NUCLEAR *DE NOVO* THYMIDYLATE BIOSYNTHESIS COMPLEX FORMATION AND COMPUTATIONAL MODELING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Xu Lan

August 2020

NUCLEAR DE NOVO THYMIDYLATE BIOSYNTHESIS COMPLEX FORMATION AND COMPUTATIONAL MODELING

Xu Lan, Ph. D.

Cornell University 2020

Folate-mediated one-carbon metabolism (FOCM) consists of a network of metabolic pathways connected through the common use of folate derivatives that carry and donate one-carbon (1C) units for *de novo* biosynthesis of purines and thymidylate (dTMP or deoxythymidine monophosphate), and homocysteine remethylation. Impairments in FOCM are associated with common pathologies, including neural tube defects (NTDs), neurodegenerations, and cancer, but the causal mechanisms are not established. This dissertation research investigated the nutritional, biochemical and genetic factors that regulate FOCM, with a focus on *de novo* thymidylate biosynthesis pathway.

Regulation of the partitioning of folate cofactors among FOCM pathways is essential to address metabolic needs that fluctuate through cell cycle progression. The first objective aimed to summarize the evidence for temporal regulation of expression, activity and cellular localization of enzymes and pathways in the FOCM network in mammalian cells through the cell cycle.

The second objective sought to investigate the structure of the nuclear *de novo* thymidylate synthesis multienzyme complex and the nature of the protein-protein interactions. Multiple SUMO1 modification sites were identified on enzymes of the *de novo* dTMP biosynthesis pathway by mass spectrometry (MS) analysis. Further experiments are needed

characterize the protein-protein interactions within the complex.

Computational modeling is useful to access the complexity of FOCM network and guide the design of biological experiments. The third objective aimed to extend the current FOCM model to include the nuclear compartment. The model confirms that accounting for the kinetic effects of nuclear multienzyme complex formation and substrate channeling is essential for the functioning of *de novo* dTMP synthesis.

The fourth objective sought to investigate the unique requirement for glycine for C2C12 myoblasts proliferation. Glycine contributed 1C units to FOCM and *de novo* thymidylate biosynthesis. However, formate, hypoxanthine or thymidine supplementation failed to rescue the growth of C2C12 myoblasts under glycine depletion, suggesting that generating 1C units is not the primary mechanism by which glycine supports C2C12 myoblasts proliferation.

BIOGRAPHICAL SKETCH

Born in 1990, Xu Lan was raised in Tianjin, China. She earned a B.S. degree in Nutrition Science at Purdue University with highest distinction. She was in the Nutrition Science Honors Program and completed a research project under the mentorship of Dr. James Fleet. Her research focused on investigating the interaction of genetics and diet on lumbar spine bone density and structure in mice using microcomputed tomography. After finishing her B.S. degree in 2013, she entered the Ph.D. degree program in Nutritional Sciences at Cornell University. Xu joined the laboratory of Dr. Patrick Stover in the Spring of 2014 and investigated the nutritional, biochemical and genetic factors that regulate folate-mediated one-carbon metabolism. In addition to research, she has served as a teaching assistant for six and a half years for several undergraduate nutrition courses at Cornell University.

Dedicated to my parents

ACKNOWLEDGMENTS

First and foremost, I would like to express my deep and sincere gratitude to my advisor Dr. Patrick Stover for constant support and giving me invaluable guidance and advice. I would also like to thank Dr. Martha Field for being a mentor and always being available to help and give advice.

I would like to thank my committee members Dr. Robert S. Weiss, Dr. John C. Schimenti, Dr. Xingen Lei, and Dr. Zhenglong Gu for their time and advice.

I would like to thank Dr. Lynn Johnson and Francoise Vermeylen at CSCU for their help with statistical analysis for Chapter 5. I would like to thank Carol Bayles at the Institute of Biotechnology for her expertise and help with flow cytometry. I would like to thank Dr. Sheng Zhang and Dr. Ruchika Bhawal at the Institute of Biotechnology for their expertise and help with mass spectrometry for Chapter 3. I would like to thank Dr. Doletha Marian Szebenyi at MacCHESS for expertise and help with protein structure modeling for Chapter 3.

I would like to thank all members of the Stover Lab for their support and help with lab techniques, especially Dr. Ashley Palmer, Dr. Denise Stover, Dr. Erica Lachenauer, Dr. James Chon, Dr. Judith Alonzo, Dr. Sajin Bae, Alex Li and Kendra Tiani.

I would like to thank my friends, especially Xu Miao, Hongyu Chen, Tongtong Li, Kho Verian, and Jui-Yun Liao, for their kindness and support. Our friendship will always grow stronger regardless of the distance that separates us.

Finally, I am extremely grateful to my parents for their unlimited love, caring, support and encouragement. I would never have made it this far without them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1C	One-carbon
5-methylTHF	5-methyltetrahydrofolate
5fTHF	5-formyltetrahydrofolate
ANOVA	Analysis of variance
AdoHcy	S-adenosylhomocysteine
AdoMet	S-adenosylmethionine
CH2F	5,10-methyleneTHF
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DSBU	Disuccinimido sulfoxide
DSSO	Disuccinimido dibutyric urea
dT	Deoxythymidine
dTMP	Thymidylate or deoxythymidine monophosphate
dU	Deoxyuridine
dUMP	Deoxyuridylate
FOCM	Folate-mediated one-carbon metabolism
GlyT	Glycine transporter
GCS	Glycine cleavage system
GLDC	Glycine decarboxylase
*h*MPCs	Human muscle progenitor cells

HCY	Homocysteine
MET	Methionine
MS	Mass spectrometry
MTHFD1	Methylene-tetrahydrofolate dehydrogenase 1
MTHFD1L	Methylene-tetrahydrofolate dehydrogenase 1-like
MTHFD2	Methylene-tetrahydrofolate dehydrogenase 2
MTHFD2L	Methylene-tetrahydrofolate dehydrogenase 2-like
MTHFR	Methylenetetrahydrofolate reductase
MTHFS	Methenyltetrahydrofolate synthetase
MTR	Methionine synthase
NTDs	Neural tube defects
PML	Promyelocytic leukaemia
SHMT	Serine hydroxymethyltransferase
SIM	SUMO-interacting motif
SUMO	Small ubiquitin-like modifier
THF	Tetrahydrofolate
TYMS	Thymidylate synthase
TK1	Thymidine kinase 1
XL/MS	Cross-linking/mass spectrometry

Chapter 1 INTRODUCTION

Folate-mediated one-carbon metabolism (FOCM) consists of a network of metabolic reactions connected through the common use of folate derivatives that carry and donate 1C units for *de novo* biosynthesis of purines and thymidylate (dTMP), and remethylation of homocysteine to methionine. FOCM plays an essential role for genome stability and methylation, and disruption of the network can be caused by genetic variants of the required enzymes or nutrient deficiencies[1], [2]. Impaired folate status and disrupted FOCM are associated with numerous human disease outcomes, including megaloblastic anemia, neural tube defects (NTDs), neurodegenerations, and cancer[3]–[9]. However, understanding the causal pathways and mechanisms underlying these pathologies has been limited, because the pathways of FOCM are tightly interconnected.

FOCM is highly complex for several reasons: 1) pathways in FOCM compete for a limiting pool of folate cofactors[10]; 2) metabolic reactions in FOCM are subjected to long-range and indirect regulatory processes, 3) multi-enzyme complexes exist in FOCM, including the *de novo* thymidylate synthesis complex and *de novo* purine synthesis complex referred to as "purinosome", 4) pathways in FOCM are compartmentalized in the cytosol, nucleus and mitochondria, 5) Genetic factors cause perturbations in FOCM through changes in coding and expression variants in folate-dependent enzymes; 6) FOCM interacts with other cellular metabolic pathways; 7) the FOCM network is sensitive to nutritional status for several vitamins that serve as enzyme cofactors, including folate, riboflavin, vitamin B6 and vitamin B12[11].

The overarching goal of this dissertation research was to investigate the nutritional, biochemical and genetic factors that regulate FOCM, with a focus on *de novo* thymidylate biosynthesis pathway. Elucidating the regulation of FOCM and nuclear *de novo* thymidylate

biosynthesis is essential to understanding the underlying mechanisms of folate associated pathologies. The results will inform future human studies and facilitate drug development for the prevention and treatment of folate associated pathologies.

The *de novo* dTMP synthesis pathway consists of four enzymes, including serine hydroxymethyltransferase (SHMT1 and SHMT2α), thymidylate synthase (TYMS), dihydrofolate reductase (DHFR), and methylene-tetrahydrofolate dehydrogenase 1 (MTHFD1). Specifically, TYMS utilizes 5,10-methyleneTHF (CH2F) as a cofactor in the conversion of deoxyuridylate (dUMP) to deoxythymidylate (dTMP), oxidizing tetrahydrofolate (THF) to dihydrofolate (DHF). DHFR catalyzes the NADPH-dependent reduction of DHF back to THF. CH2F is independently generated by SHMT (SHMT1 and SHMT2α) and MTHFD1 using serine and formate as 1C sources, respectively. Impaired *de novo* dTMP synthesis can lead to inadequate dTMP biosynthesis activity and consequently to elevated uracil misincorporation in DNA[1], [12], [13]. During DNA synthesis, either dTTP or dUTP can basepair with an adenine nucleotide base on the template strand. Therefore, proper functioning of *de novo* dTMP synthesis is important for genome stability.

During S and G_2/M phases of the cell cycle or in response to DNA damage, a fraction of enzymes involved in *de novo* dTMP synthesis pathway are SUMOylated and translocate to the nucleus[14], [15]. Enzymes in *de novo* dTMP synthesis pathway form a multienzyme complex that is associated with the replication and epigenetic machinery[16]. Formation of a multienzyme complex is essential for the functioning of the *de novo* dTMP synthesis pathway, as *de novo* dTMP synthesis activity in isolated nuclei is impaired following sonication[16]. Nuclear *de novo* dTMP synthesis complex formation may allow for channeling of folate cofactors among enzyme active sites, limiting substrate diffusion and accelerating enzymatic reaction rates[17]. Previous

studies show that SHMT acts as a scaffold protein that anchors this multienzyme complex to nuclear lamina[16], [18]; however, no direct interaction was found between SHMT and TYMS or DHFR using the yeast two hybrid assay[19]. The mechanism of nuclear *de novo* dTMP synthesis complex formation is not yet characterized.

This dissertation is divided into five chapters. Chapter 2 reviews the temporal regulation of expression, activity and cellular localization of enzymes and pathways in the FOCM network in mammalian cells through the cell cycle. Partitioning of folate cofactors among FOCM pathways is regulated to address metabolic needs that fluctuate through cell cycle progression.

Chapter 3 investigates the underlying mechanism of *de novo* thymidylate biosynthesis pathway complex formation in the nucleus. Enzymes in *de novo* dTMP synthesis pathway, including SHMT, TYMS, DHFR and MTHFD1, as well as Lamin A, can be modified by small ubiquitin-like modifier 1 (SUMO1) [14], [19]–[22]. SUMO-interacting motifs (SIMs) mediate non-covalent interactions between SUMO and SIM-containing proteins. SUMO-SIM interactions have been shown to facilitate protein complex assembly and function. We hypothesized that multiple SUMO-SIM interactions among enzymes in *de novo* dTMP synthesis pathway establish a scaffold for the complex assembly. SUMO1 modification sites on enzymes in *de novo* dTMP synthesis pathway were identified by mass spectrometry using purified recombinant SUMOylated proteins from engineered bacteria. Protein-protein interactions were investigated by *in vitro de novo* dTMP synthesis complex assembly and cross-linking/mass spectrometry (XL/MS).

Computational modeling is useful to access the complexity of FOCM and guide the design of biological experiments. In chapter 4, we extended the hybrid stochastic model of FOCM described previously by including the nuclear compartment. The effects of including

nuclear compartment and folate channeling in the FOCM network was explored with emphasis on rates of *de novo* thymidylate biosynthesis and the contribution of serine, glycine and formate as sources of one-carbons in FOCM.

Serine and glycine are non-essential amino acids that are interconvertible through FOCM. Both serine and glycine provide precursors for the biosynthesis of proteins, nucleic acids, and lipids. However, unlike cancer cells and T lymphocytes that require serine but not glycine for proliferation[23], [24], glycine alone is effective at stimulating muscle cell expansion[25]. In chapter 5, the unique requirement of glycine for C2C12 myoblasts proliferation was investigated. Specifically, we examined the role of glycine as a one-carbon source for FOCM and *de novo* thymidylate synthesis.

Chapter 2 CELL CYCLE REGULATION OF FOLATE-MEDIATED ONE-CARBON METABOLISM

Authors: Xu Lan, Martha S. Field and Patrick J. Stover

Affiliations: Cornell University, Division of Nutritional Sciences, Ithaca NY 14850

Author contributions:

Xu Lan (X.L.) wrote the original manuscript and revised it. P.J. Stover (P.J.S.), and M.S. Field (M.S.F.) revised the manuscript and provided critical feedback.

Abstract:

Folate-mediated one-carbon metabolism (FOCM) comprises a network of interconnected folate-dependent metabolic pathways responsible for serine and glycine interconversion, *de novo* purine synthesis, *de novo* thymidylate synthesis and homocysteine remethylation to methionine. These pathways are compartmentalized in the cytosol, nucleus and mitochondria. Individual enzymes within the FOCM network compete for folate cofactors because intracellular folate concentrations are limiting. Although there are feedback mechanisms that regulate the partitioning of folate cofactors among the folate-dependent pathways[26], [27], less recognized is the impact of cell cycle-regulation on FOCM. This review summarizes the evidence for temporal regulation of expression, activity and cellular localization of enzymes and pathways in the FOCM network in mammalian cells through the cell cycle.

2.1 Introduction

Folate-mediated one-carbon metabolism (FOCM) comprises a network of interconnected folate-dependent metabolic pathways responsible for serine and glycine interconversion, *de novo* purine synthesis, *de novo* thymidylate (dTMP) synthesis and homocysteine remethylation to methionine. These pathways are compartmentalized in the cytosol, nucleus and mitochondria.

They are interconnected through their dependency on one-carbon (1C) units derived primarily from serine and glycine catabolism, and through their reliance on folate cofactors, which carry and chemically activate 1C units at the oxidation states of formate, formaldehyde and methanol[11]. Intracellular folate concentrations are limiting for individual enzymatic reactions within the FOCM network, because the cellular concentration of folate-dependent enzymes exceeds by several fold the concentration of intracellular folate. Hence, the biosynthetic reactions of FOCM compete for folate cofactors[28]. Cells preferentially direct the flux of 1C units to meet changes in cellular demands due to stress conditions, including nutrient deprivation and disease states such as cancer[29], [30]. Although there are feedback mechanisms that regulate the partitioning of folate cofactors among the folate-dependent pathways[26], [27], there are also temporal differences in the activity and expression of the enzymes that constitute the various pathways within FOCM. Regulation of the partitioning of folate cofactors among FOCM pathways is essential to address metabolic needs that fluctuate through cell cycle progression, such as increased demand of dTMP for DNA replication during S phase of the cell cycle[16]. Temporal changes in FOCM metabolic inputs and outputs can be achieved through the regulation of FOCM enzyme levels at the level of transcription (table 2.1), translation and post-translational modification as well as regulation of FOCM enzyme subcellular localization. This review synthesizes the evidence for pathway-specific and cell cycle-dependent regulation of enzymes in or related to the FOCM pathway. Specifically, we focus on the temporal changes in the expression and cellular localization of these enzymes in mammalian cells as a function of the cell cycle.

Table 2.1. Genes involved in one-carbon metabolism network are regulated by transcription factors in a cell cycle-dependent manner

Gene name	Transcription factors	Cell cycle stage of transcription induction	Binding site	Reference
CAD	Myc/Max	G1/S-phase	promoter	[31]
DHFR	E2F		promoter	[32]
	Sp1	Throughout cell cycle	promoter	[33], [34]
MTHFD1	FOXM1		promoter	[35]
MTHFS	FOXM1			[35]
RRM1	E2F			[36], [37]
RRM2	E2F	S phase[38]		[36]–[39]
TK1	E2F			[36]
	FOXM1			[35]
TYMS	E2F	G_1 to S transition[40]		[36], [37], [40], [41]
	FOXM1			[35], [42]
	LSF	G_1/S	promoter and intronic regions	[43]

2.2 FOCM in the cytosol

Folate-activated 1C units in the FOCM network can be generated directly by the conversion of tetrahydrofolate (THF) and serine to glycine and 5,10-methyleneTHF catalyzed by SHMT1 or SHMT2α in the cytosol and nucleus. Alternatively, 10-formylTHF is formed from THF, ATP and formate catalyzed by MTHFD1 in the cytosol and nucleus. Formate is a major source of 1C units, and is derived primarily from THF-dependent serine and glycine catabolism in the mitochondria. Elevated plasma formate concentrations and decreased rates of *de novo* formate production were observed in the rats fed on folate-deficient diet compared with rats fed on folate-replete diet[44]. The relative contributions of endogenous formate production and exogenous formate uptake on cellular formate concentration remain unclear. Other sources of 1C units include the degradation of histidine in the cytosol[45], [46], and degradation of choline-derived methyl-glycine species, including betaine, dimethylglycine (DMG) and sarcosine (Figure

2.1). Choline can be oxidized to betaine via choline dehydrogenase and betaine aldehyde dehydrogenase. In the cytosol, betaine-homocysteine S-methyltransferase (BHMT) provides a 1C unit through the conversion of betaine and homocysteine to DMG and methionine in a pathway that is folate independent. DMG is catabolized to sarcosine and glycine in the mitochondria through dimethylglycine dehydrogenase and sarcosine dehydrogenase respectively, generating additional folate-activated 1C units. In the cytosol, formate-derived, folate-activated 1C units enter into the FOCM network and are incorporated into the end products of *de novo* purine biosynthesis, *de novo* dTMP biosynthesis, and the remethylation of homocysteine to methionine. Methionine can be converted to S-adenosyl-methionine (SAM), a cofactor that donates methyl groups for numerous methylation reactions, including methylation of DNA, RNA, proteins, phospholipids, and neurotransmitters[47], [48].

2.2.1 Methylenetetrahydrofolate dehydrogenase 1 (MTHFD1)

2.2.1.1 MTHFD1 expression

MTHFD1 is a trifunctional enzyme with 5, 10-methyleneTHF dehydrogenase, 5,10-methenylTHF cyclohydrolase and 10-formylTHF synthetase activities. In the cytosol, MTHFD1 is the source of formate-derived 1C units which are required for *de novo* purine synthesis (through 10-formylTHF synthesis) as well as *de novo* dTMP synthesis and homocysteine remethylation into methionine (through methyleneTHF synthesis). The *MTHFD1* gene is predominantly expressed in G_1/S and G_2 phase of the cell cycle in human foreskin fibroblasts[49]. Its promoter is bound by the G_2/M transcription factor FOXM1 in U2OS and Hela cells[35], although the mechanism of cell cycle regulation of *MTHFD1* gene expression remains unclear. Future study is needed to assess the changes in MTHFD1 protein levels throughout the cell cycle.

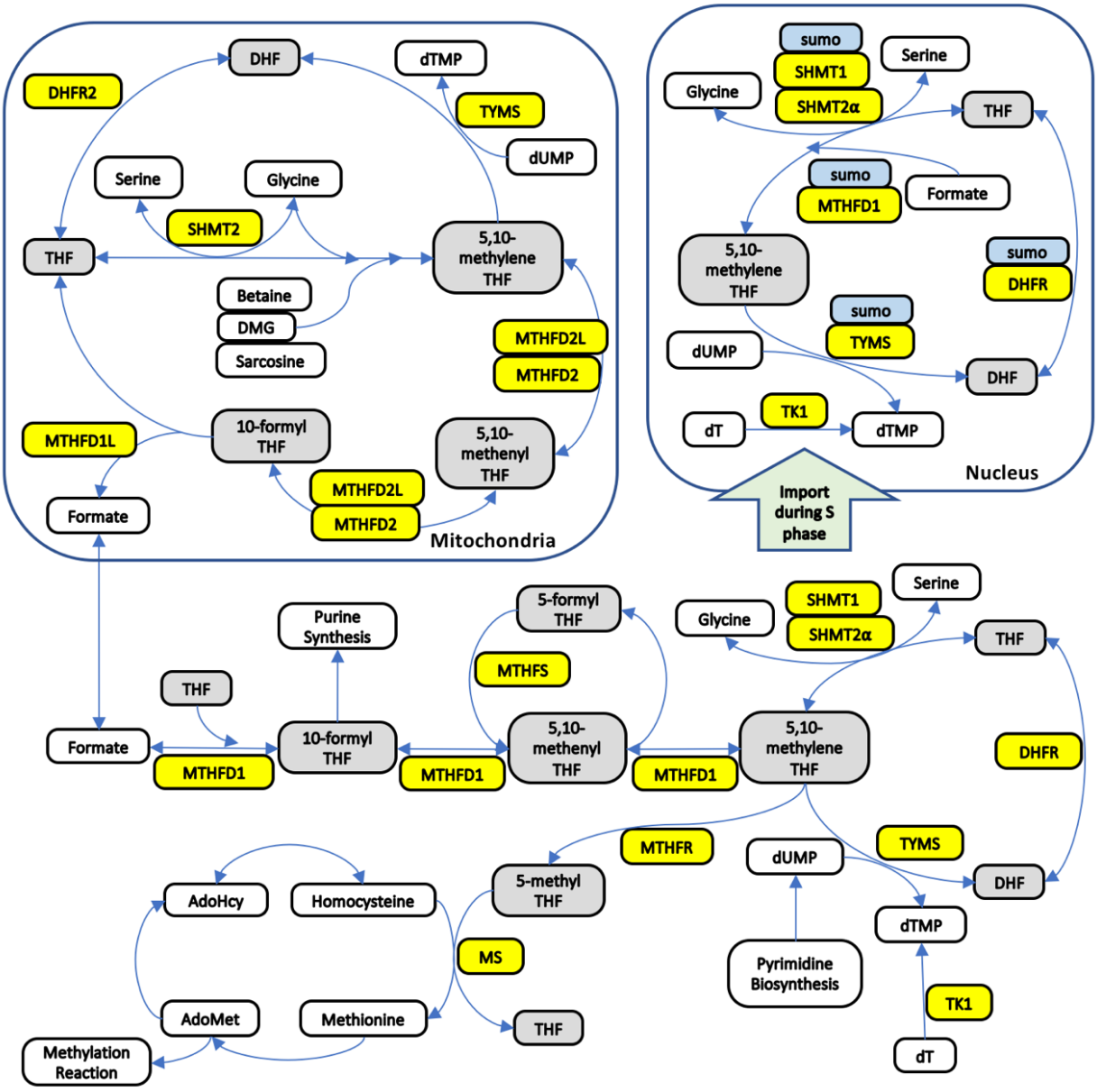

Figure 2.1. Compartmentation of folate-mediated one-carbon metabolism. One-carbon units in FOCM are derived from serine and histidine catabolism in the cytosol, and from the catabolism of serine, glycine and choline-derived methyl-glycine species in the mitochondria. FOCM in the cytoplasm is required for *de novo* purine synthesis, for *de novo* thymidylate synthesis and for homocysteine *remethylation* to methionine. FOCM in the nucleus is required for *de novo* thymidylate synthesis at sites of DNA replication. The *de novo* thymidylate synthesis pathway is also present in the mitochondria.

2.2.1.2 MTHFD1 subcellular localization

MTHFD1 protein translocates into the nucleus during S-phase, and the partitioning of MTHFD1 between the cytosol and nucleus is influenced by cellular folate levels, with increased MTHFD1 nuclear localization observed in response to folate depletion[15], [50]. A human MTHFD1-green fluorescent protein (GFP) fusion protein expressed in HeLa cells was observed in the nucleus, and the levels of MTHFD1-GFP fusion protein in the nucleus were nearly 2-fold higher in S-phase-arrested cells compared to cell arrested in either G_1 or G_2/M phase. Endogenous MTHFD1 protein nuclear localization also increased during S-phase in Hela cells[50]. In mouse liver, MTHFD1 nuclear localization was increased when mice were fed a diet lacking folate[50]. Human fibroblasts expressing an MTHFD1 variant with reduced activity exhibited increased MTHFD1 nuclear localization, suggesting MTHFD1 nuclear localization supports *de novo* dTMP biosynthesis in the nucleus at the expense of homocysteine remethylation pathway in the cytosol, especially during folate deficiency[15].

2.2.2 The *de novo* dTMP synthesis pathway

2.2.2.1 Pyrimidine biosynthesis

Pyrimidine biosynthesis is regulated through the cell cycle to meet increasing cellular demand for both cytosine and thymidine deoxyribonucleotides during DNA replication. The pyrimidine biosynthesis pathway was shown to be upregulated 2-fold prior to entry into S phase and decreased upon exit from S phase in baby hamster kidney cells[51]. Similarly, pyrimidine nucleotide concentrations are increased during growth in rat livers[52].

The carbamoyl-phosphate synthetase 2 gene (CAD) encodes a trifunctional protein that catalyzes the first 3 activities of the 6-step pyrimidine biosynthesis pathway: carbamoylphosphate synthetase (CPS II), aspartate transcarbamoylase (ATCase), and

dihydroorotase. The *cad* gene is regulated throughout the cell cycle at both the transcriptional and posttranscriptional levels. Steady state *cad* RNA levels increased 13-fold, peaking during mid to late G1 phase of the cell cycle, in serum stimulated Syrian hamster ts13 cells[53]. Newly synthesized CAD protein increased in parallel[53]. However, the rate of *cad* transcription increased only about 2-fold[53], indicating that posttranscriptional regulation plays an important role. Similarly, in serum stimulated mouse fibroblasts, steady-state levels of *cad* mRNA increased 10-fold whereas its transcription increased only 3-fold[54]. When teratocarcinoma cells were induced into differentiation, steady-state levels of *cad* mRNA decreased by 7-fold while the rate of *cad* mRNA transcription remained the same, suggesting that expression of the *cad* gene is cell-growth dependent and regulated at the posttranscriptional level[55]. Several transcription factors have been shown to regulate *cad* gene expression. Binding of Myc-Max at *cad promoter* is essential for growth-induced *cad* expression regulation[31]. The *cad* gene is also under the regulation of induction by a nonclassical ERα/Sp1-mediated pathway and repression by HIF-1α[56], [57].

The activities of enzymes in the *de novo* pyrimidine synthesis pathway are direct targets of cell cycle regulation. Increased ATCase activity was observed following increases in the *cad* transcription rate and mRNA level in serum stimulated mouse fibroblasts[54]. CPS II, ATCase and orotate phosphoribosyl transferase (OPRTase, another enzyme in pyrimidine biosynthetic pathway) have distinct peaks of enzymatic activity during S phase and diminished activity during the G_2/M phase in rat hepatoma cells. More specifically, CPS II and ATCase activities were shown to be increased rapidly during early G_1, whereas OPRTase activity did not increase until late G_1[58].

CAD activity is also regulated throughout the cell cycle by phosphorylation and

dephosphorylation signaling cascades. Phosphorylated CAD stimulates *de novo* pyrimidine synthesis and progression through S phase of the cell cycle in mammalian cells mediated by the activation of mTORC1[59]. Ribosomal protein s6 kinase 1, a mTORC1 target, phosphorylates S1859 on CAD, promoting CAD oligomerization[59], [60]. In baby hamster kidney cells, CAD was activated by MAPK phosphorylation just before the onset of S phase, and then rephosphorylated by PKA and dephosphorylated at CAD MAPK site late in S phase[51]. Furthermore, CPS II activity, which catalyzes the initial and rate-limiting step in *de novo* pyrimidine biosynthesis, is controlled by mitogen-activated protein kinase (MAPK)- and protein kinase A (PKA)-mediated phosphorylation. Changes in CAD enzymatic activity by phosphorylation and dephosphorylation was also observed in rapidly growing cells compared with quiescent cells. Baby hamster kidney cells entering the exponential growth phase exhibited an 8-fold increase in pyrimidine biosynthesis, with a 4-fold increase in CAD threonine phosphorylation and a 40-fold increase in MAPK activity. In contrast, confluent cells exhibited a 2-fold increase in CAD phosphoserine modification, a measure of PKA phosphorylation[61]. These observations suggest that CAD is activated by MAPK phosphorylation during periods of rapid growth and down-regulated by PKA phosphorylation during quiescence[61]. In normal MCF10A breast cells, pyrimidine biosynthetic pathway was up-regulated in the exponential growth phase by MAP kinase phosphorylation of CAD Thr456, and down-regulated by P~Thr456 dephosphorylation and PKA phosphorylation of CAD[62].

Enzymes in pyrimidine biosynthetic pathway are also under metabolic control by allosteric effectors (reviewed in [63]). CAD is regulated by feedback inhibition by the end product uridine 5'-triphosphate (UTP) and allosterically activated by phophoribosyl pyrophosphate(PRPP)[63]. When phosphorylated by MAP kinase or activated by epidermal

growth factor, CAD became resistant to feedback inhibition and more sensitive to activation by PRPP[64]. Increased sensitivity to PRPP and reduced inhibition by UTP in CAD were observed as baby hamster kidney cells approached S-phase of the cell cycle. These changes were reversed when cells emerged from S-phase[51]. In vitro and in vivo studies suggest that allosteric regulation of CAD controls the rate of pyrimidine biosynthesis[51], [63]. Although these activities lie upstream of folate-dependent *de novo* thymidylate synthesis, their cell-cycle regulation provides the needed levels of nucleotide precursor substrate in the form of dUMP to support folate-dependent thymidylate biosynthesis and DNA synthesis during S phase.

2.2.2.2 Dihydrofolate reductase (DHFR)

DHFR expression

DHFR catalyzes the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF), an essential reaction in the *de novo* thymidylate biosynthesis pathway. DHFR activity is also critical for the reduction of dietary folic acid, a form of folate found in fortified foods and in nutritional supplements, to DHF. DHFR transcript levels are cell cycle regulated. In synchronized human osteosarcoma U2OS cells, *DHFR* mRNA levels were elevated at the G_1/S boundary[35]. The *DHFR* gene was expressed at the beginning of S phase in Hela cells and was grouped in the "late DNA replication" cluster by hierarchical clustering analysis [65]. *DHFR* gene expression was enriched in G_2/M phase in synchronized primary human foreskin fibroblasts [49].

DHFR transcript levels are cell cycle regulated, but there is conflicting evidence regarding the underlying mechanisms. Multiple studies suggest that increased levels of *DHFR* mRNA during S phase or after serum stimulation was controlled primarily at the level of transcription [33], [66]–[68]. In methotrexate-resistant mouse sarcoma cells, *Dhfr* mRNA levels increased following growth stimulation without a change in *Dhfr* mRNA half-life[69]. In MTX-

resistant mouse 3T6 fibroblasts synchronized by mitotic selection, the transcription rate was low in G_1 phase, increased 7-fold at the beginning of S phase, decreased almost immediately thereafter, and remained low throughout the remainder of S and into G_2[70].

The cell cycle regulation of *DHFR* transcription was achieved by increasing the rate of transcription from a single promoter region[70], mediated through transcription factors and chromatin remodeling[32]. The hamster *Dhfr* gene promoter contains consensus binding sites for two eukaryotic transcription factors: Sp1 and E2F. There are four Sp1 *cis* elements in the mouse *Dhfr* promoter that play a role in *Dhfr* transcription[33] in addition to an E2F1 element that is conserved across the human, mouse and hamster *Dhfr* promoters[32]. Sp1 binds to the human *DHFR* promoter throughout the cell cycle in early passage and senescent cells, whereas the E2F binding activity was serum-inducible and was diminished in senescent cells[34]. The E2F family of transcription factors plays an important role in the regulation of gene expression at the G_1/S phase transition of the cell cycle. E2F1 expression stimulated the *DHFR* promoter 22-fold in serum-starved NIH 3T3 cells[71]. However, expression of E2F1 in quiescent cells only minimally induced *Dhfr* transcription in REF52 cells[37]. The retinoblastoma (RB) tumor suppressor and its family members, p107 and p130, repress E2F activity. There is evidence that E2F-p130 and Sp1-pRB complexes cooperate in repression of the Chinese Hamster Ovary (CHO) *Dhfr* gene when cells withdraw from the cell cycle and enter G_0[72]. In U2OS cells, HDAC1 acts through Sp1 to repress *DHFR* promoter activity, and the E2F element modulates the activity of SP1 at the *DHFR* promoter through a cis-acting mechanism[73]. *Dhfr* expression was induced early in p107-/-; p130-/- mouse embryonic fibroblasts (MEFs) after serum stimulation[39]. The mechanism of Sp1 and E2F in regulation of *DHFR* gene expression has been extensively reviewed elsewhere [32].

However, other studies failed to observe cell cycle-dependent regulation of *DHFR* transcription. One study found *Dhfr* transcription rates to be invariant throughout the cell cycle, while *Dhfr* mRNA accumulated as mouse thymocytes progressed though the cell cycle[74]. In methotrexate-resistant mouse sarcoma cells, *Dhfr* mRNA levels increase following growth stimulation without a change in *Dhfr* gene transcription rate. In proliferating cells, most *Dhfr* transcripts were shown to be converted to mRNA, whereas in quiescent cells, the majority of *Dhfr* transcripts were rapidly degraded in the nucleus[69]. Conflicting findings related to DHFR transcriptional regulation through cell cycle may be due to limitations of the nuclear run-on technique and need further investigation[68]. Regardless of the mechanism, there is strong evidence that DHFR activity is elevated during S phase to support *de novo* thymidylate biosynthesis.

DHFR activity exhibits similar cell cycle-dependent patterns as its mRNA levels. DHFR activity was observed to be at its lowest level in G_0 and G_1 in human lymphoblasts[75], but increased significantly when quiescence temperature-sensitive cells were stimulated by serum at the permissive temperature[76]. *DHFR* steady state mRNA level, protein synthesis and activity increased 2-fold as cells progressed from G_1 to G_2/M in Chinese hamster ovary cells[77]. DHFR protein has been shown to autoregulate its own translation by binding to *DHFR* mRNA and inhibiting its translation[78]. Occupation of the DHFR folate cofactor binding site by methotrexate prevented the interaction of the DHFR protein with its cognate mRNA, thereby relieving translation autoregulation and resulting in increased DHFR levels in cells[78]. This feedback regulation ensures DHFR enzyme levels do not increase in the absence of available folate cofactor availability.

DHFR subcellular localization

HeLa cells expressing a GFP-DHFR fusion protein provided evidence for cell cycle-dependent subcellular localization of DHFR. The GFP-DHFR fusion protein localized to the nucleus during S and G_2/M phases but not in G_1 phase[16]. Nuclear localization may be small ubiquitin-like modifier (SUMO) dependent. DHFR was shown to be modified by SUMO-1 *in vitro*[14] and was also identified as a target for SUMO2 modification in proteome-wide mass spectrometry analysis[79]. Potential SUMOylation modification sites were identified through mass spectrometry analysis and through sequence alignments from various species, but these observations require experimental verification[14], [79], [80].

2.2.2.3 Serine hydroxymethyltransferase (SHMT1 and SHMT2α)

SHMT1 expression

There are two cytosolic/nuclear isozymes of SHMT encoded by distinct genes; SHMT1 and SHMT2α and these two enzymes are functionally redundant[81]. They catalyze the reversible conversion of serine and THF to glycine and 5,10-methylene THF. One study found that *SHMT1* mRNA levels were elevated in G_1/S phase of the cell cycle in synchronized U2OS cells[35], whereas in another study, SHMT1 protein levels were elevated during S phase in HeLa cells blocked with hydroxyurea without changes in *SHMT1* mRNA level[50], [82]. A significant increase in SHMT activity was observed in lymphocytes when stimulated to proliferate by phytohemagglutinin treatment[83].

SHMT subcellular localization

SHMT1 and SHMT2α and their cytosolic/nuclear localization are cell cycle regulated. The *SHMT2* gene encodes an SHMT2 protein that localizes exclusively to the mitochondria, whereas it also expresses a SHMT2α isozyme that localizes to the cytosol and translocates to the

nucleus during S-phase[16], [81]. SHMT1 and SHMT2α are modified by SUMO during S-phase of the cell cycle, and SUMOylated SHMT1 is enriched and exhibits primarily nuclear localization in HeLa cells[19]. SHMT1 was localized primarily to the cytoplasm at G_1/stationary phase of the cell cycle and to the cytoplasm, nuclear periphery, and nucleus in the S-phase and G_2/M-phase of the cell cycle in blocked MCF-7 cells[19]. Evidence suggests that SUMOylation of SHMT1 occurs at the nuclear pore and is linked to its nuclear import, as mutation of either SHMT1 Lys 38 or Lys 39 to arginine impaired *in vitro* SUMOylation by SUMO1 and nuclear localization of the mutant protein. Ran-binding protein 2, a nuclear periphery/nuclear pore protein which possesses E3 SUMO ligase activity, also plays a role in SHMT1 nuclear localization[19]. It is likely that the other enzymes in the *de novo* dTMP synthesis pathway, DHFR and TYMS, are compartmentalized in the nucleus during S-phase through similar mechanisms.

Interestingly, SHMT1 is also ubiquitinated at Lys39 and is subject to ubiquitin-dependent degradation. Ub-SHMT1 was at the lowest level at S phase of the cell cycle. Lys-63 polyubiquitination of SHMT1 was enriched in the nuclear fraction in S and G_2/M phase, but absent in G_1 in both nuclear and cytosolic fractions. Mutation of the ubiquitination site increases SHMT1 stability[82]. In HeLa cells, Ubc13-mediated ubiquitination of SHMT1 is required for nuclear export and nuclear stability, whereas Ubc9-mediated modification with Sumo2/3 leads to SHMT1 degradation within the nucleus[82].

2.2.2.4 Thymidylate synthase (TYMS)

TYMS expression

TYMS catalyzes the reductive methylation of deoxyuridylate (dUMP) to deoxythymidylate (dTMP) using 5,10-methyleneTHF as a cofactor. DTMP is further

phosphorylated to dTTP for DNA synthesis. TYMS activity and protein levels are subject to cell cycle variation. Very low TYMS enzyme activity in G_0 and G_1 phase has been observed in human lymphoblasts with significant higher levels in S and G_2/M phases[75], [84]. Elevations in TYMS protein levels and activity are associated with the onset of S-phase when cells were growth arrested and then allowed to reenter the proliferating state (reviewed in [85]). In human diploid fibroblasts synchronized by serum starvation, TYMS protein and mRNA level increased through post-transcriptional regulation[86]. In MCF-7 cells synchronized with lovastatin, TYMS protein level is tightly regulated in the cell cycle with the peak of TYMS level coinciding with that of S phase, with no change in the level of *TYMS* mRNA level[87]. Similarly, other studies have shown that *TYMS* cell cycle regulation occurs through intron splicing signals, with very little change in rate of transcription during G_1-S phase transition in serum stimulated mouse 3T6 fibroblasts[88], [89].

Other evidence suggests that the TYMS gene is regulated transcriptionally and posttranscriptionally (reviewed in [85], [90]). HCT116 cells synchronized by serum starvation exhibit increased *TYMS* mRNA levels during G_1, followed by an increase in TYMS protein levels which did not decrease following S-phase completion[91]. *Tyms* mRNA levels increased 20- to 40-fold as mouse fibroblasts progressed from a resting to late S phase, as a result of increased transcription rates and other posttranscriptional regulation[92]. In synchronized HeLa cells, analysis of transcript levels and hierarchical clustering of the expression patterns classified *TYMS* in the "late DNA replication" cluster, together with other genes in nucleotide metabolism, including *DHFR*, ribonucleotide reductase (*RRM1* and *RRM2*), and additional DNA replication, repair and recombination genes[65]. Interestingly, *TYMS* expression was shown to be down-regulated during metastatic HO-1 human melanoma cells differentiation[93]. In cells that have

not been subjected to cell cycle arrest followed by release, analyses of TYMS protein and mRNA levels lead to conflicting results. Several studies found no variation in TYMS protein or *TYMS* mRNA with S phase (reviewed in [85]). On the contrary, *TYMS* mRNA has been observed to accumulate as thymocytes progress through the cell cycle[74].

The *TYMS* gene is regulated by several transcription factors. Transcription factor Late SV40 Factor (LSF) induces TYMS gene at G_1/S transition in growth-stimulated cells through binding to sites within the TYMS promoter and intronic regions[43]. The *TYMS* gene is also targeted by the transcription factor Forkhead Box M1 (FOXM1)[42]. Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) in U2OS cells identified *TYMS* gene as a target for FOXM1, which strongly activates promoters of G_2/M phase genes and weakly activates those induced in S phase[35]. *Tyms* gene expression has also been shown to be regulated by E2F in MEFs[36]. *Tyms* expression was derepressed in G_0 and G_1 in MEFs lacking both p107 and p130, two E2F activity controlling proteins[39]. DNA microarray analyses demonstrated that *Tyms* gene expression is up-regulated by E2F-1 or E2F-3, along with the functionally-related enzymes Uracil-*DNA* glycosylase (UNG), proliferating cell nuclear antigen (PCNA) and dUTP nucleotidohydrolase (dUTPase)[41]. *Tyms* gene transcription was induced by E2F1 in quiescent REF52 cells infected with E2F1 cDNA containing recombinant adenovirus[37]. E2F motifs in rat *Tyms* promoter region were found to be necessary for increased *Tyms* expression during the G_1 to S transition in serum stimulated cells[40]. However, adenovirus-mediated over-expression of E2F1 and cyclin E in three human cell lines had no effect on *TYMS* expression[91]. As seen for DHFR, TYMS protein autoregulates its synthesis by binding to its own mRNA with high affinity which represses translation (reviewed in [90]).

TYMS subcellular localization

HeLa cells expressing a GFP-TYMS fusion protein showed nuclear TYMS localization during S and G_2/M phases but not in G_1[16]. TYMS is modified by SUMO1 *in vitro*[14], and was also identified as a target for SUMO2 modification in system-wide mass spectrometry analysis [79], [80]. Several putative SUMOylation modification sites were identified through mass spectrometry analysis[14], [79], [80]. Collectively, there is strong evidence for cell cycle regulation of the enzymes that constitute the *de novo* dTMP synthesis pathway, with elevated protein levels and nuclear localization of the pathway during S phase.

2.2.3 dTMP salvage pathway

2.2.3.1 TK expression

Thymidine kinase 1 (TK1) catalyzes the conversion of thymidine to dTMP in pyrimidine salvage pathway synthesis, and its enzyme activity varies throughout the cell cycle and reflects the levels of TK1 protein[94]. TK1 activity was shown to vary through the cell cycle in synchronized mouse fibroblasts and synchronized human KB cells[95], [96]. Very little or no activity of TK1 was observed in G_0 and G_1 in human lymphoblasts[75], [84], [97], with a higher level of TK1 activity observed in S and G_2/M[84], with the maximum activity observed after the peak of DNA synthesis[97]. In contrast to increased TK1 activity in S and G_2/M phases, enzymes involved in thymidine and thymidylate catabolism, including dihydrothymine dehydrogenase, thymidine phosphorylase and dTMP phosphatase, have constant levels of activity throughout the cell cycle in synchronized Human B lymphocytes[97].

Both *TK1* mRNA and protein levels were shown to be upregulated in rapidly dividing cells. *Tk1* mRNA levels increased more than 20-fold in S-phase compared to G_0-phase of the cell cycle in BALB/c 3T3 cells[98]. Proliferating cells exhibited higher levels of TK1 mRNA and

protein compared to slowly growing, quiescent, or terminally differentiated cells. The levels of TK1 mRNA and enzyme activity increased significantly when quiescent cells were stimulated by serum or adenovirus infection [76], [99], [100].

It is not clear how TK1 protein levels are regulated throughout the cell cycle. *TK1* mRNA levels were increased in senescent human diploid fibroblasts WI38 when serum stimulated, without increased incorporation of ^{3}H thymidine[101] into DNA, indicating that TK1 activity was not increased with elevated transcript levels. However, increased TK1 activity depended on the availability of *TK1* mRNA in BALB/c 3T3 cells[98]. Pulse labeling experiments showed that TK1 protein synthesis was elevated 10-fold in S phase compared to G_1 phase, suggesting that the efficiency of translation of *TK1* mRNA increased[94]. TK1 protein synthesis ends with a post-transcriptional block at about the time human KB cells begin mitosis[96]. Conflicting data exist with regard to TK1 protein stability throughout the cell cycle. TK1 protein stability changes during the cell cycle in some circumstances, as the stability of TK1 protein was shown to decrease upon cell division in HeLa cells[94]. However, in human KB cells, the stability of the enzyme did not change significantly throughout the cell cycle, suggesting that enzyme activity is determined by rate of enzyme synthesis rather than enzyme degradation[96]. Interestingly, deletion of the carboxyl-terminal 40 amino acids or fusion of beta-galactosidase to the carboxylterminus of human TK1 completely abolishes cell cycle regulation and stabilizes the protein throughout the cell cycle. The truncated or fusion proteins have similar enzymatic activity when compared to wild-type TK1 protein[102].

TK1 mRNA levels are controlled by both rates of gene transcription and mRNA degradation, and subject to cell cycle regulation. In growth stimulated CV1 African Green Monkey kidney cells, a large increase in *TK1* mRNA levels and a relatively small increase in

transcription rates were observed, suggesting that *TK1* gene expression was controlled at both a transcriptional and post-transcriptional level[103]. In BALB/c 3T3 cells, the rate of run-on *Tk1* transcription increased 2-4 fold during S phase compared to G_0 phase [98]. The half-life of *Tk1* mRNA was longer in S and M phases then in quiescence[98].

The mechanism of *TK1* transcriptional cell cycle regulation remains elusive. In CV1 African Green Monkey kidney cells, the *TK1* cDNA alone is sufficient to encode cell cycle-regulated expression[103]. However, Travali, *et. al.* showed that the *Tk1* promoter has also an important role in cell cycle regulation of TK1 mRNA levels[104]. *Tk1* gene transcription was regulated by E2F in MEFs[36], however, *Tk1* expression was not affected in MEFs lacking E2F regulating proteins Rb, p107, p130, or both 107 and p130[39]. Similar to the *TYMS* gene, the *TK1* gene is also bound by transcription factor FOXM1 in U2OS and HeLa cells[35].

2.2.3.2 TK subcellular localization

Like the enzymes that comprise the *de novo* dTMP synthesis pathway, TK1 translocates into the nucleus during S-phase of the cell cycle, providing dTMP at site of DNA replication. In Chinese hamster cells during S-phase, TK1 co-fractionated with DNA polymerase, NDP kinase, RNR, and the *de novo* dTMP synthesis pathway enzymes, including DHFR and TYMS[105]. The mechanism for TK1 nuclear localization during S phase remains unknown.

2.2.4 Purine synthesis

Human cells produce purines in the cytosol through both salvage and *de novo* biosynthesis pathways. The purine nucleotide salvage pathway includes a one-step conversion of hypoxanthine to inosine monophosphate (IMP)[106]. IMP is further converted into adenosine and guanosine nucleotides. The more energy consuming *de novo* pathway consists of ten chemical reactions that convert phosphoribosyl pyrophosphate (PRPP) into IMP, utilizing the 1C

from two 10-formylTHF during the process (the formyl group of 10-formylTHF is incorporated as the #2 and #8 carbons of the purine ring). In order to conserve energy, purine nucleotides are synthesized preferentially through salvage pathway when hypoxanthine is available in most cells[107]. However, *de novo* purine synthesis, rather than purine salvage synthesis or pyrimidine synthesis, limits CHO cell growth rate[107].

2.2.4.1 Expression of enzymes in purine synthesis pathway

Purine nucleotide synthesis rates are growth-dependent in rat livers[52]. *De novo* purine biosynthesis was shown to be upregulated in serum-stimulated 3T6 fibroblasts, with increased levels of PRPP and formylglycinamide ribonucleotide (biosynthetic intermediates in the *de novo* purine synthesis pathway)[108]. *De novo* and salvage purine synthesis pathway increased 5- and 3.3-fold respectively, as cells progressed from mid-G_1 phase to early S phase[109]. In CHO K1 cells synchronized by serum deprivation, amidophosphoribosyltransferase (ATase) and hypoxanthine-guanine phosphoribosyltransferase (HPRT), which are key regulatory enzymes of the *de novo* and salvage pathways of purine synthesis respectively, showed distinct expression changes throughout the cell cycle. ATase activity increased from late G_1 phase to the S phase, whereas HPRT activity was nearly constant during throughout the cell cycle[107]. ATase and aminoimidazole ribonucleotide carboxylase (an enzyme in *de novo* purine synthesis pathway) mRNA levels increased approximately 5-6 fold in G_1/S phase of the cell cycle compare to G_0 in synchronized rat 3Y1 fibroblasts[110]. MTORC1 activity, which is important for S and G_2 progression[59] and M phase entry (review in [111]), was shown to increase metabolic flux through the *de novo* purine synthesis pathway and adjusted purine nucleotide pools available for nucleic acid synthesis in various mouse and human cells[112]. Similarly, purine synthesis was stimulated by mTORC1 through activating transcription factor 4 (ATF4) in both normal and

cancer cells[112].

2.2.4.2 Subcellular localization of enzymes in the *de novo* purine synthesis pathway

The six proteins that comprise the *de novo* purine synthesis pathway were shown to cluster in the cytoplasm and form a multienzyme complex known as the "purinosome" in HeLa cells under purine depletion[113]. Purinosome formation exhibited cell cycle dependence in HeLa cells cultured in purine-depleted conditions, with the highest number of cells with purinosomes observed during G_1 phase[114]. A similar observation was made in Lesch-Nyhan disease (LND) fibroblast cells that are deficient in hypoxanthine/guanine phosphoribosyl transferase (HGPRT), an enzyme in the purine salvage pathway. However, elevated levels of purinosomes in S and G_2/M phases were also observed, probably due to the fact that LND cells rely primarily on the *de novo* pathway to synthesize purine nucleotides[114].

2.2.5 Methenyltetrahydrofolate synthetase (MTHFS)

2.2.5.1 MTHFS expression

MTHFS catalyzes the irreversible and ATP-dependent conversion of 5-formylTHF (a storage form of THF cofactors) to 5,10-methenylTHF. 5-formylTHF is an inhibitor of SHMT[115] and AICAR transformylase[116] (an enzyme in the *de novo* purine synthesis pathway). MTHFS protein activity is inhibited by 10-formylTHF, which exists in chemical equilibrium with 5,10-methenylTHF[26].

MTHFS gene expression was shown to be significantly higher in G_1/S phase of the cell cycle in synchronized HacaT human keratinocytes[117]. Although the *MTHFS* gene has been observed to bind the transcription factor FOXM1[35], the mechanism of cell cycle regulated gene expression of *MTHFS* remains unknown.

2.2.5.2 MTHFS subcellular localization

MTHFS cellular localization has been shown to be cell cycle dependent. A GFP-MTHFS fusion protein colocalized with purinosomes in purine-depleted media in Hela cells, suggesting MTHFS is a component of the purinosome[118]. It has been proposed that MTHFS delivers 10-formylTHF cofactors to the purinosome, as MTHFS levels influence rates of *de novo* purine biosynthesis[118]. SUMOylation was shown to be necessary for localization of GFP-MTHFS to the purinosome. Recombinant MTHFS is modified by SUMO-1 both *in vitro* and in Hela cells, and mutation of the MTHFS SUMO1 modification sites K140 and K190 ablated the localization of a GFP-MTHFS fusion protein to the purinosome in HeLa cells cultured in purine-deficient conditions[118].

2.2.6 Homocysteine remethylation pathway

2.2.6.1 Methylenetetrahydrofolate reductase (MTHFR)

There is limited evidence for cell cycle regulation of MTHFR, an enzyme that catalyzes the conversion of 5,10-methyleneTHF to 5-methylTHF (a required cofactor for homocysteine remethylation to methionine). MTHFR is phosphorylated at T34 by CDK1/cyclin B1[119] which inhibits its catalytic activity and this phosphorylation peaked during mitosis in Hela cells[119].

2.2.6.2 Methionine synthase (MTR)

MTR is a folate- and vitamin B12-dependent enzyme that catalyzes the conversion of homocysteine into methionine, utilizing a methyl group from 5-methylTHF which is transferred to homocysteine through a methyl-cobalamin intermediate. Resting human peripheral blood lymphocytes do not efficiently take up vitamin B12 in the form of cobalamin (CbI); increased CbI uptake correlates with active cell division and DNA synthesis through increased receptor activity[120]. The *MTR* mRNA level is elevated in S phase of the cell cycle in synchronized

U2OS cells[35]. The lowest level of MTR activity was observed during G_0 and G_1 in human lymphoblasts[75]. Another study found MTR activity correlated with cell division and DNA synthesis in human lymphocytes[120].

2.2.7 Other FOCM related enzymes and transporters

2.2.7.1 Folate transporters

Folates are transported into cells through membrane-bound folate receptor alpha (FR-α), the reduced folate carrier (RFC), and the proton-coupled folate transporter (PCFT). Folate uptake is closely correlated with cell growth and S-phase of the cell cycle. FR-α expression has been shown to be cell cycle-dependent with the highest expression levels in S-phase[121]. FR-α function has been observed to decrease as cellular growth slowed in JAR, Caco-2 and MA-104 cell lines[122]. In primary human trophoblast cells, inhibition of mTORC1 or mTORC2 markedly decreased basal folate uptake and decreased the plasma membrane expression of FR-α and RFC transporter isoforms without changes in global protein expression levels[123].

2.2.7.2 Ribonucleotide reductase

Ribonucleotide reductase (RNR) converts ribonucleotides to deoxyribonucleotides, an activity that is essential for *de novo* dTMP synthesis by generating the precursor substrate dUMP. RNR consists of two subunits, RRM1 and RRM2. RRM1 protein levels remained constant throughout the cell cycle in bovine kidney MDBK cells[124], whereas RRM2 protein levels increased as cells passed from G_1 to S[125]. During G_2, RRM2 was degraded, which maintained balanced dTNP pools and genome stability[126]. RRM2 interacts with cycline F and its protein level is tightly controlled by two different ubiquitin ligases in the G_1 and G_2 phases of the cell cycle[127]. RNR activity is highest during S phase in eukaryotes, when the requirement for dNTPs is the highest (reviewed in[128]).

RNR gene expression has been shown to be cell cycle regulated[36]. Both *Rrm1* and *Rrm2* transcripts were very low or undetectable in G_0 and G_1, but increased as cells progressed into S phase, and then declined in G_2/M phase in hydroxyurea-resistant mouse mammary tumor TA3 cells[129]. The regulation of *RNR* transcription is not well understood. *Rrm1* and *Rrm2* transcription rates are invariant throughout the cell cycle in mouse thymocytes, whereas *Rrm2* mRNA levels but not *Rrm1* exhibited cyclic regulation in mouse thymocytes[74]. Another study indicated that *Rrm1* gene expression is mainly regulated at the transcriptional level during the cell cycle[130]. *Rrm1* and *Rrm2* transcription was induced by E2F1 expression in quiescent cells[37] and in synchronized MEFs[36]. *Rrm2* expression was derepressed in G_0 and G1 in MEFs lacking both p107 and p130, which control E2F activity[39]. S phase-specific transcription of the mouse *Rrm2* gene is dependent on one proximal promoter repressive element that binds E2F4, and mutation of the E2F binding site leads to premature promoter activation in G_1[38].

2.3 FOCM in the nucleus

2.3.1 Nuclear co-localization of enzymes in the *de novo* dTMP synthesis pathway

During S and G_2/M phases of the cell cycle or following DNA damage, the enzymes of the dTMP synthesis pathway, including SHMT1, SHMT2α, TYMS, DHFR, and MTHFD1 are SUMOylated and translocate into the nucleus for DNA replication or repair[16]. This pathway forms a multienzyme complex that is associated with the nuclear lamina and other enzymes of the DNA replication machinery [14], [16], [19], [81]. SHMT1 and SHMT2α independently serve as scaffold proteins that are essential for dTMP synthesis complex formation at sites of DNA replication[16]. Interestingly, in BHK cells transfected with a fluorescent CAD fusion protein, a significant fraction of CAD translocates in the nucleus when cells enter S-phase of the cell

cycle[51]. The role of CAD nuclear localization remains unknown[63]. Similarly, RNR has been observed to localize at DNA damage sites to allow production of dNTPs at sites of DNA repair[131], indicating that the nucleus may synthesize thymidylate *de novo* completely within the nuclear compartment in some cells.

The co-localization of the enzymes in the *de novo* dTMP synthesis pathway suggests that the spatial organization of the pathway enzymes play an important role in meeting cellular dTMP demand during DNA replication in S phase. The formation of the dTMP synthesis complex at sites of DNA replication potentially enables much more efficient catalysis through substrate channeling, which offers kinetic advantages over free diffusion of folate cofactors within the bulk solvent[132].

2.3.2 Uracil-DNA glycosylases (UNG)

Folate deficiency and impaired dTMP synthesis leads to uracil misincorporation into DNA and chromosome instability and breakage. Uracil-DNA glycosylases (UNG) excises uracil residues from DNA. Specifically, UNG cleaves the N-glycosylic bond and initiates the base-excision repair pathway. *UNG* gene expression is enriched in the G_1/S and G_2 phases of the cell cycle in synchronized primary human foreskin fibroblasts[49].

2.4 FOCM in the mitochondria

1C units are generated in mitochondria through catabolism of serine, glycine, histidine, betaine, dimethylglycine and sarcosine, and are released into the cytosol in the form of formate[133]. Mammalian mitochondria also contain a *de novo* dTMP synthesis pathway, with SHMT2 playing an essential role in limiting uracil misincorporation into mitochondrial DNA [134]. Confocal microscopy demonstrated that fluorescent fusion proteins of DHFR2 and TYMS localized to the mitochondria in Hela cells. Endogenous DHFR2, SHMT2 and TYMS have also

been shown to localize to the mitochondrial matrix and the inner membrane of the mitochondria in fractionated HepG2 mitochondria [134].

2.4.1 Methylenetetrahydrofolate dehydrogenase 2

Methylenetetrahydrofolate dehydrogenase (MTHFD2) is a nuclear-encoded mitochondrial bifunctional enzyme that has methyleneTHF dehydrogenase and methenylTHF cyclohydrolase activities. The multifunctional enzyme reversibly converts 5,10-methylene THF, 5,10-methenylTHF and 10-formyl THF. The formyl group of 10-formyl THF is liberated as formate, thereby generating THF via the enzyme MTHFD1L. MTHFD2 is also found in the nucleus, though its function within the nucleus has not been elucidated[135]. MTHFD2 has recently been shown to co-localize with newly synthesized DNA in the nuclei of U251, HeLa, and HCT116 cells[135]. *MTHFD2* gene expression is upregulated in S phase of the cell cycle in U2OS cells[35]. *MTHFD2* expression was stimulated by mTORC1 through activating transcription factor 4 (ATF4) in both normal and cancer cells[112]. *MTHFD2* over-expression also increased HCT116 proliferation rates, even when MTHFD2 lacking functional dehydrogenase activity was expressed[135]. This observation, taken together with the assumption that MTHFD2 dehydrogenase activity would be expected to decrease synthesis of 5,10-methylene-THF (or oppose the activity of MTHFD1 dehydrogenase activity), suggests that the function of MTHFD2 within the nucleus is distinct from its catalytic activity.

2.4.2 Serine hydroxymethyltransferase 2

The *SHMT2* gene encodes two transcripts, a mitochondrial SHMT2 isozyme and a cytoplasmic/nuclear SHMT2α isozyme. *SHMT2* expression is significantly enriched in G_2/M phase of the cell cycle in synchronized HacaT human keratinocytes[117], whereas *SHMT2* mRNA levels in MCF-7 cells did not vary during the cell cycle[136]. *SHMT2* expression is

predicted to be regulated by transcription factor E2F and Sp1[137].

2.4.3 Compartmentalization of folate derivatives

Folate derivatives are compartmentalized within the cell. Mitochondria contain roughly 40% of total cellular folate[138], [139], while the nuclear compartment contains about 10% cellular folate[138], leaving about 50% of cellular folate within the cytosol. Neither total cellular folate levels nor nuclear folate levels varied as a function of cell cycle in MCF-7 cells[50]. Interestingly, nuclear folate levels resisted depletion during S-phase in MCF-7 cells cultured in folate-depleted media[50]. In mice consuming folic acid-deficient diets for six weeks there was a 50% decrease in liver whole-cell folate concentration compared to animals consuming control diets containing folic acid, but nuclear folate levels resisted depletion[50]. These data suggest that nuclear folate levels are maintained to protect *de novo* dTMP synthesis during S-phase and in response to folate deficiency. However, protecting nuclear folate pools and dTMP synthesis occurs at the expense of cytosolic folate pools and homocysteine remethylation[15].

2.5 Conclusions

Nutrition and genetic epidemiological studies and /or randomized controlled trials implicate impaired folate status and FOCM in several pathologies, including megaloblastic anemia, neural tube defects (NTDs), neurodegenerative disease, and various types of cancer[3]–[9]. However, the causal pathways and mechanisms underlying these pathologies remain unclear due to the interconnectedness of these pathways[132]. Key to understanding the etiology of folate-associated pathologies will be elucidating the regulation of the pathways within FOCM, and how decisions are made to partition one-carbon units and other intermediates among the pathways, as well as the partitioning of folate cofactors which are a limiting resource for all pathways[3]. Cell cycle regulation of FOCM, which is necessary considering its role in DNA

synthesis, provides a mechanism to ensure cellular needs are met by assigning priority to individual pathways within the network based on temporal needs, especially deoxyribonucleotide biosynthesis. Evidence for cell cycle regulation of homocysteine remethylation and mitochondrial one-carbon metabolism is much more limited.

There is clear evidence that folate-dependent and folate independent purine and thymine deoxyribonucleotide synthesis are regulated by cell cycle. This regulation involves both an increase in enzyme activity through increased expression and enzyme activity, as well as formation of multienzyme complexes. However, the strategies for meeting cellular demands differ among the pathways. Thymidylate synthesis occurs in the nucleus at sites of DNA synthesis, with both increased expression of the pathway and nuclear import occurring at the S and G_2/M phases. Nuclear dTMP synthesis separates this pathway from FOCM in the cytosol, eliminating competition for folate cofactors with *de novo* purine biosynthesis and homocysteine remethylation, assuming that folate nuclear and cytosolic folate cofactors are not shared[50].

De novo purine biosynthesis and homocysteine remethylation both occur in the cytoplasm and are poised to compete for cofactor availability. Both of these pathways rely on MTHFD1 and cellular formate to provide folate-activated 1C units. *De novo* purine biosynthesis exhibits cell cycle regulation at the level of multienzyme complex formation, but less information is available concerning cell cycle regulation of enzyme levels. If purinosome formation, which forms at G1 and extends into S phase, is critical to meet cellular purine nucleotide needs, this may suggest that homocysteine remethylation may be most vulnerable to decreased activity during times of purinosome formation.

Understanding how metabolic decisions are made as a function of cell cycle, both in terms of enzyme expression and localization, will shed light on the regulation of FOCM network

and the role of disrupted FOCM in disease pathogenesis.

Acknowledgement: This work was support by funding from PHS R37DK58144 and 3R01 HD059120-08S1. The authors have no conflicts to declare.

Chapter 3 SUMO MODIFICATION AND COMPLEX FORMATION OF *DE NOVO* THYMIDYLATE SYNTHESIS PATHWAY

3.1 Introduction

During S and G_2/M phases of the cell cycle or following DNA damage, the enzymes of the *de novo* dTMP synthesis pathway, including serine hydroxymethyltransferase (SHMT1 and SHMT2α), thymidylate synthase (TYMS), dihydrofolate reductase (DHFR), and methylene-tetrahydrofolate dehydrogenase 1 (MTHFD1) are SUMOylated and translocate into the nucleus [16], [22], [81]. The pathway forms a multienzyme complex at sites of DNA replication with the DNA replication and epigenetic machinery[16]. Specifically, TYMS utilizes 5,10-methyleneTHF (CH2F) as the methyl donor to catalyze the conversion of deoxyuridylate (dUMP) to deoxythymidylate (dTMP) and dihydrofolate (DHF). DHFR catalyzes the reduction of dihydrofolate back to tetrahydrofolate (THF). Methylene-tetrahydrofolate can be generated through the enzyme SHMT1 and SHMT2α using serine as one-carbon (1C) source, or through the enzyme MTHFD1 using formate as 1C source.

Previous study indicates that SHMT serves as a scaffold protein that anchors *de novo* dTMP synthesis pathway as a multienzyme complex to nuclear lamina; however, no direct interaction was found between SHMT and TYMS or DHFR using the yeast two hybrid assay[16], [19]. Formation of a multienzyme complex is essential for the pathway to function, as nuclei purified from mouse liver disrupted by sonication lack *de novo* dTMP synthesis activity[81]. The structure of the nuclear *de novo* thymidylate synthesis multienzyme complex and the nature of the protein-protein interactions have not been characterized.

All enzymes in *de novo* dTMP synthesis pathway, including SHMT, TYMS, DHFR and MTHFD1, as well as Lamin A, contain SUMOylation modification sites (Table 3.1) [14], [19]–

[22]. SUMOylation is a reversible post-translational modification in which the polypeptide small ubiquitin-like modifier (SUMO) is covalently attached to its protein substrate. SUMO is attached to most substrate proteins at the lysine in a ΨKX(D/E) sequence, where Ψ is a large hydrophobic residue[140], although several proteins have been shown to be modified at other sites. Similar to ubiquitination, SUMOylation involves in three sequential enzymatic reactions, catalyzed by a SUMO-activating enzyme (E1), one and only SUMO-conjugating enzyme (E2) called Ubc9, and one of several SUMO-protein ligases (E3). The molecular consequences of SUMOylation are substrate specific, including subcellular localization, antagonizing other modifications such as ubiquitination, protein activation, interactions with other macromolecules and half-life [141]–[143].

SUMOylation of SHMT1 occurred at the nuclear pore and was linked to its nuclear import[19]. Mutation of either SHMT1 K38 or K39 to arginine residues impaired SUMOylation *in vitro*[19]. MTHFD1 was SUMOylated in HeLa cells and MTHFD1-K223R-GFP mutant proteins exhibited less SUMOylation *in vitro* compared with MTHFD1-GFP fusion protein[22]. TYMS and DHFR were substrates for UBC9-catalyzed SUMOylation *in vitro* by SUMO1[14]. Mass spectrometry analysis of SUMO2-enriched peptides also identified TYMS and DHFR as substrates[79], [80], [144]. Despite the evidences of several lysine residues being important for SUMOylation modifications or possible target, the exact SUMOylation sites on enzymes of *de novo* thymidylate biosynthesis pathway remain uncharacterized.

Table 3.1. Summary of putative SUMO modification sites and SIMs on de novo dTMP synthesis pathway enzymes. Putative SUMO modification sites are derived from Eukaryotic Linear Motif (ELM, http://elm.eu.org/) resource and previous publications. Putative SIMs were identified based on structural availability.

Enzyme	**Site**	**Positions**	**Source**
SHMT1	SIM	281-GKEILYNL-288	
		346-GYKIVSGG-353	
		237-AAGVVPSP-244	
		410-SRGLLEKD-417	
	SUMOylation	K38,K39	[19]
MTHFD1 (C/D domain)	SUMOylation	K223	[22]
		K223, K245, K246	ELM
TYMS	SUMOylation	K287, K292, K308	ELM
		K82,K93,K204,K209,K225,K253, K258,K274,K287,K292,K308	[79], [80], [144]
		K292	[14]
DHFR	SUMOylation	K81, K174	ELM
		K179	[14]
		K29,K47,K81,K99,K106,K122, K158,K174,K179,	[79], [144]

SUMO-interacting motifs (SIMs) mediate non-covalent interactions between SUMO and SIM-containing proteins. SIMs are generally characterized by a short stretch of hydrophobic amino acids, (V/I)X(V/I) (V/I), which are often flanked by acidic or serine residues. The hydrophobic core of SIMs adopts a parallel or anti-parallel β-strand conformation that interacts with a groove of SUMO formed by a beta sheet and part of the alpha helix[145]. One potential role of SIMs is to facilitate protein complex assembly and function. Evidence suggests SIMs are important in promyelocytic leukemia (PML) nuclear body formation by mediating recruitment and SUMO modification of PML body resident proteins[141].

It is possible that enzymes in *de novo* thymidylate biosynthesis pathway contain multiple SIMs that allow non-covalently interactions with SUMO modifications that facilitate complex formation (Table 3.1). We hypothesized that a combination of SUMO modifications and SIMs of

de novo dTMP synthesis pathway enzymes enables protein-protein interactions and establish a scaffold for the complex. To gain structural insight of *de novo* dTMP synthesis complex and characterize the interactions among the enzymes, SUMO1 modified proteins of *de novo* dTMP synthesis pathway were purified from bacterial lines in which SUMO1 as well as SUMO-activating and -conjugating enzymes are expressed. SUMO1 modification sites on enzymes of *de novo* dTMP synthesis pathway were identified by mass spectrometry analysis. Protein-protein interactions among enzymes were identified by cross-linking/mass spectrometry (XL/MS) using disuccinimido sulfoxide (DSSO) and disuccinimido dibutyric urea (DSBU). DSSO and DSBU are MS-cleavable crosslinkers that contain amine-reactive *N*-hydroxysuccinimide (NHS) esters that react with primary amine groups to form stable amide bonds. The ability to cleave crosslinked peptide during MS/MS considerably facilitates peptide sequencing using traditional database search engines.

Table 3.2. The SHMT1, TYMS, and MTHFD1 (C/D domain) cDNA were amplified by PCR using the primers listed.

Gene		**Restriction Enzyme**	**Primer Sequence**
SHMT1	Forward	ECOR1	5’ gtagatgaattctaatgacgatgccagtcaacg 3’
	Reverse	Hind3	5’ gcagctaagcttttagaagtcaggcaggcc 3’
MTHFD1 (C/D domain)	Forward	Sma1	5’ atatatcccgggaatggcgccagcagaaat 3’
	Reverse	Hind3	5’ gcagctaagcttctaaatcatccactttcctg 3’
TYMS	Forward	Sma1	5’ atatatcccgggaatgcctgtggccggct 3’
	Reverse	Hind3	5’ gcggctaagcttctaaacagccatttccat 3’

3.2 Methods

Plasmid construction

pE1-E2-His-SU1 vector was generously provided by Dr. Guillaume Bossis, Institute of Molecular Genetics of Montpellier, France. pCMV6-AC-TYMS-GFP, phiYEP-SHMT1 and

pET28-MTHFD1C/D were previously described [15], [16]. The SHMT1, TYMS, and MTHFD1 (C/D domain) cDNA were amplified by PCR and cloned into pCAL-n vector (Agilent) (Table 3.2). All DNA constructs were verified by DNA sequencing.

Protein expression and purification

Purification of DHFR was previously described[14]. SUMO1 modified proteins of *de novo* thymidylate synthesis pathway were purified using previously published protocols[146] with modifications. Briefly, BL21(DE3) competent bacteria (Invitrogen) were transformed with both pE1-D2-His-SU1 and pCAL-target plasmids. The transformed bacteria were plated on LB-agar plates with chloramphenicol and ampicillin. One colony was then picked up and grown overnight in 5mL of Luria-Bertani Broth (Research Products international) supplemented with antibiotics. The preculture was then diluted in 1L of terrific broth (Research Products international) with antibiotics and grown with agitation at 37°C until the OD reached 0.6-0.8. Pyridoxine was added to the culture for expression of SHMT1. Protein expression was induced by adding 500uM IPTG for 8hr at 25°C. The bacteria were harvested by centrifugation at 4°C and pellet was resuspended in 30 mL of B-PER (ThermoFisher), 1:100 diluted Protease Inhibitor Mixture (Sigma), lysozyme (1mg/mL) and 10mM β-mercaptoethanol.

Calmodulin affinity resin (Agilent) was equilibrated 3 times with binding buffer (50 mM Tris, pH 8, 150 mM NaCl, 10 mM β-mercaptoethanol, 1 mM magnesium acetate, 1 mM imidazole and 2 mM $CaCl_2$). The resuspended pellets were sonicated before centrifugation. The supernatant was collected and incubated with the equilibrated resin for 30 min at 4°C using overhead rotation. The resin was washed 5 times with binding buffer. Proteins of interest were eluted with 50 mM HEPES, pH 7.2, 150 mM NaCl, and 2 mM EGTA. Pyridoxal phosphate (PLP) was added to buffers for purification of SHMT1 as it is a PLP-dependent enzyme.

Chemical cross-linking with DSSO and DSBU

SUMO and SIMs interact with medium to low affinity with dissociation constants in the range 1-100μM[147]. Based on this information, 3nmol SHMT1, 1.5nmol TYMS, 1.5nmol MTHFD1(C/D domain), and 1.5nmol DHFR were incubated in 250μl cross-linking buffer containing 50mM HEPES, pH 7.2, 150 mM NaCl. Cross-linking reactions were also performed with 5nmol TYMS and 5nmol DHFR incubated in 250μl cross-linking buffer. Before the cross-linking reaction, the optimal concentration of DSSO and DSBU cross-linking reagent for the *in vitro* assembled *de novo* thymidylate protein complex were determined (Figure 3.1 & 3.2). Based on titration results, DSSO (50 mM stock solution in DMSO, Thermoscientific) was added to the protein solution to a final concentration of 175 μM. Similarly, DSBU (50 mM stock solution in DMSO, Thermoscientific) was added to the protein solution to a final concentration of 25 μM or 500 μM. The cross-linking reaction was incubated at 25°C for 30min. The cross-linking reaction was stopped by adding Tris-HCl pH 8.0 to a final concentration of 20 mM and incubated at 25°C for 10 min. DSSO cross-linked protein samples were subjected to in-solution digestion, and DSBU cross-linked protein samples were subjected to both in-solution digestion and SDS-PAGE followed by in-gel digestion prior to MS analysis. The cross-linked proteins were digested with trypsin or double-digested with trypsin and chymotrypsin.

Fractionation of cross-linked peptides by strong cation exchange (SCX)

SCX fractionation was performed on an UltiMate 3000 HPLC (Dionex, Sunnyvale, CA USA) using a PolySULFOETHYL A column 5um, 200Å, 2.1 x 200mm (PolyLC, Columbia, MD USA) with 10mM potassium phosphate monobasic in 25% ACN pH = 3.0 as buffer A and 10mM potassium phosphate monobasic /0.5M potassium chloride in 25% ACN pH = 3.0 as buffer B. All eluents were vacuum filtered through a 0.22um Durapore GV membrane (EMD

Millipore Corporation, Billerica, MA USA) and stored at 4° C until use. Sample (1 mg of digests) was reconstituted in 25% acetonitrile (ACN)/0.1% formic acid (FA) (v/v) (Optima LC/MS grade – Fisher Chemical, Fair Lawn, NJ USA) and filtered through a Costar Spin-X 0.22um cellulose acetate centrifuge tube filter (Corning Incorporated, Corning, NY USA), following the manufacturer's recommended protocol, prior to injection. The LC was performed using a gradient from 5-60% of buffer B in 40 min and 60-100% B in additional 10 min at a flow rate 200 µL/min. Sixty fractions were collected using 96 well plates (400ul capacity/well) at 1 min intervals monitored by UV absorbance at 220nm and 280nm, pooled into 25 fractions (ranged from 23 min to 60 min).

All 25 samples were desalted using SOLA HRP 10mg/ml cartridges (Thermo # 60109-001). Cartridges were conditioned with 90% methanol, 0.1% TFA, and then equilibrated with 0.1% TFA. Samples were dried to dryness in speed vacuum then reconstituted in 500 ul 0.1% TFA. Reconstituted samples were applied to the equilibrated cartridges and washed with 0.1% TFA. Peptides were eluted with two washes of 500ul of 50% ACN, 0.1% TFA (final volume = 1.0 ml). Eluted peptides were dried to dryness in speed vacuum and stored at -20C until LC MS/MS analysis.

NanoLC-MS/MS analysis

Peptides and cross-linked peptides were analyzed using an UltiMate3000 RSLCnano (Dionex, Sunnyvale, CA) coupled to an Orbitrap Fusion (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with a nanospray Flex Ion Source. Each sample was loaded onto an Acclaim PepMap 100 C18 trap column (5 µm, 100 µm x 20 mm, 100 Å, Thermo Fisher Scientific) at 20 µL/min of 0.5% FA. After 3 minutes, the valve switched to allow peptides to be separated on an Acclaim PepMap C18 nano column (3 µm, 75µm x 25cm, Thermo Fisher

Scientific), in a 120 min gradient of 5% to 40% B at 300 nL/min.

The Orbitrap Fusion was operating in positive ion mode with nano spray voltage set at 1.7 kV and source temperature at 275 °C. External calibration for FT, IT and quadrupole mass analyzers was performed prior to the analysis. Samples were analyzed using CID–MS2–MS3 workflow, in which peptides with charge states 4–10 were selected for CID MS2 acquisitions in Orbitrap analyzer with a resolution of 30,000 and an AGC target of 5×10^4. MS scan range was set to 375–1575 m/z and resolution was set to 60,000. The precursor isolation width was 1.6 m/z and the maximum injection time was 100 ms. The CID MS2 normalized collision energy was set to 25%. Targeted mass-difference-dependent CID–MS3 spectra were triggered for acquisition in the ion trap with CID collision energy of 35%; AGC target of 2×10^4 when a unique mass difference (Δ=31.9721 Da) was observed in the CID–MS2 spectrum. MS2 isolation window of 3 m/z with the maximum injection time set to 100 ms. All data were acquired under Xcalibur 3.0 operation software and Orbitrap Fusion Tune Application v. 2.1 (Thermo-Fisher Scientific).

Data processing, protein identification and data analysis

The precursors (MS) was isolated and fragmented (MS2), followed by selection and fragmentation of fragment peaks originating from the cross-linked peptide pair (MS3). All MS, MS2 and MS3 raw spectra from each sample were searched using Proteome Discoverer 2.2 (Thermo-Fisher Scientific, San Jose, CA) with XlinkX v2.0 algorithm for identification of cross linked peptides. The search parameters were as follow: three missed cleavages for full trypsin digestion or double digestion with fixed carbamidomethyl modification of cysteine, variable modifications of methionine oxidation, DSSO tris and DSSO hydrolyzed for lysine. The peptide mass tolerance was 10 ppm, and MS2 and MS3 fragment mass tolerance was 20 ppm

and 0.05 Da, respectively. For Xlink identification, the precursor mass tolerance was 10ppm and the fragment mass tolerance was 0.5Da for ITMS and 20ppm for FTMS. Bos Taurus NCBI refseq 2019 with added targeted protein sequences was used for PD 2.2 database search with 1% FDR for report of cross-link results. In addition, the search was also performed using the full sequences of the proteins including the recombinant tags. Identified xlinked peptides were filtered for Max. XlinkX Score >20 containing at least two identified MS3 spectra for each pair of xlinked peptides. Results of the search with were exported by the software as a spreadsheet.

3.3 Results

Expression and purification of SUMOylated proteins in *de novo* dTMP synthesis pathway using *Escherichia coli.*

Purification of endogenous SUMOylated proteins remains challenging. Most proteins subjected to SUMOylation appear to be modified at a small percentage at steady state[142]. In addition, SUMOylation is dynamic and subject to de-SUMOylating by SUMO proteases. Previously, an expression/modification system was developed to purify large amounts of SUMOylated proteins using *Escherichia coli* modified to express His-tagged human SUMO1 as well as the SUMO-activating and -conjugating enzymes[146], [148]. Here, we purified SUMO1-conjugated SHMT1, TYMS, and MTHFD1 (C/D domain) with a modified version of the previous purification system. Specifically, proteins of *de novo* thymidylate synthesis pathway were conjugated to CBP tag and purified with calmodulin affinity resin. The purified protein was a mixture of SUMOylated and non-SUMOylated target. Under native conditions, SHMT1, TYMS and MTHFD1 exist in the form of dimers or tetramers. It was technically challenging to further separate SUMOylated proteins from the non-SUMOylated proteins.

Identification of SUMOylation sites on enzymes of the *de novo* thymidylate biosynthesis pathway

To identify sites of SUMOylation, purified SUMOylated proteins were resolved by SDS-PAGE followed by in-gel digestion. SHMT1 and TYMS were digested by Trypsin and Glu-C. MTHFD1 (C/D domain) was digested by Trypsin and Chymotrypsin. Mass spectrometry analysis confirmed that K38 on SHMT1 and K223 on MTHFD1 were SUMOylated. Several novel SOMUylation sites were also identified, including K72 on SHMT1, K308 on TYMS, and multiple sites on MTHFD1 C/D domain (Table 3.3).

Table 3.3. Identification of SUMOylation site on enzymes in the de novo thymidylate biosynthesis pathway.

Protein	**SUMOylation site**
MTHFD1 (C/D domain)	K83, K175, K223, K245/K246, K262, K250
SHMT1	K38, K72
TYMS	K308

***In Vitro* XL-MS analysis of *de novo* thymidylate biosynthesis complex**

To define the spatial arrangement of the *de novo* thymidylate biosynthesis complex, XL-MS were performed on protein complexes assembled *in vitro*. Protein mixtures were cross-linked with either disuccinimido sulfoxide (DSSO) or disuccinimido dibutyric urea (DSBU). A titration of the cross-linking reagent DSSO or DSBU was performed to select the optimal concentration for each cross-linking reaction (Figure 3.1 & 3.2). XL-MS experiment was also performed using bovine serum albumin (BSA) as a positive control. The cross-linked proteins were analyzed by mass spectrometry. Unfortunately, only intra-crosslinked peptides were identified (Table 3.4).

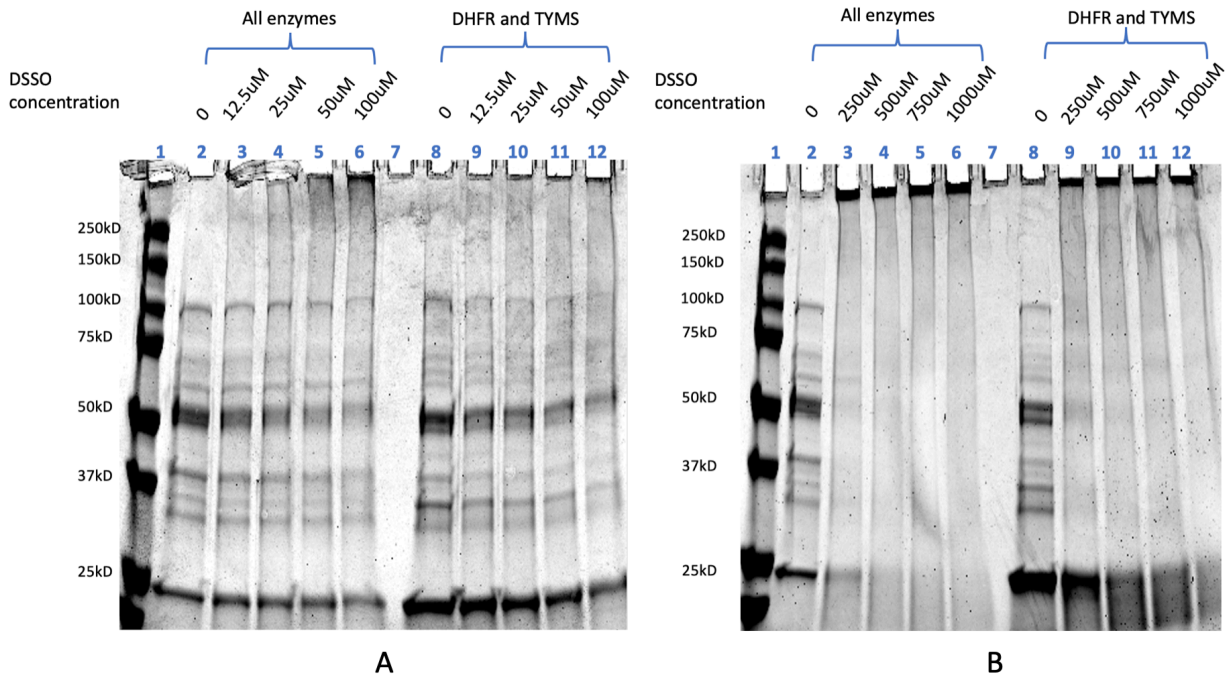

Figure 3.1. Optimization of DSSO cross-linking reaction conditions by titration. Different aliquots of protein mixtures were cross-linked with a range of DSSO concentrations. Proteins were separated by SDS-PAGE and stained with Coomassie Blue. (A & B) Lane 1, protein markers. Lane 2 and 8 correspond to the experiments when no DSSO was added. For lane 3-6 and 9-12, DSSO was added at increasing concentrations to mixtures of all four proteins, and mixtures of only TYMS and DHFR, respectively.

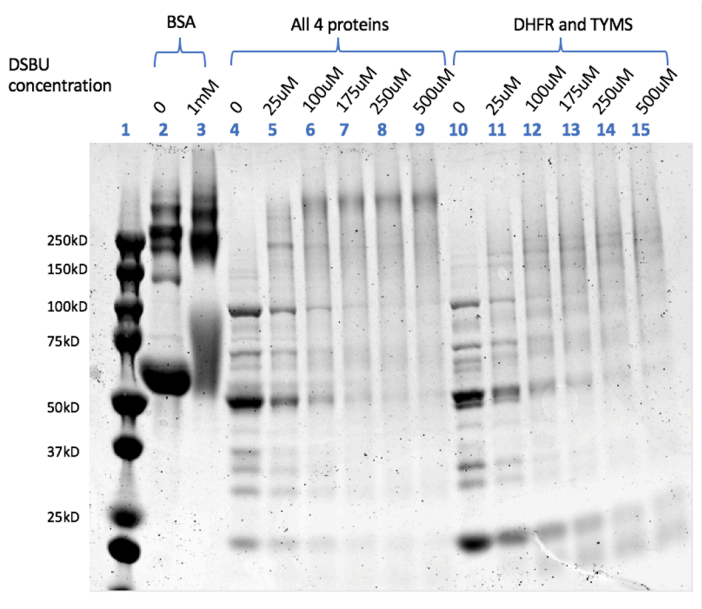

Figure 3.2. Optimization of DSBU cross-linking reaction conditions by titration. Different aliquots of protein mixtures were cross-linked with a range of DSBU concentrations. Proteins were separated by SDS-PAGE and stained with Coomassie Blue. Lane 1, protein markers. Lane 2, 4 and 10 correspond to the experiments when no DSBU was added. Lane 3 correspond to BSA cross-linking experiment with 1mM DSBU concentration. For lane 5-9 and 11-15, DSBU was added at increasing concentrations to mixtures of all four proteins, and mixtures of only TYMS and DHFR, respectively.

Table 3.4. Cross-linked lysine pairs identified among proteins in de novo thymidylate biosynthesis pathway by XL-MS.

			Peptide A		Peptide B	
Cross-linker	**Digestion**	**Peptide sequence**	**protein**	**Mod. Position**	**protein**	**Mod. Position**
DSSO	**Trypsin**	ELKEPPQGAHFLSR | QNLVIMGKK	DHFR	81	DHFR	55
		GEWIKPGAIVIDCGINYVPDDKKPNGR | KVVGDVAYDEAKER	MTHFD1 (C/D domain)	246	MTHFD1 (C/D domain)	251
		LKNQVTQLKEQVPGFTPR | AAEEIGIKATHIK	MTHFD1 (C/D domain)	28	MTHFD1 (C/D domain)	66
		AAEEIGIKATHIK | LKNQVTQLK	MTHFD1 (C/D domain)	66	MTHFD1 (C/D domain)	21
		LFHPKLIIAGTSCYSR | GVKSVDPK	SHMT1	196	SHMT1	274
		VLEACSIACNKNTCPGDR | ATLKEFKER	SHMT1	386	SHMT1	446
		VLEACSIACNKNTCPGDR | ATLKEFKER	SHMT1	386	SHMT1	449
		IMGLDLPDGGHLTHGFMTDKKK | GVKSVDPK	SHMT1	157	SHMT1	274
		GVKIWDANGSR | GSTNAKELSSK	TYMS	107	TYMS	99
	Trypsin and Chymotrypsin	ELKEPPQGAHF | QNLVIMGKK	DHFR	81	DHFR	55
		MQSTVESAKR | FKPGK	MTHFD1 (C/D domain)	292	MTHFD1 (C/D domain)	299
DSBU	**Trypsin**	AAEEIGIKATHIK | FKPGK	MTHFD1 (C/D domain)	66	MTHFD1 (C/D domain)	299
		LKNQVTQLKEQVPGFTPR | AAEEIGIKATHIK	MTHFD1 (C/D domain)	21	MTHFD1 (C/D domain)	66
		LKAAEEIGIK | ATHIKLPR	MTHFD1 (C/D domain)	58	MTHFD1 (C/D domain)	71
		LKNQVTQLKEQVPGFTPR | LKAAEEIGIK	MTHFD1 (C/D domain)	21	MTHFD1 (C/D domain)	58
		VLEACSIACNKNTCPGDR | ATLKEFK	SHMT1	386	SHMT1	446
		VLEACSIACNKNTCPGDR | EFKER	SHMT1	386	SHMT1	449
		GLLEKDFQK | SKGTDGGR	SHMT1	416	SHMT1	366
		IMGLDLPDGGHLTHGFMTDKKK | GVKSVDPK	SHMT1	157	SHMT1	274
		GTDGGRAEKVLEACSIACNK | ATLKEFKER	SHMT1	375	SHMT1	449
		VLEACSIACNKNTCPGDR | SKGTDGGR	SHMT1	386	SHMT1	366
		MLAQPLKDSDVEVYNIIK | KESNR	SHMT1	27	SHMT1	39
		LFHPKLIIAGTSCYSR | GVKSVDPK	SHMT1	196	SHMT1	274
		AEKVLEACSIACNK | ATLKEFKER	SHMT1	375	SHMT1	446
		ALSEALTELGYKIVTGGSDNHLILVDLR | SKGTDGGR	SHMT1	348	SHMT1	366
		GVKIWDANGSR | GSTNAKELSSK	TYMS	107	TYMS	99
	Trypsin and Chymotrypsin	ELKEPPQGAHF | VIMGKK	DHFR	81	DHFR	55
		MQSTVESAKR | FLEKF	MTHFD1 (C/D domain)	292	MTHFD1 (C/D domain)	297
		MQSTVESAKR | FKPGK	MTHFD1 (C/D domain)	292	MTHFD1 (C/D domain)	299
		KEQVPGFTPR | LKNQVTQL	MTHFD1 (C/D domain)	28	MTHFD1 (C/D domain)	21
		KIVTGGSDNHLIL | SKGTDGGR	SHMT1	348	SHMT1	366
		ATLKEF | AEKVL	SHMT1	446	SHMT1	375
		GVKSVDPK | LFHPKL	SHMT1	274	SHMT1	196
		VIDTIKTNPDDRR | KVEK	TYMS	169	TYMS	284

3.4 Discussion

Previous studies in our lab suggested that K38/K39 on SHMT1 and K221 on MTHFD1 were likely to be SUMOylated[19], [22]. This study confirmed that K38 on SHMT1 and K223 on MTHFD1 are SUMOylated using mass spectrometry. In addition, we identified multiple novel SUMO1 modification sites on proteins of *de novo* thymidylate synthesis pathway, including K72 on SHMT1, K308 on TYMS, and K83, K175, K245/K246, K262, K250 on MTHFD1 (C/D domain). A previous *In vitro* SUMOylation study indicated that DHFR is a substrate for UBC9-catalyzed SUMOylation in by SUMO1[14]. However, the percentage of SUMOylated DHFR produced in bacterial lines was too small to be purified and separated by SDS-PAGE for mass spectrometry analysis.

We hypothesized that complex formation of *de novo* dTMP synthesis pathway is mediated through multiple SUMO-SIM interactions among the enzymes. However, XL/MS analysis of DSSO and DSBU cross-linked *in vitro* assembled *de novo* dTMP biosynthesis complex only showed intra crosslinks and no inter crosslinks. Several limitations of this study could possibly explain why only intra crosslinks were identified. Firstly, DHFR protein used to assemble *de novo* thymidylate synthesis complex *in vitro* for XL/MS was not SUMOylated. The affinity between SUMO and SIMs is in the high micromolar range[142], meaning that additional contacts between the SUMOylated protein and its binding partner or multiple SUMO-SIMs interactions are needed for interaction in occur. Secondly, SHMTs are tetrameric proteins. DHFR, TYMS and MTHFD1 function as dimers. It was technically difficult to separate SUMO1 modified proteins from non-SUMOylated proteins during protein purification. The proteins used for cross-linking reactions were a mixture of SUMOylated and non-SUMOylated proteins, and the percentage of SUMOylated proteins in the mixture could be lower than the percentage of

endogenous SUMOylated proteins in the nucleus. Thirdly, both DSSO and DSBU react mainly with lysine residues. It is likely that there is not a sufficient number of lysine residues accessible for cross-linking where proteins of the *de novo* dTMP synthesis pathway are interacting. The length of spacer arm of DSSO and DSBU may also limit cross-linking reactions. Several novel MS-cleavable crosslinkers have been developing and are promising for future studies. One of them is Phospho-bisvinylsulfone (pBVS), a recently developed MS-cleavable cross-linker that targets multiple types of amino acid residues, including cysteine, lysine and histidine[149]. Lastly, *de novo* dTMP synthesis complex used for cross-linking reactions was assembled *in vitro* that may not be ideal for protein interactions. These proteins likely lack other post-translational modifications that are crucial for protein-protein interactions. For example, CK2 regulated phosphorylation adjacent to the hydrophobic core of SIM has been shown to mediate SUMO-SIM interaction[150].

Apart from facilitating enzyme complex formation through SUMO-SIM interactions, there are other possible roles that SUMO modification plays in the *de novo* thymidylate biosynthesis pathway. SUMO modifications of *de novo* thymidylate synthesis enzymes may change enzymatic activities or affect the protein interactions with other cellular proteins. Evidence suggest that SUMO modification facilitates nuclear import of the *de novo* thymidylate synthesis enzymes during S-phase of the cell cycle. A previous study in our lab showed that SUMOylation of SHMT1 occurred at the nuclear pore and was linked to its nuclear import [19]. K38R/K39R SHMT1 mutants were not substrates for Ubc9-mediated SUMOylation and were not translocated to the nucleus during S-phase[19]. SUMO modification of MTHFD1 was crucial for its nuclear translocation during S-Phase of the cell cycle in HeLa cells[22]. The MTHFD1-K223R-GFP mutant protein exhibited impaired nuclear translocation in S-phase blocked cells

and less *in vitro* SUMOylation compared with MTHFD1-GFP fusion protein[22]. It is possible that SUMO modifications of DHFR and TYMS have similar effect on protein nuclear translocation. The compartmentalization of *de novo* thymidylate synthesis pathway during S-phase is advantageous for cells that require increased amount of dTMP for DNA replication.

In cells, the total concentration of folate-dependent enzymes exceeds the total cellular concentration of folate[13]. The regulation of the partitioning of limited amount of folate cofactors among the pathways that constitute the FOCM network is crucial, especially under folate deficiency or other types of cellular stresses. The nuclear localization of *de novo* thymidylate synthesis pathway helps compete for methylene-tetrahydrofoalte from homocysteine remethylation, which occurs in the cytosol[16].

In conclusion, our study identified multiple SUMOylation sites on enzymes of the *de novo* dTMP biosynthesis pathway. Only intra crosslinks were identified from XL-MS analysis of *in vitro* assembled *de novo* dTMP synthesis complex using DSSO and DSBU as cross-linking reagents. Further investigations are needed to characterize the spatial structure of the *de novo* dTMP synthesis complex and the underlying mechanism of the protein-protein interactions.

Acknowledgements

We thank the Proteomic and MS Facility of Cornell University for providing the mass spectrometry data and NIH SIG grant 1S10 OD017992-01 grant support for the Orbitrap Fusion mass spectrometer.

Chapter 4 A HYBRID-STOCHASTIC MODEL OF FOLATE-MEDIATED ONE-CARBON METABOLISM: ROLE OF THE NUCLEAR COMPARTMENT AND SENSITIVITY ANALYSIS OF THE NETWORK

Authors: Karla Misselbeck[1,2+], Xu Lan[3], Luca Marchetti[1+], Corrado Priami[1,2], Martha S. Field[3]* and Patrick J. Stover[3,4*]

Affiliations: [1]The Microsoft Research - University of Trento Centre for Computational and Systems Biology (COSBI), Piazza Manifattura, 1, 38068 Rovereto, Italy

[2]Department of Mathematics, University of Trento, 38123 Povo, Italy

[3]Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853

[4]College of Agriculture and Life Sciences, Texas A&M University, College Station, TX 77843-2142

Author Contributions:

K. Misselbeck (K.M.), Xu Lan (X.L.), L. Marchetti (L.M.), C. Priami (C.P.), P.J. Stover (P.J.S.), and M.S. Field (M.S.F.) contributed to the experimental design. K.M., L.M. developed the computation model and carried out all simulations, X.L. conducted all experimental studies in cell culture models, while M.S.F. and P.J.S. provided biological insight and extracted kinetic variables from the literature. C.P., P.J.S., and M.S.F. provided overall guidance of the project.

Abstract:

Mammalian folate-mediated one-carbon metabolism (FOCM) in the cytosol is a highly interconnected network of metabolic pathways that synthesize purine bases and thymidylate (dTMP), and remethylate homocysteine to methionine. Simulations from a previously published hybrid-stochastic model of FOCM indicated that rates of *de novo* dTMP synthesis in the cytosol are insufficient to produce adequate dTMP for DNA synthesis during S-phase. Recent experimental evidence indicates that *de novo* dTMP synthesis occurs within the nuclear compartment specifically at sites of DNA replication to prevent uracil misincorporation into DNA. The existing hybrid-stochastic FOCM model has been extended to include the nuclear compartment in order to explore its effect on the FOCM network, with emphasis on rates of *de novo* dTMP biosynthesis. In addition, a key aspect of nuclear FOCM is the formation of a multienzyme complex, which permits folate channeling among the active sites of the involved enzymes. Both nuclear complex formation and substrate channeling are taken into account in this extended model and computational simulations highlight their importance for meeting cellular needs for genome replication. *In silico* experiments provide insight into the contribution of cytosolic and nuclear dTMP synthesis to overall dTMP synthesis. Moreover, model simulations provide evidence that the network is most sensitive to expression levels of TYMS and DHFR, and that folate partitioning between the cytosol and the nucleus plays an important role in cellular dTMP synthesis capacity.

4.1 Introduction

Folate-mediated one-carbon metabolism (FOCM, for a list of the abbreviations refer to Table 4.1) is required for the *de novo* synthesis of three of the four DNA bases and the remethylation of methionine (MET) to homocysteine (HCY) [151]. In the

interconnected metabolic network, folate serves as one-carbon carrier and donor for biochemical reactions associated with the anabolic pathways[17]. FOCM plays an essential role for genome stability and methylation[1], and disruption of the network can be caused by genetic and/or nutritional factors, including genetic variants of the required enzymes or folate and vitamin B12 deficiency[2]. Impairment of FOCM is associated with the pathogenesis of neural tube defects, neurodegenerative diseases and cancer[152], [153].

In the cytosol, FOCM is required for the *de novo* nucleotide biosynthesis of purines and thymidylate (dTMP), as well as for HCY remethylation to MET, which is catalyzed by the vitamin B12-dependent enzyme methionine synthase (MTR) (Figure 4.1). The enzymes that comprise the *de novo* purine biosynthesis pathway form a multienzyme complex referred to as the "purinosome" that forms during purine-deficient conditions[154]. Other studies demonstrate that FOCM also occurs in the nucleus at the site of DNA synthesis[18], [81], [155]. The enzymes present in the nucleus include serine hydroxymethyltransferase (SHMT), dihydrofolate reductase (DHFR), methylenetetrahydrofolate dehydrogenase 1 (MTHFD1), methenyltetrahydrofolate synthase (MTHFS) and thymidylate synthase (TYMS). The required co-factor for dTMP synthesis, 5,10-methyleneTHF (CH2F), is independently generated by SHMT and MTHFD1 using serine and formate as one-carbon sources, respectively. TYMS utilizes CH2F as a co-factor to catalyze the conversion of dUMP to dTMP and dihydrofolate (DHF). DHFR reduces DHF to THF in a NADPH-dependent reaction, regenerating the THF cofactor. In mammalian cells, a fraction of the enzymes involved in dTMP synthesis translocate to the nuclear compartment during S-phase of the cell cycle or in response to DNA damage after undergoing post-translational modification by the small ubiquitin like modifier (SUMO) protein[14], [15]. Nuclear enzymes involved in the *de*

novo dTMP synthesis pathway form a multienzyme complex, which is associated with the nuclear lamina and with the replication and epigenetic machinery[16]. Experimental observations show that SHMT acts as essential scaffold protein that anchors this multienzyme complex to nuclear lamina[16], [18]. Formation of the nuclear enzyme complex appears to be essential for the functioning of *de novo* dTMP synthesis, as previous studies have demonstrated that dTMP synthesis activity in isolated nuclei is reduced following sonication[16]. Furthermore, nuclear metabolic complex formation may allow for channeling of folate polyglutamate cofactors among enzyme active sites, limiting substrate diffusion and accelerating enzymatic reaction rates[17]. Interestingly, the formation of a multienzyme complex and resulting substrate channeling seems to be unique to mammalian cells[155]. Folate deficiency, anti-folate treatment or impaired assembly of the multienzyme complex responsible for nuclear *de novo* dTMP synthesis can lead to depressed dTMP biosynthesis activity and consequently to elevated uracil incorporation in DNA[1], [12], [13]. During DNA synthesis, either dTTP or dUTP can basepair with an adenine nucleotide base on the template strand; uracil misincorporation into DNA increases when dTMP becomes limiting. Therefore, adequate intracellular concentrations of folate and proper assembly of the dTMP synthesis pathway are important for genome stability.

Previously, we have reported the development of a hybrid stochastic model of FOCM that faithfully reflected experimental and clinical studies of FOCM functioning. We have shown that low levels of folate and the common C677T *MTHFR* polymorphism affect cytosolic *de novo* dTMP synthesis[132] whereas impairment of the 5-formyltetrahydrofolate (5fTHF) futile cycle introduced by decreased MTHFS activity is mainly associated to loss in purine synthesis activity[156]. The simulations also

indicate that *de novo* dTMP synthesis rates in the cytosol are insufficient to support DNA synthesis during S-phase[132]. Here, we extend the computational model of FOCM described previously by including its compartmentalization to the nucleus and by accounting for the kinetic effects of multienzyme complex formation. *In silico* experiments with the multi-compartmental model allow us to better understand the impact of enzyme complex formation and substrate channeling on overall network outcomes with a special focus on *de novo* dTMP synthesis activity. Furthermore, sensitivity analysis was conducted throughout the network to identify the impact of changes in the expression of individual enzymes on the overall functioning of FOCM.

Table 4.1. List of Abbreviations and acronyms.

5fTHF	5-formyl tetrahydrofolate
5mTHF	5-methyl tetrahydrofolate
10fTHF	10-formyltetrahydrofolate
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AICARFT	Phosphoribosylaminoimidazolecarboxamide formyltransferase
BHMT	Betaine-homocysteine methyltransferase
CHF	5,10-methenyltetrahydrofolate
CH2F	5,10-methylenetetrahydrofolate
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DNMT	DNA methyltransferase
dUMP	Deoxyuridine monophosphate
dTMP	Deoxythymidine monophosphate
FOCM	Folate-mediated one-carbon metabolism
FTD	10-formyltetrahydrofolate dehydrogenase
FTS	Formate-tetrahydrofolate ligase
GAR	Glycinamide ribonucleotide
GNMT	Glycine N-methyltransferase
HCY	Homocysteine
MAT-I	Methionine adenosyltransferase 1
MAT-III	Methionine adenosyltransferase 3
MET	Methionine
MTCH	Methenyltetrahydrofolate cyclohydrolase
MTD	Methylenetetrahydrofolate dehydrogenase
MTHFD	Methylenetetrahydrofolate dehydrogenase
MTHFR	Methylenetetrahydrofolate reductase
MTHFS	Methenyltetrahydrofolate synthetase
MTR	Methionine synthase
NADP+	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced form of Nicotinamide adenine dinucleotide phosphate
PGT	Phosphoribosylglycinamide formyltransferase
THF	Tetrahydrofolate
TYMS	Thymidylate synthase
SAH	S-adenosyl-homocysteine
SAHH	S-adenosylhomocysteine hydrolase
SAM	S-adenosyl-methionine
SHMT	Serine Hydroxymethyltransferase

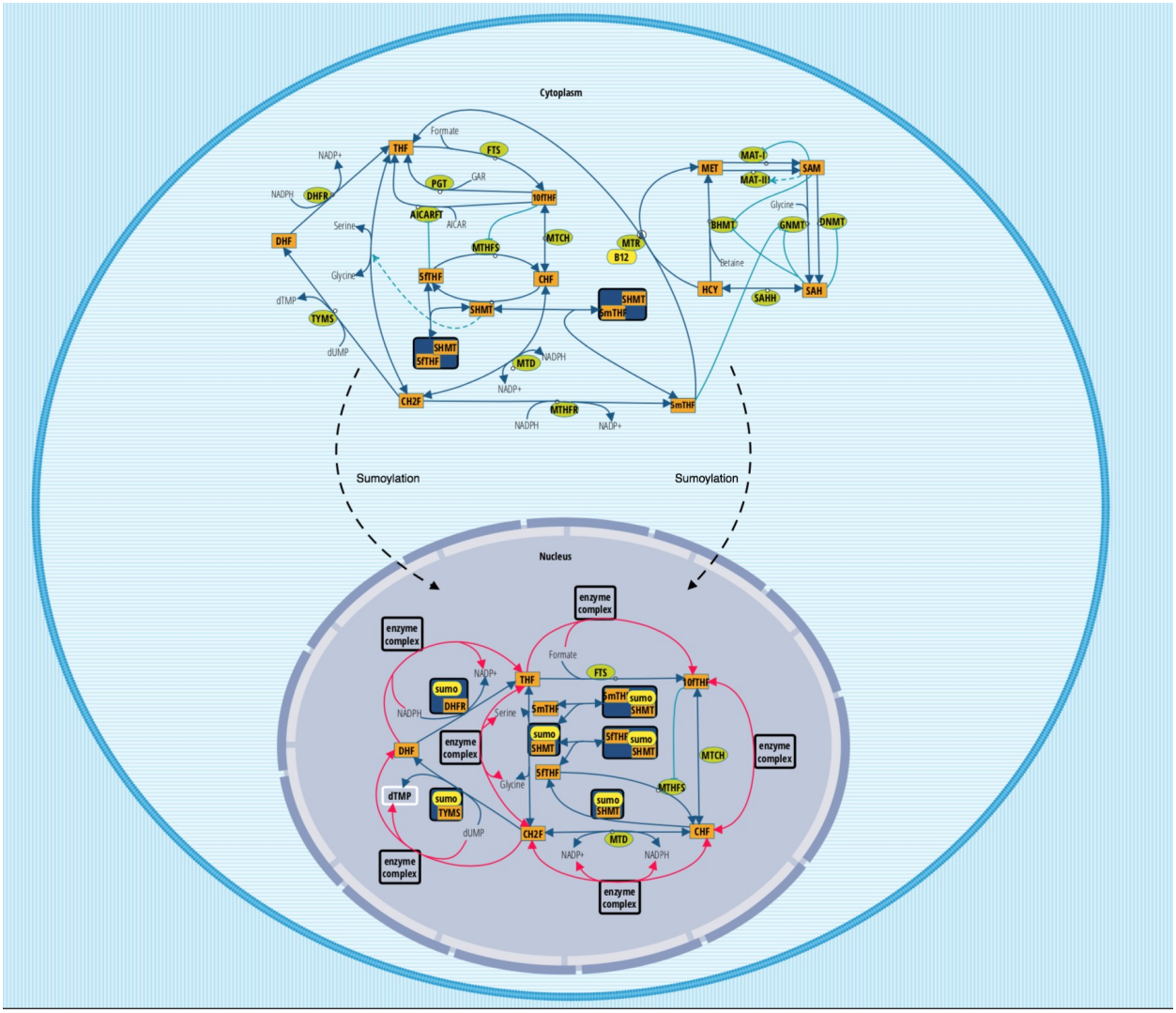

Figure 4.1. Folate-mediated one-carbon metabolism. The reaction-based specification of the model including the cytosolic and nuclear compartment according to the notation introduced previously[132]. Orange rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions).

4.2 Materials and Methods

Description of the model

The results presented herein are based on an extension of the computational model of FOCM in the cytosol described previously[132], [156]. This model provides a description of FOCM in the cytosol including its regulation of key biological processes related to *de novo* dTMP synthesis, *de novo* purine synthesis and remethylation of homocysteine to methionine. With respect to the initial model, the model herein employed has been extended to include the folate-mediated reactions occurring in the nuclear compartment (Figure 4.1, Table 4.2, 4.4, & 4.5). The model consists of 36 reversible and irreversible reactions, most of which have been parametrized by means of Michaelis-Menten kinetics with one or two substrates as described previously[156]. The parameter estimates for the folate cycle can be found in the Table 4.2, whereas the parameter estimates for the reactions related to the remethylation of methionine were as previously reported[156]. Whenever possible, physiologically relevant forms of folate polyglutamate cofactors in the nucleus have been considered, in agreement with the cytosolic model (Table 4.4). In the following we describe the technical procedure we applied to update the cytosolic model described previously[156] by including the nuclear compartment.

Updating the cytosolic compartment

The cytosolic model[156] was extended to account for enzyme translocation to the nucleus. The enzymes SHMT, MTHFD1 (trifunctional enzyme encoded as FTS, MTD and MTCH activities), DHFR, TYMS and MTHFS have been considered as variables that can change in time (some of them were encoded in the Michaelis-Menten constants in the previous model). Due to this, the Michaelis-Menten kinetics of the reactions catalyzed by the enzymes MTHFD1, DHFR and TYMS were updated to consider the turnover number and the enzyme

concentration in place of the V_{max} value, considering the relation $V_{max} = k_{cat} \cdot [Enzyme]$.

Table 4.2. Parameter estimates for the extended model including cytoplasmic and nuclear FOCM grouped by reactions. All concentrations are expressed in μm, while time is expressed in hours. The parameter estimates have been calculated by scaling the respective cytoplasmic values to the nuclear volume. The estimates for the reactions forming the remethylation of homocysteine cycle are listed in Table 4.3.

Parameter	Metabolite	Value		Cell line	Reference
		cytoplasm	nucleus		
Volume		$9.4 \cdot 10^{13}$	$2.2 \cdot 10^{13}$	HeLa	[164]
$R_{AICARFT}$: 10fTHF → THF					
V_{max}		63350		MCF-7	[165]
K_m	10fTHF	0.3		Human leukemia	[166]
K_m	AICAR	16.8		Human purH	[167]
K_i	5fTHF	3			[116]
R_{DHFR} : DHF → THF					
k_{cat}		36000	36000	Human	[168]
K_m	DHF	0.5	2.14	L1210	[169]
K_m	NADPH	4.3	18.37	L1210	[169]
R_{FTS} : THF → 10fTHF					
k_{cat}		5100	5100	L1210	[170]
K_m	THF	0.1	0.43	L1210	[170]
K_m	formate	16	68.36	L1210	[170]
R_{MTCH} : CHF → 10fTHF					
k_{cat}		324000	324000	L1210	[170]
K_m	CHF	4	17.09	L1210	[170]
R_{MTCH} : 10fTHF → CHF					
k_{cat}		324000	324000	L1210	[170]
K_m	10fTHF	20	85.45	L1210	[170]

to continue on next page . . .

Table 4.2 continued

Parameter	Metabolite	Value cytoplasm	Value nucleus	Cell line	Reference
R_{MTD} : CHF → CH2F					
k_{cat}		66000	66000	L1210	[170]
K_m	CHF	6.3	26.92	Human DC301	[196]
K_m	NADPH	10.5	44.86	Human DC301	[196]
R_{MTD} : CH2F → CHF					
k_{cat}		66000	66000	L1210	[170]
K_m	CH2F	2	8.55	L1210	[170]
K_m	NADP+	2	8.55	L1201	[170]
R_{MTHFR} : CH2F → 5mTHF					
V_{max}		120		Pig liver	[171]
K_m	CH2F	0.26		Pig liver	[171]
K_m	NADPH	125		Pig liver	[171]
R_{MTHFS} : 5fTHF → CHF					
k_{cat}		5400	5400		[26]
K_m	5fTHF	0.2	0.85		[26]
K_i	10fTHF	0.015	0.64		[26]
R_{MTR} : 5mTHF + HCY → THF + MET					
V_{max}		26			Estimated in the range 0.024 [171]-50 [172] μM/h
K_m	5mTHF	0.5		Pig liver	[171]
K_m	HCY	0.1			[173]
R_{PGT} : 10fTHF → THF					
V_{max}		6600			[173]
K_m	10fTHF	0.9		human	[174]
K_m	GAR	1.1		human	[174]
R_{SHMT} : THF → CH2F					
k_{cat}		18000	18000	L1210	[170]
K_m	Serine	600	2563.6	L1210	[170]
K_m	THF	0.2	0.85	L1210	[170]
R_{SHMT} : CH2F → THF					

to continue on next page . . .

Table 4.2 continued

Parameter	Metabolite	Value		Cell line	Reference
		cytoplasm	nucleus		
k_{cat}		45000	45000	Rabbit liver	[190]
K_m	Glycine	3000	12818	L1210	[170]
K_m	CH2F	0.2	0.85	L1210	[170]
R_{SHMT} : CHF → 5fTHF					
k_{cat}		198	198		[191]
K_m	CHF	40	170.9		[191]
R_{TYMS} : CH2F → DHF					
k_{cat}		11196	11196	Human	[192]
K_m	CH2F	4.3	18.37	Human colon	[193]
K_m	dUMP	3.6	15.38	Human colon	[193]
(un-)binding of 5mTHF and SHMT					
$k_{unbinding}$		1980	1980	Rabbit liver	[27]
k_D		0.4	1.71	Rabbit liver	[27]
$k_{binding}$		4950	1158	Rabbitliver	$k_{binding} = \frac{k_{unbinding}}{K_D}$
(un-)binding of 5fTHF and SHMT					
$k_{unbinding}$		144	144	Rabbit liver	[27]
k_D		0.2	0.85	Rabbit liver	[27]
$k_{binding}$		720	169	Rabbitliver	$k_{binding} = \frac{k_{unbinding}}{K_D}$

Table 4.3. Model parameter estimates for the homocysteine remethylation cycle grouped by reactions. All concentrations are expressed in μm, while time is expressed in hours.

Parameter	Metabolite	Value	Reference
R_{BHMT} : HCY → MET			
Vmax		2160	[172]
K_m	HCY	12	[172]
K_m	Betaine	100	[172]
R_{DNMT} : SAM → SAH			
Vmax		180	[172]
K_m	SAM	1.4	[172]
K_i	Inhibition by SAH	1.4	[172]
R_{GNMT} : SAM → SAH			
Vmax		245	[172]
K_m	SAM	32	[172]
K_m	Glycine	130	[172]
K_i	Inhibition by SAH	18	[172]
R_{MAT-I} : MET → SAM			
Vmax		260	[172]
K_m	MET	41	[172]
$R_{MAT-III}$: MET → SAM			
Vmax		220	[172]
K_m	MET	300	[172]
K_a	Activation by SAM	360	[172]
R_{SAHH} : SAH → HCY			
Vmax		320	[172]
K_m	SAH	6.5	[172]
R_{SAHH} : HCY → SAH			
Vmax		4530	[172]
K_m	HCY	150	[172]

Table 4.4. Steady state fluxes for all model reactions (in µm/h).

	cytoplasm		nucleus	
	uM	%	uM	%
THF	0.052	0.31	0.045	0.57
10fTHF	6.765	40.01	3.544	44.15
CHF	1.346	7.96	0.688	8.57
CH2F	0.381	2.25	0.109	1.35
DHF	0.003	0.02	0.003	0.04
5mTHF free	4.324	25.57	0.549	6.84
5mTHF bound	3.096	18.31	0.254	3.16
5mTHF total	7.420	43.88	0.803	10.00
5fTHF free	0.388	2.29	1.475	18.37
5fTHF bound	0.555	3.28	1.361	16.96
5fTHF total	0.943	5.58	2.836	35.33
SHMT	0.286			
HCY	3.255			
MET	38.808			
SAM	72.899			
SAH	35.559			

Table 4.5. Steady state concentration (in μm) and steady state distribution of folate (in percentage of total folate per compartment) of the cytoplasm and nucleus.

	cytoplasm	nucleus
FTS	14069.6	10980.7
MTCH (10fTHF → CHF)	713991.6	387726.9
MTCH (CHF → 10fTHF)	711138.2	376746.2
MTD (CHF → CH2F)	85758.9	27868.4
MTD (CH2F → CHF)	82905.5	16887.7
DHFR	72.5	1118.3
TYMS	72.5	1118.3
SHMT (CH2F → THF)	3223.8	12630.2
SHMT (THF → CH2F)	465.5	2767.7
SHMT (CHF → 5fTHF)	1.8	0.6
MTHFS	1.8	0.6
MTHFR	22.6	
MTR	22.6	
PGT	5247.8	
AICARFT	5968.4	
BHMT	151.7	
MAT_I	112.8	
MAT_III	61.6	
GNMT	158.9	
DNMT	119.4	
SAHH_SAH	270.5	
SAHH_HCY	96.2	
5mTHF:SHMT (unbinding)	6129.4	
5mTHF:SHMT (binding)	6129.4	
5fTHF:SHMT (unbinding)	80.0	
5fTHF:SHMT (binding)	80.0	

Adding the nucleus

The inclusion of the nuclear compartment and its interplay with the cytosol is the main novelty of the model herein presented. The nucleus has been modeled to account for the folate-dependent reactions catalyzed by the enzymes MTHFD1, SHMT, DHFR, TYMS and MTHFS and consists of 17 variables and 16 (reversible and irreversible) reactions, as indicated in Figure 4.1 and Table 4.5. The nuclear stoichiometry and regulation of these reactions have been inherited from their counterparts in the cytosol, while model parameters have been derived as explained in the following. Furthermore, the model has been developed to account for observations that the four enzymes, MTHFD1, SHMT, DHFR and TYMS, form a multi-enzyme

complex when they translocate to the nucleus[50]. This complex has the proposed structure 1: TYMS, 1: DHFR, 1: MTHFD1, 2: SHMT and allows for substrate channeling[17], which was modeled as explained in the following.

Translocation of the de novo thymidylate synthesis pathway to the nucleus

The translocation of enzymes and metabolites to the nucleus has been modeled by an abstract function representing the SUMO-dependent nuclear import process according to experimental results[81]. The function derives the enzyme and folate availability in the nucleus by subtracting a given percentage of each folate form or enzyme from the steady state of the model restricted to the cytosol. In particular, the function follows the rules/steps described below.

Nuclear kinetic parameter estimates

The parameter estimates were adapted from the values considered for the cytosol by scaling them according to the volume of the nucleus (Table 4.2).

Nuclear constant substrates

The constant substrates (serine, glycine, formate, NADPH, NADP, and dUMP) were modeled at the same constant levels used for the cytosol[132].

Nuclear folate one-carbon forms

Folate has been distributed between the two compartments by assuming that 10% of the total cytosolic folate mass is present in the nucleus[18]. The initial nuclear folate distribution has been calculated by assigning 10% and 35% of total nuclear folate to 5mTHF and 5fTHF, respectively, according to[157] (Table 4.4). The remaining 55% of nuclear folate have been assigned to THF, 10fTHF, CHF, CH2F and DHF by preserving the same proportions of their steady state distribution in the cytosol. The remaining 90% of total folate mass is redistributed

in the cytosol according to its steady state distribution.

Nuclear enzymes

The availability of MTHFS in the nucleus has been estimated to preserve 35% of nuclear folate being 5fTHF at steady state according to[157], resulting in a concentration of 0.0039 μm (Table 4.6). The 5fTHF is presumed to be bound to SHMT as a slow-binding inhibitor[27]. The enzymes MTHFD1, SHMT, DHFR and TYMS have been translocated to the nucleus by assuming that all enzymes are in a complex of the structure 1:TYMS, 1:DHFR, 1:MTHFD1, 2:SHMT. This assumption is based on knowledge that SHMT is a tetramer with 4 active sites, whereas TYMS, MTHFD1 and DHFR are dimers with 2 active sites. Among the enzymes forming the complex, TYMS was the least abundant and therefore the availability of the complex has been based on a given percentage of TYMS availability in the cytosol. In particular, 75% of cytosolic TYMS was assumed to be translocated to the nucleus. An equal amount of DHFR and MTHFD1 has been translocated to the nucleus, whereas twice the amount of SHMT was imported to maintain stoichiometry (Table 4.6).

Table 4.6. Concentrations of nuclear enzymes (in μm) before and after their translocation to the nuclear compartment. The values provided for the nucleus reflect the enzyme availability in their free form and after the multienzymatic complex formation.

Enzyme	Cytoplasm		Nucleus	
	before their translocation	after their translocation	free form	after complex formation
MTHFS	0.080	0.080	0.0039	0.0039
DHFR	0.617	0.335	1.202	-
MTHFD1	9.000	8.719	1.202	-
SHMT	4.500	3.937	2.404	-
TYMS	0.375	0.094	1.202	-
Complex	-	-	-	1.202

Reactions catalyzed by the nuclear enzymatic complex

The reactions catalyzed by the multi-enzyme complex are modeled considering

replicates of the reactions catalyzed by the single, free enzymes. Because the enzyme complex allows for substrate channeling, the reaction included in the channeling have been multiplied by an "acceleration" or scaling factor to enhance the reaction rates. To derive the scaling factor, two approaches were considered:

The first approach was based on kinetic measures of folate substrate channeling. Based on experimental evidence reported previously on the channeling of folates within the TYMS-DHFR bifunctional enzyme[158], [159], we considered a 20-fold scaling factor for the enzyme channeling. Due to the lack of other available experimental data for the other enzymes SHMT and MTHFD1 which form the complete channeling network, we assumed that a similar scaling factor would also apply to the remaining enzymes.

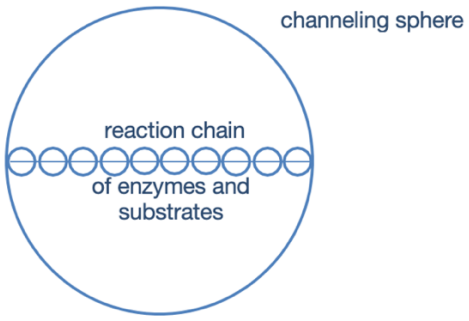

Figure 4.2. Schematic representation of the channeling reaction volume. A channeling sphere consists of a unit complex capable of *de novo* thymidylate biosynthesis consisting of 1 TYMS dimer, 1 DHFR dimer, 1 MTHFD1 dimer and 1 SHMT tetramer.

The second approach was based on the quantification of the channeling reaction volume. As an alternative approach to the one described above, we derived the scaling factor based on the estimation of the reaction volume of the multi-enzyme complex. We interpret the channeling as a set of reactions working in close proximity in a smaller spherical reaction volume. A rough estimation of the volume diameter can be given by the length of the reaction chain that includes

all the involved enzymes and substrates (Figure 4.2). The length of the reaction chain was derived by summing up the diameters of all molecules, which have been approximated as spheres (Table 4.7 & 4.8). The diameter of each sphere representing an enzyme molecule was computed by deriving the volume of the molecule according to the following formula:

$$Volume = MW \cdot PSV,$$

where MW indicates the corresponding molecular weight (Table 4.7) and PSV indicates the average protein partial specific volume (0.72 cm^3/g in all cases[160]). Once the total volume of the enzyme complexes in the nucleus was estimated, the final scaling factor was computed by the ratio of the nuclear volume and the volume of the multi-enzyme complex. These calculations resulted in a scaling factor of 25 (see Tables 4.7 & 4.8).

Table 4.7. Diameters for the molecules involved in the channeling. The structure of the molecules has been approximated as sphere and its volume has been calculated by considering the molecular weight (MW) and the average protein partial specific volume (PSV = 0.72 cm^3/g).

Enzyme	**MW (g/mol)**	**Volume** (L)	**radius** (dm)	**diameter** (dm)
MTHFD1	102000	$2.45 \cdot 10^{-19}$	$3.88 \cdot 10^{-8}$	$7.76 \cdot 10^{-8}$
SHMT	53000	$2.54 \cdot 10^{-19}$	$3.93 \cdot 10^{-8}$	$7.86 \cdot 10^{-8}$
TYMS	36000	$8.64 \cdot 10^{-23}$	$2.74 \cdot 10^{-8}$	$5.48 \cdot 10^{-8}$
DHFR	21000	$5.04 \cdot 10^{-23}$	$2.29 \cdot 10^{-8}$	$4.58 \cdot 10^{-7}$
Substrates & Products	**MW (g/mol)**	**Volume** (L)	**radius** (dm)	**diameter** (dm)
THF	445.43	$5.35 \cdot 10^{-25}$	$5.03 \cdot 10^{-9}$	$1.01 \cdot 10^{-8}$
10fTHF	473.44	$5.68 \cdot 10^{-25}$	$5.14 \cdot 10^{-9}$	$1.03 \cdot 10^{-8}$
CHF	582.53	$6.99 \cdot 10^{-25}$	$5.51 \cdot 10^{-9}$	$1.10 \cdot 10^{-8}$
CH2F	457.43	$5.49 \cdot 10^{-25}$	$5.08 \cdot 10^{-9}$	$1.02 \cdot 10^{-8}$
DHF	443.42	$5.32 \cdot 10^{-25}$	$5.03 \cdot 10^{-9}$	$1.01 \cdot 10^{-8}$
dUMP	308.18	$3.70 \cdot 10^{-25}$	$4.45 \cdot 10^{-9}$	$8.91 \cdot 10^{-9}$
dTMP	322.21	$3.87 \cdot 10^{-25}$	$4.52 \cdot 10^{-9}$	$9.04 \cdot 10^{-9}$
NADPH	833.35	$1.00 \cdot 10^{-24}$	$6.20 \cdot 10^{-9}$	$1.24 \cdot 10^{-8}$
NADP+	744.42	$8.93 \cdot 10^{-25}$	$5.97 \cdot 10^{-9}$	$1.19 \cdot 10^{-8}$
Formate	46.02	$5.52 \cdot 10^{-26}$	$2.36 \cdot 10^{-9}$	$4.72 \cdot 10^{-9}$
Serine	105.09	$1.26 \cdot 10^{-25}$	$3.11 \cdot 10^{-9}$	$6.22 \cdot 10^{-9}$
Glycine	75.07	$9.01 \cdot 10^{-26}$	$2.78 \cdot 10^{-9}$	$5.56 \cdot 10^{-9}$

Table 4.8. Calculation of the channeling speed-up factor. The diameter of the channeling sphere is estimated by the length of the reaction chain including all enzymes and substrates (diameter of involved enzymes, substrates and products, and diameter total). The standard equation for the volume of a sphere ($V = \frac{4}{3} r^3\pi$) is used to calculate the volume for one complex. The total channeling volume results from the multiplication of the number of complexes in the nucleus (Table 4.6).

Diameter of involved enzymes (dm)	$3.36 \cdot 10^{-7}$
Diameter of involved substrates and products (dm)	$1.35 \cdot 10^{-7}$
Diameter total (dm)	$4.70 \cdot 10^{-7}$
Volume of the reaction chain for one complex (L)	$5.45 \cdot 10^{-20}$
Number of complexes in the nucleus	$1.59 \cdot 10^{5}$
Volume of the reaction chain for all complexes in the nucleus (L)	$8.68 \cdot 10^{-15}$
Volume nucleus (L)	$2.20 \cdot 10^{-13}$
Scaling factor	
= Volume nucleus/Volume of the reaction chain for all complexes in the nucleus	**25**

Binding of 5fTHF and 5mTHF to SHMT within the enzymatic complex

The model was developed taking into account that SHMT binds 5mTHF and 5fTHF when in the complex; both of these forms of folate inhibit SHMT[27]. Therefore, the model considered six different complex forms according to the binding state of the two SHMT monomers included in the complex (not shown in Figure 4.1, where a generic enzyme complex has been included to simplify the network). Figure 4.3 shows the six complex configurations: 1) both SHMT monomers are unliganded, 2) one SHMT is bound to 5mTHF, while the other is unliganded, 3) one SHMT is bound to 5fTHF, while the other is unliganded, 4) one SHMT monomer is bound to 5mTHF, while the other is bound to 5fTHF, 5) both SHMT monomers are bound to 5mTHF and 6) both SHMT monomers are bound to 5fTHF.

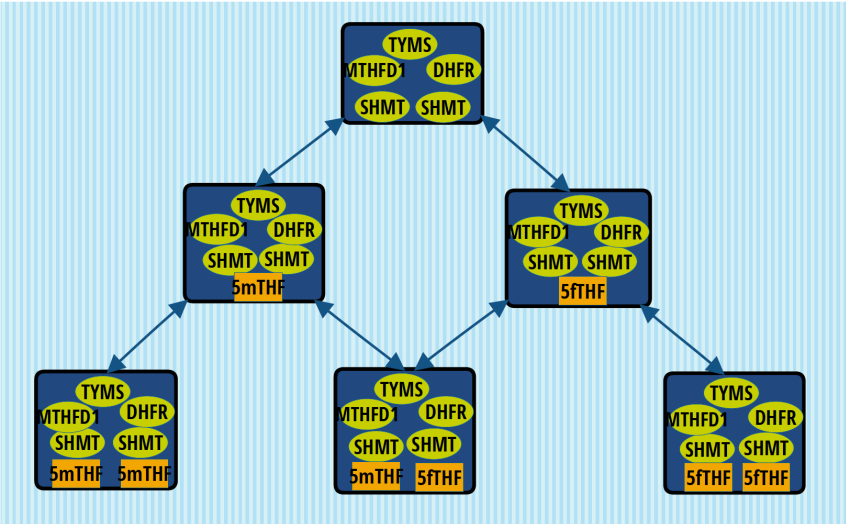

Figure 4.3. The six states of the nuclear enzymatic complex with respect to the binding of SHMT with 5mTHF and 5fTHF. Green cycles identify the enzymes of the complex, while orange rectangle boxes refer to the model variables 5mTHF and 5fTHF. The arrows between the complexes indicate the possibility to evolve from one state to the other.

Description of the computational environment and simulation procedure

The computational model has been defined as a set of ODEs implemented as a MATLAB function. All model simulations have been computed with the numerical ODE solver ode15s.

All simulations for each scenario of interest were carried out considering the following steps: 1) the cytosolic compartment was simulated until reaching steady state, 2) The translocation was carried out and initial settings for the cytosol and nuclear compartment were updated according to the rules described above, and 3) the complete model including the cytosol and the nucleus were simulated until reaching steady state.

In silico experiments

In the following a short technical description of the considered simulation scenarios is provided. If not indicated differently, we used the 25x scaling factor to produce the *in silico* experiments.

Required dTMP molecules for cell replication - Based on the assumption that 59% of the human genome consists AT base pairs[161], the number of dTMP molecules needed for the replication of the human genome can be computed as

$0.59 \cdot 3 \cdot 10^9 = 1.77 \cdot 10^9$.

Model-derived production of dTMP molecules for cell replication

The predicted number of produced dTMP molecules was derived by model simulation considering the final steady state of the system. The total number of dTMPs is given by the sum of the molecules produced in the cytosol and in the nucleus. For each one of the two compartments, the following formula has been used:

$$\#T = v_{TYMS} \cdot k_{vol} \cdot Sl \cdot N_A \cdot V\,ol,$$

where v_{TYMS} indicates the flux of the reaction catalyzed by TYMS in the corresponding compartment (cytosol or nucleus), $k_{vol} = 10^{-6} \frac{M}{\mu M}$ indicates a scaling factor used to transform the concentration from μm to m, Sl = 8h is the length of the S-phase, $N_A = 6.022 \cdot 10^{23}$ is the Avogadro constant and $V\,ol$ is the volume of the respective compartment (Table 4.2)

The role of enzymatic complex formation for sufficient dTMP production

The equation introduced above for estimating the total production of dTMP molecules was computed for four different simulation scenarios to test the contribution of nuclear enzymatic complex formation and the effect of the substrate channeling on acceleration of the pathway. In particular, the following four scenarios were considered: 1) restriction to the cytosolic model, 2) entire network, without multi-enzyme complex formation (all nuclear enzymes are in free from), 3) entire network, with multi-enzyme complex formation and 20-fold substrate channeling acceleration (literature-based approach) and 4) entire network, with complex formation and 25-fold substrate channeling acceleration (volume-based approach).

The effect of folate partitioning on dTMP production

The impact of folate partitioning between the two compartments was studied by varying the percentage of folate being translocated to the nuclear compartment. Starting from the standard scenario, in which 10% of cytosolic folate is translocated to the nucleus, the ratio of nuclear folate was increased up to 60% (considering the levels of 20%, 30%, 40% and 50%). *De novo* dTMP activity was measured for all the scenarios and expressed as number of predicted Ts synthesized, as well as percentage of dTMP synthesis rates required for genome replication. We further assessed with which folate partitioning the model would predict a sufficient dTMP activity for genome replication.

The role of nuclear MTHFS

The influence of the addition of MTHFS in the nuclear compartment was studied by comparing model steady states and the dTMP production in three different scenarios: 1) standard case with complex formation and 25-fold substrate channeling acceleration; 2) scenario 1, without MTHFS in the nucleus and 3) scenario 2 without the nuclear reaction R_{SHMT} : CHF → 5fTHF.

Sensitivity analysis of model enzymes

To study the impact of model enzymes on the network output variables a sensitivity analysis was carried out. The perturbation of the enzymes was calculated by a multiplicative scaling factor with the levels 0.25, 0.5, 1, 2 and 4. For each considered enzyme concentration we simulated the system to identify the steady state values of the following variables of interest: purine synthesis (measured as fluxes through PGT and AICARFT), thymidylate synthesis (measured as total dTMP production, and fluxes through TYMS in the cytosol and nucleus), and methionine synthesis (measured as flux through MTR). The influence of an enzyme on an

output variable was measured by the coefficient of variation (CV):

CV = mean · 100%, where a high value identifies those enzymes with a high influence on the output variable.

The contribution of NADPH and NADP to network dynamics

The contribution of NADPH and NADP to FOCM dynamics was studied by comparing model steady states of different scenarios corresponding, respectively, to the sum variation and ratio variation of these two molecules. In the model NADPH and NADP are considered to be constant over time (NADPH: 58µm, NADP: 18µm for both the cytosol and nucleus). To assess the role of NADPH and NADP two *in silico* experiments were carried out. First, to understand the influence of the overall availability of NADPH and NADP, four scenarios with respect to the sum of NADPH and NADP were considered, while their ratio stayed fixed as NADPH:NADP = 70:30. Their overall availability was varied by means of a scaling factor with the levels of 0.5, 1, 1.5, and 2, corresponding to concentrations levels of 38µm, 76µm, 114µm, and 152µm, respectively (Table 4.9). Second, the impact of the partitioning of NADPH and NADP was assessed by considering a step-wise change in the ratio NADPH/NADP, while the overall availability of the two substrates remained constant (76 µm). This resulted in eleven scenarios indicated in Table 4.10.

Table 4.9. The scenarios considered to measure the effect of the overall availability of NADPH and NADP for a fixed ratio of the two variables NADPH/NADP = 70:30. The sum of NADPH and NADP was varied according to the scaling factors: 0.5x, 1x, 1.5x, and 2x.

Scaling factor	NADPH+NADP	NADPH		NADP	
	(µM)	(% of sum)	(µM)	(% of sum)	(µM)
0.5x	38	70	26.6	30	11.4
1x	76	70	53.2	30	22.8
1.5x	114	70	79.8	30	34.2
2x	152	70	106.4	30	45.6

Table 4.10. The scenarios considered to measure the effect of the partitioning of NADPH and NADP, while their overall sum stayed fixed at 76µm. NADPH and NADP are presented as percentage of their sum, as well as the responding concentration used for model simulations in the cytosol and nucleus.

Scenario	NADPH		NADP	
	(% of sum)	(µM)	(% of sum)	(µM)
1	0	0.0	100	76.0
2	10	7.6	90	68.4
3	20	15.2	80	60.8
4	30	22.8	70	53.2
5	40	30.4	60	45.6
6	50	38.0	50	38.0
7	60	45.6	40	30.4
8	70	53.2	30	22.8
9	80	60.8	20	15.2
10	90	68.4	10	7.6
11	100	76.0	0	0.0

The contribution of glycine and serine to network dynamics

The contribution of serine and glycine to FOCM dynamics was studied following the same approach as for NADPH and NADP described in the previous paragraph. For this model steady states of different scenarios corresponding, respectively, to sum variation and ratio variation of glycine and serine were compared. In the model serine and glycine are considered to be constant over time (serine: 468µm, glycine: 1850µm for both the cytosol and nucleus). To assess the role of glycine and serine two *in silico* experiments were carried out. First, to understand the influence of the overall availability of glycine and serine, four scenarios with respect to the sum of serine and glycine were considered, while their ratio stayed fixed as glycine: serine = 80:20. Their overall availability was varied by means of a scaling factor with the levels 0.001, 0.1, 0.5, 1, 1.5, 2, 5, and 10, corresponding to concentration levels of 23.2µm, 231.8µm, 1159µm, 2318µm, 3477µm, 11590µm, and 23180µm, respectively (Table 4.11).

Second, the impact of the partitioning of glycine and serine was assessed by considering a step-wise change in the ratio glycine/serine, while the overall availability of the two substrates remained constant (2318 µm). This resulted in eleven scenarios indicated in Table 4.12.

Table 4.11. The scenarios considered to measure the effect of the overall availability of glycine and serine for a fixed ratio of the two variables glycine/serine = 80:20. The sum of glycine and serine was varied according to the scaling factors: 0.01x, 0.1x, 0.5x, 1x, 1.5x, 2x, 5x, and 10x.

Scaling factor	**Serine+Glycine**	**Serine**		**Glycine**	
	(µM)	(% of sum)	(µM)	(% of sum)	(µM)
0.01x	23.2	20	4.6	80	18.5
0.1x	231.8	20	46.4	80	185.4
0.5x	1159.0	20	231.8	80	927.2
1x	2318.0	20	463.6	80	1854.4
1.5x	3477.0	20	695.4	80	2781.6
2x	4636.0	20	927.2	80	3708.8
5x	11590.0	20	2318.0	80	9272.0
10x	23180.0	20	4636.0	80	18544.0

Table 4.12. The scenarios considered to measure the effect of the partitioning of glycine and serine, while their overall sum stayed fixed at 2318µm. Glycine and serine are presented as percentage of their sum, as well as the responding concentration used for model simulations in the cytosol and nucleus.

Scenario	**Serine**		**Glycine**	
	(% of sum)	(µM)	(% of sum)	(µM)
1	0	0.0	100	2318.0
2	10	231.8	90	2086.2
3	20	463.3	80	1854.4
4	30	695.4	70	1622.6
5	40	927.2	60	1390.8
6	50	1159.0	50	1159.0
7	60	1390.8	40	927.2
8	70	1622.6	30	695.4
9	80	1854.4	20	463.6
10	90	2086.2	10	231.8
11	100	2318.0	0	0.0

Cell culture

Human MCF7 mammary adenocarcinoma cells (ATCC, #HTB22) were passaged in minimal essential medium α-modification (Hyclone) supplemented with 10% FBS (Hyclone) and 1% penicillin/streptomycin (Corning). The MCF7GlyT cell line, which expresses murine glycine transporter GlyT1 and normalizes extracellular and intracellular glycine concentrations, was described previously[10].

Vector construction and transfection

The pCDNA-MTHFR-V5 and pCDNA-TYMS-V5 vectors were created by inserting the MTHFR and TYMS into the pCDNA 3.1/V5-His TOPO plasmid (Invitrogen) by directional TOPO cloning. Full-length MTHFR was amplified from cDNA (dharmacon, MHS6278-202833327) with the primer pair 5'-GCCACCatggtgaacgaagcc-3' and 3'-GtcatggagcctccgtttctctC-3'. Full-length TYMS was amplified from pCMV6-AC-TYMS-GFP described previously[81] with the primer pair 5'-GCCACCatgcctgtggccgg-3' and 3'-G gcgcctaaacagccatttccattttaat-3'. All DNA constructs used were verified by DNA sequencing. Cells were transfected by FuGene 6 transfection reagent (Promega) following the manufacturer's instructions.

Immunoblotting

Cellular proteins were extracted with 10 mM Tris·HCl (pH 7.4), 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 5 mM DTT, and 1:100 diluted Protease Inhibitor Mixture (Sigma). Total protein was quantified[162], resolved on 12% (vol/vol) SDS/PAGE gels (NuSep), and transferred to nitrocellulose membranes (Invitrogen). Membranes were blocked for 1 h at room temperature in blocking buffer (Rockland). Primary antibodies were diluted in blocking buffer and incubated overnight at 4°C. Secondary antibodies were diluted in 5% nonfat dry milk in

PBS with 1 % Nonidet P-40 (US Biologicals) and added to the membrane for 1 h at room temperature. MTHFR and TYMS were detected using a 1:1000 dilution of rabbit anti-MTHFR antibody (cell signaling) and a 1:1000 dilution of rabbit anti-TYMS antibody (cell signaling), respectively, followed by a 1:20000 dilution of IRDye 800CW-conjugated goat anti-rabbit IgG secondary antibody (LICOR Biosciences). As loading controls, 1:1000 mouse anti-β-Actin antibody (LICOR Biosciences) was used followed by a 1:20000 goat anti-mouse IgG secondary antibody (LICOR Biosciences). The membranes were then visualized on an Odyssey Infrared Imaging System (LICOR Biosciences).

dU suppression Assay

The contribution of the *de novo* and salvage pathways to dTMP synthesis was performed by exposing MCF7 cells to [^{14}C]-deoxyuridine and [^{3}H]-thymidine as described previously[10]. The ratio of ^{14}C/^{3}H DPM in nuclear DNA is a measure of relative capacity of *de novo* synthesis compared to salvage pathway dTMP synthesis. [^{14}C]-deoxyuridine is incorporated into DNA by TYMS in a folate-dependent manner, whereas [^{3}H]-thymidine is incorporated into DNA through the salvage pathway after its phosphorylation by Thymidine kinase 1 (TK1). Briefly, MCF7 cells were transfected with pCDNA-MTHFR-V5 or pCDNA-TYMS-V5 vector for 48 h. Cells were replated in triplicate in six-well plates (Corning) containing modified media supplemented with 20 μM methionine, 1 mg/L pyridoxine, 25 nM (6S)5-formylTHF, 500 nM [3H]-thymidine (Moravek), and 10 μM [14C]-deoxyuridine (Moravek). Cells were allowed to grow for two doublings and harvested. Genomic DNA was extracted with a DNA isolation kit with RNase A treatment (Roche), per manufacturer's instructions. Isotope levels were quantified using a Beckman LS6500 scintillation counter in dual disintegrations per minute mode. Experiments were replicated two times.

To determine the effect of serine, glycine and formate on *de novo* dTMP synthesis, MCF7 and MCF7GlyT2 cells were plated in the medium described above that was also supplemented with one of the followings in triplicate: 1) 20 μM methionine, 6 mM serine, 1.85 mM glycine, and 16 μM formate; 2) 20 μM methionine, 6 mM serine, 30 mM glycine, and 16 μM formate; 3) 20 μM methionine, 600 μM serine, 3 mM glycine, and 160 μM formate; 4) 200 μM methionine, 6 mM serine, 1.85 mM glycine, and 16 μM formate; 5) 200 μM methionine, 6 mM serine, 30 mM glycine, and 16 μM formate; 6) 200 μM methionine, 600 μM serine, 3 mM glycine, and 160 μM formate.

4.3 Results

Enzymatic complex formation is required for sufficient dTMP production

To test the impact of including the nuclear compartment and the contribution of nuclear enzymatic complex formation with substrate channeling on *de novo* dTMP synthesis, model steady states were compared across different experiments. In particular, the four following scenarios were considered: 1) restriction to the cytosolic model; 2) entire network (cytosol and nucleus), without multi-enzyme complex formation; 3) entire network, with nuclear multi-enzyme complex formation and 20-fold substrate channeling acceleration (literature-based approach); and 4) entire network, with complex formation and 25-fold substrate channeling acceleration (volume-based approach).

Table 4.13. Predicted number of produced Ts for different scenarios with respect to the compartmentalization, nuclear enzymatic complex formation and substrate channeling. Number of produced Ts was predicted during S-phase (8 hours) for the cytosol and nuclear compartment. The ratio of required activity for human genome replication (1.77 x 109) was calculated.

	Number of produced T in 8h			% of required T
	cytoplasm	nucleus	total	
only cytoplasm	$1.366 \cdot 10^9$	–	–	77.2
cytoplasm and nucleus without complex formation	$0.328 \cdot 10^9$	$0.047 \cdot 10^9$	$0.376 \cdot 10^9$	21.2
cytoplasm and nucleus with complex formation 20x scaling factor	$0.328 \cdot 10^9$	$0.948 \cdot 10^9$	$1.2768 \cdot 10^9$	72.1
cytoplasm and nucleus with complex formation 25x scaling factor	$0.328 \cdot 10^9$	$1.1858 \cdot 10^9$	$1.5138 \cdot 10^9$	85.5

The model that considered only the cytosol leads to a dTMP production rate which produces 77% of required dTMP synthesis for genome replication in a cell cycle (Table 4.13). However, based on the experimental evidence that the *de novo* dTMP pathway must localize to the nucleus to prevent uracil accumulation in DNA, this scenario can be considered as an artificial, fragmentary one. When considering the *in silico* experiments for the complete model that includes the translocation of the pathway to the nucleus without multienzyme complex formation, cytosolic dTMP rates decrease by more than 75% as a response to the decreasing enzyme and substrate availability in the compartment (Table 4.13). Nuclear dTMP synthesis rates cannot compensate for this decline and only 21% of needed dTMP is synthesized (Table 4.13), indicating that nuclear localization of the *de novo* dTMP synthesis pathway alone is not sufficient to accelerate dTMP synthesis.

When metabolic channeling and multienzyme complex formation is considered, the nuclear dTMP synthesis pathway displays an enhanced activity and the number of produced Ts increased up to 25-fold (Table 4.13, row 3 and 4). While the model with the channeling

acceleration factor of 20-fold predicts a dTMP rate which covers 72% of required dTMP synthesis, the model with the scaling factor based on the volume of the sphere (25-fold) captures 85.5% of required synthesis, thus exceeding the results of the model restricted to the cytosol.

Our observations indicate that *de novo* dTMP synthesis functions mostly in the nucleus and when the respective enzymes SHMT, TYMS, DHFR and MTHFD1 form a multienzyme complex. Accounting for the kinetic effects of nuclear multienzyme complex formation on *de novo* dTMP synthesis is therefore essential to support genome stability.

The gap between predicted (85%) and required dTMP activity can be explained by two considerations: 1) The nuclear FOCM network is quite sensitive to folate availability and only a slight increase in nuclear folate levels rises *de novo* dTMP synthesis levels remarkably (see below, Table 4.14); 2) In addition to the *de novo* dTMP synthesis pathway, some cells also contain the salvage dTMP synthesis pathway catalyzed by TK1, which recovers dTMP from thymidine and contributes to the overall needed dTMP activity. However, this process is not yet included in the model and is highly dependent upon exogenous sources of thymidine.

Table 4.14. Predicted number of produced Ts in response to the partitioning of folate between the nuclear and cytosol compartments. Number of produced Ts was predicted during S-phase (8 hours) for the cytosol and nuclear compartment. The ratio of required activity for human genome replication (1.77 x 109) was calculated.

Partitioning of folate nucleus/cytoplasm (%)	**Number of Ts produced in 8h**			**% of required T**		
	cytoplasm	**nucleus**	**total**	**cytoplasm**	**nucleus**	**total**
10/90	$3.28 \cdot 10^8$	$1.19 \cdot 10^9$	$1.51 \cdot 10^9$	18.54	66.96	85.50
11.6/88.4	$3.24 \cdot 10^8$	$1.45 \cdot 10^9$	$1.78 \cdot 10^9$	18.33	82.05	100.38
20/80	$3.02 \cdot 10^8$	$3.02 \cdot 10^9$	$3.32 \cdot 10^9$	17.06	170.48	187.54
30/70	$2.68 \cdot 10^8$	$5.01 \cdot 10^9$	$5.28 \cdot 10^9$	15.14	283.29	298.43
40/60	$2.27 \cdot 10^8$	$7.01 \cdot 10^9$	$7.24 \cdot 10^9$	12.83	396.17	409.00
50/50	$1.81 \cdot 10^8$	$8.97 \cdot 10^9$	$9.15 \cdot 10^9$	10.24	506.52	516.76
60/40	$1.33 \cdot 10^8$	$1.09 \cdot 10^{10}$	$1.10 \cdot 10^{10}$	7.54	613.59	621.13

The effect of folate partitioning on dTMP production

To assess the effect of folate partitioning between the two compartments on *de novo* dTMP synthesis, we evaluated seven scenarios. In particular, the percentage of folate translocated to the nuclear compartment was step-wise increased from 10 % (standard scenario) to 60%.

While cytosolic dTMP production rates decrease moderately in response to declining folate availability in the respective compartment, nuclear dTMP synthesis increases notably (Table 4.14). Throughout the considered folate partitioning range, the cytosolic proportion for dTMP synthesis diminishes by 60%, whereas nuclear production rate increases almost 10-fold.

We further assessed which folate partitioning allows for the model to meet *de novo* dTMP synthesis rates sufficient for genome replication. Interestingly, translocation of 11.6% of cytosolic folate, which is only 1.6% more than in the standard scenario, to the nucleus increases the flux through nuclear TYMS enough to synthesize adequate levels of dTMP. These results show that the network exhibits high sensitivity with respect to nuclear folate levels and further indicate the importance of the nuclear dTMP pathway as main contributor of overall cellular *de*

novo dTMP capacity.

The role of nuclear MTHFS activity

The impact of nuclear MTHFS activity on model steady states and dTMP production rates for three different scenarios were compared: 1) standard model with normal MTHFS activity; 2) no nuclear MTHFS activity and 3) no nuclear MTHFS activity and *SHMT* catalyzed activity $CHF \rightarrow 5f\,THF$ activity turned off. The metabolic role of 5fTHF is not yet elucidated. It does not serve as a cofactor in folate-dependent biosynthetic reactions, but rather has been proposed to serve as an intracellular storage form of folate[19], and as an inhibitor of the folate dependent enzymes SHMT and AICARFT[19]. The model demonstrates that nuclear MTHFS activity is necessary to prevent the accumulation of nuclear folate as 5fTHF. This result is consistent with the observations made for cytosolic MTHFS activity described elsewhere[156]. When the nuclear compartment was modeled without MTHFS availability, nuclear folate pools accumulated as 5fTHF (90% of total nuclear folate, Table 4.15), while the remaining 10% of nuclear folate was 5mTHF (the stoichiometry of the network implies no nuclear MTR nor MTHFR activity, therefore initial 5mTHF levels remain unchanged at the predefined availability of 10% throughout all simulation scenarios (Figure 4.1). As a consequence, no folate co-factors were available for nuclear dTMP synthesis in the absence of MTHFS activity and the overall production rate decreases to 18.5% of required TYMS activity (Table 4.15). The lethal pooling of nuclear folate as 5fTHF can be prevented, if in addition to the deletion of nuclear MTHFS activity the SHMT-catalyzed conversion of CHF to 5fTHF is also turned off (Table 4. 15).

Table 4.15. Predicted number of produced Ts during S-phase (8 hours), the ratio of required dTMP synthesis, and 5fTHF levels (as percentage of total nuclear folate) for different scenarios with respect to nuclear MTHFS and SHMT activity.

	Number of produced Ts in 8h			% of required T	% of nuclear 5fTHF
	cytoplasm	nucleus	total		
Standard model	$0.328 \cdot 10^9$	$1.185 \cdot 10^9$	$1.513 \cdot 10^9$	85.5	35.3
No nuclear MTHFS	$0.328 \cdot 10^9$	0	$0.328 \cdot 10^9$	18.5	90.0
No nuclear MTHFS and SHMT: CHF → 5fTHF activity turned off	$0.328 \cdot 10^9$	$1.190 \cdot 10^9$	$1.518 \cdot 10^9$	85.8	35.0

Sensitivity analysis of FOCM enzymes

We performed various computational experiments to summarize the dynamic effects on FOCM resulting from changes in individual model enzyme activity. For each enzyme, five simulations were carried out to perturb the enzyme activity according to a multiplicative scaling factor, and the impact of this variation was measured across the simulations by means of the coefficient of variation. A summary of the results, indicating the impact of each enzyme on network output variables of interest, is presented in Figure 4.4. The model predictions for all executed simulations are in the Table 4.16.

The steady state activity of purine synthesis depends highly on the magnitude of the two enzymes AICARFT and PGT, which is not unexpected, since AICARFT and PGT are the enzymes directly involved in *de novo* purine biosynthesis. Interestingly, changes in AICARFT and PGT also affected HCY remethylation, as measured by flux through MTR, and *de novo* dTMP biosynthesis. *De novo* dTMP was sensitive to changes in levels of purine biosynthesis enzymes, with TYMS and DHFR showing with the greatest sensitivity, followed by MTR, which exhibited the third highest influence. Indeed, TYMS and DHFR activity exhibit high values of variation for all model output variables of interest, including the relevant fluxes for

purine synthesis, MTR activity and dTMP synthesis (Figure 4.4).

	V_AICARFT	V_PGT	V_MTR	#Ts total	V_TYMS cytoplasm	V_TYMS nucleus
AICARFT	98.43	0.11	0.09	0.05	0.21	0.00
BHMT	1.86	0.88	2.82	1.30	6.10	0.00
DHFR	55.90	55.90	55.90	24.36	55.90	16.85
DNMT	0.27	0.10	0.33	0.16	0.73	0.00
GNMT	0.12	0.04	0.14	0.07	0.31	0.00
MAT-I	0.29	0.11	0.35	0.17	0.80	0.00
MAT-III	0.10	0.04	0.12	0.06	0.27	0.00
MTHFD1	1.20	9.74	19.99	6.11	31.20	0.00
MTHFR	11.71	23.96	44.96	12.49	67.43	0.00
MTHFS	11.89	0.26	0.73	11,00	1.64	13.64
MTR	11.71	23.92	47.85	12.47	67.35	0.00
PGT	0.06	98.28	0.06	0,03	0.13	0.00
SAHH	1.62	0.76	2.43	1.12	5.27	0.00
SHMT	11.01	9.55	31.68	8.27	44.31	0.00
TYMS	55.92	55.90	55.91	51.48	106.81	49.07

A

V_AICARFT	V_PGT	V_MTR	#Ts total	V_TYMS cytoplasm	V_TYMS nucleus
AICARFT	PGT	TYMS	TYMS	TYMS	TYMS
TYMS	TYMS	DHFR	DHFR	MTHFR	DHFR
DHFR	DHFR	MTR	MTHFR	MTR	MTHFS
MTHFS	MTHFR	MTHFR	MTR	DHFR	MTHFD1
MTR	MTR	SHMT	MTHFS	SHMT	PGT
MTHFR	MTHFD1	MTHFD1	SHMT	MTHFD1	MTR
SHMT	SHMT	BHMT	MTHFD1	BHMT	BHMT
BHMT	BHMT	SAHH	BHMT	SAHH	MTHFR
SAHH	SAHH	MTHFS	SAHH	MTHFS	DNMT
MTHFD1	MTHFS	MAT-I	MAT-I	MAT-I	SHMT
MAT-I	MAT-I	DNMT	DNMT	DNMT	SAHH
DNMT	AICARFT	GNMT	GNMT	GNMT	MAT-III
GNMT	DNMT	MAT-III	MAT-III	MAT-III	MAT-I
MAT-III	GNMT	AICARFT	AICARFT	AICARFT	GNMT
PGT	MAT-III	PGT	PGT	PGT	AICARFT

B

Figure 4.4. The impact of model enzymes on FOCM output. The variables of interest were purine synthesis: flux through PGT and AICARFT, methionine synthesis: flux through MTR and thymidylate synthesis: total dTMP synthesis by means of predicted number of Ts and fluxes through TYMS in the cytosol and nucleus). A) Coefficient of variation across enzyme levels for each enzyme- output combination. Colors encode the ranking of these values: red identifies the smallest CV relating to the enzyme with the least influence on the output variable, whereas green identifies the highest changes. B) Ranking of the enzyme influence on the output variables. For each variable of interest, the enzymes are listed according to their influence on this variable, starting from the one with the highest influence. Colors refer to the values in subfigure A.

Table 4.16. The impact of variation in model enzymes activity on network output variables. Enzyme levels were modified according to a scaling factor (0.5x, 1x, 2x, and 4x) and for each enzyme the resulting model predictions for the flux through the enzymes AICARFT, PGT, MTR, TYMS as well as the total number of produced Ts are provided. The variation across scenarios was assessed using the coefficient of variation (CV).

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
AICARFT						
0.25x	1491.5	5249.3	22.6	$1.51 \cdot 10^9$	72.5	1118.2
0.5x	2983.5	5248.8	22.6	$1.51 \cdot 10^9$	72.5	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	11943.1	5245.2	22.6	$1.51 \cdot 10^9$	72.5	1118.2
4x	23901.3	5235.8	22.6	$1.51 \cdot 10^9$	72.1	1118.2
CV (%)	98.43	0.11	0.09	0.05	0.21	0.00
BHMT						
0.25x	5913.7	5264.1	22.8	$1.52 \cdot 10^9$	74.2	1118.2
0.5x	5929.6	5259.5	22.8	$1.52 \cdot 10^9$	73.7	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	6050.6	5218.9	22.2	$1.50 \cdot 10^9$	69.6	1118.2
4x	6186.3	5152.1	21.3	$1.47 \cdot 10^9$	63.7	1118.3
CV (%)	1.86	0.88	2.82	1.30	6.10	0.00
DHFR						
0.25x	0.0	0.0	0.0	$0.77 \cdot 10^9$	0.0	726.5
0.5x	5966.7	5246.9	22.6	$1.51 \cdot 10^9$	72.4	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5968.5	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
4x	5968.5	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.2
CV (%)	55.90	55.90	55.90	24.36	55.90	16.85
DNMT						
0.25x	5997.9	5238.1	22.5	$1.51 \cdot 10^9$	71.5	1118.2
0.5x	5982.3	5243.3	22.5	$1.51 \cdot 10^9$	72.0	1118.3
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5961.6	5249.9	22.6	$1.51 \cdot 10^9$	72.7	1118.3
4x	5959.1	5250.7	22.7	$1.51 \cdot 10^9$	72.8	1118.2
CV (%)	0.27	0.1	0.33	0.16	0.73	0.00

to continue on next page ...

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
GNMT						
0.25x	5978.2	5244.7	22.6	$1.51{\cdot}10^9$	72.2	1118.3
0.5x	5973.4	5246.2	22.6	$1.51{\cdot}10^9$	72.3	1118.3
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5964.1	5249.1	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5961.0	5250.1	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.12	0.04	0.14	0.07	0.31	0.00
MAT-I						
0.25x	6003.5	5236.2	22.4	$1.51{\cdot}10^9$	71.3	1118.3
0.5x	5980.2	5244.0	22.6	$1.51{\cdot}10^9$	72.1	1118.3
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5963.4	5249.4	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5961.1	5250.1	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.29	0.11	0.35	0.17	0.80	0.00
MAT-III						
0.25x	5977.7	5244.8	22.6	$1.51{\cdot}10^9$	72.2	1118.3
0.5x	5973.0	5246.3	22.6	$1.51{\cdot}10^9$	72.3	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5965.0	5248.9	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5962.8	5249.6	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.10	0.04	0.12	0.06	0.27	0.00
MTHFD1						
0.25x	5831.6	4150.1	13.3	$1.31{\cdot}10^9$	28.2	1118.2
0.5x	5928.8	5255.6	22.5	$1.51{\cdot}10^9$	71.4	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5994.8	5239.4	22.7	$1.52{\cdot}10^9$	73.0	1118.2
4x	6008.8	5234.5	22.7	$1.52{\cdot}10^9$	73.3	1118.3
CV (%)	1.2	9.74	19.99	6.11	31.2	0
MTHFR						
0.25x	4749.5	5432.2	6.3	$1.62{\cdot}10^9$	96.1	1118.2
0.5x	4860.4	5422.3	12.6	$1.61{\cdot}10^9$	94.6	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3

to continue on next page ...

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
2x	6113.6	4151.5	24.0	$1.29 \cdot 10^9$	24.2	1118.2
4x	5196.7	2886.8	24.2	$1.23 \cdot 10^9$	10.0	1118.2
CV (%)	11.71	23.96	44.96	12.49	67.43	0.00
MTHFS						
0.25x	4801.0	5222.5	22.3	$1.26 \cdot 10^9$	70.3	886.8
0.5x	5458.6	5238.4	22.5	$1.40 \cdot 10^9$	71.7	1012.8
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	6309.5	5253.0	22.7	$1.60 \cdot 10^9$	72.9	1200.5
4x	6512.7	5255.8	22.7	$1.67 \cdot 10^9$	73.2	1261.0
CV (%)	11.89	0.26	0.73	11.00	1.64	13.64
MTR						
0.25x	5200.0	2890.6	6.1	$1.23 \cdot 10^9$	10.1	1118.2
0.5x	6115.6	4155.1	12.0	$1.29 \cdot 10^9$	24.2	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	4860.9	5422.3	25.2	$1.61 \cdot 10^9$	94.5	1118.2
4x	4749.7	5432.2	25.4	$1.62 \cdot 10^9$	96.1	1118.3
CV (%)	11.71	23.92	47.85	12.47	67.35	0.00
PGT						
0.25x	5966.4	1312.3	22.6	$1.51 \cdot 10^9$	72.5	1118.2
0.5x	5967.1	2624.4	22.6	$1.51 \cdot 10^9$	72.5	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5971.2	10491.1	22.6	$1.51 \cdot 10^9$	72.5	1118.3
4x	5975.4	20953.5	22.6	$1.51 \cdot 10^9$	72.3	1118.2
CV (%)	0.06	98.28	0.06	0.03	0.13	0.00
SAHH						
0.25x	6169.5	5162.4	21.4	$1.48 \cdot 10^9$	64.5	1118.2
0.5x	6031.3	5226.3	22.3	$1.50 \cdot 10^9$	70.3	1118.3
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5943.6	5255.4	22.7	$1.52 \cdot 10^9$	73.3	1118.3
4x	5934.6	5258.1	22.8	$1.52 \cdot 10^9$	73.5	1118.2
CV (%)	1.62	0.76	2.43	1.12	5.27	0
SHMT						

to continue on next page ...

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
0.25x	6658.1	5287.2	23.6	$1.55 \cdot 10^9$	79.7	1118.3
0.5x	6439.4	5277.7	23.3	$1.54 \cdot 10^9$	77.8	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5283.1	5098.1	19.7	$1.43 \cdot 10^9$	54.2	1118.2
4x	5232.0	4170.1	8.8	$1.26 \cdot 10^9$	16.0	1118.2
CV (%)	11.01	9.55	31.68	8.27	44.31	0.00
TYMS						
0.25x	5884.2	5239.1	22.4	$0.47 \cdot 10^9$	17.8	365.6
0.5x	5912.0	5242.1	22.5	$0.88 \cdot 10^9$	35.8	674.1
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	6082.4	5257.2	22.8	$2.26 \cdot 10^9$	147.8	1497.0
4x	0.0	0.0	0.0	$1.61 \cdot 10^9$	0.0	1516.2
CV (%)	55.92	55.90	55.91	51.48	106.81	49.07

Changing TYMS levels affects the variables of interest in a linear fashion, with the most pronounced effect in the scenario in which TYMS is upregulated by a 4-fold scaling factor (Table 4.16). On the other hand, the high values of variation of DHFR is solely a consequence of the most extreme scenario considering the downregulation of DHFR by a factor of 0.25, while for the remaining DHFR levels, no variation in the variables of interest can be observed (Table 4.16). The two extreme scenarios, 0.25×DHFR and 4×TYMS respectively, lead to an accumulation of DHF, which in turn leaves no folate co-factors available for *de novo* dTMP synthesis, purine synthesis or methionine synthase (Table 4.16).

Nuclear dTMP synthesis remains quite stable throughout all considered scenarios (Table 4.16 and Figure 4.4) as only changes in TYMS, DHFR or MTHFS activity affect the nuclear dTMP biosynthesis. While the impact of TYMS levels on dTMP production is obvious, the simulations suggest that 5fTHF accumulation as a result of decreasing MTHFS levels lead in

turn to a modest decrease of dTMP synthesis, because less folate co-factor is available (Table 4.16). DHFR levels influence dTMP synthesis only in the restricting scenario, when DHFR levels are downregulated by the factor of 0.25 (Table 4.16). In this scenario DHFR becomes the limiting enzyme in the cytosol and reduces the nuclear enzyme complexes by 50% and in turn nuclear dTMP synthesis declines.

The effect of the enzymes of the HCY remethylation and transmethylation pathways (MAT-I, MAT-III, GNMT, DNMT, BHMT, and SAHH) on network pathways is negligible (Figure 4.4). Indeed, only BHMT and SAHH slightly influence cytosolic dTMP activity due to diminishing HCY levels as response to downregulated SAHH levels and upregulated BHMT levels. In turn, MTR reaction flux decreases allowing an accumulation of 5mTHF, which then leaves less folate available as a co-factor for cytosolic *de novo* dTMP synthesis.

Overall, the model suggests that there is substantial robustness in the FOCM network towards smaller enzymatic variations and identifies DHFR and TYMS as the main drivers for changes in overall network dynamics.

The contribution of NADPH and NADP to network dynamics

To study the effect of changes in NADPH and NADP availability on FOCM dynamics, model steady states of different scenarios corresponding to sum and ratio variation of NADPH and NADP were considered.

Increasing overall availability of NADPH and NADP, while considering a fixed ratio (NADPH = 70%, NADP = 30%), leads to a shift in the steady state distribution of cytosolic folate with an accumulation of folate as 5mTHF (Table 4.17). This comes at the expense of the remaining folate co-factors, including 5fTHF and 10fTHF, with the exception of THF, which decreases only moderately (Figure 4.5). 10fTHF is the substrate of the purine synthesis

reactions catalyzed by AICARFT and PGT. The simulations indicate that flux through the PGT-catalyzed reaction declines as response to the lack of substrate, whereas the flux through the second purine-relevant enzyme AICARFT exhibits the opposing behavior (Figure 4.5). A possible explanation for this is the depletion of 5fTHF levels, following the increase of NADPH + NADP (Figure 4.5). 5fTHF is a known inhibitor of AICARFT[116] and therefore lower NADPH + NADP concentration levels increase the flux through the AICARFT-catalyzed reaction (Figure 4.5). These data suggest that the 5fTHF futile cycle and its link to purine synthesis could play an important role in controlling purine synthesis rates and in compensating for and/or preventing cellular and energetic consequences of 10fTHF depletion.

Table 4.17. Steady state concentrations of model variables (in µm) for different levels of NADPH+NADP (ranging from 0.5x to 2x of the standard concentration of 76 µm).

cytoplasm									
NADPH+ NADP	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.5x	0.058	9.334	1.852	0.444	0.004	0.551	0.648	1.200	2.820
1x	0.054	7.567	1.504	0.406	0.003	3.209	2.669	0.562	0.935
1.5x	0.046	4.643	0.925	0.264	0.002	7.164	3.605	0.130	0.131
2x	0.043	3.398	0.677	0.198	0.002	8.767	3.718	0.059	0.050
	SHMT	**HCY**	**MET**	**SAM**	**SAH**				
0.5x	0.470	3.725	42.631	33.416	70.748				
1x	0.333	3.331	39.863	67.433	39.893				
1.5x	0.201	3.125	36.945	80.901	29.549				
2x	0.170	3.075	36.224	83.612	27.610				
nucleus									
NADPH+ NADP	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.5x	0.033	3.558	0.696	0.079	0.003	0.550	0.252	1.489	1.476
1x	0.042	3.548	0.690	0.100	0.003	0.550	0.253	1.479	1.472
1.5x	0.047	3.543	0.687	0.111	0.003	0.549	0.253	1.473	1.470
2x	0.049	3.540	0.685	0.118	0.003	0.549	0.253	1.470	1.470

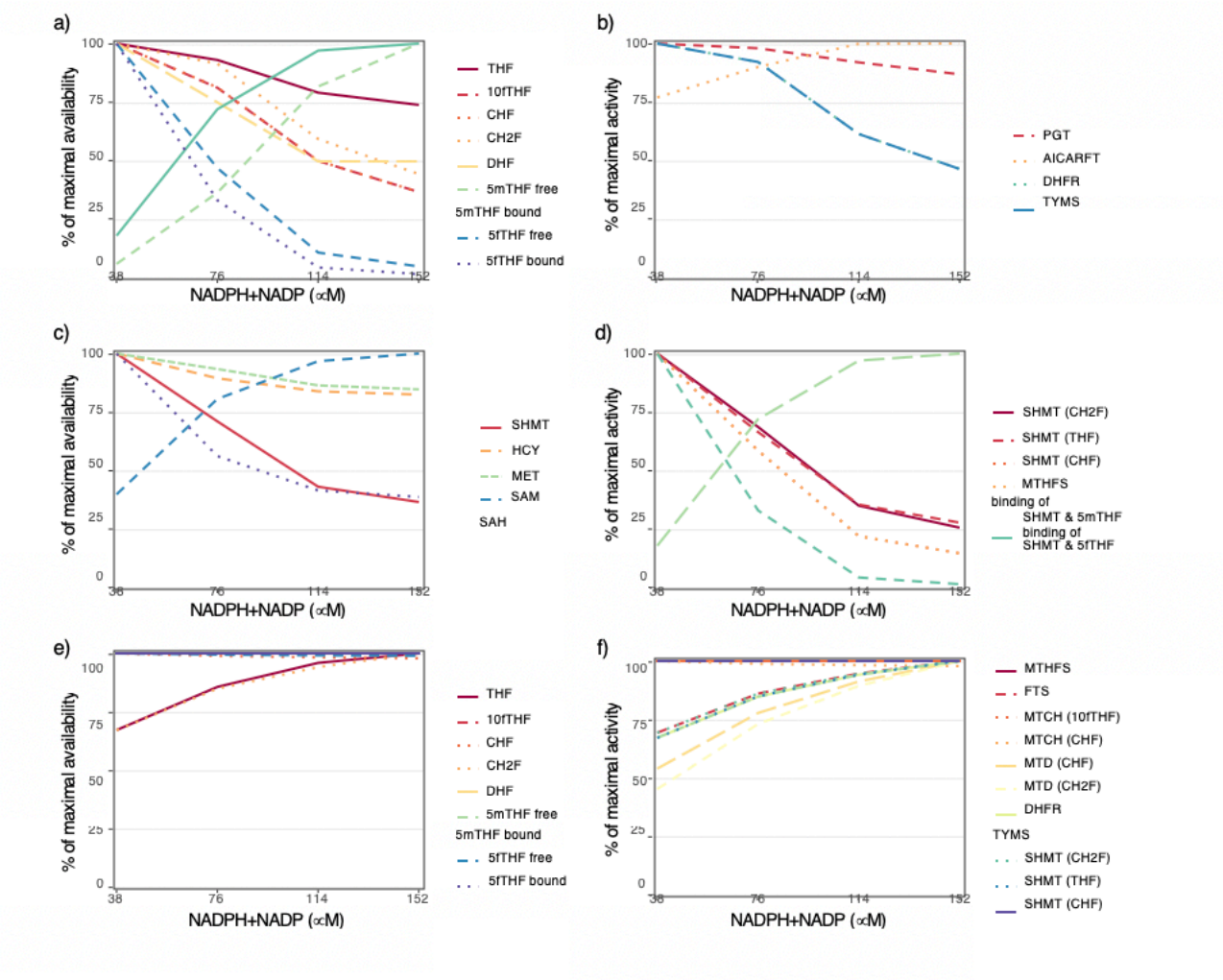

Figure 4.5. Effect of NADPH and NADP availability on model variables and fluxes. The provided values are computed by considering model steady states achieved by varying the total availability of NADPH+NADP from 38µm to 152 µm, according to Table 4.9. The results are expressed in terms of percentage of the maximal activity/availability over the considered range. Panel a) to d) display the results for cytosolic variables and fluxes, while panel e) and f) display those for the nuclear compartment. Reactions are indicated by the catalyzing enzyme, and for bidirectional reactions the direction is indicated behind the enzyme name: Enzyme (Substrate).

Even though 5fTHF depletion reduces the amount of SHMT bound to 5fTHF, fluxes through SHMT-catalyzed reactions do not increase (Figure 4.5C and d). This is a consequence of increasing 5mTHF levels, which also serves to inhibit SHMT. Therefore, increasing 5mTHF

compensates for the loss of SHMT inhibition by 5fTHF by increasing the ratio of SHMT bound to 5mTHF from 16% to 94% of overall SHMT (Figure 4.5a, c, d; Table 4.17).

The activity of *de novo* dTMP synthesis depends on the substrate CH2F of the reaction catalyzed by TYMS. When NADPH and NADP levels are increased, cytosolic CH2F levels decrease in response to accumulation of folate as 5mTHF and therefore cytosolic dTMP activity declines by more than 50% throughout the considered scenarios (Figure 4.5b; Table 4.18). However, nuclear CH2F levels increase with increasing NADPH and NADP levels, because absence of MTHFR from the nucleus averts 5mTHF accumulation in this compartment and because the flux through MTD (CHF → *CH2F*) accelerates (Figure 4.5e, f). Consequently, more folate cofactor is available for nuclear *de novo* dTMP production (Table 4.18; Figure 4.5e). The model suggests that the overall activity of dTMP synthesis is quite robust towards increases in NADPH and NADP, because nuclear productivity increases to balance the decline in cytosolic TYMS activity (Table 4.18, rows 2 to 4). Lower level of NADPH and NADP affects total dTMP synthesis and simulations suggest that the ratio of required dTMP molecules reduces from around 80% to 70% (Table 4.18, rows 1 and 2).

Table 4.18. Steady state activity of dTMP synthesis in response to the variation in cytosolic and nuclear overall availability of NADPH and NADP. The sum of NADPH and NADP was step-wise increased considering a scaling factor with the levels 0.5x, 1x, 2x and 4x, while the ratio NADPH:NADP stayed constant at 70:30. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication (1.77·109).

NADPH+ NADP	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
0.5x	83.3	815.1	$3.77{\cdot}10^{8}$	$8.64{\cdot}10^{8}$	$1.24{\cdot}10^{9}$	21.31	48.81	70.12
1x	76.8	1030.7	$3.48{\cdot}10^{8}$	$1.09{\cdot}10^{9}$	$1.44{\cdot}10^{9}$	19.65	61.72	81.37
1.5x	51.4	1144.6	$2.33{\cdot}10^{8}$	$1.21{\cdot}10^{9}$	$1.45{\cdot}10^{9}$	13.15	68.54	81.69
2x	39.1	1216.0	$1.77{\cdot}10^{8}$	$1.29{\cdot}10^{9}$	$1.47{\cdot}10^{9}$	10.00	72.81	82.82

Table 4.19. Steady state activity of dTMP synthesis in response to the variation in partitioning of NADPH and NADP. The partitioning of NADPH and NADP was modeled by step-wise increasing the percentage of NADPH from 0% to 100% (Scenario 1 to 11, Table 4.10), while overall availability of NADPH and NADP stayed fixed at 76µm. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication (1.77·109)

Scenario	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
1	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00
2	41.8	242.7	$1.89{\cdot}10^{8}$	$2.57{\cdot}10^{8}$	$4.47{\cdot}10^{8}$	10.69	14.53	25.23
3	60.1	427.5	$2.72{\cdot}10^{8}$	$4.53{\cdot}10^{8}$	$7.25{\cdot}10^{8}$	15.38	25.60	40.98
4	70.2	575.5	$3.18{\cdot}10^{8}$	$6.10{\cdot}10^{8}$	$9.28{\cdot}10^{8}$	17.96	34.46	52.42
5	76.2	699.8	$3.45{\cdot}10^{8}$	$7.42{\cdot}10^{8}$	$1.09{\cdot}10^{9}$	19.50	41.90	61.40
6	79.5	810.4	$3.60{\cdot}10^{8}$	$8.59{\cdot}10^{8}$	$1.22{\cdot}10^{9}$	20.34	48.53	68.87
7	80.2	916.4	$3.63{\cdot}10^{8}$	$9.71{\cdot}10^{8}$	$1.33{\cdot}10^{9}$	20.52	54.87	75.39
8	76.8	1030.7	$3.48{\cdot}10^{8}$	$1.09{\cdot}10^{9}$	$1.44{\cdot}10^{9}$	19.65	61.72	81.37
9	69.6	1181.2	$3.15{\cdot}10^{8}$	$1.25{\cdot}10^{9}$	$1.57{\cdot}10^{9}$	17.81	70.73	88.54
10	62.1	1472.7	$2.81{\cdot}10^{8}$	$1.56{\cdot}10^{9}$	$1.84{\cdot}10^{9}$	15.89	88.18	104.07
11	57.5	3445.2	$2.60{\cdot}10^{8}$	$3.65{\cdot}10^{9}$	$3.91{\cdot}10^{9}$	14.71	206.30	221.01

Furthermore, we carried out a second set of *in silico* experiments, in which overall availability of NADPH and NADP has been considered constant and the ratio of NADPH and NADP has been varied according to Table 4.10. The model shows that NADPH is required to prevent impairment of dTMP synthesis, while nuclear dTMP activity is enhanced with increasing NADPH availability. Cytosolic dTMP synthesis doesn't show a linear trend and reaches its maximum rate with a NADPH:NADP ratio of 60:40 (Table 4.19). Moreover, for all considered scenarios, nuclear dTMP synthesis remains the main contributor to overall dTMP activity. The depletion of NADPH levels leads to an interconvertible pooling of folate as DHF and consequently stops dTMP synthesis (Table 4.19, scenario 1). On the other hand, depletion of NADP increases nuclear dTMP synthesis notably (Table 4.19, scenarios 1 and 11). This is a consequence of the accumulation of CH2F due to the loss of R_{MTD} : CH2F $\rightarrow$ CHF activity .

The contribution of glycine and serine to network dynamics

To study the effect of changes in serine and glycine availability on FOCM dynamics, model steady states of different scenarios corresponding to sum and ratio variations were compared. Changing the overall availability of serine and glycine does not lead to significant effects on any cytosolic folate-dependent biosynthetic nucleotide synthesis pathways (Figure 4.6, AICARFT, PGT, and TYMS). Cytosolic dTMP activity remains mostly unaffected throughout the considered scenarios, because of quite constant steady state levels of CH2F (Figure 4.6b; Table 4.20). The main effect of changing total serine+glycine levels can be observed in the model dynamics of SHMT-catalyzed reactions (Figure 4.6c, d). Not unexpectedly, flux through the glycine/serine-dependent bidirectional reaction R_{SHMT} : THF $\leftrightarrow$ CH2F increases with increasing availability of glycine and serine (Figure 4.6c).

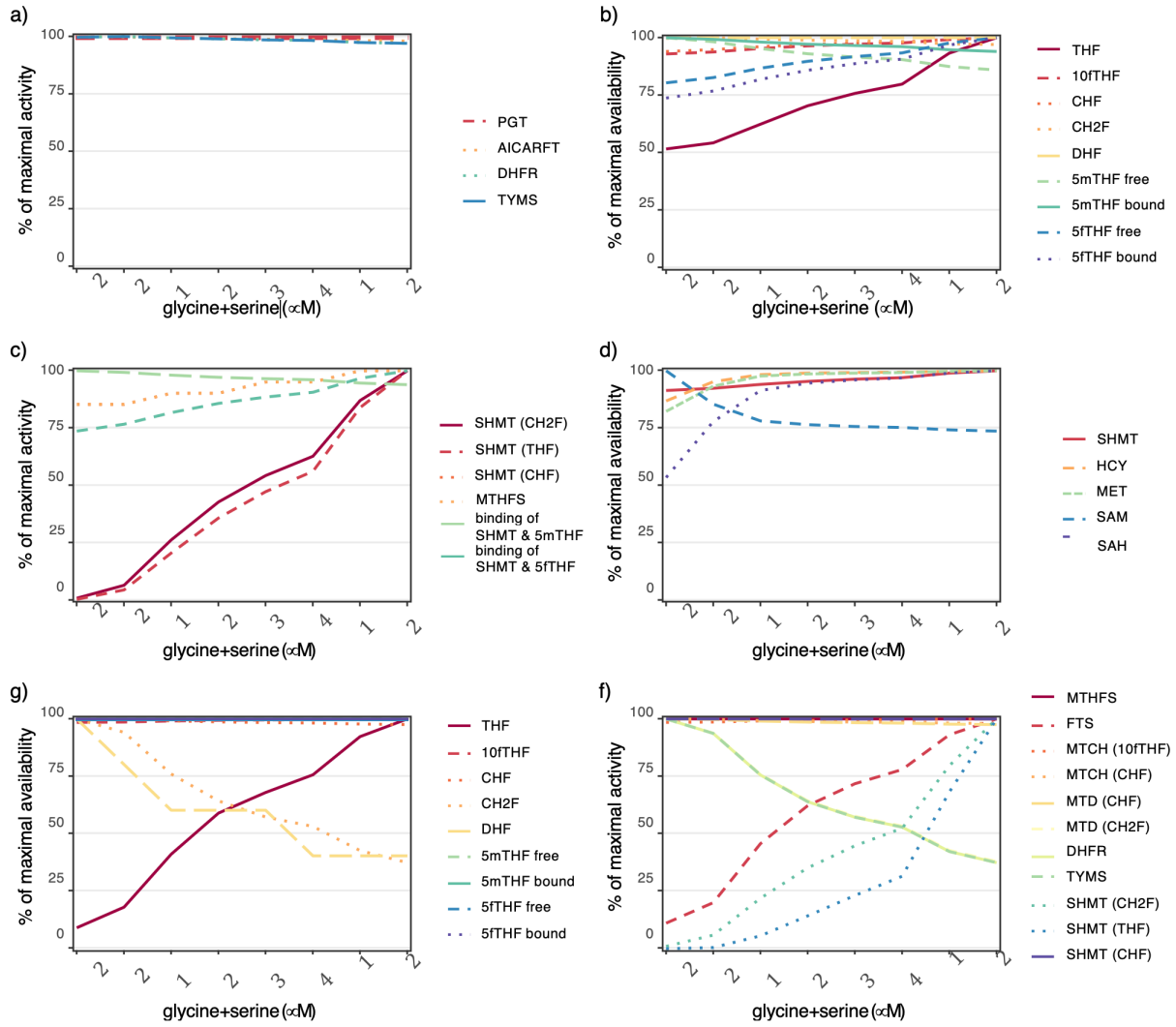

Figure 4.6. Effect of glycine and serine availability on model variables and fluxes. The provided values are computed by considering model steady states achieved by varying the total availability of glycine+serine from 23µm to 23180 µm, according to Table 4.11. They are expressed in terms of percentage of the maximal activity/availability over the considered range. Panel a) to d) display the results for cytosolic variables and fluxes, while panel e) and f) display those for the nuclear compartment. Reactions are indicated by the catalyzing enzyme, and for bidirectional reactions the direction is indicated behind the enzyme name: Enzyme (Substrate).

Table 4.20. Steady state concentrations of model variables (in µm) for different levels of glycine+serine (ranging from 0.01x to 10x of the standard concentration of 2318 µm).

cytoplasm									
glycine+ serine	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.01x	0.038	6.523	1.304	0.384	0.003	4.647	3.187	0.347	0.476
0.1x	0.040	6.584	1.315	0.385	0.003	4.565	3.164	0.357	0.496
0.5x	0.046	6.688	1.333	0.383	0.003	4.427	3.126	0.375	0.529
1x	0.052	6.765	1.346	0.381	0.003	4.324	3.095	0.388	0.555
1.5x	0.056	6.817	1.354	0.380	0.003	4.254	3.074	0.397	0.574
2x	0.059	6.855	1.360	0.379	0.003	4.204	3.059	0.404	0.587
5x	0.069	6.959	1.377	0.375	0.003	4.064	3.014	0.423	0.627
10x	0.074	7.012	1.386	0.374	0.003	3.991	2.989	0.433	0.648
	SHMT	**HCY**	**MET**	**SAM**	**SAH**				
0.01x	0.274	2.851	32.379	95.210	20.081				
0.1x	0.277	3.131	36.709	81.414	29.265				
0.5x	0.282	3.234	38.454	74.467	34.365				
1x	0.286	3.255	38.809	72.893	35.563				
1.5x	0.289	3.264	38.962	72.190	36.105				
2x	0.291	3.270	39.052	71.763	36.435				
5x	0.297	3.283	39.262	70.749	37.226				
10x	0.300	3.289	39.357	70.281	37.593				
nucleus									
glycine+ serine	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.01x	0.007	3.505	0.697	0.170	0.005	0.549	0.253	1.478	1.363
0.1x	0.014	3.512	0.696	0.160	0.004	0.549	0.253	1.477	1.363
0.5x	0.032	3.532	0.691	0.129	0.003	0.549	0.253	1.476	1.363
1x	0.046	3.544	0.688	0.109	0.003	0.549	0.253	1.475	1.363
1.5x	0.053	3.551	0.686	0.097	0.003	0.549	0.253	1.474	1.361
2x	0.059	3.555	0.685	0.090	0.002	0.549	0.253	1.473	1.361
5x	0.072	3.566	0.682	0.072	0.002	0.549	0.253	1.472	1.361
10x	0.078	3.570	0.680	0.063	0.002	0.549	0.253	1.471	1.361

For nuclear FOCM, the model suggests a more pronounced trend as a consequence to changes in glycine and serine levels. Interestingly, a reduction of overall availability of serine and glycine by 50% leads to a dTMP activity sufficient for the requirements of genome replication (Table 4.21). Indeed, increasing serine and glycine availability led to an increase of nuclear THF concentration, while CH2F decreased notably (Figure 4.6e and f). In turn, this

affected nuclear dTMP activity, as the flux through TYMS decreased by 40% of initial activity (Figure 4.6f; Table 4.20). Consequently, the percentage of produced dTMP required for genome replication decreased from 123% to 57% (Table 4.21).

Table 4.21. Steady state activity of dTMP synthesis in response to the variation in cytosolic and nuclear overall availability of glycine and serine. The sum of glycine and serine was step-wise increased considering a scaling factor with the levels 0.01x, 0.1x, 0.5x, 1x, 1.5x, 2x and 5x, and 10x while the ratio glycine: serine stayed constant at 80:20. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication (1.77·109)

glycine+ serine	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
0.01x	73.0	1747.2	$3.31{\cdot}10^{8}$	$1.85{\cdot}10^{9}$	$2.18{\cdot}10^{9}$	18.68	104.62	123.30
0.1x	73.2	1637.3	$3.31{\cdot}10^{8}$	$1.74{\cdot}10^{9}$	$2.07{\cdot}10^{9}$	18.73	98.04	116.77
0.5x	72.8	1320.9	$3.30{\cdot}10^{8}$	$1.40{\cdot}10^{9}$	$1.73{\cdot}10^{9}$	18.63	79.10	97.72
1x	72.5	1116.6	$3.28{\cdot}10^{8}$	$1.18{\cdot}10^{9}$	$1.51{\cdot}10^{9}$	18.55	66.86	85.41
1.5x	72.2	999.4	$3.27{\cdot}10^{8}$	$1.06{\cdot}10^{9}$	$1.39{\cdot}10^{9}$	18.47	59.84	78.32
2x	72.0	923.4	$3.26{\cdot}10^{8}$	$9.79{\cdot}10^{8}$	$1.30{\cdot}10^{9}$	18.42	55.29	73.71
5x	71.4	738.0	$3.23{\cdot}10^{8}$	$7.82{\cdot}10^{8}$	$1.11{\cdot}10^{9}$	18.27	44.19	62.46
10x	71.1	652.9	$3.22{\cdot}10^{8}$	$6.92{\cdot}10^{8}$	$1.01{\cdot}10^{9}$	18.19	39.10	57.29

Table 4.22. Steady state activity of dTMP synthesis in response to the variation in partitioning of glycine and serine. The partitioning of glycine and serine was modeled by step-wise increasing the percentage of serine from 0% to 100% (Scenario 1 to 11, Table 4.12, while overall availability of glycine and serine stayed fixed at 2318µm. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication (1.77·109).

Scenario	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
1	72.1	921.6	$3.27{\cdot}10^{8}$	$9.77{\cdot}10^{8}$	$1.30{\cdot}10^{9}$	18.45	55.19	73.63
2	72.3	1025.3	$3.27{\cdot}10^{8}$	$1.09{\cdot}10^{9}$	$1.41{\cdot}10^{9}$	18.50	61.39	79.89
3	72.5	1116.6	$3.28{\cdot}10^{8}$	$1.18{\cdot}10^{9}$	$1.51{\cdot}10^{9}$	18.55	66.86	85.41
4	72.6	1200.8	$3.29{\cdot}10^{8}$	$1.27{\cdot}10^{9}$	$1.60{\cdot}10^{9}$	18.57	71.90	90.48
5	72.7	1281.5	$3.29{\cdot}10^{8}$	$1.36{\cdot}10^{9}$	$1.69{\cdot}10^{9}$	18.60	76.74	95.34
6	72.8	1361.2	$3.30{\cdot}10^{8}$	$1.44{\cdot}10^{9}$	$1.77{\cdot}10^{9}$	18.63	81.51	100.13
7	73.0	1442.0	$3.31{\cdot}10^{8}$	$1.53{\cdot}10^{9}$	$1.86{\cdot}10^{9}$	18.68	86.35	105.02
8	73.1	1525.7	$3.31{\cdot}10^{8}$	$1.62{\cdot}10^{9}$	$1.95{\cdot}10^{9}$	18.70	91.36	110.06
9	73.2	1613.9	$3.31{\cdot}10^{8}$	$1.71{\cdot}10^{9}$	$2.04{\cdot}10^{9}$	18.73	96.64	115.37
10	73.4	1708.4	$3.32{\cdot}10^{8}$	$1.81{\cdot}10^{9}$	$2.14{\cdot}10^{9}$	18.78	102.30	121.08
11	73.0	1811.0	$3.31{\cdot}10^{8}$	$1.92{\cdot}10^{9}$	$2.25{\cdot}10^{9}$	18.68	108.44	127.12

Furthermore, we carried out a second set of *in silico* experiments, in which overall availability of serine and glycine has been considered constant and the ratio of serine and glycine has been varied according to Table 4.12. Simulations suggest that also varying the ratio of glycine and serine does not have a big effect on model outcomes related to cytosolic nucleotide synthesis (Table 4.22). Nuclear dTMP synthesis benefits from increasing serine availability and we observe a 2-fold increase of the flux through TYMS (Table 4.22). Consequently, overall dTMP activity increases and already at scenario 6 (glycine: serine = 50:50) we can observe dTMP synthesis rates sufficient for genome replication requirements.

Overall, the model suggests that nuclear FOCM is responsive to changes in serine and glycine availability, whereas cytosolic FOCM remains mostly unaffected. A possible explanation for this could be the substrate channeling, which occurs only in the nuclear

compartment, and which leads to increased fluxes which in turn exhibit a higher sensitivity to changes of serine and glycine.

De novo thymidylate biosynthesis was not affected in MCF7 Cells overexpressing MTHFR or TYMS

The model predictions were compared to experimental data from cultured MCF-7 cells. Specifically, the effects of changes in MTHFR and TYMS levels on *de novo* dTMP biosynthesis was investigated. MTHFR catalyzes the irreversible conversion of 5,10-methyleneTHF to 5-methylTHF, which commits one carbon units away from *de novo* dTMP synthesis and towards HCY remethylation to MET. MTHFR concentration was decreased or increased in the model to examine whether there was a dose-response relationship between MTHFR enzyme levels and various readouts of the FOCM network. Our model indicated that cellular dTMP production decreases as MTHFR enzyme concentration increases (Figure 4.4; Table 4.16). The decrease in total dTMP production resulted from decreased TYMS activity in the cytosol. To verify this observation, MTHFR was overexpressed in MCF7 cells and *de novo* dTMP synthesis activity was measured by dU suppression assay. The ratio of the incorporation of [^{14}C]-deoxyuridine (an indicator of *de novo* dTMP biosynthesis) to [^{3}H]-thymidine (an indicator of salvage pathway thymidylate synthesis) into nuclear DNA did not change in MCF7 cells overexpressing MTHFR compared with control MCF7 cells (P=0.43) (Figures 4.7 & 4.8).

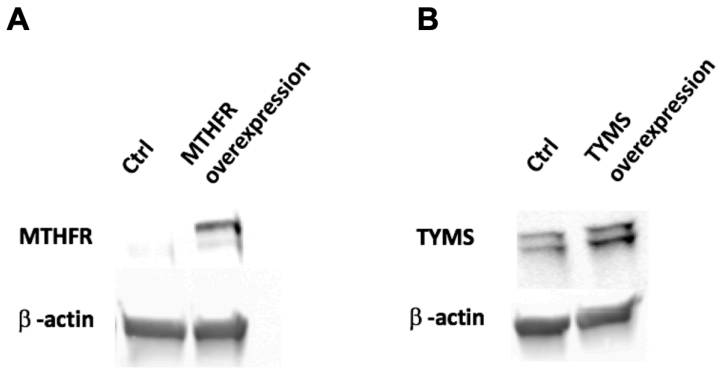

Figure 4.7. Overexpression of MTHFR or TYMS in MCF7 cells. (A) MTHFR protein levels in whole-cell lysates from control and MTHFR overexpressing MCF7 cells. Experiments were repeated two times. (B) TYMS protein levels in whole-cell lysate from control and TYMS overexpressing MCF7 cells. Experiments were repeated two times.

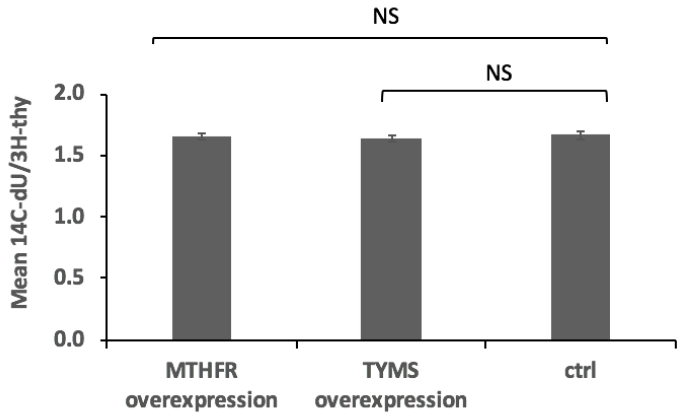

Figure 4.8. Effects of MTHFR or TYMS overexpression on de novo thymidylate biosynthesis of MCF7 cells. The dU suppression assay quantifies the capacity of *de novo* dTMP synthesis relative to the salvage pathway for dTMP synthesis. Data are means of n=3 biological replicates per condition, and the data are shown as means ± SEMs. Data were analyzed by using a paired t test with post hoc Bonferroni correction. NS $P \geq 0.05$, * $p < 0.05$. Experiments were repeated two times.

Similarly, increased or decreased TYMS enzyme concentration was modeled to examine whether there was a dose-response relationship between TYMS enzyme levels and various readouts of the network. The model showed that total dTMP production increases as TYMS enzyme concentration increases (Figure 4.4; Table 4.16). This increase in dTMP production resulted from increased flux through TYMS in both the cytosol and the nucleus. To verify this

observation, TYMS was overexpressed in MCF7 cells and dTMP synthesis activity was measured by a dU suppression assay. *De novo* dTMP biosynthesis was not increased in MCF7 cells overexpressing TYMS compared with control cells (P=0.29) (Figures 4.7 & 4.8). The model-predicted changes of total dTMP production from various MTHFR or TYMS activities may be too small to detect using dU suppression assay.

Table 4.23. Input values for the simulations to assess the contribution of serine, glycine and formate to de novo thymidylate biosynthesis

	standard concentration (uM)	Km (uM)	10x Km (uM)
Serine	468	600	6000
Formate	200	16	160
Glycine	1850	3000	30000

Table 4.24. Six comparisons were chosen to assess the contribution of serine, glycine and formate to de novo thymidylate synthesis. Comparison 1: Serine and formate concentration at their Km values; comparison 2: 10x Km serine and standard Km formate; comparison 3: Standard Km serine and 10x Km formate; comparison 4: Standard Km for serine and formate plus glycine present at Km; comparison 5: 10x Km serine and standard Km formate plus glycine at 10x Km; comparison 6: Standard Km serine and 10x Km formate plus glycine present at Km. Total, nucleus and cytosol dTMP production for various combinations of serine, glycine and formate concentrations. The simulation was carried out for high (5x standard value) MTHFS availability.

high MTHFS	# Ts			% of needed T		
	nucleus	cyto-plasm	total	nucleus	cyto-plasm	total
Reference	1.35E+09	3.32E+08	1.68E+09	76.44	18.73	95.17
Comparison 1	1.79E+09	3.31E+08	2.12E+09	101.19	18.72	119.91
Comparison 2	2.20E+09	3.33E+08	2.53E+09	124.39	18.81	143.19
Comparison 3	1.41E+09	3.32E+08	1.74E+09	79.43	18.74	98.17
Comparison 4	1.57E+09	3.30E+08	1.90E+09	88.90	18.62	107.52
Comparison 5	1.44E+09	3.25E+08	1.76E+09	81.14	18.30	99.48
Comparison 6	1.15E+09	3.30E+08	1.48E+09	64.95	18.60	83.60

De novo dTMP biosynthesis in MCF7 cells was refractory to serine, glycine and formate in culture medium

To assess the impacts of serine, glycine and formate as sources of one-carbon units on *de novo* dTMP biosynthesis, model steady states of six different combinations of serine, glycine and formate concentrations (Table 4.24) were compared to the reference standard simulation (Table 4.23). To verify if the model accurately represents the FOCM network, MCF7 cells were treated with three different combinations of serine, glycine and formate concentrations and *de novo* dTMP biosynthesis capacity was measured using the dU suppression assay (Figure 4.9). The effect of varying MET concentrations at 20 or 200 mM was investigated. Once imported into the cell, MET is converted to S-adenosylmethionine, a potent inhibitor of MTHFR. As expected from the *in silico* model, we observed that the rate of *de novo* dTMP biosynthesis was only mildly affected by different concentrations of serine, glycine and formate. MCF7 cells treated with 200 μM methionine had higher level of *de novo* dTMP biosynthesis relative to those supplemented with 20 μM methionine ($P<0.001$), regardless of serine, glycine and formate concentrations (Figure 4.9). At 200 μM methionine, MCF7 cells treated with 600 μM serine, 3 mM glycine, and 160 μM formate had significantly higher *de novo* dTMP synthesis compared to MCF7 cells treated with 6 mM serine, 30 mM glycine, and 16 μM formate ($P=0.02$). All three combinations of serine, glycine, and formate supplementations resulted in similar levels of *de novo* dTMP synthesis at 20μM methionine ($P>0.05$).

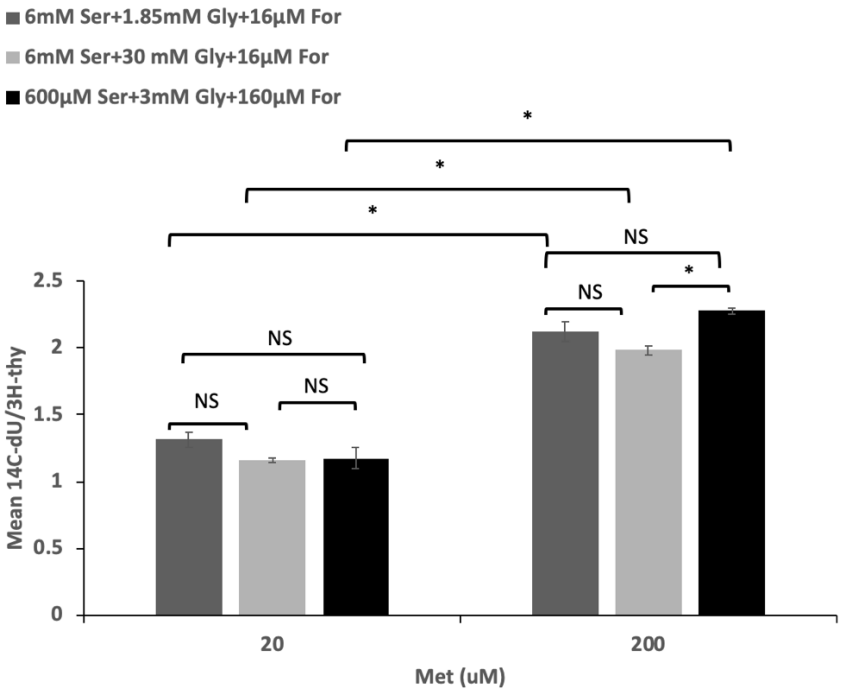

Figure 4.9. Effects of serine, glycine and formate on de novo thymidylate biosynthesis of MCF7 cells. The dU suppression assay quantifies the capacity of de novo dTMP synthesis relative to the salvage pathway for dTMP synthesis. Data are means of n=3 biological replicates per condition, and the data are shown as means ± SEMs. Data were analyzed by using a paired t test with post hoc Bonferroni correction. NS $P \geq 0.05$, * $p < 0.05$. Experiments were repeated two times. Ser, serine; Gly, glycine; For, formate; Met, methionine.

Previously, we observed that MCF7 cells expressing murine glycine transporter (*GlyT*) had altered cellular folate derivative distribution and SAM and SAH concentrations[10]. GlyT expression normalizes intracellular and extracellular glycine concentrations. MCF7GlyT cells treated with 200 μM methionine had increased level of *de novo* dTMP biosynthesis relative to those supplemented with 20 μM methionine (P<0.001) (Figure 4.10). At 200 μM methionine, MCF7GlyT cells treated with 600 μM serine, 3 mM glycine, and 160 μM formate had higher level of *de novo* dTMP synthesis compared to MCF7 cells treated with 6 mM serine, 1.85 mM glycine, and 16 μM formate(P=0.009), and MCF7 cells treated with 6 mM serine, 30 mM glycine, and 16 μM formate(p=0.008). All three combinations of serine, glycine, and formate concentrations have similar level of *de novo* dTMP synthesis at 20μM methionine (P>0.05)

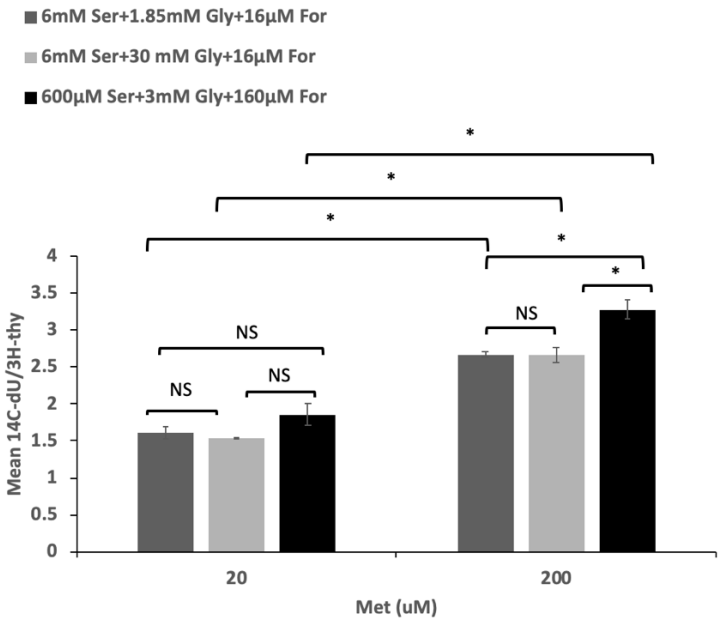

*Figure 4.10. Effects of serine, glycine and formate on de novo thymidylate biosynthesis of MCF7GlyT cells.*The dU suppression assay quantifies the capacity of de novo dTMP synthesis relative to the salvage pathway for dTMP synthesis. Data are means of n=3 biological replicates per condition, and the data are shown as means ± SEMs. Data were analyzed by using a paired t test with post hoc Bonferroni correction. NS $P \geq 0.05$, * $p < 0.05$. Experiments were repeated two times. Ser, serine; Gly, glycine; For, formate; Met, methionine.

4.4 Discussion

Recent experimental studies indicate that folate-dependent dTMP biosynthesis must also occur in the nucleus to limit genome instability[13], [81], [132]. Therefore, we extended our cytosolic FOCM model by building a multi-compartmental model that includes nuclear folate metabolism. A key aspect of the proposed model is that the enzymes constituting the *de novo* dTMP synthesis (SHMT, TYMS, DHFR, and MTHFD1) form a multi-enzyme complex in the nucleus[18], [81]. These enzymes undergo SUMOylation during the G1/S phase, which leads to their nuclear import[14], [19]. In the model, the translocation of the *de novo* thymidylate synthesis pathway to the nuclear compartment is encoded in functions and based on assumptions drawn from experimental data. The complex formation is proposed to accelerate the enzymatic reactions through folate substrate channeling (transfer of co-factors from the

active site of an enzyme to another in the absence of diffusion), a process which, to the best of our knowledge, has not yet been modeled in FOCM literature. Experimental evidence indicates that *de novo* dTMP synthesis is effective only when the enzymes are present in the multi-enzyme complex within the nucleus and is therefore important to prevent uracil misincorporation into DNA and genome instability[19], [81]. Computational models provide a feasible framework to study this aspect and *in silico* simulations enable rapid testing of assumptions related to this and other critical factors. Indeed, by including nuclear FOCM in the model we investigated factors that modify dTMP synthesis, including the effect of multienzyme complex formation, substrate channeling, nuclear MTHFS availability, enzyme expression levels, folate partitioning between the cytosol and nucleus, as well as availability of glycine, serine, NADPH and NADP.

We observed that modeling dTMP synthesis in the nucleus must include multienzyme complex formation and substrate channeling of folate substrates among the enzymes in order to produce the level of dTMP required for DNA synthesis (Tables 4.13 & 4.14). Only when both aspects were taken into account, adequate levels of dTMP synthesis activity were reached. Model simulations of the standard scenario predicted that *de novo* dTMP synthesis may produce up to 85% of needed dTMP. This portion can be increased when enhanced nuclear folate levels are considered, emphasizing the influence of folate partitioning between the two compartments on overall functionality of dTMP synthesis. Notably, even a very small increase in nuclear folate availability (from 10% to 11.6% of total folate) is sufficient to allow for adequate production of dTMP synthesis for mammalian genome replication (Table 4.14). The observed "gap" between *de novo* dTMP synthesis rates and dTMP needs for replication can also be explained by dTMP salvage pathway synthesis. The dTMP salvage pathway, which could be

included in a possible extension of the proposed model, involves the conversion of thymidine to dTMP, and occurs in the cytosol/nucleus catalyzed by TK1 and in the mitochondria by TK2. Salvage pathway synthesis of dTTP is not sufficient to sustain nuclear or mitochondrial DNA replication[163] and therefore requires folate-dependent *de novo* dTMP synthesis to maintain genomic integrity during cell division[1]. Presumably, most dTTP required for DNA replication is synthesized from dUMP by the *de novo* dTMP biosynthesis pathway[163], matching our predicted contribution of *de novo* dTMP synthesis around 85% (Table 4.13).

Sensitivity analyses confirm that the folate pathways involved in purine and dTMP biosynthesis, and HCY remethylation, are interconnected in the cytoplasm, as illustrated by changes in the activity of purine biosynthesis enzymes which then affect both dTMP synthesis and HCY remethylation (Figure 4.4, Table 4.16). Interestingly, *de novo* thymidylate biosynthesis is the most sensitive to perturbations in FOCM activity across the network when modeled in the cytosol. Importantly, *de novo* dTMP synthesis is protected from perturbations in purine biosynthesis and HCY remethylation when the dTMP synthesis enzymes are modeled in the nucleus (Figure 4.4, Table 4.16). The model also clearly shows a key role for minimal MTHFS and DHFR activity in preventing the trapping/accumulation of the folate forms 5fTHF and DHF respectively, the two forms of folate that do not serve as metabolic cofactors (Figure 4.4, Table 4.15).

The model also demonstrates that the *de novo* dTMP synthesis pathway exhibits altered sensitivity to substrate/cofactor availability depending on its location in the cytosol or nucleus. Whereas elevated NADPH + NADP levels impair the pathway in the cytosol by increasing MTHFR activity, nuclear dTMP synthesis is increased in this situation (Figure 4.5, Tables 4.18 & 4.19). In contrast, cytosolic dTMP synthesis was refractory to changes in serine + glycine

levels, whereas nuclear dTMP synthesis was impaired at elevated serine + glycine levels due to accumulation of nuclear THF, presumably due to increased conversion of glycine to serine under these conditions (Figure 4.6, Tables 4.21 & 4.22). Experiments conducted in MCF-7 cells support the relative insensitivity of *de novo* dTMP biosynthesis to exogenous serine, glycine and formate.

An open question is related to the identification of the relative contribution of nuclear and cytosolic dTMP synthesis to overall dTMP synthesis. Our model clearly highlights the importance of nuclear dTMP synthesis and predicts that during S-phase around 79% of overall dTMP molecules are synthesized in the nucleus. A possible explanation for this high contribution is given by the fact that the complex formation and substrate channeling occur solely in the nucleus. Without these two features, cytosolic *de novo* dTMP synthesis remains the main contributor to overall dTMP synthesis activity. Moreover, nuclear translocation uncouples *de novo* dTMP synthesis from *de novo* purine biosynthesis and HCY remethylation in the cytosol, thereby eliminating competition for folate cofactors[17]. The simulation results provide a computational indication that the translocation to the nuclear compartment is essential to prevent uracil incorporation in DNA by protecting dTMP biosynthesis from variability in the network, introduced e.g. by fluctuations in NADPH, NADP levels.

In summary, inclusion of the nuclear compartment in the computational model has provided new insights into the functioning of the FOCM network. The model confirms that accounting for the kinetic effects of nuclear multienzyme complex formation and substrate channeling is essential for the functioning of *de novo* dTMP synthesis. *In silico* simulations also indicate that the nuclear compartment plays an important role for regular cell replication and DNA repair.

Chapter 5 GLYCINE-DERIVED ONE-CARBON UNITS ARE NOT ESSENTIAL FOR C2C12 MYOBLASTS PROLIFERATION

Author Contributions:

Xu Lan (X.L.), Sajin Bae (S.B.), M.S. Field (M.S.F.) and P.J. Stover (P.J.S.) contributed to the experimental design. X.L. and S.B. conducted the experiments, performed data analysis, and wrote the original manuscript. X.L. and P.J.S. edited and revised the manuscript.

Abstract:

Serine and glycine are interconvertible through folate-mediated one-carbon metabolism (FOCM) and provide precursors for the biosynthesis of proteins, nucleic acids and lipids in the cell. Previous studies suggest that cancer cells and T lymphocytes required serine but not glycine for proliferation, whereas glycine alone was effective at stimulating muscle cell expansion. The underlying mechanism of the unique nutrient requirement for glycine in muscle cell proliferation is not fully understand. One function of glycine is to contribute one-carbon (1C) units to FOCM. We hypothesized that muscle cell proliferation depends on glycine derived one-carbon units in FOCM for thymidylate and purine synthesis for DNA replication. The effect of glycine on proliferation of C2C12 myoblasts and its impact on folate metabolism and *de novo* thymidylate biosynthesis were examined. Our study revealed that the addition of serine or glycine alone stimulated C2C12 myoblasts proliferation in a dose-dependent manner. Glycine supplementation increased intracellular 5-methylTHF at the expense of THF and contributed one-carbon units to *de novo* thymidylate biosynthesis. However, formate, hypoxanthine or thymidine supplementation failed to rescue the growth of C2C12 myoblasts under glycine depletion, suggesting that generating 1C units is not the primary mechanism by which glycine supports C2C12 myoblasts proliferation.

5.1 Introduction

Serine and glycine are non-essential amino acids that provide precursors for the biosynthesis of proteins, nucleic acids, and lipids that are important for cell growth and function. Although serine and glycine are interconvertible through folate-mediated one-carbon metabolism (FOCM) (Figure 5.1), cancer cells and T lymphocytes depend specifically on serine and not glycine for proliferation[23], [24]. On the contrary, serine or glycine supplementation alone was effective at stimulating human muscle progenitor cells (*h*MPCs) expansion[25], suggesting that muscle cells have a unique requirement for glycine. *h*MPCs exhibit limited capacity for serine and glycine biosynthesis in response to serine and glycine restriction[25]. The underlying mechanism of glycine supporting muscle cell proliferation remains unknown.

Mitochondria-derived formate from seine and glycine catabolism enters the cytosolic or nuclear folate-mediated one-carbon metabolism (FOCM), which is required for the *de novo* purine biosynthesis, *de novo* thymidylate biosynthesis, and for the remethylation of homocysteine to methionine[11] (Figure 5.1). Folate plays an important role in skeletal muscle cell development[175]. Folic acid deficiency induces deregulation of the cell cycle exit and many cell cycle regulatory genes, induces cellular senescence, and inhibits proliferation and differentiation in C2C12 myoblasts[176]. Folate deficiency during early-mid pregnancy affects piglet longissimus dorsi muscle fiber number and content of intramuscular triglyceride[177]. Clinical effects of impaired FOCM are linked to defects in DNA synthesis and homocysteine remethylation, which result in decreased rates of cell division and elevated plasma homocysteine[178]. Hyperhomocysteinemia has been further associated with reduced skeletal muscle function and strength[179], [180] as well as increased risk of osteoporotic fracture[181], although the underlying causal mechanisms have not been fully understood. Taken together, the

results of these studies suggest the importance of normal functioning of FOCM in skeletal muscle cell growth and function. In this study, we sought to examine the role of glycine on proliferation of C2C12 myoblasts. Specifically, our study investigated the impacts of glycine on folate metabolism and *de novo* thymidylate biosynthesis in C2C12 myoblasts.

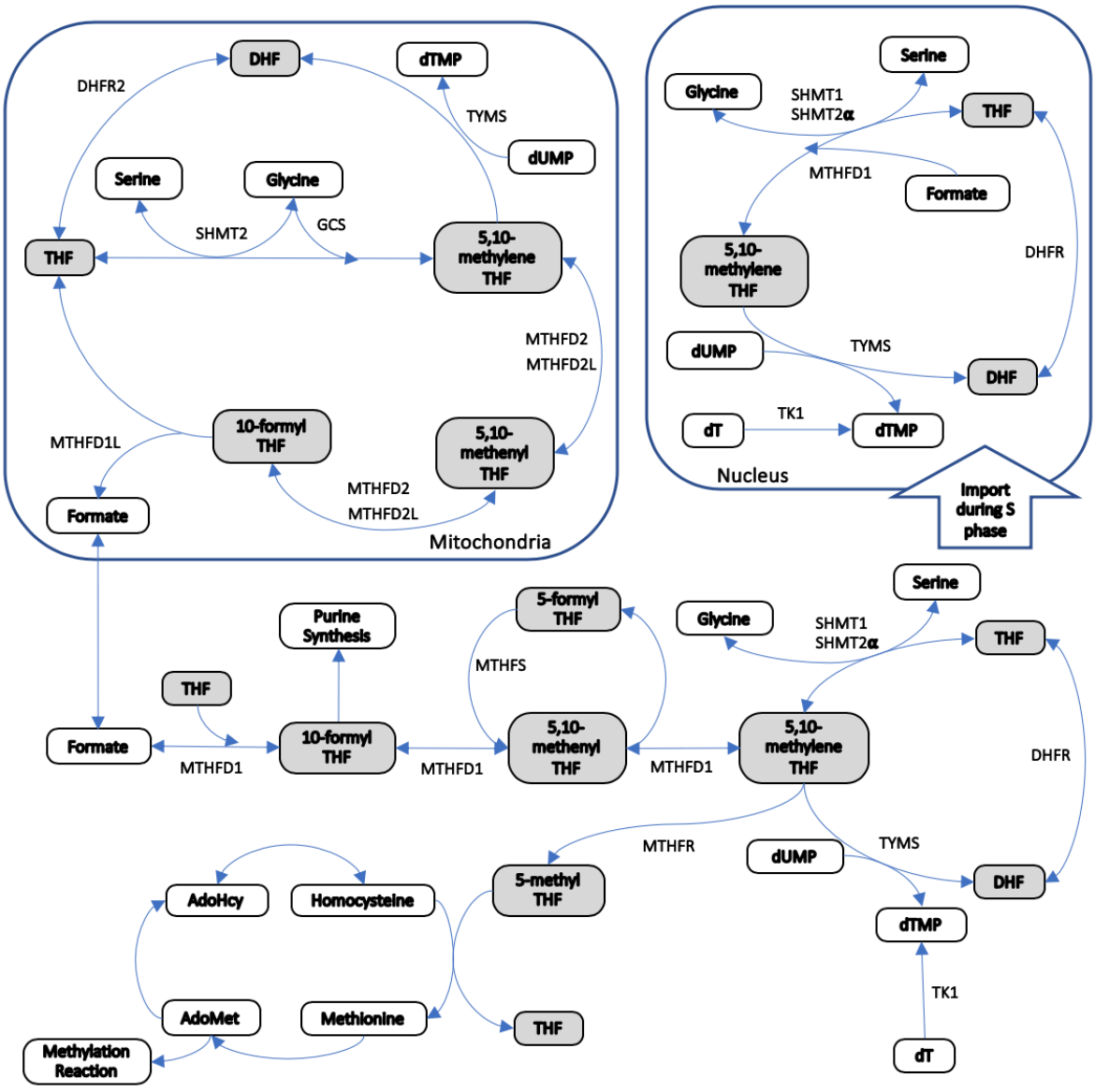

*Figure 5.1. Serine and glycine catabolism contribute 1C units to FOCM.*Serine and glycine are interconvertible through enzyme serine hydroxymethyltransferase (SHMT) in FOCM. Mitochondria-derived formate from serine and glycine catabolism enters the cytosolic or nuclear FOCM, which is required for the de novo purine biosynthesis, *de novo* thymidylate biosynthesis, and for the remethylation of homocysteine to methionine.

5.2 Methods

Cell culture.

C2C12 myoblasts were maintained in DMEM (Corning) with 10% (vol:vol) FBS (HyClone), 1% penicillin/streptomycin (Mediatech) and 4 mM L-alanyl-L-glutamine (Corning) at 37°C and 5% CO_2. For all experiments, unless otherwise indicated, cells were cultured in "modified DMEM" that consisted of DMEM lacking glycine, serine, folate, choline, methionine, and all nucleosides and nucleotides with 10% dialyzed and charcoal-treated FBS, 1% penicillin/streptomycin, and 4 mM L-alanyl-L-glutamine.

Cell proliferation.

C2C12 myoblasts were seeded in a 96-well plate in modified DMEM (as indicated in Cell Culture) supplemented with 20 μM methionine and 25 nM (6S) 5-formylTHF. To determine a dose-response effect of serine, glycine or formate on cell growth, cells were further supplemented with one of the followings in ten replicates: 1) 10 μM glycine; 2) 250 μM glycine; 3) 10 μM serine; 4) 100 μM serine; 5) 30 μM sodium formate; or 6) 150 μM sodium formate. The cells cultured in the same modified DMEM but not supplemented with glycine, serine, and formate served as a negative control, and those cultured in regular DMEM served as a positive control. To determine the effect of hypoxanthine and thymidine on cell growth, cells were further supplemented with one of the followings in six replicates: 1) 100 μM hypoxanthine; 2) 30 uM thymidine; 3) 100 μM hypoxanthine and 30 uM thymidine; 4) 500 μM glycine; 5) 500 μM glycine and 100 μM hypoxanthine; 6) 500 μM glycine and 30 uM thymidine; or 7) 500 μM glycine, 100 μM hypoxanthine, and 30 uM thymidine. The cells cultured in regular DMEM served as a positive control. Cell confluence was measured by using a Celigo Image Cytometer (Nexcelom).

Determination of folate cofactor one-carbon distribution.

C2C12 myoblasts were seeded in 10-cm tissue culture plates (Corning) in modified DMEM supplemented with 20 μM methionine. To determine the effect of glycine, serine or formate on distribution of the folate cofactor one-carbon forms, cells were further supplemented with one of the followings in duplicate: 1) 250 μM glycine; 2) 100 μM serine; 3) 150 μM sodium formate; or 4) 250 μM glycine, 100 μM serine and 150 μM sodium formate. Upon ~60% confluence, plates were rinsed twice with 1X PBS and replaced with the same culture medium but supplemented with 50 nM (6S)-[^{3}H]5-formylTHF (Moravek Biochemicals) [157]. Following a 24h incubation, the same culture medium without any folate was added to the appropriate plates, and cells continued to grow for 24h. Cells were harvested, and intracellular folates were extracted as previously described[115]. Folate one-carbon forms (i.e., 10-formylTHF, THF, 5-formylTHF and 5-methylTHF) were separated and quantified using reverse-phase HPLC as previously described [182].

Glycine metabolism and *de novo* thymidylate biosynthesis.

To determine whether 1C units derived from glycine catabolism are incorporated into dTMP, C2C12 myoblasts were seeded in 10-cm tissue culture plates in modified DMEM supplemented with 20 μM methionine and 25 nM (6S) 5-formylTHF. Following a 48h incubation, cells were plated in triplicate in the same culture medium but supplemented with 20 μM [2-^{14}C]-glycine (American Radiolabeled Chemicals) for 48h. Cells were harvested, and genomic DNA was isolated by using a High Pure PCR template preparation kit (Roche) with RNase A treatment according to the manufacturer's instructions. DNA was digested to nucleosides[183], and isotopic enrichment of thymidine in DNA was determined as previously described[184].

Statistical analysis.

Linear mixed models with Bonferroni corrections were used to assess the effect of serine, glycine or formate on cell growth over time. To examine the effects of serine, glycine, and formate on folate cofactor one-carbon distribution, a one-way ANOVA was used with post hoc Bonferroni correction. Difference in radioactivity (counts per minute) from thymidine fraction and baseline level was assessed by using a paired *t* test. Data are shown as means ± SDs of 2-10 biological replicates/condition. All statistical tests were two-sided, and significance was defined as $P < 0.05$. Analyses were performed with IBM SPSS Statistics (version 20). Linear mixed model was used to assess the effects of hypoxanthine and thymidine on cell growth over time. Analyses were performed with JMP (version 14).

5.3 Results

Serine and glycine affect the growth of C2C12 myoblasts in a dose-dependent manner

Serine or glycine, both nonessential amino acids, stimulated C2C12 growth. There were significant interactions between treatments and time ($P < 0.001$), indicating that the growth rate of C2C12 myoblasts differed by treatments (Figure 5.2). C2C12 myoblasts cultured in modified DMEM lacking all three sources of 1C units (i.e., serine, glycine and formate) exhibited significantly lower growth rates relative to those cultured in regular DMEM over 72 h ($P < 0.001$). The addition of serine (10 or 100 μM) in modified DMEM resulted in a pronounced increase in the growth of C2C12 myoblasts in a dose-dependent manner (Figure 5.2A); the cells supplemented with 10 μM serine exhibited 29%, 106%, and 191% higher cell confluence at 24, 48, and 72 h, respectively, compared with those maintained in modified DMEM lacking three sources of one-carbon units ($P < 0.001$). C2C12 myoblasts supplemented with 100 μM serine reached 100% confluence by 72 h, exhibiting 97% and 267% increases in their growth at 24 and

48 h, respectively, compared with those lacking one-carbon sources ($P < 0.001$). Compared with the cells cultured in regular DMEM, which served as a positive control, C2C12 myoblasts in modified DMEM supplemented with 100 μM serine exhibited 17% lower cell confluence at 24 and 48 h ($P < 0.001$).

Similarly, the addition of glycine (10 or 250 μM) in modified DMEM resulted in significant increases in the growth of C2C12 myoblasts in a dose-dependent manner, but to a lesser extent than the serine supplementation (Figure 5.2B); the cells supplemented with 10 μM glycine exhibited 21%, 39%, and 74% higher cell confluence at 24, 48, and 72 h, respectively, compared with those maintained in modified DMEM lacking three sources of one-carbon units ($P < 0.001$). The addition of 250 μM glycine in modified DMEM increased cell growth by 30%, 96% and 169% at 24, 48 and 72 h, respectively, relative to modified DMEM lacking one-carbon sources ($P < 0.001$). Compared with C2C12 myoblasts cultured in regular DMEM, the cells supplemented with 250 μM glycine exhibited 50% and 60% lower cell confluence at 24 and 48 h, respectively ($P < 0.001$). The addition of formate (30 or 150 μM) in modified DMEM did not affect the growth of C2C12 myoblasts compared with the condition lacking one-carbon sources ($P \geq 0.12$; Figure 5.2C).

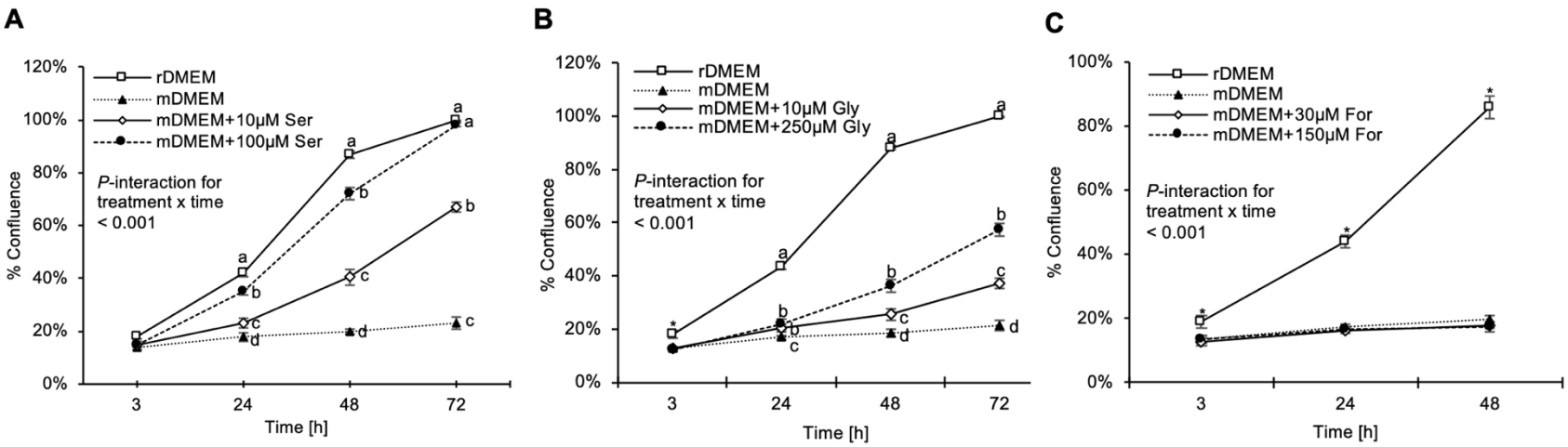

Figure 5.2. Effects of serine (A), glycine (B) and formate (C) on the growth of C2C12 myoblasts. Cell confluence was measured over 48 or 72 h. Data were analyzed using a linear mixed model with Bonferroni correction. Values are shown as means ± SDs of ten biological replicates/condition. Labeled means without a common letter at each time point differ, $P < 0.001$. *Different from the other conditions at each time point, $P < 0.001$. rDMEM, regular DMEM; mDMEM, modified DMEM lacking serine, glycine and formate; Ser, serine; Gly, glycine; For, formate

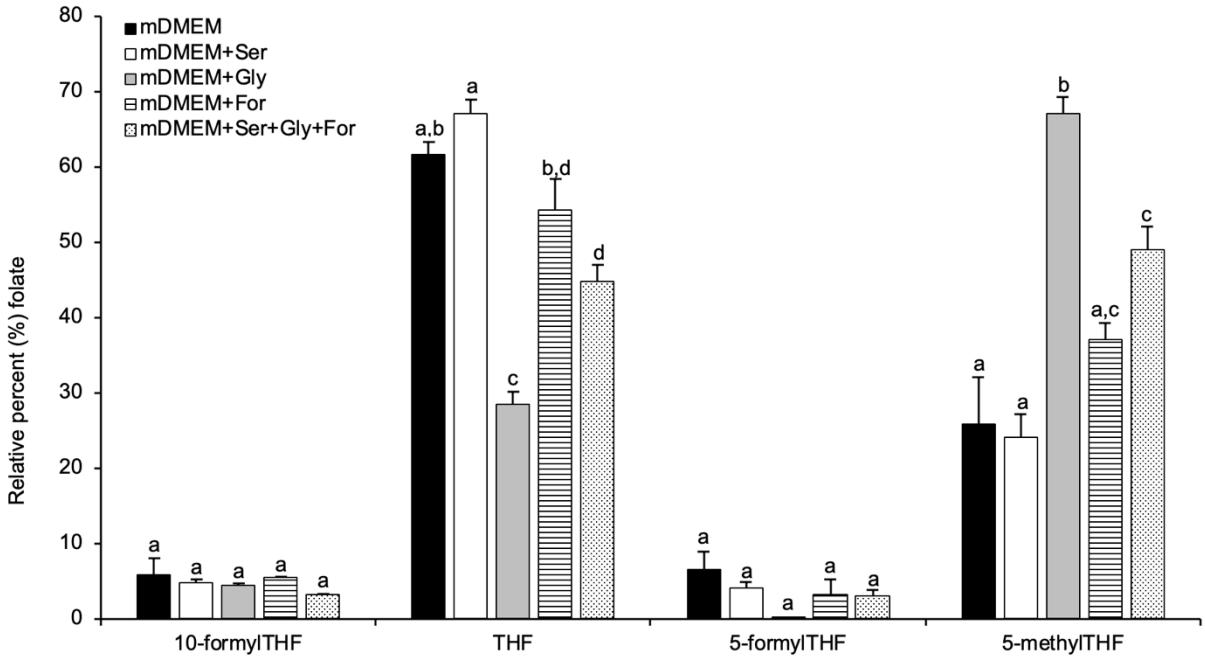

Figure 5.3. Effects of serine, glycine and formate on cellular distribution of folate one-carbon forms. The relative percentages of folate forms including 10-formylTHF, THF, 5-formylTHF and 5-methylTHF are shown for each treatment. Differences in folate forms among treatment groups were determined by using a one-way ANOVA with Bonferroni correction. Values are shown as means ± SDs of 2 biological replicates/condition. Labeled means without a common letter in each folate one-carbon forms differ, $P < 0.05$. mDMEM, modified DMEM lacking serine, glycine and formate; Ser, serine; Gly, glycine; For, formate.

Glycine supplementation increases intracellular 5-methylTHF at the expense of THF

C2C12 myoblasts cultured in modified DMEM lacking all three sources of 1C units (serine, glycine and formate) exhibited a 47% decrease in intracellular 5-methylTHF ($P = 0.01$) and 38% increase in THF levels ($P = 0.01$) compared with those supplemented with sources of 1C units including 100 μM serine, 250 μM glycine and 150 μM formate (Figure 5.3). The addition of 100 μM serine alone in modified DMEM did not reverse these changes, yielding a 51% decrease in 5-methylTHF pools ($P = 0.01$) and 50% increase in THF pools ($P = 0.003$) compared to the condition supplemented with all 1C sources. Notably, the addition of 250 μM glycine alone in modified DMEM was sufficient to reverse the changes in folate vitamer

distribution induced by the depletion 1C sources, yielding a 160% increase in 5-methylTHF pools (P = 0.001) and 54% decrease in THF pools (P < 0.001) compared with the condition lacking serine, glycine and formate. The addition of 150 μM formate alone in modified DMEM did not affect the distribution of folate one-carbon forms compared with the condition lacking 1C sources or the condition supplemented with 1C sources ($P \geq 0.12$). Overall, these findings suggest that glycine supports the generation of 5-methylTHF, with the increase in 5-methylTHF occurring at the expense of THF, in C2C12 myoblasts.

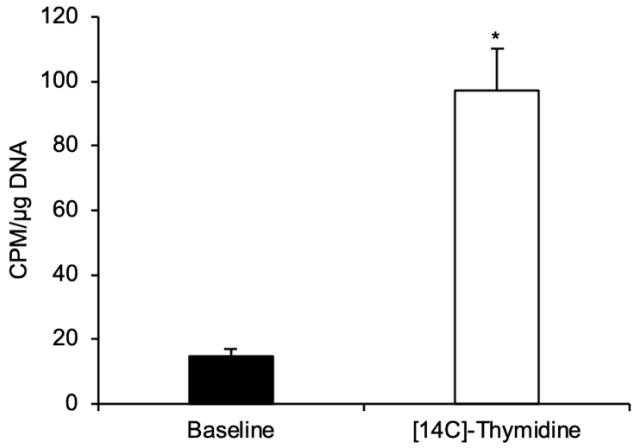

Figure 5.4. Incorporation of exogenous glycine into thymidine in DNA. The incorporation of [^{14}C]-glycine into thymidine was normalized to DNA content. Data were analyzed by using a paired *t* test. Values are shown as means ± SDs of 3 biological replicates. *Different from Baseline, P < 0.05. CPM, counts per minute.

Exogenous glycine is incorporated into thymidine in C2C12 myoblasts

We next determined whether 1C units derived from glycine contribute to *de novo* dTMP synthesis by supplementing C2C12 myoblasts with [2-^{14}C]-glycine and measuring radioactivity in thymidine present in nuclear DNA. Exogenous glycine was shown to be incorporated into thymidine as indicated by the presence of [^{14}C]-thymidine in DNA, with the levels increased by 6.5-fold compared to the baseline radioactivity levels (P = 0.009; Figure 5.4).

Hypoxanthine or thymidine does not rescue the growth of C2C12 myoblasts under glycine depletion

Next, we investigated if glycine support growth of C2C12 myoblasts by contributing 1C units in FOCM for *de novo* dTMP synthesis or *de novo* purine synthesis. To this end, cells cultured in modified DMEM were supplemented with hypoxanthine (an alternative source of purines through the purine salvage pathway) or thymidine in the presence or absence of glycine (Figure 5.5, Table 5.1). There were significant interactions between treatments and time ($P<0.001$), indicating that the growth rate of C2C12 myoblasts differed by treatment. Consistently, C2C12 myoblasts cultured in modified DMEM with glycine supplementation showed significantly higher growth rates relative to those cultured in modified DMEM lacking glycine ($p=0.0007$). Addition of hypoxanthine, thymidine, or both in modified DMEM lacking glycine did not rescue growth of C2C12 myoblasts, suggesting that generating 1C units is not the primary mechanism by which glycine supports C2C12 myoblasts proliferation ($p<0.0001$, $p=0.0094$, $p<0.0001$, respectively). Addition of hypoxanthine, thymidine, or both in modified DMEM containing glycine did not have a significant effect on cell growth ($p=0.5194$, $p=0.468$, $p=6194$, respectively).

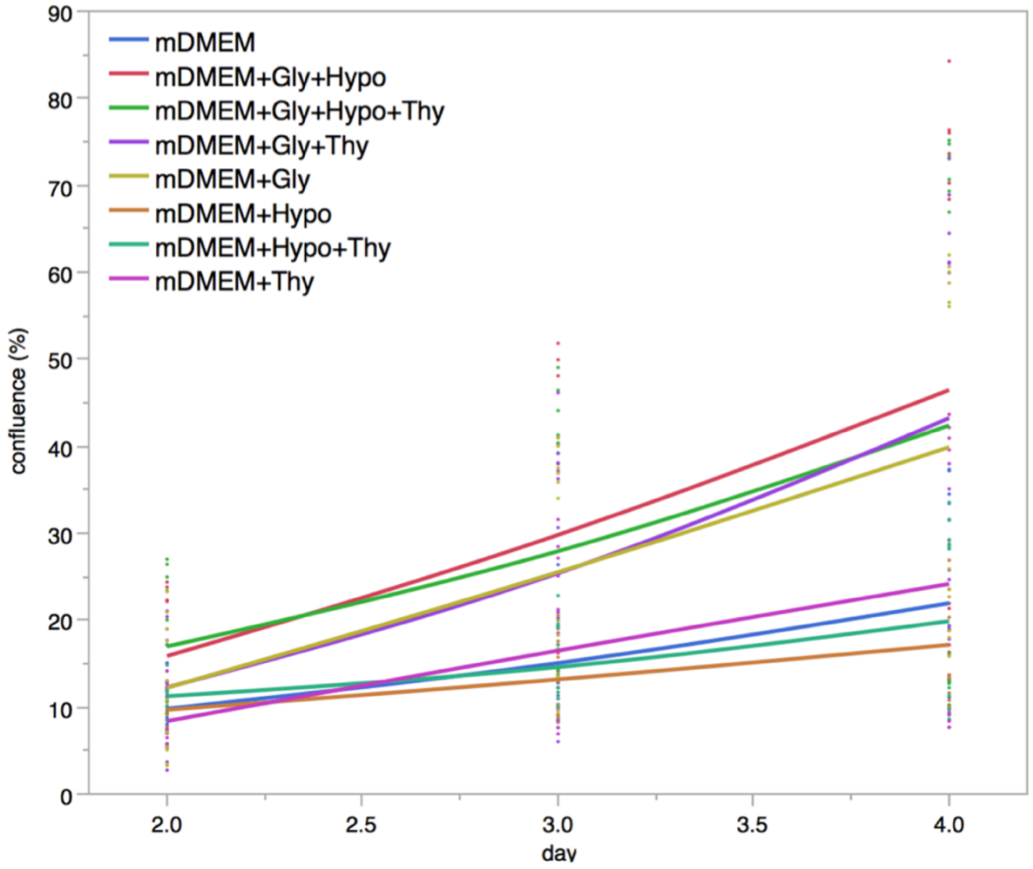

Figure 5.5. Addition of hypoxanthine or thymidine did not rescue the growth of C2C12 myoblasts in media lacking glycine. Data were analyzed using a linear mixed model. mDMEM, modified DMEM; Gly, glycine; Hypo, hypoxanthine; Thy, thymidine.

Table 5.1. Results of linear mixed model regression analysis between media and time. Confluence were measured on day 2-4. Data were combined from two replicate experiments (n=6) and were analyzed using a linear mixed model, significant differences compared to C2C12 myoblasts treated with modified DMEM plus glycine (P<0.05). The comparisons and interactions between "treatment" and "time" were statistically significant (p<0.0001). mDMEM, modified DMEM; Gly, glycine; Hypo, hypoxanthine; Thy, thymidine.

Variables	Regression coefficient	Std Error	*P*
Intercept	-15.617	11.637	0.3454
media[mDMEM]	-10.269	2.380	<.0001*
media[mDMEM+Gly+Hypo]	4.834	2.380	0.0453*
media[mDMEM+Gly+Hypo+Thy]	3.217	2.380	0.18
media[mDMEM+Gly+Thy]	1.083	2.380	0.6503
media[mDMEM+Hypo]	-12.534	2.380	<.0001*
media[mDMEM+Hypo+Thy]	-10.639	2.380	<.0001*
media[mDMEM+Thy]	-9.528	2.380	0.0001*
time	13.835	1.602	<.0001*
media[mDMEM] x time	-7.768	2.266	0.0007*
media[mDMEM+Gly+Hypo] x time	1.463	2.266	0.5194
media[mDMEM+Gly+Hypo+Thy] x time	-1.128	2.266	0.6194
media[mDMEM+Gly+Thy] x time	1.648	2.266	0.468
media[mDMEM+Hypo] x time	-10.100	2.266	<.0001*
media[mDMEM+Hypo+Thy] x time	-9.522	2.266	<.0001*
media[mDMEM+Thy] x time	-5.947	2.266	0.0094*

5.4 Discussion

In this study, we observed that serine or glycine stimulated a significant increase in growth rate of C2C12 myoblasts, which is consistent with previous finding that the addition of serine and glycine alone or in combination stimulated human muscle progenitor cells (*h*MPCs) proliferation in a dose-dependent manner [25]. In C2C12 myoblasts, glycine is essential for myotube size and addition of glycine reduces muscle wasting during growth factor deprivation or nutrient starvation [185]. *Mdx* mice treated with glycine transporter inhibitor exhibited completely compromised muscle regeneration[186]. The effects of glycine on muscle cell proliferation is unique compared to other cell types. In cancer cell lines, exogenous glycine cannot replace serine to support cancer cell proliferation, and restriction of exogenous glycine or depletion of the glycine cleavage system did not impair proliferation[24]. In T lymphocytes,

serine is essential for optimal cell proliferation and glycine cannot compensate for serine deficiency[23].

Glycine catabolism contributes 1C units to FOCM through glycine cleavage system (GCS), which can be incorporated into purines, thymidylate, as well as methyl groups for biosynthetic and regulatory methylation reactions. GCS is a multi-enzyme complex that exist in the mitochondria and consists of four constituents. One of the proteins in GCS, glycine decarboxylase (GLDC) was observed to express in a limited number of human tissues including liver, kidney and placenta, but not in skeletal muscle [187]. Consistently, the enzymatic activity of GCS was identified only in tissues where GLDC mRNA was abundantly expressed [187], [188]. Although skeletal muscles contain very limited amount of GCS activity, evidence suggest that GCS contributes 1C units to FOCM in muscle in mice[186]. Glycine-induced muscle regeneration was significantly impaired in *mdx* mice treated with GLDC-targeted short hairpin RNA (shRNA). Furthermore, treatment with formate or THF partially rescued muscle regeneration in *mdx* mice without glycine[186]. Our study confirms that glycine contributes 1Cs units to FOCM in C2C12 myoblasts. We observed that glycine supplementation increased intracellular 5-methylTHF with a corresponding decrease in THF. In addition, 1Cs units derived from exogenous glycine were incorporated into thymidine in DNA through *de novo* dTMP biosynthesis.

Despite the evidence that glycine plays a role in FOCM in C2C12 myoblasts, our results suggested that generating 1C units is not the primary mechanism by which glycine supports C2C12 myoblasts proliferation. If the requirement for glycine in C2C12 myoblasts was attributed to the role of glycine in FOCM, the proliferative arrest in glycine depleted media should be rescued by supplementation of formate or nucleotide precursors. However, formate

supplementation failed to rescue the growth of C2C12 myoblasts in media lack of serine and glycine, indicating that 1C units were not growth limiting. Furthermore, neither hypoxanthine and/or thymidine supplementation alone or in combination rescued the growth of C2C12 in glycine depleted conditions. Taken together, these findings confirmed previous observations in *h*MPCs that the glycine requirement appears to not be related to DNA synthesis as supplementation with nucleotide precursors did not rescue cell expansion[25].

In addition to the role of generating one-carbon units for FOCM, glycine participates in numerous biosynthetic and signaling pathways in the cell, including protein synthesis and glutathione production. The pathways and mechanisms through which glycine supports muscle cell growth and proliferation have been explored previously. Glycine supplementation protects muscles in a variety of wasting models, including cancer cachexia, sepsis, and reduced caloric intake, likely involve receptor-mediated responses and modulation of intracellular metabolism[189]. Recently, a study in C2C12 myoblasts indicated that mTORC1 signaling is necessary for glycine's protective effects *in vitro*[185]. Glycine stimulated satellite cell proliferation and muscle regeneration by increasing activation of mTORC1 in *mdx* mice[186]. In human muscle progenitor cells (*h*MPCs), both serine and glycine are essential for cell proliferation, likely through supporting synthesis of intracellular antioxidant glutathione, and restriction of serine and glycine was sensed in an EIF2α-dependent manner resulting in cell cycle arrest in G_0/G_1[25]. The underlying mechanisms responsible for the effects of glycine on muscle cells proliferation remain unclear and require further investigation.

In conclusion, our study observed that serine or glycine supplementation alone stimulated C2C12 myoblasts proliferation in a dose-dependent manner. Glycine contributed one-carbon units for folate one-carbon metabolism and for *de novo* thymidylate biosynthesis in C2C12

myoblasts. However, the addition of formate, hypoxanthine or thymidine did not rescue the growth of C2C12 myoblasts under glycine depleted condition, suggesting that generating 1C units is not the primary mechanism by which glycine supports C2C12 myoblasts proliferation.

Chapter 6 SUMMARY AND FUTURE WORK

This dissertation research investigated the nutritional, biochemical and genetic factors that regulate folate-mediated one-carbon metabolism (FOCM), with a focus on the *de novo* thymidylate biosynthesis pathway. Chapter 2 reviewed the cell cycle regulation of FOCM. Evidence suggests that folate-dependent and folate-independent purine and thymine deoxyribonucleotide synthesis are regulated by cell cycle. The cellular needs at each stage of the cell cycle are met by assigning priority to individual pathways within the network through regulation of enzyme expression, activity, and subcellular localization, as well as multienzyme complex formation. Specifically, enzymes in *de novo* thymidylate biosynthesis pathway are expressed at a higher level and form a complex in the nucleus during S and G_2M phase of the cell cycle. Nuclear translocation of *de novo* thymidylate biosynthesis pathway eliminates the competition for folate cofactors with *de novo* purine biosynthesis and homocysteine remethylation in the cytosol, assuming that folate nuclear and cytosolic folate cofactors are not shared. *De novo* purine biosynthesis also exhibits cell cycle regulation at the level of multienzyme complex formation, however, the cell cycle regulation of enzyme levels is less studied. Further investigations are needed to understand the cell cycle regulation of homocysteine remethylation and mitochondrial one-carbon metabolism.

In chapter 3, the role of SUMO-SIM interactions in *de novo* dTMP synthesis complex formation was investigated. Our study confirmed multiple previously predicted SUMOylation sites and identified novel SUMOylation sites on enzymes of the *de novo* dTMP biosynthesis pathway. These SUMO1 modification sites were identified using recombinant SUMOylated proteins purified from engineered bacteria, therefore, *in vivo* verifications are required. SUMOylation of SHMT1 is linked to its nuclear import. It is likely that DHFR, TYMS and MTHFD1 are also translocated into the nucleus through similar mechanisms. SUMO1

modification sites identified in Chapter 3 may be verified by investigating cellular localization of SHMT1, TYMS and MTHFD1 with mutations at SUMO1 modification sites. Polymorphisms in genes coding for proteins in *de novo* thymidylate biosynthesis pathway should also be examined, especially polymorphisms at identified SUMO1 modification sites or adjacent amino acids.

MS analysis of DSSO and DSBU cross-linking experiments with the *in vitro* assembled *de novo* dTMP biosynthesis complex only identified intra crosslinks but no interenzyme crosslinks. Further investigations are needed to characterize the spatial structure and underlying mechanism of protein-protein interactions. One limitation of the experiment is that the purified proteins were not characterized for proper folding. Additionally, DHFR proteins used to assembly *de novo* dTMP biosynthesis complex were not SUMOylated. SUMO1 modified DHFR may be obtained by optimizing the current purification system or by switching to other expression systems, such as insect cell or mammalian cell systems. Furthermore, other MS-cleavable crosslinkers may be considered. Both DSSO and DSBU crosslinkers interact with lysine, which limit the possibility of crosslinking reactions among the proteins in *de novo* dTMP biosynthesis complex. One of the promising crosslinkers is phospho-bisvinylsulfone (pBVS), a recently developed MS-cleavable cross-linker that targets multiple types of amino acid residues, including cysteine, lysine and histidine.

In vivo cross-linking experiments may be considered in future work. In the nucleus, proteins in *de novo* thymidylate biosynthesis pathway form a complex and interact with replication and epigenetic machineries. Compared to *in vitro* cross-linking of assembled protein complex using purified proteins, *in vivo* cross-linking reactions generate highly complex samples and increase difficulty in MS analysis. Proteins in *de novo* thymidylate biosynthesis pathway may be overexpressed to increase nuclear protein concentrations and therefore increase the

possibility of cross-linking and identification of interaction sites.

Aside from facilitating enzyme complex formation through SUMO-SIM interactions, other roles of SUMO1 modification of proteins of *de novo* thymidylate synthesis pathway may be explored. The enzymatic activity or protein interaction with other cellular proteins of SUMO1 modified proteins of *de novo* thymidylate synthesis pathway may be studied. Future experiments are needed to investigate if SUMO1 modifications of DHFR, TYMS and MTHFD1 have similar effect on protein nuclear translocation compared to SHMT1.

To better understand the function of the FOCM network, a previously developed hybrid-stochastic model of FOCM was extended by including the nuclear compartment which is described in Chapter 4. To the best of our knowledge, this is the first FOCM computational model to include nuclear *de novo* thymidylate biosynthesis complex formation and enzyme rate acceleration through folate substrate channeling. Inclusion of the nuclear compartment in the model has provided new insights into the functioning of the FOCM network. The model confirms that accounting for the kinetic effects of nuclear multienzyme complex formation and substrate channeling is essential for the functioning of *de novo* dTMP synthesis. Sensitivity analyses indicate that *de novo* thymidylate biosynthesis in the nucleus is protected from perturbations in purine biosynthesis and HCY remethylation, whereas cytosolic *de novo* thymidylate biosynthesis is more sensitive to changes in the activities of enzymes in FOCM. The model also demonstrates that *de novo* dTMP synthesis pathway in the nucleus exhibits altered sensitivity to substrate/cofactor availability compared to pathway in the cytosol. The model simulation suggests that nuclear translocation is essential for *de novo* thymidylate biosynthesis by eliminating competition for folate cofactors and protecting the pathway from variability in the network. Experimental data in MCF7 cells confirms that MTHFR expression levels does not

significantly affect total *de novo* dTMP synthesis.

This model is a good starting point and can be optimized in several aspects. Firstly, mitochondria one-carbon metabolism may be added to the model. Secondly, the regulation of metabolic pathways in FOCM through cell cycle progression may be added to the model. Thirdly, the model simulation of TYMS expression and serine, glycine and formate concentrations affecting *de novo* dTMP synthesis did not match the experimental results in MCF7 cells. The concentrations of the constant substrates and most model parameter estimates included in the model are from literature in L1210 cell line. The responses of FOCM output variables to changes in enzyme activities and substrate concentrations are likely to be cell type specific. The model can be modified specifically to various tissue or cell lines according to proteomics and metabolomics data. In this way, the model will account for the differential expression and regulation of enzymes associated with FOCM, as well as tissue or cell line specific metabolite concentrations.

One example of differential expression of enzymes in FOCM is glycine decarboxylase (GLDC), one of the four proteins in glycine cleavage system (GCS) in the mitochondria. Glycine catabolism through GCS contributes 1C units to FOCM. GLDC was observed to express in a limited number of human tissues including liver, kidney and placenta. In Chapter 5, the effect of glycine on proliferation of C2C12 myoblasts and its impact on folate metabolism and *de novo* thymidylate biosynthesis was investigated. Addition of serine or glycine alone stimulated C2C12 proliferation. Glycine supplementation increased 5-methylTHF at the expense of THF and contributed one-carbon units to *de novo* thymidylate synthesis. However, formate, hypoxanthine or thymidine supplementation failed to rescue the growth of C2C12 under glycine depletion, suggesting that generating 1C units is not the primary mechanism by which glycine supports

C2C12 myoblasts proliferation.

The underlying mechanism of the unique nutrient requirement for glycine in C2C12 myoblasts proliferation remains to be determined. The role of glycine in other biosynthetic and signaling pathways in the cell may be investigated, especially the mTORC1 signaling pathway and the synthesis of intracellular antioxidant glutathione.

REFERENCE

[1] J. Chon, M. S. Field, and P. J. Stover, "Deoxyuracil in DNA and disease: Genomic signal or managed situation?," *DNA Repair*. 2019.

[2] P. J. Stover, "Physiology of folate and vitamin B12 in health and disease," in *Nutrition Reviews*, 2004.

[3] J. R. Suh, A. K. Herbig, and P. J. Stover, "New perspectives on folate catabolism.," *Annu. Rev. Nutr.*, vol. 21, pp. 255–82, Jan. 2001.

[4] R. Tufi *et al.*, "Enhancing nucleotide metabolism protects against mitochondrial dysfunction and neurodegeneration in a PINK1 model of Parkinson's disease.," *Nat. Cell Biol.*, vol. 16, pp. 157–66, 2014.

[5] S. Friso, S.-W. Choi, G. G. Dolnikowski, and J. Selhub, "A method to assess genomic DNA methylation using high-performance liquid chromatography/electrospray ionization mass spectrometry.," *Anal. Chem.*, vol. 74, pp. 4526–4531, 2002.

[6] S. Friso *et al.*, "A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, pp. 5606–5611, 2002.

[7] R. Jaenisch and A. Bird, "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.," *Nat. Genet.*, vol. 33 Suppl, pp. 245–254, 2003.

[8] C. Huang, E. A. Sloan, and C. F. Boerkoel, "Chromatin remodeling and human disease.," *Curr. Opin. Genet. Dev.*, vol. 13, pp. 246–252, 2003.

[9] J. W. Crott, Z. Liu, S. W. Choi, and J. B. Mason, "Folate depletion in human lymphocytes up-regulates p53 expression despite marked induction of strand breaks in exons 5-8 of the

gene," *Mutat. Res. - Genet. Toxicol. Environ. Mutagen.*, vol. 626, pp. 171–179, 2007.

[10] K. Herbig, E. P. Chiang, L. R. Lee, J. Hills, B. Shane, and P. J. Stover, "Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses," *J. Biol. Chem.*, vol. 277, pp. 38381–38389, 2002.

[11] J. T. Fox and P. J. Stover, "Folate-mediated one-carbon metabolism.," *Vitam. Horm.*, vol. 79, pp. 1–44, 2008.

[12] J. Chakraborty and P. J. Stover, "Deoxyuracil in DNA in health and disease," *Curr. Opin. Clin. Nutr. Metab. Care*, 2020.

[13] M. S. Field, E. Kamynina, and P. J. Stover, "MTHFD1 regulates nuclear de novo thymidylate biosynthesis and genome stability," *Biochimie*, vol. 126, pp. 27–30, Jul. 2016.

[14] D. D. Anderson, C. F. Woeller, and P. J. Stover, "Small ubiquitin-like modifier-1 (SUMO-1) modification of thymidylate synthase and dihydrofolate reductase," in *Clinical Chemistry and Laboratory Medicine*, 2007, vol. 45, no. 12, pp. 1760–3.

[15] M. S. Field, E. Kamynina, D. Watkins, D. S. Rosenblatt, and P. J. Stover, "Human mutations in methylenetetrahydrofolate dehydrogenase 1 impair nuclear de novo thymidylate biosynthesis.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 2, pp. 400–5, Jan. 2015.

[16] D. D. Anderson, C. F. Woeller, E. P. Chiang, B. Shane, and P. J. Stover, "Serine hydroxymethyltransferase anchors de Novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis," *J. Biol. Chem.*, vol. 287, pp. 7051–7062, 2012.

[17] P. J. Stover and M. S. Field, "Trafficking of Intracellular Folates," *Adv. Nutr.*, 2011.

[18] M. S. Field, E. Kamynina, J. Chon, and P. J. Stover, "Nuclear Folate Metabolism," *Annu.*

Rev. Nutr., 2018.

[19] C. F. Woeller, D. D. Anderson, D. M. E. E. Szebenyi, and P. J. Stover, "Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway," *J. Biol. Chem.*, vol. 282, no. 24, pp. 17623–17631, Jun. 2007.

[20] Y. Q. Zhang and K. D. Sarge, "Sumoylation regulates lamin A function and is lost in lamin A mutants associated with familial cardiomyopathies," *J. Cell Biol.*, 2008.

[21] D. N. Simon, T. Domaradzki, W. A. Hofmann, and K. L. Wilson, "Lamin A tail modification by SUMO1 is disrupted by familial partial lipodystrophy-causing mutations," *Mol. Biol. Cell*, 2012.

[22] E. Kamynina, E. R. Lachenauer, A. C. DiRisio, R. P. Liebenthal, M. S. Field, and P. J. Stover, "Arsenic trioxide targets MTHFD1 and SUMO-dependent nuclear de novo thymidylate biosynthesis," *Proc. Natl. Acad. Sci.*, 2017.

[23] E. H. Ma, G. Bantug, T. Griss, M. J. Richer, C. Hess, and R. G. Jones Correspondence, "Serine Is an Essential Metabolite for Effector T Cell Expansion," *Cell Metab.*, vol. 25, pp. 345–357, 2017.

[24] C. F. Labuschagne, N. J. F. van den Broek, G. M. Mackay, K. H. Vousden, and O. D. K. Maddocks, "Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells," *Cell Rep.*, vol. 7, no. 4, pp. 1248–1258, May 2014.

[25] B. Gheller, J. Blum, and A. Thalacker-Mercer, "Serine and Glycine Are Essential for Human Muscle Progenitor Cell P Roliferation (P08-063-19)," *Curr. Dev. Nutr.*, 2019.

[26] M. S. Field, D. M. E. E. Szebenyi, and P. J. Stover, "Regulation of de novo purine biosynthesis by methenyltetrahydrofolate synthetase in neuroblastoma," *J. Biol. Chem.*, vol. 281, pp. 4215–4221, 2006.

[27] P. Stover and V. Schirch, “5-formyltetrahydrofolate polyglutamates are slow tight binding inhibitors of serine hydroxymethyltransferase,” *J. Biol. Chem.*, vol. 266, no. 3, pp. 1543–1550, 1991.

[28] Jae Rin Suh, E. W. Oppenheim, S. Girgis, and P. J. Stover, “Purification and properties of a folate-catabolizing enzyme,” *J. Biol. Chem.*, vol. 275, pp. 35646–35655, 2000.

[29] R. Nilsson *et al.*, “Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer,” *Nat. Commun.*, vol. 5, p. 3128, Jan. 2014.

[30] A. Paone *et al.*, “SHMT1 knockdown induces apoptosis in lung cancer cells by causing uracil misincorporation.,” *Cell Death Dis.*, vol. 5, p. e1525, 2014.

[31] K. E. Boyd and P. J. Farnham, “Myc versus USF: discrimination at the cad gene is determined by core promoter elements.,” *Mol. Cell. Biol.*, vol. 17, no. 5, pp. 2529–37, May 1997.

[32] E. E. Abali, N. E. Skacel, H. Celikkaya, and Y. Hsieh, “Chapter 9 Regulation of Human Dihydrofolate Reductase Activity and Expression,” in *Vitamins and hormones*, vol. 79, 2008, pp. 267–292.

[33] J. E. Slansky and P. J. Farnham, “Transcriptional regulation of the dihydrofolate reductase gene.,” *Bioessays*, vol. 18, no. 1, pp. 55–62, 1996.

[34] L. Good, G. P. Dimri, J. Campisi, and K. Y. Chen, “Regulation of dihydrofolate reductase gene expression and E2F components in human diploid fibroblasts during growth and senescence,” *J. Cell. Physiol.*, vol. 168, no. 3, pp. 580–588, Sep. 1996.

[35] G. D. Grant *et al.*, “Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors.,” *Mol. Biol. Cell*, vol. 24, no. 23, pp. 3634–50, Dec. 2013.

[36] S. Ishida *et al.*, “Role for E2F in Control of Both DNA Replication and Mitotic Functions as Revealed from DNA Microarray Analysis,” *Mol. Cell. Biol.*, vol. 21, no. 14, pp. 4684–4699, Jul. 2001.

[37] J. DeGregori, T. Kowalik, and J. R. Nevins, “Cellular targets for activation by the E2F1 transcription factor include DNA synthesis- and G1/S-regulatory genes.,” *Mol. Cell. Biol.*, vol. 15, no. 8, pp. 4215–24, Aug. 1995.

[38] A. L. Chabes, S. Björklund, and L. Thelander, “S Phase-specific transcription of the mouse ribonucleotide reductase R2 gene requires both a proximal repressive E2F-binding site and an upstream promoter activating region.,” *J. Biol. Chem.*, vol. 279, no. 11, pp. 10796–807, Mar. 2004.

[39] R. K. Hurford, D. Cobrinik, M. H. Lee, and N. Dyson, “pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes.,” *Genes Dev.*, vol. 11, no. 11, pp. 1447–63, Jun. 1997.

[40] Y. Lee and L. F. Johnson, “Transcriptional Control Elements of the Rat Thymidylate Synthase Promoter: Evolutionary Conservation of Regulatory Features,” *Exp. Cell Res.*, vol. 258, no. 1, pp. 53–64, Jul. 2000.

[41] S. Polager, Y. Kalma, E. Berkovich, and D. Ginsberg, “E2Fs up-regulate expression of genes involved in DNA replication, DNA repair and mitosis,” *Oncogene*, vol. 21, no. 3, pp. 437–446, Jan. 2002.

[42] K. M. Huynh, J.-W. Soh, R. Dash, D. Sarkar, P. B. Fisher, and D. Kang, “FOXM1 expression mediates growth suppression during terminal differentiation of HO-1 human metastatic melanoma cells,” *J. Cell. Physiol.*, vol. 226, no. 1, pp. 194–204, Jan. 2011.

[43] C. M. Powell, T. L. Rudge, Q. Zhu, L. F. Johnson, and U. Hansen, “Inhibition of the

mammalian transcription factor LSF induces S-phase-dependent apoptosis by downregulating thymidylate synthase expression.," *EMBO J.*, vol. 19, no. 17, pp. 4665–75, Sep. 2000.

[44] G. P. Morrow *et al.*, "In vivo kinetics of formate metabolism in folate-deficient and folate-replete rats.," *J. Biol. Chem.*, vol. 290, no. 4, pp. 2244–50, Jan. 2015.

[45] A. Solans, X. Estivill, and S. de la Luna, "Cloning and characterization of human FTCD on 21q22.3, a candidate gene for glutamate formiminotransferase deficiency.," *Cytogenet. Cell Genet.*, vol. 88, no. 1–2, pp. 43–9, 2000.

[46] G. S. Ducker and J. D. Rabinowitz, *One-Carbon Metabolism in Health and Disease*, vol. 25, no. 1. 2017, pp. 27–42.

[47] J. Selhub, "Homocysteine metabolism.," *Annu. Rev. Nutr.*, vol. 19, pp. 217–246, 1999.

[48] S. Clarke and K. Banfield, "S-adenosylmethionine-dependent methyltransferases," in *Homocysteine in Health and Disease | Hematology | Cambridge University Press*, 2001.

[49] Z. Bar-Joseph *et al.*, "Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells," *Proc. Natl. Acad. Sci.*, vol. 105, no. 3, pp. 955–960, Jan. 2008.

[50] M. S. Field *et al.*, "Nuclear Enrichment of Folate Cofactors and Methylenetetrahydrofolate Dehydrogenase 1 (MTHFD1) Protect de Novo Thymidylate Biosynthesis during Folate Deficiency.," *J. Biol. Chem.*, vol. 289, no. 43, pp. 29642–50, Oct. 2014.

[51] F. D. Sigoillot, J. A. Berkowski, S. M. Sigoillot, D. H. Kotsis, and H. I. Guy, "Cell cycle-dependent regulation of pyrimidine biosynthesis.," *J. Biol. Chem.*, vol. 278, no. 5, pp. 3403–9, Jan. 2003.

[52] R. C. Jackson, M. S. Lui, T. J. Boritzki, H. P. Morris, and G. Weber, “Purine and Pyrimidine Nucleotide Patterns of Normal, Differentiating, and Regenerating Liver and of Hepatomas in Rats,” *Cancer Res.*, vol. 40, no. 4, 1980.

[53] G. N. Rao and J. N. Davidson, “CAD gene expression in serum-starved and serum-stimulated hamster cells.,” *DNA*, vol. 7, no. 6, pp. 423–32.

[54] G. N. Rao and R. L. Church, “Regulation of CAD gene expression in mouse fibroblasts during the transition from the resting to the growing state.,” *Exp. Cell Res.*, vol. 178, no. 2, pp. 449–56, Oct. 1988.

[55] G. N. Rao, R. L. Church, and J. N. Davidson, “Posttranscriptional regulation of the expression of CAD gene during differentiation of F9 teratocarcinoma cells by induction with retinoic acid and dibutyryl cyclic AMP.,” *FEBS Lett.*, vol. 232, no. 1, pp. 238–42, May 1988.

[56] S. Khan, M. Abdelrahim, I. Samudio, and S. Safe, “Estrogen Receptor/Sp1 Complexes Are Required for Induction of *cad* Gene Expression by 17β-Estradiol in Breast Cancer Cells,” *Endocrinology*, vol. 144, no. 6, pp. 2325–2335, Jun. 2003.

[57] K.-F. Chen, Y.-Y. Lai, H. S. Sun, and S.-J. Tsai, “Transcriptional repression of human cad gene by hypoxia inducible factor-1alpha.,” *Nucleic Acids Res.*, vol. 33, no. 16, pp. 5190–8, 2005.

[58] A. D. Mitchell and N. J. Hoogenraad, “De novo pyrimidine nucleotide biosynthesis in synchronized rat hepatoma (HTC) cells and mouse embryo fibroblast (3T3) cells,” *Exp. Cell Res.*, vol. 93, no. 1, pp. 105–110, Jun. 1975.

[59] A. M. Robitaille *et al.*, “Quantitative Phosphoproteomics Reveal mTORC1 Activates de Novo Pyrimidine Synthesis,” *Science (80-.).*, vol. 339, no. 6125, pp. 1320–1323, Mar.

2013.

[60] I. Ben-Sahra, J. J. Howell, J. M. Asara, and B. D. Manning, “Stimulation of de Novo Pyrimidine Synthesis by Growth Signaling Through mTOR and S6K1,” *Science (80-.).*, vol. 339, no. 6125, pp. 1323–1328, Mar. 2013.

[61] F. D. Sigoillot, D. R. Evans, and H. I. Guy, “Growth-dependent regulation of mammalian pyrimidine biosynthesis by the protein kinase A and MAPK signaling cascades.,” *J. Biol. Chem.*, vol. 277, no. 18, pp. 15745–51, May 2002.

[62] F. D. Sigoillot, S. M. Sigoillot, and H. I. Guy, “Breakdown of the regulatory control of pyrimidine biosynthesis in human breast cancer cells,” *Int. J. Cancer*, vol. 109, no. 4, pp. 491–498, Apr. 2004.

[63] D. R. Evans and H. I. Guy, “Mammalian pyrimidine biosynthesis: fresh insights into an ancient pathway.,” *J. Biol. Chem.*, vol. 279, no. 32, pp. 33035–8, Aug. 2004.

[64] L. M. Graves *et al.*, “Regulation of carbamoyl phosphate synthetase by MAP kinase,” *Nature*, vol. 403, no. 6767, pp. 328–332, Jan. 2000.

[65] M. L. Whitfield *et al.*, “Identification of genes periodically expressed in the human cell cycle and their expression in tumors.,” *Mol. Biol. Cell*, vol. 13, no. 6, pp. 1977–2000, Jun. 2002.

[66] J.-S. R. Wu and L. F. Johnson, “Regulation of dihydrofolate reductase gene transcription in methotrexate-resistant mouse fibroblasts,” *J. Cell. Physiol.*, vol. 110, no. 2, pp. 183–189, Feb. 1982.

[67] C. Santiago, M. Collins, and L. F. Johnson, “In vitro and in vivo analysis of the control of dihydrofolate reductase gene transcription in serum-stimulated mouse fibroblasts,” *J. Cell. Physiol.*, vol. 118, no. 1, pp. 79–86, Jan. 1984.

[68] L. J. Schilling and P. J. Farnham, “Transcriptional regulation of the dihydrofolate reductase/rep-3 locus.,” *Crit. Rev. Eukaryot. Gene Expr.*, vol. 4, no. 1, pp. 19–53, 1994.

[69] E. J. Leys, G. F. Crouse, and R. E. Kellems, “Dihydrofolate reductase gene expression in cultured mouse cells is regulated by transcript stabilization in the nucleus.,” *J. Cell Biol.*, vol. 99, no. 1 Pt 1, pp. 180–7, Jul. 1984.

[70] P. J. Farnham and R. T. Schimke, “Transcriptional regulation of mouse dihydrofolate reductase in the cell cycle.,” *J. Biol. Chem.*, vol. 260, no. 12, pp. 7675–80, Jun. 1985.

[71] J. E. Slansky, Y. Li, W. G. Kaelin, and P. J. Farnham, “A protein synthesis-dependent increase in E2F1 mRNA correlates with growth regulation of the dihydrofolate reductase promoter.,” *Mol. Cell. Biol.*, vol. 13, no. 3, pp. 1610–8, Mar. 1993.

[72] Y. C. Chang, S. Illenye, and N. H. Heintz, “Cooperation of E2F-p130 and Sp1-pRb complexes in repression of the Chinese hamster dhfr gene.,” *Mol. Cell. Biol.*, vol. 21, no. 4, pp. 1121–31, Feb. 2001.

[73] K.-K. Park *et al.*, “Modulation of Sp1-dependent transcription by a cis-acting E2F element in dhfr promoter.,” *Biochem. Biophys. Res. Commun.*, vol. 306, no. 1, pp. 239–43, Jun. 2003.

[74] J. N. Feder, C. J. Guidos, B. Kusler, C. Carswell, D. Lewis, and R. T. Schimke, “A cell cycle analysis of growth-related genes expressed during T lymphocyte maturation.,” *J. Cell Biol.*, vol. 111, no. 6 Pt 1, pp. 2693–701, Dec. 1990.

[75] R. Pelka-Fleischer, W. Ruppelt, W. Wilmanns, H. Sauer, and A. Schalhorn, “Relation between cell cycle stage and the activity of DNA-synthesizing enzymes in cultured human lymphoblasts: investigations on cell fractions enriched according to cell cycle stages by way of centrifugal elutriation.,” *Leukemia*, vol. 1, no. 3, pp. 182–7, Mar. 1987.

[76] H. T. Liu, C. W. Gibson, R. R. Hirschhorn, S. Rittling, R. Baserga, and W. E. Mercer, "Expression of thymidine kinase and dihydrofolate reductase genes in mammalian ts mutants of the cell cycle.," *J. Biol. Chem.*, vol. 260, no. 6, pp. 3269–74, Mar. 1985.

[77] J. N. Feder, Y. G. Assaraf, L. C. Seamer, and R. T. Schimke, "The pattern of dihydrofolate reductase expression through the cell cycle in rodent and human cultured cells.," *J. Biol. Chem.*, vol. 264, no. 34, pp. 20583–90, Dec. 1989.

[78] E. A. Ercikan-Abali, D. Banerjee, M. C. Waltham, N. Skacel, K. W. Scotto, and J. R. Bertino, "Dihydrofolate Reductase Protein Inhibits Its Own Translation by Binding to Dihydrofolate Reductase mRNA Sequences within the Coding Region †," *Biochemistry*, vol. 36, no. 40, pp. 12317–12322, Oct. 1997.

[79] T. Tammsalu, I. Matic, E. G. Jaffray, A. F. M. M. Ibrahim, M. H. Tatham, and R. T. Hay, "Proteome-wide identification of SUMO2 modification sites," *Sci. Signal.*, vol. 7, no. 323, pp. rs2–rs2, Apr. 2014.

[80] I. A. Hendriks, R. C. D'Souza, J.-G. G. Chang, M. Mann, and A. C. O. O. Vertegaal, "System-wide identification of wild-type SUMO-2 conjugation sites," *Nat. Commun.*, vol. 6, p. 7289, Jun. 2015.

[81] D. D. Anderson and P. J. Stover, "SHMT1 and SHMT2 Are Functionally Redundant in Nuclear De novo Thymidylate Biosynthesis," *PLoS One*, vol. 4, no. 6, p. e5839, Jun. 2009.

[82] D. D. Anderson, J. Y. Eom, and P. J. Stover, "Competition between Sumoylation and Ubiquitination of Serine Hydroxymethyltransferase 1 Determines Its Nuclear Localization and Its Accumulation in the Nucleus," *J. Biol. Chem.*, vol. 287, no. 7, pp. 4790–4799, Feb. 2012.

[83] J. Thorndike, T. T. Pelliniemi, and W. S. Beck, “Serine hydroxymethyltransferase activity and serine incorporation in leukocytes.,” *Cancer Res.*, vol. 39, no. 9, pp. 3435–40, Sep. 1979.

[84] R. Pelka-Fleischer, W. Fleischer, W. Wilmanns, H. Sauer, and A. Schalhorn, “Relation between cell cycle stage and the activity of DNA-synthesizing enzymes in cultured human lymphoblasts: investigations on cells separated according to DNA content by way of a cell sorter.,” *Leukemia*, vol. 3, no. 5, pp. 380–5, May 1989.

[85] B. J. Dolnick, “Regulation of Thymidylate Synthase Gene Expression and Drug Response,” pp. 83–91, 2003.

[86] D. Ayusawa, K. Shimizu, H. Koyama, S. Kaneda, K. Takeishi, and T. Seno, “Cell-cycle-directed regulation of thymidylate synthase messenger RNA in human diploid fibroblasts stimulated to proliferate,” *J. Mol. Biol.*, vol. 190, no. 4, pp. 559–567, Aug. 1986.

[87] K. Keyomarsi, J. Samet, G. Molnar, and A. B. Pardee, “The thymidylate synthase inhibitor, ICI D1694, overcomes translational detainment of the enzyme.,” *J. Biol. Chem.*, vol. 268, no. 20, pp. 15142–9, Jul. 1993.

[88] J. Ash, W. C. Liao, Y. Ke, and L. F. Johnson, “Regulation of mouse thymidylate synthase gene expression in growth-stimulated cells: upstream S phase control elements are indistinguishable from the essential promoter elements.,” *Nucleic Acids Res.*, vol. 23, no. 22, pp. 4649–56, Nov. 1995.

[89] Y. Ke, J. Ash, and L. F. Johnson, “Splicing signals are required for S-phase regulation of the mouse thymidylate synthase gene.,” *Mol. Cell. Biol.*, vol. 16, no. 1, pp. 376–83, Jan. 1996.

[90] E. Chu and C. J. Allegra, “The role of thymidylate synthase as an RNA binding protein,”

BioEssays, vol. 18, no. 3, pp. 191–198, Mar. 1996.

[91] B. G. Le François, J. A. Maroun, and H. C. Birnboim, "Expression of thymidylate synthase in human cells is an early G(1) event regulated by CDK4 and p16INK4A but not E2F.," *Br. J. Cancer*, vol. 97, no. 9, pp. 1242–50, Nov. 2007.

[92] C. H. Jenh, P. K. Geyer, and L. F. Johnson, "Control of thymidylate synthase mRNA content and gene transcription in an overproducing mouse cell line.," *Mol. Cell. Biol.*, vol. 5, no. 10, pp. 2527–32, Oct. 1985.

[93] K. M. Huynh *et al.*, "Gene expression analysis of terminal differentiation of human melanoma cells highlights global reductions in cell cycle-associated genes," *Gene*, vol. 433, no. 1–2, pp. 32–39, Mar. 2009.

[94] J. L. Sherley and T. J. Kelly, "Regulation of human thymidine kinase during the cell cycle.," *J. Biol. Chem.*, vol. 263, no. 17, pp. 8350–8, Jun. 1988.

[95] J. W. Littlefield, "The periodic synthesis of thymidine kinase in mouse fibroblasts.," *Biochim. Biophys. Acta*, vol. 114, no. 2, pp. 398–403, Feb. 1966.

[96] L. J. Bello, "Regulation of thymidine kinase synthesis in human cells.," *Exp. Cell Res.*, vol. 89, no. 2, pp. 263–74, Dec. 1974.

[97] A. A. Piper, M. H. Tattersall, and R. M. Fox, "The activities of thymidine metabolising enzymes during the cell cycle of a human lymphocyte cell line LAZ-007 synchronised by centrifugal elutriation.," *Biochim. Biophys. Acta*, vol. 633, no. 3, pp. 400–9, Dec. 1980.

[98] D. L. Coppock and A. B. Pardee, "Control of thymidine kinase mRNA during the cell cycle.," *Mol. Cell. Biol.*, vol. 7, no. 8, pp. 2925–32, Aug. 1987.

[99] J. A. Lewis and D. A. Matkovich, "Genetic determinants of growth phase-dependent and adenovirus 5-responsive expression of the Chinese hamster thymidine kinase gene are

contained within thymidine kinase mRNA sequences.," *Mol. Cell. Biol.*, vol. 6, no. 6, pp. 2262–6, Jun. 1986.

[100] H. T. Liu, R. Baserga, and W. E. Mercer, "Adenovirus type 2 activates cell cycle-dependent genes that are a subset of those activated by serum.," *Mol. Cell. Biol.*, vol. 5, no. 11, pp. 2936–42, Nov. 1985.

[101] L. Kaczmarek, "Protooncogene expression during the cell cycle.," *Lab. Invest.*, vol. 54, no. 4, pp. 365–76, Apr. 1986.

[102] M. G. Kauffman and T. J. Kelly, "Cell cycle regulation of thymidine kinase: residues near the carboxyl terminus are essential for the specific degradation of the enzyme at mitosis.," *Mol. Cell. Biol.*, vol. 11, no. 5, pp. 2538–46, May 1991.

[103] C. J. Stewart, M. Ito, and S. E. Conrad, "Evidence for transcriptional and post-transcriptional control of the cellular thymidine kinase gene.," *Mol. Cell. Biol.*, vol. 7, no. 3, pp. 1156–63, Mar. 1987.

[104] S. Travali, K. E. Lipson, D. Jaskulski, E. Lauret, and R. Baserga, "Role of the promoter in the regulation of the thymidine kinase gene.," *Mol. Cell. Biol.*, vol. 8, no. 4, pp. 1551–7, Apr. 1988.

[105] H. Noguchi, G. Prem veer Reddy, and A. B. Pardee, "Rapid incorporation of label from ribonucleoside disphosphates into DNA by a cell-free high molecular weight fraction from animal cell nuclei.," *Cell*, vol. 32, no. 2, pp. 443–51, Feb. 1983.

[106] T. Yamaoka *et al.*, "Amidophosphoribosyltransferase limits the rate of cell growth-linked de novo purine biosynthesis in the presence of constant capacity of salvage purine biosynthesis.," *J. Biol. Chem.*, vol. 272, no. 28, pp. 17719–25, Jul. 1997.

[107] M. Kondo *et al.*, "The rate of cell growth is regulated by purine biosynthesis via ATP

production and G(1) to S phase transition.," *J. Biochem.*, vol. 128, no. 1, pp. 57–64, Jul. 2000.

[108] M. L. Smith and J. M. Buchanan, "Nucleotide and pentose synthesis after serum-stimulation of resting 3T6 fibroblasts," *J. Cell. Physiol.*, vol. 101, no. 2, pp. 293–309, Nov. 1979.

[109] A. Fridman, A. Saha, A. Chan, D. E. Casteel, R. B. Pilz, and G. R. Boss, "Cell cycle regulation of purine synthesis by phosphoribosyl pyrophosphate and inorganic phosphate," *Biochem. J.*, vol. 454, no. 1, pp. 91–99, Aug. 2013.

[110] H. Iwahana *et al.*, "Rat genomic structure of amidophosphoribosyltransferase, cDNA sequence of aminoimidazole ribonucleotide carboxylase, and cell cycle-dependent expression of these two physically linked genes.," *Biochim. Biophys. Acta*, vol. 1261, no. 3, pp. 369–80, Apr. 1995.

[111] C. G. Proud, "Chapter 8 – mTORC1 and Cell Cycle Control," in *The Enzymes*, vol. 27, 2010, pp. 129–146.

[112] I. Ben-Sahra, G. Hoxhaj, S. J. H. Ricoult, J. M. Asara, and B. D. Manning, "mTORC1 induces purine synthesis through control of the mitochondrial tetrahydrofolate cycle," *Science (80-.).*, vol. 351, no. 6274, 2016.

[113] S. An, R. Kumar, E. D. Sheets, and S. J. Benkovic, "Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells," *Science (80-.).*, vol. 320, no. 5872, pp. 103–106, Apr. 2008.

[114] C. Y. Chan *et al.*, "Purinosome formation as a function of the cell cycle.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 5, pp. 1368–73, Feb. 2015.

[115] S. Girgis, J. R. Suh, J. Jolivet, and P. J. Stover, "5-Formyltetrahydrofolate regulates

homocysteine remethylation in human neuroblastoma," *J. Biol. Chem.*, vol. 272, no. 8, pp. 4729–34, Feb. 1997.

[116] R. Bertrand and J. Jolivet, "Methenyltetrahydrofolate synthetase prevents the inhibition of phosphoribosyl 5-aminoimidazole 4-carboxamide ribonucleotide formyltransferase by 5-formyltetrahydrofolate polyglutamates," *J. Biol. Chem.*, vol. 264, no. 15, pp. 8843–6, May 1989.

[117] J. Peña-Diaz *et al.*, "Transcription profiling during the cell cycle shows that a subset of Polycomb-targeted genes is upregulated during DNA replication," *Nucleic Acids Res.*, vol. 41, no. 5, pp. 2846–2856, Mar. 2013.

[118] M. S. Field, D. D. Anderson, and P. J. Stover, "Mthfs is an Essential Gene in Mice and a Component of the Purinosome.," *Front. Genet.*, vol. 2, p. 36, 2011.

[119] B. Zhu, Z. Xiahou, H. Zhao, B. Peng, H. Zhao, and X. Xu, "MTHFR promotes heterochromatin maintenance," *Biochem. Biophys. Res. Commun.*, vol. 447, no. 4, pp. 702–706, May 2014.

[120] C. A. Hall, "The uptake of vitamin B12 by human lymphocytes and the relationships to the cell cycle.," *J. Lab. Clin. Med.*, vol. 103, no. 1, pp. 70–81, Jan. 1984.

[121] J. Tang *et al.*, "The role of the cell cycle in the cellular uptake of folate-modified poly(L-amino acid) micelles in a cell population.," *Nanoscale*, vol. 7, no. 48, pp. 20397–404, Dec. 2015.

[122] M. M. Doucette and V. L. Stevens, "Folate receptor function is regulated in response to different cellular growth rates in cultured mammalian cells.," *J. Nutr.*, vol. 131, no. 11, pp. 2819–25, Nov. 2001.

[123] F. J. Rosario, T. L. Powell, and T. Jansson, "Mechanistic target of rapamycin (mTOR)

regulates trophoblast folate uptake by modulating the cell surface expression of FR-α and the RFC," *Sci. Rep.*, vol. 6, no. 1, p. 31705, Oct. 2016.

[124] Y. Engström, S. Eriksson, I. Jildevik, S. Skog, L. Thelander, and B. Tribukait, "Cell cycle-dependent expression of mammalian ribonucleotide reductase. Differential regulation of the two subunits.," *J. Biol. Chem.*, vol. 260, no. 16, pp. 9114–6, Aug. 1985.

[125] S. Eriksson, A. Gräslund, S. Skog, L. Thelander, and B. Tribukait, "Cell cycle-dependent regulation of mammalian ribonucleotide reductase. The S phase-correlated increase in subunit M2 is regulated by de novo protein synthesis.," *J. Biol. Chem.*, vol. 259, no. 19, pp. 11695–700, Oct. 1984.

[126] V. D'Angiolella *et al.*, "Cyclin F-Mediated Degradation of Ribonucleotide Reductase M2 Controls Genome Integrity and DNA Repair," *Cell*, vol. 149, no. 5, pp. 1023–1034, May 2012.

[127] V. D'Angiolella, M. Esencay, and M. Pagano, "A cyclin without cyclin-dependent kinases: cyclin F controls genome stability through ubiquitin-mediated proteolysis," *Trends Cell Biol.*, vol. 23, no. 3, pp. 135–140, Mar. 2013.

[128] P. Nordlund and P. Reichard, "Ribonucleotide Reductases," *Annu. Rev. Biochem.*, vol. 75, no. 1, pp. 681–706, Jun. 2006.

[129] S. Björklund, S. Skog, B. Tribukait, and L. Thelander, "S-phase-specific expression of mammalian ribonucleotide reductase R1 and R2 subunit mRNAs.," *Biochemistry*, vol. 29, no. 23, pp. 5452–8, Jun. 1990.

[130] E. Johansson, K. Hjortsberg, and L. Thelander, "Two YY-1-binding proximal elements regulate the promoter strength of the TATA-less mouse ribonucleotide reductase R1 gene.," *J. Biol. Chem.*, vol. 273, no. 45, pp. 29816–21, Nov. 1998.

[131] H. Niida *et al.*, “Essential role of Tip60-dependent recruitment of ribonucleotide reductase at DNA damage sites in DNA repair during G1 phase.,” *Genes Dev.*, vol. 24, no. 4, pp. 333–8, Feb. 2010.

[132] K. Misselbeck, L. Marchetti, M. S. Field, M. Scotti, C. Priami, and P. J. Stover, “A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T MTHFR variant on de novo thymidylate biosynthesis,” *Sci. Rep.*, vol. 7, no. 1, p. 797, Dec. 2017.

[133] A. S. Tibbetts and D. R. Appling, “Compartmentalization of Mammalian folate-mediated one-carbon metabolism.,” *Annu. Rev. Nutr.*, vol. 30, pp. 57–81, 2010.

[134] D. D. Anderson, C. M. Quintero, and P. J. Stover, “Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 37, pp. 15163–15168, Sep. 2011.

[135] N. Gustafsson Sheppard *et al.*, “The folate-coupled enzyme MTHFD2 is a nuclear protein and promotes cell proliferation.,” *Sci. Rep.*, vol. 5, no. 1, p. 15029, Oct. 2015.

[136] P. J. Stover, L. H. Chen, J. R. Suh, D. M. Stover, K. Keyomarsi, and B. Shane, “Molecular cloning, characterization, and regulation of the human mitochondrial serine hydroxymethyltransferase gene.,” *J. Biol. Chem.*, vol. 272, no. 3, pp. 1842–8, Jan. 1997.

[137] A. N. Lane and T. W.-M. Fan, “Regulation of mammalian nucleotide metabolism and biosynthesis.,” *Nucleic Acids Res.*, vol. 43, no. 4, pp. 2466–85, Feb. 2015.

[138] Y. S. Shin, C. Chan, A. J. Vidal, T. Brody, and E. L. Stokstad, “Subcellular localization of gamma-glutamyl carboxypeptidase and of folates.,” *Biochim. Biophys. Acta*, vol. 444, no. 3, pp. 794–801, Oct. 1976.

[139] B. F. Lin, R. F. Huang, and B. Shane, “Regulation of folate and one-carbon metabolism in

mammalian cells. III. Role of mitochondrial folylpoly-gamma-glutamate synthetase.," *J. Biol. Chem.*, vol. 268, no. 29, pp. 21674–9, Oct. 1993.

[140] J. R. Gareau and C. D. Lima, "The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition," *Nature Reviews Molecular Cell Biology*. 2010.

[141] O. Kerscher, "SUMO junction - What's your function? New insights through SUMO-interacting motifs," *EMBO Reports*. 2007.

[142] R. Geiss-Friedlander and F. Melchior, "Concepts in sumoylation: A decade on," *Nature Reviews Molecular Cell Biology*. 2007.

[143] J. S. Seeler and A. Dejean, "Nuclear and unclear functions of sumo," *Nature Reviews Molecular Cell Biology*. 2003.

[144] I. A. Hendriks, R. C. J. D'Souza, B. Yang, M. Verlaan-De Vries, M. Mann, and A. C. O. Vertegaal, "Uncovering global SUMOylation signaling networks in a site-specific manner," *Nat. Struct. Mol. Biol.*, 2014.

[145] E. Meulmeester, M. Kunze, H. H. Hsiao, H. Urlaub, and F. Melchior, "Mechanism and Consequences for Paralog-Specific Sumoylation of Ubiquitin-Specific Protease 25," *Mol. Cell*, 2008.

[146] F. Brockly, M. Piechaczyk, and G. Bossis, "Production and Purification of Recombinant SUMOylated Proteins Using Engineered Bacteria," Humana Press, New York, NY, 2016, pp. 55–65.

[147] K. Keusekotten *et al.*, "Multivalent interactions of the SUMO-interaction motifs in RING finger protein 4 determine the specificity for chains of the SUMO," *Biochem. J.*, 2014.

[148] Y. UCHIMURA, M. NAKAMURA, K. SUGASAWA, M. NAKAO, and H. SAITOH, "Overproduction of eukaryotic SUMO-1- and SUMO-2-conjugated proteins in

Escherichia coli," *Anal. Biochem.*, 2006.

[149] R. Huang *et al.*, "A novel mass spectrometry-cleavable, phosphate-based enrichable and multi-targeting protein cross-linker," *Chem. Sci.*, vol. 10, no. 26, pp. 6443–6447, 2019.

[150] P. Stehmeier and S. Muller, "Phospho-Regulated SUMO Interaction Modules Connect the SUMO System to CK2 Signaling," *Mol. Cell*, 2009.

[151] X. Lan, M. S. Field, and P. J. Stover, "Cell cycle regulation of folate-mediated one-carbon metabolism," *Wiley Interdiscip. Rev. Syst. Biol. Med.*, p. e1426, Jun. 2018.

[152] L. B. Bailey *et al.*, "Biomarkers of Nutrition for Development—Folate Review," *J. Nutr.*, 2015.

[153] M. S. Field and P. J. Stover, "Safety of folic acid," *Ann. N. Y. Acad. Sci.*, 2018.

[154] A. M. Pedley and S. J. Benkovic, *A New View into the Regulation of Purine Metabolism: The Purinosome*, vol. 42, no. 2. 2017, pp. 141–154.

[155] J. Chon, P. J. Stover, and M. S. Field, "Targeting nuclear thymidylate biosynthesis," *Molecular Aspects of Medicine*. 2017.

[156] K. Misselbeck, L. Marchetti, C. Priami, P. J. Stover, and M. S. Field, "The 5-formyltetrahydrofolate futile cycle reduces pathway stochasticity in an extended hybrid-stochastic model of folate-mediated one-carbon metabolism," *Sci. Rep.*, 2019.

[157] A. M. Palmer, E. Kamynina, M. S. Field, and P. J. Stover, "Folate rescues Vitamin B12 depletion-induced inhibition of nuclear thymidylate biosynthesis and genome instability," *Proc. Natl. Acad. Sci. U. S. A.*, 2017.

[158] C. E. Atreya and K. S. Anderson, "Kinetic characterization of bifunctional thymidylate synthase-dihydrofolate reductase (TS-DHFR) from Cryptosporidium hominis: A paradigm shift for TS activity and channeling behavior," *J. Biol. Chem.*, 2004.

[159] P. H. Liang and K. S. Anderson, “Substrate channeling and domain-domain interactions in bifunctional thymidylate synthase-dihydrofolate reductase,” *Biochemistry*, 1998.

[160] Y. Harpaz, M. Gerstein, and C. Chothia, “Volume changes on protein folding,” *Structure*, 1994.

[161] M. R. Speicher, S. E. Antonarakis, and A. G. Motulsky, “Vogel and Motulsky’s human genetics: Problems and approaches (fourth edition),” in *Vogel and Motulsky’s Human Genetics: Problems and Approaches (Fourth Edition)*, 2010, pp. 31–53.

[162] A. Bensadoun and D. Weinstein, “Assay of proteins in the presence of interfering materials,” *Anal. Biochem.*, 1976.

[163] X. Zhou *et al.*, “Progressive loss of mitochondrial DNA in thymidine kinase 2-deficient mice,” *Hum. Mol. Genet.*, 2008.

[164] A. Fujioka *et al.*, “Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes,” *J. Biol. Chem.*, 2006.

[165] P. F. Morrison and C. J. Allegra, “Folate cycle kinetics in human breast cancer cells,” *J. Biol. Chem.*, 1989.

[166] J. Thorndike *et al.*, “Inhibition of Glycinamide Ribonucleotide Formyltransferase and Other Folate Enzymes by Homofolate Polyglutamates in Human Lymphoma and Murine Leukemia Cell Extracts,” *Cancer Res.*, 1989.

[167] E. A. Rayl, B. A. Moroson, and G. P. Beardsley, “ The Human pur H Gene Product, 5-Aminoimidazole-4-carboxamide Ribonucleotide Formyltransferase/IMP Cyclohydrolase ,” *J. Biol. Chem.*, 1996.

[168] J. P. Volpato, E. Fossati, and J. N. Pelletier, “Increasing Methotrexate Resistance by Combination of Active-site Mutations in Human Dihydrofolate Reductase,” *J. Mol. Biol.*,

2007.

[169] R. L. Seither, D. F. Trent, D. C. Mikulecky, T. J. Rape, and I. D. Goldman, “Folate-pool interconversions and inhibition of biosynthetic processes after exposure of L1210 leukemia cells to antifolates. Experimental and network thermodynamic analyses of the role of dihydrofolate polyglutamylates in antifolate action in cells,” *J. Biol. Chem.*, 1989.

[170] W. B. Strong, S. J. Tendler, R. L. Seither, I. D. Goldman, and V. Schirch, “Purification and properties of serine hydroxymethyltransferase and C1-tetrahydrofolate synthase from L1210 cells,” *J. Biol. Chem.*, vol. 265, no. 21, pp. 12149–55, Jul. 1990.

[171] R. G. Matthews, C. Ghose, J. M. Green, K. D. Matthews, and R. Bruce Dunlap, “Folylpolyglutamates as substrates and inhibitors of folate-dependent enzymes,” *Adv. Enzyme Regul.*, 1987.

[172] M. C. Reed, R. L. Thomas, J. Pavisic, S. J. James, C. M. Ulrich, and H. F. Nijhout, “A mathematical model of glutathione metabolism,” *Theor. Biol. Med. Model.*, 2008.

[173] M. C. Reed *et al.*, “A Mathematical Model Gives Insights into Nutritional and Genetic Aspects of Folate-Mediated One-Carbon Metabolism,” *J. Nutr.*, 2006.

[174] W. Manieri, M. E. Moore, M. B. Soellner, P. Tsang, and C. A. Caperelli, “Human glycinamide ribonucleotide transformylase: Active site mutants as mechanistic probes,” *Biochemistry*, 2007.

[175] S. Y. Hwang, B. Sung, and N. D. Kim, “Roles of folate in skeletal muscle cell development and functions.,” *Arch. Pharm. Res.*, vol. 42, no. 4, pp. 319–325, Apr. 2019.

[176] S. Y. Hwang *et al.*, “Folic acid is necessary for proliferation and differentiation of C2C12 myoblasts.,” *J. Cell. Physiol.*, vol. 233, no. 2, pp. 736–747, Feb. 2018.

[177] Y. Li *et al.*, “Folate deficiency during early-mid pregnancy affects the skeletal muscle

transcriptome of piglets from a reciprocal cross.," *PLoS One*, vol. 8, no. 12, p. e82616, 2013.

[178] P. J. Stover, F. Pathologies, and P. J. Stover, "One-Carbon Metabolism – Genome Interactions," *J. Nutr.*, vol. 139, no. 12, pp. 2402–2405, Dec. 2009.

[179] S. Veeranki and S. C. Tyagi, "Defective homocysteine metabolism: potential implications for skeletal muscle malfunction.," *Int. J. Mol. Sci.*, vol. 14, no. 7, pp. 15074–91, Jul. 2013.

[180] S. Veeranki, L. J. Winchester, and S. C. Tyagi, "Hyperhomocysteinemia associated skeletal muscle weakness involves mitochondrial dysfunction and epigenetic modifications.," *Biochim. Biophys. Acta*, vol. 1852, no. 5, pp. 732–41, May 2015.

[181] J. B. J. van Meurs *et al.*, "Homocysteine levels and the risk of osteoporotic fracture.," *N. Engl. J. Med.*, vol. 350, no. 20, pp. 2033–41, May 2004.

[182] M. C. Anguera, J. R. Suh, H. Ghandour, I. M. Nasrallah, J. Selhub, and P. J. Stover, "Methenyltetrahydrofolate synthetase regulates folate turnover and accumulation," *J. Biol. Chem.*, 2003.

[183] P. F. Crain, "Preparation and enzymatic hydrolysis of DNA and RNA for mass spectrometry," *Methods Enzymol.*, 1990.

[184] L. Martiniova, M. S. Field, J. L. Finkelstein, C. A. Perry, and P. J. Stover, "Maternal dietary uridine causes, and deoxyuridine prevents, neural tube closure defects in a mouse model of folate-responsive neural tube defects," *Am. J. Clin. Nutr.*, 2015.

[185] M. K. Caldow, D. J. Ham, J. Trieu, J. D. Chung, G. S. Lynch, and R. Koopman, "Glycine Protects Muscle Cells From Wasting in vitro via mTORC1 Signaling," *Front. Nutr.*, 2019.

[186] C. Lin *et al.*, "Glycine Enhances Satellite Cell Proliferation, Cell Transplantation, and Oligonucleotide Efficacy in Dystrophic Muscle," *Mol. Ther.*, 2020.

[187] S. Kure *et al.*, “Chromosomal localization, structure, single-nucleotide polymorphisms, and expression of the human H-protein gene of the glycine cleavage system (GCSH), a candidate gene for nonketotic hyperglycinemia,” *J. Hum. Genet.*, 2001.

[188] Y. Tadashi and K. Igoro, “Major pathways of serine and glycine catabolism in various organs of the rat and cock,” *J. Biochem.*, 1973.

[189] R. Koopman, M. K. Caldow, D. J. Ham, and G. S. Lynch, *Glycine metabolism in skeletal muscle: Implications for metabolic homeostasis*, vol. 20, no. 4. 2017, pp. 237–242.

[190] W. B. Strong and V. Schirch, “In Vitro Conversion of Formate to Serine: Effect of Tetrahydropteroylpolyglutamates and Serine Hydroxymethyltransferase on the Rate of 10-Formyltetrahydrofolate Synthetase,” *Biochemistry*, 1989.

[191] P. Stover and V. Schirch, “Enzymatic Mechanism for the Hydrolysis of 5, 10-Methenyltetrahydropteroylglutamate to 5-Formyltetrahydropteroylglutamate by Serine Hydroxymethyltransferase,” *Biochemistry*, 1992.

[192] L. L. Lovelace, S. R. Johnson, L. M. Gibson, B. J. Bell, S. H. Berger, and L. Lebioda, “Variants of human thymidylate synthase with loop 181-197 stabilized in the inactive conformation,” *Protein Sci.*, 2009.

[193] S. Radparvar, P. J. Houghton, and J. A. Houghton, “Characteristics of thymidylate synthase purified from a human colon adenocarcinoma,” *Arch. Biochem. Biophys.*, 1988.