Revised MS submitted to *Nitric Oxide: Biol. Chem.*

Crystal structures of two nitroreductases from hypervirulent *Clostridium difficile* and functionally related interactions with the antibiotic metronidazole

Bing Wang,^{†a,b} Samantha M. Powell,^{†a,b} Neda Hessami,^{a,b} Fares Z. Najar,^{a,b} Leonard M. Thomas,^{a,b} Elizabeth A. Karr,^{a,c} Ann H. West,^{a,b} George B. Richter-Addo^{*a,b}

^a Price Family Foundation Institute of Structural Biology, Stephenson Life Sciences Research Center, University of Oklahoma, Norman, 73019, United States.

^b Department of Chemistry and Biochemistry, University of Oklahoma, Norman 73019, United States.

^c Department of Microbiology and Plant Biology, University of Oklahoma, Norman 73019, United States.

^{*}Email: griechteraddo@ou.edu. Tel: 405-325-6401. FAX: 405-325-6111.

[†] The authors contributed equally to this work.

ABSTRACT: Nitroreductases (NRs) are flavin mononucleotide (FMN)-dependent enzymes that catalyze the biotransformation of organic nitro compounds (RNO_2 ; R = alkyl, aryl) to the nitroso $\text{RN}=\text{O}$, hydroxylamino RNHOH , or amine RNH_2 derivatives. Metronidazole (Mtz) is a nitro-containing antibiotic that is commonly prescribed for lower-gut infections caused by the anaerobic bacterium *Clostridium difficile*. *C. difficile* infections rank number one among hospital acquired infections, and can result in diarrhea, severe colitis, or even death. Although NRs have been implicated in Mtz resistance of *C. difficile*, no NRs have been characterized from the hypervirulent R20291 strain of *C. difficile*. We report the first expression, purification, and three-dimensional X-ray crystal structures of two NRs from the *C. difficile* R20291 strain. The X-ray crystal structures of the two NRs were solved to 2.1 Å resolution. Their homodimeric structures exhibit the classic NR $\alpha+\beta$ fold, with each protomer binding one FMN cofactor near the dimer interface. Functional assays demonstrate that these two NRs metabolize Mtz with associated re-oxidation of the proteins. Importantly, these results represent the first isolation and characterization of NRs from the hypervirulent R20291 strain of relevance to organic RNO_2 (e.g., Mtz) metabolism.

Keywords: nitroorganic, nitroreductase, X-ray structure, UV-vis, FMN cofactor, metronidazole

Abbreviations: CDI, *Clostridium difficile* infection; NR, nitroreductase; Mtz, metronidazole; FMN, flavin mononucleotide; NAD(P)H, nicotinamide adenine dinucleotide (phosphate); PMSF, phenylmethylsulfonyl fluoride; Ni-NTA, Ni-nitrilotriacetic acid; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; β -ME, β -mercaptoethanol; EDTA, ethylenediaminetetraacetic acid.

1. Introduction

Clostridium difficile is an anaerobic bacterium that colonizes the lower intestinal tract in humans and other mammals. *C. difficile* infections (CDIs) caused by pathogenic strains can result in diarrhea, severe colitis or even death [1]. Reported cases of CDIs exceeded 250,000 and 14,000 deaths in 2013 [2], with these numbers rising to ~500,000 with ~29,000 deaths in 2015 [3]. Concerns about rising CDIs have prompted the Centers for Disease Control (CDC) to label *C. difficile* as an urgent threat [2]. One of the reasons for this increase is the frequent use of antibiotics which creates an imbalance in the intestinal microflora, and allows *C. difficile* to thrive and colonize the gut [4].

Metronidazole (Mtz) is a common and relatively low-cost nitro-containing drug prescribed for treatment of CDIs [5, 6]. Mtz typically functions as a prodrug and is metabolized by a class of nitroreductase (NR) enzyme to produce toxic metabolites that kill the pathogen. NRs are a family of FMN-dependent, NAD(P)H-linked proteins widely distributed in bacteria that typically exist as homodimers with protomers varying in size (23-30 kDa) [7]. They function via either a two-electron (Type I) or one-electron (Type II) reduction pathway based on their response to oxygen (Scheme 1).

A rising concern is the emergence of Mtz-resistant *C. difficile* strains. Recent reports show an increased failure of treatments for CDIs with Mtz [5, 8-11]. A mutation resulting in a truncation of a NR was found in both a resistant strain and a sensitive strain of *C. difficile* with reduced susceptibility to Mtz, implicating NRs in Mtz resistance [12]. Mtz resistance has been associated with NRs in *Bacteroides fragilis* [13-15] and *Helicobacter pylori* [16, 17].

Scheme 1. Reduction pathways of nitro compounds by Type I (air-insensitive) and Type II (air-sensitive) NRs [7]. The chemical structure for metronidazole (Mtz) is shown on the right.

Interestingly, although NRs are thought to be the primary enzymes responsible for Mtz metabolism in *C. difficile*, no NRs from the hypervirulent R20291 strain have been characterized. Further, only three NR crystal structures from the less-virulent *C. difficile* 630 strain have been deposited in the Protein Data Bank (www.rcsb.org; with PDB accession codes 3GFA, 3EO8, and 3H4O/3KOQ), but no manuscripts describing the structures have been published. Importantly, in addition to lack of structural information of any NR from the hypervirulent *C. difficile* R20291 strain, no functional assays for the reaction of nitro substrates (e.g., Mtz) with NR from any *C. difficile* strain have been reported.

In this work, we determined the first X-ray crystal structures of two NRs (CDR20291_0684, 26 kDa and CDR20291_0767, 23 kDa) from the hypervirulent *C. difficile* R20291 strain. We also investigated the reaction of Mtz with these two NRs, and the results confirm the functional role of these proteins in Mtz reductive metabolism by NRs in *C. difficile*.

2. Materials and methods

2.1. Cloning and expression of NRs

Two putative NR genes (CDR20291_0684 and CDR20291_0767) were identified from the annotated genome sequence of hypervirulent *C. difficile* and confirmed by a BLAST search [18] against the conserved domain database [19]. The recombinant plasmids pNYCOMPSC-0684 and pSGC-0767 containing the NR genes were constructed by our collaborators at the Albert Einstein College of Medicine (AECOM). The expressed protein NR_0684 contained a C-terminal 10-His tag whereas the second protein NR_0767 contained an N-terminal 6-His tag extension to facilitate their purification.

The plasmids pNYCOMPSC-0684 and pSGC-0767 were transformed into *Escherichia coli* BL21 (DE3) competent cells. The NRs were expressed by autoinduction in ZY media supplemented with kanamycin (50 µg/mL), MgSO₄ (1 mM), glycerol (0.5% w/v), glucose (0.05% w/v), α-lactose (0.2% w/v), 100 µM vitamin B₁₂, and trace metals [20]. Autoinduction was performed at 37 °C for 5 h and then adjusted to 22 °C and left to incubate overnight.

2.2. Purification of the nitroreductases

The cells for NR_0684 were harvested by centrifugation, and the yellow pellets were transferred into a COY anaerobic chamber (3% H₂ in nitrogen). All purification procedures were performed in the anaerobic chamber. The collected cells were first resuspended in binding buffer (0.1 M sodium phosphate, 0.5 M NaCl, 20 mM imidazole, 2 mM β-mercaptoethanol (β-ME)) and appropriate amounts of DNase, RNase, and PMSF, and then sonicated on ice. After centrifugation, the supernatant was loaded onto a Ni-NTA affinity column (MCLAB) equilibrated with binding buffer. The column was then washed with wash buffer (binding buffer

with 50 mM imidazole). The bound His-tagged protein in the column was eluted with the elution buffer (binding buffer with 250-500 mM imidazole). The presence of protein in the collected fractions was verified by SDS-PAGE (Fig. S1A).

Pure protein fractions were pooled and concentrated using an Amicon ultra filter with a cutoff of 10 kDa (Millipore). The concentrated sample was applied to a gel filtration Superdex 200 increase 10/300GL column set up on an AKTA Pure system (GE Healthcare) inside a COY anaerobic chamber. The column was equilibrated with gel filtration buffer (20 mM Tris, pH 7, 2 mM β ME). The elution profile of the NR was recorded by monitoring the absorption at 280 nm using an AKTA (Fig. S1B). The molecular mass of the NR was estimated against the elution profiles of the standard proteins; conalbumin (75 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), and ribonuclease A (13.7 kDa). Fractions containing NR_0684, as judged by SDS-PAGE, were pooled and concentrated to ~5 mg/mL. The protein concentration of NR_0684 was calculated based on the extinction coefficient at 280 nm of $21,430 \text{ L}\cdot\text{cm}^{-1}\cdot\text{mol}^{-1}$ as determined by *ProtParam* [21].

NR_0767 was purified similarly using a binding buffer containing 50 mM Tris, 500 mM NaCl, 5 mM β -ME, 20 mM imidazole, pH 8.0. Following Ni column elution, fractions containing the protein (as determined by SDS-PAGE) were combined (Fig. S1C). A gel filtration buffer containing 50 mM Tris, 150 mM NaCl, 1 mM EDTA, and 5 mM β -ME at pH 8.0, was used. Following gel filtration, fractions containing NR_0767 were pooled and concentrated to ~10 mg/mL (Fig. S1D). The concentration was determined using a Bradford Standard Assay. Purification procedures for NR_0767 were performed in both the anaerobic chamber and in open air for comparison (Fig. S2).

2.3. Redox assay of NR activity

Titration experiments for the reduction and re-oxidation of the FMN cofactor of both NRs were performed under anaerobic conditions as previously described [22]. All buffers were purged with nitrogen gas for at least 30 min prior to being moved into the anaerobic chamber. A sealable quartz cuvette (Starna cells) with a septum was used. After each titration, the cuvette was sealed, and the absorption spectra read using a Hewlett Packard 8453 spectrophotometer.

NR_0684 (~17 μ M NR in 3 mL of 20 mM Tris buffer at pH 7) was reduced by the addition of 11 sequential aliquots of 1 μ L of a stock solution of 21 μ M dithionite in buffer to obtain the spectra shown in Fig. 1A. NR_0767 (~16 μ M NR in 3 mL of 50 mM Tris buffer at pH 8.0) was reduced by the addition of 7 sequential aliquots of 2 μ L of a stock solution of 12.5 mM dithionite in buffer in order to obtain the spectra shown in Fig. 1B.

Re-oxidation of NR_0684 was performed by first reducing the protein (see above) followed by the addition of 12 sequential aliquots of 1 μ L of a stock solution of 1 mM Mtz in buffer to the reduced protein to obtain the spectra shown in Fig. 2A. Re-oxidation of NR_0767 was performed similarly by first reducing the protein (see above) followed by the addition of 8 sequential aliquots of 1 μ L of a stock solution of 20 mM Mtz in buffer to obtain the spectra shown in Fig. 2B.

2.4. Crystallization

Broad screen crystallization trials for both NR_0684 and NR_0767 were carried out with a Mosquito robot (TTP Labtech) in the University of Oklahoma Macromolecular Crystallography Laboratory. The sitting drop vapor diffusion method was used by mixing 300 nL

of protein (4-14 mg/mL) with 300 nL of well solution to form the drop. Each crystallization well contained 60 μ L of well solution.

For NR_0684, tiny rod-shaped yellow crystals averaging approximately 20 μ m in length formed after two days in well C8 (0.2 M ammonium sulfate, 25% PEG 4000, 0.1 M sodium citrate, pH 5.6) of an MCSG1 (Microlytic Inc.) broad screen. This crystallization condition was then optimized manually in μ L sitting drops, with larger sized crystals (about 140 μ m in length) obtained from 10% PEG 1500, 0.1 M sodium citrate, at pH 5.6. The crystals were soaked in a cryosolution made of the crystallization well solution plus 25% ethylene glycol and stored in liquid N₂ for transfer to the Stanford Synchrotron Radiation Lightsource (SSRL).

For NR_0767, two peaks resulted from gel-filtration; one colorless (corresponding to a dimer) and one yellow (corresponding to a protomer) (Fig. S1D). Only the yellow pooled fractions yielded crystals (see later, Section 3.1). These yellow rod-shaped crystals (~400 μ m in length) formed after three days in well B3 (0.2 M di-ammonium hydrogen citrate, 20% PEG 3350) of an MCSG2 (Microlytic Inc.) broad screen. The crystals were briefly dipped in a cryosolution made of the crystallization well solution plus 15% glycerol, then immediately mounted on the goniometer for X-ray data collection.

2.5. Data processing, structure solution and refinement

X-ray diffraction data for NR_0684 were collected at the SSRL BL12-2 beamline equipped with a Pilatus 6M PAD detector. The data were processed using *HKL3000* [23]. The structure factors were calculated using the CCP4 program suite [24]. Initial attempts to solve the three-dimensional structure of the full-length NR by molecular replacement using *PHASER* in CCP4 [24, 25] by employing the NRs from either *Desulfiteobacterium hafniense* (PDB accession

code 3PXV) or *Bacteriodes fragilis* (PDB accession code 3EK3) were unsuccessful. However, we were successful in using an ensemble of the superimposed models for the molecular replacement. Refinements were performed using *Refmac5* [24, 26], and the model was rebuilt using *COOT* [27]. The *MolProbity* server was used for structure validation [28].

The initial 10 cycles of restrained refinement of NR_0684 were run with *Refmac5* [24, 26], and the *R* factor decreased from 0.40 to 0.33. Ligands and water were added to the model according to the $F_o - F_c$ electron density maps in the subsequent refinement cycles. The FMN cofactor coordinate file was imported from the PDB Component Dictionary and the restraint file was generated using *ACEDRG* [24]. Two FMN cofactors, one glycerol, and one phosphate anion were introduced into the model. Water molecules, a total of 138, were sequentially added using *COOT*. Two conformations of the side chains of Asp7 in protomer 1 and Ile154 in protomer 2 were modeled with 50% occupancy for each conformation. Some residues in protomer 1 (the first 5 residues in the N-terminus and 18 residues in the C-terminus) and protomer 2 (6 residues in the N-terminus and 16 residues in the C-terminus) were omitted because of lack of electron density. The final *R* factor and R_{free} are 0.17 and 0.22, respectively.

X-ray diffraction data of NR_0767 were collected at our home source using a Rigaku MicroMax 007HF microfocus X-ray generator equipped with a Dectris Pilatus 200K silicon pixel detector. The data were processed using *HKL3000* [23]. The structure factors were calculated using the CCP4 program suite [24]. The three-dimensional structure of the full-length NR was solved by molecular replacement in *PHASER* (PHENIX) using a NR from *Parabacteroides distasonis* (PDB accession code 3M5K) as a model [29]. All refinements were performed by *phenix.refine*. The model was rebuilt using *COOT* [27]. The *MolProbity* server was used for structure validation [28].

The initial 10 cycles of restrained refinement of NR_0767 were run with *phenix.refine*, and the *R* factor decreased from 0.29 to 0.25. Ligands and water were added to the model according to the $F_o - F_c$ electron density maps in the subsequent refinement cycles. Two FMN cofactors and two imidazoles (based on the electron density map, and the fact that imidazole was in the Ni column elution buffer) were introduced into the model using *phenix.ligandfit*. A total of 311 water molecules were sequentially added using *COOT*. The first 7 N-terminal residues of both protomer 1 and 2 were omitted because of the lack of electron density. The final *R* factor and R_{free} are 0.18 and 0.23, respectively. The atomic coordinates and structure factors have been deposited in the PDB for both NR_0684 (PDB accession code 5J62) and NR_0767 (PDB accession code 5J6C).

3. Results and discussion

3.1. Protein characterization

We report the first expression and purification of NRs from the hypervirulent *C. difficile* R20291 strain. Both NRs were expressed with high yield (~50 mg/L) in *E. coli*. The purified NR_0684 showed a deep yellow color consistent with the presence of oxidized FMN in the protein. The protein was obtained with high purity (>98%) as determined by SDS-PAGE (Fig. S1A). Information from gel filtration indicates that NR_0684 was isolated as a dimer (Fig. S1B); this corresponds well with previously reported bacterial NRs [30-34].

The expression and purification of NR_0767 were performed similarly to NR_0684, except that the gel filtration profile displayed two peaks (Fig. S1D). The first peak corresponds to the dimer, but the fractions were colorless due to the absence of the FMN cofactor. The second peak was yellow and corresponds to the monomer. The two sets of fractions were pooled separately for crystallization (see earlier, Section 2.4), but only the yellow pooled monomer

fractions yielded crystals. Interestingly, the monomeric form that was present in solution resulted in a dimeric form after crystallization.

We note that both NRs were initially purified anaerobically due to the anaerobic nature of *C. difficile*, but we determined that these proteins can also be purified aerobically (e.g., Fig. S2). Both proteins were crystallized aerobically. Their aerobic purification and crystallization allowed us to initially hypothesize that these NRs function as Type I (air insensitive) NRs.

3.2. Reduction of metronidazole by NRs

The two NRs were reduced anaerobically using dithionite, with a corresponding color change from yellow (oxidized FMN) to colorless (reduced FMN) (Fig. 1). Oxidized NR has a typical absorption spectrum of a flavoprotein with absorbance peaks at 370 nm and 450 nm [35]. Upon dithionite addition, these peaks decreased in intensity (Fig. 1). Reduced NRs are susceptible to reaction with oxygen, and even small amounts of air leaking into the cuvette will cause changes in the spectra due to their re-oxidation.

Fig. 1. Reduction of NR_0684 (A) and NR_0767 (B), showing absorbance changes of the oxidized FMN cofactor upon addition of dithionite. NR_0684 conditions: 20 mM Tris buffer, pH 7, [protein] = 16-17 μ M, each addition is 1 μ L of 21 μ M dithionite. NR_0767 conditions: 50 mM Tris buffer, pH 8, [protein] = 16-17 μ M, each addition is 2 μ L of 12.5 mM dithionite.

The reduced NRs (in the absence of excess dithionite) were anaerobically re-oxidized by the addition of the nitro-containing Mtz (Fig. 2) or nitrofurazone (data not shown), with a corresponding color change from colorless (reduced FMN) back to yellow (oxidized FMN) (Fig.

2). The absorbance at 450 nm increased as the reduced protein (lowest dashed line) became oxidized (top thick line). Figure 2 represents the first UV-vis spectral demonstration of the reduction of the antibiotic Mtz by NRs from any *C. difficile* strain.

The UV-vis spectral results of reduction and re-oxidation of the two NRs (Figs. 1 and 2) display no absorbance at 600 nm indicative of a semiquinone state generated by one-electron reduction of the FMN cofactor [35]. This further supports our hypothesis that both of these NRs function as Type I NRs and follow a two-electron reduction pathway with Mtz.

Fig. 2. Re-oxidation of NR_0684 (A) and NR_0767 (B) by Mtz, showing the increase in intensity of the peak due to oxidized NR. NR_0684 conditions: 20 mM Tris buffer, pH 7, [protein] = 16-17 μ M, each addition is 1 μ L of 1 mM of Mtz. NR_0767 conditions: 50 mM Tris buffer, pH 8, [protein] = 16-17 μ M, each addition is 1 μ L of 20 mM Mtz.

We are currently unsure of the identity of the primary organic product(s) resulting from the metronidazole reduction by these two *Cd* NRs. Based on literature precedent [36], we speculate that the metronidazole is reduced to an initial nitroso derivative that readily converts under our experimental conditions to the bioactive hydroxylamine product. We are pursuing the isolation and identification of the primary and secondary products.

Having purified the two proteins from the hypervirulent *C. difficile* strain and having established their NR activity, we then proceeded to crystallize and determine their three-dimensional structures by X-ray crystallography.

3.3. Overall structures of NRs from hypervirulent *C. difficile* R20291

The structures of both NR_0684 and NR_0767 were solved to 2.1 Å resolution (Fig. 3), and represent the first NR crystal structures determined from the hypervirulent *C. difficile* R20291 strain. Data collection and refinement statistics are shown in Table 1. Including the C-terminal 10-His tag and 8 extra linker residues from the plasmid, the cloned NR_0684 has 231 amino acid residues per protomer. Based on the electron density map, 208 residues were modeled into protomer 1 and 209 residues were modeled into protomer 2. Only one Ramachandran plot outlier (D158) was present in the final structure, located at the top of the

Fig. 3. (A) Overall structure of NR_0684. (B) overall structure of NR_0767; in these structures, protomer 1 is drawn in purple, protomer 2 is drawn in cyan, and FMN in orange. (C and D) FMN binding sites of NR_0684, including a phosphate in protomer 2. (E and F) FMN binding sites of NR_0767, including the imidazoles in red. The underlined residues represent those from the partner protomers in these homodimeric structures.

flexible loop between the $\alpha 6$ and $\alpha 7$ helices (Figs. 3A and 6). Similarly, the cloned NR_0767 has a total of 201 amino acid residues per protomer which includes the N-terminal 6-His tag and 16 extra linker residues from the plasmid. Based on the electron density map, 194 residues were modeled into each protomer.

The crystal structures revealed that the two NRs are homodimers (Fig. 3A-B) with each protomer binding one equivalent of FMN. Both NRs adopt a classic NR $\alpha + \beta$ fold [22, 37] within each protomer, with the FMN cofactors located near the interface of the protomers. Each

Table 1. Data collection and refinement statistics

Protein No.	NR_0684	NR_0767
PDB ID	5J62	5J6C
Data collection ^a	SSRL, beamline 12-2	MicroMax 007HF
Space group	$P4_12_12$	$P2_12_12$
Wavelength (Å)	0.98	1.54
Cell dimensions (Å)	100.31, 100.31, 99.92	67.91, 70.73, 87.88
Resolution (Å)	50.00-2.15	50.00 – 2.10
$I/\sigma[I]$	40.67 (2.67)	12.64 (2.18)
No. of reflections		
Observed	678045	86479
Unique	28386 (1388)	25340 (1260)
Multiplicity	23.9 (11.9)	3.4 (3.1)
Completeness (%)	100.0 (99.9)	99.3 (99.4)
R_{merge} ^b	0.076 (0.742)	0.109 (0.454)
CC _{1/2}	0.816	0.814
Refinement statistics		
No. of protein atoms	3268	2898
R factor ^c	0.17	0.18
R_{free} ^d	0.22	0.23
RMSD Bond length (Å)	0.017	0.007
RMSD Bond angles (°)	1.767	0.906
Overall Mean B Factor (Å ²)	45.29	25.42
Ramachandran plot (%) ^e		
most favored residues	99.3	98.3
outliers	0.24	0.00

(a) Values in parentheses correspond to the highest resolution shells. (b) $R_{merge} = \sum |I - \langle I \rangle| / \sum I$, where I is the individual intensity observation and $\langle I \rangle$ is the mean of all measurements of I . (c) $R = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structural factors, respectively. (d) R_{free} was calculated by using 5% of the randomly selected diffraction data which were excluded from the refinement. (e) As calculated with use of *MOLPROBITY*.

protomer consists of a β -sheet in the central core surrounded by α -helices. The five-strand β sheet ($\beta 5\beta 1\beta 2\beta 4\beta 3$) arrangement is formed from four antiparallel β strands ($\beta 1\beta 2\beta 4\beta 3$) complemented by a fifth β strand ($\beta 5$) from the partner protomer; specifics of the 2° structural features are discussed later.

We used a blastp search to identify homologs of the two NRs [38]. The resulting potential homologs all showed <30% sequence identity when compared to NRs from the hypervirulent *C. difficile* strain R20291. Similar results were also obtained from 3D structural comparisons employing the Dali server [39]. The potential homologs are annotated as either NRs or NAD(P)H:FMN oxidoreductases. Most of the homologs have a very similar overall protein fold (avg. RMSD 1.5-2.8 Å and Z score 21-15; calculated using the Dali server) as our NRs even with low sequence identities.

Using a blastp search, NR_0767 and NR_0684 have 24% sequence identity over 186 amino acids [38]. Despite the low sequence identity, the overall folds of the two proteins are quite similar (Fig. 4). The RMSD is 2.3 Å for protomer 1, and 2.9 Å for protomer 2. The core of

Fig. 4. Overlay of the X-ray crystal structures of NR_0684 (green) and NR_0767 (purple).

the protomers and the dimer interface overlay quite well. The largest difference between the two structures is a result of the unique segment (discussed in Section 3.5) in NR_0684, which does not align with NR_0767 in either protomer 1 or 2.

3.4. FMN and substrate binding sites

In both NRs, each protomer contains an FMN cofactor located near the dimer interface (Fig. 3C-F). In the first FMN binding site of NR_0684 (Fig. 3C), the FMN interacts primarily with residues from protomer 1, but also with one residue (S45) from protomer 2. Two water molecules underneath the FMN isoalloxazine moiety of protomer 1 are involved in the interaction network of this FMN, but there are no fixed water molecules above the FMN. Similar to that seen for the first FMN, the second FMN interacts with residues mostly from protomer 2, but S45 is joined by two residues (A43 and C47 from protomer 1) to create the FMN interaction network. Unlike that observed for the FMN binding site in protomer 1, a phosphate (from the purification buffers) is present in the FMN binding site in protomer 2 (Fig. 3D). C47 and two water molecules stabilize the phosphate group at the opening of the FMN binding pocket. Multiple water molecules above the FMN in protomer 2 participate in binding interactions, but there are no fixed water molecules underneath the FMN, probably due to an FMN positional shift after incorporation of a phosphate group.

The two FMN binding sites of the second NR (NR_0767; Fig. 3E-F) are more accessible to solvent compared with those in NR_0684. Several water molecules are involved in the interactions with the FMN cofactors and the residues above and below the FMN. Similar to that seen for the first NR (NR_0684) described above, the FMN binding site in protomer 1 of NR_0767 (Fig. 3E) is stabilized by mostly residues from this protomer, with interactions from

only one residue (S43) of protomer 2. The other FMN (Fig. 3F) is stabilized by mostly residues from protomer 2, with interactions from three residues (S42, S43, and K44) of protomer 1. Each of the FMN sites accommodates an imidazole molecule (present in the purification buffers) above the isoalloxazine group through a π - π stacking interaction (centroid to centroid distance of 3.8-3.9 Å from the imidazole ring to the isoalloxazine group). In protomer 1, S43 interacts with the imidazole through a water network, but in protomer 2, S43 is directly involved in a H-bond between O_γ(S43) and N₃(imid). The observed interaction between S43 and imidazole implies a likely role for S43 in substrate stabilization.

Our NR structures show a loop area (between α 2 and β 1) above the isoalloxazine ring of FMN; this loop area is present in some other NRs [36, 37, 40]. Both C47 of NR_0684 (underlined in Fig. 3D) and S43 of NR_0767 (underlined in Fig. 3E and 3F) are located in this loop. This loop appears to play an important role in the interaction and stabilization of the phosphate anion in NR_0684 and the imidazole in NR_0767. A structural comparison between our NRs and an *E. coli* NR complex with bound nitrofurazone (PDB accession code 1YKI) [36] showed that the position of the imidazole in NR_0767 overlaps with the semicarbazone “tail” of nitrofurazone in the *E. coli* NR (Fig. 5). The comparison suggests that the position of the imidazole in our NRs is a potential substrate binding site for Mtz when it reacts with NRs, since Mtz also contains an imidazole ring. Thus, the loop between α 2 and β 1 (residues 44-53) may play a significant role in accommodating substrates such as Mtz. Sequence alignments of NRs from different strains show that the two ends of this loop area (Fig. S3) are well conserved, leaving the center area of the loop variable. The variable region of this loop may help explain the differences in substrate binding specificity between NRs.

Fig. 5. Superposition of the FMN binding sites of NR_0767 from *C. difficile* R20291 and a NR (PDB accession code 1YKI) from *E. coli*. The FMN and imidazole from NR_0767 are represented in light green and green, respectively. The FMN and nitrofurazone from the *E. coli* NR complex are represented in light magenta and magenta, respectively.

3.5. Unique segment in NR_0684

Interestingly, a unique segment comprised of $\alpha 6$ and $\alpha 7$ (residues 136-180) was observed to fold differently in each protomer in the NR_0684 homodimeric structure. In protomer 1, this segment stretches out of the main fold, but in protomer 2, the segment folds down to wrap around $\alpha 3$ (see the arrows in the left panel of Fig. 6). The $\alpha 6$ and $\alpha 7$ helices in protomer 2 are shorter and partially unfold into a disordered structure in order to wrap around $\alpha 3$. To stabilize the disordered segment, protomer 2 forms an extra β sheet (circled area in the left panel of Fig. 6).

Structural comparisons show that the segment described above in NR_0684 is quite unique. To the best of our knowledge, no similar structural segment has been reported for any NR or other protein, as judged from our comparisons using the Dali server. This unique segment containing $\alpha 6$ and $\alpha 7$ in protomer 1 inserts into another protomer 1 from a neighboring

Fig. 6. Views of the NR_0684 structure. The structure is shown in two different orientations with a 180° rotation around the vertical axis. The purple represents protomer 1, and the cyan represents protomer 2. The motifs of the secondary structure in the protomer 1 are labeled. Arrows point to the segments with different folding in each protomer. The circled area is the extra β sheet in protomer 2.

symmetry mate, and blocks the FMN binding site resulting in restricted solvent accessibility to the FMN cofactor (Fig. 7). It is apparent that the FMN cofactor in protomer 2 of NR_0684 is exposed to the solvent (Fig. 7, left), whereas the other FMN cofactor is blocked, partially restricting its access to the solvent region (Fig. 7, right). This likely accounts for the observation that the FMN binding site in protomer 1 has no fixed water above the FMN cofactor (Fig. 3C), while the FMN binding site in protomer 2 has more water molecules involved in the FMN binding network (Fig. 3D).

Although it is not clear what the physiological role for this unique segment in NR_0684 is, we hypothesize that this segment may allow NR_0684 to interact with other proteins, and perhaps perform other roles in addition to it functioning as a NR. To test this hypothesis, it

would be interesting to design a further functional assay to find other interacting proteins, as this might help to determine the physiological role of this NR *in vivo*.

Fig. 7. Views of the interaction between symmetry mates of NR_0684. The structure is shown in two different views with a 180° rotation around the vertical axis. The upper parts of the two panels are the surfaces of NR_0684 with two FMN binding sites indicated by arrows. A symmetry mate around NR is labeled in green.

In conclusion, we report the first expression, purification and structural characterization of two *C. difficile* NRs from the hypervirulent R20291 strain. An initial functional assay shows that these two proteins indeed function as nitroreductases with the commonly prescribed Mtz antibiotic, which is the first reported observation of the interactions of Mtz and NRs from any *C. difficile* strain.

Acknowledgements

We are grateful to the Price Family Foundation for funding this research. We would like to thank Dr. Jimmy Ballard (University of Oklahoma) for *Clostridium difficile* gDNA. We would also like to thank Drs. Catherine E. Isom, Smita Menon, and Phil Bourne at University of Oklahoma, and Drs. Steve Almo, James Love, and Vern Schramm at Albert Einstein College of Medicine for very helpful discussions. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. X-ray data reported in this publication were in part collected in the OU Macromolecular Crystallography Laboratory, which is supported, in part, by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103640.

References

- [1] S.M. Poutanen, A.E. Simor, *Clostridium difficile*-associated diarrhea in adults, Can. Med. Assoc. J. 171 (2004) 51-58.
- [2] "Antibiotic resistance threats in the United States.", Centers for Disease Control and Prevention report (2013).
- [3] F.C. Lessa, Y. Mu, W.M. Bamberg, Z.G. Beldavs, G.K. Dumyati, J.R. Dunn, M.M. Farley, S.M. Holzbauer, J.I. Meek, E.C. Phipps, L.E. Wilson, L.G. Winston, J.A. Cohen, B.M. Limbago, S.K. Fridkin, D.N. Gerding, L.C. McDonald, Burden of *Clostridium difficile* infection in the United States, New Engl. J. Med. 372 (2015) 825-834.

- [4] D.E. Voth, J.D. Ballard, *Clostridium difficile* toxins: mechanism of action and role in disease, *Clin. Microbiol. Rev.* 18 (2005) 247-263.
- [5] D.A. Leffler, J.T. Lamont, Treatment of *Clostridium difficile*-associated disease, *Gastroenterology* 136 (2009) 1899-1912.
- [6] L.V. McFarland, Therapies on the horizon for *Clostridium difficile* infections, *Exp. Opin. Invest. Drugs* 25 (2016) 541-555.
- [7] M.D. Roldan, E. Perez-Reinado, F. Castillo, C. Moreno-Vivian, Reduction of polynitroaromatic compounds: the bacterial nitroreductases, *FEMS Microbiol. Rev.* 32 (2008) 474-500.
- [8] D.M. Musher, S. Aslam, N. Logan, S. Nallacheru, I. Bhaila, F. Borchert, R.J. Hamill, Relatively poor outcome after treatment of *Clostridium difficile* colitis with metronidazole, *Clin. Infect. Dis.* 40 (2005) 1586-1590.
- [9] S. Nair, D. Yadav, M. Corpuz, C.S. Pitchumoni, *Clostridium difficile* colitis: factors influencing treatment failure and relapse-a prospective evaluation, *Amer. J. Gastroent.* 93 (1998) 1873-1876.
- [10] S.D. Baines, R. O'Connor, J. Freeman, W.N. Fawley, C. Harmanus, P. Mastrantonio, E.J. Kuijper, M.H. Wilcox, Emergence of reduced susceptibility to metronidazole in *Clostridium difficile*, *J. Antimicrob. Chemoth.* 62 (2008) 1046-1052.
- [11] H. Huang, A. Weintraub, H. Fang, C.E. Nord, Antimicrobial resistance in *Clostridium difficile*, *Int. J. Antimicrob. Agents* 34 (2009) 516-522.
- [12] T. Lynch, P. Chong, J. Zhang, R. Hizon, T. Du, M.R. Graham, D.R. Beniac, T.F. Booth, P. Kibsey, M. Miller, D. Gravel, M.R. Mulvey, P. Canadian nosocomial infection surveillance; Characterization of a stable, metronidazole-resistant *Clostridium difficile* clinical isolate, *PloS One* 8 (2013) e53757.
- [13] M. Gal, J.S. Brazier, Metronidazole resistance in *Bacteroides spp.* carrying *nim* genes and the selection of slow-growing metronidazole-resistant mutants, *J. Antimicrob. Chemoth.* 54 (2004) 109-116.
- [14] G. Reyssset, Genetics of 5-nitroimidazole resistance in *Bacteroides* species, *Anaerobe* 2 (1996) 59-69.

- [15] J.M. Schapiro, R. Gupta, E. Stefansson, F.C. Fang, A.P. Limaye, Isolation of metronidazole-resistant *Bacteroides fragilis* carrying the *nimA* nitroreductase gene from a patient in Washington State, *J. Clin. Microbiol.* 42 (2004) 4127-4129.
- [16] G. Sisson, J.Y. Jeong, A. Goodwin, L. Bryden, N. Rossler, S. Lim-Morrison, A. Raudonikiene, D.E. Berg, P.S. Hoffman, Metronidazole activation is mutagenic and causes DNA fragmentation in *Helicobacter pylori* and in *Escherichia coli* containing a cloned *H. pylori* *rdxA* plus (nitroreductase) gene, *J. Bacteriol.* 182 (2000) 5091-5096.
- [17] J.Y. Jeong, A.K. Mukhopadhyay, D. Dailidiene, Y. Wang, B. Velapatino, R.H. Gilman, A.J. Parkinson, G.B. Nair, B.C. Wong, S.K. Lam, R. Mistry, I. Segal, Y. Yuan, H. Gao, T. Alarcon, M.L. Brea, Y. Ito, D. Kersulyte, H.K. Lee, Y. Gong, A. Goodwin, P.S. Hoffman, D.E. Berg, Sequential inactivation of *rdxA* (HP0954) and *frxA* (HP0642) nitroreductase genes causes moderate and high-level metronidazole resistance in *Helicobacter pylori*, *J. Bacteriol.* 182 (2000) 5082-5090.
- [18] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucl. Acids Res.* 25 (1997) 3389-3402.
- [19] A. Marchler-Bauer, M.K. Derbyshire, N.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, R.C. Geer, J. He, M. Gwadz, D.I. Hurwitz, C.J. Lanczycki, F. Lu, G.H. Marchler, J.S. Song, N. Thanki, Z. Wang, R.A. Yamashita, D. Zhang, C. Zheng, S.H. Bryant, CDD: NCBI's conserved domain database, *Nucl. Acids Res.* 43 (2015) D222-226.
- [20] F.W. Studier, Protein production by auto-induction in high density shaking cultures, *Prot. Expr. Purif.* 41 (2005) 207-234.
- [21] E. Gasteiger, C. Hoogland, A. Gattiker, S.e. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein identification and analysis tools on the ExPASy server, Springer, 2005.
- [22] F.X. Chauviac, M. Bommer, J. Yan, G. Parkin, T. Daviter, P. Lowden, E.L. Raven, K. Thalassinou, N.H. Keep, Crystal structure of reduced MsAcg, a putative nitroreductase from *Mycobacterium smegmatis* and a close homologue of *Mycobacterium tuberculosis* Acg, *J. Biol. Chem.* 287 (2012) 44372-44383.
- [23] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Meth. Enzymol.* 276 (1997) 307-326.

- [24] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr. D* 67 (2011) 235-242.
- [25] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, *J. App. Crystallogr.* 40 (2007) 658-674.
- [26] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. Section D*, 53 (1997) 240-255.
- [27] P. Emsley, K. Cowtan, *COOT*: model-building tools for molecular graphics, *Acta Crystallogr. D* 60 (2004) 2126-2132.
- [28] V.B. Chen, W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, *MolProbity*: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D* 66 (2010) 12-21.
- [29] P.D. Adams, *PHENIX*: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D* 66 (2010) 213-221.
- [30] C.A. Haynes, R.L. Koder, A.F. Miller, D.W. Rodgers, Structures of nitroreductase in three states: effects of inhibitor binding and reduction, *J. Biol. Chem.* 277 (2002) 11513-11520.
- [31] F. Hou, T. Miyakawa, N. Kitamura, M. Takeuchi, S.B. Park, S. Kishino, J. Ogawa, M. Tanokura, Structure and reaction mechanism of a novel enone reductase, *FEBS J.* 282 (2015) 1526-1537.
- [32] S. Cortial, P. Chaignon, B.I. Iorga, S. Aymerich, G. Truan, V. Gueguen-Chaignon, P. Meyer, S. Morera, J. Ouazzani, NADH oxidase activity of *Bacillus subtilis* nitroreductase NfrA1: insight into its biological role, *FEBS Lett.* 584 (2010) 3916-3922.
- [33] J. Couturier, P. Prosper, A.M. Winger, A. Hecker, M. Hirasawa, D.B. Knaff, P. Gans, J.P. Jacquot, A. Navaza, A. Haouz, N. Rouhier, In the absence of thioredoxins, what are the reductants for peroxiredoxins in *Thermotoga maritima*?, *Antioxid. Redox Signal.* 18 (2013) 1613-1622.
- [34] J.W. Choi, J. Lee, K. Nishi, Y.S. Kim, C.H. Jung, J.S. Kim, Crystal structure of a minimal nitroreductase, ydjA, from *Escherichia coli* K12 with and without the FMN cofactor, *J. Mol. Biol.* 377 (2008) 258-267.

- [35] P. Macheroux, UV-Visible Spectroscopy as a Tool to Study Flavoproteins, in: S. Chapman, G. Reid, (Eds.), Flavoprotein Protocols, Humana Press, 1999, pp. 1-7.
- [36] P.R. Race, A.L. Lovering, R.M. Green, A. Osson, S.A. White, P.F. Searle, C.J. Wrighton, E.I. Hyde, Structural and mechanistic studies of *Escherichia coli* nitroreductase with the antibiotic nitrofurazone: Reversed binding orientations in different redox states of the enzyme, J. Biol. Chem. 280 (2005) 13256-13264.
- [37] M. Martinez-Julvez, A.L. Rojas, I. Olekhovich, V. Espinosa Angarica, P.S. Hoffman, J. Sancho, Structure of RdxA--an oxygen-insensitive nitroreductase essential for metronidazole activation in *Helicobacter pylori*, FEBS J. 279 (2012) 4306-4317.
- [38] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403-410.
- [39] L. Holm, P. Rosenström, Dali server: conservation mapping in 3D, Nucl. Acids Res. 38 (2010) W545-549.
- [40] C. Chang, Xu, X., Zheng, H., Savchenko, A., Edwards, A., Joachimiak, A., Midwest Center for Structural Genomics, deposited with the PDB (PDB accession code 2H0U), (2006).