Misumi, M., & Weissbach, A. (1982) *J. Biol. Chem. 257*, 2323-2329.
Mosbaugh, D. W. (1988) *Nucleic Acids Res. 16*, 5645-5659.
Mosbaugh, D. W., & Meyer, R. R. (1980) *J. Biol. Chem. 255*, 10239-10247.
Mosbaugh, D. W., Stalker, D. M., Probst, G. S., & Meyer, R. R. (1977) *Biochemistry 16*, 1512-1518.
Sabatino, R. D., & Bambara, R. A. (1988) *Biochemistry 27*, 2266-2271.
Sala, F., Amileni, A. R., Parisi, B., & Spadari, S. (1980) *Eur. J. Biochem. 112*, 211-217.
Sala, F., Galli, M. G., Levi, M., Burroni, D., Parisi, B., Pedrali-Noy, G., & Spadari, S. (1981) *FEBS Lett. 124*, 112-118.
Schaaper, R. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8126-8130.
Scheuermann, R. H., & Echols, H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7747-7751.
Scheuermann, R., Tam, S., Burgers, P. M. J., Lu, C., & Echols, H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7085-7089.
Setlow, P., Brutlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247*, 224-231.
Spencer, D., & Whitfield, P. R. (1969) *Arch. Biochem. Biophys. 132*, 477-488.
Tewari, K. K. (1986) *Methods Enzymol. 118*, 186-201.
Wang, Z. F., Yang, J., Nie, Z. Q., & Wu, M. (1991) *Biochemistry 30*, 1127-1131.

Cloning and Sequencing of the Gene for Rubrerythrin from *Desulfovibrio vulgaris* (Hildenborough)†

Benet C. Prickril, Donald M. Kurtz, Jr.,* and Jean LeGall
Departments of Chemistry and Biochemistry, University of Georgia, Athens, Georgia 30602
Gerrit Voordouw
Department of Biological Sciences, The University of Calgary, Calgary, Alberta T2N 1N4, Canada
Received June 18, 1991; Revised Manuscript Received September 5, 1991

ABSTRACT: The gene coding for rubrerythrin from the sulfate-reducing bacterium *Desulfovibrio vulgaris* (Hildenborough) has been cloned and sequenced. Rubrerythrin is known to contain two types of iron sites: one rubredoxin-like FeS_4 center in each of the two identical subunits and one hemerythrin-like diiron site per dimer [LeGall, J., et al. (1988) *Biochemistry 27*, 1636-1642]. The gene encodes a polypeptide of 191 amino acids, and a normal ribosome binding site is located 11-6 base pairs upstream from the translational start of the gene. There is no evidence for the presence of a leader sequence, suggesting a cytoplasmic location for the protein. The rubrerythrin gene is not part of any other known transcriptional unit in the *D. vulgaris* genome. The nucleotide sequence encodes four Cys residues, the minimum required for ligation to iron in rubredoxin. The pairs of Cys residues occur in Cys-X-X-Cys sequences as they do in rubredoxin, but the 12-residue spacing between the Cys pairs in rubrerythrin is less than half that in rubredoxins. A pair of Arg residues flanking one Cys residue may contribute to the much more positive reduction potential of the rubredoxin-like site in rubrerythrin compared to that of rubredoxin. While the amino acid sequence of rubrerythrin shows no significant overall homology with that of any known protein, the C-terminal region does share some homology with rubredoxin sequences. If folding of the rubredoxin-like amino acid sequence domain in rubrerythrin is similar to that in rubredoxins, then three His residues are brought into proximity. These His residues, which have no counterpart in any rubredoxin, could, therefore, furnish ligands to the hemerythrin-like irons. This proposal is consistent with the absorption spectrum and the positive reduction potential qualitatively observed for the hemerythrin-like diiron site. Assuming that the hemerythrin-like diiron site bridges the two subunits, preliminary X-ray diffraction results on rubrerythrin [Sieker et al. (1988) *Proteins: Struct., Funct., Genet. 3*, 184-186] are consistent with this proposal.

A large number of redox-active metalloproteins have been isolated and characterized from the strictly anaerobic sulfate-reducing bacterium *Desulfovibrio vulgaris* (Hildenborough) (LeGall & Fauque, 1988). Among these proteins, rubrerythrin (Rr)[1] is unique in terms of its complement of high reduction potential non-heme iron sites (LeGall et al., 1988). Rr is a homodimer (native M_r 43 000) containing two types of iron sites: two rubredoxin-like (Rd-like) FeS_4 centers (one per subunit) with spectroscopic properties very similar to the tetrahedrally coordinated iron centers found in Rds (Watenpaugh et al., 1979; Adman et al., 1991) and one diiron site with spectroscopic similarities to those in a class of proteins for which Hr is a prototype (Holmes et al., 1991; Sanders-Loehr, 1989). This diiron site is, therefore, referred to as the Hr-like site, although it shows no ability to bind O_2. Despite

† This research was supported by grants from the National Institutes of Health (GM 40388 to D.M.K.) and the National Science and Engineering Research Council of Canada (G.V.). D.M.K. is an NIH Research Career Development Awardee during 1988-1993 (HL 02077). Amino acid sequence database searches made use of the University of Georgia Bioscience Computing Facility established under NSF Grant DIR880467.
* Address correspondence to this author at the Department of Chemistry, University of Georgia.

[1] Abbreviations: Rr, rubrerythrin; Rd, rubredoxin; Hr, hemerythrin; SDS, sodium dodecyl sulfate; IPTG, isopropyl β-D-thiogalactopyranoside; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside; TE, 10 mM Tris-HCl/0.1 mM EDTA, pH 8; SSC, 0.015 M sodium citrate/0.15 M NaCl, pH 7.2.

0006-2960/91/0430-11118$02.50/0

their spectroscopic similarity to Rd, the Rd-like centers in Rr have a standard reduction potential ($E^{\circ\prime}$) of +230 mV, more than 200 mV more positive than the centers of Rds. Liu and LeGall (1990) have reported pyrophosphatase activity associated with Rr, but it is presently unclear whether or not this activity reflects some physiological function. The high reduction potential referred to above suggests that Rr is not directly involved in the electron-transport chain of *D. vulgaris*, which has highly reducing growth requirements. Furthermore, no definitive experiments have been performed to determine the cellular localization of Rr. In this report, we describe the cloning and sequencing of the gene for Rr. Physiological and structural implications of these results are examined, and attempts to elucidate the molecular basis for some of the properties of this unique protein are also reported.

Experimental Procedures

Materials. Escherichia coli strain TG2 and plasmid pUC8 (Vieira & Messing, 1982) were used for cloning the Rr gene, and are described by Brumlik and Voordouw (1989). *E. coli* TG2 cells were grown in TY medium (Miller, 1977) containing 10 g of tryptone (Difco), 5 g of yeast extract (Difco), and 5 g of NaCl per liter at pH 7.4. Restriction endonucleases, T4 DNA ligase, DNA polymerase (Klenow fragment), T4 DNA polymerase, pUC8 cloning vector, and M13 sequencing vectors were obtained from Pharmacia. Calf alkaline phosphatase was purchased from Boehringer Mannheim Biochemicals. Radiolabeled [α-^{32}P]dATP, [γ-^{32}P]ATP, or [α-^{35}S]dATP, used for dideoxy sequencing, deoxyoligonucleotide end-labeling, and nick translation, were obtained from Amersham. Deoxynucleoside triphosphate and dideoxynucleoside triphosphate solutions and other sequencing reagents were prepared as described in Bankier and Barrell (1983). High- and low-gelling-temperature agarose was from Bethesda Research Laboratories. A synthetic 47-mer oligonucleotide probe, with a sequence derived from the N-terminal amino acid sequence of Rr, was purchased from Operon Technologies Inc., San Pablo, CA. The N-terminal amino acid sequence of Rr, which had been purified as described previously (LeGall et al., 1988), was determined by automated Edman degradation at the Molecular Genetics Instrumentation Facility of the University of Georgia.

Preparation of DNA Fragments. D. vulgaris chromosomal DNA, prepared as described by Marmur (1961), was cut with various restriction endonucleases and assayed using the Southern blotting technique (Southern, 1975) to identify the fragments which hybridized with the radioactively labeled 47-mer deoxyoligonucleotide probe. Radioactive labeling of the deoxyoligonucleotide probe was carried out as described by Sambrook et al. (1989). On the basis of this Southern blot (Figure 1), the ~7-kb *Pst*I fragment was chosen for cloning into a pUC8 cloning vector. Fragments of the *Pst*I-cut DNA in the 6–8-kb range were prepared by size-fractionation on a 0.7% low-gelling-temperature agarose gel, followed by purification of the fragments by phenol extraction, ethanol precipitation, and redissolving in TE buffer.

Cloning of DNA Fragments into pUC8. The purified *Pst*I fragments were ligated into a pUC8 vector previously cut with *Pst*I and treated with calf intestinal mucosa alkaline phosphatase. The ligation mixture was transformed in *E. coli* TG2 cells made competent by $CaCl_2$ treatment, and grown overnight at 37 °C on TY–ampicillin plates as described previously (Voordouw et al., 1985). Individual white colonies were picked and grown overnight at 37 °C in 1.5-mL cultures of TY medium containing 0.1 mg/mL ampicillin. Small amounts of the individual 1.5-mL cultures of transformed *E. coli* were kept at 4 °C for later amplification of positive clones.

Purification and Screening of Plasmids. Pooled plasmids were purified from six combined 1.5-mL cultures as described previously (Voordouw et al., 1985) and dissolved in 50 μL of TE. Purified pooled plasmids were screened as follows for the presence of the ~7-kb fragment containing the gene for Rr: 2-μL aliquots of pooled plasmid preparations, each containing six distinct plasmids, were briefly treated with sodium hydroxide, spotted onto a nitrocellulose filter, and assayed for hybridization with the 47-mer deoxyoligonucleotide probe; the filter was UV-irradiated for 3 min, prehybridized at 68 °C for 30 min in a solution containing Denhardt's reagent (Sambrook et al., 1989), 6× SSC, and 0.2% SDS, and allowed to hybridize with the ^{32}P-labeled deoxyoligonucleotide probe overnight at 60 °C. The filter was then washed in 6× SSC for 10 min at 60 °C to remove nonspecifically bound probe and autoradiographed. Screening of ~100 plasmid pools resulted in the isolation of 3 positive plasmids containing the 7-kb *Pst*I fragment (pBP93-6, pBP96-4, and pBP98-4), and 1 of these was used for subsequent steps.

Subcloning in pUC8. The plasmid vector pBP96-4 containing the 7-kb *Pst*I fragment was subcloned after digestion with *Eco*RI. The resulting fragments were separated on a 1% low-gelling-temperature agarose gel, and the 0.5-kb *Eco*RI fragment was purified by excision, phenol-extraction, and ethanol-precipitation. After ligation into *Eco*RI-cut pUC8, the plasmid was transformed in *E. coli* and amplified, purified, and nick-translated (Sambrook et al., 1989) for use as a probe in subsequent procedures. This plasmid was also used to generate larger amounts of the *Eco*RI fragment for use in producing M13 templates.

Cloning in M13. Two methods were used to obtain the M13 templates used in sequencing the Rr gene: (1) The pBP96-4 plasmid was digested either with *Eco*RI, with *Eco*RI and *Sal*I, or with *Eco*RI and *Hin*cII. The resulting fragments were ligated into, respectively, M13mp8 digested with *Eco*RI, M13mp9 digested with *Eco*RI and *Sal*I, and M13mp8 digested with *Eco*RI and *Sma*I. The ligation mixtures were then transfected into *E. coli* TG2 cells and grown overnight on TY plates treated with IPTG and X-gal, and white plaques were picked. The white plaques were grown for 6 h at 37 °C in 1.5 mL of TY medium with vigorous agitation, and the single-stranded phage DNA, purified as described elsewhere (Bankier & Barrell, 1983), was used for sequencing. (2) M13 sequencing templates were also produced by ligation with fragments derived from digestion of gel-purified fragments. Thus, the gel-purified 0.5-kb *Eco*RI fragment (Figure 2) was digested with *Taq*I, and the fragments were ligated into M13mp8 digested with *Eco*RI and *Acc*I. A similar strategy was used to produce templates after a gel-purified 3.5-kb *Sal*I fragment (Figure 2) was cleaved with *Taq*I. Ligation mixtures from these procedures were subsequently treated as described for the first method.

Nucleotide Sequencing. The single-stranded M13 templates were sequenced with the dideoxy chain termination method of Sanger et al. (1977). The 17-mer universal sequencing primer (Bankier & Barrell, 1983) and the synthetic 47-mer oligonucleotide were used as primers. The data were compiled and analyzed using the Fortran programs of Staden (1982, 1984). The nucleotide sequence of the entire 1.02-kb region containing the Rr gene has been deposited in Genbank, Accession Number M7701.

Results and Discussion

Cloning of the Rr Gene. Successful cloning of the Rr gene relied upon use of a synthetic 47-mer deoxyoligonucleotide as

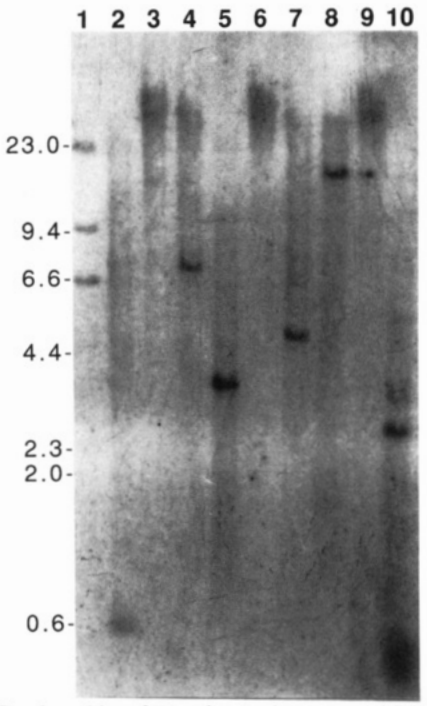

FIGURE 1: Southern blot of *D. vulgaris* chromosomal DNA digested with various restriction enzymes. A ^{32}P-labeled 47-mer deoxyoligonucleotide (cf. text) was used as probe. After hybridization with the probe at 60 °C overnight, the filter was washed 15 min at 60 °C to remove nonspecifically bound material. The film was exposed for 1 week before development. Lanes: 2, DNA digested with *Eco*RI; 3, *Hin*dIII; 4, *Pst*I; 5, *Sal*I; 6, *Bam*HI; 7, *Sma*I; 8, *Xho*I; 9, *Kpn*I; 10, *Acc*I. The positions of molecular size markers (bacteriophage λ DNA digested with *Hin*dIII, lane 1) are indicated in kilobases.

a gene-specific probe. The deoxyoligonucleotide had the sequence (5′ → 3′) TACAACTACTTCGG(CT)GG(CT)-CAGGC(CG)AAGAAGGACA*C*(CG)TTCGT(CG)G*A-(AG)AT, where the parentheses indicate degeneracies and the asterisks indicate mismatches with the ultimately determined gene sequence. The 47-mer oligonucleotide sequence was derived from the following N-terminal amino acid sequence of Rr: MKSLKGSRQ*EKNILTAFAGESQARNR-YNYFGGQAKKDT*FVE*ISDIFAETADQA*REHA, where the underlined residues indicate those used to generate the 47-mer oligonucleotide sequence and the asterisked residues indicate disagreements with the amino acid sequence determined from the gene sequence. The underlined sequence was chosen on the basis of the apparent fidelity of the protein-derived N-terminal amino acid sequence and on the relative lack of degeneracies in the corresponding DNA sequence based on known codon usage in *D. vulgaris* (Voordouw, 1988; Brumlik & Voordouw, 1989).

A Southern blot of *D. vulgaris* chromosomal DNA cut with various restriction enzymes, and probed with the ^{32}P-labeled 47-mer deoxyoligonucleotide, is shown in Figure 1. The ~7-kb *Pst*I fragment was chosen for cloning. Out of 600 plasmids screened, 3 were found to contain the ~7-kb *Pst*I fragment.

Sequencing Strategy. A restriction map of the 7-kb *Pst*I fragment containing the Rr gene, as well as the strategy used for sequencing, is shown in Figure 2. Restriction mapping of the 7-kb *Pst*I fragment revealed the presence of three *Sal*I sites within this fragment, and Southern blotting (Figure 1) placed the Rr gene within the 3.5-kb *Sal*I fragment (Figure 1, lane 5). This Southern blot also revealed the 0.5-kb *Eco*RI fragment that contains a portion of the gene (Figure 1, lane 2). Dideoxy sequencing of M13 templates containing the

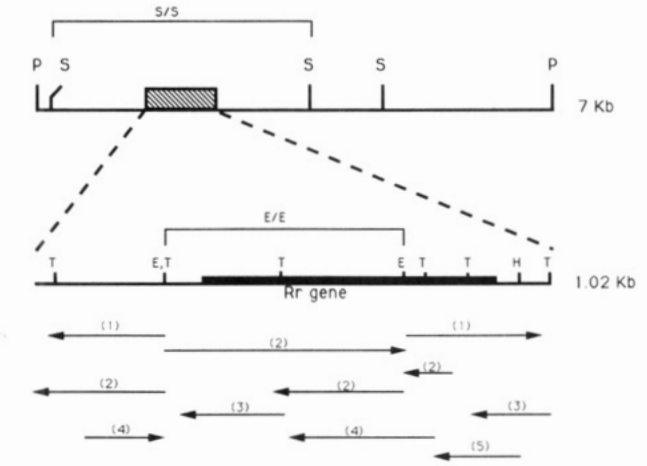

FIGURE 2: Restriction sites and sequencing scheme for the Rr gene. The 7-kb fragment containing *Pst*I (P) and *Sal*I (S) sites, as well as the 3.5-kb *Sal*I fragment (S/S), is shown at the top. The hatched area represents the 1.02-kb sequence containing the Rr gene and is expanded to show restriction sites used for sequencing. The 0.5-kb *Eco*RI fragment (E) containing most of the Rr gene is indicated (E/E). *Taq*I (T) and *Hin*cII (H) sites are also shown. Arrows indicate the direction and length of sequences generated. Fragments cloned in M13 and sequenced were from (1) *Eco*RI/*Sal*I, (2) *Eco*RI, (3) *Eco*RI/*Taq*I, (4) *Taq*I, and (5) *Eco*RI/*Hin*cII digests.

0.5-kb *Eco*RI fragment provided about 60% of the Rr gene sequence (Figure 2), and also specified the position of two *Taq*I sites, one within and one close to the coding region. Templates generated from ligation of *Eco*RI/*Sal*I fragments provided the remaining sequence for one of the strands within the coding region. Sequence information for the complementary strand was obtained from M13 templates generated from ligation of *Hin*cII/*Eco*RI and *Taq*I/*Eco*RI fragments.

The rbr Gene and the Amino Acid Sequence for Rr. The nucleotide sequence of the 1.02-kb region containing the gene for Rr (the *rbr* gene) is shown in Figure 3 together with the deduced amino acid sequence. A notable feature of the nucleotide sequence is a ribosome binding site (AGGAGG) 11–6 nucleotides upstream from the translational start of the protein. The positioning of this site indicates that the *rbr* gene does not encode a cleavable N-terminal signal sequence. The lack of an N-terminal signal sequence implies a cytoplasmic location for Rr, contrary to the periplasmic location inferred from the purification procedure (LeGall et al., 1988).

The 191 amino acid residue protein encoded by the *rbr* gene has a calculated molecular weight of 21 544, which is in excellent agreement with the independently determined subunit molecular weight of 21 900 for Rr (LeGall et al., 1988). Amino acid sequence comparisons using the Genetics Computer Group software package (Devereux et al., 1984) and associated databases revealed no strong overall homology of Rr to that of any other protein. Furthermore, the *rbr* gene has no obvious homology with the recently described *rbo* gene, which encodes the tentatively named rubredoxin oxidoreductase (Brumlik & Voordouw, 1989). The apparent protein product of the *rbo* gene has been isolated and renamed desulfoferrodoxin by Moura et al. (1990), who found it to contain both a distorted rubredoxin-type center and another mononuclear non-heme ferrous center. The *rbo* gene was found to be proximal to the *D. vulgaris rub* gene, which encodes Rd (Brumlik & Voordouw, 1989). The *rbr* gene is not located within this *rbo-rub* transcriptional unit.

Using the software package and databases cited above, amino acid sequence comparisons with other proteins were also conducted separately on N- and C-terminal portions of the Rr sequence, residues 1–152 and 153–191, respectively. This

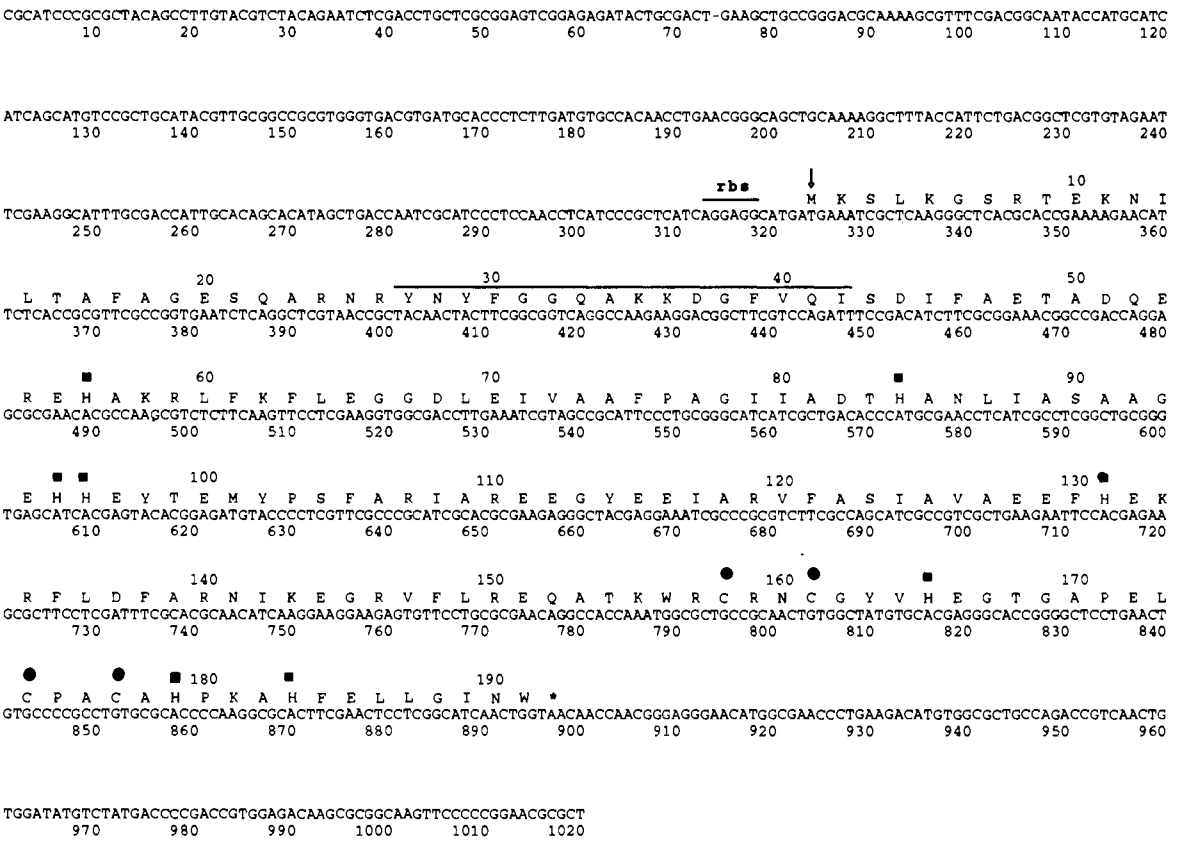

FIGURE 3: Nucleotide and amino acid sequences for *D. vulgaris* Rr. The solid line over nucleotides 313–318 indicates a probable ribosome binding site. The solid line over nucleotides 401–448 indicates the region of the amino acid sequence from which the 47-mer oligonucleotide region was derived. The stop codon is indicated by an asterisk, and the beginning of the protein is indicated by (↓). His residues (■) and Cys residues (●) are labeled.

latter search was conducted in order to examine the possibility that Rr is the result of gene fusion, as has been proposed for other *D. vulgaris* proteins (Brumlik & Voordouw, 1989). Here again, no obvious sequence homology to that of any other protein was found for the N-terminal portion. However, a significant degree of sequence homology between the C-terminal portion of Rr and Rds is readily apparent.

The Rd-like Iron Binding Site. The four cysteine codons in the nucleotide sequence of Figure 3 code for the four cysteine ligands that are presumably required for the Rd-like center in each subunit of Rr. The spacing of the Cys residues in Rr, [Cys-X-X-Cys-$(X)_n$-Cys-X-X-Cys], is identical to that in Rds, with the notable exception that $n = 12$ in Rr vs 28 or 29 (or 22 in one case) in Rds (Meyer et al., 1990; Kok et al., 1989). These regions in *D. vulgaris* Rd and Rr are compared in Figure 4. The Pro residue between the second pair of Cys residues in the Rr sequence (174 and 177) is conserved in all Rds, perhaps because it favors bending of the polypeptide chain such that both flanking Cys residues can coordinate to iron in a bidentate fashion. Similarly, the Gly residues immediately following each pair of Cys residues in nearly all Rds are thought to be sterically required for a folding geometry favoring N–H—S hydrogen bonding to the preceding cysteine sulfur atom (Adman et al., 1975). In Rr, the analogous residues are Gly162 and Ala178. The second residue preceding each Cys-X-X-Cys pair is aromatic (i.e., Trp, Phe, or Tyr) in the sequences of all Rds; these residues have recently been proposed as part of an aromatic π system extending underneath two of the cysteine sulfur ligands (Adman et al., 1991). Only one of these aromatic residues preceding the Cys pairs, namely, Trp156, appears to be conserved in Rr. Tyr163 is the only other aromatic residue in Rr that appears to have a conserved counterpart in all Rds. Thus, only some features of the con-

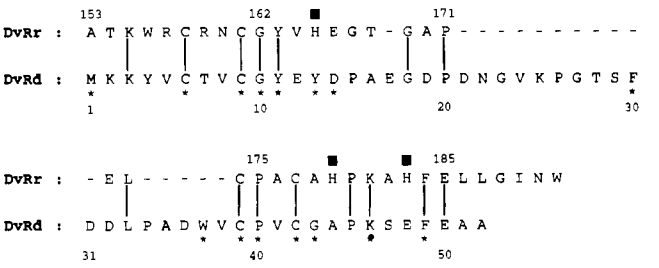

FIGURE 4: Alignment of *D. vulgaris* Rr (DvRr) and *D. vulgaris* Rd (DvRd) amino acid sequences. Residue numbers are indicated above and below the respective sequences. Gaps in the Rr sequence, when aligned with the Rd sequence, are indicated by dashes, and a vertical line connects identical residues. Asterisks indicate residues conserved in all known Rds (Meyer et al., 1990). (■) indicates His residues.

sensus Rd binding site sequence are conserved in Rr. The very similar spectroscopic properties of the Rd-like site in Rr to those of Rds (LeGall et al., 1988; Prickril et al., 1989) may, therefore, be due primarily to the Cys-X-X-Cys sequence spacing coupled with structural and energetic requirements inherent to the $[FeSR]_4^{-,2-}$ coordination unit (Stenkamp et al., 1990; Werth et al., 1989; Gebhard et al., 1991).

The features discussed up to now do not explain the ≥200 mV more positive reduction potential of the Rd-like site. However, the two Arg residues flanking Cys158, with their positively charged guanidino side chains, would provide an explanation for this more positive reduction potential. Two basic residues flanking a Cys ligand have no precedent in Rds. The flanking Arg residues also make it likely that Cys158 resides near the surface of the protein.

The region between the two Cys pairs in Rr is predicted by the algorithm of Chou and Fasman (1978) to have no propensity for α-helix formation, and could, therefore, form a

nonhelical loop, the ends of which are connected by the iron atom of the Rd-like center. Such nonhelical loops, albeit longer ones, are in fact observed in the four structurally characterized Rds (Adman et al., 1991; Frey et al., 1987; Sieker et al., 1986; Watenpaugh et al., 1979). The iron site in Rd is thought to have an electron-transfer role, and by analogy, the Rd sites in Rr could also play such a role, but with the additional distinct possibility of interaction with the Hr-like diiron site.

Potential Ligand Residues for the Hr-like Diiron Site. Known ligand residues at non-heme, non-sulfur diiron sites in proteins include His, Asp, Glu, and Tyr with μ-oxo/hydroxo and aqua ligands completing the coordination spheres (Sanders-Loehr, 1989; Nordlund et al., 1990). Tyr ligation can probably be eliminated in the case of Rr, since the absorption spectrum (LeGall et al., 1988) lacks the characteristic phenolate $\rightarrow$ Fe(III) charge-transfer transition (Que, 1983). In fact, the absorption spectrum attributed to the Hr-like diiron(III) site of Rr rather resembles that of the (μ-oxo)diiron(III) site in Hr, whose protein ligands consist of His, Asp, and Glu (Holmes et al., 1991). Similarly, Mössbauer spectroscopy of oxidized Rr supports an antiferromagnetically coupled (μ-oxo)diiron(III) site with O,N ligands (LeGall et al., 1988; Kurtz, 1990). A fairly positive reduction potential for the Hr-like site in Rr is qualitatively confirmed by the observations of a small portion of the mixed-valent form of this site in the as-isolated protein under aerobic conditions and of complete reduction of the iron centers by ascorbate ($E^{\circ\prime}$ = +58 mV) at pH 7.6 (LeGall et al., 1988). A positive reduction potential should be favored by a larger number of neutral imidazole ligands relative to negatively charged carboxylato ligands. Data for other non-heme, non-sulfur iron sites in proteins show that, for a coordination sphere consisting of carboxylato and imidazole ligands, a positive reduction potential is associated with two or more His ligands per iron (Petrouleas & Diner, 1986; Diesenhofer et al., 1985; Barette et al., 1983; Stoddard et al., 1990). The (μ-oxo)diiron(III) site in Hr has a total of five His ligands (three to one iron and two to the other) vs two carboxylato ligands and a reduction potential of +110 mV for the Fe(III,III)/(II,III) couple (Armstrong et al., 1983).[2] In contrast, the (μ-oxo)diiron(III) site in ribonucleotide reductase has a total of two His ligands (one to each iron) vs four carboxylato ligands (Nordlund et al., 1990) and a reduction potential of $\geq$-110 mV (Lam et al., 1990). Rr contains eight His residues (indicated in Figure 3), no five of which have the spacing found for the His ligands in Hr (Holmes et al., 1991). Three residues in Rr, Glu53, His56, and His179, have the same spacing as do three of the iron ligands in *E. coli* ribonucleotide reductase (Nordlund et al., 1990). However, the remaining three protein ligands in ribonucleotide reductase have no apparent counterpart in the Rr sequence, and, when coupled with the difference in reduction potentials cited above, the three matching residues may merely be coincidental. There is currently no evidence for ribonucleotide reductase activity of Rr.

Another possible set of His ligands in Rr originates from two conserved sequence motifs including the Rd-like site that were not noted above. The region including the first pair of Cys residues in Rd has a Cys-X-X-Cys-Gly-Tyr-X-Tyr consensus sequence (Meyer et al., 1990), which appears as Cys-Arg-Asn-Cys-Gly-Tyr-Val-His (residues 158–165) in Rr (cf. Figure 4). His165 in this latter sequence would lie on the short loop, mentioned above, between the pairs of Cys residues. A consensus sequence including the second pair of Cys residues in Rd, Cys-Pro-X-Cys-Gly-X-Pro-Lys-X-X-Phe-Glu, becomes Cys-Pro-Ala-Cys-Ala-His-Pro-Lys-Ala-His-Phe-Glu (residues 174–185) in Rr with 7 of the 12 residues (underlined) matching those in Rd (cf. Figure 4). The structures of native Rds (Adman et al., 1991; Frey et al., 1987; Sieker et al., 1986; Stenkamp et al., 1990; Watenpaugh et al., 1979) show that the stretch of residues immediately following the second Cys pair bends back toward the aforementioned loop between the Cys pairs. If the analogous positioning occurs in Rr, then the pair of His residues 179 and 183 would be brought into the vicinity of His165. His residues are extremely rare in Rds with only 1 occurring in the 9 known Rd sequences encompassing a total of more than 450 residues (Meyer et al., 1990). This rarity suggests that the three His residues, 165, 179, and 183, within the Rd-like sequence domain of Rr (cf. Figure 4) have a specific function, one obvious possibility being the provision of ligands to the one Hr-type iron per subunit found in the previous study (LeGall et al., 1988). The proposed positioning of the Hr-like site proximal to the Rd-like site in Rr may reflect the binding site of the putative redox partner proteins of Rds.

While this proximal positioning of iron sites is intriguing, it must be noted that a sufficient number of His, Asp, and Glu residues occur elsewhere in the sequence to accommodate one Hr-like (μ-oxo)diiron site per subunit. Therefore, the question of whether or not the Hr-like site bridges the two subunits of the Rr dimer cannot be answered unequivocally by the sequence data alone. The preliminary X-ray diffraction results on Rr indicate that both the polypeptides and metal centers in the dimer are related by 2-fold rotational symmetry (Sieker et al., 1988). If the Hr-like diiron site does indeed bridge the two subunits, then the 2-fold rotational symmetry requires identical ligand environments for the two iron atoms. This symmetry is consistent with the His ligand residues proposed above.

Summary and Conclusions. The gene coding for the non-heme iron protein Rr from *D. vulgaris* has been isolated and sequenced. The gene encodes a protein of 191 amino acids in a sequence with features that agree very well with previous chemical and spectroscopic characterizations of the isolated protein (LeGall et al., 1988). A normal ribosome binding site 11–6 base pairs upstream of the start codon indicates the absence of a leader sequence, which implies a cytoplasmic location for the protein. The four Cys residues furnishing ligands to the Rd-like site have been identified. The Cys-X-X-Cys spacing characteristic of Rds is preserved in Rr, but the 12-residue spacing between the Cys pairs in Rr is less than half that in Rds. The amino acid sequence of Rr shows no significant overall homology with that of any protein known to contain a diiron site, whereas the C-terminal region of Rr does show a significant degree of homology to Rd sequences. If folding of this Rd-like amino acid sequence domain in Rr is similar to that in Rds, then a set of three His residues can be brought into proximity. These His residues, which have no analogue in any Rd, could, therefore, furnish ligands to the Hr-like diiron site. Preliminary X-ray diffraction results are consistent with this proposal, assuming that the Hr-like diiron site bridges the two subunits.

[2] The three-His iron is believed to be the center that undergoes the redox change for the Fe(III,III) $\rightleftharpoons$ Fe(II,III) couple in Hr (Pearce et al., 1987; McCormick & Solomon, 1990).

References

Adman, E. T., Watenpaugh, K. D., & Jensen, L. H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4854–4858.

Adman, E. T., Sieker, L. C., & Jensen, L. H. (1991) *J. Mol. Biol. 217*, 337–352.

Armstrong, F. A., Harrington, P. C., & Wilkins, R. G. (1983) *J. Inorg. Biochem. 18*, 83–91.
Bankier, A. T., & Barrell, B. G. (1983) *Techniques in Nucleic Acid Biochemistry* (Flavell, R. A., Ed.) pp 1–34, Elsevier Scientific Publishers Ireland Ltd., Limerick, Ireland.
Barette, W. C., Sawyer, D. T., Fee, J. A., & Asada, K. (1983) *Biochemistry 22*, 624–627.
Brumlik, M. J., & Voordouw, G. (1989) *J. Bacteriol. 171*, 4996–5004.
Chou, P., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–147.
Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.
Diesenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1985) *Nature 318*, 618–624.
Frey, M., Sieker, L. C., Pavan, F., Haser, R., Bruschi, M., Pepe, G., & LeGall, J. (1987) *J. Mol. Biol. 197*, 525–541.
Gebhard, M. S., Koch, S. A., Millar, M., Devlin, F. J., Stephens, P. J., & Solomon, E. I. (1991) *J. Am. Chem. Soc. 113*, 1640–1649.
Holmes, M. A., Trong, I. L., Turley, S., Sieker, L. C., & Jensen, L. H. (1991) *J. Mol. Biol. 218*, 583–593.
Kok, M., Oldenhuis, R., van der Linden, M. P. G., Meulenberg, C. H. C., Kingma, J., & Witholt, B. (1989) *J. Biol. Chem. 264*, 5442–5451.
Kurtz, D. M., Jr. (1990) *Chem. Rev. 90*, 585–606.
Lam, K.-Y., Fortier, D. G., & Sykes, A. G. (1990) *J. Chem. Soc., Chem. Commun.*, 1019–1021.
LeGall, J., & Fauque, G. (1988) *Biology of Anaerobic Microorganisms* (Zehnder, A. J. B., Ed.) pp 1–34, Wiley, New York.
LeGall, J., Prickril, B., Moura, I., Xavier, A. V., Moura, J. J. G., & Huynh, B.-H. (1988) *Biochemistry 27*, 1636–1642.
Liu, M.-Y., & LeGall, J. (1990) *Biochem. Biophys. Res. Commun. 171*, 313–318.
Marmur, J. (1961) *J. Mol. Biol. 3*, 208–218.
McCormick, J. M., & Solomon, E. I. (1990) *J. Am. Chem. Soc. 112*, 2005–2007.
Meyer, J., Gagnon, J., Sieker, L. C., Van Dorsselaer, A., & Moulis, J.-M. (1990) *Biochem. J. 271*, 839–841.
Miller, J. H. (1977) *Experiments in Molecular Genetics*, 3rd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B.-H., Liu, M. Y., & LeGall, J. (1990) *J. Biol. Chem. 265*, 21596–21602.
Nordlund, P., Sjöberg, B.-M., & Eklund, H. (1990) *Nature 345*, 593–598.
Pearce, L. L., Kurtz, D. M., Jr., Xia, Y.-M., & Debrunner, P. G. (1987) *J. Am. Chem. Soc. 109*, 7286–7293.
Petrouleas, V., & Diner, B. A. (1986) *Biochim. Biophys. Acta 849*, 264–275.
Prickril, B. C., Kurtz, D. M., Jr., LeGall, J., Johnson, M. K., Huang, H.-Y., & Czernuszewicz, R. (1989) *J. Inorg. Biochem. 36*, 228.
Que, L., Jr. (1983) *Adv. Inorg. Biochem. 5*, 167–199.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanders-Loehr, J. (1989) in *Iron Carriers and Iron Proteins* (Loehr, T. M., Ed.) pp 373–466, VCH, New York.
Sanger, F., Nicklen, S., & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Sieker, L. C., Stenkamp, R. E., Jensen, L. H., Prickril, B., & LeGall, J. (1986) *FEBS Lett. 208*, 73–76.
Sieker, L. C., Turley, S., Prickril, B. C., & LeGall, J. (1988) *Proteins: Struct., Funct., Genet. 3*, 184–186.
Staden, R. (1982) *Nucleic Acids Res. 10*, 4731–4751.
Staden, R. (1984) *Nucleic Acids Res. 12*, 499–503.
Stenkamp, R. E., Sieker, L. C., & Jensen, L. H. (1990) *Proteins: Struct., Funct., Genet. 8*, 352–364.
Stoddard, B. L., Howell, P. L., Ringe, D., & Petsko, G. A. (1990) *Biochemistry 29*, 8885–8893.
Southern, E. (1975) *J. Mol. Biol. 98*, 503–517.
Vieira, J., & Messing, J. (1982) *Gene 19*, 259–268.
Voordouw, G. (1988) *Gene 69*, 75–83.
Voordouw, G., Walker, J. E., & Brenner, S. (1985) *Eur. J. Biochem. 148*, 509–514.
Watenpaugh, K. D., Sieker, L. C., & Jensen, L. H. (1979) *J. Mol. Biol. 131*, 509–522.
Werth, M. T., Kurtz, D. M., Jr., Howes, B. D., & Huynh, B.-H. (1989) *Inorg. Chem. 28*, 1357–1361.