CHAPTER 4

TRANSCRIPTOMICS VALIDATION STRATEGIES

4.1 INTRODUCTION

4.1.1 Evidence of post-transcriptional regulation in *P. falciparum*

Gene regulation of *P. falciparum* is currently a controversial issue with evidence supporting the dominant role of post-transcriptional control [119, 128, 192, 193] on the one hand and evidence demonstrating transcriptional control [123, 179, 189, 190, 198, 199] on the other. As a result, conclusions drawn on transcriptomics data alone require validation. In addition, post-transcriptional responses to perturbations can potentially be missed if only assessed by microarray. The evidence corroborating post-transcriptional control as dominant mode of gene regulation in *P. falciparum* includes the apparent paucity of recognisable promoter elements in the genome [119], the small amplitude transcriptional responses detected upon perturbation of the parasite, conflicting results obtained when the parasite is exposed to small molecules and the recent evidence of variable mRNA decay rate and translational repression [192-194, 196, 197]. These will be discussed in more detail in the following paragraphs.

Nirmalan and colleagues studied the transcriptional [197] and the translational response [242] of antifolate-inhibited parasites, but in contrast with the decreased transcript level of the drug target DHFR/TS, a marked increase was seen on the protein level. This was specific to the inhibitor used and indicated that the parasite is able to significantly relieve constraints resulting from environmental perturbation post-transcriptionally as opposed to responding at the transcriptional level [197]. Another integrated investigation of the transcriptome and proteome of *P. falciparum* after treatment with a choline analogue also indicated more pronounced changes at the protein level without specific transcriptional responses to the drug [128]. Drug-specific effects were furthermore detected on the protein level in a two-dimensional electrophoresis (2D-GE) study of the parasite response to treatment with artemether/lumefantrine [251]. In addition, most studies of environmentally perturbed *Plasmodium* that did report compensatory transcriptional responses and supported the presence of transcriptional control mechanisms in the parasite were not confirmed on the proteome/metabolome level [179, 189, 190, 198] and several others simply failed to detect such programmed transcriptional responses [194, 196, 197].

However, probably the most important evidence of post-transcriptional control of gene expression in *Plasmodium* was the demonstration of the variation in mRNA decay rate during the IDC [193] and the presence of mechanisms resulting in translational repression [192]. The mRNA decay rate increases with parasite development during the IDC and ring-stage parasites have an average mRNA half life of about 9.5

min, which is extended to an average of 65 min during the late-schizont stage [193]. These results implied that the stage-specific mode of transcription during the IDC [29] may in fact be the result of stage-specific mRNA decay [128]. Translational repression mediated via the highly conserved DEAD-box RNA helicase, DOZI, was demonstrated in *P. falciparum* [192]. Translational repression has an important role in sexual differentiation and gametogenesis in higher eukaryotes and involves the movement of specific mRNA molecules into cytoplasmic messenger ribonucleoprotein complexes (mRNPs), which prevents their translation until release from these complexes. DOZI-mutant parasites failed to store certain transcripts (p25, p28), which were therefore targeted for degradation, whereas the transcripts in the translationally repressed mRNPs from wild-type parasites were stored for translation after fertilisation [192].

The most recent addition to the debate on the role of transcriptional control in the parasite was evidence of a cascade of AP2 transcription factors that were proposed to control transcription during the IDC [123], as discussed in section 1.10.5.

4.1.2 Integrative biology from *Plasmodium* functional genomics data

There are currently no reports nearing systems biological investigations of Plasmodia, but biological questions regarding the parasitic lifestyle are starting to be addressed. The proteome data, when combined with evidence of chromosomal clusters encoding co-expressed genes, suggest a highly coordinated expression of *Plasmodium* genes involved in common biological processes [246]. Furthermore, there appears to be a positive correlation between the abundance of transcripts and their encoded proteins during the *P. falciparum* lifecycle [246, 252]. The majority of discrepancies are attributable to a delay between transcript production and protein accumulation (translational gap), which is observed as a time-shift between detection of the transcript and its protein [246].

The most comprehensive and integrated analyses of the genome, transcriptome and proteome have been reported for *P. berghei* and *P. chabaudi chabaudi* and represent state-of-the-art of functional genomics as applied to Plasmodia [30, 253]. Hall and colleagues proposed that the parasite uses four strategies/mechanisms for gene expression during its lifecycle, which includes housekeeping genes, stage-specific genes, host-specific genes (mosquito or mammalian host) and strategy-specific genes (related to invasion, asexual replication and sexual development). Their results indicated that over half of the proteins were detected solely in one stage of the lifecycle, implying a considerable degree of specialisation at the molecular level to support the demanding developmental programme. This large dataset allowed the authors to observe post-transcriptional gene silencing of some gametocyte transcripts with the implicated involvement of a 47-nucleotide sequence motif in two-thirds of the 3'-untranslated regions of these transcripts [61].

New analysis methods are being developed to integrate transcriptomics, proteomics and metabolomics datasets, such as the Partial Least Squares (PLS) method, which was used to integrate yeast transcriptome

and metabolome data [254]. Assuming that the metabolome is a function of the transcriptome, changes due to environmental perturbations were used to model the metabolic variables with PLS. This allowed the discrimination of the effects of perturbations on the transcriptome and metabolome, the modelling of the metabolome as a function of the transcriptome and the extent of similarity between these and finally the identification of transcripts that mediate changes in the metabolome [61]. However, reports of perturbations of *Plasmodium* that integrate transcriptomics, proteomics and metabolomics data are currently not available.

4.1.3 Proteomics methodologies

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE or 2D-GE) is a powerful and high-resolution protein separation method. In the first dimension proteins are separated based on their iso-electric point (pI) and in the second based on their molecular weight (Mr). Protein mixtures are resolved in this manner over the entire gel for an average of 1000 to 1250 spots (~300 to 1000 proteins) with a weight range of 10 – 250 kDa [255, 256]. This number can be improved to 10000 spots and 3000 proteins under some conditions [255]. The technique has developed considerably in recent years and the use of immobilised pH gradient (IPG) strips for iso-electric focusing (IEF) during the first dimension has tremendously improved the reproducibility. Furthermore, advances in staining technology with fluorescent stains (e.g. SyproRuby, DeepPurple and Flamingo) allow reliable quantitation of separated proteins with high sensitivity and dynamic range compared to the limitations of the traditional silver and Coomassie brilliant blue stains [255]. The fluorescent stains have linear ranges of about three orders of magnitude compared to the covalently bound fluorescent dyes (Cy2, Cy3 and Cy5) with linear ranges of greater than four orders of magnitude [257]. Most modern gel scanners are linear up to five orders of magnitude. However, proteins differ in their staining characteristics and therefore the relationship between intensity and concentrations across different proteins are not the same. Of all the protein stains, the traditional Coomassie stain is the least sensitive with a lower limit of detection of about 10 ng protein compared, to silver and fluorescent stains that have similar sensitivity of ~1 ng protein (20 fmol for a 50 kDa protein). The fluorescent dyes are currently superior and the Cy dyes have a limit of detection as low as 125 pg protein [257]. These covalently bound Cy dyes were introduced to overcome the technical variability of 2D-GE with difference gel electrophoresis (2D-DIGE). Samples are differentially labelled with these dyes (e.g. Cy2, Cy3, Cy5) were introduced to enable differentiation between samples within the same protein mixture on a single 2D gel. The fluorophores can be resolved spectrally and matched for mass and charge to allow better spot matching and normalisation across gels, which minimises gel-to-gel variation [258]. The reproducibility of 2D-GE for quantitative proteomics can also be improved by differential tagging of samples via metabolic isotope labelling by growing cultures on medium containing either heavy- or light-isotope containing components. This was successfully applied with *P. falciparum* grown on medium containing either $^{13}C,^{15}N$-isoleucine or regular isoleucine, followed with sample separation on the same gel and mass spectrometry (MS) analysis for identification and differentiation between the samples [259].

However, a significant drawback of 2D-GE is protein solubility problems, which limit the separation and detection of low-abundance and hydrophobic proteins. Multidimensional liquid chromatography (LC)-techniques have emerged to interface separated proteins/peptides directly into mass spectrometers (MS). MudPIT is an LC-based method, whereby proteins are separated depending on charge in the first dimension and hydrophobicity in the second, prior to MS analysis [260]. LC approaches also allow for differential labelling of samples via chemical tagging of proteins/peptides with stable isotopes (e.g. on the cysteine residues) known as isotope-coded affinity tags (ICAT), which enables relative quantitation during comparative studies. A more recently developed quantitative method, isobaric tags for relative and absolute quantitation (iTRAQ), employs a 4-plex set of amine reactive isobaric tags to derivatise peptides at the N-terminus and on the lysine side chains, thus labelling all the proteins in the mixture. The labelled peptides cannot be distinguished with MS (isobaric), but distinguishable signature ions are produced upon fragmentation with tandem MS (MS/MS) [261]. LC methods generally have poorer resolution than 2D-GE and can resolve only hundreds to thousands of proteins, but it can separate low molecular weights of 1 – 20 kDa and offer the additional advantage of easy automated coupling to MS for protein identification [256]. DIGE and ICAT are currently the most commonly practised techniques for gel-based and LC-based quantitative proteomics, respectively [261].

For protein identification, MS and MS/MS methodologies are most commonly used [256]. Mass spectrometers measure the mass/charge ratio (m/z) and make use of the growing knowledge of genome and protein databases where these lists of peak intensities and m/z ratios can be manipulated and compared with lists of theoretical protein digestions or fragmentions [262]. Major advances for analysing protein structure was the introduction of soft ionisation techniques such as electrospray ionisation (ESI) and matrix assisted laser desorption/ionisation (MALDI), which enable the volatilisation of biomolecules. In ESI, the sample is passed through a high-voltage needle at atmospheric pressure to produce charged droplets that desolvate prior to entrance to the high vacuum of the mass spectrometer. In MALDI, the samples are co-crystalised onto a sample plate with an organic matrix that has a conjugated ring structure to absorb at the wavelength of the lazer. ESI typically induces a range of charged states, whereas only singly charged ions are usually observed with MALDI [262]. Limitations of these techniques include the limited number of peptide ions produced by trypsin digestion for MALDI analysis that are expected to co-crystalise efficiently with the matrix; the influence of the size and composition (e.g. arginine versus lysine) of the peptides on the signal detected, ESI voltage favouring certain charged states (e.g. higher voltages favour lower charged forms, but the lower charged state may not be within the mass range of the instrument) and competition between analytes for charge as they are extruded from the ESI spray droplets [262].

The ESI and MALDI platforms are combined with various mass analysers, e.g. quadrupole, time-of-flight (TOF), quadrupole ion traps and Fourier transform ion cyclotron resonance (FTICR). Quadrupoles apply radio frequency and voltages, which allows a narrow m/z ratio range to reach the detector. The technique is usually

limited in mass range and has low resolution (separate m/z from 0 to 4000). TOF analysers accelerate ions via a short voltage gradient and measure the time ions take to traverse a field free flight tube (the flight time is proportional to the square root of m/z). TOF instruments can achieve a resolution of 10000 (separate m/z of 1000.0 from 1000.1). Quadrupole ion traps focus ions into a small space with an oscilating electric field and ions are activated and ejected by electronic manipulation of the field. Ion traps are very sensitive since ions can be trapped for varying lengths of time. FTICR uses high magnetic fields to trap and cyclotron resonance to detect and excite ions with a resolution >1000000 (separate m/z of 1000.000 from 1000.001) [262].

For the proteomics investigation of PfAdoMetDC/ODC co-inhibition, as discussed in this chapter, samples were separated with 2D-GE and stained with a fluorescent stain (Flamingo) to enable the sensitive and accurate quantitation of proteins. DIGE is currently the method of choice for differential protein abundance analysis, but the required fluorescent scanner was not available. Selected proteins were subsequently identified with MS/MS (MALDI-Q-TOF), which enabled protein identification and confirmation/validation, as discussed in section 4.3.1.1.

4.1.4 Metabolomics methodologies

Enzyme-based assays for the one-by-one measurement of metabolites have been available for decades, but methods for the simultaneous quantitation of large numbers of metabolites have only become available in recent years [263]. Efforts to quantitate multiple metabolites simultaneously have included thin-layer chromatography [264], high-performance liquid chromatography (HPLC) with UV-detection [265], nuclear magnetic resonance (NMR)-spectroscopy [116] and LC-MS [263, 266, 267]. In general, the methods without MS-detection suffer from low sensitivity and specificity [263]. Global metabolomics analysis of *Plasmodium* is still in its infancy, but recently the 2D-NMR was used to identify and quantitate more than 50 metabolites from isolated mature *P. falciparum* trophozoites [116].

The LC-MS technologies, as discussed for proteomics in section 4.1.3, can also be applied for global metabolomics analysis. NMR is an alternative method that requires minimal sample preparation and is considered to have a high throughput (hundreds of samples per day). In ^{1}H-NMR an external magnetic field aligns the nuclear spin of responding proton nuclei at specific frequencies where they occur in resonance. The radiation emission or absorbance frequency during nuclei relaxation is detected, which varies depending on the number of electrons orbiting the nucleus. ^{1}H-NMR theoretically provides unique signals for chemically distinct hydrogen nuclei, which allows analyte structure determination. However, NMR has only half the sensitivity of the discussed MS-based approaches [268].

The advantage of metabolomics compared to transcriptomics and proteomics is that a particular metabolite has the same basic chemical structure irrespective of the organism from which it was extracted. Therefore, once the technology is established and quantitation challenges are overcome, a universal approach that spans

the species barriers can be adopted [269]. However, in contrast with the transcriptome and proteome, metabolite concentrations within the cell are determined by the activities of various enzymes and as a result the components of the metabolome are generally far more complex than is the case with transcripts or proteins [270]. For example, less than 30% of metabolites are involved in only two reactions, ~12% of metabolites participate in more than 10 reactions and ~4% are involved in more than 20 reactions, resulting in a high degree of connectivity in the metabolic network. Because of these complex networks of tightly connected reactions, small perturbations in the proteome can result in significant changes in the concentration of various metabolites. Therefore, metabolomics analyses provide integrative information. However, this strength is also a drawback since data interpretation in the physiological context is generally difficult [270].

Metabolomics data analyses essentially consist of raw data processing, data mining, presentation and storage. Because of instrument inaccuracies such as chromatogram shift and mass drift, raw data processing involves the deconvolution of overlapping chromatogram peaks and chromatogram alignment to overcome the drift and to enable the comparison of different datasets. Baseline correction and noise reduction can also be advantageous, but such data smoothing may lose useful information [268]. Once quantitative datasets have been obtained the fundamental approach is to compare the level of a metabolite between experimental and control samples and to use standard statistics to assess the significance of the detected differences [271]. However, high-throughput multivariate analysis or mining of all the components within a dataset requires unsupervised [e.g. principal component analysis (PCA) and hierarchical clustering] or supervised (PLS) statistical approaches that reduce the dimensionality of the data [268] and even more advanced and detailed approaches will be required to extract the full wealth of information embedded in metabolomics datasets [272].

For the metabolomics investigation of PfAdoMetDC/ODC co-inhibition, samples were analysed with LC-ESI/MS for maximum sensitivity and specificity [263], which enabled the simultaneous determination of 172 water-soluble metabolites. In combination with the transcriptomics, the proteomics and metabolomics analyses of PfAdoMetDC/ODC co-inhibition, as discussed in this chapter, complete one of the first comprehensive functional genomics investigations of perturbed *Plasmodium*. In addition, biochemical assays were performed to investigate specific hypotheses such as the induction of LDC as compensatory response to polyamine depletion and the role of DNA hypermethylation in the mechanism of MDL73811 and in the transcriptional arrest observed during cytostasis. By following this integrative approach, two of three perturbation-specific compensatory mechanisms as detected in the transcriptome were confirmed on both the protein and metabolite level, which corroborated their biological relevance in the malaria parasite.

4.2 MATERIALS AND METHODS

4.2.1 Proteomics

3D7 *P. falciparum* culturing and DFMO/MDL73811 treatment were performed as discussed in section 2.2.1 and 3.2.2, but in order to obtain enough protein, an 180 ml culture of 8% parasitaemia and 3% haematocrit

was treated in duplicate (two biological replicates) alongside duplicate untreated controls. Sample volumes of 60 ml were harvested from these at approximately the same three time points as for the transcriptomics investigation (section 3.2.2). The parasites were released with the addition of 60 µl 10% (w/v) saponin (Merck) in PBS (0.01% final concentration) followed by incubation on ice for 5 min before centrifugation at 2500 g for 15 min. The erythrocyte lysates were aspirated and the parasite-containing pellets were washed four times with an equal volume of PBS before storage at -70°C.

4.2.1.1 Protein extraction and quantitation

A volume of 500 µl strip rehydration buffer [8 M urea, 2 M thiourea, 2% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 32.4 mM DTT with 0.7% (v/v) carrier ampholytes (IPG buffer, pH 3-10, GE Healthcare)] was added to each parasite pellet for complete solubilisation, denaturation and reduction of proteins and the two biological replicates were combined. The samples were pulse-sonicated with a Virsonic microtip sonifier for 10 cycles (a cycle consisted of 10 pulses of 1 s each at output level 2) with 1 min ice-incubation between cycles. Samples were subsequently centrifuged at 16000 g for 60 min at 4°C. The supernatant was transferred to a new microfuge tube (on ice) and the protein concentration was determined using the 2-D Quant kit (Amersham Biosciences), which uses a combination of a unique precipitant and co-precipitant (proprietary knowledge) to precipitate the sample protein quantitatively while interfering contaminants (carrier ampholytes, thiourea, detergents and reductants) remain in solution. The protein is pelleted by centrifugation and resuspended in an alkaline solution of cupric ions, which bind to the polypeptide backbone of the proteins present. A colorimetric agent reacts with the unbound cupric ions and the resulting colour density is inversely related to the protein concentration. This method is compatible with samples containing urea, thiourea and/or CHAPS as opposed to the regular Bradford protein assay [273]. A BSA dilution series was used to compile a standard curve of absorbance of the coloured reaction product at 492 nm against protein concentration, which was used to estimate the sample protein concentrations.

4.2.1.2 Iso-electric focussing (IEF)

For the first dimension separation of proteins, 400 µg of the total protein obtained (diluted to a volume of 340 µl with rehydration buffer) was applied to four 18 cm Immobiline DryStrip (Amersham Biosciences) IPG strips with a linear pH gradient (pH 3-10) inside ceramic strip holders (Amersham Biosciences). The use of IPG strips limits technical variation and is currently the reference method during IEF [255]. The strips were covered with 500 µl of strip covering oil (GE Healthcare) and active rehydration at 30 V was performed for 10 h using an Ettan IPGphor II iso-electric focussing system (Amersham Biosciences) at 20°C. The voltage was gradually increased in a step-and-hold manner to 8000 V within the next 1.5 h and kept at 8000 V (step 9) for 24000 volt hours (Vh, Eq. 4.1, 4.2).

Volt hours (Vh) = hours (h) x Volts (V) .....Equation 4.1

But during the gradient, $Vh = h \times \frac{(V_{previous\ step} + V_{new\ step})}{2}$ .....Equation 4.2

The run was terminated after a total of ~35000 Vh was reached (Table 4.1).

Table 4.1 Iso-electric focussing step-and-hold programme

S1	Step	30 V	10 h	300 Vh
S2	Gradient	200 V	0.10 h	19 Vh
S3	Step	200 V	0.20 h	66 Vh
S4	Gradient	500 V	0.20 h	116 Vh
S5	Step	500 V	0.20 h	166 Vh
S6	Gradient	2000 V	0.20 h	416 Vh
S7	Step	2000 V	0.45 h	1500 Vh
S8	Gradient	8000 V	1.40 h	8333 Vh
S9	Step	8000 V	Up to 24000 Vh	
TOTAL		~35000 Vh		

4.2.1.3 Two-dimensional polyacrylamide gel electrophoresis (2D-GE)

After completion of the IEF, the strips were briefly rinsed with MilliQ H_2O. The cysteine residues were reduced with 2% (w/v) DTT (Pharmacia Biotech) for 10 min with gentle shaking at 20°C followed by carbamidomethylation with 2.7% (w/v) iodoacetamide (FLUKA, SIGMA) for another 10 min. Both the DTT and iodoacetamide were dissolved in SDS equilibration buffer [50 mM Tris-HCl, pH 6.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 0.002% (w/v) bromophenol blue] and were prepared fresh prior to use. The strips were finally equilibrated for electrophoresis by 10 min incubation in the SDS equilibration buffer only and each strip was subsequently positioned on top of a 10% (w/v) vertical SDS-polyacrylamide gel and sealed with 0.5% (w/v) agarose for the second dimension separation. The gels were run in SDS electrophoresis buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1% SDS) at 80 mA (limits set at 500 V, 40 W), 20°C until the bromophenol blue front reached the bottom of the gels after 4 - 5 h using a Hoefer SE 600 vertical system with a water cooling unit. The gels were removed and fixed overnight in 40% (v/v) EtOH/10% (v/v) acetic acid, with gentle shaking at 20°C. The fix solution was decanted and each gel was stained overnight with gentle shaking at 20°C with 100 ml Flamingo fluorescent gel stain (Bio-Rad, California, USA) for visualisation of proteins. Fluorescent stains are light-sensitive and gels were covered with aluminium foil from this point forward.

4.2.1.4 Gel scanning and data analysis

The gels were scanned with a Pharos FX Plus molecular imager (Bio-Rad) at high and medium intensity (PMT voltage) to enable quantitation of low abundance as well as saturated spots. Twenty-one gels were scanned and the best three of four gels were selected (triplicate technical replicates) for spot analysis using PDQuest 8.0 Advanced (Bio-Rad) software. The gels were orientated by rotation, cropped to the same size, speckle filtered (median filter) and warped for optimal alignment. Roller-ball background subtraction, LOWESS normalisation and automated spot detection and matching were performed. The automated spot matching was carefully checked by visual inspection and comparison in order to achieve maximum consensus within each

replicate group. A master image was generated including all replicate groups to be compared (i.e. one master image of UT_{t1} compared to T_{t1}, T_{t2}, T_{t3} and another master image for UT_{t1} compared to UT_{t2}, UT_{t3}) since the software could only analyse 15 gels at a time. As with the microarray data analysis (section 3.2.8), differential protein abundance was calculated in comparison to UT_{t1} (relative t_0) with PDQuest 8.0 software. Proteins in at least one treated time point with an abundance greater than 2-fold in either direction compared to relative t_0, and p-values of less than 0.05 (Student's t-test), were regarded as differentially affected. The similarity between the replicate groups and relative t_0 was determined by plotting the data on the same graph and calculating the correlation coefficient of the regression line within PDQuest.

4.2.1.5 Spot excision, destaining and trypsin digestion for protein identification

Gels containing spots of interest were visualised on a UV transiluminator at 365 nm to enable the manual excision of the proteins, which was transferred to low adhesion microfuge tubes. Spots from technical replicate gels were combined where possible, dried under vacuum and stored at -20°C. The dessicated gel pieces were rehydrated twice with 200 µl MilliQ H_2O for 10 min. The gel pieces were then washed for 10 min with 200 µl of 50% (v/v) acetonitrile (AcN) followed by 50 mM ammoniumbicarbonate, 100% AcN, 50 mM ammoniumbicarbonate, 50% (v/v) AcN and finally 100% AcN and then centrifuged under vacuum for 5 – 10 min to remove all the traces of AcN. Depending on the size and intensity of the excised spot, the gel pieces were digested with 50 – 100 ng of sequencing grade trypsin (Promega) in its provided buffer and incubated overnight at 37°C. Trypsin cleaves proteins at the carboxyl side of the amino acids lysine and arginine, except when either is followed by proline. The supernatant was collected and a further 20 µl of 70% (v/v) AcN was added to the gel pieces for 30 min. The supernatants were pooled, dried under vacuum and resuspended in 10 – 20 µl of 10% (v/v) AcN/0.1% (v/v) formic acid, depending on the size of the gel pieces. The digested peptides were mixed 1:1 with the MALDI matrix, consisting of a saturated solution of α-cyano-4-hydroxy-cinnamic acid (HCCA, Bruker Daltonics) in 50% (v/v) AcN/0.1% (v/v) formic acid, and 1.5 µl volumes were spotted in duplicate onto a MALDI plate. MALDI-Q-TOF was performed using a QStar Elite instrument (Applied Biosystems) equipped with a MALDI source with the assistance of S. Stoychev at the CSIR Biosciences Division. The instrument was calibrated with a commercially available peptide calibration standard containing angiotensin II, angiotensin I, substance P, bombesin, ACTH clip 1-17, ACTH clip 18-39 and somatostatin 28 encompassing a mass-range of ~1000 - 3200 Da (Bruker Daltonics, Massachusetts, USA), which was spotted onto the plates. Peptide mass fingerprints (PMF) were obtained with the first MS and the resultant mass lists were compared to a non-redundant protein database (Swiss-Prot/TrEMBL, http://ca.expassy.org/sprot) using MASCOT (www.matrixscience.com). Oxidation (methionine) and carbamidomethylation were set as variable and fixed protein modifications respectively, only one missed trypsin cleavage was allowed and the mass tolerance was 50 parts per million [ppm = (experimental mass in daltons − theoretical mass)/(theoretical mass expressed in parts per million), equivalent to 0.1 D for a 2 kD peptide][274]. The probability based MOWSE score (-10*LogP, with P the probability of the match being a random event) was used to estimate the

significance of the identification and was set at a confidence threshold of 5% ($p<0.05$). The second mass analyser (Q-TOF) accelerated the peptides, resulting in collision with nitrogen gas particles and fragmentation, providing a higher resolution fingerprint, which was searched in a similar manner using MASCOT to confirm the initial identification.

4.2.2 Metabolomics

This experiment was performed in collaboration with K.L. Olszewski and D. Willinski, at Princeton University (Princeton, New Jersey, USA). DFMO/MDL73811-treatment of 60 ml 3D7 *P. falciparum* cultures of 9% parasitaemia and 2% haematocrit was performed in duplicate (two biological replicates) alongside duplicate 60 ml untreated controls as discussed in section 3.2.2. Samples of 20 ml were harvested at approximately the same three time points as for the transcriptomics and proteomics investigations. The metabolite extraction and LC-MS/MS methodology were previously developed and optimised for the simultaneous relative quantitation of water-soluble metabolites [263, 266].

4.2.2.1 Metabolite extraction and polyamine derivatisation

Of 20 ml cultured *P. falciparum* samples (two biological replicates), 10 ml was used for general metabolite analysis and 10 ml was derivatised for polyamine analysis. The cultures were pelleted by centrifugation at 2000 g for 5 min at room temperature and metabolites were serially extracted immediately thereafter. Four pellet volumes of -75°C 100% MeOH (e.g. 2 ml to a 500 µl pellet) were added at the start to flash-freeze all metabolic activity, followed with incubation on dry ice for 15 min with vigorous vortexing every 5 min. The sample was centrifuged for 5 min at 11000 g at 4°C and the supernantant removed and stored on ice. A volume of 80% (v/v) ice-cold MeOH was added to the pellet, followed with vortexing and sonification on ice in a water bath sonicator for 15 min. The sample was again centrifuged for 5 min at 11000 g at 4°C and the supernatant removed and added to the previous aliquot on ice. The last extraction step with 80% (v/v) MeOH was repeated and the supernatants from each extraction were pooled and centrifuged free of cell debris and protein. Culture medium was sampled as background control and these samples were treated in the same way as the culture samples by the addition of four volumes of -75°C 100% MeOH, then two volumes of 80% (v/v) MeOH to obtain the same dilution factor as the pellet extracts and subsequently centrifuged free of protein. All samples were analysed within 24 h of their generation and were analysed for 167 metabolites as described [263, 266, 267], as well as putrescine, spermidine, spermine, cadaverine and dcAdoMet. The synthetic dcAdoMet standard used for calibration was a kind gift from K. Samejima (Josai University, Japan); the other standards were obtained commercially.

For polyamine analysis, succinic anhydride derivatisation was performed to induce a negative charge on these polycationic molecules to retain them on the aminopropyl column during LC. A volume of 10 µl of triethylamine was added to 100 µl of cell extract and mixed, followed with a few crystals of solid succinic anhydride and vigorous vortex mixing. The derivatisation reaction was performed for 1 h at room temperature, with vortexing

every 10 min. After this incubation the samples were centrifuged to pellet any remaining solid succinic anhydride and the liquid portion of the sample was subjected to LC-MS/MS analysis using parameters previously determined with pure stock solutions of putrescine, cadaverine, spermidine and spermine.

4.2.2.2 LC-MS/MS metabolite analysis

LC–MS/MS was performed using an LC-10A HPLC system (Shimadzu Corporation, Kyoto, Japan) and a Luna aminopropyl column (250 mm x 2 mm with a 5 µm particle size (Phenomenex, California, USA) coupled to the mass spectrometer. Quantitation and run-to-run variability were controlled by the 1:10 addition of an internal standard mixture containing $^{13}C/^{15}N$ isotopes (alanine, aspartate, glutamate, glutamine, isoleucine/leucine, methionine, phenylalanine, serine, threonine, tyrosine and valine) encompassing a mass-range distributed over the entire chromatographic run. The LC parameters were as follows: autosampler temperature, 4 °C; injection volume, 20 µl; column temperature, 15 °C and flow rate, 150 µl/min. The LC solvents were Solvent A [20 mM ammonium acetate and 20 mM ammonium hydroxide in 5% (v/v) AcN (pH 9.45)] and Solvent B (100% AcN). The gradients were as follows: positive mode, t = 0, 85% B; t = 15 min, 0% B; t = 28 min, 0% B; t = 30 min, 85% B; t = 40 min, 85% B; and negative mode, t = 0, 85% B; t = 15 min, 0% B; t = 38 min, 0% B; t = 40 min, 85% B; t = 50 min, 85% B.

MS analyses were performed on a Finnigan TSQ Quantum Ultra triple-quadrupole MS (Thermo Electron Corporation, Massachusetts, USA), equipped with an ESI source. The ESI spray voltage was 3200 V in positive mode and 3000 V in negative mode. Nitrogen was used as sheath gas at 30 psi and as the auxiliary gas at 10 psi, and argon as the collision gas at 1.5 mTorr, with the capillary temperature 325°C. Scan time for each single reaction monitoring (SRM) event transition was 0.1 s with a scan width of 1 m/z. SRM involves selecting for ions of a specified parent molecular weight (m/z), fragmenting the parent ion at optimal collision energy to produce a particular daughter ion and then quantitating the production of ions of the daughter mass. The scanning of multiple SRM events enables the simultaneous measurement of numerous compounds in a single LC run. The LC runs were divided into time segments, with the SRM scans within each time segment limited to those compounds eluting during that time interval. For compounds eluting at the boundaries between time segments, the SRM scan corresponding to a specific compound was conducted in both time segments. The instrument control, data acquisition and data analysis were performed by Xcalibar software (Thermo Electron Corporation, version 1.4 SR1), which also controlled the chromatography system.

4.2.2.3 Metabolomics data analysis

Raw data peak quantitation was performed by Xcalibar software using 10^3 as bottom quantitation limit. Biological replicate data were averaged (arithmetic mean), background subtracted and compared to the relative t_0 in EXCEL, to calculate fold change as a result of the perturbation. Treated parasite data was also compared to their parallel time point untreated controls, to determine the validity of the relative t_0 strategy for

metabolite analysis. Metabolites with a fold change of 2 were regarded as significantly changed. Pearson correlations of the complete metabolic profiles and excluding several metabolites present in excess in the medium (arginine, asparagine, aspartic acid, cystine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine, choline, niacinamide, pantothenic acid, riboflavine, thiamine, D-glucose) were calculated in EXCEL. This measure of similarity is commonly used in global expression profiling analyses [29, 62, 91].

4.2.3 Decarboxylase activity assays

LDC activity of DFMO/MDL73811-treated and untreated culture lysates was assessed by measuring $^{14}CO_2$-release according to the methodology of Assaraf and colleagues [88], as discussed for AdoMetDC and ODC in section 3.2.1.2. L-[1-^{14}C]lysine (56 mCi/mmol, American Radiolabelled Chemicals) were incubated with the lysates from treated and untreated 3D7 cultures (10% parasitaemia, 3% haematocrit) sampled in the early (19 hpi) and mature (34 hpi) trophozoite stages. Samples of 10 - 15 ml were centrifuged at 2500 g for 5 min, the pellet washed three times with an equal volume of PBS and then 500 µl was transferred to a cryotube and stored at -70°C. Uninfected erythrocytes were sampled and processed in the same way to serve as a negative control. A volume of 1 ml of buffer A [88] or LDC buffer [0.5M sodium acetate, 1 mM DTT, 1 mM EDTA, 0.1 mM PLP, pH = 6] [275] was added, the samples were freeze-thawed three times (alternating between -70°C and 37°C) and then centrifuged at 8000 g for 20 min at 4°C. The rest of the assay was performed as described in section 3.2.1.2, but the 50 µl reaction mixture that was added to 200 µl cell lysate consisted of 40 µM ^{14}C L-lysine (100 nCi) and 40 µM PLP in either buffer A or LDC buffer. Assays were allowed to take place at 37°C for 60 min in a ZHWY-110X shaking water bath (Shanghai ZHICHENG Analytical Instruments Manufacturing Co.).

4.2.3.1 LDC induction in *E. coli* as assay positive control

BL21(DE3) *E. coli* cells were grown overnight at 37°C to saturation in Luria-Bertani (LB)-broth [1% (w/v) tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl, pH 7.5] and then diluted 1:4 with either LB-broth or modified Falkow (F-MES) medium [0.5% (w/v) peptone, 0.3% (w/v) yeast extract, 0.1% (w/v) D-glucose, 0.5% (w/v) D,L-lysine and 100 mM 2-(N-morpholino) ethanesulfonic acid, pH 5.2] [275, 276]. The cells in LB-broth were grown for 6 h with aeration (agitation at 175 rpm at 37°C) and those in F-MES for 2 h with aeration followed by 4 h without (stationary at 37°C). The cultures were centrifuged at 2000 g for 20 min, the pellets were washed once with five volumes of PBS and stored at -70°C. A cell pellet of 500 µl was resuspended in either buffer A or LDC buffer and incubated with 0.1 mg lyzozyme (Roche Diagnostics) for 30 min. The samples were pulse-sonicated with a Virsonic microtip sonifier for seven cycles as described in section 4.2.1.1, to ensure complete lysis, followed by centrifugation at 20000 g for 30 min at 4°C. The cell lysate supernatants were aspirated and only 10 µl bacterial lysate was included per 250 µl reaction.

4.2.4 Methylation status determination

4.2.4.1 CpG island analysis of the differential transcript abundance list

CpG island detection and analysis of the increased and decreased transcript abundance lists were performed in EMBOSS (http://emboss.sourceforge.net /apps/release/5.0/emboss/apps/nucleic _cpg_islands_group.html) using a batch file. The EMBOSS nucleic CpG island applications CpGplot, CpGreport and geecee-count were performed. In CpGplot, the observed number of CpGs is the count of the number of cytosines directly followed by a guanine, whereas the expected number in a window is the number of CpG dinucleotides expected based on the frequency of cytosines and guanines in that window (http://www.sacs.ucsf.edu/ Documentation/emboss/cpgplot.html). CpGplot identifies CpG islands over an average of 10 windows, where the percentage composition of cytosines and guanines is over 50% and the calculated observed/expected ratio is over 0.6 for a minimum of 200 nucleotides. CpGreport scans nucleotide sequences for regions with higher than the expected frequency of the CpG dinucleotide (http://www.sacs.ucsf.edu/ Documentation/emboss/cpgreport.html) and geecee-count calculates the fraction of guanine and cytosine nucleotides within a nucleotide sequence (http://www.sacs.ucsf.edu/Documentation/emboss/ceegee.html).

4.2.4.2 Global methylation assays

4.2.4.2.1 gDNA isolation

The methodology applied was similar to the strategy reported for investigating the methylation status of another protozoan, *Entamoeba hystolytica* [277]. Volumes of 5 ml treated and untreated *P. falciparum* culture (10% parasitaemia, 3% haematocrit) were sampled for gDNA isolation in the early (19 hpi) and mature (34 hpi) trophozoite stages, as described in section 4.2.3. The samples were pelleted by centrifugation at 2500 g for 5 min, the pellet washed with an equal volume of PBS and then stored at -70°C. gDNA was isolated with the QIAamp Blood Mini kit (QIAGEN). This kit works on the same principle as discussed in section 3.2.4 for cDNA purification. The erythrocytes were lysed by chaotropic salts and detergents and the cellular debris was removed by filtration. The DNA was purified from the soluble part of the lysate by binding to a silica matrix [211], which was then washed with alcohol-based buffers to remove the salts. The membrane-bound DNA was treated with RNase A (Fermentas Life Sciences) for 10 min to remove contaminating RNA before the final elution of the gDNA with 200 µl water. The concentration was determined by measuring the absorbance at 260 nm by UV spectrophotometry with a NanoDrop-1000. For double-stranded DNA, one absorbancy unit equals 50 ng/µl [209]. As discussed in section 3.2.3, DNA purity was estimated from the 260 nm/280 nm ratio, which should preferably be between 1.7 – 1.9 [209].

4.2.4.2.2 Methylation negative and positive controls

Synthetic DNA was prepared by amplification of the OAT (PFF0435w) gene from a pET-15b (Novagen, Merck, Darmstadt, Germany) construct provided by K. Clark, to serve as methylation-free negative control. Plasmid DNA (20 ng) was amplified with 5 U Taq DNA polymerase (New England Biolabs), 200 µM dNTPs, 20 pmoles

of primers (forward: 5'-CTCGAGGATTTCGTTAAAGAATTAAAAAG-3' and reverse: 5'-GCTCAGCCTCAGTTATAAGTTGTCATC-3') in standard buffer (New England Biolabs) in a total reaction volume of 100 µl. After denaturation at 94°C, the DNA was amplified for 30 cycles (94°C for 30 s, 56 °C for 30 s, 72°C for 2 min) with a final 5 min extension step at 72°C. Gel electrophoresis analysis of the amplified product was performed with a 1.2% (w/v) agarose in TAE (40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA), pH 8) gel containing 0.5 µg/ml ethidium bromide. A voltage of 5 - 7 V/cm was applied (~80 V) for 45 min and the gel was visualised on a UV transiluminator (Spectroline TC-312 A) at 312 nm. The amplification product was subsequently purified with the QIAquick PCR purification kit (QIAGEN) as described in section 3.6.2. OAT DNA (2.5 µg) was methylated with 12 U M.SssI CpG methyltransferase (New England Biolabs) and 20 nmoles AdoMet in 62.5µl at 37°C for 2 h to be used as a methylation positive control. The methylated OAT DNA was again purified from excess AdoMet with the QIAquick PCR purification kit (QIAGEN).

4.2.4.2.3 Restriction-enzyme digestion to assess gDNA methylation

The isolated gDNA from treated and untreated *P. falciparum* cultures was digested with methylation-sensitive restriction enzymes to reveal gDNA methylation patterns. Samples of 250 ng gDNA as well as 250 ng methylated and unmethylated OAT DNA were digested with 10 U *Hpa*II (Fermentas Life Sciences) and 20 U *Dpn*I (Fermentas Life Sciences) overnight at 37°C. Cytosine methylation within the recognition site protects against *Hpa*II digestion (C↓CGG), whereas *Dpn*I digestion only occurs when the adenine in its recognition site is methylated (GA↓TC). The digestion products were separated on a 0.8% (w/v) agarose gel in TAE buffer containing 0.5 µg/ml ethidium bromide as described in section 3.6.2 and 4.2.4.2.2.

4.2.4.2.4 South-Western immunoblotting

gDNA samples (UT_{t1}, UT_{t3}, T_{t1} and T_{t3}) were denatured by boiling for 5 min, followed by cooling on ice for 5 min. To ensure that signal intensity increased with increasing gDNA concentration, 500 ng and 1000 ng quantities were spotted onto a positively charged nylon membrane (Roche Diagnostics) along with 250 ng methylated and unmethylated OAT (prepared in section 4.2.4.2.2) and 1 µl of 0.01 M 5-methylcytidine (SIGMA) as controls. The DNA was UV cross-linked for 3 min on a UV transiluminator (Spectroline TC-312 A) at 312 nm and then blocked overnight at 4°C with 2% (w/v) BSA (Roche Diagnostics) in PBS inside a small plastic resealable bag. The blocking solution was discarded and the membrane was incubated overnight at room temperature with gentle shaking in a 0.1 µg/ml dilution of mouse anti-5-methylcytidine monoclonal IgG (AbD Serotec, Oxford, UK) in washing buffer [2% (w/v) BSA, 0.1% Tween-20 in PBS] according to the manufacturer's instructions. The membrane was washed extensively with three 10 min incubation steps in washing buffer at room temperature with gentle shaking. The membrane was subsequently incubated for 1 h with the anti-mouse hrp-conjugate followed by four 5 min washes in washing buffer, three 5 min washes with 0.1% Tween-20 in PBS and finally two 5 min washes in PBS only. All the wash steps were performed at room temperature with gentle shaking. The SuperSignal West Pico Chemiluminescence kit (Pierce, Illinois, USA)

was used as substrate for the oxidation reaction catalysed by the hrp-conjugate, as described in section 2.2.3.2.2. Usually, 4 ml of hydrogen peroxide and 4 ml of luminol provided in the kit was mixed just before use and incubated with the membrane for 5 min at room temperature with gentle shaking. Excess SuperSignal solution was removed to minimise background luminescence. The membrane was exposed to Hyperfilm ECL X-ray film for high performance chemiluminescence (Amersham Biosciences) for 5 – 6 h in the dark. The X-ray film was developed in Universal Paper Developer (Ilford) for 1.5 – 3 min until the spots became visible, briefly rinsed in MilliQ H_2O and then fixed with Rapid Paper Fixer (Ilford) for 1 min. The film was subsequently rinsed with MilliQ H_2O and left to dry at room temperature. X-ray films were scanned with a VersaDoc scanner (Bio-Rad) and densitometry analysis was performed with Quantity One (Bio-Rad) software.

4.3 RESULTS

4.3.1 Proteomics analysis of PfAdoMetDC/ODC co-inhibited *P. falciparum*

To validate the findings of the transcriptome investigation (Chapter 3), PfAdoMetDC/ODC co-inhibition was repeated and the effects of polyamine depletion and cytostasis on the parasite proteome were determined with 2D-GE. The two biological replicates of the experiment were pooled to have enough protein for three to four technical replicates, which minimised technical variability and enabled differential protein abundance analysis. A master image was generated, which represented all the spots across all the replicate groups to be compared. However, the software could analyse a maximum of 15 gels at a time and therefore one master image was prepared for UT_{t1} compared to T_{t1}, T_{t2}, T_{t3} (Fig. 4.1) and another for UT_{t1} compared to UT_{t2}, UT_{t3} (results not shown).

As was observed in the transcriptome, cytostasis resulted in a high correlation between the gels of the relative t_0 (UT_{t1}) and the three treated time points (UT_{t1} versus T_{t1}: R = 0.93; UT_{t1} versus T_{t2}: R = 0.88; UT_{t1} versus T_{t3}: R = 0.88) and a lower correlation with UT_{t2} and UT_{t3} (UT_{t1} versus UT_{t2}: R = 0.78; UT_{t1} versus UT_{t3}: R = 0.70; Table 4.2). These correlation values indicate that the cytostatic effects were more subtle in the proteome than in the transcriptome (section 3.3.4.3), but the perturbation caused an overall decrease in the number of proteins detected over the time course (UT_{t1} = 483 spots; T_{t1} = 461; T_{t2} = 409; T_{t3} = 416). A similar result was detected in the transcriptomics investigation where the majority of differentially affected transcripts were decreased, as described in section 3.3.5.2.

Table 4.2 Correlation of the 2D-GE data across replicates groups

Comparison	Correlation of regression line (R)[a]
UT_{t1}:T_{t1}	0.93
UT_{t1}:T_{t2}	0.88
UT_{t1}:T_{t3}	0.88
UT_{t1}:UT_{t2}	0.78
UT_{t1}:UT_{t3}	0.70

a. The correlation coefficient of the regression line when the respective replicate groups were plotted on the same graph in PDQuest.

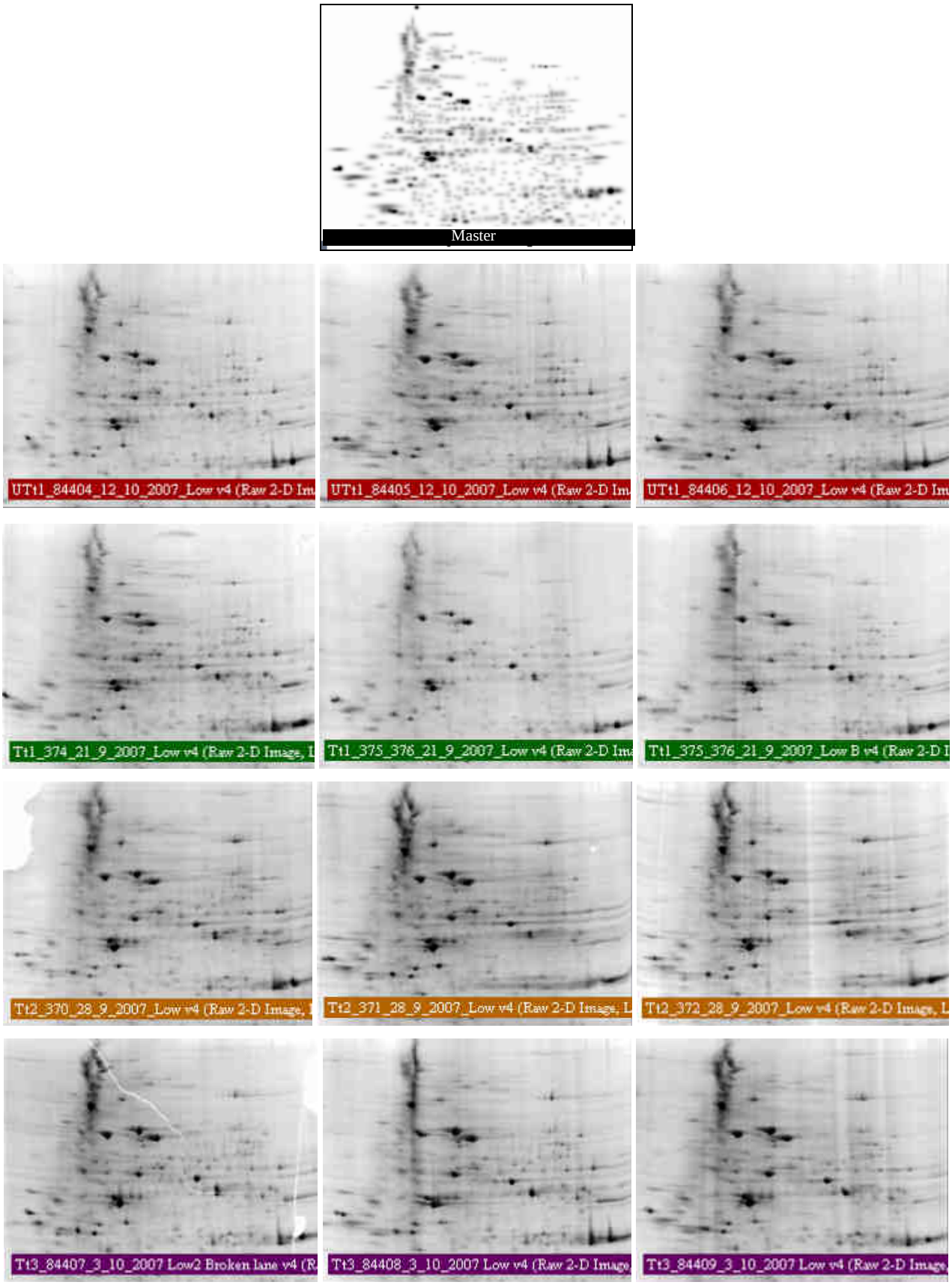


Fig. 4.1 The master image and actual images of the three best 2D-GE technical replicates used for analysis of UT_{t1} versus T_{t1}, T_{t2} and T_{t3}.

4.3.1.1 Differential protein abundance analysis and protein identification

As in the transcriptomics investigation, the differential abundance of proteins was calculated compared to UT_{t1}, which was previously defined as the relative t_0, a reference point just prior to the transcriptional arrest (section 3.3.5). Due to the translational gap or time delay between the the appearance of the transcriptional peaks and their corresponding protein peaks [246], the relative t_0 of the proteins could arguably be later than for the transcripts. However, a high resolution reference proteome dataset, similar to the IDC transcriptome [29, 91] is still lacking [187] and a similar calculation of the approximate time of the proteomics relative t_0 could not be performed. Hence, for comparison purposes, the sampling times of the transcriptomics and proteomics investigations were kept the same and fold changes were again calculated compared to UT_{t1} with the differential abundance criteria 2-fold change and $p<0.05$. By following this approach, 53 spots (26 spots with an increasing and 27 spots with a decreasing profile) were regarded as differentially affected, but only 41 spots were visible to the eye for manual excision (Fig. 4.2, Appendix D).

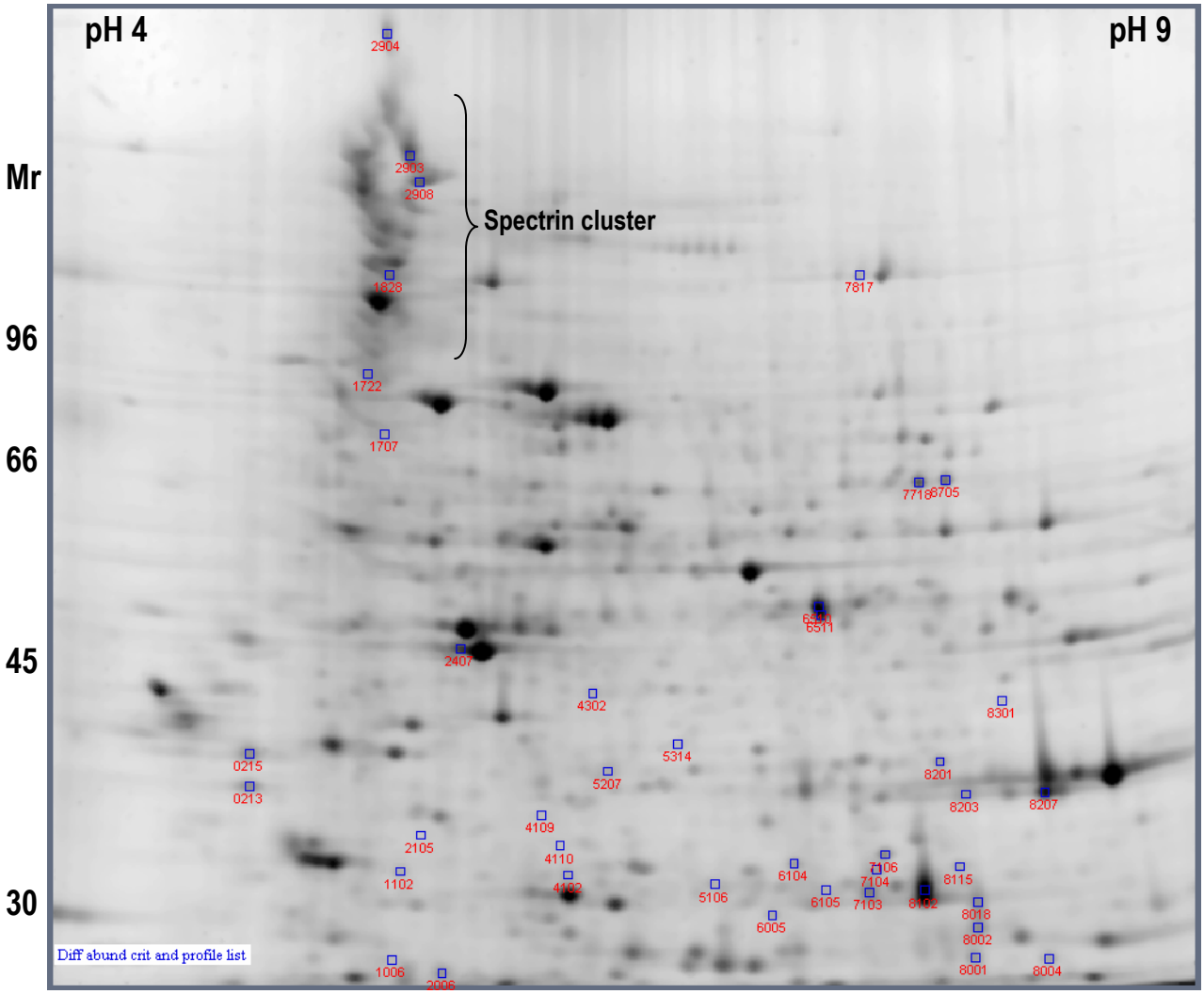


Fig. 4.2. A typical gel (UTt1_84404) indicating the 41 spots with differential abundance that were visible to the eye and could be excised and processed for identification. The molecular weight (Mr) in kDa and pH variation across the gel is displayed. Note that most spots of interest were of low molecular weight and low intensity, which hampered successful identification. The standard spot numbers (SSP) as designated by PDQuest are displayed.

The excised protein spots were cleaved into smaller peptides with trypsin and the absolute masses of these were then determined by MALDI-Q-TOF, which separates singly charged peptides based on m/z, resulting in a sequence-specific PMF (Fig 4.3A) [278]. The experimentally obtained PMF were compared *in silico* with the theoretical/calculated peptide masses of proteins stored in Swiss-Prot/TrEMBL by means of the search engine MASCOT using a threshold of 5% (MOWSE score ≥ 78 i.e. $p<0.05$). The results were statistically analysed and possible matches were indicated. However, spots excised from 2D-GE are not necessarily pure proteins, but may contain protein mixtures with the same Mr and pI characteristics, which can complicate PMF analysis. Furthermore, different peptides can sometimes have the same absolute masses, resulting in similar PMF spectra. Therefore, MS/MS validation of PMF identities was performed by fragmenting the 50 most prominent PMF peptide ions into daughter ions via acceleration and collision with nitrogen gas in the second mass analyser (Q-TOF) and again separating them according to m/z (Fig. 4.3B). These fragmented daughter ions were used to deduce the amino acid sequences of the parent peptides (Fig. 4.3B). Note that peptide fragments are indicated by a, b or c when the charge is retained on the N-terminus and x, y or z when the charge is on the C-terminus and the subscript indicates the number of amino acids in that particular fragment.

However, since the majority of the spots had low molecular weight (<30 kDa) and low intensity, only 11 proteins (27%) were successfully identified with MALDI-Q-TOF (Table 4.3). This number is in agreement with the spot identification reported in another 2D-GE plasmodial study where only 25% (50/200) of the excised proteins could be identified [251]. Although the number of spots with increasing (26 spots) and decreasing (27 spots) profiles were almost equal, 9 of the 11 proteins that could be identified showed an increase in abundance over the time course, which enhanced their identification (Appendix D).

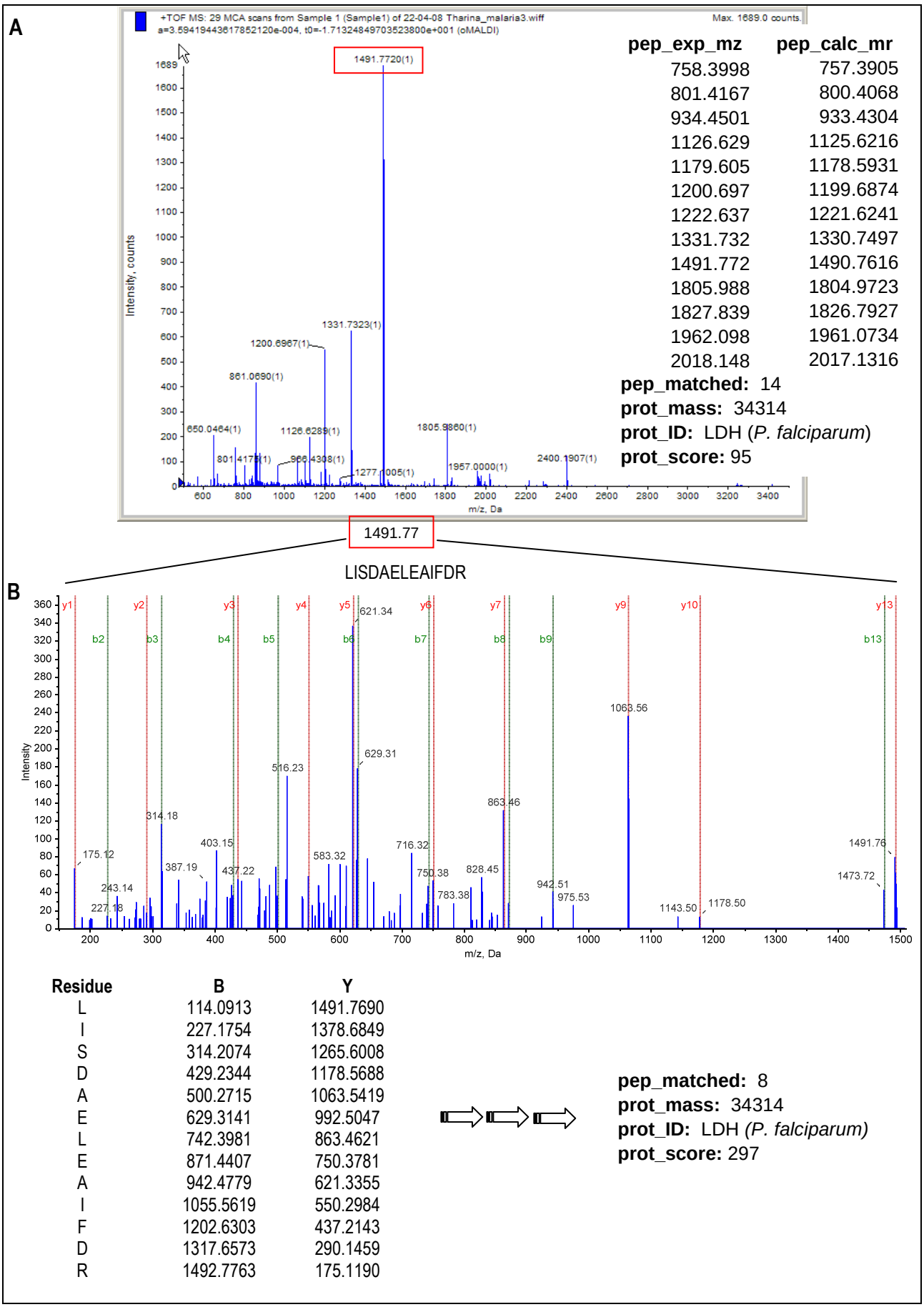


Fig. 4.3 MALDI-Q-TOF MS/MS protein identification of LDH as an example. **A)** PMF spectrum indicating the experimental and calculated molecular mass of the most prominent peptides after the tryptic digest, which was identified by MASCOT as LDH (prot_score or MOWSE score = 95). **B)** MS/MS scan indicating the amino acid sequence LISDAELEAIFDR of the parent peptide with mass 1491.77. The amino acid sequences of eight peptides were deduced in this manner, which confirmed the protein identity as LDH (MOWSE score = 297).

Table 4.3 Identification and characterisation of a subset of proteins with differential abundance

						Peptide mass fingerprinting (PMF)				MS/MS confirmation			
SSP[a]	Annotation	Fold-change[b]	M_r (kDa)	pI	PlasmoDB	Protein ID[c]	Pep.[d] no.	%Seq. cover.	MOWSE[e] score	Protein ID[c]	Pep.[d] no.	%Seq. cover.	MOWSE[e] score
1102	Falcipain-2	2.9	27.6	4.96	PF11_0165	Q3HTL5-PLAFA	11	44	71	Q3HTL5-PLAFA	4	13	134
1828	Human erythroid α-spectrin	4.7	282.0	4.98	-	-	-	-	-	SJHUA	10	5	358
2407	Human beta-actin (fragment)	2.7	41.3	5.56	-	Q96HG5_HUMAN	21	69	106	ACTB_CAMDR	8	32	266
2903	Human erythroid α-spectrin	3.3	280.9	4.98	-	Q5VYL2_HUMAN	52	30	110	Q5VYL1_HUMAN	14	8	631
2904	Human erythroid α-spectrin	5.9	280.9	4.98	-	SPTA1_HUMAN	56	24	80	SJHUA	9	5	201
2908	Human erythroid α-spectrin	2.7	280.7	4.98	-	Q5VYL2_HUMAN	46	26	100	Q5VYL1_HUMAN	17	11	652
6510	AdoMet synthetase	-2.1	45.3	6.28	PFI1090w	Q9GN14_PLAFA	15	43	119	Q9GN14_PLAFA	12	40	455
6511	OAT	2.2	46.9	6.47	PFF0435w	AAG44560	20	53	215	CR382399 NID	11	23	332
7817	Elongation factor 2	2.8	94.5	6.36	PF14_0486	Q8IKW5_PLAF7	34	32	81	Q9NDT2_PLAFA	7	13	271
8201	Pdx1 homologue	2.5	33.4	6.76	PFF1025c	Q3V7I1_PLAF7	17	48	79	Q3V7I1_PLAF7	6	31	170
8207	LDH [f]	2.8	34.3	7.12	PF13_0141	Q71T02_PLAFA	14	49	95	Q71T02_PLAFA	8	33	297

a. SSP is the standard spot number designated by PDQuest

b. Fold change calculated at the time point of maximum change

c. MASCOT search protein identifier

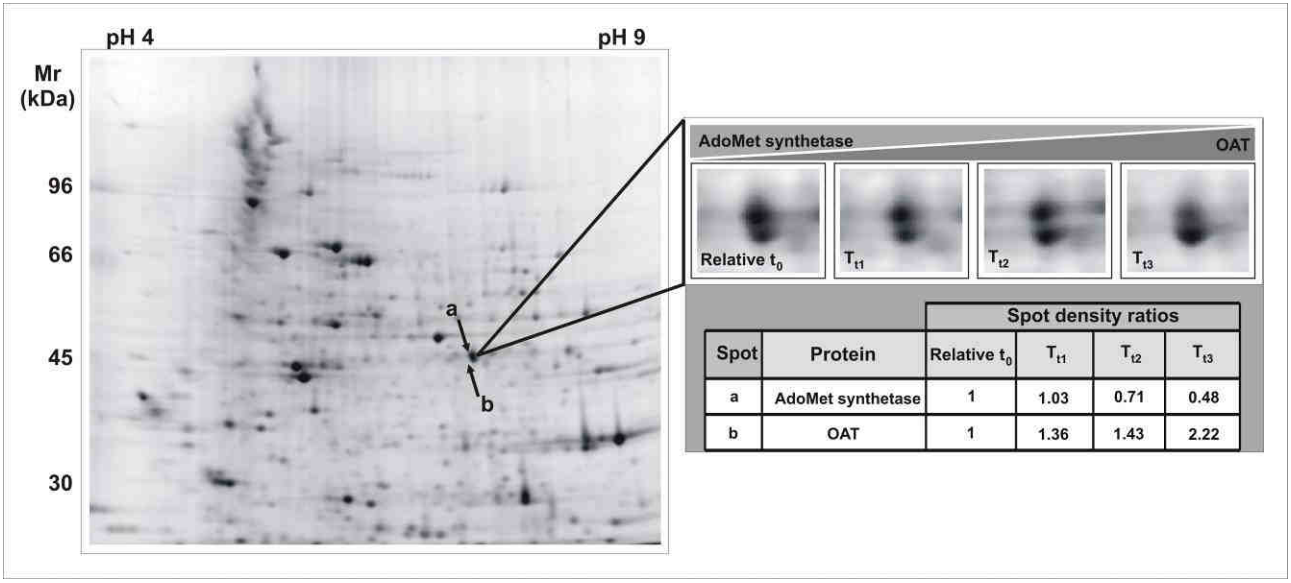
d. Number of peptides identified in the mass spectrum.

e. A MOWSE score of 78 indicates a threshold of 5% ($p<0.05$)

f. Note that LDH decreased from 2.8 to unchanged during the time course compared to the relative t_0

4.3.1.2 Perturbation-specific compensatory mechanisms confirmed in the proteome

Despite the limited number of spots with positive identities, meaningful results were obtained, including proteins involved in polyamine and methionine metabolism (Table 4.3). These include AdoMet synthetase, OAT and PLP synthase (pdx1, PFF1025c). The differential protein abundance of OAT and AdoMet synthetase (Fig. 4.4, Table 4.3) correlated with their transcript abundance after PfAdoMetDC/ODC co-inhibition (Table 3.3) and confirmed these compensatory mechanisms in the proteome. The ~2-fold increase in the OAT transcript corresponded with a 2.2-fold increase in the protein and the 2.4-fold decrease in the transcript for AdoMet synthetase corresponded with a 2.1-fold decrease in this protein (Fig. 4.4). The third compensatory mechanism proposed in the transcriptomics investigation, namely the induction of LDC, could not be assessed with 2D-GE due to the protein size (280 kDa), which prohibited gel penetration. However, pdx1 was increased 2.5-fold and synthesises PLP, which is an important co-factor for both PfAdoMetDC/ODC and LDC [279]. The change seen on the transcript level was delayed on the protein level in the case of LDH (Appendix D). The LDH transcript was decreased 1.7-fold and the protein was increased 2.8-fold in T_{t1}, but during the time course the protein gradually decreased (2.8-fold) to the same level as the relative t_0 (i.e. unchanged) in T_{t3}. LDH was also identified in another plasmodial study to show a delay between mRNA and protein accumulation due to the translational gap or time required for translation to occur [246].



Spot	Protein	Spot density ratios: Relative t_0	T_{t1}	T_{t2}	T_{t3}
a	AdoMet synthetase	1	1.03	0.71	0.48
b	OAT	1	1.36	1.43	2.22

Fig. 4.4. A typical gel (UTt1_84404) with an enlarged view of (a) AdoMet synthetase and (b) OAT over the time course, including the respective spot densities. Spot density ratios were calculated compared to UT_{t1} (relative t0) with $p<0.05$.

Other proteins that were differentially affected as a result of the perturbation include falcipain-2 (PF11_0165) and elongation factor 2 (PF14_0486). Both these proteins are involved with protein synthesis (falcipain-2 is the principal hemoglobinase in trophozoites and supplies amino acids for translation [280]) and both were significantly increased. Moreover, GO analysis in the transcriptomics investigation (section 3.3.6) determined that 10% of the transcripts with increased abundance were implicated in translation. Falcipain-2 also cleaves human ankyrin and protein 4.1 near their carboxyl termini to enable merozoite release, resulting in reduced

association of spectrin and actin with the erythrocyte membrane. Interestingly, contaminating human spectrin and actin were both found to increase over the time course, which could be related to the increased levels of falcipain-2 (Table 4.3). Human erythroid spectrin (280 kDa) contains two subunits (α and β chains) that are each composed of 21 and 16 repetitive units respectively [281]. Truncation forms were detected as a diamond-shaped pattern from 95-150 kDa (Fig. 4.2) and were also reported in another plasmodial 2D-GE study [261].

4.3.2 Metabolomics analysis of PfAdoMetDC/ODC co-inhibited *P. falciparum*

To probe and/or validate the findings of the transcriptome investigation further, another independent PfAdoMetDC/ODC co-inhibition experiment was conducted and the effects of the perturbation were assessed on the parasite metabolome. The perturbation conditions (parasite inoculum, treatment and sampling times) of the three independently performed functional genomics experiments were replicated as far as possible. The metabolite extraction and LC-MS/MS methodology were previously developed and optimised for the simmultaneous measurement of water-soluble metabolites [263, 266]. DFMO/MDL73811-treated and untreated control samples were analysed for 172 metabolites (including putrescine, spermidine, spermine, cadaverine and dcAdoMet) with LC-MS/MS [263, 266]. Reliable data were obtained for 92 metabolites (Appendix E) with the balance excluded due to levels below the detection or set quantitation limit. The latter included several relevant metabolites, such as cadaverine, spermine, dcAdoMet and PLP, that could not be detected.

As opposed to the transcriptome and the proteome, Pearson correlations revealed that the effects of cytostasis were less pronounced in the metabolome and the metabolic profiles of treated and untreated parasites were often similar. Pearson correlation between the metabolic profiles of the treated parasites and the relative t_0 (T_{t1} versus UT_{t1}: r = 0.99; T_{t2} versus UT_{t1}: r = 0.97; T_{t3} versus UT_{t1}: r = 0.96) as well as the untreated parasites at t_2 and t_3 (UT_{t1} versus UT_{t2}: r = 0.97; UT_{t1} versus UT_{t3}: r = 0.96, Table 4.4) were close to perfect correlation. However, to exclude the interference of excess amino acids, vitamins and glucose supplemented in the culture medium on the abovementioned correlations, these were recalculated without 16 amino acids, 5 co-factors, glucose-1-phosphate and glucose-6-phosphate (UT_{t1} versus T_{t1}: r = 0.99; UT_{t1} versus T_{t2}: r = 0.96; UT_{t1} versus T_{t3}: r = 0.94; UT_{t1} versus UT_{t2}: r = 0.96; UT_{t1} versus UT_{t3}: r = 0.93, Table 4.4). The close correlations between all samples, treated and untreated, indicate the maintenance of metabolic homeostasis during cytostasis.

Table 4.4 Pearson correlation of the metabolite data

Comparison	Pearson correlation (r)	Pearson correlation (r) excl. excess medium components
UT_{t1}:T_{t1}	0.99	0.99
UT_{t1}:T_{t2}	0.97	0.96
UT_{t1}:T_{t3}	0.96	0.94
UT_{t1}:UT_{t2}	0.97	0.96
UT_{t1}:UT_{t3}	0.96	0.93

4.3.2.1 Perturbation-specific compensatory mechanisms confirmed in the metabolome

Differential metabolite abundance during the time course was again quantitated compared to the relative t_0 and 24 metabolites were changed at least 2-fold after PfAdoMetDC/ODC co-inhibition (Table 4.5). Similar to the transcriptomics investigation (section 3.3.5.2), most of these (67%) were decreased.

Table 4.5 Metabolites with differential abundance (i.e. fold changes of more than 2 in either direction) in treated and untreated samples (relative t_0 comparison)

Metabolite	Fold change compared to relative t_0				
	UT_{t2}	UT_{t3}	T_{t1}	T_{t2}	T_{t3}
2-Methylglutaric acid	4.9	7.0	1.7	3.2	4.2
γ-Aminobutyrate (GABA)	29.0	59.6	7.7	10.4	11.0
α-Ketoglutarate	4.5	5.8	1.2	2.9	3.9
D-Sedoheptulose-7-phosphate	2.7	3.5	1.5	2.1	1.9
Glutathione disulfide	2.0	2.0	1.2	2.1	2.1
Orotate	2.9	4.9	1.9	2.1	2.8
Pipecolic acid	4.0	8.9	-1.2	1.2	2.6
Pyruvate	10.3	23.6	2.1	6.4	10.9
1,3-Diphosphateglycerate	1.1	-10.4	-2.3	-1.2	-2.7
5-Methylthioinosine	5.8	7.4	-4.6	-8.7	-3.9
Adenosine	-4.7	-3.3	1.1	-1.9	-4.5
D-Glyceraldehdye-3-phosphate	-8.5	-9.6	-1.9	-7.2	-7.3
Dihydroxy-acetone phosphate	-9.0	-10.3	-1.9	-6.2	-7.7
Fructose-1,6-bisphosphate	-1.3	-3.5	-1.7	-1.3	-2.1
Glutathione	-2.0	-2.2	-1.1	-2.5	-2.4
Indole	-1.7	1.1	-1.6	-2.6	-1.9
Methionine	-1.3	-1.7	-2.1	-2.1	-2.2
NADH	-28.5	-43.6	-2.0	-33.2	-209.1
Putrescine	1.1	1.8	ND	ND	-4.6
Pyridoxine	-1.6	-2.2	-1.3	-1.7	-2.0
AdoHcy	-8.4	4.6	-3.8	-1.8	-1.1
Spermidine	1.9	1.4	-5.6	-3.0	-2.6
Typtophan	-1.3	-1.3	-1.7	-2.2	-2.0
UTP	1.5	1.5	-2.4	1.2	1.6

Metabolites changed >2-fold in either direction are color-indicated (red = more than 2-fold increased, green = more than 2-fold decreased). ND = not detected

Compared to the relative t_0, the majority of the differentially affected metabolites in the treated parasites were similarly affected in the untreated controls (Table 4.5). Therefore, these changes were most likely cell cycle related peaks and not the result of the perturbation. Compellingly, the perturbation-specific effects of PfAdoMetDC/ODC co-inhibition were observed in the parasite metabolome with significantly lower levels of the polyamines (putrescine and spermidine) in the treated parasites compared to increased levels in the untreated controls (Fig 4.5A, Table 4.5), as previously described (Fig. 1.7) [88, 90]. Moreover, downstream metabolites, including 5-methylthioinosine (Fig. 4.5A, Table 4.5), were also decreased in the treated parasites only, corroborating the complete metabolic halt of polyamine metabolism after the co-inhibition. The decreased NADH levels (Table 4.5) after PfAdoMetDC/ODC co-inhibtion could be the result of decreased energy metabolism as detected in the transcriptome (Table 3.3, section 3.3.6), but NADH was also decreased in the untreated controls (although ~5 times less in T_{t3}). UTP also showed a decrease at T_{t1}, but subsequently

increased. The levels of other energy intermediates (e.g. ATP and the glycolysis metabolites) were generally maintained or similarly affected in the untreated controls (Appendix E).

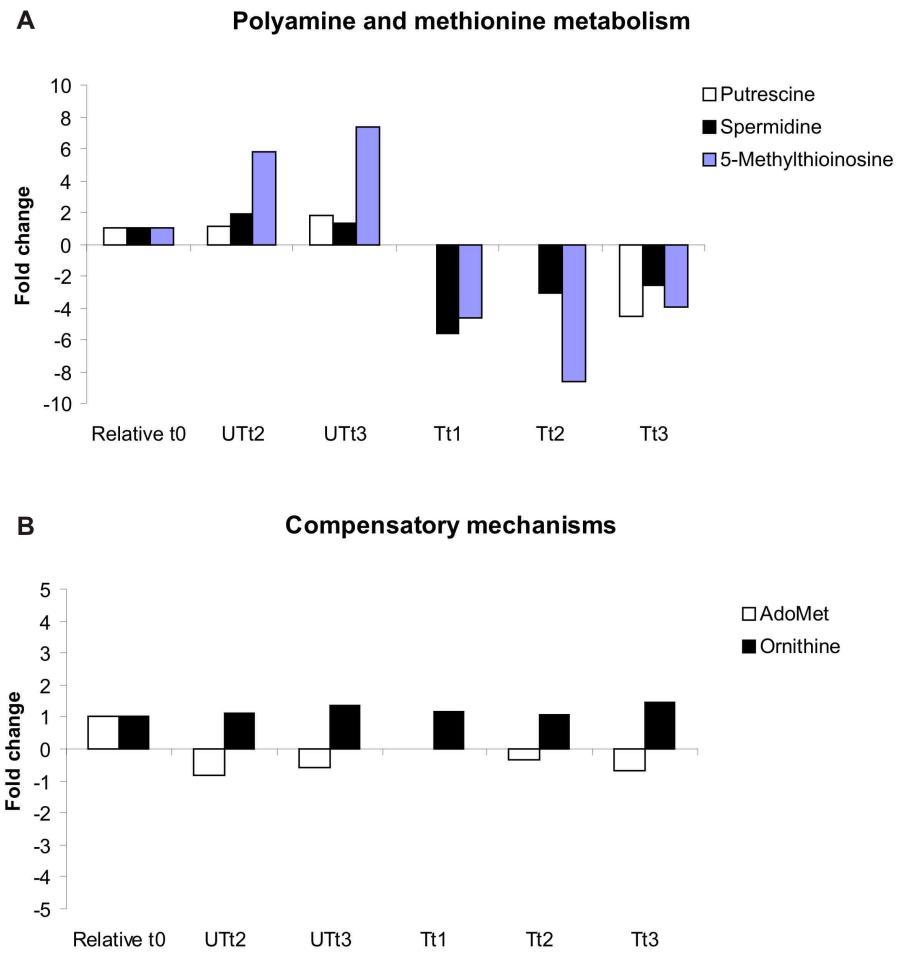


Fig. 4.5 Metabolite profiles compared to relative t_0 of **A)** putrescine, spermidine and 5-methylthioinosine downstream of PfAdoMetDC/ODC, which showed perturbation-specific decrease in the treated parasites, and **B)** ornihine and AdoMet levels directly upstream of PfAdoMetDC/ODC, which were maintained despite the complete co-inhibition. Cadaverine and spermine could not be detected. Fold changes are indicated compared to relative t_0. Thus, the fold change of relative t_0 = 1 (unchanged).

The proposed compensatory responses of LDC, OAT and AdoMet synthetase were also investigated in the metabolome. The increased abundance of LDC transcripts (Table 3.3) could indicate the functional production of cadaverine from lysine in *P. falciparum,* but cadaverine could not be detected. AdoMet and ornithine levels were altered less than 2-fold compared to the relative t_0 in all of the samples and were therefore not considered as differentially affected (therefore not present in Table 4.5). In other organisms, inhibition of AdoMetDC caused an increase of the substrate AdoMet [157]. However in *Plasmodium,* AdoMet homeostasis was apparently maintained (non-differential fold change ≈ -1, Fig. 4.5B, Appendix E). Moreover, ornithine levels also remained mostly unchanged (non-differential fold change ≈ 1, Fig. 4.5B, Appendix E) despite the complete inhibition of PfAdoMetDC/ODC, which indicates the efficiency of the compensatory AdoMet synthetase decrease and OAT increase as observed in the transcriptome (Table 3.3) and the proteome (section 4.3.1.2).

The general maintenance of metabolic homeostasis observed, caused reservations as to whether the relative t_0 strategy, as applied here and in other metabolomics investigations [267], was appropriate for differential metabolite abundance analysis, since the metabolome is more dynamic than the transcriptome and proteome. Moreover, if homeostasis exists, then an argument could be made for standard parallel time point comparison. Therefore, comparison of metabolite levels between treated and untreated parasites was in addition performed between parallel time points, which revealed 15 metabolites with differential abundance over the time course (Table 4.6).

Table 4.6 Metabolites with differential abundance (i.e. fold changes of more than 2 in either direction) after PfAdoMetDC/ODC co-inhibition (parallel time point comparison)

Metabolite	Fold change compared to parallel untreated control		
	T_{t1}/UT_{t1}	T_{t2}/UT_{t2}	T_{t3}/UT_{t3}
1,3-Diphosphateglycerate	-2.3	-1.3	3.9
γ-Aminobutyrate (GABA)	7.7	-2.8	-5.4
5-Methylthioinosine	-4.6	-50.2	-29.0
Adenosine	1.1	2.5	-1.4
Lysine	1.4	1.1	2.1
Methionine	-2.1	-1.6	-1.3
NADH	-2.0	-1.2	-4.8
Pipecolic acid	-1.2	-3.5	-3.4
Putrescine	ND	ND	-8.2
Pyruvate	2.1	-1.6	-2.2
AdoHcy	-3.8	4.7	-5.0
Spermidine	-5.6	-5.7	-3.5
S-ribosyl-L-homocysteine	1.9	-1.2	2.5
Succinate	1.9	-1.7	-2.2
UTP	-2.4	-1.3	1.1

Metabolites changed >2-fold in either direction are colour-indicated (red = more than 2-fold increased, green = more than 2-fold decreased). ND = not detected

When the results from the relative t_0 approach is compared to those of the parallel time point approach, 12/15 metabolites from Table 4.6 also appear in Table 4.5, but three metabolites uniquely appear only with parallel comparison (lysine, S-ribosyl-L-homocysteine and succinate, Table 4.6). Both analyses indicated that NADH was decreased ~5 times more after PfAdoMetDC/ODC co-inhibition in T_{t3} than in the controls (Table 4.5 and Table 4.6). However, compared to the relative t_0, most of the metabolites followed a distinct increasing or decreasing profile in the treated or untreated samples (Table 4.5), as opposed to a variable profile with the parallel strategy (Table 4.6). In addition, using the latter approach, only three metabolites (1,3-diphosphateglycerate, methionine and UTP) demonstrated a consistent increase over the time course (compared to eight with the relative t_0 approach), but levels above 2-fold were detected among a few variable profiles (e.g. adenosine, lysine, AdoHcy and S-ribosyl-L-homocysteine, Table 4.6). Moreover, using the relative t_0 approach, specific metabolites, e.g. GABA, pipecolic acid and pyruvate, were similarly increased in the treated and untreated samples, but they had a decreasing profile after PfAdoMetDC/ODC co-inhibition according to the parallel approach. Despite the observed homeostasis, this may again be due to arrest of the treated versus normal progression of the untreated parasites, since exactly the same extraction protocol were

followed for all the samples, which would result in a higher metabolite abundance of the untreated controls simply due to parasite maturity and larger size (Fig. 3.4).

Convincingly, the parallel time point comparison between treated and untreated samples also indicated the perturbation-specific effects of PfAdoMetDC/ODC co-inhibition downstream of the bifunctional complex with a maximum 8.2-fold decrease of putrescine, 5.7-fold decrease of spermidine and 50.2-fold decrease of 5-methylthioinosine (Fig. 4.6A, Table 4.6). As before, ornithine and AdoMet levels were not identified in the differential abundance set (fold change less than 2, not present in Table 4.6) and were therefore considered as unchanged, since the levels of these metabolites were maintained in both the untreated and treated parasites (Fig. 4.6B, Appendix E). This corroborates the proposed metabolic homeostasis upstream of PfAdoMetDC/ODC and the compensatory effects of OAT and AdoMet synthetase, as well as the perturbation-specific metabolic halt downstream of PfAdoMetDC/ODC.

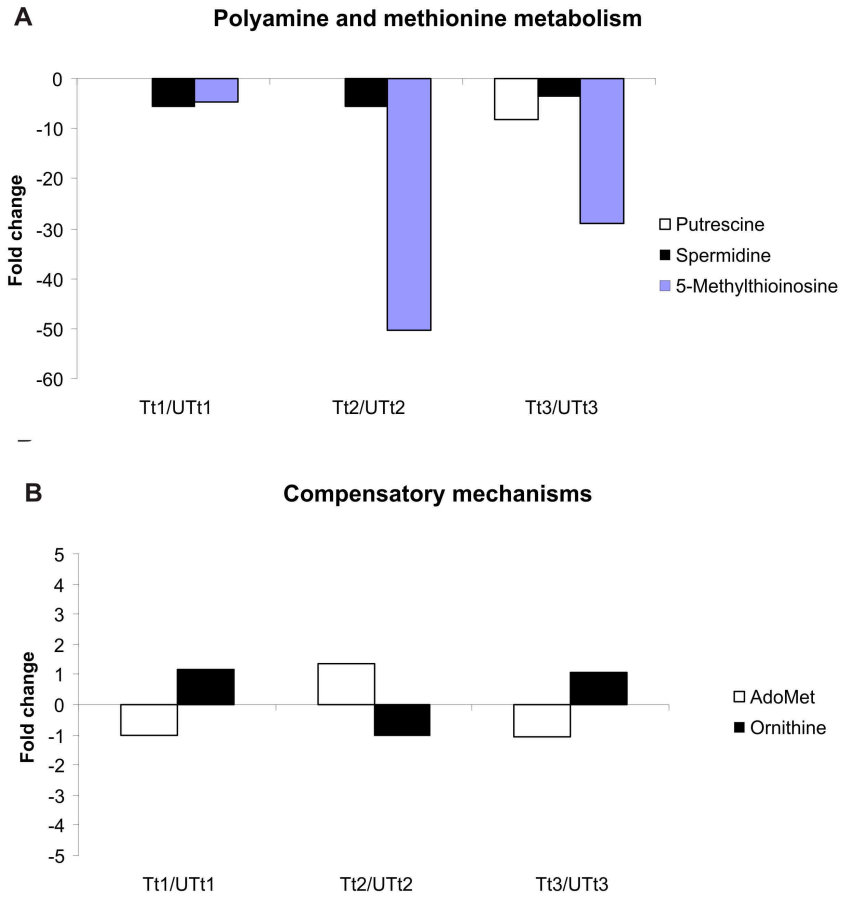


Fig. 4.6 Metabolite profiles of treated parasites directly compared to the parallel untreated controls for **A)** putrescine, spermidine and 5-methylthioinosine corroborating the perturbation-specific decrease after PfAdoMetDC/ODC co-inhibition, whereas **B)** ornihine and AdoMet levels were maintained.

4.3.3 Compensatory LDC induction during polyamine depletion investigated further

LDC activity and cadaverine accumulation reduces ethylene inhibition of arginine decarboxylase and AdoMetDC in pea seedlings [74] and cadaverine reverses DFMO-induced growth arrest of *P. falciparum* [282]. Compensatory induction of the LDC transcript was also detected after PfAdoMetDC/ODC co-inhibition of *P. falciparum* (section 3.4), but this response could not be confirmed on the protein level with 2D-GE nor could cadaverine be detected with LC-MS/MS. Therefore, to investigate the hypothesis of LDC induction as compensatory mechanism to alleviate polyamine depletion further, LDC activity of PfAdoMetDC/ODC co-inhibited parasites was biochemically determined via radio-labelled substrate (^{14}C-lysine) decarboxylase assays as described for ODC and AdoMetDC in section 3.3.1. However, no LDC activity above background could be detected in either treated or untreated *P. falciparum* cultures (Fig. 4.7). *E. coli* positive controls proved that the LDC activity assay was functional and uninfected erythrocytes were included as negative controls. Moreover, *E. coli* LDC activity was about 5-fold higher when assayed in sodium-acetate buffer (LDC buffer) [275] as opposed to the Tris-HCl (buffer A) [88], used for the AdoMetDC and ODC activity assays (section 3.3.1). This is probably due to inhibition of LDC by chloride ions as demonstrated for LDC from soybean [283]. However, no activity could be detected in the *P. falciparum* samples despite changing to the sodium-acetate buffer (without chloride ions), increasing the substrate concentration up to 100-fold (700 µM) that used for AdoMetDC and ODC activity assays or adding BSA to the reaction, as suggested for *E.. coli* [71] (results not shown).

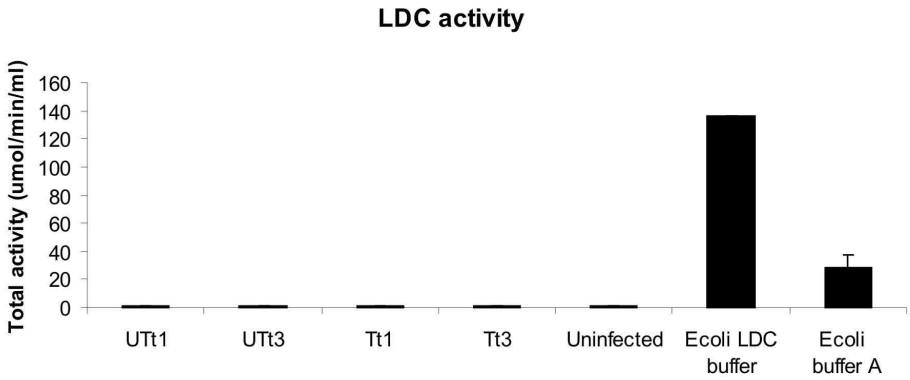


Fig. 4.7 Lack of measurable LDC activity of untreated and DFMO/MDL73811-treated parasite lysates (sampled at t_1 = 19 hpi and t_3 = 34 hpi) after incubation with L-[14C]-lysine (n=2). Uninfected erythrocytes were used as negative control and *E. coli* lysates as positive control. Note the reduced *E. coli* LDC activity due to inhibition by chloride ions [283] in buffer A compared to the recommended LDC buffer.

In a final attempt to detect plasmodial LDC activity, a mixture of co-factors at physiological concentrations (~80 µM magnesium, manganese, thiamine, ATP and NAD respectively, with 40 µM substrate), as required by other decarboxylases [284-286], was added to the reaction in addition to PLP, but enzyme activity could not be detected in *P. falciparum*.

4.3.5 gDNA Methylation status investigation

DFMO/MDL73811-treatment of trypanosomes caused a 40-fold accumulation of AdoMet [185], which resulted in speculation of hypermethylation of nucleic acids and/or proteins as the main antitrypanosomal mechanism of MDL73811 [157]. Methylation of histones [125] or 2-deoxycytosine bases within gDNA [287] can result in transcriptional repression and might explain the decreased abundance of the majority of transcripts (70%) after PfAdoMetDC/ODC co-inhibition (section 3.3.5.2). Yet, with DFMO/MDL73811-treatment of *P. falciparum*, AdoMet levels were maintained (section 4.3.2.1) by a decrease in its synthesis among others (section 4.3.1.3). However, the conversion of AdoMet to AdoHcy (resulting in methylation) could in addition have controlled AdoMet levels [241]. To investigate the potential role of DNA methylation in the mechanism of MDL73811 in *P. falciparum*, the presence of CpG islands in the genes encoding the transcripts that were differentially affected after PfAdoMetDC/ODC co-inhibition was investigated.

4.3.5.1 CpG island analysis of the differential transcript abundance list

CpG island analysis (CpGplot, CpGreport) of the genes encoding the 538 transcripts from the LIMMA dataset (377 decreased, 171 increased), including their 1000 bp upstream/downstream regions, did not detect any CpG islands. In fact, geecee-count analysis calculated the average GC content of these genes and their surrounding regions, which was found to closely represent the average composition of the *P. falciparum* genome i.e. 18.8 – 18.4% GC versus the reported 19.4% [96] (Table 4.7).

Table 4.7 Geecee-count analysis of the genes encoding the 538 differentially affected transcripts

Gene list	%GC content of the 538 genes			
	CDS	1000 bp upstream	1000 bp downstream	Total GC
Genes of increased abundance transcripts	15.67	3.13	3.67	18.8
Genes of decreased abundance transcripts	14.66	3.73	4.25	18.39

Therefore, as opposed to a gene-specific methylation approach such as bisulfite sequencing [288], a genome-wide analysis strategy was applied to determine the role of DNA methylation in the observed transcriptional suppression.

4.3.4.2 Global methylation assays

The classical method of global DNA methylation analysis using methylation-sensitive restriction endonucleases (MSRE) is based on the properties of specific endonucleases to be sensitive to the methylation status of their recognition sequence. MSRE demonstrated the high abundance of 5mC in murine DNA since different digestion profiles were obtained when using enzymes with different methylation sensitivities [289], but failed to demonstrate the genetic methylation pattern in *D. melanogaster* [290]. Instead a powerful immunological approach revealed the presence of 5mC in *D. melanogaster* DNA [290]. Both these strategies

were attempted here to determine the methylation status of *P. falciparum* gDNA isolated from samples of two time points (t_1 = 19 hpi and t_3 = 34 hpi) after PfAdoMetDC/ODC co-inhibition.

4.3.4.2.1 Methylation-sensitive restriction endonucleases

Due to the contradictory evidence of cytosine methylation [124, 126, 127] in *P. falciparum* gDNA (section 1.10.5) and the known role of adenine methylation in other low eukaryotes [291], both methylation types were investigated by including two enzymes with different methylation sensitivities: *Hpa*II, which recognises C↓CGG but does not cleave when the cytosines are methylated (5mC), and *Dpn*I, which recognises the sequence GA↓TC but only cleaves when the adenine is methylated (N^6-methyladenine). No difference in the digestion profiles of gDNA from PfAdoMetDC/ODC co-inhibited or untreated parasites could be detected after either *Hpa*II or *Dpn*I overnight digestion (Fig. 4.8).

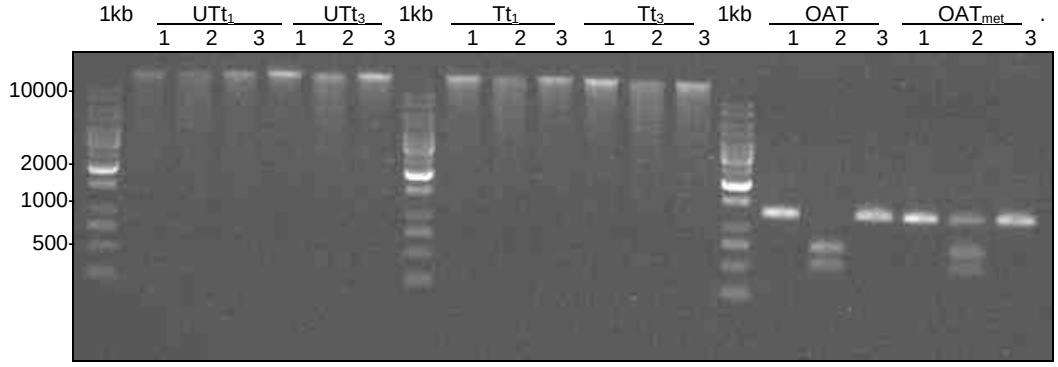


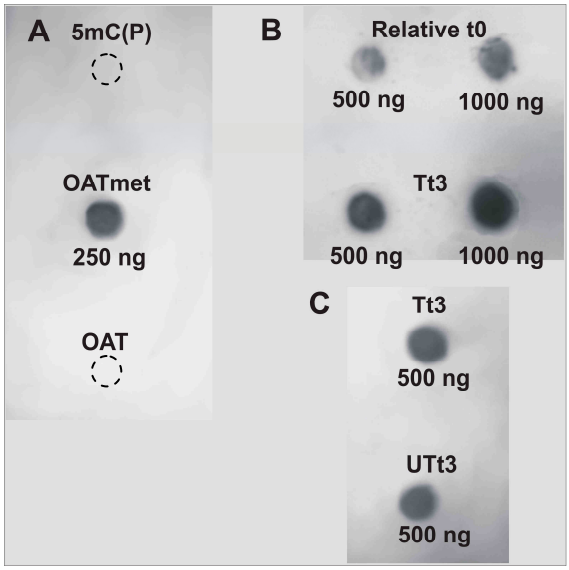
Fig. 4.8 Gel electrophoresis of 250 ng digested and undigested gDNA on 0.8% agarose, to assess methylation after PfAdoMetDC/ODC co-inhibition. Lane 1: undigested gDNA, lane 2: HpaII digested and lane 3: DpnI digested gDNA of treated (T) and untreated (UT) parasites at t_1 = 19 hpi and t_3 = 34 hpi, respectively. A synthetic unmethylated (OAT = 250 ng) and methylated control (OAT_{met} = 250 ng) DNA are included. Molecular size is indicated with a 1 kb ladder.

However, the partial protection of cytosine-methylation against *Hpa*II digestion was visible in the M.SssI (CpG methyltransferase)-treated synthetic DNA control (Fig. 4.8 OAT_{met} lane 2). Three detectable bands including the original 1245 bp OAT amplicon, as well as the 725 bp and 520 bp fragments (flanking the CCGG restriction site), are visible, compared to the unmethylated OAT *Hpa*II digest where the amplification product was completely digested with only two bands visible (Fig. 4.8 OAT lane 2). Although two *Dpn*I sites occur in OAT DNA, the amplicon was not cleaved, indicating that N^6-methyladenine was not present. These controls confirmed that MSRE can theoretically distinguish different 5mC methylation patterns, but the methodology was not sensitive enough to reveal specific methylation profiles in *P. falciparum* and gDNA methylation differences as a result of PfAdoMetDC/ODC co-inhibition could not be demonstrated.

4.3.4.2.2 South-Western immunoblotting

The low sensitivity of MSRE analysis demanded a more sensitive approach, such as immunoblotting, to establish the methylation status after PfAdoMetDC/ODC co-inhibition and to clarify the paradox concerning

5mC within the *P. falciparum* genome [124, 126, 127]. By including both 5-methylcytidine [5mC(P)] and synthetically prepared (and thus unmethylated) OAT as negative controls and OAT_{met} as positive control, the specificity of the antibodies for 5mC within gDNA was determined (Fig. 4.9A). Methylated RNA species can cause false positive results [277] and therefore contaminating RNA was removed during DNA isolation with RNase A treatment. The initial experiment (Fig. 4.9B) indicated a higher abundance of 5mC in T_{t3} compared to the relative t_0 as determined by densitometry. Moreover, the 5mC signal versus the gDNA mass (i.e. 5mC quantity) was approximately linear (Fig. 4.9 Table). However, since the putative DNA(cytosine-5)-methyltransferase (MAL7P1.151) transcript peaks between 14 and 23 hpi according to the 3D7 IDC transcriptome [91] (i.e. around t_1) and the transcript was not affected by PfAdoMetDC/ODC co-inhibition (Chapter 3), it was decided to determine the gDNA methylation of UT_{t3} (Fig. 4.9C) as well. Interestingly, the m5C content and thus gDNA cytosine-methylation of T_{t3} and UT_{t3} (based on spot density) was the same. It therefore appears as if the increased methylation detected was time-dependent (t_3) and not the result of PfAdoMetDC/ODC co-inhibition.



Experiment	Sample	gDNA mass (ng)	Total spot density[a]
B	Relative t_0	500	14068
		1000	27855
	T_{t3}	500	31255
		1000	50194
C	T_{t3}	500	31832
	UT_{t3}	500	32193

a. Total spot density = area (mm^2) x density (CNT/mm^2)

Fig. 4.9 South-Western blot of 5mC in *P. falciparum* gDNA. **A)** 1 nmole of 5-methylcytidine [5mC(P)], 250 ng artificially methylated OAT (OAT_{met}), 250 ng unmethylated OAT; **B)** 500 ng and 1000 ng relative t_0(UT_{t1}), 500 ng and 1000 ng T_{t3}; **C)** 500 ng T_{t3} and 500 ng UT_{t3}. The total spot densities of the two separate experiments (**B** and **C)** are tabled.

4.4 DISCUSSION

In this chapter, conclusions drawn from the transcriptomics investigation of PfAdoMetDC/ODC co-inhibited *P. falciparum,* resulting in polyamine depletion, were validated in the proteome and metabolome. Moreover, biochemical assays were performed to investigate specific hypotheses such as the induction of LDC as compensatory response to polyamine depletion and the role of DNA hypermethylation in the mechanism of MDL73811. As with the transcriptomics investigation, the effects of cytostasis could be detected in the proteome with the number of detectable spots decreasing in the treated samples over the time course (483 to 409). However, correlation calculations indicated that the effects of cytostasis were subtle in proteome (T_{t3}

versus UT_{t1}: R = 0.88) and even more so in the metabolome (T_{t3} versus UT_{t1}: r = 0.96) compared to the transcriptome (T_{t3} versus UT_{t1}: r = 0.57). Yet, differential abundance analysis of the treated time points compared to UT_{t1} (regarded as a relative t_0 and reference point for quantitative analysis through the whole functional genomics investigation) detected 53/500 protein spots and 24/92 metabolites compared to the 538/5332 transcripts with differential abundance as a result of the perturbation.

Technical limitations of some of the techniques applied were evident, such as the poor proteome coverage of 2D-GE [260]. Despite the excellent resolution, relatively easy, inexpensive performance and theoretical capability of detecting more than 5000 protein spots [292], only about 500 trophozoite stage proteins (including protein variants and contaminating host proteins) i.e. ~10% of the *P. falciparum* proteome, were accessible. In comparison, a recently "improved' 2D-GE method for *P. falciparum* detected only 239 spots [293] and ~300 spots were obtained from *T. cruzi* epimastigotes [294]. Yet, MudPIT confidently identified 1036 trophozoite stage proteins covering 42% of the predicted *P. falciparum* proteome [112]. The poor proteome coverage is explained among others by protein solubility constraints (e.g. membrane proteins) and exclusion due to extreme pI and/or protein size [255]. The latter caused two important polyamine metabolism proteins, PfAdoMetDC/ODC (330 kDa) and LDC (280 kDa), not to be assessable with 2D-GE since they were too large for gel penetration. In addition, the limited identification of low molecular weight and low abundance proteins resulted in the successful identification of only ~30% of the already restricted dataset. In future studies, the 2D-GE proteome coverage may be increased by fractionation [251] or by using larger gels (24 cm) with better resolution, but techniques such as MudPIT, ICAT or iTRAQ [295] should definitely be considered to improve the identification of low abundance proteins and the coverage of the *Plasmodium* proteins interrogated.

The time course approach has proven indispensable, particularly in the 2D-GE investigation. The visible increase of OAT and decrease of AdoMet synthetase, located adjacently on the 2D-image, confirmed that these changes were indeed in response to the perturbation and not due to technical artifacts. Furthermore, transcriptional responses that were delayed on the protein level (LDH) due to the translational gap [246], could have been regarded as contradictory to the transcriptomics results if only one time point had been assessed.

Pearson correlation calculations of the metabolomics data revealed the general maintenance of metabolic homeostasis and many of the metabolites that showed a 2-fold change compared to the relative t_0 in the treated parasites were similarly affected in the untreated controls. Therefore, differential metabolite abundance was also determined compared to the parallel untreated controls (standard approach). Both strategies indicated perturbation-specific effects downstream to PfAdoMetDC/ODC with the decrease of putrescine, spermidine and 5-methylthioinosine in the treated parasites. Otherwise, homeostasis was generally maintained and ornithine and AdoMet levels were not significantly altered despite the complete inhibition of PfAdoMetDC/ODC. However, the results obtained with the parallel time point comparative approach were

generally more variable over the time course and in some cases metabolite levels could have appeared to decrease after treatment simply because lesser amounts were extracted due to the smaller size of the arrested parasites compared to the more mature, untreated controls. For this reason the relative t_0 strategy is also regarded as the better approach to follow for analysis of metabolomics data during cytostasis, but if the metabolite samples could be quantitatively loaded onto the LC-MS/MS instrument, this problem could be resolved to some extent. Standardisation of the metabolite extraction efficiency/yield against an internal biomarker (internal standard) would be ideal. Another limitation of the metabolomics methodology applied here is that the sample replication encompassed only two biological replicates and no technical replicates. The variation detected with the parallel time point approach could thus have included technical variation, although the variation between biological replicates should exceed technical variation and the latter should be equal regardless of the analysis approach applied. The limited replication furthermore constrained statistical analysis, since most statistical methodologies (e.g. Student's t-test) require at least three replicate values and the statistical significance of the obtained results could therefore not be calculated. Reports on metabolomics investigations of *Plasmodium* are currently extremely scarce [116] and the data analysis performed here was limited to differential abundance analysis and Pearson correlation calculations. In other organisms intricate computational approaches and data visualisation methods (e.g. PCA, hierarchical clustering and PLS), similar to microarray data analysis [254, 296], are now being performed. In the investigation presented here the overall aim of the proteomics and metabolomics analyses was to confirm certain hypotheses resulting from the transcriptomics investigation, which was achieved with the fundamental analysis described, but a more in-depth analysis approach could certainly reveal metabolic effects of the perturbation not determined here.

Two of the three perturbation-specific compensatory transcriptional responses, namely the increased abundance of OAT, as well as the decreased abundance of AdoMet synthetase, showed coordinated responses in the proteome and metabolome. An integrated view of the results regarding polyamine and methionine metabolism from the complete functional genomics investigation is presented in Fig. 4.10. The increase in the transcript and protein levels of OAT and the maintenance of ornithine concentrations in the metabolome provides evidence for the compensatory effects of OAT during PfAdoMetDC/ODC co-inhibition and confirms its role in the regulation of ornithine in the parasite. The significant increase of the protein pdx1 could be an additional compensatory attempt to activate PfAdoMetDC/ODC (and potentially LDC) since both ODC and LDC are PLP-dependent decarboxylases [279].

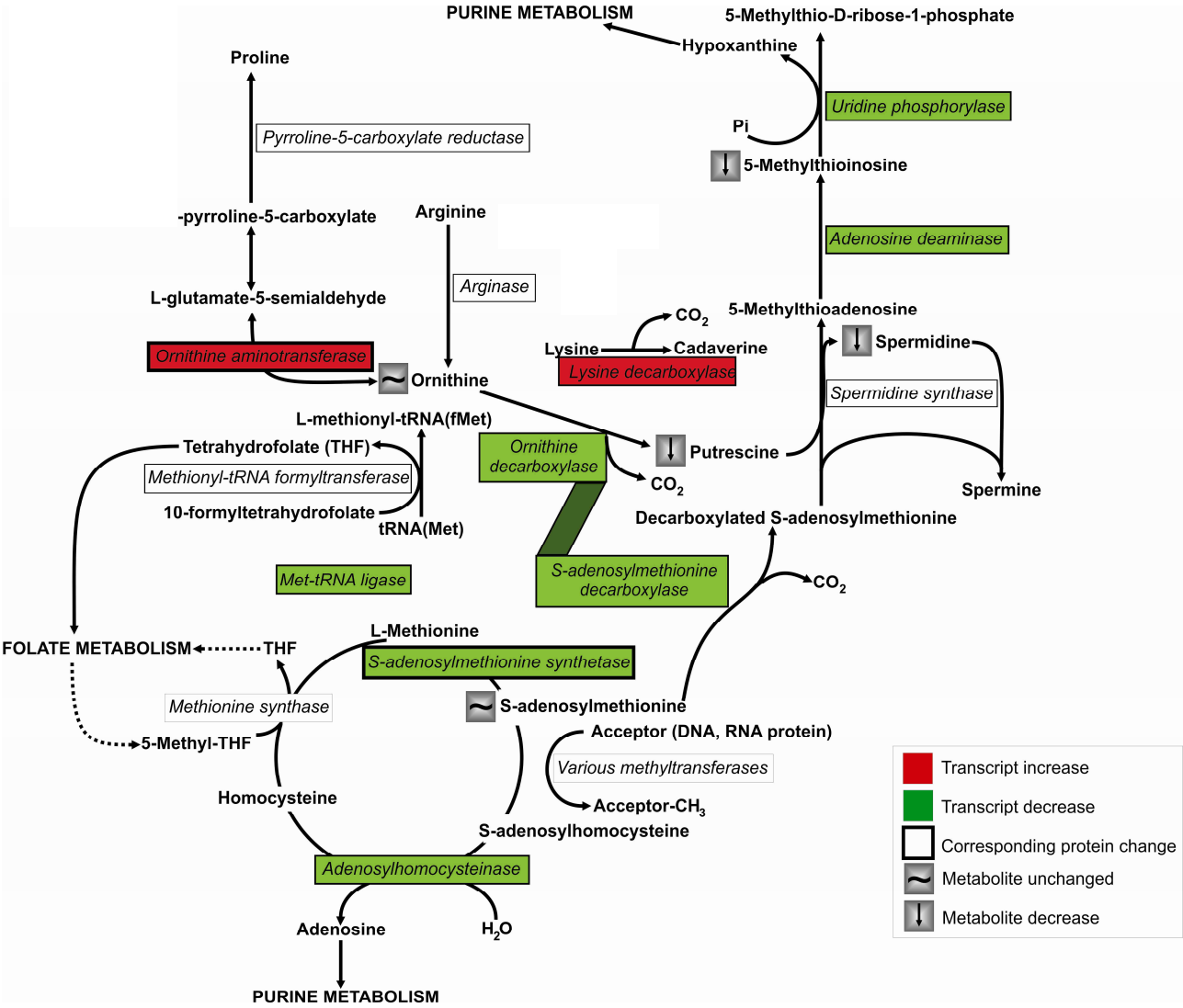


Fig. 4.10 Polyamine and methionine metabolism (adapted from MPMP at http://sites.huji.ac.il/malarial/). The minimal synthesis of spermine in *Plasmodium* is currently believed to be catalysed by spermidine synthase, as indicated [85]. Enzymes of which the transcript abundance was significantly increased are indicated in red and those significantly decreased are indicated in green, whereas proteins with confirmed corresponding abundance are framed with a thick border. Metabolites that were decreased at least 2-fold are indicated.

DFMO/MDL73811-treatment of trypanosomes caused a 40-fold accumulation of AdoMet [185], resulting in speculation on hypermethylation of nucleic acids and/or proteins being the main antitrypanosomal mechanism of MDL73811 [157]. However, in the present study of DFMO/MDL73811-treated *P. falciparum*, AdoMet levels were maintained. Several mechanisms could maintain metabolic homeostasis during a perturbation such as regulation of enzyme activity or protein production. This study revealed a decrease in both the transcript and protein of AdoMet synthetase that may act as a compensatory strategy to maintain AdoMet levels. The exact mechanism behind this regulation needs to be elucidated, but in MDL73811-treated mammalian cells the AdoMet concentration was effectively regulated through substrate feedback inhibition of AdoMet synthetase activity [157, 240]. The Plasmodial enzyme activity was recently reported not to be allosterically regulated by AdoMet [79], but based on the data presented here, AdoMet could regulate the transcript and protein levels of AdoMet synthetase. In contrast, the trypanosomal enzyme is poorly regulated, resulting in the substantial accumulation of AdoMet after AdoMetDC inhibition [157]. Moreover, the proposed difference in regulation of

AdoMet synthetase between Trypanosoma and Plasmodia, resulting in hypermethylation with polyamine depletion in the first case and possibly only polyamine depletion in the second, may be the reason for the success of MDL73811 in *T. brucei rhodesiense*-infected mice [184], and failure of MDL73811 in *P. berghei*-infected mice [153].

Hypermethylation as partial mechanism of MDL73811 in *Plasmodium* was further investigated since the methylation of histones [125] or 2-deoxycytosine bases within gDNA [287] can result in transcriptional repression, which could explain the decreased abundance of the majority of transcripts (70%) after PfAdoMetDC/ODC co-inhibition (section 3.3.5.2). Although AdoMet levels were maintained, the conversion of AdoMet to AdoHcy (resulting in methylation) could also have controlled the AdoMet concentration [241], in addition to decreased AdoMet synthesis. AdoHcy is the major by-product of AdoMet-dependent trans-methylation and a competitive inhibitor of trans-methylation reactions [239, 241]. A decrease in the AdoMet/AdoHcy ratio would result in reduced methylation [241]. AdoHcy increased from -3.8 to unchanged in T_{t1} to T_{t3} and from -8.4 to 4.6 in UT_{t1} to UT_{t3} in the data compared to the relative t_0 (Table 4.5, Appendix E). This could indicate the increased conversion of AdoMet to AdoHcy through methyltransferase or the decreased hydrolysis of AdoHcy via adenosylhomocysteinase (the transcript decreased 2.6-fold, Table 3.3). There has been contradictory evidence on cytosine-methylation as an epigenetic mechanism in *Plasmodium* [124, 126, 127], but DNA(cytosine-5)-methyltransferase homologues have been predicted in three of the human malaria parasites [*P. falciparum* (MAL7P1.151), *P. knowlesi* (PKH_021170) and *P. vivax* (Pv081670)] according to PlasmoDB 5.4. Therefore, the DNA methylation potential of the genes encoding the 538 differentially affected transcripts was determined by CpG-island bioinformatics analysis and the extent of gDNA methylation of DFMO/MDL73811-treated and untreated parasites was determined with MSRE and 5mC-antibodies. Bioinformatics analyses could not predict any CpG-islands in the differentially affected transcript data and MSRE had insufficient sensitivity to indicate differences in the digestion profiles of treated and untreated parasite gDNA. However, in contrast to other reports [124, 126, 127], the highly sensitive South-Western immunoblotting detected the presence of 5mC in *P. falciparum* gDNA of treated and untreated parasites, despite the low GC content of the genome (19.4% GC-rich) [96]. 5mC-Methylation apparently increases during development (tested from 19 hpi to 34 hpi) and was not influenced by the perturbation, i.e. the gDNA of treated and untreated parasites at t_3 was similarly methylated. This may be explained by the expression of the putative DNA(cytosine-5)-methyltransferase (MAL7P1.151) early in the trophozoite stage (transcript peaks between 14 and 23 hpi according to the 3D7 IDC transcriptome) i.e. at around t_1. The proposed stage-dependent gDNA methylation (as opposed to perturbation-dependent methylation) was further substantiated by the fact that AdoMet levels in both treated and untreated parasites were maintained during the time course. From the results obtained, it was concluded that MDL73811 does not result in hypermethylation of *P. falciparum* gDNA.

The compensatory induction of LDC to alleviate polyamine depletion remains to be confirmed in *Plasmodium,* since the protein could not be detected by 2D-GE nor cadaverine with LC-MS/MS. LDC activity assays with ^{14}C-labelled lysine were subsequently conducted to elucidate this hypothesis, but plasmodial LDC activity could not be demonstrated, whereas convincing activity of the *E. coli* positive control confirmed the integrity of the assay. It may be that LDC is not a constitutive enzyme in *P. falciparum* and that it is only induced upon polyamine depletion, which would explain the lack of activity in the untreated controls. This is supported by the fact that putrescine and spermidine were shown to inhibit LDC in a regulatory manner in *E. coli* [297], which may occur at the transcript level, resulting in increased LDC transcription only upon polyamine depletion. Experts in the field of malaria polyamine research from the Bernard Nocht Institute for Tropical Medicine, Hamburg, Germany [79], could also not demonstrate any LDC activity in *P. falciparum* cultures (I Muller, personal communication), but the recombinant enzyme was successfully expressed and lysine decarboxylation could be demonstrated [79]. However, if the plasmodial LDC is inhibited by the other polyamines, activity should have been detected in the treated/polyamine depleted parasites due to the induction of LDC protein expression. However, DFMO-inhibition of LDC activity has been described in *Selenomonas ruminantum* [73] and cannot be excluded in the case of the plasmodial LDC. Another possibility is that the plasmodial LDC is not stable under the conditions used for storage, sample preparation or assay or that it requires a co-factor which was not provided. Western blot analysis would indisputably confirm an increase in the LDC protein, but successful protein expression of this 280 kDa protein will be required to enable antibody generation for this strategy.

Similar to other studies of *Plasmodium* [246, 252], changes in transcript and protein levels as a result of PfAdoMetDC/ODC co-inhibition were generally correlated, e.g. AdoMet synthetase, OAT and LDH, though the latter was delayed. In some cases small incremental changes on the transcriptional level (e.g. falcipain-2 and pdx1) resulted in significant changes on the protein level, which could be attributed to post-transcriptional regulation of these particular enzymes. Alternatively, it may indeed be that small amplitude transcriptional responses can result in momentous alterations of the corresponding protein levels. However, transcripts and protein levels are not necessarily correlated [298], nor are all transcripts and all proteins necessarily regulated in the same way. For some genes the dominant regulation may occur at transcriptional level, whereas for others it may occur at the post-transcriptional and/or translational level. The challenge remains to be able to decipher the biological significance of the information obtained [136] via data integration from different functional genomics investigations [254]. In this investigation perturbation-specific compensatory mechanisms were detected in the transcriptome and confirmed in the proteome and the metabolome, corroborating the biological significance of OAT and AdoMet synthetase upon PfAdoMetDC/ODC co-inhibtion.