UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

# Determination of protein synthesis on a proteomic scale

Kramer, G.

**Publication date**
2011

Link to publication

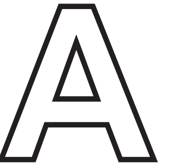
**Citation for published version (APA):**
Kramer, G. (2011). *Determination of protein synthesis on a proteomic scale*. [Thesis, fully internal, Universiteit van Amsterdam].

**General rights**
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:16 Apr 2026

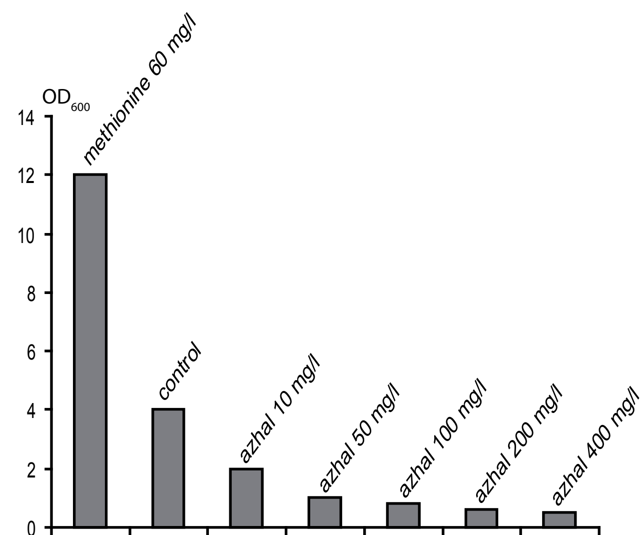


# Addenda

# Addendum Chapter 2

## Experimental Procedures

*Cell culture*— *S. cerevisiae* strain BY4741 was grown in YPD medium aerobically at 30 °C For growth experiments cells were transferred to minimal medium (276) containing 60 mg/l of all natural amino acids. Cells were inoculated at $OD_{600}$ 0.1 and allowed to grow into exponential phase before being harvested at $OD_{600}$ 1.0, by spinning down the cells for 10 minutes at 4500 rpm and 4 °C. Cells were then washed twice by resuspending the cell pellet in sterile minimal medium without additives (followed by centrifugation) to eliminate traces of methionine. After washing, cells were transferred (at $OD_{600}$ 0.1) to minimal medium in which the methionine was replaced by azhal and cells were allowed to resume growth aerobically at 30 °C. Cells were allowed to grow overnight with varying methionine or azhal conditions as indicated, before the $OD_{600}$ reached overnight was measured.




Supplemental Fig.1. **Inhibition of growth of *S. cerevisiae* by addition of azhal.** Yield of overnight growth of *S. cerevisiae* is compared for different concentrations of azhal in minimal medium compared to methionine or control (neither azhal or methionine added). Clearly addition of azhal to minimal medium inhibits growth yield below that of control and increase of azhal concentration results in a dose dependent lowering of growth yield after overnight culture

# Addendum Chapter 4

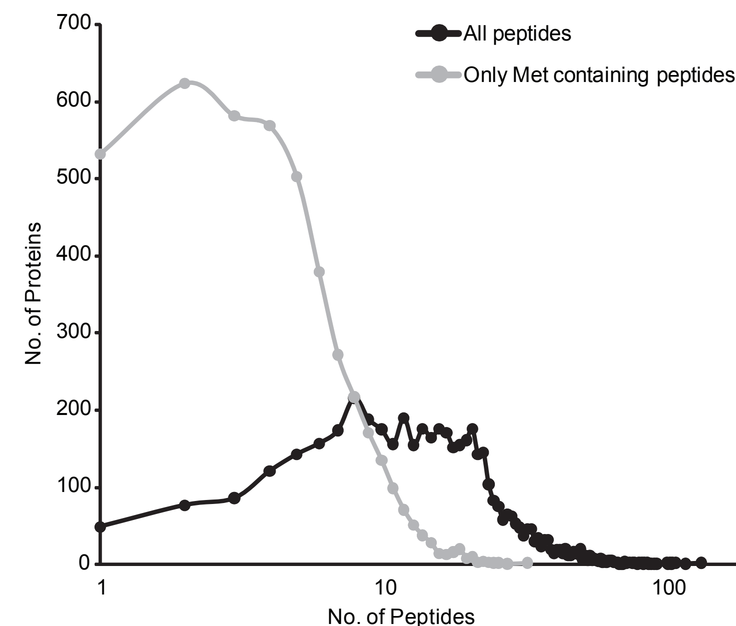
## Experimental Procedures

*In silico digests and residue content determinations*— To determine the number of proteins that can theoretically be covered by methionine containing peptides and the number of proteins that are available within a specific mass window, an *E. coli* K12 proteome database (4691 protein entries, release 2010_05, April 20, 2010, Uniprot consortium) was digested in-silico using the DBTOOLKIT version 1.4 (277). Settings: trypsin digestion, allowing 0 missed cleavages and setting the mass window for peptides from 500 to 3800 amu. Missed cleavages were not allowed as examination of MASCOT identifications of various large datasets revealed the number of assignments involving missed cleavages to be only ~2%, while the mass window was chosen because masses of MASCOT identified peptides from the QSTAR-XL fell within this range. Subsequently the digest-database was exported in FASTA-format and a filtered digest-database with only methionine containing peptides was created, using DBTOOLKIT, and exported as well.

Both filtered and unfiltered peptide sets were imported for further analysis in Excel (2007, Microsoft corporation, Redmond, USA) containing the freely available ASAP-utilities add-on (A must in every office B.V., Zwolle, The Netherlands). Because all N-terminal peptides contain a methionine and a substantial part of proteins in *E. coli* are processed modifying the N-terminus often resulting in removal of the N-terminal methionine residue, we first removed all N-terminal peptides from both digest databases created in DBTOOLKIT. Subsequently by using the pivot table function to count the number of peptides per protein and the number of proteins identified by a certain number of peptides a distribution for both sets was created. In addition the total percentage of entries theoretically covered by de different peptide sets were calculated as well by counting the number of unique entries identified by the two digest databases. To determine the methionine residue contents of identified proteins, their sequences were retrieved from Uniprot in FASTA-format (www.uniprot.org) and imported into Excel. The total number of residues and the number of methionine content were counted using the ASAPCOUNTCHAR feature of ASAP-utilities, these were used to calculate the average methionine content of the identified proteins.








Supplemental Fig.1. **Number of peptides per protein after an in silico digest of the *E. coli* proteome database.** The number of proteins for which a certain number of peptides are found within the mass range of 500-3800 amu after an in-silico digest, not allowing missed cleavages, of the *E. coli* proteome database are shown. The light grey line shows the distribution for all peptides, while the dark grey line shows the distribution for peptides containing methionine residues only. From the 4691 protein entries in the *E. coli* database 4677 (99.7%) are theoretically identified when considering all peptides while 4369 (93.1%) are theoretically identified when only considering peptides containing a methionine residue. Obviously, although the total number of proteins covered is not much reduced when only employing methionine containing peptides, the number of peptides theoretically detectable per protein declines sharply as the total number of methionine containing peptides (22585) is only 25.6% of the total number of tryptic peptides (88287).

# Addendum Chapter 5

## Kinetic Model

To extract a measure for the protein translation rate from our data set we constructed a kinetic model that predicts the number of newly synthesized proteins in the course of an azhal pulse-labelling experiment. Pulse-labelling is carried out in the exponential growth phase of *E. coli* cells in a culture flask. We assume that during the exponential growth phase the number of proteins increases proportional to the total of cellular mass and that protein synthesis ($k_{syn}$) and degradation ($k_{deg}$) rates are first-order with the number of proteins as described by Mosteller and Goldstein (278), in equation (a) where F(t) represents the time dependent increase in cellular mass.

$$\frac{\partial P_{total}}{\partial t} = k_{syn}F(t) - k_{\deg}P_{total} \qquad \text{(a)}$$

At time point t is zero minutes we have only unlabeled proteins $P_{old}$. Post azhal-labelling we have at t = 10 minutes the partly degraded old protein population $P_{old}$ and we have newly synthesized azhal-labelled proteins $P_{new}$. The degradation-rate of the old unlabeled population $P_{old}$ is given by differential equation (b).

$$\frac{\partial P_{old}}{\partial t} = -k_{\deg}P_{old} \qquad \text{(b)}$$

During labelling azhal-containing new proteins $P_{new}$ are formed with a first-order rate constant $k_{syn}$ and degrade with rate constant $k_{deg}$. We assume that the rate constant $k_{deg}$ is not influenced by azhal replacing methionine. From literature it is known that in vivo degradation-rates of some amino acid analogue-labelled proteins can differ from their natural counterparts (279, 280). However compared to these analogues, *E. coli* can sustain normal growth longer with azhal. In addition, there is evidence for unperturbed protein processing of the N-terminal azhal residue in *E. coli* (175) and for normal localization and folding of azhal-labelled proteins (148-151). Most importantly, the relative change in average synthesis-rate upon anaerobiosis found by Smith *et al.* (124) using radiolabelling for PflB is similar to that found by azhal-labelling, and similar results were found by azhal-labelling compared to radiolabelling data during heat shock, as published before (170). Altogether this suggests that the effect of azhal-labelling on protein stability, on the time-scale employed for pulse-labelling, is negligible.

We also assume that the average $k_{deg}$ is similar under aerobic and anaerobic conditions in our experiment. Evidence that this is a reasonable assumption comes from the quantitation of only methionine containing peptides. As mentioned above these peptides





represent only pre-existing material made before the pulse-labelling period. A quantitation of the protein levels derived from methionine peptide ratios is a measure for the relative stability of these pre-existing protein during the pulse-labelling time under anaerobic versus aerobic conditions. In supplemental Figure 3 the relative ratio of pre-existing protein material is presented, and it is evident that no large stability differences occur during pulse-labelling as the ratios do not differ significantly from 1 (0 on $^2$log scale) for the vast majority of proteins.

Finally we assume that biomass is equal to total protein or F(t) = $P_{total}$(t) in equation (a) and as $P_{total}(t) = P_{old} + P_{new}$ and the degradation of pre-existing proteins does not influence new protein formation, the rate of change of azhal-containing proteins $P_{new}$ can be defined by differential equation (c).

$$\frac{\partial P_{new}}{\partial t} = k_{syn}P_{old} + (k_{syn} - k_{\deg})P_{new} \qquad \text{(c)}$$

Integration of the set of equations (b) and (c) is straightforward and yields the time-dependant functions $P_{old}$(t) and $P_{new}$(t). Boundary value conditions are at t = 0: $P_{old}(t) = P_{old}(0)$ and $P_{new}(t) = P_{new}(0) = 0$.

$$P_{old}(t) = P_{old}(0)e^{-k_{\deg}t} \qquad \text{(d)}$$

$$P_{new}(t) = P_{old}(0)[e^{(k_{syn}-k_{\deg})t} - e^{-k_{\deg}t}] \qquad \text{(e)}$$

And summation of (d) and (e)

$$P_{total}(t) = P_{old}(t) + P_{new}(t) = P_{old}(0)e^{(k_{syn}-k_{\deg})t} \qquad \text{(f)}$$

In our experiments we obtain two ratios. The first is the ratio between the number of newly synthesized proteins under either aerobiosis or anaerobiosis during the pulse-labelling time interval going from one environmental condition to the other. This ratio is derived from the iTRAQ reporter ions of azhal- containing peptides that are collected off-diagonal in our COFRADIC setup.

The predicted ratio of newly synthesized proteins under anaerobiosis $P_{new,\ anaerobiosis}$(t) and aerobiosis $P_{new,\ aerobiosis}$(t) is

$$R_{newly_synthesized} = \frac{P_{new,anaerobiosis}(t)}{P_{new,aerobiosis}(t)} = \frac{P_{old,anaerobiosis}(0)[e^{(k_{syn,anaerobiosis}-k_{\deg})t} - e^{-k_{\deg}t}]}{P_{old,aerobiosis}(0)e^{(k_{syn,aerobiosis}-k_{\deg,})t} - e^{-k_{\deg}t}]} \qquad \text{(g)}$$

Which reduces to

$$R_{newly_synthesized} = \frac{P_{old,anaerobiosis}}{P_{old,aerobiosis}} \frac{e^{k_{syn,anaerobiosis}t} - 1}{e^{k_{syn,aerobiosis}t} - 1} \quad (h)$$

For small values of $k_{syn}t$ as is the case in our study we can approximate the exponential function $e^x$ by the first terms $1 + x$ of its Taylor expansion and the predicted ratio of newly synthesized proteins $P_{new}(t)$ becomes

$$R_{newly_synthesized} = \frac{P_{new,anaerobiosis}(t)}{P_{new,aerobiosis}(t)} = \frac{P_{old,anaerobiosis}(0)}{P_{old,aerobiosis}(0)} \frac{k_{syn,anaerobiosis}}{k_{syn,aerobiosis}} \quad (i)$$

For small values of $k_{syn}t$, the measured ratio of azhal-containing peptides is predicted by the ratio of synthesis-rate constants in the aerobic and anaerobic experiment. Comparison with the data directly provides insight in differences of protein synthesis or translation rates between experimental conditions.

The second experimental ratio that we extract from our data set is the copy number ratio of total protein level at the end of the labelling time between two environmental conditions. This ratio is derived from the peptide ratios of peptides that do not contain azhal or methionine. Azhal peptides are excluded as these represent exclusively newly synthesized material, while methionine containing peptides are excluded because these represent only pre-existing material. Newly formed peptides that contain azhal are found in off-diagonal fractions. Peptides that do not contain azhal or methionine are found in both on-diagonal and off-diagonal fractions; these peptides reflect the total protein level in the cell as they are made up of both pre-existing as well as new material. The non azhal/methionine containing peptide copy number equals the summation of $P_{old}(t)$ and $P_{new}(t)$ in our kinetic model.

The predicted ratio $R_{total_level}$ for non azhal/methionine containing peptides is

$$R_{total_level} = \frac{P_{old,anaerobiosis}(t) + P_{new,anaerobiosis}(t)}{P_{old,aerobiosis}(t) + P_{new,aerobiosis}(t)} = \frac{P_{old,anaerobiosis}(0)}{P_{old,aerobiosis}(0)} \frac{e^{(k_{syn,anaerobiosis} - k_{\deg})t}}{e^{(k_{syn,aerobiosis} - k_{\deg})t}} \quad (j)$$

and after Taylor polynomial expansion

$$R_{total_level} = \frac{P_{old,anaerobiosis}(0)}{P_{old,aerobiosis}(0)} \frac{[1 + (k_{syn,anaerobiosis} - k_{\deg})t]}{[1 + (k_{syn,aerobiosis} - k_{\deg})t]} \quad (k)$$

This total protein ratio reflects the overall protein expression ratio between two states. In our model we have two experimental ratios that are predicted by three reaction rate constants. To compare predicted ratios with measured ones we make the following assumptions. The aerobic culture is in the exponential growth phase. For an estimate of $k_{deg}$ we assume a protein half-life of 120 minutes. From literature it is evident that with regard to degradation-rates we broadly find two protein populations. A small rapidly degrading population with half-lives lower than our pulse time window and a more slowly degrading population with half-lives of at least a few hours up to more than 23 hours (110, 111). In our kinetic model we use the lower boundary of protein half-life $t_{½} = 120$ minutes. For larger protein half-lives the effect of degradation on kinetics can be neglected. For $t_{½} = 120$ minutes we calculate that $k_{deg} = 5.78 \times 10^{-3}$ min$^{-1}$. Next, we have to estimate $k_{syn.aerobiosis}$. During ten minutes labelling time the optical density (an approximation of biomass) of the culture increases with 6.25 %; assuming total protein content scales linearly with biomass its increase is also 6.25 %. Given our estimate of $k_{deg}$ of $5.78 \times 10^{-3}$ min$^{-1}$ and a biomass increment of 6.25% we use equation (e) to calculate an average protein synthesis-rate constant $k_{syn,aerobic} = 1.2 \times 10^{-2}$ min$^{-1}$ under aerobic conditions.

For a few selected proteins with increased rates of new formation we used our model to estimate their $k_{syn,anaerobioses}$ rate constants using MATLAB (the Mathworks, Natick, USA). As an objective criterion we minimize the summation of the absolute value of the differences between the calculated ratios $R_{newly_synthesized}$ and the measured one upon variation of $k_{syn,anaerobioses}$ To correct for differences in growth following the switch from aerobiosis to anaerobiosis we set $P_{old}(0)_{anaerobiosis}$ to 106.25 and $P_{old}(0)_{aerobiosis}$ to 100. In our method we mix absolute protein amounts of the azhal labelled anaerobic and aerobic proteomes in a one to one ratio.

The results are presented in Supplemental Table I. As is seen in this table our kinetic model predicts both protein level ratios reasonably well for the given set of first-order protein synthesis and degradation-rate constants and the estimated rate constants $k_{syn,anaerobiosis}$.

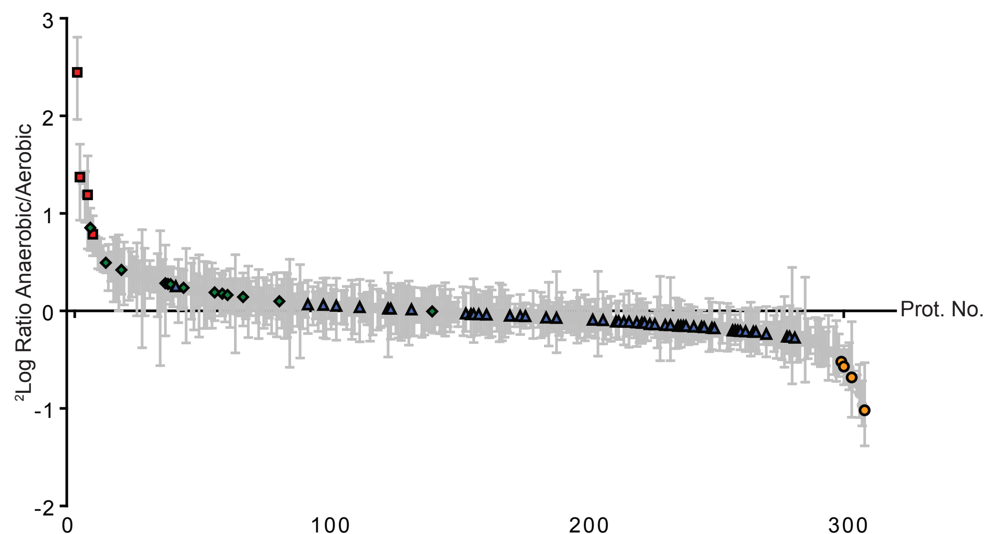
SUPPLEMENTAL TABLE I

*Predicted and measured rates of formation of selected proteins and their change in levels after a switch to an anaerobic environment in E. coli.*

| gene | \|Δratio† \| | \|Δratio‡\| | $k_{syn,\ anaerobiosis}$ (min$^{-1}$) | ratio $k_{syn,\ anaerobiosis}$ / $k_{syn,\ aerobiosis}$ |
|---|---|---|---|---|
| PfkA | 0.08 | 0.09 | $4.7 \times 10^{-2}$ | 3.9 |
| Gpml | 0.12 | 0.01 | $4.1 \times 10^{-2}$ | 3.4 |
| Pgk | 0.04 | 0.02 | $3.3 \times 10^{-2}$ | 2.8 |
| PykA | 0.09 | 0.09 | $3.1 \times 10^{-2}$ | 2.6 |
| Pgi | 0.01 | 0.06 | $3.0 \times 10^{-2}$ | 2.5 |
| GapA | 0.06 | 0.03 | $2.5 \times 10^{-2}$ | 2.1 |
| Eno | 0.09 | 0.06 | $2.4 \times 10^{-2}$ | 2.0 |
| FbaA | 0.09 | 0.04 | $2.1 \times 10^{-2}$ | 1.8 |

|Δratio†| is the absolute value of the difference between the measured and predicted ratio anaerobic/aerobic of newly formed proteins determined by quantitation of azhal-tagged proteins, |Δratio‡| is the absolute value of the difference between the measured and predicted ratio anaerobic/aerobic of protein levels determined by quantitation of non-tagged azhal peptides.

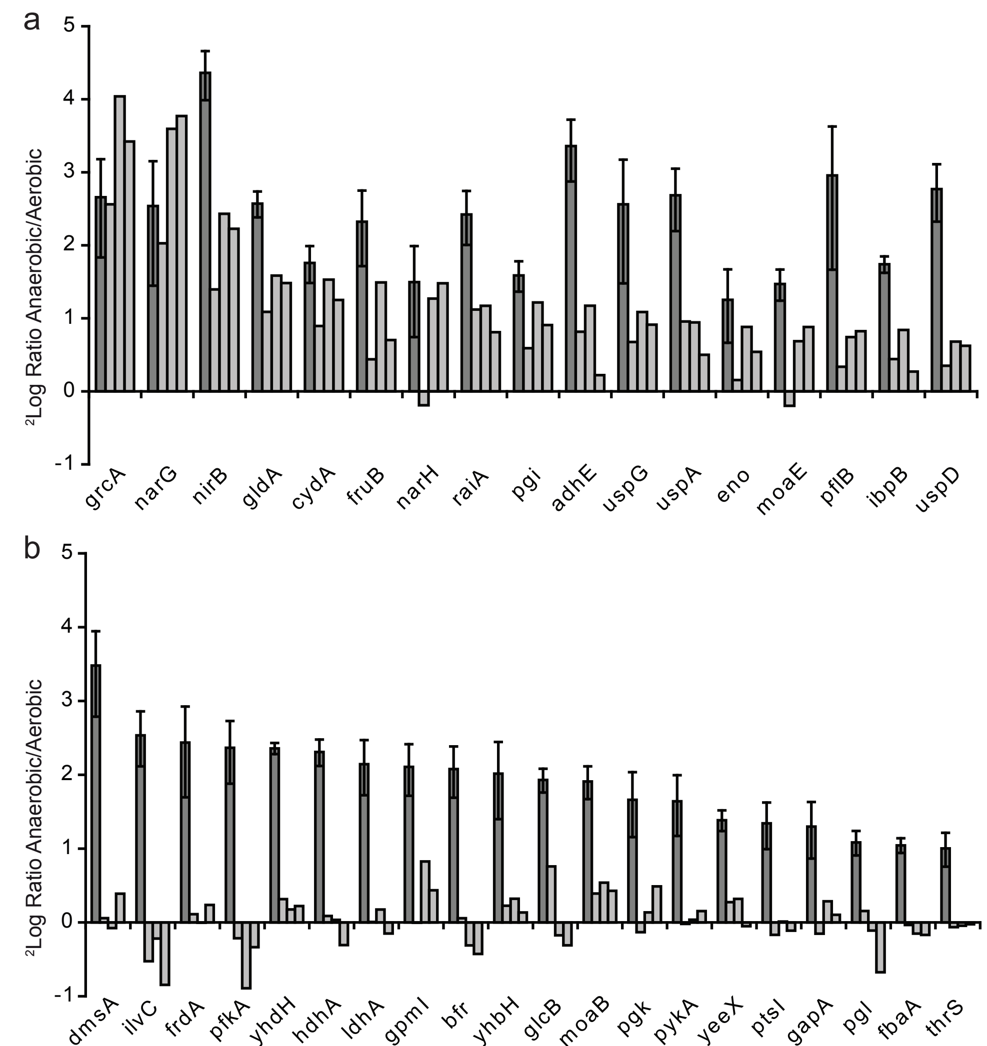
These rate constants are (as stated above) relative values that depend on initial conditions for $P_{old}(0)$, $k_{syn, aerobiosis}$ and $k_{deg}$ for both growth conditions, where $k_{syn,aerobiosis}$ is determined by the growth rate of the *E. coli* population. Although our kinetic equations provide relative rate constants for protein synthesis, the ratio $k_{syn,anaerobises}/k_{syn,aerobiosis}$ is a direct measure for regulation of protein expression by adjusting the translation rate of proteins. The calculations also show that given our premises upon transition from aerobic to anaerobic growth conditions the ratio $k_{syn,anaerobises}/k_{syn,aerobiosis}$ increases by a factor of 1.8 to 3.9 and that the rate constant for degradation ($k_{deg}$) is an order of magnitude lower than the rate constants for protein synthesis ($k_{syn}$). Consequently, increased or decreased levels of most newly synthesized proteins result from changes in translation rate.



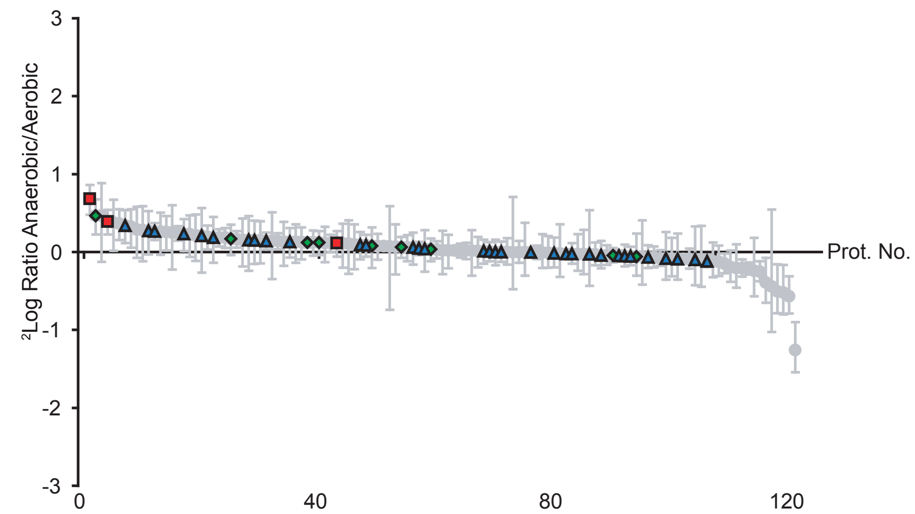

Supplemental Fig.1. **Changes in protein levels 10 minutes after the onset of anaerobiosis.** Grey dots show ²log ratio of protein levels from most up-regulated to most down-regulated under anaerobic conditions. Error bars show the ²log transformed standard deviation of the peptide ratios per protein. Red squares, levels of proteins involved in anaerobic respiration and mixed acid fermentation (adhE, pflB, grcA cydA), green diamonds, glycolytic enzymes and proteins of the PTS system (pfkA, gpmI, pgk, pykA, pgi, tpiA, gapA, eno, pykF, fbaA, gpmA, fruB and ptsI). Levels of the ribosomal proteins rpmG, rplJ, rplP, rplL, rpmD, rpmC, rpsN, rpsA, rpsO, rpsK, rplN, rpsG, rpsB, rpmA, rpsM, rplY, rplM, rpmB, rplV, rplK, rplF, rpsI, rplA, rpsP, rpsF, rplE, rplI, rpsJ, rpmF, rpsC, rpsH, rpsD, rpsE, rplR, rplO, rplW, rpsS, rplU, rpsR, rpmH, rplD, rplS, rplB, rpsQ, rpsL, rpsU, rplX, rplC, rpsT, rplQ and rplT (blue triangles). Orange dots, levels of some severe down-regulated proteins (sodA, katG, ahpC, ahpF and ompE). Listed proteins are depicted from left to right








Supplemental Fig. 2. **Relative synthesis-rates of up-regulated proteins compared with relative mRNA levels.** Proteins with a more than twofold increase of their relative synthesis-rate were compared to their mRNA levels as determined by Partridge *et al.* (267) upon micro-aerobic switch. Dark grey bars: relative ²log transformed ratio of newly synthesized proteins (as found by azhal-labeling of anaerobic over aerobically grown cultures) with their standard deviations (error bars). Light grey bars: relative transcript levels (²log-scale) at 5, 10 and 15 minutes (from left to right) in a low oxygen environment. Panel (**a**): proteins of which the sum of corresponding relative mRNAs levels at 5, 10 and 15 minutes after the onset of anaerobiosis is at least 1.4 (²log-scale). Panel (**b**): the sum of relative mRNA levels at 5, 10 and 15 minutes is 1.3 (²log-scale) or less.




Supplemental Fig. 3. **Relative stability of pre-existing protein population under anaerobic or aerobic conditions during the pulse labling period.** Grey dots show ²log ratio of relative levels of protein obtained from only methionine containing peptides ordered from highest to lowest relative ratio. Error bars show the ²log transformed standard deviation of the methionine-peptide ratios per protein. Methionine containing peptides represents protein species made before pulse-labeling commenced. The relative protein level derived from only methionine peptides ratios is an indication of the relative pre-existing protein stability under anaerobic versus aerobic conditions during the pulse-labeling. As is clear from the figure, the relative stability is the same under aerobic and anaerobic conditions as relative ratios do not differ significantly from 1 (0 on ²log scale) for the vast majority of proteins. Red squares, levels of proteins involved in anaerobic respiration and mixed acid fermentation, green diamonds, glycolytic enzymes and proteins of the PTS system Levels of the ribosomal proteins are shown by blue triangles.