INTRODUCTION

The majority of N in the biosphere is in the kinetically stable form of dinitrogen (N\textsubscript{2}). The only known biological means of accessing this vast reservoir of N\textsubscript{2} is via the activity of nitrogenase, a complex metalloenzyme that catalyzes the reduction of N\textsubscript{2} to ammonia (NH\textsubscript{3}). Nitrogenase comprises a two component enzyme whereby in an ATP-dependent process, electrons are shuttled by the nitrogenase iron protein (NifH) to the dinitrogenase reductase (NifDK) which harbors the active site (Bulen and LeCouteur, 1986).

The nitrogenase active site consists of a biologically unique iron (Fe)-sulfur (S) cluster [X–7Fe–C–9S] where “X” is either molybdenum (Mo), vanadium (V), or iron (Fe) (Figure 1) (Einsele et al., 2002; Lancaster et al., 2011; Spatzel et al., 2013). These clusters are referred to by their metal type as FeMo-co, Vnf-co, and FeFe-co, respectively. The cluster is ligated by non-protein homocitrate as well as cysteine and histidine side chains (Shah and Brill, 1977; Einsele et al., 2002). Biochemical studies have been conducted on representatives of Mo-dependent nitrogenase (Nif), V-dependent nitrogenase (Vnf), and Fe-dependent nitrogenase (Anf) (Robson et al., 1986; Eady et al., 1987; Chisnell et al., 1988; Eady, 1996).

These different classes of nitrogenase have in common the ability to catalyze the reduction of N\textsubscript{2} to produce NH\textsubscript{3} with concomitant reduction of protons to molecular hydrogen (H\textsubscript{2}). However, they differ in their catalytic rates and efficiencies including the stoichiometry of NH\textsubscript{3} and H\textsubscript{2} produced per mol N\textsubscript{2} reduced (Eady, 1996).

Phylogenetic analyses of nitrogenase amino acid sequences derived from available genome databases indicate that nitrogenase homologs cluster into five major lineages (Raymond et al., 2004; Glazer and Kechris, 2009; Boyd et al., 2011a,b; Figure 2). These lineages comprise two Nif nitrogenase groups, as well as the Vnf and Anf groups. Between the Nif and Anf lineages is a separate lineage comprised of biochemically uncharacterized nitrogenase homologs which do not cluster phylogenetically with previously characterized nitrogenases. Members of this lineage appear to be functionally competent as organisms present in the genomes of representatives of the hydrogenotrophic methanogens, methanotrophic archaea, and firmicutes. Members of this lineage appear to be functionally competent as organisms which encode homologs are capable of N\textsubscript{2} fixation (Mehta and Baross, 2006; Dekas et al., 2009). Previously, these homologs have been referred to as “uncharacterized nitrogenase” and were shown to form a lineage that branches later than Nif, Vnf, and possibly Anf, indicating that they evolved after metal differentiation in the active site cluster occurred (Boyd et al., 2011b). This observation has prompted questions regarding the metal composition and reactivity of the active site cluster present in these nitrogenase homologs (Boyd et al., 2011b).
McGlynn et al. Classification of uncharacterized nitrogenase

FIGURE 1 | The active site of nitrogenase from *A. vinelandii* (Einsle et al., 2002; Lancaster et al., 2011; Spatzal et al., 2011). The metal cofactor–homocitrate complex is shown as balls and sticks and the residues found to be within 5 Å of the cofactor are shown as lines (thin). Colors: light blue, carbon; blue, nitrogen; red, oxygen; yellow, sulfur; rust, iron; cyan, molybdenum.

In addition to the influence that active site metal composition has on reactivity, studies have shown that the peptide environment surrounding the active site cofactor also influences the enzyme activity of nitrogenase. For example, substitutions of amino acid residues in close proximity to the FeMo-cofactor (FeMo-co) N2 binding site of nitrogenase resulted in a Nif− phenotype in *A. vinelandii* and altered the electronic and substrate reduction properties of the protein (Scott et al., 1990). Other residues which surround the active site have also been shown to dramatically alter substrate range of the enzyme (Mayer et al., 2002; Benton et al., 2003; Seefeldt et al., 2009). These experimental results have been instrumental in implicating individual roles for various amino acid residues surrounding the nitrogenase active site in proton transfer, local movement of nearby side chains, substrate binding, reaction intermediate stabilization, and/or modulating the electronic state of the cluster. Thus, both the metal composition of the active site cofactor and the composition of residues lining the active site cavity influence the reactivity of nitrogenase.

In this study, phylogenetic reconstruction and structural modeling were used to compare nitrogenase homologs, specifically to gain insight into the composition of the cluster at the active site of the biochemically uncharacterized nitrogenases. Active site volumes, inferred from predicted structures, and the amino acid environment that surrounds the metallocofactor were compared among nitrogenase homologs. In addition, models and known crystal structures of nitrogenase were compared by structural relatedness. These data provide new insights into the metal cofactor binding proclivity of nitrogenases enzymes and extend our understanding of the primary sequence and structural variation of nitrogenase. Additionally, this work adds further support to the concept that the first nitrogenase was likely to contain a Mo-based active site cofactor (Boyd et al., 2011a,b).

MATERIALS AND METHODS
SEQUENCE AND PHYLOGENETIC ANALYSIS

NiF homologs were obtained from the NCBI non-redundant database through BLASTp searches using NiF, VnF, and AnF proteins from *A. vinelandii* as queries (YP_002797379, YP_002797497, and YP_002801974, respectively). In the case of NiF searches, sequences were compiled using Pattern Hit Initiated BLAST (PHI-BLAST) specifying the NiF conserved “CXRS” amino acid pattern, where X is variable. In addition, a length criterion of 300–700 amino acids was imposed by the [slen] command.

Sequences were aligned using the EMBL ClustalW2 server with default parameters ( Gonnet weight matrix) and manipulated with Jalview (Clamp et al., 2004; Waterhouse et al., 2009) and ClustalX (Larkin et al., 2007). Sequences were examined for known catalytic residues as reported in Boyd et al. (2011a). Phylogenies were generated using the PhyML webserver (Guindon et al., 2005) and calculated using the maximum likelihood method (Guindon and Gascuel, 2003; Guindon et al., 2010) using Shimodaira–Hasegawa-like aLRT supports and the LG substitution matrix (Anisimova and Gascuel, 2006). The resulting phylogram was projected using Fig Tree version 1.3.1.1, and naming accomplished using the REGEN/TREENER online web server (Leonard et al., 2009). Analysis of amino acid conservation was judged by the presence of * (conserved) or a colon above the alignment in ClustalW corresponding to amino groups of strongly similar properties (scoring >0.5 in the Gonnet PAM 250 matrix).

1 http://tree.bio.ed.ac.uk/
STRUCTURE PREDICTION
Models of NifD and NifD homologs were generated through sequence submission to the iterative threading assembly refinement (I-TASSER) server (Zhang, 2008, 2009; Roy et al., 2010). The top model based on C-score (Wu et al., 2007) was selected for further analysis. The C-scores are derived from calculating convergence of intermediate structures that are produced during the I-TASSER run and range from −5 to 2 with high scores signifying models with high confidence (Zhang, 2008). Structures were visualized using PyMol2. Models were selected in an effort to obtain a representative diversity from each nitrogenase lineage. In the case of uncharacterized nitrogenases, these included nitrogenase associated with the archaeal methanogen Methanocaldococcus infernus ME (YP_003636774.1), an uncultured archaeal methanotroph believed to belong to the ANME-2c clade (ADF27322.1; Pernthaler et al., 2008), and the firmicutes Candidatus Desulforudis audaxviator MP104C (YP_001716346.1) and Syntrophothermus thaler (YP_003703435.1). A NifD homolog from a ANME-1 phylotype (Meyerdiekis et al., 2010) was not included in this analysis as the sequence lacks ligand binding residues and is likely a "Nif-like" protein which are conserved in methanogenic archaea (Staples et al., 2007).

STRUCTURAL COMPARISONS
Inferred structures were compared using an "all against all" comparison performed by the ProCKSI server. The relatedness of structures was calculated with standardized distance matrix derived from the Vorolign V-score (Birzele et al., 2007) and hierarchical clustering accomplished using the Complete Link (furthest neighbor) method (Defays, 1977) with the Clustering Calculator server3.

Active site volumetric calculations were performed on the CASTp server with a probe radius of 1.4 Å (Dundas et al., 2006). Cavities were visually manually inspected in each case to ensure correspondence with the nitrogenase metal cofactor-binding pocket. In order to identify putative conserved active site second shell residues, representative nitrogenase structures from the Protein Data Bank (PDB, Jermain et al., 2003) and those created through homology modeling were structurally aligned with the PyMol program. The amino acids within 5 Å of the metal cofactor in the AnfD model (Einsle et al., 2002), including homocitrate, were selected for analysis.

RESULTS
NITROGENASE HOMOLOG PHYLOGENY
Phylogenetic reconstruction of representative AnfD, VnfD, NifD, and uncharacterized D homologs revealed patterns of clustering that correspond to metal utilization (Figure 2), a finding that is consistent with phylogenetic studies (Raymond et al., 2004; Boyd et al., 2011a,b; Dos Santos et al., 2012). Two lineages comprising NifD were identified, denoted here as NifD-A and NifD-B. The uncharacterized homologs, in relation to Nif, do not resolve into a monophyletic lineage indicating that their phylogenetic placement is based solely on D protein divergence should be interpreted with caution. However, the overall branching order of Nif and uncharacterized nitrogenase, whereby VnfD branch basal to AnfD, is consistent with the branching order observed for concatenated D protein homologs reported previously (Boyd et al., 2011b).

GLOBAL STRUCTURAL COMPARISON
Inferred nitrogenase homolog structures displayed modeling scores of high confidence (C-score, Table 1). These scores range from [−5,2] and models with C-score >−1.5 generally have a correct fold (Roy et al., 2010). In addition, root mean square deviation (RMSD) values obtained through comparison using the program DALI-LITE (Holm and Park, 2000) indicated that the models match very closely to the alpha chain of the high resolution NifD structure 1M1N (Einsle et al., 2002). Similarity can be visually depicted by an overlay of all A. vinelandii inferred structures (Figure 3).

Among the obtained models, the AnfD structures display the lowest I-TASSER (least confident) modeling scores (Table 1). The AnfD models were examined by superimposing structures and manually inspecting them for structural differences. This led to the identification of a short C-terminal stretch of amino acids universally present in AnfD homologs, which is absent from other nitrogenase homologs. AnfD sequences in which this segment had been removed in silico prior to submission for homology model construction displayed markedly improved scores (C-score of 1.46 and 1.40 for A. vinelandii and G. kluyveri sequences, respectively), indicating that this section of the protein is responsible for lower model fit statistics (Table 1).

Together with NifD structures obtained from the PDB [1M1N (Einsle et al., 2002), 1MIO (Kim et al., 1993), and 1QH1 (Mayer et al., 1999)], the inferred structures of AnfD, VnfD, NifD, and uncharacterized nitrogenase homologs were compared to examine whether D homologs can be resolved from other homologs at the structural level. The Vorolign V-score, which bases similarity on the conservation and evolutionary relationship of amino acid contact sets (i.e., Voronoi contacts) between proteins, resolved the various nitrogenase types into clusters by metal type without exception (Figure 4). AnfD and VnfD were resolved from NifD and uncharacterized D proteins. Uncharacterized D proteins cluster between two clades of NifD, a finding that is consistent with previous phylogenetic analyses of these proteins (Boyd et al., 2011b). This observation supports the hypothesis that the uncharacterized D protein homologs are structurally similar to that of NifD and may harbor a similar cofactor to that present in the NifD active site (i.e., FeMo-co). In addition, these observations indicate that NifD, VnfD, and AnfD can be separated on the basis of inferred structure alone in a way that is largely consistent with the branching order based on phylogenetic reconstruction. The placement of the uncharacterized D homologs between the two NifD clades is different, however, between the sequence and structural clustering methods (purple groups, Figures 2 and 4). The discrepancy between structural and phylogenetic clustering is possibly the result of the unique structural features of the NifD-B group, which harbor an internal amino acid extension that all other types lack making them structurally distinct (Kim et al., 1993).

“fmrich-03-00419” — 2013/1/29 — 11:43 — page 3 — #3
Table 1 | Protein structure modeling statistics.

| Organism                          | Accession     | Metal center | Type      | C-score       | TM-score      | RMSD±     |
|-----------------------------------|---------------|--------------|-----------|---------------|---------------|-----------|
| Geobacter metallireducens GS-15  | YP_389330.1   | FeMo-co      | NifD-A    | 1.53          | 0.93 ± 0.06   | 0.7       |
| Desulfotubacterium hafniense Y51 | YP_520503.1   | FeMo-co      | NifD-A    | 1.60          | 0.85 ± 0.08   | 0.7       |
| Uncultured archaeon (ANME-2)      | ADF27322.1    | FeMo-co#     | Unc.      | 1.91          | 0.99 ± 0.04   | 0.7       |
| Ca. Desulfotomaculum MP104C       | YP_201716346.1| FeMo-co#     | Unc.      | 1.20          | 0.88 ± 0.07   | 0.9       |
| Syntrophothermus tepidocaldarius DSM 12680 | YP_203703435.1| FeMo-co#     | Unc.      | 1.26          | 0.89 ± 0.07   | 0.6       |
| Methanocaldococcus fervida ME    | NP_00315674.1 | FeMo-co      | NifD-B    | 1.54          | 0.93 ± 0.05   | 1.6       |
| Clostridium kluyveri DSM 555     | YP_203396571  | FeMo-co      | NifD-B    | 1.59          | 0.94 ± 0.05   | 1.6       |
| Methanospirillum hungatei Go1    | NP_632746.1   | FeMo-co      | NifD-B    | 1.69          | 0.95 ± 0.05   | 1.4       |
| Azotobacter vinelandii DJ        | P16266.3      | FeFe-co      | AnfD      | −0.37         | 0.67 ± 0.13   | 0.8       |
| Clostridium kluyveri DSM 555     | YP_203193772.1| FeFe-co      | AnfD      | −0.04         | 0.71 ± 0.12   | 0.9       |
| Methanospirillum hungatei Go1    | NP_616149.1   | FeFe-co      | AnfD      | −0.40         | 0.66 ± 0.13   | 0.9       |
| Rhodobacter capsulatus DQ9333.1 | FeFe-co       | AnfD         | −0.59     | 0.64 ± 0.13   | 0.9       |
| Azotobacter vinelandii DJ        | YP_2027979671 | FeFe-co      | VnfD      | 1.66          | 0.95 ± 0.05   | 0.8       |
| Clostridium kluyveri DSM 555     | YP_2033961371 | VFe-co       | VnfD      | 1.76          | 0.96 ± 0.05   | 0.8       |
| Methanospirillum hungatei Go1    | NP_616155.1   | VFe-co       | VnfD      | 0.67          | 0.80 ± 0.09   | 0.9       |
| Rhodospseudomonas palustris CGA009| NP_346726.1   | VFe-co       | VnfD      | 1.64          | 0.94 ± 0.05   | 0.8       |

C-score and TM-scores are internal I-TASSER generated scores for estimating the quality of predicted models. C-score exist in a range of [−5, 2], where a C-score of higher value signifies a model with a high confidence (Wu et al., 2007; Roy et al., 2010). TM-score indicates topological similarity of protein structure pairs and estimated accuracy with a range of [0, 1]. A score of 1 indicates perfect structural similarity (Sali and Blundell, 1993). RMSD values in the table were calculated using DaliLite (Holm and Park, 2000) against the NifD subunit of the high resolution Nif 1M1N structure (Einsle et al., 2002). Uncharacterized nitrogenases are abbreviated as “Unc.”

*Internal I-TASSER generated score.

DaliLite calculated RMSD of model against PDB structure 1M1N.

*Cofactor content inferred in this work.

FIGURE 3 | Cartoon diagram of the A. vinelandii nitrogenase D subunits. The NifD structure, drawn from 1M1N is in black (A), and predicted structures for the A. vinelandii vanadium (VnfD) and Fe-only (AnfD) homologs are in green (B) and red (C), respectively. A structural superposition of the three as performed in PyMol is shown in (D).

*fnmcib-03-00419* — 2013/1/29 — 11:43 — page 4 — #4
ACTIVE SITE STRUCTURAL RELATIONSHIPS

A comparison of amino acids within 5 Å of the active site FeMo-co in inferred structures indicated that they comprise nine separate sequence areas or motifs (Figure 1; Table 2). For simplicity, the numbering scheme from that of A. vinelandii NifD was adopted when discussing these motifs. Two of these areas correspond to the two active site cofactor ligands Cys275 and His442 and have been discussed elsewhere (Glazer and Kechris, 2009). The remaining sequence areas analyzed here do not share bonds to the active site and can be considered to be a part of the second shell of amino acids which surround the cofactor. The amino acid distribution at these positions between the nitrogenase classes are listed in Table 2.

A number of amino acids in these regions were conserved across all nitrogenase proteins (Table 2; bold-faced entries). Unique to each class however, the different nitrogenase types exhibit distinct patterns of conservation. For example, the conserved His432 residue is flanked by either a Glu or Gln residue at position 380 in NifD, whereas this position is occupied by the positively charged Lys380 in the VnfD and AnfD. Uncharacterized nitrogenase homologs display Thr, Leu, or Met, but never a positively charged residue in this position. Another notable substitution occurs near conserved Gly424, where Lys or Arg is present at position 428 in NifD and uncharacterized nitrogenases but is present as Gly in the V- and Fe-dependent nitrogenase.

Other differences in sequence conservation in D homologs include Gly69 and Val70. In biochemical assays, Gly69 substitution has been shown to result in resistance to acetylene inhibition, while Val70 substitutions confer the A. vinelandii Mo-dependent nitrogenase the ability to reduce short chain alkynes thereby increasing the substrate size range of the enzyme (Christiansen et al., 2000; Mayer et al., 2002; Benton et al., 2003; Seefeldt et al., 2009). Both Gly69 and Val70 are conserved in the Mo-dependent nitrogenase. However, His or Leu is observed at position 69 in AnfD and VnfD, respectively. Importantly, Gly69 is conserved in the uncharacterized nitrogenase homologs, adding further support to the hypothesis that uncharacterized nitrogenases bind FeMo-co.

To further compare the relatedness of nitrogenase homologs and investigate the potential for uncharacterized nitrogenases to bind FeMo-co, active site volume calculations were performed (Dundas et al., 2006). These calculations were motivated by the untested hypothesis that some of the observed catalytic differences between the nitrogenase types could be related to the nature of the FeMo-co binding cavity. Calculations were performed on three members of each respective lineage, in the absence of the active site cluster. Representatives included homology models generated during this study as well as structures obtained from the PDB (1M1N and 1MIO). The results indicate that calculated active site volumes varied markedly depending on active site cluster composition, with NifD having active site cavities of approximately double the size of those associated with VnfD and AnfD (Table 3). Likewise, the uncharacterized nitrogenase D homologs were shown to have an active site volume that was similar to NifD. These trends were observed both when the inferred structures were compared to each other and when inferred structures were compared with those obtained through x-ray crystallography (e.g., NifD; Kim et al., 1995; Mayer et al., 1999; Einsle et al., 2002). The similarity in uncharacterized nitrogenase active site volume to that of NifD and the large differences in active site volume between uncharacterized nitrogenase D homologs and AnfD/VnfD support the hypothesis that uncharacterized nitrogenases most likely harbor an active site cofactor that is most similar to that present in Mo-dependent nitrogenase (i.e., FeMo-co).
Sequence features are identified by a residue with a raised numeral indicating the A. vinelandii sequence number of the residue directly below it. Underlined residues are those within 5Å of the A. vinelandii active site cofactor. Residues in bold are conserved among all nitrogenases. Residue positions of lower case exhibit conservation limited to “NIG” when including Delftia tsuruhatensis AAS55952.1.

Sequence conservation limited to “AG” when including Delftia tsuruhatensis AAS55952.1.

First G is V in Clostridium hungatei AAB02935.1.

C is undetermined in Candidatus Desulfovibrio autotrophicus YF00171854.1.

G replaced with S in unnamed Archaea BAF96801.1.

First R in BAF96801.1.

First G in V in Clostridium hungatei AAB02935.1.

“Deoxovii” if including Clostridium hungatei AAB02935.1.


discussion

The diversity of reactions catalyzed by metalloenzymes is a consequence of cofactor composition and the protein environment in which these cofactors are bound. This study applies computational tools to examine the relationship between active site cofactor composition, primary sequence variation, and three-dimensional structure in nitrogenase homologs. The results of the analyses of nitrogenase homologs collectively suggest that the uncharacterized nitrogenases are most similar to Mo-nitrogenase and bind nitrogenase homologs. The results of the analyses of nitrogenase homologs collectively suggest that the uncharacterized nitrogenases are most similar to Mo-nitrogenase and bind nitrogenase homologs. The results of the analyses of nitrogenase homologs collectively suggest that the uncharacterized nitrogenases are most similar to Mo-nitrogenase and bind...
the cofactor that are conserved but only specifically within each form of nitrogenase are particularly diagnostic in terms of categorizing the uncharacterized nitrogenase homologs. For example, the position of Arg88 (amino acid sequence of A. vinelandii) is conserved in NifD, whereas both the AnfD and VnfD have lys in this position. The presence of Arg at this position in unclassified nitrogenase D homologs is suggestive of FeMo-co binding. The presence of a negatively charged glutamate at position 380 in NifD and a positively charged lysine in VnFD/AnfD is intriguing, as this position is located near the variable metal binding site of the nitrogenase cofactor. While uncharacterized nitrogenase D homologs do not exhibit conservation at this position, the amino acid observed at this position is not positively charged as it is in VnFD/AnfD. Other residues that may play a role in the delineation of metal cofactor binding among nitrogenase homologs include position 65 in NifD. Here, both Mo-type and uncharacterized nitrogenase have a conserved Ala whereas both VnfD and AnfD have Cys. Near position 65, Cys of Nif and the uncharacterized lineage differs from the conserved His or Leu in Anf and Vnf, respectively. These observations are in line with the observation that in the environment of Cyta, which is a ligand of the nitrogenase cofactor, Mo-type nitrogenase exhibit the amino acid pattern Cys–Tyr–Arg–Ser, Cys–Glu–Arg–Ser, or Cys–His–Arg–Ser, whereas the VnF and Fe-only forms contain the sequence Cys–Ala–Arg–Ser (Glazer and Kechris, 2009). Together, these patterns suggest that the uncharacterized Nif bind FeMo-co.

In summary, the results presented here suggest that uncharacterized nitrogenase homologs present in the genomes of anaerobic methanogens, anaerobic methanotrophic archaea, and anaerobic fermicutes are likely to harbor FeMo-Co. Previous phylogenetic studies indicate that the ancestral nitrogenase contained FeMo-co, and that a FeMo-co containing ancestor then diversified to give rise to alternative nitrogenase and uncharacterized nitrogenase homologs. That the uncharacterized nitrogenase homologs, which have yet to be identified in the genomes of aerobes or facultative anaerobes, harbor FeMo-co supports the notion that diversification of nitrogenase toward the use of alternative metals (e.g., V and Fe) likely occurred in an anoxic environment (Biy et al., 2011b). Collectively, these observations contrast with hypotheses put forth based on the bioavailability of metals in marine environments and the evolution of nitrogenase, whereby Anf/Vnf were suggested to predate Nif (Anbar and Knoll, 2002; Raymond et al., 2004). Reconciling these observations, we suggest an important role for microenvironments and transient fluctuations in metal availability in driving the diversification of nitrogenase early during its evolutionary history.

ACKNOWLEDGMENTS

Shawn E. McGlynn is an Agouron Geobiology Option postdoctoral fellow in the Division of Geological and Planetary Sciences at Caltech and is grateful for support provided by the Agouron Institute. Additional support for this work was provided by the Department of Energy Grant DE-SC0004949 (to Victoria J. Orphan) and the NASA Astrobiology Institute (NAE) NNX08RC-N85A (to John W. Peters and Eric S. Boyd). Eric S. Boyd also wishes to acknowledge support from the National Science Foundation (EAR-1123689 and PIRE-0964241). The authors are grateful for comments from Aaron D. Goldmann, Joshua A. Steele, Wolfgang Nitschke, and members of the laboratory of Victoria Orphan.

REFERENCES

Anbar, A. D., and Knoll, A. H. (2002). Implications for prebiotic evolution: a bioinorganic bridge? Science 297, 1372–1374.

Anbar, A. D., Goldmann, O., and Knoll, A. H. (2002). Approximate likelihood-ratio test for branching: a fast, accurate, and powerful alternative. Sys. Biol. 51, 538–552.

Arnheim, N., Hummel, M., and Ruddle, F. H. (1966). Chromosomal integration of genes for a biochemical trait: the lac operon. Science 151, 882–890.

Barton, G. J. (2004). The Jalview Java Bioinformatics Moleculer System. Bioinformatics 20, 2960–2962.

Birzele, F., Gewehr, J. E., Csaba, G., and Birkenmeier, W. (2001). Psr and psr-1 encode a second alternative nitrogenase and a putative subunit of the FeMo-co synthetase in Azotobacter vinelandii. Eur. J. Bacteriol. 21, 244–250.

Biswas, B. K., Wolters, H., van der Donk, W. A., and LeComte, J. R. (1966). The vanadium nitrogenase of Azotobacter vinelandii. J. Biol. Chem. 241, 170, 27–33.

Boyd, E. S., Jobin, A. D., Miller, L. C., and Dean, D. R. (2011a). Isolation and characterization of an acetylene-resistant nitrogenase. J. Biol. Chem. 286, 11459–11464.

Bowler, G. F., Cui, J., and Ludden, P. W. (2004). Identification of Arg96 (amino acid sequence of Azotobacter vinelandii MoFe protein) as a component of the FeMo-co binding site. J. Bacteriol. 186, 9699–9706.

Bowler, G. F., Cui, J., and Ludden, P. W. (2004). Identification of Arg96 (amino acid sequence of Azotobacter vinelandii MoFe protein) as a component of the FeMo-co binding site. J. Bacteriol. 186, 9699–9706.

Bulen, W. A., and LeComte, J. R. (1966). The nitrogenase system from Azotobacter chroococcum. J. Biol. Chem. 241, 170, 27–33.

Chatterjee, R., Ludden, P. W., and Dean, D. R. (2000). Isolation and characterization of the alpha subunit of the nitrogenase FeMoco insertion–mediating protein of Azotobacter vinelandii. J. Bacteriol. 182, 5784–5790.

Chisnell, J. R., Premakumar, R., and Seefeldt, L. C. (2002). Nitrogenase MoFe protein alpha subunit. Proc. Natl. Acad. Sci. U.S.A. 99, 5789–5794.

Chisholm, D. W., and Filges, D. (1988). A late methanogen ancestor. Science 242, 1169–1171.

Christiansen, J., Cash, V. L., Seefeldt, L. C., and Dean, D. R. (2000). Isolation and characterization of an acetylene-resistant nitrogenase. J. Biol. Chem. 275, 11459–11464.

Clamp, M., Colf, S. M., and Barton, G. J. (2004). The Jamber Jera aligner edition. Bioinformatics 20, 426–427.

Dey, E. A., and LeComte, J. R. (1966). The nitrogenase system from Azotobacter chroococcum. J. Biol. Chem. 241, 170, 27–33.

Deyser, D. (1977). An efficient algorithm for a complete link method. Comput. J. 20, 364–366.

Deutscher, M. P., Pereyssat, R. C., and Orphan, V. J. (2009). Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. Science 320, 422–426.

Dinh, P. C., Fang, Z., Maen, J. W., Seto, J. C., and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. BMC Genomics 13,162. doi:10.1186/1471-2164-13-162

Dudas, J., Ouyang, Z., and Seefeldt, L. C. (2002). Molecular insights into nitrogenase FeMo-box insertion – the role of His 274 and His 451 of MoFe protein alpha subunit. J. Biol. Chem. 277, 16307–16309.

Ehlefeld, G., Konrad, S., and Seefeldt, L. C. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.

Emsley, O., Tewari, A. F., Andrade, S. L. A., Schmid, B., Toshida, M., Howard, J. R. et al. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.

Emsley, O., Tewari, A. F., Andrade, S. L. A., Schmid, B., Toshida, M., Howard, J. R. et al. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.

Emsley, O., Tewari, A. F., Andrade, S. L. A., Schmid, B., Toshida, M., Howard, J. R. et al. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.

Emsley, O., Tewari, A. F., Andrade, S. L. A., Schmid, B., Toshida, M., Howard, J. R. et al. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.

Emsley, O., Tewari, A. F., Andrade, S. L. A., Schmid, B., Toshida, M., Howard, J. R. et al. (2002). Nitrogenase MoFe protein at 1.8A resolution: a central ligand in the FeMo-co co factor. Science 297, 1066–1070.
Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Kim, C. H., Newton, W. E., and Dean, D. R. (2004). Localization of a catalytic intermediate bound to the FeMo-cofactor of nitrogenase. J. Biol. Chem. 279, 50770–50779.

Kim, J., Woo, D., and Rees, D. C. (2004). Role of the MoFe protein alpha-subunit histidine-195 residue in FeMo-cofactor binding and nitrogenase catalysis. Biochemistry 33, 7988–8000.

Kim, J., Woo, D., and Rees, D. C. (2004). X-ray crystal structure of the nitrogenase molybdenum-iron-protein from Clostridium pasteurianum at 3.0 Å resolution. J. Mol. Biol. 339, 7134–7175.

Lancaster, M., Remzek, M., Ettenhuber, P., Hsu, Y., Ribbe, M. W., Nieuw, E., et al. (2011). X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. Proc. Natl. Acad. Sci. U.S.A. 108, 9209–9214.

Lee, J.-H., Igarashi, R. Y., Larraydun, M., Dean, D. R., Des Santos, C. P. D., Deon, D. R., et al. (2004). Anisometric-goldenintermediate during alfalfa reduction by nitrogenase. J. Am. Chem. Soc. 126, 9765–9769.

Leonard, G., Stevens, J. R., and Richards, T. A. (1999). New insights into structure-function relationships in nitrogenase: a 1.6 Å resolution X-ray crystallographic study of Klebsiella pneumoniae MoFe protein. J. Mol. Biol. 292, 871–895.

Meyerdierks, A., Kube, M., Huber, P., Ribbe, M. W., Klunder, R., et al. (2010). Metagenome provided new insights into structure-function relationships in nitrogenase from an alpha-704(Mo) protein variant. J. Biol. Chem. 285, 395–399.

Michalak, G., Collin, U., Dittrich, M., Sahai-McEwan, S., and Mühle, A. (1997). Comparative biochemical characterization of the iron-only nitrogenase and the molybdenum nitrogenase from Rhodobacter capsulatus. Eur. J. Biochem. 244, 789–803.

Nicolaides, G. P., Hoepfner, L. G., and Seefeldt, L. C. (2011). Molybdenum nitrogenase catalyzes the reduction and coupling of CO to form hydrocarbons. J. Biol. Chem. 286, 19417–19423.

Peters, J. W. (2010). Insights into substrate binding at FeMo-cofactor in nitrogenase from the structure of an alpha-704(Mo) protein. Nature 467, 7052–7055.

Peters, J. W., Boyd, E. S., and Dean, D. R. (2009). Role for the nitrogenase MoFe protein alpha-subunit in FeMo-cofactor binding and catalysis. Nature 459, 188–191.

Peters, J. W., Boyd, E. S., Dean, D. R. (2009). Mechanism of Mo-dependent nitrogenases. Annu. Rev. Biochem. 78, 701–722.

Roy, A., Kucukural, A., and Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. Nucleic Acids Res. 36, W557–W560.

Roy, A., Kucukural, A., and Zhang, Y. (2013). TISSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 8, 725–738.

Sali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–779.

Sarma, R., Barney, B. M., Keable, S., Dean, D. R., Seefeldt, L. C., and Peters, J. W. (2010). Insights into substrate binding at FeMo-cofactor in nitrogenase from the structure of an alpha-704(Mo) protein variant. J. Biol. Chem. 285, 395–399.

Schlupske, K., Collin, U., Dittrich, M., Sahai-McEwan, S., and Mühle, A. (1997). Comparative biochemical characterization of the iron-only nitrogenase and the molybdenum nitrogenase from Rhodobacter capsulatus. Eur. J. Biochem. 244, 789–803.

Schulz, T., Wittinghofer, A., and Skolnick, J. (2003). Simultaneous modeling of small proteins by iterative TASSER simulations. BMC Biol. 5:7. doi: 10.1186/1471-2105-5-37.

She, Y., Tang, Y., Dean, D. R., and Seefeldt, L. C. (2011). Molybdenum nitrogenase catalyzes the reduction and coupling of CO to form hydrocarbons. J. Biol. Chem. 286, 19417–19423.

Zhang, Y. (2007). I-TASSER: fully automated protein structure prediction in CASP8. Proteins 70(Suppl. 1), 110–115.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 October 2012. Accepted: 10 November 2012. Published: 18 November 2012. Edited by: Martin Stettler, University of Lausanne, Switzerland

Copyright © 2013 McGlynn, Boyd, Peters and Orphan. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.