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. Division of Infection & Immunity, University College London

2. Department of Infectious Disease, Imperial College London

3. Institute of 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

Synopsis

Para-aminosalicylic acid (PAS) was one of the first antibiotics to be used against tuberculosis and it is still one of the drugs of last resort available to treat extensively drug-resistant disease. Despite being on the market for decades, the mechanism of action of PAS is still not completely understood. The main aim of this project was to identify new potential mechanisms of action of and resistance to PAS by performing whole genome sequencing (WGS) on PAS-resistant laboratory mutants.

PAS resistant laboratory mutants of the *Mycobacterium tuberculosis* reference strain H37Rv were spontaneously selected and WGS analysis identified different single nucleotide polymorphisms (SNPs) when compared with the sequenced reference strain. A new variant in the *folC* gene has been identified as well as some other mutations that require further studies.

Introduction

Para-aminosalicylic acid (PAS), also known as 4-aminosalicylic acid, was one of the first antibiotics to be used against tuberculosis (TB) and it is still one of the last remaining drugs available to treat extensively drug-resistant (XDR) disease. PAS is generally regarded as a bacteriostatic agent, but some studies have demonstrated a dose-related early bactericidal effect on metabolically active populations of bacilli [1]. This dose-dependent effect would be similar to the one achieved by aminoglycosides in contrast to the time-dependent effect of many other antibiotics [2].

Despite being on the market for decades, the mechanism of action of PAS 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 [3]. However, PAS may have other biological impacts, for example there is evidence that PAS can also interfere with the mycobacterial iron acquisition, causing disruption to cell replication [4].

Resistance to PAS was described shortly after its introduction into clinical use and recent studies have demonstrated that PAS resistance in *Mycobacterium tuberculosis* is probably caused by multiple mechanisms. A study using transposon mutagenesis identified mutations in the *thyA* gene that were also present in clinical isolates resistant to PAS [5]. 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 [6]. 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* [7] [8] [9]. 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 [8]. Hence, other mechanisms of resistance to the drug might exist. Efflux pumps have also been described conferring cross-resistant to PAS and other antibiotics including streptomycin [10].

The main aim of this work was to investigate potential new mechanisms of action of and resistance to PAS by performing whole genome sequencing (WGS) on PAS-resistant laboratory mutants.

Materials and methods

Selection of resistant 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) [11]. Plates without bacteria were also included as a negative control for the experiment and a Columbia blood agar plate was used to check for any contamination. After incubation at 37 °C for at least 14 days, spontaneous mutants grew on the plates and they were then selected for sequencing.

Whole genome sequencing analysis – 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 [12].

Results

Selection of resistant mutants – It was possible to grow and select 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).

Whole genome sequencing analysis – 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.

Discussion

WGS has been successfully used to determine the mechanism of action of antibiotics. 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 [13]. It is important to note that the mutations conferring resistance may involve different stages of the drug metabolism: mutations can affect enzymes that convert the compound into an active form, the target gene, the regions responsible for up- or down-regulation of targets and activating enzymes, efflux pumps or detoxifying enzymes. It is also possible that the same drug may be inactivated by different mechanisms of resistance and some of them may be common to other compounds, as already described for other anti-MTB drugs such as isoniazid and ethionamide [14]. WGS could therefore contribute to elucidating global mechanisms of action/resistance.

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 [15]. 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 [7]. 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 [11] 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 [16] 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 rising the hypothesis of an association with high level resistance as noted in other antituberculous drugs [17]. 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* [18].

Our analysis is based on only four resistant mutants. It was difficult to select PAS-resistant mutants in the laboratory, in sharp contrast with the early resistant clinical isolates described after the introduction of PAS [19]. 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. A recent paper [20] has provided a comprehensive essentiality analysis of the *M. tuberculosis* genome via saturating transposon mutagenesis and *folC* is included in the list of essential genes. At the same time, there are other limitations that need to be taken into consideration: 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, not completely understood. Further studies are still needed to confirm this and to complete the

complex puzzle of PAS resistance. Elucidating its mechanisms of action/resistance can also influence the development of new compounds and PAS analogues targeted at similar metabolic pathways.

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.					

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

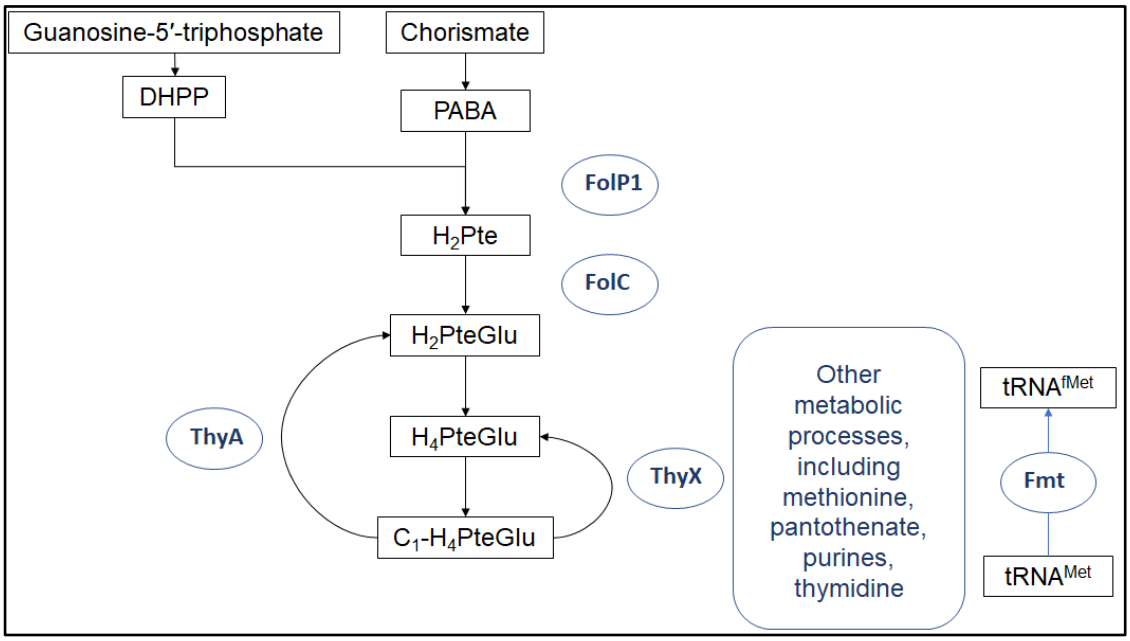

Figure 1: Folate metabolism in *M. tuberculosis*. Abbreviations: DHPPP, 7,8-dihydropterin pyrophosphate; PABA, para-aminobenzoic acid; H_2Pte, dihydropteroate; H_2PteGlu, dihydrofolate; H_4PteGlu, tetrahydrofolate; C1-H_4PteGlu, various single-carbon-modified species of H_4PteGlu; ThyX, thymidylate synthase; $tRNA^{Met}$, methionyl-tRNA; $tRNA^{fMet}$, N-formylmethionyl-tRNA; Fmt, formyl-methionine transferase [3].

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.

References

1. Xie, Z., N. Siddiqi, and E.J. Rubin, *Differential antibiotic susceptibilities of starved Mycobacterium tuberculosis isolates.* Antimicrob Agents Chemother, 2005. **49**(11): p. 4778-80.
2. Donald, P.R. and A.H. Diacon, *Para-aminosalicylic acid: the return of an old friend.* Lancet Infect Dis, 2015. **15**(9): p. 1091-1099.
3. Minato, Y., et al., *Mycobacterium tuberculosis folate metabolism and the mechanistic basis for para-aminosalicylic acid susceptibility and resistance.* Antimicrob Agents Chemother, 2015. **59**(9): p. 5097-106.
4. Ratledge, C. and K.A. Brown, *Inhibition of mycobactin formation in Mycobacterium smegmatis by p-aminosalicylate. A new proposal for the mode of action of p-aminosalicylate.* Am Rev Respir Dis, 1972. **106**(5): p. 774-6.
5. Mathys, V., et al., *Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2009. **53**(5): p. 2100-9.
6. Fivian-Hughes, A.S., J. Houghton, and E.O. Davis, *Mycobacterium tuberculosis thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid.* Microbiology, 2012. **158**(Pt 2): p. 308-18.
7. Zhao, F., et al., *Binding pocket alterations in dihydrofolate synthase confer resistance to para-aminosalicylic acid in clinical isolates of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2014. **58**(3): p. 1479-87.

8. Zhang, X., et al., *Genetic determinants involved in p-aminosalicylic acid resistance in clinical isolates from tuberculosis patients in northern China from 2006 to 2012.* Antimicrob Agents Chemother, 2015. **59**(2): p. 1320-4.
9. Wei, W., et al., *Multi-omics comparisons of p-aminosalicylic acid (PAS) resistance in folC mutated and un-mutated Mycobacterium tuberculosis strains.* Emerging Microbes & Infections, 2019. **8**(1): p. 248-261.
10. Ramon-Garcia, S., et al., *Functional and genetic characterization of the tap efflux pump in Mycobacterium bovis BCG.* Antimicrob Agents Chemother, 2012. **56**(4): p. 2074-83.
11. Heifets, L., *Role of phenotypic methods for drug susceptibility testing of M. tuberculosis isolates in the era of MDR and XDR epidemics.*, in *Tuberculosis: laboratory diagnosis and treatment strategies.*, T.D. McHugh, Editor. 2013, CABI: London, UK.
12. Satta, G., et al., *Genetic variation in Mycobacterium tuberculosis isolates from a London outbreak associated with isoniazid resistance.* BMC Med, 2016. **14**(1): p. 117.
13. Andries, K., et al., *A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium tuberculosis.* Science, 2005. **307**(5707): p. 223-227.
14. Almeida Da Silva, P.E. and J.C. Palomino, *Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs.* J Antimicrob Chemother, 2011. **66**(7): p. 1417-30.
15. Green, J.M. and R.G. Matthews, *Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives.* EcoSal Plus, 2007. **2**(2).
16. Zhao, F., et al., *Binding Pocket Alterations in Dihydrofolate Synthase Confer Resistance to para-Aminosalicylic Acid in Clinical Isolates of Mycobacterium tuberculosis.* Antimicrobial Agents and Chemotherapy, 2014. **58**(3): p. 1479.
17. Lempens, P., et al., *Isoniazid resistance levels of Mycobacterium tuberculosis can largely be predicted by high-confidence resistance-conferring mutations*, in *Sci Rep*. 2018.
18. Springer, B., et al., *Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance.* Antimicrob Agents Chemother, 2001. **45**(10): p. 2877-84.
19. Turnbull, F.W., et al., *Streptomycin resistance in patients with pulmonary tuberculosis previously treated with PAS alone.* Br Med J, 1953. **1**(4822): p. 1244-6.
20. DeJesus, M.A., et al., *Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis.* MBio, 2017. **8**(1).