Supplementary Information

Untargeted metabolomics links glutathione to bacterial cell cycle progression

Johannes Hartl[1*], Patrick Kiefer[1†], Andreas Kaczmarczyk[2†], Maximilian Mittelviefhaus[1], Fabian Meyer[1], Thomas Vonderach[3], Bodo Hattendorf[3], Urs Jenal[2], Julia A. Vorholt[1*]

[1]ETH Zurich, Institute of Microbiology, Switzerland

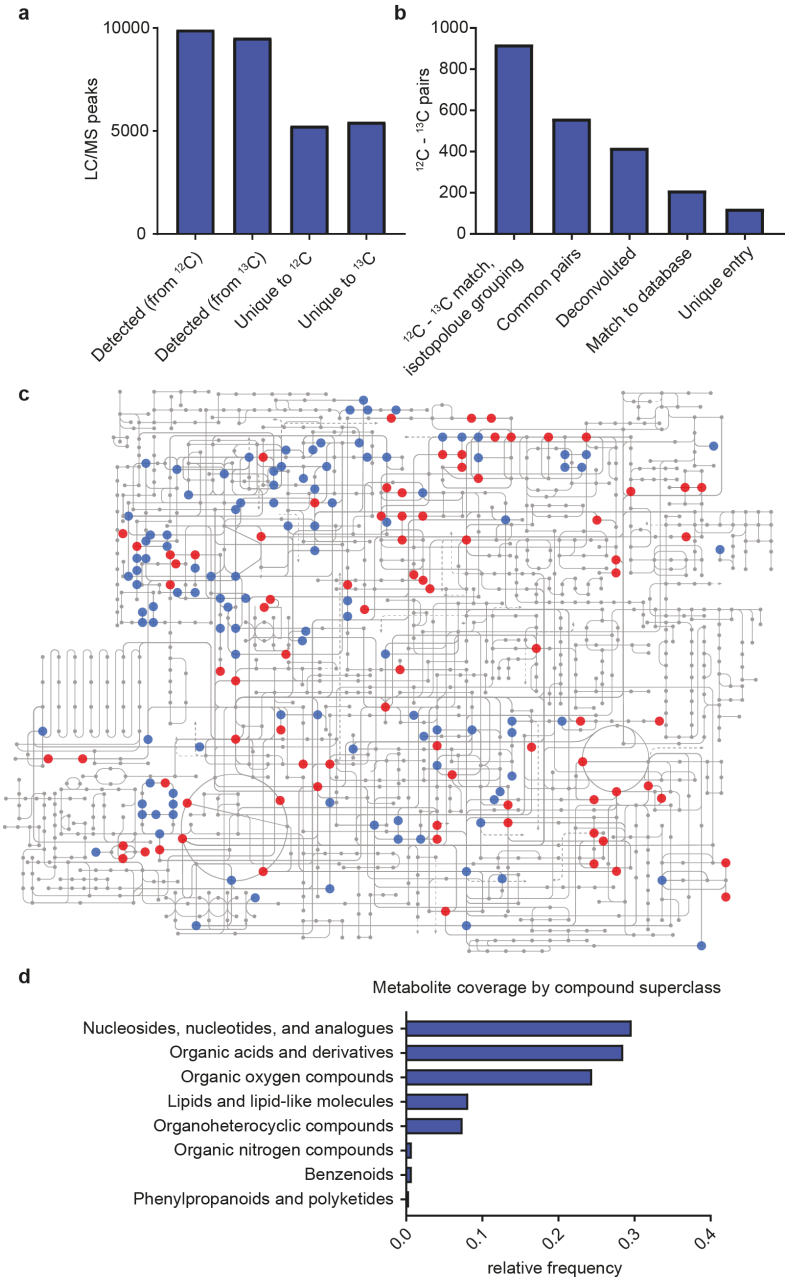
[2]Biozentrum of the University of Basel, Basel, Switzerland

[3]ETH Zurich, Laboratory of Inorganic Chemistry, Switzerland

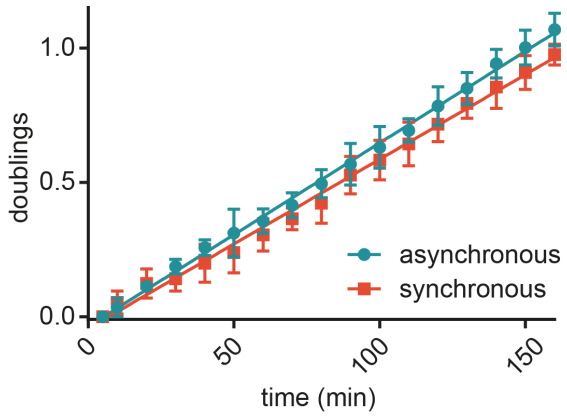
[†]Second authors contributed equally

*To whom correspondence should be addressed

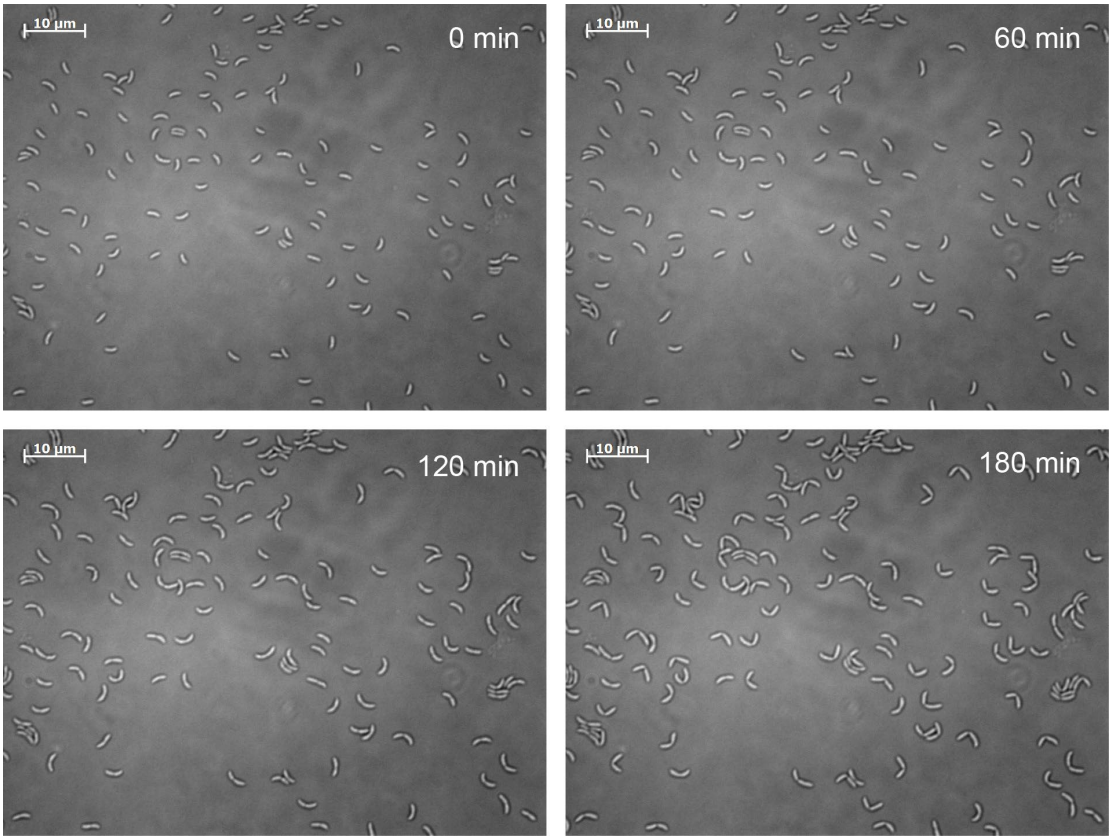
This file includes:



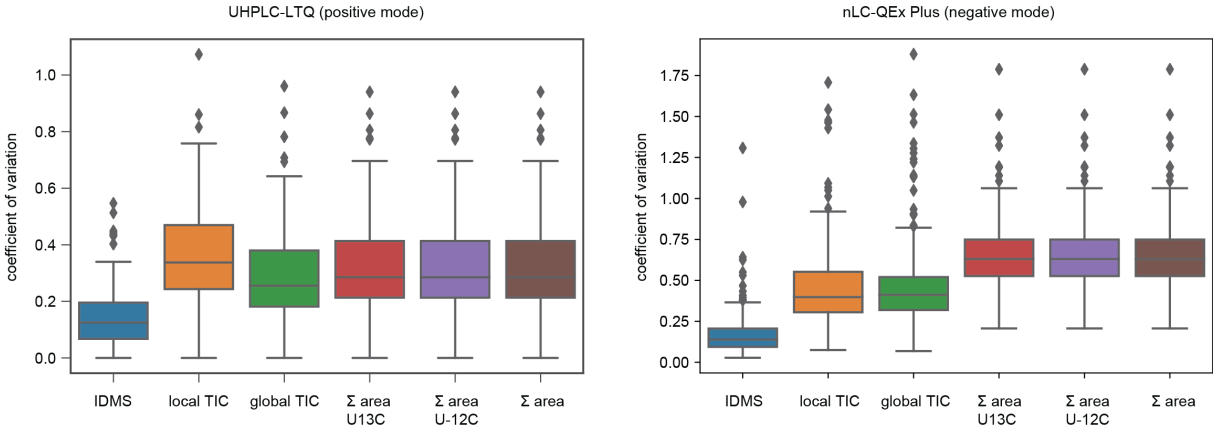
Supplementary Figure 1. Extracted metabolite features, annotation and coverage obtained from *Caulobacter crescentus* extracts. (a) Barplots depicting an overview of extracted LC/MS peaks. Shared peaks obtained from both 12C and 13C extracts were discarded (i.e. only peaks with unique spectral identity were considered). Average of six measurements are depicted. **(b)** Cheminformatic filtering and metabolite identification. Average number of features that had corresponding 12C - 13C labeling pairs; note that concomitantly isotopologues were grouped to avoid redundancy. Common pairs denote feature pairs that were obtained in 4 out of 6 processings. Deconvoluted feature pairs were ultimately used for extraction and processing of metabolites during the dynamic metabolomics approach. Metabolites were annotated by database matching as described below. For more explanations, details see Fig. 1 and Supplementary Methods. **(c)** Central metabolic map with putatively annotated metabolites. Blue circles are ambiguously annotated ions (no unique molecular formula, multiple matches returned from database query), red circles highlight unique hits (only single database match obtained). Graph was created using iPath3 (https://pathways.embl.de/)[1]. **(d)** Chemical classification of putatively identified metabolites. Retrieved KEGG identifiers were translated into *Inchi* keys using the chemical translation service tool (http://cts.fiehnlab.ucdavis.edu/)[2]; respective superclasses were annotated using Classyfire (http://classyfire.wishartlab.com/)[3].



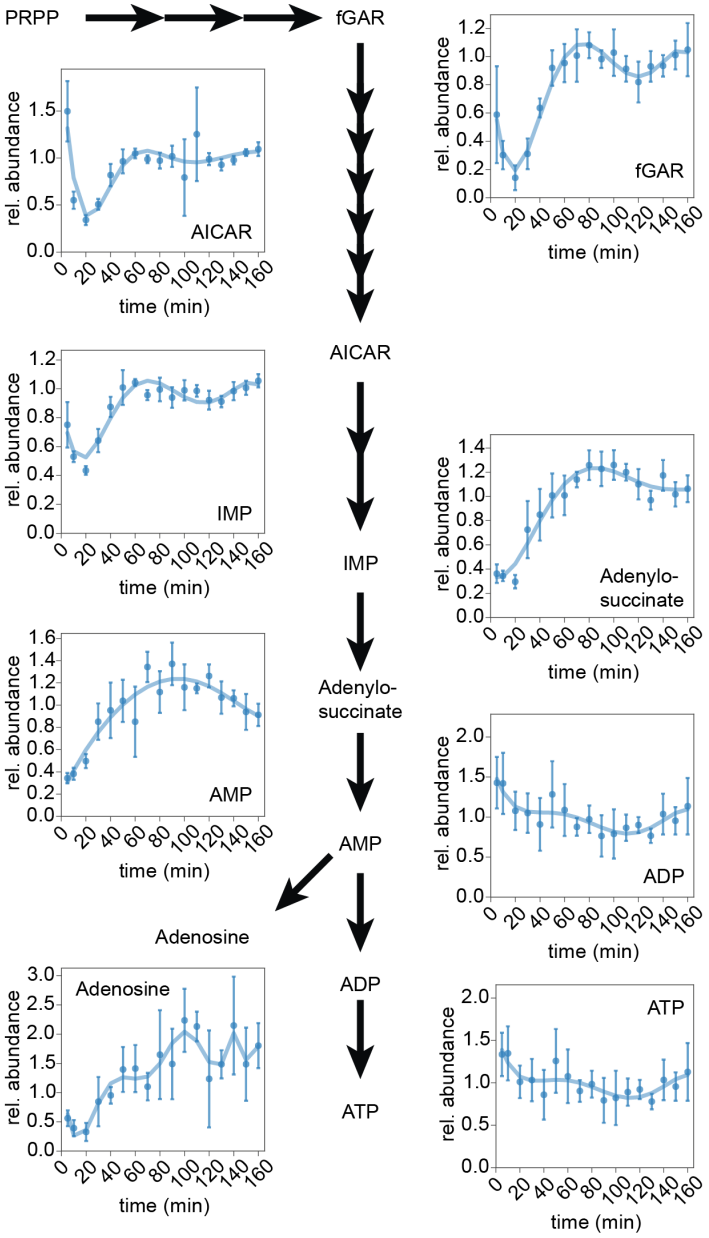
Supplementary Figure 2. Growth of synchronized and asynchronous control cultures. *C. crescentus* cultures were cell cycle synchronized or subjected to a control treatment as described in Methods. OD600 was determined over time, and values were normalized to the initial time point (for calculation of doublings, see[4]). Note that both synchronized (red) asynchronous control (turquoise) cultures grow comparably and readily retain growth (mean values +/- standard deviations are shown, n=3 biological replicates per condition)



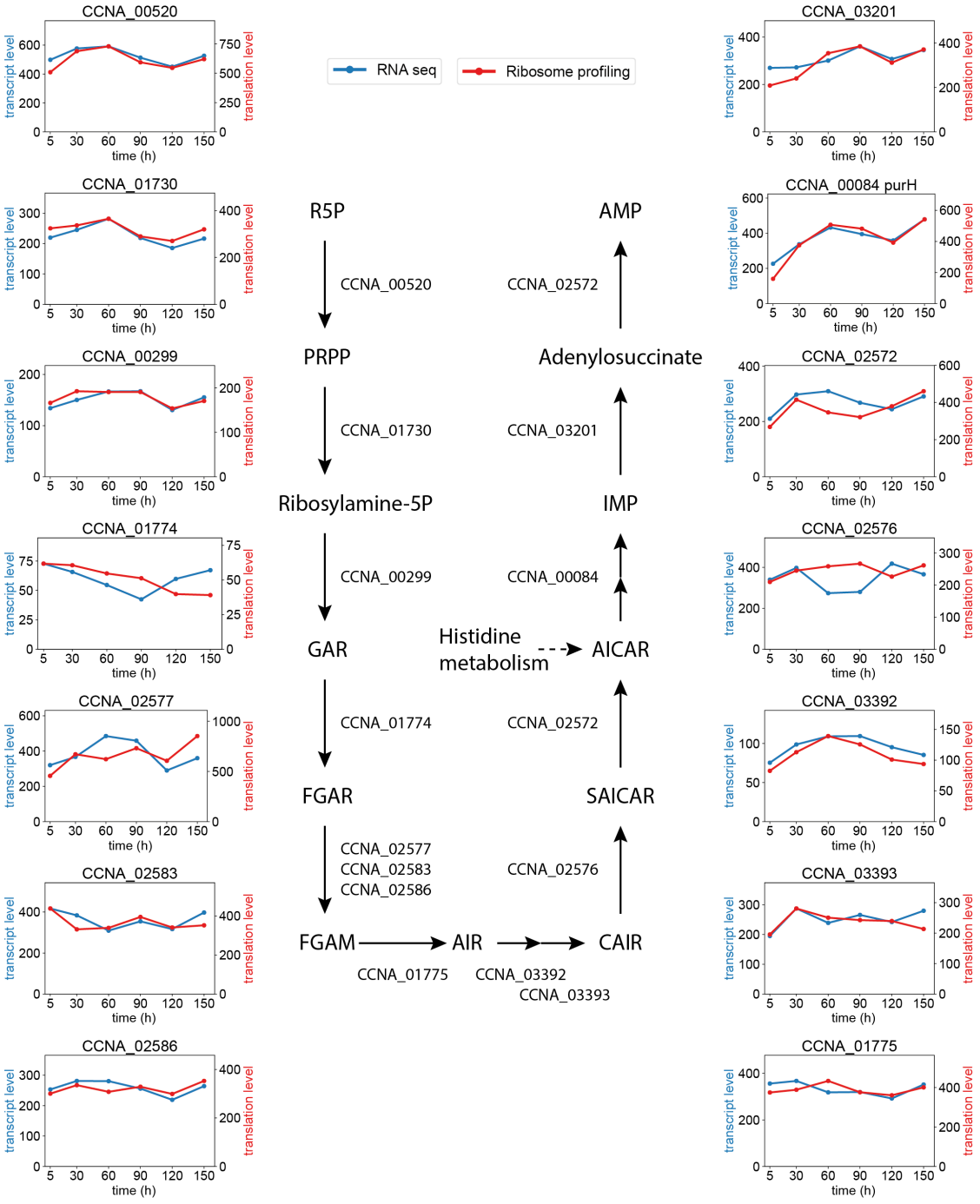
Supplementary Figure 3. Time-lapse microscopy of synchronized *C. crescentus* cells. A culture was cell cycle synchronized; isolated swarmer cells were spotted on agar pads and followed for 180 min. The initial picture (t=0min) was acquired ~15 min after the end of synchronization protocol. Note that growth-rate is reduced on oxygen-limited agar pads compared to liquid cultures. Brightness and contrast of pictures were adjusted for better visibility of individual growing cells. The experiment was independently performed once. Microscopy scale bar is 10 µm.



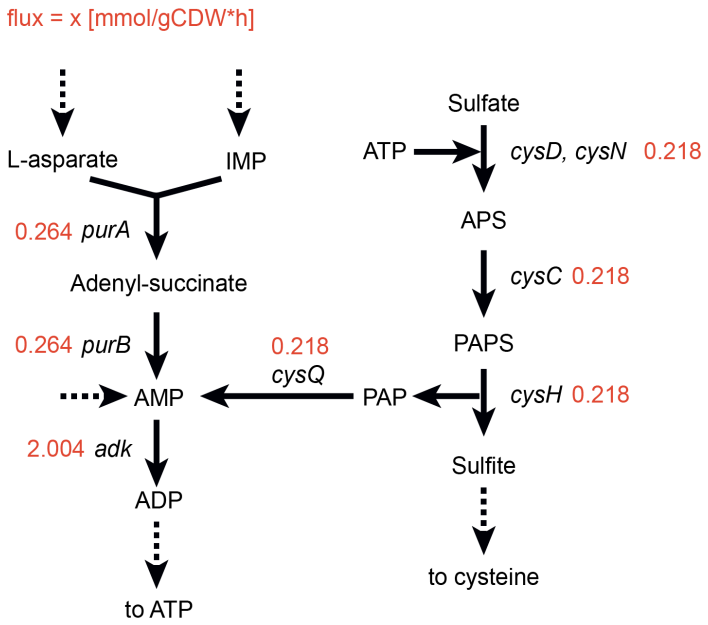
Supplementary Figure 4. Robust quantification by isotope dilution mass spectrometry. Coefficients of variation (CV) calculated from different data normalization approaches for indicated LC-MS methods. Samples were from exponentially growing (^{12}C) *C. crescentus* extracts (one for each acquired batch, respectively; i.e. 6 biological replicates, each sample was spiked with ^{13}C enriched reference extract during sampling). For each measurement, metabolites were extracted based on the retention time and m/z windows obtained by the SID-filtered library (415 metabolite features) and integrated using the trapezoid rule. CV values were calculated as standard deviation over mean for each extracted metabolite feature (127, 288 for positive and negative mode data respectively) obtained from various normalization methods: (**IDMS, blue**) isotope dilution mass spectrometry by normalization of u-^{12}C peak to corresponding u-^{13}C peak, **(local TIC, orange**) normalization of peak area to local total ion count (TIC) within extraction windows of respective peak, (**global TIC, green**) normalization of the peak areas to the global TIC of respective peakmap, (**$\underline{\Sigma}$ u-^{12}C area, red**) normalization of the peak area to the summed area of all extracted u-^{12}C peaks, (**$\underline{\Sigma}$ u-^{13}C area, purple**) normalization of the peak area to the summed area of all extracted u-^{13}C peaks, (**$\underline{\Sigma}$ area, brown**) normalization of the peak area to the summed area of all peaks. Normalization based on isotope dilution mass spectrometry, as expected, yielded lowest CVs values for both applied methods. Tukey boxplot with center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers.



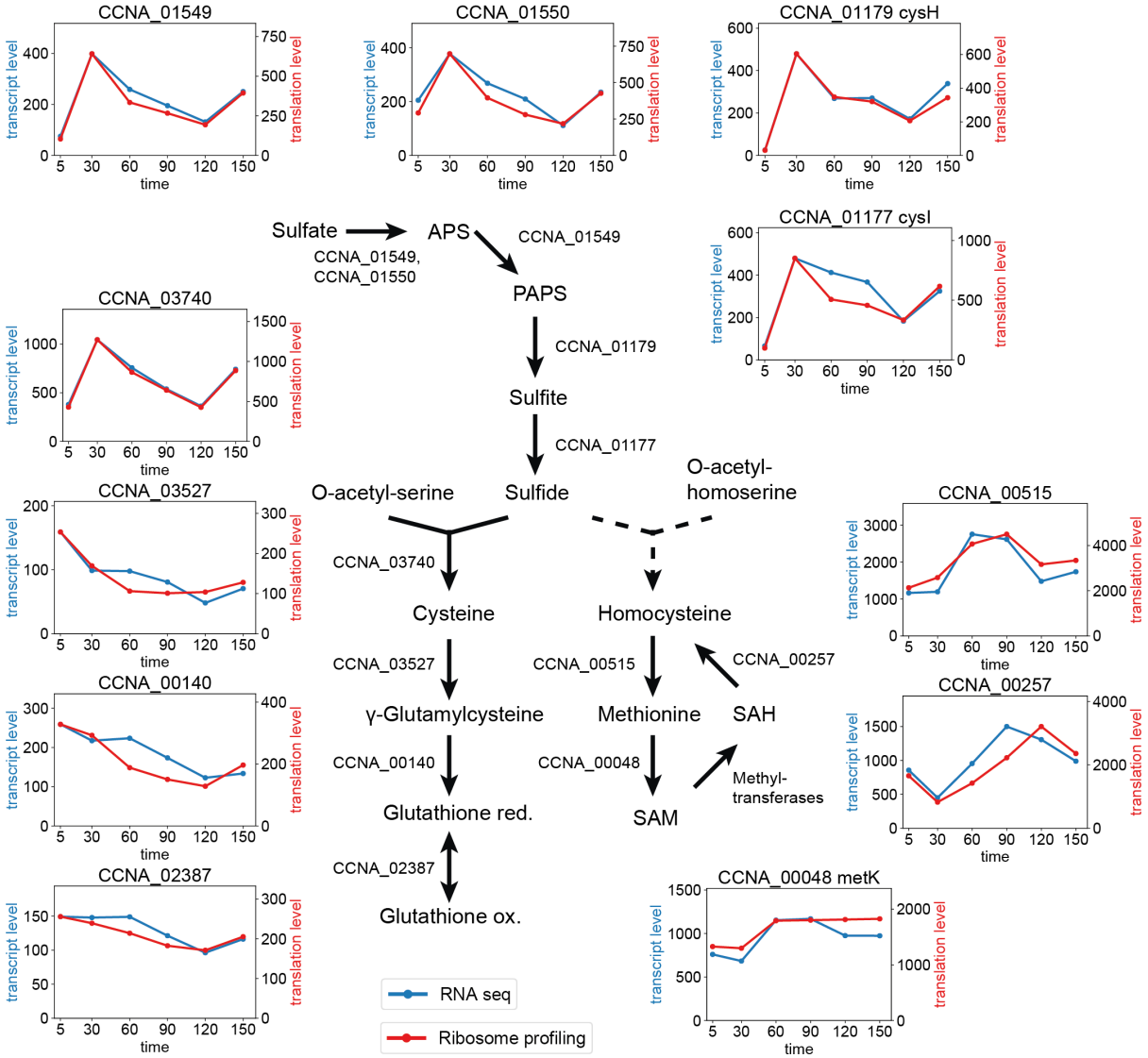
Supplementary Figure 5. Cell cycle dependency of adenylate metabolism. A simplified scheme of the purine *de novo* synthesis pathway is shown. Time series plots depict dynamic changes of respective purine *de novo* synthesis metabolites and adenylate pools. Relative abundance profiles were calculated as the ratio of the response of synchronized vs controls (mean values +/- SD; n=3 per condition; error propagation was applied). Note that guanylate pools are not affected by the low precursor pools early in the cell cycle; see Supplementary Data 1, metabolites from nLC_QEx (-) mode, id 112 (GTP), 216 (GDP) and 1201 (GMP).



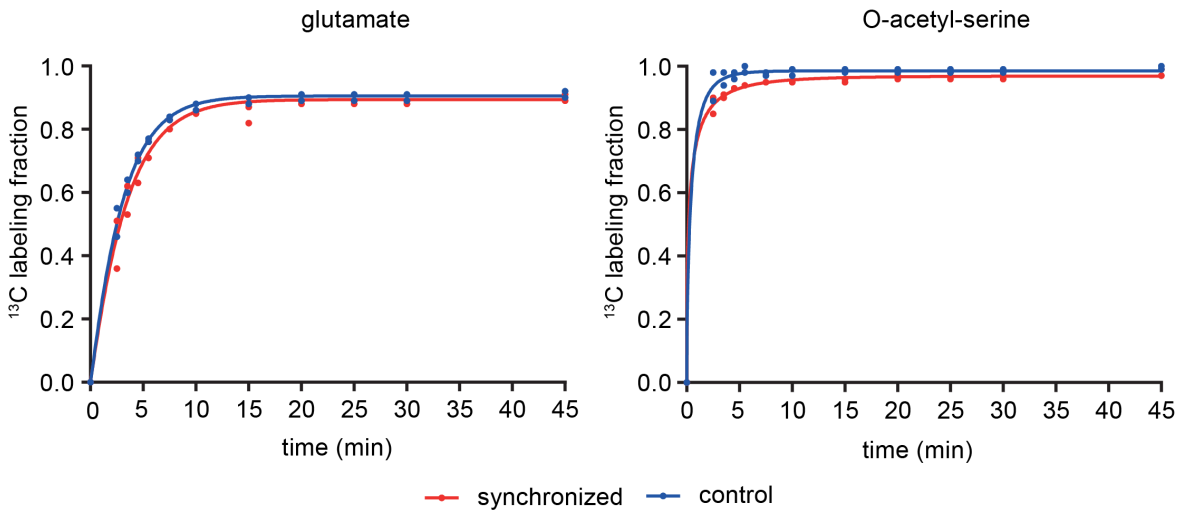
Supplementary Figure 6. Transcriptomic analysis of purine metabolism. A simplified version of the *de novo* AMP biosynthesis pathway obtained by KEGG[5] homology mapping (*C. crescentus* NA1000). Arrows indicate enzyme-catalyzed reactions, respective CCNA numbers correspond the associated genes. Time series plots show corresponding mRNA expression levels as determined by RNA seq. (blue, left y-axis) and translation levels obtained by ribosome profiling (red, right y-axis) following one cell cycle of synchronized *C. crescentus* cells. Abbreviations: R5P, ribose 5-phosphat; PRPP, 5-phosphoribosyl diphosphate; GAR, glycinamide ribonucleotide; FGAR, N-formylglycinamide ribonucleotide; FGAM, 5'-phosphoribosylformylglycinamidine; AIR, aminoimidazole ribotide; CAIR, 5'-phosphoribosyl-4-carboxy-5-aminoimidazol; SAICAR, 5'-Phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole; AICAR, 5-aminoimidazole-4-carboxamide ribotide; IMP, inosine monoshosphate; AMP, adenosine monophosphate. Data is based on single replicates per time point and was obtained from Schrader et al [6].



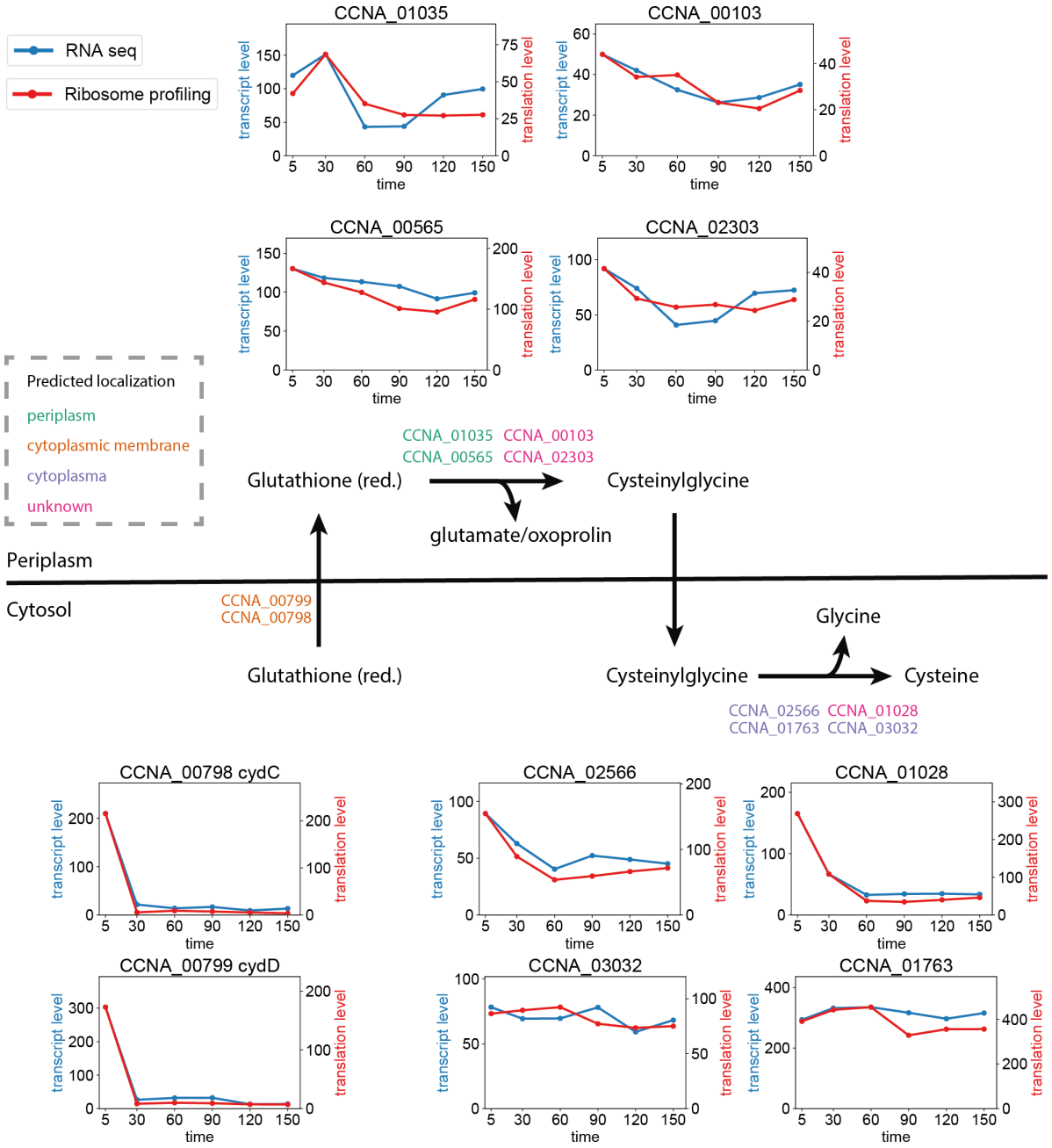
Supplementary Figure 7. Flux balance analysis of AMP *de novo* synthesis and sulfate assimilation pathways. Black arrows represent indicate enzymes, flanked by the respective gene. Red numbers represent flux over the reaction (in mM*(gCDW*h)$^{-1}$). Model is based on *Escherichia coli*[7], with a glucose uptake rate of 10 mmol*(gCDW*h)$^{-1}$. Sulfate uptake rate was not limiting. *C. crescentus* employs homologous reaction network for purine *de novo* synthesis and sulfur assimilation as depicted for *E. coli*. Note the connection between both pathways over PAP. For details, see Methods. Abbreviations: IMP, inosine monophosphate; AMP, adenosine monophosphat; ADP, adenosine diphosphate; ATP, adenosine diphosphate triphosphate; APS, adenylylsulfate; PAPS, phosphoadenylyl-sulfate; PAP, Phosphoadenosine phosphate.



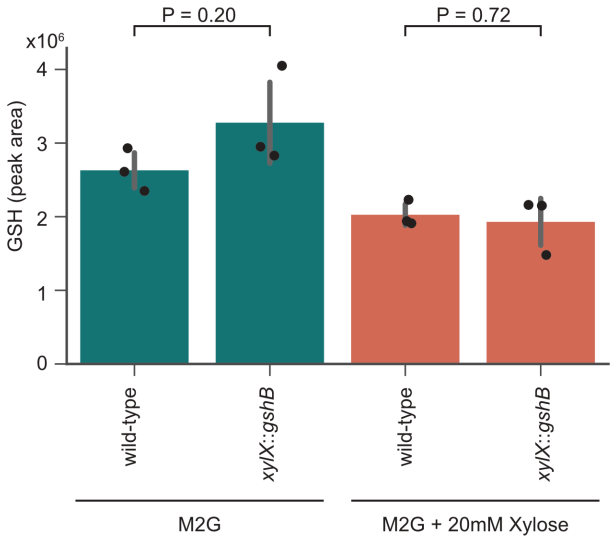
Supplementary Figure 8. Transcriptional control of sulfur metabolism. A simplified version of the sulfur metabolic network obtained from KEGG[5] homology mapping (*C. crescentus* NA1000). Arrows indicate enzyme-catalyzed reactions, respective CCNA numbers correspond to the associated genes. Dashed arrows indicate a reaction with no known enzyme homolog. Time series plots show corresponding mRNA expression levels as determined by RNA seq. (blue, left y-axis) and translation levels obtained by ribosome profiling (red, right y-axis) following one cell cycle of synchronized *C. crescentus* cells. Data is based on single replicates per time point and was obtained from Schrader et al [6].



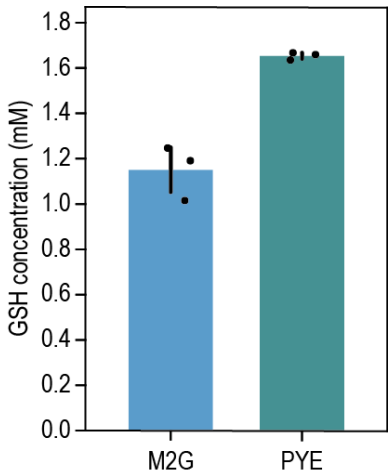
Supplementary Figure 9. Dynamic labeling incorporation into glutamate and O-acetyl-serine. Fractional ^{13}C labeling enrichment analysis into glutamate (left) of and O-acetyl-serine (right) from cell cycle synchronized (red) vs. asynchronous control (blue) cultures. Two biological replicates per condition are shown. Note that the labeling dynamics of other glutathione precursors (glycine, cysteine) could not be obtained.



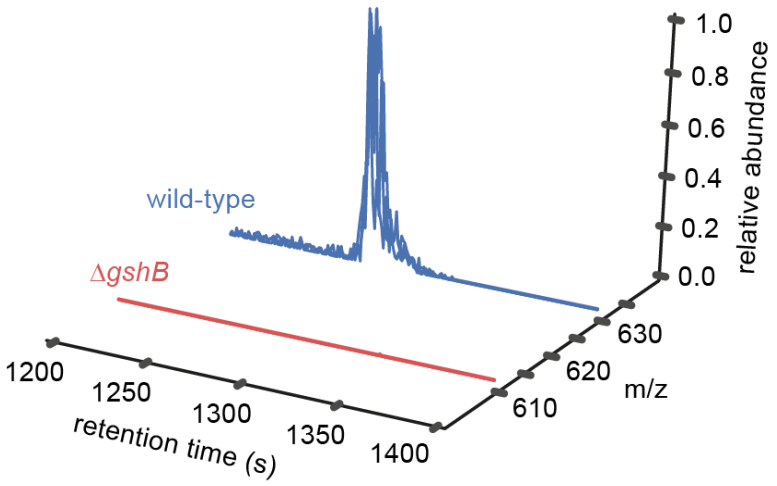
Supplementary Figure 10. Transcriptional regulation of glutathione catabolism. A simplified glutathione degradation pathway obtained by KEGG[5] homology mapping (*C. crescentus* NA1000) is shown. Arrows indicate enzyme-catalyzed reactions, respective CCNA numbers correspond to the associated genes. Gene color show predicted localization of the corresponding protein (PSORTdb predictions[8] as obtained from Caulobrowser[9]). Time series plots show corresponding mRNA expression levels as determined by RNA seq. (blue, left y-axis) and translation levels obtained by ribosome profiling (red, right y-axis) following one cell cycle of synchronized *C. crescentus* cells. Data is based on single replicates per time point and was obtained from Schrader et al[6].



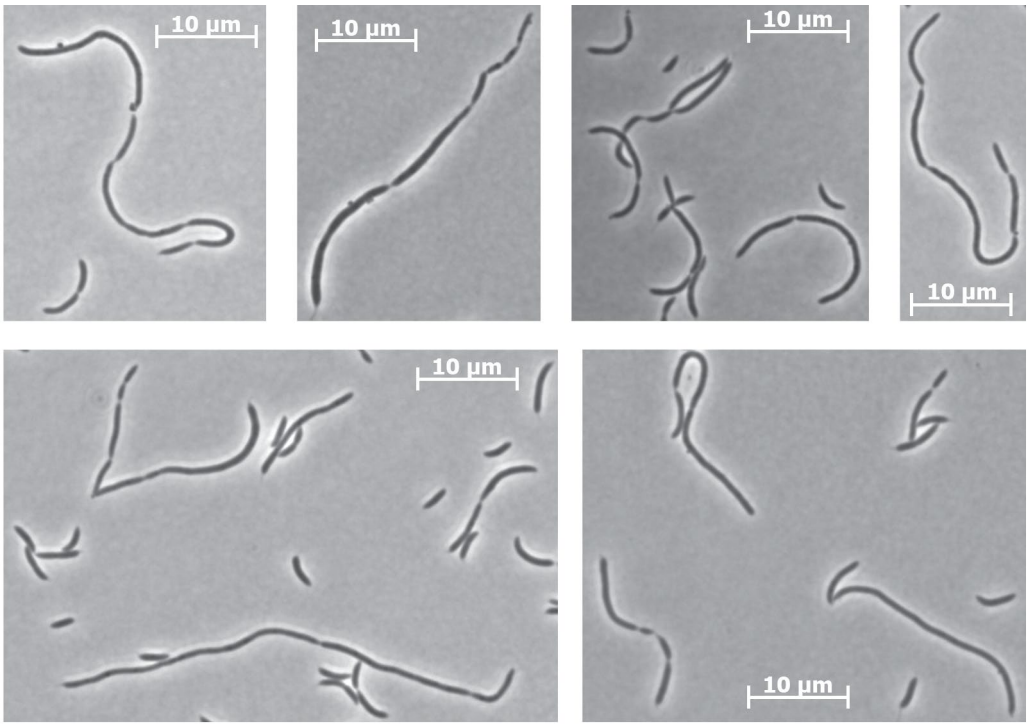
Supplementary Figure 11. Overexpression of glutathione synthetase has no major influence on glutathione levels. Wild-type cells and a strain with an additional copy of the glutathione synthetase gene (*gshB;* strain AKS439) expressed from the *xylX* locus were grown exponentially on minimal media with glucose (M2G, turquoise bars), or fully induced[10] with 20mM xylose for 6 hours (M2G+Xylose, red panel). Note that even in absence of xylose, *xylX*::*gshB* fully complements a strain lacking *gshB* at its native locus, suggesting residual expression (see main text). Shown are means +/- SD of n=3 biological replicates each. GSH was quantified by LC-MS via targeted extraction and peak integration. Statistical comparison was performed using two-sided unpaired t-tests.



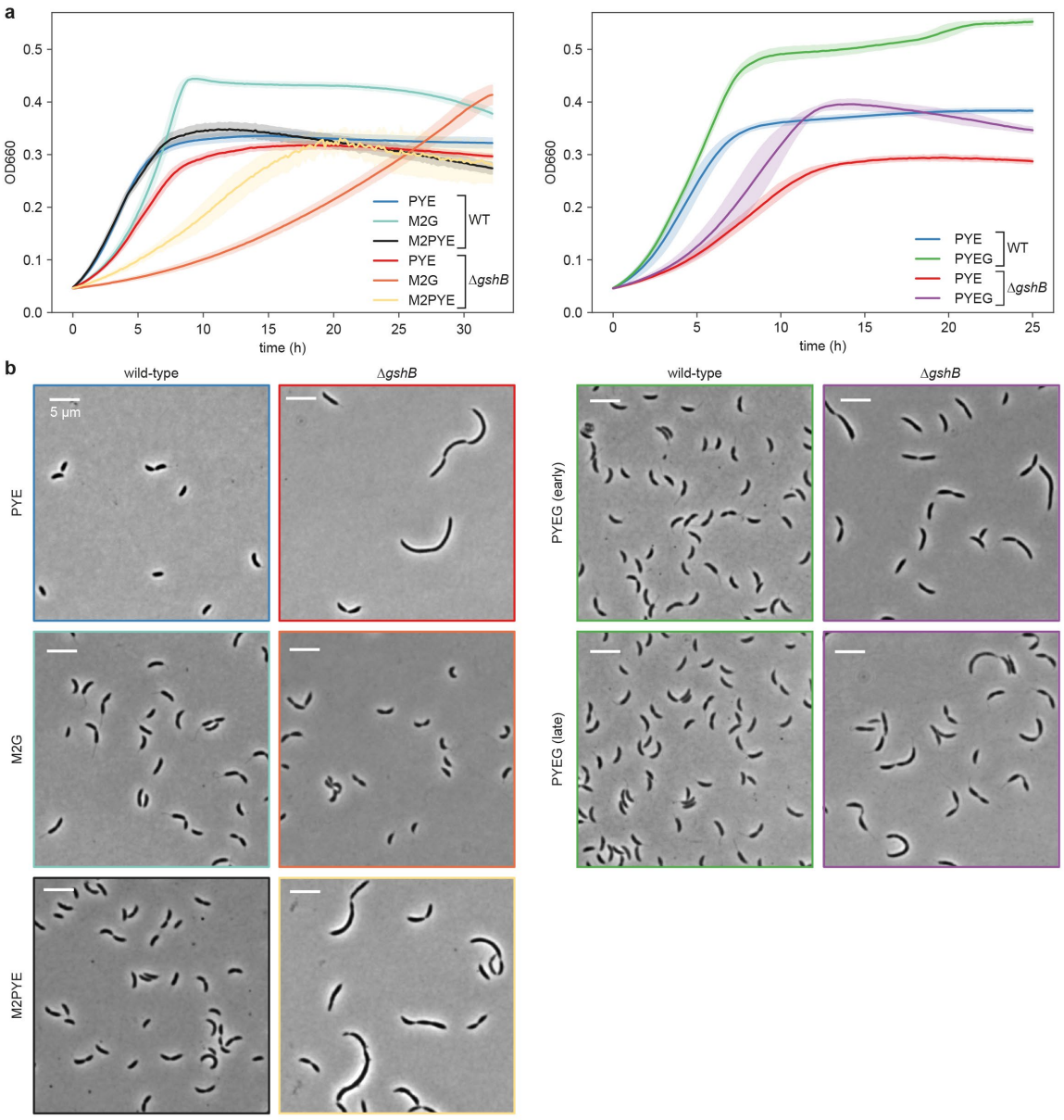
Supplementary Figure 12. Intracellular GSH concentration of *C. crescentus* wild-type cells. Cells were grown on M2G (left) or PYE (right) media. Absolute quantification was determined with a quantified ^{13}C enriched *C. crescentus* extract as an internal standard (see Methods). Shown are means +/- SD of n=3 biological replicates



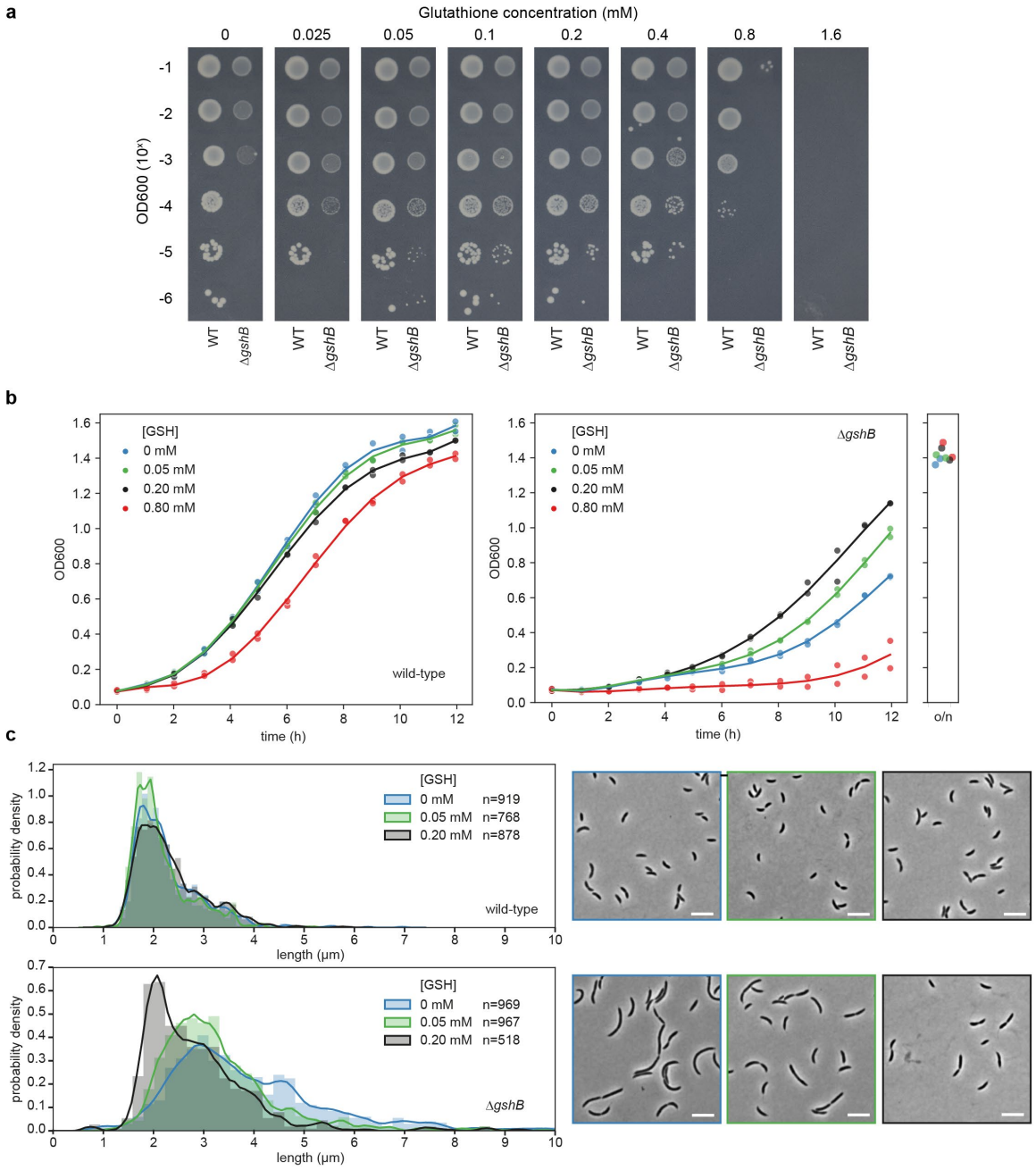
Supplementary Figure 13. No detectable oxidized glutathione in Δ*gshB* cells. Cells lacking the glutathione synthetase gene (*gshB*) have no detectable intracellular oxidized glutathione (GSSG). Shown are extracted ion chromatograms (EIC; from LC-MS data) of GSSG extracted from Δ*gshB* cells (red traces, u-^{12}C signals, m/z 611.1447) vs. a spiked wild type extract (blue traces, u-^{13}C signals, m/z 631.2118. Employed m/z tolerance for extraction was 3 milli mass units. Data was normalized to the highest spectrum intensity of the wild-type GSSG signal per measurement; n=3 biological replicates are shown.



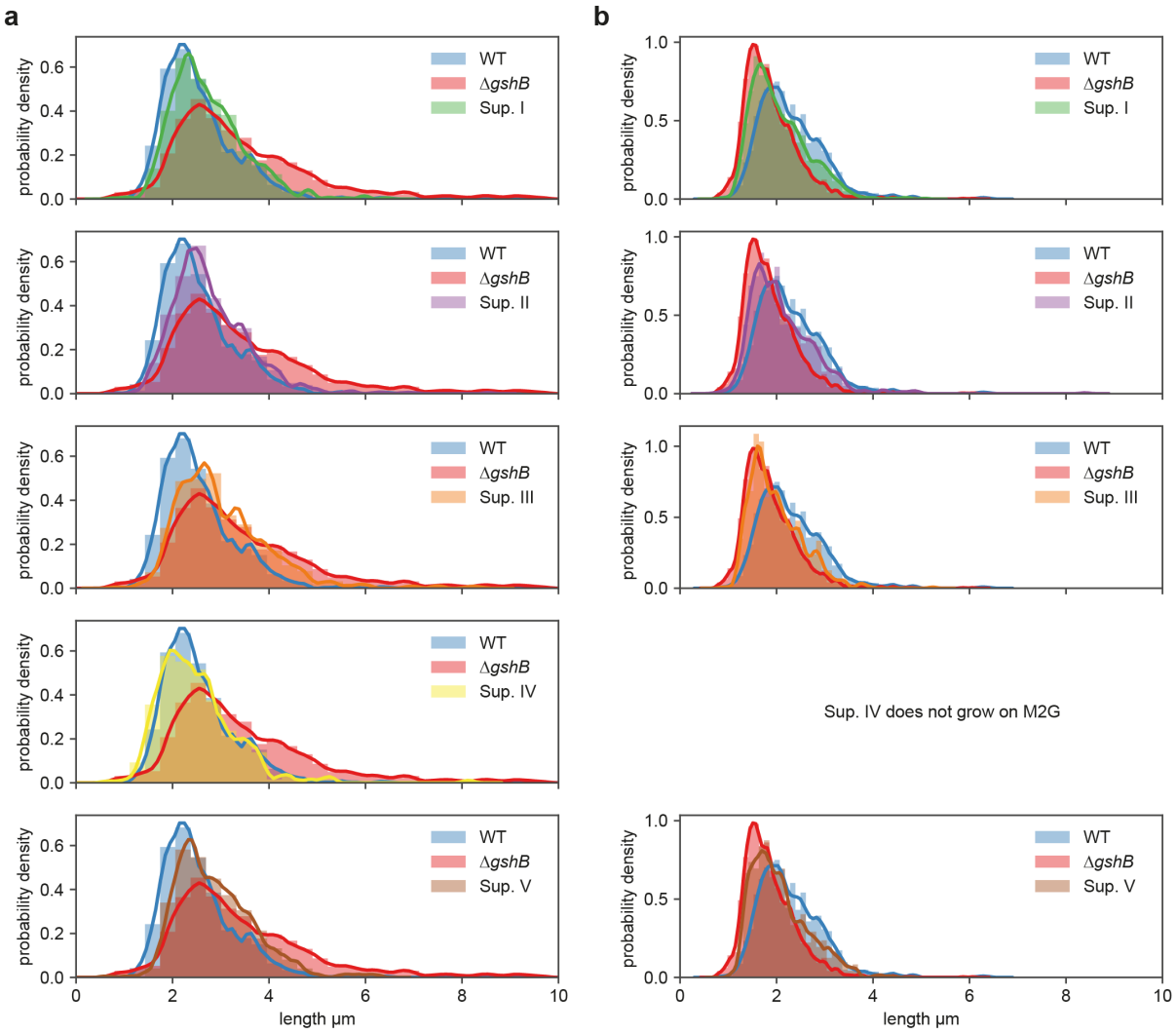
Supplementary Figure 14. Subpopulation of Δ*gshB* cells with highly aberrant cell morphology. Deletion of *gshB* in *C. crescentus* leads to a highly variable cell length distribution upon growth on complex media PYE (Fig. 4h, i). Depiction of a fraction of Δ*gshB* cells that are extremely filamented and misshaped, including partially constricted cells. Microscopy scale bar is 10 µm as indicated. Experiments were repeated at least three times with similar results.



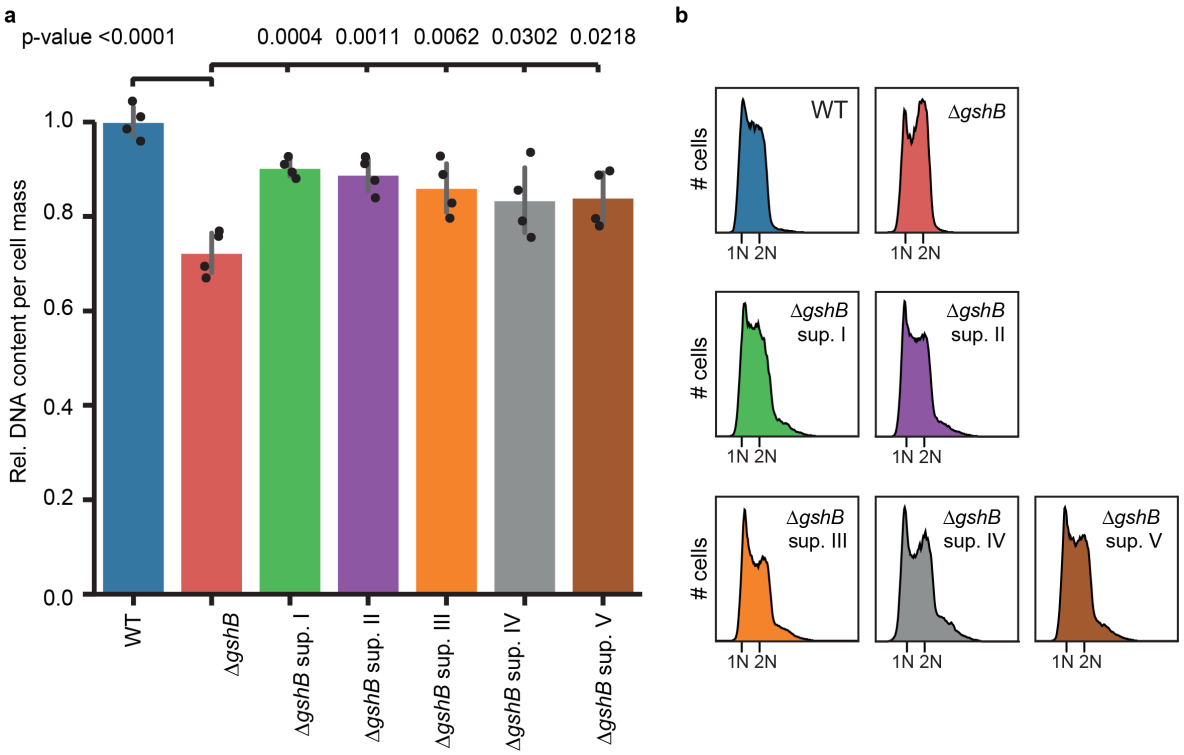
Supplementary Figure 15. Substrate dependent morphology defects in Δ*gshB* cells. (a) Growth curves of wild-type and Δ*gshB* cells on complex media with peptone + yeast extract (PYE), minimal media with M2 salts + glucose (M2G), M2 salts with peptone + yeast extract (M2PYE) or complex media supplemented with additional glucose (PYEG). Mean values +/- SD of n=5 (Δ*gshB* cells grown in M2G) or n=6 (all other) biological replicates. **(b)** Cell morphology of wild-type and Δ*gshB* grown on media compositions as described in (a); same color code. Cells grown on PYEG were imaged during early exponential phase ("early") and when cells surpassed the optical density supported by PYE alone ("late"). Indicated microscopy scale bars are 5 µm. The experiment was independently performed twice with similar results.



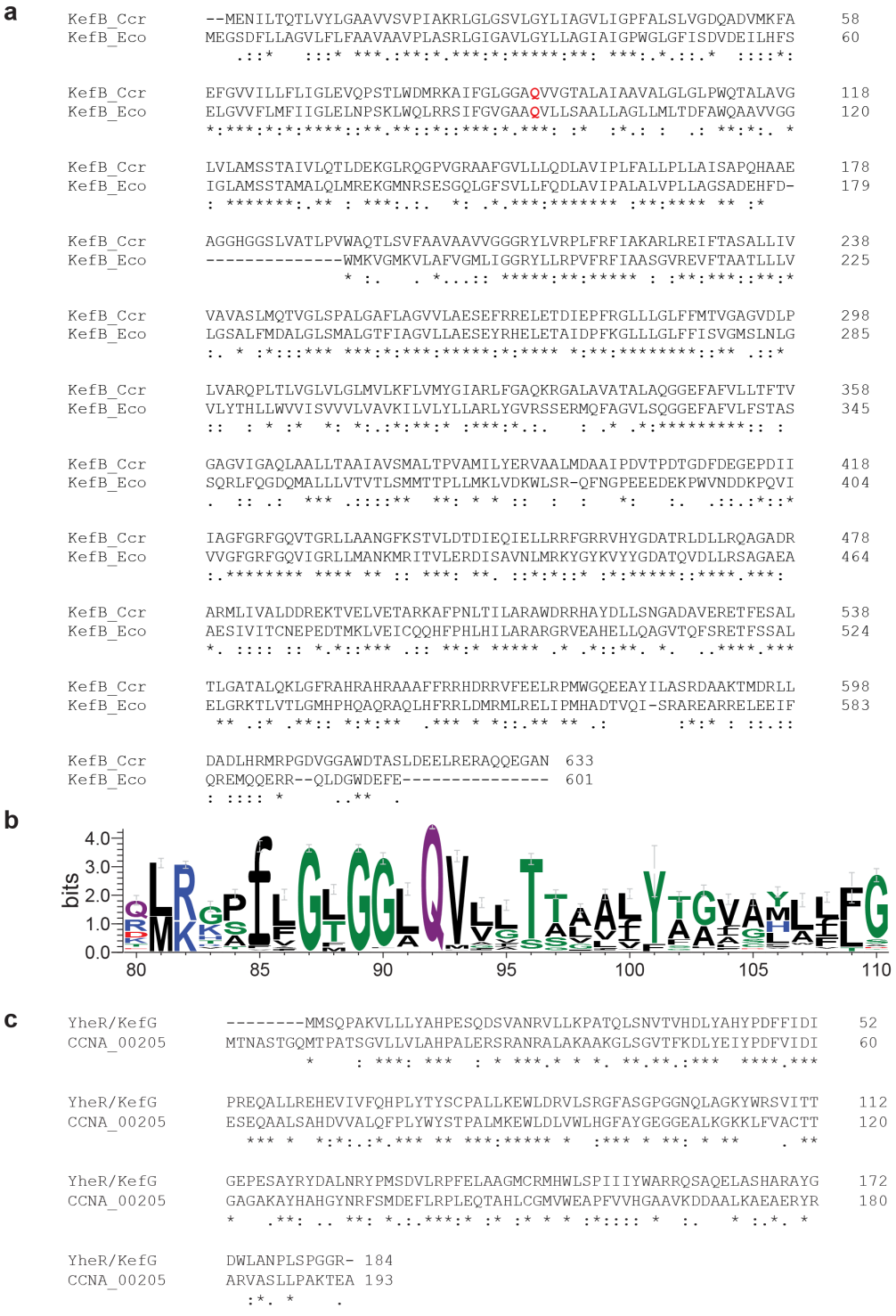
Supplementary Figure 16. Chemical complementation of *ΔgshB* cells with glutathione. (a) Serial dilutions of wild-type and Δ*gshB* cells spotted on PYE alone, or supplemented with increasing concentrations of reduced glutathione (GSH). The experiment was independently performed twice with similar results. **(b)** Growth curves of wild-type (left) and Δ*gshB* (right) cells supplemented w/o (blue), 0.05 mM (green), 0.2 mM (black) or 0.8 mM (red) GSH. The box next to the growth curve of Δ*gshB* cells indicate the OD600 after overnight (o/n) incubation. To obtain growth curves, cultures were grown to exponential phase, diluted, aliquoted, supplemented with the indicated amounts of GSH (at t=0 h), and growth was monitored. Shown are n=2 biological replicates per condition. **(c)** Cell length and morphology of wild-type (top) and Δ*gshB* (bottom) cells grown on PYE supplemented w/o (blue), 0.05 mM (green), 0.2 mM (black) GSH. Histograms (left) display the probability density of cell length distributions; outlines shows a fitted kernel estimate. Number of cells (n) per condition is indicated. Cells larger >10 µm are not included. Corresponding representative microscopy images (right); white scale bar indicates 5 µm. To obtain data, exponentially growing o/n cultures (on PYE) were diluted back to OD600 of 0.0001 (wild-type) or 0.001 (Δ*gshB*), aliquoted, and supplemented with the indicated amounts of GSH, and grown for >24 h before imaging; from n=2 biological replicates.



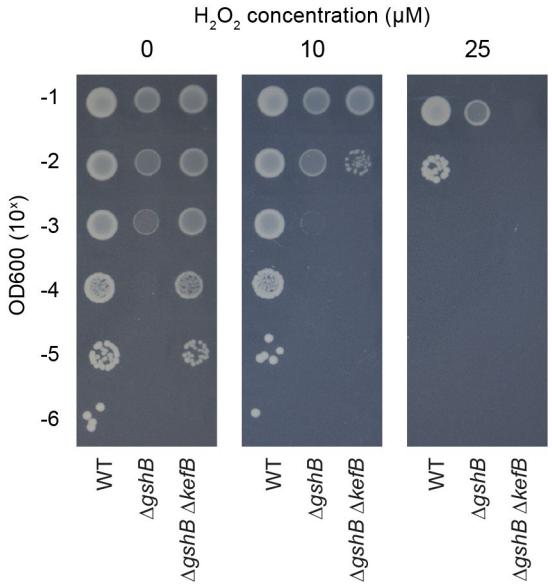
Supplementary Figure 17. Cell body length distribution of Δ*gshB* suppressor isolates. Cell length distribution of wild-type vs. Δ*gshB* and respective Δ*gshB* suppressor strains grown on **(a)** PYE or **(b)** M2G media. Cells larger >10 µm are not shown. Cell body length was estimated using MicrobeJ. Histograms show counts normalized to form a probability density; bin size was estimated on the broadest distribution (Δ*gshB* cells) following the Freedman–Diaconis rule. Outline shows a fitted kernel density estimate. Note that the distributions of wild-type and Δ*gshB* are the same as shown in Fig. 4. Plots were made using the Python based *Seaborn* and *Matplotlib* package. All data from n=4 biological replicates per genotype. Cell numbers on PYE: 1348 (WT), 1066 (Δ*gshB*), 846 (Sup. I), 980 (Sup. II), 612 (Sup. III), 617 (Sup. IV), 553 (Sup. V). Cell numbers on M2G: 754 (WT), 1785 (Δ*gshB*), 1208 (Sup. I), 928 (Sup. II), 663 (Sup. III), 666 (Sup. V).



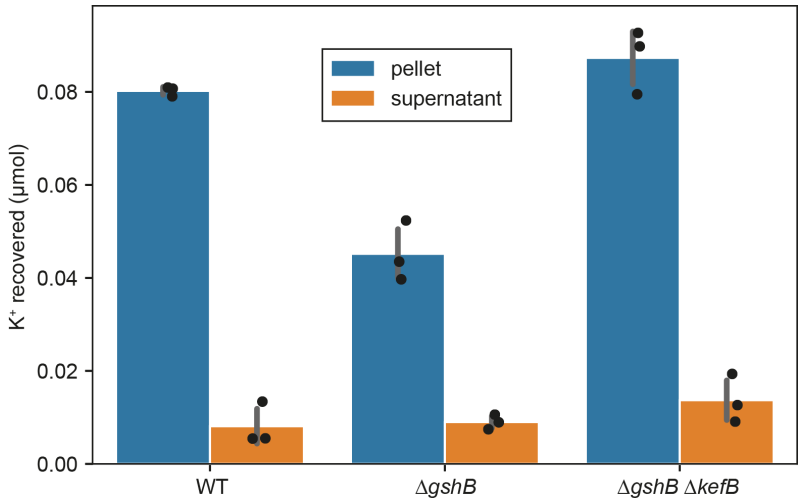
Supplementary Figure 18. Cell body length distribution of Δ*gshB* suppressor strains. (a) Flow cytometry analysis of PYE grown wild-type, Δ*gshB* and Δ*gshB* suppressor strains. Relative DNA content per estimated cell mass normalized to wild-type is shown; mean values +/- SD of 4 replicates. Indicated statistics is based on a one-way ANOVA (Δ*gshB* vs. all other strains; P<0.0001) followed by multiple comparison analysis (Dunnett) (**b)** Estimated representative chromosome equivalents of the respective suppressor strains compared to controls (n=4 biological replicates).



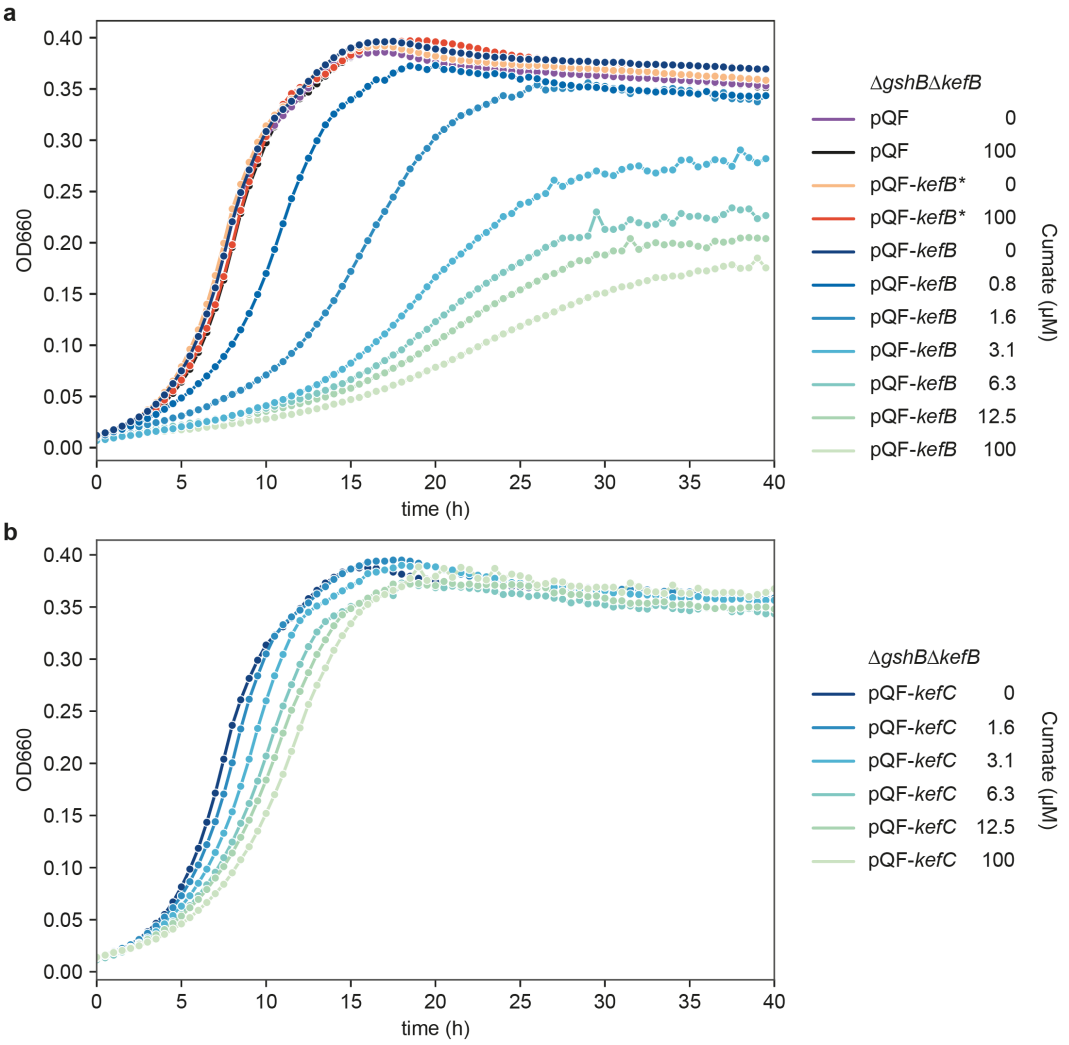
Supplementary Figure 19. Sequence alignment of CCNA_00204 (*kefB*) and CCNA_00205 with homologs of *E. coli*. **(A)** Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) sequence alignment of *C. crescentus* KefB (top) compared to *E. coli* (bottom). The conserved glutamine residue Q94/Q92 (*E. coli*/*C. crescentus*) is highlighted in red. Proteins have 43% sequence identity. **(B)** The KefB Q92 residue is highly conserved. A Weblogo 3 (http://weblogo.threeplusone.com/create.cgi) based representation of a multiple sequence alignment of KefB is shown. The multiple sequence alignment was generated in Geneious 10.2.6 (Biomatters Ltd.) using the Geneious Alignment tool with default settings on 250 sequences (Supplementary Table 11) retrieved from a standard Blastp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) using CCNA_00204 as a query (Sep 11 2019) **(C)** Clustal Omega analysis of *E. coli* KefG (top) compared to CCNA_00205 (bottom). Proteins have 44% sequence identity. Default settings of respective sequence alignment tools were used.



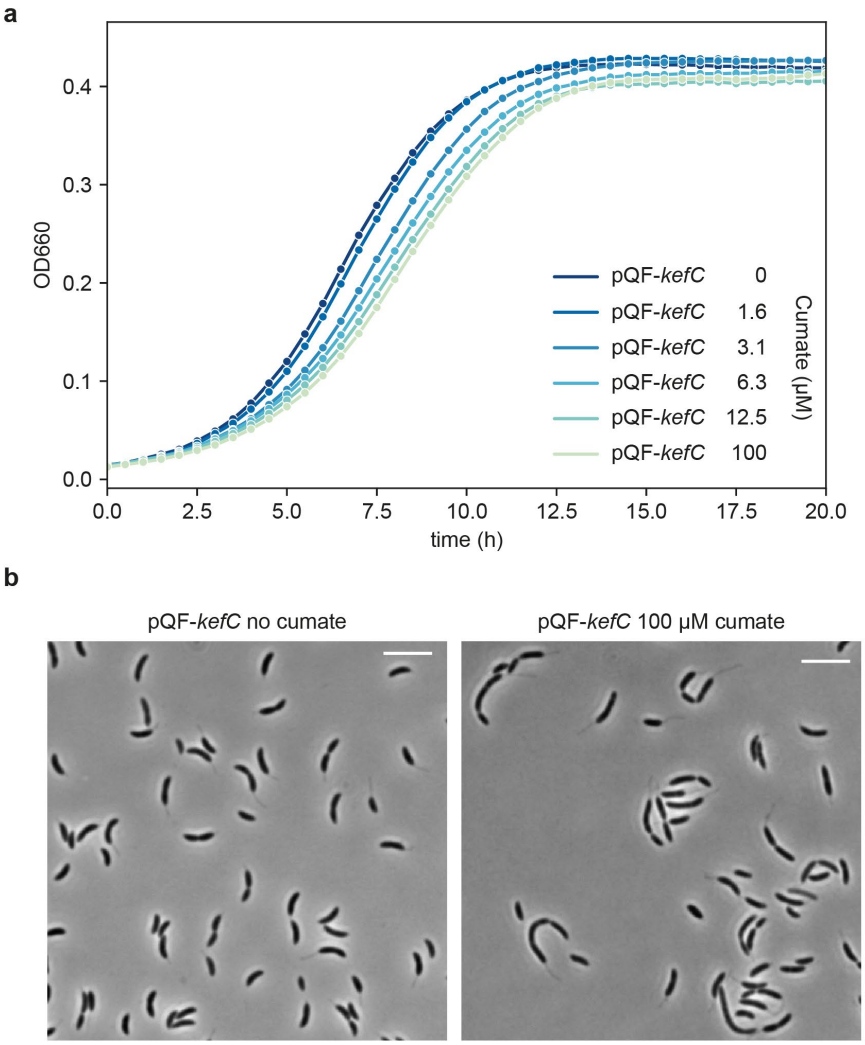
Supplementary Figure 20. H_2O_2 sensitivity of glutathione deficient strains. Serial dilutions of wild-type, Δ*gshB* and Δ*gshB*/Δ*kefB* cells spotted on PYE alone (left), or supplemented with 10 µM (middle) and 25 µM (right) H_2O_2. Note that Δ*gshB*Δ*kefB* cells are hypersensitive to H_2O_2. Experiment was performed in technical duplicates with similar results.



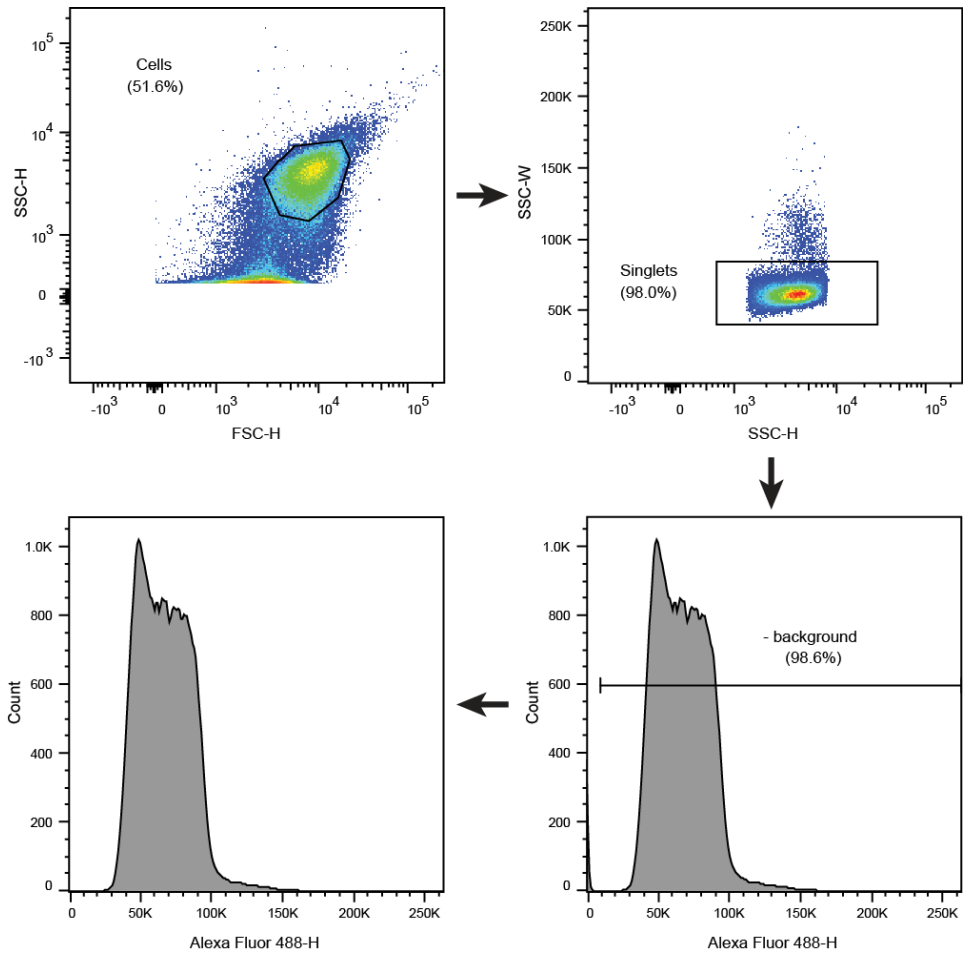
Supplementary Figure 21. Potassium recovered from cell pellets and corresponding residual leakage to supernatant. The potential for additional export or leakage from *C. crescentus* cells suspended in purified water was assayed. For details, see Methods section. Cells were washed and kept in purified water (12 min) and subsequently collected (3 min). Shown are the amount of potassium retrieved from the extracted cell pellets (blue) compared to the corresponding supernatants (orange). Note that both pellet and supernatant were subjected to the same dilutions. Data indicate that independent of the genotype, only minimal additional K leakage occurs, and that K levels in Δ*gshB* cells have already reached a plateau in the growth media. Note that data from "pellets" was used to calculate the intracellular K concentrations shown in Figure 5i. Mean values +/- SD of n=3 biological replicates per condition.



Supplementary Figure 22. Dose-response dependency of *kefB* and *kefC* induction in *ΔgshBΔkefB* strains. Growth curves of *ΔgshBΔkefB* strains harboring various complementation plasmids; induction of *kefB* alleles (top) or *kefC* (bottom) with indicated concentrations of cumate during growth on PYE media. **(a)** Expression of wild-type KefB leads to concentration dependent growth defects; induction of the loss of function allele *kefB*-Q92L (*kefB**) and the empty plasmid control has no effect. **(b)** Expression of native KefC affects growth, but not to the same extend as KefB. pQF, empty plasmid; pQF-*kefB*, wild-type *kefB*; pQF-*kefB**, mutated *kefB*-Q92L, pQF-*kefC*, wild-type *kefC*. Depicted are mean values of n=2 biological replicates.



Supplementary Figure 23. Overexpression of *kefC* leads to mild growth and morphology defects. **(a)** Growth curves of *C. crescentus* expressing *kefC* from a cumate inducible promoter. Growth on PYE is mildly impaired in a cumate concentration dependent manner. Depicted are mean values of n=2 biological replicates. **(b)** *kefC* overexpression leads to cell filamentation. The same strain as in (a) was grown exponentially for 20 hours either with fully induced kefC (right panel; 100 µM cumate) or without induction (left panel). Experiment was performed twice with similar results. Microscopy scale bar is 5 µm.



Supplementary Figure 24. General gating strategy for flow cytometry. Numbers in brackets indicate percent of events of the total population included in the gate. Gates were identical for all samples within one set of experiments and the example shown represents wild-type cells as depicted in Supplementary Fig. 17.

Supplementary Methods

A workflow for untargeted isotope dilution mass spectrometry (IDMS) feature extraction.

For each metabolite a pair of a natural labeled (12C) and a uniformly ^{13}C labeled (13C) isotopologue is present in the same sample, since the sample is a mixture of respective cell extracts. As unlabeled and labeled isotopologues have the same physicochemical properties, the corresponding LC-MS peaks will co-elute and the m/z difference equals $\Delta \mathrm{m/z} = n \cdot ({}^{13}C - {}^{12}C) \cdot z$, where n is the number of the metabolite carbon atoms, ^{13}C and ^{12}C are the isotope masses, and z is the charge state of the ions. The charge state was determined from the m/z difference of the neighboring isotopologue peaks M_0 and M_1 of the 12C sample and/or M_{UL-1} and M_{UL} of the 13C sample because M_1 and M_{UL-1} are the second most abundant isotopologues in the context of small metabolites containing the elements C, H, N, O, P, S with m < 1000 Da. Since the nominal mass difference between neighboring isotopologues equals 1, a nominal m/z difference of 1 results a charge state of 1 (a difference of 0.5 a charge state of 2, etc.). Note that untargeted metabolite extraction is not only based on analysis of an IDMS sample, but also on individual LC-MS measurements of the corresponding (12C) and the (13C) extracts, respectively. Separate analysis allows direct assignment of isotopologue identity ((12C), (13C)) in the IDMS sample. Overall, this approach significantly reduces the number of false positive metabolites and enhances metabolite annotation, as the number carbon atoms can additionally be determined from the isotopologue mass distances. To apply the workflow, metabolites were extracted from cells cultivated on natural labeled carbon source (12C) and on uniformly ^{13}C labeled carbon source (13C), respectively. Subsequently each cell extract as well as a 50: 50 mixture of both extracts (Mix) were analyzed by LC-HRMS. LC-MS data were converted into mzML files and data were analyzed with the workflow. For a more detailed workflow description, see explanations and figure below.

A. Building IDMS Pair Tables

Peak detection: For peak detection FeatureFinderMetabo tool from openMS mass spectrometry algorithm library was applied. Parameter settings are stated in Supplementary Table 9. Peak intensity threshold values were automatically estimated based on spectral signal to noise estimation over all measured spectra. To this end, baseline and noise level were determined assuming that the baseline should contain the highest point density[11]. Subsequently a histogram of all intensities over all spectra was built and the bin width of the histogram was determined by the rule of Freedman und Diaconis[12]. The mean intensity of the bin with highest abundance determines the baseline value and the full width half maximum (fwhm) of the histogram reflects the noise (more precisely fwhm / sigma (2.54)). Remark: Peak detection of the Mix sample is only required when retention time alignment is performed.

Retention time alignment. Retention time (RT) alignment was performed with an adapted algorithm of open MS pose clustering align. Adaptation of the original algorithm was necessary since it uses LC-MS peaks present in all samples, and 12C and 13C samples share only few m/z peaks due to the mass shift originating from 13C incorporation. To this end m/z peaks of the Mix sample were grouped using a previously described feature grouping algorithm which models relations between co-eluting peaks using a labeled graph whereby the label attached to an edge represents an isotopic shift explaining the mass difference between two edges. Such groups allowed mapping of m/z peaks of 12C and 13C samples where the m/z differences could be explained by 13C incorporation and hence, retention time alignment was applied to m/z groups rather to the m/z peaks themselves. To improve alignment quality, only groups of the Mix sample belonging to the top 50 percentile (in terms of total grouped peaks area) were selected for mapping procedure.

Identifying 12C and 13C peaks. 12C peaks terms peaks that are only present in the 12C sample or where the peak area was at least three times higher when compared to the same peak detected in the 13C sample. Analogous, 13C peaks in the 13C sample were determined. Subsequently corresponding 12C and 13C peaks were mapped in the Mix sample by a peak integration based targeted extraction method which adapts the RT window of each individual peak relative to a corresponding reference peak (12C

or 13C) within a predefined maximal RT window shift range. Notably, all reference peaks present in the predefined RT window range are taken into account during peak extraction, avoiding mismatches like e.g. changing the positions of partially separates peaks in the mix sample.

Finding IDMS pairs. Similar to the feature grouping applied for retention time alignment, potential IDMS pairs can be identified using labeled graph whereby the label attached to an edge represents a ^{13}C isotopic shift or a multitude explaining the mass difference between two edges. Again, resulting feature groups are computed as connected components of this graph. However, since 12C and 13C peaks are already identified prior to grouping (see above), directed instead of undirected graphs can be used which not only halves calculation time but also avoids false positives due to the reduced number of possible matches. Moreover, each 12C peak allows grouping of all corresponding 13C peaks and vice versa, and monoisotopic peak M_0 and unformly 13C labeled peak M_{UL} of each IDMS pair can now be assigned straight forward in each subgroup. Subgroups also allowed assigning charge state z if more than one isotopologue per subgroup was detected. In case of subgroups with a single peak, a local isotopologue feature finder script was used to extract potential low abundant, neighbored isotopologue peaks to enable charge state assignment. In case of ambiguous grouping or charge state assignment (e.g. more than 2 subgroups were grouped or grouped subgroups had different charge states) corresponding pairs were selected for manual curation of IDMS pair.

B. Building IDMS Metabolites Table

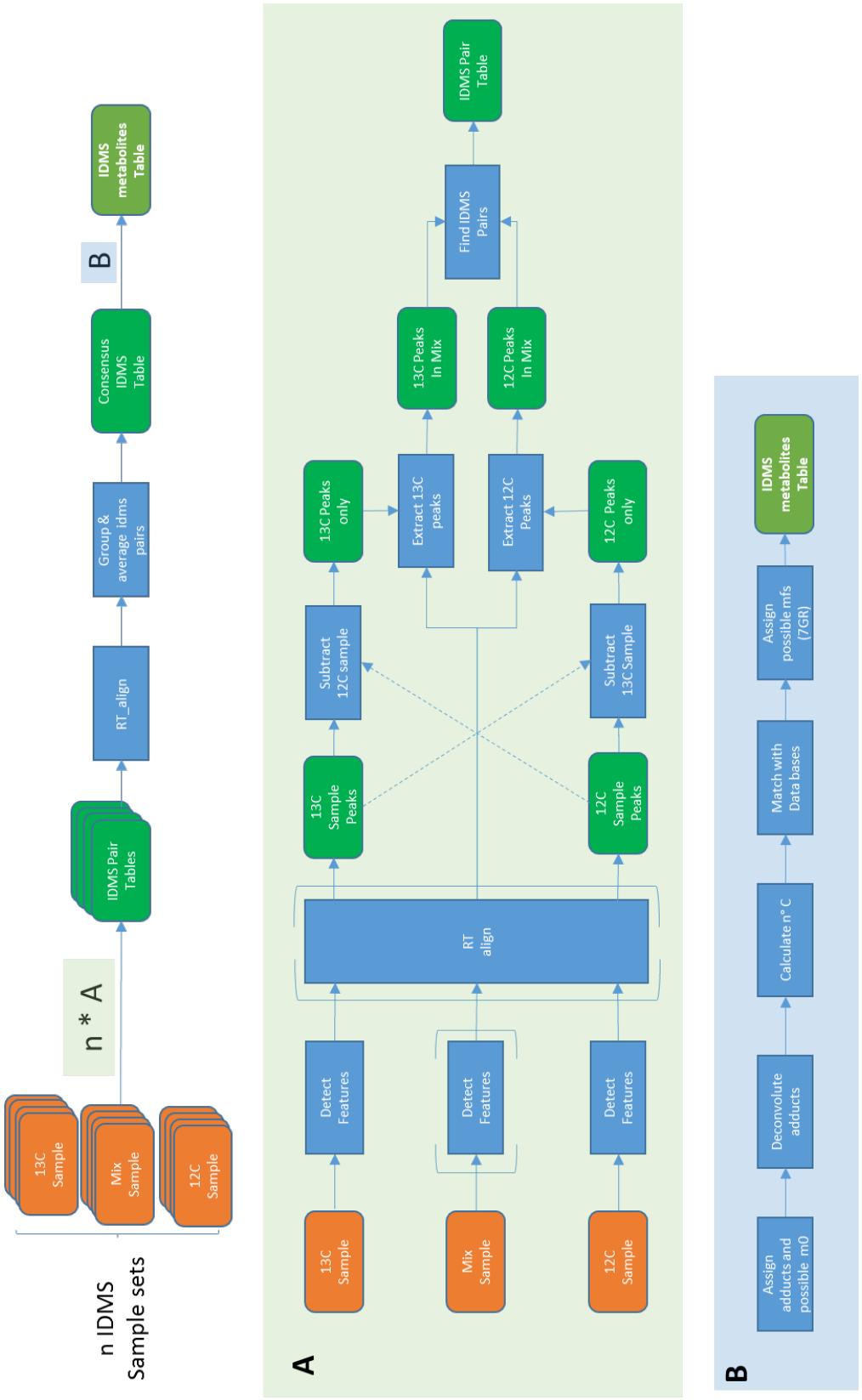
Building Reference Table. Applied parameters are stated in Supplementary Table 10. Initially, a second RT alignment was carried out to compensate for batch specific differences (each biological replicate was measured in a different LC-MS acquisition batch). Note that batch specific RT differences are in general significantly greater than those observed within an acquisition batch. Therefore, inter-batch RT alignment was performed afterwards since intra batch alignment allowed mapping of 12C and 13C peaks in mix sample with smaller RT windows than inter-batch RT alignment, which reduced mismatches during IDMS mapping process. Subsequently, all IDMS pairs, present in at least 2/3 of all IDMS tables were selected for the reference table. Peaks were set identical based on m/z on RT tolerances. Whereas m/z tolerances were predefined, RT tolerances were estimated based on observed RT variability of LC-MS peaks present in all samples and which were unique in a predefined RT window width (+/- maximal allowed RT shift) of the corresponding extracted ion chromatogram. For further processing mean m/z and RT values were calculated.

Adduct grouping, mass assignment and deconvolution. For adduct grouping and mass assignment a modified version of previously described adduct grouping algorithm was applied. The modified version is now also able to assign adducts in the positive mode. From each adduct group only the most abundant adduct was kept.

Identification. Potential metabolites were assigned to IDMS pairs by exact mass (0.003 Da) and by the number of carbon atoms deduced from the mass difference between M_0 and M_{UL} isotopologues. In case of adduct assignment missing, protonated or deprotonated ions were assumed depending on the ionization mode. In case no charge state could be assigned the charge states z = 1, 2, 3 were assumed and for each charge state the corresponding number of possible carbon atoms was calculated, respectively. Results were matched with KEGG data base[5] and with a merged database consisting of an *E. coli* model (iML1515) merged with *C. crescentus* Seed model (Seed190650_1). Merging was required since seed model database was not manually curated and known core metabolites were missing in the model. Resulting annotations by database matching are listed in Supplementary Tables 1, 2. Cell cycle dependent candidate features were further annotated using analytical standards (matching retention time), and/or tandem mass spectrometry followed by spectral matching using the *dereplication workflow* of GNPS (https://gnps.ucsd.edu/)[13]. Respective MS level 2 spectra were averaged over the elution window size of the LC-MS peak and 20 most intense spectral signals were picked for further analysis. All available spectral libraries from GNPS were initially queried with a precursor ion mass tolerance of 1 Da, a fragment ion mass tolerance of 0.5, a minimum of 5 matched peaks, and a minimum

cosine score threshold of 0.7. If no match was acquired, search was extended to analogs with a maximum parent mass difference of 250 Da. Identification results are shown in Supplementary Tables 3, 4. MS2 files are provided in Supplementary Data.

Outline of the above described workflow for untargeted isotope dilution mass spectrometry feature extraction:



Supplementary References

1 Letunic, I., Yamada, T., Kanehisa, M. & Bork, P. iPath: interactive exploration of biochemical pathways and networks. *Trends Biochem Sci* **33**, 101-103 (2008).

2 Wohlgemuth, G. *et al.* The chemical translation service--a web-based tool to improve standardization of metabolomic reports. *Bioinformatics* **26**, 2647-2648 (2010).

3 Feunang, Y. D. *et al.* ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J Cheminformatics* **8** (2016).

4 Hartl, J., Kiefer, P., Meyer, F. & Vorholt, J. A. Longevity of major coenzymes allows minimal *de novo* synthesis in microorganisms. *Nat Microbiol* **2**, 17073 (2017).

5 Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**, D199-205 (2014).

6 Schrader, J. M. *et al.* Dynamic translation regulation in *Caulobacter* cell cycle control. *P Natl Acad Sci USA* **113**, E6859-E6867 (2016).

7 Feist, A. M. *et al.* A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* **3**, 121 (2007).

8 Peabody, M. A. *et al.* PSORTdb: expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res* **44**, D663-668 (2016).

9 Lasker, K. *et al.* CauloBrowser: A systems biology resource for *Caulobacter crescentus*. *Nucleic Acids Res* **44**, D640-645 (2016).

10 Meisenzahl, A. C., Shapiro, L. & Jenal, U. Isolation and characterization of a xylose-dependent promoter from *Caulobacter crescentus*. *J Bacteriol* **179**, 592-600 (1997).

11 Horn, D. M., Zubarev, R. A. & McLafferty, F. W. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *Journal of the American Society for Mass Spectrometry* **11**, 320-332 (2000).

12 Freedman, D. & Diaconis, P. On the histogram as a density estimator - L2 Theory. *Z Wahrscheinlichkeit* **57**, 453-476 (1981).

13 Wang, M. X. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **34**, 828-837 (2016).