Deciphering the Plasmodium epitranscriptome as a mechanism of translational control

Sinha, Ameya

2019

Sinha, A. (2019). Deciphering the Plasmodium epitranscriptome as a mechanism of translational control. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/142964>

<https://doi.org/10.32657/10356/142964>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 20:20:52 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

Deciphering the *Plasmodium* epitranscriptome as a mechanism of translational control

AMEYA SINHA

SCHOOL OF BIOLOGICAL SCIENCES

2019

Deciphering the *Plasmodium* epitranscriptome as a mechanism of translational control

Ameya Sinha

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

22nd August, 2019

.....
Date

.....
AMEYA SINHA

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

22nd August, 2019

.....
Date

.....

PETER R. PREISER

Authorship Attribution Statement

This thesis contains material from two papers published in the following peer-reviewed journals in which I am listed as an author.

Chapter 3 is published as Chee Sheng Ng[†], Ameya Sinha[†], Yaw Aniweh, Qianhui Nah, Indrakanti Ramesh Babu, Chen Gu, Yok Hian Chionh, Peter C Dedon & Peter R Preiser. tRNA epitranscriptomics and biased codon are linked to proteome expression in *Plasmodium falciparum*. [†] - Co-first authors [1].

The contributions of the co-authors are as follows:

- Prof Peter Preiser and Prof. Peter Dedon designed the study and provided the initial project direction and edited the manuscript drafts.
- Dr. Chee Sheng Ng and I prepared the manuscript drafts.
- Dr. Chee Sheng Ng and I performed all the laboratory work at Singapore-MIT Alliance for Research and Technology, Singapore.
- Dr Chee Sheng Ng characterized and quantified the tRNA modifications across the lifecycle of the parasite and carried out the proteomics study.
- I carried out the mapping and quantitation of specific tRNA modifications on the wobble position.
- Dr Yaw Aniweh and Dr. Indrakanti Ramesh Babu assisted in the proteomics studies.
- Dr. Chionh Yok Hian, and Dr. Chen Gu assisted in the data collection and provide guidance in the interpretation of the mass spectrometry data.

Chapter 4 is published as Sebastian Baumgarten, Jessica M. Bryant, Ameya Sinha, Thibaud Reyser, Peter R. Preiser, Peter C. Dedon, Artur Scherf, Transcriptome-wide dynamics of extensive m⁶A mRNA methylation during *Plasmodium falciparum* blood-stage development. In press, *Nature Microbiology* (2019) [2].

The contributions of the co-authors are as follows:

- Peter Preiser, Peter Dedon and Artur Scherf designed the study and provided the initial project direction and edited the manuscript drafts
- Sebastian Baumgarten and I carried out the initial mass spectrometry-based investigation of the mRNA modification landscape in the parasite
- I carried out all the mass spectrometry-based quantitation for m⁶A in parasites.
- I carried out the cross-linking, subsequent immunoprecipitation, and mass spectrometry for the METTL3 protein complex
- Sebastian Baumgarten, Thibaud Reyser, and Jessica Bryant were involved in parasite sample preparation and mRNA extraction
- Sebastian Baumgarten carried out all the CLIP experiments, sequencing studies, and sequencing analysis
- Jessica Bryant was involved in the design of the CRISPR interference system in *Plasmodium falciparum*
- Sebastian Baumgarten, Jessica Bryant, and I wrote and revised the manuscript.

22nd August 2019

.....
Date

.....

AMEYA SINHA

Table of Contents

STATEMENT OF ORIGINALITY.....	3
SUPERVISOR DECLARATION STATEMENT.....	4
AUTHORSHIP ATTRIBUTION STATEMENT.....	5
TABLE OF CONTENTS	7
LIST OF FIGURES.....	11
ACKNOWLEDGEMENTS	13
Abstract	16
Preface	17
1. Chapter I: Introduction.....	20
1.1. Malaria: Epidemiology, Etiology, and Pathophysiology	20
1.1.1. Malaria Epidemiology.....	20
1.1.2. The lifecycle of <i>P. falciparum</i>	21
1.1.3. Pathogenesis and Clinical Symptoms.....	23
1.2. Anti-malarial Approaches and Challenges	25
1.2.1. Anti-malarial Drugs.....	25
1.2.2. Vector Control.....	26
1.2.3. Vaccines	26
1.2.4. Treatment Failure and Resistance	27
1.3. Transcriptional and Translational Control in <i>P. falciparum</i>	28
1.3.1. <i>P. falciparum</i> Genome.....	28
1.3.2. Transcriptional and Translational Profiling of the <i>Plasmodium</i> IDC	28
1.3.3. Post transcriptional control during the IDC in <i>Plasmodium</i>	29
1.4. Translational elongation and codon choice.....	30
1.4.1. Translation and codon degeneracy	30
1.4.2. Wobble Hypothesis	31
1.4.3. Codon Choice	33
1.5. Nucleotide Modifications and the Epitranscriptome	35
1.5.1. tRNA Modifications	37
1.5.2. mRNA Modifications	40
1.5.3. RNA Modifications and Stress Response.....	43
1.6. Research drive and Rationale	45
1.6.1. Specific Aims	46
CHAPTER 2.....	48

2.	Materials and Methods	48
2.1.	Culturing <i>Plasmodium falciparum</i> <i>in vitro</i>	48
2.2.	Stress Treatments for IC ₅₀ Assay	48
2.3.	RNA Extraction	49
2.4.	HPLC Purification for tRNA	50
2.5.	Oligo-based tRNA Pulldown	50
2.6.	RNA Digestion	51
2.7.	LC-MS/MS for Quantitation and Characterization of Ribonucleosides	51
2.8.	LC-MS/MS for Mapping tRNA Modifications	52
2.9.	Protein Extraction and Digestion	53
2.10.	Protein Crosslinking and Immunoprecipitation	54
2.11.	LC-MS/MS Analysis for Proteomics	54
2.12.	Data Processing and Statistical Analysis	55

CHAPTER 3 - RESULTS I - TRNA WOBBLE MODIFICATIONS MCM⁵U AND MCM⁵S²U GOVERN THE *PLASMODIUM* ‘TRANSLATOME’

3.	Overview	57
3.1.	Preliminary Data	58
3.1.1.	Limitations and Challenges	61
3.2.	Experimental Strategy	63
3.3.	Isolation of tRNA from <i>P. falciparum</i>	64
3.4.	Approach I - Oligo-based Pulldown for Isolation of Specific tRNA Isoacceptors	67
3.5.	Approach II - RNA Fingerprinting	69
3.5.1.	Computational Approaches	69
3.5.2.	Optimization of Liquid-Chromatography Mass Spectrometry (LC-MS)	70
3.5.3.	tRNA Hydrolysis and Desalting	74
3.5.4.	Mapping of Wobble Modifications	74
3.5.5.	CID fragmentation	77
3.6.	Quantification of Mapped Modifications Across the Lifecycle	82
3.7.	Discussion	83

CHAPTER 4 – RESULTS II - TRANSCRIPTOME-WIDE DYNAMICS OF EXTENSIVE M⁶A MRNA METHYLATION DURING *PLASMODIUM FALCIPARUM* BLOOD-STAGE DEVELOPMENT

4.	Overview	93
-----------	-----------------	-----------

4.1. Methodology Overview	95
4.1.1. mRNA isolation.....	96
4.1.2. Survey of mRNA modifications in <i>Plasmodium falciparum</i>	98
4.1.3. Complete set of modified ribonucleosides in <i>P. falciparum</i> mRNA.....	100
4.2. Quantitation of mRNA modification across <i>Plasmodium</i> IDC	102
4.3. m⁶A displays dynamism across <i>P. falciparum</i> IDC	106
4.3.1. Structure validation of m ⁶ A using high-resolution mass spectrometry.....	108
4.4. PfMT-A70 co-immunoprecipitation reveals the Plasmodium writer complex.....	109
4.5. m⁶A-seq	112
4.5.1. m ⁶ A-antibody specificity.....	113
4.5.2. m ⁶ A-seq summary results.....	114
4.5.3. CRISPRi knockdown of PfMT- A70	116
4.5.4. m ⁶ A-seq for PfMT-A70 knockdown using CRISPRi	118
4.6. Discussion.....	119
 CHAPTER 5 – A ROLE FOR TRNA MODIFICATIONS FOR MODULATING STRESS RESPONSE IN <i>P. FALCIPARUM</i> VIA TRANSLATIONAL CONTROL MECHANISMS.	126
5. Overview	126
5.1. Growth Inhibition Dose Response Studies.....	128
5.1.1. Flow Cytometry for Determination of Parasitemia.....	129
5.1.2. IC ₅₀ growth inhibition assays	131
5.2. tRNA Modification Profiling.....	134
5.3. A preliminary proteomics perspective – DHA and CQ Stress Response.....	137
5.3.1. Data summary.....	138
5.3.1. Proteomics data analysis and pathway enrichment	141
5.4. Patterns of synonymous codon usage in stress-dependent protein expression.....	147
5.5. Discussion.....	150
 CONCLUSION	157
 FUTURE WORK.....	163
6. Overview	163
6.1. Identify and characterize RNA-modifying enzymes important for tRNA modifications and parasite biology	163
6.2. Determine translation rates of codon biased genes during IDC	166
6.3. Discovery of new modifications in <i>Plasmodium falciparum</i> RNA.....	169
6.4. Translational control mechanisms mediating artemisinin resistance	171
 REFERENCE:	173

APPENDIX	188
7. Appendix	188
7.1. Appendix I.....	188
7.2. Appendix II	190
7.3. Appendix III.....	195

List of Figures

FIGURE 1 COUNTRIES WITH CHLOROQUINE RESISTANCE AND DOCUMENTED FAILURE FOR ARTEMISININ COMBINATION THERAPIES (ACT)	18
FIGURE 2 LIFECYCLE OF <i>P. FALCIPARUM</i>	23
FIGURE 3 SCHEMATIC DEPICTING THE WOBBLE HYPOTHESIS	32
FIGURE 4 CODON UTILIZATION AND DECONDING SCHEME IN <i>PLASMODIUM FALCIPARUM</i>	34
FIGURE 5 CHEMICAL STRUCTURES OF SELECTED RNA MODIFICATIONS.....	36
FIGURE 6 MODIFICATIONS FREQUENTLY FOUND ON tRNAs.....	39
FIGURE 7 CONTROL OF GENE EXPRESSION BY RNA MODIFICATIONS	41
FIGURE 8 PATHOGENS, RNA MODIFICATIONS AND STRESS RESPONSE	44
FIGURE 9 PRELIMINARY WORK ON <i>PLASMODIUM FALCIPARUM</i> tRNA MODIFICATIONS	60
FIGURE 10 LC-MS APPROACHES TO ANALYSING tRNA MODIFICATIONS	62
FIGURE 11 tRNA ISOLATION FROM TOTAL RNA IN <i>PLASMODIUM FALCIPARUM</i>	66
FIGURE 12 OLIGO-BASED PULLDOWN APPROACH TO ISOLATE tRNA ISOACCEPTORS.....	68
FIGURE 13 BIOINFORMATIC ANALYSIS FOR MAPPING tRNA FRAGMENTS	70
FIGURE 14 DETERMINING CHARGE STATE FOR RNA OLIGO STANDARDS	73
FIGURE 15 ELECTROSPRAY SERIES FOR UNMODIFIED AND MODIFIED tRNA SPECIES PRESENTING SUMMARY DATA FOR THE T1 DIGESTED <i>PLASMODIUM</i> tRNA.....	75
FIGURE 16 DECONVOLUTED MASSES OF tRNA FRAGMENTS FOR tRNA ISOACCEPTORS (TOP) tRNA _{GLN} CAA, (MIDDLE) tRNA _{GLY} GGA, AND (BOTTOM) tRNA _{GLU} GAA	76
FIGURE 17 COLLISION-INDUCED DISSOCIATION (CID) FRAGMENTATION OF RNA	79
FIGURE 18 MS ² FRAGMENTATION OF MODIFIED tRNA ISOACCEPTORS	81
FIGURE 19 QUANTIFYING WOBBLE-MODIFIED tRNA ISOACCEPTORS	83
FIGURE 20 tRNA GENE EXPRESSION AS MEASURED BY MICROARRAY STUDIES.....	90
FIGURE 21 ROLES OF tRNA MODIFICATIONS IN ASEXUAL BLOOD STAGE DEVELOPMENT OF <i>PLASMODIUM</i>	92
FIGURE 22 METHODOLOGY FOR DE NOVO DISCOVERY OF mRNA MODIFICATIONS	96
FIGURE 23 BIOANALYSER PLOTS SHOWING THE PURITY OF mRNA ISOLATION	97
FIGURE 24 FRAGMENTATION OF NUCLEOSIDES	99
FIGURE 25 MULTIPLE REACTION MONITORING PARAMETERS OPTIMISED FOR TARGETED DETECTION OF mRNA MODIFICATIONS IN <i>P. FALCIPARUM</i>	100
FIGURE 26 LC-MS/MS EXTRACTED ION CHROMATOGRAMS OF MODIFIED NUCLEOTIDES ANALYSED IN PARASITE mRNA COLLECTED AT 36 HPI.	102
FIGURE 27 QUANTITATION OF mRNA MODIFICATIONS ACROSS THE PARASITE LIFECYCLE	104
FIGURE 28 GLOBAL DYNAMICS OF mRNA MODIFICATIONS OTHER THAN m ⁶ A ACROSS IDC	105
FIGURE 29 EXTERNAL CALIBRATION CURVE FOR STANDARDS	106
FIGURE 30 GLOBAL DYNAMICS OF N ⁶ -METHYLADENOSINE (m ⁶ A) ACROSS PARASITE IDC	107
FIGURE 31 ANALYSIS OF mRNA-DERIVED m ⁶ A BY HIGH- RESOLUTION MASS SPECTROMETRY.	109
FIGURE 32 DECIPHERING THE PfMT-A70 PROTEIN COMPLEX IN <i>PLASMODIUM FALCIPARUM</i>	111
FIGURE 33 SCHEMATIC FOR m ⁶ A-SEQ	112
FIGURE 34 m ⁶ A ANTIBODY SPECIFICITY OF m ⁶ A-SEQ STUDIES AT THREE TIME POINTS SHOWS ENRICHMENT FOR m ⁶ A IN THE IMMUNOPRECIPITATED SAMPLE	113
FIGURE 35 m ⁶ A-SEQ SUMMARY DATA	115
FIGURE 36 CRISPRi KNOCKDOWN OF PfMT-A70.....	117
FIGURE 37 m ⁶ A-SEQ FOR PfMT-A70 KNOCKDOWN PARASITES	118
FIGURE 38 YTH-DOMAIN CONTAINING PROTEINS IN <i>P. FALCIPARUM</i>	123
FIGURE 39 tRNA MODIFICATIONS AND STRESS RESPONSE.....	128
FIGURE 40 FLOWCYTOMETRY GATING STRATEGY FOR PARASITAEMIA DETERMINATION.....	130
FIGURE 41 DOSE-DEPENDENT INHIBITION ASSAYS FOR OXIDATIVE STRESSORS	133
FIGURE 42 STRESS-INDUCED tRNA MODIFICATION CHANGES IN <i>PLASMODIUM FALCIPARUM</i>	136
FIGURE 43 PRE/ POST- NORMALIZATION OF THE RAW INTENSITIES FOR ALL PROTEINS IN EACH QUANTITATIVE SAMPLE FOR EACH REPLICATE FOR DHA (TOP) AND CQ (BOTTOM).	139
FIGURE 44 NUMBER OF PEPTIDES OR PROTEINS FOUND IN THE DIFFERENT FRACTIONS OF THE HIGH pH COLUMN FRACTIONATION FOR DHA TREATMENT	140
FIGURE 45 PRINCIPAL COMPONENT ANALYSIS (PCA) PLOTS FOR GLOBAL PROTEOMIC CHANGES UPON DHA (LEFT) AND CQ (RIGHT) TREATMENT	141
FIGURE 46 UP-REGULATED PROTEINS IN DHA TREATED-SAMPLES.....	143

FIGURE 47 DOWN-REGULATED PROTEINS IN DHA TREATED SAMPLES	144
FIGURE 48 UP-REGULATED PROTEINS IN CQ TREATED SAMPLES	145
FIGURE 49 DOWN-REGULATED PROTEINS IN CQ TREATED SAMPLES	146
FIGURE 50 PLS-R SCORES (LEFT) AND LOADINGS (RIGHT) PLOT FOR DHA (A AND B) AND CQ (C AND D) TREATED SAMPLES.....	149
FIGURE 51 DECIPHERING THE PLASMODIUM EPITRANSCRIPTOME.....	161
FIGURE 52 SCHEMATIC TO IDENTIFY AND CHARACTERIZE PUTATIVE PLASMODIUM RNA-MODIFYING ENZYMES.	165
FIGURE 53 ENRICHED SYNONYMOUS CODONS AND THEIR UNDER-REPRESENTED COUNTERPARTS IN LATE-STAGE (TP 40 AND 46) UP-REGULATED PROTEINS.	167
FIGURE 54 IMPACT OF CODON BIASED GENES ON TRANSLATION	168
FIGURE 55 QUEUOSINE IS PRESENT IN P. FALCIPARUM tRNA.....	170
FIGURE 56 SYSTEMS BIOLOGY APPROACH TO STUDYING ARTEMISININ RESISTANCE MEDIATED BY tRNA MODIFICATIONS.	172

Acknowledgements

Exceptional science is only made possible perched atop the shoulders of giants. This could not be more apt, figuratively and phenotypically, for I owe a debt of gratitude to my supervisors – **Prof. Peter Preiser** and **Prof. Peter Dedon**. Thank you **Peter**, for the time and effort that you have patiently invested in me over the years. Your scientific ideal, gentle guidance and difficult questions were instrumental in my growth as a researcher and a person. As for **Pete**, you have shown that scientific excellence is not without empathy and compassion. You have provided for every opportunity that a graduate student can dream of and then some. Your unwavering confidence in my abilities has encouraged me in times of self-doubt and for that I will always be grateful.

Navigating grad school might not have been as easy had it not been for the wonderful community I work amongst. I would like to thank all members of the **Dedon** and **Preiser Labs** and my colleagues at **SMART**. A few deserve a special shout out. **Radek**, my brother in arms, for being my scientific-agony-aunt, but more importantly for being friend *extraordinaire*. **Weijian, Donald, Aoli, Omar, Yaw, Jenny, and Regina**, for their extremely valuable inputs and continued friendship over the years. Your patience and kindness in guiding my project means the world. **Marvin**, for holding strong the malaria fort at SMART, and the sagely advice on carbs. **Michael**, for a meticulous education in operating an Orbitrap interspersed with a history of Boston and Whitey Bulger. **Daniel** and **Sabrina**, for their warmth in an otherwise unwelcoming New England weather. **Jojo** and **Faeqa**, for all of your help and also for infusing the lab with laughter and mischief. It would have been miserable had you guys not been around. **Wenhe, Boon Chong, and Cheryl** for their continued support in the lab and also keeping a timely lunch schedule. **Peiying, Megan, Farzad** and **Mien** for ensuring that I was never short

of anything - in the lab or otherwise. A quick chat with either of you was as good as therapy-hour. *Sensei* **Wilfried**, for providing lessons not only in science and on life, but also a primer on self-reflection. **Weilin**, for keeping me extremely realistic about job prospects – nevertheless, inspiring true grit and determination. **Chee Sheng, Maggie, Yok Hian, Cui Liang, Jinling, Nick**, and **Amanda** (Koch Institute), for an initiation into the cult of mass spectrometry. **Enya** and **Jiaqi**, for their diligence in learning and tireless efforts. **Rajesh** and **Trang**, for showing my novice self the ropes on parasite culture and inspiring an infatuation that took the shape of a PhD. **Dipti**, you're sorely missed, but your persistent spirit and optimism remains an inspiration. Collaborators-turned-friends, **Jen** (Fidock Lab), and **Jess** and **Sebastian** (Scherf Lab) for their constant input, encouragement, and science *par excellence*. Friends-turned-family, **Vishaka, Hannah, Ania, Michela, Wint, Militsa, Yorick** and **Eric**, for all of the love/food, and support/beer. A big thank you to **Akshay, Advait, Akshaya**, and my aunt, **Sandhya**, for keeping me well-fed in Singapore and from ever having felt home-sick.

None of this is made possible without the Sisyphus-like support and encouragement that I have received from my family. You have picked me up every time I rolled down, and then back again, without ever failing. Grudgingly, I admit that I am constantly in awe of the work ethic and scientific temperament that my sister, **Tanvi**, embodies. She encompasses a multitude of roles – from my harshest critic to my fiercest advocate and this journey would not have been possible had it not been for you. To my parents, **Mahesh** and **Rani**, you have shown me how to love and cherish the work I do. I have not known neglect in any way. You have provided me with tremendous encouragement to pursue my dreams and celebrated every achievement. I love you very much. This thesis is dedicated to the both of you.

"Though this be madness, yet there is method in 't."

- Polonius, Act II Scene II Hamlet

"The nitrogen in our DNA, the calcium in our teeth, the iron in our blood, the carbon in our apple pies were made in the interiors of collapsing stars. We are made of starstuff."

- Carl Sagan, Cosmos

Abstract

Modifications to nucleosides present in ribonucleic acid (RNA) species have been shown to influence gene expression. Using a LC-MS/MS-based platform, we have characterized the full spectrum of modified ribonucleosides on the transfer RNA (tRNA) and messenger RNA (mRNA) in *P. falciparum*. We demonstrate that modifications on the wobble position of certain tRNA isoacceptors are associated with efficient translation of corresponding codon biased genes. Of the mRNA modifications, we quantified N⁶-methyladenosine (m⁶A) across the intra-erythrocytic lifecycle and characterized the putative “writer” enzyme complex that catalyses this modification. Upon carrying out m⁶A-sequencing, we find that m⁶A-associated transcripts have shorter half-lives and lower translation efficiency. Additionally, we quantified tRNA modification levels in the parasites that were exposed to oxidative stressors like hydrogen peroxide, sodium hypochlorite and anti-malarials like artemisinin and chloroquine. The work presented here holds potential for several avenues of new discovery. For one, unique patterns of tRNA modification reprogramming can distinguish mechanistically distinct damaging agents which supports the idea the tRNA modification patterns can be used as biomarkers of exposure. Besides these stress-specific patterns of tRNA modification changes are linked to selective translation of codon-biased mRNAs for stress response proteins in other organisms. In order to address this hypothesis in *Plasmodium* we have carried out a proteomics experiment to determine if the parasite proteins that are upregulated show a marked codon bias which can be tied up with the tRNA modification data. All of these findings together provide new insights into the role of translational control mechanisms involving RNA modifications in the development and pathogenesis of malaria parasites.

Preface

This research is driven by the convergence of three elements: the unceasing malaria pandemic, the emergence of antimalarial resistance in its causative organism, *Plasmodium falciparum* (*P. falciparum*), and a developing paradigm of translational control mediated by nucleotide modifications.

Nucleic acids, the central biopolymers responsible for storing and translating the genetic code, are subject to widespread chemical modifications. In addition to the four canonical nucleosides, there exist more than 100 unique structures representing a diverse set of chemical moieties, which are “written”, “read”, and “erased” endogenously by dedicated enzymes. These modifications, found in both RNA and DNA, are present on the base, sugar and/or the backbone. Not only are they diverse in nature, but also ubiquitous among all known kingdoms of life. Of these, some modifications are shared within a phylogenetic order and others are unique to only a particular subset of organisms. Despite this widespread occurrence, the underlying function for most of these modifications remains elusive, providing an unexploited research area to be investigated.

As for malaria, unsuccessful eradication attempts saw the resurgence of the disease at the end of the 20th century. However, management of the disease through vector control approaches and the use of artemisinin-based combination therapies saw the reduction of the global burden of the disease by almost half over the past decade. And yet, this number still remains perilously high, with ~445000 deaths and ~200 million reported cases each year (Figure 1). Of particular concern, is the rise of resistance to the front-line antimalarial artemisinin, especially in South-East Asia, mimicking a historical trend that rendered the previous generation of antimalarials

like chloroquine ineffective. This underlines an overwhelming need for scientific research to tackle the spread, diagnosis, and treatment of this deadly disease.

Figure 1 Countries with chloroquine resistance and documented failure for artemisinin combination therapies (ACT)

Geographical distribution of drug-resistance Plasmodium falciparum according to varying levels. Most of south America, Indian subcontinent, and, Africa are resistant to chloroquine treatment but still respond to ACTs . On the other hand, ACTs are failing in Laos, Thailand, Cambodia, and Vietnam. The inset graphs chart out annual incidence levels and the corresponding mortality numbers from 2000 to 2015. Figure reprinted with permission from Ashley et al. [3]

Taken together, these factors form the overarching direction of this thesis – to define the role of nucleic acid modifications in mediating resistance in the human malarial pathogen, *P. falciparum*. The following introductory chapter will first provide a review of the disease landscape, then assess the nature of nucleic acid modifications, and eventually outline the specific aims of this thesis.

1. Chapter I: Introduction

1.1. Malaria: Epidemiology, Etiology, and Pathophysiology

1.1.1. Malaria Epidemiology

Transmitted by the *Anopheles* mosquito, malaria is a parasitic disease caused by genus *Plasmodium* that is found in 91 countries around the world [4]. With over a 120 species that can be hosted by mammals, birds, and reptiles, only five are commonly known to infect humans. *Plasmodium falciparum* (*P. falciparum*) manifests a high load of the blood-stage of the parasites which sequester in organ systems and lead to severe anaemia in children under five, who form the majority demographic of malaria-related deaths. *Plasmodium vivax* (*P. vivax*) produces milder symptoms since the parasites are restricted to circulating reticulocytes [5], but with a dormant liver stage that produces recurrent episodes adding to malaria-associated morbidity that is also resistant to most anti-malarials [6]. *Plasmodium malariae* (*P. malariae*) and *Plasmodium ovale* (*P. ovale*), while understudied organisms, produce symptoms quite akin to vivax malaria. *Plasmodium knowlesi* (*P. knowlesi*), which infects macaques in South East Asia (SEA), is often misdiagnosed in humans. This bonafide zoonotic infection parasite can lead to severe malaria with a similar pathophysiology as *P. falciparum* [7].

With an estimated incidence of 200 million for *P. falciparum* and 8.5 million cases for *P. vivax*. Of these, sub-Saharan Africa accounts for 90% of falciparum malaria cases with a substantial variation in occurrence within and between countries. The incidence of *P. vivax* is much lower in this region because the African population is mostly negative for the Duffy antigen which is a pre-requisite for infection [8]. In Asia, while the total incidence is much lower (accounting for 7% of the cases in the world), the proportion of *P. falciparum* and *P. vivax* cases are more or less similar in number. *P. ovale* and *P. malariae* have an even distribution globally but a

low rate of incidence. As mentioned previously, *P. knowlesi* is a zoonotic infection transmitted from the long-tailed macaques commonly found in Malaysia, Thailand, and Indonesia and is not likely to be distinguished in routine microscopic examinations from *P. malariae*. *P. simian* and *P. cynomolgi* malaria are other documented zoonotic infections, but are observed very infrequently, possibly being misdiagnosed at times with other morphologically common species [9].

1.1.2. The lifecycle of *P. falciparum*

Malaria parasites have a multifaceted lifecycle which alternates between the *Anopheles* mosquito vector and the human host (Figure 2). Broadly, the lifecycle can be divided into three distinct stages: mosquito (sexual), pre-erythrocytic (liver), and erythrocytic (asexual). Parasite male and female gametocytes are ingested when a female *Anopheles* mosquito takes a blood meal from an infected human carrier. Subsequently, ingested male and female gametocytes fuse to form a diploid zygote within the gut of the mosquito [10]. This zygote then develops into a motile ookinete that penetrates the gut lining to form an oocyst [11]. The oocyst undergoes a process of sporogony to generate sporozoites that migrate to the salivary glands of the mosquito. These sporozoites are released into the bloodstream of their human host upon being bitten by such an infected mosquito. Upon reaching the liver, they traverse the Kupffer cells and the sinusoidal layer to invade hepatocytes [12]. These sporozoites then multiply asexually for a few weeks within the hepatocytes to form liver merozoites. Upon rupture from the hepatocytes, these liver merozoites are released into the bloodstream which marks the beginning of the intra-erythrocytic developmental cycle (IDC) contributing to most of the symptoms experienced by the host.

The released merozoite interacts with a host of red blood cell receptors in a coordinated fashion to facilitate the formation of a tight junction [13]. The parasite, then, actively deforms the RBC membrane with the help of an actin-myosin motor to aid invasion into the cell [14]. Immediately upon entry, the eponymous ring-stage parasite establishes a parasitophorous vacuole and proceeds to ramp up metabolism to form the trophozoite stage [15]. During this maturation, the parasite modifies the host cell by exporting several transporter proteins to the RBC membrane, forming the new permeation pathway that aids in transporting essential nutrients in host circulation needed for parasite development [16]. Concurrently, the digestion of hemoglobin in the RBC generates toxic oxidative radicals, requiring the parasite to establish a separate digestive vacuole solely to sequester the hemazoin [17]. These trophozoites then undergo multiple rounds of DNA replication followed by cytokinesis to form 12-24 individual parasites in a process called schizogony [18]. The segmented parasites, aptly called schizonts, now destabilize the host cytoskeleton to escape and reinvade a new population of red-blood cells and propagate the cycle [19]. A sub-population of these parasites differentiate to produce male and female gametocytes to be taken up in another blood-meal by the mosquito, thus completing the life-cycle of the parasite [10].

Figure 2 Lifecycle of *P. falciparum*

(A) Original illustration by Laveran of the stages of in the life cycle of malarial parasites and their pigment (black dots), which was published in the *Société Médicale des Hôpitaux de Paris* in 1881. **(B)** Lifecycle of the malaria parasite inside the host. The cycle begins with the inoculation of the sporozoites by the mosquito. These sporozoites travel to the liver to form merosomes within hepatocytes. These liver schizonts are then released into the bloodstream and invade new red blood cells to form ring, trophozoite, and schizonts in a systematic progression. Late stage schizonts rupture and individual merozoites invade a new set of red blood cells to restart this cycle of infection. Figure reprinted with permission [3].

1.1.3. Pathogenesis and Clinical Symptoms

A certain titre of the aforementioned asexual parasites are needed in the blood stream before we are able to observe any symptoms denoting the prepatent period. This threshold sits at about 100 parasites per μL of blood. Upon being bitten by an infected mosquito, these symptoms take about 10-14 days for *Plasmodium falciparum* or *P. knowlesi*, 2-3 weeks for *P. vivax* and

P. ovale, and 18 days or longer for *P. malariae*. Classical accounts describe periodic fever spikes in an individual. This denotes the synchronisation of the intra-erythrocytic developmental stages. However, such patterns are now rarely observed in the clinical setting, especially with *P. falciparum* causing sporadic febrile illness instead of a periodic episode [20]. Symptoms of uncomplicated malaria are very non-specific and can include fever, chills, bodyache, headache, cough, and diarrhoea, thus making diagnosis a challenge. In most non-endemic areas, fevers, anaemia and thrombocytopenia associated with a travel history in the tropics is crucial to malaria diagnosis [21].

A unique feature of mature *P. falciparum* parasites containing red blood cells is that they sequester inside blood vessels and consequently avoid getting cleared by the spleen [22, 23]. This leads to microvascular obstruction and inflicts endothelial cell damage. Not only do these infected red cells become rigid, they also sequester other circulating uninfected red blood cells to form rosettes, further aggravating the obstruction in the microvasculature. The pathology of this sequestration is manifested in the organ(s) involved. For example, sequestration in the lungs inclines patients to develop respiratory failure [24]. In pregnant women, sequestration in the intervillous space of the placenta can cause low birth weight, impaired organ development in the fetus, and maternal and/or infant death [25]. Sequestration in brain capillaries leads to a complication called cerebral malaria in which patients develop an unrousable coma which may be accompanied by seizures, retinopathy, and brainstem degeneration. Whether this is caused by occlusion to the capillaries, thus depriving the brain of nutrients and oxygen, and/or is a consequence of pro- or anti-inflammatory factors is still up for debate [26-29].

1.2. Anti-malarial Approaches and Challenges

1.2.1. Anti-malarial Drugs

The therapeutic of choice is determined by the infective species of the parasite, the risk for existing drug-resistance and tolerance of the medicine by the endogenous population. Evidence-based guidelines are accessible on the WHO website for treatment of uncomplicated malaria. The past decades have seen the parasite evolve resistance to chloroquine and antifolates requiring a transition to artemisinin-based combination therapies (ACT) as a modality for treatment of *P. falciparum* malaria. ACTs generally involve a combination of an artemisinin-derivative that is fast-acting in lowering parasite load and a partner drug with a longer half-life to combat any residual parasites. Current leading formulation of ACTs exist as the following combinations: artesunate–mefloquine, artesunate–amodiaquine, artemether–lumefantrine, dihydroartemisinin–piperaquine, and artesunate plus sulfadoxine–pyrimethamine [3]. A complementary combination treatment which also serves as prophylaxis for travellers is the use of atovoquone-proguanil on a case-to-case basis. However, this is not endorsed for use on a widespread health policy level since the parasites tend to acquire resistance to atovoquone rapidly [30]. Quinine, while still effective, is poorly tolerated, requires longer treatment durations, and also requires secondary treatment with clindamycin or doxycycline. Chloroquine followed by primaquine to kill any dormant parasites is the therapy of choice for treating *vivax*, *ovale* and *malariae* species of *Plasmodium* [31]. There are several drugs in clinical trials at the moment which are mainly blood-stage schizonticides and the development of transmission-blocking drugs has garnered much interest in the field.

1.2.2. Vector Control

One of the main factors that has had an impact on keeping the malaria incidence in check was the use of long-lasting insecticide (pyrethroid) treated bed-nets (LLINs) and indoor spraying with pesticides. With more than 40 species of *Anopheles* mosquitos worldwide, the use of LLINs have considerably lessened morbidity and mortality in high transmission areas and proven to be tremendously effective against the *Anopheles gambiae* complex which is the major vector in Sub-Saharan Africa. It might be prudent to look at new pesticides since there have been a flurry of reports of resistance to insecticides in the vector [32, 33]. Orthogonal approaches to vector control involve the mass administration of Ivermectin [34], which reduces the lifespan of mosquitos or the use of transgenic mosquitos rendering their progeny to be sterile/ineffective at transmitting parasites [35].

1.2.3. Vaccines

Current vaccines focus on targeting the sporozoite homing to the liver and the subsequent transversal into hepatocytes. Consequently, the design for the epitope used for these vaccines are based on the surface displayed circumsporozoite protein. A commercial formulation termed as the RTS,S/AS01 vaccine showed effective protection against *P. falciparum* in children over 5 years in a longitudinal study spanning 4 years. The overall vaccine efficacy, however, is still under 40%, unable to provide protection across other malarial species, and a long-term reduction in overall mortality still remains to be demonstrated [36]. In contrast to using single protein epitopes, another vaccine design involves the use of irradiation-attenuated intact sporozoites. This vaccine termed PfSPZ, is now in clinical trials in Africa. The biggest challenge that we foresee is overall protection from all species of the parasite [37].

1.2.4. Treatment Failure and Resistance

Current generation artemisinins are active against all of the intra-erythrocytic developmental stages. However, their usefulness is severely limited due to the emergence of ring-stage resistance [38-40]. Clinical resistance typically manifests as a delay in clearance of the parasite from the patient bloodstream upon ACT administration. First documented in Western Cambodia, studies have identified multiple sites in Southeast Asia and South China where patients show greater than a 5.5 hours clearance half-life of the parasite which is a benchmark for artemisinin resistance by WHO standards [4]. Of critical note is that this resistance still pertains to only the ring-stage of the parasite, and thus is defined as “partial” resistance. It is still unknown whether the parasite would be able to evolve resistance to its other stages leading to “complete” resistance. By itself, artemisinin resistance rarely leads to treatment failure. However, the presence of partner drug resistant parasites can lead to treatment failure. Thus, different ACTs are currently failing in the Greater Mekong region which has both artemisinin and ACT partner drug resistance. Recently, mutations in the propeller domain of the *P.falciparum* gene encoding kelch (PfK13) have been strongly shown to be the primary determinant of this partial resistance. Up until now, 13 independent PfK13 mutations have been linked with clinical resistance. Of these, 4 (Y493H, R539T, I543T, and C580Y) have been validated *in vitro* and are in use for the widespread monitoring and surveillance of artemisinin resistance [41, 42].

1.3. Transcriptional and Translational Control in *P. falciparum*

1.3.1. *P. falciparum* Genome

The first draft of the *P. falciparum* genome was completed in 2002 from a parasite clone 3D7 [43]. The genome comprises of a 23Mb nuclear genome ordered into 14 linear chromosomes varying from 0.6Mb to 3.3Mb, a 35kb apicoplast circular plastid and a 6kb mitochondrial genome. Akin to other protists, the gene density is relatively high, with 50% of the genome predicted to code for ~5300 proteins. Of these, about half the genes contain an intron, making the average gene length longer than related organisms. Even with sophisticated predictive homology mapping 33% of the genes cannot be functionally unannotated, reflecting our limited grasp over *Plasmodium* biology [44].

1.3.2. Transcriptional and Translational Profiling of the *Plasmodium* IDC

High-throughput transcriptomic profiling suggests a tightly regulated gene expression system across the parasite life-cycle, whereby the transcriptome of the parasite reflects the physiological needs of the parasite phenotype exhibiting a “just-in-time” transcriptome [45]. This periodic transcription program is visible across the *Plasmodium* genus for at least 6 different species which includes rodent malaria for 69% for the genes [46]. Of these, peak time expression of syntectic orthologs is relatively conserved, but species-specific genes pertaining to antigenic variation, host-parasite interactions, cytoadherence, and erythrocyte aggregation differ substantially.

P. falciparum parasites, upon different drug stresses like chloroquine and several anti-folates, do not show specific transcriptional changes but rather respond by cell cycle arrest and the induction of a generic stress response pathway [45, 47-52]. This provides for a model of “transcriptional rigidity” whereby the parasite has a limited capacity for feedback mechanisms and the transcriptional regulation is “hardwired”. Contrasting, but not mutually exclusive studies show that *P. falciparum* can change its transcription upon external stimuli like glucose deprivation and heat shock which are reflective of the environmental niches it encounters in the hosts [46, 53]. In addition, a detailed transcriptional profiling of *P. falciparum* treated with 20 small molecule inhibitors indicated that functionally relevant genes in response to the growth perturbations are significantly regulated [54]. Complementing these studies are those that compare the *Plasmodium* transcriptome and proteome through its lifecycle. These studies observe that the steady-state mRNA levels do not strictly correspond with protein levels in the cell [55-57]. Not unique to *Plasmodium*, there also exists evidence of similar observations in other organisms [58]. Collectively these studies point to an important role for post-transcriptional control mechanisms in regulating gene expression during parasite development.

1.3.3. Post transcriptional control during the IDC in *Plasmodium*

Across the lifecycle of the parasite in each of its hosts, the parasite undergoes various morphological and phenotypic changes and its transcriptome and proteome broadly reflect the unique requirements for these stages. Messenger RNA (mRNA) in *Plasmodium* exists in equilibrium between de novo synthesis and degradation. These degradation pathways involve removal of the m⁷G cap and deadenylation followed by hydrolysis by a variety of ribonucleases [59]. Alternatively, it could be degraded by nonsense-mediated decay involving endonucleolytic cleavage [60]. Besides, other trans-acting splicing effectors such as

Serine/Arginine-rich Protein 1 can also influence mRNA stability [61]. Non-coding RNAs have also emerged as essential post-transcriptional control elements. However, the existence of miRNAs in *Plasmodium* is doubtful since the parasite does not appear to encompass any homologous genes required for miRNA activity nor is there evidence of any parasite encoded miRNAs [62]. There is, however, a class of small ncRNAs called natural anti-sense transcripts (NATs), which exist in prokaryotes and eukaryotes, that have been implicated in regulating gene expression in *Plasmodium* [63]. The translation of the mRNA transcript can be split into three parts: initiation, elongation, and termination. Each of these processes are under different control mechanisms. The phosphorylation of eukaryotic initiation factor 2 can lead to global translational repression in the parasite [64]. Eukaryotic initiation factor 4-binding proteins can inhibit cap-dependent ribosomal loading onto an mRNA [65]. It has also been shown that there are cis-acting elements in 5' UTRs of select genes that inhibit translation of the coding sequence [66]. A review of all mechanisms involved in post-transcriptional and translation control can be found in Vember et al [67]. However, in the interest of this thesis proposal, the emphasis is on a mechanism centering on modified ribonucleosides while bearing in mind that the *Plasmodium* genome encodes genes capable of providing such a tool.

1.4. Translational elongation and codon choice

1.4.1. Translation and codon degeneracy

During translation, a triplet of nucleotides “codons” found on the mRNA are read by transfer RNAs with complementary anticodons carrying amino acids within a ribosomal RNA-protein complex. Consecutive amino acids link to each other by amide bonds to form polypeptides which finally yield proteins. A possible permutation of four nucleotides in three positions yields $4^3=64$ codon choices on the mRNA. Accounting for three codons (UAG, UAA, UGA)

that specify for the termination of the protein sequences, this leaves 61 codons specifying for 20 possible amino acids. Thus, there exists degeneracy of the genetic code to allow for multiple codons to code for the same amino acid. Leucine, for example, is considered as a ‘six-box’ amino acid because codons UUA, UUG, CUU, CUC, CUA, and CUG, all specify for leucine. Similarly, if two codons specify for an amino acid, it is considered to be a ‘two-box’ amino acid.

1.4.2. Wobble Hypothesis

61 codons would correspondingly require tRNAs with 61 anti-codons for accurate decoding. However, most organisms encompass fewer than 45 species of tRNAs, leading us to conclude that certain tRNA anticodons pair with more than one codon. For example, mRNA codons CUC and CUU both specify for the amino acid leucine, but both of these codons are decoded by a single tRNA^{LeuGAG} (Figure 3). This is achievable only as a result of wobble-base pairings. The first two nucleotides of a codon (here “CU”) are subject to strict Watson-Crick pairing with those of the anticodon (here “GA”). The third nucleotide of an mRNA codon, however, may engage in non-standard pairing (U:G) with the first nucleotide of the tRNA anti-codon and this position is referred to as the wobble position. For example, tRNAs with ANN anticodons can undergo deamination of Adenosine to Inosine (thus forming INN) which is sufficient for wobble pairings with Uracil, Cytosine, and Adenosine and thus carry an expanded capacity for decoding multiple codons with a single tRNA isoacceptor [68]. A point about nomenclature here is that tRNA species containing different anticodon stem loops, but coding for the same amino acid are termed isoacceptors. Whereas, tRNAs species containing the same anticodon sequence and amino acids, but with variations outside the ASL, are termed isodecoders. Another common wobble pairing is G:U/U:G whereby the U ending codon of two-box codon families such as Asparagine, Aspartic acid, Cysteine, Histidine, Phenylalanine,

and Tyrosine are read by corresponding GNN anticodons instead of ANN anticodons [69]. In most organisms, including *Plasmodium*, the ANN anticodons for these amino acids are absent so as to prevent misincorporation in the event of deamination leading to the previously mentioned I:U, I:C, and I:A wobble pairings. In fact it has been shown that appearance of uridine modifying enzymes and adenosine deaminases contributed to the divergence of eukaryote and bacterial genomes from their archaeal counterparts [70]. This modification of an adenosine to inosine served to increase the decoding capacity of modified tRNAs and were thus, positively selected during evolution.

Figure 3 Schematic depicting the wobble hypothesis

Codons CUC and CUU on the mRNA both specify for the amino acid leucine. These two codons can be decoded by the same tRNA which contains GAG in the anticodon stem loop a result of non-canonical pairing (highlighted in yellow). (NB: This schematic is for the purposes of demonstration of the wobble hypothesis in general since Plasmodium

falciparum does not encode a gene for this specific tRNA^{LeuGAG}). Figure designed in Biorender and used with permission in an academic thesis.

1.4.3. Codon Choice

This concerted event of codons and anticodons during translation gives credence to the possible role played by tRNAs on codon choice. An additional point in support of this argument is that tRNA gene copy numbers often reflect the abundances of the corresponding codons in some organisms. This, however, is not the case for *Plasmodium* since all species only possess a set of 46 non-redundant, nuclear encoded tRNA genes encoding for 45 isoacceptors (where the initiator methionine gene is different from the elongator methionine gene) and no isodecoders (Figure 4).

The apicoplast tRNA on the other hand contains 35 genes encoding for 26 tRNA isoacceptors [71]. Naturally, these single gene copy numbers do not reflect the composition of the *Plasmodium* genome in terms of its codon choice which is highly skewed (Figure 4). In addition *Plasmodium* species that have been sequenced demonstrate a diverse range of GC content. Particularly *P. falciparum*, which is the focus of this thesis, has an 81% AT content in its genome, making it one of the highest of all sequenced genomes [43].

Figure 4 Codon Utilization and Decoding Scheme in *Plasmodium falciparum*

(A) Utilization of codon choice in the *Plasmodium falciparum* exome. *Lys*AAA and *Arg*AAT codons constitute for more than 20% of the total codons utilized. The histogram also depicts that codons with A+T usage are dominant in *P. falciparum*. (B) Decoding scheme illustrating possible pairings of codons and anti-codons. Red highlights indicate G:U wobbling and grey highlights depict other wobbling (I:U, I:C, and I:A). tRNA genes present in *Plasmodium falciparum* are indicated besides respective anti-codons. *Methionine is the only tRNA isoacceptor with 2 tRNA genes since the initiator methionine is different from the

elongator methionine. Reprinted with permission from [72] under Creative Commons CC BY license.

1.5.Nucleotide Modifications and the Epitranscriptome

All the canonical bases in RNA can be modified enzymatically *in vivo*. More than 120 distinct chemical modifications have been catalogued to date (Figure 5) [73]. These modifications have been studied to some depth in transfer and ribosomal RNA showing their essentiality in proper folding, stability, maturation and function of these molecules [74, 75]. An important point to note is that the enzymes catalysing these modifications are dynamic, comprising of “writers, readers, and erasers” which can add, recognize and remove the modification and thus ultimately decide the fate of the RNA molecule. Collectively these RNA modifications were first termed ‘RNA epigenetics’ which eventually gave rise to the neologism ‘epitranscriptome’ analogous to epigenetic regulation which is mediated by modifications to the DNA and histone proteins [76-79].

Figure 5 Chemical Structures of Selected RNA Modifications

Modifications derived from all 4 canonical ribonucleosides are shown. These modifications range from simple alkylation, isomerizations, sulfonation, oxidation, and deamination to complex linkages involving glycosides and amino acids. Individual chemical structures reprinted from [73] with permission.

In the past decade, while the epigenome has been studied extensively, the epitranscriptome has been neglected due to the lack of sensitive and robust tools that can detect these RNA modifications. There exist several hurdles to discovering and mapping the epitranscriptome. First, while ribosomal, transfer, and small nuclear RNA are plentiful in the parasite population, messenger RNA and long noncoding RNA are of low abundance. Second, sequencing studies

to map these modifications can only be used to determine some of the modifications and cannot be used for the discovery of new modifications [80]. Third, there is a gap in excellent quantitative and sensitive methods to determine these modifications which cannot be fulfilled by sequencing alone.

1.5.1. tRNA Modifications

Transfer RNAs (tRNAs) are adaptor molecules composed of 75-90 nucleotides in length that are charged with specific amino acids on their 3' end containing an average of 8-10 modifications. Of the known RNA species, modifications decorating the tRNA have been extensively identified and studied. The review provided in this thesis is by no means exhaustive but is comprehensive in describing the modifications and the myriad of roles they play. While these modifications can occur anywhere, most of the literature has focused on those found near the anti-codon stem loop (ASL), particularly for position 34 (wobble) and position 37 [69, 72, 81-98]. Modifications specifically at or around the anti-codon loop are implicated in translational fidelity, translational efficiency and frame-shift prevention. Wobble modifications (position 34 or 35 on the tRNA) have been demonstrated to increase the number of possible non-canonical base pairings and thus allow for organisms to require fewer tRNA isoacceptors to decode all 61 codons found in mRNAs. Conversely, certain wobble modifications can also restrict recognition of cognate codons whereby the cellular physiology might merit precise translation. Early studies focused upon the role that these modifications play in tRNA folding and stabilizing the secondary structure through base pairing and pi-stacking. For example, a sulfur-based modification s²U on the wobble position of the tRNA has been shown to improve the mRNA-tRNA duplex stability as a result of pi-pi stacking interaction [81, 99, 100]. This becomes particularly important in the context of *Plasmodium*

falciparum which has LysAAA and AspAAT repeats, to prevent frameshifting. 5-methylaminomethyl-2-thiouridine (mnm⁵s²U) at the wobble position of bacterial tRNA^{LysUUU} is able to distinguish between cognate AAA and AAG and near-cognate stop codon UAA or isoleucine codon AUA [96]. In addition, the wobble modification methylcytidine (mC) increases the efficiency of reading stretches of TTG codons. Position 34 also lends itself to some of the most complex modifications. tRNAs containing a wobble U are modified to 5-methoxycarbonylmethyl (mcm⁵) by the Elongator complex (Elp1-6) and can be subsequently hypermodified to a 2-thio derivative (mcm⁵s²U) by Ncs2 and Ncs6. This helps discriminate between amino acids Asp/Glu, His/Gln, and Asn/Lys. Mutations in these enzymes show several phenotypes, including slow growth and telomere shortening [83, 91, 92, 101].

One of the modifications that is present in archaeal and eukaryotic tRNA but absent in bacteria is wyosine (imG) and its hypermodified forms of wybutosine (yW) and hydroxywybutosine (OHyW). These modifications, usually found at position 37, have been shown to be essential in translational fidelity of tRNA^{PheGAA}. Most interestingly, yW and OHwY are found in cytosolic and mitochondrial tRNA of kinetoplastids whereas imG is strictly a mitochondrial modification. The presence of these modifications are envisaged to contribute to the coding diversity of *Trypanosoma*, another unicellular eukaryotic parasite akin to *Plasmodium* [102]. In addition, position 37 modifications play a significant role is to prevent frame-shifting (FS) in translational elongation. m¹G at position 37 in tRNA^{Pro} was found to be the major determinant of avoiding +1FS [103, 104]. Relatedly, this modification is a prerequisite for the formation of yW on tRNA^{Phe}. Modification of adenosine to i⁶A at position 37 has been shown to increase translational activity of the tRNAs by 4-fold over unmodified tRNA.

Figure 6 Modifications frequently found on tRNAs

tRNA is shown in its regular secondary structure with each circle representing a nucleotide and the lines indicating hydrogen bonds between bases. Magnesium cations stabilizing the structure are shown in yellow. The top-left inset displays the L-loop tertiary conformation for the tRNA. Commonly found modifications at each of the positions are displayed as mapped from a comprehensive literature review with the wobble position displaying the most diversity. Taken from [105] with permission.

Not only do these modifications impact tRNA:mRNA interactions but modified tRNA is also known to have two orders of magnitude higher binding capacity to the small ribosomal subunit

compared to unmodified tRNA [106, 107]. Specifically, Mg^{2+} dependent m^5C stabilizes anticodon domain structure and m^1G aids in the open loop configuration to provide a robust yeast tRNA-ribosome binding [84]. As for the amino acid charging step, aminoacyl tRNA-synthetases (aaRS) that recognize tRNAs based on its acceptor arm and the nucleoside sequence anticodon loop haven't so far been shown to be influenced by the presence or absence modifications.

1.5.2. mRNA Modifications

One of the first RNA modifications taught in undergraduate biology courses is the synthesis of the 5' N^7 -methylguanosine triphosphate cap and a 3' polyadenylated tail to the messenger RNA in the eukaryotic cell. Early studies relying on the use of a radiolabelled methyl carrier showed that internal regions in the mRNA harboured significant radioactivity [108-110]. Some elegant biochemistry revealed these sites to mostly be N^6 -methyladenosine (m^6A) with a minor proportion of N^5 -methylcytosine (m^5C). m^6A in the past few years has been shown to affect almost every stage of mRNA metabolism, from processing in the nucleus to translation and decay in the cytoplasm (as reviewed in Zhao BS et al [77]). We review and delve a little further into the enzymes catalysing the m^6A modification and its physiological role as part of Chapter 4.

Figure 7 Control of gene expression by RNA modifications

Role of m⁶A in controlling mRNA expression to induce cell differentiation. Specific transcription factors recruit the m⁶A-writer machinery (denoted by METLL3/14) to methylate transcripts that encode for cell fate factors which are translated or repressed as per m⁶A binding proteins. A stimulus to induce differentiation may lead to a methylation pattern on different genes orchestrated by the same writer-reader apparatus to aid in transcriptional switching. Reprinted with permission from [111].

In addition to this, recent evidence points to the existence of pseudouridine (Ψ) [112], which is a result of enzyme-catalysed isomerization of uridine. While it remains the most abundant modification in tRNA, the presence of Ψ in mRNA has been shown to affect base-pairing interactions and affect mRNA coding and secondary structure [113]. For example, pseudouridylation of a stop codon was previously shown to suppress translation termination whereas a contrasting study showed this to be incorrect using a combination of structural biology and biochemistry [114]. Pseudouridine-modified mRNAs used for therapeutic

applications show better higher translational activity as opposed to unmodified mRNAs [115]. A follow-up on this study shows that synthetically incorporated N¹-methylpseudouridine had much more a profound impact on protein turnover and elicited a lower immunological response as opposed to either only pseudouridine-modified or unmodified mRNA [116]. A contrasting study in response to the above two therapeutic studies tells a different story, in that neither of the modifications impact protein turnover as much as tweaking the primary coding sequence clearly depicting a gap in our understanding of the role played by pseudouridine [117]. All known eukaryotic mRNAs have been shown to have methylated ribose sugar moieties on the 2'-O-hydroxyl group to 2'-O -methylation (Xm) [118, 119]. These moieties generally sit adjacent to the 5' cap of viruses and eukaryotic mRNA which has been shown to protect mRNAs from decapping enzymes like Dxo1 in yeast [120]. Recent work by Puglisi et al show that important ribosomal-monitoring bases (G530, A1492 and A1493) are sterically blocked from reading cognate tRNAs when the coding regions in the mRNA contain 2'-O-methylation [121]. m⁷G is perhaps the best known cap modification which protects the mRNAs from exonucleases, nuclear export, and translation initiation. It has recently been shown to be present internally on mRNA coding regions and miRNAs improving translational efficiency and miRNA processing, respectively [122, 123]. N¹-methyladenosine m¹A [124, 125] was another modification reported in recent years. Of these two groups that carried out the initial methylome mapping studies, Schwartz et al published a second analysis to reconcile the disparate findings from both studies. They demonstrate that most of the sites attributed to m¹A are a result of misannotations, incorrect mapping, duplications, SNPs, and sequencing errors calling into question the physiological relevance of this modification altogether [126]. Recent work was also carried out on mapping the N⁵-methylcytosine (m⁵C) methylome in human [127, 128], mouse tissue [129], mouse ESCs [130] and plants [131]. Inconsistent results between each of these studies raise questions on the validity of techniques that were used to carry out

some of these assays. True biological variation in these varied tissues could also explain these differences but we would be prudent to hold judgement till further corroborative work has been carried out. As an emerging modification, very few functional roles have been attributed to mRNA-associated m⁵C. Since this methylation sits on the Hoogsteen edge of the base, it plays little role in dictating canonical Watson-Crick pairing. However, it can possibly have an impact on base-stacking and secondary mRNA structure. It has also been shown to interact with mRNA export proteins [128]. An oxidised form of N⁵-methylcytosine, N⁵-hydroxymethylcytosine was identified (hm⁵C) in eukaryotic mRNA but has not been attributed a functional role so far [132].

1.5.3. RNA Modifications and Stress Response

Studies carried out by the Dedon laboratory and others focused on how RNA modifications can participate in dynamic cellular processes during stress [86, 87, 90, 94, 95, 97, 133-135] (Figure 8). Specifically, for tRNA modifications, studies carried out in our lab showed that *S. cerevisiae*, a eukaryotic model organism, can operate as a system that changes dynamically in response to environmental stressors which enables its survival and subsequent adaptation. For example, the synthesis of wobble modifications like mcm⁵U is catalyzed by tRNA methyltransferase 9 (Trm9). The deletion of this gene resulted in stress-induced cell cycle and growth defects [87, 94, 136]. The mechanism proposed that this modification promoted the translation of DNA damage response genes which were enriched with codons that can be recognized by the wobble-modified anticodons. These defects were not found in normal replicative conditions although there existed a slight deficit on translation of the same codons in a reporter based assay [87]. In a similar manner, the modification to m⁵C at the wobble position of yeast tRNA^{Leu(CAA)} helped the translation of UUG-rich genes important for survival

On the mRNA front, N⁶-methyladenosine (m⁶A) is emerging as a prominent player dictating gene regulation. For example, m⁶A levels are generally low in the 5' UTRs. Heat shock treatment represses cap-dependent translation, but a single m⁶A residue in a 5' UTR leads to cap-independent translation of heat-shock-responsive mRNAs [139]. This balance is orchestrated by a variety of enzymes termed “readers, writers and erasers” which aid in the recognition, addition and removal of the modifications [139, 140]. The presence of a similar adenosine methylation, m¹A, in the 5' UTR of transcripts results in better translational efficiency as compared to unmodified transcripts. This modification has specifically been studied in yeast for a nitrogen deprivation model and mammalian cells in the context of various stimuli [124, 125] but this result needs to be approached with caution due to inconsistencies in results as mentioned previously. Pseudouridinylated codons can have differential translation dynamics, which has been implicated in stress tolerance in *Toxoplasma*, a closely-related organism to *Plasmodium* [141].

1.6. Research drive and Rationale

Precise choreography of the *Plasmodium* transcriptome with limited transcriptional control elements merits an investigation into complementary regulatory systems. Mechanisms involving RNA modifications and stress response have been demonstrated in organisms such as *E. coli*, *M. bovis* and *M. smegmatis* and eukaryotes like *S. cerevisiae* and mammalian cells. We hypothesize that these modification-based-regulatory systems are common features of all organisms. The driving rationale for this project is to investigate precisely this hypothesis in *Plasmodium* development and stress response.

1.6.1. Specific Aims

The objective of this thesis is to study translational control mechanisms that regulate gene expression in the *Plasmodium* parasite. Our hypothesis proposes that there is a dynamic shift in the RNA modifications as the parasite progresses through its developmental cycle and under stress whereby the change correlates with a corresponding change in the translation of the proteome and thereby fulfils a regulatory role.

1. To characterize and quantify wobble modifications on *Plasmodium* tRNAs across the lifecycle (featured in Chapter 3)

Research Gap: We have encouraging preliminary data that catalogues tRNA modifications found in *P. falciparum* and relates proteome-wide changes with the levels of some of these modifications. However, the identification of the specific tRNA isoacceptor containing the modification, it's relative position on the tRNA, and subsequent quantification is needed to address its significant gap in the study.

Strategy: Develop LC-MS methods for identifying, characterising, and quantifying modified tRNA oligonucleotides from bulk tRNA in *P. falciparum*.

2. To develop an experimental platform for characterization and quantification of mRNA modifications in *Plasmodium falciparum* (featured in Chapter 4)

Research Gap: There exists no study to date that has investigated mRNA modifications in *Plasmodium*, yielding a fundamental discovery question – are there any mRNA

modifications in *Plasmodium*? If so, what are they, and how do they behave through the parasite lifecycle? Can we tease out the enzymatic complexes catalysing some of these modifications?

Strategy: Develop methods for efficient isolation of *P. falciparum* mRNA, followed by identification, structural characterization and quantification of each modified ribonucleoside in purified mRNA by liquid chromatography coupled tandem mass spectrometry (LC-MS/MS). Immunoprecipitate putative tagged enzymes to investigate associated proteins.

3. To integrate ‘omics’ datasets for systems-level translational control of *Plasmodium falciparum* under stress conditions (featured in Chapter 5)

Research Gap: With limited capacity for transcriptional control, I ask the question whether *Plasmodium* stress response can be explained as function of tRNA modifications as evident in other organisms.

Strategy: Establish parasite stress response to a variety of drugs and chemical reagents. Isolate total tRNA and quantify tRNA modifications under stress conditions. Carry out quantitative proteomics across the same conditions and timepoints. Use multivariate statistics to establish a novel model of translational control in *Plasmodium* parasites.

Chapter 2

2. Materials and Methods

2.1. Culturing *Plasmodium falciparum* in vitro

In order to culture the asexual stages of *P. falciparum*, strain 3D7 was obtained from MR4 resources.. These were grown in human red blood cells (RBCs) in RPMI 1640 medium (Invitrogen) supplemented with 0.5% Albumax II (Gibco, Life Technologies), 23.81 mM (2 g/L) sodium bicarbonate (Sigma), 0.1 mM hypoxanthine (Sigma), and 104.7 μ M (50 mg/L) gentamycin (Gibco, Life Technologies). The cultures were incubated at 37 °C with 5% CO₂ and 5% O₂ [142]. The cells were routinely inspected using microscopy to check for culture contamination. Cells were synchronized by repeated treatment with 5% D-sorbitol (Sigma-Aldrich) every two to three generations to enrich for ring-form infected and uninfected RBCs. After each sorbitol treatment, the mixture was centrifuged at 600 $\times$ g (acceleration, 9; deceleration, 2) for 5 min to remove sorbitol and lysed infected RBCs in the supernatant. The pelleted RBCs were washed twice with incomplete RPMI 1640. Following this enrichment process, the synchronized cells were grown to reinvade fresh RBCs.

2.2.Stress Treatments for IC₅₀ Assay

Inhibitory concentrations (IC) assays were performed in 48-well plates with two-fold serial dilutions of dihydroartemisinin (Sigma), chloroquine (Sigma), hydrogen peroxide (Merck), spermidine-NO (Merck), SNAP (Sigma), menadione (Sigma), and sodium hypochlorite (Sigma-Aldrich) at 2.5% hematocrit and 1% parasitemia at the trophozoite stage. These cultures were incubated for 48 hours to be allowed to propagate for one life cycle for a lifecycle IC₅₀ assay

(Figure 41 red). For short term exposure, the drug treatment was performed for 3 hours at 5% parasitemia. The drugs were washed, and the parasites were diluted with fresh blood to 1% parasitemia and 2.5% hematocrit for 48 hours and allowed to propagate (Figure 41 maroon). The parasites were exposed to different stress agents for 3 hours for the tRNA modification profiling prior to harvesting as described in section 2.3 (Figure 42). The experiments were performed with three biological replicates for individual stress exposure. Parasitemia of treated and untreated cultures were quantified by flow cytometry where Hoechst 33342 (Invitrogen) and MitoTrackr (Invitrogen) were used to stain DNA and mitochondria thus distinguishing live cells from dead parasites. For plotting dose-response curve, the concentrations of the treatments were log-transformed, and the responses were expressed as the fraction (%) of treated to untreated control. Nonlinear regression curve fitting was performed with four-parameter least squares fitting method by Prism 6.0 (GraphPad). The IC_{20} , IC_{50} and IC_{80} of each stress agents determined from the fitted curves and are listed in the results section.

2.3.RNA Extraction

To minimize host RNA contamination, parasites were isolated by lysing the RBCs with 0.05% saponin (Sigma) and washed 3-times with ice-cold phosphate buffered saline (PBS). Isolated parasites were then homogenized with 5 volumes of Trizol reagent (Invitrogen), followed by a 5-min incubation at ambient temperature before freezing at -80 °C. The Trizol homogenized cell lysates were incubated at room temperature for 3 min with chloroform which was one-fifth the lysate volume. The mixture was centrifuged at 12,000 $\times$ g for 15 min at 4 °C and the aqueous phase was collected. Absolute ethanol (Merck) was added to the aqueous phase (35 % v/v) and total RNA species were extracted using a PureLink miRNA Isolation Kit (Invitrogen) according to manufacturer's instructions. A sequential isolation protocol was adopted to enrich

the yield of tRNA using 70 % ethanol for further experiments. The quality and quantity of RNA extracted from each experimental time point were determined by Bioanalyser RNA 6000 Nano chips and small RNA chips (Agilent Technologies).

2.4.HPLC Purification for tRNA

To purify tRNA to homogeneity, we performed one-dimensional size-exclusion chromatography (SEC) using a Bio SEC-3 300 Å (ID 7.8 mm, 300 mm) column (Agilent Technologies). Ammonium acetate (100 mM, pH 7.0, Sigma-Aldrich) was used as the mobile phase as a mild ion-pairing reagent. Isocratic separations were performed at 0.5 mL/min for 30 min under partially denaturing conditions at 60 °C on an Agilent 1200 HPLC system, with the RNA elution monitored by diode array absorbance spectroscopy (200-300 nm) at 10 nm intervals. The tRNA-containing peak was collected (Agilent 1260 Infinity fraction collector) and desalted and concentrated with a Vivacon 500 10K MWCO spin column (Sartorius Stedium) according to manufacturer instructions. The quality and quantity of tRNA was assessed using the Agilent Bioanalyser RNA 6000 Small RNA chip.

2.5.Oligo-based tRNA Pulldown

An oligo-based pulldown approach was used to purify tRNA Gly, Glu, and Gln from a crude sample of *P. falciparum* tRNA. Approximately 0.2 mL of NeutrAvidin agarose beads (Thermo Scientific, Waltham, MA) were bound to 18 nmol of a biotinylated DNA oligo (IDT, Coralville, IA) which was used to entrap its target, from a pool of 100 µg total tRNA. This biotinylated oligomer had a sequence complementary to the anti-codon stem loop of the target tRNA that was unique in this tRNA pool. The target tRNA was bound to its biotinylated “bait”

in 6 X NHE containing 1.2 M NaCl, 30 mM HEPES-KOH, pH 7.5, and 15 mM EDTA at 72 °C for 30 min followed by a slow cool to room temperature for 80 min. The beads bound to target tRNA were then washed three times each with 3 X NHE, 1 X NHE and 0.1 X NHE until its UV absorbance (260 nm) fell below 0.01. The target tRNA was eluted from the beads with 0.1 X NHE at 65 °C for 5 min three times.

2.6.RNA Digestion

Purified *P. falciparum* RNA was hydrolyzed enzymatically as described with a slightly modified protocol using following components in the buffer mix at final concentrations of 10 mM Tris (pH 7.9), 1 mM MgCl₂, 5 units Benzonase (99% Purity Merck), 50 µM Desferroxamine (Sigma), 372.2µM (0.1 µg/µL) Pentostatin (Sigma), 100 µM Butylated hydroxytoluene (Sigma), 2 mM (0.5µg/µL) Tetrahydrouridine (Calbiochem), 5 units Bacterial Alkaline Phosphatase (Thermo Fischer) and 0.05 units Phosphodiesterase I (Affymetrix) and [¹⁵N]₅-2'-deoxyadenosine ([¹⁵N]-dA) as the internal standard to account for variations in sample handling and MS response [143]. For the mapping studies, purified tRNA (500 µg) was digested using 2 units RNase T1 (Thermo) in RNA structure buffer for 2 hours at 37 °C followed by 60 °C for 4 hours. The enzymes were cleaned up using a 10 kDa spin filter (Omega Nanosep) and the sample was desalted and concentrated using C18 ziptips (Sigma).

2.7.LC-MS/MS for Quantitation and Characterization of

Ribonucleosides

Purified *P. falciparum* tRNA from three biological replicates of selected time points were hydrolyzed enzymatically as described previously [90, 93, 97, 133, 143]. [¹⁵N]₅-

deoxyadenosine ([¹⁵N]-dA) was added to the digested tRNA samples as an internal standard to account for sampling and machine variations. Hypersil GOLD aQ column (100 x 2.1 mm, 1.9 μm, Thermo Scientific) was used to resolve the digested ribonucleosides in a two-buffer eluent system, with buffer A consisting of water with 0.1% (v/v) formic acid and buffer B consisting of acetonitrile with 0.1% (v/v) formic acid. All solvents used were LC-MS grade. HPLC was performed at a flow rate of 300 μL/min at 25 °C. The gradient of acetonitrile with 0.1% (v/v) formic acid was as follows: 0-12 min, held at 0%; 12-15.3 min, 0%-1%; 15.3-18.7 min, 1%-6%; 18.7-20 min, held at 6%; 20-24 min, 6%-100%; 24-27.3 min, held at 100%; 27.3-28 min, 100%-0%; 28-41 min, 0%. The HPLC column was directly connected to an Agilent 6490 triple quadrupole mass spectrometer (LC-MS/MS) with ESI Jetstream ionization operated in positive ion mode. The voltages and source gas parameters were as follows: gas temperature- 50 °C; gas flow-11 L/min; nebulizer- 20 psi; sheath gas temperature 300 °C; sheath gas flow- 12 L/min; capillary voltage-1800 V and nozzle voltage- 2000 V.

2.8.LC-MS/MS for Mapping tRNA Modifications

Amide-HILIC chromatographic separation was performed on a TSK-gel Amide-80 column (2.0 mm ID x 150 mm, 3 μm particle size) using a binary solvent system consisting of 8mM ammonium acetate in ultrapure water (solvent A) and acetonitrile (solvent B). The flow rate was set to 0.1 mL/min. The gradient of solvent A was as follows: 0-2 min, held at 10% (v/v); 2-3 min, 10%-15%; 3-5.5 min, 15%-30%; 5.5-20.5 min, 30%-60%; 20.5-25 min, 60%- 70%; 25-28 min, 70%-10%. The HPLC column was maintained at 50 °C. The parameters for the QToF mass spectrometer were as follows: gas temperature 325 °C, gas flow 8L/min, nebulizer 30 psi, V_{cap} 3800 V, Fragmentor 250 V, Skimmer 165 V, octopole RF peak of 750. Targeted

MS2 performed every 5 V from collision energies from 25 to 45 V and products scanned from 100 to 1500 *m/z*.

2.9. Protein Extraction and Digestion

Parasite infected red blood cells were pelleted at 2200 $\times$ g for 4 min at ambient temperature and subsequently lysed with ice-cold 0.15% (w/v) saponin (Sigma). The lysate was thoroughly mixed using an aspiration pipette and incubated on ice for 10 min to ensure complete lysis of the RBC membrane. The parasite pellet was centrifuged at 4000 $\times$ g for 10 min at minimum deceleration and washed twice with ice cold PBS and frozen at -80 °C. The parasite pellet was resuspended in 6x volume of 8M urea containing 1 mM sodium orthovanadate and homogenized using a sonicator pulse for 3 min at 25% amplitude and 2 sec on, 3 sec off pulse time. The lysate was spun at 16,000 $\times$ g at 4 °C for 30 min to pellet the insoluble fraction and the lysate was transferred into a new tube. Protein (100 μ g) was reduced with 10 mM DTT at 56 °C for 1 h and followed by reduction using 100 mM IAA for 1 hour in the dark. This solution was diluted to 1 M urea and digested with 2 μ g Trypsin (Thermo) overnight at ambient temperature. The resulting peptides were desalted using Pierce desalting columns as per manufacturer's instructions. These peptides were reconstituted in triethylammonium bicarbonate (TEAB) and labelled using tandem-mass-tag TMT labels (Thermo) as per manufacturer's instructions. The labelled peptides were combined, dried, and reconstituted in 0.1% FA. After checking for labeling efficiency, these peptides were then further fractionated using high-pH fractionation columns (Pierce) as per manufacturer's instructions into 8 fractions.

2.10. Protein Crosslinking and Immunoprecipitation

P. falciparum PfMT-A70-HA (n=3 replicates) expressing parasites under drug pressure 2.67 nM WR99210 (Jacobus Pharmaceuticals) and wild-type *P. falciparum* 3D7 were synchronized using Sorbitol. After 36 hours, the parasites were harvested after Percoll (Sigma # P4937) enrichment, washed twice with RPMI and crosslinked with 0.5 mM dithiobissuccinimidyl propionate (DSP) (Thermo Fisher # 22585) in PBS for 60 min at 37 °C. The reaction was quenched with PBS containing 25 mM Tris-HCl. These trophozoites were lysed with RIPA buffer (10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1% SDS, 1% Triton) containing protease and phosphatase inhibitor cocktail (Thermo Fisher # 78440). The lysates were cleared by centrifugation at 16,000 $\times$ g for 10 min. Supernatants were incubated with 25 μ L of anti-HA Dynabeads (Thermo Fisher # 88836) overnight at 4 °C. Beads were magnetically isolated and washed five times with 1 mL RIPA buffer following five washes with 1 mL PBS and one wash with 1 mL of 100 mM Ammonium Bicarbonate (Sigma # 09830). The beads were reduced with 10 mM dithiothreitol (Sigma # D9779), alkylated with 55 mM iodoacetamide (Sigma # I1149) and subjected to on-bead digestion using 2 μ g of trypsin (Thermo Fisher # 90059). The resulting peptides were desalted using C18 cartridges (Thermo Fisher # 89852) and sent for MS-analysis.

2.11. LC-MS/MS Analysis for Proteomics

Peptides were separated by reverse phase HPLC (Thermo Easy nLC1000) using a precolumn (Thermo) and a self-pack 5 μ m tip analytical column (15 cm of 5 μ m C18, New Objective) over a 140 min gradient before nanoelectrospray using a QExactive HF-X mass spectrometer (Thermo). The mass spectrometer was operated in a data-dependent mode. The parameters

for the full scan MS were: resolution of 70,000 across 350-2000 m/z , AGC $3e^6$, and maximum IT 300 ms. The full MS scan was followed by MS/MS for the top 10 precursor ions in each cycle with a NCE of 28 (34 for TMT samples) and dynamic exclusion of 30 s. Raw mass spectral data files (.raw) were searched using Proteome Discoverer (Thermo) and Mascot version 2.4.1 (Matrix Science). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 15 mmu for fragment ion mass tolerance; 2 missed cleavages of trypsin; fixed modification was carbamidomethylation of cysteine; variable modifications were lysine-labelled TMT residues, peptide N-terminal TMT labels, methionine oxidation and serine, threonine and tyrosine phosphorylation. Only peptides with a Mascot score greater than or equal to 25 and an isolation interference less than or equal to 30 were included in the data analysis. For TMT samples, a minimum abundance of 500 was used a threshold in order to ensure robustness of data.

2.12. Data Processing and Statistical Analysis

Abundances of RNA modifications were normalized to canonicals rA, rU, rG, and rC to account for total RNA amount injected. These were then transformed to $\log_2$ ratios of modification levels in each time point/dosage relative to either an arbitrary average/untreated control across all samples, respectively. Unless otherwise stated, all data are represented as mean $\pm$ SD. Proteomics data for each treatment was analyzed by principal component analysis (PCA) using non-linear iterative partial least squares (NIPALS) algorithm. For interpretations of the relationships between codon usage (codon frequency) and up-regulated and down-regulated proteins at different time points, PLS-R was performed using NIPALS algorithm. The values of codon usage in synonymous codon choices of those proteins were retrieved from the pre-calculated genome-wide codon usage as provided by the Begley Lab. The fold change values from the proteomics data were used as input for the response variable in the PLS-R

analysis. The codon usage was charted out as the predictors. Outliers which could cause over-fitting were removed by manual inspection of residual sample variances and leverages, as well as Hotelling's T2 statistics. Marten's uncertainty test with optimal number of PCs was used to cross-validate the PLS model. Eigenvector based multivariate statistics was performed using UnscramblerX (v10.3, Camo).

Chapter 3 - Results I - tRNA Wobble Modifications mcm⁵U and mcm⁵s²U govern the *Plasmodium* ‘Translatome’

3. Overview

Gene regulation in *Plasmodium* parasites is poorly understood. Transcriptional profiling studies suggest a coordinated cascade of genes as the parasite progresses through its developmental lifecycle. A majority of genes display a cyclical pattern in their transcript abundance with just one expression peak. These peaks tend to correspond with phenotypic needs of the cell lending to a “just-in-time” transcriptome. However, upon different drug stresses like chloroquine and several anti-folates, *P. falciparum* parasites respond by induction of a generic stress response pathway reflecting a transcriptome that is “hardwired”, with limited capacity of feedback [45, 47-51]. Contrasting, but not mutually exclusive studies show that *P. falciparum* can change its transcriptional profile to reflect the induction of functionally relevant genes upon metabolic stress or treatment with small-molecule inhibitors [53, 144, 145]. However, this modulation of the transcriptome does not seem to be largely controlled at the transcriptional level as evidenced by the dearth of transcription factors that have been identified in other eukaryotes. Complementing all of these studies are those that compare the *Plasmodium* transcriptome and proteome through its lifecycle which find that the steady-state mRNA levels do not strictly correspond with the steady-state protein levels [55-57]. Collectively these studies point to an important role for post-transcriptional or translational control mechanisms in regulating gene expression during parasite development.

Among candidate translational control mechanisms operant in other eukaryotes, the modified ribonucleosides of the “epitranscriptome” have not been extensively explored in *Plasmodium* parasites in spite of comparative genomics analyses of tRNA-modifying enzymes. Recent efforts have described a translational regulatory mechanism in eukaryotes and prokaryotes involving growth- and stress-specific reprogramming of post-transcriptional modifications on individual tRNAs linked to selective translation of codon-biased gene families. The nuclear genome of *Plasmodium* encodes 46 tRNA genes coding for 45 tRNA non-redundant isoacceptors (initiator methionine and the elongator methionine are encoded by two different genes) [43, 71]. While these encoded tRNA genes are similar in terms of sequence and structure to other eukaryotic organisms, it is remarkable that only one gene copy exists per tRNA isoacceptor in *Plasmodium* denoting the smallest set of tRNA genes for an eukaryotic cell as opposed to multiple gene copies in other eukaryotes. This unique characteristic of *Plasmodium* highlights the potential importance of tRNA modifications in a complex regulatory network to accurately and efficiently decode 61 codons by 45 cytoplasmic tRNA genes each encoding for a single isoacceptor.

3.1.Preliminary Data

The first survey of tRNA modifications that exist in *Plasmodium* parasites was carried out by Dr. Ng Chee Sheng (Figure 9). As part of this study, it was discovered that 28 *Plasmodium* tRNA modifications are similar to the ones present in eukaryotic organisms and a complete list can be found as Appendix I. These modifications were subsequently quantified across each stage of the 48-hr *P. falciparum* IDC. Effectively, tRNA was isolated from six periodic time points representing distinctive life stages of the parasite TP 8 hours – early ring; TP 16 hours – middle ring; TP 24 hours – ring-to-trophozoite transition; TP 32 hours – trophozoite; TP 40 hours – trophozoite-to-schizont transition; TP 46 hours – mature schizont. Concurrently,

samples were harvested for the exact same stages to carry out a proteomics experiment. The main findings from these experiments can be summarized as follows.

1. As the parasite matures from ring to trophozoite to schizont stages, 22 of the 28 modifications show a synchronous rise in levels relative to the tRNA concentration in the cell across the lifecycle. Of these, mcm⁵U and mcm⁵s²U are noteworthy since they are positions found on the wobble position of the tRNA and become a central part of our hypothesis. (Figure 9A and B)
2. Multivariate statistical analysis of codon usage patterns indicates that up-regulated proteins at the late stage of the parasite lifecycle (40 h and 46 h) (Figure 9C) are significantly different in their codon utilization as opposed to proteins at earlier time points (8h-36h). Specifically, late-stage proteins show a preference for specific synonymous codons for 13 amino acids. Of these, GlyGGA, GluGAA, and GlnCAA codons are important to note since the cognate tRNA isoacceptors are known from literature to be wobble-modified: tRNA^{Gly(UCC)}, tRNA^{Glu(UUC)}, and tRNA^{Gln(UUG)} contain either mcm⁵U or mcm⁵s²U bases.(Figure 9D)

Figure 9 Preliminary work on *Plasmodium falciparum* tRNA modifications

Past work examined the landscape of the tRNA modifications as the parasite progressed through the lifecycle. (A) The heatmap depicts relative fold-change levels of all tRNA modifications to an arbitrary average over the lifecycle of the parasite. (B) tRNA modifications cluster in two different patterns as which are those that increase (mcm⁵U), decrease (m⁶A) or fall in no distinctive cluster (Am) from the ring-stage to schizonting. (C) Proteins at 46h of the parasite lifecycle with high-translational efficiency (HTE) distinguish

from low translational efficiency (LTE) proteins in with regard to their codon-usage in a principal least squares regression (PLS-R) model. (D) Proteins enriched in the late stage of the parasite have an over representation of GlyGGA, GlnCAA, and. GluGAA codons which are presumptively decoded by wobble modified tRNAs, mcm⁵ and mcm⁵s². Figure reprinted with permission from [1].

3.1.1. Limitations and Challenges

Observations from our preliminary data set up a rather tempting hypothesis whereby these modified tRNA are able to decode their cognate codons far more efficiently and can be linked to the enhanced translation of stage-specific codon-biased proteins. However, there exists a crucial gap in this study. While an excellent methodology for quantifying global tRNA modification changes, the technique relies upon the use of an indiscriminate endonuclease like Benzonase to chop up the tRNA into individual nucleotides prior to injection onto the mass spectrometer (Figure 10A). This means that we are unable to distinguish the specific tRNA isoacceptor as well as the positional information on the tRNA for that particular modification.

Figure 10 LC-MS approaches to analysing tRNA modifications

(A) Current methodology employed to catalogue and quantify tRNA modification changes involves hydrolysing tRNA into individual nucleosides and thus reflects only a global change. (B) Nucleotide mapping by base-specific endonucleases enables attributing tRNA modifications to a specific isoacceptor and also provides positional information. Adapted from [76] with permission.

There exist several challenges in being able to distinguish and quantitate individual tRNA isoacceptors and their respective modifications from a pool of tRNAs. Methods such as polymerase chain reaction (PCR) and next-generation sequencing (NGS), while extremely

sensitive in quantifying RNA, are subject to inherent biases in adaptor ligations and polymerase fall-offs as a result of modifications [76]. Microarrays are an alternative approach that yield tRNA identification and possibly modification status [146]. There also exist chemical biology approaches, such as DNA probe hybridization [147] and two dimensional gel purification [148] which require tremendous amounts of starting material, are immensely labour intensive, and still provide an indirect signal for modification status of the tRNA.

In contrast, mass spectrometry based approaches have precedence in their ability to unearth the complex modification status of tRNAs. We applied a technique which is widely used for shotgun sequencing in proteomics. A technique originally developed in the 1970s [149], ribonuclease fingerprinting, utilized endonucleases which cleave only at specific bases which is analogous to trypsin in proteomics which cleaves only at Arginine residues (Figure 10B). For example, RNase T1 which cleaves the 3' of any unmodified guanosine was used to digest a lengthier RNA molecule. The digestion products were then end-labeled using radioactive phosphorous and the resulting bands were detected by radioisotope imaging. An unusual migration pattern was indicative of a modification on a nucleoside in T1. Thirty years after this technique, McCloskey et al came up with a mass-spectrometry protocol which dispensed with the requirement of radiolabeling and became a current standard for sequencing nucleic acid molecules [150].

3.2.Experimental Strategy

There were two convergent approaches employed to resolve the modification status of a particular tRNA isoacceptor. The first approach utilized an oligo-based isolation of a specific tRNA isoacceptor. This would significantly reduce sample complexity from a pool of 46 tRNA isoacceptors. Subsequent hydrolysis of single tRNA isoacceptors into individual nucleosides

using previously established methods would yield identification and quantitation of these modifications. While sacrificing positional data, these modifications could in the very least be attributed to specific tRNAs. Complementarily, in order to map these modifications, a combination of base—specific endonucleases could be utilized to release varying oligos for the tRNA isoacceptor. These released oligonucleotides could then be subjected to amide-HILIC chromatography to resolve and collision-induced dissociation fragmentation on the mass spectrometer to conclusively ascertain the modifications and their locations along the tRNA isoacceptor. The advantage of isolating a single isoacceptor before carrying out our analysis would allow for far greater sensitivity in detection and a greater depth of analysis.

Our second approach dispenses the need for isolating individual tRNA isoacceptors with a bioinformatics strategy that is able to overcome the shortcomings of analyzing a complex tRNA pool. Manually curating oligonucleotide sequences released from T1-digested tRNAs would enable us to predict unique fragments that are non-overlapping with other tRNAs. These fragments would have specific masses and fragmentation signatures that we could experimentally verify using LC-MS. Using this approach, we were able to conclusively resolve two gaps in our study. Specifically, we were able to show that tRNA^{GluUUC} and tRNA^{GlnUUG} contain mcm⁵s²U and tRNA^{GlyUCC} contains mcm⁵U on their wobble positions.

3.3. Isolation of tRNA from *P. falciparum*

The first step to characterizing wobble modification naturally involves isolating high quality of tRNA analyte. Based on previous protocols, I established a rigorous protocol for isolation of *Plasmodium* tRNA prior to LC-MS analysis. We ensured that the cultures were free from mycoplasma and other contaminants since these would not only alter parasite biology, but also show up in our tRNA pool and carry similar associated modifications. After pelleting the

cultures, we used selective lysis with a mild detergent like saponin to release the parasites. This pellet was then homogenized in Trizol and total RNA was extracted using chloroform extraction (Figure 11 Left). The total RNA profile was indicative of quality and potential degradation. Sharp peaks for tRNA (50-90nt), 5S RNA (~120nt), 5.8S rRNA (~ 150 nt), 18S rRNA (~ 2000 nt), and 28S rRNA (~ 4000 nt) were observed as previously reported [151]. This total RNA fraction was then subjected to solid-phase extraction using ethanol gradients to selectively enrich for the small RNA fraction. A 30% ethanol concentration selectively binds RNA species to the silica cartridge that are >100nt in length to the column, with a minor percentage of tRNAs also being lost at this stage. Additional ethanol is then added to the flow through so that the resulting concentration is 70% which is then applied to another silica cartridge. Checking the RNA profile of the elution from both of these cartridges on a microfluidics-based bionalyzer chip allowed us to be stringent in our isolation. In the event that there was an appearance of 5.8s or 5S rRNA in the second cartridge, we further subjected this fraction to size-exclusion based high performance liquid chromatography (Figure 11 Right Down). For most of our isolations, I did not observe a 5.8S or 5S rRNA peak and was able to avoid the HPLC purification, which would have led to further losses in tRNA quantity.

Figure 11 tRNA isolation from total RNA in *Plasmodium falciparum*

(A) Schematic illustrating solid phase extraction for enrichment of small RNA from total RNA. Using a Trizol-Chloroform extraction, total RNA from parasite pellets is isolated in the aqueous phase. Upon addition of ethanol to a final concentration of 35%, ribosomal RNA and large RNA can be isolated using a silica-based spin column (Cartridge 1). The flow-through from this spin cartridge contains small RNA (including tRNA). At 70% ethanol, these RNA species bind to the another silica spin cartridge (Cartridge 2), thus desalting and concentrating only for small RNA species. **(B)** Signals from a ultraviolet diode array detector (UV-DAD) coupled to a HPLC machine **(C)** Bioanalyser gel analysis of tRNA isolated from Cartridge 2. Lane 1 depicts ladder. Lane 3 depicts co-isolation of 5.8S rRNA (150nt) and 5S rRNA (~100nt), which is further subjected to HPLC analysis **(B)** to collect

pure tRNA. Lane 2,5-9 do not require HPLC purification since there is no visible 5S RNA. Lane 4 depicts tRNA post-HPLC purification.

3.4.Approach I - Oligo-based Pulldown for Isolation of Specific tRNA

Isoacceptors

In order to further reduce sample complexity, I attempted to purify targeted tRNA isoacceptors using biotinylated-oligos (Figure 12A). This involved designing DNA baits 30nt in length complementary to the isoacceptors in study. These baits were designed so as not to have any complementary hybridization to the rest of the tRNA pool with a threshold of two mismatches, thus ensuring specificity (Figure 12B). The hybridization details are mentioned in the Materials and Methods Section 2.5. Briefly, the biotinylated baits were first bound to functionalized streptavidin beads. After washing off unbound beads, *Plasmodium* tRNA was hybridized in a high salt buffer at a 72 °C, followed by slow cooling to room temperature. The remaining tRNA was then transferred on to the next set of beads containing a different tRNA isoacceptor bait for hybridization. These beads, containing the hybridized bait with the specific tRNA, were then washed and eluted thrice in a low salt buffer at a high temperature. Upon checking the integrity of tRNA on the Bioanalyzer, I found that in addition to the specific tRNA band, the DNA oligo bait also shed from the beads during the elution step (Figure 12C). Besides, there were also unknown aggregates that were >150nt in length which would interfere with downstream analysis (Figure 12C). The eluted oligo amount was far in excess of the actual isolated tRNA leading us to carry out size exclusion HPLC to isolate only the tRNA. While we were able to remove the aggregates, this strategy was unsuccessful in separating the oligos from the tRNA (Figure 12C Right Lane 5-7). While single nucleoside modification analysis

might yield some qualitative data, this would affect downstream sensitivity and quantitation. Thus, we abandoned this approach and decided on a computational approach that did not require isolating individual isoacceptors.

Figure 12 Oligo-based pulldown approach to isolate tRNA isoacceptors

(A) Schematic displaying tRNA isolation using a biotinylated DNA bait. Taken from [152] under a fair use license for internal publications. While the figure shows a DNA oligo complementary to the 3' end of the tRNA, my analysis found that this region is far too conserved amongst different tRNA species. (B) Instead, our baits were designed to be complementary to the ASL of specified tRNAs. (C) Bioanalyser plots for eluted fractions of oligo-specific tRNA pulldowns. Lane 1 depicts ladder. Lanes 2-4 depict small RNA quality after the first elution. Note strong signals for aggregates >150nt and DNA oligos ~20nt and weaker signals for tRNA isoacceptor ~75nt. Lanes 5-7 20nt DNA oligos still present after subjecting pulldown tRNA to HPLC purification.

3.5. Approach II - RNA Fingerprinting

3.5.1. Computational Approaches

In order to analyse modification status of specific isoacceptors in a pool of tRNAs, we would need to utilize RNA fingerprinting and sequence analysis using LC-MS. This would require us to know the predicted masses of the released tRNA oligos *a priori* which would enable targeted analysis on the MS. I first downloaded all genomic tRNA sequences from the PlasmoDB database. These tRNA genes are listed in Appendix II. Using MongoOligo , I carried out a theoretical digest for both RNase U2 and T1 enzymes for Glutamate (Glu), Glycine (Gly), and Glutamine (Gln) tRNAs and discovered that both the enzymes would both be suitable for mapping the modification on the anticodon loop (Figure 13). U2 RNase selective cuts single stranded RNA at adenosines whereas RNase T1 cleaves at guanosine residues [149, 153, 154]. It breaks the phosphodiester bond between the 3'-guanylic residue and the 5'-OH residue of the neighbouring nucleotides. Despite initially optimising for both the enzymes, we decided to proceed with RNase T1 as a matter of resource availability since RNase U2 was no longer produced by Sigma. These digested sequences were BLASTed against the *Plasmodium* tRNA database to ensure that there was no overlap and the products generated would be unique in their mass signatures.

Figure 13 Bioinformatic analysis for mapping tRNA fragments

(A) Flowchart for targeted mapping of tRNA modifications on isoacceptors. (B) Theoretical digests using enzymes T1 and U2 yield putative unique fragments that are amenable to identification on LC-MS.

3.5.2. Optimization of Liquid-Chromatography Mass Spectrometry (LC-MS)

3.5.2.1. Chromatography Conditions

To be able to map the different RNA fragments, it was necessary to optimize LC-MS conditions using relevant standards. I purchased RNA oligos from IDT representing the unmodified digested and dephosphorylated sequence in order to optimize the LC-MS parameters. Chromatography for oligonucleotides without trityl derivatization is generally done using

reverse-phase ion-pairing reagents in the mobile phase [155]. Ion pairing reagents neutralize the negatively charged phosphate backbone of oligos and thus allow for excellent hydrophobic-based characteristics on a C18 column. Ion-pairing reagents, as per their name, are also highly charged ions which subsequently interfere with MS detection [156]. These reagents, once on an LC-MS system, are incredibly difficult to remove and cause ion-suppression leading to significant loss in sensitivity. This figured into our decision to use ammonium acetate as a salt which would serve the same purpose and is amenable to use on a mass spectrometer without compromising sensitivity. We used a TSKgel Amide-80 column where silica beads with functionalized amide groups serve as the stationary phase. The hydroxyl groups on the sample and the carbamoyl group in the stationary phase interact in a unique way based on hydrogen-bonding under normal phase conditions. 8mM ammonium acetate was used as Solvent A in the mobile phase since it has been shown that the presence of salts improves separation in HILIC chromatography. Acetonitrile was used as Solvent B.

3.5.2.2. Mass Spectrometer Optimization

The mass spectrometer used for all analysis work was either an Agilent Q-ToF 6530 or Agilent QqQ 6490. Source parameters such as sheath gas temperature and flow, nozzle voltage, gas temperature, fragmentor voltage, and nebulizer flows were optimised for maximum ionization with minimal secondary fragmentation of the analyte. I also optimized the collision energy for most efficient MS² fragmentation. I found that in order to get the full fragmentation spectra, it was necessary to use multiple collision energies and average the spectra. All of these optimizations were carried out using RNA standards representing unmodified T1 digests of the specific tRNA that were purchased from IDT. Also, there exists no a priori knowledge on the charge state on the digested oligo and could only be best estimated from running the standards. Unlike MALDI, where the ionisation results in generation of only a single charge state, ESI

results in multiple charges on each species for larger molecules. Using Mongo Oligo, we were able to calculate exact masses for different charge states of the oligos (Figure 14A). We were also able to account of sodium or potassium adducts that might form as a result of the digestion or chromatography conditions (Figure 14A). Charge state assignment for each m/z peak is possible because high-resolution MS (HRMS) is able to distinguish between different isotopic peaks. A difference of 0.5 m/z between two isotopic peaks would indicate that the oligo carries a charge of -2, whereas a mass difference of 0.33 m/z would indicate a charge state of -3. From our experiments and previous work in the lab, oligos that are 9-12 nucleotides in length would be visible with a -3 or -4 charge state where the -3 charge state is the predominant one. Oligos that are shorter (~6 nucleotides) carry either -2 or -3 charge on ESI-MS, where the -2 charge state is the predominant one (Figure 14B).

A

Sequence :

CCU GCC UUC CA

5'OH - RNA[11mer] - 3'OH

C:6 U:3 A:1 G:1

monoisotopic mass, negative mode

ELECTROSPRAY SERIES

ch	m/z	M+Na	M+K	M+2Na
-1	3359.456	3381.438	3397.412	3403.420
-2	1679.224	1690.215	1698.202	1701.206
-3	1119.147	1126.474	1131.798	1133.801
-4	839.108	844.603	848.597	850.099
-5	671.085	675.481	678.676	679.877
-6	559.069	562.733	565.395	566.396
-7	479.058	482.198	484.480	485.339
-8	419.050	421.797	423.794	424.545
-9	372.377	374.819	376.594	377.262
-10	335.038	337.236	338.834	339.435

B

Figure 14 Determining charge state for RNA oligo standards

(A) Predictive charge states for tRNA digest fragments accounting for the formation of sodium (Na) and Potassium (K) adducts. (B) Empirically ascribing charge states to RNA oligos representing T1 digested unmodified tRNA standards GlnCAA, GluGAA, and GlyGGA with a 100pmol injection on Agilent 6530 q-ToF.

3.5.3. tRNA Hydrolysis and Desalting

tRNA isolated from unsynchronized *P. falciparum* cultures was digested using T1 ribonuclease. A denaturing buffer (provided by Thermo as Sequencing Buffer) containing 7M Urea at 60 °C was used in order to avoid secondary structure formation which would be resistant to digestion by RNase T1. These released oligos were then dephosphorylated using bacterial alkaline phosphatase (BAP). This step would help with the chromatography and avoid the complex interpretation of oligos that may contain a mix of cyclic intermediates. Enzyme removal was carried out using a 10kDa centrifugal filter and the flow-through was dried and reconstituted in 1M triethylammoniumacetate (TEA-A). TEAA serves as an ion pairing reagent to neutralize the phosphate backbone and effectively bind to the C18 Ziptip, allowing for effective desalting of the samples using solid-phase extraction (SPE).

3.5.4. Mapping of Wobble Modifications

3.5.4.1. Identifying Modified Oligos at MS¹ Resolution

With all parameters optimized, we injected T1-digested *Plasmodium* tRNA onto the LC-MS. An extracted ion chromatogram (EIC) for both charge states of the parent ion for the unmodified oligos matched perfectly with our standards. Identical masses and retention times are two sufficient (but not definitive) features to identify two independent oligos as the same. Possessing no standards for the modified oligo, the EIC for predicted masses showed a slightly earlier shift in the retention time for T1-digested tRNA^{GluGAA} and tRNA^{GlnCAA}, but not for tRNA^{GlyGGA} (Figure 15). We also observed the signal intensity was higher for the modified oligos, as opposed to the unmodified oligos. With the assumption that the ionization potentials are close to similar for these oligos, these precursor abundances are a reflection of abundance

in the tRNA pool. Upon carrying out deconvolution for these precursor tRNA fragments, we were able to get an accurate mass and compare it with a calculated mass all within a 3ppm difference (Figures 15, 16). We then carried out targeted MS² for these precursors that were identified in our first preliminary MS¹ scan. This fragmentation allowed us to ascertain sequence and positional information for both the unmodified and modified oligo.

tRNA species	Sequence of RNaseT1 fragment	Calc'd Mass	Obs'd Mass	Diff'nce (Δppm)	HPLC Ret'n time (min)	Negative Ions		Charge State
		(Mono-isotopic)	(Deconvoluted)			Calc'd	Obs'd	
Gln-CAA	ACUUUG	1835.259	1835.264	2.724	14.343	610.745	610.751	-3
Modified Gln-CAA	ACUmcm ⁵ s ² UUG	1923.257	1923.262	2.599	13.64	916.621	916.635	-2
Glu-GAA	CUUUCACCCG	3055.423	3055.425	0.654	15.726	640.078	640.075	-3
Modified Glu-GAA	CUmcm ⁵ s ² UUCACCCG	3143.421	3143.429	2.545	16.248	960.621	960.623	-2
Gly-GGA	CCUCCAAG	2773.409	2773.402	-2.5239	15.623	762.848	762.951	-4
Modified Gly-GGA	CCUmcm ⁵ UCCAAG	2845.43	2845.437	2.46	15.623	1017.466	1017.474	-3
						784.847	784.851	-4
						1046.799	1046.801	-3
						923.462	923.462	-3
						692.344	692.346	-4
						947.469	947.141	-3
						710.349	-	-4

Figure 15 Electrospray series for unmodified and modified tRNA species presenting summary data for the T1 digested Plasmodium tRNA

The table depicts the theoretical mass and the observed mass of the modified and unmodified oligos obtained by deconvolution. The major charge states for these fragments was determined by the isotopic peaks. The -4 charge state for the mcm⁵U modified tRNA^{GlyGGA} was not identified in the analysis. Figure reprinted with permission from [1]

Figure 16 Deconvoluted masses of tRNA fragments for tRNA isoacceptors (top)

$tRNA_{Gln}^{CAA}$, (middle) $tRNA_{Gly}^{GGA}$, and (bottom) $tRNA_{Glu}^{GAA}$

T1 tRNA isoacceptor fragments depicted in red (left) were deconvoluted using Agilent MassHunter software to yield observed mass to compare with a theoretical mass. These observations are listed in Figure 15. The split spectral graph indicates different retention times for the modified (red) and unmodified (blue) oligo fragments. Unboxed peaks are coeluting fragments that were not characterized. Figure reprinted with permission from [1]

3.5.5. CID fragmentation

3.5.5.1. Optimization of RNA Standards

In their early work, McLuckey et al have described the nomenclature which is well in use today [134]. These are based upon the four possible cleavage sites along the phosphodiester linkage. They are illustrated by lower case letters a,b,c, and d for fragments that contain the 5'-hydroxyl group and w,x,y, and z for the fragments that contain the 3'-hydroxyl group (Figure 17A). The subscripts are indicative of the number of bases from the respective termini. With a large acquired database for CID fragmentation of RNA oligos, it has been well documented that w1, w2, c1, c2, and y2 fragments typically show upon CID fragmentation as these are the most labile bonds. The other fragments are more or less oligo-dependent and dissociation energy dependent. There also might exist internal fragments for these oligos, where a subset of the terminal nucleotides have been separated as a result of the fragmentation. A predictive set of fragment ions and their masses can be generated using MongoOligo (<https://mods.rna.albany.edu/masspec/Mongo-Oligo>) (Figure 17C). The proportionate amount of collision energy used in the dissociation cell of the MS determines the fragment ions that will be generated by an oligo of a certain length and composition. This was determined upon fragmenting the standards purchased whereby a good representation of the fragment ions are visible (Figure 17B). For example, for the -3 charge state for tRNA^{GlyGGA} having a precursor

m/z 1119.48 upon fragmentation at 20V poorly fragments the precursor ion leading to almost no product ions (Figure 17, Spectrum 4). However, at a higher collision energy of 45V (Figure 17, Spectrum 6), a gamut of product ions are visible corresponding to the theoretical fragments in the inset with w1, w2, and y2 having m/z 346.054, 651.059, and 571.129 being the most abundant fragments. For the higher charge state of -4, the precursor m/z 839.36 fragments better at a lower collision energy of 30V (Figure 17 Spectrum 1) producing some overlapping but some distinct product ions. At 45V (Figure 17 Spectrum 5), most product ions are present <400 m/z and there is no precursor ion in the spectrum, suggesting over fragmentation of the oligo for the -4 charge state.

Figure 17 Collision-induced dissociation (CID) fragmentation of RNA

(A) Nomenclature for oligonucleotide fragmentation as proposed by McLuckey et al [157]. The CID fragmentation occurs along the phosphate backbone along either side of the bridging oxygens. (B) Fragmentation of synthetic $tRNA^{GlyGGA}$ standard at multiple collision energies for both the charge states is needed to obtain a good coverage of the theoretical spectrum shown in (C) and generated from Mongo Oligo.

3.5.5.2. T1-digested *P. falciparum* tRNA

Having optimised for a well-represented product ion spectra for our standards, we proceeded to analyse the T1-digested *P. falciparum* tRNA. The focus of this analysis was the qualitative validation of the presence of predictive modifications on specific tRNA isoacceptors i.e. targeted MS analysis. The qToF MS was configured to fragmented targeted precursor masses at specific retention times as observed in Figure 15 for multiple collision energies. While seemingly close in their retention times, the chromatographic peak width for these eluting oligos is narrow enough to carry out this fragmentation for multiple precursor ions without having to carry out a second injection. Using Agilent Masshunter for Qualitative Analysis, the resulting product spectrum was manually curated and aggregated to obtain a spectrum from which we could sequence the RNA oligo. In order to aid our interpretation of these spectra and dispense with noise, all ions below either a relative threshold to the largest peak or an absolute threshold were not recorded in our analysis. From our predictive analysis, c1,c2, and w2 ions were identified in all three of our experimental spectra. We were confidently able to assign mcm⁵s²U to the wobble position of the anticodon loop for tRNA^{GlyCAA} and tRNA^{GlnGAA} and mcm⁵U to tRNA^{GluGGA}. We also observed a base loss for the mcm⁵s²U modification but not for mcm⁵U (Figure 18).

Figure 18 *MS² fragmentation of modified tRNA isoacceptors*

Collision-induced dissociation of the parent molecular ions for (Top) m/z 1046.477 from Glu^{tRNAGAA}, (Middle) m/z 960.621 from Gln^{tRNACAA}, and (Bottom) m/z 947.465 from Gly^{tRNAGGA}, produced expected -c, -w and -y ions for the sequences CU[mcm⁵s²U]UCACCCG, ACU[mcm⁵U]UG, and CCU[mcm⁵U]CCAAG respectively which matched to the prediction offered by MongoOligo. Figure reprinted with permission from [1]

3.6. Quantification of Mapped Modifications Across the Lifecycle

Taking this work forward required quantifying these mapped modifications across the lifecycle. The amount of starting tRNA material for these analysis was prohibitive to take this experiment across the 6 timepoints that was carried out in the preliminary study. As a result, We decided to broadly collect ~500 μ g of tRNA representing the three phenotypically distinct stages of the parasite lifecycle – namely ring (10-16H), trophozoite (28-34H), and schizont (38-44H) as described in methods section 2.3. We hydrolysed this tRNA from two biological replicates using RNase T1 as described in methods section 2.6. We utilized the exact LC method described in methods section 2.8 and adopted it to be used on the Agilent 6490 QQQ. We used two of the most abundant fragment ions (either of w1, w2, c1, c2, y2) as evidenced by our qToF spectra in order to quantify in the MRM mode on a Agilent 6490 QQQ. The quantitation of the modified and unmodified fragments for all three isoacceptors were carried out in a single run by calculating the area under the curve for each chromatogram. Our results show that the fraction of each tRNA isoacceptor modified with either mcm⁵U or mcm⁵s²U increases across the IDC, which is consistent with both the population increase in these modifications (Figure 19). For the ring stages, 10% of tRNA^{GlnCAA} and tRNA^{GlyGGA} are modified and increase to 33% in the trophozoite stage of the parasite. For the tRNA^{GluGAA}, a

more dramatic increase of 67% is observed. On the whole, these numbers correspond with the trends depicted in Figure 9 which is for the global tRNA population.

Figure 19 Quantifying wobble-modified tRNA isoacceptors

Ratios of modified oligonucleotides from the anticodon stem-loop of (A) GlnCAA, (B) GluGAA, and (C) GlyGGA isoacceptors are expressed as percentages of their sum total. The experiment was carried out in two biological replicates and two additional technical replicates. The error bars represent deviation from the mean values for the biological replicates. Figure reprinted with permission from [1].

3.7.Discussion

tRNA are adaptor molecules that are an integral part of a cell's machinery which aid in translating DNA-transcribed mRNA messages into proteins. While there is extensive documentation of modifications that are present on tRNAs, there is a severe dearth of functionally ascribing a role to each of these modifications [158]. An observation that arises from multiple studies is that positional information on the tRNA is just as important as the chemical moiety represented by each modification. Chemical and physical homogeneity in the secondary and tertiary structures of tRNAs make it a challenge to separate these molecules in a complex pool using available methods. Here we used an enzymatic approach to yield unique

fragments in a tRNA pool with signature masses, to identify and positionally assign modifications on specific tRNA isoacceptors.

Our first approach of isolating individual tRNA isoacceptors using biotinylated oligos was met with limited success. While we were able to isolate independent tRNAs, a large proportion of our sample contained the bait oligos that were used. Not only would these synthetic oligos interfere with RNase digestion, as they would be competing substrates, they would also interfere with subsequent MS detection leading a significant loss in sensitivity. Previous work described by Suzuki et al, have demonstrated a similar method whereby biotinylated-DNA oligos are immobilized on streptavidin columns in a daisy-chain and tRNAs are circulated through this “chaplet” in hybridization conditions [152]. Using this method, they were able to isolate several species of Bovine tRNAs from a complex pool. However, a severe drawback for this method was that they used 5 kg of bovine liver which translated into ~600 mg of tRNA (30 mL of 20 mg/mL tRNA extract). These large-scale preparations for *Plasmodium* tRNAs would be immensely difficult to obtain since a modest 3 mL of packed blood translates into ~20-50 μ g of tRNA. While there exist several other protocols which have used a similar approach, their final evaluation of tRNA species has only been by northern blots or specificity for aminoacylation, thus possibly missing the shed oligos or aggregates that might be present in the resulting eluting mixture [159].

Pioneered by McLuckey and McCloskey laboratories, our second approach adopted a ribonuclease mapping strategy coupled to LC-MS analysis. Even though the absence of modified nucleoside standards represents a potential hurdle, the manual interpretation of MS/MS spectra by an MS expert is the most accurate way of validating precursor and product ion spectra from theoretically determined ions. A limitation of this approach is that the tRNA

isoacceptor population is considered having taken only one modification into account in ASL loop. This may be potentially misleading in accounting if there are other modifications in within this tRNA fragment. There exists a significant computational and experimental challenge if we are to consider other modifications. On attempting to generate a list of putative modifications alongside the wobble modifications, I was unsuccessful in finding a promising hit from the MS¹ scan. A permutation of all possible modifications in the anticodon loop of the T1-digested fragment is an exponential task and should be considered using orthogonal approaches. There are several computation tools that have been developed to address this issue. Simple Oligonucleotide Sequencer (SOS), developed by Rozenski and McCloskey [160], OMA and OPA developed by Nyakas et al [161], and Robo Oligo developed by Limbach et al [162] are some of the computational tools which might aid in being able to handle these immense computational tasks. However, while some of them are not able to process batch data that is generated by the mass-spec, others are limited to single sequence analysis. More recently developed by Limbach et al, RNAModMapper (RAMM) represents a promising alternative to *de novo* sequencing of modified tRNA fragments [163].

While the technical challenges of this project might be of interest, the biological phenomena of tRNA biology and what they represent for parasite physiology forms a much sought-after question. This chapter takes a perspective on possible regulatory mechanisms that might function by modulating tRNA molecules. In other eukaryotes, not only are there multiple gene copies of the same tRNA isoacceptor, but there exist isodecoders. While isoacceptors decode different codons for the same amino acid, isodecoders decode the same codon for the same amino acid with small variation in the rest of tRNA sequence. This natural difference in copy numbers in tRNA genes would effectively relate to translational efficiency and genomic codon bias. However, there is sufficient evidence in literature to see that optimally codon biased genes

are not always highly expressed or alternatively sub-optimal codons have functional roles that improves translation and thus, pointing to a gap in the science or a discussion that is beyond the scope of this thesis [164]. *Plasmodium* however, possess only a single set of genes per isoacceptor and has only one set of isodecoders to distinguish the initiator methionine from the elongator methionine. This would follow a near equimolar distribution of tRNAs with a caveat accounting for equitable transcriptional potential. In a codon biased genome, the presence of tRNA modifications would make up for accurate decoding and prevent mistranslation. It has been shown that an overexpression of a single tRNA isoacceptor is enough to rescue from mistranslation effects on account of hypomodification. Yeast strains lacking the elongator complex responsible for mcm⁵s²U modification show an increase in mistranslation and can be rescued by overexpression of single tRNA^{LysUUU} isoacceptor [165]. Overwhelming presence of a certain tRNA in a present a higher opportunity for accurate decoding and not just translational efficiency. There, of course, also exist complementary studies where overexpression of several isoacceptors was not enough to rescue mistranslation defects as a result of their hypomodification [166]. In a highly codon biased *Plasmodium* genome context where there is no such evidence for control of tRNA gene expression at a transcriptional level, we proffer that this system of tRNA modifications compensates for both accurate decoding and translational efficiency by this minimalistic set of tRNA genes.

The absence of both microRNAs in *Plasmodium* and the corresponding machinery (Argonaute/DICER) [167], as identified in other eukaryotes asks the question if there are possibly other small RNA species that could serve a regulatory role. An emerging paradigm of tRNA fragments (tRFs) that could serve as regulators has been found in other organisms [137, 168]. Two papers report the presence of these cleaved tRNA molecules, but neither of the papers have been able to attribute any functionality representing an observational, nevertheless

important study [151, 169]. In other eukaryotes, proteins like DICER and Argonaute which form the miRNA silencing complex have been shown to interact with tRFs [168]. However, these proteins have not been reported in *Plasmodium* so far. There exist two papers that have attempted to study the role of human miRNAs in *Plasmodium* [62, 170]. A study by La Monte et al showed that the import of two specific miRNAs, miR451 and let7i, led to translation inhibition specifically in sickle-cell derived red blood cells. Another study by Dandewad et al show that the parasites import RBC derived miRNA-Argonaut vesicles and alter expression levels of an antigenic variation protein family *Pfempr1* but the evidence remains weak. Both of these works provide little evidence of how this might provide a regulatory framework using a system of tRNA fragments.

Another study in *Plasmodium* documented and characterized a tRNA import protein (tRIP) that aids in the import of host tRNAs into the parasite [171]. However, the authors found that this protein is not essential in the blood stages of the *P. berghei* parasite and the mutant displays a mild replication defect. While they depict the immunofluorescence co localization of tRIP in *P. berghei* parasites for sporozoites, ookinetes, hepatocytes and some of the blood stages, the study documented the import of exogenously-labelled tRNA only in the sporozoite stage of the developmental lifecycle. While the tRIP protein, unequivocally, has a tRNA binding domain and shows interaction with human tRNAs, there remains a key question in the utility and functionality of importing tRNAs into *Plasmodium*.

Two recent studies explore the relationship of codon-bias of mRNAs on gene expression in *Plasmodium* [72, 172]. An interesting observation is that even though nearly 25% of the genes contain asparagine (AAT) homorepeats, the *Plasmodium* genome does not seem to encode for any additional asparagine tRNA gene copies despite of this skewed requirement. In addition,

AsnAAT is decoded by a wobble paired tRNA^{AsnGUU}. Chan et al, using GFP reporter constructs, demonstrate that the use of these wobble decoded codons correlates negatively with protein output. An interesting direction to take this study would be to see if these reporter constructs containing the homo-repeats would have better translational outcomes if enzymes encoding for certain tRNAs or tRNA modifications could be overexpressed. For example, t⁶A as a modification is generally present at position 37 on a tRNA which is adjacent to the first Watson-Crick codon-anticodon pair [166]. Especially for an AsnAAN codon, this would help prevent frameshifting and avoid translational errors and abortions. In addition, it has also been shown that overexpressing certain tRNAs can rescue phenotypes where the modification enzyme has been knocked-out [89, 92].

To the best of our knowledge, this is the first study characterizing tRNA modifications in *Plasmodium falciparum* and lays the groundwork for all future studies [1]. This data contributes to a model where tRNA modification reprogramming coordinates with a system of synonymous codons usage to modulate stage-specific translational needs of the parasite. The wobble position on tRNAs have generally been documented to be modified to generate a sufficient selection of anticodon structures and chemistries that can be utilized for accurate decoding. Specifically, the C2 and C5 on the wobble uridine (U34) nucleoside play a critical role in mitigating translation errors. The two modifications, mcm⁵U and mcm⁵s²U, described in this study are found at the U34 position for tRNAs that decode arginine, glutamine, glutamic acid, and glycine. Wobble base pairing allows for arginine codon AGA and serine codons AGC and AGU to be decoded by tRNA^{ArgUCU}. The presence of the mcm⁵U modification of the tRNA^{ArgUCU} ensures that there isn't promiscuous arginine incorporation at serine codons. U34-modified tRNAs also have been shown to dock at the ribosome A-site more efficiently compared to the hypomodified versions ultimately accounting for the correct reading frame.

In addition to the simple modifications such as methylations, these positions can simultaneously be modified to contain carbonylations and thiolations. For the biosynthesis of mcm⁵s²U, the wobble position U is first modified to cm⁵U with the aid of the Elongator Complex (Elp1-6) and associated enzymes (Kti11-Kti13; Sit4, Sap180/Sap195; Hrr25/Kti14). This is followed by a methylation to mcm⁵U with the aid of another enzyme complex (Trm9/Trm112), and a final thiolation to form mcm⁵s²U using the sulfur transfer system (Urm1, Uba4, Ncs2, Ncs6) [83, 85, 91, 94, 173]. It is important to note, that this pathway is coordinated by distinct enzyme complexes that use different substrates. A deletion of enzymes in this sulfur transfer system show translation defects and consequently elevated proteotoxic stress as documented by ribosomal profiling studies. In a similar vein as mcm⁵U, the presence of mcm⁵s²U on the wobble position serves to restrict the decoding capabilities to A-ending codons as per our hypothesis. This observation fits well in our model where A-ending codons like GluGAA, GlyGGA, and GlnCAA, are significantly overrepresented in late-stage transcripts with higher translational efficiency and demonstrate the need for wobble modified tRNAs for accurate decoding. All of this work is taken in the context where the tRNA isoacceptor population remains relatively unchanged across the lifecycle as measured by microarray data (Figure 20) [174].

Figure 20 tRNA gene expression as measured by microarray studies

tRNA gene expression for tRNAGlu, tRNAGly, and tRNAGln remains unchanged in the three major developmental stages of the parasite lifecycle as measured by microarray expression data. Data mined from [174]. tRNA isoacceptor gene expression level values are shown as a fraction of their total abundance through the entire lifecycle. The microarray experiments were carried out in biological duplicates and the error bars represent standard deviation from the mean. Figure reprinted with permission from [1]

Based on the preliminary data, an integrated analysis of the epitranscriptomic, proteomic, and genomic data in the *Plasmodium* IDC leads to a testable model for translational modulation of gene expression. This system of 28 tRNA modifications undergoes three stage-specific reprogramming events that dictate parasite growth. (i) At the ring stage, there is an up-regulation

of modifications Am and s²U that heralds subsequent tRNA maturation at later IDC timepoints. (ii) There is a ‘rising tide’ of modifications as the parasite transitions from ring-trophozoite signifying tRNA maturation that mirrors the increase in translational and metabolic activities of parasite growth. (iii) Reprogramming of certain wobble modifications on tRNAs at the late-stage of the parasite coordinates with a biased utilization of synonymous codons that up- or down-regulate expression of important and unneeded proteins, respectively (Figure 21). These results lay down a primer for future work on tRNA modifications and display the interplay of tRNA epitranscriptomics, proteome, and genome in the developmental cycle of *Plasmodium* parasites.

Figure 21 Roles of tRNA modifications in asexual blood stage development of *Plasmodium*.

tRNA modifications choreograph parasite development at a translational level in a coherent system: (1) High levels of Am and s²U at ring stage (8H) stabilize the “hypomodified” tRNA molecules or serve as a precursor for the subsequent modifications; (2) At an early trophozoite stage (24H), a synchronous increase in 22 of the 28 modifications serves to signify tRNA maturation. This parallels up-regulation of translational and metabolic activities through the ring-to-trophozoite transition depicted by the graphs showing peaks of synthesis of DNA, RNA and protein (adapted from (de Rojas & Wasserman, 1985)); (3) At the schizont stage of the parasite (46H), increase in the levels of particular wobble modifications enhance the recognition to specific codon(s) out of other synonymous codons. This results in selective translation of codon-biased transcripts which manifests in the codon usages of the late trophozoite and schizont proteomes. Figure reprinted with permission from [1].

Chapter 4 – Results II - Transcriptome-wide dynamics of extensive m⁶A mRNA methylation during *Plasmodium falciparum* blood-stage development

4. Overview

Tightly regulated gene expression dictates parasite progression in the IDC. Each of the mRNAs transcribed reaches peak abundance only once through the life cycle, which is postulated to be reflected in the parasite need for that particular gene product [45, 57]. Chromatin structure studies show that most of the genome is in a euchromatic, transcriptionally active state throughout the IDC [175]. The exceptions are the variant surface antigen genes and those required for sexual differentiation. Up until now, comprised of 27 members, the Apicomplexan AP2 proteins are the only single transcription factor family with defined functions across all developmental stages [176]. Given the genome size, the availability of only such a small repertoire of transcriptional factors seems inadequate. It has been suggested that while transcription may dictate initial messenger RNA (mRNA) abundance, a more granular modulation of mRNA and proteins is greatly influenced by post-transcriptional and translational mechanisms [59, 67, 177]. Recent literature points to how chemical nucleotide modifications positioned internally in mRNA can choreograph post-transcriptional regulation of gene expression.

Most notably, the relatively abundant N⁶-methyladenosine (m⁶A) has been shown to affect a multitude of cellular processes [77, 119, 124, 139, 178-187]. For mammalian cells, this modification is catalyzed by core complex consisting of three proteins – a catalytically active

methyltransferase (methyltransferase-like 13, METTL3), another methyltransferase-like protein (METTL14), and an adaptor protein called Wilms-Tumor-1 associated protein (WTAP) [179]. METTL14, while catalytically inactive, has been shown to allosterically control METTL3 and provide a binding site for the mRNA substrate [188, 189]. It has been shown that m⁶A is added on the mRNA prior to nuclear export and WTAP recruits this “writer” complex to nuclear speckles [190]. There are several proteins shown to be interacting partners for this “writer” complex but METTL3, METTL14, and WTAP denote an evolutionary conserved complex in eukaryotes.

Functional studies indicate that m⁶A can alter the secondary structure of mRNA [186], thus making neighboring sites accessible for ribosomal binding proteins (RBPs). It has also been shown to recruit specific m⁶A-binding proteins. Of these, are proteins that have a YTH (YT521-B homology) domain with varied functions and eukaryotic initiation factor 3 (eIF3) [183]. YTH domain family (YTHDF) proteins have been shown to decrease mRNA stability [188] and enhance translation rates in eukaryotes [191], whereas the recruitment of eIF3 mediates translation of cap-independent mRNAs [183]. Several different cellular processes have been shown to be regulated by m⁶A that include cancer cell progression [185] and cell differentiation, meiosis in yeast [192] and sex determination in *Drosophila* [193].

In this study, I have utilized mass spectrometry to identify and quantify mRNA modifications across the developmental cycle of *P. falciparum*. In addition, our collaborators have carried out sequencing studies to investigate the dynamics of transcripts that are m⁶A-modified and demonstrate it to be an important part of post-transcriptional regulation during the IDC. Using a proteomics approach, I have also unearthed the m⁶A writer complex in conjunction with putative m⁶A-binding proteins. These results demonstrate m⁶A dynamically modulates a hard-

wired *P. falciparum* transcriptome across blood-stage development and show PfMT-A70 to be an integral component of m⁶A methylation.

4.1.Methodology Overview

The methodology employed for the *de novo* discovery of RNA modifications is based on a combination of techniques that have been employed in previous studies (Figure 22) [194]. The first step involves isolating high quality RNA. Our initial method assessed this using an unsynchronized culture of *P. falciparum* parasites for two reasons. Ring stage parasites contain very little RNA content making it extremely challenging for some of the analysis to be carried out. In addition, we did not want to omit any stage-specific modifications that might be restricted to particular stages in our initial survey. From the total RNA, ~90% of the content is comprised of transfer RNA (tRNA) and ribosomal RNAs (rRNA) which are known to be heavily modified. While optimizing for different methods we ensured stringency in our purification to the best of our ability and decided that two rounds of oligodT purification would significantly enrich mRNA with minimal loss in quantity. This mRNA was subsequently hydrolyzed to individual nucleosides and injected onto the QQQ MS for two separate analysis. An untargeted approach using a neutral loss scan (NLS) and a targeted approach using multiple-reaction monitoring (MRM) mode were employed for this discovery study. In order to ensure robustness of our identification we also carried out a high-resolution MS analysis for structure determination.

Figure 22 Methodology for de novo discovery of mRNA modifications

Discovery of mRNA modifications relies upon isolating mRNA from total RNA using oligodT dynabeads. Upon hydrolysis to individual nucleosides, we scan using a neutral loss or MRM mode on an Agilent 6490 QQQ and in parallel proceed with structure determination using high resolution mass spectrometry on an Agilent qToF 6530. Figure reprinted with permission from [2].

4.1.1. mRNA isolation

In order to ensure there was no contamination from either tRNA or rRNA species, total RNA from the parasite was subjected to two rounds of oligodT purification. This purification relies upon an oligo made up thymidine repeats that primes with the poly adenylated (polyA) tail of

eukaryotic mRNA. The purification was carried out by the collaborator at the Institut Pasteur by Dr. Sebastian Baumgarten in the laboratory of Prof. Artur Scherf. The quality for this mRNA was analyzed on the Bioanalyser and found to not contain any of the rRNA or small RNA peaks which are present in total RNA (Figure 23). This however must be interpreted with caution that a small amount of the other RNA species might still be present in the noise and can significantly contribute to the signal in our RNA modification analyses since they are known to be heavily modified. Our other attempts at mRNA purification, which relied on large RNA exclusion or the use of RNaseH which digests uncapped RNA species [195], led to significant loss in mRNA quantity and degradation.

Figure 23 Bioanalyser plots showing the purity of mRNA isolation

(Top) Bioanalyser plot depicting total RNA isolated from mixed stage parasites. 5 peaks are visible describing left to right the RNA marker (25nt), tRNA (~60-90nt), 5.8s RNA (~250nt), 18s RNA (2000nt), and 28s RNA (~3700 nt). (Bottom) mRNA isolated from total RNA using two rounds of oligoDT purification. No visible peaks for tRNA or ribosomal RNAs. Figure reprinted with permission from [2].

4.1.2. Survey of mRNA modifications in *Plasmodium falciparum*

A crucial aspect of our study involved the discovery of the complete set of ribonucleosides in an organism for which we leveraged upon liquid chromatography-coupled mass spectrometry as a formidable tool. mRNA from *P. falciparum* 3D7 strain was extracted and shipped to us by our collaborators which was subsequently hydrolyzed in the presence of anti-oxidants and deaminase inhibitors as described in the methods section. We utilized an LC system based on reversed-phase chromatography to resolve the modified nucleosides coupled to either a q-ToF or a QQQ mass spectrometer. A polar end-capped C18 stationary phase column (Hypersil Gold aq C18) provided the best retention and resolution for the analysis of ribonucleosides based on their hydrophobicity. The QQQ MS provides for a sensitive detection system for a targeted scan where the analyte is well defined, whereas the qToF MS sacrifices sensitivity for accurate mass resolution and structural characterization using CID fragmentation (Figure 22).

4.1.2.1. Neutral Loss Scan

Ribonucleosides were initially characterized using LC-MS/MS with neutral loss scan analysis (NLS). NLS essentially relies upon detecting a uncharged (hence neutral) loss of m/z 132 Da or 146 Da which is the characteristic fragmentation of the glycosidic bond between the nucleobase and either the ribose or 2'-O-methylribose moiety in a nucleoside respectively (Figure 24). Using a sample without any RNA added as a control for artefacts, this trait can be used to identify modified nucleosides. Certain nucleosides like pseudouridine (Ψ) have an unusual fragmentation pattern which cannot be detected by the above neutral loss analysis. Ψ was verified by targeted CID fragmentation of the ribose to yield the nucleobase with the ribose C1 methylene group with characteristic fragment ions having an m/z of 125 or 209. Neutral

loss scan, however, yields lower sensitivity as compared to a multiple monitoring reaction (MRM) scan.

Figure 24 Fragmentation of nucleosides

Characteristic fragmentation of the glycosidic bond of nucleosides yields a ribose (left) or 2'-O-methyl ribose moiety (right) having an m/z of 132 Da or 146 Da. This property is leveraged upon in the Neutral Loss Scan mode on QQQ-MS.

4.1.2.2. Multiple reaction monitoring (MRM)

The ions identified by this neutral loss scan were further validated and confirmed using a targeted MRM scan which allows the user to define the m/z for the precursor and fragment for the analyte for detection. For example, in order to detect the nucleoside Cytosine (rC) we would specify that the first quadrupole of the QQQ-MS selects for the specified m/z of the precursor ion which is 244.1 Da. This would be fragmented using a fixed collision energy (here 8V) which was optimized to produce the most abundant product ion without over-fragmentation. The characteristic abundant product ion is the base without the sugar moiety, with an m/z of 112.1 Da, that has been identified theoretically to be both unique and likely for this precursor ion. Figure 25 is a simplified schema of the parameters specified for the targeted analysis of

mRNA modifications on a QQQ-MS. Running standards in parallel with the same MRM method and matching retention times gave us reassurance in our data and analyte assignment.

	Chemical Name	Abbreviation	Precursor (Positive Mode) (M+H)	Product	Retention Time (m)	Collision Energy (V)	Cell Acceleration Voltage (V)	Delta (Precursor - Product)
Canonicals	Cytosine	rC	244.1	112.1	1.59	8	7	132
	Uracil	rU	245.1	113.1	2.3	4	7	132
	Guanosine	rG	284.1	152.1	5.55	16	7	132
	Adenosine	rA	268.1	136.1	4.44	16	7	132
iSTD	N15 labeled deoxyadenosine	Iso(rA)	269.1	137.1	4.44	12	7	132
		N15dA	257.1	141.1	5.33	8	7	116
	N6-methyladenosine	m6A	282.1	150.1	14	8	7	132
		Iso(m6A)	283.1	151.1	14	12	7	132
Modifications	N7-methyl-2'-guanosine	m7G	298.1	166.1	3.6	8	7	132
	Pseudouridine	Ψ	245.1	191.1	1.53	8	7	54
	2'-O-methyladenosine	Am	282.1	136.1	10.88	8	7	146
	5-methylcytosine	m5C	258.1	126.1	1.97	8	7	132
	2'-O-methylguanosine	Gm	298.1	152.1	13.1	8	7	146
	(3 or 4)-methylcytosine	mC	258.1	126.1	1.7	8	7	132
	5-methyluridine	m5U	259.1	127.1	6	8	7	132
	(?)-methyl-2'o'methyladenosine	mAm	296.1	150.1	19.4	8	7	146
	2'-O-methylcytosine	Cm	258.1	112.1	4.48	8	7	146

Figure 25 Multiple reaction monitoring parameters optimised for targeted detection of mRNA modifications in *P. falciparum*

4.1.3. Complete set of modified ribonucleosides in *P. falciparum* mRNA

The two analyses allowed us to define complete set of modifications for *Plasmodium* mRNA. In total, we identified a total of 10 candidates. Only analytes which were reproducibly detected in all biological replicates were validated. 9 out of 10 of these modifications were confirmed by corroborating compound mass, retention time, and CID fragmentation pattern to synthetic standards. Among them were pseudouridine (Ψ), N⁵-methylcytidine (m⁵C), N^(3or4)-methylcytidine (mC), N⁷-methylguanosine (m⁷G), 2'-O-methylcytidine (Cm), N⁵-

methyluridine (m^5U), 2'-O-methyladenosine (Am), 2'-O-methyladenosine (Gm), N^6 -methyladenosine (m^6A), and $N^7,2'$ -O-dimethyladenosine (mAm). Based on our column chromatography we were unable to distinguish between positional isomers of N^3 -methylcytidine and N^4 -methylcytidine since they have identical retention times and fragmentation patterns. $N^7,2'$ -O-dimethyladenosine (mAm) was based solely on selected reaction monitoring for precursor m/z 296.1 to product m/z of 150.1 and without a standard or tertiary fragmentation, the position of the methyl group is speculative. The signal intensities in the chromatograph depicted in Figure 24 are reflective of both the abundance of each modified ribonucleoside and of the mass spectrometric sensitivity on the LC-QQQ system. We were unable to detect N^1 -methyladenosine (m^1A) and N^5 -hydroxymethylcytidine (hm^5C) which are known to occur in mRNA from other organisms. The precursor to hm^5C which is m^5C was detected in our analysis albeit at a low abundance and not in all biological replicates. Inosine (I), which is known to exist on mRNAs, has only a slightly different retention time (5.0 m) compared to Adenosine (rA) (4.44m) but is unreliable in quantitation for two reasons. The characteristic fragmentation (269.1>137.1) is the same as the +1 isotopic peak for rA which is extremely abundant in mRNA. In addition, the shoulder for the rA peak overlaps with the Inosine elution masking the signal and leading to unreliable quantitation. The most abundant modifications based solely on ion count were m^6A and m^7G , which presumably comes from the mRNA cap.

Figure 26 LC-MS/MS extracted ion chromatograms of modified nucleotides analysed in parasite mRNA collected at 36 hpi.

Modified ribonucleosides were separated by reverse phase LC based on their respective polarity. LC- QQQ MS was used to perform MRM analysis for the targeted transitions. The Y – axis is scaled to reflect the total abundance of the 10 modifications. The abundance solely based on this chromatogram must be interpreted with caution since each of these molecules do not possess the exact same ionisation potential. Ion chromatograms extraction was executed by Agilent MassHunter Workstation (vB05.00). The raw figure was exported and relabelled. Figure reprinted with permission from [2].

4.2. Quantitation of mRNA modification across *Plasmodium* IDC

The next step in the study was to ascertain if these modifications displayed any dynamism through the IDC of the parasite. To this end, our collaborator shipped over mRNA that was isolated over 6 different timepoints from two biological replicates of two independent clones

each (totalling 4 biological replicates) and representing different phenotypic stages of the parasite lifecycle. After performing a quality check in order to ensure there was no contamination of ribosomal or transfer RNA, the mRNA was hydrolysed into individual nucleoside and subjected to an optimised LC-MS analysis on the QQQ. While we detected distinct signals for all of the above mentioned modifications (Signal:Noise > 3), only m⁷G, Ψ, and m⁶A had excellent peak shapes defined by symmetrical shape, tight retention times, and S:N ratio. m⁷G and m⁶A appeared to increase as the parasite developed from a ring to trophozoite, peaking at 24 hours and 30 hours respective. Ψ, mAm, and Cm peaked at the late ring-stage of the parasite at 18 hours. m⁵C, Gm, and mC were most abundant at the early ring stage (6 hours) of the parasite IDC. The signals of the modifications represented in the heat map in Figure 27 are further described in Figure 28 and Figure 30 for each individual modification.

Figure 27 *Quantitation of mRNA modifications across the parasite lifecycle*

*(A) Diagram illustrating the asexual replicative cycle of *P. falciparum* inside human RBCs including RBC invasion, host cell remodelling, replication via schizogony, and RBC egress. hpi, hours post infection. (B) Heatmap of normalized abundances (modification/canonical nucleotide) of ten mRNA modifications measured by LC-MS/MS at six-hour intervals over the course of the IDC. Average of four biological replicates for each modification and timepoint, except 30 hpi ($n = 2$). Blue indicates a high z-score while yellow indicates a low z-score. Figure reprinted with permission from [2].*

Figure 28 Global dynamics of mRNA modifications other than m^6A across IDC

Modified nucleosides were normalized to the canonical rA, rU, rG, and rC signals to account for differences in total RNA input. The ratios displayed are fold changes of these normalized values over an arbitrary average (except m^6A) as measured across 6 timepoints representing distinct developmental stages in the *P. falciparum* IDC. All values are based on the measurement of 4 biological replicates, except for TP 30 hours which is based on measurement at 2 biological replicates. Figure reprinted with permission from [2].

4.3.m⁶A displays dynamism across *P. falciparum* IDC

Of all the modifications that we detected in *Plasmodium* mRNA, m⁶A was of particular interest for several reasons. Being one of the most abundant modifications, it also displayed a synchronized dynamism across the parasite lifecycle. Not only does *Plasmodium* encode for the enzymes that are known to catalyse m⁶A, recent literature suggests an important role for this modification in post-transcriptional regulation. Using an externally calibrated curve with standards for the canonicals (rA, rU, rG, and rC) and m⁶A we were able to quantify absolute levels of m⁶A per canonical nucleoside (Figure 28). Distinct molecules have extremely different ionisation and fragmentation efficiencies and external standards help approximate these values. In general, purines (rA, rG, m⁷G, m⁶A) ionize better compared to the pyrimidines (rU, rC). Having optimized the MS source parameters for the most sensitive detection of m⁶A, both rA and m⁶A were observed to have the highest ionisation efficiencies compared to the other nucleosides. In addition we are also able to gauge a limit of detection (LOD) and a limit of quantitation (LOQ). A LOD determines the minimum amount of analyte where S:N ratio is >3, whereas a LOQ determines the minimum amount of analyte where the quantitation curve is no longer linear. The LOD and LOQ for m⁶A was determined to be 300 attomoles.

Figure 29 External calibration curve for standards

m^6A in *P. falciparum* mRNA appears to be highly developmentally regulated, with a continuous increase from 0.3% at early ring stages of the IDC up to 0.7% m^6A/A at 30 hpi representing the trophozoite stage of the parasite (Figure 30B). Even though, global m^6A/A levels in *P. falciparum* mRNA are comparable to human cells (A549), it seems that overall m^6A levels per total canonical ($m^6A/AUGC$) are much higher on average through the whole IDC. This is probably a reflection of an extremely AT rich genome composition for *P. falciparum*. It is striking to note that the putative orthologs of the writer enzymes in *P. falciparum*, PfMT-A70 (METTL3), PfMT-A70.2 (METTL14), and WTAP proteins reach peak transcription between 12 and 24 hours, just before maximal m^6A levels are reached (Figure 30C).

Figure 30 Global dynamics of N^6 -methyladenosine (m^6A) across parasite IDC

(A) MRM chromatograms for m⁶A at each timepoint of the IDC represents reproducible signals at 14.0m that correspond to the synthetic standard and no signal in the digestion matrix control. (B) Absolute quantitation of m⁶A per canonical or m⁶A per rA as calculated by the external calibration curve in Figure 29. (C) RNA-seq expression of the putative m⁶A writer complex across P. falciparum IDC as mined from [196]. Figure reprinted with permission from [2].

4.3.1. Structure validation of m⁶A using high-resolution mass spectrometry

The relative abundance of m⁶A in *Plasmodium* mRNA compared to other modifications and recent literature was an encouraging factor in our decision to pursue the dynamics of this modification and its physiological role in dictating parasite fate. Thus, in addition to using external synthetic standards, our structural assignment was carried out using C¹³-labelled *E.coli* tRNA which is well documented to contain m⁶A [197]. Intact precursor from *Plasmodium* mRNA, having a m/z of 282.10911 Da, fragmented the glycosidic bond to yield a C₆H₈N₅⁺ base ion having an m/z 150.08 Da. At the same retention time, fragmentation of the isotopic precursor with an m/z 293.15674 Da yielded a base fragment of m/z of 156.10 Da each denoting a difference of 11 and 5 C¹³ atoms respectively (Figure 31).

Figure 31 Analysis of mRNA-derived m^6A by high- resolution mass spectrometry.

*Product ion spectrum of m^6A in *P. falciparum* mRNA (grey) using heavy (C13) labelled *E. coli* tRNA (blue) as a standard. Fragmentation of m/z 282.12 (C12-red) and m/z 293.16 (C13-blue) leads to the breakage of the glycosidic bond resulting in the loss of the ribose sugar (132 Da) to form a daughter ion having a m/z 150.08 (C12-red) and m/z 156.10 (C13-blue). Figure reprinted with permission from [2].*

4.4.PfMT-A70 co-immunoprecipitation reveals the Plasmodium writer complex

A pertinent study to undertake here was to elucidate the m^6A writer complex in *Plasmodium*. Parasites expressing an episomally HA-tagged PfMT-A70 was provided to us by our collaborators at Institut Pasteur. In order to ensure full capture of the complex, I used a reversible cross-linker DSP that has been used with much success in the lab [198] (O Shariff et al, Unpublished). Late stage parasites were incubated for 30 min in PBS with DSP, subsequently quenched, washed and lysed to release the proteins (Methods). Anti-HA-coupled Dynabeads were used for a pulldown, followed by reduction, alkylation, and digestion to

individual peptides (Methods). The control sample was the background cell line 3D7 that was treated similarly. 16 proteins were identified that were specific for our pulldown sample as compared to the control sample. The top hit was our bait protein which is HA-tagged Mettl3-homolog PfMT-A70 (Q8I6Z2) with excellent coverage, alongside the two other well defined members of this complex – PfMT-A70.2 (Q8I567) and PfWTAP (Q8I5B2) (Figure 32). PfMT-A70.2 shows high homology with orthologs of the bilaterian METTL14. RNA/DNA-binding proteins like an ATP-dependent RNA helicase, and a FoP-domain containing protein, and Alba3 were also detected. Proteins like Karyopherins were indicative of nuclear localization of PfMT-A70 [199]. Three proteins that were also identified with high confidence (PF3D7_0707200, PF3D7_1036900, PF3D7_1366300) have no known functions and show no homology to other eukaryotic proteins.

Figure 32 Deciphering the PfMT-A70 protein complex in Plasmodium falciparum

(A) Comparative m^6A writer complex as elucidated in *H. sapiens* (left) and putative *P. falciparum* (right). Figure reprinted with permission from [2] (B) List of proteins detected in the co-immunoprecipitation of PfMT-A70 with unique and total peptides detected.

4.5.m⁶A-seq

In order to ascribe functionality to the presence of m⁶A on mRNA, it was necessary to identify the transcripts that contain this modification. This is carried out using a technique labelled m⁶A-seq which relies upon immunoprecipitation of fragmented mRNA using a m⁶A-specific antibody and subsequently sequencing the input and IP sample [181, 200-202]. Comparing relative enrichment allows us to define which transcripts contain the modification which is not offered by the LC-MS data. Our collaborators at Institut Pasteur carried out m⁶A-seq for three time points – 12 hours, 24 hours, and 36 hours of the *Plasmodium* IDC (Figure 33).

Figure 33 Schematic for m⁶A-seq

Schematic representing m⁶A-seq workflow where a m⁶A-antibody is used to immunoprecipitate transcripts containing the modification. Directional RNA sequencing libraries are prepared from m⁶A eluates (m⁶A-IP) and total mRNA-input (m⁶A-input). Figure reprinted with permission from [2].

4.5.1. m⁶A-antibody specificity

There exist several reports of nucleotide antibody un-specificity considering the small differential size of the modification (e.g. methylation with m/z 14 Da) as against the canonical base (adenosine with m/z 268 Da). This meant that the first step of the protocol for m⁶A-seq meant establishing specificity for the antibody using the MS method we established earlier. Upon analysing the pulldown fragments that were subsequently sent for sequencing I found an enrichment compared to an average input at every timepoint (Figure 34). I also observed a corresponding depletion in the unbound sample as compared to the original input. There was no significant enrichment or depletion of any of the other documented mRNA modifications indicating low cross-reactivity for the antibody.

Figure 34 *m⁶A antibody specificity of m⁶A-seq studies at three time points shows enrichment for m⁶A in the immunoprecipitated sample*

m⁶A levels as measured by LC-MS for the supernatant and m⁶A-IP sample normalized to total m⁶A-input depict enrichment of m⁶A at all three timepoints (12, 24, and 36 hpi). Figure reprinted with permission from [2].

4.5.2. m⁶A-seq summary results

The m⁶A-seq experiments described in section 4.5 were carried out by Sebastian Baumgarten at Institut Pasteur. The following section describes a summary of the results that were obtained from this study. Using stringent criteria for mapping from our sequencing data for the m⁶A-enriched transcripts, we found that >99% of all reads map to the protein-coding regions from mRNA and not ribosomal or transfer RNAs (Figure 35A) showing high specificity for the oligo-dT enrichment. rRNA and tRNA also contain m⁶A and this purity ensures that there is no competitive binding for the antibody used for the enrichment. Searching for sequence motifs correlated with methylated transcripts in a region ± 100 nt of the m⁶A peak found a single enriched GGACA motif (Figure 35B). We searched for anti-m⁶A pulldown transcripts over input using an enrichment criterion as described in the paper [2]. This led to the identification of 603, 873, and 996 significant m⁶A peaks at 12, 24, and 36 hpi, mirroring the increase observed using the LC/MS approach (Figure 35C and Figure 30B). Most transcripts contained only one m⁶A peak at every time point and we identified a total of 1520 non-overlapping peaks present in 1043 distinct peak transcripts (Figure 35D). We compared the mRNA half-life which refers to mRNA stability as mined from [203] and found that m⁶A methylation correlates with shorter mRNA half-life (Figure 35E). We also compared relative translational efficiencies (TE) which indicates the ratio of the footprint reads to the mRNA reads for any given gene as previously measured [204]. We observed that methylated transcripts have lower translational efficiencies (Figure 35F). Both of these measurements are more pronounced at the 12 and 24 hour timepoints as compared to 36 hpi.

Figure 35 m^6A -seq summary data

(A) m^6A -methylation is mostly observed in the coding sequence of *P. falciparum* mRNAs. (B) Consensus motif of GGACA is enriched in m^6A -methylation sites in *P. falciparum*. This compares well with the consensus motifs found in *S. cerevisiae* and *H. Sapiens* as reported in literature. (C) Proportion of m^6A containing transcripts at 3 timepoints as deduced by m^6A -seq. (D) Heatmap of average row Z-score normalized m^6A enrichments displaying changes in transcript-specific methylation throughout the IDC. Each row represents an m^6A peak in a transcript ($n = 840$). Boxplots depicting mRNA half-life (E) or translation efficiency (F) for transcripts without (grey) or with (blue) m^6A methylation sites at 12, 24, and 36 hpi. All analysis carried out by Sebastian Baumgarten (Institut Pasteur). Figure reprinted with permission from [2].

4.5.3. CRISPRi knockdown of PfMT- A70

Having already elucidated PfMT-A70 as a key component of the m⁶A writer complex using Co-IP experiments, Jessica Bryant at the Institut Pasteur designed a CRISPR-based interference system to knockdown the enzyme [205]. The technology is based on an enzymatically inactive (dead) Cas9 protein (dCas9) which is recruited to the target locus by a guide RNA but is unable to induce the double strand break. Binding to the locus of interest sterically interferes with the transcription apparatus and leads to a down-regulation of gene expression (Figure 36A) [206]. The guide RNA was designed to target the promoter of PfMT-A70 on the non-template strand 100 bp downstream of the transcription start site (gfMT-A70). Using an HA-antibody to perform a chromatin immunoprecipitation for the tagged dCas9, we observed an enrichment specific only to the PfMT-A70 locus (Figure 36B) as compared to the rest of the genome. A control cell line expressing a non-specific gRNA (gControl) was used as negative control. Upon measuring transcription levels of PfMT-A70 using RT-qPCR (reverse transcription quantitative polymerase chain reaction), it was observed that the target gene was significantly downregulated from 15% at 12 hpi to 40 % at 24 and 36 hpi (Figure 36C Top). The LC-MS measurement of global m⁶A/A levels reflect a significant decrease at all each of the time points as well, thus confirming PfMT-A70 as a key component of the m⁶A writer complex in *P. falciparum* (Figure 36C Bottom).

Figure 36 CRISPRi knockdown of PfMT-A70

(A) Schematic depicting CRISPRi strategy to disrupt target gene expression. With an unspecific guide RNA (gControl) the dCas9 protein is not recruited to a genomic locus allowing for normal transcription by RNA polymerase (RNAP). The recruitment of dCas9 to a target locus sterically blocks RNAP transcription. (B) dCas9-chromatin immunoprecipitation shows an enrichment only at the PfMT-A70 locus on chromosome 7. (C) CRISPRi targeting PfMT-A70 (gPfMT-A70) leads to reduction in both transcription (Top) and m⁶A levels (Bottom) compared to a control scrambled guide (gControl). The numbers on the x-axis represent the number of samples used for the analyses. Figure reprinted with permission from [2].

4.5.4. m⁶A-seq for PfMT-A70 knockdown using CRISPRi

Our collaborators also carried out m⁶A-seq for the PfMT-A70 knockdown parasites (gPfMT-A70) and compared it to the control cell line with a non-specific guide RNA (gControl). Summarizing these results, we find that there was a significant reduction in m⁶A peaks in the gPfMT-A70 cell line compared to the gControl at all three time points (Figure 37A) which parallels the decrease observed using LC-MS measurements (Figure 36C Bottom). We define m⁶A sensitive transcripts (m⁶A-s) as those that show $\geq$ two-fold decrease in m⁶A enrichment in gPfMT-A70 cell line and have an increased transcript abundance ($\log_2$ fold-change > 0). Looking at the mRNA half-life and translational-efficiency data we find that m⁶A-s transcripts have shorter half-lives (Figure 37B Top) and lower TE at both 12 and 24 hpi (Figure 37B Bottom) compared to either other m⁶A-containing transcripts or those that have no m⁶A.

Figure 37 m⁶A-seq for PfMT-A70 knockdown parasites

(A) gPfMT-A70 cell line displays a reduction in m⁶A enrichment levels compared to a control cell line as measured by m⁶A-seq. m⁶A-s transcripts ($\geq$ two-fold decrease in m⁶A enrichment in gPfMT-A70 cell line and an increased transcript abundance ($\log_2$ fold-change > 0)) have a shorter mRNA half-life (B Top) and lower translational efficiency (B Bottom) at 12 hpi

and 24 hpi compared to either other m⁶A-containing transcripts (m⁶A) or those that have no m⁶A (No m⁶A). Analysis carried out by Sebastian Baumgarten. Figure reprinted with permission from [2].

4.6. Discussion

Here we present the first study of mRNA modifications in *Plasmodium falciparum*. We document and quantify ten mRNA modifications across the blood stage development of the parasite. We further characterize m⁶A as a modification that it is developmentally regulated and contributes to transcriptional regulation in an otherwise hardwired cascade in the *Plasmodium IDC*. More specifically, using m⁶A-seq we are able to determine m⁶A methylation sites and investigate their translational-efficiencies and mRNA stabilities. A co-immunoprecipitation experiment identifies PfMT-A70 and its binding partners as key members of putative m⁶A-methyltransferase complex.

There are several methods that exist for identification of mRNA modifications. The oldest methods made use of a radioactive methyl substrate (SAM) and classical biochemistry to systematically identify mRNA modifications, in particular m⁷G, 2'-O methylations and m⁶A [110, 207-209]. The studies were however limited to studying modifications that necessarily contained the methyl group. In subsequent years, targeted study of mRNA modifications was performed either by derivatization which relies upon a particular enzyme or chemical reaction to induce a detectable change in the modification [210]. These derivatized mRNA also lead to predictable changes in RNA transcription, cause either a misincorporation or polymerase fall off. Pseudouridine (Ψ) is one such example which has been studied using this methodology [211]. Antibody based methods have been gaining traction in recent years but is also a targeted approach. A clever approach involves crosslinking the antibody to the mRNA and observing

polymerase fall off while preparing sequencing libraries to generate nucleotide resolution mapping of modifications akin to some of the derivatization methods [178, 182]. However, the fact remains that none of these methods are able to address an untargeted exploration of the modification landscape. The only recourse at the moment is the use of the mass spectrometry based approaches as employed in this project.

While extremely sensitive in its detection limits, mass spectrometry based approaches have their limitations as well. The biggest limitation is that we lose all transcript-specific information for the modification in question and would have to rely upon orthogonal methods to elucidate the same. In addition, the current MS approaches for *de novo* discovery are based on the characteristic fragmentation of the glycosidic bond. This shortcoming is evident for nucleosides like Pseudouridine (Ψ) [212] and Queuosine (Q) [213] that have been documented to have unusual fragmentation patterns. Another drawback for the MS based detection methods is the loss of labile and chemically reactive modifications either during the digestion or ionisation of the analytes. 5-hydroxycytidine (ho^5C) is an example of this modification as evidenced while running synthetic standards [214]. Given all of these, the consensus amongst RNA researchers still continues to put MS-based methods as one of the best for identifying new modifications. MS-based methods are amenable to absolute quantitation and also subject to far less artefacts as compared to dot-blot approaches carried out in literature. Not only are dot blots semi-quantitative, they also call into question the specificity and efficiency of the antibody. An extension of this also applies to sequencing methods which also possess a bias for highly transcribed transcripts over low abundance ones [181, 215].

Of the ten modifications identified, m^7G , m^6A , Ψ , Am and Gm have been reproducibly demonstrated to be present in eukaryotic mRNA by a variety of methods. It would be

worthwhile to investigate the localization of these modifications since m⁷G, Am, and Gm, are all known to be present in or near the cap structure of mRNAs but their existence on internal sites on the transcript remains to be elucidated. An approach here would be to use a ribonuclease like Nuclease P1 which does not digest the cap structure with a triple phosphate bond and thus would leave an ‘m⁷GpppNm’ cap intact but hydrolyse the rest of the mRNA [120]. Inosine (I) is well-known modification that we were unable to ascertain a clear quantifiable signal using our MS method. A change in solvent composition and/or column composition might help resolve Adenosine (rA) and I sufficiently enough to be detected and quantified reliably. The other modifications (m⁵C, mC, m⁵U, and mAm) identified have some reports documenting them in eukaryotic mRNA [119, 130, 180, 182], but have not been studied in much depth. A note of caution for these modifications is that they are also abundantly present in tRNA and rRNAs. Even with our stringent purification protocol, there might be a possibility that the low signals detected for these modifications may have been from these residual RNA species. Other reported modifications in mRNA, m¹A and hm⁵C, were not found in our survey, possibly indicating extremely low abundance or perhaps an issue with digestion conditions and MS ionisation parameters [126, 132]. There is literature to suggest that m¹A undergoes a Dimroth rearrangement to form m⁶A under basic conditions [216]. However, the kinetics of this rearrangement are much longer than the incubation time used in our digestion. This is demonstrated in our tRNA modification data (Chapter 3) where an m¹A signal is still abundantly detected. It is far more likely that these two modifications are altogether absent in *Plasmodium* mRNA, as the few reports of these modifications that exist in literature are not reproducible.

The quantitation of these ten modifications across the *P. falciparum* IDC showed a dynamic change across the parasite lifecycle. In particular, we focussed our studies on m⁶A which was

the relatively most abundant of all modifications measured. Using high resolution MS and external standards we were able to structurally confirm the presence of this modification. The evidence presented for the existence of this modification is more than sufficient, but secondary fragmentation of the released base using MSⁿ and NMR studies could be used for more extensive characterization. LC-MS data, in conjunction with m⁶A-seq, depicted a steady increase in this modification as the parasite progressed from ring stage parasites to late trophozoite and schizont stage parasites. This is different from some of the other studies where m⁶A levels remain relatively static in translationally demanding processes like cell differentiation [190]. Comparing m⁶A-seq results with previously published data [203], we were able to analyse mRNA stability and translational efficiencies for these transcripts that were methylated. In agreement with some of the literature in the field, m⁶A methylation is associated with transcripts that display a shorter half-life and lower translational efficiencies. This could likely be an interplay of the m⁶A-interacting proteins termed ‘reader’ proteins that contain a YTH domain. These reader proteins are capable of recruiting endoribonucleases directly or recruiting decapping enzymes (DCP1,2) which would eventually lead to mRNA decay. Mammalian cells encode for YTH domain family (YTHDF) proteins 1 to 3 and have been shown to interact with m⁶A methylated transcripts [77, 139, 186, 188, 191]. A careful observation at their primary sequence and binding domain suggests that these proteins might be partially redundant. *P. falciparum*, on the other hand, encodes two YTH proteins which are significantly divergent beyond the YTH domain and have disparate expression patterns across the asexual and sexual lifecycle (Figure 38). This suggests that they contribute to a functionally different role. A peculiarity of the *Plasmodium* transcriptome program is the translational repression of genes required for the formation of sporozoites which are the transmission intermediates between the human host and mosquito vector [67, 217]. Translationally repressed mRNAs essentially for the lifecycle of these stages are rapidly translated only upon

transmission. High levels of co-expression of writer protein PfMT-A70 and PfYTH.2 in the sporozoite stage could suggest a possible involvement of these proteins in mediating translation of repressed mRNAs (Figure 38). Another factor that could possibly induce these changes is the haemolymph circulatory temperature in the mosquito vector. A recent study shows that poly-methylated m⁶A-RNA offers a scaffold that aggregates YTHDF proteins, causing them to undergo liquid-liquid phase separations (LLPS) through interactions between their low-complexity domains [218]. This LLPS is temperature dependant whereupon lowering the temperature from 37 °C to 4 °C, the individual aggregates in a turbid suspension undergo disassembly to form a clear solution. This compartmentalization could effectively determine when translationally repressed transcripts in granules could disperse and become active again since the parasites transition from human circulatory temperature of 37 °C to mosquito hemolymph which is generally much lower at ambient temperatures of ~22 °C.

Figure 38 YTH-domain containing proteins in *P. falciparum*

(A) RNA-seq analysis of m⁶A-writer ‘PfMT-A70’ and reader ‘PfYTH.1/2’ proteins across the sexual and asexual lifecycle of the parasite. G-Gametocyte, O-Ookinete, S-Sporozoite. Sequence identity comparison of putative m⁶A reader proteins for YTH-domain (B) and full length sequence (C) from *P. falciparum*, *H. Sapiens*, and *S. cerevisiae*. Figure reprinted with permission from [2].

Of interest is the *P. falciparum* ‘writer’ complex for m⁶A. The METTL3 ortholog in *P. falciparum*, PfMT-A70 can be found in most other eukaryotes and the immunoprecipitation study identified PfMT-A70.2 and PfWTAP which reflect a conserved core m⁶A writer machinery [179]. The m⁶A-seq results point to this ‘writer’ complex recognising the motif DRACH (D = G/A/U; R = G/A; H = C/A/U) that is very similar to the one found in yeast and mammalian cells [190, 192]. However, none of the other cofactors that are well documented in plants and metazoans (eg HAKAI and VIRMA) could be identified in the PfMT-A70 pulldown [179]. The lack of VIRMA could also explain the absence of methylation in the 3’ UTR of *P. falciparum* with which it is shown to be associated previously [219]. Most of the interacting proteins show nuclear localization suggesting that m⁶A methylation takes place prior to nuclear export as also reported in other organisms. A point to note is that mammalian cells also contain an m⁶A ‘eraser’ proteins. Initially thought to be FTO, more rigorous studies have shown it to be ALKBH5 which can demethylate m⁶A on mRNA [187, 220, 221]. FTO has a 100-fold higher substrate specificity for m⁶Am as against m⁶A, and has little evidence of demethylation activity on m⁶A in vivo. Upon searching the *Plasmodium* genome, we were unable to find any orthologs for either of these demethylases. This leads to a peculiar question if there exist lineage specific demethylases that are unique to apicomplexans. As for the modulation of m⁶A levels in *Plasmodium*, this could result from differential activity of the writer complex and the subsequent rate of m⁶A-methylated transcript degradation. Using CRISPRi we were able to down-regulate the expression of PfMT-A70 which subsequently impacted the global level of m⁶A mRNA methylation and led to the increased abundance of transcripts that are m⁶A methylated during the parasites IDC.

It is undeniable that the primary factor controlling for mRNA transcript abundance at the genomic locus is RNA polymerase activity. This study, however illustrates that the fate of

certain transcripts are tuneable, either by reducing mRNA stability and/or suppressing translation as a result of m⁶A methylation. While not in its entirety, this system of RNA modifications offers for the potential to finely choreograph an otherwise imprecise, hardwired transcriptional program.

Chapter 5 – A role for tRNA modifications for modulating stress response in *P. falciparum* via translational control mechanisms.

5. Overview

Upon examining the coding sequences in the genomes of different organisms, there exists a consistent pattern in which organisms preferentially opt for one of the several synonymous codons to code for an amino acid. This is unlikely to be a random quirk of biology, but rather posits that there exists a ‘code of codons’ [222]. This has been attributed to several features of transcription and translation. For example, this choice optimization serves to regulate gene expression by triaging the translation of specific transcripts over others [223]. This is a difficult feature to identify if observed at a discrete gene transcript and protein abundance level but is grossly apparent by gene-specific codon usage of the transcriptome. In a similar vein, the biological role of tRNA modifications can only be appreciated if it is considered to function at a systems level that responds to cellular stress. Modifications to the canonical ribonucleosides interface with the translational process by altering the interactions between codons and anticodon interactions thus expanding or restricting the decoding capability of the tRNA. How would this system of tRNA modifications protect cells against stress? One hypothesis, as put forth by several papers, is that dynamic reprogramming of tRNA modifications at the wobble position either seek increase/decrease the decoding capacity of cognate/near-cognate tRNAs which in turn affect an up/down-regulation of stress-response/non-essential genes respectively (Figure 39 Left) [86, 87, 90, 94, 95, 97, 133-135]. Concisely, these modification changes enable a selective translation of genes that aid in mounting a stress response. A contrasting but not mutually exclusive hypothesis that is also starting to gain traction is “adaptive mistranslation”

(Figure 39 Right). Several different groups have shown how error-prone DNA and RNA polymerases contribute to vast heterogeneity in a population ultimately improving overall fitness of an organism [224, 225]. In a similar vein, “adaptive mistranslation” argues that errors arising during the translational process might not be so much a bug, but rather a feature of the translational program of organisms [226]. There are several ways in which this would aid the organism cope with different stresses as depicted in Figure 39. Mistranslation-induced mutagenesis could either obliterate drug binding or cause a gain-of-function that aids the cell in coping with cell stress [227]. Misincorporation of methionine (which has sulfur side-chains) function as sponges to protect against cellular reactive oxygen species (ROS) [228, 229]. General mistranslation also leads to incorrectly folded proteins which upregulate the unfolded protein response (UPR) pathway, causing the cell to ramp up the proteasomal activity and production of chaperones for folding [230]. In this chapter, we investigate how *Plasmodium* parasites respond to stresses at a systems level by carrying out a tRNA modification-based profiling as well as a proteomics study. It is important to note that at the time of this thesis, we were able to complete analysis for only two of the stressors for the proteomics study. We integrate this data using multivariate statistics to put forth a rudimentary model where RNA modifications serve to play a role in controlling parasite stress response.

Figure 39 tRNA modifications and stress response

(Left) Dynamic modification of tRNA isoacceptors leads to selective translation of codon-biased transcripts that are involved in mitigating stress-response. (Right) Adaptive mistranslation serves can contribute to increased fitness in a multitude of ways. Adapted from [226] with permission.

5.1. Growth Inhibition Dose Response Studies

We investigated tRNA modification reprogramming in *Plasmodium falciparum* using a variety of stressors. For this study, we utilized two anti-malarial drugs – chloroquine (CQ) and dihydroartemisinin (DHA), and two synthetic chemicals that mimic oxidative stress – hydrogen peroxide (H_2O_2) and sodium hypochlorite (NaOCl). CQ and DHA have been two generations of frontline antimalarial therapeutics with widespread success [28]. Despite being in use for several decades, their mechanisms of action and parasite stress response to these drugs remain a point of contention [39]. The emergence of resistance to both of these drugs further garners support of this study of malarial parasite drug response [39]. Hydrogen peroxide was chosen

to reflect the oxidative stress that is experienced by the parasite during its lifecycle. Digestion of hemoglobin inside the erythrocytes is thought to be extremely pro-oxidative due to high concentrations of iron and oxygen [231]. Most of this oxidative damage is mitigated by the parasite as it sequesters and effectively “inactivates” the free heme by sequestering it in the digestive vacuole in a crystalline form of hemazoin. However, even a small amount of this free heme can be oxidized and lead to the production of superoxide and hydroxide radicals by Fenton chemistry. One of the ways of artificially introducing a similar stress is by the use of H_2O_2 [11, 232, 233]. Besides, the parasite also encounters immune cells which effect their killing properties by the generation of similar oxidative and nitrative radicals [231]. In addition to H_2O_2 , NaOCl was chosen to reflect another radical ($\cdot\text{OCl}$) that the parasite might encounter if in proximity to natural killer cells or macrophages. Three additional compounds were chosen to mimic oxidative stress that the parasite might encounter but were not pursued for further studies due to resource constraints. We profiled S-Nitroso-N-acetyl-DL-penicillamine (SNAP) and Spermidine-NO, a slow release NONOate, which are nitric oxide (NO) donors and generate $\cdot\text{ONOO}$. Menadione has previously been used to generate intracellular reactive oxidative stress in cells through futile redox-cycling and was also characterized [234].

5.1.1. Flow Cytometry for Determination of Parasitemia

In order to ascertain parasitaemia at the end point of our drug treatments, we decided to rely on a high-throughput method that uses flow cytometry for estimation. We used a combination of two dyes to label the parasites - Hoechst (indicated here by the DAPI-A channel on the y-axis) labels for all infected red blood cells (RBCs). Since human RBCs are enucleated they contain very little to no DNA content. While Hoechst can stain leukocytes, we deplete the blood of the white blood cells by extensive washing. This is confirmed subsequently by microscopy and the leukocyte levels are $<0.0001\%$. It is important to note that Hoechst will

also label parasites that are no longer alive but are still present within the RBC. The second dye used was Mitotracker Deep Red (Invitrogen) which would stain for the mitochondria with membranes that still maintained an active redox potential thus serving as a marker for cells that are alive. While the use of either of these dyes singly is generally accepted, the combination of the two dyes gives us robustness in our measurements.

Figure 40 Flowcytometry gating strategy for parasitaemia determination

(Top) Schematic depicting the two different IC_{50} assays that were carried out. One was a ‘lifecycle’ IC_{50} which evaluated parasites over a whole 48 hours lifecycle. The other IC_{50} was a short 3h exposure of the stressor followed by wash and reseeding in fresh blood and medium. (Bottom) Gating strategies using Hoechst (y-axis) and Mitotracker Deep Red (x-axis) depict uninfected RBCs in bottom left quadrant (Q4) and live parasites in top right quadrant (Q2).

In order to define the population of cells that are infected vs non-infected, we are able to draw iterative “gates” which represent a threshold for the signals we measure. These thresholds are defined by using a negative control (uninfected red blood cells) and a positive control (infected

red blood cells) with a known parasitaemia confirmed independently by visual microscopy. The first gate is carried out on the Forward Scatter (indicative of size) and Side Scatter (indicative of cell granularity) to define a population of our cells. This helps us eliminate any signal that might come from debris or dye aggregates. The second gate is carried out on the signal intensity for wave characteristics of the side-scatter which represent the area and height. This helps determine if the cell population that was detected was single cell or an aggregate of two or more cells. The third gate has the two fluorescent dyes on the two axes, and a “double positive” signal in the top right quadrant represents the parasite population that is alive (Figure 40).

5.1.2. IC₅₀ growth inhibition assays

We performed two separate dose-response studies to evaluate an IC₅₀ which would enable us to use concentrations relevant to our stressors as described in the methods section 2.2 (Figure 41). One of these studies evaluated the dose response over the entire life cycle of *P. falciparum* for these compounds where the parasites at 1.5-2% starting parasitemia were treated at 28-36H trophozoite stages for the full 48 hours. The other study measured the potency of these compounds over a short exposure at a similar trophozoite stage but higher parasitemia of 5-7%, whereupon treatment for 3 hours, the compounds were washed, the parasites were reseeded at 1.5-2% parasitemia and allowed to progress normally through their lifecycle for ~45 hours. This short exposure study is better reflection of the final experimental conditions that would be used for the stress response experiments for two reasons. One, our model wishes to investigate translational changes that are immediate since this would facilitate a “just-in-time” response on account of the tRNA modification reprogramming. Also, all of these compounds (apart from the NONOates) are known to be effective even with small exposure since they are relatively fast-acting. The parasite growth inhibition was calculated in terms of the ratio of

parasitemia levels at different concentrations relative to a sample that was treated with vehicle (only solvent without the drug). The parasitemia at the end-point of this assay was measured using flow-cytometry as described earlier. We were unable to utilize higher doses of H₂O₂ since beyond 500mM there was significant lysis and “frothing” for the cultures. This is due to lysis of the red blood cells and the fact that hydrogen peroxide reacts quite violently at those concentrations. A four-parameter sigmoidal fitting method with variable slope was used for non-linear regression. The best fit values for IC₂₀, IC₅₀, and IC₈₀ were derived from this curve (Figure 41). The lifecycle IC₅₀ values for all compounds were within 1-log difference of their short-exposure IC₅₀ values, both of which are listed in Figure 41. We were unable to see any defect in parasite growth with Spermidine-NO for a short exposure and only a small inhibition for the lifecycle IC₅₀ assay. These doses were used as equivalent phenotypic-end points to allow for meaningful assessment amongst different stressors. A range of high to low inhibitory dosages would allow for visualization of any dose-specific tRNA modification change or proteome-wide changes.

Figure 41 Dose-dependent inhibition assays for oxidative stressors

A 11-point growth inhibition assay was carried out for 7 stressors at two-fold serial dilutions in biological triplicates to determine IC₂₀, IC₅₀, and IC₈₀ values. The concentration of these treatments were log₁₀ transformed, the %parasitemia was plotted as a ratio of treated:untreated control, and fitted using a four parameter non-linear curve regression. An IC₅₀ value represented 50% inhibition concentration and was interpolated as shown by the dotted line. These values are listed for both the lifecycle IC₅₀ assay and short exposure IC₅₀ assay in the bottom right table for all the stressors.

5.2.tRNA Modification Profiling

An established LC-MS platform allowed us to compare tRNA modification changes in parasites exposed to various stressors like dihydroartemisinin (DHA), chloroquine (CQ), hydrogen peroxide (H₂O₂), and sodium hypochlorite (NaOCl) at a range of concentrations (IC₂₀, IC₅₀, and IC₈₀). 27 modifications were monitored and quantified using an LC-MS based method operating in Multiple-Reaction-Monitoring (MRM) mode. One of the modifications, 2-thiouridine (s²U), identified in the previous study [1], is not included in the final heat map for our analysis. Being of extremely low abundance, the signal for s²U is unreliable for quantitation with a signal:noise (S:N) ratio that was less than 3. In addition, the technical replicates that were run with higher injection volumes have a high degree of variance for s²U and do not correlate with injection amounts as opposed to all of the other modifications which have a proportionate signal to the injection volume. We ensured to eliminate background signals and artefacts using a mock-digest (no tRNA) control sample that may arise due to the use of anti-oxidants or deaminase inhibitor cocktails in the hydrolysis mix. The signals were calculated as an area under the curve (AUC) and normalized to the canonical nucleosides (rA,

rU, rG, rC). The ribonucleoside levels are expressed in terms of Log₂ (fold-change) of treated sample relative to the unexposed control sample.

At first glance, it is immediately apparent that each of the individual stressors induce a characteristically different pattern of modifications (Figure 42). The fold change for all of these modifications is quite modest but in line with what has been previously reported [1, 97]. The general trend for most modifications is that they go down upon treatment with all of our stressors. Of all of these, artemisinin stress does not significantly change much of the modifications except for N⁶-threonylcarbamoyladenosine (t⁶A) which goes down upon treatment regardless of the dosage experienced by the parasite. Upon treatment with NaOCl, I, Am, mcm⁵U, m¹G, i⁶A, Um and Gm are the modifications which increase. mC, m⁶A and ac⁴C are the modifications that consistently go down regardless of the dosage. For chloroquine, while most modifications remain constant at IC₂₀ treatment, they all go down at IC₅₀ and IC₈₀ except for m⁶A, Am, and t⁶A. Of these, i⁶A is downregulated in a dose-dependent manner. For H₂O₂ treatment, IC₂₀ and IC₅₀ are similar in their responses with respect to most tRNA modifications being down-regulated whereas at IC₈₀ it seems most modifications change relatively little. The only exception is m^{6,6}A which goes up in every single treatment with H₂O₂.

Figure 42 Stress-induced tRNA modification changes in Plasmodium falciparum

*Heatmap displaying changes in 27 tRNA modifications as measured by LC-MS/MS and the data is expressed as fold change relative to the untreated control (C0, H0, N0, D0). Red – Increase; Blue – Decrease. The bottom-left color bar defines the range of $\log_2(\text{fold-change})$ values. The experiment was a result of three biological replicates. Differences between the control sample and the largest value are statistically tested with two-tailed Student's *t*-test (NS: not significant, $P < 0.01$ and $P < 0.001$ are denoted as * and **, respectively). The modifications with their full chemical names and MRM transitions are listed in Appendix I.*

5.3.A preliminary proteomics perspective – DHA and CQ Stress

Response

Based on the translational control mechanism described in Chapter I, we hypothesized that this observed reprogramming of tRNA modifications could impact stress-induced parasite protein profiles by enabling selective translation of codon-biased transcripts. To test this postulation, we carried out TMT-based quantitative proteomics [235] at the same concentrations and drug regimen at which the tRNA modifications were analyzed. We point out that the study described here reflects the proteomics data obtained from only from Dihydroartemisinin (DHA) and Chloroquine (CQ), with the other two stressors (NaOCl and H₂O₂) being processed at the time of writing of this chapter. There are several analyses that need to be carried out to ensure robustness for this dataset. The proteomics data is an aggregate value from two biological replicates and lacks statistical strong statistical validation. We have additional replicates (for which we had performed a QC on the labeling and digestion efficiency) that need to be run and should be subsequently included in the analysis. As for other analysis –a pathway enrichment analysis to decipher the proteome response to the stress would be ideal. An underlying common pathway perhaps that is triggered as a result of the different stressors can be analyzed only

upon obtaining all of the data. We would also like to preclude any proteomic changes that are a result of transcript abundance before carrying out a PLS-R analysis to test for codon-biased translation.

5.3.1. Data summary

By cross-validating protein identification between SEQUEST [236] and Mascot Servers [236], we were able to identify 117111 / 109037 peptide spectral matches (PSMs) that mapped onto 1309 / 1206 proteins (protein threshold: >99% confidence, minimum 2 peptides) by matching to *P. falciparum* protein sequence databases for treatments with DHA / CQ. The raw intensities of total quantified *P. falciparum* peptides in each channel were used for cross-sample normalization to adjust for variation in sample input (Figure 43). The intensities of the proteins in their respective samples were compared to the untreated control sample. Through the data processing, we defined the constitutive and regulated (up-regulated: fold-change relative to control >1.3; down-regulated: fold-change relative to untreated control <0.7) expression of the *P. falciparum* proteins. In addition, we analyzed peptides and proteins identified in the different high-pH fractions that was carried out to reduce sample complexity (Figure 44). While there is quite a substantial overlap for the proteins that were identified in the different high-pH fractions, the peptides that are identified in each fraction have very little overlap. This is indicative of a good fractionation profile that reduces sample complexity for mass spectrometry analysis. A principal component analysis of the fold changes for the proteins indicated that the control sample clustered separately from the treated samples for DHA (Figure 45). It was more striking however to observe that the 80% of the variation in the proteomics data was in the DHA IC₈₀ sample compared to all of the other samples indicating perhaps a discordant response. The DHA IC₂₀ and DHA IC₅₀ seem to respond in a manner that is akin to each other with a slight difference. For the CQ data, the PC1 axis accounted for majority of the variation

in the group (62.7%) and the control sample was clearly separated from the treated samples. The treated samples also displayed a progressive distance from each other in a dose-specific manner indicating a conserved response to the drug.

Dihydroartemisinin

Chloroquine

Figure 43 Pre/ Post- normalization of the raw intensities for all proteins in each quantitative sample for each replicate for DHA (top) and CQ (bottom).

Box plots displayed the values of $\log_2$ raw intensities of total proteins for corresponding sample with median and interquartile range. The cross-sample normalization was performed by Proteome Discoverer software.

Figure 44 Number of peptides or proteins found in the different fractions of the high pH column fractionation for DHA treatment

Venn diagram displays the number of peptides (left) and proteins (right) found in fractions that represent F1.1 10%ACN, F1.4 17.5%, and F1.7 25% ACN elution. While there is a large overlap in the number of proteins identified, the peptides identified in each of the fractions is quite different. This is indicative of a good fractionation profile that reduces sample complexity for MS/MS Analysis.

Figure 45 Principal component analysis (PCA) plots for global proteomic changes upon DHA (left) and CQ (right) treatment

Scores plot of PC1 vs. PC2 for DHA (Left) and CQ (Right). Mean values of biological triplicates were used for the analysis. Different treatments are presented as IC₂₀, IC₅₀, and IC₈₀; Percentages of observed variances explained by the PC1 and PC2 are indicated in parentheses along their respective axes.

5.3.1. Proteomics data analysis and pathway enrichment

I carried out a preliminary analysis on the proteins that were differentially regulated upon treatment with DHA and CQ. In addition, pathway enrichment analysis using STRING DB (<https://string-db.org/>) was performed for those up and down-regulated proteins to gain functional insights into the stress-specific proteomes. This analysis takes into account a multitude of factors while looking at the input of genes. It considers potential interacting partners based on co-immunoprecipitation experimentation, interaction data from other

organisms, intergenic distance etc. Most significantly up-regulated proteins from the DHA IC₈₀ data compared to the control sample we find an enrichment of ribosome biogenesis genes that are overrepresented. The protein showing the most change in DHA treated samples was early transcribed membrane protein 10.1 (Q8IK21) which has a peak expression in the early trophozoite stage of the parasite suggesting that the drug pressure was sufficient to halt development of the parasite from progressing further in its lifecycle. The same is true for C6KSL5 and Q8IBF2 which are Maurer's Cleft protein and EMP1-trafficking protein respectively, that also have peak expression in the early trophozoite stage of the parasite. For the CQ-treated sample, three of the top proteins - Thioredoxin (A0A144A4E0), Glutaredoxin 1 (Q9NLB2) and Peroxiredoxin (Q8IJD0), peak in the schizont stage. This would indicate that the trophozoite treated cells were trying to mount an oxidative stress response to the drug pressure [233]. A majority of the other proteins that were up-regulated with CQ treatment were ribosomal proteins. For the down-regulated proteins in the DHA treated samples, a lot of the proteins were involved in cell metabolism and vesicular trafficking. We are unable to comment on the proteins that were down-regulated in the CQ-treated samples since they mostly contained uncharacterized and unannotated proteins in the genome. Looking at the trend for fold change values there exists dose-dependency for most of the proteins that were analyzed. It is noteworthy however that the dynamic range for a proteomics experiment is much smaller (<3 fold difference) as compared to transcriptomic experiments which can yield 8-16 fold changes over the control sample. A more thorough and curated analysis of this proteomics dataset will help elucidate pathways by which the cells respond to these toxicants and if they exhibit common stress response like the transcriptional program might indicate.

Accession	Description	IC20	IC50	IC80
Q8IK21	Early transcribed membrane protein 10.1	1.825	2.619	3.438
Q8IK27	Uncharacterized protein	1.063	1.53	2.404
Q8I5N3	Uncharacterized protein	1.275	1.571	1.935
C6KSL5	Pfmc-2TM Maurer's cleft two transmembrane protein	1.34	1.6	1.769
C6KSQ8	Uncharacterized protein	1.028	1.451	1.733
Q8IBG5	Ribosome biogenesis protein BRX1 homolog,putative	1.101	1.354	1.731
C6S3B8	Uncharacterized protein	1.004	1.119	1.666
C6KT04	Uncharacterized protein	0.919	1.228	1.665
Q8I1X4	Uncharacterized protein	1.067	1.277	1.665
O96150	DNA-directed RNA polymerase II 16 kDa subunit,putative	0.921	1.26	1.657
Q8IJX2	PPPDE peptidase, putative	1.339	1.426	1.652
Q8I3V8	Pre-mRNA-splicing factor CWC2, putative	0.966	1.233	1.647
Q8IJX4	Uncharacterized protein	1.396	1.707	1.627
Q8ILB9	Dynein-related AAA-type ATPase, putative	0.878	1.154	1.627
C6KSR0	Ubiquitin-conjugating enzyme E2, putative	0.848	1.13	1.613
Q8I567	mRNA methyltransferase, putative	0.985	1.125	1.608
Q8IBF2	EMP1-trafficking protein	1.093	1.335	1.603
Q8IEU2	Gamete antigen 27/25	0.84	1.253	1.603
C0H5G9	Krox-like protein, putative	1.085	1.425	1.603
Q8I275	Small ribosomal subunit assembling AARP2 protein	0.97	1.295	1.602
Q8IJD2	U3 small nucleolar ribonucleoprotein protein MPP10, putative	0.943	1.26	1.601
O96201	Uncharacterized protein	1.027	1.213	1.598
Q8IBF6	Transcription factor with AP2 domain(S)	1.143	1.353	1.598
Q8IIB5	Ribosome biogenesis regulatory protein	1.101	1.245	1.588
Q8IAM3	Periodic tryptophan protein 2, putative	0.951	1.212	1.588
Q8II69	RNA methyltransferase, putative	0.923	1.123	1.571
C6KSX2	Nucleolar GTP-binding protein 1	1.04	1.234	1.568
Q8IC44	Sporozoite threonine and asparagine-rich protein	0.923	1.125	1.564
Q8I3H9	Nucleolar preribosomal GTPase, putative	1.057	1.262	1.557

Figure 46 Up-regulated proteins in DHA treated-samples

Accession	Description	IC20	IC50	IC80
Q8IBQ5	40S ribosomal protein S10, putative	1.176	0.908	0.757
Q8I566	Serine hydroxymethyltransferase	0.903	0.827	0.755
O77361	Parasite-infected erythrocyte surface protein	0.651	0.642	0.752
Q9TY96	Serine repeat antigen 6	0.677	0.65	0.746
Q8IDQ9	Phosphoethanolamine N-methyltransferase	1.053	0.857	0.745
Q8IIR7	Endoplasmic reticulum-resident calcium binding protein	0.629	0.65	0.745
C0H4E7	SWIB/MDM2 domain-containing protein, putative	0.681	0.74	0.744
Q8I569	CCR4-NOT transcription complex subunit 4, putative	0.995	0.967	0.744
Q8I3R0	40S ribosomal protein S9, putative	1.212	1.095	0.74
Q8I3A4	Prefoldin subunit 4	1.019	0.744	0.738
A0A144A2R4	Methyltransferase, putative	0.916	0.775	0.738
Q8I573	Uncharacterized protein	0.732	0.661	0.737
O97320	Histone H2A	0.688	0.758	0.721
Q8IJE1	Uncharacterized protein	1.057	0.989	0.718
Q8IM38	Ribonucleotide reductase small subunit	0.871	0.774	0.715
Q8I713	60S ribosomal protein L36	0.981	0.795	0.7
Q8IKU1	Co-chaperone p23	1.209	1.059	0.699
Q8IDF9	Inositol-polyphosphate 5-phosphatase, putative	0.706	0.733	0.691
C6S3F0	Acylphosphatase	0.863	0.731	0.687
Q8I488	Parasite-infected erythrocyte surface protein	0.806	0.767	0.673
C6KSY6	60S ribosomal protein L19	1.167	1.029	0.672
Q8IB51	60S ribosomal protein L22, putative	0.951	0.807	0.666
Q8I6U5	Cysteine proteinase falcipain 2b	0.473	0.621	0.66
Q8IDY0	MSP7-like protein	0.557	0.56	0.656
Q8IM34	RNA-binding protein 8A, putative	0.944	0.825	0.638
Q8I2V9	Protein disulfide isomerase	0.575	0.544	0.627
Q8I1Q2	Trafficking protein particle complex subunit	1.062	0.714	0.622
Q8ID43	Nucleoside diphosphate kinase	0.894	0.731	0.621
C0H477	Inorganic pyrophosphatase, putative	0.728	0.623	0.615
A0A144A153	RNA-binding protein, putative	0.709	0.731	0.565

Figure 47 Down-regulated proteins in DHA treated samples

Accession	Description	IC20	IC50	IC80
A0A144A4E0	Thioredoxin	2.065	2.039	2.346
Q9NLB2	Glutaredoxin 1	1.971	1.739	1.991
Q8ILK3	60S ribosomal protein L21	1.341	1.537	1.79
Q8IBY4	60S ribosomal protein L34	0.92	0.854	1.735
C6S3I6	Replication factor A protein 3, putative	1.744	1.732	1.686
Q8I3X4	Purine nucleoside phosphorylase	1.24	1.477	1.636
Q8I3T9	60S ribosomal protein L2	1.052	1.175	1.572
C6KT19	Histone H3	0.98	1.406	1.538
Q8IIV1	Histone H2B	1.15	1.513	1.537
Q8IJD0	Peroxiredoxin	1.072	1.054	1.526
C0H4A6	Ribosomal protein L15	0.975	0.861	1.513
O97248	40S ribosomal protein S23, putative	1.194	1.479	1.511
Q8IIV2	Histone H4	0.944	1.022	1.49
Q8IKK7	Glyceraldehyde-3-phosphate dehydrogenase	1.319	1.35	1.486
Q8IE82	60S ribosomal protein L23	0.882	0.869	1.485
A0A144A2R4	Methyltransferase, putative	1.162	1.427	1.467
Q8IAR3	Proteasome endopeptidase complex	1.351	1.424	1.456
Q8IB51	60S ribosomal protein L22, putative	0.927	1.021	1.449
Q8ILE8	60S ribosomal protein L14, putative	0.983	0.966	1.449
Q8IAX6	60S ribosomal protein L13	0.869	0.965	1.447
Q8IB57	Small nuclear ribonucleoprotein G	1.353	1.54	1.445
Q8IC33	Uncharacterized protein	1.668	1.483	1.443
C6KSV0	Histone H3	1.002	1.338	1.438
Q8IEU2	Gamete antigen 27/25	1.034	0.992	1.436
Q8ILV2	60S ribosomal protein L10, putative	1.158	1.165	1.428
Q8I3A3	Ubiquitin specific protease, putative	1.124	1.162	1.405
Q8IDI5	60S ribosomal protein L17, putative	1.004	0.95	1.401
O96184	60S ribosomal protein L37a	1.226	1.345	1.396
Q8IIA9	Adenosine deaminase	1.179	1.173	1.39
Q8IDR9	40S ribosomal protein S6	0.73	0.804	1.389

Figure 48 Up-regulated proteins in CQ treated samples

Accession	Description	IC20	IC50	IC80
C0H4T6	Protein phosphatase PPM5, putative	0.769	0.636	0.472
Q8I472	Mitochondrial inner membrane TIM10 associated protein, putative	0.7	0.753	0.468
Q8ILR6	Ubiquitin fusion degradation protein UFD1, putative	0.52	0.443	0.466
Q8I1X3	Uncharacterized protein	0.627	0.506	0.466
Q8IHW3	Uncharacterized protein	0.9	0.904	0.459
Q8IEE2	Uncharacterized protein	0.922	0.746	0.459
C0H535	Uncharacterized protein	0.977	0.779	0.457
Q8IIR4	Uncharacterized protein	0.647	0.703	0.457
Q8I3F9	Formin 1	0.502	0.428	0.457
Q8IKQ5	Vacuolar protein sorting-associated protein 4	0.442	0.433	0.454
Q8I4Y5	ADP-ribosylation factor GTPase-activating protein, putative	0.718	0.681	0.454
Q8IK15	EMP1-trafficking protein	0.646	0.555	0.452
Q8IK07	Uncharacterized protein	0.785	0.815	0.446
Q8I414	Uncharacterized protein	0.792	0.78	0.444
Q8IIK7	Uncharacterized protein	0.947	0.794	0.442
Q8IM55	Serine/threonine protein phosphatase CPPED1, putative	0.472	0.432	0.44
Q8ILR8	Uncharacterized protein	0.659	0.611	0.439
Q8IJJ5	RNA cytidine acetyltransferase	0.556	0.564	0.431
C0H4F1	Uncharacterized protein	0.779	0.648	0.43
Q8IDS7	Phosphate transporter	0.962	0.852	0.428
Q8I6U8	Glycophorin-binding protein	0.406	0.358	0.418
Q8I562	Clustered-asparagine-rich protein	0.557	0.644	0.413
Q8IAP4	GTPase-activating protein, putative	0.615	0.538	0.403
Q8I668	Uncharacterized protein	0.838	0.903	0.394
A0A143ZWL6	Uncharacterized protein	0.384	0.403	0.384
C6KSP3	Uncharacterized protein	0.816	0.647	0.377
Q8IB48	Uncharacterized protein	0.685	0.555	0.371
Q8IJ76	Early transcribed membrane protein 10.2	0.786	0.528	0.369
Q8IEK8	Uncharacterized protein	0.597	0.644	0.349
Q8IEC3	N2,N2-dimethylguanosine tRNA methyltransferase, putative	0.342	0.232	0.263

Figure 49 Down-regulated proteins in CQ treated samples

5.4. Patterns of synonymous codon usage in stress-dependent protein expression

After selecting the most up-regulated and down-regulated proteins at each time point (top 30 with cut-off >1.3 -fold, bottom 30 with a cut-off <0.7), we carried out a multivariate statistics analysis to investigate the relationships among the stage-specific proteomes based on their codon usage patterns (Figure 50). The gene specific codon usage count (GSCU) for the *Plasmodium falciparum* genome was provided to us by the Begley Lab. Using the codon frequency counts of the genes as the variables and the corresponding proteome expression data as the response we carried out a partial least squares regression (PLS-R). A PLS-R analysis combines a principal component analysis along with multiple regression. It is a particularly useful tool to predict a set of dependent variables from a relatively large set of independent variables. The predictors in our model are the fold change values from the proteomics results and the variables are the genome wide codon frequency values of the respective genes. This analysis helps to set up a model where the relative contribution of the codon bias of specific amino acid isoacceptors can help discern a translational model for the parasite i.e. particular isoacceptors are associated with increased or decreased protein abundance and a set of isoacceptors that do not contribute to differential translation.

Essentially, up-regulated proteins in both DHA and CQ were distinguished from those that went down with respect to their synonymous codon usage. The correlation loadings plot reveals the contributing codon usages for the discrimination DHA (PC-1: 16%, PC-2: 8%) and CQ (PC-1: 8%, PC-2: 8%) (Figure 50). As such for DHA, particular synonymous codons for 8 amino acids are enriched in up-regulated proteins in the stress-induced samples - CysTGT, TyrTAT, HisCAT, SerAGT, PheTTT, ArgCGA, AsnAAT, and ProCCT. Leaving aside

ArgCGA, note that all codons end with T in the wobble position. ArgCGA, ProCCT, and SerAGT are six box codons and have a corresponding tRNA gene with Watson-Crick anticodon pairing (T:A) for decoding. The other five codons however, are two box codons and are decoded by G wobbling in the anticodon loop of the tRNA. For a subsequent discussion, however, it is also important to note that these are also over represented (preferred) codons in the *Plasmodium falciparum* genome. For CQ however it seems to be a different distribution whereby the 7 codons that are over used in the up-regulated – AsnAAC, PheTTC, HisCAC, ProCCA, TyrTAC, ValGTT, and SerTCA. Of these, the two-box anticodons AsnAAC, PheTTC, HisCAC, TyrTAC are all ‘C’ ending codons all decoded by canonical Watson-Crick G:C pairing (Appendix III).

Figure 50 PLS-R scores (left) and Loadings (right) plot for DHA (A and B) and CQ (C and D) treated samples

(A) and (C) Scores plot and (B) and (D) loadings plot of a PLS-regression analysis of the gene-specific codon count from the most up- and down-regulated proteins in the DHA and CQ treated parasites. Proteins that are up and down regulated are in blue and red icons

respectively. The ellipse represents Hotelling T2 limit at p-value of 0.01 (F-test). The lines in the loadings plot connect anti-correlated codon pairs which distinguish the up- and down-regulated proteins.

5.5. Discussion

Recent studies undertaken to dissect out gene expression in *Plasmodium* seem to suggest that locus-specific, promoter-driven transcriptional control is not the overarching regulatory mechanism [177]. Global transcriptome and proteome studies across the life-cycle imply a tightly controlled “just-in-time” expression system [45, 47, 57, 174, 203]. With very few transcription factors identified, the mechanism underlying this tight control remains a puzzle and unconventional as compared to model eukaryotes. On the whole, our knowledge of epigenetic, transcriptional, and translational control mechanisms in *Plasmodium* and their impact on gene expression is severely limited as compared to other eukaryotes.

Several labs have attempted to investigate exogenous stress perturbations and their effect on the *Plasmodium* transcriptome. However, it is increasingly apparent that these studies detect few changes at the transcript level and often have conflicting but not mutually exclusive results. Upon treatment with WR99210, a well characterized anti-folate, very minute alterations were detected in the folate biosynthesis pathway at the RNA level [52]. In addition, treatment with 10 small molecules inhibitors at the early or late stages of the parasite seem suggest that the parasite does not have the capacity to respond specifically to these compounds [144]. Two orthogonal methods, microarray and serial analysis of gene expression (SAGE), were able to interpret only low amplitude transcriptional changes in chloroquine treated parasites [237]. Exposure of parasites to a choline analog, T4, documented a significant change in transcription

(2-69 fold changes), for genes involved in sexual differentiation but only after a 24 to 36 hours [238]. No changes were documented for genes in the lipid biosynthetic pathway but a strong arrest of genes involved in cell cycle progression was observed. Microarray based studies of synchronized cultures exposed to a variety of perturbations like depletion of hypoxanthine [239], serum, and glucose, and exposure to chloroquine and methyl methane sulfonate revealed dramatic changes in transcript abundance. These changes do not appear to be random but rather form the basis of a programmed response from the parasite. In this study we set out to characterize the toxicants that inflict oxidative damage predicated on physiological niches encountered by the parasite. We hypothesized that the cell proteome mounted in response to these toxicants is partly governed by tRNA modifications functioning as a translational control system.

The lifecycle IC_{50} values, which form the basis of most drug sensitivity assays, obtained in our study are close in agreement with literature for most of the toxicants. The deviation from literature values can be attributed to several factors. Parasite seeding densities are not uniformly observed as per the methods section for different studies by different laboratories. The medium composition for some of the studies has variation, particularly for the percentage of Albumax/serum support that is utilized which results in differential drug availability. For toxicants like hydrogen peroxide and sodium hypochlorite we observe time-dependant degradation of compounds and control for this variability by making use of a spectrophotometer to translate the absorbance value into actual concentration by using the coefficient of extinction. Menadione can also have different solvents to prepare the stock solutions which might impact the variability measured. However, all of the values obtained in our assay are well within range of reported values in literature. It was interesting to observe

that for some of the toxicants, the short exposure IC₅₀ values were less than a log fold difference from the lifecycle IC₅₀ values suggesting that these might be extremely fast acting compounds.

We used the short-exposure IC₅₀ values to form the basis of our study documenting tRNA modification profiles upon treatment with these various toxicants. As per our hypothesis, we propose that the system of tRNA modifications is dynamic and responds to these toxicants to coordinate targeted expression of a stress-response proteome. The first thing that is strikingly obvious is that the pattern of tRNA modification changes is stress-specific and shows a signature for each of these toxicants that is disparate i.e. not all oxidative stress is equal. For example, studies that employ hydrogen peroxide as a surrogate to mimic stress experienced by artemisinin might have shortcomings, as not only are their molecular mechanisms different, but the cellular response is as well. In general, we observed a drop in the values for most of the modifications for each of the treatments using the different modifications. These changes can also be explained as a result of altered growth rates. As evidenced by our data on the lifecycle tRNA modifications, there is a gradual increase in 22 of the 28 modifications as the parasite progresses from ring to schizont stage. This could explain the quantitation levels whereby the untreated control has progressed temporally, but the treated cells have either arrested or slowed down in their lifecycle progression. While we think this is unlikely based on the granular time-scale of the lifecycle tRNA modification data, we propose to use additional controls at the start of the treatment. A comparison of modification levels at the start of the treatment for the untreated control would indicate if there is a corresponding increase in the three hours as per the experimental time point. Of interest are also modifications which increase upon treatment with chloroquine and sodium hypochlorite which cannot be explained with just the stall in growth rates.

An initial school of thought portrays RNA modifications as a relatively static, with their “response” being a result of active accumulation of hypomodified tRNA and simultaneous degradation of modified mature tRNA. This contrasts with nutritional stress studies that show repression of polIII which is required for tRNA transcription and as a consequence there is no active “synthesis” of tRNAs. In addition, this view has been reworked where studies in several model organisms produce signature responses in the tRNA modification profile upon induction of a variety of stressors. For example, alkylating reagents like methyl methanesulfate cause a characteristic change in tRNA modifications in mammalian HeLa cells whereby m³C, mcm⁵U, m⁷G, and mcm⁵s²U increase relative to untreated cells. In a different signature pattern, oxidising agents induce an increase in i⁶A, m²²G, m⁵C, and nncm⁵U which remain relatively unchanged with the aforementioned alkylating agents [240].

For a lot of these modifications, the loss of enzymes catalysing their formation or “writers” might not display a phenotype in normal growing conditions. However, upon introducing these organisms to a stressed environment there is an observable fitness cost. Several studies describe deletion mutants of cytoplasmic tRNA methyltransferase which forms m⁵C at different sites on the tRNA. The loss of m⁵C writer enzyme, NSUN2, resulted in altered biogenesis of tRNA-derived fragments (tRFs) indicating that this modification is protective to endonuclease cleavage in mammalian cells. These tRFs in turn then triggered a cell stress response with a global repression of protein synthesis [241]. A different enzyme DNMT2 which carries out the same m⁵C modification on tRNAs was also shown to protect against oxidative stress induced tRNA endonuclease cleavage in *Drosophila*. The deletion of Trm4 (yeast homolog for NSUN2) sensitized *S. cerevisiae* to H₂O₂ [86]. This observation was linked to a previous study where wobble m⁵C had been shown to be important for translation of UUG codons over represented in oxidative stress response genes providing us with an alternative hypothesis for

the role of tRNA modifications in cell stress response. A recent study in our lab was able to demonstrate that in *Mycobacterium bovis* exposed to hypoxia led to an increase in the wobble modification cm⁵U at the wobble position of tRNA^{ThrUGU} [97]. This codon was found to be over represented in DosR which is the master regulator for hypoxic bacteriostasis and using codon reengineering, the author was able to ascribe a stress-specific role for this modification. An unpublished study in our lab (Cai et al) describes the role of i⁶A catalysed by the MiaA enzyme in mediating stress-induced antibiotic resistance in *Mycobacterium smegmatis*. The deletion of writer enzyme for m¹A in *Thermus thermophilus* bacteria resulted in thermosensitivity for this otherwise hyperthermophilic organism as evidenced by the name. Additionally, deletion of TrmB and TruB catalysing for m⁷G at pseudouridine led to growth defects at lower temperatures since these mutants were unable to mount an effective cold-shock response [242].

Moving on to the proteomics experiments, we used a stringent filtering criteria for the identification of peptides and obtained about 25% coverage of the ~5300 size genome for the proteomics experiment. To the best of our knowledge, there is only one other study that has studied the stress-response at the protein level for some of the toxicants employed. There exist some individual studies that take a chemoproteomics approach to artemisinin stress to find drug binding-targets but do not study the physiological basis of drug action. Of interest here is that artemisinin and chloroquine treatment seem to display opposing trends when it comes to ribosomal proteins. In DHA-treated samples these proteins seem to be down-regulated, but there are a number of ribosomal proteins that are up-regulated in the CQ-treated samples. It might be worthwhile to investigate for artemisinin if the cell machinery responds by arresting cell cycle progression, since the proteins that are enriched in the treated samples are transcribed early in the lifecycle of the parasite.

We carried out a PLS-R analysis using the codon-usage pattern for proteins that were up-regulated at the IC₈₀ treatment using the fold-change as the response variable. For DHA, synonymous codons for 8 amino acids are enriched in up-regulated proteins in the stress-induced samples - CysTGT, TyrTAT, HisCAT, SerAGT, PheTTT, ArgCGA, AsnAAT, and ProCCT. Leaving aside ArgCGA, all codons end with T in the wobble position. ArgCGA, ProCCT, and SerAGT are six box codons and have a corresponding tRNA gene with Watson-Crick anti-codon pairing (T:A) for decoding. The other five codons however, are two box codons and are decoded by G wobbling in the anticodon loop of the tRNA. These are also overrepresented (preferred) codons in the *Plasmodium falciparum* genome. If we are to go by a hypothesis where cell cycle arrest is induced upon artemisinin treatment, this would be an expected pattern and would not form sufficient basis for the involvement of tRNA modifications regulating translation. The modification that was relevant upon artemisinin treatment is t⁶A which is known to be present at position 37 for tRNAs decoding ANN codons. A reduction in this modification could lead to mistranslation by frameshifting [166]. This mistranslated aborted proteins, however, would not be picked up in our analysis as a limitation of the proteomics technique. The other possibility is that the addition of this modification to tRNA requires ATP for a substrate which could also be depleted under these stressed conditions.

For CQ however it seems to be a different distribution whereby the 7 codons that are overused in the up-regulated proteins are AsnAAC, PheTTC, HisCAC, ProCCA, TyrTAC, ValGTT, and SerTCA. Of these, the two-box anticodons AsnAAC, PheTTC, HisCAC, TyrTAC are all 'C' ending codons all decoded by canonical Watson-Crick G:C pairing. None of these codons are presumably modified as evidenced in other organisms leading us to speculate as our original

framework. A crucial point to take into consideration is that these PLS-R analyses are only preliminary and can be further refined. An RNA-seq study carried out for the same treatments will provide us the information to carry out translational efficiency calculations. This will allow us to dissect out the transcripts that are controlled at a primary transcript abundance level as opposed to a translational level and have us pursue a codon-biased translation hypothesis. An analysis of other treatment conditions and a curated list of the proteomics data would also help us in dissecting out stress response pathways that are either exclusive or hard-wired (as evidenced by the transcriptome) to each toxicant.

Conclusion

The aims of the research offered in this thesis were to examine the role of RNA modifications in general as a mechanism of translational control in *Plasmodium falciparum*. Specifically, we have surveyed the system of these modifications in the messenger RNA and transfer RNA in the parasites' intra-erythrocytic developmental cycle and response to stress.

While a first pass survey of tRNA modifications was carried out by Dr. Chee Sheng Ng, this project takes the preliminary work in a more decisive direction. The basis for the first of these studies involved setting up an analytical platform in tandem with a bioinformatics strategy for characterizing and quantifying modifications on the wobble position of specific tRNA isoacceptors (**Chapter 3**). These modifications that sit on the anticodon loop in the tRNA have the potential to directly interact and influence codon-anticodon pairing. Their hyper- or hypomodification could lead to variety of outcomes that are reflected in several translational errors in the cell – frameshifting, promiscuity in codon choice, protein folding dependent on the speed of translation, etc. Our method was confidently able to assign mcm⁵s²U to the wobble position of the anticodon loop for tRNA^{GlyCAA} and tRNA^{GlnGAA} and mcm⁵U to tRNA^{GluGGA}. We quantified how these modifications on these specific tRNA isoacceptors change as a function of the lifecycle of the parasite and were able to correlate these changes to a proteome that appeared codon-biased for these exact isoacceptors. This correlative piece of evidence has been mechanistically described in genetically tractable organisms using a combination of transcriptomics, proteomics, and codon-reporter studies. A classical genetics approach involving knockouts and overexpressor mutant lines in *Plasmodium* would help solidify our hypothesis.

A natural progression to this exercise was to set up a platform for discovery and quantitation of mRNA modifications in *P. falciparum* (**Chapter 4**). While the method setup was quite akin to our tRNA modifications platform, the challenge was working with relatively pure mRNA at nanogram quantities. Once this platform was established, we were able to define a total of 10 mRNA modification candidates. These were pseudouridine (Ψ), N⁵-methylcytidine (m⁵C), N^[(3or4)]-methylcytidine (mC), N⁷-methylguanosine (m⁷G), 2'-O-methylcytidine (Cm), N⁵-methyluridine (m⁵U), 2'-O-methyladenosine (Am), 2'-O-methyladenosine (Gm), N⁶-methyladenosine (m⁶A), and N²,2'-O-dimethyladenosine (mAm). We monitored these ten mRNA modifications across the parasite IDC across six timepoints to capture their dynamic. Combining our observations as an abundant modification and the recent literature in the field, we focused our effort in characterizing m⁶A dynamics in *Plasmodium* mRNA. A thorough structural characterization using isotope standards and high-resolution mass spectrometry definitively identified mRNA-derived m⁶A in *Plasmodium*. Using orthogonal approaches of m⁶A-seq and LC-MS, m⁶A-associated transcripts were found to have shorter mRNA half-lives and also lower translational efficiency. Whether these effects translate into phenotypic changes in the parasite would be an avenue for future work. Not only did we describe the modification, we also characterized the writer enzyme complex of PfMT-A70 (mammalian homolog of METTL3) with two other critically conserved members PfMT-A70.2 (METTL14 homolog) and PfWTAP. These writer proteins were found to be located in the nucleus suggesting that m⁶A methylation is a pre-export event and a knockdown using a CRISPR interference approach led to an observable decrease in m⁶A methylation.

This thesis also presents a perspective on the involvement of tRNA modifications in modulating stress response in the parasites (**Chapter 5**). The first observation from this study was that each oxidative stressor induced a characteristic response in tRNA modifications which

were mutually exclusive. The second part of this study carried out a proteomics-based survey of stress-induced *Plasmodium* parasites. Using a combination of proteomics and multivariate statistics we investigate the relationship of codon-bias in gene expression and tRNA modification reprogramming. We found that for DHA, synonymous codons for 8 amino acids are enriched in up-regulated proteins in the stress-induced samples - CysTGT, TyrTAT, HisCAT, SerAGT, PheTTT, ArgCGA, AsnAAT, and ProCCT. For CQ however it seems to be a different distribution from DHA whereby the 7 codons that are overused in the up-regulated proteins are AsnAAC, PheTTC, HisCAC, ProCCA, TyrTAC, ValGTT, and SerTCA. While this study remains preliminary, our observations seem encouraging in investigating codon-biased translation as one of the mechanisms for stress response in parasites.

There are several missing facets that are needed to get a clearer picture of the role of the epitranscriptome in not only *Plasmodium*, but also other organisms and a systems approach is almost indispensable to studying the epitranscriptome. Transcriptional and translational regulation involves multiple layers of control where RNA stability, localization, fidelity, translational initiation, elongation, and termination are needed to be assessed *in toto*. A rather understudied topic in tRNA modification field is the crosstalk between different modifications on the same tRNA. It has been observed that DNMT2 (which methylates cytosine) activity on the tRNA is dependent upon the presence of another modification queuosine (Q) at the wobble position [243]. This is posited to work in a three-dimensional context where queuosine would induce a conformational change making otherwise hidden sites more accessible. Not only is there crosstalk between different modifications on the same tRNA, there is also to consider the promiscuity of enzymes associated with these modifications. RlmN is a dual-specificity methyltransferase that can methylate adenosines on both tRNAs and ribosomal RNAs [244]. The same can be described for NSUN2, which can target tRNAs, rRNAs, and mRNAs [241].

Whether this is a collateral and artefactual effect of the secondary structure of the RNAs and diffused localization of the enzyme or whether this has functional implications forms the basis of future research. In addition, it almost follows a standard model of translation that tRNA modifications in the ASL must at some point interact with the modifications in the mRNA and rRNA during translation. Always studied as standalone modifications, a research question formulates at this precise interface on how these three distinct entities would influence the role of one another. What follows is also how tRNA derived fragments serve regulatory roles as small RNAs and this study requires a sophistication of bioinformatics and ‘omics technologies that we now possess. A translational aspect that has largely been ignored in this thesis but becomes increasingly relevant is the role these modifications play in designing RNA therapeutics. The involvement of some of these writer, reader, and eraser enzymes in diseases such as cancer have strong correlations and are currently being considered as therapeutic targets by several research labs and companies alike.

Figure 51 Deciphering the *Plasmodium* epitranscriptome

Specific mRNA transcripts in Plasmodium falciparum undergo methylation by the “writer” complex consisting of the PfMT70, PfMT70.2, and PfWTAP proteins in the nucleus. Whether this process is co-transcriptional is still to be elucidated. This modification binds to YTH “reader” proteins which have both a nuclear and cytoplasmic localization. In the cytoplasm, these m⁶A transcripts are associated with shorter mRNA stability and lower translational efficiencies. We speculate this is a result of YTH reader proteins interacting with themselves or other translational factors to form protein granules. On the translational

apparatus, wobble-modified tRNA decode codon-biased transcripts to respond to the needs of the cell cycle. Our work further asks whether there operates a system of tRNA modification reprogramming to preferentially translate stress response proteins. Image composed in Biorender and used with permission.

The data obtained so far strongly suggest that RNA modifications play an important role in parasite biology by regulating stage-specific gene expression at the level of translation. This implies that the parasite has evolved a finely tuned regulatory mechanism to coordinate these RNA modifications. Importantly, it also suggests a key shift in terms of regulatory control of expression from transcriptional regulation to translational regulation and may explain the relative lack of transcription factors identified in the *Plasmodium* genome to date. Altogether this study provides us with fundamental new understanding on the role of RNA modifications in the developmental lifecycle and stress response of the parasite. The results presented in this thesis aspire to offer better scientific resources in the field of parasitology and further our endeavor to eliminate a severe disease – Malaria.

Future Work

6. Overview

In this chapter, we identify the limitations of the current study and propose several experiments that could be undertaken to address these issues and take this project further to dissect out *Plasmodium* biology.

6.1. Identify and characterize RNA-modifying enzymes important for tRNA modifications and parasite biology

Our initial findings support the dynamic reprogramming of tRNA modifications in *P. falciparum* as a function of the IDC and stress responses. There are implications that these modifications may sit downstream of a cascade that enables the parasite to survive drug treatment and immune stress, thus manifesting the phenotype and altering translational activities. A way of testing these hypotheses is to approach it through classical genetic approaches. While the current *P. falciparum* 3D7 genome has about 59 genes with putative tRNA-modifying features, these annotations have not been empirically validated and the functions thus remain unknown. There also exist missing tRNA-modifying homologs for some of the modifications, which is due to low sequence homology of *Plasmodium* with other model organisms. While in the long run it would be ideal to genetically evaluate the role of all the different putative tRNA-modifying enzymes for this proposal, we aim to take a more targeted approach. The data obtained from the IDC and stress-response studies will provide us with a selection of tRNA modifications that are implicated in the regulation of developmental progression or stress response.

For example, mcm⁵U appears to be important in the parasite lifecycle based on a preliminary data. Biosynthetic pathways detailing the enzymes involved in the modification of this nucleoside can be mined from literature or using resources like Modomics [73]. In this case, Trm9 in *S. cerevisiae* modifies the canonical uridine to mcm⁵U. Homology modelling using PlasmDB indicates that this enzyme is closely related to the gene PF3D7_102800 in *P. falciparum*, whereby the *Plasmodium* gene is severely truncated at the N-terminus compared to the yeast homolog. Nevertheless, the residues that are essential in the enzymatic process are still conserved. The next step involves cross-referencing gene essentiality in *Plasmodium* by scouting the Plasmogem database [245], the transposon library, or work carried out in related organisms like *Toxoplasma gondii*. Based on the transposon library data, it appears the gene is amenable to disruption using one of the knock-out approaches. The homologous gene in the murine malarial parasite (*P. berghei*), however, indicates that there is a fitness cost that is incurred with disrupting this gene. This provides us with a putative phenotype that we may investigate.

We would like to use classical genetic approaches to disrupt these enzymes and evaluate their functionality in different contexts. As a first pass we would like to evaluate if these enzymes are essential by trying to use recently developed and characterized techniques for targeted disruption. In order to overcome the poor efficiency of homologous recombination in *P. falciparum*, Birnbaum et al developed a method permitting the selection of parasites with only genomic integrations [246]. Alternatively, we would also attempt to use a CRISPR-based strategy that was recently adapted for *Plasmodium falciparum* and has been used successfully in our lab [247, 248].

Figure 52 Schematic to identify and characterize putative *Plasmodium* RNA-modifying enzymes.

Using reference organism genomes for homology searching, we identify candidates and utilize several genetic strategies to disrupt their functionality using CRISPR or selection-linked integration. These mutants are then characterized at a molecular scale to see changes in tRNA modifications and protein homeostasis and to see larger phenotypic changes in survival, stress-response, and drug-resistance.

Upon obtaining these parasites with disrupted tRNA modifying enzymes we would proceed to characterize the phenotype of the parasite. The first would involve assessing the impact on tRNA modifications and seeing the specificity of these enzymes. Based on data from other organisms, some RNA-modifying enzymes are known to be promiscuous and it would be of important to map which tRNA modifications change and map them to their location on the

specific isoacceptor. Using LC-MS proteomics we would then assay the translational efficiency of different proteins for these parasite lines. This would enable us to conclusively tie in the specific tRNA modification to a phenotypic response by the parasite. Subsequently, we will evaluate these mutant parasite lines in the context of different stressors using dose-dependent assays at different stages of the parasite's lifecycle. In the event of a phenotype we would also generate episomal overexpressors to perform rescue experiments for the gene of interest.

6.2.Determine translation rates of codon biased genes during IDC

Our data on tRNA modifications in the *Plasmodium falciparum* IDC suggest that increased mcm⁵U and ncm⁵U levels promote the translation of late stage transcripts which have a distinctive codon usage. In order to test whether tRNA wobble modifications affect decoding capacity and translation efficiency, we propose to use the dual luciferase reporter system which has traditionally been used for quantifying promoter influences on gene expression. This system utilizes simultaneous expression and measurement of two individual reporter enzymes. One of the reporters measures the effect of specific experimental conditions, while the activity of the co-reporter provides an internal control that functions as the reference response. Consequently, we would have the “experimental” reporter synthesized as gene blocks with specifically biased and anti-correlated codons that pair specifically with our tRNA modification results. In order to ensure robustness, we would design the reference reporter gene with similarly codon biases and carry out the assay in a reciprocal manner. The table below has been mined based on the preliminary data we have obtained from the proteomics experiment of the *Plasmodium* IDC that was carried out.

Amino acid	Enriched synonymous codons in TP 40, 46			Anti-correlated (under-represented) codons		
	Synonymous codon	Cognate tRNA	Interaction	Synonymous codon	Cognate tRNA	Interaction
Arg	AGA	tRNA ^{Arg} (mcm5UGU)	Modified anticodon A3:mcm5U34	AGG	tRNA ^{Arg} (CCU)	Watson-Crick G3:C34
Gly	GGA	tRNA ^{Gly} (mcm5UCC)	Modified anticodon A3:mcm5U34	GGT	tRNA ^{Gly} (GCC)	Wobble codon U3:G34
Glu	GAA	tRNA ^{Glu} (mcm5s2UUC)	Modified anticodon A3:mcm5s2U34	GAG	tRNA ^{Glu} (CUC)	Watson-Crick G3:C34
Gln	CAA	tRNA ^{Gln} (mcm5s2UUG)	Modified anticodon A3:mcm5s2U34	CAG	tRNA ^{Gln} (CUG)	Watson-Crick G3:C34
Pro	CCA	tRNA ^{Pro} (ncm5UGG)	Modified anticodon A3:ncm5U34	CCT	tRNA ^{Pro} (AGG)	Watson-Crick U3:A34
Ser	TCA	tRNA ^{Ser} (ncm5UGA)	Modified anticodon A3:ncm5U34	AGT	tRNA ^{Ser} (GCU)	Wobble codon U3:G34
Phe	TTC	tRNA ^{Phe} (GmAA)	Modified anticodon C3:Gm34	TTT	tRNA ^{Phe} (GAA)	Wobble codon U3:G34
Thr	ACT	tRNA ^{Thr} (IGU)	Modified anticodon U3:I34	ACA	tRNA ^{Thr} (ncm5UGU)	Modified anticodon A3:ncm5U34
Ile	ATC	tRNA ^{Ile} (IAU)	Modified anticodon C3:I34	ATA	tRNA ^{Ile} (ΨAΨ)	Modified anticodon A3:Ψ34

Figure 53 Enriched synonymous codons and their under-represented counterparts in late-stage (TP 40 and 46) up-regulated proteins.

The listed codons on the left panel are more frequently used in late trophozoite/ schizont (TP 40 and 46) proteins in relative to ring/ trophozoite (TP 8 – 32) proteins. 9 out of 13 of the over-used codons are decoded by wobble modified cognate tRNA iso-acceptors (shaded in aqua). Their anti-correlated (under-represented) synonymous codons are listed in the right panel. These under-represented codons are decoded by cognate tRNA species which are not known to be modified at wobble position, with tRNA^{Thr}(ncm5UGU) and tRNA^{Ile}(ΨAΨ) as exceptions.

We plan to generate episomally transfected lines wherein both Firefly and Renilla luciferase constructs designed previously are on a single vector backbone separated by a the T2A (skip) peptide. The viral 2A-like peptide has been shown to be functional in *Plasmodium* and prevents peptide bond formation between two specific and adjacent amino acids during translation. This results in the production of two discrete proteins from a single expression cassette which would enable us to normalize the translational efficiency of both of those proteins. It would also serve

to control for transcriptional variance arising from the use of two different plasmids. A plasmid designed with the reciprocal usages of Firefly and Renilla luciferase would also serve as additional control. This approach could be taken for transient or stable transfections. An outcome of this assay would also enable us to visualize the impact of codon usage as the parasite progresses through its lifecycle. Our hypothesis would suggest that translational efficiency during the ring stage of the parasite is not significantly different for the codon-biased genes. This difference however would become manifestly larger as the parasite progresses to the trophozoite and schizont stages and the cognate tRNAs are modified.

As an orthogonal approach, we also intend to leverage an in-vitro translation system that was developed by Ah Yong et al [249]. This assay uses late-stage parasite lysates containing natively modified tRNA and independently transcribed mRNAs using a T7 polymerase. For the assay, these transcripts would be generated from our designed biased constructs encoding for Firefly and Renilla reporters. A kinetic experiment demonstrating the accumulation of translated protein from both of these reporters over time would enable us to determine the influence of codon usage patterns on translational efficiency.

Figure 54 Impact of codon biased genes on translation

Normal and codon biased firefly luciferase will be transfected into parasites and their expression will be normalized to control renilla luciferase. Expected outcome would be increased translation of codon biased firefly luciferase in late stage parasites. This approach would be similar for both transfected genes as well as those analysed in the in vitro translation system.

6.3. Discovery of new modifications in *Plasmodium falciparum* RNA

The first survey of tRNA modifications in *P. falciparum* tRNA detected 28 modifications. However, this study was carried out using an available set of standards and based on predictable fragmentation patterns. The method did not detect queuosine (Q), wybutosine (yW) and their derivatives despite the presence of their putative modifying genes in the *P. falciparum* genome. Upon optimizing for the detection of these modifications, I was able to gain data that indicated the presence the queuosine in *Plasmodium* tRNA and found it greatly enriched in gametocytes in a preliminary survey (Figure 54). Similarly, the original study was unable to find ncm⁵Um which is known to occur on eukaryotic tRNA^{LeuUUA} on the wobble position [250]. While not undermining the findings of the original study, these are the limitations of the analytical method that we had initially employed. A more rigorous analysis of both the tRNA and mRNA modification landscape could determine analytes that are *Plasmodium*-specific. Not only would these modifications be explored to study parasite specific biology, they could also be used as biomarkers for infection or design of therapeutic interventions. This strategy is also applicable to investigating modifications in the ribosomal RNA, which in some systems is well characterized to confer resistance to a class of antibiotics.

A

B

Figure 55 Queuosine is present in *P. falciparum* tRNA

(A) High resolution mass spectrometry shows CID fragmentation of the endogenous tRNA queuosine (Q) precursor (m/z 410.1877) alongside an isotopic standard (m/z 427.2302) to reveal the unusual fragmentation pattern. (B) Extracted ion chromatogram showing abundance of Q normalized to total guanosine content in the *P. falciparum* tRNA for

Gametocytes (Gam) and the 48 hour asexual lifecycle at six timepoints (TP6, TP12, TP18, TP24, TP30, TP36, TP42).

6.4. Translational control mechanisms mediating artemisinin resistance

Current efforts to reduce the global impact of malaria are threatened by the emergence of *Plasmodium falciparum* (Pf) resistance to artemisinins (ART), the core component of current first-line antimalarial therapies. ART resistance, which has compromised the efficacy of antimalarial treatments across Asia and threatens Africa, is mediated by mutations in the kelch-like 13 protein (PfK13) that protect against ART's proteotoxic and oxidative activity [38-42, 233]. We propose to test the novel hypothesis that mutant PfK13 achieves resistance by enabling the reprogramming of tRNA modifications, leading to codon-biased "just in time" changes in protein translation and parasite survival in response to ART-induced cellular stress. Our hypothesis proposes that the parasite may alter its tRNA modification profile to selectively alter translation of particular codons, via wobble modifications, allowing rapid translational changes. Alternatively, PfK13 mutant parasites may hypomodify the wobble base s²U modification, leading to ribosomal stalling and quiescence. Using a systems biology approach (Figure 55) we would like to study if and how tRNA modifications are essential in mediating a response to artemisinin resistance in ring stage parasites.

Systems Biology approach to study translational control and tRNA modifications

Reference:

1. Ng CS, Sinha A, Aniweh Y, Nah Q, Babu IR, Gu C, Chionh YH, Dedon PC, Preiser PR: **tRNA epitranscriptomics and biased codon are linked to proteome expression in *Plasmodium falciparum*.** *Mol Syst Biol* 2018, **14**(10):e8009.
2. Baumgarten S, Bryant JM, Sinha A, Reyser T, Preiser PR, Dedon PC, Scherf A: **Transcriptome-wide dynamics of extensive m(6)A mRNA methylation during *Plasmodium falciparum* blood-stage development.** *Nat Microbiol* 2019.
3. Ashley EA, Pyae Phyo A, Woodrow CJ: **Malaria.** *Lancet* 2018, **391**(10130):1608-1621.
4. Organization WH: **World Malaria Report 2016.** In. Geneva: World Health Organization; 2016.
5. Malleret B, Li A, Zhang R, Tan KS, Suwanarusk R, Claser C, Cho JS, Koh EG, Chu CS, Pukrittayakamee S *et al*: ***Plasmodium vivax*: restricted tropism and rapid remodeling of CD71-positive reticulocytes.** *Blood* 2015, **125**(8):1314-1324.
6. Cogswell FB: **The hypnozoite and relapse in primate malaria.** *Clin Microbiol Rev* 1992, **5**(1):26-35.
7. Kantele A, Jokiranta TS: **Review of cases with the emerging fifth human malaria parasite, *Plasmodium knowlesi*.** *Clin Infect Dis* 2011, **52**(11):1356-1362.
8. Miller LH, Mason SJ, Clyde DF, McGinniss MH: **The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy.** *N Engl J Med* 1976, **295**(6):302-304.
9. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, Tatem AJ, Hay SI: **A new world malaria map: *Plasmodium falciparum* endemicity in 2010.** *Malar J* 2011, **10**:378.
10. Siden-Kiamos I, Louis C: **Interactions between malaria parasites and their mosquito hosts in the midgut.** *Insect Biochem Mol Biol* 2004, **34**(7):679-685.
11. Tilley L, Dixon MW, Kirk K: **The *Plasmodium falciparum*-infected red blood cell.** *Int J Biochem Cell Biol* 2011, **43**(6):839-842.
12. Mota MM, Rodriguez A: **Migration through host cells: the first steps of *Plasmodium* sporozoites in the mammalian host.** *Cell Microbiol* 2004, **6**(12):1113-1118.
13. Cowman AF, Crabb BS: **Invasion of red blood cells by malaria parasites.** *Cell* 2006, **124**(4):755-766.
14. Keeley A, Soldati D: **The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa.** *Trends Cell Biol* 2004, **14**(10):528-532.
15. Scholtyseck EP, G. : **Elektronenmikroskopische Untersuchungen über die Merozoiten von *Eimeria stiedae* und *E. perforans* und *Toxoplasma gondii*. Zur systematischen Stellung von *T. gondii*.** *Z Parasitenk* 1962, **26**:91-115.
16. Kirk K: **Membrane transport in the malaria-infected erythrocyte.** *Physiol Rev* 2001, **81**(2):495-537.
17. Rudzinska MA, Trager W, Bray RS: **Pinocytotic uptake and the digestion of hemoglobin in malaria parasites.** *J Protozool* 1965, **12**(4):563-576.
18. Arnot DE, Gull K: **The *Plasmodium* cell-cycle: facts and questions.** *Ann Trop Med Parasitol* 1998, **92**(4):361-365.
19. Hodder AN, Drew DR, Epa VC, Delorenzi M, Bourgon R, Miller SK, Moritz RL, Frecklington DF, Simpson RJ, Speed TP *et al*: **Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of *Plasmodium falciparum* SERA5.** *J Biol Chem* 2003, **278**(48):48169-48177.

20. NJ W: **Malaria**: Elsevier; 2013.
21. Mackintosh CL, Beeson JG, Marsh K: **Clinical features and pathogenesis of severe malaria**. *Trends Parasitol* 2004, **20**(12):597-603.
22. Berendt AR, Ferguson DJ, Newbold CI: **Sequestration in Plasmodium falciparum malaria: sticky cells and sticky problems**. *Parasitol Today* 1990, **6**(8):247-254.
23. David PH, Hommel M, Miller LH, Udeinya IJ, Oligino LD: **Parasite sequestration in Plasmodium falciparum malaria: spleen and antibody modulation of cytoadherence of infected erythrocytes**. *Proc Natl Acad Sci U S A* 1983, **80**(16):5075-5079.
24. Taylor WRJ, Hanson J, Turner GDH, White NJ, Dondorp AM: **Respiratory manifestations of malaria**. *Chest* 2012, **142**(2):492-505.
25. Sharma L, Shukla G: **Placental Malaria: A New Insight into the Pathophysiology**. *Front Med (Lausanne)* 2017, **4**:117.
26. Trampuz A, Jereb M, Muzlovic I, Prabhu RM: **Clinical review: Severe malaria**. *Crit Care* 2003, **7**(4):315-323.
27. Hunt NH, Grau GE: **Cytokines: accelerators and brakes in the pathogenesis of cerebral malaria**. *Trends Immunol* 2003, **24**(9):491-499.
28. Wassmer SC, Grau GE: **Severe malaria: what's new on the pathogenesis front?** *Int J Parasitol* 2017, **47**(2-3):145-152.
29. White NJ, Turner GD, Day NP, Dondorp AM: **Lethal malaria: Marchiafava and Bignami were right**. *J Infect Dis* 2013, **208**(2):192-198.
30. Vaidya AB, Mather MW: **Atovaquone resistance in malaria parasites**. *Drug Resist Updat* 2000, **3**(5):283-287.
31. Price RN, von Seidlein L, Valecha N, Nosten F, Baird JK, White NJ: **Global extent of chloroquine-resistant Plasmodium vivax: a systematic review and meta-analysis**. *Lancet Infect Dis* 2014, **14**(10):982-991.
32. Edi CV, Djogbenou L, Jenkins AM, Regna K, Muskavitch MA, Poupardin R, Jones CM, Essandoh J, Ketoh GK, Paine MJ *et al*: **CYP6 P450 enzymes and ACE-1 duplication produce extreme and multiple insecticide resistance in the malaria mosquito Anopheles gambiae**. *PLoS Genet* 2014, **10**(3):e1004236.
33. Trape JF, Tall A, Diagne N, Ndiath O, Ly AB, Faye J, Dieye-Ba F, Roucher C, Bouganali C, Badiane A *et al*: **Malaria morbidity and pyrethroid resistance after the introduction of insecticide-treated bednets and artemisinin-based combination therapies: a longitudinal study**. *Lancet Infect Dis* 2011, **11**(12):925-932.
34. Foy BD, Alout H, Seaman JA, Rao S, Magalhaes T, Wade M, Parikh S, Soma DD, Sagna AB, Fournet F *et al*: **Efficacy and risk of harms of repeat ivermectin mass drug administrations for control of malaria (RIMDAMAL): a cluster-randomised trial**. *Lancet* 2019, **393**(10180):1517-1526.
35. Marshall JM, Taylor CE: **Malaria control with transgenic mosquitoes**. *PLoS Med* 2009, **6**(2):e20.
36. Rts SCTP: **Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial**. *Lancet* 2015, **386**(9988):31-45.
37. Sissoko MS, Healy SA, Katile A, Omaswa F, Zaidi I, Gabriel EE, Kamate B, Samake Y, Guindo MA, Dolo A *et al*: **Safety and efficacy of PfSPZ Vaccine against Plasmodium falciparum via direct venous inoculation in healthy malaria-exposed adults in Mali: a randomised, double-blind phase 1 trial**. *Lancet Infect Dis* 2017, **17**(5):498-509.

38. Menard D, Khim N, Beghain J, Adegnikaa AA, Shafiul-Alam M, Amodu O, Rahim-Awab G, Barnadas C, Berry A, Boum Y *et al*: **A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms**. *N Engl J Med* 2016, **374**(25):2453-2464.
39. Ashley EA, Dhorda M, Fairhurst RM, Amaratunga C, Lim P, Suon S, Sreng S, Anderson JM, Mao S, Sam B *et al*: **Spread of artemisinin resistance in Plasmodium falciparum malaria**. *N Engl J Med* 2014, **371**(5):411-423.
40. Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, Lwin KM, Arie F, Hanpithakpong W, Lee SJ *et al*: **Artemisinin resistance in Plasmodium falciparum malaria**. *N Engl J Med* 2009, **361**(5):455-467.
41. Arie F, Witkowski B, Amaratunga C, Beghain J, Langlois AC, Khim N, Kim S, Duru V, Bouchier C, Ma L *et al*: **A molecular marker of artemisinin-resistant Plasmodium falciparum malaria**. *Nature* 2014, **505**(7481):50-55.
42. Straimer J, Gnädig NF, Witkowski B, Amaratunga C, Duru V, Ramadani AP, Dacheux M, Khim N, Zhang L, Lam S *et al*: **Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates**. *Science* 2015, **347**(6220):428-431.
43. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S *et al*: **Genome sequence of the human malaria parasite Plasmodium falciparum**. *Nature* 2002, **419**(6906):498-511.
44. Bohme U, Otto TD, Sanders M, Newbold CI, Berriman M: **Progression of the canonical reference malaria parasite genome from 2002-2019**. *Wellcome Open Res* 2019, **4**:58.
45. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL: **The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum**. *PLoS Biol* 2003, **1**(1):E5.
46. Hoo R, Zhu L, Amaladoss A, Mok S, Natalang O, Lapp SA, Hu G, Liew K, Galinski MR, Bozdech Z *et al*: **Integrated analysis of the Plasmodium species transcriptome**. *EBioMedicine* 2016, **7**:255-266.
47. Deitsch K, Duraisingh M, Dzikowski R, Gunasekera A, Khan S, Le Roch K, Llinas M, Mair G, McGovern V, Roos D *et al*: **Mechanisms of gene regulation in Plasmodium**. *Am J Trop Med Hyg* 2007, **77**(2):201-208.
48. Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL *et al*: **A proteomic view of the Plasmodium falciparum life cycle**. *Nature* 2002, **419**(6906):520-526.
49. Gunasekera AM, Myrick A, Le Roch K, Winzeler E, Wirth DF: **Plasmodium falciparum: genome wide perturbations in transcript profiles among mixed stage cultures after chloroquine treatment**. *Exp Parasitol* 2007, **117**(1):87-92.
50. Le Roch KG, Chung DW, Ponts N: **Genomics and integrated systems biology in Plasmodium falciparum: a path to malaria control and eradication**. *Parasite Immunol* 2012, **34**(2-3):50-60.
51. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ *et al*: **Discovery of gene function by expression profiling of the malaria parasite life cycle**. *Science* 2003, **301**(5639):1503-1508.
52. Ganesan K, Ponmee N, Jiang L, Fowble JW, White J, Kamchonwongpaisan S, Yuthavong Y, Wilairat P, Rathod PK: **A genetically hard-wired metabolic transcriptome in Plasmodium falciparum fails to mount protective responses to lethal antifolates**. *PLoS Pathog* 2008, **4**(11):e1000214.
53. Horrocks P, Wong E, Russell K, Emes RD: **Control of gene expression in Plasmodium falciparum - ten years on**. *Mol Biochem Parasitol* 2009, **164**(1):9-25.

54. Hu G, Cabrera A, Kono M, Mok S, Chaal BK, Haase S, Engelberg K, Cheemadan S, Spielmann T, Preiser PR *et al*: **Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum***. *Nat Biotechnol* 2010, **28**(1):91-98.
55. Foth BJ, Zhang N, Chaal BK, Sze SK, Preiser PR, Bozdech Z: **Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum***. *Mol Cell Proteomics* 2011, **10**(8):M110 006411.
56. Foth BJ, Zhang N, Mok S, Preiser PR, Bozdech Z: **Quantitative protein expression profiling reveals extensive post-transcriptional regulation and post-translational modifications in schizont-stage malaria parasites**. *Genome Biol* 2008, **9**(12):R177.
57. Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, Yan SF, Williamson KC, Holder AA, Carucci DJ *et al*: **Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle**. *Genome Res* 2004, **14**(11):2308-2318.
58. Vogel C, Marcotte EM: **Insights into the regulation of protein abundance from proteomic and transcriptomic analyses**. *Nat Rev Genet* 2012, **13**(4):227-232.
59. Hughes KR, Philip N, Starnes GL, Taylor S, Waters AP: **From cradle to grave: RNA biology in malaria parasites**. *Wiley Interdiscip Rev RNA* 2010, **1**(2):287-303.
60. Peccarelli M, Kebaara BW: **Regulation of natural mRNAs by the nonsense-mediated mRNA decay pathway**. *Eukaryot Cell* 2014, **13**(9):1126-1135.
61. Eshar S, Allemand E, Sebag A, Glaser F, Muchardt C, Mandel-Gutfreund Y, Karni R, Dzikowski R: **A novel *Plasmodium falciparum* SR protein is an alternative splicing factor required for the parasites' proliferation in human erythrocytes**. *Nucleic Acids Res* 2012, **40**(19):9903-9916.
62. Xue X, Zhang Q, Huang Y, Feng L, Pan W: **No miRNA were found in *Plasmodium* and the ones identified in erythrocytes could not be correlated with infection**. *Malar J* 2008, **7**:47.
63. Militello KT, Refour P, Comeaux CA, Duraisingh MT: **Antisense RNA and RNAi in protozoan parasites: working hard or hardly working?** *Mol Biochem Parasitol* 2008, **157**(2):117-126.
64. Zhang M, Mishra S, Sakthivel R, Rojas M, Ranjan R, Sullivan WJ, Jr., Fontoura BM, Menard R, Dever TE, Nussenzweig V: **PK4, a eukaryotic initiation factor 2alpha(eIF2alpha) kinase, is essential for the development of the erythrocytic cycle of *Plasmodium***. *Proc Natl Acad Sci U S A* 2012, **109**(10):3956-3961.
65. Merrick WC: **eIF4F: a retrospective**. *J Biol Chem* 2015, **290**(40):24091-24099.
66. Mitchell SF, Parker R: **Modifications on Translation Initiation**. *Cell* 2015, **163**(4):796-798.
67. Vembar SS, Droll D, Scherf A: **Translational regulation in blood stages of the malaria parasite *Plasmodium* spp.: systems-wide studies pave the way**. *Wiley Interdiscip Rev RNA* 2016, **7**(6):772-792.
68. Torres AG, Pineyro D, Filonava L, Stracker TH, Batlle E, Ribas de Pouplana L: **A-to-I editing on tRNAs: biochemical, biological and evolutionary implications**. *FEBS Lett* 2014, **588**(23):4279-4286.
69. Varani G, McClain WH: **The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems**. *EMBO Rep* 2000, **1**(1):18-23.
70. Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L: **A role for tRNA modifications in genome structure and codon usage**. *Cell* 2012, **149**(1):202-213.
71. Preiser P, Williamson DH, Wilson RJ: **tRNA genes transcribed from the plastid-like DNA of *Plasmodium falciparum***. *Nucleic Acids Res* 1995, **23**(21):4329-4336.

72. Chan S, Ch'ng JH, Wahlgren M, Thutkawkorapin J: **Frequent GU wobble pairings reduce translation efficiency in *Plasmodium falciparum*.** *Sci Rep* 2017, **7**(1):723.
73. Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM *et al*: **MODOMICS: a database of RNA modification pathways--2013 update.** *Nucleic Acids Res* 2013, **41**(Database issue):D262-267.
74. Davis DR: **Stabilization of RNA stacking by pseudouridine.** *Nucleic Acids Res* 1995, **23**(24):5020-5026.
75. Motorin Y, Helm M: **tRNA stabilization by modified nucleotides.** *Biochemistry* 2010, **49**(24):4934-4944.
76. Helm M, Motorin Y: **Detecting RNA modifications in the epitranscriptome: predict and validate.** *Nat Rev Genet* 2017, **18**(5):275-291.
77. Zhao BS, Roundtree IA, He C: **Post-transcriptional gene regulation by mRNA modifications.** *Nat Rev Mol Cell Biol* 2017, **18**(1):31-42.
78. Chan C, Pham P, Dedon PC, Begley TJ: **Lifestyle modifications: coordinating the tRNA epitranscriptome with codon bias to adapt translation during stress responses.** *Genome Biol* 2018, **19**(1):228.
79. Grozhik AV, Jaffrey SR: **Distinguishing RNA modifications from noise in epitranscriptome maps.** *Nat Chem Biol* 2018, **14**(3):215-225.
80. Linder B, Jaffrey SR: **Discovering and Mapping the Modified Nucleotides That Comprise the Epitranscriptome of mRNA.** *Cold Spring Harb Perspect Biol* 2019, **11**(6).
81. Agris PF: **Wobble position modified nucleosides evolved to select transfer RNA codon recognition: a modified-wobble hypothesis.** *Biochimie* 1991, **73**(11):1345-1349.
82. Limmer S, Reif B, Ott G, Arnold L, Sprinzl M: **NMR evidence for helix geometry modifications by a G-U wobble base pair in the acceptor arm of *E. coli* tRNA(Ala).** *FEBS Lett* 1996, **385**(1-2):15-20.
83. Bjork GR, Huang B, Persson OP, Bystrom AS: **A conserved modified wobble nucleoside (mcm5s2U) in lysyl-tRNA is required for viability in yeast.** *RNA* 2007, **13**(8):1245-1255.
84. Vendeix FA, Dziergowska A, Gustilo EM, Graham WD, Sproat B, Malkiewicz A, Agris PF: **Anticodon domain modifications contribute order to tRNA for ribosome-mediated codon binding.** *Biochemistry* 2008, **47**(23):6117-6129.
85. Songe-Moller L, van den Born E, Leihne V, Vagbo CB, Kristoffersen T, Krokan HE, Kirpekar F, Falnes PO, Klungland A: **Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding.** *Mol Cell Biol* 2010, **30**(7):1814-1827.
86. Chan CT, Pang YL, Deng W, Babu IR, Dyavaiah M, Begley TJ, Dedon PC: **Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins.** *Nat Commun* 2012, **3**:937.
87. Patil A, Chan CT, Dyavaiah M, Rooney JP, Dedon PC, Begley TJ: **Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications.** *RNA Biol* 2012, **9**(7):990-1001.
88. Vendeix FA, Murphy FVt, Cantara WA, Leszczynska G, Gustilo EM, Sproat B, Malkiewicz A, Agris PF: **Human tRNA(Lys3)(UUU) is pre-structured by natural modifications for cognate and wobble codon binding through keto-enol tautomerism.** *J Mol Biol* 2012, **416**(4):467-485.

89. Rezgui VA, Tyagi K, Ranjan N, Konevega AL, Mittelstaet J, Rodnina MV, Peter M, Pedrioli PG: **tRNA tKUUU, tQUUG, and tEUUC wobble position modifications fine-tune protein translation by promoting ribosome A-site binding.** *Proc Natl Acad Sci U S A* 2013, **110**(30):12289-12294.
90. Gu C, Begley TJ, Dedon PC: **tRNA modifications regulate translation during cellular stress.** *FEBS Lett* 2014, **588**(23):4287-4296.
91. Klassen R, Grunewald P, Thuring KL, Eichler C, Helm M, Schaffrath R: **Loss of anticodon wobble uridine modifications affects tRNA(Lys) function and protein levels in *Saccharomyces cerevisiae*.** *PLoS One* 2015, **10**(3):e0119261.
92. Nedialkova DD, Leidel SA: **Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity.** *Cell* 2015, **161**(7):1606-1618.
93. Chan CT, Deng W, Li F, DeMott MS, Babu IR, Begley TJ, Dedon PC: **Highly Predictive Reprogramming of tRNA Modifications Is Linked to Selective Expression of Codon-Biased Genes.** *Chem Res Toxicol* 2015, **28**(5):978-988.
94. Deng W, Babu IR, Su D, Yin S, Begley TJ, Dedon PC: **Trm9-Catalyzed tRNA Modifications Regulate Global Protein Expression by Codon-Biased Translation.** *PLoS Genet* 2015, **11**(12):e1005706.
95. Endres L, Dedon PC, Begley TJ: **Codon-biased translation can be regulated by wobble-base tRNA modification systems during cellular stress responses.** *RNA Biol* 2015, **12**(6):603-614.
96. Rozov A, Demeshkina N, Khusainov I, Westhof E, Yusupov M, Yusupova G: **Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code.** *Nat Commun* 2016, **7**:10457.
97. Chionh YH, McBee M, Babu IR, Hia F, Lin W, Zhao W, Cao J, Dziergowska A, Malkiewicz A, Begley TJ *et al*: **tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence.** *Nat Commun* 2016, **7**:13302.
98. Ranjan N, Rodnina MV: **tRNA wobble modifications and protein homeostasis.** *Translation (Austin)* 2016, **4**(1):e1143076.
99. Kumar RK, Davis DR: **Synthesis and studies on the effect of 2-thiouridine and 4-thiouridine on sugar conformation and RNA duplex stability.** *Nucleic Acids Res* 1997, **25**(6):1272-1280.
100. Yokoyama S, Watanabe T, Murao K, Ishikura H, Yamaizumi Z, Nishimura S, Miyazawa T: **Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon.** *Proc Natl Acad Sci U S A* 1985, **82**(15):4905-4909.
101. Leidel S, Pedrioli PG, Bucher T, Brost R, Costanzo M, Schmidt A, Aebersold R, Boone C, Hofmann K, Peter M: **Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA.** *Nature* 2009, **458**(7235):228-232.
102. Sample PJ, Koreny L, Paris Z, Gaston KW, Rubio MA, Fleming IM, Hinger S, Horakova E, Limbach PA, Lukes J *et al*: **A common tRNA modification at an unusual location: the discovery of wyosine biosynthesis in mitochondria.** *Nucleic Acids Res* 2015, **43**(8):4262-4273.
103. Gamper HB, Masuda I, Frenkel-Morgenstern M, Hou YM: **Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA.** *Nat Commun* 2015, **6**:7226.
104. Bjork GR, Wikstrom PM, Bystrom AS: **Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine.** *Science* 1989, **244**(4907):986-989.
105. Koh CS, Sarin LP: **Transfer RNA modification and infection - Implications for pathogenicity and host responses.** *Biochim Biophys Acta Gene Regul Mech* 2018, **1861**(4):419-432.

106. Koval'chuke OV, Potapov AP, El'skaya AV, Potapov VK, Krinetskaya NF, Dolinnaya NG, Shabarova ZA: **Interaction of ribo- and deoxyriboanalogs of yeast tRNA(Phe) anticodon arm with programmed small ribosomal subunits of Escherichia coli and rabbit liver.** *Nucleic Acids Res* 1991, **19**(15):4199-4201.
107. Dao V, Guenther R, Malkiewicz A, Nawrot B, Sochacka E, Kraszewski A, Jankowska J, Everett K, Agris PF: **Ribosome binding of DNA analogs of tRNA requires base modifications and supports the "extended anticodon".** *Proc Natl Acad Sci U S A* 1994, **91**(6):2125-2129.
108. Perry RP, Kelley DE, LaTorre J: **Synthesis and turnover of nuclear and cytoplasmic polyadenylic acid in mouse L cells.** *J Mol Biol* 1974, **82**(3):315-331.
109. Adams JM, Cory S: **Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA.** *Nature* 1975, **255**(5503):28-33.
110. Desrosiers R, Friderici K, Rottman F: **Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells.** *Proc Natl Acad Sci U S A* 1974, **71**(10):3971-3975.
111. Frye M, Harada BT, Behm M, He C: **RNA modifications modulate gene expression during development.** *Science* 2018, **361**(6409):1346-1349.
112. Davis FF, Allen FW: **Ribonucleic acids from yeast which contain a fifth nucleotide.** *J Biol Chem* 1957, **227**(2):907-915.
113. Fernandez IS, Ng CL, Kelley AC, Wu G, Yu YT, Ramakrishnan V: **Unusual base pairing during the decoding of a stop codon by the ribosome.** *Nature* 2013, **500**(7460):107-110.
114. Svidritskiy E, Madireddy R, Korostelev AA: **Structural Basis for Translation Termination on a Pseudouridylated Stop Codon.** *J Mol Biol* 2016, **428**(10 Pt B):2228-2236.
115. Kariko K, Muramatsu H, Welsh FA, Ludwig J, Kato H, Akira S, Weissman D: **Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability.** *Mol Ther* 2008, **16**(11):1833-1840.
116. Andries O, Mc Cafferty S, De Smedt SC, Weiss R, Sanders NN, Kitada T: **N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice.** *J Control Release* 2015, **217**:337-344.
117. Thess A, Grund S, Mui BL, Hope MJ, Baumhof P, Fotin-Mleczek M, Schlake T: **Sequence-engineered mRNA Without Chemical Nucleoside Modifications Enables an Effective Protein Therapy in Large Animals.** *Mol Ther* 2015, **23**(9):1456-1464.
118. Smith JD, Dunn DB: **An additional sugar component of ribonucleic acids.** *Biochim Biophys Acta* 1959, **31**(2):573-575.
119. Sendinc E, Valle-Garcia D, Dhall A, Chen H, Henriques T, Navarrete-Perea J, Sheng W, Gygi SP, Adelman K, Shi Y: **PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression.** *Mol Cell* 2019.
120. Picard-Jean F, Brand C, Tremblay-Letourneau M, Allaire A, Beaudoin MC, Boudreault S, Duval C, Rainville-Sirois J, Robert F, Pelletier J *et al*: **2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO.** *PLoS One* 2018, **13**(3):e0193804.
121. Choi J, Grosely R, Prabhakar A, Lapointe CP, Wang J, Puglisi JD: **How Messenger RNA and Nascent Chain Sequences Regulate Translation Elongation.** *Annu Rev Biochem* 2018, **87**:421-449.

122. Pandolfini L, Barbieri I, Bannister AJ, Hendrick A, Andrews B, Webster N, Murat P, Mach P, Brandi R, Robson SC *et al*: **METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation**. *Mol Cell* 2019, **74**(6):1278-1290 e1279.
123. Zhang LS, Liu C, Ma H, Dai Q, Sun HL, Luo G, Zhang Z, Zhang L, Hu L, Dong X *et al*: **Transcriptome-wide Mapping of Internal N(7)-Methylguanosine Methylome in Mammalian mRNA**. *Mol Cell* 2019, **74**(6):1304-1316 e1308.
124. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WC *et al*: **The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA**. *Nature* 2016, **530**(7591):441-446.
125. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C: **Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome**. *Nat Chem Biol* 2016, **12**(5):311-316.
126. Schwartz S: **m(1)A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map**. *RNA* 2018, **24**(11):1427-1436.
127. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W *et al*: **5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader**. *Cell Res* 2017, **27**(5):606-625.
128. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T: **Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA**. *Nucleic Acids Res* 2012, **40**(11):5023-5033.
129. Xu L, Liu X, Sheng N, Oo KS, Liang J, Chionh YH, Xu J, Ye F, Gao YG, Dedon PC *et al*: **Three distinct 3-methylcytidine (m(3)C) methyltransferases modify tRNA and mRNA in mice and humans**. *J Biol Chem* 2017, **292**(35):14695-14703.
130. Amort T, Rieder D, Wille A, Khokhlova-Cubberley D, Riml C, Trixl L, Jia XY, Micura R, Lusser A: **Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain**. *Genome Biol* 2017, **18**(1):1.
131. David R, Burgess A, Parker B, Li J, Pulsford K, Sibbritt T, Preiss T, Searle IR: **Transcriptome-Wide Mapping of RNA 5-Methylcytosine in Arabidopsis mRNAs and Noncoding RNAs**. *Plant Cell* 2017, **29**(3):445-460.
132. Huber SM, van Delft P, Mendil L, Bachman M, Smollett K, Werner F, Miska EA, Balasubramanian S: **Formation and abundance of 5-hydroxymethylcytosine in RNA**. *ChemBiochem* 2015, **16**(5):752-755.
133. Chan CT, Dyavaiah M, DeMott MS, Taghizadeh K, Dedon PC, Begley TJ: **A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress**. *PLoS Genet* 2010, **6**(12):e1001247.
134. Jaroensuk J, Atichartpongkul S, Chionh YH, Wong YH, Liew CW, McBee ME, Thongdee N, Prestwich EG, DeMott MS, Mongkolsuk S *et al*: **Methylation at position 32 of tRNA catalyzed by TrmJ alters oxidative stress response in Pseudomonas aeruginosa**. *Nucleic Acids Res* 2016, **44**(22):10834-10848.
135. Pang YL, Abo R, Levine SS, Dedon PC: **Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number**. *Nucleic Acids Res* 2014, **42**(22):e170.
136. Gu C, Ramos J, Begley U, Dedon PC, Fu D, Begley TJ: **Phosphorylation of human TRM9L integrates multiple stress-signaling pathways for tumor growth suppression**. *Sci Adv* 2018, **4**(7):eaas9184.
137. Kumar P, Kuscü C, Dutta A: **Biogenesis and Function of Transfer RNA-Related Fragments (tRFs)**. *Trends Biochem Sci* 2016, **41**(8):679-689.

138. Fricker R, Brogli R, Luidalepp H, Wyss L, Fasnacht M, Joss O, Zywicki M, Helm M, Schneider A, Cristodero M *et al*: **A tRNA half modulates translation as stress response in *Trypanosoma brucei*.** *Nat Commun* 2019, **10**(1):118.
139. Zhou J, Wan J, Gao X, Zhang X, Jaffrey SR, Qian SB: **Dynamic m(6)A mRNA methylation directs translational control of heat shock response.** *Nature* 2015, **526**(7574):591-594.
140. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR: **Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons.** *Cell* 2012, **149**(7):1635-1646.
141. Karijolich J, Yi C, Yu YT: **Transcriptome-wide dynamics of RNA pseudouridylation.** *Nat Rev Mol Cell Biol* 2015, **16**(10):581-585.
142. Trager W, Jensen JB: **Human malaria parasites in continuous culture.** *Science* 1976, **193**(4254):673-675.
143. Cai WM, Chionh YH, Hia F, Gu C, Kellner S, McBee ME, Ng CS, Pang YL, Prestwich EG, Lim KS *et al*: **A Platform for Discovery and Quantification of Modified Ribonucleosides in RNA: Application to Stress-Induced Reprogramming of tRNA Modifications.** *Methods Enzymol* 2015, **560**:29-71.
144. Siwo GH, Smith RS, Tan A, Button-Simons KA, Checkley LA, Ferdig MT: **An integrative analysis of small molecule transcriptional responses in the human malaria parasite *Plasmodium falciparum*.** *BMC Genomics* 2015, **16**:1030.
145. Natalang O, Bischoff E, Deplaine G, Proux C, Dillies MA, Sismeiro O, Guigon G, Bonnefoy S, Patarapotikul J, Mercereau-Puijalon O *et al*: **Dynamic RNA profiling in *Plasmodium falciparum* synchronized blood stages exposed to lethal doses of artesunate.** *BMC Genomics* 2008, **9**:388.
146. Hiley SL, Jackman J, Babak T, Trochesset M, Morris QD, Phizicky E, Hughes TR: **Detection and discovery of RNA modifications using microarrays.** *Nucleic Acids Res* 2005, **33**(1):e2.
147. Weaver RF, Weissmann C: **Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15 S beta-globin mRNA precursor and mature 10 s beta-globin mRNA have identical map coordinates.** *Nucleic Acids Res* 1979, **7**(5):1175-1193.
148. Grosjean H, Keith G, Droogmans L: **Detection and quantification of modified nucleotides in RNA using thin-layer chromatography.** *Methods Mol Biol* 2004, **265**:357-391.
149. Donis-Keller H, Maxam AM, Gilbert W: **Mapping adenines, guanines, and pyrimidines in RNA.** *Nucleic Acids Res* 1977, **4**(8):2527-2538.
150. McCloskey JA, Whitehill AB, Rozenski J, Qiu F, Crain PF: **New techniques for the rapid characterization of oligonucleotides by mass spectrometry.** *Nucleosides Nucleotides* 1999, **18**(6-7):1549-1553.
151. Babatunde KA, Mbagwu S, Hernandez-Castaneda MA, Adapa SR, Walch M, Filgueira L, Falquet L, Jiang RHY, Ghiran I, Mantel PY: **Malaria infected red blood cells release small regulatory RNAs through extracellular vesicles.** *Sci Rep* 2018, **8**(1):884.
152. Suzuki T, Suzuki T: **Chaplet column chromatography: isolation of a large set of individual RNAs in a single step.** *Methods Enzymol* 2007, **425**:231-239.
153. Uchida T, Arima T, Egami F: **Specificity of RNase U2.** *J Biochem* 1970, **67**(1):91-102.
154. Irie S, Ito T, Ueda T, Egami F: **pH dependence of the RNase T1 action on nucleoside 2',- 3'-cyclic phosphates.** *J Biochem* 1970, **68**(2):163-170.

155. Huber CG, Oefner PJ, Bonn GK: **High-performance liquid chromatographic separation of detritylated oligonucleotides on highly cross-linked poly-(styrene-divinylbenzene) particles.** *J Chromatogr* 1992, **599**(1-2):113-118.
156. Annesley TM: **Ion suppression in mass spectrometry.** *Clin Chem* 2003, **49**(7):1041-1044.
157. McLuckey SA, Van Berkel GJ, Glish GL: **Tandem mass spectrometry of small, multiply charged oligonucleotides.** *J Am Soc Mass Spectrom* 1992, **3**(1):60-70.
158. Pan T: **Modifications and functional genomics of human transfer RNA.** *Cell Res* 2018, **28**(4):395-404.
159. Mohler K, Mann R, Ibba M: **Isoacceptor specific characterization of tRNA aminoacylation and misacylation in vivo.** *Methods* 2017, **113**:127-131.
160. Rozenski J, McCloskey JA: **SOS: a simple interactive program for ab initio oligonucleotide sequencing by mass spectrometry.** *J Am Soc Mass Spectrom* 2002, **13**(3):200-203.
161. Nyakas A, Blum LC, Stucki SR, Reymond JL, Schurch S: **OMA and OPA--software-supported mass spectra analysis of native and modified nucleic acids.** *J Am Soc Mass Spectrom* 2013, **24**(2):249-256.
162. Sample PJ, Gaston KW, Alfonzo JD, Limbach PA: **RoboOligo: software for mass spectrometry data to support manual and de novo sequencing of post-transcriptionally modified ribonucleic acids.** *Nucleic Acids Res* 2015, **43**(10):e64.
163. Lobue PA, Yu N, Jora M, Abernathy S, Limbach PA: **Improved application of RNAmodMapper - An RNA modification mapping software tool - For analysis of liquid chromatography tandem mass spectrometry (LC-MS/MS) data.** *Methods* 2019, **156**:128-138.
164. Kudla G, Murray AW, Tollervey D, Plotkin JB: **Coding-sequence determinants of gene expression in Escherichia coli.** *Science* 2009, **324**(5924):255-258.
165. Esberg A, Huang B, Johansson MJ, Bystrom AS: **Elevated levels of two tRNA species bypass the requirement for elongator complex in transcription and exocytosis.** *Mol Cell* 2006, **24**(1):139-148.
166. Thiaville PC, Legendre R, Rojas-Benitez D, Baudin-Baillieu A, Hatin I, Chalancon G, Glavic A, Namy O, de Crecy-Lagard V: **Global translational impacts of the loss of the tRNA modification t(6)A in yeast.** *Microb Cell* 2016, **3**(1):29-45.
167. Baum J, Papenfuss AT, Mair GR, Janse CJ, Vlachou D, Waters AP, Cowman AF, Crabb BS, de Koning-Ward TF: **Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites.** *Nucleic Acids Res* 2009, **37**(11):3788-3798.
168. Shen Y, Yu X, Zhu L, Li T, Yan Z, Guo J: **Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases.** *J Mol Med (Berl)* 2018, **96**(11):1167-1176.
169. Wang Z, Wei C, Hao X, Deng W, Zhang L, Wang Z, Wang H: **Genome-wide identification and characterization of transfer RNA-derived small RNAs in Plasmodium falciparum.** *Parasit Vectors* 2019, **12**(1):36.
170. LaMonte G, Philip N, Reardon J, Lacsina JR, Majoros W, Chapman L, Thornburg CD, Telen MJ, Ohler U, Nicchitta CV *et al*: **Translocation of sickle cell erythrocyte microRNAs into Plasmodium falciparum inhibits parasite translation and contributes to malaria resistance.** *Cell Host Microbe* 2012, **12**(2):187-199.
171. Bour T, Mahmoudi N, Kapps D, Thiberge S, Bargieri D, Menard R, Frugier M: **Apicomplexa-specific tRip facilitates import of exogenous tRNAs into malaria parasites.** *Proc Natl Acad Sci U S A* 2016, **113**(17):4717-4722.

172. Sinha I, Woodrow CJ: **Forces acting on codon bias in malaria parasites.** *Sci Rep* 2018, **8**(1):15984.
173. Kalhor HR, Clarke S: **Novel methyltransferase for modified uridine residues at the wobble position of tRNA.** *Mol Cell Biol* 2003, **23**(24):9283-9292.
174. Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, Ribas de Pouplana L, Preiser PR, Bozdech Z, Cortes A: **Transcriptional variation in the malaria parasite *Plasmodium falciparum*.** *Genome Res* 2012, **22**(5):925-938.
175. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, Cohen AM, Janssen-Megens EM, van de Vegte-Bolmer M, Selzer RR, Iniguez AL *et al*: **Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*.** *Proc Natl Acad Sci U S A* 2009, **106**(24):9655-9660.
176. Modrzynska K, Pfander C, Chappell L, Yu L, Suarez C, Dundas K, Gomes AR, Goulding D, Rayner JC, Choudhary J *et al*: **A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle.** *Cell Host Microbe* 2017, **21**(1):11-22.
177. Painter HJ, Chung NC, Sebastian A, Albert I, Storey JD, Llinas M: **Genome-wide real-time in vivo transcriptional dynamics during *Plasmodium falciparum* blood-stage development.** *Nat Commun* 2018, **9**(1):2656.
178. Grozhik AV, Linder B, Olarerin-George AO, Jaffrey SR: **Mapping m(6)A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP).** *Methods Mol Biol* 2017, **1562**:55-78.
179. Meyer KD, Jaffrey SR: **Rethinking m(6)A Readers, Writers, and Erasers.** *Annu Rev Cell Dev Biol* 2017, **33**:319-342.
180. Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR: **Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome.** *Nat Methods* 2015, **12**(8):767-772.
181. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M *et al*: **Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq.** *Nature* 2012, **485**(7397):201-206.
182. Hawley BR, Jaffrey SR: **Transcriptome-Wide Mapping of m(6) A and m(6) Am at Single-Nucleotide Resolution Using miCLIP.** *Curr Protoc Mol Biol* 2019, **126**(1):e88.
183. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, Pestova TV, Qian SB, Jaffrey SR: **5' UTR m(6)A Promotes Cap-Independent Translation.** *Cell* 2015, **163**(4):999-1010.
184. Kan L, Grozhik AV, Vedanayagam J, Patil DP, Pang N, Lim KS, Huang YC, Joseph B, Lin CJ, Despic V *et al*: **The m(6)A pathway facilitates sex determination in *Drosophila*.** *Nat Commun* 2017, **8**:15737.
185. Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, Chou T, Chow A, Saletore Y, MacKay M *et al*: **The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells.** *Nat Med* 2017, **23**(11):1369-1376.
186. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T: **N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions.** *Nature* 2015, **518**(7540):560-564.
187. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur JJ, Chen Q *et al*: **Reversible methylation of m(6)Am in the 5' cap controls mRNA stability.** *Nature* 2017, **541**(7637):371-375.

188. Wang C, Zhu Y, Bao H, Jiang Y, Xu C, Wu J, Shi Y: **A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1.** *Nucleic Acids Res* 2016, **44**(2):969-982.
189. Sledz P, Jinek M: **Structural insights into the molecular mechanism of the m(6)A writer complex.** *Elife* 2016, **5**.
190. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, Adhikari S, Shi Y, Lv Y, Chen YS *et al*: **Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase.** *Cell Res* 2014, **24**(2):177-189.
191. Patil DP, Pickering BF, Jaffrey SR: **Reading m(6)A in the Transcriptome: m(6)A-Binding Proteins.** *Trends Cell Biol* 2018, **28**(2):113-127.
192. Schwartz S, Agarwala SD, Mumbach MR, Jovanovic M, Mertins P, Shishkin A, Tabach Y, Mikkelsen TS, Satija R, Ruvkun G *et al*: **High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis.** *Cell* 2013, **155**(6):1409-1421.
193. Lence T, Akhtar J, Bayer M, Schmid K, Spindler L, Ho CH, Kreim N, Andrade-Navarro MA, Poeck B, Helm M *et al*: **m(6)A modulates neuronal functions and sex determination in Drosophila.** *Nature* 2016, **540**(7632):242-247.
194. Giessing AM, Kirpekar F: **Mass spectrometry in the biology of RNA and its modifications.** *J Proteomics* 2012, **75**(12):3434-3449.
195. Morlan JD, Qu K, Sinicropi DV: **Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue.** *PLoS One* 2012, **7**(8):e42882.
196. Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC: **Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA.** *BMC Genomics* 2015, **16**:454.
197. Shepherd J, Ibba M: **Bacterial transfer RNAs.** *FEMS Microbiol Rev* 2015, **39**(3):280-300.
198. Elsworth B, Sanders PR, Nebl T, Batinovic S, Kalanon M, Nie CQ, Charnaud SC, Bullen HE, de Koning Ward TF, Tilley L *et al*: **Proteomic analysis reveals novel proteins associated with the Plasmodium protein exporter PTEX and a loss of complex stability upon truncation of the core PTEX component, PTEX150.** *Cell Microbiol* 2016, **18**(11):1551-1569.
199. Chook YM, Blobel G: **Karyopherins and nuclear import.** *Curr Opin Struct Biol* 2001, **11**(6):703-715.
200. Wang Z, Tang K, Zhang D, Wan Y, Wen Y, Lu Q, Wang L: **High-throughput m6A-seq reveals RNA m6A methylation patterns in the chloroplast and mitochondria transcriptomes of Arabidopsis thaliana.** *PLoS One* 2017, **12**(11):e0185612.
201. Antanaviciute A, Baquero-Perez B, Watson CM, Harrison SM, Lascelles C, Crinnion L, Markham AF, Bonthron DT, Whitehouse A, Carr IM: **m6aViewer: software for the detection, analysis, and visualization of N(6)-methyladenosine peaks from m(6)A-seq/ME-RIP sequencing data.** *RNA* 2017, **23**(10):1493-1501.
202. Molinie B, Giallourakis CC: **Genome-Wide Location Analyses of N6-Methyladenosine Modifications (m(6)A-Seq).** *Methods Mol Biol* 2017, **1562**:45-53.
203. Shock JL, Fischer KF, DeRisi JL: **Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle.** *Genome Biol* 2007, **8**(7):R134.
204. Caro F, Ahyong V, Betegon M, DeRisi JL: **Genome-wide regulatory dynamics of translation in the Plasmodium falciparum asexual blood stages.** *Elife* 2014, **3**.

205. Xiao B, Yin S, Hu Y, Sun M, Wei J, Huang Z, Wen Y, Dai X, Chen H, Mu J *et al*: **Epigenetic editing by CRISPR/dCas9 in Plasmodium falciparum**. *Proc Natl Acad Sci U S A* 2019, **116**(1):255-260.
206. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al*: **CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes**. *Cell* 2013, **154**(2):442-451.
207. Wei CM, Gershowitz A, Moss B: **Methylated nucleotides block 5' terminus of HeLa cell messenger RNA**. *Cell* 1975, **4**(4):379-386.
208. Wei CM, Moss B: **Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA**. *Proc Natl Acad Sci U S A* 1975, **72**(1):318-322.
209. Perry RP, Kelley DE, Friderici K, Rottman F: **The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus**. *Cell* 1975, **4**(4):387-394.
210. Motorin Y, Helm M: **Methods for RNA Modification Mapping Using Deep Sequencing: Established and New Emerging Technologies**. *Genes (Basel)* 2019, **10**(1).
211. Wu G, Huang C, Yu YT: **Pseudouridine in mRNA: Incorporation, Detection, and Recoding**. *Methods Enzymol* 2015, **560**:187-217.
212. Rice JM, Dudek G: **Mass spectra of uridine and pseudouridine: fragmentation patterns characteristic of a carbon-carbon nucleosidic bond**. *Biochem Biophys Res Commun* 1969, **35**(3):383-388.
213. Thiaville JJ, Kellner SM, Yuan Y, Hutinet G, Thiaville PC, Jumpathong W, Mohapatra S, Brochier-Armanet C, Letarov AV, Hillebrand R *et al*: **Novel genomic island modifies DNA with 7-deazaguanine derivatives**. *Proc Natl Acad Sci U S A* 2016, **113**(11):E1452-1459.
214. Cao H, Wang Y: **Collisionally activated dissociation of protonated 2'-deoxycytidine, 2'-deoxyuridine, and their oxidatively damaged derivatives**. *J Am Soc Mass Spectrom* 2006, **17**(10):1335-1341.
215. Luo GZ, MacQueen A, Zheng G, Duan H, Dore LC, Lu Z, Liu J, Chen K, Jia G, Bergelson J *et al*: **Unique features of the m6A methylome in Arabidopsis thaliana**. *Nat Commun* 2014, **5**:5630.
216. Macon JB, Wolfenden R: **1-Methyladenosine. Dimroth rearrangement and reversible reduction**. *Biochemistry* 1968, **7**(10):3453-3458.
217. Lasonder E, Rijpma SR, van Schaijk BC, Hoeijmakers WA, Kensche PR, Gresnigt MS, Italiaander A, Vos MW, Woestenenk R, Bousema T *et al*: **Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression**. *Nucleic Acids Res* 2016, **44**(13):6087-6101.
218. Ries RJ, Zaccara S, Klein P, Olarerin-George A, Namkoong S, Pickering BF, Patil DP, Kwak H, Lee JH, Jaffrey SR: **m(6)A enhances the phase separation potential of mRNA**. *Nature* 2019, **571**(7765):424-428.
219. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, Cheng T, Gao M, Shu X, Ma H *et al*: **VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation**. *Cell Discov* 2018, **4**:10.
220. Mauer J, Sindelar M, Despic V, Guez T, Hawley BR, Vasseur JJ, Rentmeister A, Gross SS, Pellizzoni L, Debart F *et al*: **FTO controls reversible m(6)Am RNA methylation during snRNA biogenesis**. *Nat Chem Biol* 2019, **15**(4):340-347.
221. Mauer J, Jaffrey SR: **FTO, m(6) Am , and the hypothesis of reversible epitranscriptomic mRNA modifications**. *FEBS Lett* 2018, **592**(12):2012-2022.

222. Sharp PM, Matassi G: **Codon usage and genome evolution.** *Curr Opin Genet Dev* 1994, **4**(6):851-860.
223. Plotkin JB, Kudla G: **Synonymous but not the same: the causes and consequences of codon bias.** *Nat Rev Genet* 2011, **12**(1):32-42.
224. Tenaillon O, Denamur E, Matic I: **Evolutionary significance of stress-induced mutagenesis in bacteria.** *Trends Microbiol* 2004, **12**(6):264-270.
225. Lloyd SB, Kent SJ, Winnall WR: **The high cost of fidelity.** *AIDS Res Hum Retroviruses* 2014, **30**(1):8-16.
226. Ribas de Pouplana L, Santos MA, Zhu JH, Farabaugh PJ, Javid B: **Protein mistranslation: friend or foe?** *Trends Biochem Sci* 2014, **39**(8):355-362.
227. Javid B, Sorrentino F, Toosky M, Zheng W, Pinkham JT, Jain N, Pan M, Deighan P, Rubin EJ: **Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance.** *Proc Natl Acad Sci U S A* 2014, **111**(3):1132-1137.
228. Lee JY, Kim DG, Kim BG, Yang WS, Hong J, Kang T, Oh YS, Kim KR, Han BW, Hwang BJ *et al*: **Promiscuous methionyl-tRNA synthetase mediates adaptive mistranslation to protect cells against oxidative stress.** *J Cell Sci* 2014, **127**(Pt 19):4234-4245.
229. Netzer N, Goodenbour JM, David A, Dittmar KA, Jones RB, Schneider JR, Boone D, Eves EM, Rosner MR, Gibbs JS *et al*: **Innate immune and chemically triggered oxidative stress modifies translational fidelity.** *Nature* 2009, **462**(7272):522-526.
230. Kupsco A, Schlenk D: **Oxidative stress, unfolded protein response, and apoptosis in developmental toxicity.** *Int Rev Cell Mol Biol* 2015, **317**:1-66.
231. Becker K, Tilley L, Vennerstrom JL, Roberts D, Rogerson S, Ginsburg H: **Oxidative stress in malaria parasite-infected erythrocytes: host-parasite interactions.** *Int J Parasitol* 2004, **34**(2):163-189.
232. Fu Y, Tilley L, Kenny S, Klonis N: **Dual labeling with a far red probe permits analysis of growth and oxidative stress in *P. falciparum*-infected erythrocytes.** *Cytometry A* 2010, **77**(3):253-263.
233. Rocamora F, Zhu L, Liong KY, Dondorp A, Miotto O, Mok S, Bozdech Z: **Oxidative stress and protein damage responses mediate artemisinin resistance in malaria parasites.** *PLoS Pathog* 2018, **14**(3):e1006930.
234. Loor G, Kondapalli J, Schriewer JM, Chandel NS, Vanden Hoek TL, Schumacker PT: **Menadione triggers cell death through ROS-dependent mechanisms involving PARP activation without requiring apoptosis.** *Free Radic Biol Med* 2010, **49**(12):1925-1936.
235. Rauniyar N, Yates JR, 3rd: **Isobaric labeling-based relative quantification in shotgun proteomics.** *J Proteome Res* 2014, **13**(12):5293-5309.
236. Brosch M, Yu L, Hubbard T, Choudhary J: **Accurate and sensitive peptide identification with Mascot Percolator.** *J Proteome Res* 2009, **8**(6):3176-3181.
237. Gunasekera AM, Patankar S, Schug J, Eisen G, Wirth DF: **Drug-induced alterations in gene expression of the asexual blood forms of *Plasmodium falciparum*.** *Mol Microbiol* 2003, **50**(4):1229-1239.
238. Witola WH, Ben Mamoun C: **Choline induces transcriptional repression and proteasomal degradation of the malarial phosphoethanolamine methyltransferase.** *Eukaryot Cell* 2007, **6**(9):1618-1624.
239. Tewari SG, Rajaram K, Schyman P, Swift R, Reifman J, Prigge ST, Wallqvist A: **Short-term metabolic adjustments in *Plasmodium falciparum* counter hypoxanthine deprivation at the expense of long-term viability.** *Malar J* 2019, **18**(1):86.

240. Huber SM, Leonardi A, Dedon PC, Begley TJ: **The Versatile Roles of the tRNA Epitranscriptome during Cellular Responses to Toxic Exposures and Environmental Stress.** *Toxics* 2019, **7**(1).
241. Gkatza NA, Castro C, Harvey RF, Heiss M, Popis MC, Blanco S, Bornelov S, Sajini AA, Gleeson JG, Griffin JL *et al*: **Cytosine-5 RNA methylation links protein synthesis to cell metabolism.** *PLoS Biol* 2019, **17**(6):e3000297.
242. Hori H: **Regulatory Factors for tRNA Modifications in Extreme- Thermophilic Bacterium *Thermus thermophilus*.** *Front Genet* 2019, **10**:204.
243. Muller M, Hartmann M, Schuster I, Bender S, Thuring KL, Helm M, Katze JR, Nellen W, Lyko F, Ehrenhofer-Murray AE: **Dynamic modulation of Dnmt2-dependent tRNA methylation by the micronutrient queuine.** *Nucleic Acids Res* 2015, **43**(22):10952-10962.
244. Benitez-Paez A, Villarroja M, Armengod ME: **The *Escherichia coli* RlmN methyltransferase is a dual-specificity enzyme that modifies both rRNA and tRNA and controls translational accuracy.** *RNA* 2012, **18**(10):1783-1795.
245. Schwach F, Bushell E, Gomes AR, Anar B, Girling G, Herd C, Rayner JC, Billker O: **PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites.** *Nucleic Acids Res* 2015, **43**(Database issue):D1176-1182.
246. Birnbaum J, Flemming S, Reichard N, Soares AB, Mesen-Ramirez P, Jonscher E, Bergmann B, Spielmann T: **A genetic system to study *Plasmodium falciparum* protein function.** *Nat Methods* 2017, **14**(4):450-456.
247. Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC: **Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*.** *Nat Methods* 2014, **11**(9):915-918.
248. Lee MC, Fidock DA: **CRISPR-mediated genome editing of *Plasmodium falciparum* malaria parasites.** *Genome Med* 2014, **6**(8):63.
249. Ah Yong V, Sheridan CM, Leon KE, Witchley JN, Diep J, DeRisi JL: **Identification of *Plasmodium falciparum* specific translation inhibitors from the MMV Malaria Box using a high throughput in vitro translation screen.** *Malar J* 2016, **15**:173.
250. Glasser AL, el Adlouni C, Keith G, Sochacka E, Malkiewicz A, Santos M, Tuite MF, Desgres J: **Presence and coding properties of 2'-O-methyl-5-carbamoylmethyluridine (ncm5Um) in the wobble position of the anticodon of tRNA(Leu) (U*AA) from brewer's yeast.** *FEBS Lett* 1992, **314**(3):381-385.

Appendix

7. Appendix

7.1. Appendix I

Modified ribonucleosides identified by LC-MS/MS analysis of *P. falciparum* tRNA hydrolysates. Identification footnotes: **A**, MS² fragments match reported data (Chan *et al.* 2010); **B**, validated by MS² fragmentation of m⁵Cm standard eluting at 7 min, so annotated as m⁴Cm; **C**, validated by MS² fragmentation of m⁷G and m²G standards eluting at 5 and 16.8 min, respectively, so annotated as m¹G; **D**, validated by MS² fragmentation of m⁵U and Um standards; **E**, MS² fragments match reported data (Chan *et al.* 2010); **F**: Validated by MS² and MS³ fragmentations of t⁶A and ms²i⁶A standards.

Modified ribonucleosides	Short name	Retention time (min)	Theoretical [M+H] ⁺ (m/z)	Observed molecular ion (m/z)	Major fragment ion (m/z)	Neutral loss (amu)	Quantifier SRM transition (precursor → product)	Identification
Dihydrouridine	D	1.5	247.093	247.090 ± 0.005	115.048 ± 0.002	132.041	247.1 → 115.1	Standard
Pseudouridine	Ψ	1.6	245.078	245.077 ± 0.003	209.052 ± 0.004	36.034	245.1 → 209.1; 245.1 → 125.1	Standard
3-Methylcytidine	m ³ C	2.5	258.109	258.106 ± 0.004	126.064 ± 0.002	132.042	258.1 → 126.1	Standard
5-Carbamoylmethyluridine	ncm ⁵ U	2.6	302.099	302.093 ± 0.007	170.051 ± 0.002	132.042	302.1 → 170.0	A
5-Methylcytidine	m ⁵ C	3.1	258.109	258.108 ± 0.004	126.065 ± 0.002	132.042	258.1 → 126.1	Standard
1-Methyladenosine	m ¹ A	3.2	282.120	282.117 ± 0.007	150.075 ± 0.003	132.042	282.1 → 150.1	Standard
7-Methylguanosine	m ⁷ G	5.2	298.115	298.113 ± 0.005	166.071 ± 0.002	132.042	298.1 → 166.1	Standard
2'-O-Methylcytidine	Cm	5.3	258.109	258.106 ± 0.004	112.050 ± 0.002	146.056	258.1 → 112.1	Standard
Inosine	I	6.6	269.089	269.088 ± 0.007	137.047 ± 0.002	132.042	268.1 → 137.0	Standard
5-Methyluridine	m ⁵ U	6.7	259.093	259.088 ± 0.005	127.048 ± 0.002	132.042	259.1 → 127.0	Standard
2-Thiouridine	s ² U	7.4	261.055	261.051 ± 0.006	129.009 ± 0.008	132.042	261.0 → 129.0	Standard
2'-O-Methyluridine	Um	9.6	259.093	259.092 ± 0.004	113.036 ± 0.002	146.056	259.1 → 113.1	Standard
3-Methyluridine	m ³ U	10.1	259.093	259.089 ± 0.011	127.045 ± 0.005	132.044	259.1 → 127.0	Standard
5,2'-O-Dimethylcytidine	m ⁴ Cm	11.1	272.125	272.122 ± 0.009	126.065 ± 0.002	146.057	272.1 → 126.1	B
2'-O-Methyladenosine	Am	12.3	282.120	282.114 ± 0.003	136.059 ± 0.003	146.055	282.1 → 136.1	Standard
1-Methylguanosine	m ¹ G	14.5	298.115	298.111 ± 0.006	166.070 ± 0.002	132.041	298.1 → 166.1	C
5-Methoxycarbonylmethyluridine	mcm ⁵ U	14.7	317.099	317.098 ± 0.010	185.054 ± 0.003	132.044	317.1 → 185.1	Standard
2'-O-Methylguanosine	Gm	15.2	298.115	298.110 ± 0.007	152.056 ± 0.003	146.054	298.1 → 152.1	Standard
N ⁶ -Methyladenosine	m ⁶ A	16.5	282.120	282.115 ± 0.004	150.075 ± 0.002	132.04	282.1 → 150.1	Standard
N ⁴ -Acetylcytidine	ac ⁴ C	17.1	286.104	286.101 ± 0.005	154.056 ± 0.005	132.044	286.1 → 154.1	Standard
N ² -Methylguanosine	m ² G	17.7	298.115	298.112 ± 0.004	166.071 ± 0.005	132.041	298.1 → 166.1	Standard
5,2'-O-Dimethyluridine	m ⁵ Um	18.8	273.109	273.108 ± 0.007	127.047 ± 0.005	146.061	273.1 → 127.1	D
N2,N2-Dimethylguanosine	m ^{2,2} G	19.7	312.131	312.127 ± 0.005	180.085 ± 0.003	132.041	312.1 → 180.1	Standard

5-Methoxycarbonylmethyl-2-thiouridine	mcm ⁵ S ² U	20.3	333.076	333.075 ± 0.008	201.031 ± 0.005	132.044	333.1 → 201.0	<i>E</i>
N⁶,N⁶-Dimethyladenosine	m ^{6,6} A	20.5	296.136	296.131 ± 0.007	164.088 ± 0.008	132.043	296.1 → 164.1	Standard
N⁶-Threonyl-carbamoyl-adenosine	t ⁶ A	21.9	413.142	413.136 ± 0.005	136.062 ± 0.004	132.043	413.1 → 136.1; 413.1 → 281.1	Standard
2-Methylthio-N⁶-threonyl-carbamoyl-adenosine	ms ² t ⁶ A	22.2	459.120	459.127 ± 0.020	327.084 ± 0.010	132.043	459.1 → 327.1	<i>F</i>
N⁶-Isopentenyladenosine	i ⁶ A	22.5	336.167	336.159 ± 0.005	204.119 ± 0.003	132.04	336.2 → 204.1	Standard

7.2.Appendix II

List of nuclear encoded tRNA gene products.

^a. Initially annotated as Asparagine the entry was manually curated to Aspartic Acid upon examining the sequences

^b Initially annotated as Glutamic acid, the entry was manually curated to Glutamine upon examining the sequences

Gene ID	Product	Predicted RNA Sequence (introns spliced out)	Cog
	Descrip		nate
	tion		codo
			n
PF3D7_0411500	tRNA Alanine	GGGCAGGTGGTGTAGTGGTATCACGCTTGATTGCATTCAAGAGG TCCGGG GTTCAATTCCCCGTCTGTCCA	GCA
PF3D7_0620800	tRNA Alanine	GGGCTAGTAGTGTAGTGGTATCACGTTTCGCTTCGCATGCGAAAGT TCCCGG GTTCGATTCCCGGCTGGTCCA	GCG
PF3D7_0702700	tRNA Alanine	GGGCGACTAGCTCAAGTGGTAGAGCGCTCGCTTAGCATGCGAGA GGTACG GGGATCGATAACCCCGGTCGTCCA	GCT
PF3D7_0529600	tRNA Arginin	GACGGCGTAGCCTAATGGATAAGGCGTCGGTCTTCGGAACCGAA GATTGC GGGTTCGAGTCCCGTCGTCTGTTA	CGA
	e		
PF3D7_1341000	tRNA Arginin	GCCTCTGTGGCGCAATTGGATAGCGCGTTGGACTTCTAATCCAAA GGCTGC GGGTTCGAATCCCGCCAGGGGTA	AGA
	e		
PF3D7_1369800	tRNA Arginin	GGGCCGGTAGTTCAGTTGGATAGAATGCCCCGACTACGGATCGGG AGGTTCG TGGGTTCGACTCCTGCCCCGGCCTA	CGT
	e		

PF3D7_1 370000	tRNA Arginin e	GCACCAGTGGCGTAATTGGATAGCGCAATGCCTTCCTAAGGCAA AGGTTATGGGTTCGAGTCCCATCTGGTGTA	AGG
PF3D7_0 403000	tRNA Asparag ine	GGTTCCGTAGCTCAGTTGGTTAGAGCGTGCGGCTGTTAACCGCAA GGTCGT TGGTTCGATCCCAGCCGGTACCG	AAC
PF3D7_0 714700	tRNA Aspartic acid ^a	TCCGAGATAGTATAGTGGCAAGTATTTCCGCCTGTCACGCGGAAG ACCCGG GTTCAATTCCCGGTCTCGGAG	GAC
PF3D7_1 370100	tRNA Cystein e	GGGCGTGTAGCTCAGCGGTAGAGCAGCTGACTGCAGATCAGCGG GTCCAT GGTTCAAATCCGTGCGCGCCCT	TGC
PF3D7_0 411600	tRNA Glutami c acid	TCCCACGTGGTCTAGTGGCTAGGATATTCGGCTCTCACCCGAAAG GCCCCG GTTCAATTCCCGGCGTGGGAA	GAG
PF3D7_0 527700	tRNA Glutami c acid	TCCCATGTAGTCTAGGCGGTTAGGATATTCGGCTTTCACCCGAAC GACCCG GGTTTCGAGTCCCGGCGTGGGAA	GAA
PF3D7_0 203500	tRNA Glutami ne	GGTTCTGTAGTGTAGTGGTTAGCACTGCAGACTCTGACTCTGCAA ACCTGG GTTCAAATCCCAGCAGAACCT	CAG
PF3D7_1 252000	tRNA Glutami ne ^b	GGTTTCGTAGTGTAGTGGTTAGCACTGAGGACTTTGAATCCTCCA ACCCGG GTTCGAGTCCCGGCGAGACCT	CAA
PF3D7_1 103200	tRNA Glycine	GGCGTCAATAGTCTAACGGCCATGATACCTGCCTTCCAAGCAGGT GACCCG GGTTTCGACTCCCGGTTGACGCA	GGA
PF3D7_1 370200	tRNA Glycine	GCATCTGTGGTCTAGTGGTAGAATACTTCGTTGCCATCGAAGTGA CCCGGG TTCGATTCCCGGCAGATGCA	GGC

PF3D7_0 706900	tRNA Histidin e	GTCCAAATCGTCTAGTGGTTAGGACTCCACGCTGTGGACGTGGCA ACGTAG GTTCGAATCCTGCTTTGGACA	CAC
PF3D7_0 312700	tRNA Isoleuci ne	GGTCCTATAGCTCAGTTGGTTAGAGCGTACGGCTAATAACCGTAA GGTCGG CGGTTCGAGACCGCCTGGGACCAAAA	ATT
PF3D7_0 410200	tRNA Isoleuci ne	GATCCATTAGCTCAGTGGTTAGAGCGTTCGGTCTTATGTACCGAAG GTCGTG GGTTTCGAAACCCACATGGATCA	ATA
PF3D7_0 510600	tRNA Leucine	GTCAGGATGGCCGAGTGGTCTAAGGCGCAGCGTTTAGGCCGCTGT CCGTA AGGGCGTGGGTTCGAACCCCACTCCTGACA	CTA
PF3D7_0 527800	tRNA Leucine	GCACGGGTGCCCCGAGTGGTTAAGGGGGTGGACTTAAGATCCTCT GGTCAC TAGACCGCGTGGGTTCGAACCCCACTCGTGCA	TTA
PF3D7_0 620900	tRNA Leucine	GACAGAATGGCCGAGCGGTCTAAGGCGCTACAGTCAGATTGTAG TCTGTAT AAGGCGTGGGTTCGAATCCCGCTTCTGTCA	CTG
PF3D7_0 714800	tRNA Leucine	GACAGAATGGCCGAGTGGTCTAAGGCGCAACGTTAAGGCCGTTG TCCGAA AGGGCGTGGGTTCGAACCCCACTTCTGTCA	CTT
PF3D7_1 103300	tRNA Leucine	GCACGGATGGCTGAGTGGTCTAAAGCGTCAGACTCAAGATCTGA TGAACGT AAGTTCGCATGGGTTCGAACCCCATTTCTGTGCATATTTAAAAAGG CCA	TTG
PF3D7_0 707000	tRNA Lysine	GCTGCCTTAGCTCAGTCGGTAGAGCGCCAGACTCTTAATCTGGTG GTCAGG GGTTTCGAGCCCCCTAGGCAGCT	AAG
PF3D7_0 707100	tRNA Lysine	GCCTCTTTAGCTCAGTTGGTAGAGCGCCAGACTTTTAATCTGGTG GTCAGG GGTTTCGAGTCCCCTAGGAGGCT	AAA
PF3D7_1 339100	tRNA Methion ine	AGCAGCGTAGCTCAGAGGAAGAGTGGGGGGGCTCAT AACCCCCAGGACCGT GGATCGAAACCACGCGCTGCTA	ATG

PF3D7_1	tRNA	GGGTAATTGGCGCAGTTGGTTAGCGCGCGGGTCTCAT	ATG
438300	Methionine	AATCCCGAGGTCG TGAGTTCGATCCTCACATTACCCA	
PF3D7_0	tRNA	GCCGTGATAGCTCAGTTGGGAGAGCGTCAGACTGAAGATCTGAA	TTC
514400	Phenylalanine	GGTCCC TGGTTCGATCCCTGGTCACGGCA	
PF3D7_1	tRNA	GGCTACTTGATCTAGTGGTATGATTCTTGCTTCGGGTGCAAGACG	CCG
216800	Proline	TCCCGG GTTCGATTCCCGGAGTAGCCC	
PF3D7_1	tRNA	GGCTACTTAGTCTAGTGGTATGATTCTCGCTTTGGGTGCGAGAGG	CCA
339200	Proline	TCCCGG GTTCAATTCCCGGAGTAGCCC	
PF3D7_1	tRNA	GGCTACTTAGTCTAGTGGTATGATTCTCTCTTAGGGTGGGAGAGG	CCT
418400	Proline	TCCCGGGTTCGATTCCCGGAGTAGCCC	
PF3D7_0	tRNA	GACAGTGTGTCCGAGTGGTTAAGGAGTCAGACTAGAAATCTGGT	TCT
410100	Serine	AGGCTTT GCCTGCGCAGGTTTCAATCCTGCCGCTGTCTG	
PF3D7_0	tRNA	GACAGTTTGCCCGAGTGGTTAAGGGGTTGGACTTGAAATCCAATG	TCA
621600	Serine	AGCTTT GCTTGCGCAGGTTTCAATCCTGCAGCTGTCTG	
PF3D7_0	tRNA	GATAACGTGCCCGAGTGGTTAAGGGGTTGGACTGCTAATCCAATG	AGC
714900	Serine	GGTTCT GCCCGCGCAGGTTTCAATCCTGCCGTTGTCTG	
PF3D7_1	tRNA	GACAGTTTGTCCGAGTGGTTAAGGAGGTTGACTCGAAATCAACTG	TCG
337600	Serine	GGCTCT GCCCGCACAGGTTCAAATCCTGTAGCTGTCTG	
PF3D7_0	tRNA	GCCTGCATAGCTTAGTTGGTAGAGCATCCGCCTAGTAAGCGGAAG	ACT
706800	Threonine	GTCATC GGTTTCTGACTCCGGTTGTAGGCT	
PF3D7_0	tRNA	GCCACCTTAGCACAGTGGTAGTGCGTTGGTCTTGTAACCAAAGG	ACA
730700	Threonine	TCGGGA GTTCGATCCTCCCAGGTGGCT	
PF3D7_1	tRNA	GCCATCTTAGCACAGTTGGTAATGCGTTTGATTTCGTAATCAAAAAG	ACG
355400	Threonine	ATCGTGG GTTCGACTCCCACAGGTGGCT	

PF3D7_1	tRNA	GGGCCTATGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAG	TGG
369900	Tryptophan	GCTGGG GGTTCGAATCCCTCTGGGCTCA	
PF3D7_0	tRNA	CCGATGATAGCTCAGTTGGTAGAGCGGCAGACTGTAGTTGAAAT	TAC
702800	Tyrosine	GGTTATC TGTTGGTCACCGGTTTCGATTCCGGTTCATCGGA	
PF3D7_0	tRNA	GCGAGCATGGTCTAGTGGCTATGACGTTTCGCTCACACGCGAAAG	GTG
312600	Valine	ATCCC GAGTTCGATCCTCGGTGCTCGTATTAA	
PF3D7_0	tRNA	GCGGGCATGGTCTAGTGGCTATGACGCCTGCCTTACACGCAGGAG	GTA
730600	Valine	ATCCC GAGTTCGATCCTCGGTGCCCCGTA	
PF3D7_1	tRNA	GCGGGCATGGTCTAGTGGCTATGACGCCTGCCTAACACGCAGGA	GTT
251900	Valine	GATCCC GAGTTCGATCCTCGGTGCCCCGTA	

7.3.Appendix III

Amino acid	Synonymous codons (mRNA: 5' -> 3')	Cognate tRNA iso-acceptor (base 32 to 38)	Genome average codon usage	Presumptive wobble modification
Ala	GCA	TT <u>I</u> GCAT	0.425	<u>ncm</u> ⁵ UGC
	GCC	NA	0.107	NA
	GCG	TT <u>C</u> GCAT	0.052	NR
	GCU	TT <u>A</u> GCAT	0.416	<u>I</u> GC
Arg	AGA	CT <u>I</u> CTAA	0.614	<u>mcm</u> ⁵ UCU
	AGG	TT <u>C</u> CTAA	0.158	NR
	CGA	CT <u>I</u> CGGA	0.087	NR
	CGC	NA	0.018	NA
	CGG	NA	0.011	NA
	CGU	CT <u>A</u> CGGA	0.113	<u>I</u> CG
Asn	AAC	CT <u>G</u> TTAA	0.154	NR
	AAU	NA	0.846	NA
Asp	GAC	CT <u>G</u> TCAC	0.143	NR
	GAU	NA	0.857	NA
Cys	UGC	CT <u>G</u> CAGA	0.126	NR
	UGU	NA	0.874	NA
Gln	CAA	CT <u>I</u> TGAA	0.860	<u>mcm</u> ⁵ <u>s</u> ² UUG
	CAG	CT <u>C</u> TGAC	0.140	NR
Glu	GAA	CT <u>I</u> TCAC	0.862	<u>mcm</u> ⁵ <u>s</u> ² UUC
	GAG	CT <u>C</u> TCAC	0.138	NR
Gly	GGA	CT <u>I</u> CCAA	0.456	<u>mcm</u> ⁵ UCC
	GGC	TT <u>G</u> CCAT	0.043	NR
	GGG	NA	0.101	NA
	GGU	NA	0.400	NA
His	CAC	CT <u>G</u> TGGA	0.154	NR
	CAU	NA	0.846	NA
Ile	AUA	CT <u>I</u> ATGT	0.545	<u>Ψ</u> AΨ
	AUC	NA	0.066	NA

	AUU	CT <u>A</u> ATAA	0.389	<u>I</u> AU
Leu	CUA	TT <u>T</u> AGGC	0.076	NR
	CUC	NA	0.023	NA
	CUG	GT <u>C</u> AGAT	0.020	NR
	CUU	TT <u>A</u> AGGC	0.121	NR
	UUA	CT <u>T</u> AAGA	0.622	<u>ncm</u> ⁵ UmUA
	UUG	CT <u>C</u> AAGA	0.138	<u>m</u> ⁵ CAA
Lys	AAA	CT <u>T</u> TTAA	0.822	<u>mcm</u> ⁵ s ² UUU
	AAG	CT <u>C</u> TTAA	0.178	NR
Met	AUG	CT <u>C</u> ATAA	1.000	NR
Phe	UUC	CT <u>G</u> AAGA	0.162	<u>Gm</u> AA
	UUU	NA	0.838	NA
Pro	CCA	TT <u>T</u> GGGT	0.446	<u>ncm</u> ⁵ UGG
	CCC	NA	0.104	NA
	CCG	TT <u>C</u> GGGT	0.046	NR
	CCU	TT <u>A</u> GGGT	0.405	NR
Ser	AGC	CT <u>G</u> CTAA	0.059	NR
	AGU	NA	0.294	NA
	UCA	CT <u>T</u> GAAA	0.277	<u>ncm</u> ⁵ UGA
	UCC	NA	0.085	NA
	UCG	CT <u>C</u> GAAA	0.047	NR
	UCU	CT <u>A</u> GAAA	0.238	<u>I</u> GA
Thr	ACA	CT <u>T</u> GTAA	0.529	<u>ncm</u> ⁵ UCU
	ACC	NA	0.116	NA
	ACG	TT <u>C</u> GTAA	0.092	NR
	ACU	CT <u>A</u> GTAA	0.263	<u>I</u> GU
Trp	UGG	CT <u>C</u> CAGA	1.000	<u>Cm</u> CA
Tyr	UAC	CT <u>G</u> TAGT	0.114	NR
	UAU	NA	0.886	NA
Val	GUA	CT <u>T</u> IACAC	0.425	<u>ncm</u> ⁵ UAC
	GUC	NA	0.065	NA
	GUG	CT <u>C</u> ACAC	0.114	NR
	GUU	CT <u>A</u> ACAC	0.396	<u>I</u> AC