
Sapienza University of Rome
Department of Biochemical Sciences “A. Rossi Fanelli”

# Beyond intermediary metabolism: characterization of the moonlighting nucleic acid binding function of human serine hydroxymethyltransferase

Giulia Guiducci

XXXII cycle
PhD Programme in Biochemistry

| Tutor | Coordinator |
| --- | --- |
| Prof. Francesca Cutruzzolà | Prof. Stefano Gianni |

## Summary of the work

The present thesis aims to dissect the moonlighting nucleic acid binding activity of human cytosolic serine hydroxymethyltransferase (SHMT1), a double-faced metabolic enzyme.

Indeed, the work characterizes the interaction between SHMT1 and RNA molecules *in vitro* as well as in cell models, thus unraveling an unexpected complex mechanism in which the synergy among SHMT1, RNA and metabolites fine tunes the compartmentalization of serine metabolism in lung cancer cells.

On the other hand, the study of the SHMT1-DNA complexes, presented herein, set the basis to hypothesize a nuclear function of the enzyme mediated by its binding with the DNA.

Overall this manuscript is pioneer in characterizing the moonlighting nucleic acid binding activity of SHMT1 and, most of all, in determining the functional relevance of such phenomenon, thus placing SHMT1 in the network which interconnects intermediary metabolism with gene replication and gene expression.

.

# TABLE OF CONTENTS

# 1. Introduction

## 1.1 Moonlighting proteins: a fascinating mode of multitasking

The term moonlighting was coined by Costance J. Jeffery in 1999 (1), in his pioneer paper he refers this noun to proteins having two or more different functions, in analogy to moonlighting people who have multiple jobs (Fig. 1).

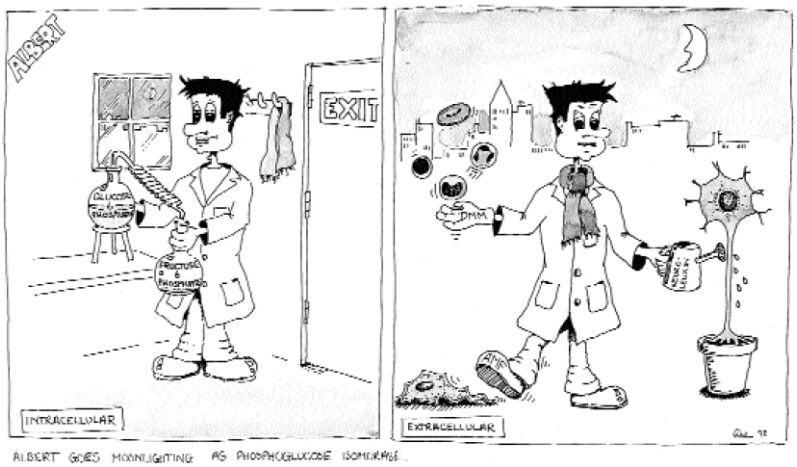


*Figure 1. Picture extracted from the pioneer paper of C.J. Jeffery (1999) in which the author makes an analogy between moonlighting proteins and moonlighting people.*

Indeed, moonlighting proteins perform multiple autonomous activities, often unrelated to each other; the definition excludes proteins resulting from gene fusions, translation of splice variants generated from the same gene and homologous counterparts.

Examples of moonlighting proteins have been described in many species including plants, animals, yeast and prokaryotes, spanning a wide range of

proteins including receptors, enzymes, transcription factors and structural proteins (2).

How a secondary function evolved, as well as how such moonlighting activity switches with the canonical one are interesting questions, which still do not have a clear answer. In some cases, proteins were adapted to a secondary function without many changes in the structural features, while in others the evolution of new binding sites, alternative conformational changes or other new features has been required (3).

The most accredited hypotheses explaining the evolution of moonlighting proteins are briefly summarized below.

The first hypothesis suggests that the rise of mutations results in a biological advantage. Mutations can occur in the coding and/or in the regulatory regions of genes encoding the moonlighting proteins. Although this scenario may be seemingly in contrast with the fact that mutations in the coding regions tend to be deleterious, some structural properties may prevent such negative effects (4). For instance, proteins having flexible and modular regions display higher probability to accumulate mutations without detrimental consequences on the ancestral structure and function, thus favoring the establishment of a moonlighting activity.

Alternatively, moonlighting functions can develop as a consequence of mutations in genes encoding for transcriptional regulators, which in this way acquire the capacity of controlling the expression of multiple proteins, enhancing the possibility that moonlighting interactions, previously nonexistent, may develop (5).

A second model involves gene duplication, such biological event could have enabled the origin of moonlighting functions. Indeed it has been observed, mainly in yeast, the existence of multifunctional proteins originated from

paralogue genes derived from whole genome duplications or from small scale duplication events (4).

Although many moonlighting activities reside in highly conserved proteins (e.g. metabolic enzymes), the presence of such functions cannot be predicted based on homology. For example, the yeast *Saccharomyces cerevisiae* synthetizes two pyruvate carboxylases sharing the 80% of sequence identity with the pyruvate carboxylase from *Hansenula polymorpha*. Despite the high similarity between these proteins, neither one of the pyruvate carboxylases from *Saccharomyces cerevisiae* display moonlighting functions, whereas *Harsenula polymorpha* pyruvate carboxylase is involved in the assembly and translocation of the peroxisomal alcohol oxidase. On the other hand, the same enzyme in *Pichia pastoris*, equally sharing 80% of sequence identity with the pyruvate carboxylase from *Harsenula polymorpha,* it's capable of fulfilling such moonlighting function (2).

Given their complex nature and the lack of common physical features, there is no straightforward method to identify moonlighting proteins, mainly because their second function is often hidden by the more evident one (that it is commonly called "primary function"), and thus hard to unveil.

The increase of newly discovered moonlighting functions prompted researchers to create a tool able to organize and store all the information. The MoonProt database was created in 2015, providing the first centralized, web-based database of moonlighting proteins (6, 7). The database contains information about the sequences, structures and functions of over 200 moonlighting proteins, thus integrating the information scattered across the literature with several links to other databases.

Most moonlighting proteins serve as connection or switch among multiple biochemical pathways, helping the cell to fine tune the response to changes in

the environment as well as to regulate the basal metabolism (3); in other cases their second function is purely structural, as described for the duck protein ε-crystallin which has structural functions in the lens, but it acts as a lactate dehydrogenase in other tissues (8). The repertoire of proteins performing additional functions will surely greatly expand in the future years.
The universe of moonlighting functions is huge and the major part of it is still currently unknown; the present thesis is focused on metabolic enzymes that moonlight as nucleic acids binding proteins.

### 1.2 A weird couple: metabolic enzymes and nucleic acids.

For years metabolic enzymes had been studied for their role in the catalysis of enzymatic reactions essential for cell metabolism. Strikingly, in the last decades it has emerged that the most part of them displays moonlighting functions; among the discovered non-canonical activities of metabolic enzymes the nucleic acid binding property is of great importance.
The evidences that cytosolic and mitochondrial enzymes are able to translocate into the nucleus (Fig. 2) (9) raised the hypothesis of the existence of moonlighting functions carried out through the interaction with the DNA. Indeed, recent papers report evidences of metabolic enzymes that, although lacking canonical DNA binding domains, are involved in the regulation of gene transcription either by interacting with specific DNA sequences or by binding to transcriptional factors (10). The existence of a metabolism-gene transcription axis is thus emerging, allowing the cell to sense metabolic changes and to rapidly respond modulating gene expression through the non canonical activity of metabolic enzymes.

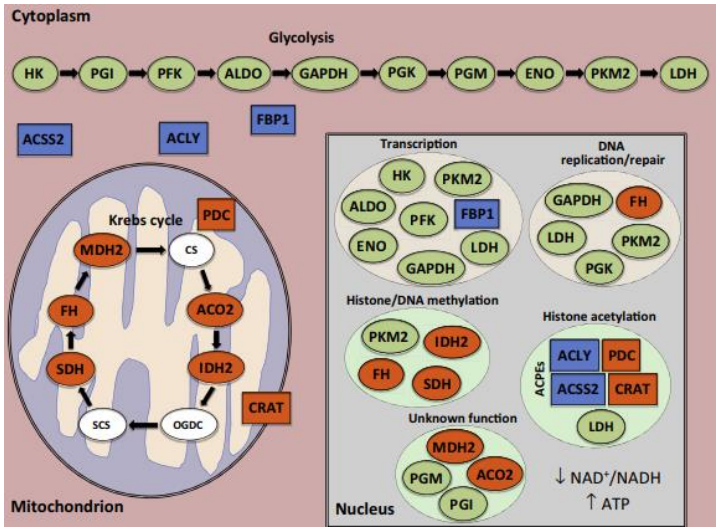


*Figure 2. Metabolic enzymes moonlighting in the nucleus. Cytoplasmic glycolytic and non-glycolytic enzymes are shown respectively in green and blue, mitochondrial enzymes are represented in orange or white. While all the essential glycolytic enzymes are found also in the nucleus, only some of the Krebs cycle enzymes have been described in the nucleus (orange) whereas others have not (white filling). In the nucleus, the translocated enzymes have been grouped together according to their proposed nuclear function, whether canonical or non-canonical. Abbreviations: ACLY, ATP citrate lyase; ACO2, aconitase 2; ACPEs, acetyl-CoA-producing enzymes; ACSS2, acetyl-CoA synthetase short-chain family member 2; ALDO, fructose bisphosphate aldolase; CRAT, carnitine acetyltransferase; CS, citrate synthase; ENO, enolase; FBP1, fructose-1,6-bisphosphatase; FH, fumarate hydratase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; HK, hexokinase; IDH2, isocitrate dehydrogenase 2; LDH, lactate dehydrogenase; MDH2, malate dehydrogenase 2; OGDC, /-ketoglutarate dehydrogenase; PDC, pyruvate dehydrogenase complex; PFK, phosphofructokinase; PGI, phosphoglycerate isomerase; PGK, phosphoglycerate kinase; PGM, phosphoglycerate mutase; PKM2, pyruvate kinase M2; SCS, succinyl-CoA synthetase; SDH, succinate dehydrogenase. The picture is taken from* (9).

An interesting example of this axis is provided by lactate dehydrogenase (LDH), which converts pyruvate into lactate under anaerobic conditions. Following the phosphorylation of Tyrosine 238, LDH translocates into the nucleus, where it binds to single stranded DNA (11) and interacts with the

DNA polymerases α, δ and ε, stimulating DNA synthesis in the post-UV radiation DNA repair (12). However, the mechanism that fine tunes this process is still not completely understood, and further investigations are required.

In parallel to the metabolism-gene transcription axis, strong evidences supporting the existence of an RNA-Enzymes-Metabolites network (recently named "REM theory") have also raised (13), thus connecting cell metabolism to RNA biology and gene expression.

The combination of *in vitro* studies and high throughput analysis has allowed to identify hundreds of novel RNA binding proteins (RBPs) (14). Among them, different enzymes involved in the tricarboxylic acid cycle, glycolysis, pentose phosphate cycle, fatty acid metabolism, pyrimidine synthesis and other metabolic processes have been found (reviewed and listed in (15)), showing RNA binding activity often with significant specificity for defined sequences and/or structural motifs. It is intriguing to notice that most of the identified proteins were not previously related to RNA biology, since they completely lack the classical well characterized RNA binding domains (RBDs) such as RNA recognition motif (RRM), hnRNP KH homology (KH) and DEAD box helicase domains (16).

A striking example is given by glyceraldehyde-3-phosphate dehydrogenase (GAPDH), a key enzyme in glycolysis that has been found to moonlight as a regulator of gene expression through its RNA binding activity. GAPDH oxidizes glyceraldehyde-3-phosphate to reduce nicotinamide adenine dinucleotide ($NAD^+$). In resting T cells, which rely on oxidative phosphorylation for energy generation, GAPDH binds to AU-rich elements in the 3' untraslated region (3'UTR) of the cytokine mRNAs, including interferon γ, inhibiting their translation. Upon activation, the metabolism of T

lymphocytes shifts from oxidative phosphorylation to aerobic glycolysis and GAPDH disengages from RNA, de-repressing the cytokine production (17). The switch between GAPDH regulatory and enzymatic activity is due to the competition between the RNA and the $NAD^+$ for the binding to the same site. Indeed, the globular Rossman-fold domain, specialized to bind nucleotide enzymatic cofactors, has been recognized as a common unconventional RBD (18). Therefore the cell metabolic state (i.e. the levels of $NAD^+$ and ATP) fine tunes the usage of the GAPDH Rossman fold between its RNA binding (inactive in glycolysis) and $NAD^+$ binding (active in glycolysis) activity *in vivo* (19).

Even if the above examples of metabolism-gene expression axis and REM regulation are intriguing, many key questions remain unanswered. Indeed, with a few exceptions, it's not clear why metabolic enzymes bind to nucleic acids as well as what are the biological consequences of these interactions and how metabolites can mediate such processes (20).

Given their biological and evolutionary differences, each single metabolic enzyme deserves a dedicated analysis of its nucleic acids binding feature. Not surprisingly, proteins belonging to the same metabolic pathway, such as LDH and GAPDH, display significant differences in their nucleic acids binding properties, highlighting the existence of various levels of regulation that need to be carefully investigated.

With the technical advances of the 'omics' age many of these questions can now be addressed far more easily and comprehensively than when most of these bifunctional proteins were firstly described.

### 1.3 Serine hydroxymethyltransferase (SHMT)

Serine hydroxymethyltransferase (SHMT) is a ubiquitous pyridoxal-5'-phosphate(PLP)-dependent enzyme, highly conserved in eukaryotes and prokaryotes. It is known to catalyze the reversible retroaldolic cleavage of L-serine into glycine, transferring the one carbon unit derived from the hydroxymethyl group of serine to (poly/mono)glutamylated tetrahydrofolate (THF), thus producing (poly/mono)glutamylated 5,10-methylenetetrahydrofolate (Me-THF) (Fig. 3).

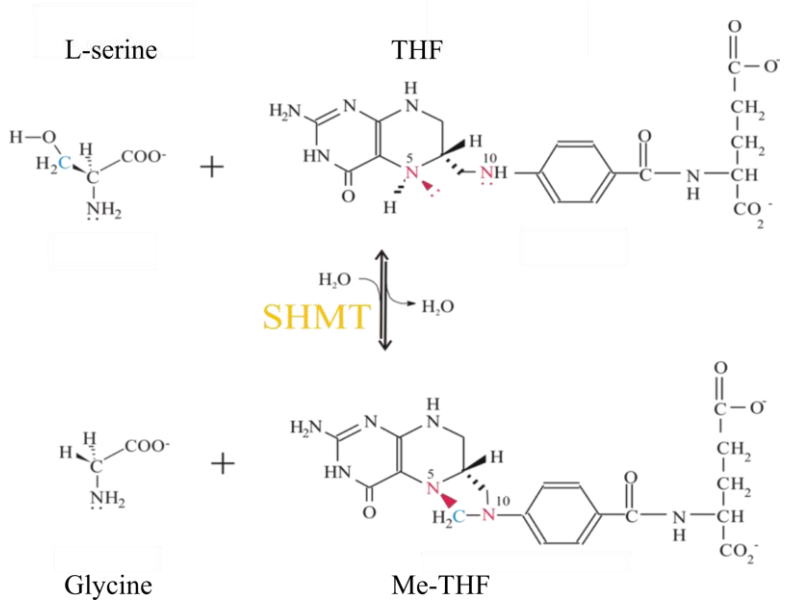


*Figure 3. Reaction catalyzed by SHMT with L-serine or glycine and monoglutamylated folates as substrates.*

Based on its fold, SHMT has been classified in the fold type I family (or aspartate aminotransferase family), indeed, as for the other members of this

group, each enzyme subunit binds a molecule of PLP and is folded into two domains. The larger domain is located at the N-terminal and is composed by seven-stranded β-sheets, while the small C-terminal domain is made by three-stranded β-sheets, covered with helixes on one side (21). These subunits associate into dimers in prokaryotes and tetramers in eukaryotes.

The human enzyme is a homotetramer and its architecture is described as "dimer of obligate dimers'' (22), where the so-called obligate or tight dimer represents the minimal catalytically active unit existing in solution (Fig. 4). Indeed, in the obligate dimer the active site is located at the monomers interface, being surrounded by residues belonging to both subunits of the dimer (23).

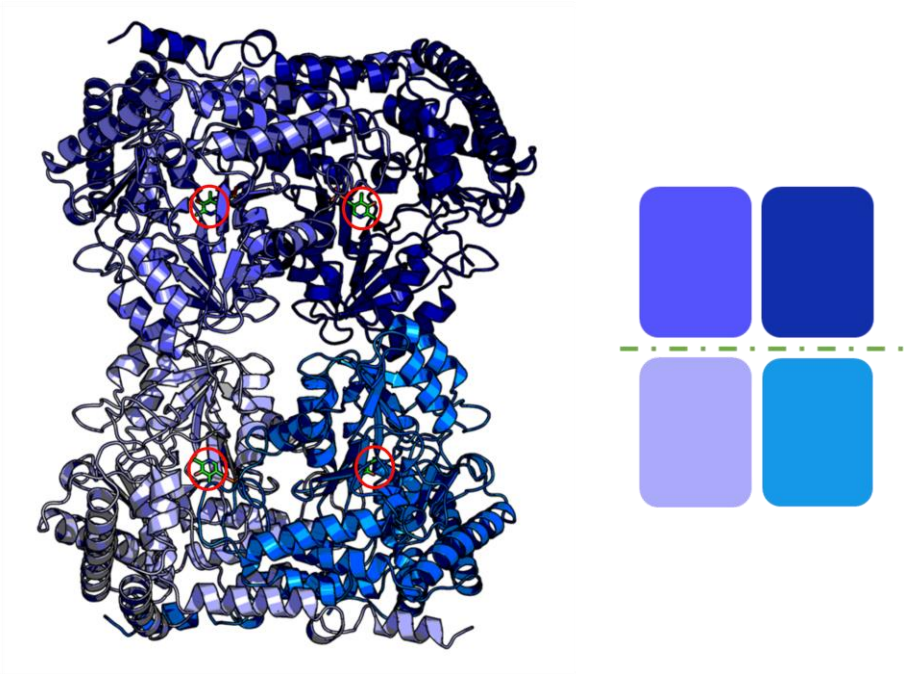

*Figure 4. Three-dimensional (left picture) and schematic (right picture) representation of the tetrameric structure of human serine hydroxymethytransferase (PDB: 1BJ4). Each monomer is colored with a different nuance of blue, one obligate dimer is composed by the darker subunits (upper part of the figures), while the other one is formed by the monomers colored in light blue (lower part of the figures). In the 3D structure PLP is colored in green and highlighted with a red circle. In the schematic representation the tetrameric interface is highlighted (green dashes), joining together the two obligate dimers.*

The human genome contains two *shmt* genes, *shmt1* and *shmt2*, encoding three main SHMT isoforms that differ in sequence and/or localization: one mitochondrial isoform, named SHMT2, and two cytosolic isoforms, SHMT1 and SHMT2α, the latter lacking the mitochondrial import signal present in SHMT2 and thus localizing in the cytosol. The isozymes are similar but not identical: they share the 66% of amino acid sequence identity, however, despite the high sequence similarity, SHMT1 and SHMT2 in the apo-form (without PLP) have different quaternary structures. Indeed, apo-SHMT1 has a

stable conformation that allows the assembly of the tetramer, whereas apo-SHMT2 is less stable, displaying a dimeric structure which only upon PLP binding converts into the holo-tetramer (23). The catalytic properties of SHMT1 and SHMT2 are similar, although their optimal reaction rate occurs at different pH; such difference may represent an adaptation of the two isoforms to their respective cellular environments (i.e. cytosol and mitochondrion) (24). SHMT has been widely studied given its central role in the one-carbon (1C) metabolism, a metabolic pathway needed to activate and transfer 1C units thus supporting nucleotide synthesis, methionine synthesis, folate homeostasis and amino acids metabolism. The 1C metabolism is compartmentalized within cytosol and mitochondrion with dedicated enzymes (Fig. 5); the reason for such compartmentalization is still puzzling, although some clues are rising thanks to the study of metabolic fluxes (25).

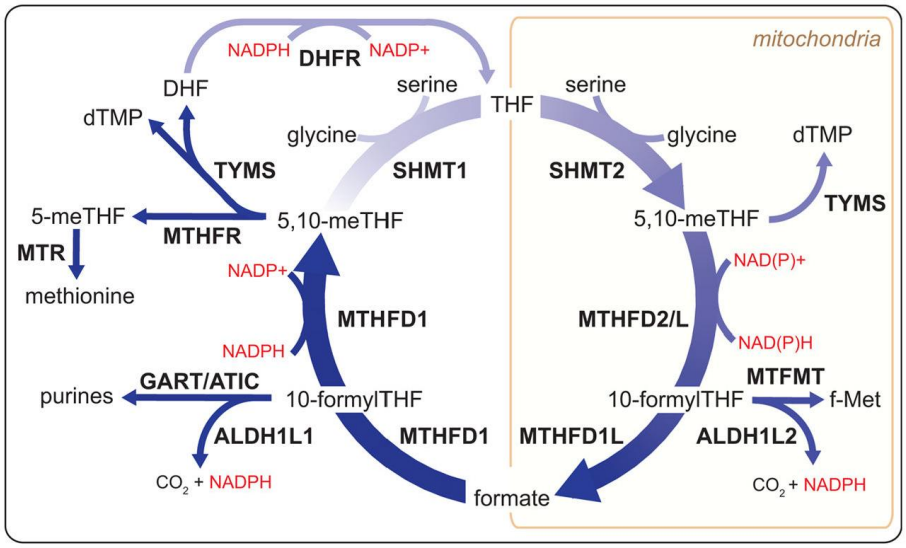


*Figure 5. Compartmentalization and products of one-carbon metabolism. All abbreviations are standard gene names. Certain descriptions utilize the common protein name for clarity. SHMT1/2, serine hydroxymethyl transferase, cytosolic(1) and mitochondrial (2); MTHFD1, methylenetetrahydrofolate dehydrogenase, cyclohydrolase, and formyltetrahydrofolate synthetase 1; MTHFD2/L, methylenetetrahydrofolate dehydrogenase 2/2-like; MTHFD1L, monofunctional tetrahydrofolate synthase, mitochondrial; MTFMT, mitochondrial methionyl-tRNA formyltransferase; TYMS, thymidylate synthetase; MTHFR, methylenetetrahydrofolate reductase; MTR, methionine synthase; DHFR, dihydrofolate reductase; GART, phosphoribosylglycinamide formyltransferase; ATIC, 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase; ALDH1L1/2, cytosolic (1) and mitochondrial (2) 10-formyltetrahydrofolate dehydrogenase. The picture is taken from* (26).

In the mitochondrion, the SHMT2 isozyme is mainly involved in the serine-derived production of both glycine and 1C units necessary to fuel the local synthesis of thymidine monophosphate (dTMP) and antioxidant molecules (NADPH and glutathione) as well as the nucleotides synthesis in the cytosol (25, 27). Indeed, 1C units are released from the mitochondrion as formate, which then fuels the cytosolic synthesis of purines and pyrimidines (28). Recent reports demonstrated that the reaction catalyzed by SHMT2 is also

required for the maintenance of specific mitochondrial transfer RNA (tRNA) pools, since it affects the levels of formylmethionyl-tRNA and other methylated tRNAs, thereby controlling the translation of mitochondrial respiratory complexes (29, 30).

By contrast, in the cytosol SHMT1 is likely to have a more versatile role, since it can switch the directionality of the reaction depending on cell types and metabolic needs. Indeed, metabolic studies on proliferating cells showed that the most frequent trend of the SHMT1 reaction is directed toward the production of serine and THF rather than glycine and Me-THF. However, it has been demonstrated that not all the cell lines follow such trend and, given the plasticity of the 1C metabolism, the reaction can be reverted in case of defects in the mitochondrial 1C metabolic flux (25, 31).

SHMT1, like many others metabolic enzymes (see section 1.2), has been found to localize also in the nucleus. Interestingly, reports by Anderson D.D. et al. (32, 33) pointed out that such localization occurs only during the S-phase of the cell cycle, allowing SHMT1 to participate to DNA replication and/or repair through nuclear thymidylate biosynthesis. However, the molecular mechanism through which this happens is still not completely clear and will be further investigated in the present thesis.

SHMT2α is the least characterized isozyme among the three, its metabolic function is believed to be redundant with the one of SHMT1, since it undergoes the same nuclear localization (33). Interestingly, recent works report a role of this isoform in the control of inflammatory cytokine signalling, as it takes part to the deubiquitylating BRCC36 isopeptidase complex (BRISC) (34). Of note, BRISC crystal structure revealed that SHMT2α participates to the complex as apo-dimer (23, 35). Overall the function of SHMT2α, either metabolic or moonlighting, needs to be further clarified.

The reaction catalyzed by SHMT isoforms is pivotal not only in cell physiology but also in cell pathology: indeed, it is at the basis of cancer metabolic reprogramming since it is essential for the rapid proliferation of neoplastic cells. Serine metabolism is often dysregulated in tumors (36) and, not surprisingly, cancer cells often overexpress one or more SHMT isoforms, indeed both isozymes are considered chemotherapic targets (37, 38). A lot of efforts have been done to develop small molecules able to selectively inhibit the cytosolic or the mitochondrial SHMT isoform; to date, a couple of effective small molecules having a nanomolar affinity for the enzymes have been found, although they are still far from the clinical application (39, 40).

SHMT1 and SHMT2 are subjected to multiple layers of regulatory events inside the cell. At the transcriptional level both isozymes are regulated by the c-Myc transcriptional factor, which binds to their promotors in normal and transformed cells (41); moreover, given the involvement of SHMT2 in the redox cell balance, the expression of this isoform is induced by hypoxic stress through the hypoxia-inducible factor 1 (HIF-1) to counteract the production of reactive oxygen species (42).

Different post-translational modifications (PTMs) have been characterized for both proteins as well. SHMT1 is subjected to an interplay between ubiquitination and SUMOylation: the basal homeostasis of the cytosolic protein is regulated by an ubiquitination on Lys48, which leads to a proteasome-mediated degradation of the enzyme. During the S phase and in response to UV damage, the enzyme is sumoylated (SUMO-1) and imported in the nucleus where, following the removal of such modification, SHMT1 can be either stabilized by ubiquitination and exported again into the cytoplasm or, if the export is impaired, the protein undergoes a nuclear sumoylation (SUMO-2/3) which is thought to lead to nuclear degradation (Fig.6 and (43)).

A peculiar PTM has been recently described for mitochondrial SHMT2: the succinylation. When succinylated, the dimeric form of the enzyme is mostly populated, resulting in a very low enzymatic activity; when cells are in proliferation or under metabolic stress, the enzyme is desuccinylated by SIRT5 thus allowing the tetramer formation which is required for the activation of the catalysis (44).

Overall human serine hydroxymethytransferase is a good example of how metabolic enzymes involved in the intermediary metabolism are subjected to a multi-layered regulation, needed to fine tune their multiple functions in the crowded cellular environment.

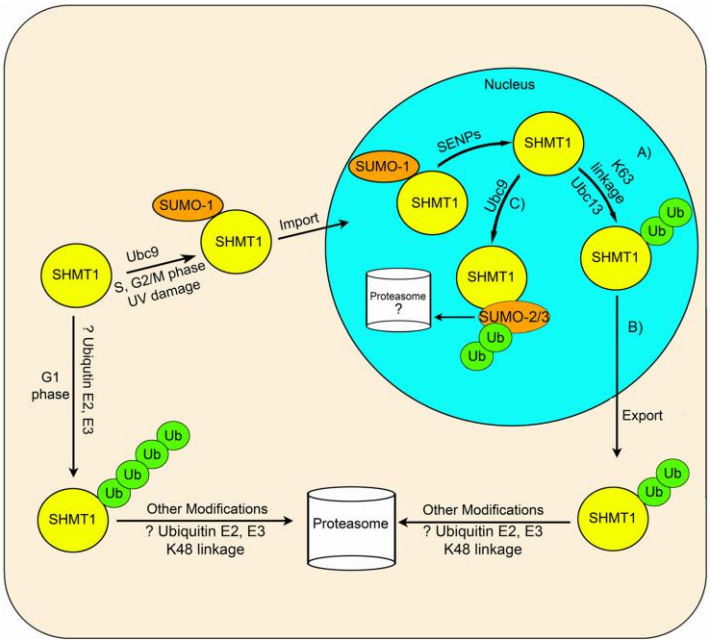


*Figure 6. Proposed model for the interplay between ubiquitination (Ub) and SUMOylation (SUMO) of SHMT1. When known, the name of the enzymes that perform the modifications are reported. The picture is taken from* (43).

### 1.4 The crosstalk between cytosolic and mitochondrial SHMT

A peculiar link between SHMT1 and SHMT2 expression has been observed in some normal and tumor cells.

Literature reports a general trend in which, among different cancer types, SHMT2 is usually the most overexpressed isozyme with respect to SHMT1 (42). However there are some exceptions: in lung cancer tissues and in lung cancer tumor initiating cells it has been observed also a significant upregulation of the SHMT1 expression (37, 45), suggesting a major role of this isoform in such models.

In this background, Paone et al. reported for the first time a crosstalk between SHMT1 and SHMT2 in lung cancer cells (37). Indeed, it was observed that the knockdown of SHMT1 in H1299 triggered apoptosis and induced a compensatory increase in the SHMT2 protein expression (Fig. 7), whereas the knockdown of the mitochondrial isoform did not have any effects on the cytosolic counterpart (37). It was clear that the expression of SHMT1 and SHMT2 was somehow interconnected, but the molecular mechanism at the basis was completely unknown.

Two years later a cross regulation between SHMT isozymes was observed also in a normal cell model (i.e. HEK293T), although with a completely different trend (Fig. 8). In fact, in this case, the deletion of SHMT1 decreased SHMT2 protein levels, whereas the deletion of SHMT2 increased SHMT1 expression (25).


*Figure 7. Western blot analysis of SHMT2 protein expression in H1299 cell line. iscr and iSHMT1 indicate, respectively, the samples treated with scrambled RNA interference (RNAi) or with SHMT1 RNAi. The picture is taken from* (37).

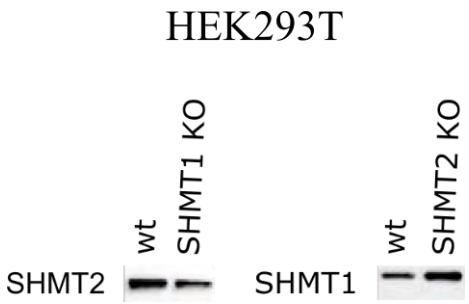


*Figure 8. Western blot analysis of SHMT2 and SHMT1 protein expression in HEK293T WT or knockout (KO) for SHMT1 or SHMT2. The picture is taken from* (25).

How the cell is able to sense the variation of each isozyme and how it modulates in response the expression of the other counterpart still remains an open question.

## 1.5 The hypothesis: SHMT1 as a putative nucleic acid binding protein

Two decades ago an interesting report showed the first evidences that SHMT1 was able to bind RNA, and that it could potentially behave as an RNA binding protein inside the cell. In their paper *Liu et al.* performed an *in vitro* analysis, showing that the cytosolic isozyme binds to the 5'UTR of its own mRNA, inhibiting its translation (46).

In line with these observations, a similar phenomenon was observed for other two enzymes that, together with SHMT1, participate to the thymidylate cycle. Indeed, it was demonstrated that human thymidylate synthase (TYMS) and dihydrofolate reductase (DHFR) can bind to the coding region of their mRNA *in vitro*, repressing their translation (47, 48). Therefore for all the three enzyme it has been proposed a mechanism of post-transcriptional autoregulation; however, since all the mentioned studies were performed *in vitro*, a physiological characterization of such interactions *ex vivo* or *in vivo* is still missing.

It may seem weird to notice that in the paper mentioned above SHMT1 was shown to have non-specific RNA binding properties, since it was able to bind also to luciferase mRNA (46). Although the terms “specific” and “non-specific” are widely used, the specificity, or the lack thereof, is a considerably more nuanced problem than suggested by this binary classification (49). In fact it’s common to see *in vitro* RBPs bound to random RNA sequences, and the same phenomenon has been observed also for DNA binding proteins (DBPs) (50). To better describe the entire spectrum of the existing affinities for a given DNA or RNA binding protein it is useful to employ the affinity distributions, i.e. the distribution of affinities of a DBP or an RBP for all the substrate variants *in vitro* (Fig. 9).

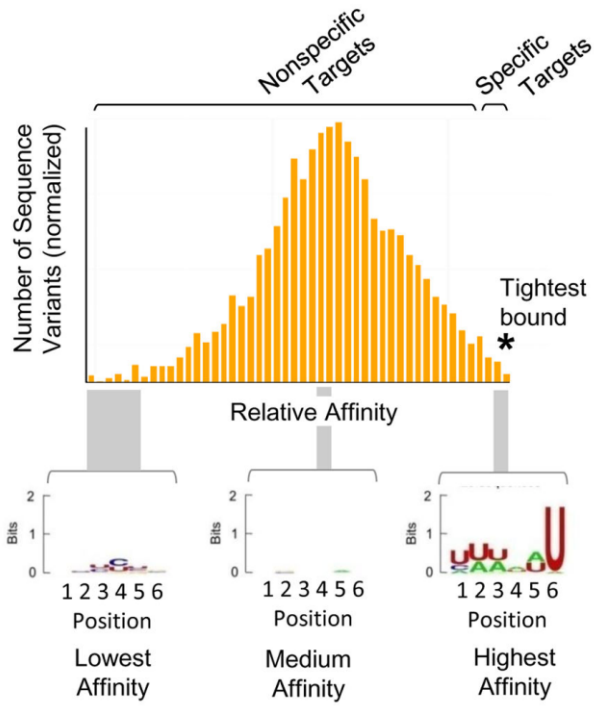


*Figure 9. Affinity distribution of an RBP (C5 from E.coli) with a binding site for 6 nucleotides. Histogram of relative affinities (log scale) for 4096 sequence variants (all the possible combinations made of 6 nucleotides). Relative affinities are calculated in relation to a standard variant, which can be chosen freely. "Specific" RNA targets are marked by the asterisk and cluster in the high affinity region of the distribution, producing a binding consensus sequence, shown as a logo underneath the plot. The remainder of the distribution consists of "non-specific" RNA variants, which do not produce a consensus motif. The picture is taken from* (49).

For proteins with canonical specificity, cellular substrates appear to fall mainly into the high affinity region of the affinity distribution (tail), instead the so called non-specific nucleic acid binding proteins have physiological substrates which fall in a distinct part of the affinity distribution (near the median). In other words, this means that a considerable fraction of nucleic acid binding proteins could potentially bind a wide array of different sequences *in vitro*, but *in vivo* the cellular environment mostly favors only the interaction with their

physiological targets.

The interaction between a protein and a RNA/DNA sequence in the cell is dictated by multiple factors: the affinity of the protein for the RNA/DNA site, the concentration of the protein, the RNA or DNA accessibility, the competition from other sequences for the binding site, the need for additional interacting partners, the metabolic control of the interaction and so on.

Indirect evidences led to hypothesize that SHMT1 could be able to bind to the DNA as well. As mentioned in Paragraph 1.3, SHMT1 has been found to localize into the nucleus in the S phase of the cell cycle and, interestingly, TYMS and DHFR display the same behavior. Small scale tandem purification experiments have shown that nuclear TYMS and DHFR coprecipitate with SHMT1, but only when the DNA is intact (32). Moreover, among SHMT1 nuclear interacting partners, several proteins involved in the DNA replication machinery such as proliferating cell nuclear antigen (PCNA), replication factor C activator 1 (RFC1) and DNA polymerases ε and δ have been found and, interestingly, there are preliminary evidences which suggest the enrichment of SHMT1 on the mammalian origins of replication (32, 33).

All these evidences support the idea that a ternary complex formed by the enzymes of the thymidylate cycle (SHMT1-TYMS-DHFR) may exist and it may localize on specific DNA sites thanks to the DNA binding capability of SHMT1, probably in order to localize *de novo* thymidylate biosynthesis at the sites of DNA synthesis.

It may seem unusual to find a protein that binds to both nucleic acids, as RNA and DNA binding are typically considered distinct functions and thus studied as different properties. However, the consideration of the DNA and RNA binding functions within proteins as separate entities is becoming outdated, since a lot of DNA and RNA binding proteins have recently been characterized

(namely DRBPs): these proteins bind to DNA and RNA through a mutual exclusive mechanism or simultaneously (51, 52).

### 1.6 Aim of the thesis

It has become evident that SHMT1 is a multitasking protein subjected to a complex regulation which can vary in different cellular milieu and/or metabolic conditions.

Among the multiple properties of SHMT1, the present thesis aims to focus on the RNA and DNA binding activity (Fig. 10). We are facing two moonlighting functions that need a deep biochemical, molecular and cellular characterization to understand their physiological and/or pathological relevance inside the cell, to date unknown.

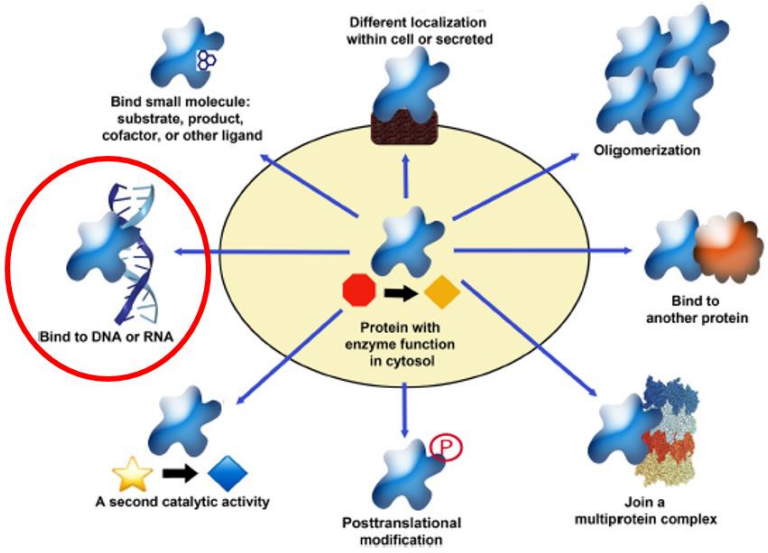


*Figure 10. Schematic representation of the multiple properties of human cytosolic serine hydroxymethytransferase (SHMT1), the subject of the present thesis is circled in red.*

In such background, the present work firstly investigates the molecular mechanism at the basis of the crosstalk between SHMT isozymes observed by us in lung cancer cells, understanding if it is mediated by the moonlighting RNA binding activity of SHMT1.

The second goal is a deep biochemical and molecular characterization of the SHMT1-RNA interaction focused on: (i) the discovery of the non-canonical RNA binding domain, (ii) the effects of the SHMT1-RNA interaction on the catalytic activity of the enzyme and (iii) the metabolic effects of such interaction in lung cancer cell lines. In addition, to understand the biological meaning of the moonlighting activity of SHMT1, as well as why the protein has evolved such capacity, this work tries to go deeper into the discovery of all the sequence variants bound, and so potentially regulated, by the cytosolic isozyme in lung cancer models, in this way widening the knowledge on lung cancer metabolic regulation.

In parallel, the present thesis aims to shed light also on the DNA binding properties of SHMT1, mainly focusing on the biochemical characterization of such activity together with the analysis of the DNA sequence variants bound by the protein.

Overall the present work fits well in the emerging network which interconnects intermediary metabolism, gene transcription and gene expression through its main players: metabolic enzymes, nucleic acids and metabolites.

# 2. Materials and Methods

## 2.1 Analysis of SHMT1-nucleic acids interaction

### 2.1.1 Protein expression and purification

Wild-type SHMT gene was cloned into a pET22b(+) vector (Novagen) and expressed as N-terminal histidine-tagged fused protein in *Escherichia coli* (BL21-DE3). Mutant SHMT constructs were obtained by site-directed mutagenesis of the WT gene (see section 2.1.2). Bacterial cultures were grown at 37 °C in Luria-Bertani (LB) liquid medium supplemented with 30 μg/ml kanamycin. When $OD_{600}$ reached about 0.8, the temperature was lowered to 22 °C and the protein expression was induced by adding 0.1 mM IPTG (isopropyl B-d-thiogalactoside). Cells were harvested 20 hours later by centrifugation and stored at -20 °C. Bacterial pellets were then resuspended in lysis buffer (see Table 1) and lysed by sonication. After centrifugation, the soluble protein was purified by affinity chromatography on a 5 ml HP-chelating column (GE Healthcare) loaded with nickel ions equilibrated with buffer A (see Table 1). The elution was carried out with an imidazole gradient, SHMT proteins were eluted with 300 mM imidazole. Fractions containing pure protein were pooled, and imidazole was removed with desalting columns PD10 (GE Healthcare), the histidine tag was then removed by proteolytic digestion with 1U/mg of thrombin at 16 °C for 15 h. The digestion mix was then loaded again on the HP-chelating column and thrombin was eluted in the flow thought, whereas SHMT, which now binds with low affinity, was eluted with 100 mM imidazole. The protein was then concentrated and injected into a Superdex 200

column (either 10/300 or 16/600; GE Healthcare) and eluted with buffer A.

*Table 1. Buffers utilized in the purification of SHMT1 (WT and mutants).*

| | |
|---|---|
| Lysis Buffer | 20 mM Hepes, 100 mM NaCl, 5% glycerol, 1 mM PMSF |
| Buffer A | 20 mM Hepes, 100 mM NaCl*, 5% glycerol |

*The purification of SHMT1 K157S/K158S was carried out with 200 mM NaCl.

Protein concentration was determined measuring the absorbance at 280 nm and applying the Beer-Lambert law (SHMT1 $\varepsilon_{280}$ 47565 $M^{-1}cm^{-1}$). Samples were snap freezed and stored at -20 °C.

2.1.2 Site-directed mutagenesis

SHMT mutants were produced by site-directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). The pET28b vector containing the wild-type SHMT1 gene was used as template DNA. Two complementary oligonucleotides, synthesized by Metabion (Steinkirchen, Germany), were used as primers for the mutagenesis reactions (primers listed in Table 2). The DNA was then transformed in *Escherichia coli* Neb5α cells for amplification, plasmids were purified with QIAprep Spin Miniprep kit (Qiagen). Mutagenesis was then verified by sequencing.

*Table 2. Primers used in the mutagenesis reactions (written 5' to 3').*

| | |
|---|---|
| SHMT1<br>K157S/K158S | Fw ccatgggttcatgacagacagcagcaaaatctctgccacgtccatc<br>Rv gatggacgtggcagagattttgctgctgtctgtcatgaacccatggg |
| SHMT1 Y82A | Fw ggtacccgggccagagagcctatggcgggact<br>Rv agtcccgccataggctctctggcccgggtacc |

### 2.1.3 *In vitro* transcription and RNA biotinylation

pGL3 plasmids (pGL3 Control) containing the specific UTRs cloned upstream of the luciferase gene were provided by GeneArt/Invitrogen (Ratisbonne, Germany).

UTR sequences were amplified by polymerase chain reaction (PCR) from pGL3 plasmids using primers designed to include the T7 promoter sequence upstream of the UTR in the PCR product (Table 3). PCR products were run on a 1% agarose gel in 1x TAE buffer (Tris-Acetic acid-EDTA) and then purified using the Nucleospin Gel and PCR Clean-up Kit (Macherey-Nagel). The DNA fragments generated with this process were used as templates to produce the corresponding RNA sequences with the Ribomax Large Scale RNA Production System-T7 Kit (Promega). The RNA was purified according to the manufacturer instructions.

The UTR sequences were labelled at the 3' end with a molecule of cytidine-5'-phosphate-3'-(6-aminohexyl)phosphate labelled with biotin (pCp-biotin) using the Pierce™ RNA 3'End Biotinylation Kit (Thermo Scientific) and purified according to the manufacturer instructions.

*Table 3. Primers used in the PCR reactions (written 5' to 3'), the forward primers contain the T7 promoter sequence at the 5' terminal.*

| | |
|---|---|
| UTR1 | Fw taatacgactcactatagggcctggcgcgcagagtgcacc<br>Rv tgcactggttcgaagctgcc |
| UTR2 | Fw taatacgactcactatagggataaagaaaaaagcggtgag<br>Rv cgcaactcggaagtcgcaggag |
| UTR2α short | Fw taatacgactcactatagggacggactgctaaaggtctcccc<br>Rv cctgaccagctgcccacatctctg |
| UTR2α intermediate | Fw taatacgactcactatagggtttcacgtggcattaggggag<br>Rv cctgaccagctgcccacatctctg |
| luciferase | Fw taatacgactcactatagggatggaagacgccaaaaacataaag<br>Rv gtggtttgtccaaactcatc |
| UTR2-luciferase | Fw taatacgactcactatagggataaagaaaaaagcggtgag<br>Rv gtggtttgtccaaactcatc |

### 2.1.4 Electrophoretic mobility shift assay (EMSA)

Gel mobility shift assays were conducted by incubating a fixed amount of RNA or DNA with fixed or increasing concentrations of recombinant human SHMT1. For competition assays increasing concentrations of unlabeled competitors UTRs were added to the reaction mixture. To determine the effect of metabolites upon RNA binding, the enzyme was pre-incubated with an excess of serine or glycine and 5-CHO-THF-$Glu_n$ (monoglutamylated: (6S)-5-formyl tetrahydrofolate and pentaglutamylated: (6R,S)-5-formyl-5,6,7,8-tetrahydropteroypenta-γ-L-glutamic acid lithium salt) for 5 minutes at room temperature, the latter being a substrate analogue stable in an oxidative environment.

All the components were incubated at room temperature for 30 minutes in 12 µl of binding buffer (20 mM HEPES pH 7.4, 100 mM NaCl) containing 20 µg/ml bovine serum albumin (BSA), 1 µg of tRNA, 1 U/µl of recombinant Rnasin ribonuclease Inhibitor and 8% (v/v) glycerol. The reaction mixtures were separated onto 4% nondenaturing polyacrylamide gels in 0.5× TBE buffer (45 mM Tris-Borate, 1 mM ethylenediaminetetraacetic acid pH 8.6). For the visualization of unlabeled sequences, gels were stained with SYBR Safe (Invitrogen) in 30 ml of 0.5× TBE. To detect biotin labelled RNA, gels were transferred onto Immobilon-Ny+ positive charged nylon membranes (0.45 µm) in 0.5× TBE, the RNA was then crosslinked to the membrane using a hand-held UV lamp equipped with 314 nm bulbs. The biotin moiety was detected by chemiluminescence using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific). In both experiments, images were acquired using Chemidoc MP Imaging System (Bio-Rad). Densitometric measurements of the free RNA bands were transformed into percentages, the amount of bound RNA (% bound) was then calculated by subtracting the percentage of free RNA from the total, the result was plotted as a function of protein concentration. The apparent dissociation constants ($K_d^{app}$) were estimated fitting the data with Equation 1, in which $B_{max}$ corresponds to the maximum binding (100%) and [P] corresponds to the concentration of SHMT1 in the reaction mixture. All experiments were run in triplicate.

Equation 1

$$\%\, Bound = \frac{B_{max}[P]}{K_d + [P]}$$

### 2.1.5 RNA Immunoprecipitation (RIP)

Cells were grown in complete RPMI-1640 medium (see section 2.3.1) and harvested at 80% of confluence spinning at 1000 rpm for 5 minutes at 4 °C, pellets were then snap freezed and stored at -80 °C.

Each pellet was lysated in 1 mL of RIP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 0.5% NP40, 1mM DTT and 100U/ml Ribolock Rnase Inhibitor), 10 U of TURBO Dnase (Invitrogen) were added to each lysate to avoid DNA contamination. Cellular debris were removed by centrifugation and protein quantification was carried out following the absorbance at 280 nm. In parallel, 2 μg of antibody, either mouse monoclonal anti-SHMT1 purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA) or mouse polyclonal IgG (Sigma), were conjugated to 100 μl of Protein G Dynabeads (Invitrogen), rotating for 1 hour at 4°C. The immunoprecipitation was performed incubating 500 μg of protein lysate with the beads conjugated to the antibody, rotating the samples o/n at 4 °C.

Three stringent washing were performed the day after with RIP wash buffer (25 mM Tris-HCl pH 7.4, 150 NaCl, 5 mM EDTA, 1 mM DTT and 0.5% NP40) followed by one last wash in phosphate saline buffer (PBS). The beads were then resuspended in 1 ml of QIAzol reagent (Qiagen) and processed according to the manufacturer instructions to extract RNA.

One part of the extracted RNA was used to prepare libraries suitable for next generation sequencing (see next paragraph).

On the other hand, to assess the enrichment of the SHMT2 mRNA, the same volume of each sample (8 μl) was retrotranscribed with NZY First-Strand synthesis kit (Nzytech) following the manufacturer instructions.

The cDNA (2 μl) was then used as template to perform a semi-quantitative

PCR using NZYTaq II 2x Green Master Mix (Nzytech) according to the manufacturer instructions.

2.1.6 Libraries preparation for next generation sequencing (NGS)

To sequence all the RNAs recovered with the RNA Immunoprecipitation (RIP-Seq) it is necessary to generate a library of cDNAs having adaptor sequences containing indexes suitable for sequencing. For this purpose it was used the TrueSeq RNA Library Prep Kit v2 provided by Illumina (San Diego, CA, USA), the libraries preparation was performed according to the manufacturer instructions. The libraries quality control was done with Agilent 2100 Bioanalyzer using a DNA-specific chip (Agilent DNA 1000).

**2.2 Analysis of SHMT1 enzymatic activity**

2.2.1 *In vitro* enzymatic assays

All assays were carried out at 30 °C in 20 mM KPi buffer at pH 7.2, when working with RNA the buffer was treated with diethyl pyrocarbonate (DEPC). The serine cleavage reaction was measured using 0.2 μM enzyme with L-serine and THF as substrates by means of a spectrophotometric coupled assay, in which the 5,10-$CH_2$-THF produced by the reaction was oxidized by the NAD-dependent *E. coli* 5,10-$CH_2$-tetrahydrofolate dehydrogenase (53). The possible inhibition of methylenetetrahydrofolate dehydrogenase by RNA was checked in an experiment in which the activity of the enzyme (1 μM) with 50 μM 5,10-$CH_2$-THF and 250 μM $NADP^+$ was measured in the presence and absence of tRNA and UTR2. No inhibition was observed under these

conditions.

Two assays were used to analyze the inverse reaction. The first assay uses glycine and 5-CHO-THF as substrates and is based on the spectrophotometric measurement of the quinonoid intermediate that develops when both ligands bind to SHMT, forming an enzyme–glycine–folate ternary complex. The quinonoid intermediate, which yields an intense absorption band with a maximum at 502 nm, derives from deprotonation of glycine, but it accumulates to a measurable extent only when a folate ligand is also bound to SHMT and a ternary complex is formed (54). Therefore, the absorbance at 502 nm is proportional to the fraction of enzyme present as the ternary complex. The second assay is a coupled assay carried out in the presence of 5,10-$CH_2$-THF and glycine substrates. The L-serine product of this reaction is oxidized by a $NAD^+$-dependent serine dehydrogenase.

The inhibition curves were obtained keeping fixed concentration of substrates (10 mM glycine, 10 mM L-serine, 5 μM 5-CHO-THF and 80 μM THF, 90 μM 5,10-$CH_2$-THF) while varying the RNA concentration. All obtained inhibition curves were fitted to Equation 2, in which [I] corresponds to the RNA concentration.

Equation 2

$$\%Activity = 100\left(1 - \frac{[I]}{[I] + IC_{50}}\right)$$

The characterization of the tRNA inhibition mechanism was carried out using serine and THF as substrates. In a first series of experiments, L-serine concentration was held constant at 10 mM, while THF was varied between 10 and 488 μM. The single saturation curves obtained varying THF at various

RNA concentrations were fitted to a modified Michaelis–Menten equation that accounts for substrate inhibition (Equation 3).

Equation 3

$$v_i = v_{max}\frac{[THF]}{K_m + [THF]\left(1 + \frac{[THF]}{Ki_{THF}}\right)}$$

Moreover, all the saturation curves were globally fitted to a modified Michaelis–Menten equation that accounts for both substrate inhibition and hyperbolic competitive inhibition (Equation 4).

Equation 4

$$v_i = v_{max}\frac{[THF]}{K_m\left(\frac{1 + \frac{[tRNA]}{K_i}}{1 + \frac{[tRNA]}{\alpha K_i}}\right) + [THF]\left(1 + \frac{[THF]}{Ki_{THF}}\right)}$$

In these equations, $v_i$ stands for initial velocity, $V_{max}$ is the maximum velocity, $K_m$ is the Michaelis–Menten constant for THF, $K_{iTHF}$ is the substrate inhibition constant, $K_i$ is the RNA inhibition constant and α is the factor by which *K*i changes when THF is bound to the enzyme. In a second series of experiments, THF concentration was held constant at 80 μM, while L-serine was varied between 0.156 and 10 mM. The saturation curves obtained were fitted to a Michaelis–Menten equation.

All spectrophotometric measurements were performed using a Hewlett-Packard 8453 diode-array spectrophotometer. Fitting of data to equations was

carried out with the PRISM software (GraphPad, La Jolla, CA, USA).

### 2.2.2 Enzymatic assays in living cells

SHMT activity measurements in living cells were performed using a radioisotope assay based on the capability of SHMT to catalyze the exchange of the pro-2S proton of glycine with solvent (55, 56). After 24 hours from transfection, H1299 cells were harvested using trypsin and washed twice with 2 ml of PBS buffer to eliminate growth medium. Aliquots (100 μl) of cells suspension were incubated with [2−3H]glycine ($2 \times 109$ dpm $mol^{-1}$) (23 μM) for 4 hours at 37 ℃ . Subsequently, samples were centrifuged to remove the cells and reactions were stopped by the addition of 3% (w/v) trichloroacetic acid to remove radiolabeled glycine and measure radioactivity in the solvent. Control reactions were performed to correct for background exchange (these samples did not contain cells). All measurements were performed in duplicate and were normalized on the protein concentration.

## 2.3 Analysis of the cellular effects of SHMT1-RNA interaction

### 2.3.1 Cell lines

H1299 and A549 cancer cells were grown in RPMI-1640 medium supplemented with 2 mM L-glutamine, 100 IU/ml penicillin/streptomycin and 10% fetal bovine serum (FBS; Biowest USA). HAP1shmt2KO cells were obtained from Horizon Discovery Ltd., Cambridge, UK and maintained in IMDM medium (Gibco) supplemented with 100 IU/ml penicillin/streptomycin and 10% FBS (Biowest USA). The composition of the minimal medium

(MEM) is the following: Minimum Essential Medium Eagle (Sigma-Aldrich) supplemented with 1X MEM Vitamin Solution, 2 mM L-glutamine, 25 mM glucose, 10% fetal bovine serum dialyzed, 100 IU/ml penicillin/streptomycin; the supplements for the specific media are 0.4 mM serine, 0.4 mM glycine and 0.5 mM formate.

### 2.3.2 Cells transfection

H1299 and HAP1-SHMT2KO cells were transiently transfected with interference RNA sequences or DNA plasmids. In detail, $1x10^5$ cells were seeded and simultaneously incubated for 18 hours with the following components:

- Qiagen AllStars RNAi Controls (scrambled sequences),
- shmt1 siRNA sequences (1_6 CTGCTGTAAATCAGAAGTGTA, 1_7 CTGACGGAGCTGGGCTACAAA, 1_8 CTCCCGTAATCAGGAA GCCAA),
- shmt2 siRNA sequences (2_6 CAGGCGCAGCAAATTCAATTT, 2_7 CCGGGAGATCCCTTACACATT, 2_10 CCCAGCCAACCTGGCCGTCTA) (GeneSolution, Qiagen)
- pGL3 plasmids.

For the rescue experiments, 24 hours after the transfection, the medium was replaced with fresh medium containing the specified supplementations. For the fluorescent microscopy experiments, $1x10^5$ cells were seeded and transfected with 1 μg of SHMT1-GFP plasmid. The transfection medium was replaced with complete medium after 12 hours. The fluorescent signal was detected 24 hours after the transfection using an AXIO vert.A1 fluorescence

Zeiss microscope (Carl Zeiss Microscopy, Switzerland).
In all the set of experiments, JetPrime (Polyplus-transfection® SA) was used as transfection reagent following the manufacturer instructions.

### 2.3.3 RNA extraction and real-time qRT-PCR analysis

Total RNA was extracted from cells using TRIzol reagent (Invitrogen) following the manufacturer instructions. Reverse transcription reactions were performed on 1 μg of total RNA with SuperScript First-Strand Synthesis System Kit (Invitrogen). The obtained complementary DNA was used for the subsequent real-time PCR reaction: each PCR well contained 5 μl 2× KAPA Sybr Fast universal premix, 0.4 μl of forward and reverse primer, 1 μl cDNA (1:5 diluted) and 3.2 μl water (final volume 10 μl). Reactions were performed using Stratagene MX3000P (Stratagene, La Jolla, CA, USA). Each sample was analyzed in triplicate. Primer sequences are shown in Table 4.

*Table 4. Primers used in the Real Time qRT-PCR analysis (written 5' to 3').*

| | |
|---|---|
| SHMT1 | Fw aggaaaggagtgaaaagtgtggat<br>Rv gacaccagtgtcgctctggatctg |
| UTR2 | Fw ggcgaactacaattcccaaa<br>Rv actctggtctagggcagcaa |
| β-actin | Fw aggatggcaagggacttcctg<br>Rv aatgtggccgaggactttgat |

### 2.3.4 *Ex vivo* and *in vitro* luciferase assays

*Ex vivo* luciferase activity was determined after transfecting the pGL3 plasmids into the H1299 cells. Cells were cotransfected with 0.5 μg Renilla luciferase plasmid as control of transfection together with 1 μg of the pGL3, pGL3-UTR2 or pGL3-UTR1 plasmid plus scrambled RNA sequences or the

specific interference RNA (iRNA) against *shmt1*. The expression of Firefly and Renilla luciferases was analyzed 48 hours after transfection using the Dual-Luciferase® Reporter Assay System (Promega) in a VICTOR™ Multilabel Counter (Perkin Elmer, USA). Firefly luciferase was standardized to the value of Renilla.

*In vitro*, translation was measured as the synthesis of luciferase protein from the corresponding mRNAs using Flexi Rabbit Reticulocytes Lysate System kit (Promega). The mRNAs employed in the *in vitro* translation assays were produced from pGL3 plasmids as described in section 2.1.3 using the primers listed in Table 3. The luciferase mRNA with or without the upstream UTR2 sequence was heated 3 minutes at 65°C and immediately cooled down in an ice-water bath. The transcripts were incubated 5 minutes at room temperature with SHMT1 in presence of recombinant Rnasin Ribonuclease Inhibitor (40 U); the other reaction components were then added to obtain a final mixture containing 125 ng RNA, 0.6 μM SHMT1, 17.5 μl rabbit reticulocyte lysate, 20 μM amino acid mixture and 70 mM KCl. The translation reactions were incubated 30 minutes at 30°C and then stopped in ice. Luminescence was detected according to the manufacturer's protocol using VICTOR™ Multilabel Counter (Perkin Elmer, USA) equipped with 96 well plate reader.

### 2.3.5 Serine consumption assay

Sample extraction and derivatization. Serine was analyzed according to the method of Gao *et al.* (57) with slight modifications. Cells transfected as indicated were cultured in MEM originally supplemented as described before and specifically with 400 μM serine. After 48 hours, cells were harvested and

centrifuged to eliminate cellular debris. Then, 450 µl cell culture medium was spiked with 20 µl of 0.7 mM 2-(4-chlorophenyl)ethylamine as an internal standard. The first step of derivatization was performed by adding 350 µl of ethanol/pyridine (6:1) and 50 µl of ECF. The reaction mixture was then vortexed for 30 seconds and ultrasonicated for 60 seconds to increase the speed of reaction at room temperature. The derivatization products were extracted with 700 µl of n-hexane for 60 seconds and centrifuged for 5 minutes at 3000 rpm. The organic layer was removed and the aqueous phase was adjusted to pH ≥12 with 0.1 ml 7M NaOH. The second derivatization step was performed by adding a further 50 µl of ECF and 700 µl of n-hexane. The reaction mixture was again vortexed and ultrasonicated as before, and then centrifuged at 3000 rpm for 10 minutes. The organic layers were combined, dried under N2, resuspended in dichloromethane and analyzed by GC/MS gascromatography-mass spectrometry (GC/MS).

GC-M analysis was performed using an Agilent 6850A gas chromatograph coupled to a 5973N quadrupole mass selective detector (Agilent Technologies, Palo Alto, CA, USA). Chromatographic separations were carried out with an Agilent HP5ms fused-silica capillary column (30 m × 0.25 mm i.d.) coated with 5%-phenyl-95%-dimethylpolysiloxane (film thickness 0.25 µm) as stationary phase as follows: injection mode: splitless at a temperature of 280°C; column temperature program: 80 °C (2 minutes) ramped to 140 °C at a rate of 10 °C/min, to 240 °C at a rate of 4 °C/min, to 280 °C at a rate of 10 °C/min and held for 5 min. The carrier gas was helium at a constant flow of 1.0 ml/min. The spectra were obtained in the electron impact mode at 70 eV ionization energy; ion source 280°C; ion source vacuum $10^{-5}$ Torr.

MS analysis was performed simultaneously in TIC (mass range scan from 50 to 600 m/z at a rate of 0.42 scans/s) and SIM mode. GC-SIM-MS analysis was

performed selecting the following ions: 102 m/z for glycine, 114 m/z for serine and 227 m/z for 2-(4-chlorophenyl)ethylamine (internal standard).

## 2.4 Statistical analysis

The mean and the standard deviation of data are computed from at least three independent biological experiments, each repeated in three technical replicates. Paired samples of real time PCR, *in vitro* translation and RIP experiments were analyzed with the two tailed *Student's t-test*; all the others statistical analyses were performed using one-way ANOVA followed by the *Bonferroni post-hoc* comparison test. A p-value $< 0.05$ was considered significant.

# 3. Results

## 3.1 Expression of SHMT transcripts in lung cancer

There are many examples of regulatory proteins and enzymes interacting with the 5' untranslated regions of messenger RNAs (5'UTRs) (58). Notably, also the mechanism of SHMT1 autoregulation has been shown to be mediated by a 5'UTR sequence (see section 1.5). For these reasons, we raised the hypothesis that the post-transcriptional regulation of SHMT2 may involve a direct interaction between SHMT1 and the 5'UTR of the SHMT2 transcripts. To choose the candidate 5'UTRs to employ in the *in vitro* assays we performed, in collaboration with Prof. Alessandro Paiardini (Sapienza University of Rome), an analysis of the expression of SHMT transcripts in samples of lung adenocarcinoma (LUAD) cell lines, obtained from the TCGA database (59). Several transcripts arising from the *shmt1* or the *shmt2* gene were identified, and a representative pool of differently expressed 5'UTR sequences was chosen (the transcripts with the corresponding 5'UTRs are listed in Table 5). We selected three SHMT2 transcripts (one encoding the SHMT2 isoform and the other two encoding the SHMT2α isoform) and two SHMT1 transcripts, the latter two share the same 5'UTR. The features of the chosen 5'UTR sequences employed in the binding experiments reported in the next paragraph, respectively named UTR2, UTR2αint, UTR2αsh and UTR1, are shown in Figure 11.

*Table 5. Expression of the SHMT1 and SHMT2 transcripts selected for this study in Healthy (H) and LUAD (T) cells.*

| **Isoforms*** | ***TPM*** | | ***Percentage*** | |
|---|---|---|---|---|
| | *Healthy* | *Tumor* | *Healthy* | *Tumor* |
| **SHMT1** *(Gene: ENSG00000176974)* | | | | |
| **NM_004169**[a] (*uc002gta.3*) 5'UTR 191nt | *13.9* | *14.8* | *73.9%* | *67.9%* |
| **NM_148918**[a] (*uc002gtb.3*) 5'UTR 191nt | *3.2* | *4.8* | *17.1%* | *22.1%* |
| **SHMT2** *(Gene: ENSG00000182199)* | | | | |
| **NM_005412** (*uc001snf.2*) 5'UTR 205nt | *26.5* | *63.1* | *95.9%* | *92.5%* |
| **NM_001166358** (*uc001sni.2*) 5'UTR 230nt | *0.3* | *2.4* | *1.1%* | *3.5%* |
| **NM_001166359** (*uc010srh.2*) 5'UTR 123nt | *<0.01* | *<0.01* | - | - |

TPM: Transcripts-per-million.

[a] The two SHMT1 transcripts share the same 5'UTR.

*Minor SHMT1 isoforms - **uc002gtc.1**:H, 0.2; 1.3% / T, 0.5; 2.0% - **uc002gtd.1**:H, 0.9; 5% / T, 0.8; 3.9% - **uc002gsz.3**:H, 0.4; 2.1% / T, 0.7; 3.4% -**uc010vxu.1**:H, 0.1; 0.6% / T, 0.1; 0.6%.

Other minor SHMT2 isoforms - **uc001snj.2**:H, 0.15; 0.6% / T, 0.7; 1.1% - **uc010srj.2**:H, -; - / T, 1.2; 1.8% - **uc001snh.2**: H, 0.7; 2.5% / T, 0.8; 1.1%.

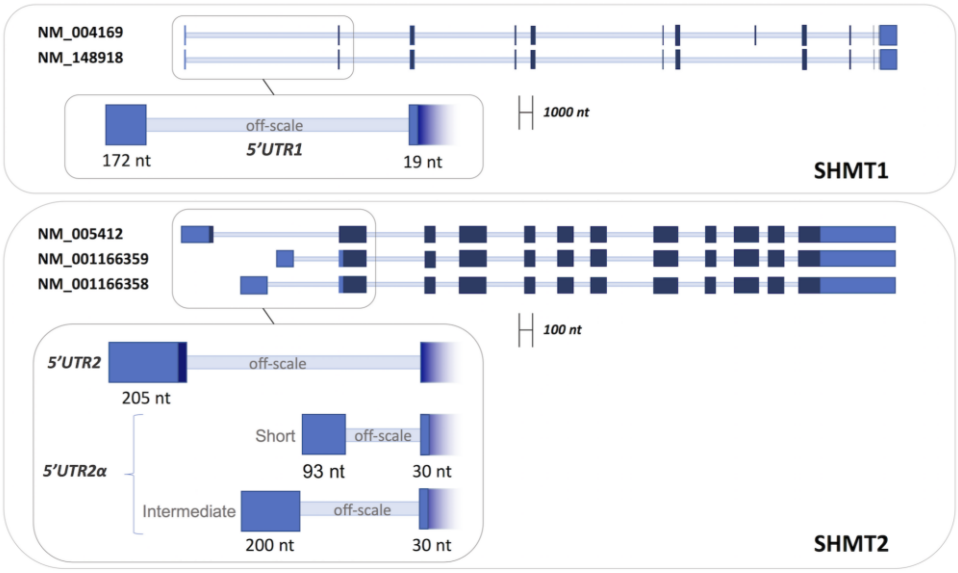


*Figure 11. Schematic representation of the pre-mRNAs of different SHMT1 and SHMT2 transcripts. Non-coding exons are colored in light blue, coding exons in dark blue and introns are represented as thin light blue lines. The 5'UTR sequences selected for the experiments, composed by the indicated exons, are shown in the insets:*
***UTR1** = (NM_004169-NM148918: Ex1_21-192+Ex2_1-19). **UTR2** = (NM005412:Ex1_1-205). **UTR2αsh** = (NM_001166359: Ex1_1-93+Ex2_1-30). **UTR2αint** = (NM_001166358: Ex1_1-200+Ex2_1-30).*

### 3.2 SHMT1 binds to the 5'UTR of *shmt2* transcripts

Electrophoretic mobility shift assays (EMSA) have been employed to assess if SHMT1 is capable to interact *in vitro* with the selected 5'UTRs. Given the propensity of SHMT1 to bind to multiple unrelated sequence variants (see section 1.5), it was essential to add an excess of tRNA in the experimental setup to eliminate the low-affinity interactions with the UTRs.

In a first set of experiments, the UTRs were incubated with increasing concentrations of SHMT1 (1 - 9.6 μM); as shown in Figure 12, the enzyme is

able to retard the migration of all the UTRs except the one of UTR1. Although the lack of binding between SHMT1 and UTR1 may be in contrast with what mentioned in section 1.5 and reported in (46), it's worth noticing that the experimental setup employed in the present work is far more stringent compared to the one chosen by Liu et al. (46), aiming to unravel only the high affinity interactions.

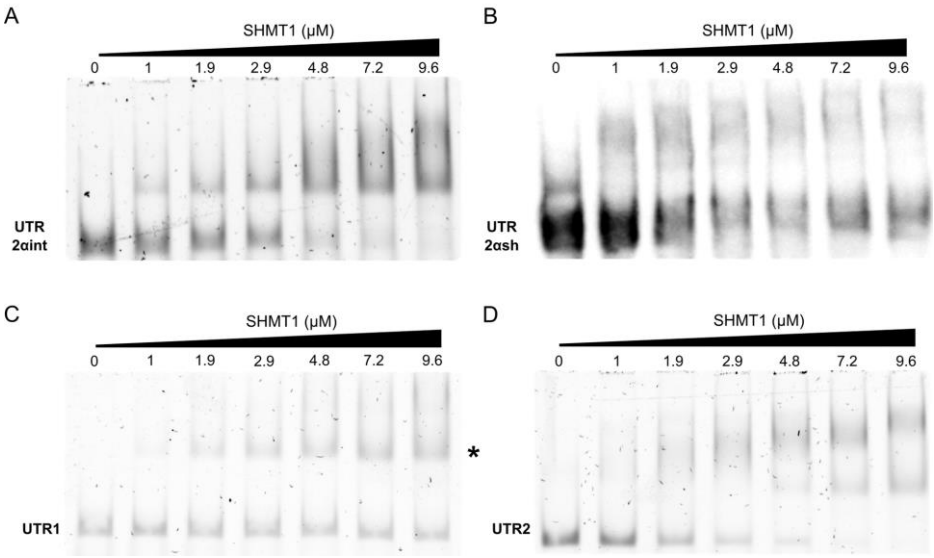


*Figure 12. Electrophoretic mobility shift assays. Migration of 0.24 μM unlabeled UTR2αint (A), UTR1 (C), UTR2 (D) and of 0.08 μM biotin-labeled UTR2αsh (B) in the presence of the indicated concentrations of SHMT1. The labelling of UTR2αsh was necessary to visualize the fragment, since it comigrates with the tRNA (which is cut out of the picture). The protein-RNA complexes visualized in (C), highlighted with a star (*), are composed by SHMT1 and tRNA; densitometric analysis of the UTR1 bands (unbound) confirms that the fragment it is not bound by the protein.*

A competition assay was then employed to assess which 5'UTR sequence binds with higher affinity to the protein. At this purpose, UTR2 was biotinylated and

incubated with SHMT1 in the presence of two different concentrations of each unlabeled UTR sequence. As shown in Figure 13, UTR2 and UTR2αint were the only fragments that displaced the biotinylated UTR2 from SHMT1, suggesting a stronger binding affinity with respect to UTR2αsh and UTR1.

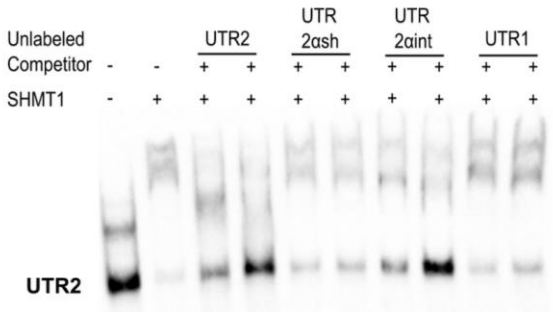


*Figure 13. Competition assay performed by incubating 0.08 μM of biotin-labeled UTR2 with 4.8 μM SHMT1 and, respectively, 0.6 and 1.2 μM of unlabeled competitors (UTR2, UTR2αsh, UTR2αint and UTR1).*

**SHMT1-UTR2 interaction: post-transcriptional and translational effects**

The investigation of the post-transcriptional and translational effects of the SHMT1-UTRs interaction has been focused only on UTR2 binding since, in addition to having an high affinity for the enzyme, it is far more expressed in LUAD cells compared to the other UTRs (Table 5).

At first it was investigated if the presence of SHMT1 is able to modulate the levels of UTR2-containing transcripts. The H1299 cell line was transfected with an RNA interference (RNAi) against SHMT1 (ishmt1) or with a scramble RNA sequence (iscr) and then the expression of UTR2 was studied by qRT-PCR. As shown in Figure 14, SHMT1 knockdown resulted in a significant

increase of UTR2 levels.

Then, to evaluate the effects on translation, two luciferase assays were employed. The first one, performed using rabbit reticulocytes extracts, was carried out to evaluate whether the SHMT1 protein is able to affect *in vitro*

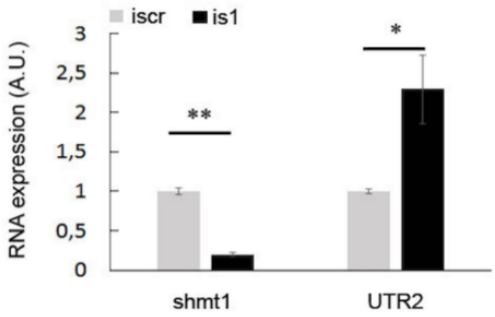


*Figure 14. Effect of SHMT1 on UTR2 expression in H1299 lung cancer cells. Expression of shmt1 and UTR2 RNAs, measured by qRT-PCR 48 hours after the transfection with the indicated RNAi (iscr, scrambled RNA sequences or is1, RNAi against shmt1). Statistical analysis is performed on three independent experiments, *P ≤ 0.05 ** P ≤ 0.01.*

the translation of the luciferase mRNA cloned downstream of the UTR2 sequence. Results show that SHMT1 inhibits to a greater extent (20%) the luciferase translation when the mRNA is fused to UTR2 (Fig. 15), highlighting a regulative role of the SHMT1-UTR2 interaction on translation.

The second assay was intended to confirm the same phenomenon *ex vivo* by employing a plasmid containing UTR2 cloned upstream of the firefly luciferase gene.

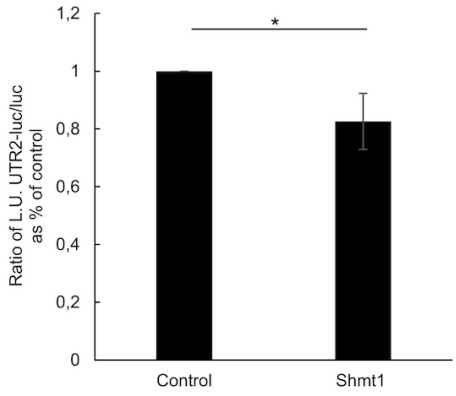


*Figure 15. In vitro translation experiments. Translation of luciferase mRNA with (UTR2-luc) or without (luc) the upstream UTR2 sequence was measured in the absence (Control) or presence (Shmt1) of 0.6 μM SHMT1. The results are plotted as the ratio of luminescence units (L.U.) between UTR2-luc and luc samples normalized on control reactions carried out in the absence of SHMT1. Statistical analysis is performed on five independent experiments, *P ≤ 0.05.*

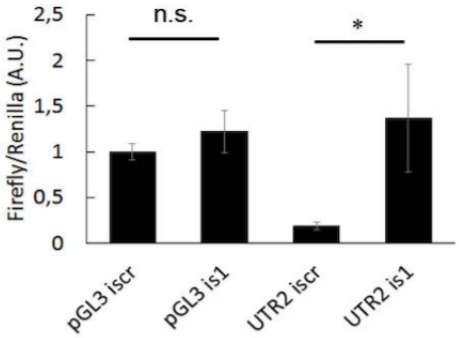


*Figure 16. Ex vivo translation experiments in H1299 cells. The panel shows the ratio between firefly and renilla luciferase evaluated 48 hours after the transfection with the indicated plasmids (pGL3, empty vector or UTR2, vector containing UTR2) and the indicated RNAi (iscr, scramble sequences or is1, RNAi against shmt1). Statistical analysis is performed on three independent experiments, *P ≤ 0.05.*

A significant increase (5-fold) in the luciferase signal was observed only in those cells co-transfected with the UTR2-containing plasmid and ishmt1 (Fig. 16). These data demonstrate that in the H1299 lung cancer line the presence of SHMT1 reduces the luciferase expression only when the latter is controlled by UTR2, the results confirm that SHMT1 affects the translation of UTR2-containing transcripts also in the cellular system.

### 3.4 Modulation of the RNA binding through SHMT substrates

To investigate the relationship between the catalytic and the RNA binding activity of SHMT1, we started by assessing if the binding of substrates could interfere with RNA binding. EMSA assays were carried out in the presence of SHMT substrates and/or analogues: the protein was incubated with its amino acid and/or folate substrates prior to the addition of UTR2. As shown in Fig. 17A, pentaglutamylated 5-formyl-tetrahydrofolate (5-CHO-THF-$Glu_5$) competes with the formation of the RNA-protein complex in a concentration dependent manner. The presence of saturating concentrations of glycine in the reaction mixture significantly lowers the minimal folate concentration needed to compete with UTR2 (Fig. 17B). The latter phenomenon is due to the higher affinity of the folate moiety for the enzyme-glycine complex than for the free enzyme (54). Interestingly, even at the highest folate concentration, the RNA moiety was not fully displaced. Being the polyglutamylated folate binding stoichiometry 0.5 per subunit (2 per tetramer) (60, 61), and assuming that the RNA randomly binds to each monomer, it is likely that the folate-mediated displacement affects only two subunits and thus that at least 50% of UTR2 remains bound to the other sites (Fig.18), in accordance with our data.

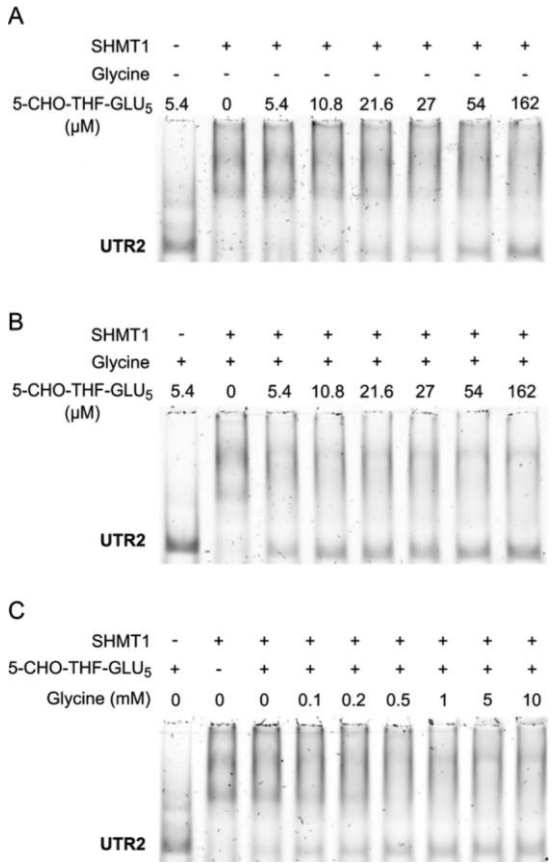


*Figure 17. Effect of SHMT substrates and/or analogues on UTR2-SHMT1 interaction. EMSA carried out by incubating 5.4 μM of SHMT1 with the indicated amounts of 5-CHO-THF-Glu$_5$ in absence (A) or presence (B) of 10 mM glycine prior to the addition of 0.18 μM unlabeled UTR2. Panel C shows the effect of glycine concentration on the folate-mediated displacement of UTR2. In the latter case, EMSA was performed incubating 5.4 μM of SHMT1 with the indicated amounts of glycine and a fixed concentration of 5-CHO-THF-Glu$_5$ (27 μM) prior to the addition of 0.18 μM unlabeled UTR2.*

To further confirm the essential role of glycine in the competition between folate and RNA, glycine concentration was varied from 0.1 to 10 mM while maintaining a fixed concentration of 5-CHO-THF-Glu$_5$. As expected, the folate-mediated displacement of the RNA from its binding site depends on

glycine levels, reaching a maximum at a glycine concentration ≥ 0.5 mM (Fig. 17C). Notably, in the absence of folate, saturating concentrations of glycine or serine do not interfere with the formation of the enzyme-RNA complex (Fig. 19A).

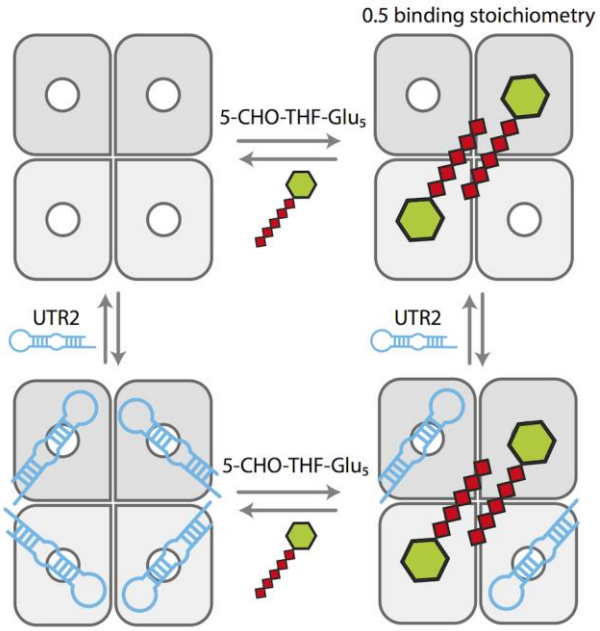


*Figure 18. Hypothesized mechanism for the folate-mediated displacement of UTR2. In the picture, each SHMT1 monomer is represented as a grey rectangle with a white circle in the center which indicates the position of the cofactor, UTR2 is drawn in blue and the folate is represented in green together with its polyglutamylated tail colored in red.*

The same experiments were also performed using 5-CHO-THF-$Glu_1$. Interestingly we did not observe any competition between this folate specie and the RNA for binding to SHMT1 (Fig. 19B); moreover, in this case the presence of saturating concentration of glycine only slightly affects the formation of the SHMT1-UTR2 complex (Fig. 19C). This suggests that the

polyglutamylated tail of the folate is mainly responsible for the displacement of the RNA molecule, leading us to hypothesize at least a partial overlapping between the RNA and the folate binding site.

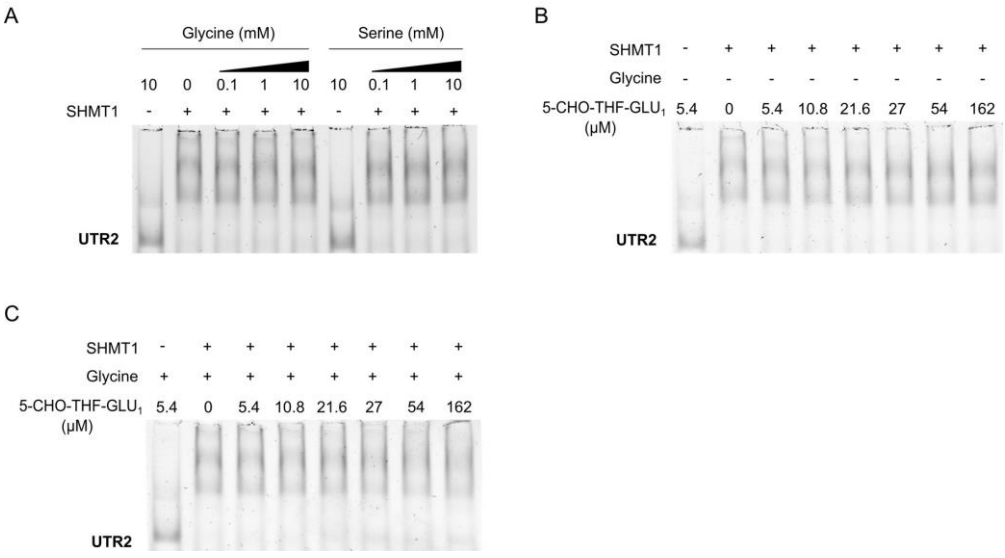


*Figure 19. EMSA carried out by incubating 5.4 μM of SHMT1 with only glycine or serine (A) or with the indicated amounts of 5-CHO-THF-Glu$_1$ in absence (B) or presence (C) of 10 mM glycine prior to the addition of 0.18 μM unlabeled UTR2.*

### 3.5 Identification of the residues involved in the RNA binding

Given the evidences shown in the previous paragraph, to investigate the residues involved in the RNA binding we started considering the amino acids in proximity to the folate binding site and to a positively charged cleft of the enzyme, which we thought could be suitable for nucleic acids binding.

Indeed, it is known that the formation of a protein-nucleic acid complex is often mediated by positively charged side chains (i.e. arginine and/or lysine

side chains) and by aromatic residues that make stacking interactions with the nucleobases (62).

As first candidates we chose two lysine residues (K157 and K158) that flank the folate binding site; notably, one of the two makes a salt bridge with the glutamic end of the folate. A tyrosine (Y82') that forms an aromatic stacking interaction with the aminobenzoic moiety of the folate was selected as well (Fig. 20A and 20B), this latter residue is located on the other subunit of the obligate dimer with respect to the lysine residues.

We characterized the behavior of the K157S/K158S double and the Y82A single SHMT1 mutants to understand if such residues are involved in the RNA binding. As shown in Figure 20 (panels C-E), the double lysine mutation almost completely abolishes the binding of SHMT1 to UTR2, whereas, the Y82A substitution has a milder effect (2-fold). The apparent dissociation constants ($K_d^{app}$) of the protein-UTR2 complexes were measured for the WT and for the Y82A mutant and are (1.39 ± 0.30) μM and (3.0 ± 0.68) μM, respectively. The $K_d^{app}$ for the K157S/K158S mutant could not be calculated given the low affinity of this protein for UTR2 in the concentration range explored.

These data confirm the hypothesis of a partial superimposition between the folate and the RNA binding sites, highlighting an important role of the positively charged region surrounding the folate binding cleft in the RNA binding event (Fig. 20F).

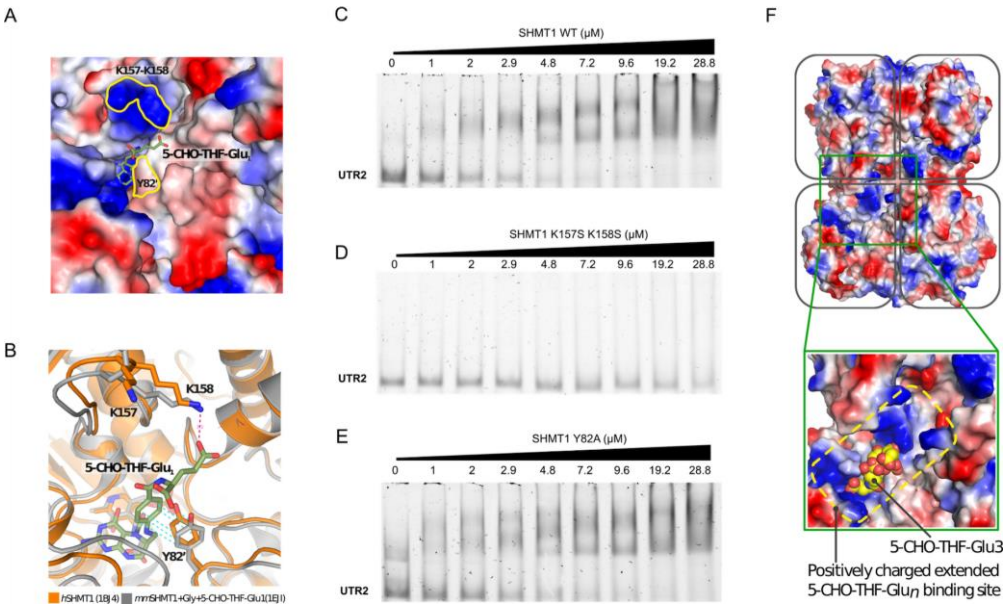


*Figure 20. SHMT1 amino acid residues involved in folate and RNA binding. (A) Surface representation of the folate-binding site of human SHMT1 (hSHMT, PDB id: 1BJ4* (22)*) colored by electrostatic potential (in red are the acidic and in blue the basic residues). The substrate (5-CHO-THF-Glu$_1$) from mouse SHMT1 (mmSHMT1, PDB id: 1EJI* (63)*) is shown as sticks and superposed to the human enzyme. The location of K157, K158 and Y82' is highlighted by the yellow contours. (B) Superposition of the human and mouse crystal structures (in orange and grey, respectively) showing the interaction of K158 and Y82' with the folate substrate. (C-E) EMSA assays carried out by incubating 0.24 μM of unlabeled UTR2 with the indicated concentrations of SHMT1 wild-type (WT) (C), K157S-K158S (D) and Y82A (E) mutants. (F) Surface charge distribution of the human SHMT1 tetramer (PDB id: 1BJ4* (22)*), in the blow up, 5-CHO-THF-Glu$_3$, in complex with rabbit SHMT1 (PDB id: 1LS3* (61)*), is shown as yellow spheres superposed to the human enzyme. In the rabbit structure, the 5-CHO-THF-Glu$_3$ is bound to two subunits (0.5 stoichiometry), with the poly-glutamylated tail making extensive and non-specific electrostatic interactions with the positively charged residues that surround the folate binding cleft.*

### 3.6 Effects of the RNA binding on SHMT1 activity

The effect of the RNA binding on SHMT1 enzymatic activity was evaluated employing biochemical assays in the presence of the selected UTR sequences; in addition, yeast tRNA, considered as a random sequence variant, was chosen as control.

The serine cleavage activity of the enzyme, whereby L-serine and THF are converted into glycine and Me-THF, was measured by incubating SHMT1 (0.2 μM) with a saturating concentration of serine (10 mM) and 80 μM of monoglutamylated THF while varying the concentration of the RNAs.

The ability of SHMT1 to catalyze this reaction is inhibited by all the RNAs, although to a different extent (Fig. 21A and 21B). UTR2 shows the highest inhibitory efficacy, coherently with its binding propensity (section 3.2), UTR2αint has an intermediate inhibitory efficacy, while UTR1, UTR2αsh and tRNA have the lowest.

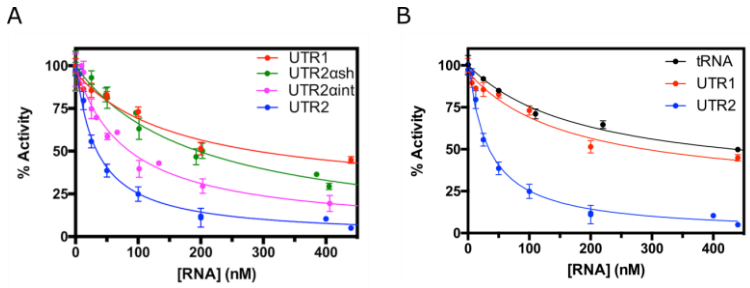


*Figure 21. RNA-mediated inactivation of the serine cleavage reaction. The inactivation of SHMT1 by the RNAs was followed as described in section 2.2.1. Panel A shows the inactivation of SHMT1 by UTR1 (red), UTR2αsh (green), UTR2αint (magenta) and UTR2 (blue). In panel B the RNAs having the highest (UTR2, blue) and the lowest (UTR1, red) inhibitory efficacy are compared to the inhibition given by a random sequence variant (tRNA, black). Experimental data, which are the average of three independent experiments, were fitted to Eq. 2 (see section 2.2.1).*

The estimated apparent inhibition constants (i.e. the RNA concentration required to obtain 50% activity inhibition, $IC_{50}$) are reported in Table 6.

*Table 6. $IC_{50}$ values calculated from the inhibition curves shown in Figure 21.*

| **RNA** | **$IC_{50}$ (nM)** |
|---|---|
| UTR1 | 170 ± 58 |
| UTR2 | 33 ± 5 |
| UTR2αsh | 197 ± 38 |
| UTR2αint | 81 ± 17 |
| tRNA | 192 ± 40 |

The inhibitory effect of RNA on the serine synthesis reaction (i.e. the reverse reaction) was measured employing two parallel experimental approaches. The first one follows the conversion of glycine and Me-THF into L-serine and THF, while the other one measures the formation of the quinonoid intermediate that develops when glycine and 5-CHO-THF bind to the enzyme forming a ternary complex (54). Strikingly, a milder inhibitory effect of all the RNAs was observed with both kinetic assays, even at the highest RNA concentration employed (Fig. 22 shows only UTR2 and tRNA as representatives). In these experiments, UTR2 remains the most efficient inhibitor, nevertheless, the $IC_{50}$ is higher by two orders of magnitude (~ 4 μM) with respect to the one measured for the serine cleavage reaction.

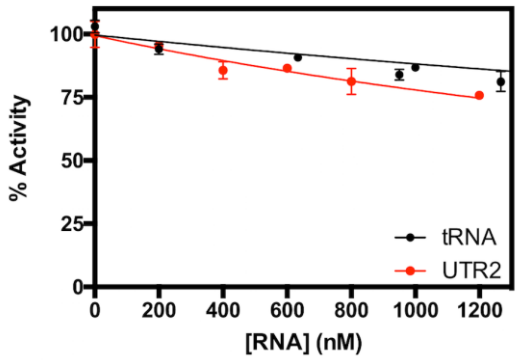


*Figure 22. RNA-mediated inactivation of the serine synthesis reaction. The inactivation of SHMT1 by UTR2 (red) and tRNA (black) was followed as described in section 2.2.1. The figure combines the data obtained in three independent experiments with both assays described in the text, the data were fitted to Eq. 2 (see section 2.2.1).*

In agreement with the results obtained by EMSA analyses, the activity of the K157S/K158S mutant was unaffected by UTR2, confirming the pivotal role of the two lysines in the RNA binding (Fig. 23).

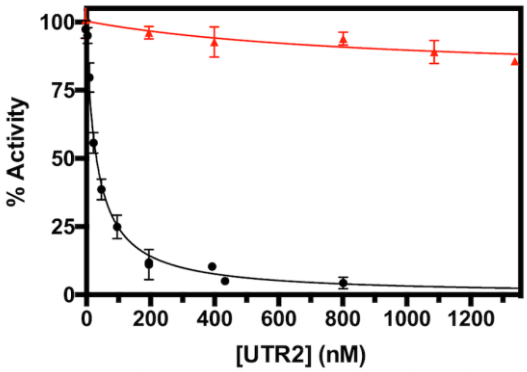


*Figure 23. Inactivation of SHMT1 WT (black) and SHMT1 K157S/K158S mutant (red) by UTR2 (serine cleavage reaction). The enzymatic activity was followed by measuring the initial velocity of the hydroxymethyltransferase reaction using L-serine (10 mM) and THF as substrates. The experimental data, obtained in three independent experiments, were fitted to Eq.2 (see section 2.2.1).*

Finally, it is important to point out that the analysis of the kinetic parameters of the K157S/K158S mutant, conducted before the inhibition assays, revealed a $K_m$ for the monoglutamylated THF about 5-fold higher with respect to the wild type isoform, whereas the $K_m$ for serine and the $k_{cat}$ are similar for both proteins (Table 7), confirming the involvement of the lysines in the THF binding.

*Table 7. Steady-state kinetic parameters of the serine cleavage reaction catalyzed by SHMT1 WT and SHMT1 K157S/K158S.*

| Parameter | WT | K157S/K158S |
|---|---|---|
| $K_m^{THFGlu1}$ (μM) | 30 ± 8 | 146 ± 18 |
| $K_m^{Ser}$ (μM) | 122 ± 24 | 264 ± 26 |
| $k_{cat}$ ($min^{-1}$) | 433 | 429 |

The next step was to move from *in vitro* to *ex vivo*: we thus tested if the RNA-mediated inhibition of SHMT1 activity occurs also in cell models. At this purpose we employed the H1299 cell line, the model in which we have characterized the interaction. The SHMT enzymatic activity was measured in living cells using the radioisotope assay described in section 2.2.2. Since the assay measures the total SHMT activity (SHMT1 + SHMT2), to evaluate the contribution only of the cytosolic isoform, the cells were cotransfected with an RNA interference against shmt2 (is2), thereby silencing the SHMT2 expression, and with a plasmid containing UTR1, UTR2 or the luciferase gene alone (pGL3, empty vector). Under these conditions, SHMT1 activity

decreased by 25% in the cells overexpressing UTR2 with respect to those overexpressing UTR1 or the control (Fig. 24). These results nicely parallel the data obtained with the *in vitro* assays, where UTR2 displays the highest inhibitory efficacy.

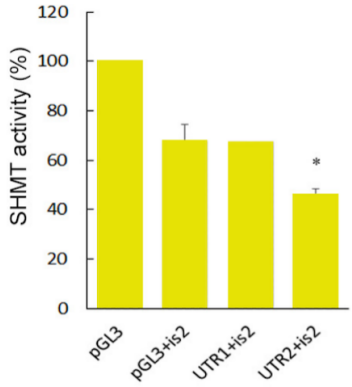


*Figure 24. SHMT activity in H1299 cell line evaluated by tritium-exchange radioisotopic assay. The SHMT enzymatic activity was measured 48 hours after the transfection with the indicated plasmids (empty pGL3 or pGL3 containing UTR1 or UTR2) and iRNAs.*

### 3.6.1 Mechanism of inhibition

A detailed study of the SHMT1 inhibition kinetics was performed to characterize the molecular mechanism of the RNA-mediated inhibition. Since yeast tRNA inhibits the enzyme's activity (Fig. 21B), although less efficiently than UTR2, and is easily available in the large amounts necessary to perform these experiments, we employed tRNA instead of UTR2 for the mechanistic analysis. As a first step, the initial velocity of the reaction was measured at different tRNA concentrations, maintaining the L-serine concentration fixed (10 mM) and varying the THF concentration. Results show a typical THF

substrate inhibition (64) (Figure 25A), which became less evident with the increase of tRNA concentration. Fitting of data to Equation 3 (see section 2.2.1), a modified Michaelis-Menten equation that accounts for substrate inhibition, showed that while tRNA binding increased the $K_m$ for THF, it did not affect the $V_{max}$. After exclusion of the substrate inhibition portion of data, a double reciprocal plot confirmed this observation, which is typical of competitive inhibition (Figure 25B). However, slopes ($K_m/V_{max}$) obtained by fitting the data in the double reciprocal plot did not show the characteristic linear dependency on inhibitor concentration, instead they exhibited a hyperbolic behavior (Figure 25C). Moreover, plots of initial velocity measured at a fixed THF concentration as a function of tRNA clearly showed that the limiting velocity as the inhibitor concentration tends to infinity was not zero (Fig. 21B). All these observations indicate that RNA is not a pure competitive inhibitor and suggest a 'hyperbolic' or 'partial competitive' inhibition (65).

In such type of inhibition, the substrate and the inhibitor bind to partially overlapping or totally different sites and have higher affinity for the free enzyme than for the enzyme-inhibitor and enzyme-substrate complexes, respectively. When the enzyme forms a complex with the first ligand, the affinity for the other ligand is decreased by an equal $\alpha$ factor (Figure 25G). Nevertheless, the enzyme-substrate and the enzyme-substrate-inhibitor complexes yielded the product with equal rate constants. Global fitting of velocity data to Equation 4 (see section 2.2.1), which combines hyperbolic competitive inhibition to substrate inhibition, gave estimates of $K_i$ for tRNA ($0.031 \pm 0.005$ μM) and of $\alpha$ factor ($18.8 \pm 4.6$). Notably, at the highest THF concentration used in the experiments, tRNA attenuated THF substrate inhibition (Figure 25A). This paradoxical situation is caused by the decrease of affinity for the folate substrate induced by tRNA. In a second set of

experiments, the initial velocity was measured in the presence of increasing tRNA and maintaining a fixed THF concentration (0.08 mM), while varying L-serine (Figure 25D). In this case, the double reciprocal plot of data (Figure 25E) together with the behavior of the obtained slopes and intercepts (Figure 25F) indicated a pure mixed inhibition, showing that tRNA mainly competes with folate but not with L-serine.

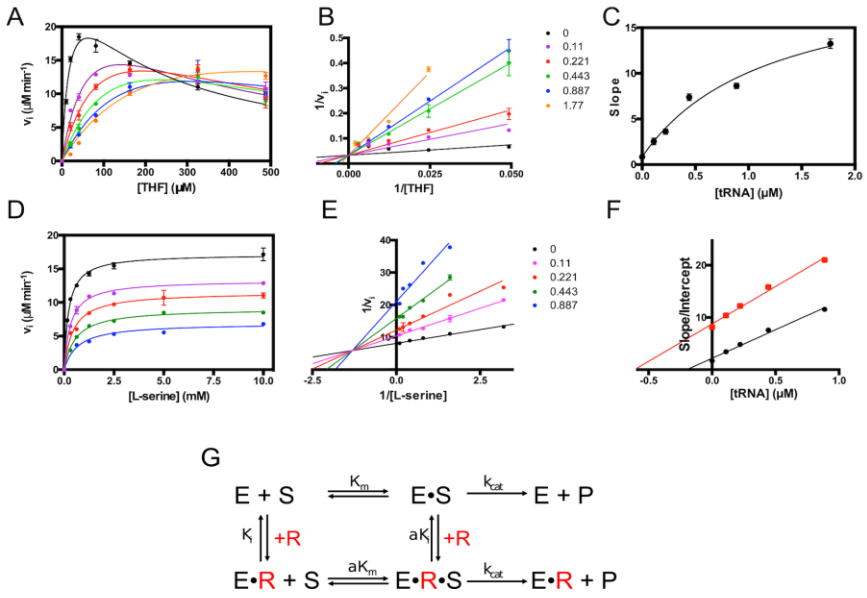


*Figure 25. Mechanism of the RNA-mediated inhibition of SHMT1 catalysis. (A) Initial velocity of the SHMT reaction was measured at different THF concentrations, in the presence of saturating L-serine (10 mM) and fixed concentrations of tRNA: 0, 0.11, 0.22, 0.44, 0.88 and 1.77 μM. The continuous lines through the experimental points (which are the average of three independent experiments) are those obtained upon the least square fitting of data to Equation 3 (see section 2.2.1). (B) Double reciprocal plots of experimental points from A, after exclusion of the substrate inhibition portion of data. The continuous straight lines through the experimental points were obtained from the independent linear fitting of data. (C) Dependence of the slope, calculated in double reciprocal plot (B), on tRNA concentration. The continuous line through the experimental points was obtained from fitting to a saturation curve. (D) Initial velocity of the SHMT reaction was measured at different L-serine concentrations, in the presence of 80 μM THF and fixed concentrations of tRNA: 0, 0.11, 0.22, 0.44, 0.88 and 1.77 μM. The continuous lines through the experimental points (average of three independent experiments) are those obtained upon the least square fitting of data to Michaelis-Menten equation. (E) Double reciprocal plots of experimental points from D. The continuous straight lines through the experimental points were obtained from the independent linear fitting of data. (F) Replots of data taken from the double reciprocal plot (E): slope ($K_m/V_{max}$; in black) and intercept ($1/V_{max}$; in red) versus tRNA concentration. The continuous straight lines through the experimental points were obtained from linear fitting. (G) Scheme describing hyperbolic competitive inhibition between SHMT1 and RNA (data shown in panels A-C). In this case, a saturating amino acid substrate concentration (either L-serine or glycine) is present and therefore E corresponds to the enzyme-amino acid complex. E•S is the ternary complex with the folate substrate (either THF or Me-THF). If RNA (R) is present, it binds to both E and E•S forming E•R and E•R•S complexes, respectively.*

### 3.7 Metabolic effects of UTR2

Given the striking data on the inhibition of the cytosolic isozyme, we decided to study the metabolic effects of the SHMT1-UTR2 interaction *ex vivo*. For this purpose we employed the HAP1 cell line, a human haploid line currently used as metabolic model for the 1C metabolism (28, 40, 66). Our group has recently reported that the SHMT2 knockout line (HAP1-SHMT2KO) can survive and proliferate in minimal medium (MEM) supplemented with serine, indicating that, in such line, SHMT1 is able to perform the serine cleavage reaction (i.e. serine to glycine), thus complementing the function of the missing mitochondrial enzyme. Notably, the HAP1-SHMT2KO cells were produced by the CRISPR/CAS9 system which introduced a frameshift of shmt2 mRNA in exon 2, thus, in this line, the SHMT2 protein is not produced, while its mRNA, comprising the UTR2 region, is still expressed.

In a first set of experiments, we have eliminated the shmt2 mRNA by RNAi (is2) and then, at 24 hours after transfection, the complete medium was exchanged for MEM medium supplemented with 400 μM serine; finally, 48 hours after the medium exchange, the residual serine concentration was evaluated by gas chromatography/mass spectrometry. Interestingly, the cells lacking shmt2 mRNA (and therefore UTR2) consumed twice as much serine as the cells transfected with scramble sequences (Fig. 26A). In a second set of experiments, the opposite strategy was adopted: in order to overexpress the UTR portion, the knockout line was transfected with pGL3 containing UTR2, UTR1 or the empty vector, and the serine residual concentration in the medium was evaluated as described above. The overexpression of UTR2 caused a higher residual serine concentration in the medium compared to the controls (UTR1 an pGL3) (Fig. 26B), in agreement with the first set of experiments.

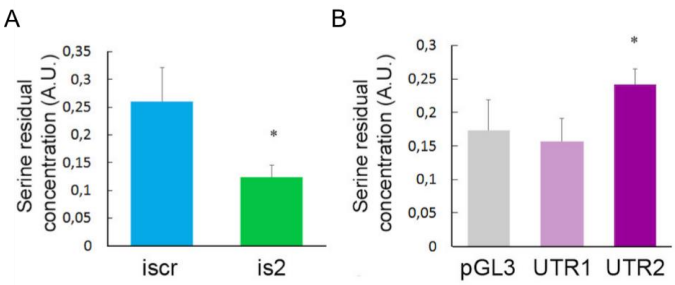


*Figure 26. SHMT1 activity in HAP1-SHMT2KO cells evaluated measuring serine residual concentration in the medium 72 hours after transfection with the indicated RNAi (A) or plasmids (B). Statistical analyses are performed on three independent experiments, *P ≤ 0.05.*

Altogether the results confirmed the inhibition of the SHMT1 serine cleavage activity in cell models and, importantly, highlighted that the observed biological effect is UTR2-specific, since the overexpression of UTR1 didn't affect serine catabolism.

In parallel, to show that SHMT1 activity is correlated to cell viability, the effects of the inhibition of the serine cleavage reaction on cell viability have been analyzed. At this aim, we transfected HAP1-SHMT2KO cells with the same constructs described for Fig. 26B (i.e. UTR2 or UTR1 containing pGL3 or the empty vector), and then evaluated cell viability using the trypan blue exclusion assay. Only the UTR2 overexpression induced a significant decrease in cell viability (~40%) when cells were grown in MEM supplemented with serine (Fig. 27A), whereas no effects were observed in the medium supplemented with glycine and formate (Fig. 27B). Such data indicate a correlation between impaired serine consumption and cell viability, notably, once again, the effect is mediate only by UTR2.

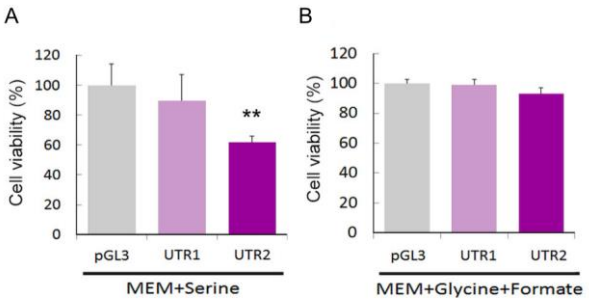


*Figure 27. Cell viability evaluated in HAP1-SHMT2KO cells cultured in MEM supplemented for 48 hours with 0.4 mM serine (A) or with 0.4 mM glycine + 0.5 mM formate (B) the day after the transfection with the indicated plasmids. Statistical analyses are performed on three independent experiments, $**P \leq 0.01$.*

### 3.8 Discovering other target RNAs

RNA immunoprecipitation experiments (RIP) on cell lysates have been employed to validate the SHMT1 interaction with the SHMT2 mRNA definitively and, moreover, to broaden the analysis in order to unveil other RNAs bound by the cytosolic enzyme in H1299 cell line.

The SHMT2 mRNA was found enriched by 3-fold in the SHMT1 immunoprecipitate fractions (SHMT1 IPs) compared to the control (IgG IPs), whereas no enrichment was detected for the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA (Fig. 28), a sequence that we chose as negative control, supposing that it does not bind to SHMT1 in the cells.

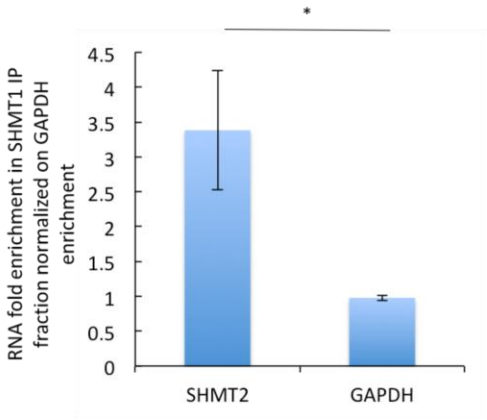


*Figure 28. RNA immunoprecipitation assays in H1299 cells. Fold enrichment of the indicated mRNAs in the SHMT1 immunoprecipitate fraction (SHMT1 IP) relative to the negative control (IgG IP), data were normalized on glyceraldehyde 3-phosphate dehydrogenase (GAPDH) enrichment. The relative quantification of the RNA was carried out by semi-quantitative PCR performed on the cDNA obtained from the RNA extracted from the IP samples. The bands resulting from the amplified products were quantified by densitometric analyses. Statistical analysis is performed on three independent experiments, *P ≤ 0.05.*

The latter results have confirmed the interaction *ex vivo* and, importantly, represent a good starting point for the setup of RNA immunoprecipitation and sequencing experiments (RIP-seq). To this purpose, the RNA extracted from the IP fractions was used to generate libraries suitable for Next Generation Sequencing (NGS), the kit employed was the "TruSeq RNA Sample Preparation V2", provided by Illumina (workflow shown in Fig. 29). The kit is intended to generate DNA libraries flanked by adapters suitable for NGS, each adapter contains a specific index sequence (composed by a combination of 6 nucleotides), necessary to identify the samples during the sequencing.

All the IPs were processed equally and the libraries were subjected to a quality control (Fig. 30). All the samples displayed the attended size (~ 270 bp ) and didn't contain contaminating adapter-dimers.

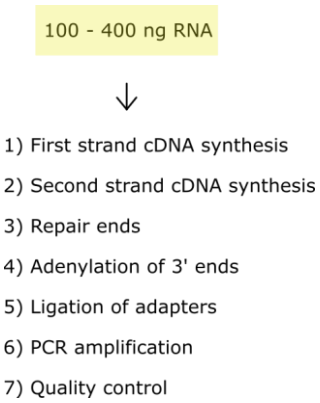


*Figure 29. Workflow of the TruSeq RNA Sample Preparation V2 kit provided by Illumina, the detailed protocol is available on the manufacturer's website (www.illumina.com).*

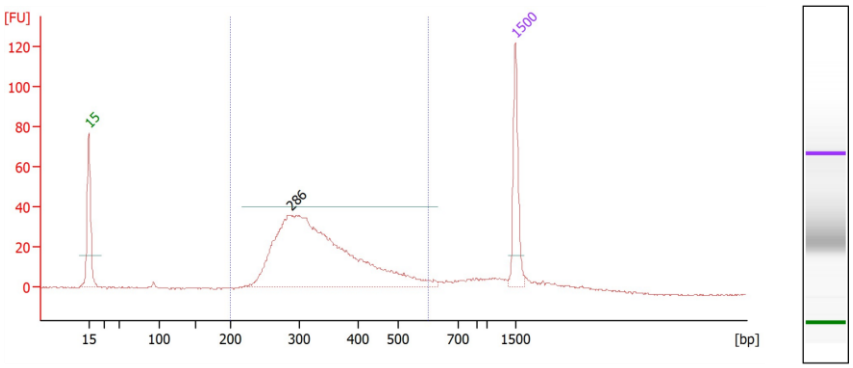


*Figure 30. Quality control of the DNA libraries carried out with Agilent 2100 Bioanalyzer provided with a DNA-specific chip (Agilent DNA 1000). On the left, electropherogram showing the fragments size, fluorescent units (FU) are reported as function of the DNA length expressed as base pairs (bp), the major peak, indicating the fragments size, is highlighted with a horizontal line; on the right is shown the corresponding gel-like image. The lower and upper markers are represented in green and purple, respectively. Since all the samples gave a similar Bioanalyzer profile, the figure is representative for all the libraries.*

Finally, indexed libraries were pooled and sequenced using 125 bp paired-end reads on an Illumina HiSeq2500 instrument (Centre for Genomic Regulation, Genomic core facility). The sequencing was successful and the reads are currently being analyzed.

### 3.9 Not only RNA: insights into the DNA binding ability of SHMT1

The transient translocation of SHMT1 in the cell nucleus, together with the evidences indicating a putative DNA binding activity of this isozyme (32, 33), prompted us to shift the analysis from RNA to DNA and to investigate whether SHMT1 behaves as a DRBP.

To get insights in this phenomenon we decided to employ EMSA assays, which allow the direct study of the interaction *in vitro*. Since we had no clues on a possible sequence specificity of the enzyme, to setup the assay a random double stranded DNA sequence (a bacterial promoter, 160 bp) was chosen and incubated with increasing fold molar excesses of protein (0 to 75 SHMT1:DNA fold excess). As evident in Figure 31A, the protein interacts with the DNA when its concentration is greater equal to 10-fold molar excess with respect to the DNA. As expected, the addition of an aspecific competitor, the same employed in the RNA binding assays (tRNA), almost completely abolished the binding (Fig. 31B), indicating that the interaction shown in panel A is of low affinity nature.

The SHMT1 mutants Y82A and K157S/K158S, already used for the investigation of the RNA binding site, were also assayed to evaluate their involvement in the DNA binding event. The DNA binding ability of SHMT1 was strongly affected by all the mutations (Fig. 32), with the double lysine

mutant completely unable of interacting with the nucleic acid (Fig. 32B). The results strongly suggest that the RNA and the DNA share, at least partially, the same binding site.

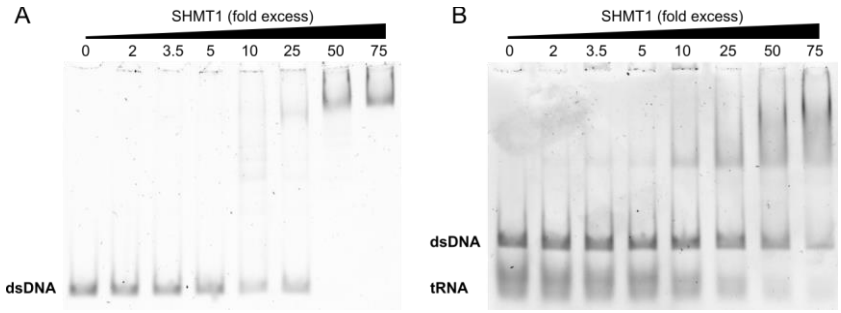


*Figure 31. Electrophoretic mobility shift assay. Migration of 0.08 μM dsDNA (160 bp) incubated with the indicated fold molar excess of SHMT1 in absence (A) or presence (B) of 0.5 ug of yeast tRNA. The SHMT1 fold molar excess is considered as excess of tetramer with respect to the DNA.*

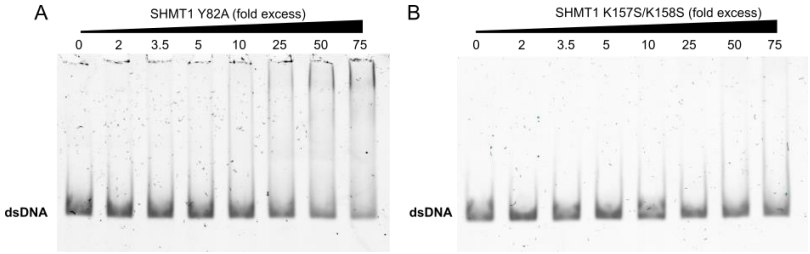


*Figure 32. Electrophoretic mobility shift assay. Migration of 0.08 μM dsDNA (160 bp) incubated with the indicated fold molar excess of SHMT1 Y82A (A) or SHMT1 K157S/K158S (B). The fold molar excess is considered as excess of tetramer with respect to the DNA.*

To assess if SHMT1 discriminates between dsDNA and ssDNA, shorter DNA fragments (24 nt), either double or single stranded, have been employed in EMSA assays. In this set of experiments, to better appreciate possible

differences, we explored more intervals of fold molar excess in the lower range (0-10 fold excess). Figure 33 shows that SHMT1 has a higher affinity for the ssDNA with respect to dsDNA, binding all the single stranded oligonucleotide already at one-fold molar excess. Given these preliminary experiments, it is very likely that ssDNA represents the physiological ligand of SHMT1 in the nucleus.

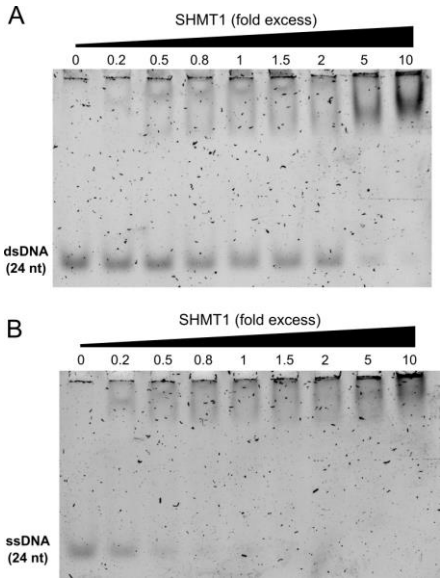


*Figure 33. Migration of 1 μM ds DNA (24 bp) (A) or 2.1 μM ssDNA (24 nt) (B) in the presence of the indicated fold molar excess of SHMT1 (SHMT1:DNA). The fold excess is considered as excess of tetramer with respect to the DNA.*

The intriguing evidence that only higher organisms synthetize tetrameric SHMTs led us to investigate the importance of the oligomeric structure in the DNA binding ability of the human enzyme.

It is known that the tetrameric structure of eukaryotic SHMTs is necessary for the binding of polyglutamylated THF (67, 68), however this might not be the

only reason for their difference with the bacterial homologues, which are dimeric. To assess if the tetrameric structure is essential to bind the DNA, a dimeric catalytically active mutant of human SHMT1 has been produced and assayed in our laboratory (SHMT1 H135N/R137A, (69), illustrated in Fig. 35). EMSA assays were employed to compare the DNA binding ability of human dimeric SHMT1 with the one of dimeric SHMT from *E. coli* (eSHMT). The double stranded conformation of the nucleic acid, mostly populated in the cells, was chosen for an initial analysis.

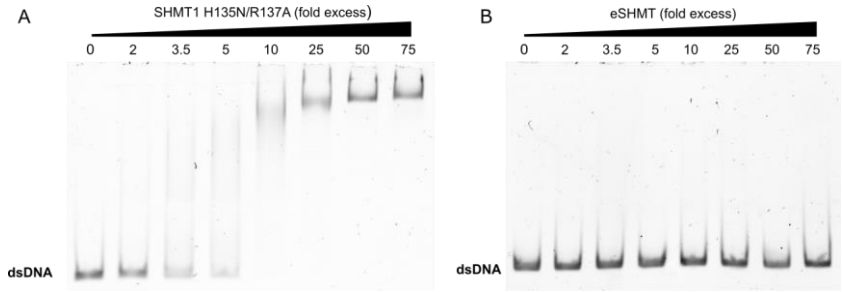


*Figure 34. Migration of 0.08 μM dsDNA (160 bp) incubated with the indicated fold excess of human dimeric SHMT1(SHMT1 H135N/R137A) (A) or SHMT from E. coli (eSHMT) (B). The fold excess is considered as excess of dimer with respect to the DNA.*

Strikingly, eSHMT is completely unable to bind the dsDNA, whereas the human dimeric form has maintained such capacity (Fig. 34). Comparable results were obtained employing RNA instead of dsDNA (data not shown). Such data not only suggest that the nucleic acids binding activity of the human enzyme appears to be independent on the oligomeric structure, but also suggests that this property is not evolutionary conserved in all organisms.

### 3.10 The nuclear function of SHMT1

Studies performed in different cell lines (HeLa, MCF-7 and SH-SY5Y) support the existence of two main nuclear roles for SHMT1: (i) the biosynthesis of thymidylate and (ii) a scaffolding function that likely allows to anchor the thymidylate synthesis metabolic complex (i.e. TYMS, DHFR and SHMT1 itself) to the DNA (32, 70).

We set up *ex vivo* experiments to analyze whether SHMT1 nuclear function is dependent on the oligomeric state of the enzyme; given the results presented in the previous paragraph, we would expect that this is not the case. In parallel, we have also evaluated the importance of the SHMT1 catalytic activity in the nuclear compartment.

A set of SHMT1 mutants aimed at uncoupling the oligomeric state of the protein from the enzymatic activity was produced and assayed. Such mutants were assayed in two models of LUAD cancer, H1299 and A549 cells.

Figure 35 illustrates the three mutants that have been generated:

1. SHMT1 H135N/R137A, a catalytically Active Dimer (AD), already mentioned in section 3.9;
2. SHMT1 K257Q/Y82A/Y83F, a catalytically Inactive Tetramer (IT), unable to bind PLP and THF;
3. SHMT1 H135N/R137A/E168N, an Active Unstable Tetramer (AUT), an active mutant with a less stable tetrameric interface.

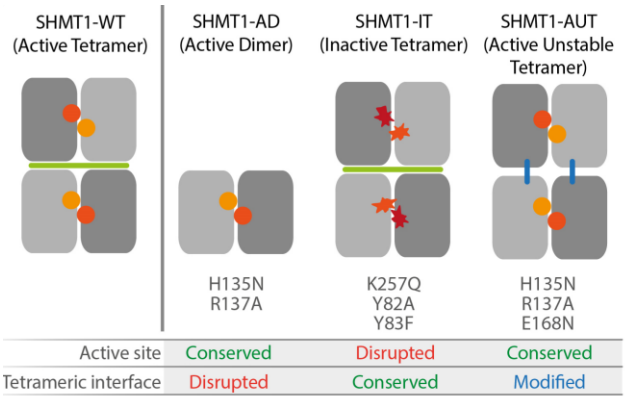


*Figure 35. Scheme representing SHMT1 WT in comparison with the three SHMT1 mutants designed to uncouple the oligomeric state from the catalytic activity. The mutations, together with the corresponding effects on the oligomeric state and catalysis, are indicated below each mutant. The active sites are indicated with circles, the native tetrameric interface is represented with a green line, while the modified one is indicated in blue.*

The biophysical and kinetic characterization of all the mutants has been performed and published by our group (69), since it is not in the aim of the present thesis, it has been excluded from the results section.

In the cellular environment, all the mutants are expected to behave as dominant negative mutants with respect to the endogenous wild type enzyme, as previously reported for SHMT1-IT (71).

Firstly, the ability of the mutated enzymes to enter the nucleus has been investigated through fluorescence microscopy. Indeed, the SHMT1 mutants were produced as GFP fusion proteins, and their localization was then assed following the fluorescent signal.

For all the proteins, including the wild type, a low number of cells showed a sharp nuclear fluorescent signal, suggesting, as expected, that the nuclear accumulation occurs in a specific phase of the cell cycle (i.e. S phase). Despite this, a nuclear accumulation was observed for all the mutants (Fig. 36), and the number of cells displaying this phenomenon was comparable to the one of cells

transfected with the wild type construct.

These results suggest that the oligomeric state and the catalytic activity of the enzyme are not relevant for the nuclear translocation of SHMT1.

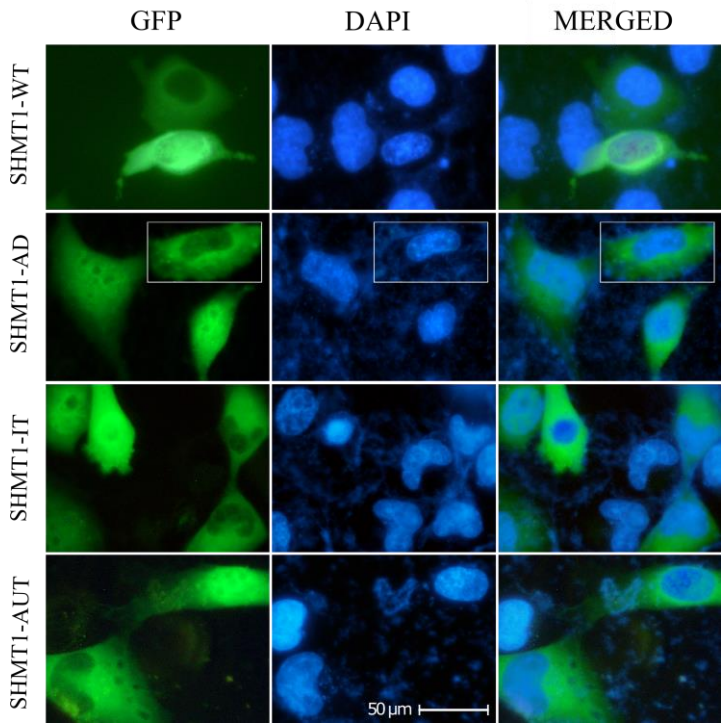


*Figure 36. SHMT1 nuclear localization. GFP (green), DAPI (blue) and merged staining representing the localization of SHMT variants transfected in A549 cells. The scale bar applies to all samples.*

We then studied the effect of the mutants expression in H1299 and also in another lung cancer model, the A549 line, focusing the analysis on cell survival and proliferation.

In both cell types, the expression of SHMT1-WT, SHMT1-AD, SHMT1-AUT does not cause significant modifications in the number of live and dead cells with respect to the samples transfected with the empty plasmid (pcDNA).

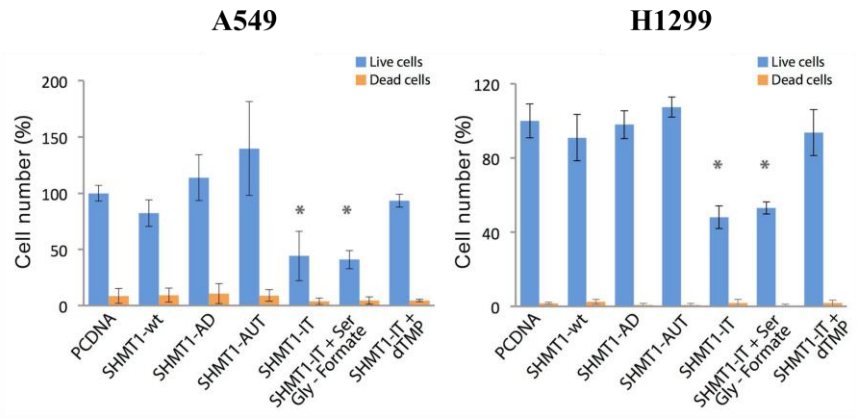


*Figure 37. Cell viability of A549 and H1299 lung cancer cells 48 hours after transfection with the control plasmid (pcDNA) or with the plasmids that induce the expression of SHMT1 variants in complete medium or complete medium supplemented as indicated. Statistical analyses are performed on three independent experiments, *P ≤ 0.05.*

On the other hand, the transfection with SHMT1-IT induces a dramatic decrease in the cell number after 48 hours, indicating that the inactive tetramer causes defects in the proliferation but, given the low quantity of dead cells, it does not affect cell viability (Fig. 37). To confirm that the impaired proliferation is due to a block in the thymidylate cycle, we supplemented the cells with deoxythymidine monophosphate (dTMP), the end product of the pathway (Fig. 38). Indeed this is the case, since the proliferation is completely rescued by the supplementation of dTMP in the culture medium, while supplementation with serine, glycine and formate (donor of 1C units) has no significant effects.

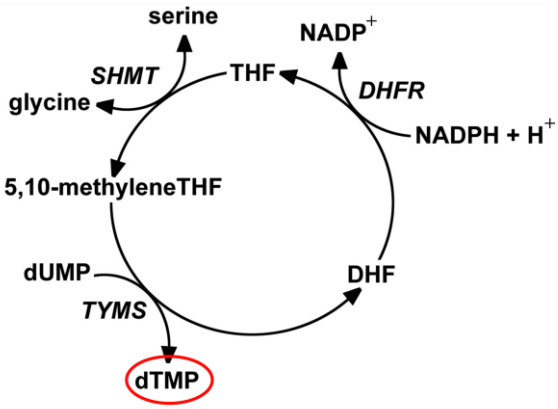


*Figure 38. Scheme illustrating the thymidylate cycle, the end product is circled in red. Abbreviations: DHF, dihydrofolate; DHFR, dihydrofolate reductase; dTMP, deoxythymidine monophosphate; dUMP, deoxyuridine monophosphate; NADP$^+$, nicotinamide adenine dinucleotide phosphate; NADPH reduced nicotinamide adenine dinucleotide phosphate; SHMT, serine hydroxymethyltransferase; THF, tetrahydrofolate; TYMS, thymidylate synthase.*

Overall the data suggest that, in these lung cancer models, the catalytic activity of SHMT1 is required for cell proliferation, to carry out the nuclear thymidylate biosynthesis. Given that neither the dimer nor the unstable tetramer, which are both catalytically active, impaired cell proliferation, apparently the oligomeric state of the protein does not affect the nuclear functions of SHMT1.

# 4. Discussion

## 4.1 The discovery of a nucleic acid binding protein

The work presented herein puts a spotlight on the nucleic acid binding property of SHMT1, so far overshadowed by the more characterized enzymatic activity. Being able to bind to RNA and DNA, we claim that human cytosolic serine hydroxymethyltransferase (SHMT1) can be ascribed to the family of DRBPs. Concerning this class of proteins, it is not easy to discriminate between "biological specificity" (i.e. the binding of sequence variants *in vivo*, often correlated to a functional effect) and "intrinsic specificity" (i.e. the binding of sequence variants *in vitro*, not necessarily occurring in cells); thus the present thesis was intended to characterize the SHMT1 RNA and DNA binding property and to investigate the nucleic acid targets having a functional relevance in lung cancer cells when bound to SHMT1.

The data demonstrates that SHMT1 forms a high affinity complex with the 5'UTR region of the SHMT2 transcript (namely UTR2) both *in vitro* and in H1299 cells, finally unraveling the molecular mechanism at the basis of the crosstalk between SHMT isozymes, previously observed in lung cancer (section 1.4). Furthermore, to go beyond the identification of a single interaction, RIP-seq analyses are currently being finalized to find the complete spectrum of RNA targets bound and regulated by SHMT1 in lung cancer lines. On the other hand, given the latest report on the strict association of SHMT1 with chromatin (72), we have also begun to address the question regarding the biological meaning of the SHMT1-DNA interaction; our preliminary analysis shows that the enzyme has an higher affinity for ssDNA with respect to

dsDNA, although specific DNA targets, if any, have not been identified so far. Given the lack of canonical nucleic acid binding domains, it is important to clarify how SHMT1 can interact with such molecules. The experiments on site directed mutants, together with the inhibition kinetics indicating the RNA-mediated hyperbolic inhibition and the EMSA assays showing a competition between folates and UTR2, have strongly suggested that the unconventional RNA binding site involves a part of the folate binding cleft. In addition, the different displacement effects observed with the mono- and pentaglutamylated 5-CHO-THF indicate that the polyglutamylated tail has a major role in the RNA displacement, leading to hypothesize that the nucleic acid binding site likely extends beyond the folate binding cleft, involving the positively charged area that surrounds the substrate binding site.

Given the involvement of K157, K158 and Y82 in the binding of both nucleic acids, we hypothesize that RNA and DNA may share a common binding region; however, we cannot exclude at this stage that the specificity for unique sequence variants is further mediated by dedicated residues.

Interestingly, SHMT1 is not the only moonlighting enzyme having a RNA/DNA binding site that overlaps, at least partially, with the substrate binding cleft (16, 73). The direct consequence of such phenomenon is a cellular multilayered regulation in which the levels of metabolites and nucleic acids modulate the fine balance between the catalytic and the nucleic acid binding activity of the enzyme (13). Since the amount of glycine and folate necessary to modulate *in vitro* the binding of UTR2 falls within the physiological range (74), we predict that in the cells polyglutamylated folates (the prevalent form of folates) can be effective competitors of the RNA and, most likely, also of the DNA.

Apparently the RNA binding function has been conserved among eukaryotic

SHMTs, since also Plasmodium SHMT (PfSHMT) is able to bind to specific RNA sequences (75), while the DNA binding activity has not been tested in other organisms yet. Considering that eukaryotic SHMTs have a common oligomeric state (i.e. tetrameric) which is different from the prokaryotic one (i.e. dimeric), the present work investigated if the tetrameric structure is responsible for the observed moonlighting property. Surprisingly, the results show that the SHMT1 moonlighting activity, and so possibly also the one of PfSHMT, is independent from the oligomeric structure of the enzyme; indeed the human dimeric mutant maintains such capability, whereas the prokaryotic dimeric protein from *E. coli* does not display any nucleic acid binding property. We thus conclude that the different behavior can be attributed to the amino acid sequence composition and to the surface charge distribution, significantly different between human and *E. coli* SHMT (76).

### 4.2 RNA-protein interaction: who affects whom?

The current perception of protein-RNA interactions is strongly biased by a protein-centric view derived from the past decades, in which proteins regulate the RNA expression, translation or any other process related to RNA biology. This work, together with many others emerging in the scientific literature (14, 77), witnesses that the RNA itself is able to act as a potent regulator. Hence there's the need to replace the old view with a new one in which proteins and RNAs lie on the same level.

The data presented herein are a good example of how RNA binding proteins partner with Protein binding RNAs to fine tune complex metabolic processes in the cell.

Indeed, the results show that the formation of the metabolite-sensitive UTR2-

SHMT1 complex exerts, at the same time, a double regulatory role: (i) it fine tunes the expression and translation of the SHMT2 mRNA and (ii) it strongly inhibits the enzymatic activity of SHMT1.

On one side, we have validated the initial hypothesis (section 1.6), showing that SHMT1 fine tunes the expression and translation of SHMT2 mRNA by binding to UTR2 in H1299 cells. While the translational inhibition is a biological event that frequently follows a protein-RNA interaction (17, 46, 78), on the other side, the data concerning the selective inhibition of the enzymatic activity were unprecedented. Indeed, the work shows that the cleavage of serine is strongly inhibited by the SHMT1-RNA interaction, with UTR2 being the most efficient inhibitor *in vitro* ($IC_{50}$ ~33 nM), while the serine synthesis is unexpectedly only slightly affected by all the RNAs tested.

We can explain these data considering that the SHMT-serine and SHMT-glycine complexes have a different conformational status, in particular L-serine binding induces an open-to-close conformational change that rearranges a protein region close to the active site, conversely, the binding of glycine doesn't induce such change (79). Therefore, it is conceivable that the interaction of the RNA with such dissimilar conformations may differentially affect the binding of folates and the subsequent catalytic steps: in detail, the binding of the RNA to the SHMT-serine complex might hamper the binding of folate substrate more that the binding of the RNA to the SHMT-glycine complex.

A detailed mechanistic explanation of our hypothesis is illustrated in the scheme shown in Figure 25G: assuming the value of the α factor to be higher in the forward SHMT reaction (serine to glycine) than in the reverse SHMT reaction (glycine to serine).

Coherently, we have previously reported a similar behavior for

3-bromopyruvate and the pyrazolopyran compound 2.12 (40, 80), small molecules inhibitors of SHMT1, both having a stronger inhibitory effect on the serine cleavage reaction with respect to the serine synthesis one. Importantly, the UTR2-mediated inhibition of SHMT1 was validated in living cells. The data presented herein show that UTR2 causes specific biological effects in two different cell models. Firstly, when overexpressed, it lowers SHMT activity, an effect that is not reported upon UTR1 transfection. Secondly, in a SHMT2-KO background, where the cytosolic isozyme must catalyze the serine cleavage reaction, serine consumption increases only upon UTR2 downregulation, whereas its overexpression leads to an accumulation of the amino acid in the medium; overall these data indicate that the serine to glycine reaction is inhibited and that such effect is UTR2-specific.

Interestingly, the UTR2-mediated inhibition of SHMT1 correlates with a significant reduction of cell viability in serine-containing medium, which can be attributed to the inability of the cytosolic enzyme to catabolize serine, given the formation of the SHMT1-UTR2 complex in absence of the mitochondrial counterpart.

The validation of the *in vitro* data in cell models was fundamental to infer that such biochemical events actually occur also in living cells, thus allowing to enlarge our view and to consider them as physiological mechanisms.

## 4.3 Understanding the biological meaning of SHMT1 moonlighting activity

The moonlighting activity of metabolic enzymes has been widely studied biochemically, and great efforts have also been made to understand the biological relevance of such properties; unfortunately, in most cases a clear picture is still missing (13, 20).

To get insights into the biological sense of the SHMT1 moonlighting activity it is crucial to consider the background in which the enzyme exerts its multiple functions: the one-carbon metabolism (1C-metabolism). Being compartmentalized between cytosol and mitochondrion, the 1C-metabolism involves proteins belonging to both cellular subfractions, linking the two compartments through the exchange of substrates (25). Cytosolic SHMT1 and mitochondrial SHMT2 are among the essential players of the 1-C metabolism in normal and tumor cells (26). As mentioned in section 1.3, in most cell lines SHMT1 is reported to shift the directionality of the reaction in accordance with the cellular needs, while SHMT2 is committed to the serine-derived synthesis of glycine (25). Nevertheless, the mechanism regulating this metabolic flux as well as why the cell has adopted such tasks distribution is still unclear.

This work presents a novel regulatory mechanism which interconnects the human SHMT isozymes in lung cancer cells: as mentioned in the previous paragraph, the formation of the UTR2-SHMT1 complex fine tunes the expression and translation of SHMT2 and, at the same time, modulates the enzymatic activity of SHMT1. Therefore, we hypothesize that when SHMT2 mRNA is highly expressed (e.g. in tumor cells), SHMT1 binds to it and the cytosolic conversion of glycine into serine is favored over serine cleavage, consequently, all the available serine (both exogenous and endogenous) is

catabolized in the mitochondrion through SHMT2 activity. Conversely, when the levels of SHMT2 mRNA are low, SHMT1 can partially work in the opposite direction, since the RNA-mediated inhibition of the serine cleavage reaction is partially relieved, in this way complementing the mitochondrial counterpart (Fig. 39). In parallel, the reduction of SHMT1 protein may affect the amount of SHMT2 mRNA available for translation, favoring such process.

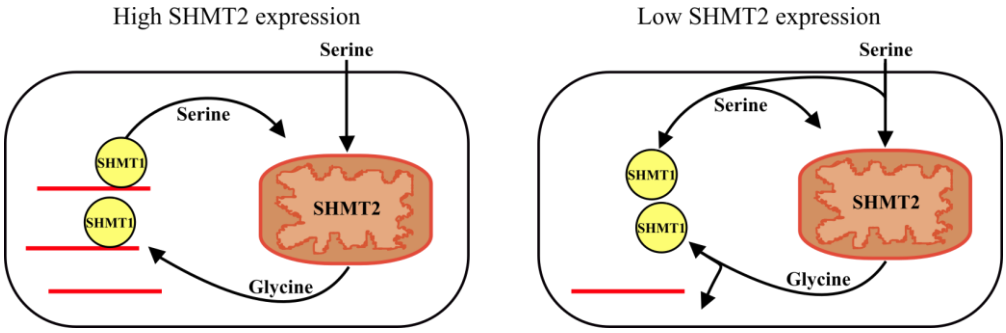


*Figure 39. Model illustrating the hypothesized regulatory mechanism in presence of high (left) and low (right) levels of SHMT2 mRNA. SHMT1 is represented with a yellow sphere while SHMT2 mRNA is shown as a red line.*

If we widen the view and consider the presence of many other RNAs bound to SHMT1 in the cytosol, it is reasonable to hypothesize that the synergic modulation of the enzyme's activity by such sequences might be a strong tool to limit the cytoplasmic consumption of serine, directing the major amount of this amino acid in the mitochondrion, where it is catabolized exclusively by SHMT2. At the same time, the stability and/or translation of all these RNA targets could be affected by SHMT1 binding. Certainly, other metabolic factors, such as the distribution and/or the levels of metabolites (folate species, glycine and NADH/NADPH ratio), and the expression of the RNA sequences contribute to maintain such regulatory mechanism. As a matter of fact, if the

activity of SHMT1 was not somehow directed toward serine synthesis, the enzyme would perform both reactions in the cell; however, it is important to point out that, given the extreme variability of the serine-glycine 1C metabolism (SGOCM) among different lines (25), and also between cells and tissues belonging to the same histotype (81), the specific traits of this regulation need a dedicated investigation.

A preliminary list of SHMT1-specific RNA targets, coming from our analysis of the RIP-seq data, is already available for both cell lines (H1299 and A549) and has revealed the enrichment of mRNAs encoding for proteins belonging to different cellular pathways. A gene ontology analysis is underway to clarify if the SHMT1 post-transcriptional regulation of mRNAs affects only specific pathways, or rather if it follows other criteria.

On the other side, the understanding of the biological relevance of the DNA binding ability of SHMT1 is still at its infancy. The present work is the first direct proof stating such capability *in vitro*, moreover, the higher affinity for ssDNA with respect to dsDNA leads to speculate on a possible nuclear function localized on the partially denaturated strands of DNA. In line with our speculation, literature reports indirect evidences claiming the presence of SHMT1 at the replication forks (single stranded) (32) and also the involvement of SHMT1 in the repair of UV-induced DNA damage, which manifests as strand breaks (82).

Surely, in our lung cancer models, the catalytic activity of SHMT1 is essential for nuclear *de novo* thymidylate biosynthesis, as shown by the experiments involving the catalytically inactive mutant. Once entered in the nucleus, the enzyme provides the substrates needed for dTMP synthesis, thus sustaining cell proliferation; given the ability of SHMT1 to bind to ssDNA, it is conceivable to think that such reaction could take place directly *in situ*, at the

replication forks or at the DNA strand breaks.

Overall the moonlighting activity of SHMT1 is a perfect example of how the human cells regulate metabolism by establishing a unique axis which links gene replication, gene expression and enzymes activity; in this view, it is striking how each element of such axis is a regulator of the others (RNA, metabolites, enzymes, probably also DNA) to ultimately ensure the survival of the cell.

# 5. Conclusions and Perspectives

The present study is pioneer in dissecting the nucleic acid binding features of human cytosolic serine hydroxymethyltransferase. In doing so, a novel and unpredicted layer of complexity has been added to the regulation of SHMT isozymes and serine metabolism in lung cancer.

All the findings presented herein fit perfectly with the current theory supporting the existence of a link among gene replication, gene expression and intermediary metabolism. Such interconnections allow the cell to rapidly respond to the environmental and/or intracellular metabolic changes. The deep knowledge of the multiple layers of regulation in specific cancer types is fundamental to understand, and consequently exploit, the metabolic vulnerabilities of each class of tumors in order to develop dedicated therapies. The affinity of SHMT1 for RNA sequences, reported in the present study, can be exploited to develop RNA-based metabolic therapies. Indeed, the treatment of cancer cells with aptamers having low nanomolar affinity for SHMT1 would strongly unbalance tumor intermediary metabolism, thus leading to a block of cell proliferation.

Yet, a lot of nodes still need to be loosened: the incoming data from RIP-seq analyses will soon provide a global picture of the network interconnecting SHMT1 and RNA regulation. On the other hand, since a wide range of metabolic enzymes have been shown to enter the nucleus (9), and for some of them the spectrum of DNA targets is already known (72, 83), it is essential to explore the enrichment of the SHMT1 protein upon specific DNA sequences, if any, in this way deepening the knowledge on the functional relevance of the DNA binding ability related to the nuclear SHMT functions, to date still mostly unknown.

From a structural point of view, future analysis will be needed to fully identify the non-canonical nucleic acid binding domain as well as its complexes with RNA or DNA. The identification of such unconventional site could ultimately lead to the discovery of other novel unconventional DRBPs, indeed, since some of the proteins belonging to the fold type I pyridoxal phosphate-dependent enzymes family have a quaternary structure resembling the one of SHMT1 (84, 85), it is reasonable to think that the non-canonical nucleic acid binding domain could be conserved among some of them.

We shall not forget about SHMT2 and SHMT2α, encoded by the same gene and localized respectively in mitochondrion and cytosol. A moonlighting function in the assembly of a deubiquitylating complex has been recently reported for the cytosolic isoform (SHMT2α); such activity is carried out through the population of a dimeric state of the protein (35), feasible thanks to the PLP dependent regulation which modulates the SHMT2 oligomeric state (23). Notably, SHMT2 has also been found in RNA interactome capture analysis conducted in HeLa cells (86), employed to study the mRNA interactome composition. Overall, such data concerning the SHMT2 isoforms raise the question of whether these proteins display different moonlighting functions as well. Importantly, their putative RNA binding capacity needs to be carefully investigated, in analogy to what has been already done in the present thesis for SHMT1.

The fascination of the complex picture of the intermediary metabolism resides in looking at it from different angles, in doing so, the observer can capture the dynamic coexistence of different nuances which are invisible if observed from a single point.

# 6. References


1. Jeffery,C.J. (1999) Moonlighting proteins. *Trends Biochem. Sci.*, **24**, 8–11.
2. Huberts,D.H.E.W. and van der Klei,I.J. (2010) Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta*, **1803**, 520–5.
3. Jeffery,C.J. (2015) Why study moonlighting proteins? *Front. Genet.*, **6**, 211.
4. Espinosa-Cantú,A., Ascencio,D., Barona-Gómez,F. and DeLuna,A. (2015) Gene duplication and the evolution of moonlighting proteins. *Front. Genet.*, **6**, 227.
5. Copley,S.D. (2014) An evolutionary perspective on protein moonlighting. *Biochem. Soc. Trans.*, **42**, 1684–91.
6. Mani,M., Chen,C., Amblee,V., Liu,H., Mathur,T., Zwicke,G., Zabad,S., Patel,B., Thakkar,J. and Jeffery,C.J. (2015) MoonProt: a database for proteins that are known to moonlight. *Nucleic Acids Res.*, **43**, D277–D282.
7. Chen,C., Zabad,S., Liu,H., Wang,W. and Jeffery,C. (2018) MoonProt 2.0: an expansion and update of the moonlighting proteins database. *Nucleic Acids Res.*, **46**, D640–D644.
8. Hendriks,W., Mulders,J.W., Bibby,M.A., Slingsby,C., Bloemendal,H. and de Jong,W.W. (1988) Duck lens epsilon-crystallin and lactate dehydrogenase B4 are identical: a single-copy gene product with two distinct functions. *Proc. Natl. Acad. Sci.*, **85**, 7114–7118.
9. Boukouris,A.E., Zervopoulos,S.D. and Michelakis,E.D. (2016) Metabolic Enzymes Moonlighting in the Nucleus: Metabolic Regulation of Gene Transcription. *Trends Biochem. Sci.*, **41**, 712–730.

10. Yu,X. and Li,S. (2017) Non-metabolic functions of glycolytic enzymes in tumorigenesis. *Oncogene*, **36**, 2629–2636.
11. GROSSE,F., NASHEUER,H.-P., SCHOLTISSEK,S. and SCHOMBURG,U. (1986) Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-alpha-primase complex. *Eur. J. Biochem.*, **160**, 459–467.
12. Popanda,O., Fox,G. and Thielmann,H.W. (1998) Modulation of DNA polymerases alpha, delta and epsilon by lactate dehydrogenase and 3-phosphoglycerate kinase. *Biochim. Biophys. Acta*, **1397**, 102–17.
13. Hentze,M.W. and Preiss,T. (2010) The REM phase of gene regulation. *Trends Biochem. Sci.*, **35**, 423–426.
14. Beckmann,B.M., Castello,A. and Medenbach,J. (2016) The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. *Pflügers Arch. - Eur. J. Physiol.*, **468**, 1029–1040.
15. Cieśla,J. (2006) Metabolic enzymes that bind RNA: yet another level of cellular regulatory network? *Acta Biochim. Pol.*, **53**, 11–32.
16. Hentze,M.W., Castello,A., Schwarzl,T. and Preiss,T. (2018) A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.*, 10.1038/nrm.2017.130.
17. Chang,C.-H., Curtis,J.D., Maggi,L.B., Faubert,B., Villarino,A. V, O'Sullivan,D., Huang,S.C.-C., van der Windt,G.J.W., Blagih,J., Qiu,J., *et al.* (2013) Posttranscriptional control of T cell effector function by aerobic glycolysis. *Cell*, **153**, 1239–51.
18. Hentze,M.W., Castello,A., Schwarzl,T. and Preiss,T. (2018) A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.*, **19**, 327–341.

19. Nagy,E. and Rigby,W.F.C. (1995) Glyceraldehyde-3-phosphate Dehydrogenase Selectively Binds AU-rich RNA in the NAD $^{+}$ -binding Region (Rossmann Fold). *J. Biol. Chem.*, **270**, 2755–2763.
20. Albihlal,W.S. and Gerber,A.P. (2018) Unconventional RNA-binding proteins: an uncharted zone in RNA biology. *FEBS Lett.*, **592**, 2917–2931.
21. Florio,R., di Salvo,M.L., Vivoli,M. and Contestabile,R. (2011) Serine hydroxymethyltransferase: A model enzyme for mechanistic, structural, and evolutionary studies. *Biochim. Biophys. Acta - Proteins Proteomics*, **1814**, 1489–1496.
22. Renwick,S.B., Snell,K. and Baumann,U. (1998) The crystal structure of human cytosolic serine hydroxymethyltransferase: a target for cancer chemotherapy. *Structure*, **6**, 1105–16.
23. Giardina,G., Brunotti,P., Fiascarelli,A., Cicalini,A., Costa,M.G.S., Buckle,A.M., di Salvo,M.L., Giorgi,A., Marani,M., Paone,A., *et al.* (2015) How pyridoxal 5′-phosphate differentially regulates human cytosolic and mitochondrial serine hydroxymethyltransferase oligomeric state. *FEBS J.*, **282**, 1225–1241.
24. Tramonti,A., Nardella,C., di Salvo,M.L., Barile,A., Cutruzzolà,F. and Contestabile,R. (2018) Human Cytosolic and Mitochondrial Serine Hydroxymethyltransferase Isoforms in Comparison: Full Kinetic Characterization and Substrate Inhibition Properties. *Biochemistry*, **57**, 6984–6996.
25. Ducker,G.S., Chen,L., Morscher,R.J., Ghergurovich,J.M., Esposito,M., Teng,X., Kang,Y. and Rabinowitz,J.D. (2016) Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway. *Cell Metab.*, **23**, 1140–1153.

26. Ducker,G.S. and Rabinowitz,J.D. (2017) One-Carbon Metabolism in Health and Disease. *Cell Metab.*, **25**, 27–42.
27. Anderson,D.D., Quintero,C.M. and Stover,P.J. (2011) Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proc. Natl. Acad. Sci.*, **108**, 15163–15168.
28. Meiser,J., Schuster,A., Pietzke,M., Voorde,J. Vande, Athineos,D., Oizel,K., Burgos-Barragan,G., Wit,N., Dhayade,S., Morton,J.P., *et al.* (2018) Increased formate overflow is a hallmark of oxidative cancer. *Nat. Commun.*, **9**.
29. Minton,D.R., Nam,M., McLaughlin,D.J., Shin,J., Bayraktar,E.C., Alvarez,S.W., Sviderskiy,V.O., Papagiannakopoulos,T., Sabatini,D.M., Birsoy,K., *et al.* (2018) Serine Catabolism by SHMT2 Is Required for Proper Mitochondrial Translation Initiation and Maintenance of Formylmethionyl-tRNAs. *Mol. Cell*, **69**, 610-621.e5.
30. Morscher,R.J., Ducker,G.S., Li,S.H.J., Mayer,J.A., Gitai,Z., Sperl,W. and Rabinowitz,J.D. (2018) Mitochondrial translation requires folate-dependent tRNA methylation. *Nature*, **554**, 128–132.
31. Pacold,M.E., Brimacombe,K.R., Chan,S.H., Rohde,J.M., Lewis,C.A., Swier,L.J.Y.M., Possemato,R., Chen,W.W., Sullivan,L.B., Fiske,B.P., *et al.* (2016) A PHGDH inhibitor reveals coordination of serine synthesis and one-carbon unit fate. *Nat. Chem. Biol.*, **12**, 452–458.
32. Anderson,D.D., Woeller,C.F., Chiang,E.P., Shane,B. and Stover,P.J. (2012) Serine hydroxymethyltransferase anchors de Novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis. *J. Biol. Chem.*, **287**, 7051–7062.

33. Anderson,D.D. and Stover,P.J. (2009) SHMT1 and SHMT2 are functionally redundant in nuclear de novo thymidylate biosynthesis. *PLoS One*, **4**.

34. Rabl,J., Bunker,R.D., Schenk,A.D., Cavadini,S., Gill,M.E., Abdulrahman,W., Andrés-Pons,A., Luijsterburg,M.S., Ibrahim,A.F.M., Branigan,E., *et al.* (2019) Structural Basis of BRCC36 Function in DNA Repair and Immune Regulation. *Mol. Cell*, **75**, 483-497.e9.

35. Walden,M., Tian,L., Ross,R.L., Sykora,U.M., Byrne,D.P., Hesketh,E.L., Masandi,S.K., Cassel,J., George,R., Ault,J.R., *et al.* (2019) Metabolic control of BRISC–SHMT2 assembly regulates immune signalling. *Nature*, **570**, 194–199.

36. Mattaini,K.R., Sullivan,M.R. and Heiden,M.G. Vander (2016) The importance of serine metabolism in cancer. *J. Cell Biol.*, **214**, 249–257.

37. Paone,A., Marani,M., Fiascarelli,A., Rinaldo,S., Giardina,G., Contestabile,R., Paiardini,A. and Cutruzzolà,F. (2014) SHMT1 knockdown induces apoptosis in lung cancer cells by causing uracil misincorporation. *Cell Death Dis.*, **5**, 1–11.

38. Woo,C.C., Chen,W.C., Teo,X.Q., Radda,G.K. and Lee,P.T.H. (2016) Downregulating serine hydroxymethyltransferase 2 (SHMT2) suppresses tumorigenesis in human hepatocellular carcinoma. *Oncotarget*, **7**, 53005–53017.

39. Ducker,G.S., Ghergurovich,J.M., Mainolfi,N., Suri,V., Jeong,S.K., Hsin-Jung Li,S., Friedman,A., Manfredi,M.G., Gitai,Z., Kim,H., *et al.* (2017) Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 11404–11409.

40. Tramonti,A., Paiardini,A., Paone,A., Bouzidi,A., Giardina,G., Guiducci,G., Magnifico,M.C., Rinaldo,S., McDermott,L., Menendez,J.A., *et al.* (2018) Differential inhibitory effect of a pyrazolopyran compound on human serine hydroxymethyltransferase-amino acid complexes. *Arch. Biochem. Biophys.*, 10.1016/J.ABB.2018.07.001.
41. Nikiforov,M.A., Chandriani,S., O'Connell,B., Petrenko,O., Kotenko,I., Beavis,A., Sedivy,J.M. and Cole,M.D. (2002) A functional screen for Myc-responsive genes reveals serine hydroxymethyltransferase, a major source of the one-carbon unit for cell metabolism. *Mol. Cell. Biol.*, **22**, 5793–800.
42. Ye,J., Fan,J., Venneti,S., Wan,Y.-W., Pawel,B.R., Zhang,J., Finley,L.W.S., Lu,C., Lindsten,T., Cross,J.R., *et al.* (2014) Serine catabolism regulates mitochondrial redox control during hypoxia. *Cancer Discov.*, **4**, 1406–17.
43. Anderson,D.D., Eom,J.Y. and Stover,P.J. (2012) Competition between Sumoylation and Ubiquitination of Serine Hydroxymethyltransferase 1 Determines Its Nuclear Localization and Its Accumulation in the Nucleus. *J. Biol. Chem.*, **287**, 4790–4799.
44. Yang,X., Wang,Z., Li,X., Liu,B., Liu,M., Liu,L., Chen,S., Ren,M., Wang,Y., Yu,M., *et al.* (2018) Shmt2 desuccinylation by SIRT5 drives cancer cell proliferation. *Cancer Res.*, **78**, 372–386.
45. Zhang,W.C., Shyh-Chang,N., Yang,H., Rai,A., Umashankar,S., Ma,S., Soh,B.S., Sun,L.L., Tai,B.C., Nga,M.E., *et al.* (2012) Glycine Decarboxylase Activity Drives Non-Small Cell Lung Cancer Tumor-Initiating Cells and Tumorigenesis. *Cell*, **148**, 259–272.

46. Liu,X., Reig,B., Nasrallah,I.M. and Stover,P.J. (2000) Human cytoplasmic serine hydroxymethyltransferase is an mRNA binding protein. *Biochemistry*, **39**, 11523–11531.
47. Chu,E., Takimoto,C.H., Voeller,D., Grem,J.L. and Allegra,C.J. (1993) Specific binding of human dihydrofolate reductase protein to dihydrofolate reductase messenger RNA in vitro. *Biochemistry*, **32**, 4756–4760.
48. Liu,J., Schmitz,J.C., Lin,X., Tai,N., Yan,W., Farrell,M., Bailly,M., Chen,T.M. and Chu,E. (2002) Thymidylate synthase as a translational regulator of cellular gene expression. *Biochim. Biophys. Acta - Mol. Basis Dis.*, **1587**, 174–182.
49. Jankowsky,E. and Harris,M.E. (2015) Specificity and nonspecificity in RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.*, **16**, 533–44.
50. Guenther,U.-P., Yandek,L.E., Niland,C.N., Campbell,F.E., Anderson,D., Anderson,V.E., Harris,M.E. and Jankowsky,E. (2013) Hidden specificity in an apparently nonspecific RNA-binding protein. *Nature*, **502**, 385–8.
51. Cassiday,L.A. and Maher,L.J. (2002) Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Res.*, **30**, 4118–4126.
52. Hudson,W.H. and Ortlund,E.A. (2014) The structure, function and evolution of proteins that bind DNA and RNA. *Nat. Rev. Mol. Cell Biol.*, **15**, 749–760.
53. Fu,T.F., Rife,J.P. and Schirch,V. (2001) The role of serine hydroxymethyltransferase isozymes in one-carbon metabolism in MCF-7 cells as determined by (13)C NMR. *Arch. Biochem. Biophys.*, **393**, 42–50.

54. Schirch,L. and Ropp,M. (1967) Serine transhydroxymethylase. Affinity of tetrahydrofolate compounds for the enzyme and enzyme-glycine complex. *Biochemistry*, **6**, 253–7.

55. Stover,P.J., Chen,L.H., Suh,J.R., Stover,D.M., Keyomarsi,K. and Shane,B. (1997) Molecular cloning, characterization, and regulation of the human mitochondrial serine hydroxymethyltransferase gene. *J. Biol. Chem.*, **272**, 1842–8.

56. Marani,M., Paone,A., Fiascarelli,A., Macone,A., Gargano,M., Rinaldo,S., Giardina,G., Pontecorvi,V., Koes,D., McDermott,L., *et al.* (2016) A pyrazolopyran derivative preferentially inhibits the activity of human cytosolic serine hydroxymethyltransferase and induces cell death in lung cancer cells. *Oncotarget*, **7**, 4570–4583.

57. Gao,X., Pujos-Guillot,E., Martin,J.-F., Galan,P., Juste,C., Jia,W. and Sebedio,J.-L. (2009) Metabolite analysis of human fecal water by gas chromatography/mass spectrometry with ethyl chloroformate derivatization. *Anal. Biochem.*, **393**, 163–75.

58. Mignone,F., Gissi,C., Liuni,S. and Pesole,G. (2002) Untranslated regions of mRNAs. *Genome Biol.*, **3**, REVIEWS0004.

59. Tomczak,K., Czerwińska,P. and Wiznerowicz,M. (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Poznan, Poland)*, **19**, A68-77.

60. Huang,T., Wang,C., Maras,B., Barra,D. and Schirch,V. (1998) Thermodynamic analysis of the binding of the polyglutamate chain of 5-formyltetrahydropteroylpolyglutamates to serine hydroxymethyltransferase. *Biochemistry*, **37**, 13536–42.

61. Fu,T.-F., Scarsdale,J.N., Kazanina,G., Schirch,V. and Wright,H.T. (2003) Location of the pteroylpolyglutamate-binding site on rabbit cytosolic serine hydroxymethyltransferase. *J. Biol. Chem.*, **278**, 2645–53.
62. Cléry,A., Blatter,M. and Allain,F.H.-T. (2008) RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.*, **18**, 290–8.
63. Szebenyi,D.M.E., Liu,X., Kriksunov,I.A., Stover,P.J. and Thiel,D.J. (2000) Structure of a Murine Cytoplasmic Serine Hydroxymethyltransferase Quinonoid Ternary Complex: Evidence for Asymmetric Obligate Dimers †. *Biochemistry*, **39**, 13313–13323.
64. Amornwatcharapong,W., Maenpuen,S., Chitnumsub,P., Leartsakulpanich,U. and Chaiyen,P. (2017) Human and Plasmodium serine hydroxymethyltransferases differ in rate-limiting steps and pH-dependent substrate inhibition behavior. *Arch. Biochem. Biophys.*, **630**, 91–100.
65. Yoshino,M. (1987) A graphical method for determining inhibition parameters for partial and complete inhibitors. *Biochem. J.*, **248**, 815–20.
66. Cuyàs,E., Fernández-Arroyo,S., Verdura,S., García,R.Á.-F., Stursa,J., Werner,L., Blanco-González,E., Montes-Bayón,M., Joven,J., Viollet,B., *et al.* (2018) Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism. *Oncogene*, **37**, 963–970.
67. Stover,P. and Schirch,V. (1991) 5-Formyltetrahydrofolate polyglutamates are slow tight binding inhibitors of serine hydroxymethyltransferase. *J. Biol. Chem.*, **266**, 1543–50.

68. Besson,V., Rebeille,F., Neuburger,M., Douce,R. and Cossins,E.A. (1993) Effects of tetrahydrofolate polyglutamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. *Biochem. J.*, **292**, 425–430.

69. Giardina,G., Paone,A., Tramonti,A., Lucchi,R., Marani,M., Magnifico,M.C., Bouzidi,A., Pontecorvi,V., Guiducci,G., Zamparelli,C., *et al.* (2018) The catalytic activity of serine hydroxymethyltransferase is essential for *de novo* nuclear dTMP synthesis in lung cancer cells. *FEBS J.*, 10.1111/febs.14610.

70. Woeller,C.F., Anderson,D.D., Szebenyi,D.M.E. and Stover,P.J. (2007) Evidence for Small Ubiquitin-like Modifier-dependent Nuclear Import of the Thymidylate Biosynthesis Pathway. *J. Biol. Chem.*, **282**, 17623–17631.

71. Zanetti,K.A. and Stover,P.J. (2003) Pyridoxal phosphate inhibits dynamic subunit interchange among serine hydroxymethyltransferase tetramers. *J. Biol. Chem.*, **278**, 10142–9.

72. Sdelci,S., Rendeiro,A.F., Rathert,P., You,W., Lin,J.-M.G., Ringler,A., Hofstätter,G., Moll,H.P., Gürtl,B., Farlik,M., *et al.* (2019) MTHFD1 interaction with BRD4 links folate metabolism to transcriptional regulation. *Nat. Genet.*, **51**, 990–998.

73. Hentze,M.W. (1994) Enzymes as RNA-binding proteins: a role for (di)nucleotide-binding domains? *Trends Biochem. Sci.*, **19**, 101–3.

74. Chen,L., Ducker,G.S., Lu,W., Teng,X. and Rabinowitz,J.D. (2017) An LC-MS chemical derivatization method for the measurement of five different one-carbon states of cellular tetrahydrofolate. *Anal. Bioanal. Chem.*, **409**, 5955–5964.

75. Pang,C.K.T., Hunter,J.H., Gujjar,R., Podutoori,R., Bowman,J., Mudeppa,D.G. and Rathod,P.K. (2009) Catalytic and ligand-binding characteristics of Plasmodium falciparum serine hydroxymethyltransferase. *Mol. Biochem. Parasitol.*, **168**, 74–83.
76. Scarsdale,J.N., Radaev,S., Kazanina,G., Schirch,V. and Wright,H.. (2000) Crystal structure at 2.4 Å resolution of E. coli serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate 1 1Edited by I. A. Wilson. *J. Mol. Biol.*, **296**, 155–168.
77. Balcerak,A., Trebinska-Stryjewska,A., Konopinski,R., Wakula,M. and Grzybowska,E.A. (2019) RNA–protein interactions: disorder, moonlighting and junk contribute to eukaryotic complexity. *Open Biol.*, **9**, 190096.
78. Klausner,R.D. and Rouault,T.A. (1993) A double life: cytosolic aconitase as a regulatory RNA binding protein. *Mol. Biol. Cell*, **4**, 1–5.
79. Schirch,V., Shostak,K., Zamora,M. and Guatam-Basak,M. (1991) The origin of reaction specificity in serine hydroxymethyltransferase. *J. Biol. Chem.*, **266**, 759–64.
80. Paiardini,A., Tramonti,A., Schirch,D., Guiducci,G., di Salvo,M.L., Fiascarelli,A., Giorgi,A., Maras,B., Cutruzzolà,F. and Contestabile,R. (2016) Differential 3-bromopyruvate inhibition of cytosolic and mitochondrial human serine hydroxymethyltransferase isoforms, key enzymes in cancer metabolic reprogramming. *Biochim. Biophys. Acta - Proteins Proteomics*, **1864**, 1506–1517.

81. Fan,T.W.M., Bruntz,R.C., Yang,Y., Song,H., Chernyavskaya,Y., Deng,P., Zhang,Y., Shah,P.P., Beverly,L.J., Qi,Z., *et al.* (2019) De novo synthesis of serine and glycine fuels purine nucleotide biosynthesis in human lung cancer tissues. *J. Biol. Chem.*, 10.1074/jbc.RA119.008743.

82. Fox,J.T., Shin,W.K., Caudill,M.A. and Stover,P.J. (2009) A UV-responsive internal ribosome entry site enhances serine hydroxymethyltransferase 1 expression for DNA damage repair. *J. Biol. Chem.*, **284**, 31097–108.

83. Feo,S., Arcuri,D., Piddini,E., Passantino,R. and Giallongo,A. (2000) ENO1 gene product binds to the c- *myc* promoter and acts as a transcriptional repressor: relationship with Myc promoter-binding protein 1 (MBP-1). *FEBS Lett.*, **473**, 47–52.

84. Palioura,S., Sherrer,R.L., Steitz,T.A., Söll,D. and Simonovic,M. (2009) The human SepSecS-tRNASec complex reveals the mechanism of selenocysteine formation. *Science*, **325**, 321–5.

85. Fesko,K., Suplatov,D. and Švedas,V. (2018) Bioinformatic analysis of the fold type I PLP-dependent enzymes reveals determinants of reaction specificity in l-threonine aldolase from Aeromonas jandaei. *FEBS Open Bio*, **8**, 1013–1028.

86. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M., *et al.* (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, **149**, 1393–1406.

# Publications

**Guiducci G**, Paone A, Tramonti A, Giardina G, Rinaldo S, Bouzidi A, Magnifico MC, Marani M, Menendez JA, Fatica A, Macone A, Armaos A, Tartaglia GG, Contestabile R, Paiardini A, Cutruzzolà F "The moonlighting RNA-binding activity of cytosolic serine hydroxymethyltransferase contributes to control compartmentalization of serine metabolism". *Nucleic Acids Res 2019 May 7;47(8):4240-4254.*

Giardina G, Paone A, Tramonti A, Lucchi R, Marani M, Magnifico MC, Bouzidi A, Pontecorvi V, **Guiducci G**, Zamparelli C, Rinaldo S, Paiardini A, Contestabile R, Cutruzzolà F "The catalytic activity of serine hydroxymethyltransferase is essential for de novo nuclear dTMP synthesis in lung cancer cells".
*FEBS J 2018 Sep;285(17):3238-3253.*

Tramonti A, Paiardini A, Paone A, Bouzidi A, Giardina G, **Guiducci G**, Magnifico MC, Rinaldo S, McDermott L, Menendez JA, Contestabile R, Cutruzzolà F "Differential inhibitory effect of a pyrazolopyran compound on human serine hydroxymethyltransferase-amino acid complexes". *Arch Biochem Biophys 2018 Sep 1;653:71-79.*

**Ringraziamenti**

Caro lab Cutruzzolà
mi hai portata fino a qua
corre veloce questo tempo
più di un gel al 4%.

Tante cose ho imparato
ad osservar da più di un lato
perchè la scienza è un po' balorda
ti dà certezze e poi le smonta.

Giorgio Alessio e Serena
sono come un fiume in piena
che scorre forte, e anche se affanna
c'è l'aiuto di Giovanna.

Gli apulini sempre presenti
son passati al CU20
per me sono dei gioielli
ad un passo dal Fanelli.

Manca solo il CRG
Gian, Elias e quelli lì
Catalogna, vista mare
con Michele e con Ale.

Una chiave appende fuori
nonostante lei lavori
legge lampo, sarà esperienza
si diverte con la scienza,

è la Prof. Cutruzzolà
in quell'ufficetto là
una guida assai vicina
come la linea blu in piscina.

Con questi versi vi ringrazio
per la fiducia e per lo spazio
per il tempo condiviso
che mi lascia un gran sorriso.

Un grande grazie anche:

A Federico Mantoni, Maria Chiara Magnifico, Amani Bouzidi, Marina Marani, Roberta Lucchi, Chiara di Lucente, Valentina Zollet e Sharon Spizzichino, per la piacevole compagnia dentro e fuori il laboratorio.

A tutti i dottorandi del Dipartimento di Biochimica, per l'aiuto e il supporto che ci siamo sempre scambiati.

Al Prof. Francesco Malatesta, per la sua dedizione e il suo costante supporto nei confronti di noi studenti di dottorato.

Ad Ivano Amelio e Andrea Cerase per la revisione della tesi.