Covalent modification of the *Mycobacterium tuberculosis* FAS-II dehydratase by Isoxyl and Thiacetazone

Anna E. Grzegorzewicz, Nathalie Eynard, Annaïk Quémard, E. Jeffrey North, Alyssa Margolis, Jared J. Lindenberger, Victoria Jones, Jana Korduláková, Patrick J. Brennan, Richard E. Lee, Donald R. Ronning, Michael R. McNeil and Mary Jackson*

*E-mail: Mary.Jackson@colostate.edu

Supporting Information

Table of Contents

Bacterial strains and growth conditions S1
Mutagenesis of HadA S1
Expression of *hadA<sup>WT</sup>, hadA<sup>C61S</sup>, hadA<sup>C105A</sup> and hadC<sup>WT</sup> in *Mycobacterium bovis* BCG S1
Co-expression of *ethA* and *hadA<sup>C105A</sup>BC* in *E. coli* S1
Synthesis of the urea analog of Isoxyl S1
In-gel digestion, extraction of peptides and LC-MS/MS S1
Analysis of LC-MS/MS data S2
Treatment of purified HadAB with diamide S2
Dehydratase assay S2
Effect of Isoxyl and analogs on mycolic acid biosynthesis in *Mtb* S3
Structural modeling S3
Supplementary References S5
Supplementary Figures S6
Bacterial strains and growth conditions

*Escherichia coli* DH5α, the strain used for cloning, and *E. coli* BL21(DE3) were grown in LB Lennox (BD, Difco) medium at 37°C. *M. bovis* BCG Pasteur (strain 1173 P2) and *Mtb* H37Ra (ATCC 25177) were grown in Middlebrook 7H9 broth supplemented with OADC (BD, Difco) and 0.05 % Tween 80 or on solid Middlebrook 7H11 medium supplemented with OADC at 37°C. Kanamycin was used at a final concentrations of 20 to 50 µg/ml and ampicillin at a final concentration of 100 µg/ml.

Mutagenesis of HadA

The mutated hadA allele hadA-G182C which results in the mutation of Cys61 to serine was PCR-amplified from genomic DNA prepared from the ISO- and TAC-spontaneous resistant mutant BCG-R5 [Grzegorzewicz et al., 2012]. HadA<sup>C105A</sup> was generated by site-directed mutagenesis using the two single primer reaction method [Edelheit et al., 2009].

Expression of hadA<sup>WT</sup>, hadA<sup>C61S</sup>, hadA<sup>C105A</sup> and hadC<sup>WT</sup> in *Mycobacterium bovis* BCG

The wild-type (WT) and mutated hadA alleles were PCR-amplified using primers hadA.1 (5’-GCCGCCGCATATGGCGTTGAGGCGAGCATCG-3’) and hadA.2 (5’-TAAAGCTTCGAGCTTCGCAGGCGCATC-3’) and the amplicons were cloned in the NdeI and HindIII restriction sites of the mycobacterial expression plasmid pVV16 [Kordulákova et al., 2002] to yield pVV16-hadA<sup>WT</sup>, pVV16-hadA<sup>C105A</sup> and pVV16-hadA<sup>C61S</sup>. The WT hadC gene from *M. bovis* BCG (Pasteur strain 1173P2) are 100 % identical to those of *Mtb* H37rv. pVV16 allows for the constitutive expression of genes under control of the hsp60 promoter and the production of recombinant proteins harboring a hexahistidine tag at their C-terminus.

Co-expression of ethA and hadA<sup>C105A</sup>-hadB-hadC in *E. coli*

The ethA gene was PCR-amplified from *Mtb* H37Rv genomic DNA using primers ethA.1 (5’-CGCCCGCATATGACCCAGACCTCGACGTTG-3’) and ethA.2 (5’-CGCAAGCTTAACCCCGCAGCAGGCC-3’) and cloned in the NdeI and HindIII sites of pET29b (EMD Biosciences). *Escherichia coli* BL21(DE3) was co-transformed with pET29b-ethA and pEXP5 TOPO-hadA<sup>C105A</sup>BC (a derivative of pEXP5 TOPO-hadABC [Sacco et al., 2007] harboring the C105A-mutated form of hadA) and grown in LB media supplemented with 100 µg/ml ampicillin and 50 µg/ml kanamycin at 37°C. Gene expression was induced at Abs<sub>600nm</sub> = 0.6 with 0.2 mM IPTG for 4 hours at 37°C.

Synthesis of the urea analog of Isoxyl

The synthesis of the urea analog of ISO, compound 41 [Figure S2], was reported earlier [Brown et al., 2011].

In-gel digestion, extraction and LC-MS/MS analysis of peptides

In-gel trypsin digestion was performed using a standard protocol with sequencing-grade trypsin (Promega). The reduction treatment with DTT was omitted in samples prepared for drug-protein analysis. Samples were resuspended in 5 % acetonitrile and 0.1 % formic acid. Peptides were
purified and concentrated using an on-line enrichment column (Thermo Scientific 5mm, 100 mm ID x 2cm C18 column). Subsequent chromatographic separation was performed on a reverse phase nanospray column (Thermo Scientific EASYnano-LC, 3mm, 75 mm ID x 100mm C18 column) using a 30-minute linear gradient from 10 % - 30 % buffer B (100 % acetonitrile, 0.1 % formic acid) at a flow rate of 400 nanoliters/min. Peptides were eluted directly into the mass spectrometer (Thermo Scientific Orbitrap Velos) and spectra were collected over a m/z range of 400-2000 Da using a dynamic exclusion limit of 2 MS/MS spectra of a given peptide mass for 30 s (exclusion duration of 90 s). The instrument was operated in Orbitrap-LTQ mode where precursor measurements were acquired in the orbitrap (60,000 resolution) and MS/MS spectra (top 20) were acquired in the LTQ ion trap with a normalized collision energy of 35 kV. Compound lists of the resulting spectra were generated using Xcalibur 2.2 software (Thermo Scientific) with a S/N threshold of 1.5 and 1 scan/group.

**LC-MS analysis of proteins**

The LC/MS (electrospray ionization) analysis of proteins was done on an Agilent 6220 TOF (time-of-flight) mass spectrometer equipped with a multimode source (positive mode) and an Agilent 1200 binary pump. The column used was a Zorbax (2.1 × 150 mm 300 SB-C8 3.5 µM) and separation was done using a gradient of solvent A (water, 1 % formic acid) and solvent B (acetonitrile, 1 % formic acid). The drying gas temperature was 310 °C, and the vaporizer temperature was set at 200 °C. Flow rate was 0.25 ml/min, and the total run time was 27 min. The fragmentor voltage was set to 1180 V. The runs were later analyzed using the Agilent’s Mass Hunter program.

**Treatment of purified HadAB with diamide**

Purified HadA^{WT}B from *E. coli* (13.36 µg) [Sacco *et al.*, 2007] was incubated with 2 mg/ml of diazenedicarboxylic acid bis(N,N-dimethylamide) (diamide) for 1 hour at room temperature. Samples were then analyzed by LC-MS/MS and SDS-PAGE as described above and tested for enzymatic activity as described in the next section.

**Dehydratase assay**

The enzymatic activity of HadA^{C105A}B partially purified from untreated and ISO- or TAC-treated *M. bovis* BCG and that of the oxidized form of HadAB generated *in vitro* upon treatment with diamide was measured in the presence of trans-2-dodecenoyl-CoA (C12:1-CoA) as the substrate since (R)-specific enoyl hydratase/hydroxyacyl dehydratases family enzymes have been shown to preferentially catalyze the hydration reaction when isolated from their enzymatic complex [Sacco *et al.*, 2007]. Kinetic assays in the presence of C12:1-CoA were monitored at 263 nm (ΔA of 0.67 for a variation of 100 µM in substrate concentration) using a SAFAS UVmc² spectrophotometer (Monaco). Reactions were performed in a quartz cuvette in a total volume of 700 µl at 25°C in 100 mM sodium phosphate buffer (pH 7.0) containing 10 µM C12:1-CoA. After equilibration of the baseline, reactions were started by adding 80 nM of enzyme preparation and measurements were performed for 3 min. The initial velocities were determined by linear fitting. C12:1-CoA was synthesized as previously described [Quémard *et al.*, 1995]. MALDI-TOF mass spectrometry analyses of the substrate and product of the reaction were performed as described previously [Sacco *et al.*, 2007].
Effect of Isoxyl and its urea analog on mycolic acid biosynthesis in *Mtb*

Radiolabeling of whole *Mtb* H37Ra (ATCC 25177) with [1,2-14C]acetic acid (0.5 to 1 µCi/ml; specific activity, 57 Ci/mol; NEN Radiochemicals) was performed for 4 hr with shaking. The radioactive was added to the cultures at the same time as 7.5 µg/ml ISO (3 x MIC) or 20 µg/ml compound 41. Preparation of fatty acid methyl esters (FAMEs) and mycolic acid methyl esters (MAMEs) from whole cells followed procedures described earlier [Phetsuksiri et al., 1999]. FAMEs and MAMEs were analyzed by TLC on aluminum-backed silica gel 60–precoated plates F254 (Merck) using n-hexane:ethyl acetate (95:5 by vol.; 3 developments) as the eluent. Radiolabeled products were visualized by exposing the TLC plates to Kodak X-Omat AR films at −80 °C.

**Structural modeling**

The HadA and HadB sequences were fused to mimic a single chain hydratase enzyme to promote the prediction of an intact enzyme model. The linker sequence (RGQGGFGGARGPAAPEFPDR) from a putative single chain dehydratase encoded by the Rv3389c gene of *M. tuberculosis* H37Rv (RCSB accession 3KHP) was placed between the HadA and HadB protein sequences and submitted to the Phyre2 server [Kelley and Sternberg, 2009]. The ISO and TAC molecules were built in PHENIX using the REEL ligand restraints and editor module [Adams et al., 2010]. The bond lengths and angles were minimized using the eLBOW module. Using the program Coot, the ligands were fit manually within the modeled HadAB active site based on the knowledge that a covalent linkage is formed between Cys61 of HadA and the thiourea moieties of ISO and TAC [Emsley et al., 2010]. The protein preparation wizard in Glide (Schrodinger) was used to regularize the geometry of the disulfide bond forming the covalent adduct, assign bond order, add hydrogen atoms, and define the most likely protonation states for the covalent adducts. Geometric and energy minimization was performed at pH 7.0 using the OPLS_2005 force field. Molecular dynamic minimizations on the final covalent models were performed in PHENIX using the simple dynamics module. The resulting HadAB structural model gave a 100 % confidence score for 95 % of the submitted HadAB sequence. For further confirmation of accuracy and biological relevance, the HadAB model was manually compared with deposited dehydratase structures possessing the highest sequence similarity to *M. tuberculosis* HadB (RCSB accession numbers 3WEW and 3IR3). This comparison confirmed the superposition of the conserved active site aspartate and histidine residues essential for dehydratase activity.

The predicted location of HadA-C61 is at the end of an internal channel that extends from the enzyme surface near active site residues HadB-D36 and HadB-H41 [Figure 6A]. This channel accommodates all of TAC and one of the two phenyl isopentyl ether arms of ISO as well as affording disulfide bond formation between the thiol of HadA-Cys61 and the tested compounds [Figure 6B and 6C]. Multiple interactions between amino acid side chains forming this pocket appear to promote binding of either drug to the HadAB heterodimer. The polar thiourea moiety is accommodated in the channel via hydrogen bonding interactions with the side chain of HadA-D88. Since HadA-D88 is immediately across the internal channel from HadA-C61, these interactions may also play a role in positioning the activated forms of ISO or TAC near HadA-C61 to afford the formation of a disulfide bond. Additionally, the aromatic moieties of HadA-Y65 and HadB-F81 form direct interactions with the aromatic moiety of both ISO and TAC. The models suggest that HadA-Y65 interacts with either drug through π-π interactions; however, a slight shift in the position of either drug toward the protein surface would allow a similar
interaction between HadB-F81 and the phenyl moiety of either ISO or TAC. Therefore, it is challenging to assess which of these two aromatic side chains would dominate the positioning of the drug prior to disulfide bond formation. The side chain of HadB-M60 also forms van der Waals interactions with the hydrophobic portion of both drugs.

The relative bulk of the second phenyl isopentyl ether arm of ISO, relative to the corresponding hydrogen atom of TAC, is more challenging to model. Additionally, the disulfide link between HadA-C61 and ISO necessitates that one of the two nitrogen atoms is sp²-hybridized, which thereby limits rotational freedom and flexibility near the newly formed disulfide bond. For this modeling, it is assumed that the second arm of ISO will possess the sp³-hybridized nitrogen since it will allow slightly more flexibility in positioning the second arm of ISO. In this model, arm 2 binds within a pocket formed exclusively by hydrophobic residues HadA-I60, HadA-L103, HadA-V127 and HadA-V135 [Figure 6C]. While this pocket nicely accommodates the second phenyl isopentyl ether arm of ISO, other conformations are possible.
Supplementary References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221

Brown, J. R.; North, E. J.; Hurdle, J. G.; Morisseau, C.; Scarborouugh, J. S.; Sun, D.; Korduláková, J.; Scherman, M. S.; Jones, V.; Grzegorzewicz, A.; Crew, R. M.; Jackson, M.; McNeil, M. R.; Lee, R. E. *Bioorg. Med. Chem.* **2011**, **19** (18), 5585-5595.

Edelheit, O.; Hanukoglu, A.; Hanukoglu, I. *BMC biotechnology* **2009**, 9, 61.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501

Grzegorzewicz, A. E.; Korduláková, J.; Jones, V.; Born, S. E.; Belardinelli, J. M.; Vaquié, A.; Gundi, V. A.; Madacki, J.; Slama, N.; Laval, F.; Vaubourgeix, J.; Crew, R. M.; Gicquel, B.; Daffé, M.; Morbidoni, H. R.; Brennan, P. J.; Quémard, A.; McNeil, M. R.; Jackson, M. *J. Biol. Chem.* **2012**, **287** (46), 38434-41.

Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-371

Korduláková, J.; Gilleron, M.; Mikusova, K.; Puzo, G.; Brennan, P. J.; Gicquel, B.; Jackson, M. *J. Biol. Chem.* **2002**, **277** (35), 31335-31344.

Phetsuksiri, B.; Baulard, A. R.; Cooper, A.; Minnikin, D. E.; Douglas, J. D.; Besra, G. S.; Brennan, P. J. *Antimicrob. Agents Chemother.* **1999**, **43** (5), 1042-1051.

Quémard, A.; Sacchettini, J. C.; Dessen, A.; Vilcheze, C.; Bittman, R.; Jacobs, W. R., Jr.; Blanchard, J. S. *Biochemistry* **1995**, **34** (26), 8235-41.

Sacco, E.; Suarez Covarrubias, A.; O'Hare, H. M.; Carroll, P.; Eynard, N.; Jones, A. J.; Parish, T.; Daffé, M.; Backbro, K.; Quémard, A. *Proc. Natl. Acad. Sci. USA* **2007**, **104** (37), 14628-14633.
Supplementary Figures

Figure S1: Characterization of the oxidized form of HadA<sup>WT</sup>

(A) HadA<sup>WT</sup>-His was partially purified from recombinant <i>M. bovis</i> BCG cells that were either untreated or treated with ISO or TAC and analyzed by LC-MS. The ESI mass spectra show the mass of the reduced form of HadA<sup>WT</sup>-His in the untreated control (top left panel). After treatment with either drug, and consistent with the SDS-PAGE analysis, HadA<sup>WT</sup> reduced was still present (MS not presented) but, at a slightly earlier retention time on the LC column, a new protein with a mass approximately 2 amu’s lower than HadA<sup>WT</sup> (middle and bottom panels left side) was found suggesting the formation of a disulfide bond during drug treatment. Incubation of the HadA<sup>WT</sup>-His drug-treated samples with DTT prior to analysis by LC-MS revealed only the HadA<sup>WT</sup> (middle and bottom panels right side) confirming the disulfide bond by reducing it back to the untreated form. Calculated masses: HadA<sup>WT</sup>-His (reduced form) = 18,413.51; HadA<sup>WT</sup>-His (oxidized form) = 18,411.49; HadB = 14,772.82. An unidentified protein at a molecular mass of 16,951.7 was observed in the non-drug-treated samples.
Figure S1 (continued)

(B) HadA<sup>WT</sup>-His purified from ISO-treated BCG cells was in-gel trypsin digested and the peptides analyzed by LC-MS. The ESI spectra show the mass of the dipeptide formed by the Cys61- and Cys105-containing peptides upon reaction of the two cysteines (covalently linked through a disulfide bond). The dipeptide was not present in the non drug-treated sample (data not shown). Calculated masses for the dipeptide at charge states (+3) and (+4) = 1352.37 and 1014.53, respectively.

(C) Structure of the dipeptide

| Mass (m/z) | Counts |
|-----------|--------|
| 1015.04   | 1014.78| 1015.30| 1015.55| 1015.79| 1016.05|
| 1352.71   | 1353.03| 1353.38| 1353.71| 1354.05|

FEKPIVAGDKLYCDVVVDSVR

GLLAPLTFCVFGYK
Figure S2: The urea analog of ISO (compound 41) does not interact with HadA$^{C105A}$-His.

(A) Structure of compound 41.

(B) HadA$^{C105A}$-His was partially purified from recombinant E. coli cells co-expressing hadA$^{C105A}BC$ and ethA that were treated with compound 41 and analyzed by LC-MS. The unmodified form of HadA$^{C105A}$-His was the only form of HadA$^{C105A}$-His detected in the sample. Calculated mass for HadA$^{C105A}$-His = 18,268.24

(C) Effect of compound 41 on mycolic acid biosynthesis in Mtb. Mtb H37Ra was treated with either no drug (0) or 7.5 µg/ml ISO (3 x MIC) or 20 µg/ml compound 41 for 4 hours. [1,2-$^{14}$C]acetate was added to the cultures at the same time as the inhibitors. Drug treatments were done in duplicate. Preparation and TLC analysis of radiolabeled fatty acids and mycolic acid methyl esters from whole cells followed earlier procedures [Phetsuksiri et al., 1999]. The same volume of samples was loaded per lane and the plates were developed using n-hexane:ethyl acetate (95:5 by vol.; 3 developments) as the eluent. Radiolabeled products were visualized by exposing the TLC plates to Kodak X-Omat AR films at −80 °C. FAMEs, fatty acid methyl esters; MAMEs, mycolic acid methyl esters.
Figure S3: Analysis of the product of the hydratase assay by MALDI-TOF mass spectrometry.

When isolated from the FAS-II complex, the HadA<sup>C105A</sup>-HadB enzyme partially purified from untreated <i>M. bovis</i> BCG cells catalyzes a hydration reaction in the presence of trans-2-enoyl-CoA. The peaks at <i>m/z</i> 988, 1,010, 1,032, and 1,054 correspond, respectively, to the [M+Na]<sup>+</sup>, [M-H+2Na]<sup>+</sup>, [M-2H+3Na]<sup>+</sup>, and [M-3H+4Na]<sup>+</sup> ions of 3-hydroxydodecanoyl-CoA (hydration product). The minor peaks at <i>m/z</i> 992, 1,014, and 1,036 stand for the di-, tri- and tetra-sodium adducts, respectively, of the residual substrate, trans-2-dodecenoyl-CoA.
**Figure S4:** Treatment of purified HadAB with diamide leads to the formation of a disulfide bond between the two Cys residues of the HadA subunit.

Purified HadAB from *E. coli* [Sacco *et al.*, 2007] was treated with diamide [see (SI) for details] to induce the formation of a disulfide bond between Cys61 and Cys105 of HadA and analyzed by SDS-PAGE and LC/MS.

Similar to the situation in Figure 2, diamide-treated HadA\textsuperscript{WT} migrates with an apparent slightly lower MW compared to the untreated sample.

The ESI mass spectrum shows the mass of the oxidized form of HadA\textsuperscript{WT}-His in the diamide-treated sample. Calculated masses: HadA\textsuperscript{WT}-His (reduced form) = 18,300.3; HadA\textsuperscript{WT}-His (oxidized form) = 18,298.28. UNTR, non diamide-treated sample.
**Figure S5:** Analysis of HadBC purified from ISO- and TAC-treated *M. bovis* BCG cells.

HadC-His was partially purified from recombinant *M. bovis* BCG cells (expressing a C-ter His-tagged form of HadC from pVV16-*hadC*WT) that were either untreated or treated with ISO or TAC as in Figure 2. Shown on the SDS-PAGE are the flow-through and last two elution fractions for each sample. HadC contains three Cys residues (Cys25; Cys34 and Cys127). HadB has no Cysteine. HadB and HadC-His show no shift in migration in the drug-treated samples.