)** and PM **(A)** cells treated with 1μM 5daza for 24 hrs. Sequences of 6-10 colonies were determined for each amplified region across the P1 promoter in P and PM cells. At the position of each CpG dinucleotide in the genomic sequence, the percentage of clones containing a methylated cytosine is represented in the graph, plotted relative to the positions of exons A1a and A1b (vertical bars).

A.


% Methylated Clone
100
80
60
40
20
0
-500
-250
0
250
500
750
1000
Distance from the P1 transcriptional start site, nt

B.

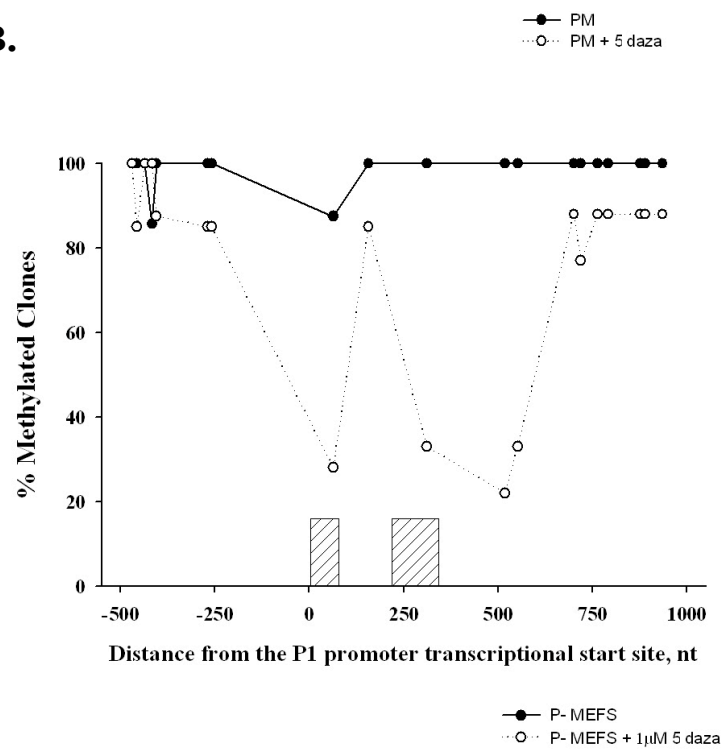
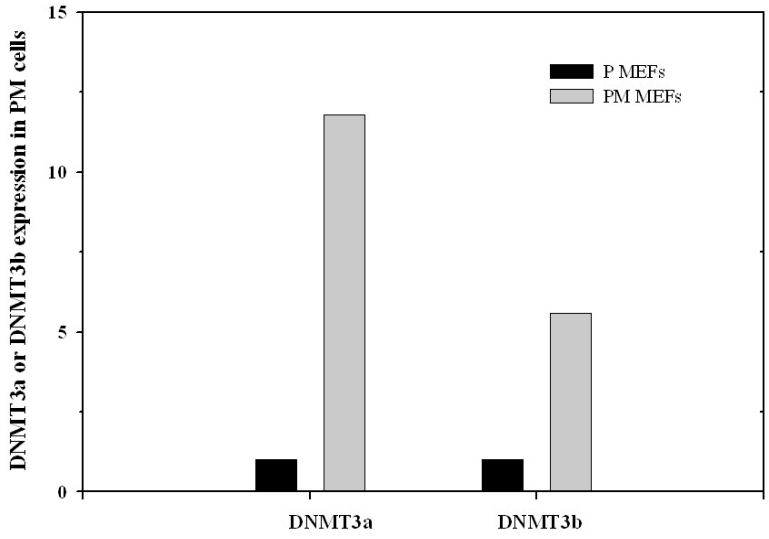
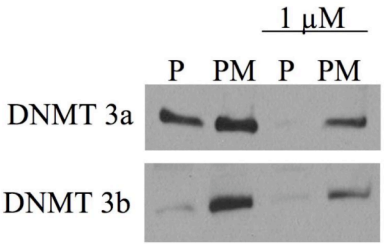

PM
PM + 5 daza
% Methylated Clones
100
80
60
40
20
0
-500
-250
0
250
500
750
1000
Distance from the P1 promoter transcriptional start site, nt
P- MEFS
P- MEFS + 1μM 5 daza

Figure 3-13. DNMT3a and DNMT3b are overexpressed in PM cells compared to P cells. (A) RNA was isolated from P and PM cells and cDNA was synthesized as described in Figure 3-5. The level of GAPDH cDNA was measured as a reference and the ΔCt method was used to assess the amount of gene-specific cDNA in each cell line relative to the amount of GAPDH. (B and C) Whole cell lysates were prepared using 2% SDS lysis buffer. 20µg of total protein were run on a 4-15% SDS-PAGE gel. Proteins were transferred to Immobilon-P PVDF membrane using a sem-dry apparatus. PVDF membranes were blocked in Blotto solution, probed with primary antibody (1:500) overnight, washed with TBST, incubated in with secondary antibody for 1hr, and developed using Pierce chemiluminescence kit.

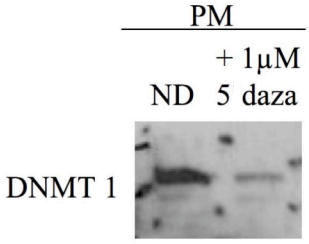
A.


15
10
5
0
DNMT3a or DNMT3b expression in PM cells
P MEFs
PM MEFs
DNMT3a
DNMT3b

B.


1 μM
P
PM
P
PM
DNMT 3a
DNMT 3b

C.


PM
+ 1μM
ND
5 daza
DNMT 1

Chapter 4: Antifolates active against *de novo* purine synthesis have a marked spillover effect on inhibition of mTORC1 and activation of AMPK.

INTRODUCTION

The concept of molecular based therapeutics is founded on the idea that genetic alterations between normal and neoplastic cells are distinguishing features, that when targeted by chemotherapeutic agents will affect only those cells in a cancerous state. In theory, this drug development strategy should result in the generation of compounds that cause substantial regression of tumors, while having little, if any, secondary affects on the surrounding normal tissue or normal stem cells. The cell signaling pathways responsible for the coordination of cellular proliferation depend on the effector molecule, the mammalian target of rapamycin (mTOR), and are rich in proteins commonly dysregulated in human cancers and other proliferative disorders. As such, mTOR is often found unchecked in a number of human pathologies, and a great deal of effort has been directed towards generating direct and indirect inhibitors of this protein as strategies for molecularly targeted therapeutics.

Cell growth, translation, and mTOR

Protein translation is a substantial energy burden to the cell and several signaling pathways responding to growth factor stimulation and nutrient availability converge on the

mTOR protein and coordinate mTOR activity, to ensure that the energy demands of protein synthesis are met by the available cellular resources (149). mTOR is a 289 kD serine/threonine kinase that is a member of the phosphatidylinositol 3-kinase related kinase family (PIKK), sharing structural similarity to ATM, ATR, and DNA protein kinase (78). The amino-terminal domain of the protein is composed of HEAT repeats, which are thought to enhance protein-protein interactions (78). The remaining portions of mTOR include a FAT domain (signature of all PIKK kinases) at the C-terminal end, a FKBP12-rapamycin-binding domain (FRB), and a kinase domain (78). In the cell, mTOR exists in two large multiprotein complexes, mTORC1 and mTORC2, with functionally distinct consequences (Figure 4-1). mTORC2 is composed of mTOR, mLST8, and rictor (208) and its activity has been linked to the phosphorylation of Akt at Ser473 (209) and the control of the actin cytoskeleton (116). The mTORC1 complex is activated in response to abundant cellular nutrients and growth factors and stimulates protein translation; this complex includes the subunits mTOR, Raptor, and mLST8 (129) and is specifically inhibited by rapamycin (see below) (Figure 4-1). Raptor facilitates the positioning of the substrates of mTOR to enhance phosphorylation (174). PRAS40 (Proline-rich AKT substrate) has also recently been shown to bind to and inhibit mTORC1 prior to its phosphorylation by Akt, which causes the release of PRAS40 from the complex (248). The events leading to mTORC1 activation represent an integration of growth factor signaling and nutrient availability (149) (Figure 4-1).

Downstream effectors of mTORC1:

mTORC1 activates translational initiation through the phosphorylation of two key downstream targets 4E-binding protein 1 (4EBP1) and the 40S ribosomal protein S6 kinase (S6K1) (Figure 4-1) (78). Several residues within 4EBP1 (T37, T45, T70, S65) serve as substrates for mTORC1 (78). The hypophosphorylated form of 4EBP1 prevents the assembly of the translation initiation complex through its binding with the elongation initiation factor 4 E (eIF4E) when it is positioned at the 5' cap of mature mRNA (78). Activation of mTOR phosphorylates 4EBP1, disrupting its interaction with eIF4E, and facilitates the activation of cap-dependent protein translation (78) (Figure 4-1). eIF4E is an important mediator of translation of mRNAs involved in cell cycle progress, angiogenesis, and survival such as mRNAs for cyclin D1, Bcl-2, Bcl-xl, and vascular endothelial growth factor (VEGF) (151).

Under basal, non-stimulated conditions the hypophosphorylated form of S6K1 interacts with elongation initiation factor 3 (eIF3) (105). Activation of mTOR leads to the phosphorylation of T389 and the dissociation of the two proteins. The free S6K1 interacts with phophoinositide-dependent kinase 1 (PDK1), which phosphorylates S6K1 at T229, and fully activates the kinase activity of this protein (149). The active S6K1 phosphorylates several targets including insulin receptor substrate-1 (IRS-1), the 40S ribosomal subunit S6, eukaryotic initiation factor 4B (eIF4B), and programmed cell death 4 (PDC4), which is accompanied by the enhancement of translational initiation, ribosome biogenesis and cellular proliferation (149). In particular, the phosphorylation of eIF4B stimulates its interaction with eIF3 and recruitment to the translational pre-initiation

complex (105). The translocation of eIF4B to the initiation complex stimulates the RNA helicase activity of eIF4A, which enhances the translation of transcripts with 5' untranslated regions with high secondary structure; mRNAs of this nature typically encode proteins involved in cell-cycle progression and proliferation (257)(Figure 4-1).

Upstream regulators of mTORC1:

TSC1/TSC2 complex and Rheb

The mTORC1 complex associates with and is directly activated by the GTP-bound form of the small G protein, Rheb (152)(Figure 4-1). The interconversion between a GTP and GDP bound state of Rheb is determined by the GTPase (GAP) activity of the upstream negative regulator of mTORC1, TSC2 (152)(Figure 4-1). TSC2 exists in the cell as a heterodimer with TSC1 and together these proteins are the primary determinant of the cellular activity of mTORC1. Multiple upstream signaling inputs phosphorylate the TSC1/TSC2, dictating the GAP activity of TSC2, and thus mTORC1 (Figure 4-1, see below).

Activators of mTORC1:

A major upstream regulator of mTORC1 activity is the serine/threonine kinase Akt. Akt activity is stimulated by growth factors, such as insulin and epidermal growth factor (EGF), and by the activity of phosphatidylinositol 3-kinase (PI3K) (Figure 4-1) (48). A balance of Akt activity is reached in the cell by the activity of the phosphatase and tensin homolog (PTEN), which dephorylates phosphatidylinositol-3,4,5-trisposphate (PIP3), to phosphatidylinositol-4,5,-bisphosphate (PIP2), preventing the recruitment of Akt to the membrane and activation of the protein by phosphorylation by PDK1 (Figure 4-1). Akt

has several cellular targets that coordinate cell growth and proliferation, one of which is TSC2. The GAP activity of the heterodimer is inactivated by the Akt-mediated phosphorylation of three residues, S939, S981, and T1462, within TSC2 (114). As a consequence, Rheb-GTP levels are increased and mTORC1 activity is stimulated. As discussed previously, recent studies have uncovered an additional link between mTORC1 and Akt through the Akt-mediated phosphorylation of PRAS40 (248). In normal cells, Akt mediated activiation of mTORC1 is tightly controlled by a negative feedback inhibitory loop mediated by the kinase activity of S6K1 (246). S6K1 phosphorylates IRS1 and stimulates the degradation of this protein, limiting the activation of Akt through the PI3K pathway (Figure 4-1) (93, 98). As a result, in some instances drug-induced mTORC1 inhibition has led to the activation of Akt due to the loss of this negative feedback loop (57). There are several additional signaling pathways that stimulate mTORC1 and the major players involved are outlined in Figure 4-1.

Inhibitors of mTORC1

Since the downstream targets of mTORC1 lead to the activation of processes requiring substantial levels of cellular energy it is no surprise that several mechanisms exist to inhibit mTORC1 in cells depleted of energy stores. Original studies suggested that mTOR itself sensed levels of ATP and thus served as a metabolic sensor (52). However, several lines of investigation determined that in a low energy state mTORC1 is inhibited through the activation of the AMP-activated protein kinase (AMPK), also known as the energy sensor of the cell (36). AMPK is activated in response to cellular stress, which is often signaled by an increase in cellular levels of AMP (96). One mechanism of AMPK-

mediated inhibition of mTORC1 is accomplished by phosphorylation of TSC2 on position S1345, which activates the GAP activity of this protein and reduces the activation of mTORC1 through Rheb-GTP (113, 115). Recent studies have discovered that AMPK also directly targets mTORC1 by phosphorylating the Raptor component of the complex (89). This post-translational modification enhances the interaction of Raptor with protein binding partner 14-3-3 and decreases overall mTORC1 activity (89). The importance of the AMPK/TSC/mTORC1 axis in maintenance of cellular survival under stress conditions has been highlighted in studies where cells deficient in either TSC2 or resistant to AMPK-mediated phosphorylation underwent rapid induction of apoptosis in response to conditions of energy depletion (115).

Regulation of AMPK

AMPK exists in the cell as a heterotrimeric complex composed of a α-catalytic subunit and two regulatory subunits, β and γ. Each subunit is encoded by at least two genes (α1, α2, β1, β2, γ1, γ2, γ3), creating the potential for a large number of AMPK isoforms to exist mammal tissues (95). A number of these isoforms have been identified but the functional consequences of the different complexes are not entirely understood. The α-subunits are composed of a conserved serine/threonine kinase domain at the N-terminus and the C-terminal portion facilitates binding to the other two subunits (95). The β subunits complex with α and γ and bind glycogen through the C-terminal domain and central domain of the protein, respectively (95). AMP and ATP bind to the γ subunits through two bateman domains, regions of protein defined to bind molecules containing adenosine, located at the carboxy-terminal end of the protein (95).

LKB1 is the major upstream kinase coordinating activation of AMPK in intact cells through the phosphorylation of AMPK at the residue T172 within the α-catalytic subunit (100, 220). Phosphorylation of this residue is a post-translational modification that is required for the activation of AMPK (101). LKB1 functions in a heterotrimeric complex as a tumor suppressor and is required for the AMPK-mediated inhibition of mTORC1 (100, 219, 220). As such, mouse embryo fibroblasts with mutant LKB1 have elevated levels of mTORC1 activity (219). AMPK activation occurs as a consequence of both phosphoryaltion of the α-subunit and the binding of AMP to the γ-subunit (95). *In vitro* enzyme assays have determined that AMP binding to the γ-subunit both allosterically activates AMPK and prevents dephosphorylation of the catalytic subunit by cellular phosphatases, e.g. PP2Cα (95, 204, 231). Aminoimidazole-4-carboxamide ribonucleotide (ZMP) has also been shown to stimulate AMPK activity in a similar manner to AMP, and administration of aminoimidazole-4-carboxamide ribonucleoside to cells has been used extensively to define the role of AMPK in energy metabolism (49, 231). A second class of upstream kinases of AMPK has been identified, the calmodulin-dependent protin kinases (CaMKKs) (94). However, the CaMKKs are expressed in a highly tissue-restricted manner and appear to respond to levels of Ca^{+2} and not AMP. This is supported by the observation that when Hela cells, deficient in LKB1 but expressing CaMKK, are treated with agents that increase cellular AMP levels the phosphorylation of AMPK is not enhanced (94).

In addition to the affects on mTORC1 activity, AMPK activation is primarily responsible for shifting the metabolic focus of the cell from anabolism to catabolism. As

such, glucose uptake, glycolysis, and fatty acid oxidation are all stimulated, while fatty acid synthesis, gluconeogenesis, and lipolysis are largely inhibited (94). Additionally, another very important consequence of AMPK activation is the induction of a p53-dependent metabolic check-point, which results in cells arrested in the G1 phase of the cell cycle (124). This effect appears to be coordinated by the phosphorylation of p53 at serine 15 and induction of two cyclin-dependent kinase inhibitors, p27 and p21 (124, 145). Interestingly, studies in HCT116 cells suggest that a link between LKB1-AMPK and p53 may be required for the p53-mediated transactivation (268). These data argue that AMPK activation may have substantial impact on the control of rapidly dividing tumor cells, independent of its inhibitory effects on mTORC1.

The mTOR pathway and disease

The upstream and downstream proteins involved in the mTOR pathway are common tumor suppressors or oncogenes that are dysregulated in a number of proliferative disorders and human cancers. For the sake of clarity, the proteins and their role in transformation have been left out of the discussion above and are presented in Table 4-1.

Consequences of mTOR inhibition *in vitro* and *in vivo*

Inhibitors of mTOR have been extensively sought as ideal candidates for molecular targeted therapeuitic agents because of the common phenotype of mTOR dysregulation found in several human cancers. Rapamycin is the prototypical mTORC1 inhibitor that was originally identified as an antifungal agent (25, 217). This agent is bound

intracellularly to the immunophilin, FK506 binding protein-12 (FKBP12), which binds to the FKBP12-rapamycin binding (FRB) domain of mTOR, leading to the selective inhibition of mTORC1 activity and downstream signaling events, with no apparent effects on mTORC2 (78). Currently, rapamycin is clinically used as an immunosuppressant to prevent kidney transplant rejection and is being tested for efficacy in the use of treatment of graft versus host disease (240).

In the early 1980's, rapamycin was evaluated by the developmental therapeutic branch of the national cancer institute and was found to be "a non-cytotoxic agent that had cytostatic activity against several human cancers *in vitro* and *in vivo*"(63). However, the insoluble nature of rapamycin limited the utility of this compound as an anticancer drug and developmental efforts were halted. Over the course of the past fifteen years, second-generation water-soluble rapamycin analogues, termed rapalogs, have been designed and their utility as chemotherapeutic agents explored both *in vitro* and *in vivo* (63). The lessons learned from the pre-clinical and clinical studies using direct mTOR inhibitors have shed light on the cell signaling pathways linked to mTOR activity and the anticancer potential of mTOR inhibitors beyond those directly targeting the enzyme.

In vitro experiments have suggested that a variety of tumor types including glioblastoma, rhabdomyosarcoma, small cell carcinoma, prostate, and breast may be sensitive to mTOR inhibition *in vivo* (56). The determinants of a particular cellular response as an affect of mTOR inhibition appear to depend largely on the genetic background of the treated cell-type (25). A clear example described where a particular genotype predicted the anticancer activity of mTOR inhibition has been shown in studies

in mice and human tumors that are deficient in PTEN; the activity of Akt was elevated in cells where PTEN was deficient and the growth of cells with this genotype both *in vivo* and *in vitro* were preferentially blocked by inhibition of mTOR (171, 185, 221). Similarly, unchecked PI3K signaling, either as a result of PTEN loss or dysregulated growth receptor activity, has been shown to sensitize breast cancer cells to mTOR inhibition (173). These data argue that screening for PTEN status and aberrant PI3K/Akt signaling, two derangements commonly found in human cancers, may identify a subpopulation of patients that will respond to treatment based on mTOR-inhibition.

The cellular response to rapamycin is largely the result of the hypophosphorylation of the two major targets of mTORC1, 4EBP1 and S6K1. Cells exposed to rapamycin display a minor decrease in overall protein translation, however the predominant cellular affect is substantial G1 growth arrest (63). One mechanism involved in coordinating this signature block involves the decrease in cellular proteins tightly linked to cell cycle progression, e.g. cyclin D1 and c-myc, in response to loss of phosphorylation of S6K and 4EBP1 and their downstream signaling (56, 76, 99). This decline in protein levels is accompanied by the stabilization of p27, which inhibits the activity of the cyclin-dependent kinases 4 and 6 and reinforces the G1/S-phase block (127, 176). Likewise, the presence of p53 and p21 has also been linked to the rapamyacin induced growth affects, and cell lines mutant in p53 have been shown to undergo apoptosis, rather than cell-cycle arrest following treatment with an mTOR inhibitor (109).

Tumor growth suppression has also been shown in a number of cases to be a consequence of anti-angiogenesis effects in response to mTOR inhibition. The

microenviornment of tumors tends to be hypoxic and tumor survival and metastasis is often dependent upon ongoing processes of angiogenesis. Mouse models with endothelial cell-specific loss of PTEN have dysfunctional angiogenesis and developing embryos die at E 11.5 as a consequence of bleeding and cardiac failure, linking the PTEN/PI3K/mTOR pathway to the control of vascular development (91). The connection between the mTOR pathway and angiogenesis is defined by the substantial cross-talk between the two processes; mTOR activation induces the expression of HIF1α mRNA that leads to the induction of vascular endothelial growth factor (VEGF), a critical regulator of angiogenesis (85); likewise, VEGF has been shown to stimulate the P13/AKT/mTOR pathway (63, 85). The consequences of this connection on tumor progression were realized when the observation was made that rapamycin treatment inhibited metastatic tumor growth and angiogenesis *in vivo* (85).

Several cases have been described where mTOR inhibition has resulted in the induction of apoptosis in human cancer cell lines (25), but the causative events are not entirely defined. In spite of this fact, it is very clear that the genetic backgrounds of the affected cells play a major role (56). For example, in cells with mutant p53 the tendency towards a G1 arrest following rapamycin treatment is disrupted and apoptosis is initiated (109); this p53 mutant-mediated apoptosis is dependent upon expression of 4EBP1 and loss of p21 (110). Additional studies have also correlated null-PTEN status and Akt hyperactivity with a predisposition to apoptosis following treatment with rapamyacin (56). Further understanding of the signaling events and genetic backgrounds driving cells toward

apoptosis are critical to identifying cancer types where single agent therapy with mTOR inhibitors will cause tumor regression rather than slowed tumor progression.

Potential of mTOR inhibition in the treatment of human cancers

The efficacies of three rapamycin analogs temsirolimus, everolimus, and AP23573 as single agent therapy in the treatment of a number of different cancer types and several proliferative disorders are currently being tested in clinical trials. Success of the rapalogs has been found in the treatment of a variety of lymphomas including refractory mantle-cell lymphoma and non-mantle-cell-non Hodgkins lymphoma, as well as, in the control of facial angiofibromas and renal angiolypomas, common features of profilerative disorders (160). Recently, mTOR inhibitors were approved by the FDA as first-line therapy for poor-prognosis renal cell carcinoma (RCC); phase III clinical trial data showed that the number of patients responsive to treatment with tesmirolimus was double that effected by treatment interferon, which is the current therapy for RCC (160). It has been surprising to clinicians that rapalogs have not shown clinical activity in a substantial number of the tumor types tested. This may be the result of the dominant cytostatic cellular response to mTOR inhibition observed in pre-clinical studies, and it may suggest that the use of these compounds may be best in combination with cyotoxic agents. Pre-clinical studies have shown potential for mTOR inhibitors in combination with radiation (CR) and clinical trials are currently underway to test their utility in a variety of different clinical diseases (160). The full clinical benefit of mTOR inhibitors has yet to be determined, but current data suggests that the responsive tumor types may define certain genetic features, including

PTEN loss and cyclin D1 overexpression, as predictors of sensitivity to mTOR inhibition (160). Thus, the anticancer potential of mTOR inhibitors identified in pre-clinical studies appears to hold true for certain cancers with specific genetic disruptions.

Antifolates targeting *de novo* purine synthesis

As discussed in chapter 1, drug development efforts to identify highly selective inhibitors of folate metabolism lead to the synthesis of the two diasteromers 6R- and 6S- 5, 10-dideaza-5,6,7,8-thetraydrofolate (DDATHF) (237)(Figure 1-4). These compounds were determined to inhibit the first-folate dependent step of *de novo* purine synthesis, glycinamide ribonucleotide formyltransferase (GART) (18, 165)(Figure 1-3). DDATHF was found to be an excellent substrate for FPGS (18) and polyglutamated form targeted GART much more efficiently than the parent compound (205). Mouse leukemic cells treated with DDATHF were growth inhibited and the ATP and GTP levels within these cells were substantially reduced (18); these effects were completely reversed by the addition of a purine source, such as hypoxanthine, to the medium (18). Further studies, determined that the effects of DDATHF on *de novo* purine synthesis caused substantial cytotoxicity on human colorectal carcinoma cells exposed to drug (226). In spite of the potential suggested through pre-clinical data of DDATHF as an anticancer drug, the clinical utility of this compound was limited because it caused severe thrombocytopenia in patients during phase I and II clinical trials (191).

The Eli Lilly compound, 231514 (pemetrexed) (Figure 1-4), was originally designed as a fourth generation *de novo* purine synthesis inhibitor, was identified as a

potent cytotoxic agent and was advanced as clinical candidate following the failure of DDATHF in the clinical setting (224, 238). Pemetrexed is the first antifolate drug approved by the FDA in the past 20 years for the treatment of cancer and is currently approved for first line therapy in combination with cisplatin for non-small cell lung cancer and mesothelioma (61, 148, 211).

As previously discussed, *in vitro* experiments determined that the primary target of this compound was thymidylate synthase, but also found that the folate-dependent enzymes in *de novo* purine synthesis, GART and to a much lesser extent AICART, were inhibited by the drug (224). These studies concluded that the polyglutamated forms of pemetrexed were 50-200 times more potent inhibitors of thymidylate synthase than of either GART or AICART (224). In agreement with this data, the ATP and GTP pools found in CEM cells following treatment with pemetrexed were not depleted, as they were after treatment with DDATHF (45). Taken together, these data imply that the inhibition of the secondary targets by pemetrexed play a minor role in the cellular effects mediated by this compound. However, pemtrexed has acted as an effective cytotoxic agent towards cells that have been confirmed to be resistant to ralitrexed, a pure thymidylate synthase inhibitor, supporting the idea that the multiple targets of pemetrexed are involved in mediating its cellular effects (45). It is clear that the mechanism of action of pemetrexed is more complicated than antifolate compounds targeting a single folate-dependent enzyme, and understanding the cellular consequences has become increasingly important with the recent success of this compound in the clinic. The studies presented in this chapter have

discovered that pemetrexed, through its action on *de novo* purine synthesis, leads to the activation of AMPK and the inhibition of mTOR.

Objectives

We have designed a series of experiments to determine if a link between antifolate therapy and mTOR inhibition exists. More specifically, we hypothesized that changes in nucleotide pools within the cell induced by antifolate compounds inhibiting either of the two folate-dependent enzymes in *de novo* purine synthesis, GART or AICART, result in the inhibition of mTORC1 through the activation of AMPK. To address this question we performed cell culture experiments using DDATHF, the prototypical GART inhibitor, and a pemetrexed analogue found in our laboratory to primarily inhibit AICART. Furthermore, we expanded our findings by including pemetrexed in our experiments and determined a new potential mechanism explaining the antitumor activity of this compound for lung carcinomas, an unusual activity of antifolates. As we will demonstrate in this chapter, pemetrexed exerts a strong inhibitory effect on mTOR as an indirect effect of AICART inhibition. This is the first time such an effect has been found.

MATERIALS AND METHODS

Materials

Antibodies were obtained from Cell Signaling Technology (Beverly, MA) and will be discussed individually in the sections to follow. Compound C (P54996-[4-(2-piperidin-1-yl-etoxy)-phenyl)]-3-pyridin-4-yl-pyrrazolo[1,5-a]pyrimidine) was purchased from Sigma-Aldrich. Starting Block Buffer (37542) was purchased from Pierce. 10x Tris-glycine transfer buffer was obtained from Bio-rad (161-0734). Goat anti-rabbit secondary antibody was purchased from Thermo-scientific (35560). SuperSignal West Pico and Est Dura Chemiluniescent Substrate Kits were purchased from Pierce. Pemetrexed, 6R-DDATHF, and Lilly AICART inhibitor (LCA) were obtained from Eli Lilly and Co (Indianapolis, IN). All other culture reagents were from Sigma Aldrich and were of highest available quality.

Cell culture

CCRF-CEM human lymphoblastic leukemia cells were maintained at a density between 10^5-10^6 cells/ml in RPMI 1640 medium (Gibco/Invitrogen) supplemented with 10% dialyzed fetal bovine serum. HCT116 human colorectal carcinoma cells were grown in RPMI 1640 medium (Gibco/Invitrogen) supplemented with 10% dialyzed fetal bovine serum. All cells were grown at 37^0C in 15% CO_2.

Drug treatment of HCT116 cells or CEM cells.

HCT116 cells were plated at 2.5 x 10^6 cellsl in 10 mls in 100 mm dish and were to allowed cells to adhere to the plate overnight. Ten milliliters of fresh media were added to the plates and drug was added simultaneously with the appropriate amount of reversal agents. Pemetrexed, DDATHF, LCA, thymidine, AICA, AICAR, and inosine were prepared in PBS and sterilized filtered prior to use. Hypoxanthine was prepared in 75 mM HCl. Compound C was dissolved in DMSO and diluted into PBS. The cells were exposed to drug for 24 or 48 hours and harvested as described in chapter 3. CEM cells were treated at a density of 1 x 10^5 cells/ml and, during treatment, the media was supplemented with 20 mM HEPES and 40 mM MOPS. Since CEM cells are grown in suspension, at the time of harvesting they were pelleted by centrifugation and then washed and lysed as described in chapter 3.

Reversal experiments using HCT116 cells.

Cells were plated in 12-well plates at a concentration of 2 x 10^4/ml in 2 ms of media. The cells were incubated overnight to allow cells to become adherent. The following day fresh media was added to the plates with the appropriate amount of drug and rescue agents. The plates were examined every 24 hours microscopically cell morphology and death were noted. After 96-hours of treatment, the supernatant was aspirated and morphology washed the wells with 1 ml of 1x PBS. Each well was trypsinized and the cells were pelleted by centrifugation. The cell pellets were resuspendend in 1 ml of PBS and 500 µl of cell suspension was added to 9.5 ml of PBS for counting using the Beckman

Coulter Counter. Each condition was performed in duplicate and the average cell number was plotted.

Immunoblot analysis.

Cells were lysed in buffer containing 62.5mM Tris, pH 6.5, 5% glycerol, 2% SDS, 5% 2-mercapothethanol, 50mM NaF, 0.2 mM Na_3VO_4 and 1x protease inhibitor complete mixture (Roche Applied Science). Protein concentrations were determined using the Bradford assay, against a standard of BSA using reagents from Bio-rad laboratories. Total cellular protein CEM cells (40 μg) and HCT116 cells (20 μg) were resolved on 7.5%, 4-15%, and 15% SDS-polyacrylamide gels and were transferred to an Immobilon-P polyvinylidene fluoride membrane (Millipore) using a protocol for wet-transfer in Tris/Glycine transfer buffer Biorad. Western blot analysis using some antibodies and cellular lysates obtained from CEM cells were challenging, and we found that increasing the total level of protein loaded to 40 μg helped in obtaining successful blots. Prior to transfer, the protein gels and PVDF membrane (which was dipped first in methanol) were rinsed in water and then soaked in cold 1x transfer buffer prepared using the manufacturer's protocol (Biorad) for 30 minutes. The pre-cast mini-gel transfer apparatus from Biorad was used to carry out the wet transfer. The "transfer sandwich" was prepared as follows: a sponge pre-soaked in transfer buffer was placed on the white plate, followed by three pieces of whatman paper, the PVDF membrane, the gel, three additional sheets of Whatman, and a second sponge. All items in the sandwich were soaked in transfer buffer prior to assembly. The transfer was run at 100 V for 1 hour. Membranes were blocked with either 5% milk or

Starting Block Buffer (Pierce), washed, and probed with antibodies against AMPKα (1:1000) (cell signaling, #2532), phospho-AMPKα (T172) (cell signaling, #2531)(1:250), 4EBP1 (9452) (1:1000), phospho-4EBP1 (T70) (cell signaling, #9455) (1:1000), S6 kinase (1:1000) (9202), phospho-S6 kinase (T389) (cell signaling #9205) (1:250 and 1:500), acetyl CoA carboxylase (1:1000) (cell signaling #3662), Raptor (cell signaling #24C12) and phospho-Raptor (S792) (cell signaling #2083) (1:1000), phospho-acetyl-CoA carboxylase (S79) (cell signaling #3661) (1:1000) at 4^0 C overnight. All antibodies were purchased from Cell Signaling. Washes were carried out as described in chapter 3, but used 0.1% TBST. This change usually decreased the background on these blots. Membranes were incubated with goat anti-rabbit secondary antibody with horseradish peroxidase conjugate for 1hr at room temperature (Pierce). Chemiluminescence was detected using the SuperSignal West Pico and West Dura Chemiluminescent Substrate Kits (Pierce). Washing with these antibodies is critical and if upon developing the blots were dirty we repeated washing 3x for 5 minutes or longer and re-developed. Below are the conditions we found to work best for these antibodies. Unfortunately, changes in cell lines may result in the need for re-optimization. If these do not work, for another cell line, it is suggested than an optimization experiment should be performed with several conditions, varying primary and secondary concentrations, as well as diluting solutions. The differences that small changes make in the signal obtained from a particular antibody are remarkable. For optimization of antibodies, I always started out with the least sensitive developing agent, i.e. West Pico, and then went to more sensitive reagents as needed. West dura is very strong and diluting it 1:5 has proven to also be very useful.

Table 4-1: Antibody conditions for the total and phosphorylated forms of AMPK, 4EBP1, S6K1, and ACC

Antibody	Blocking	Primary	Secondary	ECL
ACC and phospho-ACC	5% milk	1:1000 5% BSA	1:2000 5% milk	Pico
Phospho-S6K1	Startblock	1:250 Startblock	1:2000 Startblock	Pico
S6K1	Startblock	1:1000 Startblock	1:2000 Startblock	Pico
Phospho-4EBP1	5% milk	1:1000 5% BSA	1:10,000 5% milk	Pico
4EBP1	Startblock	1:1000 Startblock	1:2000 Startblock	Pico
Phospho-AMPK	5% milk	1:250 5% BSA	1:10,000 5% milk	Dura
AMPK	5% milk	1:1000 5% BSA	1:2000 5% milk	Pico 30 sec-1 min

RESULTS

The direct link between mTOR dysregulation and a larger number of human cancers has generated a great deal of interest in identifying therapeutic strategies that target the mTOR complex either directly or indirectly. The studies described in this chapter were initially designed to test if *de novo* purine synthesis inhibition by antifolate compounds targeting the folate-dependent enzymes, GART or AICART, caused inhibition of the mTOR pathway. Treatment of mammalian cells with inhibitors of *de novo* purine synthesis has been shown to cause depletion of ATP and GTP pools and as a result induced rapid growth arrest (18). Previous studies have identified AMPK, a kinase whose activity leads to the inhibition of the mTORC1 complex, as a sensor of cellular energy that becomes activated when ATP levels decline. We hypothesized that nucleotide imbalance induced by a *de novo* purine synthesis blockade may lead to the activation of AMPK and inhibition of the mTORC1 complex, which would represent a novel and clinically relevant consequence of antifolate compounds targeting this biochemical pathway.

Targets of 6R-DDATHF and LCA by end-product reversal experiments.

End-product growth-inhibitory reversal experiments have been used in several instances to define the folate-dependent enzymes targeted by antifolate compounds. Growth rescue by inclusion of thymidine (TdR) or hypoxanthine (Hx) distinguishes between antifolates that primarily target thymidylate synthase and *de novo* purine synthesis, respectively. In order to determine whether the first or second folate dependent enzyme of *de novo* purine synthesis, GART or AICART, is affected by drug, the effects of

aminoimidazolecarboxamide (AICA) on the growth of cells exposed to drug is assessed. AICA is metabolized to the corresponding ribonucleotide, AICAR monophosphate (also known as ZMP) by adenine phosphoribosyltransferase (APRT) and, thus, inclusion of AICA in growth medium introduces a purine pathway intermediate into the cell that is downstream of GART, but upstream of AICART (Figure 4-3). If GART is inhibited, inclusion of AICA in the medium should reverse growth inhibition (18, 165) whereas, if AICART is inhibited, no effect on growth would be expected.

We were interested in testing the effect of GART and AICART inhibition on the mTOR pathway. Prior studies have shown that the major site of action of 6R-DDATHF is the first folate-dependent step in *de novo* purine synthesis, GART (18, 165). We had several antifolate compounds in house and Dr. Moran performed a series of end-product reversal experiments to identify compounds targeting AICART. In these experiments the data generated using 6R-DDATHF represents the results expected of a pure GART inhibitor (Figure 4-2). Growth inhibition by 6R-DDATHF was not affected by the addition of thymidine to the medium, but was completely reversed by the addition of hypoxanthine (Figure 4-2). Cells treated with 6R-DDATHF supplemented with AICA (320 μM) are substantially less growth inhibited than cells grown in drug alone (Figure 4-2), in agreement with concept that this compound is a GART inhibitor. The inclusion of thymidine in the medium of cells treated with the compound Lilly AICART inhibitor (LCA) did not affect cell growth, but hypoxanthine partially rescued the growth inhibitory effect (Figure 4-2). However, the combination treatment of hypoxanthine and thymidine completely rescued cells treated with LCA (Figure 4-2). Taken together, these data

suggest that this compound targets primarily *de novo* purine synthesis and secondarily thymidylate synthase. The addition of AICA with or without thymidine in growth medium had no effect on cell growth, identifying AICART as the enzyme targeted by LCA in *de novo* purine synthesis.

Recent studies performed by a current student in the laboratory, Scott Rothbart, have supported the idea that LCA inhibits AICART by showing that abundant levels of ZMP, the substrate of the AICART reaction (Figure 4-3), accumulates to high levels (1-4 mM) in both CEM and HCT116 cells treated with 1μM of LCA. The accumulation of ZMP behind the AICART block added an additional link between LCA and mTOR inhibition outside of the realm of nucleotide depletion; the expansion of cellular ZMP pools following treatment with the ribonucleoside analoge of AICA, AICAR, has been shown to activate AMPK, resulting in mTORC1 inhibition, and substantial cytotoxic effects (49, 88). ZMP is capable of binding to the γ-subunit of AMPK similar to AMP, which supports the activation of AMPK by enhancing the phosphorylation of T172 within the α-catalytic subunit of the protein (95). Thus, it appeared that LCA had potential to activate AMPK kinase through two mechanisms: depletion of ATP and accumulation of ZMP pools within cells.

We had in our hands two compounds, 6R-DDATHF and LCA, that targeted the first and second-folate dependent steps in *de novo* purine synthesis, respectively. We were in a unique position to question if treatment with these compounds impacted the mTOR pathway and if cellular effects of nucleotide depletion and ZMP accumulation were

different. The steps of de novo purine synthesis blocked by 6R-DDATHF and LCA are illustrated in Figure 4-3.

End-product reversal studies of HCT116 cells treated with LCA.

In our initial studies we repeated end-product reversal studies with LCA using HCT116 cells and the data from these experiments are shown in Figure 4-4. The growth of HCT116 cells was partially inhibited by LCA at a concentration of 10 nM and was completely blocked when treated with 100 nM or higher concentrations of drug (Figure 4-4). The growth inhibition observed in cells treated with LCA was not affected by the addition of thymidine (TdR) (5.6 μM) to the medium, which would be expected to reverse the cellular affects attributed to thymidylate synthase inhibition. However, inclusion of inosine (IR) (100 μM), a nucleoside source of purines, substantially lessened the level of growth inhibition observed when compared to the growth of cells treated with drug alone or with TdR (Figure 4-4). In contrast to the studies performed in CEM cells, the addition of both TdR and IR to the medium did not enhance the reversal of growth inhibition by IR alone, suggesting that any potential secondary effects of LCA on thymidylate synthase were minimal and not contributing to the growth patterns observed of the treated HCT116 cells (Figure 4-4).

Inhibition of *de novo* purine synthesis blocks mTOR activation.

HCT116 colorectal carcinoma cells were exposed to 6R-DDATHF and LCA to question if mTORC1 inhibition was a consequence of the *de novo* purine synthesis

blockade mediated by these compounds. We studied the activity of mTOR by monitoring the phosphorylation status of two primary targets of the mTORC1 complex S6K1 and 4EBP1 (Figure 4-5). mTOR-mediated phosphorylation of S6K1 and 4EBP1 results in an overall increase in protein translation and cellular growth (148). A 1μM concentration of 6R-DDATHF and LCA inhibited mTORC1 activity in HCT116 cells, demonstrated by the reduced level of phosphorylation detected in western blot analysis using phospho-specific antibodies against T389 of S6K1 and T70 of 4EBP1 (Figure 4-6). The effect of LCA on mTOR appeared to be more substantial than that of 6R-DDATHF, since phosphorylation of both S6K1 and 4EBP1 were completely eliminated by the former and only partially reduced by the latter treatment (Figure 4-6). Likewise, the lower migration of total 4EBP1 found in cells treated with LCA suggested that the phosphorylation of the additional mTOR-sensitive residues in this protein may also be more affected by LCA than 6R-DDATHF (Figure 4-6). Interestingly, the effect of LCA on mTORC1 activity was also more potent than that observed when HCT116 cells were treated with AICAR, a known inhibitor of mTOR through the activation of AMPK (49).

mTORC1 inhibition is specific to the effects of 6R-DDATHF and LCA on *de novo* purine synthesis.

We tested if the effects of LCA and 6R-DDATHF observed on the mTORC1 pathway were reversed by the addition of a purine source, such as inosine, to the medium containing drug (Figure 4-7); inosine is converted to inosine mono phosphate (IMP) in cells through the action of inosine kinase and serves as the substrate for synthesis of

adenosine mono-phosphate (AMP) and guanosine mono-phosphate (GMP) (Figure 4-3). Since the growth reversal studies showed that LCA inhibited thymidylate synthase as a secondary target (Figure 4-2), we also included thymidine in the media of cells treated with this compound to separate out any effects of thymidylate synthase on the mTOR pathway (Figure 4-7). The hypophosphorylation of 4EBP1 and S6K1 in response to 6R-DDATHF treatment was completely eliminated when cells were grown in the presence of inosine (Figure 4-7); likewise LCA-mediated mTORC1 inhibition appeared relieved by the presence of purines in the medium (Figure 4-7). Initially, we took this to suggest that the contribution of ZMP accumulation towards the inhibition of mTORC1 in cells treated with LCA was minimal, since repletion of nucleotide pools was sufficient to rescue the observed effects. However, studies discussed later in this chapter demonstrated that the addition of a purine source to the medium of cells treated with an AICART inhibitor prevented the accumulation of ZMP behind the block, most likely as a consequence feedback inhibition mediated by AMP, GMP, or IMP on the rate-limiting enzyme in the *de novo* purine synthesis pathway, *glutamine PRPP amidotransferase*. Thus, the effects observed on mTORC1 activity in cells exposed to LCA were either due to ATP depletion, ZMP accumulation, or a combination of the two. Substantial hypophosphorylation of the mTORC1 targets S6K1 and 4EBP1 was retained in cells treated with LCA and thymidine (Figure 4-7), suggesting that the effects observed on mTORC1 were specific to *de novo* purine synthesis inhibition. These studies confirmed our hypothesis that a *de novo* purine synthesis block results in mTORC1 inhibition, and suggested that treatment with LCA has more potent consequences on mTORC1 activity than 6R-DDATHF. We had yet to

determine the mechanism of this cellular affect, but considered the activation of AMPK as a likely candidate.

AMPK is activated when *de novo* purine synthesis is inhibited in HCT116 cells.

We assessed if the proposed mechanism of AMPK activation was responsible for the inhibition of mTORC1 observed when HCT116 cells were treated with either 6R-DDATHF or LCA (Figure 4-8) AMPK activation in these studies was monitored by assessing the phosphorylation of the α-subunit at position T172, a requisite post-translational modification for the induction of the kinase activity of this protein (101). The phosphorylation of AMPK was stimulated when cells were exposed to either 6R-DDATHF or LCA for a period of 24-hours (Figure 4-8). In agreement with the rescue of mTORC1 activity (Figure 4-7) observed in cells grown in the presence of both inosine and 6R-DDATHF, cellular levels of phospho-AMPK remained low under these conditions (Figure 4-8). Interestingly, this was not the case in cells exposed to LCA and inosine; the phosphorylation of AMPK was detected at equal levels in cells grown in medium containing LCA with or without the addition of thymidine or inosine (Figure 4-8). We are not yet clear how mTOR activity persists in spite of the fact that AMPK is phosphorylated when cells are treated with LCA and inosine, but measurement of AMPK activity under these conditions may shed light on this mechanism.

mTOR inhibition occurs in cells treated with pemetrexed, a novel consequence of this antifolate compound.

The impact of treatment with 6R-DDATHF and LCA on the mTOR pathway was very interesting to us because of the therapeutic benefits direct inhibitors of mTOR have had in the treatment of several human cancer types in clinical trials (160). The notion that antifolate compounds targeting *de novo* purine synthesis could lead to the inhibition of mTOR had not previously been described in the literature, but the clinical implications of this observation on the use of this class of antifolates in the treatment of several human malignancies was potentially huge. At this point in our experiments, we focused our attention on determining if treatment with pemetrexed, a multi-targeted antifolate, currently used as first line therapy for mesothelioma and non-small cell lung cancer caused the inhibition of mTOR. The primary target of this drug is thymidylate synthase, however *in vitro* enzyme assays and cell culture experiments have shown that the secondary target of this compound is *de novo* purine synthesis. Understanding the potential mechanisms responsible for the anti-tumor activity of a drug in widespread clinical use is of obvious importance.

The studies described in the sections to follow were performed in collaboration with two current students in the laboratory, Scott Rothbart and Cortney Heyer. In our initial experiments, we questioned if pemetrexed-mediated inhibition of *de novo* purine synthesis caused activation of AMPK and inhibition of mTORC1, similar to the effects mediated by 6R-DDATHF and LCA on this cell-signaling pathway (Figure 4-9). The activation of AMPK by pemetrexed and thymidine was assessed in CEM cells after 24 and

48 hours of exposure. The western blot data presented in Figure 4-9 were generated from whole cell lysates harvested after 48 hours of treatment. We monitored the phosphorylation of the α-subunit of AMPK as an indication of the kinase activity of this protein, but we also used the phosphorylation status of a direct target of AMPK, acetyl-CoA carboxylase (ACC), as a direct indicator of the activity of AMPK in intact cells (258). A 1μM concentration of pemetrexed promoted activation of AMPK as demonstrated by both direct phosphorylation at T172 and by an enhanced phosphorylation of ACC at residue S79 (Fig. 4-9). Activation of AMPK is known to cause inhibition of mTOR by phosphorylation of TSC2, upstream of mTOR, as well as a direct phosphorylation of the Raptor component of the mTORC1 complex (Figure 4-5 and 9) (89, 115). Accordingly, a robust hypophosphorylation of 4EBP1 and S6K1 was seen in pemetrexed-treated CEM cells (Fig 4-9), as well as an increase in phosphorylation of S792 of Raptor (Figure 4-9). As in our previous experiments using LCA, we observed a lower and broader migration pattern of total 4EBP1 in pemetrexed-treated cells, suggesting that phosphorylation status of this protein was severely affected by drug treatment (Figure 4-9). Thymidine was included in these experiments to separate the effects of inhibition of thymidylate synthase from any effect caused by *de novo* purine synthesis inhibition. The effects of pemetrexed on activation of AMPK and inhibition of mTOR, as observed by the phosphorylation of AMPK and the hypophosphorylation of the downstream targets of mTOR were dose dependent starting at concentrations as low as 0.1 μM of pemetrexed (Figure 4-9). These studies determined that treatment with pemetrexed in the presence of thymidine resulted in AMPK activation and mTORC1 inhibition.

The second target of pemetrexed is the second folate-dependent step in *de novo* purine synthesis, AICART and results in ZMP accumulation.

We sought an understanding of the mechanism behind the AMPK activation observed and to prove if the mTORC1 inhibition detected was, in fact, a direct consequence of AMPK activity. In order to address these processes we needed to know if pemetrexed inhibited GART, and was most likely causing nucleotide depletion, or AICART, resulting in accumulation of ZMP levels. The field was conflicted on this point: i*n vitro* enzyme assays suggested that the second target of pemetrexed was GART (224). Dr. Moran performed growth inhibition reversal experiments in CEM cells to determine whether the second target of pemetrexed was GART or AICART (Figure 4-11). This growth inhibition of CEM cells by pemetrexed was not reversed by hypoxanthine (Hx), but inclusion of thymidine in the culture medium did reverse the growth inhibitory effects of pemetrexed, shifting the concentrations of drug needed to affect growth by about 12-fold, as shown for the CEM human leukemia cell line in Figure 4-11. Inclusion of both thymidine and hypoxanthine reversed the effects of pemetrexed at even high concentrations (Figure 4-11), in agreement with previous literature (224, 238). The critical piece of data in this experiment was that the addition of AICA (320 μM) in the presence of thymidine did not reduce the growth inhibitory effects of pemetrexed. We concluded that the secondary target of pemetrexed was not GART, and was most likely AICART.

These observations were further expanded when studies performed by Scott Rothbart determined that the ZMP pools increased substantially in pemetrexed-treated cells, further supporting the idea that an effect of pemetrexed or its metabolites was

predominantly on AICART (Figure 4-11). If inhibition of GART played a role in the secondary effects of pemetrexed, ZMP accumulation would not have been observed (Figure 4-11). To this point, ZMP accumulation did not occur in 6R-DDATHF-treated CEM cells (Figure 4-11). The level of ZMP in pemetrexed treated cells accumulated in a dose-dependent fashion and was maximal after 48 hours of treatment, reaching cellular levels approximating 1.5 mM.

ZMP accumulates in pemetrexed treated cells correlates with the degree of AMPK activation and mTOR inhibition observed in treated cells.

We initially considered the possibility that the AMPK activation observed in cells treated with pemetrexed and thymidine was a consequence of the depletion of nucleotide pools by AICART inhibition. However, direct measurment of ATP levels in studies performed in our laboratory by Scott Rothbard and others by hplc indicated that the levels of ATP did not decrease, but actually increased slightly in cells treated with pemetrexed (data not shown). In light of these data, it appeared that the activation of AMPK observed in the presence of pemetrexed and thymidine was solely linked to the level of ZMP accumulation forced by the pemetrexed-induced AICART block.

We took advantage of the fact that AICA enters the *de novo* purine synthesis pathway after being converted to ZMP through the activity of APRT (Figure 4-3) to further increase the degree of ZMP accumulation found under conditions where the AICART enzyme was inhibited by pemetrexed. The administration of AICA (320 μM) for 48 hours caused the levels of ZMP in pemetrexed treated carcinoma cells to rise to concentrations

approaching 4 mM, whereas the levels in cells exposed to AICA in the abscense of an AICART block were undetectable (data not shown). We generated whole cell lysates from cells treated under these conditions for 48 hours and found that the expansion of the ZMP pool observed appeared to enhance the effects of pemetrexed on AMPK activation and mTORC1 inhibition (Figure 4-12). The phosphorylation of the direct targets of AMPK, ACC and to a greater extent Raptor, were increased by the addition of AICA to the medium containing pemetrexed and thymidine (Figure 4-12). Likewise, the hypophosphorylation of 4EBP1 was markedly reduced in AICA treated cells, as evidenced by the low level of protein detected using the phospho-specific antibody generated against the T70 residue, and the broader migrating species detected with the pan-4EBP1 antibody (Figure 4-12). Since the hypophosphorylation of S6K1 was substantial in pemetrexed treated cells, a further loss of phosphorylation as a consequence of an increase in ZMP levels was difficult to detect (Figure 4-12). These data greatly supported the notion that the activation of AMPK and the inhibition of mTORC1 observed in pemetrexed-treated CEM cells were directly correlated with the degree of ZMP accumulation behind the AICART block.

The effect of pemetrexed on the AMPK-mTOR pathway is prevented when the levels of ZMP are forced to remain low.

In our studies using 6R-DDATHF and LCA we determined that the addition of purines reversed the inhibition of mTORC1 (Figure 4-7). We questioned if a source of preformed purines, such as hypoxanthine, would also reverse the effects of pemetrexed on

the AMPK-mTOR pathway (Figure 4-13). Cells were treated with pemetrexed and thymidine with or without hypoxanthine (32 μM) for 48 hours. The level of phosphorylation of ACC and Raptor detected in cells treated with pemetrexed and thymidine with the addition of hypoxanthine were not increased above the levels found in the control samples treated with thymidine and hypoxanthine only (Figure 4-13), indicating that the activation of AMPK induced by pemetrexed was prevented when hypoxanthine was present in the medium. As would be predicted, the abscense of detectable AMPK activity was coincident with loss of mTORC1 inhibition (Figure 4-13); the hypophosphorylation caused by pemetrexed of S6K1 and 4EBP1 was completely prevented by the inclusion of hypoxanthine in the medium (Figure 4-13). ZMP measurements confirmed that the mechanism of this hypoxanthine reversal of the secondary target of pemetrexed was most likely due to the prevention of the accumulation of ZMP, presumably due to the known feedback effects of an expanded purine pool on early steps in purine synthesis (39, 260).

mTOR inhibition observed in pemetrexed specific cells is dependent upon AMPK activation.

The accumulation of ZMP in pemetrexed-treated cells and the activating effects of ZMP on AMPK suggested that the inhibition of mTOR by pemetrexed was mediated by ZMP-dependent activation of AMPK. To determine the intermediacy of activation of AMPK in this mechanism, the effects of an inhibitor of AMPK, compound C (20), on the pemetrexed-induced mTOR inhibition was studied (Figure 4-14). The experiment was

performed at two concentrations of compound C, 10 μM and 1 μM. Cell treated with 10 μM of compound C and pemetrexed were substantially growth arrested, and at this concentration compound C appeared to have substantial effects on the mTOR pathway as a single agent, as judged by the hyperphosphorylation of S6K1 (Figure 4-14). When a 1 μM dose of compound C was used, secondary effects on cell growth and the mTOR pathway were undetectable (Figure 4-14). As expected, compound C inhibited the enhanced AMPK activity in cells treated with pemetrexed and thymidine, as judged by an inhibition of the phosphorylation of ACC (Figures 4-14). The effects of 10 μM and 1 μM compound C on AMPK blocked the inhibition of mTOR by pemetrexed, as judged by the hyperphosphorylation of S6K1 in the presence of pemetrexed and compound C (Figure 4-14), a striking contrast to the marked hypophosphorylation of S6K1 caused by pemetrexed. Interestingly, the phosphorylation of AMPK in the presence of pemetrexed and compound C was greater than that seen with pemetrexed, in spite of the fact that compound C diminished the level of ACC phosphorylation; this suggested that compound C was blocking the activity of AMPK while allowing phosphorylation of the α-subunit of AMPK. Overall, we concluded that the effects of pemetrexed on the mTOR signaling pathway were caused by the activation of AMPK following the accumulation of ZMP.

DISCUSSION

The search for new therapeutic agents useful against cancer has focused on molecularly-targeted small molecules. Such strategies have rarely been identified, although the Gleevec-sensitive BCR-Abl and Iressa-selective mutations in the EGFR protein are two

well-studied cases (55, 182). One of the most promising approaches seems to be the design of therapeutic agents that affect pathways dependent on tumor suppressor genes, whose function is often eliminated or dramatically altered during transformation. The mTOR pathway responsible for the balance of energy metabolism, protein and lipid synthesis, and growth involves a series of upstream controlling proteins recognized as tumor suppressor proteins, including LKB1, PTEN, TSC1 and 2, and others recognized as cellular oncogenes, such as AKT and PI3 kinase (48).

In the studies presented in this chapter we identified a new action of a select class of antifolates that inhibit *de novo* purine synthesis: inhibition of mTORC1. Our studies showed that a block of either the first or second folate-dependent steps of purine synthesis results in an inhibition of mTORC1 that is prevented by supplementation with a purine source or inhibition of AMPK activity. Thus, the effect on mTORC1 observed appears to be a direct consequence of AMPK activation in response to changes in purine nucleotide pools following the inhibition of the GART or AICART enzymes. Furthermore, we determined that the activation of AMPK under GART and AICART inhibition occurred by seemingly different mechanisms. Previous studies have determined that GART inhibition results in a reduction of ATP and GTP levels in CEM leukemic cells (18) and HCT116 cells (Woodard and Moran, unpublished data), and we attribute the stimulation of AMPK following treatment with 6R-DDATHF, a pure GART inhibitor, to this consequence. On the other hand, AICART inhibition results in the accumulation of its substrate ZMP, which itself is an activator of AMPK by mimicking the effect of AMP on the γ-subunit of AMPK (94, 231). A rather striking observation was made with the discovery that pemetrexed, a multi-targeted

antifolate currently in clinical use for the treatment of mesothelioma and non-small cell lung cancer, has as one of its targets mTORC1. The ramifications of this observation in the clinic have yet to be fully determined, but those studies are currently underway.

GART inhibition vs. AICART inhibition

We were able to observe the effects of both ATP depletion and ZMP accumulation on the AMPK-mTORC1 pathway through the use of GART and AICART inhibitors, respectively. Previous studies have determined that ZMP is an allosteric activator of AMPK, in a manner similar to AMP, but is substantially (20x) less potent than the adenine nucleotide (49). This supports the idea that an inhibitor causing ATP depletion and a change in the AMP:ATP cellular ratios would be a more potent activator of AMPK than compounds causing the expansion of ZMP pools. Rather surprisingly, we found that mTORC1 inhibition as a consequence of an AICART block appeared more effective than that observed following GART inhibition, as evidenced by a more substantial decrease in phosphorylation of 4EBP1 and S6K1 under the former conditions (Figure 4-6). We took this as an indication that AMPK activation was higher in AICART treated cells. It was possible that treatment with the AICART inhibitor resulted in depletion of ATP levels, and that the enhanced activation of AMPK was a consequence of both effects. However, ongoing experiments by Scott Rothbart in this laboratory, as well as a published study (45), indicate that ATP pools do not decrease after pemetrexed. We propose that the levels of ZMP detected in cells treated with the AICART inhibitor (2-4 mM) were sufficiently high to make any effect of changes in AMP:ATP levels on the activity of AMPK minimal.

These data suggest that ZMP pool expansion as a consequence of an AICART block is a more effective mechanism of mediating mTORC1 inhibition than ATP depletion by GART inhibitors.

Our analysis raises the question of whether GART or AICART would be the ideal target of a new generation of antifolates specifically selected as indirect inhibitors of mTORC1. While GART inhibitors showed substantial potential as anticancer drugs in pre-clinical studies, their use in the clinical setting was limited by the complication of severe thrombocytopenia in treated patients (191). This and the other side effects were most likely attributable to the multitude of consequences resulting from substantial depletion in purine nucleotide pools in normal highly proliferative cells, a common type of problem faced when treating patients with non-targeted agents. In contrast, while secondary effects of ZMP have been described on mitochondrial function (88), current thought is that the cellular effects of ZMP are mediated predominately through AMPK. As such, ZMP accumulation as a result of an AICART block may represent a method of targeting transformed cells that are hypersensitive to mTORC1 inhibition, e.g. cells deficient in PTEN, while preserving genetically normal cells. We propose that new drug development efforts towards the design of antifolates inhibitory of mTORC1 will have effects that are more selective and specific through a mechanism of action based on ZMP accumulation, rather than ATP depletion. Thus, the screening and selection of novel compounds in this class of antifolates should be based on their activity against the AICART enzyme.

Does AMPK activation impact the cytotoxicity of de novo purine synthesis inhibitors?

Previous studies have determined that antifolates targeting *de novo* purine synthesis have cytotoxic effects, but they appear to be less potent than antifolates targeting additional folate-dependent steps, such as thymidylate synthase (226). GART inhibition in human colorectal carcinoma cells was found to be cytotoxic, but the commitment to cell death was substantially lower than the cells treated with an inhibitor of thymidylate synthase (226). Interestingly, antifolate GART inhibitors only killed 1.5-2 logs of WiDR cells, whereas inhibitors of thymidylate synthase, i.e. raltitrexid, killed down to mutant frequencies. Likewise, these investigators noted that the morphology of the cells under the two conditions were remarkably dissimilar, suggesting that the mechanisms driving the cytotoxic effects were different (226). In those same studies, it was found that the cytotoxicity induced by methotrexate treatment in was substantially enhanced when purines were added to the media, eliminating any affects on purine synthesis (226). The polyglutamated derivatives of methotrexate have been shown to inhibit AICART in addition to DHFR, and it appears from these studies that the purine effect was protective to the cells grown in drug (4). Why would purine depletion protect cells from methotrexate? The activation of AMPK often results in the induction of autophagy or growth arrest to prevent cells from undergoing apoptosis (34, 115, 124). Authophagy is a process induced by nutrient and energy depletion and results in the lysosomal degradation of cellular material to promote cell survival in a time of stress (8). In a recent study, p27, a cyclin-dependent kinase inhibitor, was shown to promote autophagy in response to nutrient withdrawal. p27 is phosphorylated in an AMPK-dependent manner, which results in the

stabilization and accumulation of this protein within the cell (145). The investigators determined that under conditions of energy depletion knockdown of p27 drove cells towards an apoptotic state rather than the autophagic state observed in cells under the same conditions with wild-type levels of p27 (145). The induction of autophagy in response to anticancer agents has been described and how this affects the efficacy of a particular therapeutic strategy is currently being explored (8). It is clear, however that the cellular decision to induce growth arrest, autophagy, or apoptosis depends on the activity of several proteins within the mTOR pathway, and AMPK has been suggested to be a major player. With this in mind, it seems likely that activation of AMPK in response to *de novo* purine synthesis inhibition caused the induction of autophagy in the methotrexate and GART inhibited cells. From these studies, one may predict AICART inhibitors to have more of a cytostatic rather than cytotoxic effect on tumor cells.

Compound C effect on AMPK phosphorlyation at T172

In the experiments performed using compound C, an AMPK inhibitor (270), we made some surprising observations that should be further discussed. We used this inhibitor to test if the effects of pemetrexed on mTORC1 activity were a direct consequence of AMPK activity. The hypophosphorylation of S6K1 and 4EBP1 resulting from pemetrexed treatment were prevented when cells were exposed to 1 μM or 10 μM compound C. Likewise, the phosphorylation of ACC, a direct target of AMPK, was also prevented by the addition of compound C to the medium, supporting the idea that inhibition of AMPK activity by compound C blocked the effect of pemetrexed on the

mTOR pathway. Rather strikingly, in cells exposed to either concentration of compound C and pemetrexed a substantial increase in the phosphorylation of AMPK at T172 was observed. How the phosphorylation of AMPK was enhanced while allowing a block of AMPK enzyme activity remains unclear. Recent studies have determined that the binding of AMP leads to the accumulation phospho-AMPK within the cell by preventing the dephosphorylation of the α-subunit of AMPK by cellular phosphatases (204). One possible explanation for the increase in phospho-AMPK observed after treatment with pemetrexed and Compound C is that the AMPK-inhibitor further prevents dephosphorylation of the α-subunit by phosphatases through a conformational change or by direct inhibition of the phosphatases themselves. Interestingly, when cells were treated with AICAR and compound C the increase in phospho-AMPK was not observed (data not shown). Since ZMP was responsible for the activation of AMPK in both pemetrexed and AICAR treated cells, these observations suggest the affects observed were specific to pemetrexed treatment.

AMPK as a drug target

We have identified that the class of antifolates targeting *de novo* purine synthesis causes the activation of AMPK. The positioning of AMPK upstream of several key signaling pathways involved in cellular metabolism and proliferation has brought forth the idea that AMPK may be a potential drug target for the treatment of diseases such as type II diabetes, atherosclerosis, and most recently human cancers (94). Current data has uncovered that the activity of AMPK is central to p53-mediated checkpoints, as well as,

the induction of autophagy and apoptosis under conditions of cellular stress (34, 124, 145). These affects are both the consequences of the AMPK-mediated inhibition of mTORC1 and represent additional mechanisms of AMPK that are not necessarily tied to mTOR.

The cross-talk between AMPK, mTORC1, and p53 is evident by recent studies showing that p53 supports the activation of AMPK and the inhibition of mTORC1 by the transactivation of target genes whose products either lead to the enhancement of AMPK activaton, i.e. sestrins, or the up-regulation of proteins responsible for reducing mTORC1 activity, such as PTEN (33, 68, 69). Likewise, AMPK activity has been shown to induce the phosphorylation of p53 at the serine 15 position, a post-translational modification required for the induction of p53-mediated growth arrest (124). These data place the AMPK regulator at the center of p53 and mTORC1 mediated cellular events, two factors whose dysregulation account for most of the pathologies causing human cancers. Thus, exploring the efficacy of AMPK activation as an anticancer target is of substantial importance.

The vast number of tumor suppressors and oncogenic proteins (Figure 4-1 and table 4-1) linked to AMPK predicts that the cellular response and sensitivity to AICART inhibition will depend on the genetic background of the tumors exposed to drug. A direct correlation between PTEN deficiency, as well as, PI3K hyperactivity and the cytotoxic effects of rapamycin has been described in mouse tumors *in vivo* and in breast cancer cells lines *in vitro* (171, 173, 185, 221). As such, we predict that cells with mutations in the PTEN/PI3/Akt pathway will be particularly sensitive to treatment with *de novo* purine synthesis inhibitors. Proliferative disorders often result from loss of function of either

LKB1 or TSC1/TSC2 and the morbidities associated with these diseases have been reduced when treated with direct mTORC1 inhibitors (48, 160). It is of interest to test the efficacy of AICART inhibitors in the treatment of these pathologies. Likewise, the link between AMPK-mTOR activation and the induction of HIF-1α, in response to hypoxia (see introduction) also suggests that AICART inhibitors may have anti-angiogenesis properties.

The potential for AMPK activation in the treatment of human malignances has begun to be addressed both *in vitro* and *in vivo*. The treatment of breast cancer cells with metformin, an activator of AMPK by mechanisms that remain unknown, resulted in AMPK-dependent growth inhibition (265). *In vivo* studies demonstrated that tumor development in mice heterozygous for a PTEN deficiency, were affected by LKB1 levels and the onset of tumor formation in PTEN +/- mice was delayed when activation of AMPK was induced pharmacologically (111).

New targets for the multi-targeted antifolate, pemetrexed

Our studies have suggested that a drug currently in clinical use as a chemotherapeutic agent may be working through mechanisms linked to AMPK activation. Pemetrexed was shown to inhibit the second-folate dependent enzyme in *de novo* purine synthesis, AICART. The resulting metabolic block causes the accumulation of ZMP leading to the activation of AMPK and the inhibition of mTORC1. The clinical utility of pemetrexed has largely been attributed to its effect on thymidylate synthase, but *in vitro* experiments have highlighted the importance of the secondary targets of pemetrexed in

mediating its cytotoxic effects (45). We are very interested in determining the extent of how AMPK activation and mTORC1 inhibition has played a role in this compounds clinical utility. Interestingly, the combination of pemetrexed and cisplatin was more efficacious in the treatment of non-small cell lung cancer than gemcitabine and cisplatin except when patients had squamous cell carcinoma (61). This finding is remarkable and investigators have postulated that this observation was a consequence of the over-expression of thymidylate synthase in squamous cell tumors (61). The studies presented in this chapter offer an alternative explanation, whereby the distinction between squamous cell carcinoma and the other non-small cell cancers is the status of the mTORC1 pathway.

The assessment of the phosphorylation status of 4EBP1 and S6K1 has been used to assess mTORC1 inhibition in the clinical trials of direct mTOR inhibitors (160). Retrospective analysis of the status of these proteins in tumor samples from patients treated with pemetrexed will be useful in understanding if our observations are relevant to the clinical setting. Additionally, pre-clinical and clinical data obtained from studies using rapaloges may assist in identifying tumor types likely to be sensitive to treatment with pemetrexed. To this point, current rapalogs have also shown activity against human lung cancers in clinical trials (84, 181). Likewise, Rad001, a direct mTOR inhibitor, has recently been approved as first line therapy for the treatment of renal cell carcinoma (160). It may very well be the case that the effect of pemetrexed on the AMPK-mTOR pathway has been largely responsible for the selective use of this compound in the treatment of NSLC.

Genetic alterations that lead to proliferative disorders

Gene	Presentation	Syndrome	Aetiology
PTEN	Hamartomatous tumor syndromes	Cowden syndrome	PTEN loss, hyperactive PI3K
TSC1	Hamartomas	Tuberous sclerosis complex	TSC1 mutation, hyperactive mTORC1
TSC2	Abnormal growths in lungs (smooth muscle)	Lymphangio-leiomyomatosis	TSC2 mutation, hyperactive mTOR
AMPK	Cadiomyopathy	Familial hypertrophic cardiomyopathy	Loss of AMPK
LKB1	Gastrointestinal polyps	Peut-Jeghers syndrome	Loss of LKB1

Tumor suppressor/oncogenic proteins involved in cancer

Gene/protein	Genetic alteration	Cancer type
PTEN	Deletion, silencing, loss of function	Endometrial, glioblastoma, hepatocellular carcinoma
PI3K	Increased PI3K activity	Endometrial, glioblastoma, hepto-carcinoma, thyroid, prostate, lung
eIF4E	Overexpression	Breast, squamous cell, adenocarcinoma
LKB1	Mutation, silencing	Non-small cell lung cancer
Akt	Gene amplification, overexpression	Breast, ovarian, colon, ovarian, breast
S6K1	Gene amplification	Breast, ovarian

Table 4-2. Common mutations within the mTOR pathway that lead to proliferative disorders or cancer. Adapted from Faivre et. al, vol. 5, pages 671-688, August 2006, Nature Reviews Drug Discovery. Italicized text represent tumor suppressors and bold represent oncogenic proteins.

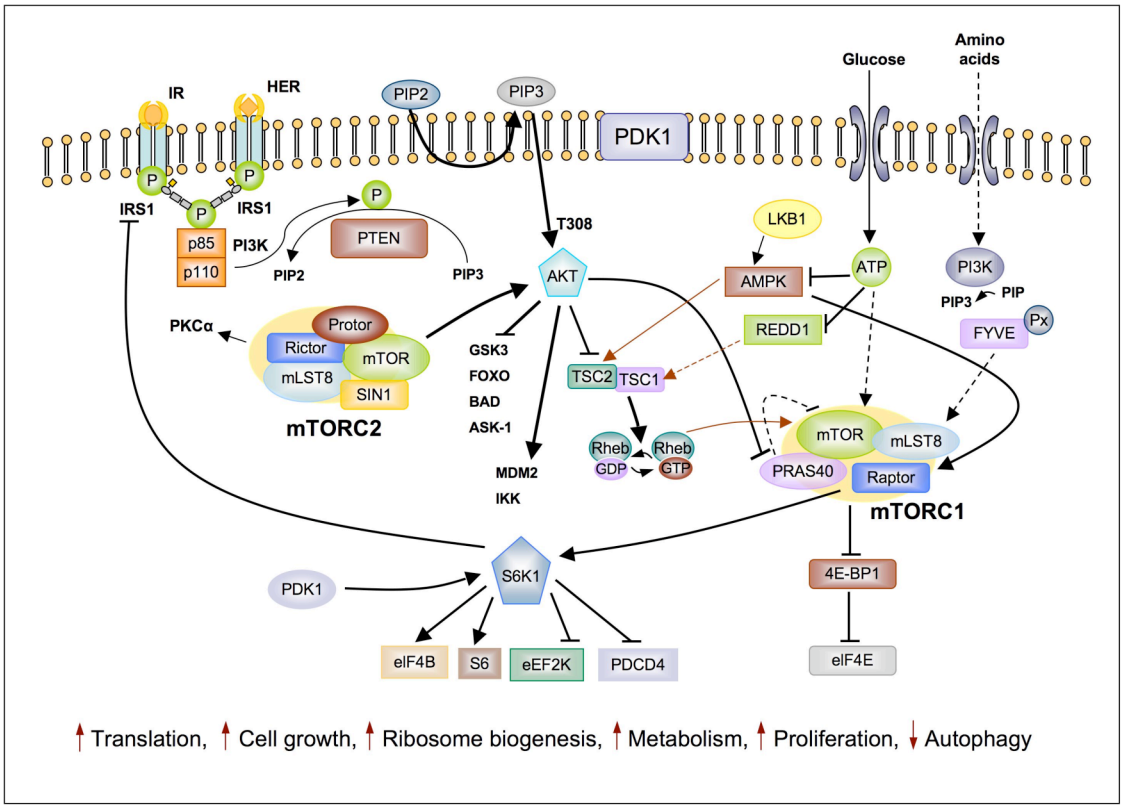


Figure 4-1. mTOR signaling network. Adapted from Merc-Bernstam et. al, Journal of Clinical Oncology, vol. 27, no. 13, May 2009

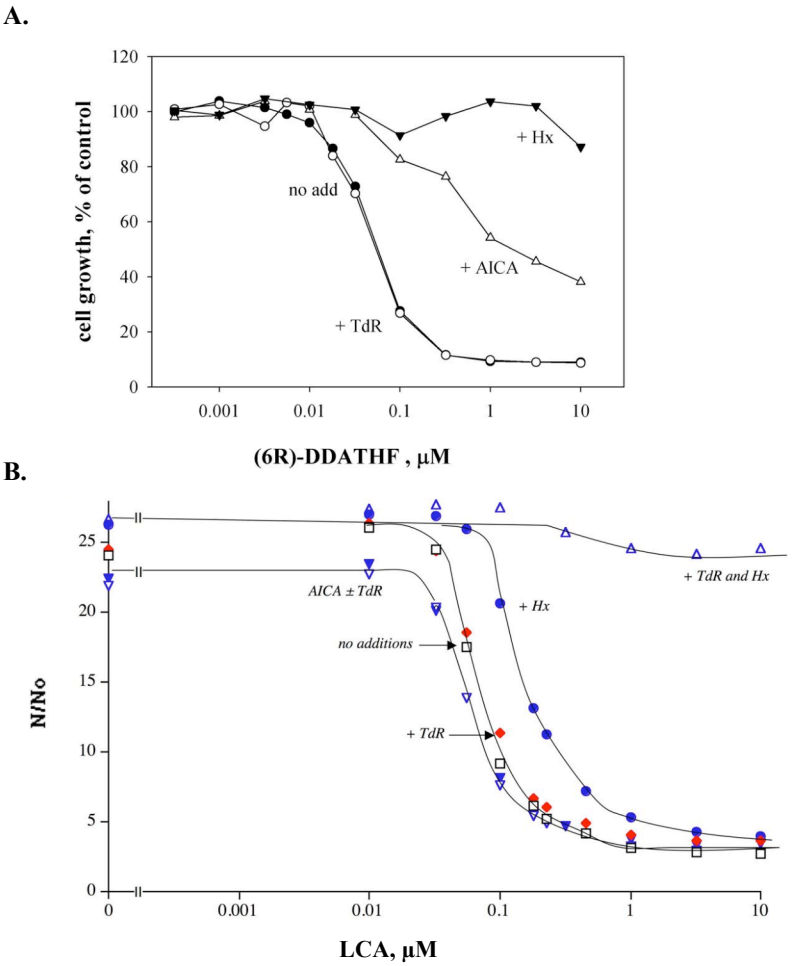


Figure 4-2. End product reversal studies in CEM cell growth treated with the *de novo* purine synthesis inhibitors 6R-DDATHF (A) and LCA (B). CEM cells were treated with the indicated concentrations of 6R-DDATHF (A) and LCA (B) alone (no addition) or in the presence of Thymidine (TdR, 5.6 µM), Hypoxanthine (Hx, 32 µM), or a combination of TdR with either Hx or AICA (320 µM). Drug and modifying agents were added simultaneously and cells were incubated in media for 96 hours. This work was performed by Dr. Richard Moran.

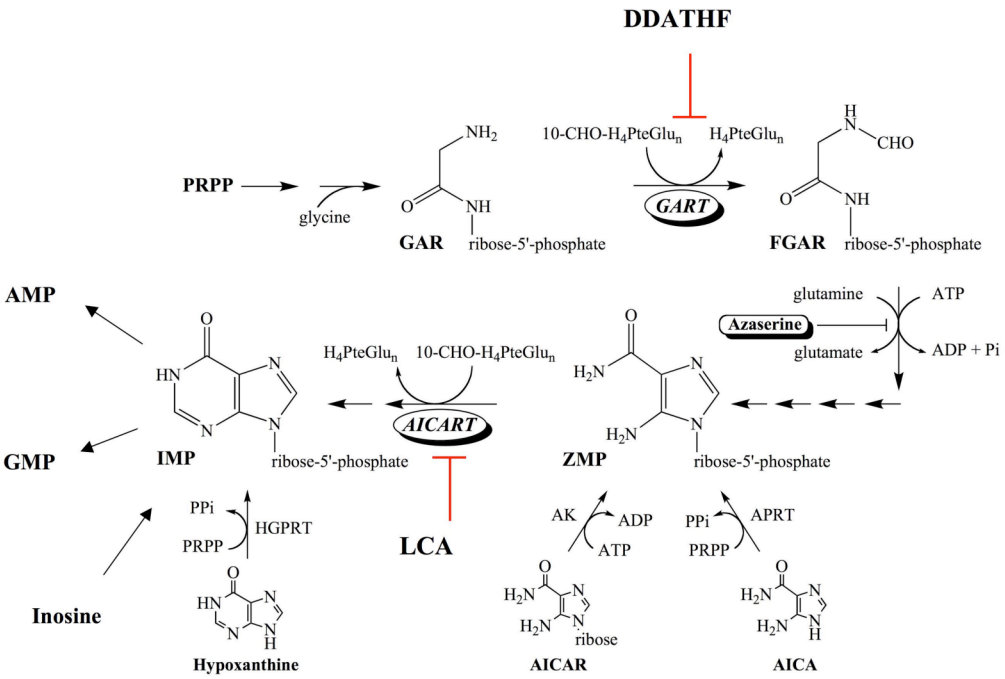


Figure 4-3. The folate dependent steps of *de novo* purine synthesis and the targets of DDATHF and LCA. *De novo* purine synthesis consists of ten sequential enzymatic reactions starting with 1-phosphoribosyl-5-pyrophosphate (PRPP) of which two, GART and AICART, are folate-dependent. The figure depicts the substrate of the GART reaction, glycinamide ribonucleotide (GAR), and the product of this reaction formylglycinamide ribonucleotide (FGAR), as well as the salvage of AICA, AICAR, and hypoxanthine catalyzed by adenine phosphoribosyltransferase (APRT), adenosine kinase (AK), and hypoxanthine-guanine phosphoribosyltransferase (HGPRT), respectively. Inosine mono-phosphate (IMP) is converted to AMP and GMP.

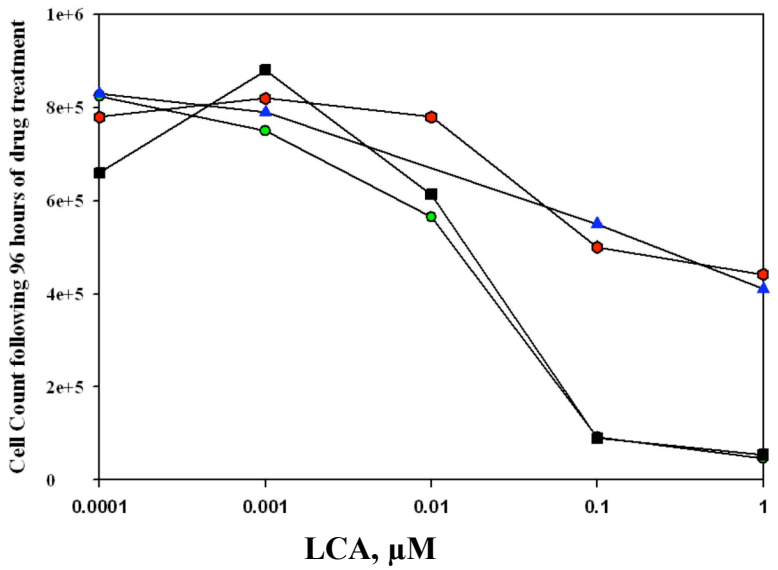


Figure 4-4. Reversal of HCT116 cell growth inhibition by addition of purines supported the observation that the target of LCA is *de novo* purine synthesis not thymidylate synthase. HCT116 cells were treated with the indicated concentrations of LCA alone (no addition, green small circles) or in the presence of Thymidine (TdR, 5.6 µM, black squares), Inosine (IR, 100 µM, blue triangles), or a combination of TdR and IR (red large circles). Drug and modifying agents were added simultaneously and cells were incubated in media for 96 hours. Cell growth was determined after 96 hours and cell number is expressed.

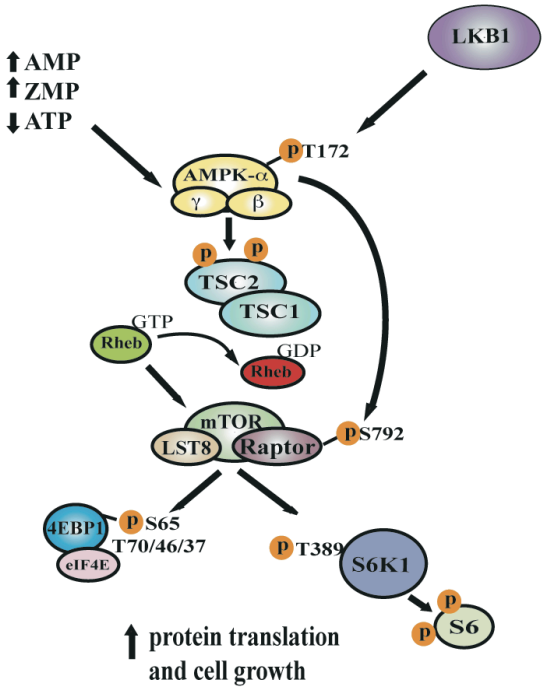


Figure 4-5. AMPK/TSC/mTORC1 axis studied.

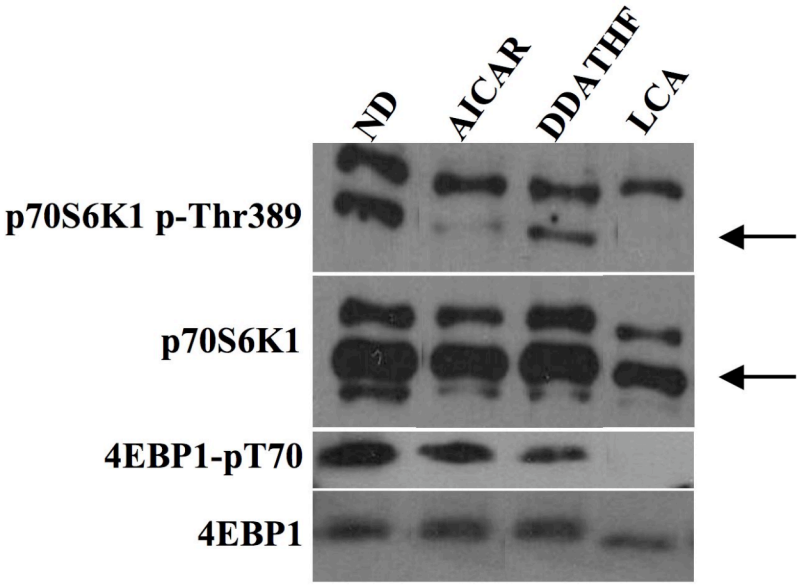


Figure 4-6. *De novo* purine synthesis inhibition causes a decrease in phosphorylation of 4EBP1 and S6K1, two major downstream targets of mTORC1. Western blots of total and phosphorylated S6K1 and 4EBP1 protein levels following treatment with 1 μM of DDATHF or LYLCA. The molecular masses of these bands were 62 kD (AMPK), S6K1 (70 kD), and 15-20 kD (4EBP1). Equal levels of total HCT116 protein (20 μg) was loaded in each lane; use of actin as a control demonstrated equal loading between lanes in this and the following figures (not shown). AICAR was used at 320 μM of 1 μM in the 4EBP1 or S6K1 blot, respectively.

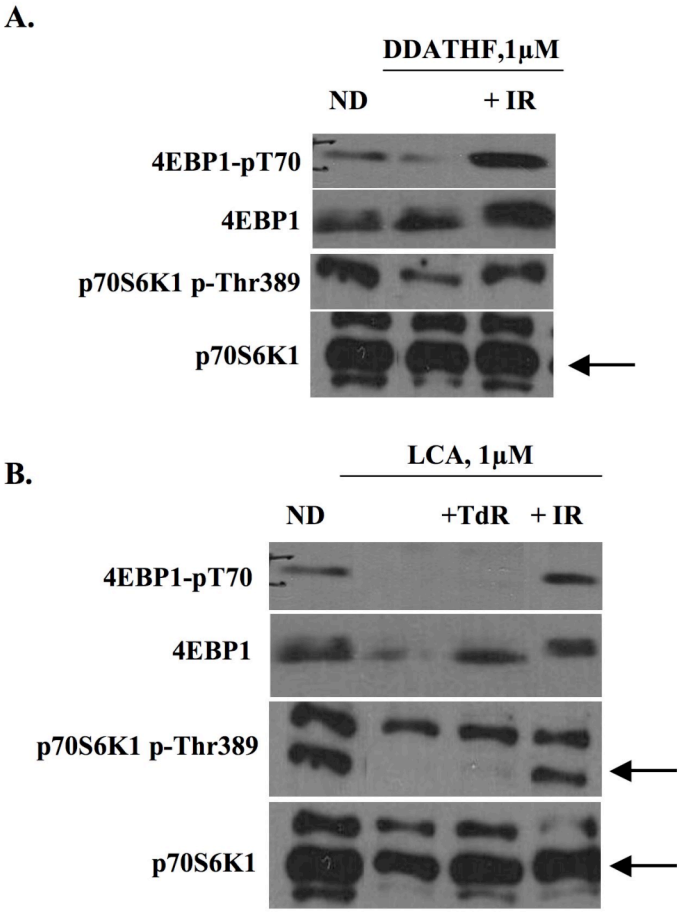


Figure 4-7. Reversal of mTOR inhibition induced in HCT116 cells by DDATHF and LCA by the addition of purines. Cellular activity of mTOR was assessed by the phosphorylation of S6K1 and 4EBP1 after treatment with DDATHF (A) and LYLCA (B). HCT116 cells were exposed to drug and modifying agents for 24 hours and IR was used at 100 μM and TdR was used at 5.6 μM.

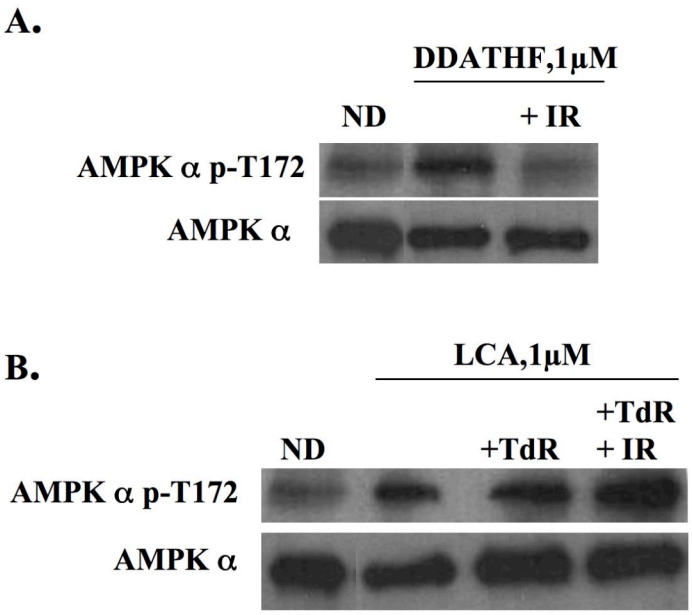


Figure 4-8. mTORC1 inhibition following treatment with *de novo* purine synthesis inhibitors correlates with activation of AMPK by phosphorylation at position T172. Phosphorylation of AMPK at the Thr172 position was assessed by western blot analysis following treatment of HCT116 cells with 1μM of DDATHF **(A)** and LCA **(B).** Cells treated with DDATHF were also exposed to 100 μM of IR as a rescue agent. Cells treated with LCA were exposed to 5.6 μM of TdR or 5.6μM of TdR and 100 μM of IR. The phosphorylation of AMPK appeared to be reversed using IR in cells treated with DDATHF, but not in those exposed to LCA. HCT116 cells were exposed to drug and modifying agents for 24 hours as in Fig. 4-6.

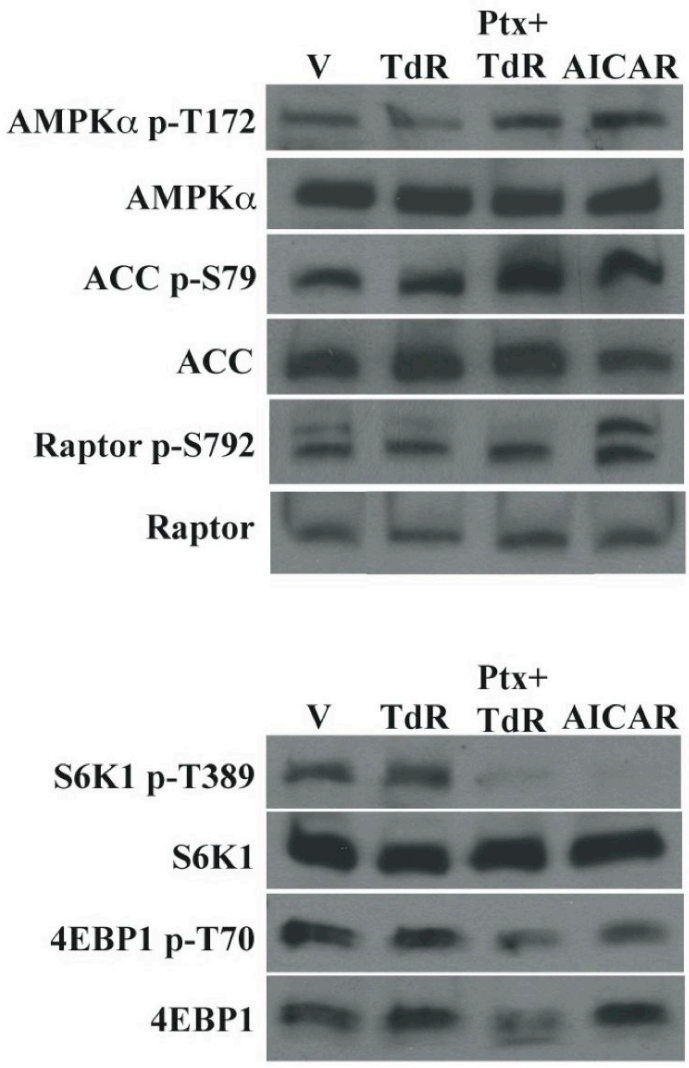


Figure 4-9. Effects of pemetrexed on activation of AMPK and inhibition of mTOR. (A and B) Western blot analysis of total and phosphorylated AMPK, ACC, Raptor, S6K1, and 4EBP1. The molecular masses of these bands were 62 kD (AMPK), 280 kD (ACC), 150 kDa (Raptor), S6K1 (70 kD), and 15-20 kD (4EBP1). Equal levels of total protein (40 μg) was loaded in each lane; use of actin as a control demonstrated equal loading between lanes in this and the following figures (not shown). Vehicle was PBS, pemetrexed was used at 1 μM, AICAR was at 250 μM, and TdR was 5.6μM; drug exposure was 48 hours.

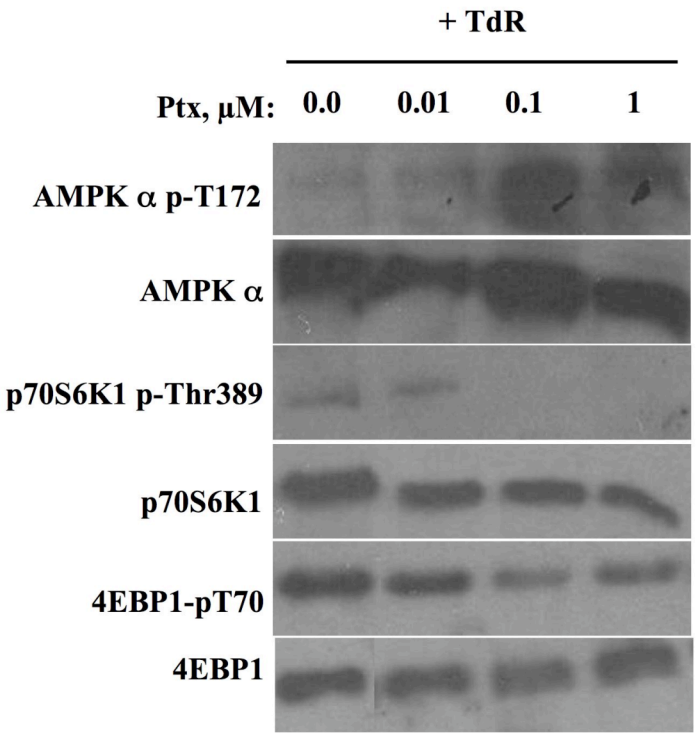
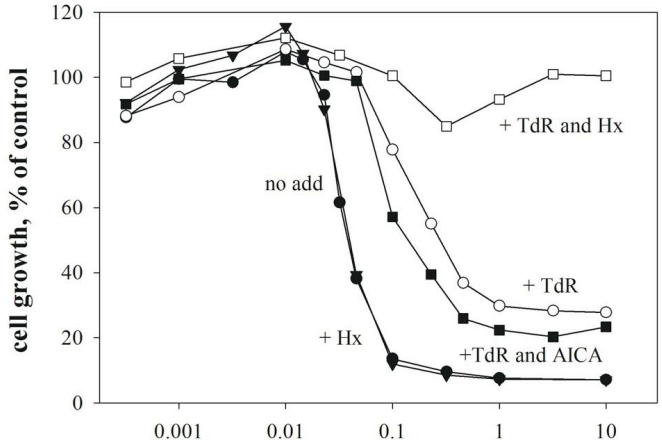


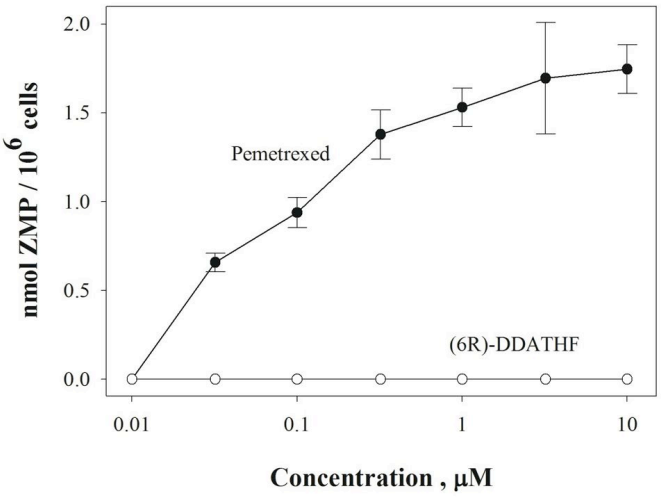
Figure 4-10. Dose-dependent activation of AMPK and inhibition of mTOR upon treatment of CEM cells with pemetrexed. AMPK activity was assessed by western blot analysis of phosphorylation of ACC and Raptor, and the activity of mTOR was indicated by the phosphorylation of S6K1 and 4EBP1. CEM cells were exposed to drug at .01 μM, .1 μM, and 1.0 μM with 5.6 μM TdR for 48 hours; conditions for western blots were as in Fig. 4-9.

Figure 4-11. Reversal of CEM cell growth inhibition by AICA and accumulation of cellular levels of ZMP indicated that the second target of pemetrexed is AICART, not GART. (A) CEM cells were treated with the indicated concentrations of pemetrexed alone (no add) or in the presence of TdR (5.6 μM), Hx (32 μM), AICA (320 μM), or a combination of TdR with either Hx or AICA. Drug and modifying agents were added simultaneously and drug-containing medium was changed at 48 hours. Cell growth was determined after 96 hours and cell number is expressed relative to controls without drug. (**B**) Increased cellular ZMP pools in pemetrexed-treated but not (6R)-DDATHF-treated CEM cells. Cells were exposed to drugs for 48 hours, nucleotides extracted and ZMP separated by hplc. **This work was performed by Dr. Moran (A) and Scott Rothbart (B).**

A.


cell growth, % of control
120
100
80
60
40
20
0
0.001
0.01
0.1
1
10
+ TdR and Hx
no add
+ TdR
+ Hx
+TdR and AICA

B.


nmol ZMP / 10^6 cells
2.0
1.5
1.0
0.5
0.0
0.01
0.1
1
10
Pemetrexed
(6R)-DDATHF
Concentration , μM

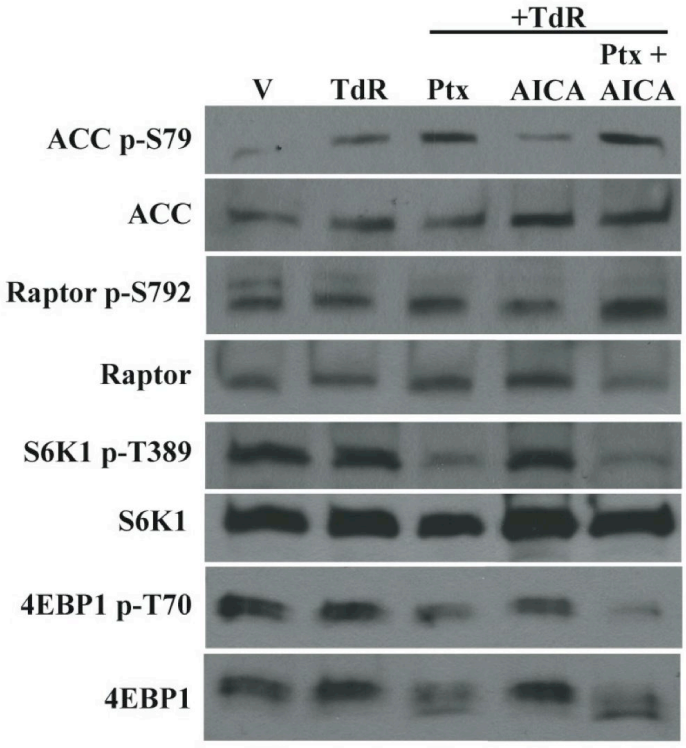


Figure 4-12. Expansion of the ZMP pool by AICA enhances the effect of pemetrexed on the AMPK-mTOR pathway in CEM cells. AMPK activity was assessed by western blot analysis of phosphorylation of ACC and Raptor, and the activity of mTOR was indicated by the phosphorylation of S6K1 and 4EBP1. CEM cells were exposed to drug and modifying agents for 48 hours as in Fig. 4-9 and AICA was used at 320 μM; conditions for western blots were as in Fig. 4-9.

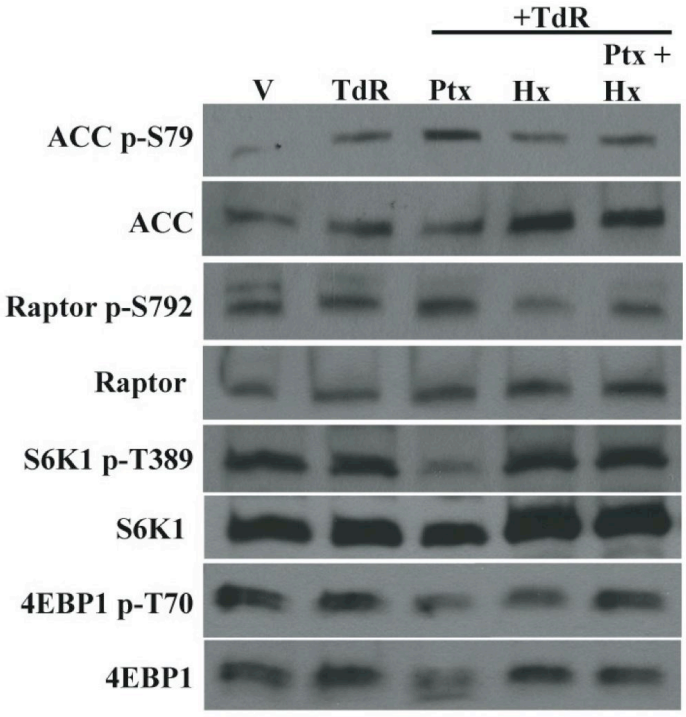


Figure 4-13. Reversal of secondary effects of pemetrexed on the AMPK-mTOR pathway by the addition of purines. Cellular activity of AMPK was assessed by western blot analysis of phosphorylation of ACC and Raptor, and the activity of mTOR was indicated by the phosphorylation of S6K1 and 4EBP1. CEM cells were exposed to drug and modifying agents for 48 hours as in Fig. 4-9 and Hx was used at 32 μM; conditions for western blots were as in Fig. 4-9.

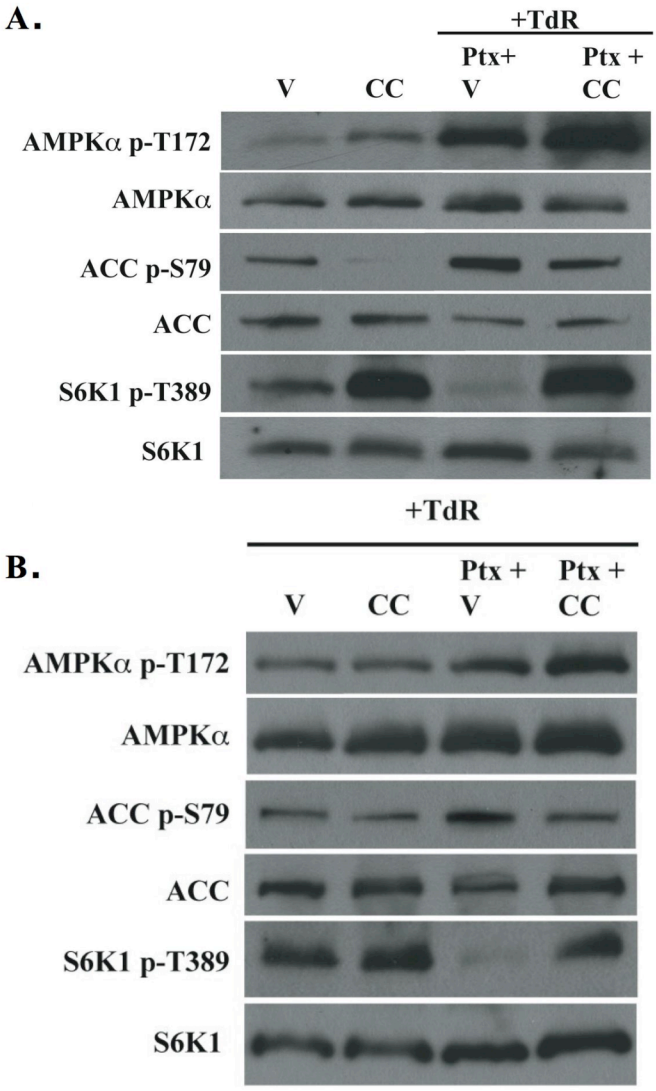


Figure 4-14. Prevention of the secondary effects of pemetrexed on the mTOR pathway by inhibition of AMPK. When CEM cells were exposed to 1 μM pemetrexed and 5.6 μM thymidine for 48 hours, 10 μM **(A)** or 1 μM **(B)** of Compound C (CC) in the medium prevented the cellular activity of AMPK by pemetrexed without interfering with the phosphorylation of AMPK. Vehicle (V) was DMSO.

Chapter 5: Perspectives

Throughout the course of this dissertation, I have addressed the epigenetic, transcriptional, and cell signaling mechanisms involved in the actions of folate anti-metabolites. In chapter 2, I presented evidence that the control of the multi-promoter mouse *fpgs* gene involves epigenetic, transcriptional interference, and promoter-proximal pausing mechanisms to coordinate the tissue-specific expression patterns detected in adult differentiated tissues and normal and neoplastic proliferative cells. In chapter 3, I perturbed those controls to study the response of the integrated system. In chapter 4, I presented evidence to suggest that antifolates inhibiting the second folate-dependent enzyme of *de novo* purine synthesis, AICART, may represent a class of folate anti-metabolites that have potential as molecularly targeted therapies in the treatment of human cancer. Our studies have lead to several conclusions and generated a number of new questions that remain to be answered, as discussed below.

The mouse fpgs gene: a model of tissue-specific transcriptional control mechanisms

The field of transcriptional control mechanisms is in a time of great expansion. The realization that a relatively small number of eukaryotic genes are controlled to generate the vast number of diverse proteins found in mammals has called into question the relevance

of basic transcriptional mechanisms once considered to be fundamental. Genome-wide technologies have identified the overwhelming prevalence of multi-promoter genes in mammals and greatly expanded the fields of microRNAs and epigenetics. During the course of this dissertation a major paradigm shift was caused by the discovery that a large proportion of mammalian genes are regulated not at the level of transcriptional initiation, but at a stage in transcriptional elongation (87). In this new framework, our studies at the mouse *fpgs* gene have highlighted the importance of several recently identified transcriptional processes in tissue-specific patterns of regulation, and the multitude of control mechanisms coordinating this gene continues to surprise us.

We consider the mouse *fpgs* gene an excellent model system to expand the genome-wide observations regarding tissue-specific and multi-promoter gene regulation to a mechanistic understanding of the interplay of these events at a single genetic locus. The work presented in this dissertation and generated from previous members of the laboratory have dissected several interesting aspects of this gene, and with our knowledge base we are in a unique position to address highly relevant and current topics using an *in vivo* model system. The areas of interest include tissue-specific patterns of DNA methylation and histone post-translational modifications, the mechanisms responsible for establishing poised elongation complexes, and the fundamental differences between CpG-sparse and CpG-island promoter regulation.

To further address these fields at the *fpgs* gene the next phase of *in vivo* studies will benefit from observations made that suggested that the *fpgs* gene undergoes promoter switching during the course of mouse developmental. The premise for this hypothesis

rests on data presented in Dr. Fiona Turner's dissertation where she analyzed the use of the two *fpgs* promoters in developing mouse tissues. Interestingly, in neonatal mouse liver *fpgs* mRNA was detected from both the P2 and P1 promoters (243). Likewise, following a partial hepatectomy, performed to induce hepatocyte cellular division, P2 was also found to be active. A separate study determined that detectable levels of FPGS protein were found in developing mouse brain, but not in differentiated tissue, which suggested that the P2 promoter is active in brain during development but silent in adult animals (12). Our studies have determined the P2 promoter in brain is silenced through a mechanism of promoter-proximal pausing. Tracking changes in promoter usage, DNA methylation status, and histone post-translational modifications throughout the course of development may provide insight into the sequence of events dictating the transcriptional, epigenetic, and poised states at the *fpgs* promoters in differentiated tissues. The rules defined will most likely be applicable to additional tissue-specific multi-promoter genes.

Of particular interest is the establishment of DNA methylation patterns at the P1 promoter in mouse liver. Is the P1 promoter silent and methylated in developing hepatocytes and if so, when is methylation lost? Tremendous insight has been gained towards understanding the meaning of certain histone modifications and DNA methylation through genome-wide studies using mammalian pluripotent stem cells (159, 161). However, patterns of DNA methylation across CpG-sparse promoters are difficult to obtain with the current strategies used to analyze genome-wide methylation. As such, monitoring methylation at *fpgs* P1 during development may be uniquely informative. Additionally, the field addressing processes of DNA demethylation is rapidly changing.

Current thought suggests that active de-methylation may occur through processes linked to nucleotide excision repair (213). The concept that DNA demethylation may define tissue-specific patterns of methylation has only been proposed and pursuing studies focused at the P1 promoter may shed light on this topic.

Literature precedent shows that primary events controlling genes are often defined from studies *in vitro*, but the promoter switching and poising occurring during development at the mouse *fpgs* may allow us to distinguish mechanisms *in vivo*.

Antifolates as molecularly targeted therapies

The use of antifolate drugs as chemotherapeutic agents was initiated in the early 1940's and remains first line therapy for several human pathologies including acute lymphocytic leukemia, rheumatoid arthritis, and sarcoidosis. These compounds target enzymes involved in nucleotide synthesis and amino acid metabolism and thus have proven to be very effective at eliminating highly proliferative tumor cells that depend on these reactions for cellular survival. Previously, a place for antifolates in targeted chemotherapy had not been realized, but our studies have defined a potential for folate antimetabolites in this arena. AICART inhibitors indirectly target mTOR, a protein central to several signaling pathways aberrantly regulated in human cancers. Thus, these agents are akin to direct mTOR inhibitors, which are currently used as immunosuppressant agents following kidney transplant and have recently been approved as first-line therapy in the treatment of renal cell carcinoma (160, 240). Several questions remain regarding the utility of AICART inhibitors as molecularly targeted therapies. What tumor types will be

sensitive to AICART inhibition? How will this class of compounds compare to direct mTORC1 inhibitors? Will these compounds be useful in combination with cytotoxic agents?

We propose that AICART inhibitors will have clinical success in the treatment of human cancers in combination therapy. Direct mTOR inhibitors have shown either additive or synergistic anti-tumor activity in combination with paclitaxel, cisplatin, doxorubicin, and carboplatin (160). Pre-clinical studies have suggested that mTOR inhibition may be beneficial in circumstances where an aberrant PI3/Akt pathway causes tumor resistance to first-line chemotherapeutic strategies. For example, the cytotoxic effects of trastuzumab, a monoclonal antibody blocking the Erb2 tyrosine kinase receptor, on breast cancer tumors are eliminated when cellular PTEN levels are lost, resulting in high levels of Akt activity (147). Co-administration of trastuxumab and Rad001, a mTORC1 inhibitor, has been shown to increase the sensitivity of resistant breast cancer cells to trastuxumab therapy by eliminating the downstream consequence of Akt stimulation, mTORC1 activation (147).

A clinically relevant consequence of mTORC1 inhibition is the activation of Akt, which occurs through the prevention of the feedback inhibition of Akt mediated by S6K1 activity on IRS-1 (Figure 4-1) (57). This has been shown to attenuate the anti-tumor effects of mTORC1 inhibition. Pre-clinical studies have shown that the combination of direct mTORC1 and IGF-R inhibitors alleviates this unwanted secondary consequence of mTORC1 inhibition and enhances antitumor activity (223). IGF-R and mTOR inhibitors are currently being tested in clinical trials (160). Akt activation will most likely also occur

in response to mTORC1 inhibition induced by an AICART block and the combination strategies successful with direct mTORC1 inhibitors should be considered when selecting agents to combine with this class of antifolates.

Beyond the use of AICART inhibitors as anticancer drugs these compounds should now be tested in the treatment of non-neoplastic pathologies sensitive to AMPK activation, such as type II diabetes, and proliferative disorders caused by loss of tumor suppressors, i.e. LKB1, TSC2, and PTEN, that are upstream regulators of mTORC1. Recent *in vivo* experiments have shown that pharmacological activation of AMPK improves blood glucose homeostasis, lipid profiles, and blood pressure in rodent models of type II diabetes, supporting a potentially serious role for AMPK activation in the treatment of this common metabolic disease (251). Likewise direct mTORC1 inhibitors in phase I and II clinical trials have shown promise in the reduction of angiomyolipomas, facial angiolypomas, and lymphangiomyomatosis, which are common consequences of proliferative disorders due to TSC1 and TSC2 dysregulation (149, 160). Clearly, in light of our recent data testing the use of antifolates targeting AICART as single agent therapy or in combination with direct mTORC1 inhibitors in the treatment of metabolic and proliferative disorders will be of great interest.

Literature Cited

Literature Cited

1. **Adhya, S., and M. Gottesman.** 1982. Promoter occlusion: transcription through a promoter may inhibit its activity. Cell **29:**939-44.
2. **Ahn, S. H., M. Kim, and S. Buratowski.** 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell **13:**67-76.
3. **Allegra, C. J., B. A. Chabner, J. C. Drake, R. Lutz, D. Rodbard, and J. Jolivet.** 1985. Enhanced inhibition of thymidylate synthase by methotrexate polyglutamates. J Biol Chem **260:**9720-6.
4. **Allegra, C. J., J. C. Drake, J. Jolivet, and B. A. Chabner.** 1985. Inhibition of phosphoribosylaminoimidazolecarboxamide transformylase by methotrexate and dihydrofolic acid polyglutamates. Proc Natl Acad Sci U S A **82:**4881-5.
5. **Allegra, C. J., K. Hoang, G. C. Yeh, J. C. Drake, and J. Baram.** 1987. Evidence for direct inhibition of de novo purine synthesis in human MCF-7 breast cells as a principal mode of metabolic inhibition by methotrexate. J Biol Chem **262:**13520-6.
6. **Allfrey, V. G., R. Faulkner, and A. E. Mirsky.** 1964. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A **51:**786-94.
7. **Andreassi, J. L., 2nd, and R. G. Moran.** 2002. Mouse folylpoly-gamma-glutamate synthetase isoforms respond differently to feedback inhibition by folylpolyglutamate cofactors. Biochemistry **41:**226-35.
8. **Apel, A., H. Zentgraf, M. W. Buchler, and I. Herr.** 2009. Autophagy-A double-edged sword in oncology. Int J Cancer.
9. **Appling, D. R.** 1991. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. FASEB J **5:**2645-51.
10. **Assaraf, Y. G.** 2007. Molecular basis of antifolate resistance. Cancer Metastasis Rev **26:**153-81.
11. **Bannister, A. J., R. Schneider, F. A. Myers, A. W. Thorne, C. Crane-Robinson, and T. Kouzarides.** 2005. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem **280:**17732-6.
12. **Barredo, J., and R. G. Moran.** 1992. Determinants of antifolate cytotoxicity: folylpolyglutamate synthetase activity during cellular proliferation and development. Mol Pharmacol **42:**687-94.
13. **Barrera, L. O., Z. Li, A. D. Smith, K. C. Arden, W. K. Cavenee, M. Q. Zhang, R. D. Green, and B. Ren.** 2008. Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. Genome Res **18:**46-59.
14. **Barreto, G., A. Schafer, J. Marhold, D. Stach, S. K. Swaminathan, V. Handa, G. Doderlein, N. Maltry, W. Wu, F. Lyko, and C. Niehrs.** 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature **445:**671-5.

15. **Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao.** 2007. High-resolution profiling of histone methylations in the human genome. Cell **129:**823-37.
16. **Baugh, C. M., and C. L. Krumdieck.** 1971. Naturally occurring folates. Ann N Y Acad Sci **186:**7-28.
17. **Baugh, C. M., C. L. Krumdieck, and M. G. Nair.** 1973. Polygammaglutamyl metabolites of methotrexate. Biochem Biophys Res Commun **52:**27-34.
18. **Beardsley, G. P., B. A. Moroson, E. C. Taylor, and R. G. Moran.** 1989. A new folate antimetabolite, 5,10-dideaza-5,6,7,8-tetrahydrofolate is a potent inhibitor of de novo purine synthesis. J Biol Chem **264:**328-33.
19. **Bentley, D. L., and M. Groudine.** 1986. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature **321:**702-6.
20. **Bernstein, B. E., M. Kamal, K. Lindblad-Toh, S. Bekiranov, D. K. Bailey, D. J. Huebert, S. McMahon, E. K. Karlsson, E. J. Kulbokas, 3rd, T. R. Gingeras, S. L. Schreiber, and E. S. Lander.** 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell **120:**169-81.
21. **Bestor, T. H., and V. M. Ingram.** 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc Natl Acad Sci U S A **80:**5559-63.
22. **Bird, A.** 2002. DNA methylation patterns and epigenetic memory. Genes Dev **16:**6-21.
23. **Bird, A.** 1992. The essentials of DNA methylation. Cell **70:**5-8.
24. **Bird, O. D., V. M. McGlohon, and J. W. Vaitkus.** 1965. Naturally occurring folates in the blood and liver of the rat. Anal Biochem **12:**18-35.
25. **Bjornsti, M. A., and P. J. Houghton.** 2004. The TOR pathway: a target for cancer therapy. Nat Rev Cancer **4:**335-48.
26. **Blackwood, E. M., and J. T. Kadonaga.** 1998. Going the distance: a current view of enhancer action. Science **281:**60-3.
27. **Bottomley, M. J.** 2004. Structures of protein domains that create or recognize histone modifications. EMBO Rep **5:**464-9.
28. **Boussadia, O., F. Amiot, S. Cases, G. Triqueneaux, H. Jacquemin-Sablon, and F. Dautry.** 1997. Transcription of unr (upstream of N-ras) down-modulates N-ras expression in vivo. FEBS Lett **420:**20-4.
29. **Boyer, L. A., K. Plath, J. Zeitlinger, T. Brambrink, L. A. Medeiros, T. I. Lee, S. S. Levine, M. Wernig, A. Tajonar, M. K. Ray, G. W. Bell, A. P. Otte, M. Vidal, D. K. Gifford, R. A. Young, and R. Jaenisch.** 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature **441:**349-53.
30. **Brigle, K. E., M. J. Spinella, E. H. Westin, and I. D. Goldman.** 1994. Increased expression and characterization of two distinct folate binding proteins in murine erythroleukemia cells. Biochem Pharmacol **47:**337-45.

31. **Bronder, J.** 2003. A defect in P53 response pathway induced by GART inhibition. Virginia Commonwealth University, Richmond.
32. **Bruniquel, D., and R. H. Schwartz.** 2003. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. Nat Immunol **4:**235-40.
33. **Budanov, A. V., and M. Karin.** 2008. p53 target genes sestrin1 and sestrin2 connect genotoxic stress and mTOR signaling. Cell **134:**451-60.
34. **Buzzai, M., R. G. Jones, R. K. Amaravadi, J. J. Lum, R. J. DeBerardinis, F. Zhao, B. Viollet, and C. B. Thompson.** 2007. Systemic treatment with the antidiabetic drug metformin selectively impairs p53-deficient tumor cell growth. Cancer Res **67:**6745-52.
35. **Calvert, A. H., D. L. Alison, S. J. Harland, B. A. Robinson, A. L. Jackman, T. R. Jones, D. R. Newell, Z. H. Siddik, E. Wiltshaw, T. J. McElwain, and et al.** 1986. A phase I evaluation of the quinazoline antifolate thymidylate synthase inhibitor, N10-propargyl-5,8-dideazafolic acid, CB3717. J Clin Oncol **4:**1245-52.
36. **Carling, D.** 2004. The AMP-activated protein kinase cascade--a unifying system for energy control. Trends Biochem Sci **29:**18-24.
37. **Carninci, P., A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. Semple, M. S. Taylor, P. G. Engstrom, M. C. Frith, A. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume, and Y. Hayashizaki.** 2006. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet **38:**626-35.
38. **Carrozza, M. J., B. Li, L. Florens, T. Suganuma, S. K. Swanson, K. K. Lee, W. J. Shia, S. Anderson, J. Yates, M. P. Washburn, and J. L. Workman.** 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell **123:**581-92.
39. **Caskey, C. T., D. M. Ashton, and J. B. Wyngaarden.** 1964. The Enzymology Of Feedback Inhibition Of Glutamine Phosphoribosylpyrophosphate Amidotransferase By Purine Ribonucleotides. J Biol Chem **239:**2570-9.
40. **Cedar, H.** 1988. DNA methylation and gene activity. Cell **53:**3-4.
41. **Chabner, B. A., C. J. Allegra, G. A. Curt, N. J. Clendeninn, J. Baram, S. Koizumi, J. C. Drake, and J. Jolivet.** 1985. Polyglutamation of methotrexate. Is methotrexate a prodrug? J Clin Invest **76:**907-12.
42. **Chattopadhyay, S., R. G. Moran, and I. D. Goldman.** 2007. Pemetrexed: biochemical and cellular pharmacology, mechanisms, and clinical applications. Mol Cancer Ther **6:**404-17.
43. **Chen, J., P. Hayes, K. Roy, and F. M. Sirotnak.** 2000. Two promoters regulate transcription of the mouse folylpolyglutamate synthetase gene three tightly

clustered Sp1 sites within the first intron markedly enhance activity of promoter B. Gene **242:**257-64.

44. **Chen, T., S. Hevi, F. Gay, N. Tsujimoto, T. He, B. Zhang, Y. Ueda, and E. Li.** 2007. Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet **39:**391-6.
45. **Chen, V. J., J. R. Bewley, S. L. Andis, R. M. Schultz, P. W. Iversen, C. Shih, L. G. Mendelsohn, D. E. Seitz, and J. L. Tonkinson.** 1998. Preclinical cellular pharmacology of LY231514 (MTA): a comparison with methotrexate, LY309887 and raltitrexed for their effects on intracellular folate and nucleoside triphosphate pools in CCRF-CEM cells. Br J Cancer **78 Suppl 3:**27-34.
46. **Cheng, B., and D. H. Price.** 2007. Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. J Biol Chem **282:**21901-12.
47. **Cook, R. J., and J. A. Blair.** 1979. The distribution and chemical nature of radioactive folates in rat liver cells and rat liver mitochondria. Biochem J **178:**651-9.
48. **Corradetti, M. N., and K. L. Guan.** 2006. Upstream of the mammalian target of rapamycin: do all roads pass through mTOR? Oncogene **25:**6347-60.
49. **Corton, J. M., J. G. Gillespie, S. A. Hawley, and D. G. Hardie.** 1995. 5-aminoimidazole-4-carboxamide ribonucleoside. A specific method for activating AMP-activated protein kinase in intact cells? Eur J Biochem **229:**558-65.
50. **Cunningham, D., J. Zalcberg, I. Smith, M. Gore, R. Pazdur, H. Burris, 3rd, N. J. Meropol, G. Kennealey, and L. Seymour.** 1996. 'Tomudex' (ZD1694): a novel thymidylate synthase inhibitor with clinical antitumour activity in a range of solid tumours. 'Tomudex' International Study Group. Ann Oncol **7:**179-82.
51. **de la Mata, M., C. R. Alonso, S. Kadener, J. P. Fededa, M. Blaustein, F. Pelisch, P. Cramer, D. Bentley, and A. R. Kornblihtt.** 2003. A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell **12:**525-32.
52. **Dennis, P. B., A. Jaeschke, M. Saitoh, B. Fowler, S. C. Kozma, and G. Thomas.** 2001. Mammalian TOR: a homeostatic ATP sensor. Science **294:**1102-5.
53. **Dong, K. B., I. A. Maksakova, F. Mohn, D. Leung, R. Appanah, S. Lee, H. W. Yang, L. L. Lam, D. L. Mager, D. Schubeler, M. Tachibana, Y. Shinkai, and M. C. Lorincz.** 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J **27:**2691-701.
54. **Dover, J., J. Schneider, M. A. Tawiah-Boateng, A. Wood, K. Dean, M. Johnston, and A. Shilatifard.** 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem **277:**28368-71.
55. **Druker, B. J., S. Tamura, E. Buchdunger, S. Ohno, G. M. Segal, S. Fanning, J. Zimmermann, and N. B. Lydon.** 1996. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. Nat Med **2:**561-6.
56. **Easton, J. B., and P. J. Houghton.** 2006. mTOR and cancer therapy. Oncogene **25:**6436-46.

57. **Easton, J. B., R. T. Kurmasheva, and P. J. Houghton.** 2006. IRS-1: auditing the effectiveness of mTOR inhibitors. Cancer Cell **9:**153-5.
58. **Eckhardt, F., J. Lewin, R. Cortese, V. K. Rakyan, J. Attwood, M. Burger, J. Burton, T. V. Cox, R. Davies, T. A. Down, C. Haefliger, R. Horton, K. Howe, D. K. Jackson, J. Kunde, C. Koenig, J. Liddle, D. Niblett, T. Otto, R. Pettett, S. Seemann, C. Thompson, T. West, J. Rogers, A. Olek, K. Berlin, and S. Beck.** 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet **38:**1378-85.
59. **Egan, M. G., S. Sirlin, B. G. Rumberger, T. A. Garrow, B. Shane, and F. M. Sirotnak.** 1995. Rapid decline in folylpolyglutamate synthetase activity and gene expression during maturation of HL-60 cells. Nature of the effect, impact on folate compound polyglutamate pools, and evidence for programmed down-regulation during maturation. J Biol Chem **270:**5462-8.
60. **Egger, G., S. Jeong, S. G. Escobar, C. C. Cortez, T. W. Li, Y. Saito, C. B. Yoo, P. A. Jones, and G. Liang.** 2006. Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc Natl Acad Sci U S A **103:**14080-5.
61. **Esteban, E., M. Casillas, and A. Cassinello.** 2009. Pemetrexed in first-line treatment of non-small cell lung cancer. Cancer Treat Rev **35:**364-73.
62. **Eszterhas, S. K., E. E. Bouhassira, D. I. Martin, and S. Fiering.** 2002. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. Mol Cell Biol **22:**469-79.
63. **Faivre, S., G. Kroemer, and E. Raymond.** 2006. Current development of mTOR inhibitors as anticancer agents. Nat Rev Drug Discov **5:**671-88.
64. **Farber, S.** 1949. Some observations on the effect of folic acid antagonists on acute leukemia and other forms of incurable cancer. Blood **4:**160-7.
65. **Farber, S., E. C. Cutler, J. W. Hawkins, J. H. Harrison, E. C. Peirce, 2nd, and G. G. Lenz.** 1947. The Action of Pteroylglutamic Conjugates on Man. Science **106:**619-621.
66. **Farber, S., and L. K. Diamond.** 1948. Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid. N Engl J Med **238:**787-93.
67. **Fatemi, M., A. Hermann, H. Gowher, and A. Jeltsch.** 2002. Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. Eur J Biochem **269:**4981-4.
68. **Feng, Z., W. Hu, E. de Stanchina, A. K. Teresky, S. Jin, S. Lowe, and A. J. Levine.** 2007. The regulation of AMPK beta1, TSC2, and PTEN expression by p53: stress, cell and tissue specificity, and the role of these gene products in modulating the IGF-1-AKT-mTOR pathways. Cancer Res **67:**3043-53.
69. **Feng, Z., H. Zhang, A. J. Levine, and S. Jin.** 2005. The coordinate regulation of the p53 and mTOR pathways in cells. Proc Natl Acad Sci U S A **102:**8204-9.

70. **Foley, K. P., and J. D. Engel.** 1992. Individual stage selector element mutations lead to reciprocal changes in beta- vs. epsilon-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. Genes Dev **6:**730-44.
71. **Freemantle, S. J., and R. G. Moran.** 1997. Transcription of the human folylpoly-gamma-glutamate synthetase gene. J Biol Chem **272:**25373-9.
72. **Freemantle, S. J., S. M. Taylor, G. Krystal, and R. G. Moran.** 1995. Upstream organization of and multiple transcripts from the human folylpoly-gamma-glutamate synthetase gene. J Biol Chem **270:**9579-84.
73. **Fry, D. W., J. C. Yalowich, and I. D. Goldman.** 1982. Rapid formation of poly-gamma-glutamyl derivatives of methotrexate and their association with dihydrofolate reductase as assessed by high pressure liquid chromatography in the Ehrlich ascites tumor cell in vitro. J Biol Chem **257:**1890-6.
74. **Fuks, F.** 2005. DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev **15:**490-5.
75. **Futscher, B. W., M. M. Oshiro, R. J. Wozniak, N. Holtan, C. L. Hanigan, H. Duan, and F. E. Domann.** 2002. Role for DNA methylation in the control of cell type specific maspin expression. Nat Genet **31:**175-9.
76. **Gera, J. F., I. K. Mellinghoff, Y. Shi, M. B. Rettig, C. Tran, J. H. Hsu, C. L. Sawyers, and A. K. Lichtenstein.** 2004. AKT activity determines sensitivity to mammalian target of rapamycin (mTOR) inhibitors by regulating cyclin D1 and c-myc expression. J Biol Chem **279:**2737-46.
77. **Gilchrist, D. A., S. Nechaev, C. Lee, S. K. Ghosh, J. B. Collins, L. Li, D. S. Gilmour, and K. Adelman.** 2008. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev **22:**1921-33.
78. **Gingras, A. C., B. Raught, and N. Sonenberg.** 2001. Regulation of translation initiation by FRAP/mTOR. Genes Dev **15:**807-26.
79. **Goldman, I. D., N. S. Lichtenstein, and V. T. Oliverio.** 1968. Carrier-mediated transport of the folic acid analogue, methotrexate, in the L1210 leukemia cell. J Biol Chem **243:**5007-17.
80. **Goll, M. G., F. Kirpekar, K. A. Maggert, J. A. Yoder, C. L. Hsieh, X. Zhang, K. G. Golic, S. E. Jacobsen, and T. H. Bestor.** 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science **311:**395-8.
81. **Gomes, N. P., G. Bjerke, B. Llorente, S. A. Szostek, B. M. Emerson, and J. M. Espinosa.** 2006. Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. Genes Dev **20:**601-12.
82. **Gowher, H., K. Liebert, A. Hermann, G. Xu, and A. Jeltsch.** 2005. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. J Biol Chem **280:**13341-8.

83. **Greger, I. H., F. Demarchi, M. Giacca, and N. J. Proudfoot.** 1998. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res **26:**1294-301.
84. **Gridelli, C., A. Rossi, F. Morgillo, M. A. Bareschino, P. Maione, M. Di Maio, and F. Ciardiello.** 2007. A randomized phase II study of pemetrexed or RAD001 as second-line treatment of advanced non-small-cell lung cancer in elderly patients: treatment rationale and protocol dynamics. Clin Lung Cancer **8:**568-71.
85. **Guba, M., P. von Breitenbuch, M. Steinbauer, G. Koehl, S. Flegel, M. Hornung, C. J. Bruns, C. Zuelke, S. Farkas, M. Anthuber, K. W. Jauch, and E. K. Geissler.** 2002. Rapamycin inhibits primary and metastatic tumor growth by antiangiogenesis: involvement of vascular endothelial growth factor. Nat Med **8:**128-35.
86. **Guccione, E., F. Martinato, G. Finocchiaro, L. Luzi, L. Tizzoni, V. Dall' Olio, G. Zardo, C. Nervi, L. Bernard, and B. Amati.** 2006. Myc-binding-site recognition in the human genome is determined by chromatin context. Nat Cell Biol **8:**764-70.
87. **Guenther, M. G., S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young.** 2007. A chromatin landmark and transcription initiation at most promoters in human cells. Cell **130:**77-88.
88. **Guigas, B., N. Taleux, M. Foretz, D. Detaille, F. Andreelli, B. Viollet, and L. Hue.** 2007. AMP-activated protein kinase-independent inhibition of hepatic mitochondrial oxidative phosphorylation by AICA riboside. Biochem J **404:**499-507.
89. **Gwinn, D. M., D. B. Shackelford, D. F. Egan, M. M. Mihaylova, A. Mery, D. S. Vasquez, B. E. Turk, and R. J. Shaw.** 2008. AMPK phosphorylation of raptor mediates a metabolic checkpoint. Mol Cell **30:**214-26.
90. **Hahn, S.** 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol **11:**394-403.
91. **Hamada, K., T. Sasaki, P. A. Koni, M. Natsui, H. Kishimoto, J. Sasaki, N. Yajima, Y. Horie, G. Hasegawa, M. Naito, J. Miyazaki, T. Suda, H. Itoh, K. Nakao, T. W. Mak, T. Nakano, and A. Suzuki.** 2005. The PTEN/PI3K pathway governs normal vascular development and tumor angiogenesis. Genes Dev **19:**2054-65.
92. **Hampsey, M., and D. Reinberg.** 2003. Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell **113:**429-32.
93. **Harata, K., N. Ishii, N. Rashid, M. Morikawa, and T. Imanaka.** 2000. Crystallization and preliminary X-ray study of Pk--REC from a hyperthermophilic archaeon, Pyrococcus kodakaraensis KOD1. Acta Crystallogr D Biol Crystallogr **56:**648-9.
94. **Hardie, D. G.** 2007. AMP-activated protein kinase as a drug target. Annu Rev Pharmacol Toxicol **47:**185-210.
95. **Hardie, D. G.** 2007. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. Nat Rev Mol Cell Biol **8:**774-85.

96. **Hardie, D. G., and S. A. Hawley.** 2001. AMP-activated protein kinase: the energy charge hypothesis revisited. Bioessays **23:**1112-9.
97. **Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse, and S. M. Tilghman.** 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature **405:**486-9.
98. **Harrington, L. S., G. M. Findlay, A. Gray, T. Tolkacheva, S. Wigfield, H. Rebholz, J. Barnett, N. R. Leslie, S. Cheng, P. R. Shepherd, I. Gout, C. P. Downes, and R. F. Lamb.** 2004. The TSC1-2 tumor suppressor controls insulin-PI3K signaling via regulation of IRS proteins. J Cell Biol **166:**213-23.
99. **Hashemolhosseini, S., Y. Nagamine, S. J. Morley, S. Desrivieres, L. Mercep, and S. Ferrari.** 1998. Rapamycin inhibition of the G1 to S transition is mediated by effects on cyclin D1 mRNA and protein stability. J Biol Chem **273:**14424-9.
100. **Hawley, S. A., J. Boudeau, J. L. Reid, K. J. Mustard, L. Udd, T. P. Makela, D. R. Alessi, and D. G. Hardie.** 2003. Complexes between the LKB1 tumor suppressor, STRAD alpha/beta and MO25 alpha/beta are upstream kinases in the AMP-activated protein kinase cascade. J Biol **2:**28.
101. **Hawley, S. A., M. Davison, A. Woods, S. P. Davies, R. K. Beri, D. Carling, and D. G. Hardie.** 1996. Characterization of the AMP-activated protein kinase kinase from rat liver and identification of threonine 172 as the major site at which it phosphorylates AMP-activated protein kinase. J Biol Chem **271:**27879-87.
102. **Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz-Bourget, H. Liu, X. Zhang, R. D. Green, V. V. Lobanenkov, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis, and B. Ren.** 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature **459:**108-12.
103. **Hermann, A., H. Gowher, and A. Jeltsch.** 2004. Biochemistry and biology of mammalian DNA methyltransferases. Cell Mol Life Sci **61:**2571-87.
104. **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:**405-11.
105. **Holz, M. K., B. A. Ballif, S. P. Gygi, and J. Blenis.** 2005. mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. Cell **123:**569-80.
106. **Hsu, M., R. Mabaera, C. H. Lowrey, D. I. Martin, and S. Fiering.** 2007. CpG hypomethylation in a large domain encompassing the embryonic beta-like globin genes in primitive erythrocytes. Mol Cell Biol **27:**5047-54.
107. **Hu, X., S. Eszterhas, N. Pallazzi, E. E. Bouhassira, J. Fields, O. Tanabe, S. A. Gerber, M. Bulger, J. D. Engel, M. Groudine, and S. Fiering.** 2007. Transcriptional interference among the murine beta-like globin genes. Blood **109:**2210-6.
108. **Hu, Y. G., R. Hirasawa, J. L. Hu, K. Hata, C. L. Li, Y. Jin, T. Chen, E. Li, M. Rigolet, E. Viegas-Pequignot, H. Sasaki, and G. L. Xu.** 2008. Regulation of

DNA methylation activity through Dnmt3L promoter methylation by Dnmt3 enzymes in embryonic development. Hum Mol Genet **17:**2654-64.

109. **Huang, S., L. N. Liu, H. Hosoi, M. B. Dilling, T. Shikata, and P. J. Houghton.** 2001. p53/p21(CIP1) cooperate in enforcing rapamycin-induced G(1) arrest and determine the cellular response to rapamycin. Cancer Res **61:**3373-81.
110. **Huang, S., L. Shu, M. B. Dilling, J. Easton, F. C. Harwood, H. Ichijo, and P. J. Houghton.** 2003. Sustained activation of the JNK cascade and rapamycin-induced apoptosis are suppressed by p53/p21(Cip1). Mol Cell **11:**1491-501.
111. **Huang, X., S. Wullschleger, N. Shpiro, V. A. McGuire, K. Sakamoto, Y. L. Woods, W. McBurnie, S. Fleming, and D. R. Alessi.** 2008. Important role of the LKB1-AMPK pathway in suppressing tumorigenesis in PTEN-deficient mice. Biochem J **412:**211-21.
112. **Hutchins, A. S., A. C. Mullen, H. W. Lee, K. J. Sykes, F. A. High, B. D. Hendrich, A. P. Bird, and S. L. Reiner.** 2002. Gene silencing quantitatively controls the function of a developmental trans-activator. Mol Cell **10:**81-91.
113. **Inoki, K., Y. Li, T. Xu, and K. L. Guan.** 2003. Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. Genes Dev **17:**1829-34.
114. **Inoki, K., Y. Li, T. Zhu, J. Wu, and K. L. Guan.** 2002. TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. Nat Cell Biol **4:**648-57.
115. **Inoki, K., T. Zhu, and K. L. Guan.** 2003. TSC2 mediates cellular energy response to control cell growth and survival. Cell **115:**577-90.
116. **Jacinto, E., R. Loewith, A. Schmidt, S. Lin, M. A. Ruegg, A. Hall, and M. N. Hall.** 2004. Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. Nat Cell Biol **6:**1122-8.
117. **Jackman, A. L., P. R. Marsham, R. G. Moran, R. Kimbell, B. M. O'Connor, L. R. Hughes, and A. H. Calvert.** 1991. Thymidylate synthase inhibitors: the in vitro activity of a series of heterocyclic benzoyl ring modified 2-desamino-2-methyl-N10-substituted-5,8-dideazafolates. Adv Enzyme Regul **31:**13-27.
118. **Jackman, A. L., G. A. Taylor, W. Gibson, R. Kimbell, M. Brown, A. H. Calvert, I. R. Judson, and L. R. Hughes.** 1991. ICI D1694, a quinazoline antifolate thymidylate synthase inhibitor that is a potent inhibitor of L1210 tumor cell growth in vitro and in vivo: a new agent for clinical study. Cancer Res **51:**5579-86.
119. **Jenuwein, T., and C. D. Allis.** 2001. Translating the histone code. Science **293:**1074-80.
120. **Jones, P. A., and S. B. Baylin.** 2002. The fundamental role of epigenetic events in cancer. Nat Rev Genet **3:**415-28.
121. **Jones, P. A., and D. Takai.** 2001. The role of DNA methylation in mammalian epigenetics. Science **293:**1068-70.
122. **Jones, P. A., and S. M. Taylor.** 1980. Cellular differentiation, cytidine analogs and DNA methylation. Cell **20:**85-93.

123. **Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, and A. P. Wolffe.** 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet **19:**187-91.
124. **Jones, R. G., D. R. Plas, S. Kubek, M. Buzzai, J. Mu, Y. Xu, M. J. Birnbaum, and C. B. Thompson.** 2005. AMP-activated protein kinase induces a p53-dependent metabolic checkpoint. Mol Cell **18:**283-93.
125. **Jones, T. R., A. H. Calvert, A. L. Jackman, S. J. Brown, M. Jones, and K. R. Harrap.** 1981. A potent antitumour quinazoline inhibitor of thymidylate synthetase: synthesis, biological properties and therapeutic results in mice. Eur J Cancer **17:**11-9.
126. **Kangaspeska, S., B. Stride, R. Metivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon, and G. Reid.** 2008. Transient cyclical methylation of promoter DNA. Nature **452:**112-5.
127. **Kawamata, S., H. Sakaida, T. Hori, M. Maeda, and T. Uchiyama.** 1998. The upregulation of p27Kip1 by rapamycin results in G1 arrest in exponentially growing T-cell lines. Blood **91:**561-9.
128. **Keogh, M. C., S. K. Kurdistani, S. A. Morris, S. H. Ahn, V. Podolny, S. R. Collins, M. Schuldiner, K. Chin, T. Punna, N. J. Thompson, C. Boone, A. Emili, J. S. Weissman, T. R. Hughes, B. D. Strahl, M. Grunstein, J. F. Greenblatt, S. Buratowski, and N. J. Krogan.** 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell **123:**593-605.
129. **Kim, D. H., D. D. Sarbassov, S. M. Ali, J. E. King, R. R. Latek, H. Erdjument-Bromage, P. Tempst, and D. M. Sabatini.** 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. Cell **110:**163-75.
130. **Kizer, K. O., H. P. Phatnani, Y. Shibata, H. Hall, A. L. Greenleaf, and B. D. Strahl.** 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol **25:**3305-16.
131. **Klose, R. J., and A. P. Bird.** 2006. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci **31:**89-97.
132. **Koch, F., F. Jourquin, P. Ferrier, and J. C. Andrau.** 2008. Genome-wide RNA polymerase II: not genes only! Trends Biochem Sci **33:**265-73.
133. **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:**2452-60.
134. **Kornberg, R. D.** 1974. Chromatin structure: a repeating unit of histones and DNA. Science **184:**868-71.
135. **Kouzarides, T.** 2007. Chromatin modifications and their function. Cell **128:**693-705.
136. **Krogan, N. J., J. Dover, A. Wood, J. Schneider, J. Heidt, M. A. Boateng, K. Dean, O. W. Ryan, A. Golshani, M. Johnston, J. F. Greenblatt, and A.**

Shilatifard. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell **11:**721-9.

137. **Lande-Diner, L., J. Zhang, I. Ben-Porath, N. Amariglio, I. Keshet, M. Hecht, V. Azuara, A. G. Fisher, G. Rechavi, and H. Cedar.** 2007. Role of DNA methylation in stable gene repression. J Biol Chem **282:**12194-200.
138. **Lee, C., X. Li, A. Hechmer, M. Eisen, M. D. Biggin, B. J. Venters, C. Jiang, J. Li, B. F. Pugh, and D. S. Gilmour.** 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila. Mol Cell Biol **28:**3290-300.
139. **Lee, T. I., R. G. Jenner, L. A. Boyer, M. G. Guenther, S. S. Levine, R. M. Kumar, B. Chevalier, S. E. Johnstone, M. F. Cole, K. Isono, H. Koseki, T. Fuchikami, K. Abe, H. L. Murray, J. P. Zucker, B. Yuan, G. W. Bell, E. Herbolsheimer, N. M. Hannett, K. Sun, D. T. Odom, A. P. Otte, T. L. Volkert, D. P. Bartel, D. A. Melton, D. K. Gifford, R. Jaenisch, and R. A. Young.** 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell **125:**301-13.
140. **Lehnertz, B., Y. Ueda, A. A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenuwein, and A. H. Peters.** 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol **13:**1192-200.
141. **Li, B., M. Carey, and J. L. Workman.** 2007. The role of chromatin during transcription. Cell **128:**707-19.
142. **Li, E., T. H. Bestor, and R. Jaenisch.** 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69:**915-26.
143. **Li, H., S. Ilin, W. Wang, E. M. Duncan, J. Wysocka, C. D. Allis, and D. J. Patel.** 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature **442:**91-5.
144. **Liang, G., J. C. Lin, V. Wei, C. Yoo, J. C. Cheng, C. T. Nguyen, D. J. Weisenberger, G. Egger, D. Takai, F. A. Gonzales, and P. A. Jones.** 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proc Natl Acad Sci U S A **101:**7357-62.
145. **Liang, J., S. H. Shao, Z. X. Xu, B. Hennessy, Z. Ding, M. Larrea, S. Kondo, D. J. Dumont, J. U. Gutterman, C. L. Walker, J. M. Slingerland, and G. B. Mills.** 2007. The energy sensing LKB1-AMPK pathway regulates p27(kip1) phosphorylation mediating the decision to enter autophagy or apoptosis. Nat Cell Biol **9:**218-24.
146. **Lorincz, M. C., D. R. Dickerson, M. Schmitt, and M. Groudine.** 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat Struct Mol Biol **11:**1068-75.
147. **Lu, C. H., S. L. Wyszomierski, L. M. Tseng, M. H. Sun, K. H. Lan, C. L. Neal, G. B. Mills, G. N. Hortobagyi, F. J. Esteva, and D. Yu.** 2007. Preclinical testing

of clinically applicable strategies for overcoming trastuzumab resistance caused by PTEN deficiency. Clin Cancer Res **13:**5883-8.

148. **Ma, C. X., S. Nair, S. Thomas, S. J. Mandrekar, D. A. Nikcevich, K. M. Rowland, T. R. Fitch, H. E. Windschitl, S. L. Hillman, S. E. Schild, J. R. Jett, C. Obasaju, and A. A. Adjei.** 2005. Randomized phase II trial of three schedules of pemetrexed and gemcitabine as front-line therapy for advanced non-small-cell lung cancer. J Clin Oncol **23:**5929-37.
149. **Ma, X. M., and J. Blenis.** 2009. Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol **10:**307-18.
150. **Mabaera, R., C. A. Richardson, K. Johnson, M. Hsu, S. Fiering, and C. H. Lowrey.** 2007. Developmental- and differentiation-specific patterns of human gamma- and beta-globin promoter DNA methylation. Blood **110:**1343-52.
151. **Mamane, Y., E. Petroulakis, L. Rong, K. Yoshida, L. W. Ler, and N. Sonenberg.** 2004. eIF4E--from translation to transformation. Oncogene **23:**3172-9.
152. **Manning, B. D., and L. C. Cantley.** 2003. Rheb fills a GAP between TSC and TOR. Trends Biochem Sci **28:**573-6.
153. **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:**27176-83.
154. **Marshall, N. F., and D. H. Price.** 1992. Control of formation of two distinct classes of RNA polymerase II elongation complexes. Mol Cell Biol **12:**2078-90.
155. **Martens, J. A., L. Laprade, and F. Winston.** 2004. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature **429:**571-4.
156. **Martens, J. A., P. Y. Wu, and F. Winston.** 2005. Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes Dev **19:**2695-704.
157. **McBurney, M. W., and G. F. Whitmore.** 1974. Characterization of a Chinese hamster cell with a temperature-sensitive mutation in folate metabolism. Cell **2:**183-8.
158. **McCarthy, E. A., S. A. Titus, S. M. Taylor, C. Jackson-Cook, and R. G. Moran.** 2004. A mutation inactivating the mitochondrial inner membrane folate transporter creates a glycine requirement for survival of chinese hamster cells. J Biol Chem **279:**33829-36.
159. **Meissner, A., T. S. Mikkelsen, H. Gu, M. Wernig, J. Hanna, A. Sivachenko, X. Zhang, B. E. Bernstein, C. Nusbaum, D. B. Jaffe, A. Gnirke, R. Jaenisch, and E. S. Lander.** 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature **454:**766-70.
160. **Meric-Bernstam, F., and A. M. Gonzalez-Angulo.** 2009. Targeting the mTOR signaling network for cancer therapy. J Clin Oncol **27:**2278-87.
161. **Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T. K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch,**

C. Nusbaum, E. S. Lander, and B. E. Bernstein. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature **448:**553-60.
162. **Miller, D. R.** 2006. A tribute to Sidney Farber-- the father of modern chemotherapy. Br J Haematol **134:**20-6.
163. **Mito, Y., J. G. Henikoff, and S. Henikoff.** 2007. Histone replacement marks the boundaries of cis-regulatory domains. Science **315:**1408-11.
164. **Mohn, F., and D. Schubeler.** 2009. Genetics and epigenetics: stability and plasticity during cellular differentiation. Trends Genet **25:**129-36.
165. **Moran, R. G., S. W. Baldwin, E. C. Taylor, and C. Shih.** 1989. The 6S- and 6R-diastereomers of 5, 10-dideaza-5, 6, 7, 8-tetrahydrofolate are equiactive inhibitors of de novo purine synthesis. J Biol Chem **264:**21047-51.
166. **Moran, R. G., and P. D. Colman.** 1984. Measurement of folylpolyglutamate synthetase in mammalian tissues. Anal Biochem **140:**326-42.
167. **Moran, R. G., W. C. Werkheiser, and S. F. Zakrzewski.** 1976. Folate metabolism in mammalian cells in culture. I Partial characterization of the folate derivatives present in L1210 mouse leukemia cells. J Biol Chem **251:**3569-75.
168. **Mujtaba, S., L. Zeng, and M. M. Zhou.** 2007. Structure and acetyl-lysine recognition of the bromodomain. Oncogene **26:**5521-7.
169. **Muse, G. W., D. A. Gilchrist, S. Nechaev, R. Shah, J. S. Parker, S. F. Grissom, J. Zeitlinger, and K. Adelman.** 2007. RNA polymerase is poised for activation across the genome. Nat Genet **39:**1507-11.
170. **Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird.** 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature **393:**386-9.
171. **Neshat, M. S., I. K. Mellinghoff, C. Tran, B. Stiles, G. Thomas, R. Petersen, P. Frost, J. J. Gibbons, H. Wu, and C. L. Sawyers.** 2001. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. Proc Natl Acad Sci U S A **98:**10314-9.
172. **Nielsen, P. R., D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina, and E. D. Laue.** 2002. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature **416:**103-7.
173. **Noh, W. C., W. H. Mondesire, J. Peng, W. Jian, H. Zhang, J. Dong, G. B. Mills, M. C. Hung, and F. Meric-Bernstam.** 2004. Determinants of rapamycin sensitivity in breast cancer cells. Clin Cancer Res **10:**1013-23.
174. **Nojima, H., C. Tokunaga, S. Eguchi, N. Oshiro, S. Hidayat, K. Yoshino, K. Hara, N. Tanaka, J. Avruch, and K. Yonezawa.** 2003. The mammalian target of rapamycin (mTOR) partner, raptor, binds the mTOR substrates p70 S6 kinase and 4E-BP1 through their TOR signaling (TOS) motif. J Biol Chem **278:**15461-4.
175. **Noronha, J. M., and V. S. Aboobaker.** 1963. Studies on the folate compounds of human blood. Arch Biochem Biophys **101:**445-7.
176. **Nourse, J., E. Firpo, W. M. Flanagan, S. Coats, K. Polyak, M. H. Lee, J. Massague, G. R. Crabtree, and J. M. Roberts.** 1994. Interleukin-2-mediated

elimination of the p27Kip1 cyclin-dependent kinase inhibitor prevented by rapamycin. Nature **372:**570-3.

177. **Okano, M., D. W. Bell, D. A. Haber, and E. Li.** 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell **99:**247-57.
178. **Okano, M., S. Xie, and E. Li.** 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet **19:**219-20.
179. **Okano, M., S. Xie, and E. Li.** 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res **26:**2536-40.
180. **Palii, S. S., B. O. Van Emburgh, U. T. Sankpal, K. D. Brown, and K. D. Robertson.** 2008. DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. Mol Cell Biol **28:**752-71.
181. **Pandya, K. J., S. Dahlberg, M. Hidalgo, R. B. Cohen, M. W. Lee, J. H. Schiller, and D. H. Johnson.** 2007. A randomized, phase II trial of two dose levels of temsirolimus (CCI-779) in patients with extensive-stage small-cell lung cancer who have responding or stable disease after induction chemotherapy: a trial of the Eastern Cooperative Oncology Group (E1500). J Thorac Oncol **2:**1036-41.
182. **Pao, W., V. Miller, M. Zakowski, J. Doherty, K. Politi, I. Sarkaria, B. Singh, R. Heelan, V. Rusch, L. Fulton, E. Mardis, D. Kupfer, R. Wilson, M. Kris, and H. Varmus.** 2004. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A **101:**13306-11.
183. **Perchiniak, E., S. A. Lawrence, S. Kasten, B. A. Woodard, S. M. Taylor, and R. G. Moran.** 2007. Probing the mechanism of the hamster mitochondrial folate transporter by mutagenesis and homology modeling. Biochemistry **46:**1557-67.
184. **Peterlin, B. M., and D. H. Price.** 2006. Controlling the elongation phase of transcription with P-TEFb. Mol Cell **23:**297-305.
185. **Podsypanina, K., R. T. Lee, C. Politis, I. Hennessy, A. Crane, J. Puc, M. Neshat, H. Wang, L. Yang, J. Gibbons, P. Frost, V. Dreisbach, J. Blenis, Z. Gaciong, P. Fisher, C. Sawyers, L. Hedrick-Ellenson, and R. Parsons.** 2001. An inhibitor of mTOR reduces neoplasia and normalizes p70/S6 kinase activity in Pten+/- mice. Proc Natl Acad Sci U S A **98:**10320-5.
186. **Pokholok, D. K., C. T. Harbison, S. Levine, M. Cole, N. M. Hannett, T. I. Lee, G. W. Bell, K. Walker, P. A. Rolfe, E. Herbolsheimer, J. Zeitlinger, F. Lewitter, D. K. Gifford, and R. A. Young.** 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. Cell **122:**517-27.
187. **Proudfoot, N. J.** 2003. Dawdling polymerases allow introns time to splice. Nat Struct Biol **10:**876-8.
188. **Proudfoot, N. J.** 1986. Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation. Nature **322:**562-5.

189. **Pui, C. H., and W. E. Evans.** 2006. Treatment of acute lymphoblastic leukemia. N Engl J Med **354:**166-78.
190. **Qiu, A., M. Jansen, A. Sakaris, S. H. Min, S. Chattopadhyay, E. Tsai, C. Sandoval, R. Zhao, M. H. Akabas, and I. D. Goldman.** 2006. Identification of an intestinal folate transporter and the molecular basis for hereditary folate malabsorption. Cell **127:**917-28.
191. **Ray, M. S., F. M. Muggia, C. G. Leichman, S. M. Grunberg, R. L. Nelson, R. W. Dyke, and R. G. Moran.** 1993. Phase I study of (6R)-5,10-dideazatetrahydrofolate: a folate antimetabolite inhibitory to de novo purine synthesis. J Natl Cancer Inst **85:**1154-9.
192. **Rea, S., F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis, and T. Jenuwein.** 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature **406:**593-9.
193. **Reik, W., W. Dean, and J. Walter.** 2001. Epigenetic reprogramming in mammalian development. Science **293:**1089-93.
194. **Rhee, I., K. W. Jair, R. W. Yen, C. Lengauer, J. G. Herman, K. W. Kinzler, B. Vogelstein, S. B. Baylin, and K. E. Schuebel.** 2000. CpG methylation is maintained in human cancer cells lacking DNMT1. Nature **404:**1003-7.
195. **Roberts, J. D., E. A. Poplin, M. B. Tombes, B. Kyle, D. V. Spicer, S. Grant, T. Synold, and R. Moran.** 2000. Weekly lometrexol with daily oral folic acid is appropriate for phase II evaluation. Cancer Chemother Pharmacol **45:**103-10.
196. **Roberts, L. J.** 2004. New drugs for rheumatoid arthritis. N Engl J Med **351:**2659-61; author reply 2659-61.
197. **Roenigk, H. H., Jr., R. Auerbach, H. Maibach, G. Weinstein, and M. Lebwohl.** 1998. Methotrexate in psoriasis: consensus conference. J Am Acad Dermatol **38:**478-85.
198. **Roh, T. Y., S. Cuddapah, and K. Zhao.** 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. Genes Dev **19:**542-52.
199. **Rosenblatt, D. S., V. M. Whitehead, N. Vera, A. Pottier, M. Dupont, and M. J. Vuchich.** 1978. Prolonged inhibition of DNA synthesis associated with the accumulation of methotrexate polyglutamates by cultured human cells. Mol Pharmacol **14:**1143-7.
200. **Roy, K., K. Mitsugi, and F. M. Sirotnak.** 1997. Additional organizational features of the murine folylpolyglutamate synthetase gene. Two remotely situated exons encoding an alternate 5' end and proximal open reading frame under the control of a second promoter. J Biol Chem **272:**5587-93.
201. **Roy, K., K. Mitsugi, and F. M. Sirotnak.** 1996. Organization and alternate splicing of the murine folylpolyglutamate synthetase gene. Different splice variants in L1210 cells encode mitochondrial or cytosolic forms of the enzyme. J Biol Chem **271:**23820-7.

202. **Roy, K., Mitsugi, K., Sirotnak, F.M.** 1997. Additional organizational features of the murine folylpolyglutamate synthetase gene. Two remotely situated exons encoding an alternate 5' end and proximal open reading frame under the control of a second promoter. J Biol Chem **272:**5587-5593.
203. **Rupon, J. W., S. Z. Wang, K. Gaensler, J. Lloyd, and G. D. Ginder.** 2006. Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human betaYAC transgenic mice. Proc Natl Acad Sci U S A **103:**6617-22.
204. **Sanders, M. J., P. O. Grondin, B. D. Hegarty, M. A. Snowden, and D. Carling.** 2007. Investigating the mechanism for AMP activation of the AMP-activated protein kinase cascade. Biochem J **403:**139-48.
205. **Sanghani, S. P., and R. G. Moran.** 1997. Tight binding of folate substrates and inhibitors to recombinant mouse glycinamide ribonucleotide formyltransferase. Biochemistry **36:**10506-16.
206. **Santos, F., B. Hendrich, W. Reik, and W. Dean.** 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol **241:**172-82.
207. **Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. Emre, S. L. Schreiber, J. Mellor, and T. Kouzarides.** 2002. Active genes are tri-methylated at K4 of histone H3. Nature **419:**407-11.
208. **Sarbassov, D. D., S. M. Ali, D. H. Kim, D. A. Guertin, R. R. Latek, H. Erdjument-Bromage, P. Tempst, and D. M. Sabatini.** 2004. Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. Curr Biol **14:**1296-302.
209. **Sarbassov, D. D., D. A. Guertin, S. M. Ali, and D. M. Sabatini.** 2005. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. Science **307:**1098-101.
210. **Sarraf, S. A., and I. Stancheva.** 2004. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell **15:**595-605.
211. **Scagliotti, G. V., P. Parikh, J. von Pawel, B. Biesma, J. Vansteenkiste, C. Manegold, P. Serwatowski, U. Gatzemeier, R. Digumarti, M. Zukin, J. S. Lee, A. Mellemgaard, K. Park, S. Patil, J. Rolski, T. Goksel, F. de Marinis, L. Simms, K. P. Sugarman, and D. Gandara.** 2008. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol **26:**3543-51.
212. **Schilsky, R. L., B. D. Bailey, and B. A. Chabner.** 1980. Methotrexate polyglutamate synthesis by cultured human breast cancer cells. Proc Natl Acad Sci U S A **77:**2919-22.
213. **Schmitz, K. M., N. Schmitt, U. Hoffmann-Rohrer, A. Schafer, I. Grummt, and C. Mayer.** 2009. TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. Mol Cell **33:**344-53.

214. **Schones, D. E., K. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. Wang, G. Wei, and K. Zhao.** 2008. Dynamic regulation of nucleosome positioning in the human genome. Cell **132:**887-98.
215. **Schrem, H., J. Klempnauer, and J. Borlak.** 2004. Liver-enriched transcription factors in liver function and development. Part II: the C/EBPs and D site-binding protein in cell cycle control, carcinogenesis, circadian gene regulation, liver regeneration, apoptosis, and liver-specific gene regulation. Pharmacol Rev **56:**291-330.
216. **Schubeler, D., D. M. MacAlpine, D. Scalzo, C. Wirbelauer, C. Kooperberg, F. van Leeuwen, D. E. Gottschling, L. P. O'Neill, B. M. Turner, J. Delrow, S. P. Bell, and M. Groudine.** 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev **18:**1263-71.
217. **Sehgal, S. N., H. Baker, and C. Vezina.** 1975. Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. J Antibiot (Tokyo) **28:**727-32.
218. **Shahbazian, M. D., and M. Grunstein.** 2007. Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem **76:**75-100.
219. **Shaw, R. J., N. Bardeesy, B. D. Manning, L. Lopez, M. Kosmatka, R. A. DePinho, and L. C. Cantley.** 2004. The LKB1 tumor suppressor negatively regulates mTOR signaling. Cancer Cell **6:**91-9.
220. **Shaw, R. J., M. Kosmatka, N. Bardeesy, R. L. Hurley, L. A. Witters, R. A. DePinho, and L. C. Cantley.** 2004. The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. Proc Natl Acad Sci U S A **101:**3329-35.
221. **Shi, Y., J. Gera, L. Hu, J. H. Hsu, R. Bookstein, W. Li, and A. Lichtenstein.** 2002. Enhanced sensitivity of multiple myeloma cells containing PTEN mutations to CCI-779. Cancer Res **62:**5027-34.
222. **Shi, Y., F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, and R. A. Casero.** 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell **119:**941-53.
223. **Shi, Y., H. Yan, P. Frost, J. Gera, and A. Lichtenstein.** 2005. Mammalian target of rapamycin inhibitors activate the AKT kinase in multiple myeloma cells by up-regulating the insulin-like growth factor receptor/insulin receptor substrate-1/phosphatidylinositol 3-kinase cascade. Mol Cancer Ther **4:**1533-40.
224. **Shih, C., V. J. Chen, L. S. Gossett, S. B. Gates, W. C. MacKellar, L. L. Habeck, K. A. Shackelford, L. G. Mendelsohn, D. J. Soose, V. F. Patel, S. L. Andis, J. R. Bewley, E. A. Rayl, B. A. Moroson, G. P. Beardsley, W. Kohler, M. Ratnam, and R. M. Schultz.** 1997. LY231514, a pyrrolo[2,3-d]pyrimidine-based antifolate that inhibits multiple folate-requiring enzymes. Cancer Res **57:**1116-23.
225. **Shilatifard, A.** 2006. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem **75:**243-69.

226. **Smith, S. G., N. L. Lehman, and R. G. Moran.** 1993. Cytotoxicity of antifolate inhibitors of thymidylate and purine synthesis to WiDr colonic carcinoma cells. Cancer Res **53:**5697-706.
227. **Song, F., J. F. Smith, M. T. Kimura, A. D. Morrow, T. Matsuyama, H. Nagase, and W. A. Held.** 2005. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc Natl Acad Sci U S A **102:**3336-41.
228. **Spath, G. F., and M. C. Weiss.** 1997. Hepatocyte nuclear factor 4 expression overcomes repression of the hepatic phenotype in dedifferentiated hepatoma cells. Mol Cell Biol **17:**1913-22.
229. **Stover, P. J.** 2004. Physiology of folate and vitamin B12 in health and disease. Nutr Rev **62:**S3-12; discussion S13.
230. **Struhl, K.** 1998. Histone acetylation and transcriptional regulatory mechanisms. Genes Dev **12:**599-606.
231. **Suter, M., U. Riek, R. Tuerk, U. Schlattner, T. Wallimann, and D. Neumann.** 2006. Dissecting the role of 5'-AMP for allosteric stimulation, activation, and deactivation of AMP-activated protein kinase. J Biol Chem **281:**32207-16.
232. **Takai, D., and P. A. Jones.** 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A **99:**3740-5.
233. **Takai, D., and P. A. Jones.** 2004. Origins of bidirectional promoters: computational analyses of intergenic distance in the human genome. Mol Biol Evol **21:**463-7.
234. **Tamaru, H., and E. U. Selker.** 2001. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature **414:**277-83.
235. **Tate, P. H., and A. P. Bird.** 1993. Effects of DNA methylation on DNA-binding proteins and gene expression. Curr Opin Genet Dev **3:**226-31.
236. **Taverna, S. D., B. M. Ueberheide, Y. Liu, A. J. Tackett, R. L. Diaz, J. Shabanowitz, B. T. Chait, D. F. Hunt, and C. D. Allis.** 2007. Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A **104:**2086-91.
237. **Taylor, E. C., J. M. Hamby, C. Shih, G. B. Grindey, S. M. Rinzel, G. P. Beardsley, and R. G. Moran.** 1989. Synthesis and antitumor activity of 5-deaza-5,6,7,8-tetrahydrofolic acid and its N10-substituted analogues. J Med Chem **32:**1517-22.
238. **Taylor, E. C., D. Kuhnt, C. Shih, S. M. Rinzel, G. B. Grindey, J. Barredo, M. Jannatipour, and R. G. Moran.** 1992. A dideazatetrahydrofolate analogue lacking a chiral center at C-6, N-[4-[2-(2-amino-3,4-dihydro-4-oxo-7H-pyrrolo[2,3-d]pyrimidin-5- yl)ethyl]benzoyl]-L-glutamic acid, is an inhibitor of thymidylate synthase. J Med Chem **35:**4450-4.
239. **Taylor, R. T., and M. L. Hanna.** 1977. Folate-dependent enzymes in cultured Chinese hamster cells: folypolyglutamate synthetase and its absence in mutants auxotrophic for glycine + adenosine + thymidine. Arch Biochem Biophys **181:**331-4.

240. **Thomson, A. W., H. R. Turnquist, and G. Raimondi.** 2009. Immunoregulatory functions of mTOR inhibition. Nat Rev Immunol **9:**324-37.
241. **Titus, S. A., and R. G. Moran.** 2000. Retrovirally mediated complementation of the glyB phenotype. Cloning of a human gene encoding the carrier for entry of folates into mitochondria. J Biol Chem **275:**36811-7.
242. **Trasler, J. M., D. G. Trasler, T. H. Bestor, E. Li, and F. Ghibu.** 1996. DNA methyltransferase in normal and Dnmtn/Dnmtn mouse embryos. Dev Dyn **206:**239-47.
243. **Turner, F. B.** 2002. Expression patterns of mouse and human folylpolyglutamate synthetase. Virginia Commonwealth University
244. **Turner, F. B., J. L. Andreassi 2nd, J. Ferguson, S. Titus, A. Tse, S. M. Taylor, and R. G. Moran.** 1999. Tissue-specific expression of functional isoforms of mouse folypoly-gamma-glutamae synthetase: a basis for targeting folate antimetabolites. Cancer Res **59:**6074-9.
245. **Turner, F. B., S. M. Taylor, and R. G. Moran.** 2000. Expression patterns of the multiple transcripts from the folylpolyglutamate synthetase gene in human leukemias and normal differentiated tissues. J Biol Chem **275:**35960-8.
246. **Um, S. H., F. Frigerio, M. Watanabe, F. Picard, M. Joaquin, M. Sticker, S. Fumagalli, P. R. Allegrini, S. C. Kozma, J. Auwerx, and G. Thomas.** 2004. Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. Nature **431:**200-5.
247. **Vakoc, C. R., M. M. Sachdeva, H. Wang, and G. A. Blobel.** 2006. Profile of histone lysine methylation across transcribed mammalian chromatin. Mol Cell Biol **26:**9185-95.
248. **Vander Haar, E., S. I. Lee, S. Bandhakavi, T. J. Griffin, and D. H. Kim.** 2007. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. Nat Cell Biol **9:**316-23.
249. **Vettese-Dadey, M., P. A. Grant, T. R. Hebbes, C. Crane- Robinson, C. D. Allis, and J. L. Workman.** 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J **15:**2508-18.
250. **Vilkaitis, G., I. Suetake, S. Klimasauskas, and S. Tajima.** 2005. Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. J Biol Chem **280:**64-72.
251. **Viollet, B., L. Lantier, J. Devin-Leclerc, S. Hebrard, C. Amouyal, R. Mounier, M. Foretz, and F. Andreelli.** 2009. Targeting the AMPK pathway for the treatment of Type 2 diabetes. Front Biosci **14:**3380-400.
252. **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:**343-56.

253. **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:**7395-403.
254. **Walsh, C. P., and T. H. Bestor.** 1999. Cytosine methylation and mammalian development. Genes Dev **13:**26-34.
255. **Wang, J., S. Hevi, J. K. Kurash, H. Lei, F. Gay, J. Bajko, H. Su, W. Sun, H. Chang, G. Xu, F. Gaudet, E. Li, and T. Chen.** 2009. The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. Nat Genet **41:**125-9.
256. **Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Paabo, M. Rebhan, and D. Schubeler.** 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet **39:**457-66.
257. **Wilkie, G. S., K. S. Dickson, and N. K. Gray.** 2003. Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. Trends Biochem Sci **28:**182-8.
258. **Winder, W. W., and D. G. Hardie.** 1996. Inactivation of acetyl-CoA carboxylase and activation of AMP-activated protein kinase in muscle during exercise. Am J Physiol **270:**E299-304.
259. **Wu, C. H., Y. Yamaguchi, L. R. Benjamin, M. Horvat-Gordon, J. Washinsky, E. Enerly, J. Larsson, A. Lambertsson, H. Handa, and D. Gilmour.** 2003. NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev **17:**1402-14.
260. **Wyngaarden, J. B., and D. M. Ashton.** 1959. Feedback control of purine biosynthesis by purine ribonucleotides. Nature **183:**747-8.
261. **Xiao, T., C. F. Kao, N. J. Krogan, Z. W. Sun, J. F. Greenblatt, M. A. Osley, and B. D. Strahl.** 2005. Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol Cell Biol **25:**637-51.
262. **Xiao, T., Y. Shibata, B. Rao, R. N. Laribee, R. O'Rourke, M. J. Buck, J. F. Greenblatt, N. J. Krogan, J. D. Lieb, and B. D. Strahl.** 2007. The RNA polymerase II kinase Ctk1 regulates positioning of a 5' histone methylation boundary along genes. Mol Cell Biol **27:**721-31.
263. **Xie, L. Y.** 2008. Creation of a Mouse with a Humanized fpgs Gene Compatible with Normal
Development. Virginia Commonwealth University, Richmond, Va.
264. **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.
265. **Zakikhani, M., R. Dowling, I. G. Fantus, N. Sonenberg, and M. Pollak.** 2006. Metformin is an AMP kinase-dependent growth inhibitor for breast cancer cells. Cancer Res **66:**10269-73.
266. **Zaret, K. S.** 2000. Liver specification and early morphogenesis. Mech Dev **92:**83-8.

267. **Zeitlinger, J., A. Stark, M. Kellis, J. W. Hong, S. Nechaev, K. Adelman, M. Levine, and R. A. Young.** 2007. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet **39:**1512-6.
268. **Zeng, P. Y., and S. L. Berger.** 2006. LKB1 is recruited to the p21/WAF1 promoter by p53 to mediate transcriptional activation. Cancer Res **66:**10701-8.
269. **Zhao, R., L. H. Matherly, and I. D. Goldman.** 2009. Membrane transporters and folate homeostasis: intestinal absorption and transport into systemic compartments and tissues. Expert Rev Mol Med **11:**e4.
270. **Zhou, G., R. Myers, Y. Li, Y. Chen, X. Shen, J. Fenyk-Melody, M. Wu, J. Ventre, T. Doebber, N. Fujii, N. Musi, M. F. Hirshman, L. J. Goodyear, and D. E. Moller.** 2001. Role of AMP-activated protein kinase in mechanism of metformin action. J Clin Invest **108:**1167-74.

Appendix A

Primer pairs are listed together

Chapter 2

Primers used across the length of *fpgs* gene: 5' to 3':

Primer 1 forward:	5'TCCCTCCGCAGCTTCCTG3'
Primer 1 reverse:	5'ACTGACTTGCCATCTCCCC3'
Exon A1a 1a2:	5'CTTGATGTCATTTGCAGTCCAA3'
Exon A1a 2b:	5'GTTCGCATTTCCATCCTGTG3'
Primer 3 forward:	5'GACTCAGGGGTGTTGCTCT3'
Primer 3 reverse:	5'CAGGTTCAACAAGAGATCGAGAG3'
Primer 4 forward:	5'GCGTCTTGGGTTAGTGGCAG3'
Primer 4 reverse:	5'CCATAAGTCACGTAGGCAATGC3'
Primer 5 forward:	5'GGTTCAGATGCACTCGCTTG3'
Primer 5 reverse:	5'CTGGGCAGTGGTAGCACA3'
Primer 6 forward:	5'GAGAAAGCCTGTGGCAGG3'
Primer 6 reverse:	5'AGGAATACAGTTGTAAGCCAGG3'
Exon 1 (3a):	5'GCTTCTCTGCCAAGGAGTCG3'
Exon 1 (4b):	5'GCCCTGACGCCATCCTAA3'
Primer 8 forward:	5'GCACAGGTTTCTCAGAATGTAGGG3'
Primer 8 reverse:	5'CTCTCTACCACTATACTGGACACC3'
Primer 9 forward:	5'AGTAGTGGGATTAAAGGCGTGC3'
Primer 9 reverse:	5' ATGGGTACTGCTGGTTGTCG3'
Primer 10 forward:	5'TGCCAGGGATTAGCTGTGG3'
Primer 10 reverse:	5'CTGCCCAGCTTCCATGTCTTTA3'
Primer 11 forward:	5'GACTGTGGCAAGTCATTACCAGG3'
Primer 11 reverse:	5'CACATGAGGAACCCAAAGC3'
Primer 12 forward:	5'CTACACATCCCAGTGCAGG3'
Primer 12 reverse:	5'AGTCCCTGTTCCTCACTGAGC3'

Primer used for Chip walking across P1 and P2:

P1 promoter

Exon A1a (1a2)	5'CTTGATGTCATTTGCAGTCCAAG3'
Exon A1a (1b)	5' CTGGGAGATGTCCAAGTCCTG3'
Exon A1a (2a)	5'GTCTCACCTGGGCAGCAAG3'
Exon A1a (2b)	5'GTTCGCATTTCCATCCTGTG3'
Exon A1a (3a)	5'CAGCTAGCTAGCCCTGTAGTGG3'
Exon A1a (3b)	5'GGTAGGGCACAAATGACTCCAAG3'
Exon A1a (4c)	5'TGCAGTCCCAGTCCAGC3'
Exon A1a (4d)	5'AGGCAGTCTTAGCTTCGTAAGC3'
Exon A1a (5e)	5'GTGGAAGGACCCATCAGGGA3'
Exon A1a (5f)	5'AGTAGGGTTCCTAACATCACCTGC3'
Exon A1a (6a2)	5'GTGTGCCAGTCTGATGCACATTAG3'
Exon A1a (6b)	5'GGAATGAGGCGACCTTGACC3'

P2 promoter

Exon (1c)	5'CGATAGGGGTGTCATTCTTCC3'
Exon (1d)	5'GCAGCCTTTTCGGTCAAG3'
Exon 1 (2a)	5'AGCCTGGAGAGACGGGGT3'
Exon 1 (2b)	5'CCAATCCGAGCCAGCAGTTC3'
Exon 1 (3a)	5'GCTTCTCTGCCAAGGAGTCG3'
Exon (3b)	5'GCTGCTCTCGCCATGCTTC3'
Exon 1 (4a)	5'CTCCCAGGAACACCCTCTTC3'
Exon 1 (4b)	5'GCCCTGACGCCATCCTAAG3'
Exon 1 (5c)	5'CGCTGCTTCTCATTGGTC3'
Exon 1 (5d)	5'GCGGGAACGGTCAGAAA3'
Exon 1 (6a)	5'GAGCCGACTGTGCTCGACTC3'
Exon 1 (6b2)	5'ACGTGTCAGGACGGGACC3'

Control primers for histone and RNAPII studies:

B-globin Set #2 forward (coding region)	5'AGGCTGCTGGTTGTCTACC3'
B-globin Set#2 reverse (coding region)	5'ACACTCCACACACAGTCATGG3'
GAPDH Set#1 forward (promoter)	5'GTCCTCGATGTCCTTGGTGC3'
GAPDH Set#1 reverse #2 (promoter)	5'TGGAACAGGGAGGAGCAGAG3'
GAPDH Set #2 forward (coding region)	5'TCCAGTATCACTCCACTGACG3'
GAPDH Set #2 reverse (coding region)	5'CAGGTTGCACCATATCAAGGG3'

Chapter 3

Control primers for P and PM cells lines:

Cryaa Forward (Exon 1-2)	5'ATCTCTGAGGTCCGATCTAGCC3'
Cryaa Reverse (Exon 2-3)	5'GTAGCCATGGTCATCCTGCCTC3'
Slpi Forward (Exon 1)	5'TCACGGTGCTCCTTGCTCTG3'
Slpi Reverse (Exon 2-3)	5'CTTCCTCCACACTGGTTTCG3'

Expression control primers (Rt-PCR):

mGAPDH forward (Lin)	5'AACTTTGGCATTGTGGAAGGGCTC3'
mGAPDH reverse (Lin)	5'TGGAAGAGTGGGAGTTGCTGTTGA3

P1 expression primers (Rt-PCR):

A1b forward	5'TGTGGCTGAGAAGTTCTGGTGG3'
A1b reverse	5'CGCACAGCATCCTGAAAGG3'
Exon 3b reverse	5'CTTCCCTTTGGTCCCAGTGACATGAA3'
Exon 6R reverse	5'AGACATGGCTGTCGTCCTTGAACT3'
Exon 1 forward	5'CTAGCTTGCGGTGGCATTAAGACT3'

De novo methyltransferase primers (Rt-PCR):

DNMT3a forward	5'GTAACCTTCCTGGCATGAACA3'
DNMT3a reverse	5'ACTCTTGCAGCTCCAGCTTATC3'
DNMT3b forward	5'CTGAATTACACGCAGGACATGAC3'

DNMT3b reverse	5’TCAGAGCCATTCCCATCATCTAC3’

Bisulfite primers:

Sense: 201:	5’TAGGGTTTTTTTGAAGTTTATTTTGTA3’
Antisense BSS3:	5’CATCTAACTAACATCCTACCTAACAA3’
Sense 202C:	5’AGTTTTATTTGGGTAGTAAGGTTAATTTTAG3
Antisense BSS4:	5’TCCTAATAAATAAAACACAAATAACTCC3’
Sense 203A1A:	5’AGAAGTTTTGGTGGGAAGTAG3’
Antisense BSS11:	5’CTAATATCTCACCTACCTCCTATATAACAC3’
Sense 203C:	5’ATTTTTGTTGATAGTTATTTGGAGTAGGTG3’
Antisense BSS10	5’CCCCAACCTAACTATAATTAATCATCTA3’

VITA

Alexandra Christin Racanelli was born on September 15, 1979 in Brightwaters, NY. She graduated from St. Anthony's High School in 1997 and attended the University of Mary Washington for her undergraduate education. In 2001, she earned her Bachelor of Science in Biology with Honors after completing her thesis work entitled: "The effects of lipotechoic acid on ME-180 and P69 carcinoma cells". Following her graduation she attended the Pre-medical basic health certificate program at Virginia Commonwealth University in the department of Microbiology and Immunology. Subsequently, she received a research assistant position at Memorial Sloan-Kettering Cancer Center where she worked under the guidance of Dr. Gary K. Schwartz in the department of medicine. After her acceptance into the physician-scientist program at Virginia Commonwealth University she moved back to Richmond, VA where she has spent the past six years. For the first two years she was enrolled in the medical program and for the past four years she has been a graduate student of the Molecular Biology and Genetics program under the Pharmacology and Toxicology Department under the tutelage of Richard G. Moran, PhD. She was inducted into Phi Kappa Beta in 2007 and she has presented posters of her work at the American Association of Cancer Research in 2006 and the Gordon Research Conference on Chromatin Structure and Function in May 2008. In the fall of 2009, she will return to the medical program for her clinical rotations and plans to finish her medical degree in May 2011.

Manuscripts resulting from the present dissertation research

1. Racanelli, A.C., Turner, F.B., Xie, L.Y., Taylor, S.M. and Moran, R.G. (2008) A mouse gene that coordinates epigenetic controls and transcriptional interference to achieve tissue-specific expression. *Mol Cell Biol*. Jan; 28(2):836-48.

2. Racanelli, A.C., Rothbart, S.B., Heyer, C.H., and Moran, R.G. (2009) Therapeutics by cytotoxic accumulation: pemetrexed causes ZMP accumulation, AMPK activation, and mammalian target of rapamycin inhibition. *Cancer Res.* 2009 Jul 1;69(13):5467-74.