Trudinger, P. (1970) *J. Bacteriol. 104*, 158–167.
Vega, J. M., & Kamin, H. (1977) *J. Biol. Chem. 252*, 896–909.
Walker, F. A., & Bolke, V. L. (1984) *J. Am. Chem. Soc. 106*, 6888–6898.
Yonetani, T., Drott, H. R., Leigh, J. S., Reed, G. H., Waterman, M. R., & Asakura, T. (1970) *J. Biol. Chem. 245*, 2998–3003.
Yonetani, T., Yamamoto, H., Erman, J. E., Leigh, J. S., & Reed, G. H. (1972) *J. Biol. Chem. 247*, 2447–2455.

Distinct Structural Features of the α and β Subunits of Nitrogenase Molybdenum–Iron Protein of *Clostridium pasteurianum*: An Analysis of Amino Acid Sequences†

Shu-Zhen Wang, Jiann-Shin Chen,* and John L. Johnson

Department of Anaerobic Microbiology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
Received October 6, 1987; Revised Manuscript Received December 17, 1987

ABSTRACT: Nitrogenase is composed of two separately purified proteins, a molybdenum–iron (MoFe) protein and an iron (Fe) protein. Structural genes (*nifD* and *nifK*) encoding α and β subunits of the MoFe protein of *Clostridium pasteurianum* (Cp) have been cloned and sequenced. The deduced amino acid sequences were analyzed for structures that could be related to the unique properties of the Cp protein, particularly its low capacity to form an active enzyme with a heterologous Fe protein. Cp *nifK* is located immediately downstream from Cp *nifD*, with the start codon of *nifK* overlapping by one base with the stop codon of *nifD*. An open reading frame following *nifK* was identified as *nifE*. The amino acid sequence deduced from *nifK* encompasses the partial amino acid sequences previously reported from the isolated β subunit. Cp *nifK* encodes a polypeptide of 458 amino acid residues (M_r 50 115) whose amino-terminal region is about 50 residues shorter than the otherwise conserved corresponding polypeptides from four other organisms. In contrast, Cp α subunit (*nifD* product) contains an additional stretch of 50 amino acid residues in the 380–430 region, which is unique to the Cp protein. It therefore appears that the combined size of the α and β subunits could be important to nitrogenase function. An analysis of the predicted secondary structure from the amino acid sequence of each subunit from three species (*C. pasteurianum*, *Azotobacter vinelandii*, and *Rhizobium japonicum*) further revealed structural features, including regions adjacent to some of the conserved cysteine residues, differentiating the Cp MoFe protein from others. These different regions may be further tested for correlation with distinct properties of Cp nitrogenase.

The reduction of N_2 to ammonia (nitrogen fixation) is carried out by a number of free-living and symbiotic bacteria, and the reaction is catalyzed by the enzyme nitrogenase [for a recent review, see Orme-Johnson (1985)]. Some organisms contain at least two types of nitrogenase (Bishop et al., 1986; Dilworth et al., 1987), but both types are composed of two separable protein components. The extensively characterized *conventional* nitrogenase (hereinafter referred to as nitrogenase) consists of an iron protein (Fe protein, component II, or dinitrogenase reductase) and a molybdenum–iron protein (MoFe protein, component I, or dinitrogenase), whereas the recently purified *alternative* nitrogenase from *Azotobacter* consists of a different iron protein (Robson et al., 1986; Hales et al., 1986a) and a vanadium–iron protein (Hales et al., 1986b; Eady et al., 1987).

All nitrogenase activities require the presence of both components, and the N_2 reduction site is believed to be located on the MoFe protein. The polypeptide of the Fe protein (a dimer) is encoded by *nifH* (or *nifH1* in *Clostridium pasteurianum*), whereas the polypeptides of the MoFe protein (an $\alpha_2\beta_2$ tetramer) are encoded by *nifD* (α subunit) and *nifK* (β subunit), respectively. Both the Fe protein and the MoFe protein contain iron–sulfur cluster(s), whereas the MoFe protein also contains molybdenum–iron–sulfur clusters that can be extracted as the iron–molybdenum cofactor [see Orme-Johnson (1985)].

In studies concerning the function of nitrogenase, an important aspect is the interaction between the Fe protein and the MoFe protein. Such interactions affect the enzyme's efficiency and relative reactivity toward alternative substrates (Smith et al., 1976; Emerich et al., 1981; Wherland et al., 1981). The physicochemical properties of Fe and MoFe proteins are highly conserved among all N_2-fixing bacteria so far studied, but significant differences also exist, especially between nitrogenase of *C. pasteurianum* (Cp) and the others. Specifically, nitrogenase components (the MoFe protein in particular) of *C. pasteurianum* have a distinctly low capacity to form an active hybrid enzyme with complementary components from other organisms (Smith et al., 1976; Emerich & Burris, 1978; Tsai & Mortenson, 1978; Emerich et al., 1978). *C. pasteurianum* nitrogenase is also less sensitive to H_2 as an inhibitor (Guth & Burris, 1983) and shows a higher specificity for nucleotides (Weston et al., 1983). Recently, a difference in both the relative reduction sequence of the redox

† This work was supported by USDA Competitive Grant 86-CRCR-1-2073 and by Project 6124400 from the Commonwealth of Virginia.
* Address correspondence to this author.

0006-2960/88/0427-2800$01.50/0 © 1988 American Chemical Society

centers and their measured midpoint potentials between *C. pasteurianum* and *A. vinelandii* MoFe proteins was observed when redox titrations were monitored by electron paramagnetic resonance (Morgan et al., 1987). These differences contrast the highly conserved subunit and metal compositions of nitrogenase from these organisms.

A more detailed structural comparison of nitrogenase from *C. pasteurianum* with those nitrogenases that are highly cross-reactive, such as the enzymes from *A. vinelandii* (Av) and *Bradyrhizobium japonicum* [*Rhizobium japonicum* (Rj)], could reveal structural features distinguishing Cp nitrogenase from the others and indicate regions potentially involved in interaction between component proteins. Amino acid sequences of the three polypeptides of nitrogenase are available for two highly cross-reactive organisms, *A. vinelandii* (Hausinger & Howard, 1982; Brigle et al., 1985) and *R. japonicum* (Fuhrmann & Hennecke, 1984; Kaluza & Hennecke, 1984; Thony et al., 1985). Complete amino acid sequences of the Cp Fe protein and the α subunit of Cp MoFe protein are also available (Tanaka et al., 1977; Chen et al., 1986; Hase et al., 1984; Wang et al., 1987). In this paper, we report the complete Cp *nifK* sequence as well as present a comparison of amino acid sequences from five organisms (six for the α subunit) and the predicted structures of the α and β subunits of MoFe proteins from *C. pasteurianum*, *A. vinelandii*, and *R. japonicum*.

Materials and Methods

Clostridium pasteurianum strain W5 was used throughout this study. Restriction enzymes were either from Bethesda Research Laboratories (BRL) or International Biotechnologies, Inc. (IBI). Calf intestinal alkaline phosphatase was from Boehringer Mannheim. *Exo*III and *Exo*VII nucleases were obtained from BRL. ^{125}I was from Amersham, and ^{35}S-labeled dATP was from New England Nuclear Corp.

The recombinant plasmid pCP124 (Figure 1) was constructed as follows: The *Aha*III–*Eco*RI fragment (about 0.5 kb) containing a portion of *nifD* in pCP114 was digested out by restriction enzymes *Aha*III (which cuts after base 14 of the *nifD* coding region) and *Eco*RI (which is the cloning site of pCP114). The digestion mixture was separated by electrophoresis in 0.7% agarose gel. The region containing the 0.5-kb *nifD* fragment was cut from the gel, the *nifD* DNA was isolated by electroelution following the procedure provided by IBI. The purified DNA fragment was labeled with ^{125}I (Selin et al., 1983) and used as a probe to screen for the rest of the *nifD* gene.

High molecular weight DNA from *C. pasteurianum* W5 was digested with several restriction enzymes, and the resultant fragments were analyzed by agarose gel electrophoresis followed by Southern transfer (Southern, 1975) and hybridization with the 0.5-kb *nifD* probe. A 2.3-kb *Aha*III band reacted positively with the probe. DNA fragments in the 2.3-kb region were electroeluted from a gel slice and cloned into the plasmid pUC18 that had been digested with *hin*cII and treated with calf intestinal alkaline phosphatase. The recombinant plasmid was used to transform *Escherichia coli* strain DH1 following the procedure of Hanahan (1983). Positive clones were obtained by membrane hybridization of transformants with ^{125}I-labeled *nifD* DNA probe (described above).

For DNA sequencing, the 2.3-kb DNA fragment was digested out of a positive clone (pCP124) and subcloned into M13 mp18 and mp19. Replicating form (RF) DNA was isolated from M13 phages containing 2.3-kb inserts. The RF DNA was subjected to sequential digestion with *Exo*III and *Exo*VII nucleases as described by Yanisch-Perron et al. (1985). Phages with different sizes of inserts were selected to generate templates for sequencing. DNA sequencing was accomplished by using the dideoxy method (Sanger et al., 1977) and ^{35}S-labeled dATP as described by Biggin et al. (1983).

DNA sequencing revealed that the 2.3-kb insert in pCP124 contains the first 229 amino acids of the *nifK* gene. This piece of DNA was digested out of pCP124 by the restriction enzyme *Hin*dIII, which cuts 38 nucleotides from the beginning of the *nifK* gene. The DNA generated was labeled with ^{125}I and used to screen for DNA fragments containing the rest of the *nifK* gene. The cloning procedures used were the same as those for the *nifD* gene. Positive clones with 2.2-kb inserts were isolated. The strategy used for sequencing the *nifDK* region in pCP124 was also used to sequence the *nifK* region in pCP134.

The predicted secondary structures (Chou & Fasman, 1978, 1979) and the hydropathic index (Kyte & Doolittle, 1982) of amino acid sequences were obtained by using the PEPPLOT (Chou/Fasman) and PEPPLOT (Hydrophobic Moment) programs of the Sequence Analysis Software Package (Version 4) of the University of Wisconsin Genetics Computer Group.

Results and Discussion

Organization of C. pasteurianum nifH1, nifD, and nifK Genes. In *C. pasteurianum*, the nitrogenase structural genes (*nifH1*, *nifD*, and *nifK*) are contiguous (Figure 1). This organization is conserved among all N_2-fixing organisms except the slow-growing rhizobia in which *nifH* is separated from *nifDK* (Fischer & Hennecke, 1984) and the vegetative cells of *Anabaena* in which *nifHD* are separated from *nifK* (Golden et al., 1985).

The 2.3-kb *Aha*III fragment carried on pCP124 contains coding regions for nearly the entire *nifD* and the first 229 amino acids of *nifK*. The assignment of *nifK* to this region was based on a comparison of the deduced amino acid sequence (Figure 2) with the partial sequence (underlined in Figure 2) determined from the isolated β subunit of Cp MoFe protein (Hase et al., 1984). The remaining portion of *nifK* is located on a 2.2-kb *Hin*dIII fragment carried on pCP134 (Figure 1).

The coding region of *nifK* started from the third base of the stop codon for *nifD*, giving a junction sequence or 5′-UAAUG-3′ (Figure 2). A putative ribosome-binding site (GAGG) was present between −13 and −10 nucleotides (in the coding region of *nifD*) from the putative start codon (AUG) for *nifK*. The putative ribosome-binding sequence GAGG is also present in a similar region preceding *C. pasteurianum nifH1* and *nifD* (Chen et al., 1986).

A translational coupling mechanism (Oppenheim & Yanofsky, 1980) has been proposed for overlapping translational stop and start signals, which ensures the coordinated production of polypeptides contained in equimolar amounts in an oligomeric enzyme or enzyme complex. However, an overlap between *nifD* and *nifK* has not been found in the other organisms studied so far (see references in Figures 3 and 4 for nucleotide sequences). The distance between *nifD* and *nifK* of the other organisms ranges from 18 nucleotides in *R. japonicum* (Kaluza & Hennecke, 1984) to 199 nucleotides in *Anabaena* 7120 (Golden et al., 1985) between the first stop codon (up to three in tandem) and the start codon. It is interesting to note that in *A. vinelandii* the apparent ribosome-binding site for *nifN* is partially overlapped by the stop codon for *nifE*, and *nifEN* products are proposed to function as a complex (Brigle et al., 1987a). [Sequence homologies are present between *nifD* and *nifE* and between *nifK* and *nifN* (Brigle et al., 1987a).] The proximity between *nifD* and *nifK*

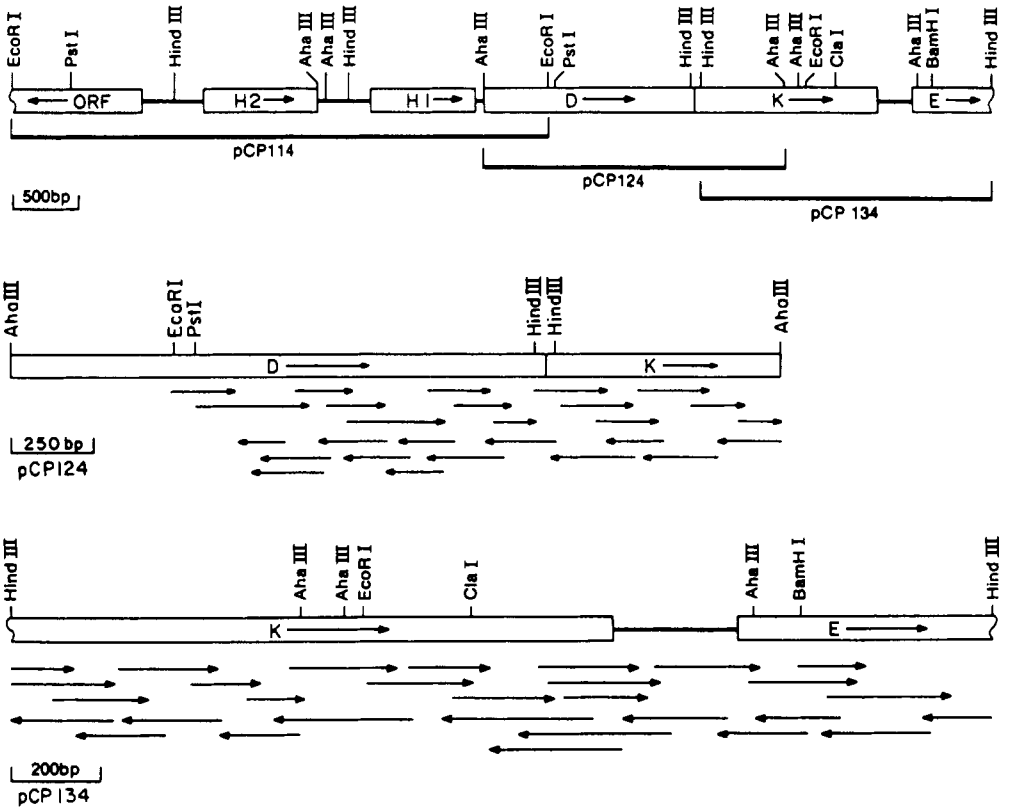

FIGURE 1: Maps of *C. pasteurianum nifH2H1DKE* region, showing coding regions, restriction sites, and sequencing strategy. These *nif* genes are located on three overlapping restriction DNA fragments: a 4-kb *Eco*RI fragment carried on pCP114 (Chen et al., 1986), a 2.3-kb *Aha*III fragment carried on pCP124, and a 2.2-kb *Hin*dIII fragment carried on pCP134. *nifH1* encodes the isolated iron protein. *nifD* and *nifK* encode, respectively, the α and β subunits of the MoFe protein. Arrows in the boxes show directions of transcription. Other arrows indicate the extent of sequencing from each *Exo*III- and *Exo*VII-generated fragment from the insert in pCP124 and pCP134 and the strand on which sequencing was performed. ORF, open reading frame; bp, base pairs.

in *C. pasteurianum* may be related to the shortness of the N-terminal region of the Cp β subunit (see below).

An open reading frame (ORF) was observed 279 bp downstream from *nifK* (Figure 1). Two regions of dyad symmetry (underlined with arrows, with the center marked by a dot) were found between Cp *nifK* and the following ORF. These stem and loop structures could serve as a transcription termination signal (Rosenberg & Court, 1979) for the *nifH1DK* operon. The cloned portion of the ORF encodes 188 amino acid residues (data not shown) that are highly homologous to the *A. vinelandii* (Av) *nifE*-encoded sequence (Brigle et al., 1987a). Between residues 51 and 150 of the deduced Cp sequence, 48 residues are identical with those of the deduced Av *nifE* product. Because of the sequence homology and the relative position of the ORF to *nifK*, we have assigned this ORF as Cp *nifE*. Thus, in *C. pasteurianum*, *nifK* and *nifE* are not separated by the *nifY* gene (*nifY* may be absent in *C. pasteurianum* or is located elsewhere in the genome).

Nucleotide Sequence and the Encoded Amino Acid Sequence of nifK. The complete nucleotide sequence of the noncoding strand of the *nifK* region is shown in Figure 2. The deduced amino acid sequence for the *nifK* product (β subunit) has 458 residues, which gives a M_r of 50 115. The deduced N- and C-terminal residues are consistent with the partial sequences determined from the isolated protein (Hase et al., 1984), indicating that no peptide bond is cleaved following translation. The C-terminal arginine was not detected in an earlier study using hydrazinolysis (Chen et al., 1973), which is consistent with the fact that arginine is completely destroyed during the hydrazinolytic reaction (Braun & Schroeder, 1967). In *A. vinelandii*, the methionine from the start codon is not present in the isolated β subunit (Lundell & Howard, 1978).

Amino acid residues that were previously determined from the protein are underlined (Figure 2) and show a good agreement with the DNA-derived sequence. There are, however, eight discrepancies between the DNA-derived sequence and the protein-derived partial sequence (with the protein-derived residue in parentheses): residue 89, Lys (Pro); 136, His (Met); 141, Ser (Gly); between 152 and 153 (Ser); 324, Ile (Ala); 326, Ala (Leu); 327, Leu (Lys); 351, Gln (Glu).

C. pasteurianum nifH1DK genes consist of 1267 codons (Chen et al., 1986; Wang et al., 1987; Figure 2), and an asymmetric pattern of codon usage is seen (data not shown). The most asymmetric ones are Glu, 99 (97%) of the 102 residues using GAA; Cys, 21 (96%) of the 22 residues using UGU; and Arg, 37 (93%) of the 40 residues using AGA. Codons used in Cp *nifH1DK* generally avoid the presence of two or more G and/or C in a coding triplet, which is consistent with the low G + C content (26–28%) of Cp DNA. Since nitrogenase is an abundantly expressed enzyme (Zumft & Mortenson, 1973), this codon usage pattern likely reflects the distribution of tRNAs in this organism (Ikemura, 1981).

Comparison of Amino Acid Sequences and the Predicted Secondary Structures of MoFe Proteins. The elucidation of a structure–function relationship is an important aspect in studies of biologically active molecules. Useful information concerning the structure–function relationship can emerge when sequences of a group of highly conserved proteins are compared and where distinct structural features of a protein can be correlated with unique biological properties of that protein.

The amino acid sequence of the α subunit has been deduced from *nifD* gene sequences determined for six organisms (see

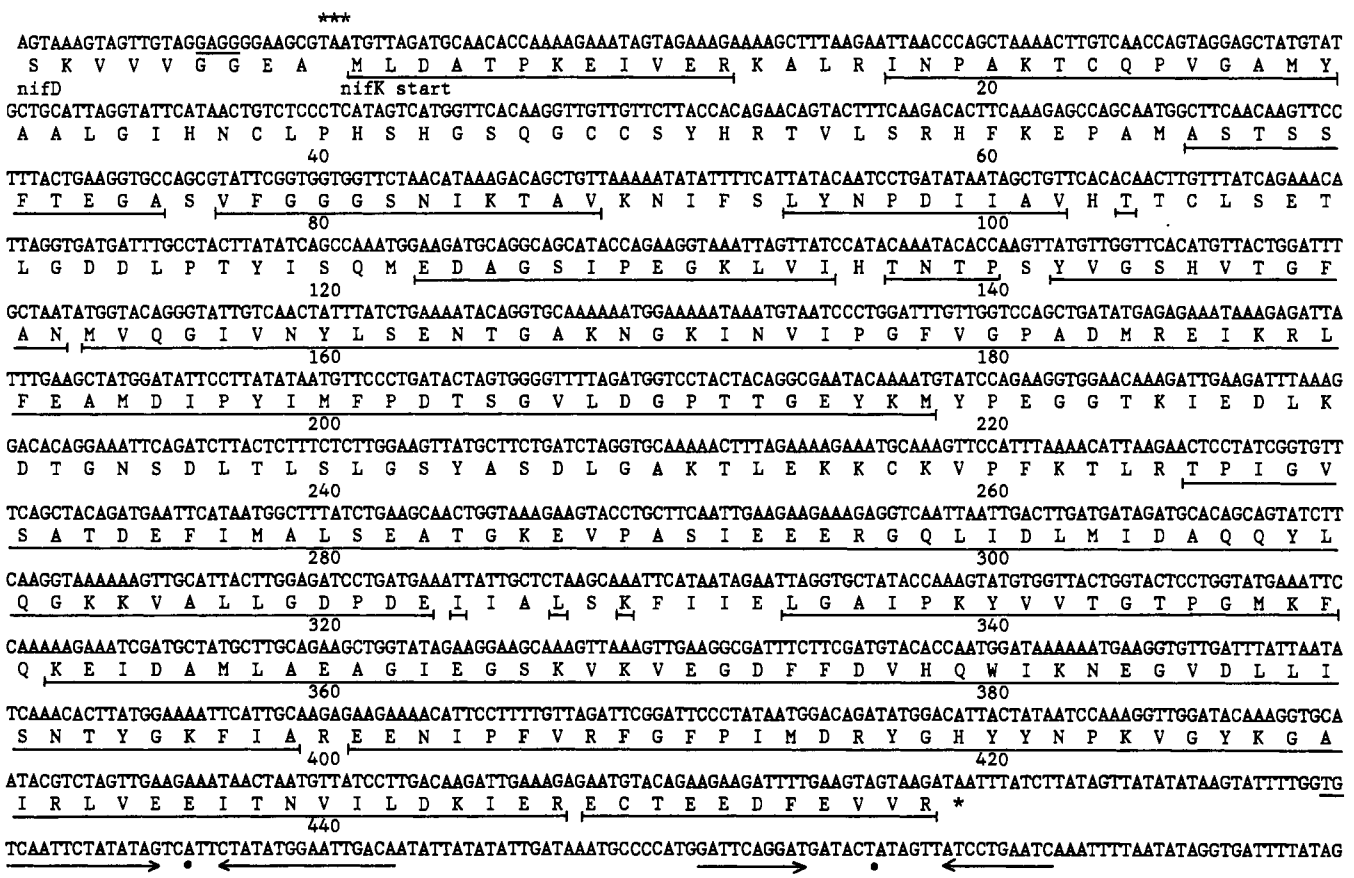

FIGURE 2: Nucleotide sequence and the encoded amino acids of *nifK* and its flanking regions of *C. pasteurianum*. The DNA strand shown is that identical with mRNA. The stop codon (TAA) for *nifD* and the start codon (ATG) for *nifK* overlapped by one base. The underlined amino acid residues have been determined from the isolated protein (Hase et al., 1984). The inverted repeats are indicated by arrows, with the center of symmetry marked by a dot.

Figure 3 and references cited therein), whereas the amino acid sequence of the β subunit has been deduced from *nifK* gene sequences determined for five organisms (see Figure 4 and references cited therein). They encompass symbiotic organisms [*R. japonicum* (Rj), *Parasponium Rhizobium* (PR), and *Rhizobium* IRc78], an aerobic, free-living organism [*A. vinelandii* (Av)], an anaerobic, free-living organism [*C. pasteurianum* (Cp)], and a photosynthetic organism (*Anabaena*). They include both Gram-positive (Cp) and Gram-negative organisms and span a range of DNA G + C contents from about 28 mol % (Cp) to near 70 mol %. The amino acid sequences of both subunits from this wide range of organisms are highly conserved (see below). Thus, the conserved regions are likely required for general nitrogenase functions (catalysis, component interaction, and stability).

C. pasteurianum nitrogenase is uniquely different from the others, especially in its low activity in heterologous reconstitutions, which provides an approach for identifying structures that are involved in determining the compatibility between nitrogenase components. Unique structural features are expected to be present in both Fe protein and MoFe protein of *C. pasteurianum* to account for their distinct specificity for a compatible complement. Distinct structural features of the Cp Fe protein have been reported (Hausinger & Howard, 1982; Chen et al., 1986), which include a shorter C-terminal region and the lack of a positive charge (Lys) that is conserved in the extended C-terminal region of Fe proteins from eight other organisms.

A recent study showed that the EPR and redox properties of the protein-associated iron–molybdenum cofactor (FeMoco) differ between *C. pasteurianum* and *A. vinelandii* (Morgan et al., 1987), which was attributed to differences in the protein environment for FeMoco in the two organisms because the isolated FeMoco from both organisms showed similar biological activity and spectroscopic properties. Thus, some distinct structures of Cp MoFe protein could be related to the binding of or interaction with FeMoco, which are further related to the functional differences.

We have analyzed the primary and the predicted secondary structures of MoFe protein subunits of *C. pasteurianum* and the others. The MoFe protein subunits of *A. vinelandii* (Av) and *R. japonicum* (Rj) were chosen for a detailed comparison with the Cp subunits because the Av and Rj proteins are highly active in reciprocal heterologous reconstitutions (Emerich & Burris, 1978). The regions not conserved in the α and β subunits of *A. vinelandii* and *R. japonicum* are presumed not critical to component interaction, i.e., where significant differences (e.g., changes causing charge reversions or changes in secondary structures) do not prevent effective component interactions. We can thus focus on (1) amino acid residues and secondary structures that are conserved in all organisms other than *C. pasteurianum* and (2) amino acid residues and secondary structures that are conserved at least in Av and Rj but not in Cp.

In the following, the conserved structures among all MoFe protein subunits are described first, which is followed by a description of structures unique to the *C. pasteurianum* MoFe protein. An apparent homology between the α and β subunits is also discussed. In order to establish the proposed structure–function relationship, site-directed mutagenesis [e.g., Brigle et al. (1987b)] can be very useful when applied to selected regions that are predicted to be relevant.

(*a*) *Conserved Features of α Subunits.* Amino acid sequences of the α subunit of six organisms are aligned in Figure

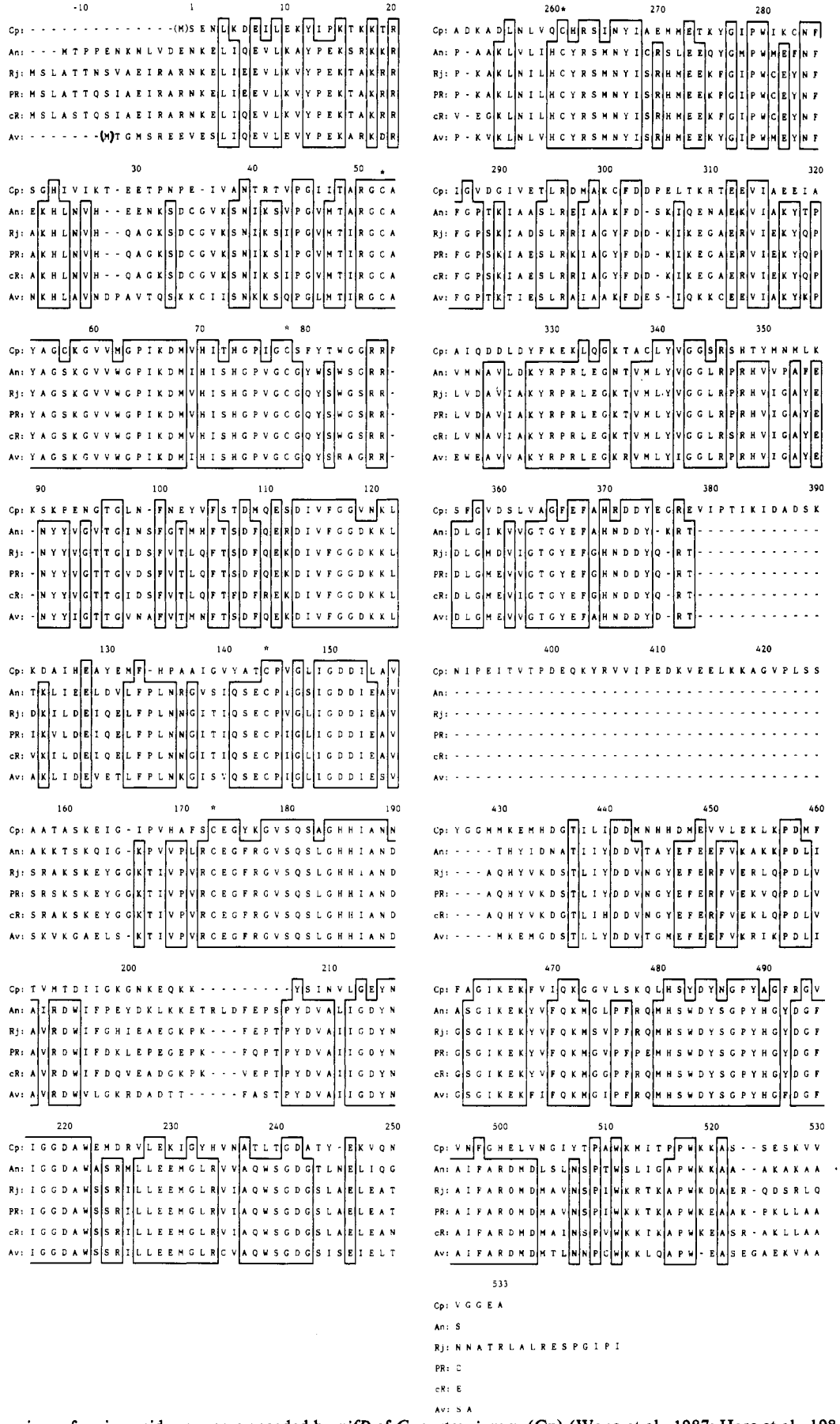

FIGURE 3: Comparison of amino acid sequences encoded by *nifD* of *C. pasteurianum* (Cp) (Wang et al., 1987; Hase et al., 1984), *Anabaena* 7120 (An) (Golden et al., 1985; Lammers & Haselkorn, 1983), *R. japonicum* (Rj) (Kaluza & Hennecke, 1984), *Parasponia Rhizobium* (PR) (Weinman et al., 1984), *Rhizobium* strain IRc78 (cR) (Yun & Szalay, 1984), and *A. vinelandii* (Av) (Brigle et al., 1985; Lundell & Howard, 1978, 1981; D. R. Dean, personal communication). The numbering refers to *C. pasteurianum* α subunit. Amino acid residues conserved in all six organisms or in the five organisms excluding *C. pasteurianum* are boxed. The five conserved cysteine residues are marked by asterisks above them. Dashes signify gaps introduced to align the sequences. Parenthesized methionines are not present in the mature protein.

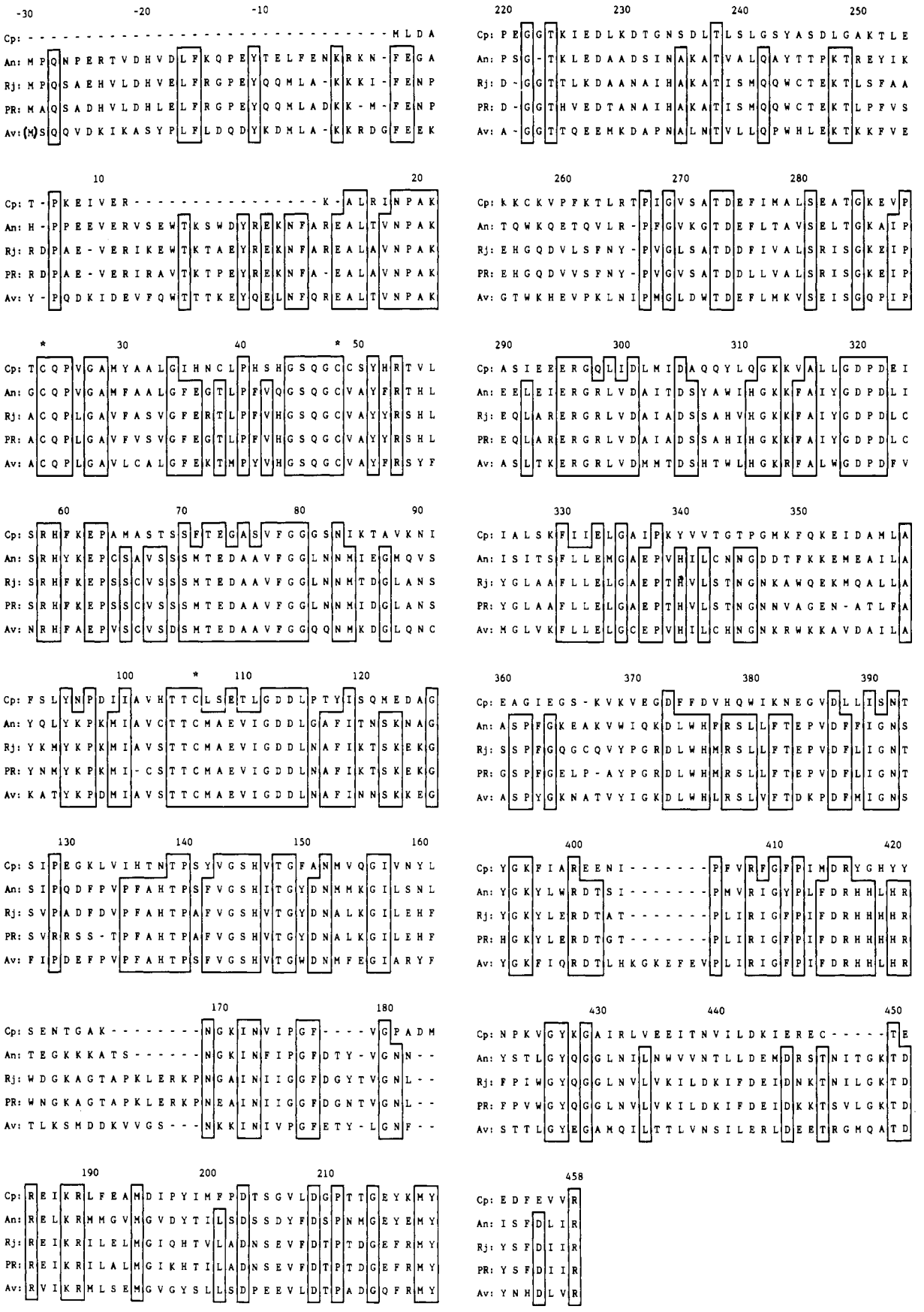

FIGURE 4: Comparison of amino acid sequences encoded by *nifK* of *C. pasteurianum* (Cp) (this study; Hase et al., 1984), *Anabaena* 7120 (An) (Mazur & Chui, 1982), *R. japonicum* (Rj) (Thony et al., 1985), *Parasponia Rhizobium* (PR) (Weinman et al., 1984), and *A. vinelandii* (Av) (Brigle et al., 1985; Lundell & Howard, 1979, 1981). The numbering refers to *C. pasteurianum* β subunit. Amino acid residues conserved in all five organisms or in the four organisms excluding *C. pasteurianum* are boxed. The three conserved cysteine residues are marked by asterisks above them. See Figure 3 for other notations.

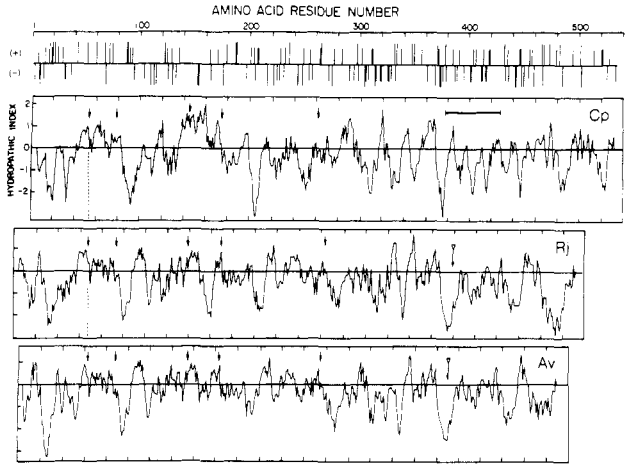

FIGURE 5: Distribution of conserved positive and negative charges in *nifD*-encoded polypeptide of six organisms (see Figure 3 for organisms) and a comparison of the hydropathic index of *nifD*-encoded polypeptides from *C. pasteurianum* (Cp), *R. japonicum* (Rj), and *A. vinelandii* (Av). The numbering refers to the *C. pasteurianum* protein. In the graph showing charge distribution, longer bars indicate positions for conserved positive charges (from Arg, Lys, or His) or negative charges (from Asp or Glu), whereas shorter bars indicate those present in *C. pasteurianum* but not conserved. In hydropathy plots, arrows show the location of the five conserved Cys residues, and the dashed line marks the reference point for alignment. The position of the stretch of 50 additional amino acid residues found only in Cp (between residues 380 and 430; see Figure 3) is indicated by a horizontal bar, and the corresponding positions in Rj and Av sequences are indicated by arrows with open triangles.

3. To achieve a better alignment, gaps were introduced into these sequences, most notably in the 380–430 region where an extra stretch of 50 amino acids is present only in Cp. (Structural features of this stretch of 50 residues are discussed later.) Among the six sequences, 159 amino acid residues (29.8% of the Cp sequence) are conserved, whereas an additional 132 amino acid residues are conserved, resulting in a 56–59% homology, in the α subunit of the five Gram-negative organisms. However, the predicted secondary structure of these proteins is more conserved than the amino acid sequences because of the many conservative substitutions. For instance, despite relatively low sequence conservation between residues 30 and 50 (numbering refers to the Cp sequence unless otherwise specified), the predicted secondary structure of this region is largely homologous among Cp, Rj, and Av, where the α helix is absent and the positions of β turns (31–34, 49–52) and β sheets (33–37, 43–47) are conserved (data not shown).

Figure 5 shows the distribution of acidic (−) and basic (+) amino acids in the Cp α subunit and a comparison of hydropathicity of Cp, Av, and Rj α subunits. The conserved acidic and basic amino acids (Asp, Glu, Arg, His, Lys) are indicated by longer bars, and the arrows indicate the positions of five conserved cysteine residues (residues 52, 78, 144, 173, 261). When the 380–430 region of the Cp sequence (marked by a bar) is excluded, the hydropathy plots show a remarkable similarity among the three α subunits. The hydropathy plots suggest that residues 1–38 and 82–100 constitute very hydrophilic regions. Residues 39–81 constituted a hydrophobic region, but the two conserved cysteine residues (52 and 78) were located at points that were nearly amphiphilic. In oligomeric proteins, relatively hydrophobic regions can be located at the surface of subunits when these surfaces are within subunit interfaces (Holbrook et al., 1975). If the two conserved thiols within this region serve as ligands for a prosthetic group, such as an FeS cluster, the prosthetic group may need to be located near the surface of the subunit, which would facilitate electron transfer between this redox center and another redox center or a mediating group situated on another subunit. The presence of three conserved prolines (residues 44, 64, and 75; Pro-44 and Pro-75 are in predicted turns) near Cys-52 and Cys-78 also suggest that these residues are located near the surface of the subunit. Thus, although some potential ligating thiols occur in a generally hydrophobic region, they could still be located near the subunit surface, but they will be within the subunit–subunit interface. Such a location could shield this type of prosthetic group of the native protein from solvents [see a pertinent observation in reference 107 of Orme-Johnson (1985)].

The abundance and distribution of basic amino acids in the α subunits could be significant. The deduced *nifD* products have an excess of basic amino acids, with the excess ranging between 2 (*C. pasteurianum*) and 16 (*Anabaena* 7120). The abundance of basic amino acids in the MoFe protein α subunit distinguishes it from related redox proteins, i.e., Fe proteins (Chen et al., 1986), ferredoxins (Yasunobu & Tanaka, 1973), and flavodoxins (Drummond, 1985), which all have a large excess of acidic amino acids. However, it is the spatial distribution of these ionizable residues that is more pertinent to the interaction between subunits or nitrogenase components. It is noteworthy that within the first 100 amino acid residues, there is an excess of 10 basic residues among conserved residues (9–13 when all residues are counted). Between residues 41 and 87 (encompassing Cys-52 and Cys-78), the charge property of all nine ionizable residues (with a potential net charge of +8) is conserved. The positive charges could have a role in orienting the region toward a negatively charged region containing additional ligands for the prosthetic group. The cluster of charged groups in the first 30 residues suggests a surface location of the N-terminal region, and the patch of positive charges between residues 11 and 27 may serve to orient an interacting protein, such as another subunit of MoFe protein or Fe protein. In this regard, there is an excess of negative charges around Cys-94 and Cys-129 in the Fe protein; the two Cys residues have been proposed as ligands for an FeS cluster (Hausinger & Howard, 1983).

(*b*) *Conserved Features of β Subunits.* The amino acid sequences of the β subunit of five organisms are aligned in Figure 4. Extensive homologies can be seen throughout the entire length of the molecule, except for the first 13 residues of the Cp sequence that cannot match the first 60 residues of the other four sequences. There are 121 conserved residues among the five sequences (26.4% of the Cp sequence). An additional 88 residues are conserved in the four organisms excluding *C. pasteurianum*. No sequence homology occurs in the 250–260 and the 350 regions; however, homologous α helices are predicted for these regions of the Cp, Rj, and Av sequences (data not shown).

Figure 6 shows the distribution of acidic (−) and basic (+) amino acids in Cp β subunit and a comparison of hydropathicity of Cp, Rj, and Av β subunits. The conserved acidic and basic residues are indicated by longer bars, and the arrows indicate the position of three conserved cysteine residues (residues 23, 48, 106). The hydropathy plots of the three β subunits, like those of the α subunits, show a remarkable similarity between Cys-23 (first conserved Cys residue) and the C terminus. A similar folding of the polypeptides may thus be anticipated. Nevertheless, the predicted α-helical, β-sheet, and β-turn regions are not as conserved in β subunits as in α subunits of these three organisms (data not shown).

For the three conserved Cys residues, Cys-23 is near the beginning of a predicted α helix (26–32) and Cys-48 and

Table I: Comparison of Size of α and β Subunits of Nitrogenase MoFe Protein from Five Organisms

organism	α subunit		β subunit		α + β	
	no. of residues	M_r	no. of residues	M_r	no. of residues	M_r
C. pasteurianum	533	58 990	458	50 115	991	109 105
Anabaena 7120[a,b]	497	55 902	512	57 583	1009	113 485
R. japonicum[a,b]	515	57 918	518	57 428	1033	115 346
P. Rhizobium[a,b]	500	56 142	513	56 538	1013	112 680
A. vinelandii[a]	491	55 062	522	59 325	1013	114 387

[a] See Figures 3 and 4 for references. These amino acid sequences were deduced from DNA. [b] The N-terminal methionine residue from the start codon is included in the calculation.

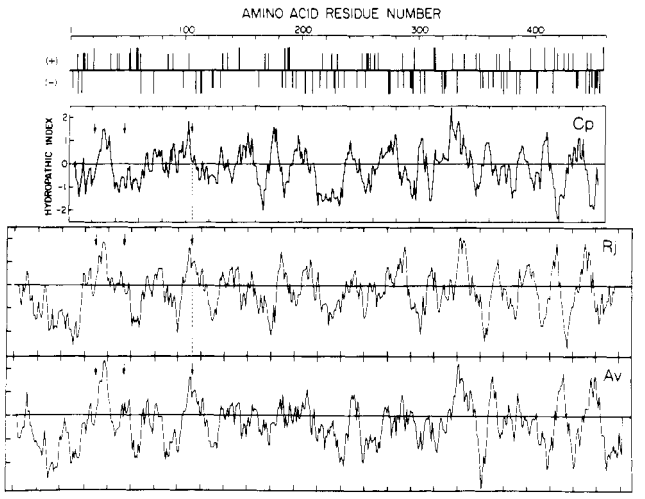

FIGURE 6: Distribution of conserved positive and negative charges in *nifK*-encoded polypeptide of five organisms (see Figure 4 for organisms) and a comparison of the hydropathic index of *nifK*-encoded polypeptides from *C. pasteurianum* (Cp), *R. japonicum* (Rj), and *A. vinelandii* (Av). The numbering refers to the *C. pasteurianum* protein. See Figure 5 for other notations.

Cys-106 are, respectively, at the beginning (48–55) and the end (97–105) of two predicted β-sheets (data not shown). All three conserved Cys residues occur in hydrophilic or amphiphilic regions (Figure 6) and are close to conserved proline residues, which are usually located at the surface of proteins. If these Cys residues serve as ligands for redox prosthetic groups, these redox centers are likely to be located near the surface of the subunit and also near the surface of the holo MoFe protein.

A near-surface location of the prosthetic group(s) in the holo MoFe protein would suggest that the Fe protein interacting site is on the β subunit. However, Zumft et al. (1980) have suggested, on the basis of an immunologic study, that the binding site for the Fe protein might be located on the α subunit *or* that the antibody-binding site on the α subunit is near the catalytic site. It seems possible that redox centers of both subunits of the MoFe protein are spatially close to each other, and when electron transfer occurs between the Fe protein and the MoFe protein, the Fe protein may be in contact with both the α and the β subunits, although only one subunit (such as the β subunit) is the electrn acceptor.

Distinct Features of Nitrogenase MoFe Protein of C. pasteurianum. The subunits of *C. pasteurianum* MoFe protein are distinct in their size, their charge properties (especially a sensitivity to pH above 7 and the occurrence of unique histidine residues near the conserved cysteine residues), and other differences in the primary and the predicted secondary structures. These differences are described in the following sections.

(*a*) *Size of α and β Subunits.* Among sequenced α and β subunits (Table I), the *C. pasteurianum* α subunit is the largest (larger by 18–41 amino acid residues), whereas the Cp β subunit is the smallest (smaller by 54–65 residues). In *C. pasteurianum* the α subunit (533 residues) is larger than the β subunit (458 residues) by 75 residues (about 18% by M_r), whereas in the other four organisms the α subunit is 3–31 residues smaller than the β subunit. (The M_r of the deduced Rj α subunit is actually 1% higher than that of Rj β subunit, although the α subunit is three residues shorter.) That Cp α subunit is larger than Cp β subunit was previously demonstrated (Hase et al., 1984). Despite variations in the size of the two subunits of MoFe protein in these organisms, the combined size of the α subunit plus the β subunit is fairly constant (1009–1033 residues for the other four organisms; 991 for *C. pasteurianum*), which is consistent with the measured M_r of MoFe proteins. Perhaps the size of MoFe protein is conserved in order to accommodate the Fe protein in a productive way. It was observed that Cp Fe protein can form a tight but inactive complex with the Av MoFe protein, but Av Fe protein does not bind Cp MoFe protein (Emerich et al., 1978). Since both Cp Fe protein (Chen et al., 1986) and Cp MoFe protein (this study) are smaller than the corresponding proteins from other organisms, it may be postulated that in Cp MoFe protein, the Fe protein binding site(s) is (are) a smaller domain that can only accommodate the smaller Cp Fe protein. (With a reversed size relationship between the α and β subunits of Cp MoFe protein, the spatial arrangement of functional groups surrounding the domain the MoFe protein could also differ between *C. pasteurianum* and the other organisms. Parallel changes must also occur in Cp Fe protein to accomodate the situation.) Thus, the smaller Cp Fe protein could fit into the conceivably larger Av MoFe protein domain but the interaction could be such (e.g., too strong) that the normally ensuing reactions cannot occur, whereas the larger Av Fe protein simply cannot fit into the smaller domain of the Cp MoFe protein.

(*b*) *Gross Charge Properties.* The calculated pI's of the α and β subunits of Cp MoFe protein are 6.06 and 5.86, respectively, which are much lower than the corresponding pI's of the Rj and Av subunits (7.17 and 6.34 for the Rj subunits; 6.55 and 6.34 for the Av subunits). The lower pI's of the Cp subunits appear to correlate with the acidic intracellular pH (about 6.5 when the external pH is 6) of *C. pasteurianum* (Riebeling et al., 1975) and the lower optimal pH (6.6–6.8) for Cp nitrogenase than for nitrogenases of other organisms (Emerich & Burris, 1978). The activity of Cp nitrogenase is about 50% lower at pH 7.4 than at pH 6.6. Interestingly, combinations of Cp MoFe protein with Fe protein of *Bacillus polymyxa* or *Klebsiella pneumoniae* showed little or no activity at pH 7.4, whereas 8–38% of the homologous activity was observed at pH 6.6 (Emerich & Burris, 1978). The results indicate that the *function* of Cp MoFe protein is sensitive to pH above 7, and the decreased activity could result from decreased binding between nitrogenase components because of a change in the net charge of group(s) at the binding site (Emerich & Burris, 1978).

Between pH 6.5 and 7.5, histidine is responsible for changes

α subunit

```
                          52                  78                        144
                          *                   *                         *
Cp  H2N-16aa-KKTR-29aa-RxCxxxG-6aa-G-9aa-HGxxGC-7aa-RR-23aa-ExxxVFGG-25aa-CxxxxIGDDI

Rj  H2N-31aa-AKRR-29aa-RxCxxxG-6aa-G-9aa-HGxxGC-7aa-RR-22aa-ExxxVFGG-26aa-CxxxxIGDDI

Av  H2N-23aa-RKDR-31aa-RxCxxxG-6aa-G-9aa-HGxxGC-7aa-RR-22aa-ExxxVFGG-26aa-CxxxxIGDDI

                        T    T             T    T                    T    T
                      (-3)  (+2)           (-3)  (+2)                 (0)  (+6)
```

β subunit

```
                          23                  48                        106
                          *                   *                         *
Cp          H2N-20aa-KxCxxxG-6aa-G-8aa-HGxxGC-9aa-RH-13aa-ExxxVFGG-25aa-CxxxxLGDDL

Rj  H2N-25aa-KKK-38aa-KxCxxxG-6aa-G-8aa-HGxxGC-9aa-RH-13aa-ExxxVFGG-25aa-CxxxxIGDDL

Av  H2N-24aa-KKR-39aa-KxCxxxG-6aa-G-8aa-HGxxGC-9aa-RH-13aa-ExxxVFGG-25aa-CxxxxIGDDL

                       T    T             T                    T         T
                     (-5)  (+1)           (-3)                 (-3)      (+6)
```

FIGURE 7: Homologous regions between α and β subunits of nitrogenase MoFe protein. The comparison is based on amino acid sequences of subunits from *C. pasteurianum* (Cp), *R. japonicum* (Rj), and *A. vinelandii* (Av). The conserved cysteine residues are marked by asterisks above them, with the numbering referring to the *C. pasteurianum* protein. aa or x denotes an amino acid residue. T marks the first position of a four-residue β turn, with its relative position to the nearest cysteine noted in parentheses.

in the charge properties of proteins. In this respect, the α and β subunits of Cp MoFe protein have a number of unique His residues occurring near the conserved Cys residues: His residues 127, 134, 169, and 262 for the α subunit and His residues 36, 41, 52, and 103 for the β subunit. (Others occur at positions 234, 434, 445, 446, and 500 in the α subunit and positions 136 and 378 in the β subunit.) Some of these histidine residues and their positive charges could be an important determinant for compatibility between Cp MoFe protein and Fe proteins. The location of these His residues probably encompasses the Fe protein binding region(s) of Cp MoFe protein.

(*c*) *Differences Surrounding Conserved Cysteine Residues.* Among the five conserved Cys residues in α subunit, Cys-52 and Cys-78 occur in a region with very conserved amino acid sequences (hence, the charge property and the predicted secondary structure in this region are essentially identical among the six organisms compared in Figure 3). However, the environment of the remaining three conserved Cys residues in Cp α subunit differs between Cp and the others: In Cp, Cys-144 occurs in a hydrophobic region and Cys-173 occurs in an amphiphilic region, whereas in Rj and Av they occur in amphiphilic and hydrophobic regions, respectively. The residue immediately preceding Cys-144 or Cys-173 is charged in the other organisms, whereas in Cp it is not (Figure 3). In Cp, Cys-261 has a different charge environment: Asp-255 replaces a Lys, Gln-260 replaces a His, His-262 replaces a Tyr, and Glu-270 replaces an Arg (Figure 3). Also, a predicted β turn begins at His-262 instead of at Ser-264. If Cys-261 (Cys-275 in Av) is a ligand for FeMoco as suggested by Brigle et al. (1987a), the different charge environment and the closer β turn may contribute to the observed different redox and spectroscopic properties of FeMoco in Cp MoFe protein (Morgan et al., 1987).

In the β subunit of *C. pasteurianum*, Cys-23 and Cys-48 occur in a region with 12 potential positive charges and no interspersing negative charges, whereas negative charges (−4 to −5) are present in this region in the other four sequences. Between Cys-23 and Cys-48, a negative charge from a Glu is conserved in the other four organisms, whereas *C. pasteurianum* has a basic residue (His-36) in its place. Cys-106

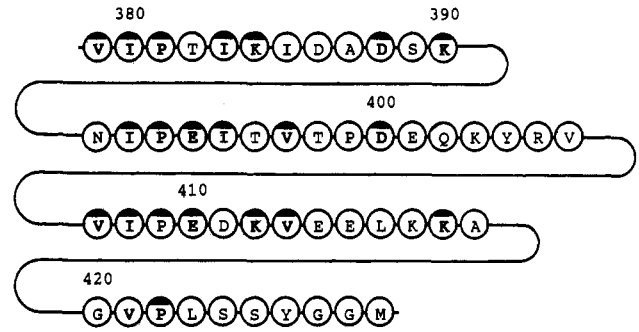

FIGURE 8: Apparent periodicity in the additional stretch of amino acid sequence unique to *C. pasteurianum* α subunit. The amino acid sequence is aligned to show residues occurring in apparently repetitive sequences. Conserved residues are marked by partially filled circles.

occurs in an amphiphilic region in *C. pasteurianum*, whereas in Rj and Av the corresponding cysteine residues occur in hydrophobic regions (Figure 6). It may be significant that a concentration of conserved positive charges is present in a region 30–40 residues upstream from the first conserved Cys residue in all α and β subunits except for the Cp β subunit whose N terminus does not extend into this length (Figures 3, 4, and 7).

These differences near the conserved Cys residues in Cp α and β subunits may be related to the distinct biochemical properties of Cp MoFe protein, assuming that these thiols are ligands for the prosthetic groups.

(*d*) *Unique Amino Acid Sequences.* The unique sizes of subunits of Cp MoFe protein can be attributed to two particular regions of the polypeptides: One is a stretch of about 50 amino acids in the 380–430 region found only in the Cp α subunit (Figure 3), whereas the other is the shortened N-terminal region (by about 50 amino acids) within the Cp β subunit (Figure 4).

It is interesting to note that this additional stretch of about 50 amino acids in Cp α subunit shows a periodicity that can be seen in the aligned amino acid sequence (Figure 8) and in the hydropathy plot (Figure 5). The hydropathy plot (Figure 5) indicates that this region is highly hydrophilic, which suggests a surface location for this sequence.

Among β subunits, the most noticeable difference is in the N-terminal region where the Cp sequence has 13 residues, whereas the other four sequences have about 60 residues. The N-terminal 60 residues of the β subunit of the other four organisms are moderately conserved (Figure 4) and are highly hydrophilic (Figure 6). Also, the 60 residues can form four α-helical regions (Av residue number): 1–9, 14–26, 37–44, and 48–61 (data not shown). The N-terminal 13 residues of the Cp β subunit may form one α-helical region (4–14), and they are hydrophilic (Figure 6), which suggests a surface location. The additional stretch of 50 amino acids unique to Cp α subunit (Figure 3) can form three α-helical regions (382–387, 407–418, and 427–433), which are hydrophilic (Figure 5). One of the possible explanations for the current situation might be that the event which placed the Cp *nifD* and Cp *nifK* genes in proximity caused the deletion of about 50 amino acids from the N-terminal region of the Cp β subunit and that a size constraint on the MoFe protein somehow caused the Cp α subunit to gain about 50 amino acids. X-ray crystallographic studies should reveal whether the folded structure of this Cp α subunit region resembles that of the N-terminal region of the other β subunits and whether this Cp α subunit region is spatially near the N-terminal region of Cp β subunit in MoFe protein.

There are additional regions in Cp α and β subunits that differ from the other organisms in terms of α helices, β sheets, or β turns. These are not discussed here because these regions do not have other known features and the current accuracy of structural predictions from amino acid sequences alone does not justify a detailed comparison in those regions.

Apparent Homology between α and β Subunits. A certain degree of similarity between the α and β subunits of MoFe protein was implicated from a low difference index between the amino acid compositions (Chen et al., 1973; Lundell & Howard, 1978) and from the position of the three conserved Cys residues in the N-terminal third of both subunits (Lammers & Haselkorn, 1983). Yamane et al. (1982) observed a low-resolution homology between the α and β subunits of the MoFe protein from X-ray crystallographic data. On the basis of limited sequence data, Thony et al. (1985) were able to detect homology between the α and β subunits in the region covering the three conserved Cys residues proximal to the N termini. With the sequence data now available, a refined alignment between the α and β subunits of *C. pasteurianum*, *A. vinelandii*, and *R. japonicum* is shown in Figure 7. The amino acid sequences clearly show a conservation of the position of the three Cys residues and a number of other amino acids between the two subunits in this region.

The homology is lower at the nucleotide level, which is expected. However, it is intriguing that the predicted α-helical and β-sheet structures (not shown) are largely different between the two subunits within this region (between the first two Cys residues; also around the third Cys residue). Nevertheless, the predicted positions of most β turns seem conserved. X-ray crystallographic data should reveal whether the difference in predicted secondary structures affects polypeptide folds or the spatial positions of the three Cys residues. Whether or not the three Cys residues in the α subunit are structurally equivalent to the three Cys residues in the β subunit is pertinent to our understanding of the function of nitrogenase and the evolution of these polypeptides.

Concluding Remarks. During the reduction of N_2 to ammonia by nitrogenase, electron transfer occurs between a physiological electron donor (ferredoxin, flavodoxin, or other low-potential reductants) and the Fe protein and then between the Fe protein and the MoFe protein. During the electron-transfer process, an effective interaction must occur between the donor protein and the acceptor protein to allow pertinent redox centers and mediating groups to be within a proper distance and in a correct orientation. A proper interaction must be important to the maintenance of an electron flux that allows the allocation of electrons into the desired substrates to achieve a high efficiency of N_2 fixation (Emerich et al., 1981). It is thus important to understand the nature of the interaction between nitrogenase components to maximize the efficiency of nitrogenase. This paper reports the identification of regions that are likely involved in component interaction, and these regions may be further tested by site-directed mutagenesis (Brigle et al., 1987b).

ACKNOWLEDGMENTS

We thank Katherine C.-K. Chen for computer programming and assistance with data analysis.

REFERENCES

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Bishop, P. E., Premakumar, R., Dean, D. R., Jacobson, M. R., Chisnell, J. R., Rizzo, T. M., & Kopczynski, J. (1986) *Science (Washington, D.C.) 232*, 92–94.

Braun, V., & Schroeder, W. A. (1967) *Arch. Biochem. Biophys. 118*, 241–252.

Brigle, K. E., Newton, W. E., & Dean, D. R. (1985) *Gene 37*, 37–44.

Brigle, K. E., Weiss, M. C., Newton, W. E., & Dean, D. R. (1987a) *J. Bacteriol. 169*, 1547–1553.

Brigle, K. E., Setterquist, R. A., Dean, D. R., Cantwell, J. S., Weiss, M. C., & Newton, W. E. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7066–7069.

Chen, J.-S., Multani, J. S., & Mortenson, L. E. (1973) *Biochim. Biophys. Acta 310*, 51–59.

Chen, K. C.-K., Chen, J.-S., & Johnson, J. L. (1986) *J. Bacteriol. 166*, 162–172.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Chou, P. Y., & Fasman, G. D. (1979) *Biophys. J. 26*, 367–384.

Dilworth, M. J., Eady, R. R., Robson, R. L., & Miller, R. W. (1987) *Nature (London) 327*, 167–168.

Drummond, M. H. (1985) *Biochem. J. 232*, 891–896.

Eady, R. R., Robson, R. L., Richardson, T. H., Miller, R. W., & Hawkins, M. (1987) *Biochem. J. 244*, 197–207.

Emerich, D. W., & Burris, R. H. (1978) *J. Bacteriol. 134*, 936–943.

Emerich, D. W., Ljones, T., & Burris, R. H. (1978) *Biochim. Biophys. Acta 527*, 359–369.

Emerich, D. W., Hageman, R. V., & Burris, R. H. (1981) *Adv. Enzymol. Relat. Areas Mol. Biol. 52*, 1–22.

Fischer, H.-M., & Hennecke, H. (1984) *MGG, Mol. Gen. Genet. 196*, 537–540.

Fuhrmann, M., & Hennecke, H. (1984) *J. Bacteriol. 158*, 1005–1011.

Golden, J. W., Robinson, S. J., & Haselkorn, R. (1985) *Nature (London) 314*, 419–423.

Guth, J. H., & Burris, R. H. (1983) *Biochemistry 22*, 5111–5122.

Hales, B. J., Langosch, D. J., & Case, E. E. (1986a) *J. Biol. Chem. 261*, 15301–15306.

Hales, B. J., Case, E. E., Morningstar, J. E., Dzeda, M. F., & Mauterer, L. A. (1986b) *Biochemistry 25*, 7251–7255.

Hanahan, D. (1983) *J. Mol. Biol. 166*, 557–580.

Hase, T., Nakano, T., Matsubara, H., & Zumft, W. G. (1981) *J. Biochem. (Tokyo) 90*, 295–298.
Hase, T., Wakabayashi, S., Nakano, T., Zumft, W. G., & Matsubara, H. (1984) *FEBS Lett. 166*, 39–43.
Hausinger, R. P., & Howard, J. B. (1982) *J. Biol. Chem. 257*, 2483–2490.
Hausinger, R. P., & Howard, J. B. (1983) *J. Biol. Chem. 258*, 13486–13492.
Holbrook, J. J., Liljas, A., Steindel, S. J., & Rossmann, M. G. (1975) *Enzymes (3rd Ed.) 11*, 191–269.
Ikemura, T. (1981) *J. Mol. Biol. 151*, 389–409.
Kaluza, K., & Hennecke, H. (1984) *MGG, Mol. Gen. Genet. 196*, 35–42.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Lammers, P. J., & Haselkorn, R. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4723–4727.
Lundell, D., & Howard, J. B. (1978) *J. Biol. Chem. 253*, 3422–3426.
Lundell, D., & Howard, J. B. (1981) *J. Biol. Chem. 256*, 6385–6391.
Mazur, B. J., & Chui, C.-F. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6782–6786.
Morgan, T. V., Mortenson, L. E., McDonald, J. W., & Watt, G. D. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 46*, 2241.
Oppenheim, D. S., & Yanofsky, C. (1980) *Genetics 95*, 785–795.
Orme-Johnson, W. H. (1985) *Annu. Rev. Biophys. Biophys. Chem. 14*, 419–459.
Riebeling, V., Thauer, R. K., & Jungermann, K. (1975) *Eur. J. Biochem. 55*, 445–453.
Robson, R. L., Eady, R. R., Richardson, T. H., Miller, R. W., Hawkins, M., & Postgate, J. R. (1986) *Nature (London) 322*, 388–390.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–353.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Selin, Y. M., Harich, B., & Johnson, J. L. (1983) *Curr. Microbiol. 8*, 127–132.
Smith, B. E., Thorneley, R. N. F., Eady, R. R., & Mortenson, L. E. (1976) *Biochem. J. 157*, 439–447.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.
Tanaka, M., Haniu, M., Yasunobu, K. T., & Mortenson, L. E. (1977) *J. Biol. Chem. 252*, 7093–7100.
Thony, B., Kaluza, K., & Hennecke, H. (1985) *MGG, Mol. Gen. Genet. 198*, 441–448.
Tsai, L. B., & Mortenson, L. E. (1978) *Biochem. Biophys. Res. Commun. 81*, 280–287.
Wang, S.-Z., Chen, J.-S., & Johnson, J. L. (1987) *Nucleic Acids Res. 15*, 3935.
Weinman, J. J., Fellows, F. F., Gresshoff, P. M., Shine, J., & Scott, K. F. (1984) *Nucleic Acids Res. 12*, 8329–8344.
Weston, M. F., Kotake, S., & Davis, L. C. (1983) *Arch. Biochem. Biophys. 225*, 809–817.
Wherland, S., Burgess, B. K., Stiefel, E. I., & Newton, W. E. (1981) *Biochemistry 20*, 5132–5140.
Yamane, T., Weininger, M. S., Mortenson, L. E., & Rossmann, M. G. (1982) *J. Biol. Chem. 257*, 1221–1223.
Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103–119.
Yasunobu, K. T., & Tanaka, M. (1973) in *Iron–Sulfur Proteins* (Lovenberg, W., Ed.) Vol. 2, pp 27–130, Academic, New York.
Yun, A. C., & Szalay, A. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7358–7362.
Zumft, W. G., & Mortenson, L. E. (1973) *Eur. J. Biochem. 35*, 401–409.
Zumft, W. G., Hase, T., & Matsubara, H. (1980) in *Molybdenum Chemistry of Biological Significance* (Newton, W. E., & Otsuka, S., Eds.) pp 59–72, Plenum, New York.