Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/2/10/e1601273/DC1

Supplementary Materials for

Serine one-carbon catabolism with formate overflow

Johannes Meiser, Sergey Tumanov, Oliver Maddocks, Christiaan Fred Labuschagne, Dimitris Athineos, Niels Van Den Broek, Gillian M. Mackay, Eyal Gottlieb, Karen Blyth, Karen Vousden, Jurre J. Kamphorst, Alexei Vazquez

Published 28 October 2016, *Sci. Adv.* **2**, e1601273 (2016)
DOI: 10.1126/sciadv.1601273

This PDF file includes:

Supplementary Materials

Supplementary Text

Mass isotopomer distribution (MID)

Given a metabolite of interest, the mass isotopomer distribution (MID) was calculated by first dividing the peak area corresponding to a given mass isotopologue for that metabolite by the total peak area associated with that metabolite. Second, MIDs were corrected for natural abundance of ^{13}C as described previously (*21*).

Exchange fluxes (serine, glycine)

For the estimation of exchange fluxes fresh, expend and reference medium samples were analysed with LC-MS and the peak area of selected metabolites were quantified. Expend medium represents the cell culture medium after a specified time interval Δt of cell culture. Fresh medium represents the cell culture medium in the absence of cells that expend the same time interval between addition to the cell culture wells and sample collection as the expend medium. Unless specified, the reference medium was DMEM plus 8 mM Lactate. This reference medium is only used to obtain an absolute quantification of metabolite concentrations. It is not used for the cell cultures. The complete medium was diluted 1:2 and used as a second reference.

Concentrations: The peak areas were converted to absolute concentrations using the reference medium samples for calibration. Concentrations (C) were related to peak areas (A) by the linear model $C=aA$. The conversion factor a was estimated from a least-squares fit to the data for the reference medium at 1 and ½ dilution.

Cell count and population growth rate: The population growth rate (μ) was estimated from the cell count at time zero (N_0) and at time Δt (N), using the equation $\mu=\ln(N/N_0)/\Delta t$.

Exchange flux model: The change of a metabolite concentration was modelled by the equation

$$\frac{dC}{dt} = \frac{Ne}{V} \tag{E1}$$

where e is the metabolite exchange flux per cell per unit of time and V is the cell culture volume. Integrating equation (E1) over time and assuming an exponential population growth with rate μ ($N=N_0\exp(\mu t)$) we obtain the working equation for exchange flux estimation

$$e = \frac{C_{\exp end} - C_{fresh}}{V(N - N_0)}\mu \tag{E2}$$

Statistics: For each independent experiment, three wells were used for the expend medium and C_{fresh} was estimated from the average of three fresh medium samples, resulting in 3 technical replicates for e.

Formate exchange fluxes

The exchange flux of formate was calculated as described above, with the exception of the quantification of absolute concentrations of formate in the cell culture medium.

Concentrations: Formate concentrations were estimated from peak areas using [^{13}C,^{2}H]-formate as a standard and the following model. The samples analysed contain M+0 and M+1 formate at unknown concentrations C_0 and C_1. The added standard contributes to the M+2 formate at it has a known concentration C_s. In addition there is a background of formate present extraction fluid with an unknown concentration C_B. We took into account that benzyl formate, the compound obtained after derivatisation, contains 7 carbon atoms (in addition to formate) that could each be ^{12}C and ^{13}C depending on the natural abundance of these isotopes. Putting together all these contributions we obtain the following equations for the expected peak areas of M+0, +1 and +2 formate

$$\begin{aligned} A_0 &= a\left(p_{70}C_0 + p_{80}C_B\right) \\ A_1 &= a\left(p_{71}C_0 + p_{70}C_1 + p_{81}C_B\right) \\ A_2 &= a\left(p_{72}C_0 + p_{71}C_1 + p_{82}C_B + p_{70}C_s\right) \end{aligned} \tag{F1}$$

where *a* is a conversion factor from areas to concentrations and

$$p_{mn} = \binom{n}{m} c^m (1-c)^{n-m} \tag{F2}$$

is the probability that a molecule *of n* carbon atoms contains *m* ^{13}C atoms given the frequency of ^{13}C in nature (*c*=0.011). Equations (F1) for peak areas can be converted to equations for isotopomer fractions, obtaining

$$\begin{aligned}
&(p_{70} - b_0 x_0) C_0 + (\quad - b_1 x_0) C_1 = (b_B x_0 - p_{80}) C_B + (b_s x_0) C_s \\
&(p_{71} - b_0 x_1) C_0 + (p_{70} - b_1 x_1) C_1 = (b_B x_1 - p_{81}) C_B + (b_s x_1 - p_{70} s_1) C_s \\
&(p_{72} - b_0 x_2) C_0 + (p_{71} - b_1 x_2) C_1 = (b_B x_2 - p_{82}) C_B + (b_s x_2 - p_{71} s_1 - p_{70} s_2) C_s
\end{aligned} \tag{F3}$$

where

$$x_n = \frac{A_n}{A_0 + A_1 + A_2}, \quad n = 1, 2, 3 \tag{F4}$$

is the fraction of M+n formate and

$$\begin{aligned}
b_0 &= p_{70} + p_{71} + p_{72} \\
b_1 &= p_{70} + p_{71} \\
b_B &= p_{80} + p_{81} + p_{82} \\
b_S &= p_{70}
\end{aligned}$$

For the analysis of cell culture medium expend samples, we used a standard of 20 nmol/sample. For the analysis of plasma formate, where formate is found in lower concentrations, we used standards in the range C_s=0.4-1 nmol/sample. The absolute amounts of nmol/sample were converted to mmol/l (mM) after dividing by the volume of the biological sample (40 μL). The background formate concentration was estimated from the analysis of water samples. In this case C_0=C_1=0 and we can estimate C_B from the solution of (F3). Estimates from independent runs gave values around C_s~20 μM.

Numerical solution: The values of C_0 and C_1 were obtained from the least-squared solution with nonnegative variables of the system of linear equations (F3).

Statistics: The M+0, +1 and +2 formate peak areas all come from the same sample. Therefore, we obtain an independent estimate of the M+0 and M+1 formate concentration for each sample analysed.

Thymidylate synthesis

The thymidylate synthesis rate of a proliferating cell is given by $f_{dTMP}=\mu b x_T$, where $\mu=\ln 2/T_D$ is the proliferating rate, T_D is the doubling time, b is the number of DNA bases in the cell genome and x_T is the fraction of base T in the genome. In general x_T~1/4. The doubling time of the investigated cells is T_D~24 hours. HCT116 cells are nearly diploid and IMR90 cells are diploid, with an estimated DNA content of b=10 fmol/cell [(*29*), Fig. 2E]. Based on these numbers we estimate f_{dTMP}~0.07 fmol/cell/h for HCT116 and IMR90 cells. This value is one order of magnitude lower than the estimated purine synthesis rates in these cells (HCT116: 1.5 fmol/cell/h, IMR90: 0.33 fmol/cell/h) and it is within the error of our measurements (generally at two decimal in units fmol/cell/h).

Serine synthesis

The rate of serine synthesis from glucose was estimated as described before (*20, 21, 30*). Specifically, cells were switched from regular culture medium to culture medium containing [$^{13}C_3$]-L-serine (*tracer*). We took into account that serine exhibits a steady-state isotopomer distribution in a time window around 8 hours after addition of the tracer. Serine can be produced from glucose (f_s), imported from the media (u_s) or produced from the reverse flux of SHMT (f_{gs}). The serine ^{13}C fractions satisfy the metabolic flux analysis (MFA) equations

$$\begin{aligned}
S_{i0}\left(u_s+f_s+f_{gs}\right)&=u_sS_{e0}+L_0G_0f_{gs}+f_s\\
S_{i1}\left(u_s+f_s+f_{gs}\right)&=u_sS_{e1}+L_1G_0f_{gs}\\
S_{i2}\left(u_s+f_s+f_{gs}\right)&=u_sS_{e2}+L_0G_2f_{gs}\\
S_{i3}\left(u_s+f_s+f_{gs}\right)&=u_sS_{e3}+L_1G_2f_{gs}
\end{aligned} \qquad \text{(S1)}$$

where the subscripts *i* and *e* stands for intracellular and extracellular serine isotopomer fractions, respectively. Since the ^{13}C fractions sum to 1, this is a system of three equations with three variables, f_{cs}, f_{gs} and L_0. A convenient way to solve this system of equations is to introduce the auxiliary variables

$$x=\frac{f_s}{u_s} \qquad \text{(S2)}$$

$$y=\frac{f_{gs}}{u_s} \qquad \text{(S3)}$$

$$z = \frac{L_0 f_{gs}}{u_s} \tag{S4}$$

obtaining the linear system of equations

$$\begin{aligned} S_{io} - S_{e0} &= (1 - S_{e0})x - S_{e0}y + G_0 z \\ S_{i1} - S_{e1} &= (1 - S_{e1})x + (G_0 - S_{e0}y + G_0 z \\ S_{i2} - S_{e2} &= (1 - S_{e2})x - S_{e0}y + G_2 z \\ S_{i3} - S_{e3} &= (1 - S_{e3})x + (G_2 - S_{e0})y + G_2 z \end{aligned} \tag{S5}$$

Numerical solution: The values of *x,y,* and *z* were obtained from the least-squared solution with nonnegative variables of the system of linear equations above. Substituting the estimated *x* and the serine exchange flux in equation (S2) we obtain the rate of serine synthesis from glucose.

Statistics: The samples for the intracellular MIDs, the fresh medium serine MIDs, and the serine uptake were always obtained from the same experiment (plate), from the same replicate (well). However, since the intracellular and extracellular LC-MS analysis was run independently, we consider the resulting MIDs independent. Furthermore, the estimate of serine uptake uses conditioned medium data and, therefore, we consider the resulting quantification independent of the fresh medium MIDs. To estimate the mean and standard deviation of the serine synthesis estimate, we sampled all possible combinations of replicates for the intracellular MIDs, fresh medium serine MIDs, and the serine uptake.

Purines synthesis

The rate of purine synthesis was estimated as described before (*21*). Specifically, we switched cells from regular culture medium to culture medium containing [$^{13}C_3$]-L-serine (*tracer*).

Purine concentrations (AMP, ADP, ATP, GMP, GDP, GTP): The peak areas were converted to absolute concentrations using calibration samples. Concentrations (*C*) were related to peak areas (*A*) by the linear model *C=aA*. The conversion factor *a* was estimated from a least-squares fit to the data for calibration samples with spiked standards. The concentrations of a given purine in calibration samples was modelled by the equations $C_0 + C_{si} = aA_i$, where C_0 is the purine concentration in the sample

and C_{si} is the concentration of the spiked standard (ATP for adenines and GTP for guanines). The conversion factor *a* was estimated from the slope (1/*a*) of A_i versus C_{si}, using four calibration samples with standard concentrations of 0, 800, 1600 or 4000 pmol ATP and GTP.

Purine synthesis model: At the time of intracellular sample collection the purine pool will be divided into the pool of purine remaining from the addition of the tracer (residual) and *de novo* synthesized purine during the time interval between tracer addition and sample collection (Δt). The newly synthesized purine pool will be partially labelled depending on the MIDs of the purine precursors glycine and 10-formyl-tetrahydrofolate. When the concentration of purines is constant and the cell count follows an exponential growth with rate μ, the rate of purine synthesis is given by [(*21*), Additional file 1, equation 1.22]

$$f = (1-r)n(k+\mu) \tag{P1}$$

where 1-*r* the *de novo* synthesized purine fraction, *r* is the residual purine fraction, n_{pur} is the purine content per cell, and

$$k = \frac{1}{\Delta t}\ln\frac{1}{r} \tag{P2}$$

is the purines turnover rate [(*21*), Additional file 1, equation 1.20]. The residual purine fraction *r* can be estimated from the purines isotopomer distribution. If we denote by P_i, G_i and F_i the purines, glycine and formate MIDs, then the purine MIDs are modelled by the equations

$$\begin{aligned}
P_0 &= (1-r)G_0F_0^2 + r \\
P_1 &= (1-r)G_0 2F_0F_1 \\
P_2 &= (1-r)G_2F_0^2 \\
P_3 &= (1-r)G_3 2F_0F_1 \\
P_4 &= (1-r)G_2F_2^2 \\
P_i &= 0, \quad i = 5,6,7,8,9,10
\end{aligned} \tag{P3}$$

From these equations we obtain

$$F_0 = \frac{y - G_0}{1 + yG_0 - 2G_0} \tag{P4}$$

$$r = 1 - \frac{x_1}{1 - F_0 G_0} \quad \text{(P5)}$$

where

$$y = \frac{x_2}{x_1} \quad \text{(P6)}$$

$$x_1 = 1 - P_0 - \frac{1}{2} P_3 \quad \text{(P7)}$$

$$x_2 = \frac{1}{2} P_1 + P_2 + \frac{1}{2} P_3 \quad \text{(P8)}$$

Numerical solution: Substituting the measured purine and glycine MIDs into equations (P4)-(P8) we estimated the residual purine fraction *r*. Substituting the estimated *r* in equation (P2) we estimated the rate of purine turnover. Finally, substituting the estimated values of the purine residual fraction (*r*), purine content per cell (*n*), purine turnover rate (*k*) and population growth rate (μ) in equation (P1) we estimated the rate of purine synthesis.

Statistics: In part the experimental estimates of MIDs and concentrations are dependent, because they both used as input the measured peak areas from the same samples. However, to calculate the concentration we also used the peak areas of the calibration samples. Therefore, to estimate the variability of the purine synthesis estimates, we considered the MIDs and concentration to be independent. To obtain the mean and standard deviation we sampled over all replicates of MIDs measurements and concentrations measurements.

Gluthatione synthesis

The estimation of the glutathione synthesis rate was similar to that for the purine synthesis rate (*21*), except for that the model for GSH MIDs (H_i) was different from that for purines.

GSH concentration: Same procedure as for purines, spiking GSH in the calibration samples.

GSH synthesis model: In experiments using [$^{13}C_3$]-L-serine as a tracer the only GSH precursor that is labelled is glycine. Therefore, the expected MIDs are

$$\begin{aligned} H_0 &= (1-r)G_0 + r \\ H_1 &\approx 0 \\ H_2 &= (1-r)G_2 \end{aligned} \qquad \text{(H1)}$$

where *r* now denotes the GSH residual fraction. From equation (H1) we obtain

$$r = \frac{H_0 - G_0}{1 - G_0} \qquad \text{(H2)}$$

Numerical solution: Substituting the measured GSH and glycine M+0 fractions into equation (H2) we estimated the residual GSH fraction *r*. Substituting the estimated *r* in equation (P2) we estimated the rate of GSH turnover. Finally, substituting the estimated values of the GSH residual fraction (*r*), GSH content per cell (*n*), GSH turnover rate (*k*) and population growth rate (μ) in equation (P1) we estimated the rate of GSH synthesis.

Statistics: Same approach as used for purines.

Serine derived formate, in vivo

The dynamics of ^{13}C formate in blood was modelled by the differential equation

$$\frac{dX_1}{dt} = s_1(t)f - x_1(t)g \qquad \text{(FI1)}$$

where X_1 is the concentration of ^{13}C formate in blood, x_1 is the ^{13}C formate fraction in blood, s_1 is the ^{13}C serine fraction in blood, *f* is the rate of formate production from serine and *g* is the rate of formate consumption, which matches the total rate of formate production from all sources. Integrating (FI1) over a given time interval (t_i,t_{0i}) we obtain

$$b_i = a_{\mathrm{f1}}f + a_{i2}g \qquad \text{(FI2)}$$

where

$$b_i = X_1(t_i) - X_1(t_{0i}) \qquad \text{(FI3)}$$

$$a_{i1} = \int_{t_{0i}}^{t_i} dt s_1(t) \approx \frac{t_i - t_{0i}}{2}\left[s_1(t_{0i}) + s_1(t_i)\right] \qquad \text{(FI4)}$$

$$a_{i2} = -\int_{t_{0i}}^{t_i} dt x_1(t) \approx -\frac{t_i - t_{0i}}{2}\left[x_1(t_{0i}) + x_1(t_i)\right] \quad \text{(FI5)}$$

where the integrals has been approximated by the trapezoidal rule. Given two or more time intervals: i=1,…,n (n>1), we can solve the linear system of equations (FI2) to determine f and g.

Numerical solution: The values of f and g were obtained from the least-squared solution with nonnegative variables of the system of linear equations (FI2), with the time intervals: (0,15 min), (15,30 min) and (30,60 min). The values of X_1 and s_1 at 15, 30 and 60 min were obtained from measurement. At time zero there is no ^{13}C formate (X_1=0). At time zero the ^{13}C serine fraction was estimated from the known amount of ^{13}C serine in the bolus (10 μmol, calculated as 100 μL of 100 mM solution) and total ^{12}C serine in blood (blood volume × serine concentration in blood). The mice analysed in these experiments have an average weight of 22.3 g that, assuming a blood volume to body weigh mass of 58.5 ml/kg (NC3Rs), corresponds to 1.3 ml of blood per mouse. The serine concentration in mice is about 0.1 mM (*31*), which corresponds to 0.13 μmol of serine in the total blood volume. Based on these values, the ^{13}C serine fraction at the time the bolus was added (time 0) is about 0.99. The flux estimates are obtained in units of mol/L of blood. They are converted to mol/kg of body weight assuming the blood volume to body weigh mass of 58.5 ml/kg (NC3Rs).

Statistics: The X_1, x_1 and s_1 measurements all correspond to the same sample. However, since the blood sampling at different was done in parallel, i.e. different mice for each time point, they are considered as independent samples. To estimate the mean and standard deviation of f and g, we sampled all possible combinations of for X_1, x_1 and s_1 across different time points and technical replicates at each time point.

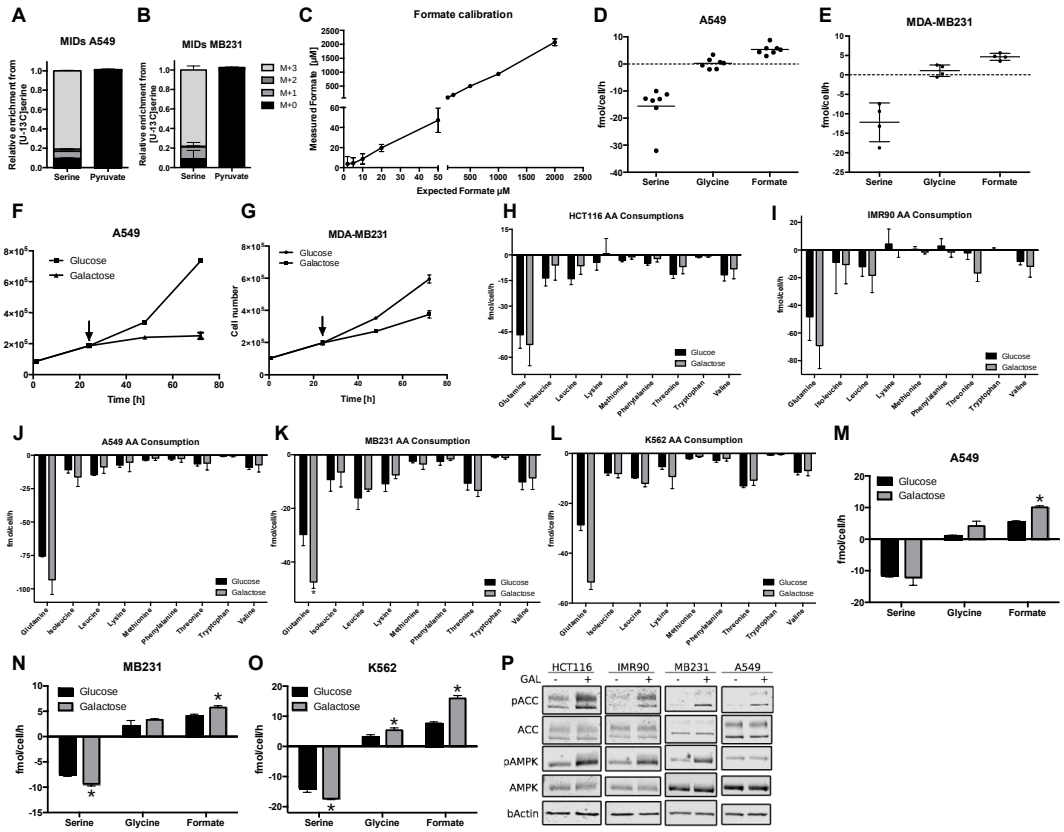

fig. S1 Serine catabolism is induced upon energy stress. (**A**-**B**) MID of pyruvate and serine after labelling with [U-^{13}C]-serine as a tracer in (B) A549 cells (8h labelling) and (B) MDA-MB231 cells (16h labelling). (**C**) Measured formate concentrations of given formate standards using the newly developed GC-MS method using M+2 formate as internal standard. (**D**-**E**) Absolute exchange rates of serine, glycine and formate in (D) A549 lung cancer cells, (E) MDA-MB231 breast cancer cells. Each dot indicates an independent experiment (performed with three cultures per experiment). (**F**-**G**) Proliferation rates of (F) A549 and (G) MDA-MB231 cells upon galactose. The arrow indicates start of galactose treatment. (**H**-**L**) Exchange rates of glutamine and essential amino acids in (H) HCT116 (I) IMR90 (J) A549 (K) MDA-MB231 and (L) K562 cells upon glucose or galactose. We note a significant difference for glutamine in (L). However, this was not reproducible in an independent repetition. (**M**-**O**) Absolute exchange rates of serine, glycine and formate in (M) A549 cells, (N) MDA-MB231 and (O) K562 leukemia cells upon galactose. (**P**) Western Blot against phosphorylated ACC (pACC) and AMPK (pAMPK), ACC, AMPK and Actin in HCT116, IMR90, A549 and MDA-MB231 cells upon galactose.

Data are presented as mean ± SD n = 3 cultures representative of at least two independent experiments (except of D-E (see above)). * $P < 0.05$ by Welch's t-test.

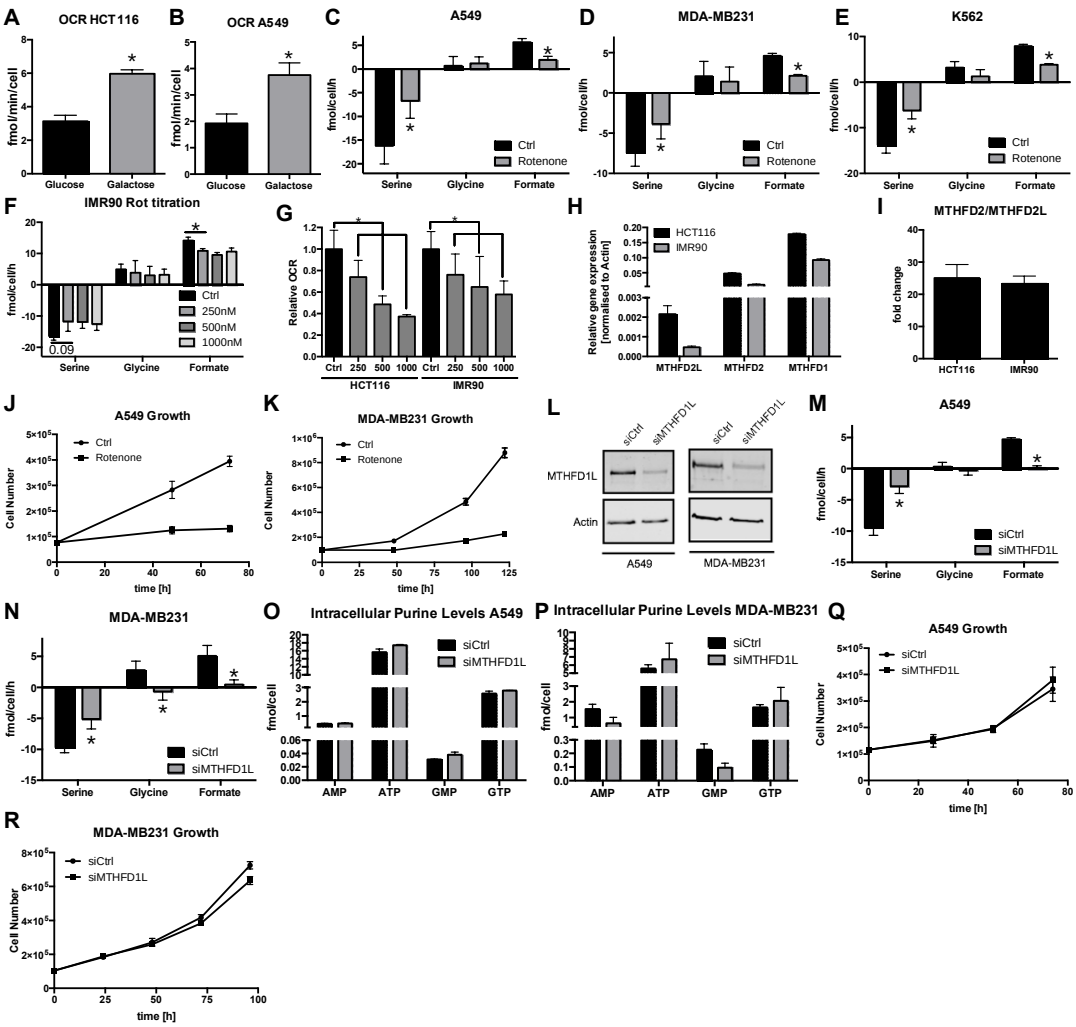

fig. S2. Serine catabolism is linked to mitochondria. (**A**-**B**) Oxygen consumption rates (OCR) in (A) HCT116 and (B) A549 cells upon galactose. (**C**-**E**) Absolute exchange rates of serine, glycine and formate in (C) A549 cells, (D) MDA-MB231 cells and (E) K562 cells upon rotenone treatment. (**F**) Absolute exchange rates of serine, glycine and formate in IMR90 cells upon different rotenone concentrations. (**G**) Relative oxygen consumption rate in HCT116 and IMR90 cells upon 250 nM, 500 nM or 1000 nM of rotenone ($n \geq 6$). Cells were cultured at baseline conditions, different concentrations of rotenone were injected and differences in OCR were determined relative to the respiration before rotenone injection. (**H**) Relative gene expression of MTHFD2L, MTHFD2 and MTHFD1 in HCT116 and IMR90 cells. (**I**) Ratio of MTHFD2/MTHFD2L gene expression levels in HCT116 and IMR90 cells. (**J**-**K**) Proliferation rates in (J) A549 and (K) MDA-MB231 cells upon rotenone. (**L**) Western Blot confirming knockdown of MTHFD1L in A549 and MDA-MB231 cells. (**M**-**N**) Absolute exchange rates of serine, glycine and formate in (M) A549 and (N)

MDA-MB231 cells upon MTHFD1L knockdown. (**O**-**P**) Absolute intracellular purine levels in (O) A549, and (P) MDA-MB231 cells upon MTHFD1L knockdown. (**Q**-**R**) Proliferation rates in (Q) A549 and (R) MDA-MB231 cells in respect to MTHFD1L knockdown. Data are presented as mean ± SD n = 3 cultures representative of at least two independent experiments (except of L (one culture) and H, I: one experiment with three cultures). * $P < 0.05$ by Welch's t-test.

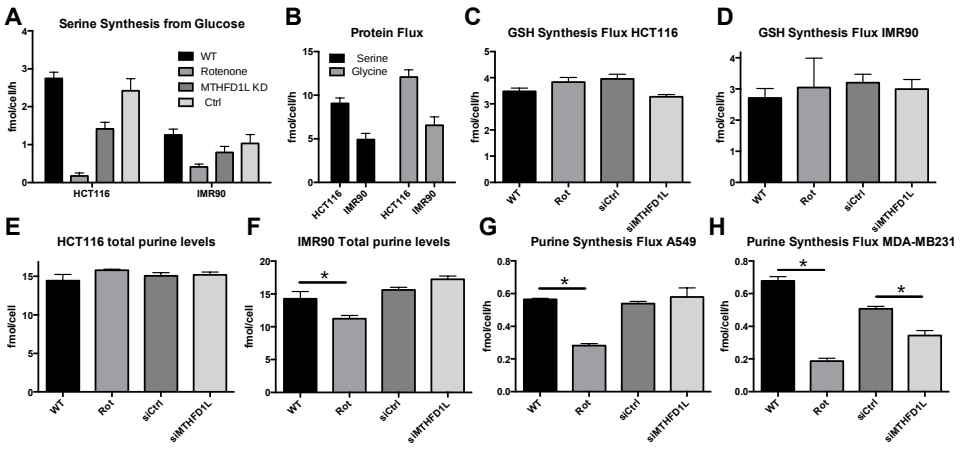

fig. S3. Formate efflux exceeds anabolic one-carbon demands. (**A**) Serine synthesis flux from glucose in HCT116 and IMR90 cells upon rotenone and MTHFD1L knockdown. (**B**) Flux of serine and glycine to proteins in HCT116 and IMR90 cells. (**C**-**D**) Absolute GSH synthesis flux in (C) HCT116 and (D) IMR90 cells. (**E**-**F**) Total purine levels upon rotenone and MTHFD1L knockdown in (E) HCT116 and (F) IMR90 cells. (**G**-**H**) Absolute purine synthesis flux in (G) A549 and (H) MDA-MB231 cells. Data are presented as mean ± SD n = 3 cultures representative of at least two independent experiments. * $P < 0.05$ by Welch's t-test.

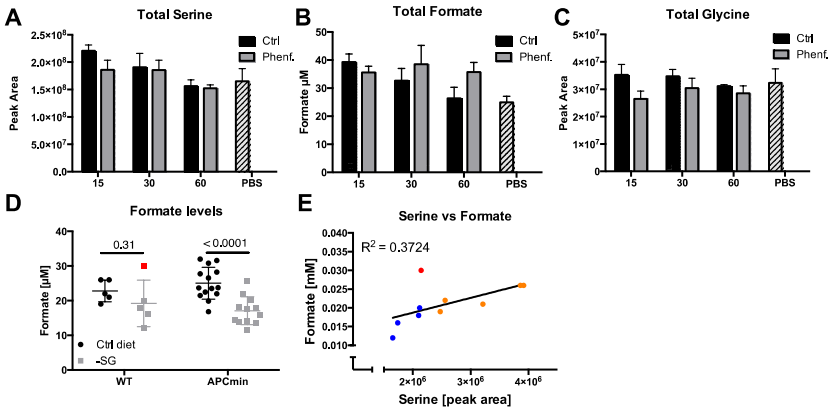

fig. S4. Serum formate depends on serine catabolism in vivo. (**A**) Total plasma serine levels (peak area) in C57BL6 mice 15, 30 and 60 minutes after intraperitoneal injection of [3-$^{13}C_1$]serine with or without Phenformin. Pure PBS injection (same volume) served as baseline control. (**B**) Absolute plasma formate concentration as in (A). (**C**) Total plasma glycine levels (peak area) as in (A). Data are presented as mean ± SD n = 4 mice. (**D**-**E**) Same data as in Fig. 4F,G but including the outlier. The outlier that has been excluded for technical reasons is indicated in red.

table S1. GC temperature program for formate analysis.

Start temperature (°C)	Ramp (°C/min)	End temperature (°C)	Hold time (min)
60	-	60	0.5
60	38	230	1