NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS

A metabolomics approach for characterising tuberculosis

Ilse Olivier

Thesis submitted for the degree Philosophiae Doctor in Biochemistry at the Centre for Human Metabonomics, School for Physical and Chemical Sciences, North-West University, Potchefstroom campus

Promotor: Prof. Du Toit Loots

May 2012

Potchefstroom

Education is not the filling of a bucket, but the lighting of a fire.
˜ W.B. Yeats ˜

TABLE OF CONTENTS

CHAPTER 4: The effect of rifampicin-resistance conferring mutations on the fatty acid metabolome of *M. tuberculosis*

CHAPTER 5: **Further development and comparison of a simpler fatty acid and total metabolome extraction procedure for differentiating various *Mycobacterium* species and *P. aeruginosa***

ACKNOWLEDGEMENTS

The contributions of the following institutions and individuals made towards the successful completion of this study are hereby acknowledged:

The Royal Tropical Institute (KIT), Amsterdam, The Netherlands for providing access to all the bacterial samples used, Hans de Ronde for performing bacterial culture, and Dr. Richard Anthony and Prof. Paul Klatser for useful discussions.

AMPATH laboratories, Pretoria, South Africa (especially Dr. Jennifer Coetzee and Eileen Grabowski), for providing the patient collected sputum samples used in this investigation.

Dr. Gerhard Koekemoer (Statistical Consultation Services, North-West University, Potchefstroom, South Africa) for his guidance with regards to the statistical data analyses.

The National Research Foundation (NRF) of South Africa, the Technology Innovation Agency (TIA) and the North-West University (NWU) for the research grants provided.

Mr. Peet Jansen van Rensburg for assistance with the analytical equipment used.

Prof. Du Toit Loots, my study supervisor, the other members of the NWU Infectious Disease Metabolomics Laboratory (including Enrike Weyers, Reinart Meissner-Roloff, and Nelus Schoeman), and all the other students and staff members of the Centre for Human Metabonomics, for sharing their expertise and friendship with me during this time.

My parents, Hein and Thelma, for their ongoing physical, emotional and financial support during the course of this Ph.D. study.

My fiancé, Dupi, for all your love, patience and support.

SUMMARY

In 2001, the WHO declared tuberculosis (TB) a global emergency, as one third of the world's population suffered from latent *M. tuberculosis* infection. Today, a decade later, millions of people still die worldwide as a result of this disease. This growing TB incidence may be ascribed to a variety of reasons, including, amongst others, the inadequacies associated with the currently available diagnostic methods and TB treatment regimes, especially when considering the growing MDR-TB and HIV epidemics.

This study investigated the potential of metabolomics as a tool for characterising TB and various TB-causing bacteria, allowing for a better understanding of TB disease mechanisms, which may ultimately lead to improved diagnostic and treatment regimens.

Firstly, we investigated the potential of a fatty acid, metabolomics approach to characterise various cultured *Mycobacterium* species. For this exploration, three fatty acid extraction procedures, prior to GC-MS analyses, were compared based on their respective repeatability and extraction capacities. Using the data obtained from the analyses done with the most optimal extraction approach (the modified Bligh-Dyer method), multivariate statistical analyses were able to differentiate between the various *Mycobacterium* species at a detection limit of 1 x 10^3 bacterial mL^{-1}, in 16 hours. Subsequently, the compounds best describing the variation between the sample groups were identified as potential metabolite markers and were discussed in the light of previous studies.

The most optimal GC-MS, fatty acid metabolomics approach, mentioned above, was then applied to analyse and characterise a wild-type *M. tuberculosis* parent strain and two rifampicin-resistant conferring *rpoB* mutants (S522L and S531L). Due to the variation in their fatty acid profiles, a clear differentiation was achieved between these *M. tuberculosis* sample groups, and those metabolites contributing most to this variation were identified as metabolite markers characteristic for rifampicin-resistance. The altered metabolite markers detected in the *rpoB* mutants propose a decreased synthesis of various 10-methyl branched-chain fatty acids and cell wall lipids, and an increased use of the shorter-chain fatty acids and alkanes as alternative carbon sources. Furthermore, the *rpoB* S531L mutant, previously reported to occur in well over 50% of all clinical rifampicin-resistant *M. tuberculosis* strains, showed a better capacity for using these alternative energy sources, in comparison to the less frequently detected *rpoB* S522L mutant.

The developed fatty acid GC-MS metabolomics approach was then successfully adapted in order to improve its speed, cost and complexity. This improved fatty acid extraction method

was furthermore compared to another, similar approach (total metabolome extraction method), developed for the extraction of a much wider variety of compounds, prior to GC-MS and statistical data analyses. Although both these methods show promise for bacterial characterisation using matabolomics, the total metabolome extraction method proved the better of the two methods because it is comparatively simpler, faster (taking less than 4 hours), more repeatable, better differentiates between sample groups due to less within group variation, has a lower detection limit, and isolates a wider variety of biologically relevant metabolites (as opposed to fatty acids alone). We, furthermore, identified and described the occurrence of those compounds, extracted by both methods, which contribute most to the variation between the bacterial groups, in order to validate these methods for metabolomic applications and the isolation of compounds with biological relevance.

In order to evaluate the potential of this developed metabolomics approach for application to biological samples other than bacteriological cultures, it was adapted for the direct analyses of complex sputum samples. For this application, four sputum pre-extraction preparation methods, including three standard *Mycobacterium* cell isolation procedures (Sputolysin, NALC-NaOH, and NaOH) and a fourth, applying only a simple ethanol homogenisation step, prior to direct sputum extraction, were compared. Of these methods, the ethanol homogenisation method proved to have the best comparative extraction efficiency, repeatability and differentiation capacity, when used in combination with the previously developed metabolomics methods. Subsequently, when applying this approach to patient collected sputum samples, a set of metabolite markers, differentiating the TB-positive from the TB-negative samples, were identified. These markers could directly be linked to: 1) the physical presence of the *M. tuberculosis* in these samples; 2) changes in the bacterial metabolome due to *in vivo* growth conditions and; 3) changes in the human metabolome due to pulmonary *M. tuberculosis* infection.

In addition to the proposal of a number of new hypotheses, explaining various mechanisms of TB and drug-resistant TB, the mapping of the newly identified metabolite markers to known metabolic pathways led to the confirmation of various previously suggested metabolic pathways and alterations thereof due to an assortment of perturbations. Therefore, this study significantly contributes to the characterisation of various TB causing bacteria, rifampicin-resistant *M. tuberculosis* strains and the TB disease state, which may in future lead to the development of innovative TB vaccination, diagnostic and treatment protocols.

Key words: Metabolomics; *Mycobacterium*; Rifampicin-resistance*;* Tuberculosis.

OPSOMMING

AFRIKAANSE TITEL:

'n Metabolomika benadering om tuberkulose te karakteriseer

In 2001 het die Wêreld Gesondheid Organisasie tuberkulose (TB) as 'n globale noodtoestand verklaar aangesien een derde van die wêreld se bevolking aan latente *M. tuberculosis* infeksie gely het. Vandag, 'n dekade later, sterf miljoene mense steeds wêreldwyd as gevolg van hierdie siekte. Hierdie groeiende voorkoms van TB kan toegeskryf word aan 'n verskeidenheid van redes wat onder andere die oneffektiwiteit van die huidig beskikbare diagnostiese metodes en TB-behandelings regimes, veral wanneer die groeiende voorkoms van medisyne weerstandbiedende TB en MIV-epidemies in ag geneem word, insluit.

Hierdie studie het die potensiaal van metabolomika vir die karakterisering van TB en verskeie TB veroorsakende bakterieë, om sodoende 'n beter begrip van TB meganismes te bekom, ondersoek. Hierdie karakterisering kan uiteindelik lei tot die ontwikkeling van beter diagnostiese en behandeling regimes.

Eerstens het ons die potensiaal van 'n vetsuur, metabolomika benadering ondersoek om verskillende gekweekte *Mycobacterium* spesies te karakteriseer. Vir hierdie verkenning is drie vetsuur ekstraksie prosedures, voor GC-MS analisering, vergelyk op grond van hul onderskeie herhaalbaarheid en ekstraksie vermoëns. Deur gebruik te maak van die GC-MS data verkry na die analisering van die bakteriële ekstrakte, vanaf die mees optimale ekstraksie benadering (die gewysigde Bligh-Dyer metode), kon ons onderskeid tref tussen die verskillende *Mycobacterium* spesies by 'n deteksie limiet van 1×10^3 bakterie mL^{-1}, in 16 ure. Vervolgens is die verbindings wat die variasie tussen die groepe die beste beskryf geïdentifiseer as potensiële metaboliet merkers en bespreek in die lig van vorige studies.

Die bogenoemde, mees optimale GC-MS, vetsuur metabolomika benadering, was daarna aangewend om 'n wilde-tipe *M. tuberculosis* ouer en twee rifampisien-weerstandbiedende *rpoB* mutante (S522L en S531L) te ontleed en te karakteriseer. As gevolg van die variasie in hul vetsuur profiele kon 'n duidelike onderskeid getref word tussen hierdie *M. tuberculosis* monster groepe, en die metaboliete wat die meeste bydra tot hierdie variasie is geïdentifiseer as metaboliet merkers wat kenmerkend is vir rifampisien-weerstandbiedendheid. Die metaboliet merkers wat geïdentifiseer was vir die *rpoB* mutante dui op 'n afname in die sintese van verskeie 10-metiel vertakte ketting vetsure en selwand lipiede, en 'n toename in die gebruik van die korter ketting vetsure en alkane as alternatiewe koolstof bronne. Verder was dit

waargeneem dat die *rpoB* S531L mutant, voorheen gerapporteer om voor te kom in meer as 50% van alle kliniese rifampisien-weerstandbiedende *M. tuberculosis* monsters, 'n beter kapasiteit het om hierdie alternatiewe energiebronne te gebruik, in vergelyking met die *rpoB* S522L mutant, wat minder klinies voorkom.

Die ontwikkelde vetsuur GC-MS metabolomika benadering is hierna suksesvol aangepas om sy spoed, koste en kompleksiteit te verbeter. Hierdie verbeterde vetsuur ekstraksie metode is verder vergelyk met ’n ander, soortgelyke benadering (totale metaboloom ekstraksie metode), wat ontwikkel is om 'n veel wyer verskeidenheid van verbindings te ekstraheer, voor GC-MS en statistiese data ontleding. Alhoewel albei hierdie metodes beloofde resultate getoon het vir bakteriële karakterisering met behulp van metabolomika, het die totale metaboloom ekstraksie metode beter gevaar, want dit: 1) is relatief makliker, 2) is vinniger (neem minder as 4 ure), 3) is meer herhaalbaar en onderskei dus beter tussen die bakteriële monster groepe as gevolg van minder binne groep variasie 4) het 'n laer deteksie limiet en 5) isoleer 'n groter verskeidenheid van biologies relevante metaboliete (in teenstelling met vetsure alleen). Ons het voorts die metaboliete wat die meeste bydra tot die variasie tussen die bakteriële groepe, geëkstraheer deur beide metodes, geïdentifiseer en die voorkoms van hierdie verbindings bespreek, om sodoende hierdie metodes te evalueer vir metabolomika toepassings en hul vermoë om biologiese relevante verbindings te isoleer.

Om ten einde die potensiaal van hierdie ontwikkelde metabolomika benadering te evalueer vir toepassing op ander biologiese monsters as bakteriologiese kulture, is dit aangepas vir die direkte ontleding van komplekse sputum monsters. Vir hierdie aanpassing is vier sputum pre-ekstraksie voorbereiding metodes, insluitende drie standaard *Mycobacterium* sel isolasie prosedures (Sputolysin, NALC-NaOH, en NaOH) en 'n vierde, 'n eenvoudige etanol homogenisasie stap, voor sputum ekstraksie, vergelyk. Tussen hierdie metodes het die etanol homogenisasie metode die beste vergelykende ekstraksie doeltreffendheid, herhaalbaarheid en differensiasie kapasiteit getoon wanneer dit gebruik was in kombinasie met die voorheen ontwikkelde metabolomika metodes. Daarna is hierdie benadering toegepas op pasiënt versamelde sputum monsters om ’n stel metaboliet merkers, wat die meeste verskil tussen die TB-positiewe van die TB-negatiewe monsters, te identifiseer. Hierdie merkers kon direk gekoppel word aan: 1) die fisiese teenwoordigheid van *M. tuberculosis* in hierdie monsters; 2) veranderinge in die bakteriële metaboloom as gevolg van *in vivo* groei toestande en 3) die veranderinge in die menslike metaboloom as gevolg van pulmonêre *M. tuberculosis* infeksie.

Benewens die voorstel van 'n aantal nuwe hipoteses, wat verskeie meganismes van TB en medisyne weerstandbiedende TB verklaar, het die karakterisering van die nuut geïdentifiseerde metaboliet merkers gelei tot die bevestiging van verskeie voorheen voorgestelde metaboliese

bane en wysigings van bekende metaboliese bane as gevolg van 'n verskeidenheid van eksterne faktore. Daarom dra hierdie studie betekenisvol by tot die karakterisering van verskeie TB veroorsaak bakterieë, rifampisien-weerstandbiedende *M. tuberculosis* en die TB siekte toestand, wat dalk in die toekoms kan lei tot die ontwikkeling van innoverende TB inenting-, diagnostiese- en behandelingsprotokolle.

Sleutelterme: Metabolomika; *Mycobacterium*; Rifampisien-weerstandbiedendheid; Tuberkulose.

LIST OF TABLES

Chapter 6

Chapter 7

LIST OF FIGURES

Chapter 6

Chapter 7

LIST OF ABBREVIATIONS

AIDS	acquired immunodeficiency syndrome
Alr	D-alanine racemase
AMDIS	automated mass spectral deconvolution and identification system
ATD/GC-MS	automated thermal desorption, gas chromatography and mass spectroscopy
ATP	adenosine triphosphate
BCG	bacille Calmette-Guerin
CDA	canonical discriminant analysis
CF	cystic fibroses
CFP	culture filtrate protein
CFU	colony forming unit
CoA	coenzyme-A
CTP	cytidine triphosphate
CV	coefficient of variance
CXR	chest X-rays
Ddl	D-alanine:D-alanine ligase
DNA	deoxyribonucleic acid
DOTS	directly observed treatment, short-course
DST	drug susceptibility testing
E-MTD	enhanced *Mycobacterium tuberculosis* direct
ESAT	early secreted antigen target
FAD	flavin adenine dinucleotide
FAS	fatty acid synthase
FMN	flavin mononucleotide
GABA	γ-aminobutyric acid
GC-MS	gas chromatography mass spectrometry
GCxGC-TOFMS	two-dimensional gas chromatography time of flight mass spectroscopy
GidB	7-methylguanosine methyltransferase
GTP	guanosine triphosphate
HCl	hydrochloric acid
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography
IGRA	INF-γ release assay
INF-γ	interferon-gamma
KOH	potassium hydroxide
LAM	lipoarabinomannan
MADD	multiple acyl-CoA dehydrogenase defect
MACP	mycolic acid cleavage product
mAG	mycolylarabinogalactan
MDG	millennium development goal

MDR	multidrug-resistant
MGIT	mycobacterial growth indicator tube
MODS	microscopic observation drug susceptibility
MSTFA	N-methy-N-(trimethylsilyl) trifluoroacetamide
MTBSTFA	N-methyl-N-(tert-butyldimethylsilyl) trifluoroacetamide
NaOH	sodium hydroxide
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NALC	N-acetyl-L-cysteine
NAT	nucleic acid amplification test
NTB	non-tuberculosis
PBS	phosphate buffered saline
PC	principal component
PCA	principal component analysis
PCR	polymerase chain reaction
PLS-DA	partial least squares discriminant analysis
POA	pyrazinoic acid
PPD	purified protein derivative
PZase	pyrazinamidase
RD	regions of difference
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
ROS	reactive oxygen species
RRDR	rifampicin-resistance determining region
SAM	*S*-adenosylmethionine
SDA	strand displacement amplification
SV	similarity value
TB	tuberculosis
TBSA	tuberculostearic acid
TMA	transcription mediated amplification
TMCS	trimethylchlorosilane
TST	tuberculin skin test
UTP	uridine triphosphate
VIP	variable influence on the projection volatile organic compound
VOC	volatile organic compound
WHO	World Health Organization