Determination of protein synthesis on a proteomic scale

Kramer, G.

Publication date
2011

4

Immediate changes in both protein levels and newly synthesized proteins following a change in growth temperature in *E. coli*

Summary

Pulse-labelling with the methionine analogue azhal has been used for the identification of hundreds of newly synthesized proteins in *E. coli* on a short time-scale. However a quantitative mass-spectrometric approach is essential to enable research into how different environmental conditions affect new protein formation in *E. coli*. We applied azhal pulse-labelling in combination with iTRAQ for quantitation and COFRADIC isolation of labelled peptides to determine changes in newly synthesized proteins immediately following a heat shock in *E. coli*. In addition by extension of our analytical strategy to determine changes in total protein levels on the same time-scale stable or labile proteins can be identified.

Measurement of the relative amounts of 344 proteins newly synthesized in 15 minutes upon a switch of growth temperature from 37 ºC to 44 ºC showed that nearly 20% in- or decreased more than two-fold. Amongst the most up-regulated proteins many were chaperones and proteases, in accordance with the cells response to unfolded proteins due to heat stress. While collation of changes in protein levels with changes in newly synthesized proteins showed that the vast majority of proteins were stable, only a subset of 5 proteins (PspA, IbpB, σ^{32}, AhpC, and CysK) were found to have a higher turnover rate. Finally, comparison of our data with results from previous microarray experiments revealed the importance of regulation of gene expression at the level of transcription of the most elevated proteins under heat shock conditions and enabled identification of several candidate genes whose expression may predominantly be regulated at the level of translation. This work demonstrates for the first time the use of a bioorthogonal amino acid for proteome-wide detection of changes in the amounts of proteins synthesized during a brief period upon variations in cellular growth conditions.

Introduction

Changes in protein levels in cells during adaptation from one environmental condition to another may be regulated both by transcription and translation. However, little is known about the contribution of translational regulation. This requires information about genome-wide changes in translation rates and mRNA levels upon environmental perturbations. Measuring protein synthesis and degradation-rates on a proteomic scale is an important step towards modelling the kinetics in complicated cellular response networks. Pulse-labelling with the methionine analogue azhal enables proteome-wide identification of proteins that are newly formed during the pulse-labelling period, while different enrichment schemes can facilitate short pulse-labelling times, increasing temporal resolution over stable-isotope based pulse-labelling approaches (see *Chapter 1 and 3*). In *Chapter 3* we demonstrated how a selective reaction against the azide-moiety of azhal in combination with an enrichment scheme based on an induced retention-time shift enables a short pulse-labelling time window of only 15 minutes in *E. coli*. However to enable investigations into how different environmental conditions affect the formation of new proteins on a proteome-wide scale, a quantitative mass spectrometric approach is essential to compare differences in the synthesis of new proteins between different samples. In this chapter we show that through combining azhal pulse-labelling with relative quantitation by means of iTRAQ labelling (183) to compare different samples, relative changes in newly synthesized proteins after a 15 minute pulse-labelling period can be measured.

In addition we show that the approach is easily extended to also determine changes in total protein levels on the same time-scale as changes in newly synthesized proteins. Total protein levels are determined by reporter ions from peptides that do not contain azhal or methionine, as these represent the sum amount of newly synthesized and pre-existing material after the labelling period. Comparison of relative changes in total protein levels to those in newly synthesized proteins allows identification of stable and labile proteins for those proteins that undergo significant regulation upon heat shock. Information on the stability of proteins is an interesting extension. Although there is data that the majority of abundant proteins in *E. coli* cells is relatively stable (109) and has a half-life of at least a few hours (110, 111), while a small pool is rapidly degraded (90), no proteome-wide information on protein turnover for individual proteins is available.

To study relative changes in newly synthesized proteins labelled with azhal and the effects on total protein levels, we chose the heat shock response in *E. coli,* as it has been studied in detail before it therefore seems a good system to validate our approach (48, 103-105). Heat shock in *E. coli* is defined as the cellular response to an increase in growth temperature, and is accompanied by the upregulation of a defined set of 'heat shock' proteins. The increase of heat shock proteins is controlled by *rpoH* which encodes the heat shock transcription sigma factor σ^{32} (226-228, 233). Nearly 100 genes have been identified to be part of the σ^{32}-regulon (229, 230), of which some 18 genes encode chaperones and proteases. The upregulation of both chaperones and proteases seems to be aimed at restoring impaired

protein folding at higher temperature and to degrade misfolded proteins. The intracellular concentration of σ^{32} shows a rapid transient increase upon heat shock followed by a decrease to reach a new steady state within 10-15 minutes (50, 106, 107). This affects transcription of σ^{32} regulated genes concomitantly and also induces expression of heat shock proteins within this time-frame (106, 107). In addition the transcription of a large number of other genes is rapidly up- or down-regulated upon heat shock (234, 235).

The early changes in protein synthesis, upon the change in growth temperature, were examined and quantitative data is presented for 344 newly synthesized proteins. In addition the changes in total protein levels during the initial 15 minutes after the temperature switch were quantified for 292 proteins. Comparison of the changes in newly synthesized proteins with relative mRNA levels enables assessment of the separate contributions of transcription and translation to the regulation of gene expression and allows identification of candidate genes that could be subject to post-transcriptional regulation. This demonstrates for the first time a proteome-wide, bioorthogonal approach for relative quantitation of proteins synthesized during a small time-window, upon a change in growth conditions.

Results

Quantitation of newly synthesized proteins induced by heat shock— To study the immediate changes in newly synthesized proteins following heat shock, quantitative azhal pulse-labelling to measure changes in the amount of newly synthesized protein after an up-shift in growth temperature from 37 ºC to 44 ºC. Two cultures (biological replicates) were grown into exponential phase before being harvested and washed prior to pulse-labelling, subsequently cells were pulse-labelled in minimal medium containing azhal for 15 minutes at either 37 ºC or 44 ºC. We used iTRAQ (183) to detect changes in the amounts of proteins synthesized during this brief period by quantitative mass spectrometry. After the 15 minute pulse with azhal at either 37 °C or 44 °C, cells were harvested and proteins were extracted and digested as described above. The different digests were treated with different iTRAQ labels and then mixed in equivalent amounts. Newly synthesized proteins were identified by azhal-labelled peptides only, enriched by COFRADIC.

Because only methionine containing peptides can be labelled with azhal and used for the identification of newly synthesized proteins, the number of available peptides per protein that can be used for identification and quantitation sharply declines as shown in the supplemental data using an in-silico digest (Supplemental Figure 1). Consequently the contribution of single peptide protein identifications increases, as compared to the standard approach when all tryptic peptides can be used for identification (Supplemental Table I). However, a large part of the *E. coli* proteome (93.1%) is still predicted to be represented by this subset of peptides (Supplemental Figure 1) and proteins identified do not seem to be biased towards higher methionine content (Supplemental Table I).

A total of 394 newly synthesized proteins were identified after the 15 minute pulse. Of the identified newly synthesized proteins, 344 could be quantified using the iTRAQ reporter

ions, according to the criteria formulated in the experimental procedures. Upon an elevation in temperature the relative abundance of 64 newly synthesized proteins significantly ($p< 0.05$) in- or decreased more than a factor of two, (1 on ^{2}log scale) while 65 changed significantly ($p< 0.05$) by only a factor of 1.5-2 (0.58-1 on ^{2}log scale). The relative abundance of the remaining 216 newly synthesized proteins changed less than by a factor of 1.5 or did not change significantly at all during the 15 minute period after temperature switch (Figure 1).

Expression of heat shock-related proteins is induced by the *rpoH* gene product σ^{32}. This alternate sigma-factor is induced by heat shock and growth at higher temperatures during exponential aerobic growth (226-228). The level of newly formed σ^{32} was found to be up-regulated more than twofold within the first 15 minutes after the temperature change from 37 °C to 44 °C (Figure 1). Large scale transcriptomics studies with respect to σ^{32}-inducible genes have been carried out in which 97 genes were identified to be part of the σ^{32}-regulon (229, 230). From this σ^{32}-regulon, 28 gene products were identified, among these a number have chaperone functions, and aid in refolding proteins that are misfolded due to the temperature increase. The levels of newly synthesized proteins increased dramatically for HtpG, ClpB, IbpB, GroEL (CH60), DnaK and GrpE immediately following the temperature increase. Together with chaperones that aid in (re)folding of proteins at higher temperature, proteases are another important class of proteins induced during heat shock. Proteases degrade misfolded proteins and aid chaperones in (re)folding proteins during growth at higher temperatures. From the 28 heat shock-inducible proteins identified that are part of the σ^{32}-regulon, eight have protease functions. Of these four were found to be up-regulated in their respective levels of new protein formation more than two-fold (DegP, HslU, HtpX, LoN), while newly synthesized protein levels of the four others (ClpP/ClpA, FtsH and HflK) did not change when growth temperature increases (Figure 1).

Apart from the heat shock proteins that are part of the σ^{32}-regulon, we also found that levels of newly synthesized PspA increased significantly (3.3 fold ^{2}log scale). PspA is part of the phage shock PspABCDE regulon, which is induced by filamentous phage infection, and various environmental stresses including severe heat shock (236-239). Induction of PspA by increased temperature is independent of σ^{32} but is mediated by σ^{54}. However σ^{32}-mutants have a prolonged increase of PspA levels following heat shock (236, 237, 240), probably due to the lack of a proper heat shock response, suggesting cross-talk between the two pathways. Finally there is also a large group of proteins which have significantly lower levels of newly synthesized proteins after the increase in temperature. This is a group containing proteins with diverse functions, such as ribosomal proteins, cysteine biosynthesis/sulphur metabolism.

Among our dataset of 344 proteins, we could identify 15 species for which relative synthesis-rates under heat shock conditions have been measured previously (50, 106, 107, 122, 241). In all cases our data are remarkably similar with this previous work The relative amounts of three proteins of which the relative synthesis-rates increased dramatically during heat shock, namely GroEL (CH60) (122, 241), chaperone protein ClpB (106, 122) and σ^{32} (RP32) (50, 106, 107) where also highly increased under our assay conditions. In general

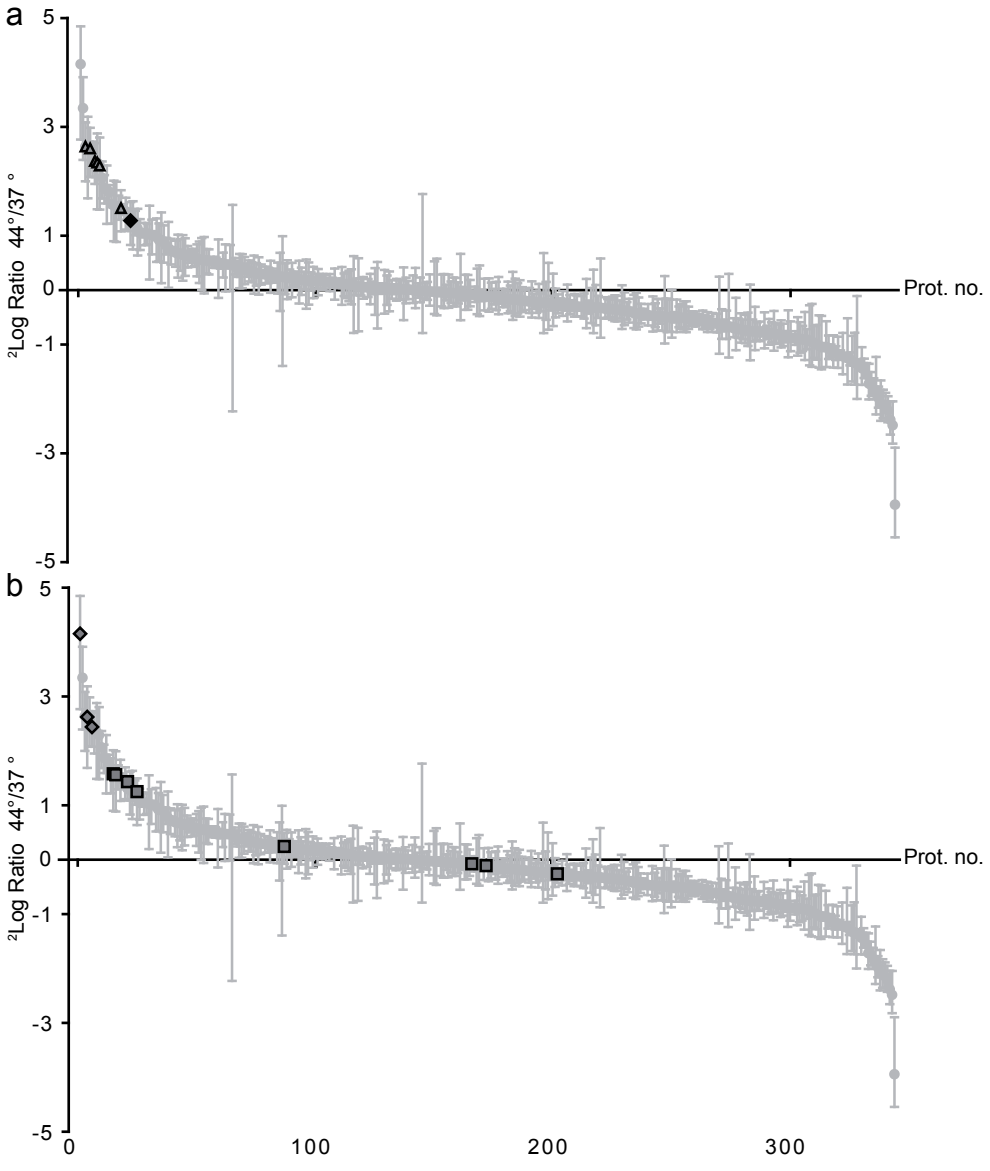

FIGURE 1. **Relative abundance of newly synthesized proteins 15 minutes after an increase in growth temperature**. Plot of relative abundance of newly synthesized proteins of cells grown at 44 °C compared to those grown at 37 °C during 15 minute pulse-labelling period with azhal. In panel (**a**) the increase in newly synthesized heat shock related proteins with a chaperone function (dark grey triangles), HtpG, ClpB, DnaK, IbpB, GroEL (CH60) and GrpE and the heat shock sigma factor, σ^{32} (black diamond) are shown. In panel (**b**) heat shock proteins with protease function (dark grey squares) DegP, HslU, HtpX, LoN, FtsH, ClpA, HflK, ClpP and three proteins (UxuA, UxuB, UxaC) involved in hexuronide and hexuronate degradation pathway (dark grey diamonds) are shown. Proteins are ordered from most to least increased in new synthesis after an increase in growth temperature, error bars denote standard deviation of peptide ratios obtained from biological replicates.

the relative amounts of newly synthesized species of the remaining 12 proteins, all involved in protein biosynthesis, were slightly decreased or not changed at all, in agreement with the slightly decreased synthesis-rates measured previously under heat shock conditions (122). These results strongly indicate that pulse-labelling with azhal is a reliable method to detect changes in the amounts of protein synthesized in a brief time frame upon changes in growth conditions.

Quantitation of immediate changes in total protein levels in response to heat shock— To estimate how the changes in newly synthesized proteins affect total protein levels on the same time-scale we also measured peptides that do not contain methionine or azhal. Just as azhal labelled peptides represent exclusively newly formed proteins, peptides that do not contain azhal or methionine represent the total protein content as they are made up of both pre-existing and newly synthesized proteins. To measure these peptides we analyzed the 'un-shifted' fractions obtained during COFRADIC by LC tandem-MS. Peptides that do not contain azhal or methionine are found in off-diagonal fractions as well, their presence is due to tailing of main chromatographic peaks that contain the bulk of unlabeled non-shifting peptides. The combined database searches from data acquired from shifted and non-shifted pooled fractions yielded a total of 1060 peptides identifying 435 proteins of which 292 could be quantified by the iTRAQ reporter ions. Upon an elevation in temperature the total protein levels of 8 proteins significantly ($p< 0.05$) in- or decreased more than a factor of two, (1 on ^{2}log scale) while 5 changed significantly ($p< 0.05$) by only a factor of 1.5-2 (0.58-1 on ^{2}log scale). The relative abundance of the remaining 279 proteins changed less than by a factor of 1.5 or did not change significantly at all during the 15 minute period after temperature switch (Figure 2).

In accordance with the elevated newly synthesized proteins, the proteins which have increased levels following the elevation of growth temperature mostly carry out chaperone and protease functions and are part of the σ^{32}-regulon. The heat shock sigma-factor increased its levels more than two-fold in the culture grown at 44 °C during pulse-labelling, in accordance with earlier observations (50, 106, 107). Concomitantly most of the chaperones IbpB, IbpA, ClpB, HtpG, DnaK, GroES (CH10), GroEL (CH60) and GrpE under its transcriptional control that were detected had increased levels from 1.4 to almost seven-fold (Figure 2). In contrast only one of the proteases had significantly changed total protein levels (DegP), while four others (HslU, ClpA, ClpX and FtsH) did not significantly change within the first 15 minutes upon heat shock. Total levels of PspA also increased significantly upon heat shock in accordance with it the increased levels of newly synthesized proteins described in the above. Conversely there were only two proteins (SyW and CysK) that had significantly decreased protein levels 15 minutes after the increase in growth temperature.

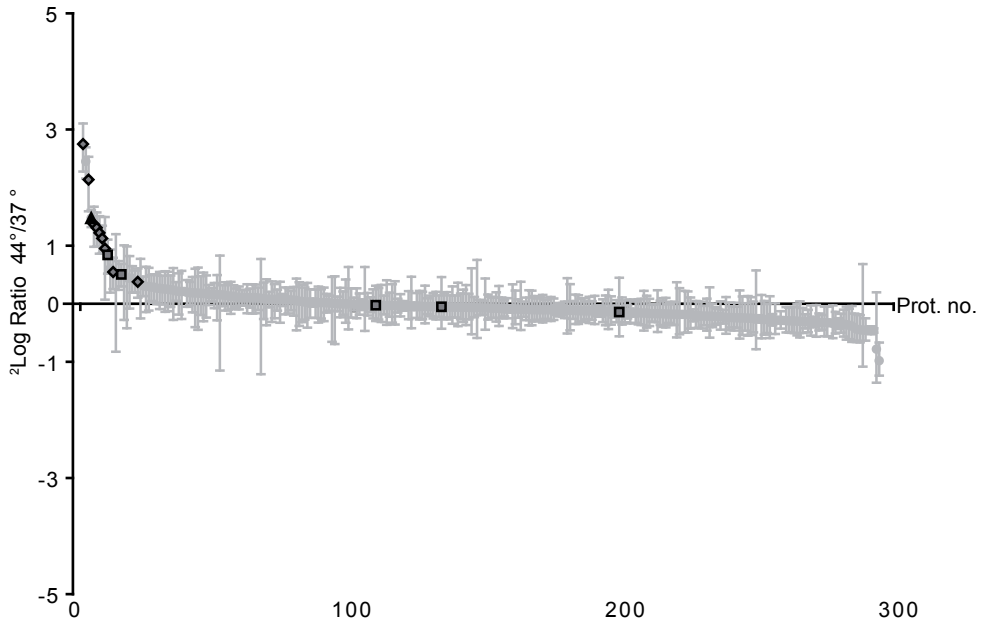

FIGURE 2. **Changes in total protein levels 15 minutes after an increase in growth temperature**. Total protein levels are shown for the heat shock sigma factor, σ^{32} (black triangle), heat shock proteins with chaperone function: IbpB, IbpA, CpbB, HtpG, DnaK, GroES (CH10), GroEL (CH60), GrpE and DnaJ (dark grey diamonds) or protease function: DegP, HslU, ClpA, ClpX and FtsH (dark grey squares). Proteins are ordered from most to least increased in total levels after an increase in growth temperature, error bars denote standard deviation of peptide ratios obtained from biological replicates.

Data on both changes in total protein levels and newly synthesized proteins during pulse-labelling allows identification of labile and stable proteins—By measurement of both changes in protein levels and newly synthesized proteins through the iTRAQ reporter ions stemming from azhal/methionine lacking- and azhal-containing peptide populations, our dataset contains 176 identified proteins for which both changes in total protein levels as well as changes in newly synthesized proteins were determined. In Figure 3 the data are ordered according to the relative levels of newly synthesized proteins (dots), while corresponding total levels are represented by diamonds at the same position along the X-axis. Overall, it is clear that the protein levels change to a lesser extent than the newly synthesized proteins newly synthesized proteins during the first 15 minutes after the change in growth temperature. However, we also identified five proteins, PspA, IbpB, σ^{32}, AhpC, and CysK, of which the relative total levels changed significantly upon the change in growth temperature, practically as much as the relative levels of the newly synthesized species (Figure 3a). Apparently, almost all pre-existing protein molecules of this group have been replaced by newly synthesized polypeptides during the pulse. So, the half-life of these proteins lies well within the pulse-labelling time used. For these proteins synthesis and degradation-rates are large compared to their total cellular levels during exponential growth at 37 °C. As a consequence a change in new protein formation will also affect the total cellular levels concomitantly

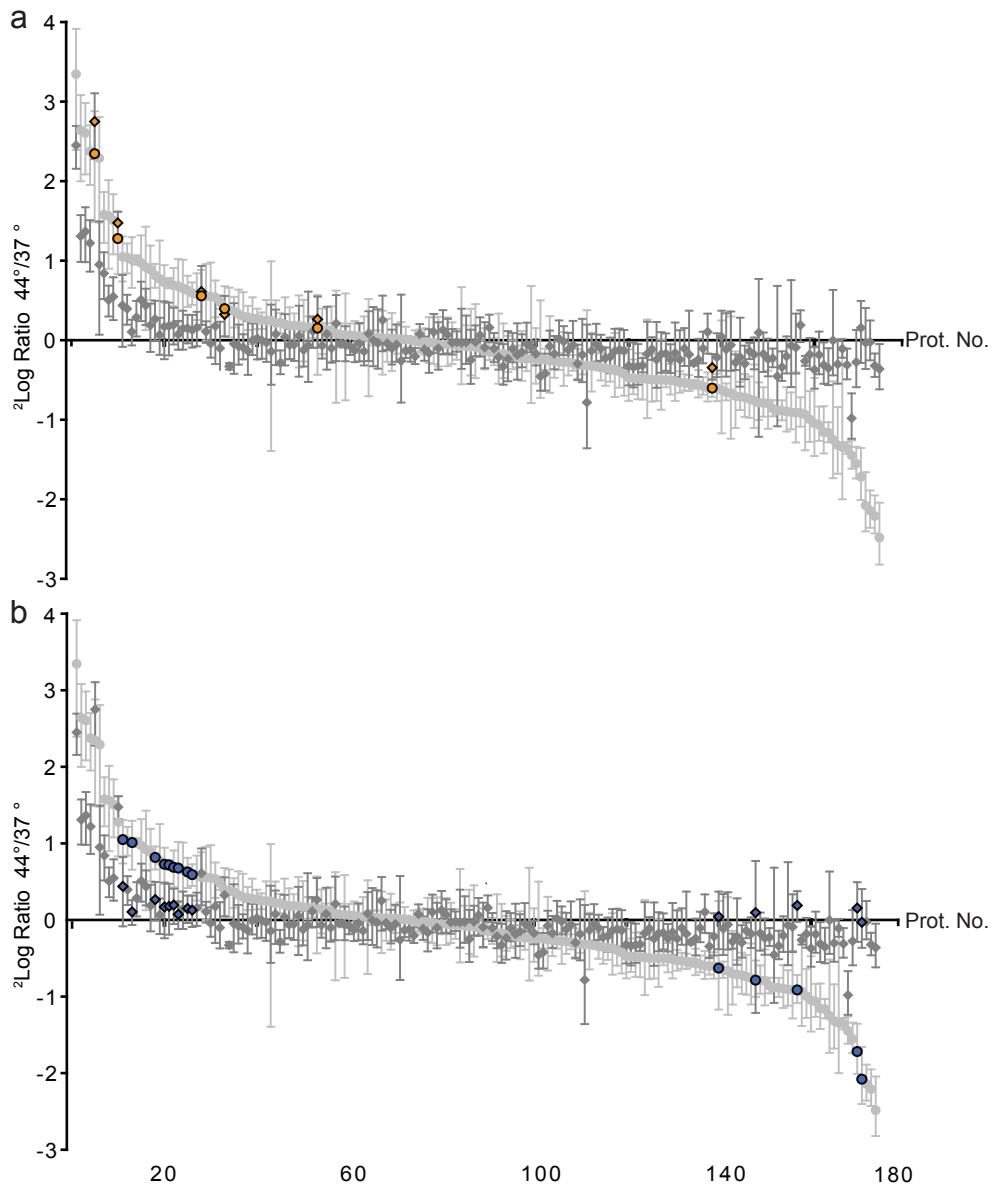

FIGURE 3. **Changes in protein synthesis compared to changes in protein levels upon onset of a change in growth temperature.** Light grey or coloured dots, relative change of newly synthesized proteins; dark grey or coloured diamonds at the same position on the X-axis as a dot are relative total protein levels of the same protein 15 minutes after the switch from 37 to 44 °C. Proteins are ordered from most induced to most repressed synthesis. In panel (**a**) proteins having a high turnover rate are marked, namely IbpB , σ^{32}, AhpC, AhpF and KatG (orange dots and diamonds). Coloured symbols in panel (**b**) represent proteins with a low turnover rate. Blue dots and diamonds, ribosomal proteins.

on the short time-scale used for pulse-labelling. The high turnover of σ^{32} has been reported before; it has a half-life of less than a minute during growth at normal temperatures, and is transiently stabilized during temperature increase (106, 107). This is thought to occur due to the an increase of unfolded protein as alternative substrate for FtsH, which decreases the degradation of σ^{32} (105). IbpB, a member of the family of small heat shock proteins, also has been found to have a high turnover rate (242). Surprisingly enough PspA on the other hand has been reported to be a stable protein (237), however the chase-experiment was conducted at 37 °C following induction of PspA synthesis by a temperature increase from 37 °C to 50 °C during pulse-labelling. So it is possible that following heat shock PspA degradation-rate is decreased in analogy to σ^{32}. If the half-life under normal growth conditions is short, then even though it is stabilized during heat shock, the increase in protein levels will mirror the increase in newly synthesized proteins as is the case for σ^{32}. No information on turnover was available for the other two proteins.

In contrast to the proteins for which the relative total levels change to almost the same degree as the relative newly synthesized levels upon heat shock, a significant difference between relative total levels and relative levels of newly synthesized material after a pulse of 15 minutes was noted for the majority of proteins. This is shown in Figure 3b by blue symbols for ribosomal proteins, but also the case for various chaperones and proteases. The difference between levels of total and newly synthesized protein can be explained by assuming that these proteins have a half-life that exceeds the pulse-labelling time used. Consequently, a large change in synthesis will not affect the total protein levels to the same extent in the short time frame used for pulse-labelling; only a continued altered rate of new formation will eventually change the protein levels to the same extent.

The observation that the protein levels for the majority of the proteins does not change significantly during the 15 minutes of elevated growth temperature also for most stable proteins for which protein new formation decreases, suggests that pre-existing proteins are mostly stable and not affected by the increased temperature. This is illustrated by the relative levels of methionine containing peptides in Figure 4. Just as azhal labelled peptides represent newly formed proteins, and non-azhal/methionine containing peptides the total protein levels, the iTRAQ reporter ions from these methionine containing peptides represent the relative levels of pre-existing proteins at 37 °C and 44 °C. As is clear from Figure 4 the relative levels of pre-existing proteins at 44 °C hardly change relative to those at 37 °C during the 15 minutes of pulse-labelling for the 87 proteins for which methionine containing peptides were found. This means that the elevated temperature does not change the half-life of the pre-existing proteins to such an extent that it is appreciable on the 15 minute period used for pulse-labelling. Apparently newly synthesized proteins are affected most with respect to stability, as these still need to be folded, while already folded proteins are not denatured and degraded at the temperature increase used here.

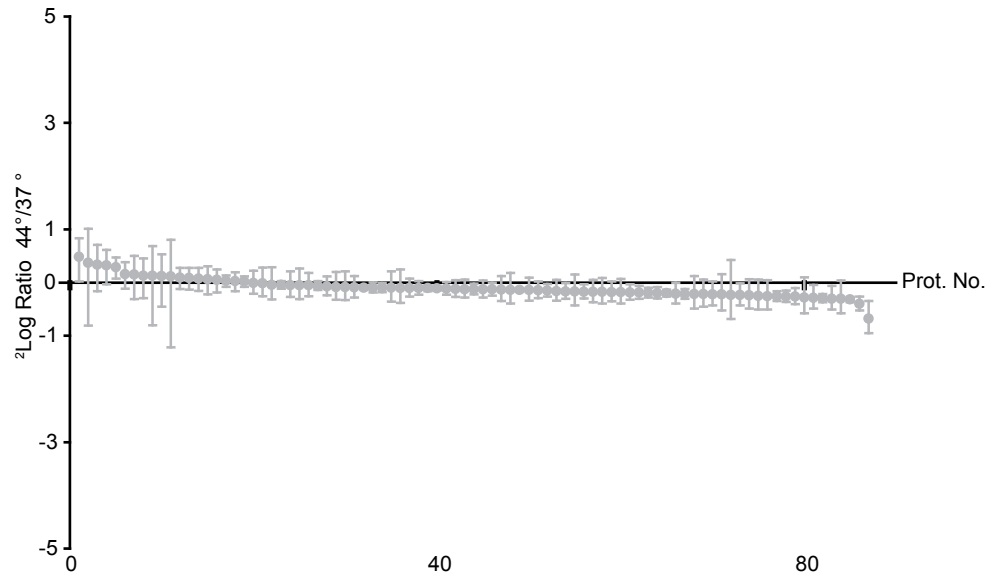

FIGURE 4. **Relative stability of pre-existing proteins during 15 minutes after an increase in growth temperature.** Plot of relative abundance of pre-existing proteins in cells grown at 44 °C compared to those grown at 37 °C during 15 minute pulse-labelling period with azhal. Proteins are ordered from most to least increased after an increase in growth temperature, error bars denote standard deviation of peptide ratios obtained from biological replicates. No major differences in the relative abundance of pre-existing proteins is obvious, which suggest that on the time-scale used for pulse-labelling, the degradation of pre-existing proteins does not play an important role for those proteins which were measured.

Comparison of changes in newly synthesized proteins and transcript levels suggest transcriptional regulation upon heat shock—Among the 24 proteins of which the relative amounts synthesized during the pulse were increased by at least a factor 2.4 (1.3 on ^{2}log scale), 14 belonged to the group of σ^{32}-regulated heat shock proteins of which the corresponding transcripts increased considerably during heat shock as measured by Richmond *et al.* (235) and Harcum *et al.* (234) (Table I). Remarkable is the presence of three proteins (UxuA, UxaC and UxuB), involved in hexuronide and hexuronate catabolism in the top six of proteins, of which the amounts synthesized during the pulse of azhal is increased, along with a corresponding increase in transcript level. Of the remaining seven proteins in the top 24, the corresponding mRNA levels are likewise increased in six cases and not measured in one case. The strong correlation between our proteomic data and transcriptomic data available in literature (234, 235), although measured at different temperature shifts, underscores the importance of regulation of gene expression at the level of transcription of the most elevated proteins under heat shock conditions.

Interestingly, the strong correlation between transcript levels and levels of newly synthesized proteins in the group of most up-regulated proteins did not exist in the group of proteins of which the amounts synthesized during the azhal pulse was decreased by a factor of 2.4 or more (Table II).

TABLE I
Highly up-regulated proteins upon heat shock

Gene name	Protein	Protein ratio†	Transcript ratio‡	Transcript ratio‡‡
uxuA	Mannonate dehydratase	3.8	7.6	1.29
pspA	Phage shock protein A	3.3	28.2	1.82
htpG	Chaperone protein htpG	2.6	33.8	4.26
uxaC	Uronate isomerase	2.6	1.8	0.77
clpB	Chaperone protein clpB	2.6	36.5	4.32
uxuB	D-mannonate oxidoreductase	2.4	3.2	0.47
dnaK	Chaperone protein dnaK	2.4	58.5	3.05
ibpB	Small heat shock protein ibpB	2.3	327.5	5.73
mopA	60 kDa chaperonin GroEL	2.3	37.9	3.83
ybeD	UPF0250 protein ybeD	2.1	1.6	2.93
relB	Antitoxin RelB	1.9	3.7	1.47
hdhA	7-alpha-hydroxysteroid dehydrogenase	1.9	2.5	0.17
ybdQ	Universal stress protein G	1.8	1.6	2.78
htrA	Protease do	1.6	9.6	3.11
hslU	ATP-dependent hsl protease ATP-binding subunit hslU	1.6	10.3	2.50
yfiA	Ribosome-associated inhibitor A	1.5	2.3	1.77
sdaA	L-serine dehydratase 1	1.5	23.6	3.56
grpE	Protein grpE	1.5	24.1	4.13
yibT	Uncharacterized protein yibT	1.5	n.d.	n.d.
htpX	Probable protease htpX	1.4	36.1	5.34
recN	DNA repair protein recN	1.3	1.7	1.08
rpoH	RNA polymerase sigma-32 factor	1.3	4	2.33
trxC	Thioredoxin-2	1.3	2.4	1.43
lon	ATP-dependent protease La	1.3	20.3	3.51
rplD	50S ribosomal protein L4	1.0	- 4.1	-1.27
yfgB	Ribosomal RNA large subunit methyltransferase N	1.0	- 1.2	-0.64
gltA	Citrate synthase	1.0	- 3	-1.52
rpsH	30S ribosomal protein S8	1.0	- 1.4	-1.93

† Relative protein ratio 44 ºC/37 ºC of proteins synthesized during 15 minute labelling period upon a change in growth temperature. ‡ Relative transcript ratio 50 ºC/37 ºC as reported by Richmond *et al.* (235), ‡‡ Relative transcript ratio 50 ºC/37 ºC as reported by Harcum *et al.* (234) n.d. transcript ratio not determined. (^{2}logscale)

Of the 19 proteins in this group, 7 are reported to have increased transcript levels upon an elevation in growth temperature by Richmond *et al* (235), while in the study of Harcum *et al.* (234) 6 of these are actually reported to be significantly down-regulated under the same conditions. There were 3 proteins (NlpA, YfdZ and SerA) which had significant decreased levels of newly synthesized proteins, while transcript hardly changed upon an increase in temperature reported by both transcriptomic-studies (234, 235) (Table II). In addition four proteins were identified of which the synthesis increases ~2-fold (1 fold on ^{2}log scale), whereas transcript levels change in the opposite direction. This suggests regulation at a post-transcriptional level for these proteins (Table I).

TABLE II
Highly down- regulated proteins upon heat shock

Gene name	Protein	Protein ratio†	Transcript ratio‡	Transcript ratio‡‡
cysM	Cysteine synthase	-1.3	- 1.6	-0.9
purA	Adenylosuccinate synthetase	-1.3	- 4.4	-1.2
b1680	Cysteine desulfurase	-1.3	2.5	-4.1
rho	Transcription termination factor rho	-1.3	- 7.3	-1.4
serA	D-3-phosphoglycerate dehydrogenase	-1.4	n.d.	0.2
tyrB	Aromatic-amino acid aminotransferase	-1.4	- 2.4	-1.1
cysK	Cysteine synthase A	-1.4	- 3.8	-3.3
yihK	GTP-binding protein typA/BipA	-1.6	- 4.6	-0.7
rplN	50S ribosomal protein L14	-1.7	- 1.7	-2.3
cysN	Sulfate adenylyltransferase subunit 1	-1.7	2.3	-3.5
gcd	Quinoprotein glucose dehydrogenase	-1.8	2.2	-0.9
cysJ	Sulfite reductase flavoprotein alpha-component	-1.9	1.3	-3.1
nlpA	Lipoprotein 28	-2.0	0	-0.1
rplY	50S ribosomal protein L25	-2.1	0.1	-1.7
cysD	Sulfate adenylyltransferase subunit 2	-2.1	2.9	-3.6
ydfY	Protein gnsB	-2.1	- 1.2	0.6
hlpa	Chaperone protein skp	-2.2	- 4.8	-1.3
b2379	Uncharacterized aminotransferase yfdZ	-2.4	1.3	-0.1
oppA	Periplasmic oligopeptide-binding protein	-2.5	- 4.6	-3.2
cysP	Thiosulfate-binding protein	-3.9	1.9	-2.4

† Relative protein ratio 44 ºC/37 ºC of proteins synthesized during 15 minute labelling period upon a change in growth temperature. ‡ Relative transcript ratio 50 ºC/37 ºC as reported by Richmond *et al.* (235), ‡‡ Relative transcript ratio 50 ºC/37 ºC as reported by Harcum *et al.* (234) n.d. transcript ratio not determined. (^{2}logscale)

Discussion

We used pulse-labelling with azhal to determine the relative abundance of proteins synthesized during the transition period after a change in growth temperature from 37 to 44 ºC. By the use of iTRAQ for relative quantitation of newly synthesized proteins, the relative abundance of 344 newly synthesized proteins was determined upon a change in growth temperature. Amongst proteins highly up-regulated after the temperature switch there were many heat shock inducible chaperones and proteases, all part of the σ^{32}-regulon. In addition, PspA which has been previously reported to be induced after heat shock independent of σ^{32} (237) was also identified amongst the most up-regulated newly synthesized proteins after the temperature switch. Proteins, for which radiolabelling data was available, showed high similarity to levels of newly synthesized proteins determined by azhal-labelling. This further validated that the protein expression measured here is due to temperature-induced changes in gene expression, and not the result of labelling cells with azhal.

It should be noted however that there may already be a significant contribution of degradation to the amount of each protein formed in 15 minutes. This is dependent on the particular protein's half-life. We also determined the changes in total levels on the same time-scale of pulse-labelling by iTRAQ reporter ions stemming from non-azhal/methionine containing peptides. Comparison of total protein levels and newly synthesized proteins for those proteins that showed significant changes upon heat shock revealed two distinct groups of proteins. The first group comprised of five proteins (PspA, IbpB, σ^{32}, AhpC, and CysK) had protein levels that changed as much as the levels of newly synthesized proteins upon heat shock. This can be explained by assuming these are labile proteins that have a short half-life within or not much longer than the pulse-labelling time used. This means that a change in new protein formation, which can be the result of a change in protein synthesis-rate, protein degradation-rate or a combination of the two, will also significantly affect protein levels on the short time-scale used for pulse-labelling.

The large majority of proteins that were significantly regulated upon heat shock showed a marked discrepancy between changes in protein levels and newly synthesized proteins. In contrast to the group of labile proteins mentioned in the above, these are stable proteins with a half-life that exceeds the pulse-labelling time used, as a change in new protein formation does not affect the total protein levels on a short time-scale to the same degree. Among these stable proteins were chaperones and proteases of the σ^{32}-regulon and the ribosomal proteins. The change in newly synthesized protein levels for these proteins could be the result of an increase or decrease of protein synthesis, as protein half-life should not play a role on the short-time-scale of pulse-chase labelling. This is demonstrated for most of the pre-existing proteins, which did not significantly change their levels within the pulse-labelling time with respect to the cells grown at 37 °C. However, newly formed proteins are probably more affected by the increase in temperature than pre-existing ones as these still have to be folded, and are thus more likely to be degraded if they misfold in spite of chaperone to aid in this process. Consequently the relative levels found at these two different

temperatures probably do not directly reflect the relative synthesis-rates for each protein found, this depends on whether the relative stability of the newly synthesized polypeptides are also similar under the different temperatures.

Comparison of the results with transcript data from literature revealed that many proteins that are highly up-regulated upon a temperature switch seem to be regulated at the transcriptional level, as increased transcript levels corresponded with an increase of protein synthesis, as determined by azhal incorporation. Examples to illustrate this are the genes *uxuA*, *uxuB* and *uxaC* that encode enzymes that catalyze different steps in the catabolism of hexuronides and hexuronates to 2-keto-3-deoxy-gluconate (KDG). KDG is metabolized by the Entner-Doudoroff pathway and enters the lower part of glycolysis (243). There are no previous reports about temperature induction at the protein level of members of this pathway. However UxuA, UxuB and UxaC were found to be highly up-regulated upon heat shock in this study, in good correlation with altered transcript levels reported before (235). A possible explanation for the measured increase in transcription could be instability of the transcriptional repressors of these genes (ExuR, UxuR) at higher temperatures.

In contrast to the highly up-regulated heat shock proteins, there was poor correlation between transcript data and relative levels of newly synthesized proteins in a group of three highly down-regulated proteins. Increased turnover of these proteins at higher temperature offers one explanation. While for a group of six other proteins that were highly down-regulated with respect to newly synthesized proteins, transcriptomic studies contradicted each other, whether proteins were up or down-regulated upon heat shock. For a group of four proteins found to be up-regulated, with transcript levels going down significantly, upon elevation of the growth temperature, post-transcriptional regulation is a possible explanation. This shows that azhal-labelling is very suitable to identify candidates that may be subject to post-transcriptional regulation.

The first proteome-wide approach to quantitation of newly synthesized proteins by azhal-labelling is demonstrated here using COFRADIC. Altogether, the fact that no severe azhal-related disturbances were obvious, combined with the large number of newly synthesized proteins identified and quantified, makes azido-peptide isolation by COFRADIC in combination with iTRAQ an excellent tool for both identification and quantitation of transient changes in protein expression. While the measurement of total protein levels on the same time-scale allows identification of stable and labile proteins among those proteins that are significantly regulated by the environmental stress. The pulse-labelling technique described is uniquely suited to follow an adaptation to changes in the environment of *E. coli.* Furthermore comparison with transcript data allows for screening for different types of regulation in response to a change in environment. The presented method can open up new avenues in systems biology research, by filling the gap of information between transcriptomics and proteomics and allow for new input into advanced modelling of cellular networks.

Supplemental Data—Supplemental Figures and Tables can be found in the addendum section on pages 136-137. Protein identification and quantitation data can be found online at http://www.mcponline.org as supplemental information for references (141) and (169).

Experimental Procedures

Synthesis of L-azhal— L-azhal was synthesized from L-Boc-2,4-di-aminobutyric acid (L-Boc-DAB, Chem-Impex, Wood Dale, USA) by diazotransfer (198) using Triflic azide (TfN_3) as previously described (ref).

Cell culture— The methionine-auxotrophic *E. coli* strain MTD123 (180) was grown in M9 minimal medium containing 6.8 µM $CaCl_2$, 1.0 mM $MgSO_4$, 59.3 µM thiamine·HCl, 57.0 nM Na_2SeO_3, 5.0 µM $CuCl_2$, 10.0 µM CoCl, 5.2 µM H_3BO_3, 99.9 µM $FeCl_3$, 50.5 µM $MnCl_2$, 25.3 µM ZnO, 0.08 µM Na_4MoO_4, 111 mM glucose and 60 mg/l for each of the 19 natural amino acids and 40 mg/l for tyrosine (Sigma-Aldrich, St Louis, USA). For temperature switch experiments two cultures (A and B) were grown aerobically at 37 °C in M9 minimal medium as described above, in order to have a biological replicate. After overnight culture, cells were inoculated at OD_{600} 0.01 and allowed to grow into exponential phase before being harvested at OD_{600} 1.0. These cells were washed at room temperature with complete M9 minimal medium (but lacking methionine) to prevent osmotic shock during washing and then transferred to M9 minimal medium in which the methionine was replaced by 400 mg/l azhal. The cultures were split and then transferred to water-bath shakers set to either 37 or 44 °C. The four cultures were allowed to resume growth aerobically for 15 minutes before being harvested.

Sample preparation— Samples were essentially prepared as described before (141), in short: azhal-labelled cells were harvested by centrifugation, pellets were resuspended in lysis buffer and lysed by sonication after which cellular debris was removed by centrifugation. Samples were dialyzed overnight and their protein content was determined. Samples were then subjected to overnight digestion with trypsin.

iTRAQ labelling— Samples (125 µg protein per sample) were lyophilized after digestion and redissolved in 40 µl of 125 mM tri-ethyl-ammonium bicarbonate pH 8.5, and labelled with iTRAQ (244) according to the manufacturer's protocol (Applied Biosystems, Toronto, Canada), with the exception that two vials of iTRAQ reagent were used per sample to ensure complete labelling. Samples were incubated for two hours at room temperature after which the reaction was quenched by adding 300 µl of 0.1% formic acid. The digests from the cultures A and B grown at 37 ºC during pulse-labelling were labelled with iTRAQ 114 and 116 respectively, while digests from the cultures A and B grown at 44 ºC were labelled with iTRAQ 115 and 117. The four labelled samples were mixed in a 1:1:1:1 (w/w) ratio, resulting in 500 µg iTRAQ labelled digest. To remove the excess of iTRAQ reagent the sample was diluted three times to a final volume of 6 ml 20% acetonitrile in 0.1% formic acid and loaded on an ICAT cation exchange cartridge (Applied Biosystems, Toronto, Canada). The cartridge was washed with 500 µl 20% acetonitrile in 0.1% formic acid, before the digest was eluted with 2 M ammonium formate buffer pH 6.8 containing 20% acetonitrile and lyophilized. Samples were redissolved in 50 mM Hepes pH 8.0 and reduced and alkylated as described above, before ~200 µg was loaded for the primary run of diagonal chromatography as described below.

COFRADIC and mass spectrometric analysis—Azhal-containing peptides were enriched by COFRADIC (213), using TCEP to selectively modify target peptides between the primary and secondary chromatographic runs. TCEP induces a set of competing reactions in azhal-containing peptides present in primary fractions, i.e., conversion of the azido-group to an amine or hydroxyl group and cleavage of the peptide bond at the C-terminal side of azhal residues (141, 143, 223). The subsequent enrichment is based on a difference in retention-time during the secondary chromatographic runs between TCEP-induced reaction products and the bulk of unmodified peptides that are present in the particular primary fraction subjected to TCEP treatment (141). Three fractions 16 minutes apart in the primary run are pooled and reinjected after TCEP treatment. Fractions of secondary runs enriched in TCEP-induced reaction products from azhal-containing peptides, were analyzed by LC tandem-MS. These 'Off-diagonal' pooled

fractions were redissolved in 10 µl 0.1% TFA with the addition of 150 pmol human [Glu1]-Fibrinopeptide B (Sigma-Aldrich, St Louis, USA) for internal calibration, 10 µl sample was separated and analyzed as described in detail previously (141). In addition pooled fractions containing non-shifted material were also collected, resuspended in 400 µl of 0.1% TFA with the addition of 150 pmol of human [Glu1]-Fibrinopeptide B (Sigma-Aldrich) for internal calibration. Of these samples 5 µl was injected and analysed as described for the fractions enriched with labelled material. Assessment of the relative quantity of each protein by analysis of tandem-MS spectra is based on the signal intensities of reporter ions derived from the iTRAQ-moieties of the respective peptides.

Identification and Quantitation— The tandem-MS runs were first internally recalibrated on the fragmentation spectrum of [Glu1]-Fibrinopeptide B, before being exported by the mascot.dll as described before (141). Generated peak lists were submitted to MASCOT to identify newly synthesized proteins using the following parameters: Cleavage after lysine or arginine unless followed by proline plus cleavage after methionine, allowing up to 2 missed cleavages, fixed carbamidomethyl cysteine, iTRAQ (K) modifications. Variable modifications used were iTRAQ (N-terminal) modification and modifications induced by reaction of TCEP with azhal-containing peptides as described before (141). Peptide mass tolerance was set at 0.1 Da and MS/MS tolerance was set at 0.05 Da. The significance threshold was set to 0.01 resulting in a threshold score of 34. Mudpit scoring and 'require bold red' were applied with an ion-score cut-off of 35, in order to have all peptide matches identified at a p-value of <0.01. Mascot performed fragment ion searches with the above settings in a local database of the *E. coli* K12 proteome (4506 proteins; 1426768 residues, release 14.4 04/11/08, Uniprot consortium, http://beta.uniprot.org/). In addition peak lists from non-shifted fractions was searched with the following parameters: cleavage after lysine or arginine unless followed by proline, allowing up to one missed cleavage, fixed carbamidomethyl cysteine modification, iTRAQ (K) modification. Variable modification used was iTRAQ (N-terminal) modification. Peptide mass tolerance was set at 0.1 Da, and MS/MS tolerance was set at 0.1 Da. The significance threshold was set to 0.01 resulting in a threshold score of 28. Multidimensional protein identification technology scoring and "require bold red" were applied with an ion score cut-off of 35 to have all peptide matches identified at a *p* value of <0.01. MASCOT, False positive rates were estimated using a decoy database as described before (141) and were found to be less than three percent. TCEP-induced reaction products were selected manually and selected queries were used to recalculate protein coverage and protein score based on azhal-labelled peptides only. The resulting MASCOT data-file of this search was imported into Quant (245) (http://www.protein-ms.de), for quantitation using the iTRAQ reporter ions using only labelled peptides unique to each protein. For identification purposes both searches were also exported as csv-files. Quant settings were as follows: all four iTRAQ reporters on, report peak areas on, reporter tolerance set at 0.1 Da, intensity range turned of, peak dimensions at 0.025 Da, absolute intensity error set at zero, experimental error set at 0%, use of unique peptides on, p-value cut-off set at 0.01 and macro language set at English with the macro parameter separator set to comma. The correction factors were put in for iTRAQ kit no. 080591.

Quant output is a tab-delimited text file containing both reporter ion ratios per peptide as well as mean-protein ratios derived from these. To assess the combined effect of technical and biological variance the average and standard deviations of all protein ratios of 116/114 and 117/115 for the data-set were calculated. These ratios represent replicate B/replicate A at 37 ºC and replicate B/replicate A at 44 ºC and should theoretically be one. The 116/114 ratio yielded an average of 1.01 with standard deviation (s.d.) 0.19 and the 117/115 ratio an average of 1.03 (s.d. 0.20), for the azhal-labelled proteins, while for the unlabeled proteins the 116/114 ratio yielded an average of 1.07 (s.d. 0.17) and the 117/115 ratio an average of 1.11 (s.d. 0.18), which shows that no large systemic error was made during mixing of samples. Subsequently tandem-MS spectra were inspected manually. Peptides which did not have signals for all four reporter ions or peptides which showed inconsistent 115/114 and 117/116 ratios (biological plus technical replicates) were discarded. Box-plots per protein were checked to identify outliers in the peptide ratios per protein as described in (245). In addition for quantitation of relative total protein level changes induced by temperature, methionine containing peptides were removed from the non azhal-containing peptide set as well.

Due to the incorporation of replicates into one quantitation-experiment, each peptide can yield four relevant reporter ratios. First of all: 115/114 and 117/116 for each pre-labelling culture split into two different growth temperatures during labelling. Furthermore 117/114, 115/116 for one growth temperature of one pre-labelling culture compared to the other growth temperature of the other pre-labelling culture during labelling. For each protein

the mean expression ratio was determined by calculating the mean of the peptide ratios 115/114, 117/114, 115/116 and 117/116 per protein (mean ratio of expression between the biological replicates). The accuracy of the ratio per protein is expressed by calculating the standard deviation of the peptide ratios for all the peptides measured per protein (maximum standard deviation between the biological replicates). To ascertain if up- or down-regulation was significant, a double sided Welch's T-test for each protein was performed, using the Welch-Satterthwaite equation (246) to determine the degrees of freedom, to see whether the mean protein ratio differed significantly from the mean ratio obtained for unregulated proteins ($\mu = 1.02$, $\sigma = 0.16$, n=343 for newly synthesized proteins and $\mu = 1.08$, $\sigma = 0.17$, n=291 for protein levels; from ratio of 116/114 and 117/115 reporter ions). Proteins which changed more than 1.5 fold and for which the p-value was adjusted for a false discovery rate of less than <0.05 due to multiple testing (247) were considered to have a significantly altered expression level.