Identification of a Mycothiol-Dependent Nitroreductase from *Mycobacterium tuberculosis*

Ana Negri, † Prisca Javidnia, ‡ Ran Mu, §∥¶ Jeremie Vendome, † Ben Gold, † Julia Roberts, † Dipti Barman, ○ Thomas Ioerger, † James C. Sacchettini, § Xiuju Jiang, ∥ Kristin Burns-Huang, † Thulasi Warrier, † Yan Ling, † J. David Warren, ○ Deena A. Oren, □ Thijs Beuming, † Hongyao Wang, § Jie Wu, § Haitao Li, § Kyu Y. Rhee, §, † Carl F. Nathan, † Gang Liu, and Selin Somersan-Karakaya †

†Schrödinger, Inc., 120 West 45th Street, New York, New York 10036, United States
‡Department of Medicine, Division of Infectious Diseases, Weill Cornell Medical College, 1300 York Avenue, New York, New York 10065, United States
§Tsinghua-Peking Joint Center for Life Sciences, §School of Pharmaceutical Sciences, and †Beijing Advanced Innovation Center for Structural Biology, School of Medicine, Tsinghua University, Beijing 100084, P. R. China
\( \text{V} \)Department of Microbiology & Immunology, ○Department of Chemistry and CLC Milstein Chemistry Core Facility, Weill Cornell Medical College, 1300 York Avenue, New York, New York 10065, United States

◆Department of Computer Science and ∆Department of Biochemistry/Biophysics, Texas A&M University, College Station, Texas 77843, United States
∇Structural Biology Resource Center, Rockefeller University, 1230 York Avenue, New York, New York 10065, United States

**ABSTRACT:** The success of *Mycobacterium tuberculosis* (Mtb) as a pathogen depends on the redundant and complex mechanisms it has evolved for resisting nitrosative and oxidative stresses inflicted by host immunity. Improving our understanding of these defense pathways can reveal vulnerable points in Mtb pathogenesis. In this study, we combined genetic, structural, computational, biochemical, and biophysical approaches to identify a novel enzyme class represented by Rv2466c. We show that Rv2466c is a mycothiol-dependent nitroreductase of Mtb and can reduce the nitro group of a novel mycobactericidal compound using mycothiol as a cofactor. In addition to its function as a nitroreductase, Rv2466c confers partial protection to menadione stress.

**KEYWORDS:** Rv2466c, nitroreductase, oxidoreductase, *Mycobacterium tuberculosis*, mycothiol, menadione, mrx-2, nitrofuranylcalanolide

After coexisting with humans for an estimated 70,000 years, † *Mycobacterium tuberculosis* (Mtb) has emerged as the single leading cause of death from an infectious disease.‡ With no other reservoir but humans, its only naturally transmitting host, Mtb must be adept at resisting stresses imposed by host immunity, including reactive oxygen species (ROS), reactive nitrogen species (RNS), and restricted access to oxygen, iron, and metabolizable sources of carbon. During adaptation to such stresses, Mtb alters its physiology and enters a slowly replicating or nonreplicating state §∥¶ in which it becomes phenotypically relatively tolerant to most TB drugs. §∥¶ Both host immunity § and conventional antimycobacterial chemotherapy ‡ might be rendered more effective by inhibition of Mtb’s antioxidant and antinitrosative defenses.

A coordinated attack on Mtb’s defense system requires a deeper understanding of its components and their interactions. The goal is complicated both by the potential redundancy of such defenses and by their diversity. Mtb’s antioxidant and antinitrosative enzymes range from conventional dismutases and peroxidases (catalase and superoxide dismutases (SODs)) to NADH-dependent and α-ketoacid-dependent systems that co-opt enzymes of central carbon metabolism.

Among the least explored classes of Mtb’s enzymatic defenses against exogenous small molecules are the nitroreductases. This class of enzymes is known or predicted to draw reducing power from NADH or NADPH via FAD or FMN.

Received: July 25, 2017
Published: February 21, 2018

DOI: 10.1021/acsinfecdis.7b00111
ACS Infect. Dis. 2018, 4, 771−787

**Supporting Information**
Recently, a candidate nitroreductase was identified in Mtb on the basis that mutations in \textit{rv2466c}, the gene encoding it, conferred resistance to a nitro-containing antimycobacterial compound.\textsuperscript{13} However, the specific reaction chemistry catalyzed by Rv2466c was left undefined in that study or in a subsequent report of the redox state of Rv2466c.\textsuperscript{14} and the impact of deleting this gene on pathogenesis was not tested in either of these reports. Here, we present evidence that Rv2466c is a nitroreductase, which transforms nitro-containing compounds to their corresponding amine and whose cofactor is not a flavin but a low molecular weight “sugar” thiol, mycothiol.

Furthermore, Rv2466c is also important in resisting menadione induced redox stress. On the basis of its sequence, structure, activity profile, cofactor dependency, and phylogeny, Rv2466c appears to be the founding member of a new class of mycothiol-dependent reductases. Conservation of the homologous gene in \textit{Mycobacterium leprae}, which has undergone extensive gene reduction, suggests that enzymes of this class can subsume critical functions in bacterial pathogenesis when other components of the stress resistance machinery are lacking.

\textbf{RESULTS}

Rv2466c Activates Nitrofuranyl Calanolides (NFCs). Natural product-inspired nitrofuranylcalanolides (NFCs) were recently described that have a remarkable constellation of properties: potent bactericidal activity against both replicating and nonreplicating Mtb and the ability to eliminate detectable Mtb from primary human macrophages in vitro, coupled with...
lack of cidal activity against the other Gram-negative and Gram-positive bacteria tested and against human cells in vitro.15 Such profound cidal activity against such a narrow range of organisms invited examination of their mechanism of action. Accordingly, we incubated Mtb with the NFCs termed 18 or 30a15 (Figure 1A) and isolated resistant clones of Mtb to identify relevant mutations.

We obtained NFC-resistant clones with an unusually high frequency of $\sim 1 \times 10^{-5}$, compared to what was published for isoniazid (INH) ($1 \sim 3.5 \times 10^{-8}$) and rifampicin ($1 \sim 10^{-8}$).16 The frequency of resistance (FOR) to NFCs was similar in two strains of Mtb (H37Rv or Erdman) and in three bacteriologic media (7H9-agar, 7H10, and 7H11). The selected clones were as sensitive as the wild-type strain to the TB drugs rifampicin, streptomycin, and ethambutol (Figure S1A). This suggested that enhanced nonspecific drug efflux did not account for resistance to NFCs. Whole genome resequencing of two NFC-resistant Mtb clones identified rv2466c as the only gene mutated in common in these two strains. Nineteen additional resistant clones were selected in Mtb and M. bovis BCG for sequencing of PCR products of rv2466c. Eighteen of these clones had nonsynonymous point or frameshift mutations, all but one in a different codon among the 207 codons in the rv2466c coding region (Figure 1B). The remaining clone had a mutation in the predicted upstream regulatory region of rv2466c. The consistent presence of nonsynonymous coding, frameshift, or presumptive regulatory mutations in rv2466c in 21 of 21 clones strongly suggested that loss of function of rv2466c conferred resistance to NFCs. To confirm the impact of rv2466c, we generated a knockout of rv2466c in Mtb by homologous recombination (Figure S1B). Mtb lacking rv2466c was as resistant to NFCs as the isolated point mutants (Figure S1C).

Conversely, transformation of NFC-resistant Mtb rv2466c W181C (mutant 1) with a wild-type (WT) copy of rv2466c driven by the constitutive hsp60 promoter sensitized the strain to NFCs (Figure 1C). Moreover, overexpression of the WT copy of rv2466c in WT bacteria (confirmed by a Western blot with antibody raised against recombinant Rv2466c (rRv2466c), Figure 1D) rendered the cells 2- to 4-fold more susceptible to...

Figure 2. C19 and C22 are critical residues for the activity of Rv2466c. (A) Δrv2466c was transformed with pMV261 plasmid carrying either a WT copy of rv2466c or rv2466c expressing the following mutations: C19S, C22S, or both C19S and C22S. Log phase bacteria were exposed to varying concentrations of 30a for 7 days and % inhibition was calculated on the basis of growth in vehicle (DMSO) wells. Data are means ± SDs of triplicate wells. (B) Structural superposition of one subunit of Rv2466c C22S mutant (pink, current publication) with that of the published structure of Rv2466c (blue, PDB ID: 4NXI; green, PDB ID: 4ZIL) and the M. leprae homologue (gray, PDB ID: 4WKW) represented in ribbon format. Subunits are superimposed and shown as ribbon. Dashed circles denote regions that undergo significant conformational changes. Side chains of C19 and C22S residues of the C22S mutant are depicted as spheres with yellow for sulfur, red for oxygen, and pink for carbon. (C) Conformational rearrangement of the α2−α3 region to generate a putative substrate pocket around C19. The pocket is shown by a black arrow; side chain of C19 is depicted as spheres. The “open state” α2−α3 region of wild-type Rv2466c is shown in green ribbon form.
NFCs (Figure 1C) but did not impact sensitivity to rifampicin (results not shown). These data, together with our finding that the loss of function mutants had no survival defect in standard growth conditions, are consistent with Rv2466c activating prodrug NFCs to form mycobactericidal product(s).

The rv2466c frameshift mutants (Figure 1B, mutants 2 and 8) were predicted to produce truncated proteins (70 and 6 aa’s, respectively) that we anticipated to be degraded. The point mutants might likewise generate a loss of function phenotype by destabilizing the protein so that it undergoes accelerated degradation or by interfering with its enzymatic function without altering its expression. To distinguish among these possibilities, we used our anti-Rv2466c antibody to assess protein levels in 10 resistant clones that were selected in Mtb under NFC pressure. In log phase Mtb, Rv2466c was expressed at levels that were too low for Western blot detection. Exposure of Mtb to the ROS-generating compound diamide induced transcription of rv2466c in a SigH-dependent manner. We extended this observation to the protein level and confirmed that Rv2466c was highly expressed in WT Mtb after sublethal exposure to diamide (Figure 1E). We evaluated diamide-exposed bacteria to determine the impact of different rv2466c mutations on protein expression. As expected, ∆rv2466c had no induction of Rv2466c upon exposure to diamide (Figure 1E). Diamide induced expression of Rv2466c in mutant 6 (W21R) and mutant 7 (E195Q) (Figure 1F) but not in the remaining 8 point mutants tested. The latter mutations presumably affected the translation or stability of Rv2466c. We concluded that our analysis of Rv2466c’s enzymatic function would have to account for prominent functional rather than structural roles of W21 and E195.

Rv2466c was annotated by bioinformatics analysis to encode a thioredoxin domain, including the canonical 19CXXC22 motif critical for enzyme activity. Mutations in C22 partially, and C19 completely decreased the ability of Rv2466c to transform a nitroso-thienopyrimidine to unidentified products during in vitro enzymatic reactions. 14 To confirm the role of Rv2466c’s cysteines within the bacteria, we overexpressed Rv2466c with C22S or C19S substitutions in ∆rv2466c Mtb and tested the activity of 30a in these strains. Mtb overexpressing Rv2466c C19S was resistant to 30a, presumably due to its inability to activate the compound. However, Mtb overexpressing Rv2466c with C22S remained sensitive to 30a (Figure 2A) although to a slightly lesser degree, with a 2- to 4-fold increase in the minimum inhibitory concentration (MIC) when compared to bacteria overexpressing WT protein. The three strains (WT, C19S, C22S) overexpressed Rv2466c to similar levels (results not shown).

The published crystal structure of Rv2466c indicated that it is a homodimeric protein. We were unable to crystallize WT Rv2466c, but we crystallized the C22S mutant and resolved the structure to a resolution of 2.0 Å. As noted above, this mutation sustained partial enzymatic activity. Compared with published crystal structures of Rv2466c (PDB ID: 4NXI, 4ZIL) and the M. leprae homologue (PDB ID: 4WKF), there were significant conformational changes in one subunit of our structure (Figure 2B) that encompassed region 1 (amino acids 46–66) and region 2 (amino acids 130–154). Region 1 is an α2–α3 helical hairpin proximal to the active site C19. In our crystal structure, region 1 takes on a “closed” conformation and crouches over C19 to form a narrow and deep pocket speculated to bind substrate (Figure 2C), whereas in the published structures, the conformation of this hairpin left the hydrophobic groove exposed. Region 2 is located across from region 1 and lies adjacent to the C terminal end of the active site. By swinging away from the thioredoxin domain, the α8 helix becomes partially disordered and pulls the subsequent cis-Pro54 loop further from the C19 thiol group than what had been previously observed. These observations imply that mutations in C22 can impact catalytic efficiency, rationalizing the partial loss of activity observed during the in vitro enzymatic reaction and within the bacteria.

Rv2466c Has a Role in Oxidative Stress Defense. DsbA is an oxidizing protein of the thioredoxin family that catalyzes disulfide bond formation. On the basis of bioinformatics prediction of a DsbA-like thioredoxin domain in Rv2466c, we expected that this enzyme would be important in the face of oxidative stress. Menadione is a redox cycling agent that generates ROS. Mtb Rv2466c mutant 1 (W181C) was more sensitive to menadione than WT Mtb, and its resistance was restored upon introduction of the wild-type rv2466c allele (Figure 3A), demonstrating that Rv2466c can contribute in a nonredundant fashion to resistance to this form of oxidative stress. Additional stresses tested included hydrogen peroxide and diamide, to which a lack of Rv2466c did not confer a change in susceptibility, likely as a result of redundant defense mechanisms against these agents (Figure S2A). Similarly, bacteria lacking a functional copy of Mtb Rv2466c behaved like WT Mtb in primary human macrophages (Figure S2B) and
during acute and chronic infection of mice (Figure 3B), consistent with the extensive redundancy of Mtb’s antioxidant defenses.18,19

Rv2466c Is a Nitro-Reductase. Because frame-shift mutations and deletion of rv2466c conferred high-level resistance to NFC 30a, we hypothesized that Rv2466c uses 30a as a substrate and converts it to an active form, presumably by exploiting the nitro group based on its significance for compound activity in the published structure–activity relationship studies.15 To test this biochemically, we compared recombinant WT Rv2466c with Rv2466c proteins containing mutations in the dimerization fold (mutant 7), the active site groove (mutant 6), or in a residue in the thioredoxin domain with a predicted role in protein dynamics (mutant 4). Reactions containing any of these forms of Rv2466c and candidate cofactors NAD(H), NADP(H), or ATP did not modify 30a.

Since M. smegmatis is inherently resistant to NFCs and overexpression of WT Mtb Rv2466c makes M. smegmatis NFC-susceptible (results not shown), we inferred that M. smegmatis must make the requisite Rv2466c cofactors. Accordingly, we coincubated WT or mutant Rv2466c with 30a in the presence of 1 mg protein/mL M. smegmatis lysate for 3–18 h. The peak absorption of 30a at A_{390} decreased ≤50% (range, 30–50%) with WT Rv2466c but not with mutant

Figure 4. Rv2466c reduces 30a into its amine. (A) 400 µg/mL WT or mutant rRv2466c was coincubated with 1 mg/mL M. smegmatis lysate in the presence of DTT and 30a for 4.5 h and OD_{390} was measured at indicated time points in relation to the absorbance at t = 0. (B, C) 400 µg/mL WT or mutant rRv2466c was coincubated with 1 mg/mL M. smegmatis lysate in the presence of DTT and 30a, and the reaction mixture was analyzed by LC-MS. (D) LC/MS/MS analysis of synthetic AN1 and isolated reaction product. (E) Synthetic AN1 was tested against replicating and nonreplicating WT Mtb at varying concentrations, and growth was assessed using absorbance at 580 nm.
proteins 4, 6, or 7 (Figure 4A) or in the absence of the recombinant protein. We then used liquid chromatography-coupled high resolution mass spectrometry (LC-MS) to identify specific reaction products formed upon incubation of WT Rv2466c with lysate and 30a. Parent compound 30a (m/z 358.13 [M + H]) decreased in abundance when lysate was coincubated with the WT protein but not with mutant protein (Figure 4B), while new ions appeared: a consistently observed major peak with m/z of 328.15 [M + H] (Figure 4C) and several additional peaks with 2-fold or lower ion counts that were not further characterized.

The newly formed m/z 328.15 [M + H] species was consistent with a reduced form of 30a in which the −NO₂ group had undergone conversion to −NH₂ (amine-NFC, AN1) (Figure 4C). We then synthesized AN1 and confirmed that it had LC-MS characteristics (Figure S3) and an MS/MS profile (Figure 4D) similar to that of the major reaction product of m/z 328.15. Thus, WT but not mutant Rv2466c acted as a nitroreductase.

AN1, the major product formed by Rv2466c when coincubated with 30a, was poorly mycobactericidal against replicating and nonreplicating Mtb when supplied in the

Figure 5. Mycothiol is required for activity of Rv2466c. (A) M. smegmatis protein lysates were parted into two fractions by size cutoff spin columns and coincubated with 400 μg/mL of Rv2466c, 50 μg/mL 30a, and 1 mM DTT. Values are means ± SD of duplicate wells, representing two similar experiments. (B) Bacteria were coincubated with varying doses of 30a or PA-824, and growth was evaluated at the end of 7 days by absorbance at 580 nm. Values are means ± SD of triplicate wells. (C) 400 μg/mL WT was coincubated with 1 mg/mL lysate from WT or ΔmshA M. smegmatis in the presence of DTT and 30a, and the reaction mixture was analyzed by LC-MS. (D) Bacteria were coincubated with varying concentrations of 30a, and growth was evaluated at the end of 7 days by absorbance at 580 nm. Values are means ± SDs of triplicate wells. (E) Rv2466c amount was determined by immunoblot of Mtb lysates from log phase bacteria which were normalized to protein content. Mtb Dihydrolipoamide acyltransferase (DlaT) served as loading control. (F) rRv2466c and 30a were coincubated with synthetic reduced mycothiol (MSH) or mycothione (oxidized MSH) (MSSM), with or without DTT, and fluorescence (Ex/Em 370/470) was measured at the end of 90 min. Results are mean ± SD of duplicate wells. (G) 30a, 0.5 mM MSH, and 1 mM DTT were coincubated with or without rRv2466c and analyzed by LC-MS.
extracellular medium (Figure 4E) but may not have entered the cells efficiently or may not have entered into the appropriate microcompartment.

Rv2466c’s Nitroreductase Activity Is Mycothiol Dependent. Given that Rv2466c required mycobacterial lysate for transformation of 30a, we concluded that it requires cofactor(s). Because 30a’s bactericidal impact is unique to mycobacteria,15 we reasoned that cofactors necessary for its activation may not be present in other bacterial species, such as E. coli. Consistent with this hypothesis, an E. coli protein lysate failed to support Rv2466c-mediated conversion of 30a (Figure SSA). To identify mycobacterial factors required for the reaction, we co-incubated WT rRv2466c and 30a with fractions of M. smegmatis lysate from a size exclusion spin column. Rv2466c’s enzymatic activity required factors from two fractions: <10 and >10 kDa (Figure SA).

Most canonical nitroreductases rely on flavin-based cofactors such as FAD or FMN.20 To evaluate if FAD or FMN can be associated with Rv2466c function, we tested whether Rv2466c copurifies with FAD. Absorbance spectra of rRv2466c did not contain any pattern consistent with FAD absorption. In contrast, the mycobacterial protein lipoadime dehydrogenase (Lpd) was purified in a similar manner, which is a protein known to bind FAD,21 and showed absorbance spectra consistent with the presence of FAD (Figure SSB,C).

Furthermore, addition of FAD or FMN to the reaction did not significantly improve mycothiol-dependent, Rv2466c-mediated transformation of 30a (data not shown). The Mtb deazaflavin-dependent nitroreductase (Ddn) activates the antimycobacterial nitromidazole PA-824 using deazaflavin F430 as a cofactor.22 To test if F430 might be needed for Rv2466c mediated transformation of 30a, we tested the activity of 30a on WT Mtb and on an Mtb strain deficient in F430 synthesis due to a transposon insertion in fbiC.23 As judged by resulting inhibition of Mtb’s replication, the F430-deficient strain failed to activate PA-824 but activated 30a, while the WT strain activated both (Figure SB).

Analysis of the secondary structure of Rv2466c using iTASSER24−28 revealed homology to kappa glutathione transferases such as 2-hydroxychromene-2-carboxylic acid isomerase (HCCA) (PDB ID: 2meA) from Pseudomonas putida.27 We compared the sequence identity and homology of Rv2466c to two HCCA’s, one from Pseudomonas putida and the other from Rhodobacter sphaeroides. The amino acid (AA) sequence identity was low: 13.02% and 14.02% for P. putida and R. sphaeroides, respectively, and homology was 22.79% and 17.29% (Figure S6A). However, when we compared the crystal structure of Rv2466c with the HCCA’s of P. putida and R. sphaeroides (PDB ID: 21ME and 3FZ5, respectively), the alignments matched well with a score of 0.416 and 0.358 and RMSD of 3.146 and 2.961 Å, respectively (Figure S6B−D). In contrast, Rv2466c’s crystal structure did not align well when compared to the structures of Mtb nitroreductases, in addition to having poor identity and homology with these protein sequences (Figure S6E).

Mtb lacks glutathione and instead produces mycothiol, a low molecular weight thiol that consists of a cysteine moiety linked to two sugar moieties, inositol and N-glucosamine.28 The requirement for a small molecule cofactor for activity of Rv2466c, together with the prediction of structural and ligand binding similarity with glutathione (GSH) transferases such as HCCA, led us to predict that mycothiol was the small molecular weight cofactor needed for activation of 30a. In support of this hypothesis, lysates from M. smegmatis deficient in MshA, a glycosyltransferase required for mycothiol synthesis,29 did not support the Rv2466c-dependent biotransformation of 30a (Figure SC). Moreover, 30a failed to kill MshA-deficient Mtb (ΔmshA) but did kill the complemented strain (Mtb ΔmshA::mshA) (Figure SD). Deficiency of mycothiol (MSH) led to an increase in levels of Rv2466c, and this apparently compensatory response was reversed by restoration of the wild-type allele of mshA (Figure SE).

Using the intrinsic fluorescence of AN1, we developed a sensitive, lysate free in vitro assay to test the activity of Rv2466c in the presence of reduced mycothiol (MSH), mycothione (oxidized MSH) (MSSM), or glutathione (GSH). Co-incubation of Rv2466c with 30a led to an increase in fluorescence in the presence of MSH but not in the presence of MSSM (Figure SF) or GSH (Figure S7A). Fluorescence increased further when dithiothreitol (DTT) was included in addition to reduced or oxidized mycothiol but not when DTT was added along with GSH. The major product formed in the purified reaction had the LC-MS retention time and m/z (328.15) of AN1 (Figure S5G). These data confirmed that Rv2466c can function as a mycothiol-dependent nitroreductase.

Using the fluorescence-based assay described above, we showed that the reaction is saturable with time-dependent formation of product (Figure S7B) and determined the kinetic parameters of Rv2466c in this assay. Our enzymatic reaction includes DTT at a fixed concentration. In order to take into account any potential impact this can have on enzyme kinetics, we have referred to $K_M$ calculations as $K_{app}$ ($K_{app}$). 30a had $K_{app}$ of 25 μM, a $k_{cat}$ of 37 min⁻¹, and $k_{cat}/K_{app}$ as 1.5 μM⁻¹. In comparison, mycobacterial nitroreductase Ddn has a $K_M$ of 28.6 μM, $k_{cat}$ of 4.7 min⁻¹, and $k_{cat}/K_{M}$ of 0.164 μM⁻¹ measured for PA-824.30 These values are also similar to the Rv2466c kinetic values measured by Rosada et al. ($K_M$ of 12 μM, $k_{cat}$ of 0.016 s⁻¹).31 To calculate enzyme parameters for mycothiol, we used the linear component of the plot of rate (min⁻¹) vs substrate concentration (mM). The slope of this line was used to approximate $K_{cat}/K_{app}$, which was 163.93 min⁻¹ mM⁻¹ or 0.164 min⁻¹ μM⁻¹.

To confirm that a similar conversion occurs within mycobacteria, we incubated WT, Δrv2466c, and Δrv2466c::rv2466c with 30a and showed that the amount of fluorescence correlated with the amount of Rv2466c present in the cell (Figure 6A). We then examined LC-MS spectra of lysates from bacteria exposed to 30a and confirmed that m/z 328.15 [M + H] was detected (Figure 6B), and its abundance correlated with Rv2466c levels in the cell. In this assay, we detected an additional species with m/z 344.14 [M + H] of similar abundance (Figure S8). This mass had LC-MS characteristics that were consistent with the intermediate 30a−NOHOH (data not shown). These findings confirm that Rv2466c mediated nitroreduction of 30a occurs within Mtb.

In Silico Modeling of Mycothiol’s Interaction with Rv2466c. Using the covariant docking protocol CovDock, we modeled the pose of mycothiol in the active site of Rv2466c32 assuming a disulfide bond formation between the mycothiol cysteine sulfhydryl and the sulfhydryl of Rv2466c C19. This model predicted the following noncovalent ligand stabilizing interactions: (1) the oxygen close to the glucosamine moiety establishes a hydrogen bond with the side chain of W21. (2) The guanidinium nitrogen of R54 and the hydroxyl group of Y61 stabilize the acetylated amino group of the ligand’s cysteine residue. (3) The side-chain oxygen of Q205 maintains a
hydrogen bond interaction with one of the hydroxyls of the glucosamine. (4) One of the nitrogens of the imidazole of H104 stabilizes the inositol group through a hydrogen bond interaction with a hydroxyl group (Figure 7A). Moreover, our model identified several mutations (P20A, W21A, R54A, Y61A/Y61F, H104A, and Q205A) that were predicted to decrease the binding affinity of mycothiol. Indeed, mutants C19S, W21A, and Y61A lost all ability to convert NFCs, whereas other mutations had a lesser impact (Figure 7B). The strong impact of the C19S mutation could plausibly be attributed to loss of binding of mycothiol. The significant impact of the W21A and Y61A mutations suggests that these residues play a role in stabilizing the Michaelis Rv2466c–mycothiol complex. The modest effect of the Y61F mutation could be rationalized as follows: loss of the predicted hydrogen bond interaction of the tyrosine hydroxyl with the acetylated amino group of mycothiol could be compensated with an internal hydrogen bond between the nitrogen of the acetylated amine and the hydroxyl group of the inositol. The pronounced but partial impact of R54A, H104A, and Q205A mutations may reflect the solvent-exposed location of these polar residues, which when mutated leads to a less stable mycothiol interaction. Finally, P20A is located in the conserved catalytic domain of mycothiol, and the substitution of this rigid amino acid with a more flexible one could affect the conformation of the catalytic domain, leading to disruption of mycothiol binding.

**Direct Evidence for Binding of Mycothiol to Rv2466c.** Interaction of mycothiol with Rv2466c was further tested using differential scanning fluorimetry (Figure 7C). The melting temperature of WT Rv2466c decreased in the presence of 2 mM MSH but not in the presence of equivalent concentrations of GSH or DTT. MSSM also caused a significant shift in the melting temperature, pointing to interaction of WT Rv2466c with both MSH and MSSM. The C19S mutant form of Rv2466c had a lower melting temperature than WT protein, suggestive of a conformational change, but that melting temperature was unchanged in the presence of MSH or MSSM, indicating the inability of mycothiol to interact with the mutant protein (Figure 7C).

**Phylogeny Analysis Shows Rv2466c as a Sequentially Divergent Class of Enzyme.** Given the activities we discovered, we sought to identify how Rv2466c relates to other enzymes with comparable function, structure, or sequence. For this, we did a phylogeny analysis of Rv2466c and select proteins. In addition to proteins that come up after a BLAST search of Rv2466c, we evaluated proteins from the following classes: nitroreductases, HCCA, DsbA from Gram negative and Gram positive organisms, oxidoreductases, and thioredoxins. These sequences were enriched using a BLAST search for each category. Phylogeny analysis using PhyML showed that Rv2466c is part of a family of DsbA-like enzymes within the Actinobacteria class that is sequentially divergent from the selected nitroreductases, HCCAs, thioredoxins, and oxidoreductases. Interestingly, mycobacterial DsbA-like protein Rv2466c seems, to an extent, also divergent from *E. coli* and *S. aureus* DsbA (Figure 8A).

Collectively, these genetic, biochemical, structural, computational, biophysical, and phylogenetic studies show that Rv2466c belongs to a novel DsbA-like class of enzymes, reduces substrate in the presence of mycothiol (Figure 8B), can act as a nitroreductase, and plays a largely redundant role in redox homeostasis.

**DISCUSSION**

Within granulomas and macrophages, *Mtb* encounters and resists host defenses such as ROS and RNS. For example, *Mtb* faces nitrosative stress from host immunity and its own respiration of nitrate and possibly from exposure to metabolites produced by the host’s microbiota. Improving our understanding of the mycobacterial nitro-oxidative defense systems may identify points of vulnerability in *Mtb*. Bacterial and eukaryotic nitroreductases reduce polyaromatic nitro groups. Their exact role in *Mtb* pathogenesis is unclear, but presumably, they contribute to the antinitrosative defenses on which *Mtb* relies by detoxifying chemicals encountered in the host. *Mtb* has at least four predicted nitroreductases, including Rv3547 (deazaflavin-dependent nitroreductase) which activates the TB drug PA-824; Rv2032 (acon), which has a role in *Mtb* virulence; Rv3131, which activates TLR2; Rv3127, whose role is unknown. Typical nitroreductases depend on flavin cofactors and reduce nitrogen species in a one- or two-electron transfer mechanism in an oxygen-
dependent or oxygen-independent manner. Using a chemical biology approach, we have identified an Mtb protein without a canonical nitroreductase signature or flavin dependency that relies on mycothiol while serving as a nitroreductase and also affords protection against menadione. The activity of 30a against nonreplicating bacteria in the 4-stress model (hypoxia, low pH, acidified nitrite, and fatty acid as the main carbon source) suggests that Rv2466c can be expressed and is active under these conditions as well. We have not observed hyper-susceptibility of the Δrv2466c strain to diamide or hydrogen peroxide, in contrast to what was published by Rosado et al. We presume the differences in genetic makeup of parental strain or media used for testing underlie this difference. We have not evaluated the specific biochemical mechanism by which Rv2466c handles oxidative stress but presume it is by reducing oxido groups during redox cycling upon menadione exposure. Alternatively, Rv2466c may reduce mixed disulfides induced by menadione exposure. To our knowledge, this is the first description of a mycothiol-dependent nitroreductase, although the enzyme has a previously described role in resistance to oxidative stress.

Figure 7. Mycothiol binds Rv2466c through interactions with the C19 of the active site. (A) Details of the substrate-binding site of the Rv2466c–mycothiol complex after mycothiol was docked into the proposed binding site of Rv2466c. The ligand shown represents the best-scoring pose from the CovDock algorithm. C atoms of mycothiol are labeled green, and Rv2466c is displayed as a ribbon in pink. Only relevant residues discussed in the text are shown as sticks and labeled. The black dashes represent the proposed hydrogen bonds involved in the binding of mycothiol. The disulfide bond formed between mycothiol and C19 is shown in yellow. (B) The impact of point mutations on the function of Rv2466c’s ability to convert 30a to AN1 in the presence of mycothiol and DTT. (C) WT or C19S mutants of rRv2466c were incubated with reduced mycothiol (MSH), oxidized mycothiol (MSSM), glutathione (GSH), or dithiothreitol (DTT); and the effect of this interaction was measured by change in peak melting temperature via differential fluorimetry.
Several antimycobacterial compounds require bioactivation, including PA-824, whose bactericidal effect on Mtb requires the mycobacterial nitroreductase Ddn. The transformation of 30a by Rv2466c in mycobacteria also leads to its activation, but an explanation for the resulting mycobactericidal activity remains unclear. The major in vitro conversion product of 30a, AN1, does not kill Mtb when supplied exogenously. Either AN1 fails to enter the organism, or cidality is due to an intermediate reaction product, perhaps a radical, to one of the minor conversion products or to a combination of conversion products.

Figure 8. Rv2466c is a phylogenetically divergent mycothiol-dependent nitrooxidoreductase. (A) Unrooted phylogenetic tree for the sequences of proteins functionally or structurally related to Rv2466c, including HCCAs (in blue) and NAD(P)H nitroreductases (in green), as well as diverse oxidoreductases including thioredoxins (in brown), short-chain dehydrogenases (in gray), NAD(P)/FAD-dependent oxidoreductases (in purple), ferredoxin reductases (in pink), and other proteins annotated as DsbA in other Actinobacteria (in red) or Gram positive or negative bacteria (in orange). The references of the sequences used to generate the tree are reported in the Experimental Methods section. (B) The model of how Rv2466c alters NFCs. The left side of the reaction emphasizing the role of the mtr/NADPH system, and its kinetics were worked out by Rosado et al. In the majority of our experiments, we utilized DTT as a surrogate reductant.
products. Conversion products of 30a that react covalently with macromolecules in mycobacterial lysate would not be detected in our LC-MS analysis.

The association of Rv2466c with mycothiol physiology was predicted from the coregulation of Rv2466c and mycothiol synthesis genes in the SigH regulon.54,56 Mycothiol may sequester 30a to the active site of Rv2466c by forming a mycothiol-30a adduct.37 Alternatively, MSH may alter the active site of Rv2466c to accommodate 30a. The latter scenario would be consistent with our inability to detect a direct interaction of Rv2466c with 30a.

Most nitroreductases use flavin-based cofactors such as FAD, which has a redox potential of \(\sim -340 \text{ mV}\). In contrast, mycothiol has a redox potential of \(-230 \text{ mV}\) in vitro and \(-275 \text{ to } -280 \text{ inside Mtb}\.8 Unlike the more negative redox potential needed for nitroimidazoles such as metronidazole \((-480 \text{ mV})\), nitrofurans in general have less negative redox potentials.59 For example, nitrofurazone has a redox potential of \(-257 \text{ mV}\), and select substitutions can change this to higher or lower values (ranging from \(-211 \text{ to } -350 \text{ mV}\)), depending on the electron affinity of the substitution,50 where increased conjugation correlated to less negative redox potentials. Although we could not directly measure the redox potential of 30a, we predict that the high conjugation of this molecule will in general lead to less negative redox potentials than unconjugated nitrofurans, suggesting the feasibility of mycothiol’s role in the nitroreduction of 30a. Moreover, thiol mediated reduction of nitroaromatics is illustrated by Andres et al.51 and further supports our findings.

It is intriguing that we found such a high frequency of mutations conferring resistance to 30a, all of them affecting the expression or amino acid composition of Rv2466c. A 10-fold lower frequency of resistance due to \(r-v2466c\) mutations was seen after selection with a nitroso-thienopyrimidine.13 This means that the high frequency of resistance we observed to 30a is not likely to be attributable to pre-existent mutations in \(r-v2466c\) but instead may represent mutations induced by 30a or its transformation products. Whole genome resequencing of 30a-exposed Mtb did not reveal a notable increase in mutations in other genes besides \(r-v2466c\), nor were we able to isolate kanamycin-resistant mutants more frequently in WT Mtb in the presence of 30a than in its absence (results not shown). We speculate that coupled transcription and translation allow a product of 30a activation to mutagenize \(r-v2466c\) selectively. Alternatively, there may be some other form of compound–protein–gene functional interaction leading to localized hypermutation.

Most of the Rv2466c mutations we detected upon exposure to NFGs led to depletion of the enzyme due to lack of translation of Rv2466c or early degradation of the protein. The two exceptions to this were the W21R and E195Q mutants. The translation of Rv2466c or early degradation of the protein. The former may sequester 30a to the active site of Rv2466c by forming a mycothiol–30a adduct. Alternatively, MSH may alter the active site of Rv2466c to accommodate 30a. The latter scenario would be consistent with our inability to detect a direct interaction of Rv2466c with 30a.

The impact of the E195Q mutant would maintain the hydrogen bond interaction of Rv2466c with 30a, leading to a nonfunctional but intact enzyme. The impact of the W21R mutation would result in a grossly intact but nonfunctional enzyme. Our

mycothiol docking studies used the available crystal structure of the reduced form of the protein, which differs conformationally from the oxidized form.14 We cannot definitively rule out that the residues we determined to impact protein function based on mycothiol interaction may have alternate reasons for impacting protein function.

Rv2466c is transcriptionally upregulated in conditions where Mtb faces nitrosative and oxidative stresses, including during hypoxia and within macrophages.38 Aside from resistance to 30a, we detected only one phenotype of \(r-v2466c\) deletion: increased sensitivity to menadione. Nonetheless, the significance of the enzymatic role of \(r-v2466c\) is suggested by the retention of a homologous gene in \(M. leprae, ml1485\), during the extensive reductive evolution of that species53 that led to elimination of many genes considered important in pathogenesis and stress resistance of Mtb, such as catalase-peroxidase and the dosR (dormancy survival regulator) regulon. Furthermore, the four predicted nitroreductases of Mtb do not have homologues in \(M. leprae\), highlighting the significance of preservation of the \(r-v2466c\) homologue \(ml1485\).

While this paper was being submitted, Rosado et al.31 reported Rv2466c as a mycothiol-dependent reductase involved in Mtb’s oxidative stress response, which also plays a role in mixed disulfide reduction, and named this protein Mrx-2 (mycoredoxin). This recently published work supports our findings regarding the importance of mycothiol in activity of Rv2466c and the role of Rv2466c in oxidative stress by alternate oxidizing agents. In addition to showing the role of Rv2466c upon exposure to the redox cycling agent, menadione, our findings further highlight the significance of this enzyme by showing its ability to transform the nitro group of a novel mycobactericidal compound, nitrofuranyl calanolide, to its amine, using mycothiol as a cofactor, thus acting as a nitroreductase. Furthermore, our findings expose the in vivo redundancy of this enzyme using macrophage and mouse infection models.

On the basis of our structural, functional, and phylogenetic analysis, we believe that Rv2466c represents a novel enzyme class that is DsbA-like and mycothiol dependent, acts as a nitroreductase, and plays a role in oxidative stress resistance. Ongoing work aims to highlight the physiological role of Rv2466c and distinguish the redundant genes with the goal to manipulate this machinery for identifying novel and species-specific target pathways.

■ EXPERIMENTAL METHODS

**Bacterial Strains and Growth Conditions.** Mtb H37Rv and \(M. bovis\) BCG were grown in 7H9 medium (Middlebrook 7H9 supplemented with 0.2% glycerol, 0.02% tyloxapol, and 10% OADC (oleic acid, albumin, dextrose, catalase)). For replicating (R) conditions, Mtb was grown at 37 °C with 20% \(O_2\) and 5% \(CO_2\). For nonreplicating (NR) conditions, bacteria were incubated at 37 °C with 1% \(O_2\) and 5% \(CO_2\) in a Sauton’s-based medium ("NR medium") consisting of 0.5 g of \(MgSO_4\), 0.05 g of ferric ammonium citrate, 0.5 g of \(KH_2PO_4\), 0.5% BSA, 0.085% NaCl, 0.02% tyloxapol, 50 μM butyrate, and 0.5 mM NaNO3 at pH 5.0.34 \(M. bovis\) BCG strain was grown in Middlebrook 7H9 supplemented with 0.2% glycerol, 0.02% tyloxapol, and 10% ADN (albumin, dextrose, NaCl). \(ΔmshA\) and their complemented \((ΔmshA::mshA)\) strains were a kind gift from Bill Jacobs and were grown in Middlebrook 7H9 supplemented with 0.2% glycerol, 0.02% tyloxapol, and 10% OADC (albumin, dextrose, NaCl). \(M. smegmatis\) was grown in...
Middlebrook 7H9 supplemented with 0.2% glycerol and 0.02% tyloxapol, with shaking at 37 °C. *E. coli* was grown in Luria Broth (LB), with shaking at 37 °C. Where indicated, plasmids were transformed into *Mtb* by electroporation and grown with selection markers (50 μg/mL kanamycin, 50 μg/mL hygromycin B).

**Antibacterial Activity Evaluation.** For R activity, bacteria in mid log growth were diluted to an OD_{580} of 0.01 and dispensed into 96 well plates containing 200 μL/well. Two μL of 100× compound stocks in DMSO was added to each well. After 7 days of coincubation, the optical density was determined at 580 nm. For NR activity, log phase bacteria were washed twice with PBS/0.02% tyloxapol, resuspended in NR medium, and diluted to an OD_{580} of 0.1. 200 μL/well of this cell suspension was dispensed into 96 well plates, and 2 μL of test agent was added per well. After 7 days of coincubation at 37 °C with 1% O_{2} and 5% CO_{2}, a 1:21 dilution into new 96 well plates containing 200 μL of R media was done, and OD_{580} of this new plate was determined after 10–14 days. For colony-forming unit (CFU) assays, bacteria were initially plated into 96 well plates of test conditions as a single-cell suspension made by centrifuging the bacterial suspension at 800 rpm for 10 min without a break. The supernatant of this spin was diluted to an OD_{580} of 0.01 and exposed to test agents. At the end of incubation, bacteria were enumerated by serially diluting in 7H9-ADN (10-fold dilutions), plating onto Middlebrook 7H11 agar plates containing 0.5% glycerol and 10% OADC, and incubating for 3 weeks at 37 °C.

**Animal Work.** 8-wk-old C57BL/6 female mice were infected with *Mtb* via an aerosol route using an Inhalation Animal Work. 8-wk-old C57BL/6 female mice were infected with *Mtb* via an aerosol route using an Inhalation Exposure System (Model 099C A4212; Glascol). *Mtb* strains were grown in mid log phase in 7H9 medium and washed once with PBS + 0.05% Tween 80, and a single cell suspension was prepared at an OD 0.2 for aerosol infection. At the indicated time points, lungs, spleens, and livers were harvested and homogenized in PBS for determination of colony-forming units. This protocol was approved by the Weill Cornell Medical College Institutional Animal Care and Use Committee and performed in accordance with all local guidelines and regulations.

**Isolation of Mtbb and *M. bovis* BCG Clones Resistant to 30a and Whole Genome Sequencing.** Log phase *Mtb* H37Rv or *M. bovis* BCG were plated on 7H11 agar plates containing 5 μg/mL 30a. After 3–4 weeks of incubation at 37 °C, colonies were counted to determine frequency of resistance. Select colonies were picked for culturing in 7H9 medium and treated with 30a for confirmation of activity. DNA from resistant colonies were extracted as described previously using cetyltrimethylammonium bromide and lysozyme. Whole genome sequencing was performed using an Illumina HiSeq 2500 and paired-end data were collected. Genome sequences were assembled as described previously using a comparative assembly method with custom-developed scripts for mapping reads and building contigs to identify indels. Targeted sequencing (Macrogen Corp. New York, USA) of rv2466c was performed after PCR amplification using the following primers: S′ GTCCCTGCGATTTGCTGGGCGGTGT- ACTC 3′ and S′ GTCCGTGCAGTCACTATCGGCCGC- TAGTCG 3′.

**Plasmid Constructs and Protein Purification.** WT rv2466c was introduced into pET28B (Novagen), and the resulting plasmid was transformed into BL21(DE3) competent cells (Life technologies) with kanamycin selection (50 μg/mL).

Bacteria were grown initially at 37 °C while shaking, and protein expression was induced using 0.75 mM IPTG at 16 °C for 18 h. Cells were pelleted and lysed using a sonicator and His tagged protein was purified with nickel chromatography and dialyzed into a 50 mM Tris, pH 7.5, 50 mM NaCl solution. On some occasions, protein was further purified by FPLC. Protein concentration was measured using a BCA kit (Thermo). Purity and size of the protein was confirmed using SDS-PAGE gel and Coomassie staining (Biorad) (Figure S8). Mutant proteins were constructed by introduction of point mutations into the pET28b plasmid that contains the WT rv2466c using the Quick Change II site directed mutagenesis kit (Agilent). For expression in mycobacteria, rv2466c was cloned into pMV261 (for episomal expression) or pMV306 (genomic expression) plasmids under the control of the hsp60 promoter for constitutive expression and selected by kanamycin.

**Generation of KO and Complemented Strains.** The native copy of rv2466c was inactivated in the *M. tuberculosis* H37Rv strain harboring a plasmid encoding the enzymes required for recombinering, pNI-Rec-ET, and using homologous recombination to replace the rv2466c with a hygromycin resistance cassette as the antibiotic marker. The selected clones were cured of pNI-Rec-ET by passaging of the bacteria. The recombination was verified by specific PCR and Southern blot as shown in Figure S1B and called Δrv2466c. The Δrv2466c strain was complemented with plasmids containing a copy of either WT or mutant rv2466c as detailed above.

**Enzymatic Reaction and LC-MS Analysis of Reaction Products.** *M. smegatis* lystate was prepared from late log phase bacteria which was lysed mechanically using 0.1 mm Zirconia beads. Protein concentration was measured using the BCA kit (Bio-Rad), and lysates were brought to 0.5 mg/mL protein lysate. For enzymatic reactions, 5–1 μg/mL protein lysate was coincubated with 50 μg/mL 30a, 0.5–1 mM DTT, and 400 μg/mL recombinant protein. At the end of the reaction, the OD_{590} was measured and/or the reaction mix was prepared for LC-MS analysis with a mixture of acetonitrile/methanol/water (40:40:20) as described. The entire 14 min of the elution spectrum of LC-MS was scanned to evaluate for conversion products. In some experiments, Profinder software (Agilent) was used to find newly formed masses. Given the variation between different runs and to minimize confusion, m/z 328.15 was reported in the manuscript to represent m/z 328.15 ± 0.03 and 358.13 was reported to represent 358.13 ± 0.03. For enzymatic reactions without lystate, 0.1–5 mM mycothiol was incubated with 200–400 μg/mL rrv2466c in the presence of 0–50 μg/mL 30a, 1 mg/mL BSA, and 0.5–1 mM DTT. Fluorescence Ex/Em 370/470 was measured every 10 min over 3 h or every 1 min over 60 min to evaluate for reaction product formation.

**Characterization of Conversion Products by LC-MS and MS/MS.** All samples were extracted with dichloromethane and then filtered to remove any particulates. Solvent was removed by rotary evaporation, and each sample was diluted with 500 μL of acetonitrile. LC/MS separations were performed using a Waters ACQUITY UPLC system equipped with an ACQUITY PDA (diode array) detector, a Waters Micromass SQD 2000 spectrometer, and a Waters ACQUITY BEH C18 column (1.7 μm, 2.1 × 100 mm column length). Chromatography was performed at a flow rate of 0.3 mL/min using the following method: 10% B to 95% B over 5 min (linear). Mobile phase A consisted of water with 0.1% 

DOI: 10.1021/acsinfecdis.7b00111

ACS Infect. Dis. 2018, 4, 771–787

ACS Infectious Diseases

Article

782
trifluoroacetic acid (TFA); mobile phase B consisted of acetonitrile with 0.1% TFA. Mass detection was performed in ES+ mode at a cone voltage of 30 V. LC/MS/MS measurements were performed on an Agilent 1290 Infinity II LC/MS/MS equipped with an Agilent 1290 Infinity II Flexible Pump, Agilent 1290 Infinity II Multisampler, Agilent 1290 Infinity II Multicolumn Thermostat, Agilent 6495 Triple Quadrupole mass spectrometer, and Agilent ZORBAX Eclipse Plus C18 column (2.1 × 50 mm, 1.8 μm). Chromatography was performed at a flow rate of 0.5 mL/min using the following method: 65% B to 95% B over 1.5 min (linear). Mobile phase A consisted of water with 0.1% formic acid; mobile phase B consisted of acetonitrile with 0.1% formic acid. Mass detection was performed in ES− mode, using 326.1 → 270.0 for the MRM transition and a collision energy of 21.

**Western Blots.** Rabbit anti-Rv2466c polyclonal rabbit antibody was generated at Thermo Scientific Pierce Protein Research Department. Post immune serum was used at 1:10 000 dilution for immunoblot assays. Concentration of bacterial lysates was measured by Bradford protein assay (Bio-Rad). Lysates were diluted with buffer containing 50 mM Tris-HCl, pH 7.5, 50 mM NaCl to similar protein concentrations, boiled for 5 min at 100 °C, and separated on a 10% SDS-PAGE gel. Proteins were then transferred to a nitrocellulose membrane and blocked overnight at 4 °C in Tris-buffered saline (TBS) solution containing 5% milk. Blocked membranes were incubated with 1:10 000 dilution of rabbit anti-Rv2466c or 1:5000 dilution of rabbit anti-DlaT for 1 h at room temperature, followed by four washings with TBS containing 0.1% Tween 20 (TBST). Anti-rabbit HRP was used at a 1:10 000 dilution as the secondary antibody for 1 h at RT followed by five washes with TBST. Membranes were developed using Pierce ECL Western Blotting Substrate.

**Differential Fluorimetry.** The method described previously was followed for differential scanning fluorimetry of Rv2466c and its C19S mutant. Briefly, protein was diluted to a final concentration of 200 μg/mL in buffer containing 50 mM Tris–HCl, pH 7.5, and 50 mM NaCl. Where indicated, protein was coinubicated with 2 mM MSH, MASM, or GSH on ice for 30 min. Each sample was then transfered to a 96-well PCR plate and mixed with a 1:1000 dilution of SYPRO Orange (Sigma) to a final volume of 20 μL per well. The temperature of the plates was gradually raised from 25 to 95 °C while measuring the fluorescence of each well. Data was analyzed using BioRad CFX Manager version 3.1.

**Synthesis of AN1 and Mycothiol.** Synthesis of AN1 is described in detail in the Supporting Information. Mycothiol was synthesized according to the published protocol, and briefly, it included linking of 2-camphanoyl-3,4,5,6-tetra-O-benzyl-β-D-xylo-inositol and 2-deoxy-2-azido-3,4,6-tri-O-benzyl-1-p-toluene-thio-D-glucoside which was then coupled with N-Boc-S-acetyl-l-cysteine.

**Molecular Modeling.** In the system preparation, mycothiol and the Rv2466c crystal structure (PDB ID: 4NXI) were prepared using LigPrep and Protein Preparation Wizard, respectively, in Maestro v10.7 (Schrodinger, Inc.: Maestro v10.7, 2016) and docked with covalent docking protocol CovDock. Binding of mycothiol to the active binding site of Rv2466c was explored in silico using the covalent docking protocol CovDock. This protocol follows a two-step process; first, the prereactive ligand is docked to determine suitable poses that will bring the reactive warhead of mycothiol in close proximity with the reactive Rv2466c residue (C19). In a second step, the covalent bond is formed for the top scored complex structures. The covalently bound mycothiol is then sampled, and the complexes are scored using OPLS3 force field and VSGB2.0 implicit solvent model. A receptor grid generation was performed, using default outer and inner box sizes (30 and 10 Å, respectively) and as a center of the grid the following coordinates, X = 0.744560, Y = 11.286497, and Z = 93.792589. The covalent reaction “disulfide formation” and the “thorough” pose prediction mode were applied, and 5 output poses per ligand reaction site were reported for further analysis.

**Protein Preparation, Crystallization, Data Collection, and Structural Determination.** The full-length gene of Rv2466c was cloned into pET22b vector (Novagen). The C22S mutant was generated by two-step overlap PCR, verified by sequencing, and overexpressed with C-terminal 6× His tag in E. coli BL21 (DE3) strain (Novagen) in the presence of 0.2 mM isopropyl β-D-thiogalactoside at 16 °C. Overnight induced cells were harvested by centrifugation and suspended in buffer containing 20 mM Tris-HCl, pH 7.5, and 200 mM NaCl and then disrupted by an EmulsiFlex-C3 homogenizer (Avestin). The lystate was further cleared by centrifugation, and the supernatant was loaded onto a HisTrap nickel column (GE Healthcare). The bound protein was eluted with a 20–100 mM imidazole gradient, and the resultant peak was pooled. The protein was concentrated and further purified by size-exclusion chromatography on a Superdex 75 10/300 column (GE Healthcare) in elution buffer containing 20 mM Tris-HCl, pH 7.5, 100 mM NaCl. The peak fractions were pooled and concentrated to approximately 12 mg/mL for crystallization.

Crystallization was performed via the sitting drop vapor diffusion method at 18 °C by mixing 1 μL of protein with 1 μL of reservoir solution. Crystals were obtained using the reservoir condition: 25% (w/v) PEG 3350, 0.1 M Bis-Tris, pH 5.6, and 0.1 M ammonium sulfate. Crystals were flash frozen in liquid nitrogen under cryo-protectant condition consisting of the reservoir solution with 10% supplemented glycerol. Diffraction data were collected at beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF) at the wavelength of 0.9792 Å. All data were indexed, integrated, and merged using the HKL2000 software package (www.hkl-xray.com). Detailed data collection statistics are summarized in Table S1. The structure was solved by molecular replacement using the MOLREP program from the CCP4 software suite (www.ccp4.ac.uk) with the oxidized Rv2466c (PDB ID: 4NXI) as a search model. Structural refinement was performed using the PHENIX software suite (www.phenix-online.org), with iterative manual model building using COOT (http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot). Final structural refinement statistics are summarized in Table S1. The structural figure was created using Chimera (www.cgl.ucsf.edu/chimera). Coordinates of Rv2466c C22S mutant have been deposited in the Protein Data Bank with the accession number of 5XUR.

**Phylogenetic Tree Preparation.** Sequences of proteins structurally or functionally related to Rv2466c were first selected. For the functionally related proteins, 5 sequences were selected: two nitroreductases (both from Mycobacterium tuberculosis (WP_031578693.1 and AA46370.1)), one thioredoxin (Mycobacterium tuberculosis (WP_006242285.1)), and two other proteins annotated as DsbA (one from Gram-negative bacteria (Escherichia coli, WP_000725336.1) and one from Gram-positive bacteria (Staphylococcus aureus, WP_000162811.1)). For the structurally related proteins, we selected two HCCA sequences (from Pseudomonas putida 378570.1).
database (nr) similar sequences in the performed using the NCBI server to retrieve the 250 most similar sequences as well as the following 53 sequences: using MUSCLE, and the phylogenetic tree was generated sequences selected through the CD-Hit runs were then aligned using PhyML. The tree was then visualized and edited for clarity using the iTOL server. The final tree includes the sequence of Rv2466c as well as the following 53 sequences:

WP_083065252.1 DsbA [Mycobacterium aroiniense]
WP_068152034.1 DsbA [Mycobacterium sp. E740]
WP_000725336.1 DsbA [Escherichia coli]
WP_000162811.1 DsbA [Staphylococcus aureus]
WP_002464706.1 DsbA [Staphylococcus simiae]
WP_002486768.1 glutaredoxin [Staphylococcus epidermidis]
SKT77536.1 DsbA [Mycobacterium abscessus]
WP_082895166.1 DsbA [Rhodococcus kroppenstedtii]
WP_083773791.1 DsbA [Tsukamurella paurometabola]
WP_052596160.1 DsbA [Luteipulveratus mongolensis]
SEL64728.1 Predicted DsbA [Rhodococcus maanshanensis]
WP_084524139.1 DsbA [Nocardia inohanensis]
WP_070830232.1 DsbA [Staphylococcus sp. HMSC036D05]
COW61688.1 NAD(P)H nitroreductase [Mycobacterium tuberculosis]
WP_065509101.1 NAD(P)H nitroreductase [Mycobacterium ulceris]
WP_064401540.1 NAD(P)H nitroreductase [Mycobacterium sp. IS-1590]
WP_062829046.1 NAD(P)H nitroreductase [Mycobacterium brisbanense]
WP_069405370.1 NAD(P)H nitroreductase [Mycobacterium holsaticum]
WP_085323471.1 NAD(P)H nitroreductase [Mycobacterium sp. TBL 1200985]
WP_020786359.1 NAD(P)H nitroreductase [Mycobacterium sp. 012931]
XP_007625011.1 Glutathione S-transferase kappa 1 [Cricetulus griseus]
XP_006861250.1 Glutathione S-transferase kappa 1 [Cricetulus griseus]
XP_005205928.1 Glutathione S-transferase kappa 1 [Bos taurus]

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.7b00111.

KO generation and antibiotic susceptibilities; stress phenotypes; LC-MS and TLC evaluations; cofactor evaluation and FAD spectra; homology studies; enzymatic assays; intrabacterial measurement of reaction product 30a-NHOH; Coomassie gel for protein purity;
methods for chromatography; data collection statistics for crystal structure; chemical synthesis of compounds (PDF)

AUTHOR INFORMATION

Corresponding Author
*E-mail: ses9022@med.cornell.edu.

ORCID

Ben Gold: 0000-0002-9562-8241
Selin Somersan-Karakaya: 0000-0003-2000-2700

Author Contributions
*A.N., P.J., R.M., and X.Z. are cofirst authors.

Author Contributions
#G.L. and S.S.-K. are cosenior authors.

The authors contributed as follows: conceptualization (S.S.-K.); investigation (S.S.-K., A.N., T.B., J.V., P.J., X.Z., D.W., D.B., B.G., K.B.-H., R.M., H.L., D.A.O.); formal analysis (J.V., A.N., J.C.S., T.I.); resources (G.L., R.M., H.W., J.W., S.S.-K.); writing, original draft (S.S.-K.); writing, review and editing (S.S.-K., B.G., K.B.-H., T.W., C.F.N., K.Y.R., D.A.O.); supervision (C.F.N., K.Y.R., S.S.-K., G.L., H.L., D.W.); funding acquisition (S.S.-K., C.F.N., K.Y.R., G.L., H.L.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Bill Jacobs for providing the ΔmskA and its complemented strains. We thank Steve J. Brickner (SJ Brickner Consulting, LLC) for suggesting the structure of the amino calanolide. We thank Dr. Carolina Adura at the Rockefeller High Throughput Screening Resource Core facility and Dr. Tania Lupoli for guidance with fluorimetric binding assays. We thank Mayako Michino and colleagues at Tri-Institutional Therapeutics Discovery Institute for discussion and analysis of the mycothiol interaction data. We thank the staff members at beamline BL17U of the Shanghai Synchrotron Radiation Facility for their assistance in data collection. We thank Prof. Nat Brot for guidance in enzymatic assay development. We thank Dr. Ruslana Bryk for DlaT antibody. PA-824 was a kind gift from the Global Alliance for TB Drug Development. S.S.-K. is a Nan and Stephen Swid Research Scholar in Medicine. This research was in part supported by the Bill & Melinda Gates Foundation/TB Drug Accelerator Phase II/Hit generation for tuberculosis drug discovery (OPP1024029) (C.F.N., B.G, J.R, T.W, K.B.-H.) and by the Milstein Program in Chemical Biology and Translational Medicine (C.F.N.). S.S.-K. and P.J were supported by the following grants: NIH K08AI108799, Bohmfalk Charitable Trust, and the Potts Memorial Foundation. G.L. and R.M. were funded by Natural Science Foundation of China, No. 81573289. H.L and X.Z. were funded by the National Basic Research Program of China, No. 2015CB910503.

ABBREVIATIONS

Mtbc, Mycobacterium tuberculosis; NFC, nitrofuranylicalanolid; SOD, superoxide dismutase; HCCA, hydroxychromene-2-carboxylic acid isomerase; Mxr, mycoredoxin; desR, dormancy survival regulator; Ddn, deazaflavin-dependent nitroreductase; GSH, glutathione; MSH, mycotothiol; MSSM, mycotothione (oxidized MSH)

REFERENCES

(1) Comas, L., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S. R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., and Gagneux, S. (2013) Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat. Genet. 45 (10), 1176–1182.
(2) WHO. (2016) WHO Global Tuberculosis Report, World Health Organization, Geneva.
(3) Russell, D. G. (2007) Who puts the tubercle in tuberculosis? Nat. Rev. Microbiol. 5 (1), 39–47.
(4) Nathan, C. (2012) Fresh approaches to anti-infective therapies. Sci. Transl. Med. 4 (140), 140es2.
(5) Zhang, Y., Permar, S., and Sun, Z. (2002) Conditions that may affect the results of susceptibility testing of Mycobacterium tuberculosis to pyrazinamide. J. Med. Microbiol. 51 (1), 42–49.
(6) Hu, Y., Coates, A. R., and Mitchison, D. A. (2006) Sterilising action of pyrazinamide in models of dormant and rifampicin-tolerant Mycobacterium tuberculosis. Int. J. Tuberc Lung Dis 10 (3), 317–322.
(7) Gold, B., and Nathan, C. (2017) Targeting Phenotypically Tolerant Mycobacterium tuberculosis, in Tuberculosis and the Tubercle Bacillus, pp 317–360, DOI: 10.1128/microbiolspec.TTB2-0031-2015.0.
(8) Bryk, R., Gold, B., Venugopala, A., Singh, J., Sany, R., Pupek, K., Cao, H., Popescu, C., Gurney, M., Hotha, S., Cherian, J., Rhe, K., Ly, L., Converse, P. J., Ehrt, S., Vandal, O., Jiang, X., Schneider, J., Lin, G., and Nathan, C. (2008) Selective killing of nonreplicating mycobacteria. Cell Host Microbe 3 (3), 137–145.
(9) Nakadukam, M., Nathan, C., and Rhee, K. Y. (2014) Isocitrate lyase mediates broad antibiotic tolerance in Mycobacterium tuberculosis. Nat. Commun. 5, 4306.
(10) Bryk, R., Lima, C. D., Erdjument-Bromage, H., Tempst, P., and Nathan, C. (2002) Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. Science 295 (5557), 1073–1077.
(11) Venugopala, A., Bryk, R., Shi, S., Rhee, K., Rath, P., Schnapper, D., Ehrt, S., and Nathan, C. (2011) Virulence of Mycobacterium tuberculosis depends on lipomide dehydrogenase, a member of three multienzyme complexes. Cell Host Microbe 9 (1), 21–31.
(12) Maksymiuk, C., Balkrishnan, A., Bryk, R., Rhee, K. Y., and Nathan, C. F. (2015) E1 of alpha-ketoglutarate dehydrogenase defends Mycobacterium tuberculosis against glutamate anaplerosis and nitric-oxide stress. Proc. Natl. Acad. Sci. U. S. A. 112 (43), E5843–E5843.
(13) Albesa-Jove, D., Chiarelli, L. R., Makarov, V., Pasca, M. R., Uresti, S., Mori, G., Salina, E., Vocat, A., Comino, M., Mohorko, E., Ryabova, S., Plieffer, B., Lopes Ribeiro, A. L., Rodrigo-Unzueta, A., Tersa, Z., Zanoni, G., Buroni, S., Altman, K. H., Hartkoon, R. C., Glocshuber, R., Cole, S. T., Riccardi, G., and Guerin, M. E. (2014) Rv2466c mediates the activation of TP053 to kill replicating and nonreplicating Mycobacterium tuberculosis. ACS Chem. Biol. 9 (7), 1567–1575.
(14) Albesa-Jove, D., Comino, N., Tersa, M., Mohorko, E., Uresti, S., Dainese, E., Chiarelli, L. R., Pasca, M. R., Manganelli, R., Makarov, V., Riccardi, G., Svergun, D. I., Glockshuber, R., and Guerin, M. E. (2015) The Redox State Regulates the Conformation of Rv2466c to Activate the Antitubercular Prodrug TP053. J. Biol. Chem. 290 (52), 31077–31089.
(15) Zheng, P., Somersan-Karakaya, S., Lu, S., Roberts, J., Pingle, M., Warrier, T., Little, D., Guo, X., Brickner, S. J., Nathan, C. F., Gold, B., and Liu, G. (2014) Synthetic calanolides with bactericidal activity against replicating and nonreplicating Mycobacterium tuberculosis. ACS Chem. Biol. 9 (7), 1567–1575.
(16) Johnson, R., Streicher, E. M., Louw, G. E., Warren, R. M., van Helden, P. D., and Victor, T. C. (2006) Drug resistance in Mycobacterium tuberculosis. Curr. Issues Mol. Biol. 8 (2), 97–111.
(17) Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs, W. R., Jr., and Husson, R. N. (2001) The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in Mycobacterium tuberculosis. J. Bacteriol. 183 (20), 6119–6125.
(18) Jaeger, T., Budde, H., Flohe, L., Menge, U., Singh, M., Trujillo, M., and Radi, R. (2004) Multiple thioredoxin-mediated routes to detoxify hydroperoxides in Mycobacterium tuberculosis. Arch. Biochem. Biophys. 423 (1), 182–191.

(19) Nathan, C., and Shiloh, M. U. (2000) Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. Proc. Natl. Acad. Sci. U. S. A. 97 (16), 8841–8848.

(20) Roldan, M. D., Perez-Reinaedo, E., Castillo, F., and Moreno-Vivian, C. (2008) Reduction of poly(3-hydroxyalkanoic acid) by the bacterial nitroreductases. FEMS Microbiol. Rev. 32 (3), 474–500.

(21) Hopkins, N., and Williams, C. H. Jr (1995) Characterization of lipoamide dehydrogenase from Escherichia coli lacking the redox active disulfide: C44S and C49S.

(22) Singh, R., Manjunatha, U., Bosshoff, H. I., Ha, Y. H., Niyomrattanakit, P., Ledwise, R., Dowd, C. S., Lee, I. Y., Kim, P., Zhang, L., Kang, S., Keller, T. H., Jiricek, J., and Barry, C. E. (2008) PA-842 kills nonreplicating Mycobacterium tuberculosis by intracellular NO release. Science 322 (5906), 1392–1395.

(23) Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N., and Nathan, C. F. (2003) The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. Science 302 (5625), 1963–1966.

(24) Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5 (4), 725–738.

(25) Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12 (1), 7–8.

(26) Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinf. 9, 40.

(27) Thompson, L. C., Ladner, J. E., Codreanu, S. G., Harp, J., Gillilland, G. L., and Armstrong, R. N. (2007) 2-Hydroxychromene-2-carboxylic acid isomerase: a kappa class glutathione transferase from Pseudomonas putida. Biochem. 46 (23), 6710–6722.

(28) Rawat, M., and Av-Gay, Y. (2007) Mycothiol-dependent proteins in actinomycetes. FEMS Microbiol. Rev. 31 (3), 278–292.

(29) Vilcheze, C., Av-Gay, Y., Attarhan, R., Liu, Z., Hazbon, M. H., Colangeli, R., Chen, B., Liu, W., Alland, D., Sacchettini, J. C., and Jacobs, W. R., Jr. (2008) Mycothiol biosynthesis is essential for tuberculosis. Proc. Natl. Acad. Sci. U. S. A. 105 (14), 5246–5246.

(30) Gold, B., Pingle, M., Brickner, S. J., Shah, N., Roberts, J., Rundell, M., Bracken, W. C., Warrier, T., Somerosan, S., Venugopal, A., Darby, C., Jiang, X., Warren, J. D., Fernandez, J., Ouferelli, O., Nuermerberger, E. L., Cunningham-Bussel, A., Rath, P., Chidawayanka, T., Deng, H., Reublut, R., Glickman, J. F., and Nathan, C. F. (2012) Nonsteroidal anti-inflammatory drug sensitizes Mycobacterium tuberculosis to endogenous and exogenous antimicrobials. Proc. Natl. Acad. Sci. U. S. A. 109 (40), 16004–16011.

(31) Hu, Y., and Coates, A. R. (2011) Mycobacterium tuberculosis acg gene is required for growth and virulence in vivo. PLoS One 6 (6), e20958.

(32) Peddreddy, V., Doddam, S. N., Qureshi, I. A., Yerra, P., and Ahmed, N. (2016) A putative nitroreductase from the DosR regulon of Mycobacterium tuberculosis induces pro-inflammatory cytokine expression via TLR2 signaling pathway. Sci. Rep. 6:24535.

(33) Rodrigue, S., Provvedi, R., Jacques, P. E., Gaudreau, L., and Manganelli, R. (2008) The sigma factors of Mycobacterium tuberculosis. FEMS Microbiol. Rev. 30 (6), 926–941.

(34) Paget, M. S., Kang, J. G., Roe, J. H., and Buttner, M. J. (1998) sigmaR, an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in Streptomyces coelicolor. Mol. Microbiol. 29 (4), 1316–1329.

(35) Niesen, F. H., Berglund, H., and Vedadi, M. (2002) Enzymes associated with reductive activation and action of nitrofurans. Annu. Rev. Pharmacol. Toxicol. 42, 493–515.

(36) Neill, P., Fielden, E. M., and Adams, G. E. (1990) Anthranilate 3-ribosyltransferase. E. coli. EMBO J. 9 (22), 6137–6140.

(37) Sisson, G., Goodwin, A., Aundoni, A., Hughes, N. J., Mukhopadhyay, A. K., Berg, D. E., and Hoffman, P. S. (2002) Synthesis and evolution of novel electrophilic nitrofurin arylamides and carbamates as radiosensitizers and bioreductively activated cytotoxins. Bioorg. Med. Chem. 10 (3), 5208–5213.

(38) Andres, T., Eckmann, L., and Smith, D. K. (2013) Voltammetry of nitrobenzene with cysteine and other acids in DMSO. Implications of the stoichiometry of the reaction.
for the biological reactivity of reduced nitroaromatics with thiols. *Electrochim. Acta* 92, 257–268.

(52) Rustad, T. R., Harrell, M. I., Liao, R., and Sherman, D. R. (2008) The enduring hypoxic response of *Mycobacterium tuberculosis*. *PLoS One* 3 (1), e1502.

(53) Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Douthey, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J. R., and Barrell, B. G. (2001) Massive gene decay in the leprosy bacillus. *Nature* 409 (6823), 1007–1011.

(54) Gold, B., Warrier, T., and Nathan, C. (2015) A multi-stress model for high throughput screening against non-replicating *Mycobacterium tuberculosis*. *Methods Mol. Biol.* 1285, 293–315.

(55) Larsen, M. H., Biermann, K., Tandberg, S., Hsu, T., and Jacobs, W. R. (2007) Genetic Manipulation of *Mycobacterium tuberculosis*, in *Current Protocols in Microbiology*, Vol. 6, pp 10A.2.1–10A.2.21, DOI: 10.1002/97804701729259.mc10a02s6.

(56) Ioerger, T. R., Feng, Y., Ganesula, K., Chen, X., Dobos, K. M., Fortune, S., Jacobs, W. R., Jr., Mizrahi, V., Parish, T., Rubin, E., Sassetti, C., and Sacchettini, J. C. (2010) Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories. *J. Bacteriol.* 192 (14), 3645–3653.

(57) Murphy, K. C., Papavinasasundaram, K., and Sassetti, C. M. (2015) Mycobacterial recombineering. *Methods Mol. Biol.* 1285, 177–199.

(58) Eoh, H., and Rhee, K. Y. (2013) Multifunctional essentiality of succinate metabolism in adaptation to hypoxia in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 110 (16), 6554–6559.

(59) Liu, Z., Wu, Y., and Liu, G. (2015) Efficient and convenient total synthesis of mycothiol on a large scale. *J. Chin. Pharm. Sci.* 24 (6), 347–355.

(60) Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* 27 (3), 221–234.

(61) Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R., and Friesner, R. A. (2016) OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* 12 (1), 281–296.

(62) Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., and Friesner, R. A. (2011) The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins: Struct., Funct., Genet.* 79 (10), 2794–2812.

(63) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410.

(64) Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26 (5), 680–682.

(65) Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 (5), 1792–1797.

(66) Letunic, I., and Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44 (W1), W242–245.