THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 34, Issue of August 22, pp. 32150–32156, 2003

Printed in U.S.A.

The Three-dimensional Structure of the Core Domain of NafY from *Azotobacter vinelandii* determined at 1.8-Å Resolution*

Received for publication, April 23, 2003, and in revised form, May 14, 2003
Published, JBC Papers in Press, May 16, 2003, DOI 10.1074/jbc.M304264200

David H. Dyer‡, Luis M. Rubio‡, James B. Thoden, Hazel M. Holden, Paul W. Ludden§, and Ivan Rayment¶

From the Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

The *Azotobacter vinelandii* NafY protein (nitrogenase accessory factor Y) is able to bind either to the iron molybdenum cofactor (FeMo-co) or to apodinitrogenase and is believed to facilitate the transfer of FeMo-co into apodinitrogenase. The NafY protein has two domains: an N-terminal domain (residues Met[1]–Leu[98]) and a C-terminal domain (residues Glu[99]–Ser[232]), referred here to as the "core domain." The core domain of NafY is shown here to be capable of binding the FeMo cofactor of nitrogenase but unable to bind to apodinitrogenase in the absence of the first domain. The three-dimensional molecular structure of the core domain of NafY has been solved to 1.8-Å resolution, revealing that the protein consists of a mixed five-stranded β-sheet flanked by five α-helices that belongs to the ribonuclease H superfamily. As such, this represents a new fold capable of binding FeMo-co, where the only previous example was that seen in dinitrogenase.

Nitrogenase, the bacterial enzyme responsible for nitrogen fixation, catalyzes the reduction of nitrogen gas (N_2) to ammonium in an ATP-dependent manner. Nitrogenase has two components: dinitrogenase (otherwise known as the MoFe protein, NifDK, or component I) and dinitrogenase reductase (also known as the Fe protein, NifH, or component II). Dinitrogenase is a 240-kDa $\alpha_2\beta_2$ tetramer of the *nifD* and *nifK* gene products (1) that contains two interesting types of metal clusters: an iron-molybdenum cluster conjugated with homocitrate (FeMo-co)[1] and an 8Fe-7S center (P-cluster). Dinitrogenase reductase is a 60-kDa α_2-dimer of the *nifH* gene product with a single 4Fe-4S cluster coordinated between the two subunits (2).

The assembly of FeMo-co *in vivo* is a complex biosynthetic process that requires the products of at least six nitrogen fixation (*nif*) genes: *nifB*, *nifE*, *nifH*, *nifN*, *nifQ*, and *nifV* (3). In addition, the *nafY* product is required under conditions of molybdenum starvation (4), and *nifX* is required for *in vitro* FeMo-co synthesis (3). The component polypeptides of mature dinitrogenase (NifDK) are not required for FeMo-co synthesis. In a *nifB* deletion strain of *Azotobacter vinelandii*, the synthesis of FeMo-co is interrupted, and cells accumulate a hexameric form of dinitrogenase ($\alpha_2\beta_2\gamma_2$), known as apodinitrogenase (5). The hexameric apodinitrogenase has a complete P-cluster assembly but lacks FeMo-co. Substrate reduction can be achieved *in vitro* after activation of apodinitrogenase by the addition of purified FeMo-co. Upon activation, the γ-subunits disassociate, and mature dinitrogenase ($\alpha_2\beta_2$) is formed.

In *Klebsiella pneumoniae*, the apodinitrogenase γ-subunit has been identified as the product of the *nifY* gene (6). Surprisingly, whereas *A. vinelandii* possesses a *nifY* gene of unidentified function, its apodinitrogenase γ-subunit is encoded in a non-*nif* gene termed *nafY* (4). An analysis of the amino acid sequence of NafY reveals that NafY is part of a family of proteins that also includes NifY, NifX, VnfX, and NifB (Fig. 1). All members of this family of proteins have been implicated in the biosyntheses of FeMo-co or FeV-co from the molybdenum- or vanadium-containing dinitrogenases, respectively (see Ref. 7 for a recent review). NafY can bind independently to FeMo-co and to apodinitrogenase. It has been suggested that NafY protects FeMo-co before the cofactor is inserted into apodinitrogenase (8). In addition, NafY may function as a molecular prop that maintains apodinitrogenase in a conformation that facilitates FeMo-co insertion.

In an effort to understand the manner in which NafY interacts with apodinitrogenase and its role in FeMo-co incorporation, a structural investigation was initiated. During crystallization trials, it was discovered that the full-length NafY (26 kDa) undergoes adventitious cleavage between residues Leu[98] and Glu[99] and residues Ser[232] and Val[233] so that the crystals contained an autonomous "core domain" of 14 kDa that extends from Glu[99] to Ser[232]. In this paper we report the purification, crystallization, and determination of the three-dimensional structure of this domain and its ability to bind FeMo-co.

MATERIALS AND METHODS

Expression of NafY (Full-length and Core Domain)—Full-length NafY from *A. vinelandii* was expressed as a glutathione *S*-transferase (GST) fusion protein. The chimera was constructed in the pGEX-4T-3 plasmid (Amersham Biosciences), which incorporates a thrombin cleavage site to remove GST from the fusion partner. *nafY* was PCR-amplified from plasmid pRHB25 (4) using oligonucleotides 5′-TAACATATGGTAACCCCCGTGAACATGAG-3′ and 5′-TCGGGATCCTCATGCCCTGGCCGCCTCGTCC-3′ as primers. The *nafY* gene cartridge was then digested with *Nde*I and *Bam*HI, made blunt by Klenow treatment, and ligated into the Klenow-treated *Bam*HI site of plasmid pGEX-4T-3 to

* This work was supported in part by National Institutes of Health Grants GM35332 (to P. L.) and AR35186 (to I. R.). Use of the Argonne National Laboratory Structural Biology Center beam-lines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Research, under Contract No. W-31-109-ENG-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1P90) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ These authors contributed equally to this work.

§ To whom correspondence may be addressed: Dept. of Plant & Microbial Biology, 211 Koshland Hall, University of California, Berkeley, CA 94720-3102. Tel.: 510-642-7171; Fax: 510-642-4612; E-mail: pludden@nature.berkeley.edu.

¶ To whom correspondence may be addressed: Dept. of Biochemistry, 433 Babcock Dr., Madison, WI 53706. Tel.: 608-262-0437; Fax: 608-262-1319; E-mail: Ivan_Rayment@biochem.wisc.edu.

[1] The abbreviations used are: FeMo-co, iron-molybdenum cofactor; GST, glutathione *S*-transferase; PEG, polyethylene glycol.

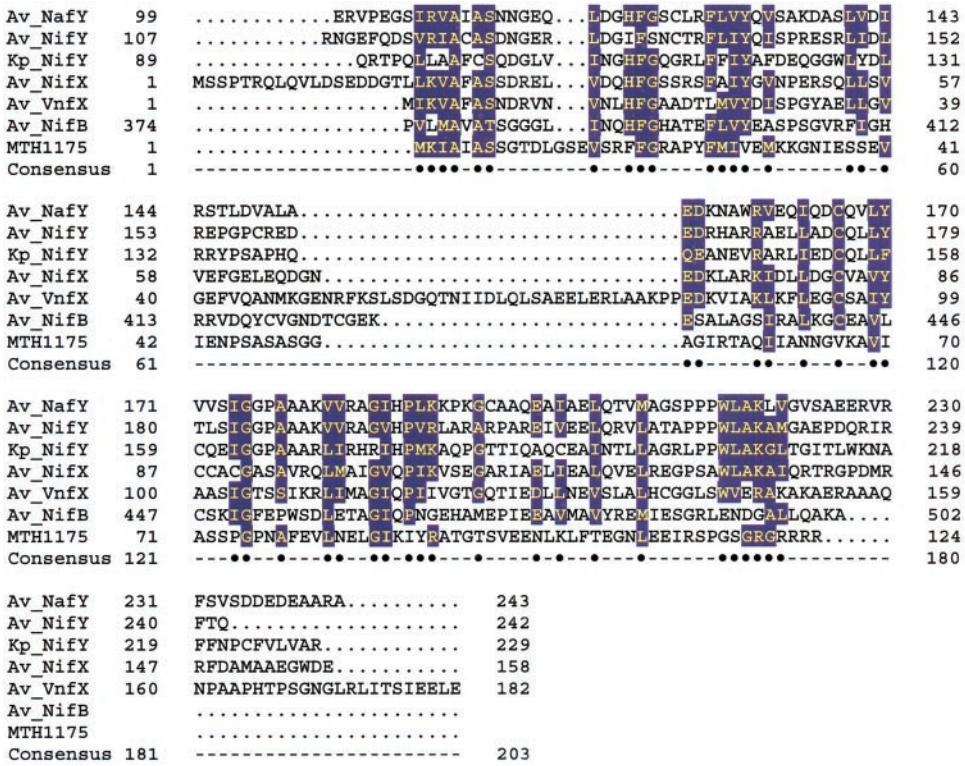

FIG. 1. **Amino acid sequence alignment of members of the nitrogen fixation family and the hypothetical protein MTH1175.** The entire sequence is shown for Av_NifX (P14887), Av_VnfX (C41660), and MTH1175 (1EO1). Alignments of the core domain are shown for Av_NafY (Q9F5X9), Av_NifY (P14886), Kp_NifY (P09135), and Av_NifB (P11067). The *circles* indicate amino acid residues identical or similar in five or more of the aligned sequences. Amino acid sequences are indicated by the single letter code. Gaps were introduced for optimal alignments. *Av*, *A. vinelandii; Kp*, *K. pneumoniae; MT*, *M. thermoautotrophicum*.

generate plasmid pRHB62. The fidelity of the construction was confirmed by sequencing. Once it was discovered that the crystals contained a proteolytic cleavage product, a second GST fusion construct was created that encoded only the residues of the core domain of NafY (residues Glu^{99}–Ser^{232}). Protein prepared from the truncated construct was used to grow crystals of the NafY core domain necessary for preparation of heavy atom derivatives. The overproduction of both the full-length and the truncated GST-NafY fusion proteins was carried out in *Escherichia coli* BL21 (pREP-4). *E. coli* cells were grown in 250-ml cultures in Luria-Bertani medium supplemented with ampicillin (50 μg/ml) and kanamycin (25 μg/ml) at 30 °C. Overexpression was induced by the addition of isopropyl-β-D-thiogalactoside to a final concentration of 0.2 mM when the cells reached an optical density of 0.6 at 550 nm. *E. coli* cells from 2.5 liters of culture were harvested at 5 h postinduction and resuspended in 15 ml of phosphate-buffered saline (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.4). Glycerol was added to a final concentration of 10% (v/v), and the cells were stored at −80 °C.

Purification of NafY—Prior to cell lysis, the cell suspension was supplemented with the protease inhibitors leupeptin (1 μg/ml) and phenylmethylsulfonyl fluoride (0.2 mM). Cells were lysed with a French press at 15,000 p.s.i., followed by centrifugation at 37,000 × *g* for 30 min to remove cellular debris. Purification of the GST-NafY fusion protein was accomplished within 6 h at room temperature. The clarified cell extract was passed over a 2-ml prepacked glutathione-Sepharose 4B column (Amersham Biosciences) according to the manufacturer's recommendations. Both the full-length NafY and the core domain were separated from the GST moiety while bound to the column by incubation for 4 h with 100 units of thrombin (Amersham Biosciences) and then eluted in 6 ml of phosphate-buffered saline. To remove the thrombin, the sample was then loaded onto a 1-ml benzamidine-Sepharose column (Amersham Biosciences) equilibrated in phosphate-buffered saline. NafY was retained by the benzamidine-Sepharose resin but could be eluted with 8 ml of phosphate-buffered saline supplemented with NaCl to a final concentration of 1.14 M. The absence of thrombin contamination in pure preparations of NafY was verified using the chromogenic substrate S-2238 (Chromogenic). S-2238 or H-D-phenylalanyl-L-pipecolyl-L-arginine-*p*-nitroaniline dihydrochloride is cleaved by thrombin between the arginine residue and *p*-nitroaniline, liberating *p*-nitroaniline, which in the free form absorbs at 405 nm. No protease was detected with this assay after passage over the benzamidine-Sepharose column. The purified protein was concentrated to 20 mg/ml with a Centricon YM-10 device (Amicon) and then dialyzed in a Slide-A-Lyzer dialysis cassette (Pierce) for 4 h against 25 mM KCl, 50 mM Tris-HCl, pH 8.5 at 4 °C. A typical purification procedure yielded 15 mg of protein from 500 ml of cell culture. Initially, the protein was either stored with 10% glycerol and frozen in liquid nitrogen or drop-frozen directly into liquid nitrogen when used in crystallization trials.

Anoxic Native Gel Electrophoresis and Immunoblot—Anoxic native gel electrophoresis was performed as described previously (9). Purified proteins were resolved on 7–16% acrylamide (37.5:1 acrylamide/bisacrylamide) and 0–20% sucrose gradient gels. After electrophoresis for 20 h at 100 V, gels were either stained with Coomassie Brilliant Blue or transferred to nitrocellulose membranes for immunoblot as described before (10). For all samples, 4 μg of NafY protein was used per lane, and when required, 1.5 nmol of FeMo-co was added. Native gels were stained for iron as previously described (11).

Purification of Other Components—Apodinitrogenase was purified in its NafY-deficient $\alpha_2\beta_2$ form by affinity chromatography utilizing the NafY-GST fusion protein bound to glutathione-Sepharose 4B resin.[2] The resultant protein was dialyzed anaerobically against 50 mM Tris-HCl, pH 7.5, prior to use in the native gel experiments.

Crystallization of NafY Core Domain—Crystals of NafY core domain were grown by hanging drop vapor diffusion at room temperature. Typically, 3 μl of NafY protein at 10 mg/ml was combined with an equal volume of 13% PEG 10,000, 2.5% glycerol, 100 mM Tris-HCl, pH 8.5. Crystals grew to a size of 0.2 × 0.2 × 0.5 mm within 6 days if they were obtained from the full-length construct, for which adventitious proteolysis was necessary for crystal formation, whereas crystals from the core domain construct would grow to these dimensions after 2 days. The crystals belonged to space group $P3_221$ and had cell dimensions of a = 49.7 Å, b = 49.7 Å, c = 101.3 Å, and γ = 120°.

Structural Analysis of NafY—The initial structural analysis was performed with x-ray data collected to 2.6-Å resolution with CuK_α radiation at 4 °C with a Bruker HISTAR area detector equipped with

[2] L. M. Rubio and P. W. Ludden, unpublished results.

TABLE I
X-ray data collection statistics

Data set	Resolution		Independent reflections	Completeness	Redundancy	Avg I/Avg $\sigma(I)$	R_{merge}[a]
	Å		%				
Native no. 1	30.0	2.60	4785	93.4	4.2	7.1	7.5
	2.72	2.60[b]	575	93.6	2.6	1.7	26.1
Gold	30.0	2.60	4378	85.5	3.5	5.2	10.5
	2.72	2.60	452	73.6	1.9	1.5	32.9
Lead	30.0	2.60	3496	77.0	3.1	4.4	12.6
	2.72	2.60	378	61.6	1.4	1.4	44.5
Native no. 2	30.0	1.80	24,908	100.0	7.2	43.7	5.7
Synchrotron data APS 19ID	1.86	1.80	2531	100.0	5.1	5.6	34.6

[a] $R_{merge} = (\Sigma I_{hkl} - I \times 100)/(\Sigma I_{hkl})$, where the average intensity I is taken over all symmetry equivalent measurements and I_{hkl} is the measured intensity for a given reflection.
[b] Statistics for the highest resolution bin.

Supper long mirrors. The x-ray source was a Rigaku RU200B x-ray generator operated at 50 kV and 90 mA. The x-ray data were processed with XDS (12, 13) and internally scaled with XSCALIBRE.[3] X-ray data collection statistics are presented in Table I.

Crystals used for the preparation of heavy atom derivatives and the initial native data set were stabilized in synthetic mother liquor containing 20% PEG 10,000, 100 mM NaCl, 100 mM Tris, pH 8.5. Two isomorphous heavy atom derivatives were prepared by soaking native crystals in either 12 mM $KAu(CN)_2$ for 3 days or 10 mM trimethyllead acetate for 7 days (14). X-ray data for the heavy atom derivatives were collected to 2.6-Å resolution. The isomorphous R-factors between the native, gold, and lead derivative data sets were 16.5 and 9.5%, respectively. Heavy atom binding sites (two per derivative) were determined with SOLVE (15). The positions, occupancies, and temperature factors of these sites were refined with SOLVE, yielding an overall figure of merit of 0.32. The phases were improved by solvent flattening with RESOLVE (16) to give a figure of merit of 0.56 for the solvent-flattened electron density map. The polypeptide chain was built into the electron density map with the program TURBO (17) and refined with TNT (18).

High Resolution X-ray Data Collection and Least Squares Refinement—Prior to high resolution data acquisition, the crystals were flash-frozen. The crystals were transferred sequentially from the synthetic mother liquor to a cryoprotective solution in five equal steps starting from 20% PEG 10,000, 100 mM Tris-HCl, pH 8.5, to a final cryoprotectant concentration of 20% PEG 10,000, 500 mM NaCl, 100 mM Tris-HCl, pH 8.5, and 17% ethylene glycol and flash-frozen in a nitrogen stream at −160 °C. The unit cell parameters, after freezing, for proteolyzed native NafY were as follows: a = 49.2 Å, b = 49.2 Å, c = 99.2 Å, and γ = 120°. A native x-ray data set was collected to 1.8-Å resolution at the Structural Biology Beamline 19-ID at the Advanced Photon Source and processed and scaled with HKL2000 (19). This structure was solved by molecular replacement with the program AMORE (20) starting from the refined structure determined at 2.6-Å resolution. Iterative cycles of least-squares refinement and manual model building reduced the R-factor to 19.4% for all measured x-ray data from 30.0- to 1.8-Å resolution. The R-free was 26.3% for 10% of the data that were excluded from the refinement. Least-squares refinement statistics are presented in Table II. Analysis of the coordinates with the program PROCHECK (21) revealed that 88.6% of the residues lie in the most favored regions of the Ramachandran plot, whereas the remaining 11.4% of the residues lie in additionally allowed areas. No residues are located in the disallowed regions. A section of representative electron density is shown in Fig. 2.

TABLE II
Refinement statistics

Parameter	Value
Resolution Limits (Å)	30.0–1.80
R-factor[a] (overall) %/no. of reflections	19.4/13,182
R-factor (working) %/no. of reflections	18.8/11,853
R-factor (free) %/no. of reflections	26.3/1347
No. of protein atoms	932[b]
No. of heteroatoms	184[c]
Bond lengths (Å)	0.011
Bond angles (degrees)	2.47
Trigonal planes (Å)	0.005
General planes (Å)	0.013
Torsional angles (degrees)[d]	16.7

[a] R-factor = $(\Sigma||F_o| - |F_c|| / \Sigma|F_o|) \times 100$, where F_o is the observed structure factor amplitude, and F_c is the calculated structure factor amplitude.
[b] These include multiple conformations for Asp^{137}, Ser^{145}, Asp^{154}, and Lys^{194}.
[c] These include 1 ethylene glycol and 180 water molecules.
[d] The torsional angles were not restrained during the refinement.

RESULTS AND DISCUSSION

In Situ Generation of the NafY Core Domain—The initial goal of this study was to determine the structure of the full-length NafY protein from *A. vinelandii* (243 amino acids and a molecular weight of 26,141). However, it was discovered during crystallization trials that NafY protein is highly susceptible to adventitious proteolytic cleavage, which yields a truncated protein that crystallizes readily. To date, the full-length NafY protein has not been crystallized.

The presence of the full-length NafY protein in the initial material was confirmed by SDS-PAGE and by mass spectroscopy. No obvious contaminating proteins or degradation products were observed (data not shown). However, the hanging drop crystallization experiments yielded diffraction quality crystals of a truncated product of the NafY protein. NafY crystals were isolated, dissolved in Tris buffer; and analyzed by SDS-PAGE, mass spectroscopy, and N-terminal amino acid sequencing. The results of these analyses demonstrated that NafY crystals contained a stable domain composed of 134 amino acids with a molecular weight of 14,337 that extends from residue Glu^{99} to Ser^{232} in the sequence of the full-length NafY. We have termed this domain the core domain of NafY because it is found conserved in all members of the NafY/NifY/NifX/VnfX/NifB family of proteins. One explanation for the production of this truncated product of NafY would be the presence of a trace amount of an unidentified protease that would co-elute during the purification of NafY. An obvious candidate is the thrombin used to separate NafY from the GST moiety. However, this is unlikely, because the preparations of pure NafY were shown to be free of thrombin activity when assayed with the chromogenic substrate S-2238. Moreover, purified GST-NafY fusion protein, which was not thrombin-treated, was also cleaved after being 48 h at room temperature (data not shown). Attempts to prevent the adventitious cleavage of NafY with benzamidine were unsuccessful.

In order to reproduce the results in a controlled way, a truncated version of *nafY* was constructed, starting at Glu^{99}

[3] I. Rayment and G. Wesenberg, unpublished data.

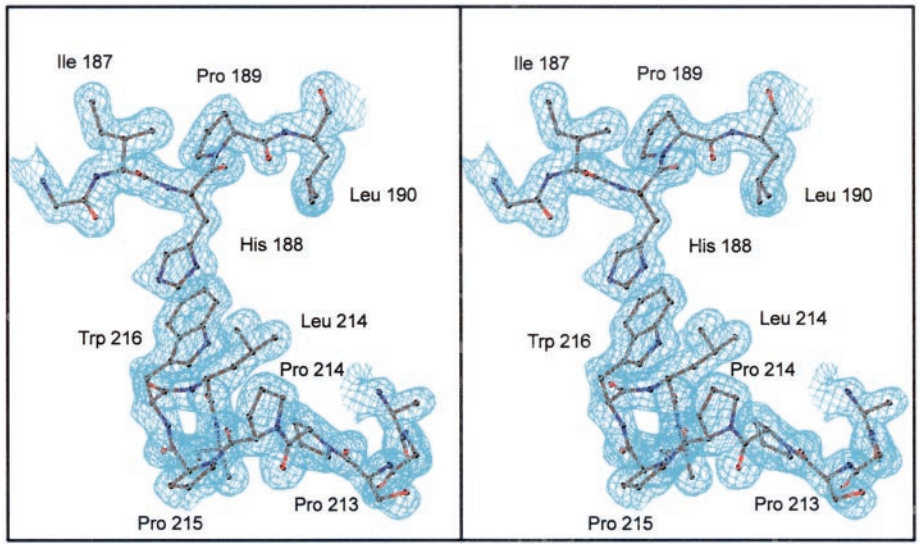

FIG. 2. **Stereo view of a representative region of electron density from the final model for the NafY core domain.** The map was calculated with coefficients of the form $2F_o - F_c$ map and was contoured at 1 σ. The region shown is from β-strand 5 (residues Ile^{187}–Leu^{190}) and α-helix 4 (residues Pro^{213}–Leu^{217}). Note the unusual stretch of three prolines and parallel π stacking between His^{188} and Trp^{216}. Figs. 2–4 were prepared with the programs Molscript and Bobscript (33, 34).

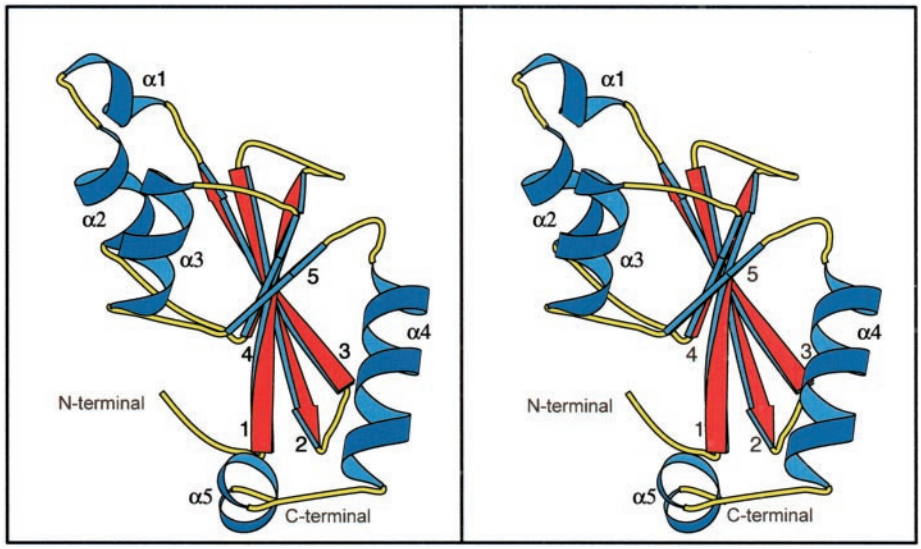

FIG. 3. **Stereo ribbon representation of NafY core domain.** The α-helices and β-sheets are numbered according to their occurrence in the structure.

and ending at Ser^{232} of the native sequence. This fragment was fused to the GST gene, overexpressed in *E. coli*, and purified in the same manner as the full-length NafY. Identical crystals were obtained from the engineered truncated version of NafY and the *in situ* proteolyzed NafY. This observation shows that the fragment of NafY that extends from Glu^{99} to Ser^{232} is able to fold independently and thus constitutes a true domain.

Structural Description—The electron density map for the core domain of NafY is continuous and well defined for almost all of the residues from Glu^{99} to Val^{221}, where the remaining 11 residues, Gly^{222}–Ser^{232}, are disordered but present within the crystal. The coordinate set also includes 180 water molecules and one molecule of ethylene glycol. The polypeptide consists of a single domain that is dominated by a mixed five-stranded β-sheet flanked on both sides by α-helices, three on one side and two on the other (Fig. 3). The strand order within the sheet is 3-2-1-4-5, where β-strand 2 is antiparallel to the other strands. The α-helices on one side of the β-sheet occur between β-strands 3 and 4, and between strands 4 and 5. The two α-helices on the opposite side of the β-sheet occur after β_5. Thus, the order of the secondary structural elements from the N to C termini is β_1-β_2-β_3-α_1-α_2-β_4-α_3-β_5-α_4-α_5. The secondary structural elements are defined as follows: β-strand 1, residues Ser^{105}–Asn^{113}; β-strand 2, residues Leu^{126}–Ser^{134}; β-strand 3, residues Asp^{137}–Ser^{145}; α-helix 1, residues Thr^{146}–Ala^{150}; α-helix 2, residues Ala^{152}–Ile^{163}; β-strand 4, residues Gln^{167}–Val^{172}; α-helix 3, residues Gly^{176}–Gly^{186}; β-strand 5, residues Ile^{187}–Lys^{192}; α-helix 4, residues Ala^{197}–Ala^{210}; α-helix 5, residues Pro^{214}–Val^{221}. There are two large loops positioned at the same side of the NafY core domain, one between β_1 and β_2 and the other between β_3 and α_1. This arrangement of secondary structural elements is similar to the ribonuclease H motif as defined in SCOP (22).

A search of the protein structure data base (23, 24) with the program DALI (25, 26) reveals that the structure of the NafY core domain is most similar to that of the 124-amino acid protein MTH1175 from the thermophilic archeobacterium *Methanobacterium thermoautotrophicum* (27) (RCSB accession number 1EO1). The *Z*-score for the comparison of NafY core domain and MTH1175 is 10.0, whereas the next highest match (Z = 6.0) corresponds to the Holliday junction resolvase (RuvC), which belongs to the ribonuclease H fold superfamily (the *Z*-score represents the strength of the structural similarity in S.D. values above the expected value for a random match). The overall root mean square difference in the positions of the α-carbons when NafY and MTH1175 are superimposed with the program ALIGN (28) is 1.4 Å for 92 structurally equivalent residues, which represent 75% of the total number of residues in the NafY core domain (Fig. 4). These two polypeptides have

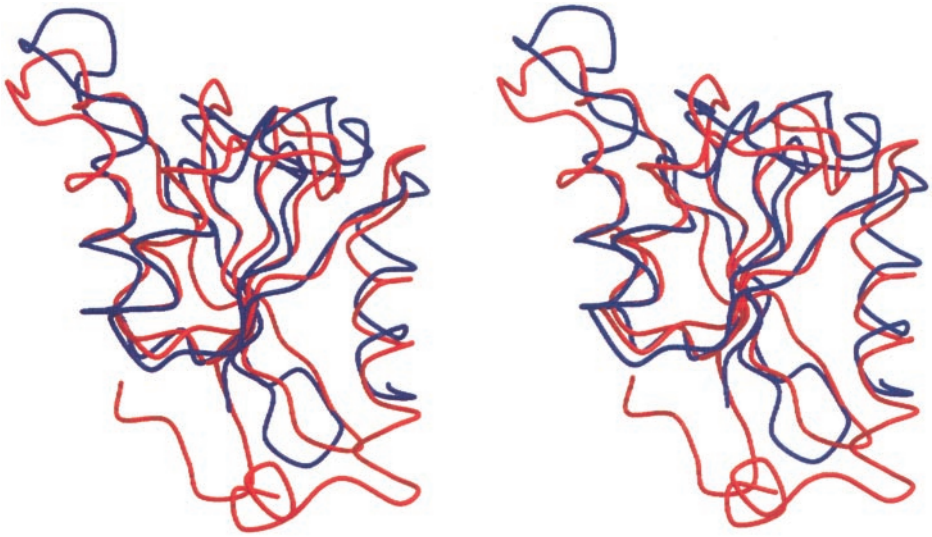

FIG. 4. **Superposition of the structures of protein MTH1175 from M. thermoautotrophicum (*blue*) (27) and the NafY core domain (*red*).** The coordinates for MTH1175 were taken from the RCSB with accession number 1EO1 and superimposed on NafY core domain with the program ALIGN (28).

a remarkably similar structure, considering that the overall sequence identity between them is only ~15% (Fig. 1).

Although the function of MTH1175 is unknown, it also belongs to the ribonuclease H fold superfamily (22), whose members are involved in binding to DNA or hybrid DNA-RNA complexes (27). As a consequence, proteins in the ribonuclease H fold class frequently exhibit distinct patches of surface positive charges that interact with the negatively charged phosphate backbones of the nucleic acid strands. Interestingly, NafY also exhibits a very nonrandom charged surface that can be roughly described as negatively charged on one side and positively charged upon the other (Fig. 5).

Although there is unequivocal evidence that NafY interacts with dinitrogenase and with FeMo-co, there is no experimental evidence to suggest that NafY interacts with DNA or RNA molecules. However, NafY belongs to a family of proteins whose members are involved at different steps in the biosynthesis of FeMo-co, including NifX and NifY proteins, and it has been reported that both NifX and NifY from *K. pneumoniae* can influence the stability of the *nifHDK* mRNA, which encodes the structural components of dinitrogenase and dinitrogenase reductase (29, 30). The deletion of *nifY* or *nifX* genes in *K. pneumoniae* increases the stability of the *nifHDK* mRNA, which accumulates to higher levels than in the wild-type strain under conditions of nitrogenase expression. Conversely, when *nifY* or *nifX* are overexpressed, the stability of the *nifHDK* mRNA is reduced, and it accumulates at lower levels than in the wild type. Given their amino acid sequence similarity to NafY, it is likely that NifX and the C-terminal domain of NifY (Fig. 1) have similar structures to that of the NafY core domain and, in turn, to ribonuclease H. This extrapolation would give structural credence to the reports by Roberts and collaborators (29, 30) assigning a role for NifX and NifY as negative regulators of nitrogenase expression.

The Different Roles of NafY Rely on Different Domains of the Protein—The primary role of NafY appears to be that of a molecular chaperone that maintains apodinitrogenase in a conformation that facilitates receipt of FeMo-co, thereby acting as a metallo-insertase (8). In addition, NafY can bind *in vitro* to FeMo-co, suggesting that this protein has two distinct activities. The presence of multiple domains within NafY, as suggested by the facile generation and stability of the core domain and by amino acid sequence alignments with NafY homologs, raises the question of whether the molecular chaperone and FeMo-co binding functions of NafY are mediated by different modules within the polypeptide chain. To address this question, the ability of the core domain to bind FeMo-co was investigated by anoxic native gel electrophoresis. As shown in Fig. 6, the addition of FeMo-co to the full-length NafY results in faster gel migration of the resulting species. Native gel shift mobility of NafY upon FeMo-co binding has been extensively used to assay for the presence of a NafY-FeMo-co complex. It suggests that the protein is either more compact in the presence of FeMo-co or that the net charge on the protein has changed. Similarly, the addition of FeMo-co to the core domain of NafY results in a shift of mobility in native gels. This experiment

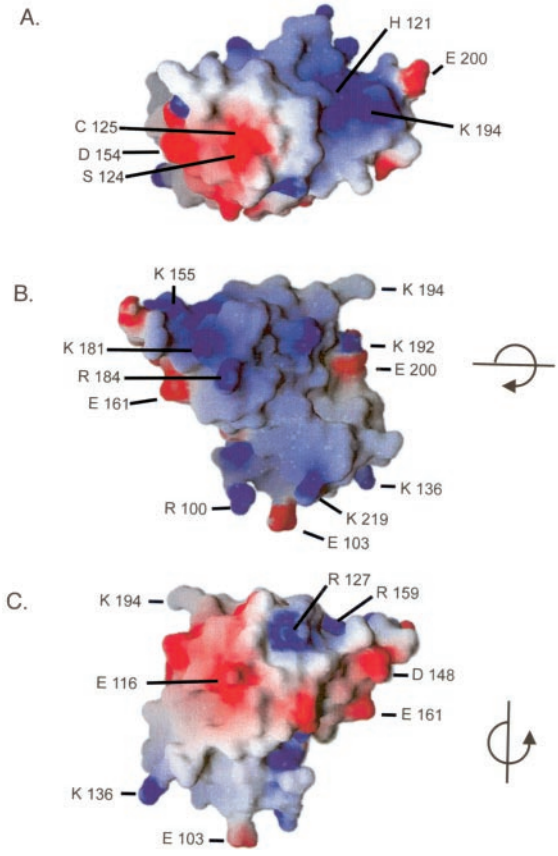

FIG. 5. **Three orthogonal views of the surface electrostatic features of NafY core domain.** The surface potentials were calculated and depicted with the program GRASP, where *blue* and *red* coloration indicate +6.00 and −6.00 kT, respectively (35).

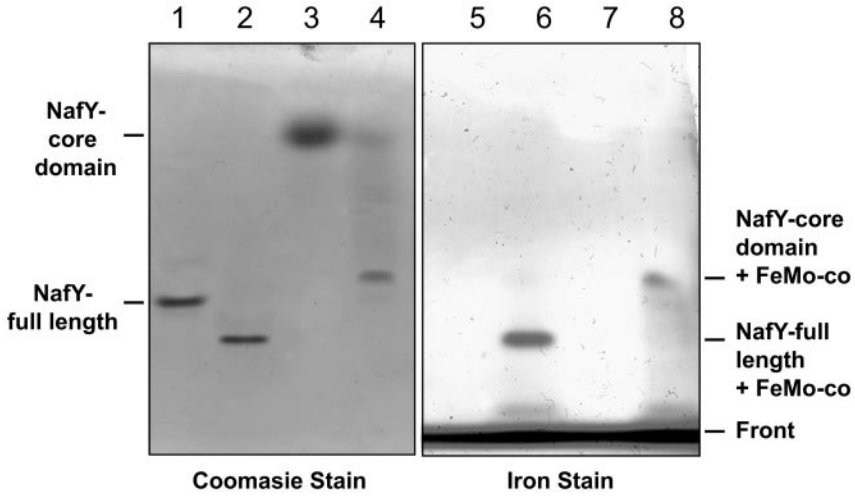

FIG. 6. **Gel shift analysis demonstrating interaction of the full-length and core domain of NafY with FeMo-co.** The samples were prepared in anaerobic 9-ml serum vials and contained 4 μg of purified NafY (full-length or core domain) in 50 μl of anaerobic buffer (25 mM Tris, 20% glycerol, pH 7.5). When indicated, NafY was incubated with 1.5 nmol of purified FeMo-co for 5 min at room temperature in order to allow interaction between NafY and FeMo-co. The samples were subsequently electrophoresed for 20 h in an anoxic native gel (7–16% acrylamide) that was finally stained with Coomassie Brilliant Blue (*lanes 1–4*) or iron stain (*lanes 5–8*). *Lanes 1* and *5*, full-length NafY; *lanes 2* and *6*, full-length NafY plus FeMo-co; *lanes 3* and *7*, NafY core domain; *lanes 4* and *8*, NafY core domain plus FeMo-co.

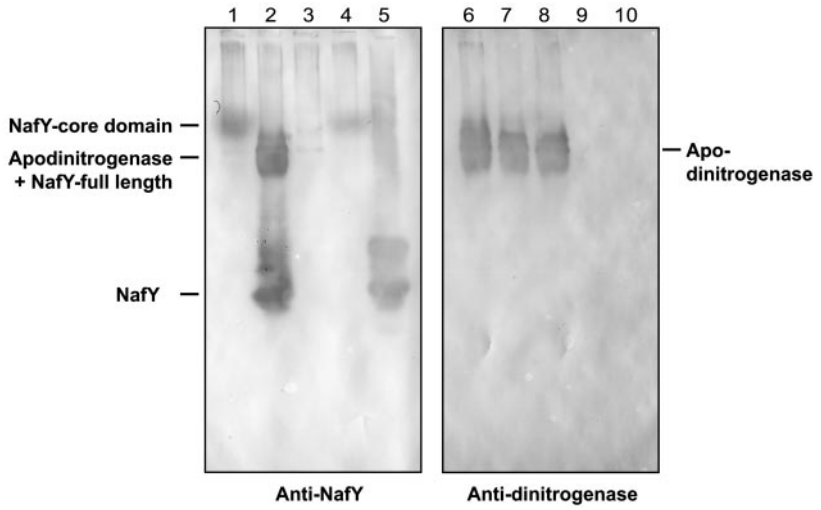

FIG. 7. **Immunoblot analysis of the interaction between NafY (full-length and core domain) and the NafY-deficient $\alpha_2\beta_2$ apodinitrogenase.** The samples were prepared in anaerobic 9-ml serum vials and contained 4 μg of purified NafY (full-length or core domain) in 50 μl of anaerobic buffer (25 mM Tris, 20% glycerol, pH 7.5). When indicated, 8 μg of purified apodinitrogenase was added, and samples were incubated for 5–10 min at room temperature to promote interaction between NafY and apodinitrogenase. Samples were then subjected to electrophoresis for 20 h in anoxic native gels (7–16% acrylamide) and transferred to nitrocellulose membranes, and immunoblots were developed with antibodies to NafY (*lanes 1–5*) or to dinitrogenase (*lanes 6–10*). *Lanes 1* and *6*, apodinitrogenase plus NafY core domain; *lanes 2* and *7*, apodinitrogenase plus full-length NafY; *lanes 3* and *8*, apodinitrogenase; *lanes 4* and *9*, NafY core domain; *lanes 5* and *10*, full-length NafY.

implies that FeMo-co can bind to both proteins as is confirmed by staining the gels for iron, which only binds at the gel-shifted positions corresponding to the NafY-FeMo-co complexes. These experiments establish that the FeMo-co binding capacity of NafY resides within the core domain. The identification of a specific binding site for FeMo-co on the NafY core domain is under investigation.

The ability of the core domain to interact with a purified NafY-deficient $\alpha_2\beta_2$ form of apodinitrogenase was also examined. This form of apodinitrogenase was purified from an *A. vinelandii* mutant strain with deletions in *nifB* and *nafY* genes. It has been previously shown that the full-length NafY binds to apodinitrogenase in crude extracts of *A. vinelandii* (8). Anoxic native gel electrophoresis, followed by immunoblotting with antibodies to either NafY or to dinitrogenase, revealed that full-length NafY has the ability to bind and comigrate with purified apodinitrogenase (Fig. 7, *lanes 2* and *7*). However, the NafY core domain does not co-electrophorese with apodinitrogenase, indicating that it binds weakly or not at all under these conditions (Fig. 7, compare *lanes 1* and *6*).

It is noteworthy that the full-length NafY has a low isoelectric point (pI = 4.6) when compared with the calculated isoelectric point for the NafY core domain (pI = 7.2), indicating that the N-terminal region of NafY (Met^1–Leu^{98}) is negatively charged. Interestingly, the structure of a NafY-deficient $\alpha_2\beta_2$ apodinitrogenase reveals a solvent-exposed region rich in positive charges in the proximity of the FeMo-co binding site (31). That region is found buried inside the protein in the three-dimensional structure of mature holodinitrogenase, once the FeMo-cofactor has been inserted (32). This observation, together with the fact that the core domain of NafY is not able to bind to apodinitrogenase by itself, suggests that the role of the N-terminal region of NafY might be to recognize the apo state of dinitrogenase.

Conclusions—The present study establishes the modular nature of NafY from *A. vinelandii*. It shows that the C-terminal region of NafY (Glu^{99}–Ser^{232}; referred to here as the core domain) folds autonomously and is sufficient for FeMo-co binding, whereas the N-terminal region of the protein (Met^1–Leu^{98}) is required for interaction with apodinitrogenase. The determination of the three-dimensional structure of NafY is of interest because it is the first FeMo-co-binding protein, other than dinitrogenase, that has been solved, and it represents a new fold for FeMo-co binding. The core domain of NafY has a β_1-β_2-β_3-α_1-α_2-β_4-α_3-β_5-α_4-α_5 fold and a polarized surface charge distribution that are surprisingly similar to those of the ribonuclease H family. NafY is part of a conserved family of proteins, whose members are involved at different steps in the biosynthesis of FeMo-co in *A. vinelandii* and *K. pneumoniae*, and we predict that the structure presented here represents a general protein fold for the carriage of FeMo-co or FeMo-co precursors.

Acknowledgments—We thank Dr. Scott Lovell for assistance in recording the high resolution native data, Andrew Hoyord for assistance during crystallization trials, and Carolyn Brown for help growing bacterial strains. We also thank Dr. Que Lan for helpful discussions.

REFERENCES

1. Kim, J., and Rees, D. C. (1992) *Nature* **360,** 553–560
2. Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J., and Rees, D. C. (1992) *Science* **257,** 1653–1659
3. Shah, V. K., Rangaraj, P., Chatterjee, R., Allen, R. M., Roll, J. T., Roberts, G. P., and Ludden, P. W. (1999) *J. Bacteriol.* **181,** 2797–2801
4. Rubio, L. M., Rangaraj, P., Homer, M. J., Roberts, G. P., and Ludden, P. W. (2002) *J. Biol. Chem.* **277,** 14299–14305
5. Paustian, T. D., Shah, V. K., and Roberts, G. P. (1990) *Biochemistry* **29,** 3515–3522
6. Homer, M. J., Paustian, T. D., Shah, V. K., and Roberts, G. P. (1993) *J. Bacteriol.* **175,** 4907–4910
7. Rubio, L. M., and Ludden, P. W. (2002) in *Nitrogen Fixation at the Millenium* (Leigh, G. J., ed) pp. 101–136, Elsevier Science, Amsterdam, The Netherlands
8. Homer, M. J., Dean, D. R., and Roberts, G. P. (1995) *J. Biol. Chem.* **270,** 24745–24752
9. Allen, R. M., Chatterjee, R., Ludden, P. W., and Shah, V. K. (1995) *J. Biol. Chem.* **270,** 26890–26896
10. Brandner, J. P., McEwan, A. G., Kaplan, S., and Donohue, T. J. (1989) *J. Bacteriol.* **171,** 360–368
11. Kuo, C. F., and Fridovich, I. (1988) *Anal. Biochem.* **170,** 183–185
12. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 67–71
13. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 916–924
14. Holden, H. M., and Rayment, I. (1991) *Arch. Biochem. Biophys.* **291,** 187–194
15. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D* **55,** 849–861
16. Terwilliger, T. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1937–1940
17. Roussel, A., and Cambillau, C. (1991) *Silicon Graphics Geometry Partners Directory*, Vol. 86, Silicon Graphics, Mountain View, CA
18. Tronrud, D. E. (1997) *Methods Enzymol.* **277,** 306–319
19. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
20. Navaza, J. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57,** 1367–1372
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
22. Lo Conte, L., Ailey, B., Hubbard, T. J., Brenner, S. E., Murzin, A. G., and Chothia, C. (2000) *Nucleic Acids Res.* **28,** 257–259
23. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542
24. Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H., and Westbrook, J. (2000) *Nat. Struct. Biol.* **7,** 957–959
25. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480
26. Holm, L., and Sander, C. (1999) *Nucleic Acids Res.* **27,** 244–247
27. Cort, J. R., Yee, A., Edwards, A. M., Arrowsmith, C. H., and Kennedy, M. A. (2000) *J. Struct. Funct. Genomics* **1,** 15–25
28. Cohen, G. H. (1997) *J. Appl. Crystallogr.* **30,** 1160–1161
29. Simon, H. M., Gosink, M. M., and Roberts, G. P. (1999) *J. Bacteriol.* **181,** 3751–3760
30. Gosink, M. M., Franklin, N. M., and Roberts, G. P. (1990) *J. Bacteriol.* **172,** 1441–1447
31. Schmid, B., Ribbe, M. W., Einsle, O., Yoshida, M., Thomas, L. M., Dean, D. R., Rees, D. C., and Burgess, B. K. (2002) *Science* **296,** 352–356
32. Kim, J., and Rees, D. C. (1992) *Science* **257,** 1677–1682
33. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
34. Esnouf, R. M. (1999) *Acta Crystallogr. Sect. D* **55,** 938–940
35. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296