The Crystal Structure of Mycobacterium tuberculosis NrdH at 0.87 Å Suggests a Possible Mode of Its Activity

ARTICLE *in* BIOCHEMISTRY · MAY 2013

Impact Factor: 3.02 · DOI: 10.1021/bi400191z · Source: PubMed

2 AUTHORS:

Swastik Phulera
National Centre For Cell Science, Pune
5 PUBLICATIONS **6** CITATIONS
SEE PROFILE

Shekhar C Mande
National Centre For Cell Science, Pune
95 PUBLICATIONS **1,956** CITATIONS
SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

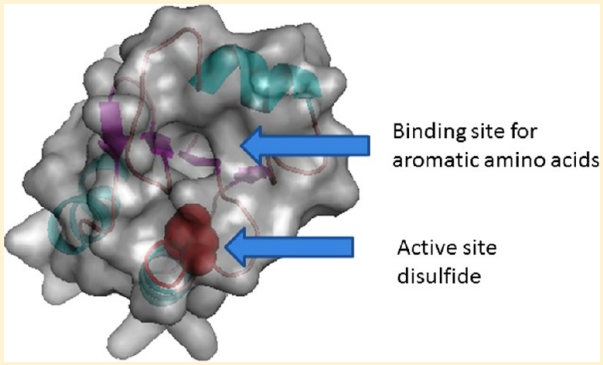
The Crystal Structure of *Mycobacterium tuberculosis* NrdH at 0.87 Å Suggests a Possible Mode of Its Activity

Swastik Phulera and Shekhar C. Mande*

National Centre for Cell Science, NCCS Complex, University of Pune Campus, Ganeshkhind, Pune 411007, Maharashtra, India

Supporting Information

ABSTRACT: Members of the NrdH family of redox proteins, which consists of small glutaredoxin-like proteins with thioredoxin-like activity, serve as the reducing partners of class Ib ribonucleotide reductases. Here, we report the crystal structure of NrdH from *Mycobacterium tuberculosis*, refined to a crystallographic *R* factor of 14.02% (R_{free} = 15.53%) at 0.87 Å resolution. The tertiary structure of *M. tuberculosis* NrdH has a typical thioredoxin fold as expected. The extremely high resolution of the structure allows us to dissect the functionality of the protein in great depth. Structural superimposition of *M. tuberculosis* NrdH and thioredoxin reductase over the *Escherichia coli* thioredoxin reductase−thioredoxin complex suggests the ability of NrdH to accept electrons from *M. tuberculosis* thioredoxin reductase. This raises the important question of why glutaredoxins are unable to accept electrons from thioredoxin reductases and why thioredoxins are unable to reduce ribonucleotide reductases. Furthermore, forms of NrdH from other organisms have been shown to be a specific reductant of class Ib ribonucleotide reductases. We attempt to explain this substrate specificity by modeling the C-terminal peptide of a ribunucleotide subunit, NrdE, in the active site of NrdH using the already available Grx-NrdA-Cter-peptide structure. Statistical coupling analysis of NrdH, glutaredoxins, and thioredoxins reveals different sets of co-evolving contiguous clusters of amino acid residues, which might explain the differences in the biochemical properties of these structurally similar yet functionally distinct subclasses of proteins.



Intracellular redox homeostasis in prokaryotes is maintained by low-molecular weight thiols, such as glutathione, and proteins involved in disulfide exchange, such as thioredoxins.[1] *Mycobacterium tuberculosis*, like most other actinomycetes, possesses low-molecular weight mycothiol instead of glutathione as the redox buffer. The proteins involved in maintaining the redox balance in *M. tuberculosis* include three thioredoxins and three glutaredoxin-like proteins. Prior work from our laboratory had shown that of the three thioredoxins, one might be a pseudogene as evidenced by the absence of any detectable mRNA transcript under many different experimental conditions and its inability to accept electrons from NADPH via thioredoxin reductase.[2] Among the three glutaredoxin-like proteins, a homologue of one has recently been shown to possess mycoredoxin activity in *Mycobacterium smegmatis*.[3] Cytoplasmic targets of some of these proteins are likely to be overlapping, allowing *M. tuberculosis* sufficient redundancy for coping with redox stress. At the same time, some of them might have evolved specifically as reducing partners of specialized enzymes.

The NrdH protein of *M. tuberculosis* encoded by open reading frame Rv3053c (MtNrdH) is a 79-residue protein with a glutaredoxin-like sequence. However, unlike glutaredoxins, MtNrdH can accept electrons from thioredoxin reductase. MtNrdH is thought to supply electrons required for the essential biochemical reaction of ribonucleotide reduction (RR). RR has been suggested to be one of the most fundamental biochemical processes required for the existence of a DNA-based life form.[4,5] In any given cell, ribonucleotide reductases (RNRs) use ribonucleotides to make deoxyribonucleotides that then act as precursors for DNA synthesis. RNRs are thus essential for both DNA replication and repair.

There are three main classes of RNRs (classes I−III), which differ in their subunit composition, cofactor usage, and oxygen requirements. The catalytic reaction mechanism of RNR is conserved among all these classes. Among these three classes, class I enzymes constitute the most widely studied RNRs and are present in a large number of organisms ranging from viruses to prokaryotes to higher eukaryotes. Class I enzymes differ from class II and class III enzymes in several respects. Most importantly, all the class I enzymes require oxygen to function. The class I enzymes reduce ribonucleotide 5′-diphosphates, whereas most of the class II and all of the class III enzymes reduce ribonucleotide 5′-triphosphates (NTPs).[6] On the basis of differences in allosteric regulation and utilization of different electron donors, class I RNRs are further subdivided into classes Ia and Ib. Members of class Ia are present in eukaryotes,

Received: February 14, 2013
Revised: May 14, 2013
Published: May 15, 2013

eubacteria, bacteriophages, and viruses, while members of class Ib are present in eubacteria alone. Thus, despite similarities in their reaction mechanisms, subtle differences exist among the three classes of RNRs.

At the end of each cycle of RR, RNR needs to be reduced to be primed for the next cycle of reduction. An external cofactor, a glutaredoxin/thioredoxin, usually performs this step for the class Ia RNR. The oxidized glutaredoxin/thioredoxin is then reduced by either the glutathione or thioredoxin reductase (TrxR), which can then be reduced by NADPH.[7] In the case of a class Ib RNR, this cofactor is a protein called NrdH with a glutaredoxin-like sequence. Interestingly, NrdH, despite appearing to be Grx-like in its sequence, functionally behaves like a Trx and can accept electrons from TrxR. Recently, it has also been shown that NrdH from *Bacillus anthracis* and *Bacillus cerius* can have up to 10-fold higher activity when transferring electrons to Mn-NrdE than when transferring electrons to Fe-NrdE.[8]

It has been shown that NrdH is not an essential gene under normal conditions in organisms such as *Escherichia coli*[9] and *Staphylococcus aureus*[10] because these organisms have an alternate class of RNRs to maintain the required levels of ribonucleotide reduction. *In vitro* experiments with *E. coli* NrdH have shown that it can *in vitro* reduce *S*-sulfocysteine to L-cysteine and its overproduction leads to an increase in the level of L-cysteine production.[8] The physiological relevance of this in the normal cellular physiology is not clear. Moreover, expression analysis in *E. coli* has indicated that class Ib RNR operon genes consisting of nrdHIEF might be important under oxidative stress conditions.[11] In *M. tuberculosis*, it has been shown that the genes encoding nrdE and nrdF2 are essential for growth *in vitro*, and thereby, ribonucleotide reduction might be an attractive biochemical pathway for antimycobacterial drug discovery.[12] It may therefore be argued that in organisms depending solely on class Ib RNR, nrdH might be an essential gene, and a potential drug target.

In this work, we evaluate the paradoxical behavior of NrdH, i.e., its sequence similarity with Grxs but functional resemblance to the Trxs. We report the crystal structure of MtNrdH at 0.87 Å resolution, analysis of which yields useful insights. Phylogenetic reconstructions show that NrdHs are evolutionarily closer to Grxs than to Trxs. Further, using statistical coupling analysis, we find that several amino acid positions are evolutionarily coupled and that this pattern differs significantly among the three classes of proteins, i.e., NrdH, Trx, and Grx. An intricate network of hydrogen bonding along with evolutionarily coupled residue positions might be the answer to this difference in behavior observed in these proteins.

MATERIALS AND METHODS

Cloning and Expression. Open reading frame Rv3053c in the *M. tuberculosis* genome has been annotated as NrdH. Primers were designed to yield an expression construct with a C-terminal hexahistidine tag. This gene was amplified using polymerase chain reaction (PCR) from *M. tuberculosis* H37Rv genomic DNA. The PCR product was digested with *Nde*I and *Hin*dIII and then cloned in pET21a (Novagen). For overexpression, the cloned construct was transformed into BL21-(DE3) cells (Novagen). Cultures were grown at 37 °C while being shaken, and induction was conducted at an OD_{600} of 0.6 by the addition of IPTG (1 mM). Cells were harvested by centrifugation and stored at −20 °C until they were needed.

Purification. Cells were kept on ice and suspended in lysis buffer [50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, and 5% glycerol (pH 8.0)] followed by disruption by sonication on ice. All subsequent steps were conducted at 4 °C. The lysate was centrifuged at 12000 rpm for 30 min. The resulting supernatant was loaded on a glass column containing Ni-NTA resin (Qiagen). After being extensively washed with buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, and 5% glycerol (pH 8.0), the protein was then eluted using a linear gradient of imidazole up to 500 mM imidazole. Fractions containing the protein were analyzed by Tricine−SDS−PAGE. The purified protein was finally concentrated using a Centricon-10 (Amicon) and stored in a buffer with 10 mM Tris, 50 mM NaCl, and 1% glycerol.

Biochemical Activity Assay. The thioredoxin-like activity assay was performed by using insulin as a substrate in a 96-well plate using a UV−vis microplate spectrophotometer. The assay was performed using two different electron donors (DTT and NADPH) in the presence of TrxR as described previously.[13] Each assay was conducted in triplicate. Reduction of disulfide bonds in insulin leads to the precipitation of insulin chains, which was monitored by an increase in turbidity. We tested the ability of NrdH to catalyze this reaction using DTT as a reductant. The reaction mix consisted of NrdH (2 mM), insulin (130 mM), and DTT (1 mM). *E. coli* Trx (2 μM) was used in place of NrdH as a positive control.

A similar assay was repeated with MtTrxR and NADPH in place of DTT as the electron source. Briefly, *M. tuberculosis* TrxR (MtTrxR) was first incubated with NADPH (100-fold excess). The final reaction mix consisted of insulin (130 mM), NrdH (2 mM), MtTrxR (2 μM), and NADPH (200 μM). The reaction was initiated by addition of TrxR. Measurement of insulin precipitation was done 4 h after the reaction had been initiated. No blank subtractions, normalization, etc., were made to the data to avoid any ambiguity in the analysis.

Crystallization. MtNrdH (7 mg/mL) was crystallized using the sitting drop vapor diffusion method at 4 °C. One microliter of a protein solution was mixed with 1 μL of the reservoir solution. The optimized condition consisted of 0.2 M ammonium nitrate and 22% PEG3350. The dimensions of the crystals, which typically grew in 2−5 days, were approximately 0.4 mm × 0.3 mm × 0.2 mm. Prior to data collection, the crystal was placed in reservoir buffer supplemented with 20% glycerol.

Data Collection and Reduction. X-ray diffraction data were recorded from flash-frozen crystals (on liquid nitrogen) at −178 °C first on a home rotating anode source on a MAR345 image plate detector and later using a MarCCD detector at beamline BM14 of ESRF (Grenoble, France). The data were processed using iMOSFLM.[14]

Structure Solution and Refinement. To determine the structure using molecular replacement, a model sought to represent the monomeric form of Protein Data Bank (PDB) entry 1R7H was made by joining the structure from the two chains after residue 48 (i.e., residue 1−48 from chain A and residue 49 onward from chain B followed by the remaining chain). The structure was first determined using home source data at 1.67 Å with Phaser, AutoMR, AutoBuild, and Phenix.refine as implemented in Phenix[15] with intermittent rounds of model building using COOT.[16] Then, this model was refined against the high-resolution synchrotron data. Initial refinement was conducted at 1.1 Å for 20 cycles. Thereafter, the resolution was increased stepwise in 0.05 Å steps. Hydrogens

were added to the model at 0.9 Å, and the final refinement step included data up to 0.87 Å. A few rounds of refinement were also performed using Refmac[17] as implemented in the CCP4 program suite.[18] The final refined structure was also subjected to full matrix refinement using ShelxL.[38]

Phylogenetic and Statistical Coupling Analysis. Several different sequence alignments for various subfamilies of Grx, Trx, and NrdH were generated using *E. coli* GrxA, *E. coli* GrxB, *E. coli* GrxC, *E. coli* TrxA, *E. coli* TrxB, and different well-characterized sequences of NrdH. These sequences were used as queries against the NR database using an *E* value of 1^{-10}, with an appropriate alignment length as a cutoff. Redundant sequences that were >90% identical were removed using CD-HIT.[19]Statistical coupling analysis was performed on these different subfamilies using Matlab codes and the method described by Ranganathan et al.[20,21]

To perform combined phylogenetic analysis, all these alignments were combined and realigned. Thereafter, the alignment was manually curated to remove positions consisting mainly of gaps. The bootstrap consensus tree inferred from 50 replicates is taken to represent the evolutionary history of the taxa analyzed.[22] Branches with partitions in <50% of bootstrap replicates were collapsed. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with a superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites [five categories (+G, parameter = 2.0011)]. The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 5.8189% of sites). The analysis involved 302 amino acid sequences. All positions with gaps and missing data were eliminated. There were a total of 51 positions in the final data set. Evolutionary analyses were conducted in MEGA5.[23]

RESULTS AND DISCUSSION

Purification of MtNrdH. *E. coli* cell lysates expressing MtNrdH, when purified on Ni-NTA columns, gave a single prominent band on a 10% Tris-Tricine−SDS−PAGE gel. This eluted protein when loaded on a Superdex 75 10/300 GL column pre-equilibrated with 20 mM Tris (pH 8), 50 mM NaCl, and 1 mM DTT and eluted with a single major peak. We further performed dynamic light scattering experiments and found that in the presence of 50 mM NaCl most of the protein existed in a dimeric form. When the same experiments were repeated at higher salt concentrations and we found that at 1 M NaCl, the protein existed in monomeric form (Table 1 of the Supporting Information).

Insulin Precipitation Assay with DTT as the Electron Source. Having successfully purified the protein to near homogeneity, we performed biochemical characterization of the same. With insulin reduction being a well-known test of disulfide reduction capability, we first tested whether NrdH possesses the capability to reduce disulfide bonds in insulin using DTT as the electron source. As expected on the basis of orthology arguments from *E. coli* NrdH,[24] MtNrdH can efficiently reduce disulfide bonds present in insulin (Figure 1A).

Insulin Precipitation Assay with NADPH as the Electron Source. Having shown that MtNrdH is capable of reducing insulin in the presence of DTT, we then sought to test whether it can accept electrons from thioredoxin reductase. The electron transfer chain from NADPH to thioredoxin reductase to NrdH has been well characterized in other organisms but has been paradadoxical as NrdH is evolutionarily closer to the glutaredoxins, which do not possess the ability to participate in this chain. Once again, just like *E. coli* NrdH,[24] MtNrdH can also accept electrons from NADPH via the thioredoxin reductase pathway (Figure 1B).

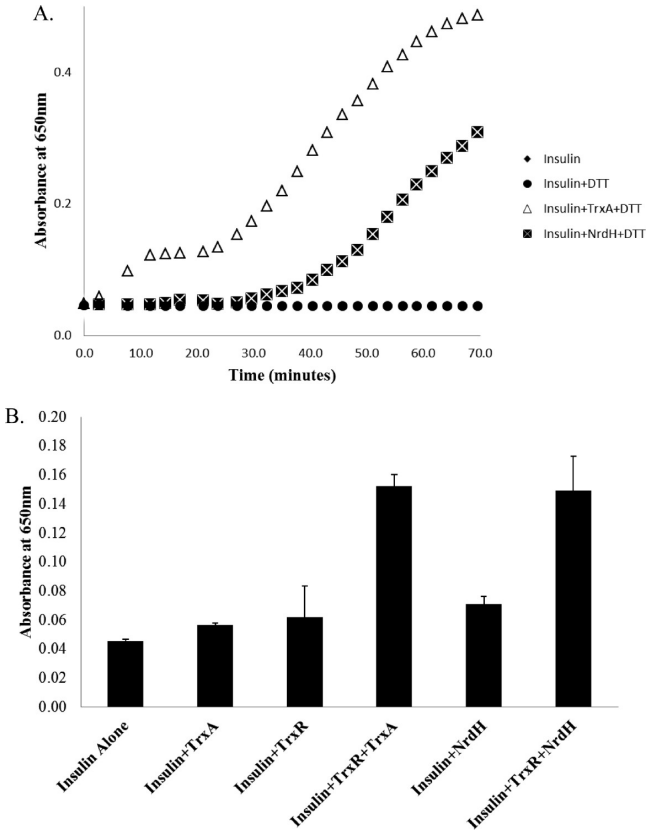


Figure 1. MtNrdH and EcTrxA mediated reduction of insulin disulfides. The assay was performed in 0.1 M KH_2PO_4 (pH 6.5) and 2 mM EDTA with (A) 1 mM DTT as the reductant or (B) NADPH/TrxR as the reductant. The absorbance at 650 nm was measured. The assay performed in the presence of DTT (A) or NADPH/TrxR (B) shows that MtNrdH is as active as EcTrxA. The data for insulin alone and those for insulin with DTT in the graph mask each other, thereby making it difficult to visualize the two distinct data points.

Crystallization, Structure Determination, and Refinement. MtNrdH crystallized in space group $P2_12_12$. Cell parameters along with Matthews parameters suggested a single monomeric protein molecule per asymmetric unit (Table 1). The initial map obtained after molecular replacement allowed auto building of a molecular model with clear density for all residues. The final structure contains, in addition to the single polypeptide chain, 98 water molecules. Two nitrate ions and one chloride ion were also clearly visible in the $2F_o - F_c$ and $F_o - F_c$ density maps and were included in the refinement. Moreover, the $F_o - F_c$ difference maps showed large connected density at the >5σ level, for which water molecules could not account. Because the crystallization buffers contained PEG and glycerol, one glycerol and one PEG molecule were modeled in this density in the later stages of refinement. The *R* factor and R_{free} of the final model are 0.128 and 0.146, respectively (Table 1).

Overall Structure of MtNrdH. The final refined structure consists of residues 2−85 in one chain designated A. We did not observe electron density for the first N-terminal residue.

dx.doi.org/10.1021/bi400191z | *Biochemistry* 2013, 52, 4056−4065

Table 1. Data Collection and Refinement Statistics

Data Collection[a]			
PDB entry	4F2I	4HS1	
collection	rotating anode generator using Cu Kα radiation	BM14 at ESRF	
wavelength (Å)	1.541	0.71255	
space group	$P2_12_12$	$P2_12_12$	
unit cell parameters (Å)	52.0, 48.1, 26.2	51.6, 47.7, 26.1	
resolution limits (Å)	26.23−1.67 (1.76−1.67)	23.85−0.87 (0.92−0.87)	
total no. of observed reflections	56236	492282	
no. of unique reflections	8023 (1065)	53538 (7648)	
$\langle I/\sigma(I)\rangle$	17.73 (3.74)	16.4 (1.9)	
completeness (%)	99.2 (92.3)	99.8 (99.0)	
multiplicity	5.7 (6.7)	9.2 (6.5)	
Wilson B factor (Å^2)	17.57	5.9	
Refinement[b]			
R factor	0.176	0.128	0.140
R_{free}	0.212	0.146	0.155
no. of atoms	1376	1897	1635
no. of macromolecules	647	762	769
no. of ligands	8	22	22
no. of waters	83	98	99
no. of protein residues	83	84	84
rmsd for bonds (Å)	0.010	0.021	0.018
rmsd for angles (deg)	1.16	2.18	2.86
Ramachandran favored (%)	99	99	99
Ramachandran outliers (%)	0	0	0
Clash score	6.98	14.35	16.75
average B factor	16.90	10.9	12.1
macromolecule B factor	15.60	9.7	10.5
solvent B factor	26.30	17.8	22.0

[a]Statistics for the highest-resolution shell are shown in parentheses. [b]The last column of refinement statistics lists those of the structure refined using ShelxL.

This residue is probably disordered or cleaved during the heterologous expression in *E. coli*. As anticipated, MtNrdH has a typical thioredoxin fold, similar to those of the EcNrdH and CaNrdH molecules (Figure 2).[25,26]

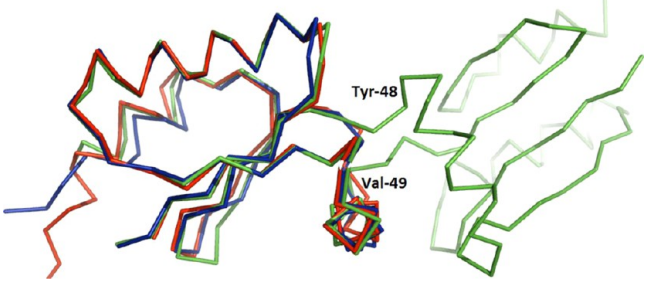


Figure 2. Superimposed structures of EcNrdH (blue, PDB entry 1H75), CaNrdH (green, PDB entry 1R7H), and MtNrdH (red). CaNrdH is a domain-swapped dimer; nonetheless, the three molecules superimpose well on each other.

Several side chains have alternate conformations, as seen in many other high-resolution structures, and have been modeled accordingly.[27] The high resolution of our diffraction data also allowed us to refine occupancies of the side chains that are observed in alternate conformations. The six histidine residues arising from the affinity tag were also seen to be present in the final structure. These residues appear to interact with the neighboring molecule, thereby promoting crystal lattice formation.

One of the interesting observations made on NrdH structures has been on possible domain swapping forming a stable dimer. Although the EcNrdH structure is monomeric, CaNrdH has been reported to be dimeric in crystals.[26] Despite the observed dimeric form of CaNrdH, biochemical evidence of dimerization of NrdH has not been convincing. Dimeric CaNrdH is a domain-swapped dimer, with residues 1−48 forming one part of the tertiary structure and residues 49−74 forming part of the other monomer in the dimer (Figure 2). In our structure, we found no evidence of similar domain swapping. Moreover, our results for the monomeric form are supported by size exclusion chromatography, dynamic light scattering, and finally the monomeric crystallographic structure. Thus, we believe that the domain swapping of CaNrdH might be specific to that species.

Among the many side chains that exhibited alternate conformations, Thr 7 presents an interesting case study (Figure 1 of the Supporting Information). Thr 7 is present in two alternate conformations with occupancies of 0.4 (conformation A) and 0.6 (conformation B). In the A form, the OH group of Thr 7 faces the redox active disulfide of the CVQC motif and there is little space to accommodate the SH group of Cys 14 once it is reduced, while in the B form, it is away from the same. The static disorder of this residue suggests that the reduced sulfhydryl group of Cys 14 will occupy the void created by Thr 7 in its B conformation. We therefore hypothesize that in the oxidized state of NrdH this threonine is free to remain in any of the two conformations while in the reduced state only the B form shall be possible (Figure 3).

We observe a kink in helix α3, which seems to be due to the disordered Lys 70 that exists in two alternative conformations. In one of the forms (A, occupancy of 0.6), the carbonyl oxygen

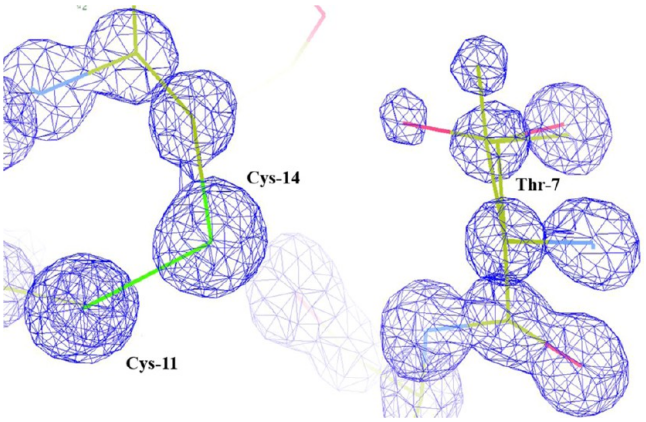


Figure 3. Electron density maps ($2F_o - F_c$) showing MtNrdH's redox active cysteine pair and alternative conformations of Thr 7. The map is contoured at the 2σ level. It is clear that only one conformation of Thr 7 will be possible once the disulfide between Cys 11 and Cys 14 is reduced, and the SH group of Cys 14 will point toward the space created by the B conformation of Thr 7.

dx.doi.org/10.1021/bi400191z | *Biochemistry* 2013, 52, 4056−4065

forms hydrogen bonds with the peptidyl nitrogen of Ala 73, whereas in the second instance (form B), this hydrogen bonding does not occur. In either case, the canonical helical hydrogen bond that should have otherwise existed between Lys 70 and Gly 74 does not exist because of the kink (Figure 2 of the Supporting Information). A kink in the helix therefore appears to introduce strain into the local structure where hydrogen bonding between the peptidyl nitrogen of Ala 73 and the carbonyl oxygen of Lys 70 is satisfied in only one of the two alternate conformations.

Aromatic−Aromatic Interaction and Cation−π Interaction. Tyr 6 and Tyr 42 are involved in aromatic−aromatic interaction. To check whether similar interactions are present in glutaredoxins and thioredoxins, we superinposed *E. coli* TrxA (PDB entry 2TRX) and *E. coli* Grx-1 (PDB entry 1EGO) on the MtNrdH structure. Upon analysis of the superimposed structures, similar aromatic−aromatic interactions are also seen in Grx and Trx. Moreover, in *E. coli* TrxA, a much more intricate network of aromatic aromatic intractions exists compared to those of both *E. coli* Grx and MtNrdH. Similarly, Arg 68, though it exists in two alternate conformations, seems to satisfy the criteria needed to form cation−π interactions with Trp 61 in the two forms. Aromatic−aromatic interactions have been shown to be important for proper protein folding, ligand binding, and stability;[28] similarly, cation−π interactions are also important in the maintenance of protein stability.[29] Although we cannot at present identify the exact relevance of these interactions, their presence and evolutionary conservation might be indicative of structurally or functionally relevant amino acid side chain interaction.

Members of the NrdH family of proteins are characterized by their CVQC and WSGFRP sequence motifs. CXXC is one of the best characterized redox motifs present in a vast array of Trx and Trx-like proteins.[30] The residues between the two cysteines are known to affect the redox potentials and pK_a values and to regulate functions by means of changing their target proteins.[31] It is known that the N-terminal Cys acts as a nucleophilic attacking group while the C-terminal Cys acts as the resolving cysteine. In the CVQC motif of MtNrdH, the amide oxygen of the glutamine is firmly hydrogen bonded with the peptidyl nitrogen of Phe 44, while the amide nitrogen of glutamine is available for further hydrogen bonding. Furthermore, the carbonyl oxygen of Val 12 hydrogen bonds with the peptidyl nitrogen of Ala 16. The local hydrogen bonding network therefore lends stability to the redox active site of MtNrdH.

WSGFRP is the other highly conserved sequence motif in all NrdH sequences, and this structural motif is seen to be stabilized by Gln of the CVQC motif. The Phe of WSGFRP is exposed to solvent; similarly, Val 13 of the CVQC motif is also exposed to solvent. Together, Phe 64 and Val 12 along with Ala 16 and Ala 20 create a distinct hydrophobic patch that is exposed to solvent (Figure 4). This patch is adjacent to the free amide nitrogen of glutamine of the CVQC motif. The unusual nature of the exposed hydrophobic patch is suggestive of the functional significance that might be interaction with TrxR or the C-terminus of RNR.

Arg 68 is a highly conserved residue in all members of the NrdH family of proteins. In our structure, we observe this Arg to be in an alternate conformation. In both conformations, it forms hydrogen bonds with the main chain carbonyl oxygen of His 60 (Figure 5). Interestingly, His 60 precedes the highly conserved sequence WSGFRP motif, suggesting that this

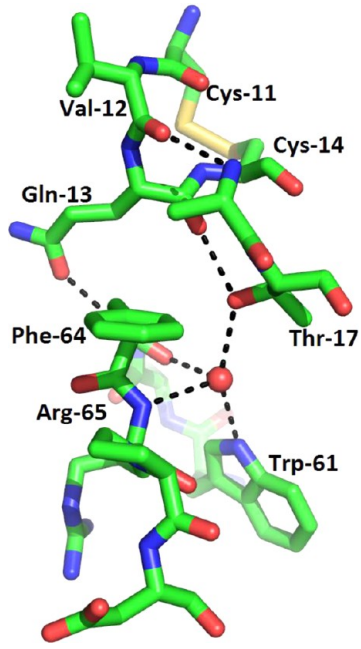


Figure 4. Image showing the two conserved motifs, i.e., CVQC and WSGFRP, along with the conserved water molecule. The network of hydrogen bonds stabilizes the redox active site of MtNrdH and also presents a distinct hydrophobic patch for its interaction with substrates.

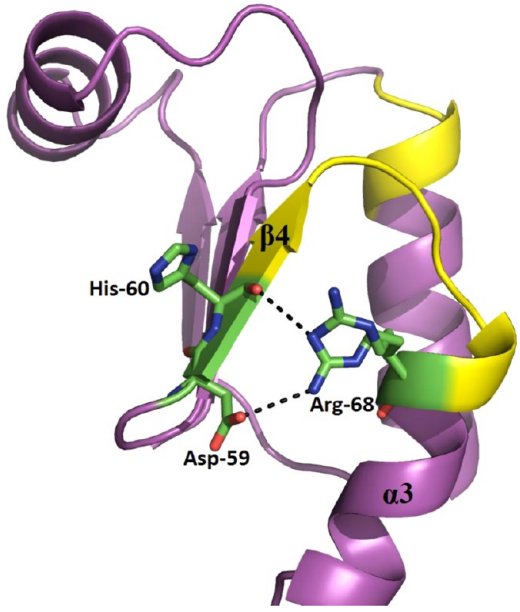


Figure 5. Image showing hydrogen bonding by Arg 68. The yellow regions are the conserved regions, i.e., CVQC and WSGFRP.

interaction between Arg 68 and His 60 might be structurally important. In one of the alternate conformations of Arg 68, its guanidinyl group also forms a salt bridge with Asp 59 (Figure 5). At the identical location, a salt bridge occurs between Arg 65 and Asp 67 in the EcNrdH structure,[25] while in the CaNrdH structure, a hydrogen bond between NH1 of Arg 68 and the ε-oxygen of Glu 59 is reported to be present.[26] Together, these two interactions stabilize the following WSGFRP motif (Figure 4).

dx.doi.org/10.1021/bi400191z | *Biochemistry* 2013, 52, 4056−4065

Table 2. Superimpositions of Representative Members of the Trx, Grx, and NrdH Family Using Superpose[36] from the CCP4 Package[a]

		1EGO[b]	1FOV[c]	31R4[d]	MtNrdH	2TRX[e]	3P2A[f]
glutaredoxins	1EGO	0 (85)	2.6 (68)	2.8 (63)	1.9 (64)	2.3 (70)	2.3 (66)
	1FOV	2.6 (68)	0 (82)	2.7 (69)	1.9 (74)	2.8 (70)	2.3 (60)
	31R4	2.8 (63)	2.7 (69)	0 (218)	1.8 (63)	2.9 (61)	3.6 (76)
NrdH	MtNrdH	1.9 (64)	1.9 (74)	1.8 (63)	0 (84)	2.2 (70)	2.8 (70)
thioredoxins	2TRX	2.3 (70)	3.0 (69)	2.8 (56)	2.3 (70)	0 (216)	1.2 (113)
	3P2A	2.3 (66)	3.1 (65)	– (114)	2.8 (70)	1.2 (113)	0 (512)

[a]The rows indicate the reference structure, while the columns indicate the moving structure. Values indicating the rmsd are given. The values in parentheses are the numbers of equivalent residues. [b]*E. coli* glutaredoxin 1. [c]*E. coli* glutaredoxin 2. [d]*Salmonella typhimurium* glutaredoxin 2 (representative structure from the glutaredoxin 3 family). [e]*E. coli* thioredoxin A. [f]*Yersenia pestis* thioredoxin 2 (representative structure from the Trx B family).

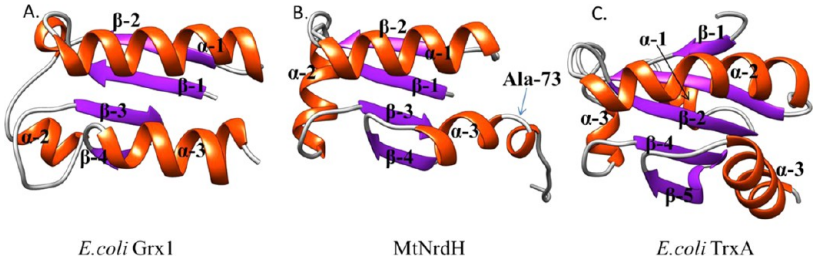


Figure 6. Cartoon representations of Grx (A), NrdH (B), and Trx (C). The coloring is based on secondary structure (helix, orange; sheet, purple; loop, white), and the kink in helix α3 caused by the disordered Lys 70 of MtNrdH is clearly visible.

A buried water that hydrogen bonds with the highly conserved WSGFRP motif is observed with a geometry similar to that observed in the EcNrdH and CaNrdH structures. This water thus seems to be evolutionarily conserved (Figure 4), and therefore, we believe that it is one of the structural signatures of all NrdH proteins. Together with the CVQC and WSGFRP motifs, the extensive hydrogen bonding network also involves this water molecule, suggesting a crucial role for this region in the evolution of NrdH.

Comparison with *E. coli* Thioredoxin and Glutaredoxin. Structural superimpositions were performed to analyze the similarity and dissimilarity of MtNrdH with Trx and Grx (Table 2). Upon specific examination of NrdH with respect to *E. coli* TrxA, the following arguments can be made. (a) *E. coli* TrxA has an additional N-terminal β-strand followed by an α-helix that is absent in NrdH and Grx structures (Figure 6C). Furthermore, the loop regions in Trx are in general longer than those in NrdH and Grx (Figure 6A–C). (b) Helix α3 of *E. coli* Trx, helix α2 of NrdH, and helix α2 of Grx are differently juxtaposed with respect to the central β-sheet in the three structures. This helix is similarly oriented in Trx and NrdH, while in Grx, it seems to be approximately perpendicular to both Trx and NrdH. (c) The loop between strand β4 and helix α3 in MtNrdH is much longer than the structurally equivalent loop between strand β5 and helix α4 in *E. coli* TrxA. Therefore, the α-helix in Trx is positioned at an acute angle with respect to the β-sheet. In the case of *E. coli* Grx-1, this loop region is smaller than NrdH, and this places the helix approximately perpendicular to the helix in Trx. There are a series of sequence changes leading to a change in the tertiary structure, although one cannot comment on the evolutionary direction of changes being from Trx to NrdH to Grx or vice versa. The reason for these particular difference is manifested in the final tertiary structures, i.e., the hydrogen bonding network present in the WSGFRP region and bordering residues, and as discussed earlier the structural connectivity of this region to the CVQC motif.

Modeling Interactions with the C-Terminal Peptide of RNR. When we attempted to model NrdH over the glutaderoxin–class Ib RNR C-terminal peptide complex, it appeared that the mode of peptide binding can be slightly altered in the case of class Ib; nonetheless, it provides a good explanation for the specificity of NrdH for NrdE. On the basis of the modeling, we expect that Glu 143 of NrdE would form a hydrogen bond with the glutamine of the CVQC motif. The WSGFRP motif seems to be important for the formation of interactions with NrdE, with residues 141–143 interacting with the same.

Similarly, modeling *M. tuberculosis* thioredoxin reductase–MtNrdH interactions explains the mode of interaction of MtNrdH with *M. tuberculosis* TrxR (MtTrxR). We attempted to model MtTrxR over the structure of the *E. coli* TrxR–TrxA complex by structural superimposition. Previous results from our laboratory indicate that MtTrxR has two domains, and considerable motion between these occurs during various stages of the redox cycle.[32] Because the already determined structure of MtTrxR represents the oxidized state of the molecule, we first try to model a reduced MtTrxR. This was done by splitting the structure into two domains (domain 1, residues 1–126 and 250–322; domain 2, residues 127–250). These two domains were then superimposed over the *E. coli* TrxR complex structure[33] (PDB entry 1F6M) and later joined together. MtNrdH was superimposed over *E. coli* Trx in the complex structure. Similar modeling studies have been performed previously, i.e., superimposition of EcNrdH[25] and CgNrdH[26] over Trx in complex with EcTrxR. We offer two major inferences from this modeled complex. (a) Two of the differences observed previously, i.e., differences in the orientation of helices α2 and α3 (MtNrdH), would affect the NrdH–TrxR interactions when compared to the Trx–TrxR interaction. (b) In the *E. coli* TrxA–TrxR complex structure,

dx.doi.org/10.1021/bi400191z | *Biochemistry* 2013, 52, 4056–4065

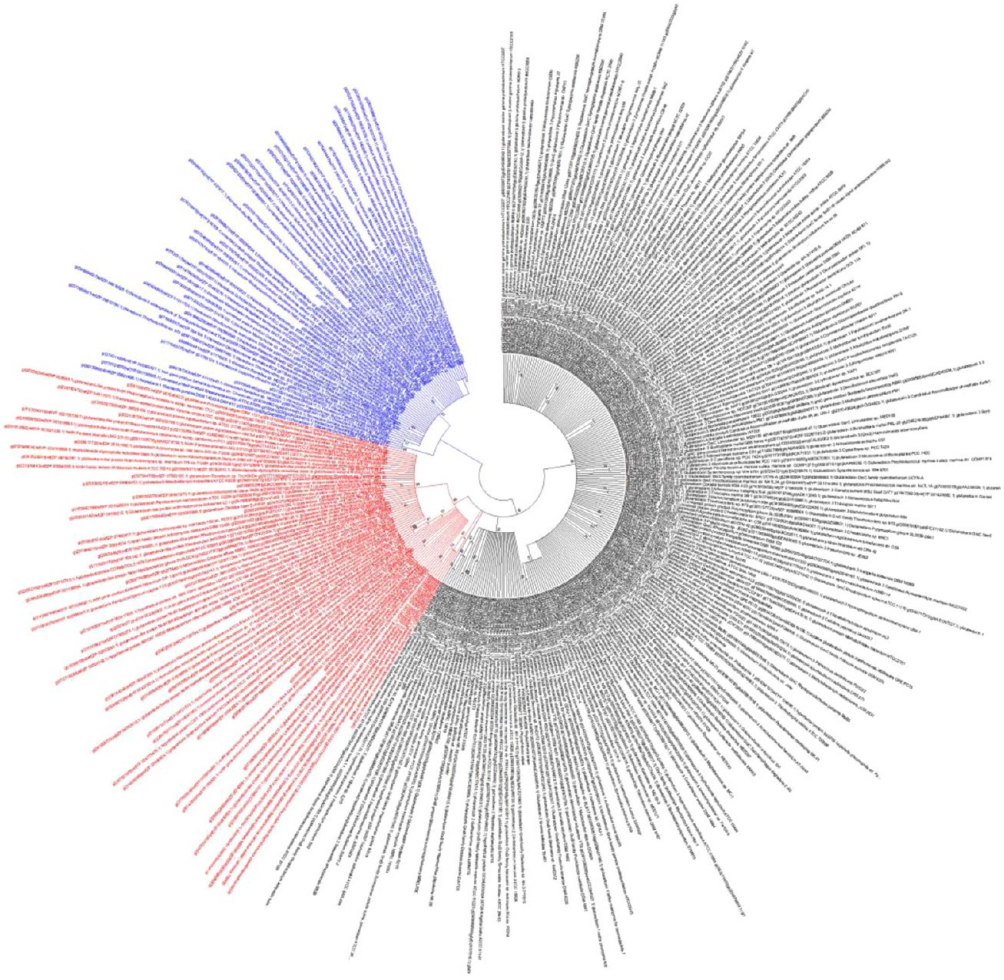


Figure 7. Phylogenetic representation of Trx (blue), Grx (black), and NrdH (red) sequences. The tree shows that the NrdH family clusters with the Grx family and that the Trx family might have diverged from the two earlier in evolution.

Asp 99 interacts with the N-termini of helix α4. Because in MtNrdH this helix is at an angle, we expect Asp 92 of MtTrxR to mimic this interaction. In the structure presented here, we observe that a nitrate ion is present at the N-termini of the corresponding helix. This nitrate ion seems to be stabilized by the ionic interactions of the helix dipole. We believe that the nitrate ions here serve as an excellent mimic of the aspartic acid side chain carboxyl group and that similar interactions are possible when MtTrxR interacts with MtNrdH.

Phylogenetic Analysis. Phylogenetic reconstruction of NrdH, thioredoxins, and glutaredoxins shows that NrdH proteins cluster with glutaredoxins as anticipated (Figure 7). Ribonucleotide reduction is thought to mark the beginning of a DNA world that was preceded by the presence of predominantly RNA-based life forms. We can then hypothesize that RNRs and their electron donors were among the first set of redox proteins to have evolved. Currently, the pattern of RNR evolution is not understood in great detail. Arguably, the first RNRs to evolve were anerobic and resembled either class III RNRs[34] or class II RNRs.[5] Further down the evolutionary trail, one would expect aerobic RNRs to evolve, ultimately giving rise to class I RNRs. The development of overall activity regulation by means of an N-terminal ATP cone possibly evolved much later. This N-terminal ATP cone is absent in members of class Ib, while among other members of class I, it is present in nearly all of them. Thus, among the various classes of modern-day RNRs, class Ia is supposed to be the most evolved. Interestingly, this is also the RNR that is present in human beings. In humans and other higher forms of life, both Grxs and Trxs serve as efficient electron donors of this class of RNRs. We believe that the absence of an ATP cone in class Ib suggests that it is a more primitive class of RNR. By similar arguments, one can hypothesize that the NrdH family is probably an ancestor to both Grxs and Trxs. If such a hypothesis were true, one would believe that NrdH or NrdH-like proteins would have been the first to evolve along with TrxR, and these proteins later on gave rise to paralogous proteins, some of which later evolved to be reduced by small molecules like glutathione, mycothiol, etc., while others diversified with respect to their substrate specificity and later evolved to modern-day Trxs.

Statistical Coupling Analysis. Analysis of coupled residue positions from the NrdH family indicates that the coupled positions are structurally and mechanistically linked. These residues seem to form a physically contiguous cluster of interactions when the interacting pairs of residues are viewed from a structural perspective (Figure 8). Evolutionary covariance among positions in a multiple-sequence alignment might be indicative of either structurally, functionally, or energetically relevant interactions between pairs of amino acids.[20,35]

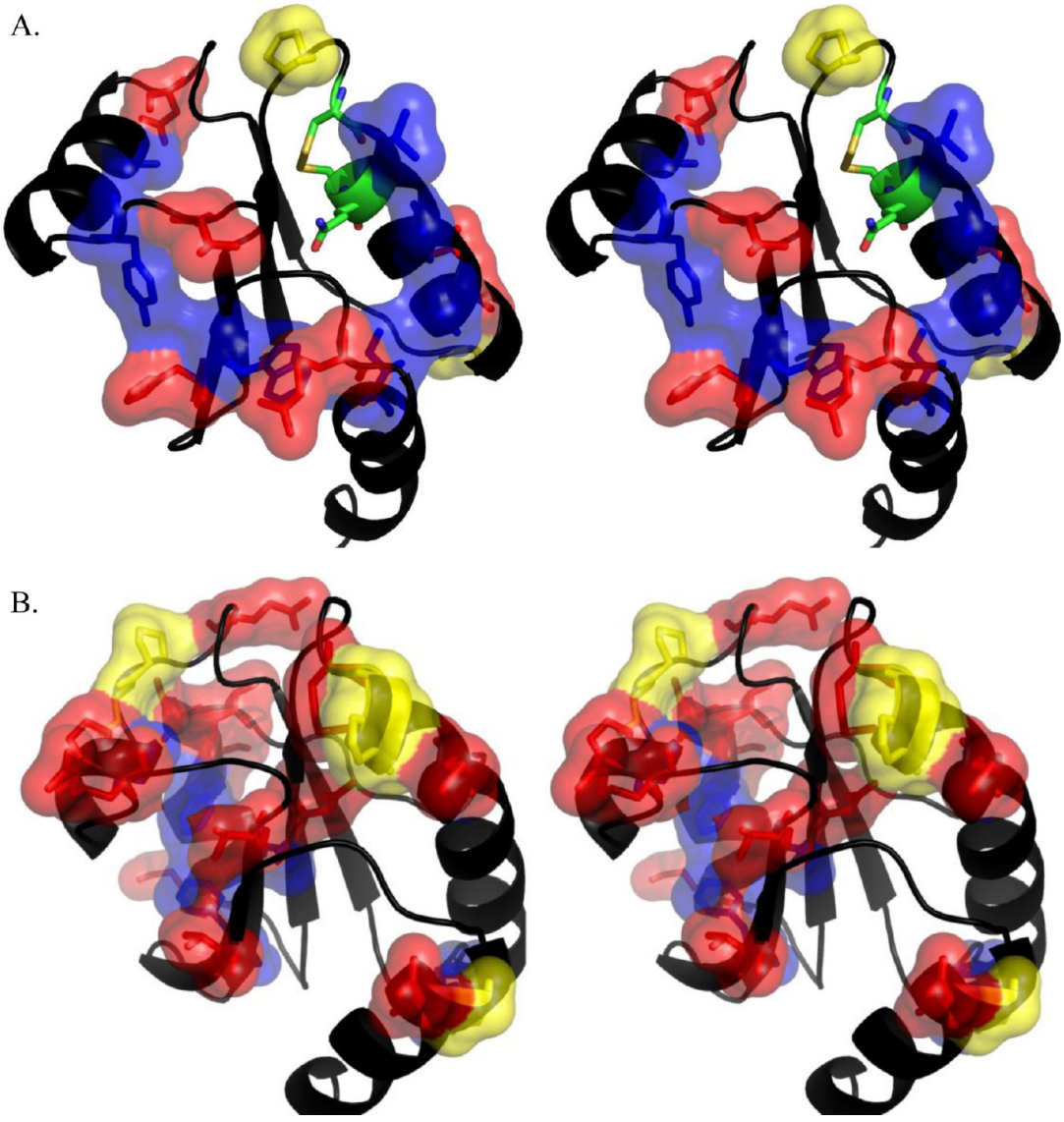


Figure 8. Statistically coupled residues mapped onto the crystal structure of (A) MtNrdH or (B) *E. coli* Trx (PDB entry 2TRX). Statistically coupled residues are shown as sticks with surface display. Among the coupled residues, proline and glycine are colored yellow, hydrophobic amino acids blue, and charged amino acids red.

Among the two conserved motifs, i.e., CVQC and WSGFRP, we observe that Val 12 from the CVQC motif and the Trp 61 and Arg 65 from the WSGFRP motif along with other residues such as Pro 9, which is important for the turn before helix α1 (thus keeping CVQC in its place), Glu 50, which forms a part of the TrxR binding pocket, His 60, which forms a salt bridge with Arg 68 (thus keeping WSGFRP in the required conformation), and Ala 20 and Ile 69, which clearly form a hydrophobic patch between helices α1 and α3, show covariance with each other in a statistically significant manner.

Prediction of NrdH Functional Linkages. Previous results from our laboratory indicate that a combination of genome context methods, namely, phylogenetic profile, gene distance, and operonic co-occurrence, along with high-throughput gene expression studies can be used successfully to predict genome-wide functional linkages.[36] The predicted functional interactions of NrdH are listed in Table 3. As expected, NrdH seems to functionally interact with the other components of ribonucleotide reduction, i.e., NrdE, NrdF, and NrdI. Furthermore, as expected, NrdF (Rv3048c) is featured among the list of genes possibly mediating the dormancy signals between up- and downregulated modules during dormancy. Among the predicted proteins, the functions of all are already annotated, and their functional interactions with NrdH are understandable considering the fact that most of the processes in which they participate would be either regulated or

Table 3. Predicted Protein Functional Linkages of MtNrdH (Rv3053c)

gene	protein	function
Rv0480c	Rv0480c	amidohydrolase
Rv1210	Taga	DNA-3-methyladenine glycosylase I
Rv1523	Rv1523	methyltransferase
Rv2763c	Dfra	dihydrofolate reductase DFRA (DHFR) (tetrahydrofolate dehydrogenase)
Rv3048c	Nrd-F	ribonucleotide-diphosphate reductase subunit β
Rv3051c	Nrd-E	ribonucleotide-diphosphate reductase subunit α
Rv3052c	NrdI	ribonucleotide reductase stimulatory protein
Rv3242c	Rv3242c	hypothetical protein

dx.doi.org/10.1021/bi400191z | *Biochemistry* 2013, 52, 4056−4065

affected by the redox state of the *M. tuberculosis* cytoplasm. The only unannotated gene that appears in the list is Rv3242c. Microarray data available at TbDatase[37] show that this gene is significantly upregulated in the presence of DTNB and thus possibly has an uncharacterised redox-related function. We propose that Rv3242c participates in reactions related to ribunucleotide reductases.

We have thus been able to determine the structure of this important protein at an extremely high resolution. Although the structures determined using in-house diffraction data at 1.67 Å resolution and that collected on a synchrotron at 0.87 Å resolution share most broad features, some of the structural aspects are tantalizingly clear at the ultrahigh resolution. For example, many side chains exhibit alternate conformations in the refined structure at 0.87 Å. We were able to refine occupancies of such side chains at atomic resolution. Such alternate conformations represent static disorder in the crystal lattice but can be visualized only at atomic resolution. One special feature that we were able to capture at 0.87 Å resolution is that related to Thr 7, but could not refine the occupancies of the alternate conformation of side chains at 1.67 Å resolution. The occupancies of the two conformations are 0.6 and 0.4, where one of the conformations would be able to accommodate the reduced Cys side chain and the other would not. Moreover, on the basis of the structure and its models with different electron transfer proteins, we are able to clearly identify residues that might play important roles in mediating interactions among the partners. We are also able to explain the importance of the two conserved sequence motifs, to accept electrons from MtTrxR and pass them to NrdE. Interestingly, although our protein is monomeric in crystals, we find conflicting evidence using dynamic light scattering at low salt concentrations. Whether the dimeric form observed in solution is a domain-swapped dimer remains to be probed further. The *in vivo* status of MtNrdH as a monomer or dimer is therefore not clear.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary figures and table. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*National Centre for Cell Science, NCCS Complex, University of Pune Campus, Ganeshkhind, Pune 411007, Maharashtra, India. Phone: +91-20-257080121. Fax: +91-20-25692259. E-mail: shekhar@nccs.res.in.

Author Contributions

S.P. and S.C.M. designed the study. S.P. performed all the experiments and computational analysis. S.P. and S.C.M. analyzed the results and wrote the paper.

Funding

The work was supported by DBT Centre of Excellence Grant BT/01/COE/07/02 and New Indigo Project BT/In/New-Indigo/18/SM/2010. S.P. is a Senior Research Fellow supported by the Council for Scientific and Industrial Research (New Delhi, India).

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

We thank Abhijit A. Sardesai from the Center for DNA Fingerprinting and Diagnostics (CDFD, Hyderabad, India) for several useful discussions. We also thank DBT for providing support in facilitating data collection at ESRF and CDFD for providing the in-house data collection facility.

■ ABBREVIATIONS

NrdH, NrdH-redoxin; RNR, ribonucleotide reductase; EcNrdH, *E. coli* NrdH; CaNrdH, *Cornybacterium ammoniagenes* NrdH; Trx, thioredoxin; Grx, glutaredoxin; TrxR, thioredoxin reductase; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; rmsd, root-mean-square deviation.

■ REFERENCES

(1) Holmgren, A. (2000) Antioxidant function of thioredoxin and glutaredoxin systems. *Antioxid. Redox Signaling 2* (4), 811−820.

(2) Akif, M., et al. (2008) Functional studies of multiple thioredoxins from *Mycobacterium tuberculosis*. *J. Bacteriol. 190* (21), 7087−7095.

(3) Van Laer, K., et al. (2012) Mycoredoxin-1 is one of the missing links in the oxidative stress defence mechanism of Mycobacteria. *Mol. Microbiol.*, DOI: 10.1111/mmi.12030.

(4) Lundin, D., et al. (2010) Ribonucleotide reduction: Horizontal transfer of a required function spans all three domains. *BMC Evol. Biol. 10*, 383.

(5) Hofer, A., et al. (2012) DNA building blocks: Keeping control of manufacture. *Crit. Rev. Biochem. Mol. Biol. 47* (1), 50−63.

(6) Kolberg, M., et al. (2004) Structure, function, and mechanism of ribonucleotide reductases. *Biochim. Biophys. Acta 1699* (1−2), 1−34.

(7) Nordlund, P., and Reichard, P. (2006) Ribonucleotide reductases. *Annu. Rev. Biochem. 75*, 681−706.

(8) Nakatani, T., et al. (2012) Enhancement of thioredoxin/glutaredoxin-mediated L-cysteine synthesis from S-sulfocysteine increases L-cysteine production in *Escherichia coli*. *Microb. Cell Fact. 11*, 62.

(9) Kim, J., and Copley, S. D. (2007) Why metabolic enzymes are essential or nonessential for growth of *Escherichia coli* K12 on glucose. *Biochemistry 46* (44), 12501−12511.

(10) Rabinovitch, I., et al. (2010) *Staphylococcus aureus* NrdH redoxin is a reductant of the class Ib ribonucleotide reductase. *J. Bacteriol. 192* (19), 4963−4972.

(11) Monje-Casas, F., et al. (2001) Expression analysis of the nrdHIEF operon from *Escherichia coli*. Conditions that trigger the transcript level in vivo. *J. Biol. Chem. 276* (21), 18031−18037.

(12) Mowa, M. B., et al. (2009) Function and regulation of class I ribonucleotide reductase-encoding genes in mycobacteria. *J. Bacteriol. 191* (3), 985−995.

(13) Holmgren, A., and Bjornstedt, M. (1995) Thioredoxin and thioredoxin reductase. *Methods Enzymol. 252*, 199−208.

(14) Battye, T. G., et al. (2011) iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D67* (Part 4), 271−281.

(15) Adams, P. D., et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66* (Part 2), 213−221.

(16) Emsley, P., et al. (2010) Features and development of Coot. *Acta Crystallogr. D66* (Part 4), 486−501.

(17) Murshudov, G. N., et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D67* (Part 4), 355−367.

(18) Winn, M. D., et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D67* (Part 4), 235−242.

(19) Li, W., and Godzik, A. (2006) Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics 22* (13), 1658−1659.

(20) Lockless, S. W., and Ranganathan, R. (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science 286* (5438), 295−299.
(21) Halabi, N., et al. (2009) Protein sectors: Evolutionary units of three-dimensional structure. *Cell 138* (4), 774−786.
(22) Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution 39*, 8.
(23) Tamura, K., Peterson N., P. D., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol. 28*, 2731−2739.
(24) Jordan, A., et al. (1997) Characterization of *Escherichia coli* NrdH. A glutaredoxin-like protein with a thioredoxin-like activity profile. *J. Biol. Chem. 272* (29), 18044−18050.
(25) Stehr, M., et al. (2001) Structural basis for the thioredoxin-like activity profile of the glutaredoxin-like NrdH-redoxin from *Escherichia coli*. *J. Biol. Chem. 276* (38), 35836−35841.
(26) Stehr, M., and Lindqvist, Y. (2004) NrdH-redoxin of *Corynebacterium ammoniagenes* forms a domain-swapped dimer. *Proteins 55* (3), 613−619.
(27) Walsh, M. A., et al. (1998) Refinement of triclinic hen egg-white lysozyme at atomic resolution. *Acta Crystallogr. D54* (Part 4), 522−546.
(28) Lanzarotti, E., et al. (2011) Aromatic-aromatic interactions in proteins: Beyond the dimer. *J. Chem. Inf. Model. 51* (7), 1623−1633.
(29) Gallivan, J. P., and Dougherty, D. A. (1999) Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A. 96* (17), 9459−9464.
(30) Ritz, D., and Beckwith, J. (2001) Roles of thiol-redox pathways in bacteria. *Annu. Rev. Microbiol. 55*, 21−48.
(31) Chivers, P. T., Prehoda, K. E., and Raines, R. T. (1997) The CXXC motif: A rheostat in the active site. *Biochemistry 36* (14), 4061−4066.
(32) Akif, M., et al. (2005) Conformational flexibility of *Mycobacterium tuberculosis* thioredoxin reductase: Crystal structure and normal-mode analysis. *Acta Crystallogr. D61* (Part 12), 1603−1611.
(33) Lennon, B. W., Williams, C. H., Jr., and Ludwig, M. L. (2000) Twists in catalysis: Alternating conformations of *Escherichia coli* thioredoxin reductase. *Science 289* (5482), 1190−1194.
(34) Torrents, E., et al. (2002) Ribonucleotide reductases: Divergent evolution of an ancient enzyme. *J. Mol. Evol. 55* (2), 138−152.
(35) Suel, G. M., et al. (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol. 10* (1), 59−69.
(36) Hegde, S. R., et al. (2012) Understanding communication signals during mycobacterial latency through predicted genome-wide protein interactions and boolean modeling. *PLoS One 7* (3), e33893.
(37) Reddy, T. B., et al. (2009) TB database: An integrated platform for tuberculosis research. *Nucleic Acids Res. 37* (DatabaseIssue), D499−D508.
(38) Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr. A64*, 112−122.