Title: The role of whole genome sequencing in characterizing the mechanism of action of para-aminosalicylic acid and its resistance

Authors:

Giovanni Satta[1,2], Adam A. Witney[3], Neelu Begum[1], Julio Ortiz Canseco[1], Andrew N. Boa[4], Timothy D. McHugh[1]

Affiliations:

1. Centre for Clinical Microbiology, University College London
2. Department of Infectious Disease, Imperial College London
3. Institute for Infection and Immunity, St George's University of London
4. Department of Chemistry, University of Hull

Corresponding Author:

Dr Giovanni Satta FRCPath MBBS MSc MBA PhD DTM&H
Consultant in infectious diseases and medical microbiology
Imperial College Healthcare NHS Trust, St. Mary's Hospital, Salton House, Level 3 - Room 15 London W2 1NY
Mobile: +447747814281 – Email: giovanni.satta@nhs.net

Abstract

Para-aminosalicylic acid (PAS) remains one of the drugs of last resort for the treatment of tuberculosis, but its mechanism of action is not completely understood yet. The main aim of this project was to identify new potential mechanisms of action and resistance to PAS by performing whole genome sequencing (WGS) on PAS-resistant laboratory mutants. A new variant in the *folC* gene has been identified as well as some other mutations that require further studies.

Antimicrobial Agents and Chemotherapy

AAC

Antimicrobial Agents and Chemotherapy

Para-aminosalicylic acid (PAS), also known as 4-aminosalicylic acid, was one of the first chemotherapeutic agents to be used against tuberculosis (TB) and it is currently an orphan drug only available to treat extensively drug-resistant disease[1]. Despite being used for decades, its mechanism of action is not completely understood. It has been proposed that, being an analogue of para-amino benzoic acid (PABA), PAS competes with PABA for dihydropteroate synthase, interfering in the process of folate synthesis[2]. A study using transposon mutagenesis identified mutations in the *thyA* gene that were also present in clinical isolates resistant to PAS [3]. The gene *thyA* encodes for a thymidylate synthase enzyme (essential for DNA replication and repair) and its deletion has been demonstrated to confer resistance to PAS[4]. Other studies have also identified various missense mutations in *folC* (encoding a dihydrofolate synthase) and *ribD* (alternative dihydrofolate reductase) that conferred resistance to PAS in laboratory and clinical isolates of *M. tuberculosis* [5] [6] [7]. Nevertheless, mutations in *folC* were detected in only 34.8% of resistant clinical isolates, whilst mutations of *thyA* and *ribD* were detected in 26.0% and 5.8%, respectively [6]. Hence, other mechanisms of resistance to the drug might exist. Efflux pumps have also been described conferring cross-resistance to PAS and other chemotherapeutic agents including streptomycin [8]. The main aim of this work was to investigate potential new mechanisms of action and resistance to PAS by performing whole genome sequencing (WGS) on PAS-resistant laboratory mutants.

PAS resistant laboratory mutants of reference strain H37Rv were spontaneously selected by growth on 7H10 medium with the addition of PAS. *M. tuberculosis* was cultured in 20 mL aliquots of sterile 7H9 broth and incubated for 14 days, achieving

log phase and a colony count of 0.5-1 McFarland (150-300X10^6/mL). Following the incubation period, the bacteria were concentrated by centrifugation at 10,000g for ten minutes and the entire sediment was inoculated onto pre-prepared 7H10 plates containing different concentrations of PAS, 2 µg/ml and 4 µg/ml (the critical concentration for PAS is 2 µg/ml) [9]. After incubation at 37 °C for at least 14 days, spontaneous mutants grew on the plates and they were then selected for sequencing. All selected mutants and the parent reference H37Rv were sub-cultured on Lowenstein-Jensen (LJ) slopes, DNA was extracted using the CTAB (cetyl trimethylammonium bromide) method and WGS analysis performed as previously described [10].

It was possible to grow only one PAS mutant for each critical concentration (2 µg/ml and 4 µg/ml) from a culture containing 150-300X10^6/mL bacteria. As the experiment was repeated in two separate occasions, we were able to select four mutants in total. These resistant mutants were designated PAS2 (1st and 2nd) and PAS4 (1st and 2nd).

When compared to the sequenced reference strain H37Rv, a total of seven non-synonymous single nucleotide polymorphisms (SNPs) affecting four different genes were identified in the four PAS resistant mutants. Both PAS4 mutants showed variants in the *folC* gene, a known mutation in position 2747141 and a new mutation in position 2747195 (Table 1). There was evidence of a V58I variant in the *Rv3218* gene in both the PAS2 and one of the PAS4 mutants, but reads matching both the reference and variant base were found at this site and this call is uncertain. At low coverage, a total of twenty-seven SNPs were also found in the *rrs* and *rrl* genes (16s

and 23s RNA genes) of all four mutants. However, these SNPs remain unconfirmed due to the low coverage.

WGS has been previously used to determine the mechanism of action of antituberculous agents. In the case of bedaquiline (BDQ), the authors selected and sequenced BDQ resistant *Mycobacterium smegmatis* strains and identified mutations in the proton pump of adenosine triphosphate (ATP) synthase associated with resistance [11]. In the case of PAS-resistant mutants, WGS analysis indicated that it is necessary to revisit the folate metabolic pathway to fully understand our data. The folate biosynthetic pathway starts when the aromatic precursor chorismate is converted to *p*-aminobenzoic acid (PABA) and coupled with pteridine to generate dihydropteroate. The protein encoded by *folC*, dihydrofolate synthetase, adds glutamate to the dihydropteroate forming dihydrofolate [12]. Mutations in the *folC* gene have been associated with PAS resistance but in fewer than 35% of the cases, whilst the same authors could not find any mutation in the *folP1* gene [4]. Here, both PAS4 resistant mutants gained a SNP affecting the *folC* gene (Table 1). The mutation E153A has been previously reported to confer resistance to PAS [12] and it is already included in some online database for WGS analysis. The new mutation D135G has not been previously associated with PAS resistance. However, it has been shown to be important for the linkage of α-helices in the *folC* protein structure [13] and it could represent an additional mechanism of resistance to PAS. It is interesting to note that such variants in the *folC* gene did not develop in the PAS2 mutants at lower concentration raising the hypothesis of an association with high level resistance as noted in other antituberculous drugs [14]. This could have some clinical relevance in that higher levels of resistance could be managed and potentially prevented by optimizing the dosage as previously described by other

authors [15]. It is also worth highlighting the fascinating finding (albeit unconfirmed) of *rrs* and *rrl* genes mutations. These genes are linked to the ribosomes, including the conversion from tRNAMet to tRNAfMet (Figure 1) and hence the synthesis of proteins in general. This would hypothetically resemble the mechanism of resistance of streptomycin, even if streptomycin resistance is more frequently associated with mutations in the *rpsL* gene, rather than the *rrs/rrl* [16]. As PAS and streptomycin have been companion drugs for decades and at a time when next generation sequencing technologies were not available, this mechanism of action/resistance may have been missed and it is certainly worth further research. Our analysis is based on only four resistant mutants as it was difficult to select PAS-resistant mutants in the laboratory. This could be due to the essentiality of genes involved and the negative impact that a deletion and/or an altered gene function would have on cell survival [17]. In addition, the presence of mutations in the genome needs additional confirmation that such mutations encode for significant metabolic changes.

After nearly 70 years of clinical use of PAS, WGS analysis may help in elucidating its mechanism of action, but further studies are still needed.

Acknowledgements

None

Funding

None to declare. The experiments described in this manuscript were performed as part of Dr Satta's PhD at University College London using internal departmental funds.

Transparency declarations

All authors have no relevant conflict to declare.

Data availability

The sequence data generated has been deposited in the European Nucleotide Archive database hosted by The European Bioinformatics Institute under project accession PRJEB36463 (ERP119659).

References

1. Donald PR, Diacon AH. 2015. Para-aminosalicylic Acid: The Return of an Old Friend. Lancet Infect Dis 15(9):1091-1099.

2. Minato Y, Thiede JM, Kordus SL, McKlveen EJ, Turman BJ, Baughn AD. 2015. Mycobacterium tuberculosis folate metabolism and the mechanistic basis for para-aminosalicylic acid susceptibility and resistance. Antimicrob Agents Chemother 59(9):5097-5106.

3. Mathys V, Wintjens R, Lefevre P, Bertout J, Singhal A, Kiass M, Kurepina N, Wang XM, Mathema B, Baulard A, Kreiswirth BN, Bifani P. 2009. Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of Mycobacterium tuberculosis. Antimicrob Agents Chemother 53(5):2100-2109.

4. Fivian-Hughes AS, Houghton J, Davis EO. 2012. Mycobacterium tuberculosis thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid. Microbiology 158(Pt 2):308-318.

5. Zhao F, Wang XD, Erber LN, Luo M, Guo AZ, Yang SS, Gu J, Turman B J., Gao YR, Li DF, Cui ZQ, Zhang ZP, Bi LJ, Baughn AD, Zhang XE, Deng J Y. 2014. Binding pocket alterations in dihydrofolate synthase confer resistance to para-aminosalicylic acid in clinical isolates of Mycobacterium tuberculosis. Antimicrob Agents Chemother 58(3):1479-1487.

6. Zhang X, Liu L, Zhang Y, Dai G, Huang H, Jin Q. 2015. Genetic determinants involved in p-aminosalicylic acid resistance in clinical isolates

158 from tuberculosis patients in northern China from 2006 to 2012. Antimicrob
159 Agents Chemother 59(2):1320-1324.

160 7. Wei W, Yan H, Zhao J, Li H, Li Z, Guo H, Wang X, Zhou Y, Zhang X, Zeng
161 J, Chen T, Zhou L. 2019. Multi-omics comparisons of p-aminosalicylic acid
162 (PAS) resistance in folC mutated and un-mutated Mycobacterium
163 tuberculosis strains. Emerging Microbes & Infections 8(1):248-261.

164 8. Ramón-García S, Mick V, Dainese E, Martín C, Thompson CJ, De Rossi E,
165 Manganelli R, Aínsa JA. 2012. Functional and genetic characterization of
166 the tap efflux pump in Mycobacterium bovis BCG. Antimicrob Agents
167 Chemother 56(4):2074-2083.

168 9. Heifets L. Role of phenotypic methods for drug susceptibility testing of M.
169 tuberculosis isolates in the era of MDR and XDR epidemics. In: McHugh
170 TD, ed. Tuberculosis: laboratory diagnosis and treatment strategies.
171 London, UK: CABI; 2013.

172 10. Satta G, Witney AA, Shorten RJ, Karlikowska M, Lipman M, McHugh TD.
173 2016. Genetic variation in Mycobacterium tuberculosis isolates from a
174 London outbreak associated with isoniazid resistance. BMC Med 14(1):117.

175 11. Andries K, Verhasselt P, Guillemont J, Göhlmann HW, Neefs JM, Winkler
176 H, Van Gestel J, Timmerman P, Zhu M, Lee E, Williams P, de Chaffoy D,
177 Huitric E, Hoffner S, Cambau E, Truffot-Pernot C, Lounis N, Jarlier V. 2005.
178 A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium
179 tuberculosis. Science 307(5707):223-227.

12. Green JM, Matthews RG. 2007. Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives. EcoSal Plus 2(2).

13. Zhao F, Wang XD, Erber LN, Luo M, Guo AZ, Yang SS, Gu J, Turman BJ, Gao YR, Li DF, Cui ZQ, Zhang ZP, Bi LJ, Baughn AD, Zhang XE, Deng JY. 2014. Binding Pocket Alterations in Dihydrofolate Synthase Confer Resistance to Para-Aminosalicylic Acid in Clinical Isolates of Mycobacterium tuberculosis. Antimicrobial Agents and Chemotherapy 58(3):1479.

14. Lempens P, Meehan CJ, Vandelannoote K, Fissette K, de Rijk P, Van Deun A, Rigouts L, de Jong BC. 2018. Isoniazid resistance levels of Mycobacterium tuberculosis can largely be predicted by high-confidence resistance-conferring mutations. Sci Rep 19;8(1):3246.

15. Sy SK, de Kock L, Diacon AH, Werely CJ, Xia H, Rosenkranz B, van der Merwe L, Donald PR. 2015. N-acetyltransferase Genotypes and the Pharmacokinetics and Tolerability of Para-Aminosalicylic Acid in Patients With Drug-Resistant Pulmonary Tuberculosis. Antimicrob Agents Chemother 59(7):4129-38.

16. Springer B, Kidan YG, Prammananan T, Ellrott K, Bottger EC, Sander P. Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. 2001. Antimicrob Agents Chemother 45(10):2877-2884.

17. DeJesus MA, Gerrick ER, Xu W, Park SW, Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrt S, Fortune SM, Sassetti CM, Ioerger TR. 2017.

204 Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis
205 Genome via Saturating Transposon Mutagenesis. mBio 8(1):e02133-16.

206

207

208

209 **Tables and figures**

Gene	Function	SNP position			
		PAS2 1st	**PAS2 2nd**	**PAS4 1st**	**PAS4 2nd**
Rv1392 (metK)	S-adenosylmethionine synthetase			1566981	
Rv2447c (folC)	folylpolyglutamate synthase			2747195 (D135G)	2747141 (E153A)
Rv3218	hypothetical protein	3594639* (V58I)	3594639* (V58I)	3594639 (V58I)	
Rv3759c (proX)-Rv3760 intergenic	Possible osmoprotectant binding lipoprotein; conserved membrane protein				4205442
At low coverage, a total of twenty-seven SNPs were found in the *rrs* and *rrl* genes (16s and 23s RNA genes) of all four mutants. Data not shown due to low coverage.					

210

211 **Table 1: List of SNPs in the PAS resistant mutants.** Table above shows the genes involved in the PAS-resistant

212 mutants, with the respective function and SNP position in the genome. Only high quality non-synonymous and

intergenic SNPs were considered (*mixed base calls and therefore lower quality evidence in these mutants for these sites). The hypothetical amino acid changes caused by the SNP are also shown in brackets.

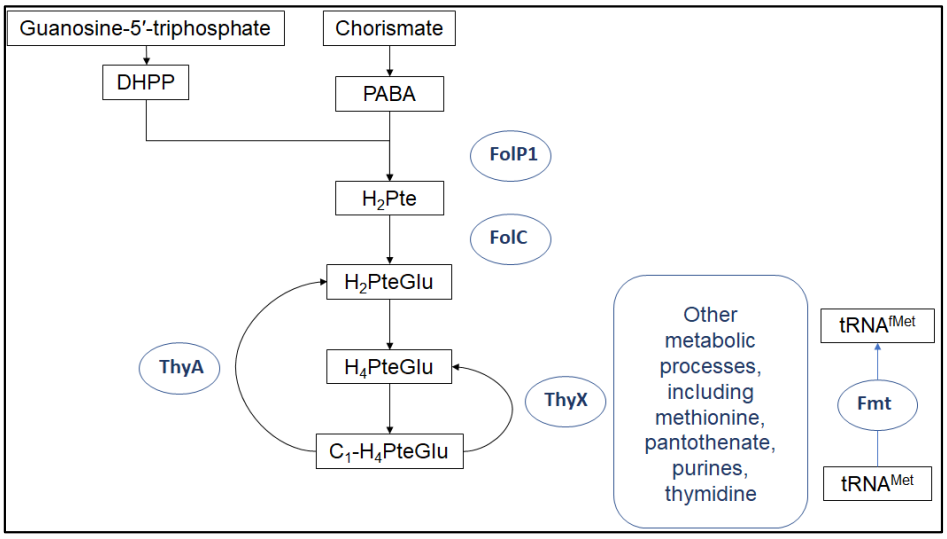

Figure 1: Folate metabolism in *M. tuberculosis*. Abbreviations: DHPPP, 7,8-dihydropterin pyrophosphate; PABA, para-aminobenzoic acid; H_2Pte, dihydropteroate; H_2PteGlu, dihydrofolate; H_4PteGlu, tetrahydrofolate; C1-H_4PteGlu, various single-carbon-modified species of H_4PteGlu; ThyX, thymidylate synthase; $tRNA^{Met}$, methionyl-tRNA; $tRNA^{fMet}$, N-formylmethionyl-tRNA; Fmt, formyl-methionine transferase.

Antimicrobial Agents and Chemotherapy