The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
The class III ribonucleotide reductase from *Neisseria bacilliformis* can utilize thioredoxin as a reductant

Yifeng Wei,1, Michael A. Funk, Leonardo A. Rosado, Jiyeon Baek, Catherine L. Drennan, and JoAnne Stubbe

Departments of *Chemistry* and *Biology*, and *Howard Hughes Medical Institute*, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by JoAnne Stubbe, July 29, 2014 (sent for review July 13, 2014)

The class III anaerobic ribonucleotide reductases (RNRs) studied to date couple the reduction of ribonucleotides to deoxyribonucleotides with the oxidation of formate to CO$_2$. Here we report the cloning and heterologous expression of the *Neisseria bacilliformis* class III RNR and show that it can catalyze nucleotide reduction using the ubiquitous thioredoxin/thioredoxin reductase/NADPH system. We present a structural model based on a crystal structure of the homologous *Thermotoga maritima* class III RNR, showing its architecture and the position of conserved residues in the active site. Phylogenetic studies suggest that this form of class III RNR is present in bacteria and archaea that carry out diverse types of anaerobic metabolism.

The generic mechanism of nucleotide reduction by all three classes of RNRs can be divided into two half reactions: the radical initiation process and the reduction process (Fig. 1) (20, 21). In all RNR classes, nucleotide reduction is initiated by generating a 3'-nucleotide radical (22–24) (Fig. 1A, 2) via a transient, conserved, top face thiyl radical (25) (I) on the Cys loop in the active site.

Ribonucleotide reductases (RNRs) catalyze nucleotide reduction via complex radical chemistry, providing deoxyribonucleotides for DNA synthesis in all domains of life. Many anaerobic bacteria and archaea contain the class III O$_2$-sensitive RNR, and those that have been studied to date couple nucleotide reduction to formate oxidation. Here we report the characterization of a second class III RNR subtype that couples nucleotide reduction to the oxidation of thioredoxin. Because of the central role of formate and thiols in many anaerobic processes, the distribution of class III RNRs among different organisms may shed light on aspects of anaerobic biochemistry.


The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4U3E).

1To whom correspondence may be addressed. Email: yifwei@mit.edu, cdrennan@mit.edu, or stubbe@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414396111/-/DCSupplemental.

The formate-hydrogenylase pathway in which it is converted to the waste products H$_2$ and CO$_2$ by formate dehydrogenase (FDH) and hydrogenase (16). However, there are many proteins annotated as class III RNRs present in diverse bacteria and archaea (17, 18) which do not possess PFL or generate formate as an intermediate or end product in their primary metabolism (19), suggesting that an alternative reducing system for class III RNRs might be involved. This variability in the presence of formate-producing pathways is in contrast to the ubiquitous distribution of thioredoxin-like proteins used by the class I and II RNRs. This observation prompted us to carry out a bioinformatics search for candidate class III RNRs that use disulfide chemistry similar to the class I and II enzymes.

The class III ribonucleotide reductases (RNRs) are glycyl radical enzymes present in many strict and facultative anaerobes that catalyze the conversion of nucleotides to deoxyribonucleotides (1, 2) via a mechanism involving complex free radical chemistry and are largely responsible for providing the balanced pool of deoxyribonucleotides required for DNA synthesis and repair (3). The class III RNRs that have been characterized thus far obtain the reducing equivalents required to make deoxyribonucleoside triphosphates (dNTPs) from the oxidation of formate to CO$_2$ (4). Here we report a second subtype of class III RNR from *Neisseria bacilliformis*, which can obtain its reducing equivalents from the thioredoxin (TrxA)/thioredoxin reductase (TrxB)/NADPH system.

RNRs provide the only pathway for de novo biosynthesis of dNTPs (5). They share a structurally homologous active site architecture in the α subunit and a partially conserved, radical-based reduction mechanism. RNRs have been isolated and characterized from all kingdoms of life and, based on the characterization of these proteins thus far, are divided into three classes (I, II, and III) according to the metallo-cofactor used to generate a thyl radical that initiates the radical-dependent reduction chemistry (6). The class I RNRs use cofactors generated by the reaction of reduced metals (Fe, Mn, and Fe/Mn) and O$_2$ and are present only in aerobic organisms. The class II RNRs use adenosylcobalamin in an O$_2$-independent reaction and are present in both aerobes and anaerobes. The class III RNR uses an O$_2$-sensitive glycyl radical (G•) (2) situated in the α protein (NrdD), which is generated by a separate activating enzyme (NrdG) via radical S-adenosylmethionine (SAM)-[Fe₄S₄]⁺⁺ chemistry (7, 8). The class III RNRs are only found in facultative and obligate anaerobes. A second distinction between the three classes has been the source of the reducing equivalents for nucleotide reduction. In the class I and II RNRs, they are provided by a redoxin (thioredoxin, glutaredoxin, or NrdH), which is reduced by thioredoxin reductase and NADPH (9–11). In contrast, for the bacteriophage T4 (12), its Gram-negative host *Escherichia coli* (1), and the Gram-positive *Lactococcus lactis* (13), the only class III RNRs characterized in detail to date, nucleotide reduction is coupled to the oxidation of formate to CO$_2$ (4).

Formate in *E. coli* and *L. lactis* is provided by carrying out the fermentation of sugars to acetate and formate via a pathway involving pyruvate-formate lyase (PFL) (14, 15). For *E. coli* growing in the absence of electron acceptors, formate induces...
occurs via disulfide exchange with a pair of conserved cysteines on the C-terminal tail of α (Fig. 1D) (27).

In the class III RNR, only one of the disulfide-forming cysteines in the active site is conserved (21, 29, 30), and we recently showed (31) that reduction of the ketyl radical to the 3′-keto-deoxy-nucleotide is accompanied by the formation of a thiosulfuranyl radical (Fig. 1C, 7) between the bottom face cysteine thyl radical and a methionine residue. The thiosulfuranyl radical, in equilibrium with the thyl radical, then oxidizes formate to a •CO₂⁻ radical (21) (8), proposed to serve as the reductant for the 3′-keto-deoxy-nucleotide. In all classes, the 3′-keto-deoxynucleotide intermediate (Fig. 1B, 4, or C, 7) is proposed to be reduced by proton-coupled electron transfer (32), with the source of the proton being the conserved glutamate in the class I and II RNRs, and unknown in the class III RNR.

Because of the role of this methionine residue in the reaction with formate in the E. coli class III RNR, we expected that it would be conserved in all formate-dependent NrdDs. However, just as the pathways for formate production are not conserved, our search in the RNRdb (33) showed that this methionine is missing in a set of NrdD sequences. In addition, we were able to clone, express, isolate, and crystallize a related NrdD from the deep-branching thermoophilic bacterium Thermotoga maritima (TmNrdD, 30% sequence identity with NbNrdD; Fig. S1). The mesophilic NbNrdD proved more amenable to biochemical studies, and we demonstrate that it is a G• enzyme and show that, like the previously characterized class III RNRs, NTPs are substrates. We also show that formate is unable to provide the reducing equivalents to make dNTPs. This organism possesses a TrxA that has 61% sequence identity with E. coli TrxA, and activity can be reconstituted in vitro using the E. coli TrxA/TxkB/NADPH system. The X-ray structure of TmNrdD reported here provides our framework for modeling the conserved residues in this newly discovered class III RNR subtype and supports the hypothesis for the NbNrdD that three cysteines and a glutamate are located in the active site region where they can play a role in catalysis. The distribution and significance of this form of class III RNR are discussed.

Results

Our bioinformatics analysis, leading to the identification of a previously unidentified class III RNR subtype that couples nucleotide

![Fig. 1. Mechanistic model for nucleotide reduction by RNRs. (A) First half reaction common to all RNRs. (B) Second half reaction of E. coli class Ia and N. bacilliformis class III RNR. (C) Second half reaction of E. coli class III RNR. M382 in E. coli class III RNR is located two residues from the top face thyl radical (C384), in a position similar to the conserved N437 in E. coli class Ia RNR, which makes a hydrogen bond to the 2'-OH group of the substrate. (D) Mechanistic model for rereduction of the active site disulfide in class I and II RNRs via a pair of conserved cysteines on the C-terminal tail of α.](image-url)
reduction to disulfide bond formation, is described below. To test our hypothesis, we attempted to clone, express, and purify six of the class III RNRs of this subtype, including the enzymes from *N. bacilliformis*, *T. maritima*, and *Shewanella sediminis*, *Pyrococcus furiusus*, *Pseudomonas aeruginosa*, and *Schizosaccharomyces japonicas*. The NbNrdD and NbNrdG were soluble and could be obtained in reasonable amounts and thus became the focus of our attention.

**NbNrdD is a G Enzyme.** To generate active NbNrdD for biochemical studies, we incubated NbNrdD with NbNrdG and SAM in the presence of the acriflavine/bicine photoreduction system (31), resulting in the generation of a radical with an electron paramagnetic resonance (EPR) signal shown in Fig. 2A. The spectrum reveals a dominant hyperfine coupling constant of 40 MHz, proposed to be associated with the H\_2 of glycine, consistent with its assignment as G\_2. Uniform labeling of NrdD with \[^{2}H\]-glycine resulted in the collapse of the signal into a singlet (Fig. 2C), establishing the assignment. Additional spectral features (indicated with arrows in Fig. 2A), which persist even when the activation reaction is carried out in D\_2O (Fig. 2B), are also visible for the PFL G\_2, although less well resolved (36). These features are attributed to hyperfine coupling with additional nonexchangeable protons, likely the \(\alpha\)-protons of the two adjacent amino acids in the sequence (36). These interactions are predicted to be affected by the conformation of the peptide backbone, leading to variations in the G\_2 EPR spectra between different G\_2 enzymes. The lack of exchange of the H\_2 of glycine with D\_2O, previously shown to occur with PFL (37), is similar to observations for the G\_2 of *E. coli* class III RNR (38).

**NbNrdD Catalyzes CTP Reduction Using TrxA/TrxB/NADPH.** Our hypothesis is that reducing equivalents for nucleotide reduction by NbNrdD are delivered by a redoxin, similar to the class I and II RNRs (Fig. 1A and B). To identify a candidate redoxin for NbNrdD, we first carried out a BLAST search using the NbNrdD sequence. This search yielded a set of related sequences with over 50% pairwise identity in diverse organisms, including *Shewanella sediminis* (gammaproteobacteria), *Bacteroides ovatus* (Sphingobacteria), and *Clostridium citroniae* (Firmicutes). Examination of the redoxins present in these organisms revealed that only TrxA/TrxB is conserved. The high sequence identity between *N. bacilliformis* TrxA and *E. coli* TrxA (61% identity; Fig. S2), which has been used in assays for other class I and II RNRs (39–41), suggested that *E. coli* TrxA could be used in our activity assays.

The assays were thus carried out with NbNrdD (0.25 G\_2/a) and *E. coli* TrxA/TrxB/NADPH, and the results are summarized in Table 1. NbNrdD was active for reduction of CTP to dCTP with ATP as an effector, but nearly inactive for CDPP reduction (−3% of the activity for CTP reduction; Table 1). Catalytic activity was dependent on the presence of G\_2 and TrxA (Table 1). Formate failed to produce any dCTP. Unexpectedly the activity was the same in the absence or presence of allosteric effectors (ATP or dATP; Table 1). NbNrdD lacks the ATP cone domain that controls the activity of many RNRs by binding the activator (ATP) or the inactivator (dATP) (42). Thus, in NbNrdD, both ATP and dATP would be predicted to bind to the specificity site and activate nucleotide reduction. The activity that we have obtained with NbNrdD is ∼0.24 s\(^{-1}\) per G\_2, which is 20-fold lower than that of *E. coli* NrdD, which is ∼4 s\(^{-1}\) per G\_2 (43). We hypothesize that the low activity and insensitivity to allosteric effectors is a result of *E. coli* TrxA functioning as a suboptimal reductant, making reoxidation of the active site disulfide, rather than nucleotide reduction, rate limiting (44). Further study of the allosteric regulation of this enzyme will likely be facilitated by cloning, expressing, and using the *N. bacilliformis* Trx system in our assays.

The number of dCTPs formed per NADPH in the reaction mixture is 0.97 (Fig. 3), suggesting a 1:1 stoichiometry, in agreement with the proposal that the reducing equivalents are provided by the TrxA/TrxB/NADPH system (Fig. 1A and B). The ~26 turnovers per G\_2 that occur without addition of NADPH are attributed to the reduction of TrxA by residual DTT carried over from the NbNrdD storage buffer (~27 μM), which is essential for maintaining enzymatic activity.

**NbNrdD(C301A) is Inactive, and Reaction of NbNrdD(C300A) with CTP Generates Cytosine (Cyt).** To test our hypothesis that C301 forms the top face thyl radical that initiates nucleotide reduction (Fig. 1A and B), the NbNrdD(C301A) mutant was generated. This mutant is inactive in dCTP and Cyt formation, despite having 0.25 G\_2/a, consistent with our model. We propose that C300 in NbNrdD, which is adjacent to the top face thyl radical residue C301 on the Cys loop (Fig. 2B and Fig. 6A), plays a role analogous to that of C462 in the *E. coli* class Ia α protein (NrdA) (Fig. 1B), donating reducing equivalents for nucleotide reduction by generating a disulfide with C99 (Fig. 9B). To test this hypothesis, we generated the NbNrdD(C300A) mutant and reacted it with CTP. The analogous C462A mutation in *E. coli* NrdA results in the generation of a 3′-keto-deoxycytidine intermediate (Fig. 4B, 16) that decomposes to release Cyt (17) (27).

The C\_2 of the NbNrdD(C300A) has an EPR spectrum identical to that of the WT-NbNrdD and is generated with a similar efficiency. Reaction of NbNrdD(C300A) with 5-[\(^{2}H\)]-CTP leads to the time-dependent release of ~5.5 equivalents of 5′-[\(^{2}H\)]-Cyt per G\_2 (Fig. 4C), identified by HPLC (Fig. S3), with an initial rate of

<table>
<thead>
<tr>
<th>Reaction conditions</th>
<th>Activity (U/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete (CTP, ATP)</td>
<td>49</td>
</tr>
<tr>
<td>−SAM (no G_2 formed)</td>
<td>ND</td>
</tr>
<tr>
<td>−TrxA</td>
<td>ND</td>
</tr>
<tr>
<td>−TrxA, + formate (10 mM)</td>
<td>ND</td>
</tr>
<tr>
<td>−ATP</td>
<td>47</td>
</tr>
<tr>
<td>−ATP, +dATP (0.1 mM)</td>
<td>51</td>
</tr>
<tr>
<td>−CTP, +CDP (1 mM)</td>
<td>1.5</td>
</tr>
</tbody>
</table>

ND, activity not detected, less than three turnovers per G\_2 over 10 min.

![Fig. 2. X-band EPR spectra of the NbNrdD G\_2.](image-url)
9.3 U/mg (~2.5 min⁻¹ per G•). No dCTP is detected, and the same amount of Cyt is produced in the presence or absence of the Trx system. A control with WT-NbNrdD shows no Cyt release either in the presence or absence of TrxA/TrxB/NADPH.

Crystal Structure of TmNrdD Allows Modeling of Active Site Residues in Redoxin-Dependent NrdDs. Bioinformatics analysis (described subsequently) suggested that in addition to the three cysteines in the active site, a glutamate will also be present. To determine if these residues are located in the active site of this class of redoxin-dependent NrdDs, we wanted to obtain a structure of a representative of this NrdD subtype. Selenomethionine (SeMet)-labeled NrdD from T. maritima (TmNrdD), which is related to NbNrdD (30% sequence identity; Fig. S1), was successfully crystallized (Table S1), and the structure was solved by single-wavelength anomalous dispersion to 1.64-Å resolution. We observe a (β/α)₁₀ barrel architecture similar to the T4 bacteriophage class III RNR (29) [root mean square deviation (RMSD), 2.5 Å], including the C-terminal G• domain and Zn-binding site (Fig. S4A) but with four additional helices at the N terminus (Fig. S4A and S4B). The Cys loop containing essential cysteines C329 and C330 (equivalent to C300 and C301 in NbNrdD), however, is not present in its expected conformation within the barrel. In fact, residues 330–349 are not visible in the electron density map in either monomer, and SDS/PAGE analysis of the crystals indicates that protein cleavage occurs within this segment of amino acids (Fig. S365, which show no sequence or structural similarity to the Cys loop to initiate catalysis (Fig. 6B)). Despite the absence of the Cys loop in our structure, the barrel architecture is intact, and the G• domain is ordered. The C terminus beyond the G• loop is also ordered but adopts two distinct conformations in the two different molecules in the asymmetric unit (Fig. 5C).

Using the TmNrdD structure and T4 bacteriophage NrdD structure [Protein Data Bank (PDB) ID code 1H79] (45), we constructed a hybrid structural model (Fig. 6A). The Cys loop was modeled from T4 NrdD (T4 residues 287–293; equivalent to TmNrdD 327–333). Because there are no structures of a nucleoside triphosphate-bound RNR, CTP was modeled in the active site based on the position of CDP in the T. maritima class II RNR (NrdJ) (PDB ID code 1XJN) (46). Structural superpositions based on aligning residues in the β strands surrounding the active site yield all-atom RMSD values between TmNrdD and T4 NrdD of 0.9 Å and between TmNrdD and T. maritima NrdJ of 2.0 Å.

In the TmNrdD model, C330 is positioned at the tip of the Cys loop to initiate catalysis (Fig. 6A). The modeled C329 is directly adjacent to C125 (C–C–C distance, 3.8 Å), sufficiently close to form the proposed disulfide (Fig. 1B), as occurs in E. coli class Ia (C–C–C distance, 4.1 Å) and T. maritima class II enzymes (C–C–C distance, 4.2 Å) (Fig. 6B). The precise orientation of the Cys loop and the placement of C329 in relation to the ribose of the CTP cannot be determined from this model, but the general locations of C329 and C125 are consistent with their proposed mechanistic role. In the class Ia RNR from E. coli, reduction of the ketyl radical intermediate (Fig. 1A, 3) is proposed to occur by hydrogen atom transfer from C225 from the bottom face of the nucleotide, followed by eventual disulfide formation with C462 (Fig. 1B), which is deeply buried on the innermost side of the active site. TmNrdD C125 is observed in the crystal structure to be positioned equivalently to E. coli class Ia C225, consistent with it playing the same catalytic role. However, TmNrdD does not have a cysteine equivalent to E. coli class Ia C462 in the rear of the active site; S368 occupies this space in the structure. Instead, the location of C329 on the Cys loop requires formation of the disulfide at the front of the active site, in an orientation distinct from the class I/II enzymes, possibly in a position to facilitate its rereduction by redoxins (discussed below).

The reduction of the 3'-keto-deoxynucleotide (Fig. 1B, 4) requires a proton in addition to the electron from the disulfide radical anion. We proposed from sequence alignments that residue E495 could function as the proton source, taking the place of E441 in the E. coli class Ia RNR. This glutamate is not conserved in the E. coli-type NrdD, but is conserved in the T. maritima-type NrdD (Fig. 7 and Discussion). Here we find that the location of E495 in the TmNrdD structure overlaps quite well with the position of E441 (Fig. 6A). Unfortunately, the mutant of the corresponding residue in NbNrdD, E438Q, displayed low G• content, and the enzyme was inactive in forming either dCTP or Cyt, which prevented us from carrying out experiments analogous to the E. coli NrdA(E441Q) mutant (28, 47).

A major puzzle with respect to this newly discovered class III RNR subtype is how the active site disulfide, which must be formed on each round of catalysis, is rereduced by a redoxin. The class I and II RNRs require five cysteine residues for catalysis: three in the active site (Fig. 1B) and two located in the C-terminal tail (Fig. 1D) that are involved in rereduction of the disulfide generated during dNDP (dNTP) formation so that multiple turnovers can occur. All class III RNRs lack a C-terminal tail containing a pair of cysteines (27) that could function in this capacity. The TmNrdD structural model shows that C329 and C125 are found at the outer edge of the active site, in contrast to the deeply buried cysteine pair found in the class I/II enzymes. Nonetheless, a large conformational...
change would seem to be required to allow the direct access of TrxA to the oxidized cysteine pair. The active site in our structure is buried primarily by the presence of the G• domain and two N-terminal helices specific to TmNrdD (blue). Although the active site appears buried, the (β/α)₁₀ barrel architecture conserved in all classes of RNRs and all G• enzymes is known to undergo conformational changes that can either close or expose the active site. In the class II RNR from *Lactobacillus leichmannii*, coenzyme B₁₂ binding causes a shift in a small number of residues exterior to the barrel, which closes the barrel and shields the cofactor from solvent (48). In PFL, the G• domain must exit the active site to become accessible to its partner activating enzyme for radical generation, and circular dichroism spectroscopy shows that binding of PFL to its activase is accompanied by enhanced enzyme conformational flexibility (49, 50). The recent structure of the G• enzyme benzylsuccinate synthase (51) reveals a clamshell-like opening of the barrel, allowing release of the G• domain from the interior of the protein, which would allow for G• formation. In this work, support for the proposal that the G• domain is flexible comes from the observation that the domain position in TmNrdD is shifted by 9 Å (~15° rotation) relative to that of T4 bacteriophage (Fig. 5D). Although the relevance, if any, of this shift is not established at this time, it illustrates the wide range of motion available to the G• domain in class III RNRs. Also notable for understanding active site access in TmNrdD is the observation that the G• domain C-terminal helix has two distinct conformations (Fig. 5C): in monomer A, the helix is extended and completely blocks active site access; in monomer B, the helix kinks and wraps around the protein, similar to other G• enzymes, revealing an opening to the active site. Although this active site cavity is not large enough to accommodate thioredoxin, the flexibility observed in the C- and N-terminal helices, as well as in the G• domain, indicates the likelihood that further opening of the active site that would permit recombination by thioredoxin.

**Bioinformatics Analysis Suggests Three Chemically Distinct NrdD Subtypes.** Further support for the existence of distinct subtypes of NrdDs was obtained by a bioinformatics analysis. A phylogenetic tree of NrdDs (Fig. 7B) was constructed using 59 sequences from the RNRdb (33), chosen to include phylogenetically and metabolically diverse bacteria and archaea. Analysis of the conserved residues proposed to be important for chemistry [the three active site cysteines Cys(1–3), Met, and Glu in Fig. 7A] led us to propose that there are three chemically distinct NrdD subtypes that we label NrdD₁, NrdD₂, and NrdD₃. The top face Cys(1) on the Cys loop and the bottom face Cys(2) are conserved in all NrdDs, and Cys(3), Met, and Glu are variously conserved in the different NrdD subtypes. The phylogeny of these NrdDs suggests that horizontal gene transfer plays a large role in accounting for their distribution among the diverse bacteria and archaea, as postulated for all classes of RNR (18).

NrdD₁ includes the previously studied bacteriophage T4, *E. coli*, and *L. lactis* enzymes that use formate as a reductant (4). Met, thought to play a role in the reaction with formate (Fig. 1A and C) (31), is conserved in all NrdDs. None of the NrdD₁ sequences examined have the Cys(3) and the Glu. NrdD₁s are present almost exclusively in fermentative bacteria, which have a PFL, and in hydrogenotrophic methanogens, where the F420-dependent formate dehydrogenase provides a pathway for formate generation (52). In certain fermentative bacteria, like *E. coli* and *L. lactis*, PFL plays a role in energy metabolism by converting pyruvate to acetyl-CoA, which leads to ATP production via acetyl phosphate, and formate, which is a waste product (16). In some bacteria, PFL also plays a role in providing C1 intermediates for biosynthesis (53, 54).

NrdD₂ includes the enzymes from *N. bacilliformis* and *T. maritima*, which are proposed to obtain reducing equivalents from a redoxin. In these proteins, three Cys(1–3) and a Glu, which enable nucleotide reduction via disulfide chemistry (Fig. 1A and B), are conserved. In these enzymes, Met is often present but is not conserved.
In NrdD2s lacking Met, this position is most commonly occupied by Ser, as in NbNrdD and TmNrdD, but also Ala, Glu, or Phe in some deeply rooted bacteria and archaea. The residue corresponding to the catalytic E441 in E. coli NrdA (Fig. 1A and B), which is expected to be close to the formate binding site, is replaced with either Ser or Thr in NrdD1s. However, in NrdD2s, it is replaced with Ile, Leu, or Phe, which may hinder the access of formate. In contrast to the restricted distribution of NrdD1, NrdD2 is present in non-methanogenic archaea and in bacteria with diverse types of anaerobic metabolism. In support of the nature of the proposed reductant, several organisms have NrdD2s with associated redoxins, including Acetobacterium woodii (55), in the form of a C-terminal fusion to NrdD2, and Desulfarculus bauersii (56) and Methanomalophilicoccus luminyensis (57), which have a thioredoxin-like protein in the operon.

NrdD3, present only in certain methanogens and the closely related Archaeoglobi (58), shows important differences with respect to NrdD1 and NrdD2. Although they share up to 50% pairwise sequence identity with NrdD1s from other methanogens, they lack Met. Also, although they contain all three Cys (1–3) like the NrdD2s, they lack Glu. Additional observations support the designation of a third NrdD subtype and that these proteins will possess RNR activity. NrdD3s all contain the G• consensus sequence, and some of them (e.g., Archaeoglobus veneficus) contain an N-terminal ATP cone domain. In some sequenced organisms, NrdD3 is the only annotated RNR (e.g., Archaeoglobus profundus, Methanospirillum hungatii), although other organisms also contain a NrdD. Interestingly, some NrdD3s (Archaeoglobus veneficus, Methanosarcina barkeri) are found in the same operon as a thioredoxin-like protein.

The designation of the three NrdD subtypes largely correspond to the phylogeny of the protein (Fig. 7B), with the exception of a deeply rooted branch containing M. luminyensis, Aminomonas paucivorans, and Kuenenia stuttgartiensis NrdD2. An additional observation suggesting the existence of NrdDs with distinct types of chemistry is that, although it is uncommon for a single organism to contain two copies of the same NrdD subtype, there are many organisms that contain both NrdD1 and NrdD2 (e.g., Shewanella sediminis, Bacteroides ovatus, and Elusimicrobium minutum). We hypothesize that these class III RNR variants are used under different anaerobic growth conditions, similar to the case of the class Ia and Ib RNRs in E. coli.

**Discussion**

Because of their essential role in the de novo production of deoxynucleotides, RNRs are ubiquitous enzymes in nearly all cellular organisms and many viruses (18, 33, 59). The complex chemistry involved in nucleotide reduction requires initial generation of a transient thiyl radical. The class of RNR used by an organism reflects the mechanism of this radical formation and a combination of factors including the presence or absence of oxygen (60) and availability of metals (61). Here we report the discovery of a subtype of class III RNR in N. bacilliformis where, like the class I and II RNRs, nucleotide reduction is facilitated by a redoxin, which is a ubiquitous protein found in all organisms. This proposed reliance on a redoxin is unlike the class III RNRs studied to date that couple nucleotide reduction to the oxidation of formate, a metabolite produced by some but not all organisms as part of their primary metabolism.

**Biochemistry, Structure, and Bioinformatics Support the Existence of a Second NrdD Subtype.** The cloning of N. bacilliformis class III RNR was motivated by a bioinformatics analysis, which led to the identification of a potential second NrdD subtype containing three cysteines necessary to carry out nucleotide reduction via a mechanism analogous to that of the class I and II RNRs (Fig. 1 A and B). Our biochemical investigations showed that NbNrdD catalyzes the reduction of NTPs, which are also the substrates of formate-dependent class III RNRs. DeoxyNTP formation required the presence of the TrxA/TrxB/NADPH system but not formate, and the 1:1 stoichiometry of NADPH consumption and dCTP production demonstrates that the reducing equivalents for dCTP generation are provided by the Trx system.

As predicted, the NbNrdD(C301A) mutant is inactive because no initiating thiyl radical can be produced. In support of the active site disulfide between C99 and C300 (Fig. 1B), our structural modeling using the related TmNrdD enzyme as the molecular scaffold is consistent with these cysteines being close enough to form a disulfide bond. Furthermore, the NbNrdD (C300A) mutant behaves like the E. coli class Ia RNR C462A mutant in that reaction with substrate (CTP in the case of NbNrdD) results in the formation of Cyt (Fig. 4B), showing that the mutant is competent for formation of a 3′-keto-dCTP intermediate but unable to carry out nucleotide reduction. Thus, our assays establish the presence of a second NrdD subtype that is distinct from E. coli NrdD in being able to use thioredoxin rather than formate as a reductant for nucleotide reduction.

**Distribution of NrdD Subtypes Correlates with Metabolism.** Our bioinformatics study revealed that there are in fact three NrdD subtypes (NrdD1, NrdD2, and NrdD3), the distribution of which shows a striking correlation with the organism’s anaerobic metabolism. Among bacteria, NrdD1 is localized almost exclusively in fermentative bacteria that use PFL. An interesting exception is the deep branching Elusimicrobium minutum (termite group 1; Fig. 7B) (62), which has two NrdD subtypes (NrdD1 and NrdD2) but no PFL: its NrdD1 operon contains a 10N-formyl-tetrahydrofolate synthetase as a possible source of formate. In contrast, NrdD2 is present in bacteria carrying out a diverse range of anaerobic metabolism.

Although it is tempting to explain the distribution of NrdD subtypes by the availability of intracellular formate, a counterexample is provided by acetogens: although formate is an
intermediate of acetogenesis (63), acetogens invariably contain NrdD2. Instead we hypothesize that the energy metabolism of the organism and the redox window that it inhabits determines whether it is more thermodynamically efficient to use formate or other reductants for nucleotide reduction. In fermentative bacteria, formate is produced by PFL as a waste product; thus, NrdD1 may be preferred. In contrast, in bacteria that carry out anaerobic respiration, formate is not known to be produced as part of their energy metabolism, and oxidation of any available formate from the environment can be coupled to the generation of ATP, which rationalizes why NrdD2 may be preferred. In acetogens, coupling of ribonucleotide reduction by NrdD2 to the TrxA/TrxB system could provide oxidized pyridine nucleotides that are substrates for energy conservation (55).

All archaea examined contain exclusively NrdD2, except for methanogens and their relatives, which contain either NrdD1 or NrdD3. In the case of the methanogens, the NrdD subtype used correlates with its mechanism for energy conservation (64), which is in turn related to its formate metabolism. NrdD1 is present only in type I hydrogenotrophic methanogens that carry out energy conservation by means of the cytosolic electron-bifurcating heterodisulfide reductase (64). The type I methanogens examined contain the F420-dependent formate dehydrogenase (52) as a possible means of synthesizing formate, and many can carry out methanogenesis with formate as a substrate. Apart from ribonucleotide reduction, the other known role of formate in the primary metabolism of type I methanogens is in purine biosynthesis, in an ATP-dependent reaction with formate to form formylphosphate catalyzed by PurT (glycinamide ribonucleotide synthetase) (65, 66), suggesting the presence of intracellular formate. A case in point is Methanocella paludicola (rice cluster I) (67), a methanogen with cytochromes that nevertheless uses the type I pathway for methanogenesis (64) and contains NrdD1.

Other methanogens, including methylotrophic and acetoclastic methanogens that carry out energy conservation by other means, and their relatives, the sulfate-reducing Archaeoglobi, use NrdD3 instead. Unlike type I methanogens, these organisms are unable to use formate as a substrate for methanogenesis, and several sequenced members (e.g., Methanosaeta acetivorans and Methanothermobacter mazei) do not contain formate dehydrogenase (68). They carry out purine biosynthesis in a manner dependent not on formate, but on \( ^{10}\text{N}\)-formyl-tetrahydrofolate, catalyzed by PurN (glycinamide ribonucleotide synthase) (65), suggesting that formate may not be present in the cell. Although many NrdD3 operons contain a redoxin, methanogens lack a conserved thioredoxin reductase, and the origin of the reducing equivalents for NrdD3 is unclear, but could be linked to the heterodisulfide reductase system common to these organisms.

Several interesting methanogens and relatives also possess NrdD2. The obligate methylotroph Methanococcales burtonii (69) has both NrdD2 and NrdD3, suggesting that the source of reducing equivalents for these two RNRs may be different. The methanogen relative ANME-1 (70, 71), which carries out anaerobic methane oxidation, has NrdD2 but no NrdD3, in conjunction with its unique metabolism. Methanomassiliicoccus luminyensis (57), which carries out methanogenesis from methanol but is a relative of the nonmethanogenic Thermoplasmatales,

![Diagram](image_url)

**Legend:**
- **A:** Model of a generic NrdD active site showing residues proposed to be important for chemistry. Cys(1) and Cys(2) are conserved in all NrdDs, and Met, Cys(3), and Glu are variously conserved in the different NrdD subtypes.
- **B:** Phylogenetic tree of NrdDs. NrdD subtypes are indicated by highlighting (green = NrdD1, blue = NrdD2, red = NrdD3). Presence of Met, Cys(3), and Glu are indicated by colored bars (green, red, and blue, respectively) along the circumference of the plot. F, bacteria with PFL; B, other bacteria; M1, type I methanogens that use the cytosolic electron-bifurcating heterodisulfide reductase; M2, other methanogens; A, other archaea.
contains a deeply rooted NrdD2 with an associated redoxin in the operon. Because of the central role of formate and thios in many anaerobic processes, the distribution of class III RNRs among different organisms may shed light on aspects of anaerobic biochemistry, particularly in organisms that are unculturable.

A Clue Regarding the Ancestral NrdD. RNR has been proposed to provide the link between the RNA and DNA worlds (72, 73), with the class III RNR proposed to precede the class I and II enzymes. Given the central role of formate and thios in many anaerobic processes, the identity of the original NrdD has implications regarding the metabolism of the first organism with a DNA genome. Unfortunately, the NrdD1 and NrdD2 sequences are too divergent to allow us to convincingly root the NrdD phylogenetic tree. However, we note that the residue Met (Fig. 7A) proposed to be involved in formate chemistry in NrdD1 (31) is also present in many but not all NrdDs, possibly as an evolutionary relic, suggesting that NrdD1 may precede NrdD2. If true, this suggests that formate is present in the ancestral organism, possibly as an intermediate in an ancient Wood–Ljungdahl pathway (55).

Materials and Methods

For additional details, see SI Text.

Cloning, Expression, and Purification of NbNrdD, NbNrdG, and TmNrdD. Details are provided in SI Text. N. bacilliformis nrdD and nrdG were obtained by PCR from genomic DNA (gift from Xiang-Yang Han, University of Texas M. D. Anderson Cancer Center, Houston), and T. maritima nrdD was obtained by PCR from genomic DNA (gift from Kenneth Noll, University of Connecticut, Storrs, CT). The genes were cloned into pET28a-NbNrdD, pET28a-NbNrdG, and pET28a-TmNrdD vectors (Novagen) with primers described in Table S2. To increase soluble expression, N. bacilliformis nrdG was later cloned into pOZ27 to give pMBP-NbNrdG, which contains an N-terminal maltose-binding protein (MBP) fusion with an N-terminal His6-tag. The procedure for expression of [SeMet]-labeled TmNrdD for crystallography was adapted from existing protocols (74). All proteins were purified by TALON (Clontech) chromatography (Fig. S6). The final yield was ~10 mg g−1 cells for NbNrdD (ε280 = 73,410 M−1 cm−1), ~1 mg g−1 cells for MBP-NbNrdG (ε280 = 95,760 M−1 cm−1), and ~2 mg g−1 cells for [SeMet]-TmNrdD (ε280 = 108,830 M−1 cm−1).

Reconstitution of the NbNrdG [4Fe4S] Cluster. The procedure was carried out in a MBraun anaerobic chamber. Solutions of Na2S and of Fe(NH4)2(SO4)2 in water (100 mM) were freshly prepared in the anaerobic chamber. A solution of MBP-NbNrdG (200 μM, 0.3 mM) was made anaerobic on a Schlenk line and brought into the glovebox. A solution of DTT (1 M) was added to 10 mM, for [SeMet]-TmNrdD (∼0.24 s−1 per Ga). Stoichiometry of NADPH Consumption and dCTP Production. The assay mixture was divided into 20-μL aliquots containing NbNrdD (4 μM, ~1 μM Ga), dATP (0.1 mM), 5′-[3H]-CTP (1 mM, 4,170 cpm/nmol), E. coli Trx (5 μM), E. coli TrxB (1 μM), and NADPH (1 mM) in assay buffer (30 mM Tris, pH 7.5, 30 mM KCl, and 10 mM MgSO4) and was incubated at 30 °C. Aliquots (20 μL) were removed at 2-min intervals and quenched with 2% (vol/vol) perchloric acid (20 μL). Subsequent to removal of the precipitates using calf intestine alkaline phosphatase (0.1 mg/mL), the mixture was analyzed by passing the mixture through an anion exchange column to remove the nucleoside triphosphates as previously described (31).

To a 7-mL portion of the eluate of the Dowex-1-x-borate column was added carrier Cyt and dC (10 nmol each). The mixture was concentrated by lyophilization, redissolved in water, and cooled on ice, and the precipitated borate salts were removed by centrifugation. The supernatant was analyzed by HPLC using an Alltech Econosil column (C18, 10 μm, 250 × 4.6 mm) on a Waters 515 HPLC system equipped with a 2,996 photodiode array detector. The compounds were eluted with KPi (20 mM, pH 6.8) at a flow rate of 1.0 mL/min. Fractions were collected (0.5 mL) and analyzed by scintillation counting. Cyt was identified by coelution with a standard at 5 min (Fig. S3).

Crystallography and Crystal Structure of TmNrdD. Crystals of SeMet-TmNrdD were grown aerobiqally by sitting drop vapor diffusion at 21 °C. Protein, with the His-tag intact, at 13 mg/mL in a buffer containing 25 mM Hepes, pH 7.6, 15 mM MgCl2, 20 mM KCl, 0.5 mM Tris-[2-carboxyethyl]phosphine (TCEP), 5 mM dGTP, and 5 mM ATP, was screened against commercial screens (Hampton Research, Microlytic, and Qiagen) at a 1:1 ratio of protein to precipitant. A Phoenix pipetting robot (Art Robbins Instruments) was used for dispensing 150-nL drops in screening trays. Diffraction quality crystals grew over several weeks in wells containing 0.085 M trisodium citrate, pH 5.6, 0.17 M ammonium acetate, 25.5% (wt/vol) PEG 4000, and 15% (vol/vol) glycerol. Crystals were flash frozen in liquid nitrogen without additional cryoprotection. The structure of TmNrdD was solved by single-wavelength anomalous dispersion. A 1.64 Å resolution Se peak anomalous dataset (12664.1 eV) was collected at the Advanced Photon Source using a miniKap goniometer to collect Friedel pairs on a single image. Data were collected on a Pilatus 6M detector (Dectris). The data were indexed and scaled with HKL2000 (78) in space group P21 with unit cell constants a = 78.3, b = 98.8, c = 86.6, and β = 111.7. The resulting unit cell volume of 622,000 Å3 is consistent with two monomers in asymmetric unit with a solvent content of ~40% (statistics found in Table S1). Forty-five initial Se sites were found with SHELXD in the package HKL2MAP (79) with a resolution cutoff of 2.1 Å (d(2/3) = 0.8).
resulting sites were refined using data to 1.8-Å resolution, and density modification was performed in Phenix AutoSol (80), yielding the phased statistics shown in Table S1. The resulting maps at 1.8-Å resolution were adequate for chain tracing and manual building of the entire protein model and surrounding water molecules in Coot (81). The resolution was increased to 1.64 Å, and iterative refinement of the model was performed with phenix. refine. Compose omit maps were used to verify residue and ligand placement. All native residues are present at the N and C termini of chain A; chain B is missing three C-terminal residues. Both chains have additional density at the N terminus corresponding to eight residues of the thrombin cleavage site and linker; no residues of the His-tag are observed, although the His-tag was not cleaved before crystallization. Residues 54–62 of chain A are unstructured. Residues 330–349 of chain A and 328–350 of chain B are not observed in the density at all as a consequence of peptide bond cleavage near residue 330 (Fig. S5). The final model contained 98.3% of residues in the favored region of the Ramachandran plot, with 0.0% in additionally allowed regions and 0.08% (1 residue, G621) disallowed. G621 is converted to the G and thus is expected to adopt a slightly strained conformation in the unactivated protein. Figures were created in PyMOL Version 1.4.1 (Schrodinger, LLC). A structural model for TmNrdD Cys loop was constructed by superimposing the structures of T4 bacteriophage (1H79) (45) and T. maritima class III NRR using residues in the β-barrel surrounding the active site loops. Because there are no structures of a nucleoside triphosphate-bound NRR, Cys modeled in the active site based on the position of CDP in the class II T. maritima NrdJ (1XNJ) (46). The resulting hybrid model should be considered purely as a tool for guiding discussions of the possible chemistry occurring in TmNrdD and was not subjected to energy minimizations or dynamics.

Phylogenetic Analysis of NrdDs. To determine the existence of different NrdD subtypes, 59 amino acid sequences were chosen to include phylogenetically divergent representatives from well-studied systems with MEGAS software (82). Subsequently, a neighbor-joining phylogenetic tree (83) was generated, and the Jones-Taylor-Thornton matrix method (84) was used to determine the evolutionary distances. The unequal rate of variation among amino acid sites was modeled with a γ distribution with shape parameter of 4 (85). Additionally, a complete deletion of the sites containing gaps was used to palliate the length and divergence variation, focusing on more conserved amino acid clusters. The level of confidence for the branches was determined based on 2,000 bootstrap replicates (86). The resulting consensus tree was rendered using the web-based program iTOL (87) (Fig. 7B, bootstrap values presented in Fig. S7).

ACKNOWLEDGMENTS. We thank Dr. Xiang-Yang Han for the gift of N. bacilliformis genomic DNA, Prof. Kenneth Noll for the gift of T. maritima genomic DNA, and Dr. Greg Fournier, Dr. Silvan Scheller, and Prof. Rudolf K. Thauer for helpful discussions. This work was supported in part by ASTAR Singapore (Y.W.) and the National Science Foundation Graduate Research Fellowship under Grant 0649560 (to M.A.F.). C.L.D. is a Howard Hughes Medical Institute Investigator. This work is based on research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by National Institutes of Health (N.I.H.) Award GM103403 from the National Center for Research Resources. Use of the Advanced Photon Source, a U.S. Department of Energy, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357. This work was supported by NIH Grant GM25959 (to J.S.).

15. Knapp J, Sawers G (1990) A radical-chemical route to acetyl-CoA: The anaerobically crystallized residues 54 of chain A are unstructured. Residues 330–349 of chain A and 328–350 of chain B are not observed in the density at all as a consequence of peptide bond cleavage near residue 330 (Fig. S5). The final model contained 98.3% of residues in the favored region of the Ramachandran plot, with 0.0% in additionally allowed regions and 0.08% (1 residue, G621) disallowed. G621 is converted to the G and thus is expected to adopt a slightly strained conformation in the unactivated protein. Figures were created in PyMOL Version 1.4.1 (Schrodinger, LLC). A structural model for TmNrdD Cys loop was constructed by superimposing the structures of T4 bacteriophage (1H79) (45) and T. maritima class III NRR using residues in the β-barrel surrounding the active site loops. Because there are no structures of a nucleoside triphosphate-bound NRR, Cys modeled in the active site based on the position of CDP in the class II T. maritima NrdJ (1XNJ) (46). The resulting hybrid model should be considered purely as a tool for guiding discussions of the possible chemistry occurring in TmNrdD and was not subjected to energy minimizations or dynamics.