SHMT1-RNA interaction dynamically regulates serine and glycine concentration in cells

Michele Monti*†,[1,2] Giulia Guiducci*,[3,4] Alessio Paone,[3] Serena Rinaldo,[3] Giorgio Giardina,[3] Francesca Cutruzzola†,[3] and Gian Gaetano Tartaglia†[1,5,3]

[1] *Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Dr. Aiguader 88, 08003 Barcelona*

[2] *RNA System Biology Lab, Department of Neuroscience and Brain Technologies, Istituto Italiano di Tecnologia (IIT), Via Morego 30, 16163, Genoa, Italy*

[3] *La Sapienza University, Rome Piazzale Aldo Moro 5*

[4] *Barts Cancer Institute, Centre for Cancer Cell and Molecular Biology, John Vane Science Centre, Charterhouse Square, Queen Mary University of London, London EC1M 6BQ, UK.*

[5] *RNA System Biology Lab, Department of Neuroscience and Brain Technologies (IIT), Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genoa, Italy*

Serine hydroxymethyltransferase (SHMT) regulates the serine-glycine one carbon metabolism and plays a role in cancer metabolic reprogramming. Two SHMT isozymes are acting in the cell: SHMT1 encoding the cytoplasmic isozyme (SHMT1) and SHMT2 encoding the mitochondrial one (SHMT2). Here we present a molecular model built on experimental data reporting the interaction between SHMT1 protein and SHMT2 mRNA, recently discovered in lung cancer cells. Using a stochastic dynamic model, we show that RNA moieties dynamically regulate serine and glycine concentration, shaping the system behaviour. For the first time we observe an active functional role of the RNA in the regulation of the serine-glycine metabolism and availability, which unravels a complex layer of regulation that cancer cells exploit to fine tune aminoacids availability according to their metabolic needs. The quantitative model presented herein exploits RNA molecules as metabolic switches of the SHMT1 activity. Our results pave the way for the development of RNA-based molecules able to unbalance serine metabolism in cancer cells.

Introduction

The last decade has seen the rise of an RNA-centric view of many biological processes, which now flanks the more classical protein-centric view [1] [2]. The transcriptome is no longer seen just in between genome and proteome and can be now regarded as a key player in many cellular processes such as differentiation, proliferation, apoptosis and many others [3] [4],[5] [6].

Many non-canonical RNA binding proteins (RBPs) have been shown to interact with transcripts through novel RNA binding domains (RBDs) [3] [7]. More than 100 metabolic enzymes have been recognized as unconventional RBPs [3] [8] [9], and so far many efforts have been made to understand the hidden connections among intermediary metabolism, RNA biology and gene expression [10] [11] [12]. The non-canonical activity of numerous metabolic enzymes, also called 'moonlighting' [13], together with the emerging key roles of RNA molecules in several biological processes, could represent a complex layer of regulation in which metabolites, RNA and metabolic enzymes co-ordinately control important biochemical pathways such as glycolysis, nucleotide biosynthesis, aminoacid metabolism and many others.

Human serine hydroxymethyltransferase (SHMT) is a central enzyme of the serine-glycine-one-carbon metabolism (SGOC). It catalyses the reversible conversion of serine into glycine, transferring a one-carbon (1C) unit to tetrahydrofolate (THF) to generate 5,10-methylentetrahydrofolate (Me-THF). Two different SHMT isoforms have been characterized in humans: SHMT1, cytosolic, and SHMT2, mitochondrial [14] [15][16]. Indeed, unlike prokaryotic organisms, the SGOC metabolism is mainly compartmentalized between cytoplasm and mitochondrion in all eukaryotes [17], however how the pool of enzymes, folates and aminoacids are co-ordinately regulated as well as how the two compartments communicate is still unknown. In this work we show that the recently discovered RNA-binding ability of SHMT1 [18] links and regulates the cytoplasmic and mitochondrial SGOC metabolism, thus shaping the concentration of serine and glycine in the cell. In agreement with the evolutionary development of a compartmentalized SGOC metabolism in eukaryotes, we show that the bacterial SHMT isoform does not have a strong RNA binding propensity. We herein propose that the RNA binding property of SHMT1 has evolved to finely coordinate the cytoplasmic and micothondrial SGOC metabolism. Indeed, we show a neat example of how the RNAs are not only exploited to synthetise proteins, but also play an active functional role into the biochemical networks. It is possible to speculate that the choice of using an already existing element of the interaction networks, without the need to recruit another protein to

*These authors contributed equally to the present work.

†Corresponding authors: michele.monti@iit.it (computational parts), francesca.cutruzzola@uniroma1.it (experimental part), gian.tartaglia@iit.it (computational parts)

perform an allosteric regulation of human SHMT1, can be seen as a sign of optimization driven by the evolution.

SHMT isozymes are also key players in the metabolic reprogramming of human tumours and therefore considered attractive targets for cancer therapy. However, the complex dynamic structure of the SHMT active site has so far hindered the development of effective inhibitors [19][20].

SHMT1 is a non-canonical RBP characterised *in vitro* and studied in lung cancer models [21] [18]. SHMT1 binds to SHMT2 mRNA through a RBD that, interestingly, partially overlaps with the folates binding cleft, potentially leading to a metabolite-mediated control of the interaction [18]. What is striking about such discovery is that the RNA affects the directionality of the SHMT1 reaction, favoring the cytosolic synthesis of serine rather than its cleavage [18]. Experiments carried out in multiple cancer lines indicate that SHMT1 is mainly committed to synthesize serine rather than glycine, whereas the mitochondrial counterpart is involved only in serine catabolism, although the drivers of such metabolic compartmentalization have not been clarified [15].

Although it is already known that RNA moieties are able to inhibit or boost enzymatic reactions [22] [23] , effects on the reaction equilibria have never been reported before. Exploiting a stochastic dynamic model which reliably reproduces a SHMT-centric biochemical system, the present work aims to investigate whether RNA molecules can dynamically regulate serine and glycine concentration by modulating the SHMT1 reaction. Our results show that the RNA shapes the system's behaviour, allowing to explore multiple states in which the amounts of serine and glycine vary qualitatively. Our model proves that the SHMT1-interacting RNAs (i.e. SHMT2 mRNA and other potential interactors) link cytosolic and mitochondrial serine metabolism, fine tuning the amount of serine that can be directed into the mitochondrion.

In a broader perspective, the upregulation or downregulation of the SHMT1-interacting RNAs can represent an unexpected tool through which the cell fine tunes serine metabolism according to its metabolic needs. Importantly, given the pivotal role of serine in the proliferation of certain types of tumors [24] [25] [26], it is conceivable that cancer cells could exploit the SHMT1-interacting RNAs to reprogram their metabolism in order to get higher levels of available serine.

Results

Our model aims to reliably reproduce the intricate system of biochemical reactions which gives rise to the cross-talk between SHMT isozymes in human cells [27] [18], thus linking cytosolic and mitochondrial serine-glycine one carbon metabolism.

Setting the SHMT isozymes at the centre of the model, we calculated its interactions propensity with RNA molecules and developed a chemical rate equations system that implements the concentration dynamic of all chemical species involved over time. Importantly, in our model we have taken into account the compartmentalization of the enzymes (cytosolic and mitochondrial). The diffusion among compartments is modelled as a stochastic process and rather than following the exact position of the molecules inside the cell we consider their localization into cytosol and/or mitochondrion.

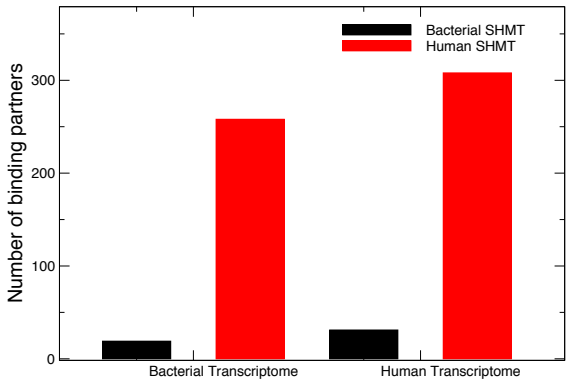


FIG. 1: RNA intearctome of human and bacterial SHMT predicted by *cat*RAPID. Bar plot of the predicted RNA interactome by *cat*RAPID relative to the human and bacterial SHMT isoforms. The predicted number of interacting RNAs of the human and bacterial SHMT are represented in red and black, respectively. The two pairs of bars are related to the calculation between the SHMTs and the human transcriptome (right) or the bacterial transcriptome (left).

A. Protein-RNA interactions

In order to understand the evolutionary role of the RNAs in the one carbon metabolism network of eukaryotic cells, we have compared the ability of the human and bacterial SHMT isoforms to bind RNAs. In Fig 1 using *cat*RAPID [28] we have predicted the number of possible RNA binding partners of the human and bacterial SHMT. The plot shows two types of interactions, one relative to the calculations using the bacterial RNAs and one relative to human RNAs. The human isoform clearly shows a much higher affinity to bind RNAs, this can be interpreted as an evolutionary drift of the enzyme towards the possibility to be be regulated allosterically by RNAs.

To describe the metabolic effects caused by the recently reported RNA-mediated regulation of the SHMT1 activity [16] [18], we simulated the binding of SHMT2 mRNA to the cytosolic isozyme, taking into account the competition between RNA and folates for SHMT1 [18], Fig. 2.

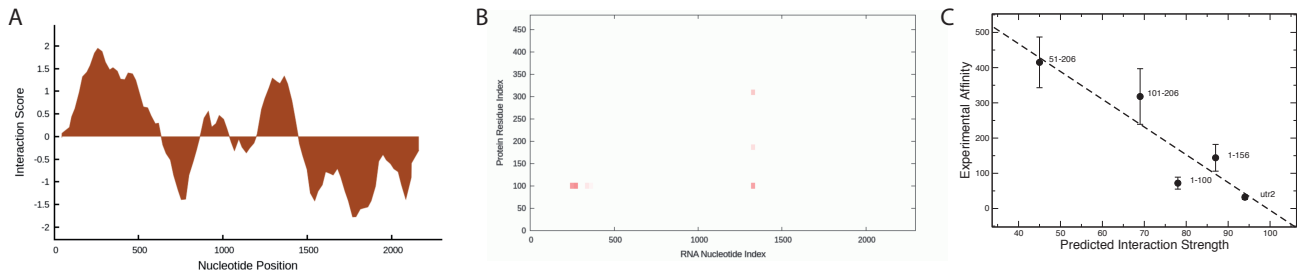


FIG. 2: A) *cat*RAPID prediction of SHMT1 and SHMT2 mRNA interaction. B) The N-term of SHMT1, also involved in folate binding, contacts SHMT2 mRNA in the 5' UTR (UTR2). C) Prediction of SHMT1 interaction with different UTR2 fragments indicate a large variability of experimental affinities.The N-term of SHMT1, also involved in folate binding, contacts SHMT2 mRNA in the 5' UTR (UTR2).

Using *cat*RAPID [28], we confirmed that SHMT1 (Uniport P34896) binds to the 5ÚTR of SHMT2 (Fig. 2 A) through a non-canonical RBD that overlaps with the folates binding cleft (N-term of the protein; Fig 2 A,B, C). Our results indicate that RNA and folate binding events are in competition. Relying on *cat*RAPID calculations [29], we could estimate the physical constants involved in the protein-RNA binding reaction (Fig. 2 C), which are input of our dynamical model.

We also have considered the SHMT1-mediated post-transcriptional regulation of SHMT2 translation. Indeed, as previously reported in [18] and herein confirmed in the A549 cell line Fig. (4 D), the cytosolic enzyme is able to inhibit the SHMT2 translation when highly expressed.

Worthy of note, the SHMT2 mRNA plays a double role in the system: the coding region is translated into the mitochondrial enzyme, while the 5-untranslated region (herein referred as UTR2) binds to SHMT1 and modulates its activity. The double function of the SHMT2 mRNA leads the system to dynamically change the number of amino acid molecules as a function of the SHMT2 mRNA levels (which positively correlate with the mitochondrial enzyme levels).

B. Modeling Reactions

In our model described in Fig. 3 the dynamic interconversion of serine and glycine, as well as of THF and Me-THF, is guided only by SHMT1 and SHMT2, according to biochemical parameters experimentally measured [16]. In the initial condition, the cytosolic enzyme catalyses the serine cleavage as well as the serine synthesis reaction, whereas the mitochondrial counterpart can only drive serine cleavage. Such asymmetry has been reported in cellular models [15] and, together with the exchange of metabolites between cytosol and mitochondrion, confers a complex dynamic to the system.

Given the affinity of SHMT1 for multiple RNA moieties [18], we assumed that SHMT1 can bind to a wide range of RNA molecules in the cell. We built a theoretical model to explore the metabolic effects of such phenomenon: accordingly, multiple interactions have been modelled considering the presence of a cytosolic pool of RNA molecules that compete with each other for binding to the RNA binding domain of SHMT1 again, as seen for SHMT2 mRNA, affecting the directionality of the reaction.

The system dynamic as well as most physical constants used in our model are governed by the chemical equations derived from [16] and [18]. We have integrated the RNA molecules in the system and introduced the following dynamic to describe the effect of this new component:

$$\begin{aligned}\frac{dR_i}{dt} &= (k_T - k_d R_i) \\ &- \left(k_i R_i E + k_i^- R_i[s\cdot E] + k_i R_i[g\cdot E]\right) \\ &+ \left(k_{Est}^-[R_i\cdot EST] + k_{Egm}^-[R_i\cdot EGM]\right) \\ &+ \left(k_g[R_i\cdot E] + k_g[R_i\cdot s\cdot E] + k_g[R_i\cdot g\cdot E]\right) \\ &+ \left(k_{Est}^+[R_i\cdot s\cdot E]Tf + k_{Egm}^+[R_i\cdot g\cdot E]Mt\right)\end{aligned} \tag{1}$$

Where we have identified the RNA molecules with R_i, SHMT1: E, serine: s, glycine: g, tetrahydrofolate: Tf, methylentetrahydrofolate: Mt, the catalytic complexes SHMT1 - Serine - Tetrahydrofolate : EST and SHMT1 - Glycine - Methylentetrahydrofolate: EGM. When represented inside squared brackets, the molecules are considered as a complex. The first line describes the synthesis and the degradation of the RNA, we consider simply a birth-dead process with no other components involved. The second line contains the terms representing the dissociation of the RNA molecule from the SHMT-substrates complexes. The third line contains terms describing the formation of the complexes with RNA molecules (Supplementary Materials and Methods).

Given the spectrum of binding affinities Fig. 2 A, we have considered the propensity of an RNA to interact with SHMT1 to follow a gaussian distribution centered on the average k_0 associated with the experimental value measured in [16][18], moreover it matches very well also

the predicted value from the *cat*RAPID algorithm. In this way the i-th RNA in the cytosol will bind to SHMT1 with an affinity k_i taken from the following Gaussian distribution:

$$P(k_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k_i - k_{SR})^2}{2\sigma^2}} \quad (2)$$

The standard deviation of the distribution σ has been inferred from the distribution of affinity values in Fig. 2 C, however in this model the value of σ does not affect the results (data not shown). The biochemical reactions set in the system are described in the Supplementary Information, with a detailed table of rates and constants. All the parameters are derived from previous experiments [16] where available, whereas the missing ones have been chosen arbitrarily to properly fit experimental conditions.

We also took into account the translation of the SHMT2 mRNA; such process occurs in the cytosol and is followed by the translocation of the enzyme into the mitochondrion. We consider the transport of the molecules from and to the other cell compartment as a simple stochastic process, indeed, discerning between active and passive transport systems and looking at their effects on our biochemical model is out of the aim of this paper.

The stochastic reactions involved in the translation of the SHMT2 enzyme are the following:

$$R_0 \rightarrow E_2^* + R_0 \quad (3)$$
$$E_2^* \rightarrow E_2 \quad (4)$$
$$E_2 \rightarrow 0 \quad (5)$$

The first reaction describes the translation of the mRNA (R_0). Importantly, we assume the ribosomes at a constant concentration, such to not being necessary to directly include their direct contribution into the model. It is important to notice that the R_0 is not degraded right after its translation, implying that the concentration of R_0 is not being affected by the translation. The E_2^* refers to the SHMT2 enzyme when still in the cytosol, then it is transported into the mitochondrion (second reaction) becoming the effective mitochondrial isoform here called E_2. The SHMT2 molecules is degraded with a rate constant. The equations that regulate the mean behaviour of all the system, as mentioned already above, are reported in the SI.

As schematically described in Fig.3 the peculiarity of this biochemical network is that it evolves in parallel between different cellular compartment. Indeed, SHMT1 and SHMT2 are present in the cytosol and in the mitochondrion, respectively. As reported in [15] and schematically illustrated in 3 A, SHMT1 can catalyse both serine synthesis and cleavage, whereas the mitochondrial counterpart can drive mainly the serine cleavage reaction. This asymmetry, together with the diffusion of amino-acids and metabolites between cytosol and mitochondrion, gives to the system a complex behaviour. As already introduced above, we quantify here the biochemical effects of the RNAs on the dynamic of the system. We aim to study the effects of the SHMT2 mRNA and the pool of RNAs present in the cytosol, understanding how such RNA species can drive the system to fine tune the amount of amino-acids available in the cell.

We simulated the system using the stochastic Gillespie algorithm [30]. Such well-established simulation method has the power to be an exact method for simulating networks of stochastic processes and it fits well the configuration of this system. This algorithm considers that all the chemical reactions depend on the concentration of the components in order to react .

Previous *in vitro* experiments have revealed the biochemical features and the dynamic of the interaction between SHMT isozymes and their substrates (either metabolites or RNA) [16], however their regulation in a biological context is puzzling. Most of all, it is not clear why in some cell types the SHMT1 reaction is reversible, while in others it is only directed towards serine synthesis [15].

C. Simulation Results

In the biochemical network simulated here we consider the reactions occurring inside the cell, neglecting the transport of metabolites from the environment. To draw quite general conclusions, we have chosen the units of the amino acids to be considered as the percentage of free serine and glycine in the system with respect to their total concentration. Indeed, we use relative measurement also because the exact volume of the cell and its inner organelles it is not easily available, thus limiting the possibility to quantify exactly the value of the amino acids concentration.

To describe how the system funnels the reaction fluxes and regulates the relative amount of amino acids, we chose to set the initial amount of serine and glycine equal to the quantity generally present in most cell mediums (i.e. serine 0.4 mM, glycine 0.4 mM) [31]. Stochastic simulations show that, in absence of RNA molecules, the amount of serine and glycine fluctuates around a mean value with negative correlation (Fig. 3 B). Indeed, given the flux balance of chemical reactions, serine and glycine interchange each other in a stochastic manner, chasing an equilibrium that is impossible to keep due to the intrinsic noise of the system. Figure 3 panel B shows a scatter plot describing serine and glycine concentration behaviour; of note, Pearson correlation computed between the temporal profiles of serine and glycine is $p \simeq -0.8$, this occurs because the system is continuously interchanging serine and glycine without any further regulation. With this setting, the biochemical system can explore a very limited space of amino acids concentration.

Next step was to evaluate the effect of the SHMT1-$\mathrm{RNA_{SHMT2}}$ complex on the dynamics of the system. It is important to point out that the behaviour of glycine

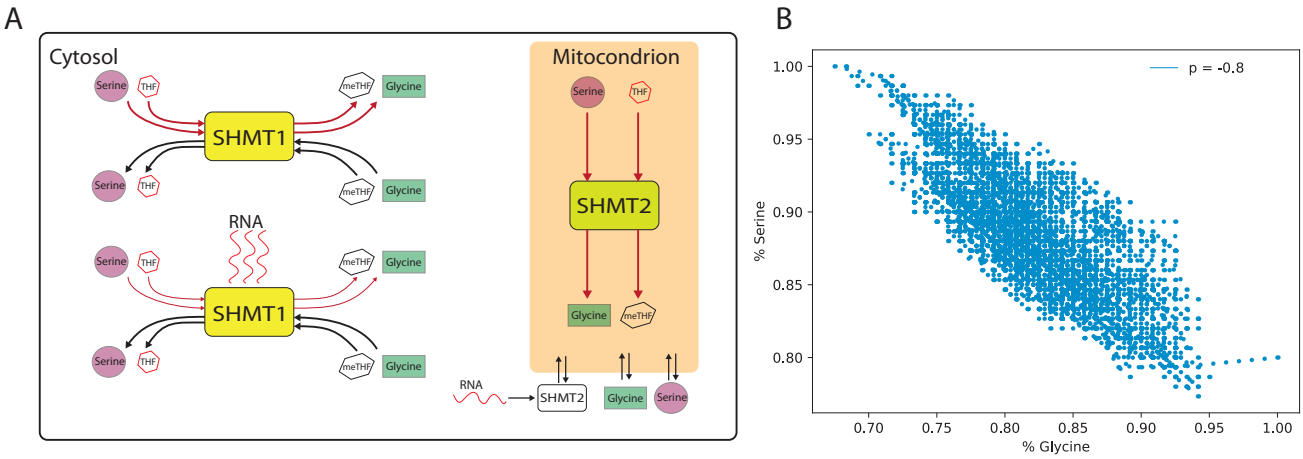


FIG. 3: (A) Scheme representing the reactions modelled in our theoretical approach. In marked red are highlighted the most favorable reactions driven by these enzymes. The effect of RNA binding on SHMT1 is to reduce the Serine to Glycine reaction in the cytosol compartment. (B) Scatter plot of Serine and Glycine relative free amount in solution. The plot comes from the stochastic simulation of the system with a fix amount of RNAs. Important to notice is that the stochastic oscillations around the mean values of the concentrations are always distributed with a negative correlation. The region of the phase space that the system explore stochastically during a simulation is finite.

and serine reported herein refers to the amount of free molecules in the system, hence there is a fraction of amino acids that at the steady state is complexed with the other species of the system (e.g. SHMT isozymes), and it is not reported in our simulations. For the latter reason, the sum of glycine and serine appears not to be constant along the simulations.

As shown in Fig 4 A, after an initial lag phase, the levels of serine and glycine vary as a function of RNA_{SHMT2} molecules. Serine displays a bi-phasic behaviour, experiencing first a rise and then a decrease before levelling off to a constant value at high RNA_{SHMT2}, while glycine level undergoes a decrease and reaches a constant amount, which does not vary further upon varying SHMT2 mRNA. This result clearly indicates that the serine/glycine levels indeed dynamically change, responding to the presence of RNA_{SHMT2} molecules, in agreement with our initial hypothesis.

Considering that SHMT1 will have multiple target RNAs in the cytoplasm, we simulated the presence of other RNAs interacting with the enzyme; such assumption implies that the pool of RNAs will compete with folates substrates as well, contributing to inhibit the serine cleavage reaction while slightly affecting serine synthesis. We investigated how the presence of these RNA moieties (named RNA_{pool} in Fig. 4B) affects amino acid levels. Fig. 4B shows how serine and glycine amounts change as a function of RNA_{SHMT2} molecules in the system, different coloured curves are then referred to different values of RNA_{pool} molecules considered in the simulation. As before, we quantified the amount of amino acids presents in the system as a percentage of their maximum availability in the system.

We identified two regimes: relative to low (dark and light blue curves in Fig. 4B) and high (red and green curves in Fig 4B) amount of RNA_{pool} molecules. Each point of the curves represents the mean value of the steady state amino acids concentration over different simulations. In the first couple of curves (i.e. light and dark blue), we can see again the very interesting and non-monotonic behaviour of the serine and glycine concentration, also seen in the absence of RNA_{pool} (4A). When the amount of RNA_{pool} is high (red and green curves), a more symmetric system with similar starting levels of serine and glycine is observed. In Fig 4 C it is plotted the mean percentage of serine and glycine for a given amount of RNA_{SHMT2} and RNA_{pool}. Being a different representation of the data already plotted in Fig 4 B, such curves aim to drive the attention on a different aspect of the system. Indeed, they highlight the fact that varying the RNA molecules in the system, it is possible to explore and sample a wide region of the phase space of glycine and serine concentrations. Worthy of note, moving through this space in a controlled manner implies that the cell could be able to regulate the relative amino acid concentrations simply by modulating the amount of RNAs bound to SHMT1. For instance, for low amounts of RNA_{pool} is possible to have an higher percentage of free serine combined with a lower concentration of glycine, on the other hand, for high values of RNA_{pool} molecules, the system explores a lower serine concentration combined with a higher amount of glycine.

Discussion

The integration of an alpha-proteobacterial endosymbion in the cell [32] allowed the development of a com-

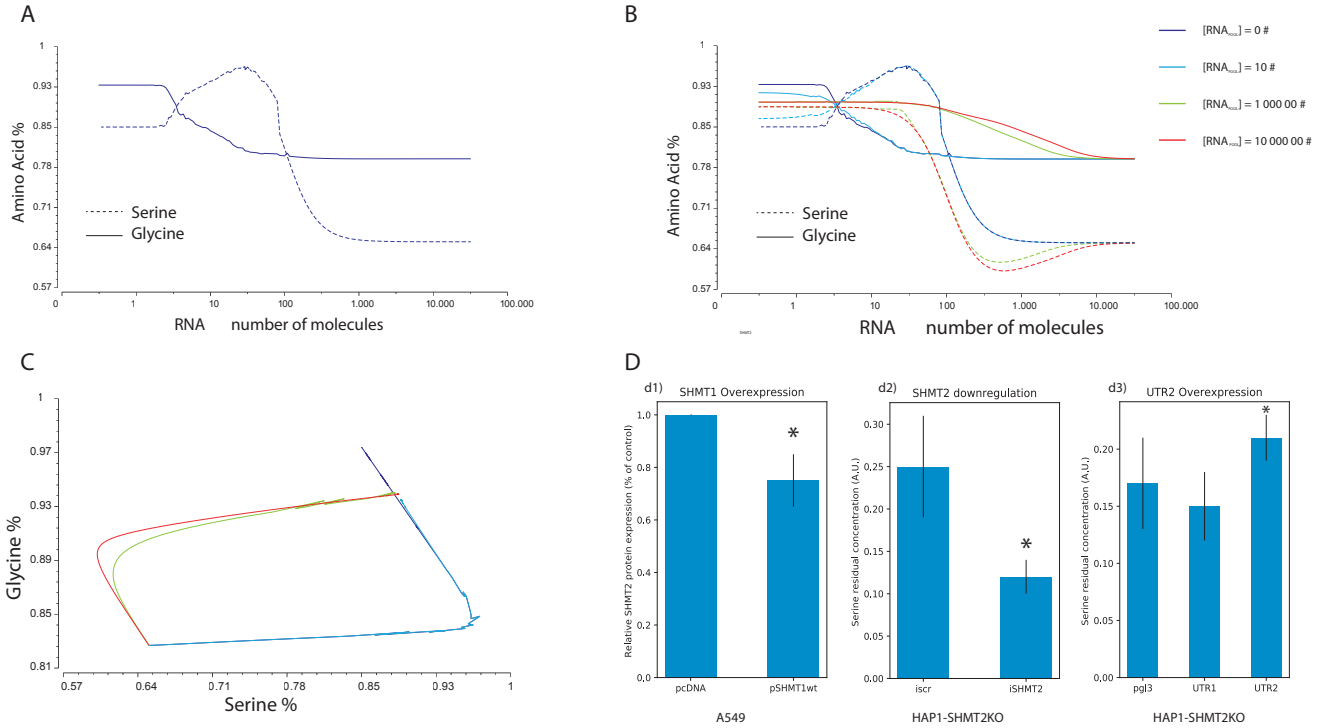


FIG. 4: Behaviour of the final steady state of the system for different amount of RNA molecules inserted in the simulations. Panel A describes the behaviour of serine and glycine as a function of the SHMT2 mRNA, considering the RNA_{pool} equal to 0. Panel B shows the same dynamic but considering different numbers of RNA pool molecules present in the system. Panel C shows the phase space of glycine serine for the final steady states. Interesting here it is possible to appreciate how different class of RNAs can drive the system to explore different ranges of the serine glycine concentrations. D) Experimental evidence of SHMT1 crosstalk with SHMT2. d1) Densitometric analysis of western blot data obtained in the A549 cell line transfected with an empty vector (pcDNA) or with a vector overexpressing SHMT1 (pSHMT1wt). d2) Data taken from [18]. SHMT1 activity assessed evaluating serine residual concentration in the growth medium of the HAP1SHMT2KO line, which expresses SHMT2 mRNA but not its protein product. d2) Serine has been measured 72 h after transfection with RNAis to knockdown SHMT2 mRNA (iscr = scramble RNA and iSHMT2 = shmt2 interference) or d3) with plasmids to overexpress UTR2 (pgl3 = empty vector, UTR1 = 5' UTR of SHMT1 (UTR1)-containing plasmid, UTR2 = UTR2-containing plasmid). Statistical analyses are performed on three independent experiments, *P¡ 0.05.

plex compartmentalized metabolic network that requires the interplay of enzymes present in both the mitochondrion and the cytoplasm [33]. Here we used state-of-the-art theoretical approaches to demonstrate that RNA molecules link the activities of human cytosolic hydroxymethyltransferase (SHMT1) and its mitochondrial counterpart (SHMT2). The dynamic modulation of serine and glycine orchestrated by SHMT enzymes is fundamental for cell growth and plays an important role in cancer [34]. Intriguingly, while bacterial SHMT does not have RNA binding ability (Fig. 1), human cytosolic SHMT1 has evolved to interact with RNA [18]. The biochemical effects of RNA interactions with SHMT1 and their impact on SHMT2 activity represent one of the missing link between cytosolic and mitochondrial SGOC metabolism, having important repercussions on cell proliferation. Suffice to say that mitochondrial SHMT2 and the immediately downstream mitochondrial enzyme 5,10-methylene-tetrahydrofolate dehydrogenase MTHFD2 are the most consistently overexpressed metabolic enzymes in cancer [35].

The data presented herein suggest that an RNA-mediate control of amino acid concentrations is operative inside the cell, exerted through the moonlighting activity of human cytosolic SHMT1. The cell can explore different range of amino acids concentration simply by tuning the expression of RNAs interacting with the cytosolic enzyme SHMT1, acting as a sensor of the metabolic status of the cell.

The observed behaviour allows us to interpret previous results obtained in cancer cell models . On the one side, RNA is able to allosterically modulate SHMT1 activity, by binding to the 5' region of the RNA_{SHMT2} and selectively inhibiting serine cleavage. On the other side, RNA_{SHMT2} is translated into the mitochondrial enzyme which controls serine to glycine conversion in the or-

ganelle. The combined action of the sensor (SHMT1) and the multifunctional RNA effector (RNA_{SHMT2}) allows the system to access states of serine and glycine concentrations impossible to explore otherwise.

For low SHMT2 mRNA levels, serine and glycine amounts are similar, with a slightly higher amount of glycine due to the fact that the system has an intrinsic asymmetry (Fig. 4 A). As soon as SHMT2 mRNA levels increase, serine raises to reach a maximum, whereas glycine levels lower. Remarkably, this dynamics accurately describes the modulation of serine levels experimentally determined in leukaemia HAP cells containing a frameshift mutation in the RNA_{SHMT2} and thus blocking synthesis of SHMT2 protein [18]. By lowering SHMT2 mRNA levels, a decrease in serine extracellular levels has been measured, in agreement with predictions (Fig. 4D). The overexpression of the 5' UTR, which corresponds to the SHMT2 mRNA allosteric region alone, increases the levels of extracellular serine (Fig. 4D); under these conditions, SHMT1 serine cleavage is indeed inhibited and thus extracellular serine is consumed less, well in agreement with the results of the simulation. In a wild type background, where the translation of RNA_{SHMT2} in the functional mitochondrial protein acts co-ordinately with the inhibitory effect of the same mRNA, the accumulation of serine (Fig. 4A) is the result of an insufficient amount of mitochondrial enzyme that transforms the substrate into glycine. Thus, in this range of RNA values, the main metabolic effect results to be the inhibition of the SHMT1 mediated serine cleavage. As the amount of RNA raises, the RNA binding site of SHMT1 becomes saturated, and more SHMT2 mRNA molecules become available to be translated into the mitochondrial protein, which is catalytically active in the mitochondria and thus starts to affect serine levels; indeed, as expected, serine starts to be rapidly consumed, as shown by the drop right after the peak (Fig. 4 A). Importantly, glycine does not decrease further, because the saturation effect driven by the RNA mediated inhibition of SHMT1 is compensated by a higher concentration of mitochondrial enzyme, which converts serine into glycine. Finally, when the levels of SHMT2 mRNA are extremely high, the translation of mitochondrial enzyme reaches a plateau, explaining why the serine curve gets stabilized.

In the presence of a more complex population of allosteric RNA effectors of SHMT1 (Fig. 4B), the intermediate serine peak disappears because SHMT1 is synergistically saturated by RNA_{SHMT2} and pool of RNAs. Interestingly, under these conditions, the amino acids concentration remains constant for a wide range of RNA_{SHMT2}, until the point in which both serine and glycine start to drop. The increase of RNA_{SHMT2} causes the increase of SHMT2 in the mitochondrion, meaning that the SHMT2 reaction becomes favoured. The pronounced drop of serine can be explained by the fact that the reaction driven by SHMT2 consumes serine more quickly; such effect is hardly compensated by the cytosolic reaction, consequently a significative amount of serine gets trapped into RNA-protein complexes, slowing down the transformation of serine into glycine. This explains why the drop in serine concentration is not compensated by a raise of glycine levels. In addition, in the presence of a pool of RNAs, these molecules behave as competitors of RNA_{SHMT2} for the binding to SHMT1; a left shift of the curve is observed in the region where serine levels drop with respect to the behaviour recorded in the absence of the RNA_{pool} (Fig. 4 B). Such phenomenon can readily be explained by envisaging that the number of free RNA_{SHMT2} molecules able to yield active SHMT2 enzyme is higher under these conditions, allowing serine to be more efficiently consumed. In the final part of the curve, a critical concentration of RNA_{SHMT2} is reached, where the expression of the mitochondrial protein is saturated by the finite number of ribosomes set in the system. Indeed, the mitochondrial reactions are saturated as well, whereas in the cytosol the increasing levels of RNA_{SHMT2} will increase inevitably the amount of amino acids trapped into complexes. Thus, glycine starts to compete with serine complex with SHMT1 and RNA, letting the amount of free serine to increase until it reaches the equilibrium.

In summary, tuning amino-acids concentrations is crucial for the cell behaviour and cell fate, and we showed that RNA can act a key regulator in this scenario. These predictions and observations offer an unprecedented and perspective view of how RNA dynamics might directly control metabolite dynamics through key protein sensors. Intriguingly, many of the proteins classified as non-conventional RNA binding proteins (NC-RBP) not only are indeed metabolic enzymes [8], [36] [37], [38] , but also have isoforms localized in different subcellular compartments, such as SHMT in the present paper, which might undergo a differential allosteric regulation by RNA molecules depending on their location. Our results open the possibility of designing new treatments employing RNA molecules for diseases where the dysregulation of serine and glycine plays an important role, including cancer.

Material and methods

We have used the COPASI software [30] in order to implement the Gillespie algorithm for the stochastic simulations. Protein-RNA binding predictions have been computed using the *cat*RAPID software [28]. The experiments have been performed using Western Blot Analysis analysis and serine consumption assays.. Further information are provided in the Supporting Information.

Acknowledgment

We acknowledge Alexandros Armaos for fruitful discussion and for the *cat*RAPID guidelines. MM and GGT thank the IH2020project IN-FORE 825080 for funding

this research. GG is supported by the Barts Charity Grant. Funding from AIRC under IG2019-ID23125 project– PI FC and from Sapienza University of Rome (Grants n. RG11816430AF48E1, RM11916B46D48441, RP11715C644A5CCE, GA116154C8A94E3D to F.C.) are gratefully acknowledged.

Author contributions

MM and GG conceived the theoretical approach. MM developed the metabolic model together with GG. GGT implemented the catRAPID analysis. MM, GG and GGT wrote the paper with the help of FC. AP, SR and GioG designed and performed the experiments. FC designed and supervised the experiments. MM and GG are to be considered co-first authors. To whom correspondence may be addressed. E-mail: MM michele.monti@iit.it and GGT gian.tartaglia@iit.it for the computational parts; FC francesca.cutruzzola@uniroma1.it for the experimental part.

[1] E. J. Deeds and E. I. Shakhnovich, Advances in Enzymology and Related Areas of Molecular Biology **75**, 133 (2007), ISSN 0065258X.

[2] A. Perino, A. Ghigo, J. D. Scott, and E. Hirsch, Circulation Research **111**, 482 (2012), ISSN 15244571.

[3] M. W. Hentze, A. Castello, T. Schwarzl, and T. Preiss, Nature reviews. Molecular cell biology **19**, 327 (2018), ISSN 1471-0080, URL http://www.nature.com/articles/nrm.2017.130http://www.ncbi.nlm.nih.gov/pubmed/29339797.

[4] R. W. Yao, Y. Wang, and L. L. Chen, *Cellular functions of long noncoding RNAs* (2019).

[5] K. Sampath and A. Ephrussi, Development (Cambridge) **143**, 1234 (2016), ISSN 14779129.

[6] F. Hubé and C. Francastel, *Coding and non-coding RNAs, the frontier has never been so blurred* (2018).

[7] A. Castello, B. Fischer, K. Eichelbaum, R. Horos, B. M. Beckmann, C. Strein, N. E. Davey, D. T. Humphreys, T. Preiss, L. M. Steinmetz, et al., Cell **149**, 1393 (2012), ISSN 00928674, URL http://www.ncbi.nlm.nih.gov/pubmed/22658674http://linkinghub.elsevier.com/retrieve/pii/S0092867412005764.

[8] A. Castello, B. Fischer, C. K. Frese, R. Horos, A. M. Alleaume, S. Foehr, T. Curk, J. Krijgsveld, and M. W. Hentze, Molecular Cell **63**, 696 (2016), ISSN 10974164.

[9] C. He, S. Sidoli, R. Warneford-Thomson, D. C. Tatomer, J. E. Wilusz, B. A. Garcia, and R. Bonasio, Molecular Cell **64**, 416 (2016), ISSN 10974164.

[10] M. W. Hentze and T. Preiss, Trends in biochemical sciences **35**, 423 (2010), ISSN 0968-0004, URL https://linkinghub.elsevier.com/retrieve/pii/S0968000410000976http://www.ncbi.nlm.nih.gov/pubmed/20554447.

[11] J. Cieśla, Acta biochimica Polonica **53**, 11 (2006), ISSN 0001-527X, URL http://www.ncbi.nlm.nih.gov/pubmed/16410835.

[12] A. Castello, M. W. Hentze, and T. Preiss, Trends in Endocrinology and Metabolism **26**, 746 (2015), ISSN 18793061, URL http://dx.doi.org/10.1016/j.tem.2015.09.012.

[13] C. J. Jeffery, Biochemical Society Transactions **42**, 1679 (2014), ISSN 14708752.

[14] G. S. Ducker and J. D. Rabinowitz, Cell metabolism **25**, 27 (2017), ISSN 1932-7420, URL http://www.ncbi.nlm.nih.gov/pubmed/27641100http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5353360.

[15] G. S. Ducker, L. Chen, R. J. Morscher, J. M. Ghergurovich, M. Esposito, X. Teng, Y. Kang, and J. D. Rabinowitz, Cell Metabolism **23**, 1140 (2016), ISSN 15504131, URL http://www.ncbi.nlm.nih.gov/pubmed/27211901http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4909566http://linkinghub.elsevier.com/retrieve/pii/S1550413116301681.

[16] A. Tramonti, C. Nardella, M. L. di Salvo, A. Barile, F. Cutruzzolà, and R. Contestabile, Biochemistry **57**, 6984 (2018), ISSN 0006-2960, URL http://www.ncbi.nlm.nih.gov/pubmed/30500180http://pubs.acs.org/doi/10.1021/acs.biochem.8b01074.

[17] A. S. Tibbetts and D. R. Appling, Annual Review of Nutrition **30**, 57 (2010), pMID: 20645850, https://doi.org/10.1146/annurev.nutr.012809.104810, URL https://doi.org/10.1146/annurev.nutr.012809.104810.

[18] G. Guiducci, A. Paone, A. Tramonti, G. Giardina, S. Rinaldo, A. Bouzidi, M. C. Magnifico, M. Marani, J. A. Menendez, A. Fatica, et al., Nucleic Acids Research **47**, 4240 (2019), ISSN 0305-1048, https://academic.oup.com/nar/article-pdf/47/8/4240/28534948/gkz129.pdf, URL https://doi.org/10.1093/nar/gkz129.

[19] G. Giardina, P. Brunotti, A. Fiascarelli, A. Cicalini, M. G. Costa, A. M. Buckle, M. L. Di Salvo, A. Giorgi, M. Marani, A. Paone, et al., FEBS Journal **282**, 1225 (2015), ISSN 17424658, URL https://pubmed.ncbi.nlm.nih.gov/25619277/.

[20] D. M. Szebenyi, X. Liu, I. A. Kriksunov, P. J. Stover, and D. J. Thiel, Biochemistry **39**, 13313 (2000), ISSN 00062960, URL https://pubmed.ncbi.nlm.nih.gov/11063567/.

[21] X. Liu, B. Reig, I. M. Nasrallah, and P. J. Stover, Biochemistry **39**, 11523 (2000), ISSN 00062960.

[22] C.-H. Chang, J. D. Curtis, L. B. Maggi, B. Faubert, A. V. Villarino, D. O'Sullivan, S. C.-C. Huang, G. J. W. van der Windt, J. Blagih, J. Qiu, et al., Cell **153**, 1239 (2013), ISSN 1097-4172, URL http://linkinghub.elsevier.com/retrieve/pii/S0092867413005825http://www.ncbi.nlm.nih.gov/pubmed/23746840http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3804311.

[23] F. P. Marchese, E. Grossi, O. Marín-Béjar, S. K.

Bharti, I. Raimondi, J. González, D. J. Martínez-Herrera, A. Athie, A. Amadoz, R. M. Brosh, et al., Molecular Cell **63**, 397 (2016), ISSN 10974164.

[24] M. Yang and K. H. Vousden, *Serine and one-carbon metabolism in cancer* (2016).

[25] K. R. Mattaini, M. R. Sullivan, and M. G. Vander Heiden, *The importance of serine metabolism in cancer* (2016).

[26] I. Amelio, F. Cutruzzolá, A. Antonov, M. Agostini, and G. Melino, Trends in Biochemical Sciences **39**, 191 (2014), ISSN 09680004.

[27] A. Paone, M. Marani, A. Fiascarelli, S. Rinaldo, G. Giardina, R. Contestabile, A. Paiardini, and F. Cutruzzolà, Cell Death and Disease **5**, 1 (2014), ISSN 20414889.

[28] K. P. M. D. R. S. B. B. T. G. Cirillo D., Agostini F., RNA **19 2**, 129 (????).

[29] B. B. T. G. Agostini F., Cirillo D., Nucleic Acid Research (2013).

[30] D. T. Gillespie, The Journal of Physical Chemistry **81**, 2340 (1977), https://doi.org/10.1021/j100540a008, URL https://doi.org/10.1021/j100540a008.

[31] A. Salazar, M. Keusgen, and J. Von Hagen, *Amino acids in the cultivation of mammalian cells* (2016), URL /pmc/articles/PMC4833841/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4833841/.

[32] E. V. Koonin, Genome biology **11**, 209 (2010), ISSN 1474-7596, URL https://europepmc.org/articles/PMC2898073.

[33] A. A. Pittis and T. Gabaldón, Nature **531**, 101 (2016), ISSN 1476-4687, URL https://doi.org/10.1038/nature16941.

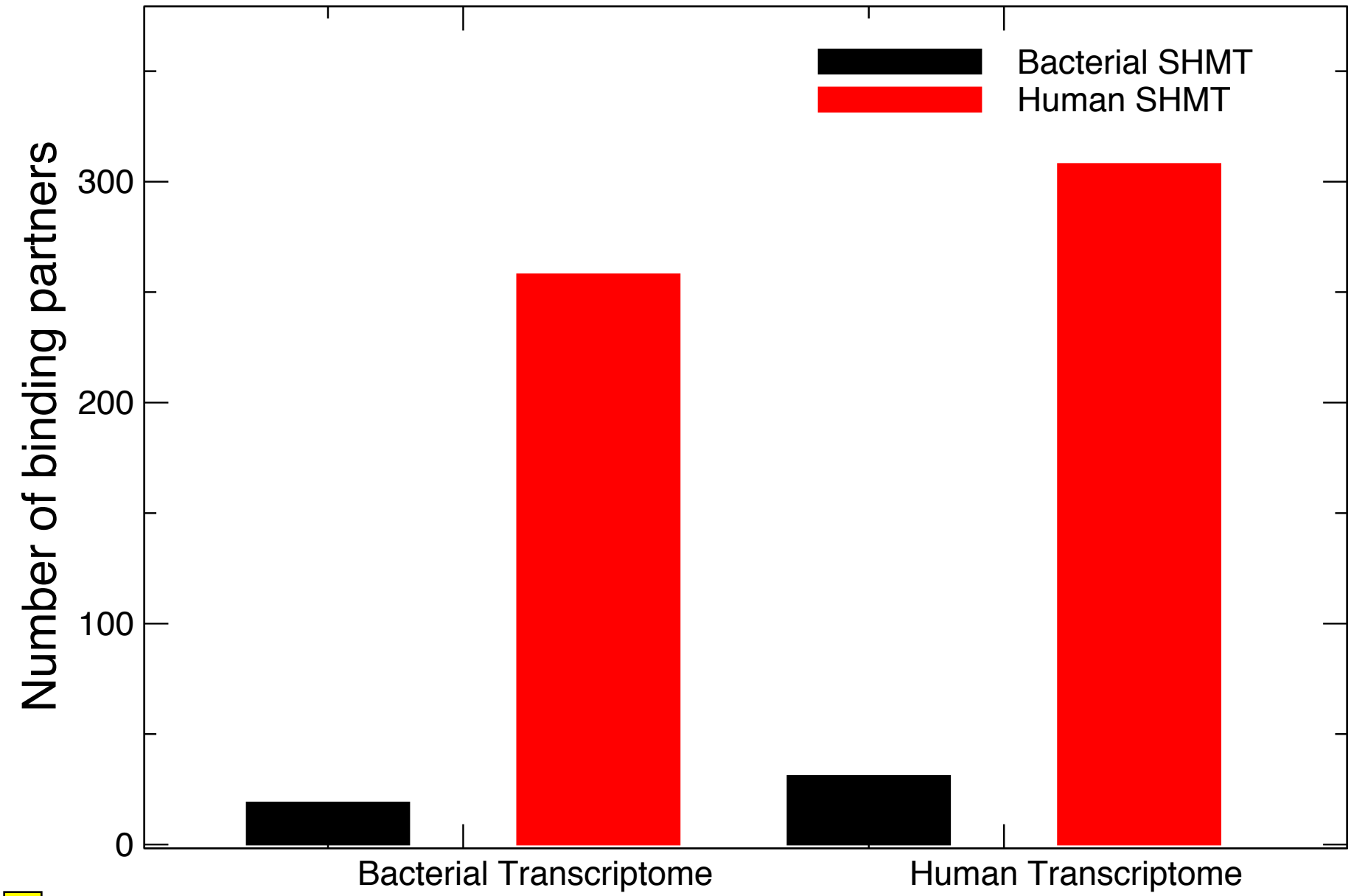
[34] T. Clarke, Journal of clinical pathology **44**, 87 (1991), ISSN 0021-9746, URL https://europepmc.org/articles/PMC497032.

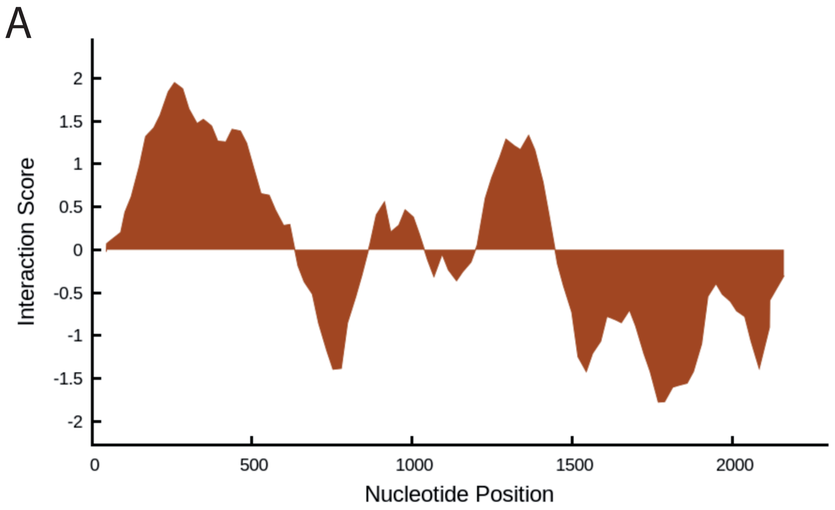
[35] G. S. Ducker, J. M. Ghergurovich, N. Mainolfi, V. Suri, S. K. Jeong, S. Hsin-Jung Li, A. Friedman, M. G. Manfredi, Z. Gitai, H. Kim, et al., Proceedings of the National Academy of Sciences **114**, 11404 (2017), ISSN 0027-8424, https://www.pnas.org/content/114/43/11404.full.pdf, URL https://www.pnas.org/content/114/43/11404.

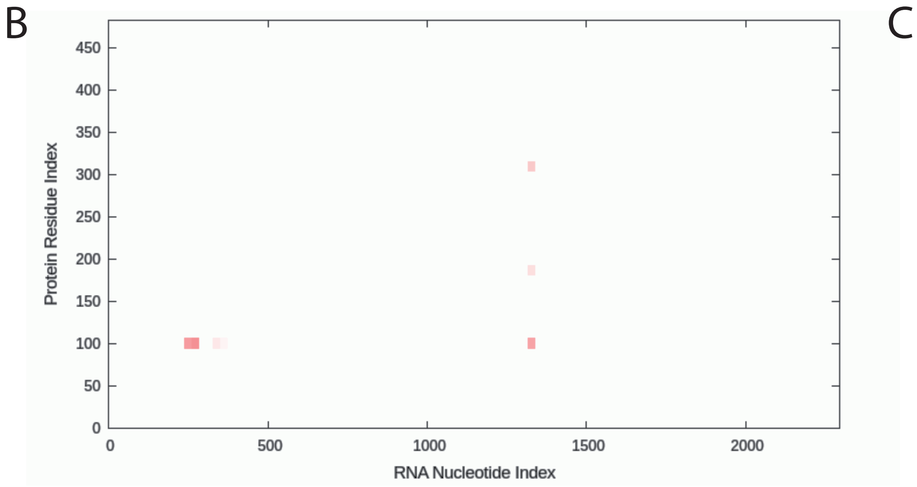
[36] X. G. S. K. Yuan Lv, Muqddas Tariq and M. A. Esteban, J Mol Cell Biol **11**, 813 (2019).

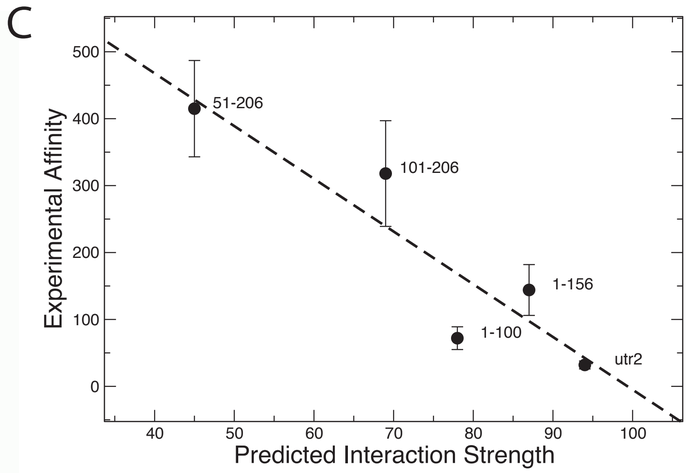
[37] Y. Lv, M. Tariq, X. Guo, S. Kanwal, and M. A. Esteban, Journal of molecular cell biology **11**, 813—815 (2019), ISSN 1674-2788, URL https://europepmc.org/articles/PMC6884707.

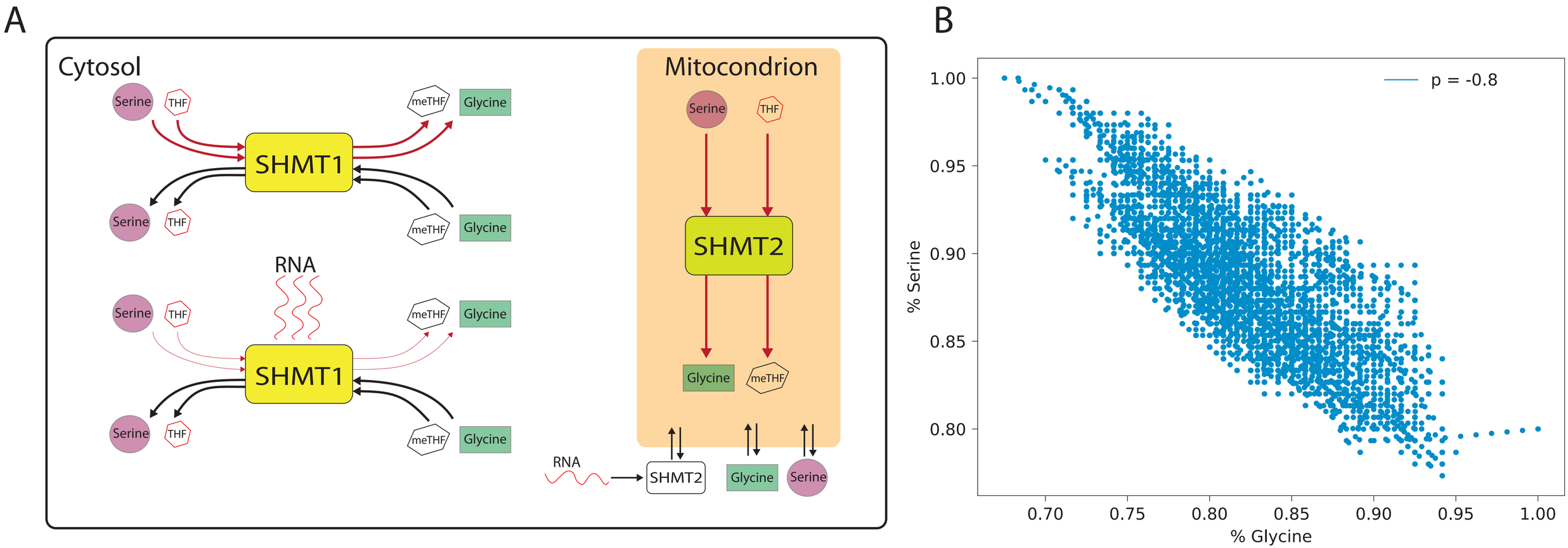
[38] I. Huppertz, J. I. Perez-Perri, P. Mantas, T. Sekaran, T. Schwarzl, L. Dimitrova-Paternoga, J. Hennig, P. A. Neveu, and M. W. Hentze, bioRxiv (2020), https://www.biorxiv.org/content/early/2020/10/14/2020.10.14.3 URL https://www.biorxiv.org/content/early/2020/10/14/2020.10.14.337444.

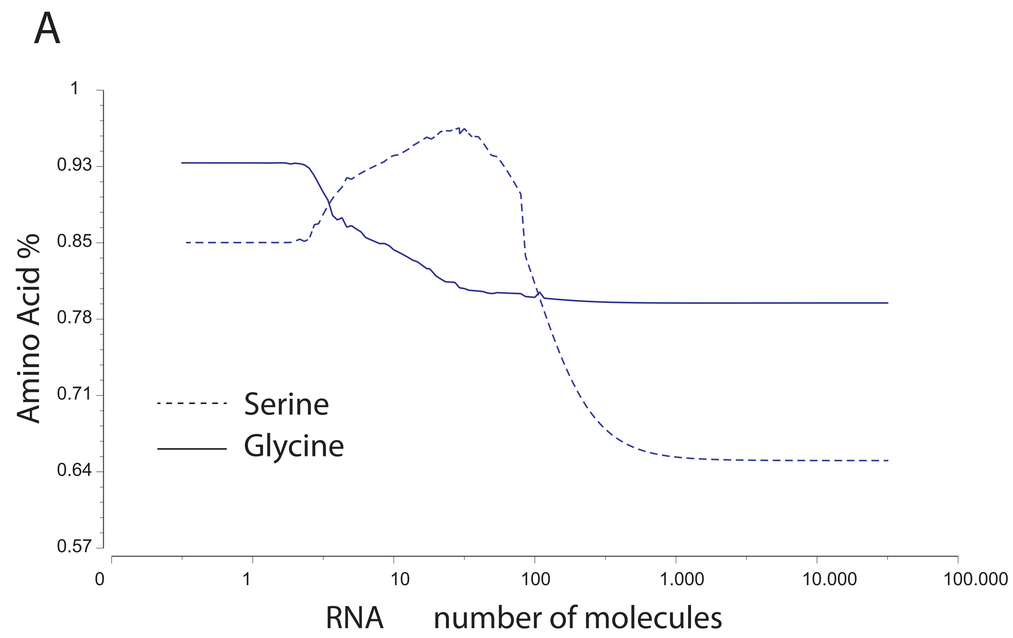

Bacterial SHMT
Human SHMT
Number of binding partners
300
200
100
0
Bacterial Transcriptome
Human Transcriptome

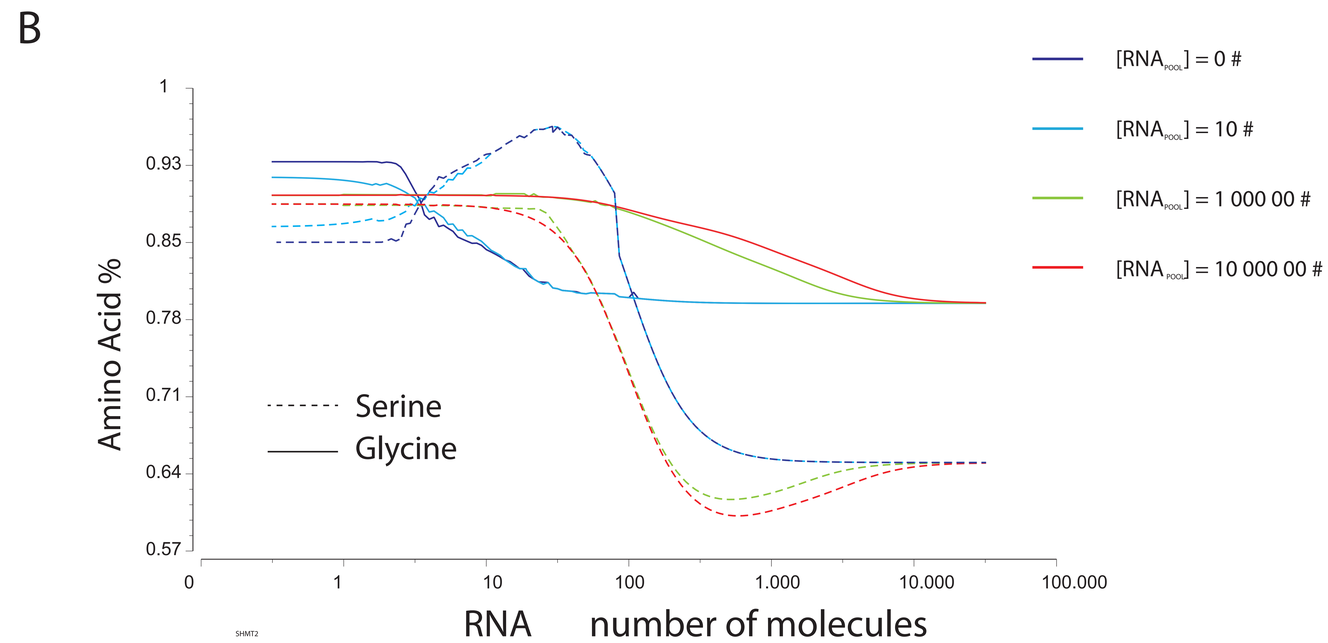

A
Interaction Score
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
500
1000
1500
2000
Nucleotide Position

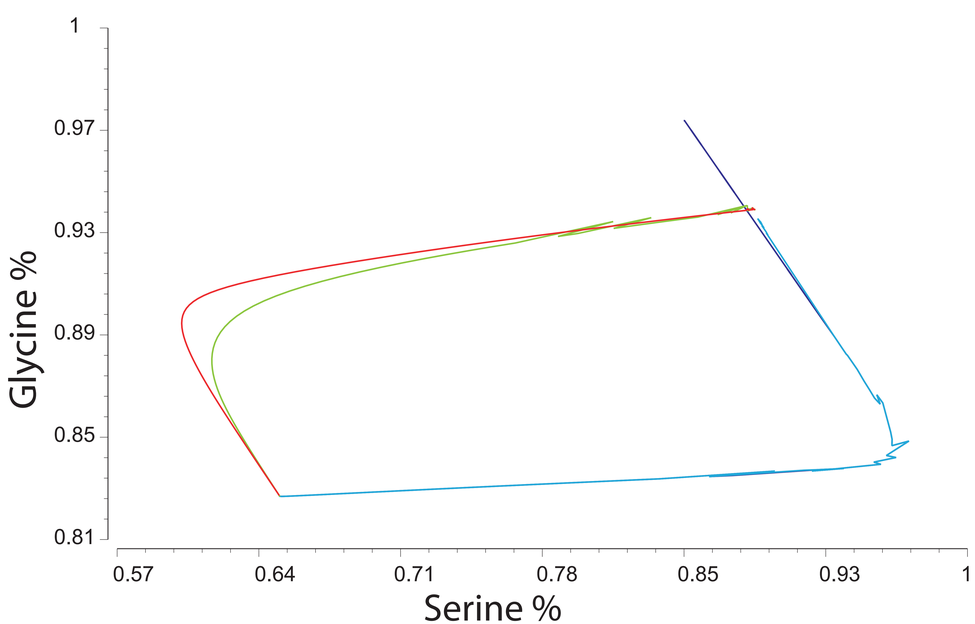

B
Protein Residue Index
450
400
350
300
250
200
150
100
50
0
0
500
1000
1500
2000
RNA Nucleotide Index

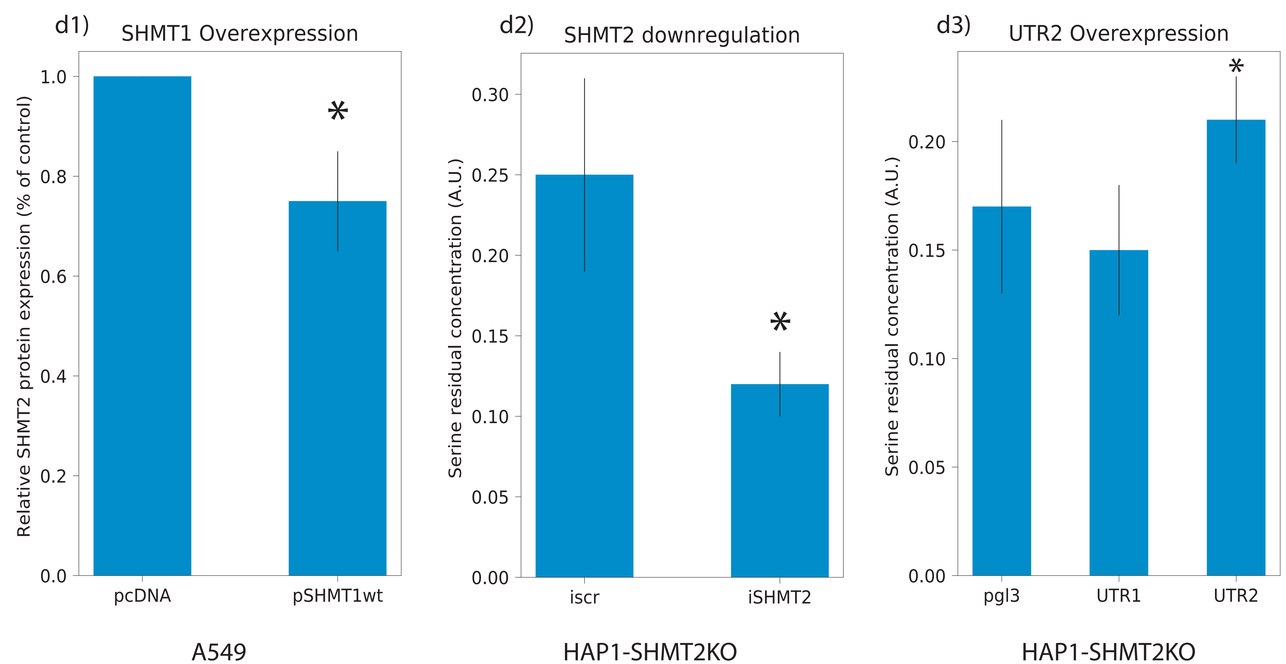

C
Experimental Affinity
500
400
300
200
100
0
51-206
101-206
1-156
1-100
utr2
40
50
60
70
80
90
100
Predicted Interaction Strength


A
Cytosol
Serine
THF
meTHF
Glycine
SHMT1
Serine
THF
meTHF
Glycine
RNA
Serine
THF
meTHF
Glycine
SHMT1
Serine
THF
meTHF
Glycine
Mitocondrion
Serine
THF
SHMT2
Glycine
meTHF
RNA
SHMT2
Glycine
Serine
B
p = -0.8
% Serine
% Glycine
1.00
0.95
0.90
0.85
0.80
0.70
0.75
0.80
0.85
0.90
0.95
1.00


A
Amino Acid %
1
0.93
0.85
0.78
0.71
0.64
0.57
Serine
Glycine
0
1
10
100
1.000
10.000
100.000
RNA number of molecules


B
[RNA POOL] = 0 #
[RNA POOL] = 10 #
[RNA POOL] = 1 000 00 #
[RNA POOL] = 10 000 00 #
Amino Acid %
1
0.93
0.85
0.78
0.71
0.64
0.57
Serine
Glycine
0
1
10
100
1.000
10.000
100.000
RNA number of molecules


C
Glycine %
1
0.97
0.93
0.89
0.85
0.81
0.57
0.64
0.71
0.78
0.85
0.93
1
Serine %


D
d1)
SHMT1 Overexpression
Relative SHMT2 protein expression (% of control)
1.0
0.8
0.6
0.4
0.2
0.0
*
pcDNA
pSHMT1wt
A549
d2)
SHMT2 downregulation
Serine residual concentration (A.U.)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
*
iscr
iSHMT2
HAP1-SHMT2KO
d3)
UTR2 Overexpression
Serine residual concentration (A.U.)
0.20
0.15
0.10
0.05
0.00
*
pgl3
UTR1
UTR2
HAP1-SHMT2KO