CHAPTER 4

Transcriptional responses of *P. falciparum* to inhibition of AdoMetDC with MDL73811

4.1 Introduction

Microarray technologies consist of thousands of transcripts or even the whole genome on a single chip or array and enable expression profiling of differentially expressed transcripts that may be induced by a certain perturbation. This global overview of the response of organisms to any perturbation makes transcriptomic investigations by microarrays extremely promising technologies to deduce the mode-of-action of drugs (Brazas & Hancock, 2005). The ultimate aim of functional genomics is to increase the number of validated drug targets (Chanda & Caldwell, 2003). An advantage associated with transcriptomic investigations are the large number of data points, which is more than that of gel-based proteomics. Therefore, transcriptomics offer a larger data set and in combination with proteomics can provide a global picture of both the transcriptome and proteome.

4.1.1 Transcriptomic perturbation studies in other organisms

The wealth of information gained with the use of microarrays has been demonstrated in several research fields. Transcriptional responses have been determined for the effect of rapamycin on the immune response of human cell lines (Grolleau *et al.*, 2002), as well as the response of *Staphylococcus aureus* to glycopeptides (Scherl *et al.*, 2006). Similarly, the mode-of-action has also been determined for anti-fungal agents against *Saccharomyces cerevisiae* (Agarwal *et al.*, 2003), and various anti-microbial peptides against *S. aureus* (Pietiainen *et al.*, 2009). The information gained from these transcriptomic studies can then be used in the elucidation and design of new anti-microbial peptides and compounds that will not infer resistance to infection (Pietiainen *et al.*, 2009).

Tuberculosis research has employed microarrays under a variety of conditions to determine the transcriptional response of *Mycobacterium tuberculosis* to various drugs. Transcriptional investigations using Affymetrix array microarrays confirmed the mode-of-action of the tuberculosis drug isoniazid and ethionamide against *M. tuberculosis* (Wilson *et al.*, 1999, Fu, 2006). Signature profiles of the gene expression of isoniazid, thiolactomycin, and triclosan treated *M. tuberculosis*, elucidated that the expression of 21 transcripts were able to distinguish between the mode-of-action of these 3 drugs (Betts *et al.*, 2003). To aid in the identification of gene expression signature profiles of tuberculosis drugs a large scale microarray study was performed on *M. tuberculosis* to

determine the mode-of-action of tuberculosis drugs. A total of 430 microarrays were done for 75 known and unknown tuberculosis drugs to validate their mode-of-action and to provide fingerprint profiles of the transcriptomic response of *M. tuberculosis* to these drugs (Boshoff *et al.*, 2004). Ongoing studies employed the use of Affymetrix oligonucleotide GeneChips to determine the transcriptional response of highly resistant *M. tuberculosis* strains against isoniazid in order to determine the mode-of-resistance (Fu & Shinnick, 2007).

4.1.2 Microarray platforms

Various microarray platforms exist all with their own advantages and disadvantages. NimbleGen has several multiplex arrays which include a 4-plex format (4×72K) with 72,000 probes per array and a 12-plex format (12×135K). Ultra-high density NimbleGen arrays can contain between 385 000 and 2.1 million probes, resulting in the presence of multiple, unique probes. Eppendorf provides pathway-focused DualChip® microarrays that contains two identical microarrays, printed side-by-side, and makes use of Xmer probe technology that contains long sense DNA (200-400 nucleotides) to provide maximum signal and minimal background. Affymetrix has quartz Gene Chips that provide whole genome coverage for humans and contains 28,869 genes each represented by approximately 26 probes spread across the full length of the gene, ultimately resulting in a total of 764 885 distinct probes. These Gene Chips are produced by *in situ* manufacturing of short (25-mer) oligo's on glass by photolithography (Kreil *et al.*, 2005). Agilent provides custom printing of 60-mer oligo's on a slide in a base-by-base manner by a combination of inkjet technology and phosphoramidite chemistry (Wolber *et al.*, 2006). The principle of phosphoramidite chemistry relies on the reactive sites of the nucleotides that are blocked with chemical groups that can then be selectively removed with the progression of synthesis. This process therefore allows the addition of one base at a time in a controlled manner. Overall, this process allows more spots to be printed on an array due to the precision of inkjet printing, as well as better shaped spots (Wolber et al., 2006). Agilent slides were compared to custom microarrays using cDNA long length probes (800-2000 bp) (Hockley *et al.*, 2009). Although great overlap of genes were detected for both arrays, more differentially expressed genes were found using the Agilent slides (Hockley *et al.*, 2009). Overall, Agilent provides greater specificity and sensitivity due to the Agilent 60-mer design, and also provides the opportunity to investigate various organisms like *Plasmodium* with the custom array design. Comparisons between platforms are constantly being investigated together with validation of these results by quantitative real-time polymerase chain reactions (qRT PCR) (Hester *et al.*, 2009, Baumbusch *et al.*, 2008, Arikawa *et al.*, 2008, Wang *et al.*, 2006).

4.1.3 Experimental design and normalisation methods

Microarrays provide a technology for comparing the expression profiles of genes across the entire genome. The arrow annotation (Figure 4.1) used, was proposed by Kerr and Churchill in 2001(Kerr & Churchill, 2001). Each arrow is representative of a single microarray. The point of the array is indicative of the sample labelled with Cy3 (green channel), while the base of the arrow is representative of the sample labelled with Cy5 (red channel). The symbols are representative of the samples used for analysis (Kerr & Churchill, 2001).Various designs are possible for two colour arrays (Figure 4.1). The direct design is used to make direct comparisons between two samples (Figure 4.1 a). Dye swaps are often used in direct designs to compensate for possible dye bias that may exist. A variety of reference designs does exist (Figure 4.1 d, e, h) that include a single reference or a combined reference design.

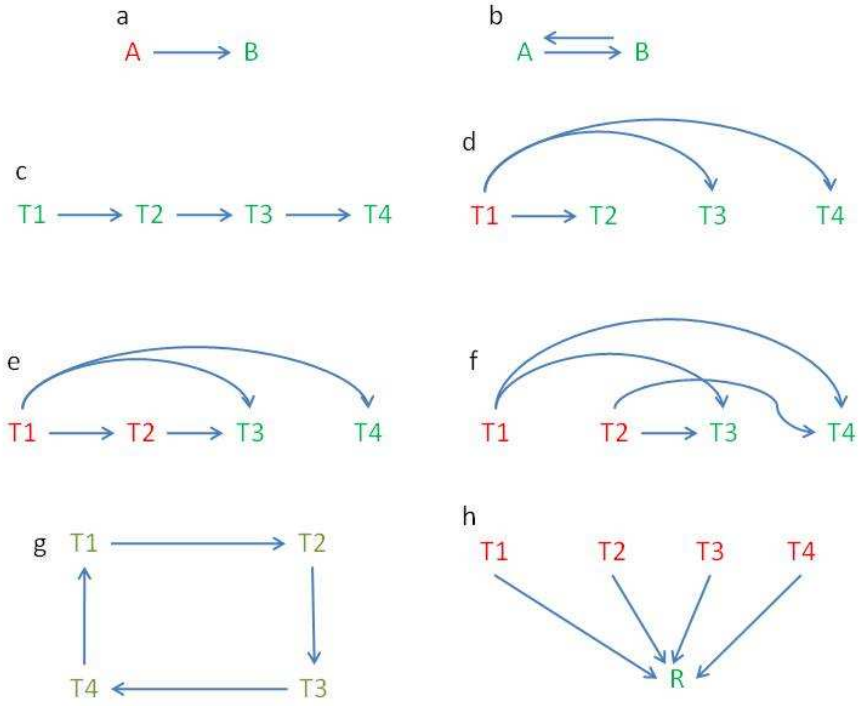

Figure 4.1: Microarray designs for time course experiments (Yang & Speed, 2002).
(a)Common design to use only one microarray, (b) dye swap design, (c) direct sequential design, (d) T1 as common reference, (e) T1 as common reference, (f) direct mixed design, (g) direct loop design, (h) common reference

The reference design is the most commonly used design (Dombkowski *et al.*, 2004, Churchill, 2002), which has the advantage of simplicity and the ease to add arrays, although to its disadvantage is limited experimental design (Kerr, 2003). The use of a reference design allows easy expansion of an experiment as long as each additional sample was included in the reference sample. Reference designs have higher variability than direct designs, but have the added advantage that all the comparisons that are made in reference designs are made with equal efficiency since the

samples are always compared to the same reference sample (Churchill, 2002). Therefore, in a reference design the differential abundance of any sample can be determined in relation to the other samples as long as all the samples that are compared were included within the reference sample.

Normalisation of microarray data is used to remove systematic bias and variation that is introduced into the sample by technical artefacts, though it is important to still maintain the important biological variations, therefore creating unbiased microarray variation between the samples that are to be analysed (Quackenbush, 2002, Oshlack *et al.*, 2007). Normalisation of microarray slides may also be used to compensate and correct for differences that exist in the microarray experimentation rather than the biological data (Smyth & Speed, 2003). Various normalisation methods exist for correcting microarrays and can be applied in two classes which include within-array normalisation which is normalisation of the M-values (log transformation of Cy3/Cy5) and between-array normalisation which is normalisation of the intensities ($\log_2$-ratios) to be comparable between all the arrays within the dataset (Smyth *et al.*, 2008). The M-value is defined as the log transformation of Cy3/Cy5 while the A-value is defined as the log transformation of the squared root of Cy3×Cy5. Linear model for microarray data (LIMMA) is a package for the analysis of microarray data and are used to obtain differentially regulated transcripts from the microarray data. Print-tip Loess is the default normalisation method used by LIMMA, but is generally unreliable when less than 150 spots per print-tip are used. Global Loess normalisation assumes that the majority of the oligo's is not differentially expressed, but does not assume that the number of up- and down-regulated genes is equal (Smyth *et al.*, 2008). Another option is Robust Spline normalisation which is an empirical Bayes compromise between print-tip and global Loess normalisation. It makes use of a 5-parameter regression spline that is used in place of the Loess curves (Smyth *et al.*, 2008). Within-array normalisation only affects the M-values and not the A-values. Normalisation of the A-values, which result in similar distribution across all arrays, makes use of quantiles that can correct for the individual red and green channels (Smyth *et al.*, 2008). R-quantile and G-quantile normalisation is useful for reference designs since R-quantile will normalise the samples labelled with red (Cy5), while G-quantile will normalise the samples in green (Cy3). Therefore, if the reference sample was labelled with Cy5 throughout in a reference design, it will be useful to use R-quantile for normalisation in order to correct the reference sample similarly across all arrays.

4.1.4 Minimum information about a microarray experiment (MIAME)

Microarray technology provides a large scale high-throughput method to investigate the transcriptional response of any organism to any stimuli. One of the drawbacks of microarray

analysis is the generation of large volumes of data and to maintain high standards of microarray analysis and to enable the comparison of microarray data between laboratories, it is necessary to set a standard for microarray data. The minimum information about a microarray experiment (MIAME) was established to ensure equivalent data to be reported for various microarrays (Brazma *et al.*, 2001). Three main levels of data reporting are important for microarray experiments and include the scan images which is the raw data, the quantitative output of the data, and the derived measurements (Brazma *et al.,* 2001). Detailed information regarding experimental design, sample preparation, hybridisation conditions, array type, array manufacturer, the number and size of the spots printed on the array as well as all the information regarding the generation of data should be given (Brazma *et al.,* 2001). For the reporting of such information various databases have been created to deposit microarray data. One such data repository is the National Centre for Biotechnology Information Gene Expression Omnibus (NCBI GEO, www.ncbi.nlm.nih.gov/geo) which provides an accession number that can be searched for easy access to all the relevant data.

4.1.5 Transcriptomic perturbation studies in Plasmodial parasites

The transcriptional response of Plasmodial parasites under CQ pressure analysed by serial analysis of gene expression (SAGE) revealed 100 regulated transcripts that included groups of transcripts involved in oxidative stress, hemoglobin digestion and proteins synthesis (Gunasekera *et al.*, 2003). Further, microarray investigations of Plasmodial parasites under CQ pressure revealed differential expression of 600 transcripts of which 41% were cell-cycle related (Gunasekera *et al.*, 2007). The transcriptome of doxycyclin-treated *P. falciparum* revealed a delayed death effect in which parasites were able to invade erythrocytes after the first cycle but died soon thereafter (Dahl *et al.*, 2006). This effect was most likely due to loss of the apicoplast function with loss of mitochondrion function as a secondary effect (Dahl *et al.,* 2006). Stress responses on the Plasmodial parasite elicited by febrile temperature (41^{o}C) commonly associated with malaria infections, identified 336 transcripts that were differentially regulated of which 162 (49%) had increased abundance and 173 (51%) decreased abundance (Oakley *et al.,* 2007). Severely affected transcripts were involved in stress responses, cell surface adhesion and, a large number of regulated transcripts containing a PEXEL sequence associated with protein export to the erythrocyte, therefore indicating a possible extrusion to the erythrocyte or even erythrocyte remodelling and parasite sequestration (Oakley *et al.,* 2007). Perturbation with artesunate resulted in 398 transcripts identified as differentially expressed of which 244 had increased abundance and 154 decreased abundance. The majority of the latter were classified as chaperones, transporters, kinases, and transcription activating proteins, oxidative stress and cell cycle regulation (Natalang *et al.*, 2008). The transcripts that displayed an

increase in transcript abundance contained a large number of export proteins and transporter transcripts which may result in drug resistance (Natalang et al., 2008). In one of two histone-related microarray studies, histone acetyltransferase was inhibited using anacardic acid which resulted in hypo-acetylation of histone H3 (Cui *et al.*, 2008). Of the 271 transcripts that were differentially expressed 207 were decreased and only 64 increased in abundance. This major decrease in transcript abundance is probably as a result of the hypo-acetylation which will lead to gene silencing and therefore have a pronounced effect on transcription (Cui *et al.*, 2008). In the other study, histone deacetylase activity was inhibited with apicidin which resulted in profound transcriptional changes within the parasites (Chaal *et al.*, 2010). Transcription factors were affected by the inhibition and especially schizont-stage genes were severely affected. Overall the inhibition of histone deacetylase resulted in complete de-regulation of the IDC transcriptional cascade (Chaal *et al.*, 2010). The most recent microarray study, the expression of genes in response to 20 different compounds resulted in arrest in the schizont stage (Hu *et al.*, 2010). A total of 3000 differentially affected transcripts were identified demonstrating that functionally related genes share transcription profiles even with all the different perturbations, since they may also share similar regulatory mechanisms that are associated with transcription rather than mRNA decay (Hu *et al.*, 2010).

4.1.6 Polyamine perturbation studies on Plasmodial parasites

The first polyamine depletion study on *Plasmodium* was done using suppression subtractive hybridization (SSH) in which pre-selected libraries were created and subsequently used for microarray analysis (Clark *et al.*, 2008). Plasmodial parasites were treated with DFMO to inhibit ODC and consequent putrescine depletion. Interesting polyamine-specific responses included the increased transcript abundance of OAT and hypoxanthine phosphoribosyltransferase (HPPRT) (Clark *et al.*, 2008). A follow-up study, the co-inhibition of AdoMetDC/ODC with DFMO and MDL73811, respectively induced total polyamine depletion (van Brummelen *et al.*, 2009), and revealed the differential regulation of 538 transcripts of which 171 had increased abundance and 377 decreased abundance. Treated parasites were arrested in the trophozoite stage while untreated parasites progressed through their life cycle, which therefore prompted the use of a t_0 reference time point to which the treated parasites best correlate. Analysis of the differentially regulated transcripts revealed a polyamine-related response with increased transcript abundance of OAT and lysine decarboxylase (LDC) and decreased abundance of AdoMet synthase (van Brummelen *et al.*, 2009). Inhibition of spermidine synthase (SpdS) with cyclohexylamine was also investigated at 3 time points and revealed the differential regulation of various polyamine-dependent transcripts (Becker *et al.*, 2010).

Microarray technologies enable the expression profiling of transcripts induced by perturbation. Given the advances in Plasmodial transcriptomics and the promising data obtained from 3 other polyamine depletion studies, the transcriptional response to inhibition of AdoMetDC with the irreversible inhibitor MDL73811, will be investigated in this chapter.

4.2 Methods

4.2.1 Culturing of parasites for transcriptomics

*Pf*3D7 parasites were maintained *in vitro* in human O^+ erythrocytes in culture media and monitored daily through light microscopy of Giemsa stained thin blood smears as described in Chapter 2 section 2.2.3. Before treatment could commence the parasites were always synchronised for 3 consecutive cycles (6 times in total, always 8 h apart once in the morning and later in the afternoon) as described in section 2.2.4. A starting parasite culture (in the schizont stage) at 2% parasitemia, 5% hematocrit was treated with 10 µM MDL73811 ($10\times IC_{50}$) at invasion after which the parasitemia increased to 10% in both the treated and untreated samples in the ring stage. Treatments were always done in duplicate cultures with 2 biological replicates for both treated and untreated samples. Cultures were split into 4 separate flasks of which 2 were the untreated parasites (control) and 2 were the MDL73811 treated parasites. Ten millilitres of parasite cultures at 10% parasitemia and 5% hematocrit were used per sample. A small scale morphology study was always conducted at the same time, and used as a positive control to ensure that complete inhibition (cell cycle arrest) of the *Pf*3D7 parasites did occur with the use of the MDL73811. At 3 different time points (t_1 = 16 hours post-invasion (HPI), t_2 = 20 HPI and t_3 = 26 HPI) cells were harvested by centrifugation at 2500×*g* for 5 min, followed by washing of the cell pellet twice with PBS. These time points were chosen due to the peak transcript production of AdoMetDC that occur between 18-40 HPI. The erythrocyte pellet containing the parasites was snap frozen and stored at -80°C until use.

4.2.2 RNA isolation from cultured parasites

RNA was isolated from untreated and treated *Pf*3D7 parasites for the 3 time points in an RNAse free environment using a combined RNeasy Mini Kit (QIAGEN) and TRI-Reagent (Sigma) method, with the incorporation of DNase I on-column digestion (QIAGEN). The RNA isolation procedure does not make use of saponin lysis of the erythrocytes, in order to prevent possible contamination of human and bovine RNA that may be released upon lysis of the erythrocytes. The tubes containing the frozen infected erythrocyte pellets were removed from -80°C and thawed. The pellet was loosened by flicking the tube, before the addition of 600 µl RLT lysis buffer (Proprietary, QIAGEN) to the pellet and mixed by vortexing. The mixture was transferred onto a QIA-Shredder column (QIAGEN) and centrifuged at 15 700×*g* for 2 min. The eluate of each QIA-Shredder column was divided into 2 equal parts and transferred to clean microfuge tubes. TRI-Reagent (600 µl) was added and mixed by vortexing after which the mixture was incubated at room temperature for 5 min. TRI-reagent contains phenol/guanidine that denatures proteins and therefore inhibits

possible RNAse activity. This was followed by the addition of 400 µl chloroform to each tube and vigorous vortexing. The chloroform separates the homogenate into an upper aqueous phase and a DNA interphase with the lower organic phase containing the denatured proteins. The chloroform containing mixture was incubated at room temperature for 10 min followed by centrifugation at 15 700×*g* for 15 min. The upper aqueous phase of each tube was transferred to clean microfuge tubes to which 700 µl of 70% (v/v) ethanol was added to precipitate the RNA. The mixtures in the two tubes that were split earlier were combined and 700 µl was loaded onto an RNeasy column and centrifuged at 8000×*g* for 15 s. This was repeated until all the sample was loaded onto a single column. Wash buffer RW1 (350 µl, proprietary, QIAGEN) was added and centrifuged at 8000×*g* for 15 s to wash the membrane containing the RNA. For the on-column DNase I digestion, 70 µl Buffer RDD (Proprietary, QIAGEN) was combined with a 10 µl aliquot of DNase I before the addition of this mixture directly onto the membrane. The membrane containing the DNase I mixture was incubated at room temperature for 15 min. The membrane was washed by the addition of 350 µl wash buffer RW1, centrifuged at 8000×*g* for 15 s followed by another two wash steps with 500 µl wash buffer RPE, and centrifuged at 8000×*g* for 15 s to remove all residual ethanol. The RNeasy column was placed in a clean 2 ml collection tube and centrifuged again at 15 700×*g* for 1 min to ensure that the membrane was dried completely and that absolutely no residual ethanol is present. The RNeasy column was transferred to another clean 2 ml microfuge tube, and 30 µl RNase free water was added directly onto the membrane and incubated for 2 min before centrifugation at 8000×*g* for 1 min to elute the RNA. The RNA concentration and purity was determined on the Gene Quant (GE Healthcare) and stored at -80°C until cDNA synthesis.

4.2.3 RNA integrity determination

The RNA integrity and purity was assessed using the Experion automated electrophoresis system (Bio-Rad). The RNA was prepared and run according to the manufacturer's instructions on a Lap-Chip microfluidic separation technology using a fluorescent sample detection method. All the gel-based steps which include sample preparation, staining, and destaining, imaging, band detection, and data analysis are automatically performed by the system (Delibato *et al.*, 2009).

4.2.4 cDNA synthesis from RNA

Due to the difficulty in isolating mRNA from *P. falciparum* parasites that is representative of the complete sample an approach was followed in which RNA was isolated from the parasites for cDNA synthesis (Bozdech *et al.*, 2003). A RNA reference pool was constructed by using 2 µg total RNA from each of the twelve samples resulting in a representative RNA reference pool. First strand

cDNA synthesis was initiated using 2 µg RNA of the reference or individual samples, 775 pmol random primer nonamer (Inqaba), 250 pmol OligodT (dT_{25}) (Inqaba) and incubated at 70°C for 10 min, followed by 10 min at 4°C. After this incubation step, 1.7 µl of the 10× aminoallyl dNTP mixture (10 mM dATP, 5 mM dCTP, 5 mM dGTP, 5 mM dTTP, 5 mM aminoallyl-dUTP), 6µl of the 5× SuperScript First-strand buffer, 10 mM DTT, 40 U rRNasin (RNAse Inhibitor, ProMega) and 340 U Superscript III Reverse Transcriptase (Invitrogen) were added, mixed and incubated at 42°C for 18-20 h for cDNA synthesis. The high amount of dNTPs allowed the synthesis of the A+T-rich cDNA of the Plasmodial genome. Contaminating RNA was removed by hydrolysis with the addition of 1 M NaOH, and 0.5 M EDTA, pH 8 to the reaction mixture and incubating at 65°C for 15 min. The cDNA were purified using the PCR Clean-up kit (Macherey-Nagel) and is based on the principle that the DNA binds to the silica matrix in the presence of choatropic salts. These salts are then removed by the addition of alcohol based buffers after which the DNA is eluted in water. In short, ten volumes of the binding buffer NT was added to the cDNA mixture and then transferred to a Nucleospin extract II column and incubated for 4 min on the column before centrifugation at 13 000×*g* for 1 min, followed by washing of the silica membrane. The cDNA was subsequently eluted by the addition of 30 µl pre-heated RNAse-free SABAX water (Adcock) (37°C) directly onto the membrane and incubated for 4 min before centrifugation at 13 000×*g* for 90 s to elute the cDNA. The cDNA concentration and purity was measured on a Nanodrop-1000 (Thermo).

4.2.5 cDNA labelling for hybridisation

cDNA (1.2–2 µg) for each of the individual samples were dried *in vacuo* and then resuspended in 2.5 µl SABAX water followed by the addition of 5 µl 0.2 M Na_2CO_3, pH 9.0 and 2.5 µl of the respective dye to each sample. The aminoallyl-dUTP incorporated during cDNA synthesis was used to couple the Cy-dyes to the samples. The reference pool was labelled with Cy3 (green) and the samples for each of the time points were labelled with Cy5 (red) using a common reference design (Figure 4.2).

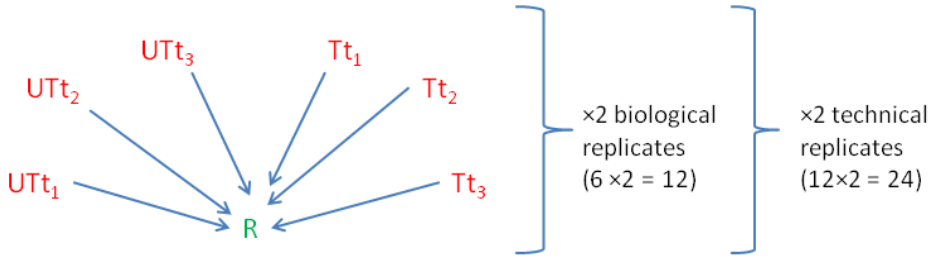

Figure 4.2: Common reference design used for the inhibition of AdoMetDC.
The reference was labelled with Cy3 (green) and all the samples were labelled with Cy5 (red). The reference sample contains equal amounts of all the samples and is therefore representative of all samples. 24 arrays were done in total since two biological replicates and two technical replicates of each sample were done.

The samples were incubated in a desiccator in the dark for 2 h at room temperature. This is done since the Cy-dyes used for labelling is sensitive to ozone degradation and because it is a fluorophore it is light sensitive. Excess dye was removed with the QIAquick PCR purification kit (QIAGEN). In short, 10 volumes of buffer PBI were added to the sample, mixed and applied directly onto a QIAquick column and incubated for 4 min before centrifugation for 1 min at 13 000×*g*. Three wash steps were performed by adding 500 µl Buffer PE followed by centrifugation at 13 000×*g* for 1 min, after which the membrane were dried. 30 µl pre-heated water (37°C) were applied directly onto the membrane and incubated for 4 min followed by elution at 13 000×*g* for 90 s. The dye incorporation and concentration were determined using the microarray setting on the Nanodrop-1000. The coupling efficiency was calculated (Equation 4.1), and should be at least 10 labelled nucleotides per 1000 nucleotides to proceed with hybridisation.

The coupling efficiency for each reaction was determined using the following formulae:

$$\text{Efficiency} = \frac{\text{pmol dye} \times 324.5\ \text{pg/mol}^*}{\text{ng DNA}} \qquad \text{Equation 4.1}$$

*324.5 pg/mol is the average mass of a dNTP

4.2.6 Slide assembly and sample preparation for oligonucleotide hybridisation

Equal amounts of the Cy5 labelled sample and Cy3 labelled reference pool (20 pmol each) were added in a PCR tube (40 pmol in total). For the 8×15K *P. falciparum* Agilent slides, 5 µl of the 10× Blocking buffer (Proprietary, Agilent), 1 µl of the 25× Fragmentation buffer (Proprietary, Agilent) and finally SABAX water to a total volume of 25 µl was added to the Cy-labelled samples. The mixture was incubated at 60°C for 30 min to fragment any remaining RNA. This was followed by the addition of 25 µl 2× GE hybridisation buffer (Proprietary, Agilent) to the sample mixture. It is important not to vortex this mixture as this will introduce bubbles, and should therefore be mixed by careful pipetting and then put on ice during array loading. 40 µl of each sample was loaded onto

each array to obtain a final sample concentration of 20 pmol. Care should be taken not to introduce bubbles as this might introduce problems during hybridisation. Before sample loading the slide is assembled by loading a gasket slide into the assembly chamber. The 40 µl sample was then loaded onto the gasket slide. The slide containing the printed arrays was placed on top of the gasket slide containing the samples. The arrays were sealed by tightening the screw onto the chamber. The chamber containing the arrays was placed in the hybridisation oven at 65°C for 17 h at a rotational speed setting of 10.

4.2.7 Post-hybridisation, washing and slide scanning

After hybridisation the array slide was removed from the chamber and disassembled in wash buffer 1 (Proprietary, Agilent). The slide was washed twice in pre-heated (37°C) wash buffer 2 (Proprietary, Agilent) for 1 min each. Finally the slide was dried in a centrifuge for 1 min after which it was scanned on an Axon GenePix 4000B scanner (Molecular Devices).

4.2.8 Data analysis

The original *P. falciparum* Operon Array containing 8089 70-mer gene probes was adapted to the 60-mer Agilent platform by Mr J Verlinden (J. Verlinden, MSc thesis in preparation). In short, all the unnecessary 'NULL' annotations and corresponding oligonucleotide sequences, used as controls specific to the Operon platform, were removed. The remaining 7004 oligonucleotides were adapted to the 60-mer Agilent platform by shortening their sequences using a 10-mer scanning window from both 3' and 5' ends and keeping the annealing temperature close to 65°C using the following equation:

$$*Tm = 64.9°C + 41°C\ (GC-16.4)/N \qquad \text{Equation 4.2}$$

*Tm is the annealing temperature for microarray hybridisation, GC is the number of G and C nucleotides in a target sequence, N is the total length of the sequence

The shortened sequences were validated by submitting the target sequence to NBLAST (www.ncbi.nlm.nih.gov/BLAST). All sequences submitted to NBLAST analysis had E-values below 10^{-6}. In addition to adapting the *P. falciparum* Operon Array, the most recent annotated form of the *P. falciparum* (strain 3D7) genome from PlasmoDB 5.4 (www.plasmodb.org) was used to design a new 60-mer based array to overcome ambiguities in previous annotations used for the Operon array dataset. The Agilent 60-mer probes were designed by submitting the FASTA files into ArrayOligoSelector (AOS) in order to design the 60-mer probes from the various target sequences

(http://arrayoligosel.sourceforge.net/). The designed target sequences were once again validated using NBLAST. The designed probes were then submitted to Agilent for printing of the slides.

Each array was analysed with Axon GenePix Pro 6.0 software (Molecular Devices). The spots of each array were analysed according to the criteria of the five parameters in Table 4.1 (parameter and function for the cut-off values). Spots not fulfilling the specified criteria were flagged and received a zero weight value (Table 4.1).

Table 4.1: Parameters set for automated spot detection using GenePix.

Parameter	Function[a]	Flag[b]	% of flagged spots[c]
Circularity of spots	[Circularity] < 40 Or [F Pixels] < 50	Bad	0
CV of scan channels	[F532 CV] > 400 Or [F635 CV] > 400	Bad	0.2 – 3.8
Intensity	[F532 Mean] < 150	Absent	15.2 – 36.9
Saturation	[F532 % Sat.] > 20 Or [F635 % Sat.] > 20	Bad	0.3 – 2.9
Signal to noise ratio	[SNR 532] < 3 And [SNR 635] < 3	Bad	19.8 – 45.4

Automated spot detection parameters used for GenePix to flag spots if they did not qualify according to the criteria set. [a]is the function used to set the criteria in GenePix. [b]flag is how the spot is marked as not useable. [c]the percentage range for each of the parameters set for the 24 arrays and is indicative of the percentage of spots that were flagged for each of the parameters. F Pixels is the minimum number of pixels for an intensity measurement. 532 nm is the red channel and 635 nm is the green channel. CV is the coefficient of variation. Circularity is a measure of the shape of the spot.

For further data analysis and identification of differentially expressed transcripts, the LIMMA (Smyth *et al.*, 2005b) and LIMMA-GUI (Wettenhall & Smyth, 2004) packages were used. This included the mArray package that provides alternative functions for reading and normalising spotted microarray data and overlaps with the LIMMA package, while LIMMA-GUI provides more graphical user interfaces. All these packages are from Bioconductor (www.bioconductor.com) and are freely available. Background correction was performed on all arrays with a subset of 50 (Ritchie *et al.*, 2007). Within-array normalisation made use of robust spline normalisation, followed by between-array normalisation making use of Gquantile normalisation due to the common reference that was used and always labelled with Cy3 in the green channel (Smyth & Speed, 2003). Pearson correlations were calculated for each of the time points. The differentially expressed transcripts were determined by comparing UTt_3 to Tt_3 and making use of the linear modelling approach (lmFit) and the empirical Bayes statistics (eBayes) (Smyth *et al.*, 2005a). All transcripts with a $\log_2$ ratio $\geq$ 0.75 or $\leq$ -0.75 and a p-value of less than 0.05 were considered as significant. This related to a fold change (FC) of 1.7 and -1.7. The FC is calculated from the $\log_2$ ratio in Microsoft EXCELL™ 2007 using the equation: FC = POWER(2, $\log_2$-value) Equation 4.3

Therefore, a $\log_2$-ratio of 1 is similar to a FC-value of 2 and a $\log_2$-ratio of 0.75 is equal to a FC-value of 1.7. The list of differentially regulated transcripts was submitted to PlasmoDB 6.0 (www.plasmodb.org) to obtain the GO terms of each of the transcripts. The transcripts were then

manually sorted according to their biological function and GO term and grouped together. Each group were submitted to MADIBA (www.bi.up.ac.za/MADIBA) to verify the GO terms and groupings. To determine the biological pathway most affected by the perturbation study all the differentially expressed transcripts were submitted to MADIBA to determine the significant metabolic pathway affected, together with a p-value as calculated by Fishers test (Fisher, 1935).
Hierarchical clustering was performed on all the data using CLUSTER 2.1.1 (http://rana.stanford.edu.software). Only data that had expression values in all of the time points were used for clustering to avoid blank spots upon clustering. Clustering of data was performed using average linkage clustering which indicates that the distances between transcripts are calculated on average vales and uncentered symmetric correlation which assumes that the average is zero. Clustering data was visualised in TREEVIEW 1.6 (www.EisenSoftware/ClusterTreeView/TreeView).

The STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) 8.0 database covers 2.5 million proteins from 630 organisms (http://string.embl.de) (Jensen *et al.*, 2009). The curated database is able to provide a comprehensive view of protein-protein interactions (http://string.embl.de) (Jensen *et al.*, 2009) by experimental repositories, computational prediction methods and public text collections. STRING scores and weighs protein interactions. The basic interaction unit is the functional association which is specific and meaningful between two proteins that jointly contribute to the same functional process. AdoMetDC was submitted to the web-based programme to determine possible protein-protein interactions.

Finally, the differentially affected transcripts from the AdoMetDC-inhibited parasites were submitted to the *P. falciparum* interactome (www.plasmomap.org) (Date & Stoeckert, 2006, Wuchty *et al.*, 2009). The *P. falciparum* interactome was constructed *in silico* using Bayesian frameworks (Date & Stoeckert, 2006, Wuchty *et al.*, 2009).

Comparisons made between the AdoMetDC-inhibited data and the artesunuate, CQ, and febrile temperature studies as well as the comparisons made between the AdoMetDC-inhibited data and the co-inhibition study and spermidine synthase inhibition study was done using Microsoft EXCELL™ 2007. The PlasmoID identifiers of each study were submitted to the EXCELL™ worksheet. The VLOOKUP function in EXCELL™ was used to compare all the different studies at once and considered only if the PlasmoID was present or not.

4.2.9 Validation of microarray results with qRT-PCR

The treated and untreated cDNA from all the samples were diluted to 0.65 ng/µl with SABAX water for use in qRT-PCR. A standard curve was constructed from a dilution series of UTt_1 samples that contained the following dilutions: an undiluted sample, 1/10, 1/20, 1/50 and 1/100 dilutions. Cyclophillin was used as household transcript and used to construct the standard curve. The reactions were performed in a 384-well plate using the Lightcycler 480 (Roche). The total reaction volume was 10 µl and consisted of 5× KAPA SYBR FAST qPCR reaction mixture, 0.1625 ng/µl cDNA and 1 pmols each of both the forward and reverse primers for each individual reaction. The reaction mixture was pre-incubated at 95ºC for 10 min. This was followed by 48 amplification cycles that each consisted of denaturation at 95ºC for 10 s, annealing at 55ºC for 5 s and extension at 72ºC for 7 s. Fluorescence was detected at the end of each cycle. Amplification was followed by melting curve analysis to detect possible primer-dimers and to determine the specific melting temperature (TM) of each product. The thermal profile of melting curve analysis consisted of incubation at 95ºC for 5 s, 65ºC for 5 s and 95ºC for denaturation with continuous measurement of fluorescence. Finally the reaction mixture was cooled to 40ºC for 30 s. The fold change was calculated for each sample by comparing the untreated samples to the treated samples for each specific time point and then normalised to cyclophillin that remained unchanged in all 3 time points.

4.3 Results

4.3.1 RNA quality assessment

Parasite culture, *Pf*3D7 (10 ml) at 10% parasitemia were sampled at 3 time points (t_1: 16 HPI, t_2: 20 HPI, and t_3: 26 HPI) for RNA isolation and subsequent transcriptomic investigation of MDL73811-treated parasites (Figure 4.3). Figure 4.3 is a morphological representation of the parasites harvested at the 3 time points used for the transcriptomic investigation.

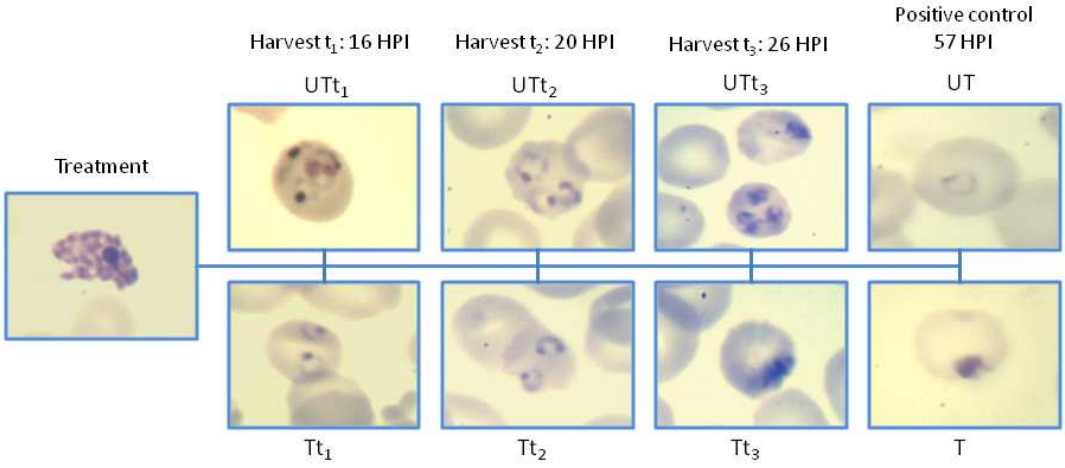

Figure 4.3: Transcriptomic sampling points.
Treatment with MDL73811 was done at invasion. Three time points were taken for the transcriptomic analysis (t_1: 16 HPI, t_2: 20 HPI, and t_3: 26 HPI). The positive control is included to ensure arrest of the MDL73811-treated parasites while the untreated parasites have progressed to a new life cycle with rings being formed.

The RNA integrity number (RIN) or RNA Quality Indicator (RQI) provides the most accurate and reproducible account of RNA integrity, and requires very small amounts of sample per run (Imbeaud et al., 2005). Three of the 12 RNA samples (UTt_3, Tt_3, and Tt_1) were therefore chosen randomly to run on the Experion system (Bio-Rad). All 3 samples (Ut_3, Tt_3 and Tt_1) had RQI values in excess of 9 indicating high quality, non-degraded RNA (Figure 4.4). The RQI number is based on 3 regions of the electropherogram and should be between 1 and 10, with 1 the most degraded, and 10 the most intact RNA. RQI values between 7 and 10 are considered as good intact RNA samples, but the compatibility of the RQI number with down-stream applications should be determined by the user (Buhlmann *et al.*, 2004).

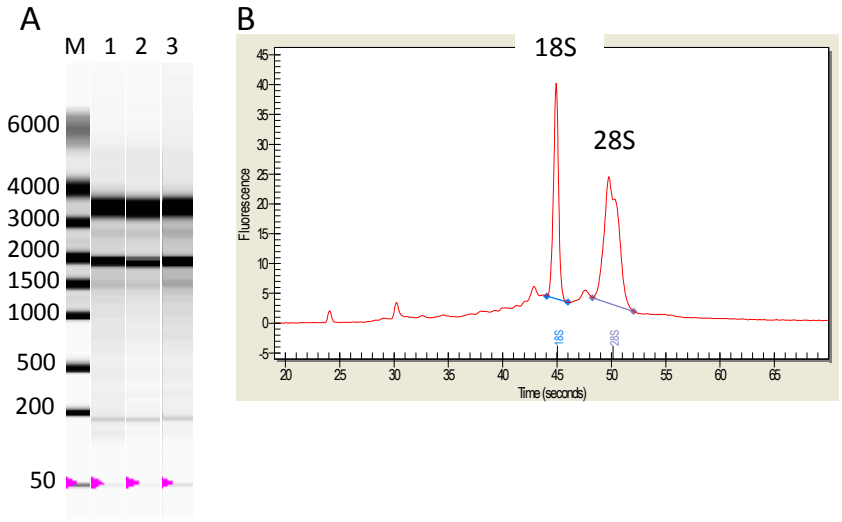

Figure 4.4: Assessment of RNA purity and integrity from the *P. falciparum* RNA that will be used for microarray analysis.

(A) Virtual gel image of RNA samples with lane 1: UTt_3(9.1; 1.47), lane 2: Tt_3(9.4; 1.5), lane 3: Tt_1(9.0; 1.49). RQI number is the first number given in brackets, followed by the 28S/18S ratio. (B) is representative of the electropherogram indicating the 18S and 28S rRNA subunits. No RNA degradation is visible since no peaks is found except the rRNA subunits as indicated.

The RNA electropherogram in Figure 4.4 illustrates the clearly distinguishable bands, the lack of smears, and the 28S rRNA subunit that has higher intensity than the 18S rRNA subunit therefore indicating good RNA quality (Copois *et al.*, 2007). Good RNA yields (ranging from 3 µg in the ring stage to 18 µg in the untreated trophozoite stages) were obtained and none of the RNA had any indication of either protein or DNA contamination and was therefore used for the microarray analysis to follow.

4.3.2 Microarray preparation

A Plasmodial Agilent microarray platform was used that enabled the simultaneous analysis of 8 different samples on a single slide (J. Verlinden, MSc thesis in preparation). Each of the slides used consisted of 8×15 K hybridisation chambers that could each be prepared for a different sample. The Agilent arrays required only 4 µg RNA per sample for a complete experiment. Figure 4.5 is an example of a typical Agilent array from *Pf*3D7 representing an untreated and treated array from time point 3. The 4 corners of each of the arrays contained several control spots that was used for assessment of hybridisation of the samples (Figure 4.5). The treated array has an overall yellowish colour that is associated with differences between the treated and untreated arrays (Van Brummelen, 2009). Also notice the differences on the enlarged inserts between the Tt_3 array which is yellowish

or green whereas the UTt_3 array has several yellow, green and red spots associated with the differential expression of transcripts as the parasite progress through its life cycle. Due to the use of the inkjet technology of the Agilent system all the spots are exactly the same size, which was previously tedious to achieve. Overall, the use of the Agilent arrays resulted in better quality microarray data, and confidence in analysis.

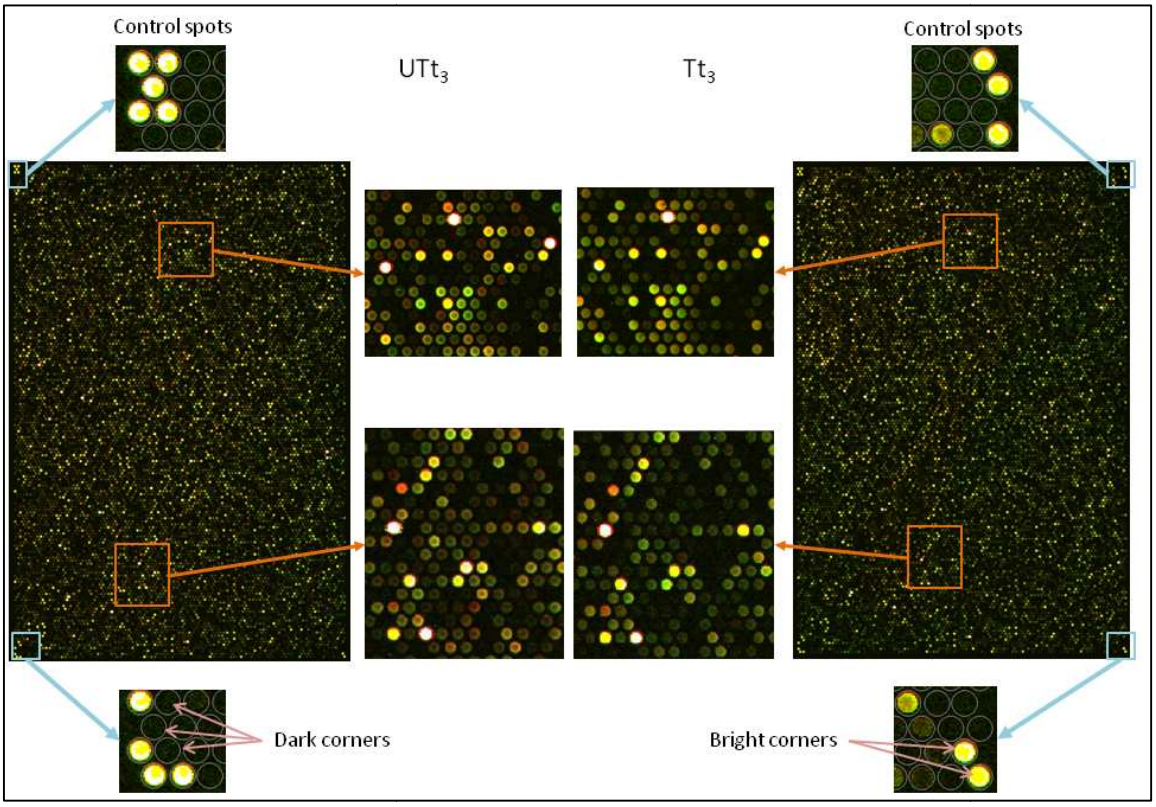

Figure 4.5: The 60-mer Agilent array for one UTt_3 and one Tt_3.

The control spots are incorporated in the corners of each array. Enlarged images from UTt_3 and Tt_3 (images in the middle indicate spot colour differences between the treated and untreated samples). Control spots are also added to each of the arrays and are in the four corners of each array as indicated. The yellow spots in the control boxes are indicative of the bright corners that should always be yellow and dark corners in which no cDNA have been hybridised to the slide.

4.3.3 Normalisation of data

The dataset comprised of 24 individual hybridisations and the first step is usually background correction, if necessary. As shown in Figure 4.6, some localized artefacts (at the top and at the edges) may occur on the arrays that only affect a specific channel and therefore needs correction. Artefacts localised in the green channel (indicated with blue arrows) are not seen in the red channel. The red channel has more speckling (indicated with grey arrow), probably due to insufficient washing and therefore also needed background correction.

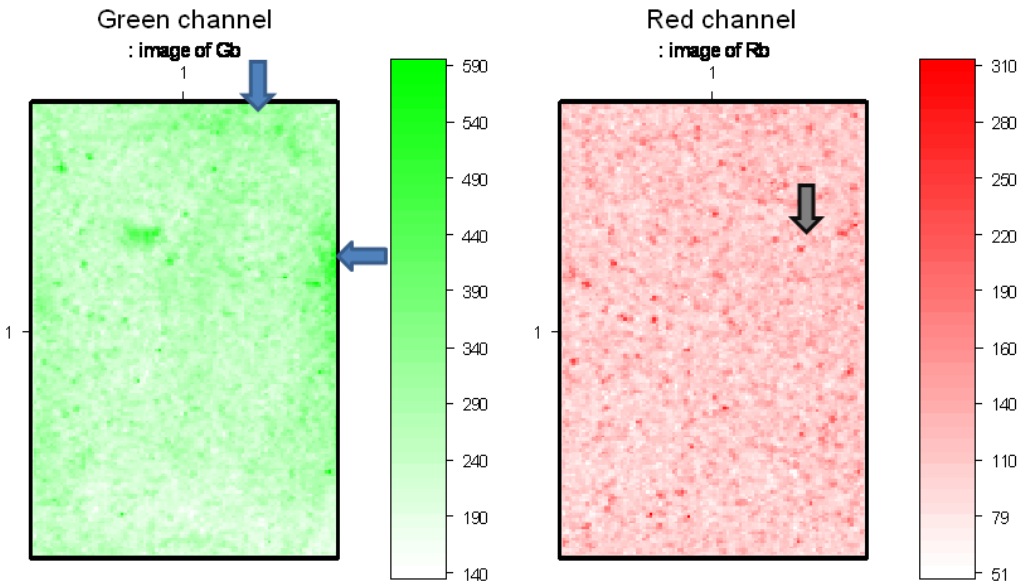

Figure 4.6: Red and green background images of slide Tt_3 array8.

An offset of 50 was used for background subtraction, which is the preferred methodology when microarrays are to be used for differential expression analysis (Smyth & Speed, 2003). Noise is reduced for each slide in every step that is performed and accordingly adjusts the foreground intensity for the background intensity. An important aspect of background subtraction necessary for differential expression analysis is the representation of only positive values on the array with low intensity spots and negative values that are converted to zero resulting in lower log-ratio variation as well as the stabilisation of the variability of the M-value as a function of intensity (Smyth et al., 2005b). The offset method for background subtraction has the benefit of providing small variability and more reliable data. Two normalisation strategies exist, which include within-array normalisation in which the M-value are normalised for each array separately to correct for dye effects, and between-array normalisation in which the intensities or log-ratios are normalised in order to be comparable across all the arrays that are compared (Chiogna *et al.*, 2009). Print-tip Loess normalisation is the standard within-array normalisation method, but is not applicable to Agilent slides, since these slides do not have print tips, and therefore Global Loess or Robust Spline would be more appropriate. Global Loess normalisation assumes that the majority of probes are not differentially regulated but inefficient normalisation was evident since various outliers were still visible within the data (Figure 4.7 A and B, indicated with arrows) and had a relative large amount of noise (Figure 4.7 C).

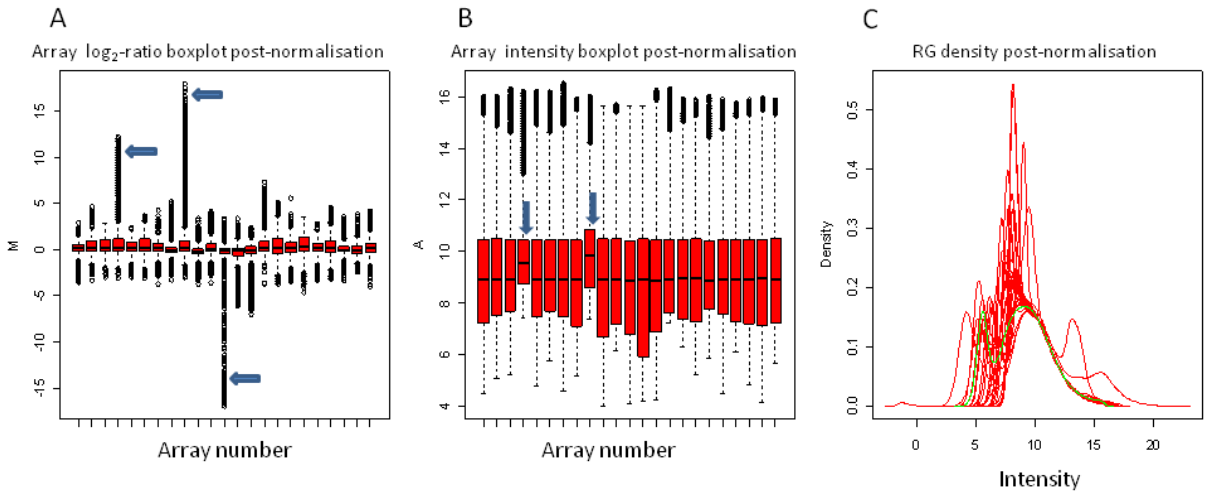

Figure 4.7: Boxplot data after Loess normalisation and Gquantile.

A: Boxplot of the log2-ratios after normalisation, B: Boxplot of the intensities for each array, C: Density plot after normalisation. Visible outliers are indicated with arrows.

The M-values of 3 slides were clear outliers (with values of 10, 15 and -15), while the other 21 slides had M-values ranging between -5 and 5 (Figure 4.7 A, marked with arrows). The RG density plot (Figure 4.7 C) appears erratic and not smooth after normalisation. Figure 4.7 B also indicates the large box sizes that will skew the dataset. The variation in box sizes poses a problem since the bigger boxes have a larger influence on the data than the smaller boxes (Smyth & Speed, 2003). Inefficient normalisation with Loess prompted the use of Robust Spline within-array normalisation method. Robust Spline is an empirical Bayes compromise between print-tip and global Loess normalisation, with 5-parameter regression splines used in place of the Loess curves. Robust Spline analysis resulted in more stable M-values between 4 and -4 and stable A-values (Figure 4.8 A and B), which is in stark contrast to the M-values of 15 obtained for Loess normalisation. Robust Spline filters were originally designed for surface texture analysis and are able to deal with outliers in data without affecting the mean data (Krystek, 2005).

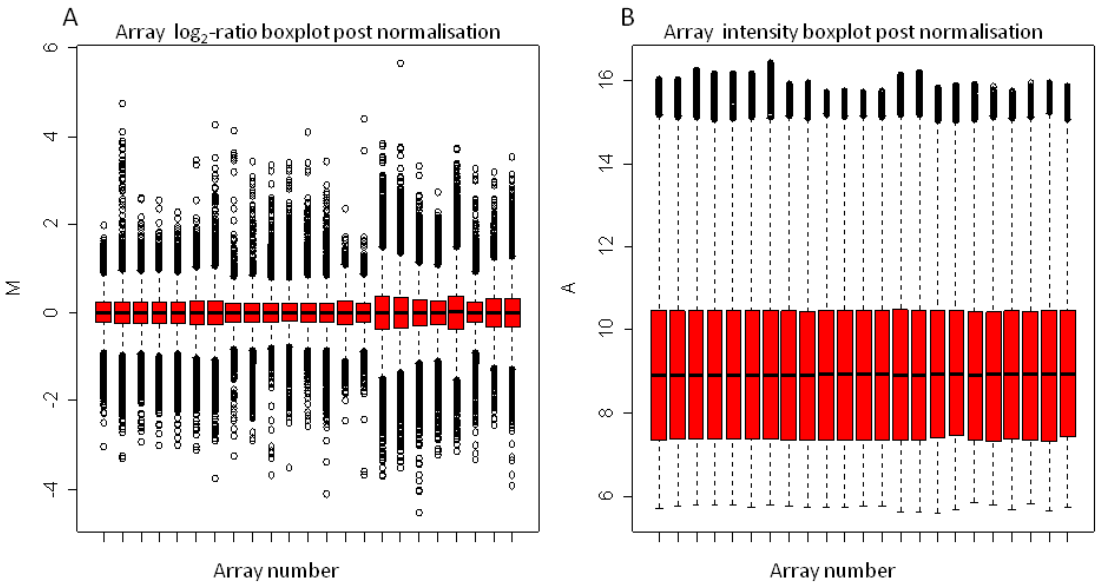

Figure 4.8: Boxplots of Robust Spline normalisation

A: All the boxes are centred on zero and are of similar sizes after normalisation indicating that all the boxes have a similar influence on the data. B: the spot intensities are all similar since all the boxes has the same A-value and are of equal size.

Dye differences may also play an important role and needs to be corrected. The dyes have various differences that include their chemistry, half-life, dynamic range, and susceptibility to degradation by ozone. Taken together all these differences may result in signal discrepancies and therefore it needs normalisation to produce similar signals from both dyes since it is assumed that the starting material for both dyes are similar (Meiklejohn & Townsend, 2005). Between-array normalisation normalises the individual (red and green) intensity values rather than the $\log_2$-ratios. It is also important for between-array normalisation that the background has been corrected to provide quality data.

Gquantile was used since all the reference samples were always labelled with Cy3 (green) and the different samples were always labelled with Cy5 (red). The use of a reference pool and labelling of the reference pool with only one dye may aid in normalisation (Kreil *et al.*, 2005). Therefore, the reference samples were all normalised to a single green line (Figure 4.9 B) since it is essentially the same reference used on all of the slides. As shown in Figure 4.9 the combination of Robust Spline and Gquantile produces a smooth density plot post-normalisation.

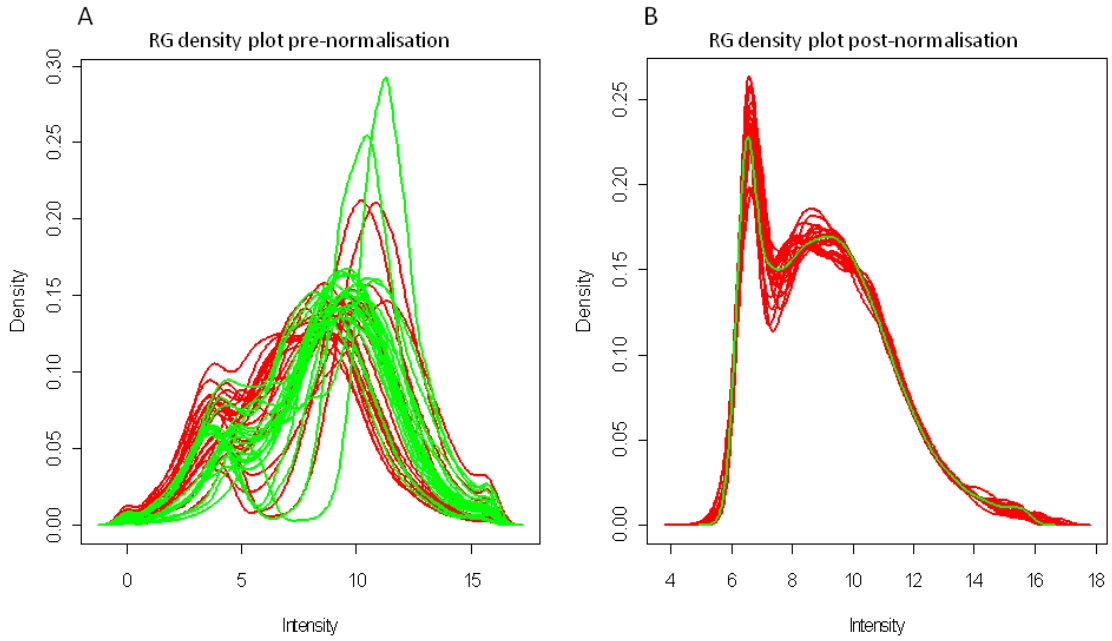

Figure 4.9: RG density plots after Robust Spline and Gquantile normalisation.
A: RG density plot pre-normalisation, B: RG density plot post normalisation with Robust Spline and Gquantile.

MA plots give an indication on the quality of the microarray data before and after normalisation. On MA plots the M-value should ideally be around zero and the ratios should ideally not be dependent on intensity (A-value). Upon normalisation of the data the MA-plots changed into more desirable data as the spots were concentrated around zero and is parallel to the intensity axis (A-value) (Figure 4.10 A and B). It is expected that there should be no variance between M and A, but this is not the case with small variances being detected in the relationship with M versus A. The variance of M is larger for small A-values, stable for the middle A-values and once again a slight reduction in variance in the larger A-values. This is one of the reasons that the MA plot does not give a straight line on M = 0.

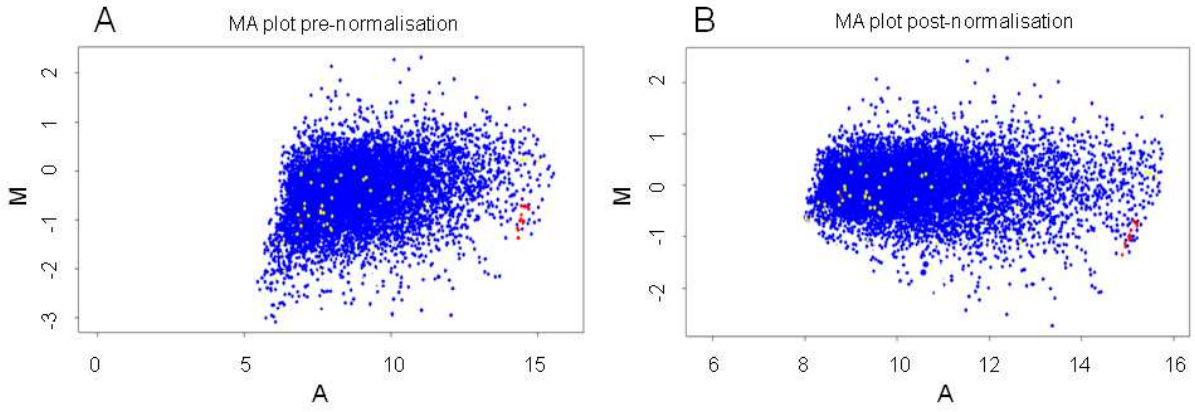

Figure 4.10: MA plot of Tt_3 array 6 before and after normalisation of the data.
(A) MA plot before any normalisation methods were employed, (B) MA plot after normalisation of the data.

4.3.4 Pearson correlations of the three time points

Spot finding was performed on all the arrays using GenePix 6.0, in which poor quality and saturated spots were removed from the dataset. Pearson correlations of all the valid spots present within the arrays were grouped into the 6 groups and compared (Table 4.2). The Pearson correlations were calculated after normalisation in LIMMA-GUI.

Table 4.2: Pearson correlations of the *Pf*AdoMetDC inhibited transcriptome data.

Comparison	Correlation (r)
UTt_1 : Tt_1	0.865
UTt_2 : Tt_2	0.608
UTt_3 : Tt_3	0.584
UTt_1 : Tt_2	0.396
UTt_1 : Tt_3	-0.327
UTt_2 : Tt_1	0.096
UTt_2 : Tt_3	0.212
UTt_1 : UTt_2	0.188
UTt_1 : UTt_3	-0.529
Tt_1 : Tt_3	-0.312
UTt_3 : Tt_1	-0.531

A Pearson correlation is observed between UTt_1 and Tt_1 (0.865) but progressively reduces to 0.584 between UTt_3 and Tt_3. This, as well as the anti-correlation detected between UTt_1 and Tt_3 (-0.327) and UTt_1 compared to UTt_3 (-0.529), indicate the progression of the parasite from the initial ring to the late trophozoite stages. Therefore, the early time points that were used within this study negate the use of the t_0 strategy due to the early time points taken (van Brummelen *et al.*, 2009) and allowed a direct comparison between time points for data analysis.

4.3.5 Data analysis of differentially expressed transcripts

A volcano-plot displays the fold changes as a measure of the statistical significance of the change (Smyth *et al.*, 2003). The volcano-plots for the data are given in Figure 4.11 A-C. Analysis of the first treated time point (UTt1:Tt1) did not result in the identification of any differentially expressed transcripts, and UTt_2:Tt_2 resulted in only a few differentially expressed transcripts that were significant ($p<0.05$). Time point 3 (UTt_3:Tt_3) resulted in the identification of 549 differentially expressed transcripts that consisted of 143 transcripts that had increased abundance and 406 transcripts with decreased abundance (Figure 4.11). The distribution of the $\log_2$-ratios for the differentially expressed transcripts from t_3 (UTt_3:Tt_3) indicated that the majority of the transcripts have a $\log_2$-ratio of about 1 (similar to a FC of 2).

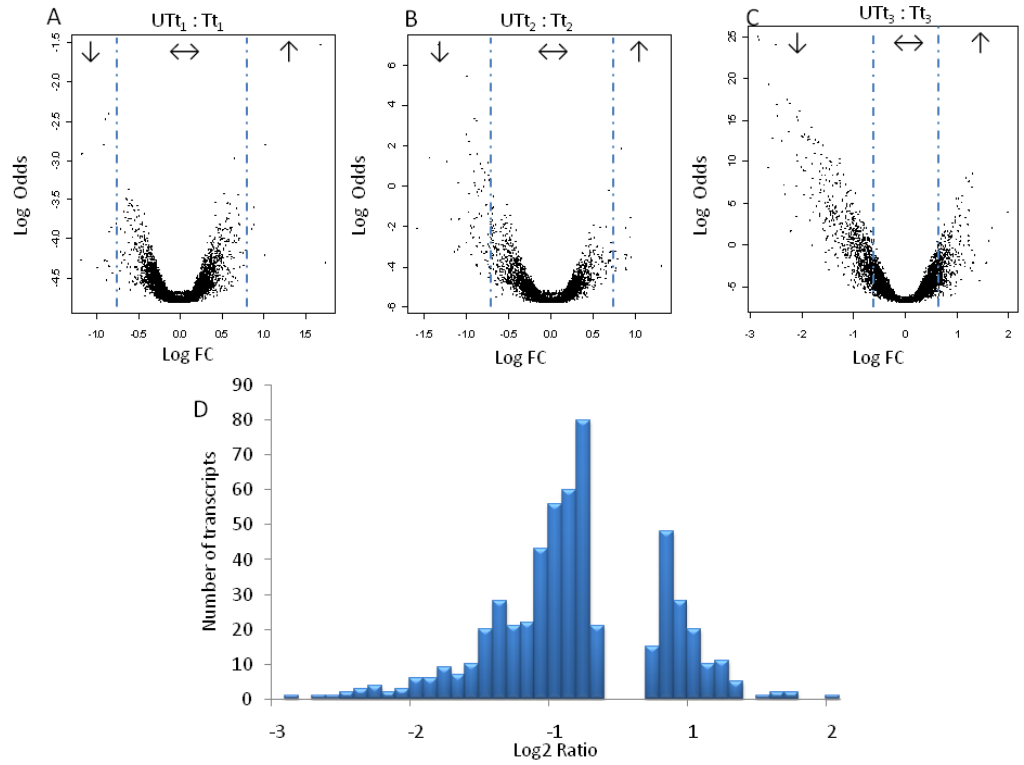

Figure 4.11: Log_2-distribution ratios and volcanoplots of the 3 time points investigated upon inhibition of AdoMetDC.

(A) t_1 (UTt_1:Tt_1), UTt_1 compared to Tt_1, resulted in no differentially expressed transcripts, (B) t_2 (UTt_2:Tt_2), UTt_2 compared to Tt_2, resulted in few differentially expressed transcripts, (C) t_3 (UTt_3:Tt_3), UTt_3 compared to Tt_3 resulted in 549 differentially expressed transcripts. (D) The distribution of log ratios for the differentially expressed transcripts from t_3. The log_2 ratio cut-off was ± 0.7, which relates to a fold change of ± 1.7.

The transcripts with decreased abundance represent ~74% (406/549) of the differentially expressed transcripts, while the transcripts with increased abundance are representative of ~24% (143/549) of the differentially regulated transcripts. The 25 most profoundly affected transcripts with a decrease in abundance are given in Table 4.3 and have a maximum fold change of -7.3 for α-tubulin which progressively decrease to -3.8 for a putative transporter. The increased abundance transcripts were less profound with the 25 most abundant transcripts having a fold change of 4.0 decreasing to a fold change of 2.2 (Table 4.3). A complete list of all 549 differentially regulated transcripts is available in Appendix B.

Table 4.3: The 25 most increased and decreased transcripts for AdoMetDC inhibited parasites.

Nr	PlasmoDB ID	Product Description	FC[a]	adj.P.Val[b]
		Decreased abundance		
1	PFI0180w	Alpha tubulin	-7.3	1.1E-11
2	PF11_0282	Deoxyuridine 5'-triphosphate nucleotidohydrolase, putative	-6.3	5.8E-06
3	PFI0905w	Probable protein, unknown function	-6.3	1.6E-09
4	PF13_0328	Proliferating cell nuclear antigen	-5.7	1.5E-11
5	PFI0135c	Serine repeat antigen 9 (SERA-9)	-5.6	1.2E-08
6	PF10_0154	Ribonucleotide reductase small subunit, putative	-5.4	1.1E-07
7	PF10_0084	Tubulin beta chain, putative	-5.2	3.5E-07
8	MAL13P1.214	Phosphoethanolamine N-methyltransferase	-5.1	2.5E-05
9	PFD0830w	□□□□nctional dihydrofolate reductase-thymidylate synthase	-4.9	1.1E-08
10	PF14_0443	Centrin-2	-4.9	8.8E-09
11	PFL1720w	Serine hydroxymethyltransferase	-4.8	1.6E-06
12	PF07_0065	Zinc transporter, putative	-4.8	1.8E-07
13	PFA0520c	Chromatin assembly factor 1 protein WD40 domain, putative	-4.7	1.1E-08
14	PF13_0032	Hydrolase, putative	-4.7	2.2E-03
15	PFF0510w	Histone H3	-4.3	5.2E-04
16	PFB0835c	Conserved Plasmodium protein, unknown function	-4.3	1.6E-06
17	PF13_0192	Conserved Plasmodium protein, unknown function	-4.2	5.5E-08
18	MAL13P1.303	Polyadenylate-binding protein, putative	-4.1	2.2E-08
19	PF10_0020	Alpha/beta hydrolase, putative	-4.1	4.8E-06
20	PFL2005w	Replication factor C subunit 4	-4.0	2.7E-07
21	PFF0630c	Conserved Plasmodium protein, unknown function	-4.0	2.7E-07
22	PF14_0053	Ribonucleotide reductase small subunit	-3.9	7.0E-06
23	PFL1670c	Conserved Plasmodium protein, unknown function	-3.9	7.2E-05
24	PF14_0309	Protein-L-isoaspartate O-methyltransferase beta-aspartate	-3.9	2.3E-07
25	PFA0245w	Transporter, putative	-3.8	7.0E-06
		Increased abundance		
1	PF13_0314	Conserved Plasmodium protein, unknown function	4.0	4.6E-04
2	PFB0115w	Conserved Plasmodium protein, unknown function	3.2	1.9E-03
3	PF14_0182	Hypothetical protein	3.2	7.4E-03
4	PFB0475c	Conserved Plasmodium protein, unknown function	3.1	7.2E-03
5	PFB0815w	Calcium-dependent protein kinase 1	3.0	4.0E-02
6	MAL13P1.328	DNA topoisomerase VI, B subunit, putative	2.8	1.3E-06
7	PFB0923c	Plasmodium exported protein, unknown function	2.5	1.8E-02
8	PF14_0698	Conserved Plasmodium protein, unknown function	2.5	3.4E-02
9	PF14_0015	Aminopeptidase, putative	2.5	3.3E-02
10	PFD0285c	Lysine decarboxylase, putative	2.5	8.6E-06
11	PFI1780w	Plasmodium exported protein (PHISTc), unknown function	2.5	3.1E-05
12	PF14_0017	Lysophospholipase, putative	2.4	3.4E-04
13	PFB0920w	DNAJ protein, putative	2.4	8.6E-03
14	PFD1175w	Serine/Threonine protein kinase, FIKK family	2.4	1.8E-05
15	PF10_0307	Conserved Plasmodium protein, unknown function	2.4	3.0E-04
16	MAL8P1.330	Conserved Plasmodium protein, unknown function	2.3	5.2E-03
17	PF10_0034	Conserved Plasmodium protein, unknown function	2.3	9.6E-06
18	PFA0130c	Serine/Threonine protein kinase, FIKK family, putative	2.3	1.0E-03
19	PF14_0018	Plasmodium exported protein (PHISTb), unknown function	2.3	1.2E-05
20	PFL1885c	Calcium/calmodulin-dependent protein kinase 2	2.2	3.2E-03
21	PF08_0118	Conserved Plasmodium protein, unknown function	2.2	3.4E-02
22	PF08_0060	Asparagine-rich antigen	2.2	4.5E-02
23	PF14_0703	Conserved Plasmodium protein, unknown function	2.2	7.3E-03
24	PFE0340c	Rhomboid protease ROM4	2.2	3.3E-02
25	PF08_0001	Plasmodium exported protein, unknown function	2.2	7.7E-05

[a]FC is representative of the fold change calculated from the $\log_2$-ratios given in LIMMA. [b]Adjusted p-value calculated by LIMMA to determine the statistical significance of transcripts and to avoid false positives with $p<0.05$ considered as significant.

4.3.6 Biological classification of differentially expressed transcripts

The 549 differentially regulated transcripts were sorted according to their biological functions and grouped according to their Gene Ontology (GO) annotations (Figure 4.12).

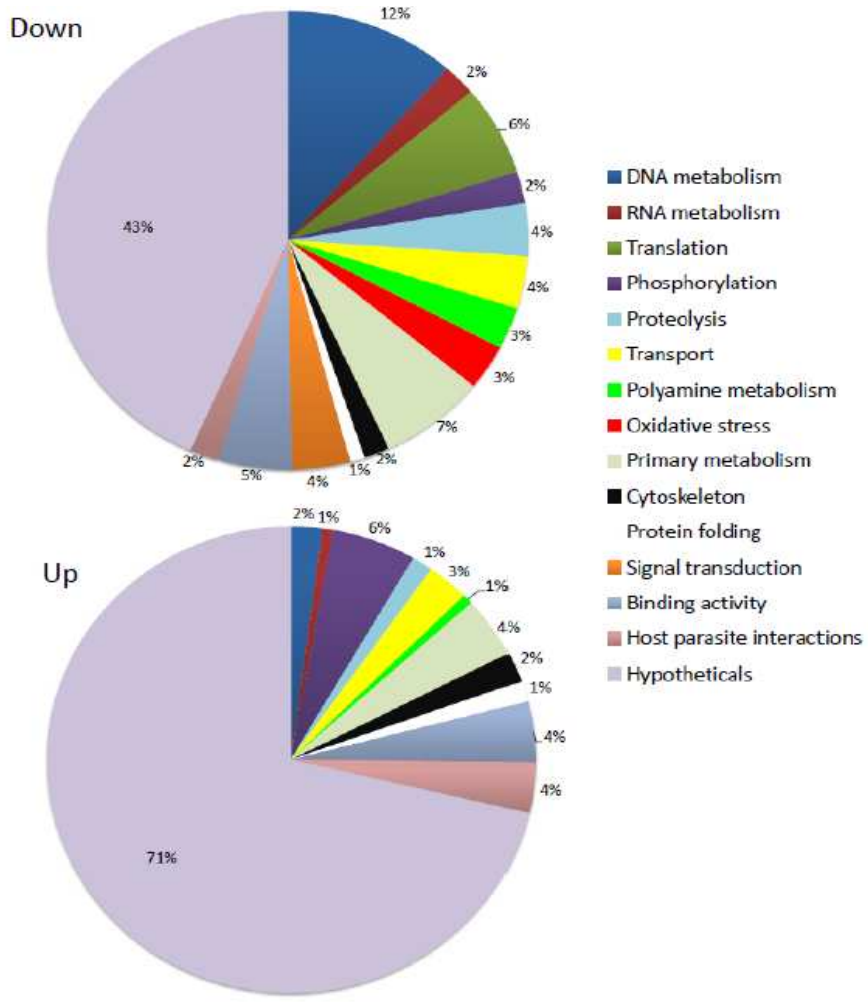

Figure 4.12: Functional classification of regulated transcripts according to their GO terms.
GO terms of differentially regulated transcripts were obtained from PlasmoDB 6.0 and classified according to their biological functions. Percentages were calculated from the total number of up- or down-regulated transcripts.

Of the 143 transcripts that had increased abundance, 6% of the transcripts are involved in phosphorylation. Primary metabolism, host-parasite interactions and binding activity each represented 4% of the biological functions, with transport represented by 3% of the differentially affected transcripts. The majority (71%) of the transcripts with increased abundances were hypothetical transcripts. The 406 transcripts that had decreased abundance consisted of 43% hypothetical transcripts. DNA metabolism (12%), translation (6%) and RNA metabolism (2%). Polyamine metabolism (3%) and oxidative stress (3%) also represented transcripts with decreased abundance. Transcripts associated with phosphorylation included 2% of the decreased transcripts.

Polyamine and methionine metabolism included the increased abundance (2.5-fold) of lysine decarboxylase (PFD0285c)(Table 4.4). Transcripts directly involved in methionine and AdoMet metabolism, AdoMet synthetase (PFI1090w; -2.3-fold) and adenosylhomocysteinase (PFE1050w; -2.1-fold), all had decreased abundance. Five methyltransferase transcripts associated with polyamine metabolism had decreased abundance of which phosphoethanolamine N-methyltransferase (-5.1-fold) were the most affected. The abundance of the transcript associated with polyamine metabolism, calcium/calmodulin-dependent protein kinase 2, had increased abundance (2.2-fold). Thirteen transcripts involved in oxidative stress and redox metabolism had decreased abundances (Table 4.4). Transcripts involved in folate and pyrimidine biosynthesis also had decreased abundances. Interestingly, the transcript of AdoMetDC was not differentially affected by inhibition with MDL73811, in contrast to the co-inhibition of AdoMetDC/ODC which resulted in 2-fold decreased abundance of the transcript (van Brummelen et al., 2009).

Table 4.4: Biological functions of some of the differentially regulated transcripts for AdoMetDC inhibited parasites according to their GO annotations.

PlasmoDB ID	Name	FC UTt_3:Tt_3	Min exp time HPI	Max exp time HPI
	Polyamine and Methionine metabolism			
PF10_0121	Hypoxanthine phosphoribosyltransferase	-1.7	4	26
PF10_0289	Adenosine deaminase, putative	-3.1	42	27
PFD0285c	Lysine decarboxylase, putative	2.5	1	17
PFE0660c	Purine nucleotide phosphorylase, putative	-3.0	4	22
PFE1050w	Adenosylhomocysteinase	-2.1	1	30
PFI1090w	S-adenosylmethionine synthetase	-2.3	12	32
PFL1475w	Sun-family protein, putative	-1.8	36	21
	Methyltransferases			
MAL13P1.214	Phosphoethanolamine N-methyltransferase	-5.1	10	33
PF14_0309	Protein-L-isoaspartate O-methyltransferase beta-aspartate methyltransferase, putative	-3.9	14	31
PF14_0526	Conserved Plasmodium protein, unknown function	-3.1	10	25
PF13_0016	Methyl transferase-like protein, putative	-1.9	1	26
PFL1775c	S-adenosyl-methyltransferase, putative	-1.7	8	27
	Potential polyamine associated effects			
PFL1885c	Calcium/calmodulin-dependent protein kinase 2	2.2	24	43
	Oxidative stress and redox metabolism			
PF08_0071	Fe-superoxide dismutase	-2.0	1	28
PF08_0131	1-cys peroxiredoxin	-2.7	8	26
PF14_0187	Glutathione S-transferase	-1.8	37	21
PF14_0192	Glutathione reductase	-2.2	8	22
PF14_0545	Thioredoxin, putative	-3.1		
PFL0595c	Glutathione peroxidase	-2.2	47	35
PF13_0353	NADH-cytochrome B5 reductase, putative	-2.1	14	33
PF14_0248	Ubiquinol-cytochrome c reductase hinge protein, putative	-1.8	14	37
PF14_0597	Cytochrome c1 precursor, putative	-3.2	11	35
PFI1170c	Thioredoxin reductase	-1.9	10	31
PFI1250w	Thioredoxin-like protein 2	-1.7		
PFL1550w	Lipoamide dehydrogenase	-2.3	15	38
PF11_0352	Protein disulfide isomerase	-1.8	12	32

	Folate and Pyrimidine metabolism			
PFD0830w	Bifunctional dihydrofolate reductase-thymidylate synthase	-4.9	47	35
PFL1720w	Serine hydroxymethyltransferase	-4.8	48	35
PF13_0140	Dihydrofolate synthase/folylpolyglutamate synthase	-1.8	10	31
PF13_0349	Nucleoside diphosphate kinase b, putative	-3.5	11	28
PF10_0154	Ribonucleotide reductase small subunit, putative	-5.4	11	32
PF14_0053	Ribonucleotide reductase small subunit	-3.9	10	32
PF14_0352	Ribonucleoside-diphosphate reductase, large subunit	-2.1	10	32
PFA0555c	UMP-CMP kinase, putative	-2.9		
MAL13P1.218	UDP-N-acetylglucosamine pyrophosphorylase, putative	-1.7	4	26
	Glycolysis			
PF10_0155	Enolase	-2.7	38	16
PF13_0141	L-lactate dehydrogenase	-1.9	42	26
PF14_0378	Triosephosphate isomerase	-1.7	1	26
PFF1300w	Pyruvate kinase	-1.7	41	26
	DNA replication			
PF11_0061	Histone H4	-3.4	14	37
PF11_0062	Histone H2B	-2.9	42	27
PF11_0117	Replication factor C subunit 5, putative	-1.8	10	28
PF11_0282	Deoxyuridine 5'-triphosphate nucleotidohydrolase, putative	-6.3	12	28
PF13_0095	DNA replication licensing factor MCM4-related	-3.1	12	37
PF13_0149	Chromatin assembly factor 1 subunit, putative	-2.9	11	35
PF13_0291	Replication licensing factor, putative	-2.5	11	32
PF14_0177	DNA replication licensing factor MCM2	-2.0	11	32
PF14_0254	DNA mismatch repair protein Msh2p, putative	-1.9	11	30
PFB0840w	Replication factor C, subunit 2	-3.4	13	30
PFD0685c	Chromosome associated protein, putative	-2.0	10	28
PFE0270c	DNA repair protein, putative	-3.5	12	31
PFE0450w	Chromosome condensation protein, putative	-2.6	10	30
PFE0675c	Deoxyribodipyrimidine photolyase, putative	-2.8	9	28
PFF0510w	Histone H3	-4.3		
PFF0865w	Histone H3	-1.8		
PFF1225c	DNA polymerase 1, putative	-2.0	10	30
PF14_0314	Chromatin assembly factor 1 P55 subunit, putative	2.0	24	40
PFI0235w	Replication factor A-related protein, putative	-2.1	8	32
	Transcription factors			
PF11_0241	Myb-like DNA-binding domain, putative	1.7	32	1
PFL0465c	Zinc finger transcription factor (krox1)	1.8		
PF14_0374	CCAAT-binding transcription factor, putative	1.7	20	43
	Translation			
PF14_0289	Mitochondrial ribosomal protein L17-2 precursor, putative	-3.6	10	26
PF14_0606	Mitochondrial ribosomal protein S6-2 precursor, putative	-2.2	15	45
PF14_0709	Mitochondrial ribosomal protein L20 precursor, putative	-2.1	36	44
PFB0645c	Mitochondrial large ribosomal subunit, putative	-2.3	11	24
PFC0675c	Mitochondrial ribosomal protein L29/L47 precursor, putative	-1.9	41	46
PFC0701w	Mitochondrial ribosomal protein L27 precursor, putative	-2.5	12	1
PFD0675w	Apicoplast ribosomal protein L10 precursor, putative	-2.9	11	35
PFI0890c	Organelle ribosomal protein L3 precursor, putative	-2.2	10	35
PFI1240c	Prolyl-t-RNA synthase, putative	-2.8	11	31
PFI1575c	Peptide release factor, putative	-2.7	10	28
	Cell cycle and cytokinesis			
PF13_0328	Proliferating cell nuclear antigen	-5.7	11	28
PFL1330c	Cyclin-related protein, Pfcyc-2	-2.7	14	36
PF11_0478	Kinesin-like protein, putative	2.1	21	35
PFE0165w	Actin-depolymerizing factor, putative	-2.2	14	36
PFI0180w	Alpha tubulin	-7.3	11	35

PFI1565w	Profilin, putative	-3.0	15	38
PFL0925w	Formin 2, putative	2.0		
PFL2215w	Actin I	-2.5	15	39
	Signal transduction			
MAL13P1.19	Peptidase, putative	-2.4	10	28
MAL13P1.205	Rab11b, GTPase	-1.9	48	35
PFE0690c	PfRab1a	-1.8	10	28
PFI0215c	Signal peptidase, putative	-1.8	10	33
PFI1005w	ADP-ribosylation factor-like protein	-2.0	21	41

4.3.7 Hierarchical clustering of the AdoMetDC inhibited transcripts

Hierarchical clustering of the AdoMetDC-inhibited transcripts was done using Gene Cluster to cluster the transcripts of UTt_3:Tt_3. Hierarchical clustering is able to join transcript data and cluster the data in groups that can be easily visualised. It compares the expression profile of each transcript and then form groups which represent transcripts that have similar expression profiles. This process of comparing all the transcripts within the dataset will continue until only one large cluster is present. Transcripts within close distance on the dendogram have similar expression profiles, while transcripts with larger distances between them on the dendrogram are less similar. Transcripts that are within a specific single cluster are then assumed to be co-regulated and functionally related to each other. A specific tight cluster (correlation of 0.949) containing 4 polyamine-related transcripts was revealed (Figure 4.13), adenosine deaminase (PF10_0289), adenosylhomocysteinase (PFE1050w), S-adenosylmethionine synthetase (PFI1090w) and phosphoethanolamine N-methyltransfease (MAL13P1.214). All of these transcripts increased in abundance from UTt_1 to UTt_3 under unchallenged conditions, but after inhibition of AdoMetDC these transcripts revealed decreased abundances in Tt_3 compared to UTt_3. Due to the lack of significant regulation of the transcript of AdoMetDC (PF10_0322), its transcript clustered separately from the other polyamine related transcripts and is therefore not shown in Figure 4.13.

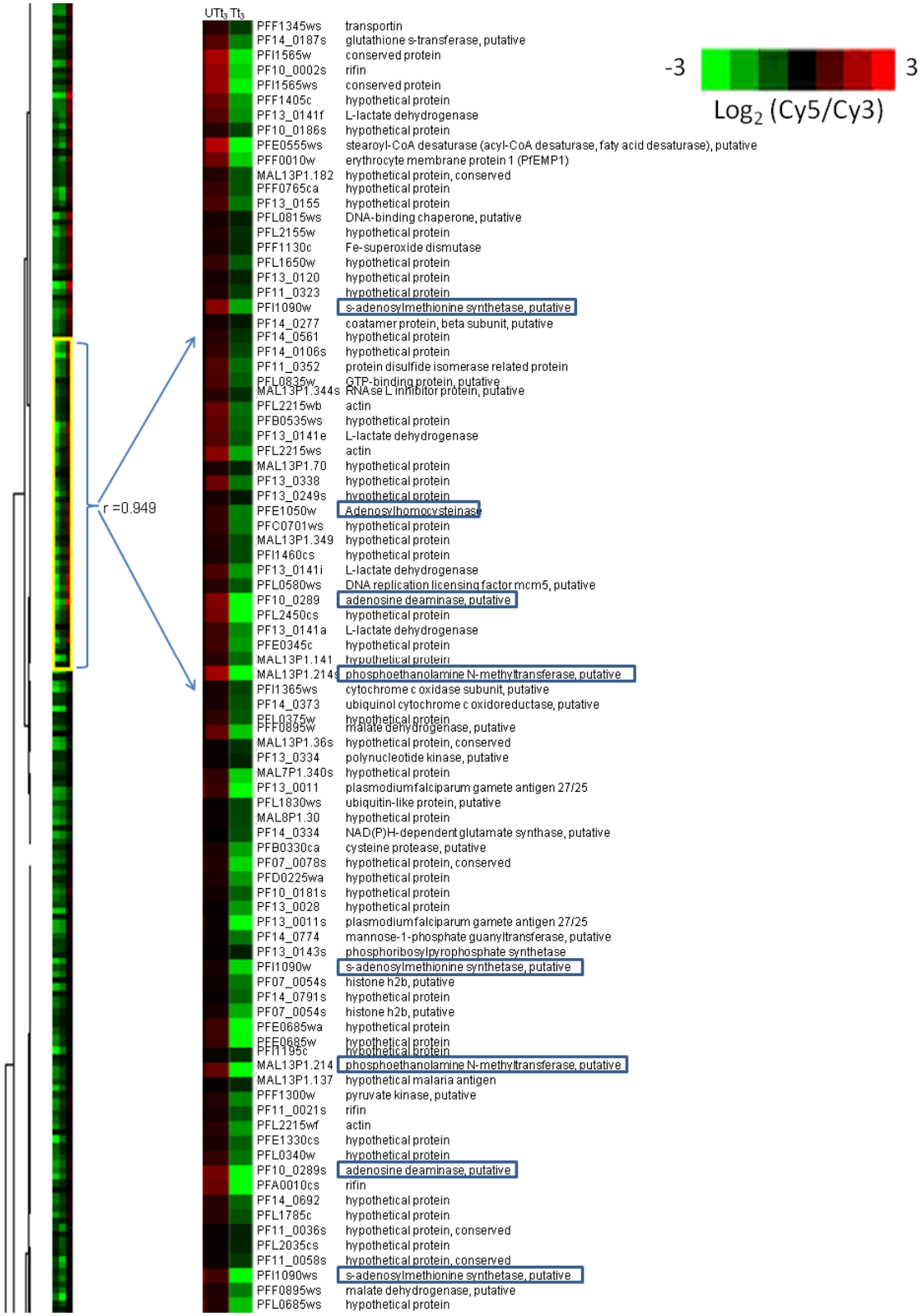

Figure 4.13: A tight cluster (r =0.949) containing polyamine-related transcripts.

A correlation coefficient of 0.949 was produced by hierarchical clustering of AdoMetDC-inhibited parasites and contained 8 polyamine-related transcripts that related to 4 unique transcript groups. These transcripts are blocked on the picture in blue. This cluster contains 91 transcripts that clustered together out of the total of 9966 transcripts that were detected on the slides.

Hierarchical clustering was performed on the polyamine-specific differentially regulated transcripts that were identified with the inhibition of AdoMetDC as well as some oxidative stress-related transcripts that were identified in Table 4.4 (Figure 4.14). Three tight clusters were identified from these transcripts [1-3]. The first cluster ([1] correlation of 0.88) resulted in clustering of mostly the methyltransferase-related transcripts. These transcripts have low expression in UTt_1, which then increased in abundance over time in the untreated samples (UTt_3). With the inhibition of AdoMetDC the transcripts from this cluster [1] remained low in abundances (Tt_3). The second [2] and third group [3] contained the majority of the polyamine and methionine-related transcripts. The second cluster ([2] correlation of 0.83) contained transcripts that had slightly increased expression in UTt_1, which increased even further in abundances at the untreated time points (UTt_3). Similar to the first cluster the treated sample (Tt_3) were indicative of transcriptional arrest since the transcripts in Tt_3 had lower abundances. The last cluster ([3] correlation of 0.88) contained the transcripts with low abundances in both Utt_1 and UTt_3, but increased abundances in Tt_3.

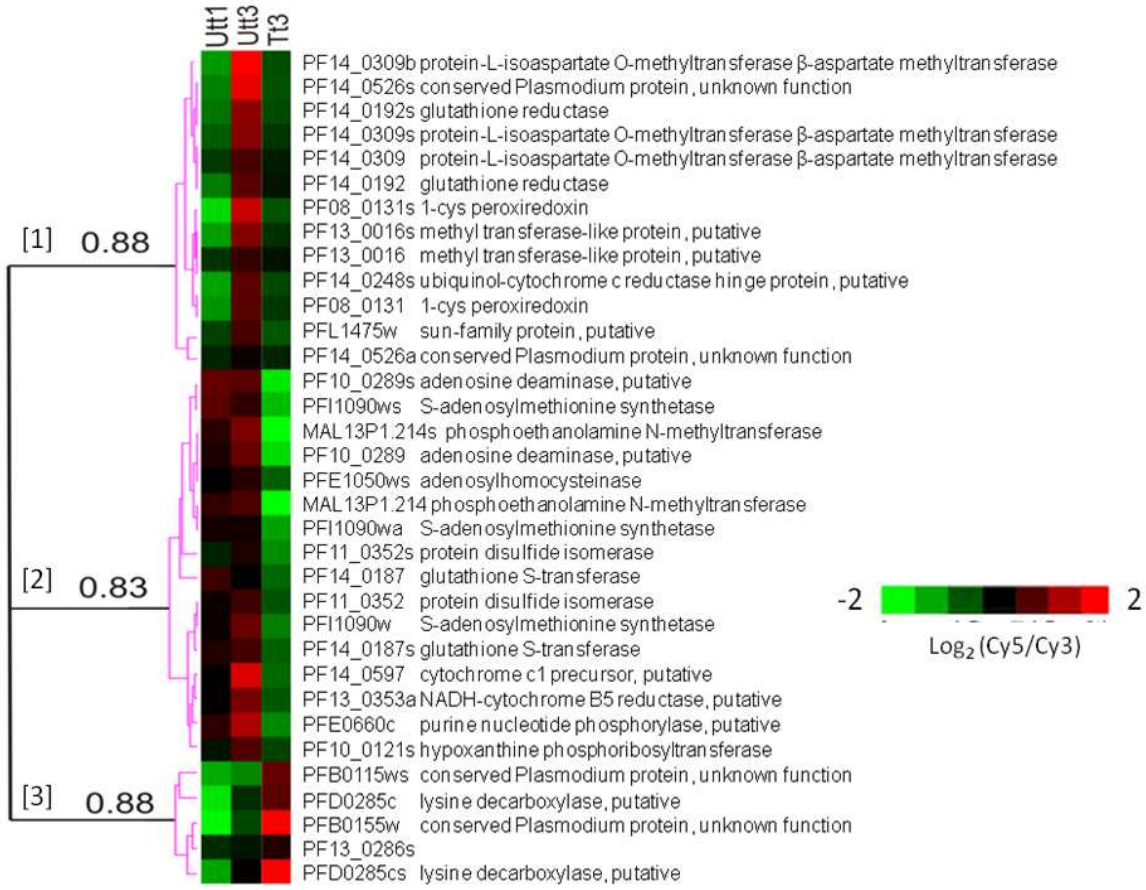

Figure 4.14: Hierarchical clustering of polyamine-specific and oxidative stress transcripts.
Three tight clusters exist for the polyamine and oxidative stress related transcripts.

4.3.8 Transcript regulation of polyamine-specific transcripts followed over all 3 time points

The differential regulation of polyamine-specific transcripts were analysed over the 3 time points by comparison of their FC over time (Figure 4.15). The fold change values obtained for UTt_1:Tt_1 and UTt_2:Tt_2 was not significantly differentially affected, but these fold changes could still be used to determine the trend of transcript abundances over time (Figure 4.15). At Tt_1 the majority of the transcripts are completely unaffected by the inhibition of AdoMetDC with MDL73811. At Tt_2 the majority of the transcripts were slightly affected by either a small increase or decrease in transcript abundance. In Tt_3 the differential regulation of the transcripts were more pronounced. Phosphoethanolamine N-methyltransferase (MAL13P1.214), methyl transferase-like protein (PF13_0016), hypoxanthine phosphoribosyltransferase (PF10_0121), S-adenosylmethionine synthetase (PFI1090w), sun-family protein (PFL1475w), S-adenosyl-methyltransferase (PFL1775c) all had decreased transcript abundance at Tt_1 which decreased further over time, and are therefore some of the transcripts that seem to be more severely affected by AdoMetDC inhibition. Adenosine deaminase (PF10_0289), protein-L-isoaspartate O-methyltransferase beta-aspartate methyltransferase (PF14_0309), adenosylhomocysteinase (PFE1050w), purine nucleotide phosphorylase (PFE0660c), conserved *Plasmodium* protein (PF14_0526) all had a positive fold change at Tt_1 which then gradually decreased over time and are therefore affected over time. Lysine decarboxylase (PFD0285c) and calcium/calmodulin-dependent protein kinase 2 (PFL1885c) transcript abundances both increased over time with AdoMetDC inhibition.

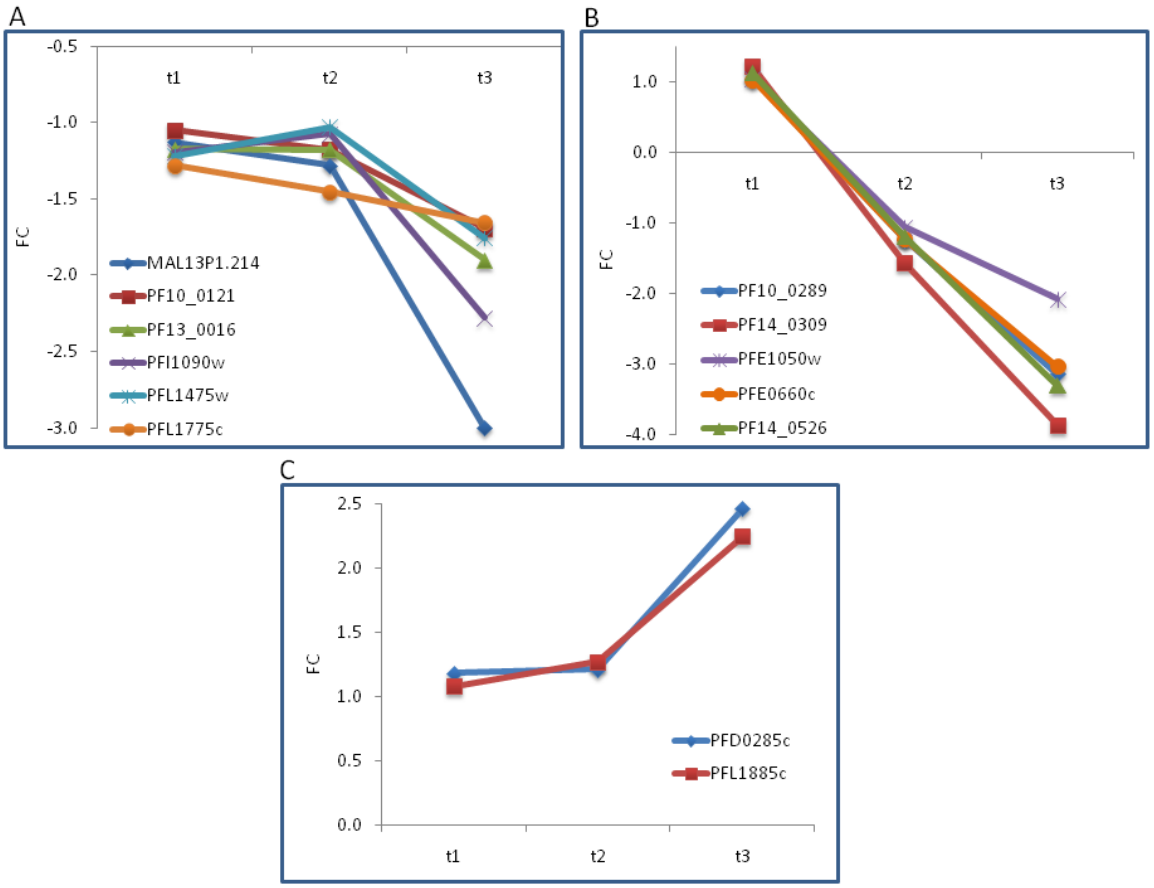

Figure 4.15: Fold change of polyamine-specific transcripts over the 3 time points.

(A)Include transcripts that have decreased transcript abundance which is the even further decreased over the 3 time points and include: MAL13P1.214: phosphoethanolamine N-methyltransferase, PF10_0121: hypoxanthine phosphoribosyltransferase, PF13_0016: methyl transferase-like protein, PFI1090w: S-adenosylmethionine synthetase, PFL1475w: sun-family protein, PFL1775c: S-adenosyl-methyltransferase, (B) Include transcripts which is high in abundance initially in Tt_1 but then decreases over time and include: PF10_0289: adenosine deaminase, PF14_0309: protein-L-isoaspartate O-methyltransferase beta-aspartate methyltransferase, PFE1050w: adenosylhomocysteinase, PFE0660c: purine nucleotide phosphorylase, PF14_0526: conserved *Plasmodium* protein, (C) Include transcripts that increase in abundances: PFD0285c: lysine decarboxylase, PFL1885c: calcium/calmodulin-dependent protein kinase 2.

4.3.9 Identification of uniquely affected Plasmodial pathways as a result of AdoMetDC inhibition

All 549 differentially expressed transcripts identified in the study were subjected to metabolic pathway identification in MADIBA (Law *et al.*, 2008). P-values for each of these pathways containing unique enzymes were calculated according to Fishers test (Fisher, 1935), therefore $p<0.05$ is representative of significance within the results. A unique enzyme according to MADIBA is defined as an enzyme that can only be classified into one specific pathway and not multiple pathways. A total of 49 pathways containing unique enzymes were detected, of which 27 pathways had only 1 unique enzyme, 8 pathways had 2 unique enzymes, 4 pathways had 3 unique enzymes, 5 pathways had 4 unique enzymes, and 5 pathways had 5 or more unique enzymes per pathway. Only

1 pathway, methionine and polyamine metabolism had a significant p-value of 0.0183 with 7 unique enzymes (Table 4.5 and Figure 4.16).

Table 4.5: Unique metabolic pathway identification of the data from the inhibition of AdoMetDC.

Pathway	p-value[a]	Nr of unique enzymes found[b]
Glycine, serine and threonine metabolism	0.9990	3
Methionine metabolism	0.9575	3
Selenoamino acid metabolism	0.7400	3
Pantothenate and CoA biosynthesis	0.8573	3
Citrate cycle (TCA cycle)	0.5503	4
Pyruvate metabolism	0.9982	4
Oxidative phosphorylation	0.2286	4
Glycerophospholipid metabolism	0.9827	4
Glutathione metabolism	0.6851	4
Glycolysis / Gluconeogenesis	0.8263	5
Lysine degradation	0.9600	5
Pyrimidine metabolism	0.9095	7
Purine metabolism	0.9929	9
Methionine and Polyamine Metabolic Pathway	0.0183*	7

[a]p-value is calculated by MADIBA according to the Fisher test and considered as significant if $p < 0.05$. [b]Unique enzymes determined for a specific pathway. * The only pathway with $p < 0.05$ that were considered as significant.

MADIBA identified methionine and polyamine metabolism as a pathway that was significantly affected with AdoMetDC inhibition. The differentially regulated transcripts associated with methionine and polyamine metabolism are given in Figure 4.16. The transcripts with decreased transcript abundance are indicated in green, while increased transcripts are indicated in red. Transcripts that are both up-stream and down-stream of AdoMetDC were affected including transcripts in the methionine cycle that had decreased transcript levels.

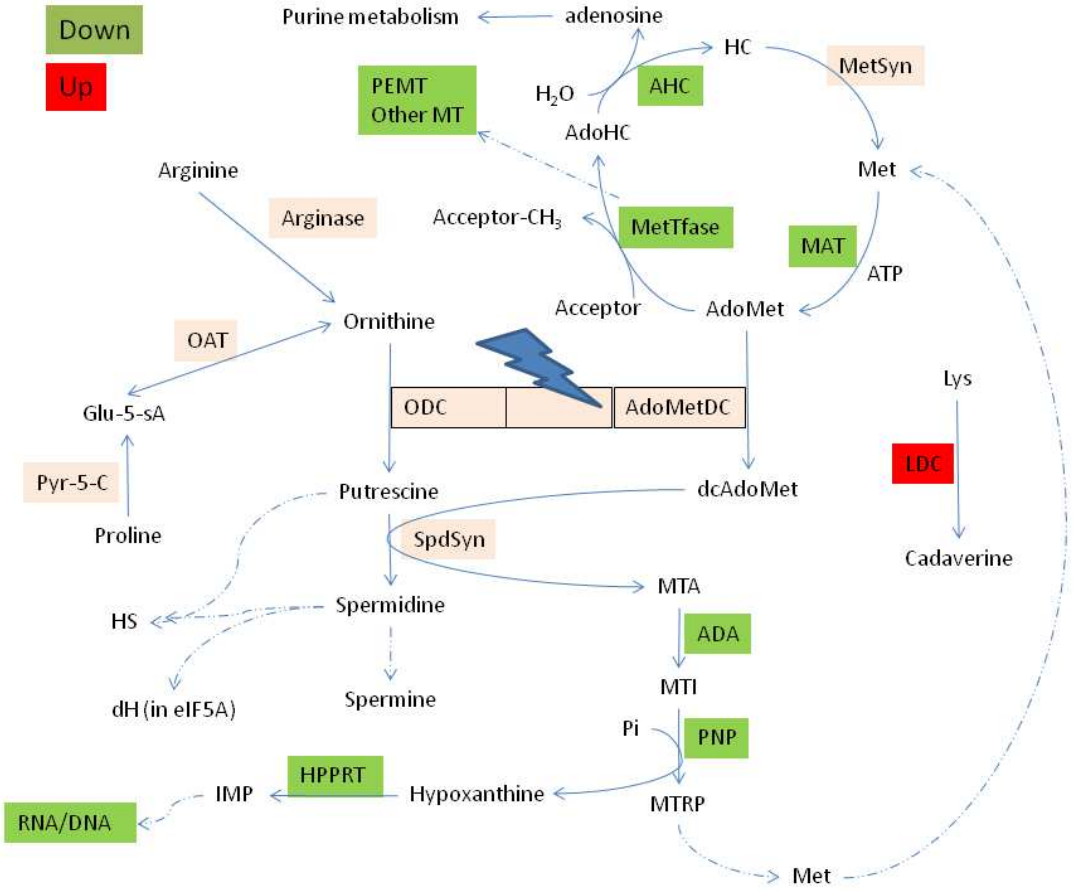

Figure 4.16: Polyamine and methionine metabolism affected by AdoMetDC inhibition.
Green is indicative of transcripts that have decreased abundance, while red is indicative of increased abundance of the transcripts. All other transcripts not affected by AdoMetDC inhibition within this pathway is marked in pink.

4.3.10 Interactions of the AdoMetDC inhibited transcriptome

The *P. falciparum* interactome was constructed *in silico* using Bayesian frameworks (Date & Stoeckert, 2006, Wuchty *et al.*, 2009). Submission of the AdoMetDC inhibited dataset to MADIBA indicated methionine and polyamine metabolism as a pathway that was significantly affected by AdoMetDC inhibition (Table 4.5). Due to the fact that the target was known within this study and to further iterate the enrichment for polyamine metabolism with AdoMetDC inhibition, the AdoMetDC inhibited dataset was investigated for possible AdoMetDC interacting partners (Van Brummelen, 2009). The interactome for the AdoMetDC inhibited dataset was determined *in silico* by comparison of the interactome database (PlasmoMAP)(Date & Stoeckert, 2006) and the AdoMetDC inhibited dataset. A total of 147 potential interacting partners for AdoMetDC were determined of which 41 (28%) were present in the AdoMetDC inhibited dataset (Table 4.6). The top 20 interacting partners for AdoMetDC included 11 transcripts (55%) from the AdoMetDC inhibited dataset of which 1-cys peroxiredoxin was one. Three other transcripts involved in

oxidative stress, glutathione reductase (PF14_0192), disulfide isomerase precursor, putative (MAL8P1.17), and ferredoxin (MAL13P1.95) were also present within the interactome of AdoMetDC, therefore establishing a possible link between AdoMetDC inhibition and oxidative stress. The complete list of interacting partners for AdoMetDC is given in Appendix C.

To determine if the 55% obtained for the top 20 interacting partners for the AdoMetDC inhibited transcriptome dataset was random, the interactome data from another unrelated bifunctional enzyme, dihydroopteroate synthase/hydroxymethylpterin pyrophosphokinase (DHPS/HPPK) was also compared to the data from the AdoMetDC inhibited transcriptome dataset (Appendix C). Unlike the interacting partners of AdoMetDC, only 19 out of a total of 164 (12%) possible interacting partners of DHPS/HPPK were present within the AdoMetDC inhibited transcriptome dataset. Interestingly, the DHPS/HPPK interactome only included hypoxanthine phosphoribosyltransferase (PF10_0121) that was also present within the AdoMetDC inhibited transcriptome dataset and is a polyamine related transcript. Therefore, these results indicate that the AdoMetDC inhibited transcriptome dataset is specific to proteins that may interact with AdoMetDC (Table 4.6 and Appendix C).

According to STRING 8.2 analysis functional binding partners of AdoMetDC included SpdS (PF11_0301), AdoMet synthase (PFI1090w) and a putative modification methylase-like protein (MAL7P1.151). Therefore these 3 functional binding partners of AdoMetDC were also subjected to *in silico* interactome analysis (PlasmoMAP) to determine if the high percentage of interacting partners that was determined for AdoMetDC was specific to AdoMetDC or if it is a polyamine-related process. Comparison of the AdoMetDC inhibited transcriptome dataset with the interactome for putative modification methylase-like protein (MAL7P1.151) revealed a total of 325 interacting partners of which only 20 (6%) were present within the AdoMetDC inhibited transcriptome dataset. Similarly, the comparison of the interactome data for SpdS and the AdoMetDC inhibited transcriptome dataset revealed that only 9 out of 95 (9%) interacting partners could be identified in the AdoMetDC inhibited transcriptome dataset. However, comparison of the interacting partners of AdoMet synthase to the AdoMetDC inhibited transcriptome dataset revealed that 84 out of 257 (33%) interacting partners were present within the AdoMetDC inhibited transcriptome dataset (Table 4.6 and Appendix C). The interactome of AdoMet synthase contained 6 oxidative stress transcripts that included NADH-cytochrome b5 reductase (PF13_0353), putative protein disulfide isomerase related protein (PF11_0352), putative disulfide isomerase precursor (MAL8P1.17), thioredoxin reductase (PFI1170c), glutathione peroxidase (PFL0595c), putative thioredoxin-related protein (PF13_0272). Although the interactome (PlasmoMAP) is an *in silico* datbase and needs

experimental verification the results may reveal a possible link between AdoMetDC and AdoMet synthase. The complete set of interactome binding partners for AdoMetDC, DHPS/HPPK and AdoMet synthase is given in Appendix C.

Table 4.6: The top 20 interacting partners for AdoMetDC, and AdoMet synthase.

Nr	PlasmoDB ID	Name	Score[x]	Present in diff affected transcriptome
		AdoMetDC[a]		
1	PF11_0317	Structural maintenance of chromosome protein, putative	9.53	
2	PFE0195w	P-type ATPase, putative	8.31	
3	PFA0390w	DNA repair exonuclease, putative	7.98	
4	MAL8P1.99	Hypothetical protein	6.62	Yes
5	PF11_0427	Dolichyl-phosphate b-D-mannosyltransferase, putative	6.62	
6	PF07_0129	ATP-dept. acyl-coa synthetase	6.62	Yes
7	PFA0590w	ABC transporter, putative	6.62	Yes
8	PF10_0260	Hypothetical protein	5.90	
9	PF13_0348	PfRhop148,Rhoptry protein	5.90	Yes
10	PF14_0053	Ribonucleotide reductase small subunit	5.70	Yes
11	PFD0685c	Chromosome associated protein, putative	4.71	Yes
12	PFC0125w	ABC transporter, putative	4.71	Yes
13	PF14_0709	Ribosomal protein L20, putative	4.71	Yes
14	PF08_0131	1-cys peroxidoxin	4.71	Yes
15	PF11_0117	□eplication factor C subunit 5, putative	4.71	Yes
16	PF11_0181	Tyrosine --tRNA ligase, putative	4.71	Yes
17	PFB0180w	5'-3' exonuclease, N-terminal resolvase-like domain, putative	4.71	
18	PFL2180w	50S ribosomal protein L3, putative	4.71	
19	PF14_0097	Cytidine diphosphate-diacylglycerol synthase	4.71	
20	PF14_0081	DNA repair helicase, putative	4.71	
		AdoMet synthase[b]		
1	PFE1345c	Minichromosome maintenance protein 3, putative	11.69	Yes
2	PFB0895c	Replication factor C subunit 1, putative	11.69	Yes
3	PFL0835w	GTP-binding protein, putative	8.31	
4	PFI1575c	Peptide release factor, putative	8.31	Yes
5	PF13_0095	DNA replication licensing factor mcm4-related	8.31	Yes
6	PF14_0177	DNA replication licensing factor MCM2	8.31	Yes
7	PFB0795w	ATP synthase F1, alpha subunit, putative	8.31	
8	PFE0450w	Chromosome condensation protein, putative	7.98	Yes
9	PFD0420c	Flap exonuclease, putative	7.98	
10	MAL13P1.96	Chromosome segregation protein, putative	7.98	
11	PFD0590c	DNA polymerase alpha	7.98	Yes
12	PFC0745c	Proteasome component C8, putative	6.62	
13	PF13_0061	ATP synthase gamma chain, mitochondrial precursor, putative	6.62	
14	PF07_0023	DNA replication licensing factor mcm7 homologue, putative	6.62	Yes
15	MAL8P1.128	Proteasome subunit alpha, putative	6.62	
16	PF13_0353	NADH-cytochrome b5 reductase, putative	5.96	Yes
17	MAL8P1.101	Hypothetical protein	5.96	Yes
18	PF14_0063	ATP-dependent Clp protease, putative	5.96	
19	PFI0240c	E1-E2_ATPase/hydrolase, putative	5.96	Yes
20	PF11_0249	Hypothetical protein	5.96	

[a]AdoMetDC top 20 interacting partners resulted in 11/20 hits (55%). [b]AdoMet synthase top 20 interacting partners resulted in 11/20 hits (55%). [x]Probability score predicted by the interactome database. Interactome data obtained from PlasmoMAP (http://www.cbil.upenn.edu/cgi-bin/plasmoMAP).

4.3.11 Comparison of AdoMetDC inhibited transcriptome dataset to the transcriptomes of inhibited AdoMetDC/ODC and inhibited spermidine synthase

Comparison of the AdoMetDC inhibited transcriptome dataset (549 transcripts) to the co-inhibition of AdoMetDC/ODC (dataset of 538 transcripts), revealed that 154 transcripts were shared between these two datasets (Figure 4.17). Of these 154 transcripts, 21% (34/154) had increased transcript abundance and 79% (122/154) had decreased transcript abundance with regard to the AdoMetDC inhibited transcriptome dataset. Comparison of the AdoMetDC inhibited transcriptome dataset to the differentially affected transcriptome of inhibited SpdS (dataset of 708 transcripts) revealed that 194 transcripts were shared between these 2 perturbation studies. Within these 194 shared transcripts, 76% (148/194) had decreased transcript abundance while 24% (46/194) had increased transcript abundance.

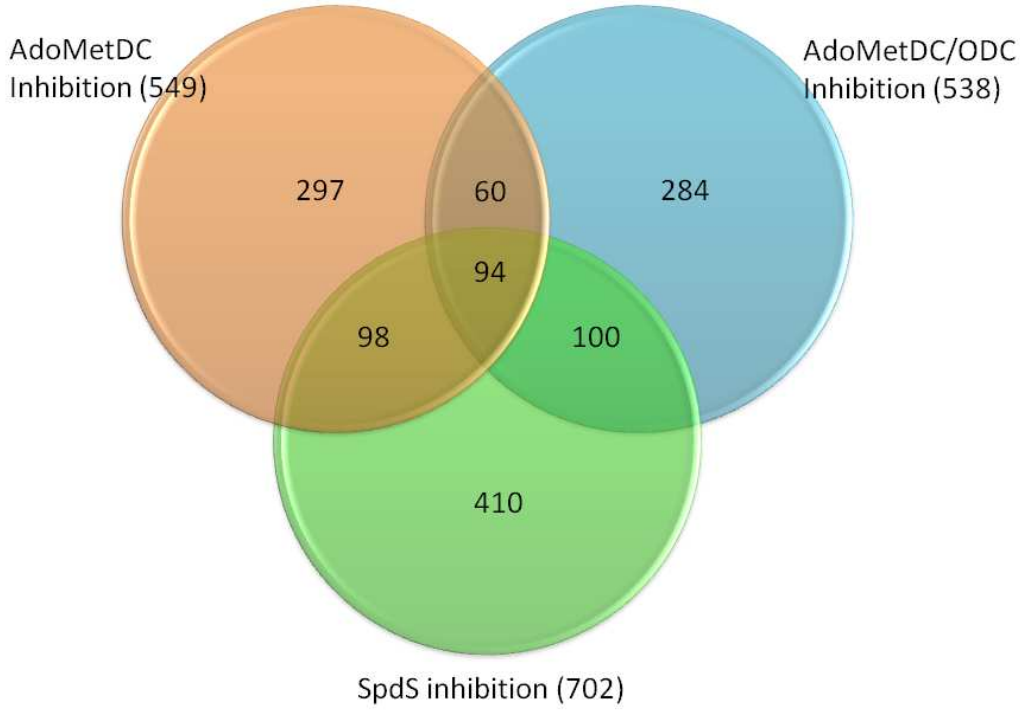

Figure 4.17: Correlation between transcript data from the AdoMetDC inhibited transcriptome dataset, co-inhibition of AdoMetDC/ODC and SpdS inhibition.

The total number of differentially affected transcripts for each of the 3 different studies is given in brackets next to the study name. In total, 94 transcripts were shared between all 3 studies.

In total, 94 transcripts were shared between all 3 polyamine-perturbation studies mentioned (Table 4.7 and Appendix D for full list of shared transcripts). Submission of these 94 shared transcripts to MADIBA did not reveal any metabolic pathway to be significantly affected according to Fishers

test. Of the transcripts that were shared between the AdoMetDC inhibited transcriptome dataset and at least one of the other polyamine-perturbation studies, 6 polyamine-related transcripts were similarly affected in all 3 polyamine-perturbation studies (Table 4.7). Adenosine deaminase (PF10_0289), purine nucleotide phosphorylase (PFE0660c), lysine decarboxylase-like protein (PFD0670c), phosphoethanolamine N-methyltransferase (MAL13P1.214), and pyridoxal 5'-phosphate synthase (PF14_0570) had decreased transcript abundances in all 3 polyamine-perturbation studies. AdoMet synthetase was only shared between the AdoMetDC inhibited transcriptome dataset and the co-inhibition of AdoMetDC/ODC, and was not regulated in the SpdS inhibited transcriptome. Only 2 of the polyamine-related transcripts that were shared in all 3 polyamine-perturbation studies had increased transcript abundances and included LDC (PFD0285c) and calcium/calmodulin-dependent protein kinase 2 (PFL1885c). Various oxidative stress-related transcripts were shared between all 3 polyamine-perturbation studies and all of these transcripts had decreased abundance. Similar results were obtained for folate metabolism, cell cycle regulation, transcription factors and the majority of transporters shared between all 3 polyamine-perturbation studies.

Table 4.7: Shared transcripts from the AdoMetDC inhibited transcriptome dataset, the co-inhibited AdoMetDC/ODC dataset and the inhibition of SpdS.

PlasmoDB ID	Name	Fold change		
		AdoMetDC	AO	SpdS
	Polyamine and Methionine metabolism			
PF10_0289	Adenosine deaminase, putative	-3.1	-2.4	-2.3
PFD0285c	Lysine decarboxylase, putative	2.5	2.8	2.4
PFE0660c	Purine nucleotide phosphorylase, putative	-3.0	-2.7	-3.6
PFE1050w	Adenosylhomocysteinase	-2.1	-1.5	
PFI1090w	S-adenosylmethionine synthetase	-2.3	-1.5	
PFD0670c	Lysine decarboxylase-like protein, putative	-2.0	-1.6	-3.1
	Methyltransferases			
MAL13P1.214	Phosphoethanolamine N-methyltransferase	-5.1	-2.7	-3.4
PF14_0309	Protein-L-isoaspartate O-methyltransferase beta-aspartate methyltransferase, putative	-3.9	-1.8	
PF14_0526	Conserved Plasmodium protein, unknown function	-3.1	-2.1	-2.0
	Potential polyamine associated effects			
PFL1885c	Calcium/calmodulin-dependent protein kinase 2	2.2	2.3	2.4
PF14_0570	Pyridoxal 5'-phosphate synthase, putative	-2.3	-2.2	-2.5
PF14_0200	Pantothenate kinase, putative	-1.7		-2.2
	Oxidative stress and redox metabolism			
PF08_0131	1-cys peroxiredoxin	-2.7	-2.8	-4.5
PF14_0187	Glutathione S-transferase	-1.8	-1.5	-2.1
PF14_0192	Glutathione reductase	-2.2	-1.7	-2.6
PF13_0353	NADH-cytochrome B5 reductase, putative	-2.1		-2.4
PFI1170c	Thioredoxin reductase	-1.9		-3.1
	Folate and Pyrimidine metabolism			
PFD0830w	Bifunctional dihydrofolate reductase-thymidylate synthase	-4.9	-1.6	-2.2
PF13_0349	Nucleoside diphosphate kinase b, putative	-3.5	-1.9	-4.4

PF10_0154	Ribonucleotide reductase small subunit, putative	-5.4	-1.5	-4.2
PF14_0053	Ribonucleotide reductase small subunit	-3.9	-1.4	-5.0
	Glycolysis			
PF13_0141	L-lactate dehydrogenase	-1.9	-1.5	-2.4
	DNA replication			
PF11_0117	Replication factor C subunit 5, putative	-1.8	-1.8	-4.5
PF11_0282	Deoxyuridine 5'-triphosphate nucleotidohydrolase, putative	-6.3	-2.9	-2.8
PF13_0291	Replication licensing factor, putative	-2.5	-1.2	-3.5
PF14_0254	DNA mismatch repair protein Msh2p, putative	-1.9	-1.4	-2.0
PFD0685c	Chromosome associated protein, putative	-2.0	-2.0	-2.0
PFE0675c	Deoxyribodipyrimidine photolyase, putative	-2.8	-1.5	-4.3
PFI0235w	Replication factor A-related protein, putative	-2.1	-1.8	-4.0
	Transcription factors			
PF11_0241	Myb-like DNA-binding domain, putative	1.7	1.8	2.1
PFL0465c	Zinc finger transcription factor (krox1)	1.8	2.0	
PF14_0374	CCAAT-binding transcription factor, putative	1.7	2.1	
PFL1900w	Transcription factor with AP2 domain(s), putative	-2.7	-1.4	-4.9
PF11_0404	Transcription factor with AP2 domain(s), putative	1.9		2.0
PF13_0097	Transcription factor with AP2 domain(s), putative	1.7		2.3
PFI1665w	Transcription factor with AP2 domain(s), putative	-1.9		-2.2
	Translation			
PF14_0709	Mitochondrial ribosomal protein L20 precursor, putative	-2.1	-1.9	-2.4
PFI0890c	Organelle ribosomal protein L3 precursor, putative	-2.2	-1.6	-2.7
	Cell cycle and cytokinesis			
PF13_0328	Proliferating cell nuclear antigen	-5.7	-1.9	-4.2
PFE0165w	Actin-depolymerizing factor, putative	-2.2	-2.0	-2.5
PFI0180w	α-tubulin	-7.3	-1.5	-4.5
PFI1565w	Profilin, putative	-3.0	-2.0	-2.2
	Transporters			
PFC0125w	ABC transporter, (TAP family), putative	-1.9	-1.9	-2.1
MAL13P1.23	CorA-like Mg^{2+} transporter protein, putative	1.8		2.0
PF14_0211	Ctr copper transporter domain containing protein, putative	-2.3		-3.4
PFI0240c	Cu^{2+} -transporting ATPase,	-1.9		-4.4
MAL8P1.32	Nucleoside transporter, putative	-2.8	-1.5	-3.0
PF14_0662	Nucleoside transporter, putative	1.8	1.8	
PFA0245w	Transporter, putative	-3.8	-1.3	
PFE0410w	Triose phosphate transporter	-1.7	-1.9	-2.5
PF07_0065	Zinc transporter, putative	-4.8	-2.2	-3.4

4.3.12 Comparison of AdoMetDC inhibited transcriptome dataset to other *P. falciparum* perturbation data

The 549 differentially affected transcripts from the AdoMetDC inhibited transcriptome dataset was compared to the transcriptomes of 5 other perturbation studies which included CQ inhibition (Gunasekera *et al.*, 2007, Gunasekera *et al.*, 2003), febrile temperature perturbation (Oakley *et al.*, 2007), artesunuate inhibition (Natalang *et al.*, 2008), anti-folate inhibition (Ganesan *et al.*, 2008) and the effect of 20 individual compounds on the transcriptome (Hu *et al.*, 2010). Microarray analysis of artesunate inhibition resulted in the differential regulation of 398 transcripts (Natalang et al., 2008). Of these, only 62 transcripts were shared between artesunate inhibition (398 transcripts)

and the AdoMetDC inhibited transcriptome dataset (549 transcripts). Anti-folate inhibition resulted in only 54 differentially regulated transcripts (Ganesan *et al.*, 2008) of which 9 were shared with the AdoMetDC inhibition transcriptome dataset. CQ inhibition resulted in 601 differentially affected transcripts (Gunasekera et al., 2007) of which 65 were shared with the AdoMetDC inhibited transcriptome dataset. The effect of febrile temperature perturbation on parasites resulted in 336 differentially affected transcripts (Oakley et al., 2007) of which 66 were shared with the AdoMetDC inhibited transcriptome dataset (Figure 4.18).

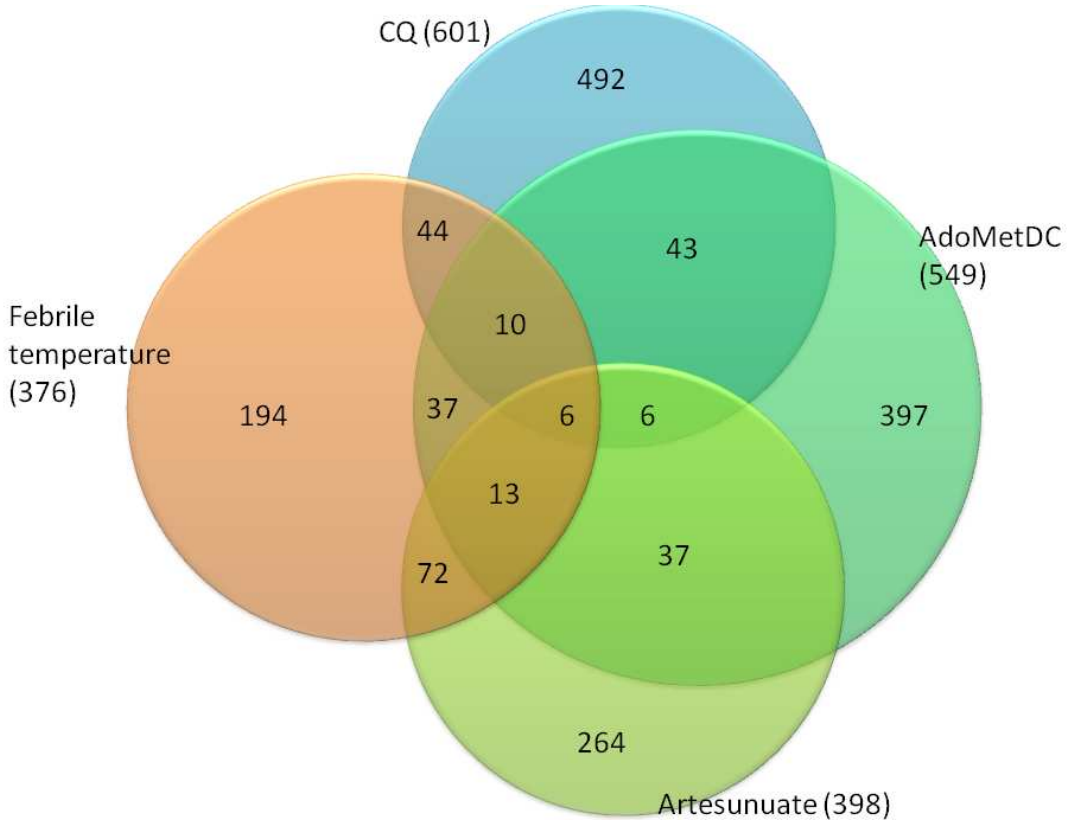

Figure 4.18: Comparisons between the differentially affected transcriptomes of the AdoMetDC inhibited transcriptome dataset, febrile temperature perturbation, CQ inhibition and artesunuate inhibition.

The total number of differentially affected transcripts for each of the 4 different studies is given in brackets next to the study name. In total, 6 transcripts were shared between all 4 studies. 6 transcripts were shared between the AdoMetDC inhibited transcriptome dataset, artesunuate inhibition, and CQ inhibition. 10 transcripts were shared between the AdoMetDC inhibited transcriptome dataset, febrile temperature perturbation and CQ inhibition. 13 transcripts were shared between the AdoMetDC inhibited transcriptome dataset, artesunuate inhibition, and febrile temperature perturbation. 37 transcripts were shared between the AdoMetDC inhibited transcriptome dataset and febrile temperature perturbation. 37 transcripts were shared between the AdoMetDC inhibited transcriptome dataset and artesunuate inhibition. 43 transcripts were shared between the AdoMetDC inhibited transcriptome dataset and CQ inhibition.

Parasites inhibited by 20 individual compounds in the schizont-stage resulted in the differential transcript regulation of 3125 transcripts (Hu, 2010) of which 430 were shared with the AdoMetDC inhibited transcriptome dataset. In total 466 transcripts from the AdoMetDC inhibited transcriptome

dataset were shared with at least one of the studies mentioned. Comparison of all the data from the above mentioned studies resulted in the identification of 5 transcripts shared between the AdoMetDC inhibited transcriptome dataset, artesunate inhibition, CQ inhibition, febrile temperature perturbation, and the 20 individual compounds data (Table 4.8). These 5 transcripts are therefore indicative of a general stress response by the parasite regardless of the perturbation. Of these 5 transcripts, only 4 were shared with the co-inhibition of AdoMetDC/ODC and inhibition of SpdS. The only transcript that was not shared within these studies was protein disulfide isomerase (PF11_0352).

Table 4.8: Five transcripts shared between all of the perturbation studies.

PlasmoID	Name	A	20 Comp	Temp	Artes	CQ
PF08_0060	Asparagine-rich antigen	2.2	n/a	2.1	2.4	n/a
PF11_0352	Protein disulfide isomerase	-1.8	n/a	-2.7	-1.9	n/a
PF14_0631	Conserved Plasmodium protein, unknown function	1.7	n/a	3.2	2.8	n/a
PF14_0758	Plasmodium exported protein (hyp17), unknown function	1.7	n/a	5.2	2.2	n/a
PFC0085c	Plasmodium exported protein, unknown function	1.8	n/a	4.3	2.7	n/a

A: AdoMetDC inhibition transcriptome dataset, 20 Comp: is the 20 individual compounds dataset, Temp: is the febrile temperature-perturbation dataset, Artes: Artesunuate inhibited dataset, CQ: CQ inhibited dataset. n/a is not available since the dataset provided only transcripts and not the fold changes.

Eighty three transcripts were identified that were unique to the AdoMetDC inhibited transcriptome dataset and not shared with any other published transcriptome analysis of *P. falciparum* after any other perturbation (Appendix E). These transcripts were sorted according to their GO terms (Table 4.9). Of the 83 unique transcripts, 32% (27/83) had increased transcript abundance and 68% (57/83) had decreased transcript abundance. Transcripts associated with DNA metabolism with increased transcript abundance included putative DNA topoisomerase VI B subunit (MAL13P1.328; 2.8-fold). Another transcript that was increased was kinesin-like protein (PF11_0478; 2.1-fold) and is associated with the cytoskeleton of the parasite. Various translation associated transcripts were identified that were all decreased, similarly for RNA metabolism and signal transduction. Interestingly, 3 polyamine-associated transcripts were unique and included phosphoetanolamine N-methyltransferase (MAL14P1.214; -5.1-fold), methyl transferase-like protein (PF13_0016; -1.9-fold) and a conserved protein (PF14_0526; -3.1-fold). Pyridoxal 5'-phosphate synthase (PF14_0570; -2.3-fold), associated with Vitamin B synthesis was also unique to the AdoMetDC inhibited transcriptome dataset and had decreased transcript abundance. Various exported proteins were also identified as unique for AdoMetDC perturbation and were mostly increased in transcript abundance.

Table 4.9: Unique transcripts associated with AdoMetDC perturbation.

PlasmoDB ID	Product Description	FC[a]
	DNA metabolism	
MAL13P1.328	DNA topoisomerase VI, B subunit, putative	2.8
PF14_0053	Ribonucleotide reductase small subunit	-3.9
PFL1180w	Chromatin assembly protein (ASF1), putative	-2.2
	Proteolysis	
PF13_0084	Ubiquitin-like protein, putative	1.7
PF14_0348	ATP-dependent Clp protease proteolytic subunit, putative	-2.0
	Translation	
PFC0675c	Mitochondrial ribosomal protein L29/L47 precursor, putative	-1.9
PFF0495w	Mitochondrial ribosomal protein L19 precursor, putative	-2.0
PF11_0113	Mitochondrial ribosomal protein L11 precursor, putative	-2.0
PFC0701w	Mitochondrial ribosomal protein L27 precursor, putative	-2.5
PFD0675w	Apicoplast ribosomal protein L10 precursor, putative	-2.9
	Phosphorylation	
PFC0485w	Protein kinase, putative	-1.7
PFF0260w	Serine/threonine protein kinase, Pfnek-5	-1.7
	Polyamine methionine metabolism	
PF13_0016	Methyl transferase-like protein, putative	-1.9
PF14_0526	Conserved Plasmodium protein, unknown function	-3.1
MAL13P1.214	Phosphoethanolamine N-methyltransferase	-5.1
	Primary metabolism	
PFI0960w	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase, putative	-1.7
MAL13P1.220	Lipoate synthase, putative	-1.7
PFB0505c	3-oxoacyl-(acyl carrier protein) synthase III, putative	-2.1
	Cytoskeleton organization and biogenesis	
PF11_0478	Kinesin-like protein, putative	2.1
	RNA metabolic process	
MAL8P1.72	High mobility group protein	-1.7
PF13_0043	CCAAT-binding transcription factor, putative	-1.8
PFD0750w	Nuclear cap-binding protein, putative	-1.8
PF10_0313	Mitochondrial preribosomal assembly protein rimM precursor, putative	-1.9
	Signal transduction	
PF14_0317	Microsomal signal peptidase protein, putative	-1.7
PFI1005w	ADP-ribosylation factor-like protein	-2.0
	Coenzyme metabolic process	
MAL7P1.130	3-demethylubiquinone-9 3-methyltransferase, putative	-1.7
PF14_0570	Pyridoxal 5'-phosphate synthase, putative	-2.3
	Host parasite	
PFF0020c	Erythrocyte membrane protein 1 (PfEMP1)-like protein	1.7
PF07_0138	Rifin	-2.1
MAL7P1.23	RAP protein, putative	-1.7
PF11_0046	CPW-WPC family protein	-1.8
PF14_0297	Apyrase, putative	-2.0
MAL8P1.216	Rifin	-2.4
	Hypotheticals	
PFI1780w	Plasmodium exported protein (PHISTc), unknown function	2.5
PF14_0760	Plasmodium exported protein, unknown function	2.0
PF11_0514	Plasmodium exported protein (PHISTa), unknown function	2.0
PFF1535w	Plasmodium exported protein (hyp5), unknown function	1.9
PFF0075c	Plasmodium exported protein (PHISTb), unknown function	1.8
PFB0970c	Plasmodium exported protein, unknown function	1.7
MAL7P1.230	Hypothetical protein, pseudogene	1.7

Unique differentially expressed transcripts associated with the perturbation of AdoMetDC when compared to the Bozdech data, Artesunate, CQ, antifolate, and febrile temperatures. [a] Fold change of AdoMetDC perturbation.

To determine if the comparisons between the AdoMetDC perturbation and all the other perturbation studies mentioned had any significance, the 83 transcripts unique to the AdoMetDC inhibited transcriptome dataset were analysed with MADIBA to identify significant metabolic pathways. Methionine and polyamine metabolism were once again identified as significant ($p = 0.0638$ with $p<0.01$) for the 83 transcripts unique to the AdoMetDC inhibited transcriptome dataset. This is similar to the complete AdoMetDC inhibited transcriptome dataset in which methionine and polyamine metabolism were identified as the only significant ($p = 0.018$) metabolic pathways affected. Of the 83 unique transcripts that were identified for the AdoMetDC inhibited transcriptome dataset only 9 were shared with the co-inhibition of AdoMetDC/ODC and SpdS inhibition (Table 4.10) and may be unique to parasites in which polyamine metabolism were affected.

Table 4.10: Nine of the unique transcripts only found in polyamine-regulated parasites.

PlasmoID	Name	Fold Change[a]		
		A[b]	AO[c]	SpdS[d]
MAL13P1.214	Phosphoethanolamine N-methyltransferase	-5.1	-2.7	-3.4
MAL7P1.33	Conserved Plasmodium protein, unknown function	-2.6	-1.6	-2.2
MAL7P1.61	Hypothetical protein	1.7	-1.8	2.4
PF14_0053	Ribonucleotide reductase small subunit	-3.9	-1.4	-5.0
PF14_0297	Apyrase, putative	-2.0	-1.5	-2.1
PF14_0526	Conserved Plasmodium protein, unknown function	-3.1	-2.1	-2.0
PF14_0570	Pyridoxal 5'-phosphate synthase, putative	-2.3	-1.7	-2.5
PFB0953w	Plasmodium exported protein (hyp15), unknown function	-1.8	-1.3	-3.9
PFE0685w	Hypothetical protein	-2.6	-1.9	-3.5

[a]Fold change for each of the transcripts in each of the individual studies. [b]A: AdoMetDC inhibited transcriptome dataset. [c]AO: AdoMetDC/ODC co-inhibition. [d]SpdS: spermidine synthase inhibition.

4.3.13 Validation of microarray results with real-time PCR

The MicroArray Quality Control (MAQC) consortium has been established to evaluate the performance of several microarray and qRT-PCR platforms, which is crucial for comparisons between microarray data (Arikawa et al., 2008). Microarray validation consists of a selection of genes on the microarray that is not differentially regulated in order to use these genes as housekeeping genes for normalisation purposes (Abruzzo *et al.*, 2005).

To validate the microarray data from the AdoMetDC inhibited transcriptome dataset, 6 transcripts were selected and used for qRT-PCR. Cyclophillin was used as “housekeeping gene” since its transcript abundance remains relatively unchanged within the AdoMetDC inhibited transcriptome dataset. The qRT-PCR was done over the 3 time points and the fold change calculated to compare to the microarray results. Comparison of the FC for both the microarray Tt_3 and the qRT-PCR Tt_3

revealed similarities in the values obtained. The qRT-PCR data also show that during Tt_1 and Tt_2 the transcripts were not yet affected by the inhibition of AdoMetDC, which is similar to the data obtained from the microarray experiment for Tt_1 and Tt_2, and that the affected transcripts progress over time. SpdS was included in the validation process since it is a polyamine-related transcript, but was not differentially affected by AdoMetDC perturbation (Table 4.11).

Table 4.11: Comparison of microarray data with real-time PCR data.

Name	qRT-PCR FC			Micro-array FC
	Tt_1	Tt_2	Tt_3	Tt_3
PEMT	1.1	-1.7	-3.9	-5.0
AHC	1.0	-1.4	-1.9	-2.1
SpdS	1.1	1.1	1.2	nd
AdoMet synthase	-1.1	-1.8	-1.5	-2.3
LDC	1.2	1.1	2.4	2.5
HH4	-1.0	-1.3	-2.6	-3.4

PEMT: phosphoethanolamine N-methyltransferase, AHC: adenosylhomocysteinase, SpdS: spermidine synthase, AdoMet synthase: S-adenosylmethionine synthestase, LDC: lysine decarboxylase, HH4: histone H4. Nd is representative of a transcript not detected as regulated in the microarray data.

4.4 Discussion

Three time points were selected for RNA extraction for microarray analysis, in an attempt to span the maximum life stages in which the transcript for *Pf(adometdc/odc)* is available. According to the IDC the transcript of *Pf(adometdc/odc)* is produced from 12 to 36 HPI with maximum transcript expression of *Pf(adometdc/odc)* at 24 HPI (Bozdech *et al.*, 2003). The morphology study conducted in Chapter 3 also determined that morphological arrest occurs at 30 HPI and therefore earlier time points were needed before visible morphological arrest of the parasites occur.

Microarrays enable the analysis of thousands of genes on a single slide, and offer the promise of a wealth of information on the transcriptomic state of an organism at any particular moment in time. The use of the Plasmodial Agilent platform enables the simultaneous analyses of 8 samples on a single slide therefore reducing labour intensive hours and improving reproducibility. Overall, the use of the Agilent arrays resulted in better quality microarray data, and confidence in analysis (Hester *et al.*, 2009). This increased spot quality and the subsequent increased quality in the microarray data was also seen in our lab when the Agilent hybridised spots were compared to previously used in-house spotted Plasmodial arrays. The A+T-richness of the Plasmodial genome negates the use of amplification methods for microarrays (Bozdech *et al.*, 2003), therefore the 3-fold reduction in sample size needed per array on the Agilent platform is highly advantageous. The transcriptomic investigation of the AdoMetDC inhibited transcriptome employed a reference design (Figure 4.1), which enabled valid comparisons of data. The reference is a representative sample of equal amounts of each of the treated and untreated samples used at the 3 different time points. Synchronised parasite cultures were used during the transcriptomic investigation in order to determine drug-specific responses of the parasite and not life cycle-related responses.

Determination of the Pearson correlations at the 3 time points investigated for the AdoMetDC inhibited transcriptome revealed that Tt_1 and Tt_2 had only a few differentially expressed transcripts. The 2 early time points were harvested before any observable morphological difference between the treated and untreated parasite (Chapter 3). The later time point (Tt_3: 26 HPI) which correlates with the predicted maximum transcript expression of AdoMetDC (Bozdech *et al.*, 2003) did result in the identification of differentially affected transcripts as a result of AdoMetDC inhibition. The Pearson correlations indicated that the 3 time points used within the AdoMetDC inhibited transcriptome was early enough to enable direct comparisons between treated and untreated parasites and deemed the use of a t_0 strategy unnecessary. The relative t_0 strategy determines the point of transcriptional arrest within the parasite and then uses this point as reference for comparisons made between treated and untreated parasites. This strategy compensates for life and stage specific responses, and ensures that

only drug-specific responses are detected (Van Brummelen, 2009). The co-inhibition of AdoMetDC/ODC used 3 time points taken at 19 HPI, 27 HPI and 34 HPI. This is later than the time points used within this study for the AdoMetDC inhibited transcriptome, of which all 3 time points were in the early life cycle stages. Pearson correlations between all the time points confirmed the validity of directly comparing UTt_3 with Tt_3 for the AdoMetDC inhibited transcriptome. Further analyses of the data for UTt_3:Tt_3 subsequently revealed the differential regulation of 549 transcripts with AdoMetDC inhibition. MDL73811 is an irreversible inhibitor of AdoMetDC, and it is assumed to inhibit its effects from 24 HPI when the transcript for *Pf(adometdc/odc)* is maximally expressed at 24 HPI. Hierarchical clustering of all 3 time points of the AdoMetDC inhibited transcriptome confirmed that transcriptional arrest occurs at Tt_3 and that only a few differentially affected transcripts are present in Tt_1 and Tt_2. Analysis of the data with LIMMA-GUI identified 549 differentially expressed transcripts of which 143 transcripts (24%) had an increase in abundance and 406 transcripts (74%) had a decrease in abundance. Transcript differential expression levels ranged from 7-fold decreased to 4-fold increased transcript abundance, which is in agreement with most other microarray perturbation studies on *P. falciparum*

Methionine and polyamine metabolism were the only significantly affected pathways with 7 unique transcripts present within the AdoMetDC inhibited transcriptome dataset when analysed with MADIBA. This result clearly indicates the specificity of the AdoMetDC inhibited transcriptome dataset for polyamine-related responses of the parasite under AdoMetDC inhibition. All the transcripts associated with polyamine biosynthesis had decreased transcript abundances with the exception of lysine decarboxylase (PFD0285c) that had an increased transcript abundance (2.5-fold). Five methyltransferase transcripts associated with polyamine metabolism were also decreased in transcript abundance.

A comparison was made between the microarray data for the AdoMetDC inhibited transcriptome data (this study), AdoMetDC/ODC co-inhibition (Van Brummelen, 2009) and the inhibition of SpdS (Becker *et al.*, 2010). Comparison of these transcriptomic datasets revealed 154 transcripts that were shared between the AdoMetDC inhibited transcriptome dataset and the AdoMetDC/ODC co-inhibition study and 194 transcripts that were shared between the AdoMetDC inhibited transcriptome dataset and inhibition of SpdS. Ninety-four transcripts were shared between the 3 polyamine-affected transcriptomic studies. Transcripts with decreased abundance in these transcriptomic studies included adenosine deaminase, purine nucleotide phosphorylase, lysine decarboxylase-like protein, phosphoethanolamine N-methyltransferase and pyridoxal 5'-phosphate synthase. These transcripts may therefore be considered as signature transcripts for polyamine-

affected parasites. LDC and calcium/calmodulin-dependent protein kinase 2 had increased transcript abundance in all 3 polyamine-affected transcriptomic studies, suggesting that these 2 transcripts may be involved in compensatory mechanisms of the parasite to cope with polyamine-depletion.

The transcript of AdoMetDC/ODC was not differentially affected by inhibition with MDL73811, which corresponds to the cyclohexylamide perturbation of SpdS that also resulted in the transcript of AdoMetDC/ODC remaining unchanged. This is different to the co-inhibition data in which the transcript abundance of AdoMetDC/ODC decreased 2-fold (van Brummelen *et al.*, 2009) as well as the mono-functional inhibition of ODC, which also resulted in the transcript of AdoMetDC/ODC to decrease in abundance (K. Clark, unpublished data). The inhibition of AdoMetDC and SpdS would result in increased putrescine levels and the transcript abundance of AdoMetDC/ODC is maintained. With the mono-functional inhibition of ODC and the co-inhibition of AdoMetDC/ODC putrescine levels decrease, resulting in the abundance of the AdoMetDC/ODC transcript being decreased (Van Brummelen, 2009). It appears that putrescine may act as a transcriptional stabiliser for the transcript of AdoMetDC/ODC. Evidence for the regulation of *Pf*ODC activity by putrescine does exist (Wrenger *et al.*, 2001), but to date there is no evidence as to the transcriptional regulation of *Pf*AdoMetDC/ODC transcript by putrescine. Spermidine synthase was not differentially affected in any one of the 3 polyamine-affected transcriptomic studies. Therefore the transcript abundance of SpdS is not affected by the absence of any one of the polyamines. The regulatory role of SpdS within the polyamine pathway therefore needs further investigation.

Another measure of the specificity of the AdoMetDC inhibited transcriptome dataset was obtained by the comparison of the differentially affected transcripts from the AdoMetDC inhibited transcriptome, with those of the other Plasmodial perturbation studies. In total 466 transcripts from the AdoMetDC inhibited transcriptome dataset were shared with at least one of the studies mentioned. Only 5 transcripts were shared between all the different perturbation studies and are therefore indicative of general stress responses elicited by the parasites under exposure to unwanted external stimuli.

Comparison of the AdoMetDC data to the total of 4513 differentially expressed transcripts from the inhibition studies mentioned above revealed a total of 83 transcripts unique to only AdoMetDC inhibition (Table 4.9). Three transcripts were unique to polyamine metabolism and were all decreased in abundance, identifying polyamine and methionine metabolism as uniquely-affected pathways for the 83 unique transcripts submitted to MADIBA. The importance of these 83 unique

transcripts provides a specific transcriptomic signature profile for the AdoMetDC inhibited transcriptome. These specific transcriptomic signature profiles therefore can be utilised to determine the mode-of-action of unknown compounds, which has already been successfully applied in antimicrobial (Pietiainen *et al.*, 2009) and tuberculosis transcriptomic studies (Boshoff *et al.*, 2004).

The inhibition of AdoMetDC with MDL73811 resulted in the decreased transcript abundance of adenosine deaminase (PF10_0289), HPPRT (PF10_0121) and PNP (PFE0660c), and decreased dcAdoMet metabolite levels and conversely also in decreased spermidine and MTA levels and subsequently MTI. The absence of MTA as result of AdoMetDC-inhibition may result in the decreased transcript abundances observed for adenosine deaminase, PNP and HPPRT. These results indicate a co-dependency between polyamine metabolism and down-stream MTA and MTI metabolism.

The transcript of AdoMet synthase was decreased with the inhibition of AdoMetDC. AHC also had decreased transcript abundance and together with AdoMet synthase play an essential role in regulation of the methionine levels. AdoMet synthase produces AdoMet from methionine and is therefore an important link between polyamine biosynthesis and methionine metabolism. In Trypanosomes treated with MDL73811 there is a significant increase in AdoMet levels, which ultimately results in hypermethylation and parasite death (Goldberg *et al.*, 1997a). Trypanosomal AdoMet synthase is not inhibited by its own product (Goldberg *et al.*, 2000, Yarlett *et al.*, 1993). Similar to Trypanosomes, AdoMet synthase in Plasmodial parasites is also not feedback regulated by its product, AdoMet (Muller *et al.*, 2008), which may also lead to the expectation of increased AdoMet levels and consequent hypermethylation within the Plasmodial parasite as a mode-of-action for MDL73811. Methylation is dependent on the AdoMet:AdoHcy levels. An increase in AdoMet will result in hypermethylation, while an increase in AdoHcy will result in hypomethylation. However, the effect of the decreased transcript abundance observed for both AdoMet synthase and AHC may restore the balance of AdoMet and AdoHcy within methionine metabolism (Van Brummelen, 2009), but it does seems that in *P. falciparum* polyamine metabolism and methionine metabolism is closely linked. The effect of the decreased transcript abundance of both AdoMet synthase and AHC on the methylation status of Plasmodial parasites is unclear. Closely linked to the methionine metabolism and methylation cycle is methyltransferases. The majority of methyltransferases had decreased transcript abundances. In particular, PEMT had decreased transcript abundance and is uniquely-affected by AdoMetDC inhibition, the co-inhibition of AdoMetDC and inhibition of SpdS, but was not affected with the mono-functional inhibition of ODC or any other perturbation. Therefore, in the absence of spermidine and spermine, PEMT is

affected. These possible methylation regulatory mechanisms will be investigated further in Chapter 5.

Polyamine-related transcripts that remained unaffected by the inhibition of AdoMetDC include arginase and OAT. The transcript regulation of OAT could not be determined with microarray due to saturation of the spot on the microarray. Subsequent RT-PCR revealed that the transcript of OAT remained unchanged in abundance with AdoMetDC inhibition. Ornithine levels are also homeostatically regulated through the action of arginase (Olszewski *et al.*, 2009). Therefore, ornithine homeostasis is independent of polyamine metabolism. It should however be noted that it did seem like the protein abundance of OAT increased although this could not be measured in Tt2 of the proteomic study due to the spot for OAT and AdoMet synthase that overlapped and were saturated (Chapter 3).

Upon inhibition of AdoMetDC with MDL73811, 9 folate and pyrimidine metabolism transcripts displayed a decrease in transcript abundance. Folate and methionine metabolism is linked by the recycling of N5-methyl THF to methionine (Bistulfi *et al.*, 2009). In human prostate and colon cell lines, folate and polyamine pathways are connected since N-5,10-methylene THF can be converted to N-5-methyl THF, which plays a role in methionine metabolism in the conversion of homocysteine back into methionine (Bistulfi *et al.*, 2009). Folate depletion in prostate cells resulted in an imbalance of AdoMet levels due to the link between folate and methionine metabolism (Bistulfi *et al.*, 2009). Given the evidence in prostate cells, and the fact that 9 folate-related transcripts were decreased with AdoMetDC inhibition it is postulated that a link does exist between polyamine metabolism and folate metabolism within the Plasmodial parasite. This will be investigated further in Chapter 5.

Thirteen transcripts associated with oxidative stress and redox status of the parasite all had a decrease in transcript abundance (Appendix B). Some of these transcripts include 1-cys peroxiredoxin (PF08_0131), glutathione S-transferase (PF14_0187), glutathione reductase (PF14_0192), thioredoxin (PF14_0545) and thioredoxin reductase (PFI1170c). The parasite is exposed to a constant risk of oxidative stress since it resides in a pro-oxidant environment. The constant exposure to oxygen and iron from hemoglobin within the erythrocyte enables the formation of reactive oxygen species (ROS) via the Fenton reaction. It is therefore essential that the parasite needs an efficient anti-oxidant system to be able to deal with all these threats (Muller, 2004). The role of polyamines to protect against ROS is different to that of glutathionine probably as a result of the close association of polyamines with DNA, lipids and proteins (Rider *et al.*, 2007).

Polyamines are able to provide protection by several mechanisms that include possible direct scavenging of ROS, induction of DNA conformational changes and primary protection of DNA by close association with DNA (Rider *et al.*, 2007). The relationship between polyamines and oxidative stress presents two possible scenarios. The first is that polyamine depletion reduces anti-oxidant capacity thus promoting oxidative stress (Assimakopoulos *et al.*, 2010). The second scenario is that ROS act as messengers that regulate expression of enzymes implicated in polyamine synthesis (Assimakopoulos *et al.*, 2010). *E. coli* and yeast strains with reduced polyamines have increased sensitivity towards oxidative damage (Chattopadhyay *et al.*, 2003, Jung *et al.*, 2003, Chattopadhyay *et al.*, 2006). Various studies have shown that polyamines are able to protect against superoxide (Lovaas & Carlin, 1991), radiation (Chiu & Oleinick, 1997, Chiu & Oleinick, 1998) and fenton radicals (Ha *et al.*, 1998). As such, the high polyamine content of Plasmodial parasites during normal proliferation may therefore provide protection against oxidative stress.

The transcript abundance of thioredoxin (PF14_0545) and thioredoxin reductase (PFI1170c) were decreased. The thioredoxin interactome revealed that OAT, AdoMet synthase and AHC are all interacting partners of thioredoxin (Sturm *et al.*, 2009). Similarly, thioredoxin reductase is also a binding partner of AdoMet synthase (Sturm et al., 2009, Wuchty *et al.*, 2009). This provides a further link between polyamine biosynthesis and oxidative stress. AdoMet synthase, AHC and thioredoxin reductase all had decreased transcripts levels with AdoMetDC inhibition. Therefore, due to the absence or decreased abundance of the binding partners, the decreased transcript abundances may result in the deregulation of the redox status of the parasite. This also establishes a link between AdoMet synthase and consequent polyamine regulation and the redox status of the parasite.

Similarly, decreased spermidine and spermine with increased putrescine levels in rat brain resulted in the production of ROS (Assimakopoulos *et al.*, 2010). Putrescine in *E. coli* protected DNA against oxidative stress and also resulted in increased survival rates of the bacteria, with an increase in ODC and LDC activity to combat the effects of the oxidative stress exposure (Tkachenko & Shumkov, 2004 (b), Tkachenko, 2004 (a)). As such, this provides a direct link between polyamine biosynthesis and oxidative stress.

The transcript abundance of LDC was increased with polyamine depletion in the Plasmodial parasites in all polyamine-affected transcriptome studies. Induced LDC may result in lysine being converted to cadaverine. The functional importance of cadaverine to the survival of Plasmodial parasites in AdoMetDC-inhibited parasites is unclear. However, LDC is able to increase in response

to oxidative stress while its product; cadaverine acts as a radical scavenger of superoxide in *Vibrio vulnificus* (Kim *et al.*, 2006, Kang *et al.*, 2007). In a polyamine-depleted environment, mammalian SpdS is able to utilise cadaverine due to the presence of the diamine group (Pegg *et al.*, 1981). This was similarly shown in *E. coli* cells in which spermidine was subsequently available for eIF5A synthesis (Park *et al.*, 1991). The induction of LDC and consequent increased cadaverine levels have been determined to alleviate AdoMetDC and arginine decarboxylase inhibition in pea seedlings (Icekson *et al.*, 1986) as well as alleviate ODC inhibition in *P. falciparum* (Assaraf *et al.*, 1987). Therefore, the induction of LDC is a possible compensatory mechanism in response to AdoMetDC inhibition.

Three cyclin-associated transcripts with decreased abundance were identified and included putative cyclin related protein (PFL1335w; -2.4-fold), Pfcyc-2 cyclin-related protein (PFL1330c; -2.6-fold) and proliferating cell nuclear antigen (PF13_0328; -5.7-fold). Polyamines play a role during cell cycle progression by being able to degrade cyclin B1 mRNA in the eukaryotic G1-phase, therefore enabling cells to enter the S-phase of the cell cycle (Thomas & Thomas, 2001). Polyamine-depletion can also result in cell cycle arrest due to decreased stabilisation of cyclin D1 (Wallace *et al.*, 2003). Therefore, polyamines and the cell cycle have an involvement in cell proliferation, since cyclin-associated transcripts decreased in abundance upon AdoMetDC inhibition.

Recently, a link was established between polyamines and microtubules (Savarin *et al.*, 2010). Interestingly, various transcripts involved in cytoskeleton organisation and biogenesis was also uniquely affected in the AdoMetDC inhibited transcriptome dataset. The transcript abundance of actin and tubulin were severely decreased with AdoMetDC perturbation. The disruption of tubulin results in G2/M cell cycle arrest and the induction of apoptosis in tumour cells (Chen *et al.*, 2007). It may therefore be assumed that polyamines may play a role in stabilisation of these transcripts. Microtubules in Plasmodial parasites have an essential role in cell division, motility and preserving the structural integrity of the parasite (Naughton *et al.*, 2008). The transcripts of centrin-3 (PF10_0271; -2.8-fold) and centrin-2 (PF14_0443; -4.9-fold) were also decreased with AdoMetDC perturbation. These centrins play a role in the cell cycle and cell proliferation within the malaria parasite. The decreased transcript abundances of the centrins in the erythrocytic stages may result in attenuation of sporozoite stage and erythrocyte stage parasites as well as result in transmission-defiecient Plasmodial strains (Mahajan *et al.*, 2008). Therefore, the decreased transcript abundances of the various cytoskeleton associated transcripts as a result of AdoMetDC inhibition are indicative of cell cycle arrest in the Plasmodial parasite.

In this chapter the transcriptomic response of the Plasmodial parasite was investigated after AdoMetDC inhibition. Evidence was provided that the 549 differentially regulated transcripts provide information on drug-specific responses by the parasite. Polyamine biosynthesis was a pathway uniquely associated with the AdoMetDC inhibited transcriptome dataset, which revealed unique links between polyamine biosynthesis and methionine metabolism. Other interesting consequences of AdoMetDC inhibition included the link between polyamine-regulation (spermidine and spermine depletion) and oxidative stress, folate metabolism, cytoskeleton biogenesis and phosphorylation that may impact on regulation of the *P. falciparum* cell cycle.

Due to the "just-in-time" production of transcripts, the transcripts are only expressed as they are needed (Bozdech *et al.*, 2003). It is also a general notion that the protein levels will mimic transcript levels within the Plasmodial parasite (Daily *et al.*, 2004). In the following chapter the correlation between the transcript and proteins of the AdoMetDC inhibited datasets are compared. Furthermore, it should be noted that the transcript and protein expression levels does not necessarily indicate the activity of that specific enzyme, therefore further biological investigations are presented in Chapter 5.