Poulsen, L. L., Taylor, K., Williams, D. E., Masters, B. S. S., & Ziegler, D. M. (1986) *Mol. Pharmacol. 30*, 680-685.
Sabourin, P. J., Smyser, B. P., & Hodgson, E. (1984) *Int. J. Biochem. 16*, 713-720.
Smith, M. J., & Koch, G. L. E. (1989) *EMBO J. 8*, 3581-3586.
Treves, S., DeMattei, M., Landfredi, M., Villa, A., Green, N. M., MacLennan, D. H., Meldolesi, J., & PozzAn, T. (1990) *Biochem. J. 271*, 473-480.
Tynes, R. E., Sabourin, P. S., & Hodgson, E. (1985) *Biochem. Biophys. Res. Commun. 126*, 1069-1075.
Walls, F. C., Baldwin, M. A., Falick, A. M., Gibson, B. W., Gillece-Castro, B. L., Kaur, S., Maltby, D. A., Medzihradszky, K. F., Evans, S., & Burlingame, A. L. (1990) in *Biological Mass Spectrometry* (Burlingame, A. L., & McCloskey, J. A., Eds.) pp 197-216, Elsevier Science Publishers, Amsterdam.
Williams, D. E. (1991) in *N-Oxidation of Drugs. Biochemistry, Pharmacology, Toxicology* (Hlavica, P., & Damani, L. A., Eds.) Chapter 6, pp 91-105, Chapman and Hall, New York.
Williams, D. E., Ziegler, D. M., Nordin, D. J., Hale, S. E., & Masters, B. S. S. (1984) *Biochem. Biophys. Res. Commun. 125*, 116-122.
Williams, D. E., Hale, S. E., Muerhoff, A. S., & Masters, B. S. S. (1985) *Mol. Pharmacol. 28*, 381-390.
Ziegler, D. M. (1980) in *Enzymatic Basis of Detoxication* (Jacoby, W. B., Ed.) pp 201-227, Academic Press, New York.
Ziegler, D. M. (1988) *Drug Metab. Rev. 19*, 1-32.
Ziegler, D. M., & Petit, F. H. (1964) *Biochem. Biophys. Res. Commun. 15*, 188-193.
Ziegler, D. M., & Poulsen, L. L. (1977) *Trends Biochem. Sci. 2*, 79-81.
Ziegler, D. M., Poulsen, L. L., & York, B. M. (1983) in *Functions of Glutathione* (Larsson, A., Orrenius, S., Holmgren, A., & Mannervik, B., Eds.) pp 297-305, Raven Press, New York.

# Primary Structure of the Assimilatory-Type Sulfite Reductase from *Desulfovibrio vulgaris* (Hildenborough): Cloning and Nucleotide Sequence of the Reductase Gene[†,‡]

Jian Tan,[§] Larry R. Helms,[∥] Richard P. Swenson,[∥] and J. A. Cowan*[,§]

*Evans Laboratory of Chemistry, 120 West 18th Avenue, and Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210*

*Received April 8, 1991; Revised Manuscript Received June 7, 1991*

ABSTRACT: The nucleotide sequence encoding the structural gene (651 bp) and flanking regions for the assimilatory-type sulfite reductase from the sulfate-reducing bacterium *Desulfovibrio vulgaris* (Hildenborough) was determined after cloning a 1.4 kb *Hin*dIII/*Sal*I genomic fragment possessing the gene into Bluescript pBS(+)KS. The primary structure of the protein was deduced, and the molecular mass of the apoprotein was estimated as 24 kDa. The amino acid sequence of the polypeptide shows some similarities at putative $[Fe_4S_4]$ cluster binding sites in comparison with the heme protein subunit of the larger *Escherichia coli* and *Salmonella typhimurium* sulfite reductases and spinach nitrite reductase. This is the first reported sequence of a member of a new class of low molecular weight assimilatory sulfite-reducing enzymes recently identified in a number of anaerobic bacteria [Moura, I., Lina, A. R., Moura, J. J. G., Xavier, A. V., Fauque, G., Peck, H. D., & Le Gall, J. (1986) *Biochem. Biophys. Res. Commun. 141*, 1032-1041].

Biological redox chemistry is a diverse field that encompasses both electron transfer proteins and oxidoreductase enzymes (Cowan et al., 1989; Barber, 1984; Witt et al., 1986; Papa, 1983; Hatefi et al., 1985; Mortenson & Thorneley, 1979). Typically, the latter class possesses redox cofactors that serve to bind substrate and facilitate the formation and cleavage of chemical bonds. Important examples are found in the biological sulfur cycle. For instance, many bacteria possess assimilatory enzymes that reduce oxyanions of sulfur to the oxidation level of sulfide, which is the redox state of sulfur in a large number of cellular compounds (Le Gall & Postgate, 1973; Peck & Le Gall, 1982; Siegel, 1975; Vega & Kamin, 1977; Lancaster et al., 1979; Huynh et al., 1984; Murphy et al., 1974). Bacteria that use sulfate as the terminal electron acceptor in anaerobic respiration (dissimilatory reduction) also possess a sulfate ($SO_4^{2-}$) reducing system[1] and sulfite ($SO_3^{2-}$) reductases (Le Gall & Postgate, 1973; Peck & Le Gall, 1982; Siegel, 1975).

Our laboratory is engaged in the elucidation of the molecular details of enzymatic sulfite reduction as carried out by the assimilatory-type sulfite reductase from *Desulfovibrio vulgaris* (Hildenborough). The six-electron reduction of $SO_3^{2-}$ to $S^{2-}$ is catalyzed by a single enzyme possessing a siroheme chromophore that is essential for sulfite binding and reduction (Young & Siegel, 1988). On the basis of spectroscopic evi-

[†] This work was supported by NSF Grant CHE-8921468.
* To whom correspondence should be addressed.
[‡] GenBank Accession Number J05344.
[§] Evans Laboratory of Chemistry.
[∥] Department of Biochemistry.

[1] The enzyme ATP sulfurylase activates $SO_4^{2-}$ by formation of the adenosine phosphosulfate (APS) adduct. Subsequent reduction of APS to $SO_3^{2-}$ and AMP is catalyzed by APS reductase (Postgate, 1984).

0006-2960/91/0430-9900$02.50/0 

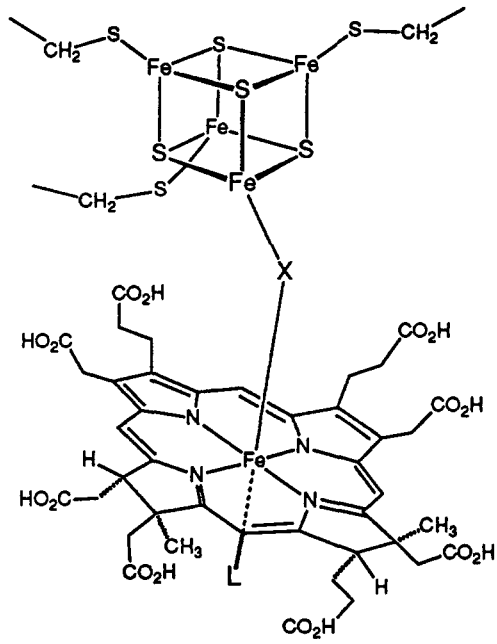

FIGURE 1: Schematic representation of the proposed exchange coupled $[Fe_4S_4]$–siroheme in the active site of assimilatory sulfite reductase (*D. vulgaris*). The bridging unit X may be either inorganic sulfide or cysteine. L is most probably histidine (Cowan & Sola, 1990).

dence (Huynh et al., 1984; Moura et al., 1986) and with reference to the low-resolution crystallographic analysis of the *Escherichia coli* sulfite reductase subunit (McRee et al., 1986), the proposed active site is an exchange-coupled $[Fe_4S_4]$–siroheme complex (Figure 1) that would appear to be ubiquitous to all sulfite-reducing enzymes, although this viewpoint has been challenged (Pierik & Hagen, 1991). A cluster–siroheme complex has also been identified in several nitrite reductases (Vega & Kamin, 1977; Lancaster et al., 1979; Huynh et al., 1984; Murphy et al., 1974; Crowe et al., 1983a,b). A number of these enzymes were found to possess high-spin pentacoordinate sirohemes; however, a novel set of low molecular weight assimilatory sulfite reductases, containing hexacoordinate low-spin sirohemes, have recently been isolated from three anaerobic bacteria [viz., *Methanosarcina barkeri*, *Desulfuromonas acetoxidans*, and *D. vulgaris* (Hildenborough)] (Moura et al., 1986). There appears to be at least two quite distinct classes of assimilatory enzyme: a high molecular weight protein such as found in the *E. coli* heme subunit ($M_r$ ca. 57000) and the low molecular weight enzymes typified by that from *D. vulgaris* ($M_r$ ca. 24000). Our work is focused on the latter enzyme (Huynh et al., 1984). On the basis of chemical analysis data, inorganic sulfide may serve as the bridging ligand between the siroheme and the $[Fe_4S_4]$ center in the *D. vulgaris* reductase (Huynh et al., 1984). This stands in contrast to the cysteine thiolate bridge proposed for the *E. coli* enzyme (Cline et al., 1986; Madden et al., 1989; McRee et al., 1986; Ostrowski et al., 1989). Recent NMR studies suggest histidine as the sixth ligand (Figure 1; Cowan & Sola, 1990).

Herein we describe the identification and cloning of the gene encoding the assimilatory-type sulfite reductase from *D. vulgaris*. The primary structure of the protein has been deduced directly from the nucleotide sequence, and a comparison is made with the sequences of the larger assimilatory sulfite reductases from *E. coli* and *Salmonella typhimurium* and the nitrite reductase from spinach (Ostrowski et al., 1989). This is the first reported sequence of one of the new class of low molecular weight sulfite-reducing enzymes noted earlier (Moura et al., 1986). By comparing the primary structures of these proteins, we hoped to identify common residues neighboring putative binding domains for the prosthetic centers, identify structural variations at the active sites, and rationalize the implications of homologies for the understanding of enzymatic reaction chemistry.

## Materials and Methods

*Materials.* [$\gamma$-$^{32}$P]ATP (>3000 Ci/mmol) and [$\alpha$-$^{32}$P]ATP (3000 Ci/mmol) were purchased from Amersham Corp. Tetramethylammonium chloride was obtained from Aldrich Co. The deionized water used in all procedures was obtained by passing distilled water through a Barnstead nanopure system. Sephadex gel filtration materials were obtained from Sigma Chemical Co. Unless otherwise stated, restriction endonucleases were purchased from Bethesda Research Laboratories. *E. coli* strain DH5$\alpha$MCR was obtained from Bethesda Research Laboratories. An Accumet pH meter 910 equipped with a Ross combination pH electrode (Orion Research Inc.) was used for routine pH measurements.

*Cell Growth and Protein Isolation.* *D. vulgaris* (Hildenborough) was cultured in modified Baars medium at the pilot fermentation facility in the Department of Biochemistry (Madison, Wisconsin). Assimilatory sulfite reductase was isolated by following literature procedures (Huynh et al., 1984).

*Amino-Terminal Sequencing.* An aliquot of the *D. vulgaris* reductase (80 $\mu$g) in potassium phosphate (50 mM, pH 7.6) was exchanged with $NH_4OAc$ (1 mM) by Centricon-10 ultrafiltration (Amicon) and lyophilized. The amino-terminus was sequenced by the Ohio State University Biochemical Instrument Center on an Applied Biosystems model 470A gas-phase sequencer using the standard 03RPTH Edman degradation sequencing protocol. Phenylthiohydantoin derivatives were identified by reverse-phase HPLC[2] using an Applied Biosystems model 120A instrument that was interfaced directly to the sequencer.

*Tryptic Digestion and Peptide Isolation.* To 350 $\mu$L of *D. vulgaris* sulfite reductase (1 $\mu$g/$\mu$L in 0.2 M NaPi, pH 7.0) was added 4.5 $\mu$L of a 1% solution (w/v in 1 mM HCl) of trypsin (TPCK-treated, Worthington Enzymes), and the mixture was incubated at room temperature for 4 h. A second 4.5-$\mu$L aliquot of trypsin was added and incubated for an additional 4 h. The reaction was terminated by addition of HCl (2 M, ca. 2 drops) to bring the pH to ca. 2. The resulting precipitate was dissolved by addition of solid urea to give a final concentration of 4 M, and the pH was adjusted to ca. 7.3. To ensure complete digestion, two additional aliquots (4.0 and 3.5 $\mu$L) of trypsin stock solution were added and incubated each time for 4 h. The pH was again lowered to ca. 2–3 as before, and 200 $\mu$L of the solution was injected onto a Phenomenex reverse-phase HPLC column ($\mu$-Bondapak 10C18, 300 × 3.9 mm) connected to a LKB HPLC. The peptide

[2] Abbreviations: bp, base pair; DSiR, *Desulfovibrio vulgaris* assimilatory-type sulfite reductase; EDTA, ethylenediaminetetraacetic acid; ESiR, *Escherichia coli* assimilatory sulfite reductase (heme subunit); HPLC, high-pressure liquid chromatography; $Me_4NCl$, tetramethylammonium chloride; NaPi, sodium phosphate; $NH_4OAc$, ammonium acetate; NiR, spinach nitrite reductase; PAGE, polyacrylamide gel electrophoresis; RT, room temperature; SDS, sodium dodecyl sulfate; SSC, sodium chloride/sodium citrate (Sambrook et al., 1989); SSiR, *Salmonella typhimurium* assimilatory sulfite reductase (heme subunit); TE, Trishydroxymethylamine (Trizma)–ethylenediaminetetraacetic acid buffer (Sambrook et al., 1989); TEAB, triethylammonium bicarbonate; TFA, trifluoroacetic acid; TPCK, 1-*p*-tosyl-amino-2-phenylethyl chloromethyl ketone.

fragments were eluted with a linear gradient from 0.1% TFA to 50% acetonitrile in 0.1% TFA over a period of 50 min. The elution profile was monitored at both 220 and 280 nm. A peak eluting after 30 min was isolated. The combined fractions from two runs were lyophilized prior to sequencing as described above. The sequence data suggested that the sample contained two major peptide fragments at two quite different concentrations, thus the sequence of each could be determined independently from the same data set.

*Amino Acid Analyses.* Amino acid compositions were determined by the Biochemical Instrument Center after hydrolysis of purified protein (60 μg) in hydrochloric acid vapors (phenol was added as an antioxidant) at 110 °C in an evacuated vial for 24, 48, and 72 h, in order to account for degradation of unstable amino acids. Cysteine (/cystine) was determined as cysteic acid after oxidation of the protein in performic acid for 24 h at −20 °C, followed by acid hydrolysis for 24 h as above. Amino acid analyses were performed after precolumn derivatization with phenylisothiocyanate on a Waters Associates Pico Tag amino acid analysis system according to standard manufacturer protocols.

*Preparation of Oligonucleotide Probes.* Mixed dodeca- and heptadecanucleotide probes were synthesized by the Biochemical Instrument Center on an Applied Biosystems model 380B DNA synthesizer using standard phosphoramidite chemistry. The oligonucleotides were deblocked by incubation in concentrated ammonium hydroxide at 55 °C for 6–12 h, desalted by elution (TEAB, 10 mM, pH 7) through a Sephadex G-50 column (0.8 × 10 cm), and lyophilized in a Savant speedvac. The sequences of the two mixed probes were as follows:

P1: 5′-GA(C,T)GA(A,G)CC(T,G,C)AA(A,G)GG (A,G,C,T)GC(A,G,C,T)AT-3′

(derived from the N-terminal sequence)

and

P2: 5′-GA(A,G)CC(G,C)ATG(A,T)(G,C) (G,C)GA(A,G)GA-3′

(derived from the tryptic peptide)

The probes were radiolabeled by 5′-phosphorylation using [γ-$^{32}$P]ATP and polynucleotide kinase, following standard procedures (Sambrook et al., 1989). Unincorporated [γ-$^{32}$P]ATP was removed by passage over a Sephadex G-50 column equilibrated in TE buffer.

*Construction of a Subgenomic Library.* Unless otherwise stated all recombinant DNA procedures were taken from Maniatis (1982) or Sambrook et al. (1989). A subgenomic library was constructed by using the Bluescript vector pBS(+)KS (Strategene) as follows. Genomic DNA (Krey et al., 1988) was digested to completion with *Hin*dIII and *Sal*I restriction enzymes. The *Hin*dIII/*Sal*I restriction fragments were separated by agarose gel (0.8%) electrophoresis and then transferred to Genescreen Plus membrane (DuPont) by a modified Southern blotting procedure (Southern, 1975; Sambrook et al., 1989). Membranes were prehybridized (48 °C, 5 h) in the following solution (50 mL): 6× SSC, 0.7% SDS (45.5 mL); Denhart's (2.5 mL, 50× Denhart's solution); boiled (10 min), sonicated salmon sperm DNA (500 μL, 10 mg/mL); polyA (250 μL, 10 mg/mL). Hybridization with radiolabeled oligonucleotide probes P1 or P2 (48 °C, 18 h) was performed in a $Me_4NCl$ solution (50 mL; Wood et al., 1985) containing 3 M $Me_4NCl$, 50 mM Tris (pH 8), 2 mM EDTA, and 0.7% SDS (44 mL); boiled, sonicated salmon sperm DNA (500 μL, 10 mg/mL); 50× Denhart's solution (2.5 mL); $^{32}$P-labeled probe (4 × $10^7$ cpm). Washing was carried out in 3 M $Me_4NCl$ wash solution (3 M $Me_4NCl$, 50 mM Tris, pH 8, 2 mM EDTA, 0.7% SDS) as follows. The membranes were rinsed twice with the wash solution (RT) and incubated in the same solution (150 mL, RT, 15 min). The incubation step was repeated, and the membranes were then washed in the 3 M $Me_4NCl$ solution (180 mL, 50 °C, 1 h) with agitation and rinsed once with the wash solution (60 mL, RT) and twice with 2× SSC, 0.2% SDS, followed by incubation in 2× SSC, 0.2% SDS (150 mL, RT, 15 min) with occasional shaking. The last step was repeated, and the membranes were subjected to autoradiography (−70 °C).

The fragments (0.9–2.0 kb) that putatively contain the gene were size-selected by preparative agarose gel electrophoresis (Krey et al., 1988) and recovered by electroelution. These fragments were ligated into the *Hin*dIII/*Sal*I restriction sites in the pBS(+)KS vector under conditions described by Grundstrom et al. (1985). The ligation mixture was used to transform *E. coli* $DH_5\alpha$MCR cells by established procedures. The transformation mixture was then plated onto GeneScreen Plus membranes (DuPont) that had been placed on LB/amp plates. All plates were then replica plated onto GeneScreen Plus (duplicate copies were made for each plate), and colonies were allowed to grow at 37 °C to a diameter of 1–2 nm for the colony hybridizations described below (Helms et al., 1990).

*Screening of the Subgenomic Library.* The membranes were treated with various wash solutions to affect cell lysis and denaturation of DNA (Maniatis, 1982; Sambrook et al., 1989). The membranes were rinsed in 3× SSC and 1% SDS buffer (500 mL), air-dried, and baked in a 90 °C oven for 2 h. Subsequently, they were washed three times in 3× SSC and 1% SDS (100 mL) and each time scrubbed thoroughly with a Kimwipe. After soaking in 3× SSC and 1% SDS (500 mL, 61 °C, 17 h), the membranes were rinsed with fresh 3× SSC/1% SDS solution (200 mL, 61 °C). Membranes were prehybridized and hybridized under the same conditions described above for the restriction blots except that only probe P2 was used in identifying positive clones. Positive recombinant clones were rescreened by hybridization at 51 °C of Southern blots of enzyme digested (*Hin*dIII/*Sal*I) recombinant plasmid by using the two oligonucleotide probes as described above. Plasmid(s) containing inserts that hybridized strongly with both probes were characterized by restriction mapping and the structural gene localized within the insert by Southern blot analysis.

*DNA Sequencing.* Nucleotide sequences were determined by the dideoxy chain-termination method of Sanger et al. (1977). Single-stranded templates were prepared according to the protocol provided by Stratagene in which interference resistance helper phage VCS-M13 (Stratagene) was employed. Double-stranded DNA sequencing was performed following the Stratagene protocol. The sequencing primers used were universal primer, reverse primer, and eight synthetic 17-mer oligonucleotide primers that were derived from the sequences obtained by using the commercially available primers. Each primer was prepared on the Applied Biosystems model 380B DNA synthesizer described above.

*Determination of Molecular Weight and Isoelectric Point.* The molecular weight of the sulfite reductase was estimated by using molecular weight standards from Bethesda Research Laboratories (high range). SDS–PAGE was run on a Phast electrophoresis system (Pharmacia/LKB) at 15 °C using precast 20% homogeneous gels (Pharmacia/LKB). A sample of enzyme (3 μL of a 19.2 mM solution) was prepared, applied, and electrophoresed under recommended preprogramed operating conditions described by the manufacturer. Analytical

IEF gels were also run on the Phast electrophoresis system using precast 20% homogeneous gels (pH 3.5–9.0) at 15 °C, and isoelectric pH values were referenced to oxidized and reduced Mb [pI 8.1 and 7.4, respectively (Cowan et al., 1989)] and high-potential iron protein (*Chromatium vinosum*) [pI 3.6 (Bartsch, 1971, 1978)]. Typically 1 μL of a 1 mM protein solution was applied to each lane and electrophoresed under recommended preprogramed operating conditions.

*Sequence Comparison Analysis.* Alignment of the sulfite reductase sequences from *D. vulgaris*, *E. coli*, and *S. typhimurium* and the nitrite reductase from spinach, was carried out with a number of computer programs from the University of Wisconsin Genetics Computer Group, including SEQED, COMPARE, DOTPLOT, and BESTFIT. Sequences were entered manually with the SEQED program. A dot matrix alignment of the shorter sequence from the *D. vulgaris* reductase with the larger data files from the *S. typhimurium* and spinach reductases was performed. Sequence alignments were refined with the BESTFIT program and finally by visual inspection to select those alignments that retained the maximum sequence continuity. The various stringency parameters employed in the fits are reported either in the legend to Figure 4 (dot matrix) or in the text below (BESTFIT).

## Results and Discussion

*Isolation of the Gene Encoding the Assimilatory-Type Sulfite Reductase (D. vulgaris).* Our general strategy for identification of the reductase gene was patterned after that previously described (Krey et al., 1988; Helms et al., 1990): (1) digestion of the genomic DNA with various combinations of restriction enzymes; (2) screening of digestion fragments with radiolabeled oligonucleotide probes, which were constructed on the basis of limited peptide sequence data, to identify small restriction fragments potentially containing the gene encoding the reductase; (3) cloning of size-selected restriction fragments to form a subgenomic library; (4) screening the subgenomic library and identification of transformants containing the sulfite reductase gene by use of the radiolabeled probes noted above; (5) subcloning of appropriate restriction fragments for sequence analysis.

To provide greater control in the process of isolating the gene, we chose to utilize two synthetic oligonucleotide probes in the identification and cloning of the *D. vulgaris* sulfite reductase gene. The first probe (P1) was designed after sequencing the initial 16 residues of the N-terminus: $H_2N$-Ser-Asp-Glu-Pro-Lys-Gly-Ala-Ile-Leu-Glu-Arg-Asp-Lys-Leu-Thr-Tyr. The underlined residue was later identified as glutamine from the nucleotide sequence. Hydrolysis of the amide in Gln and Asn residues is commonly observed during amino acid sequencing reactions. A portion of this sequence, Asp-Glu-Pro-Lys-Gly-Ala-Ile, was used to derive the oligonucleotide sequence of a 20-mer mixed probe (P1) that has a 384-fold redundancy (taking account of the preferred codon usage by this organism; Krey et al., 1988). Two internal tryptic peptide fragments isolated by HPLC were also sequenced:

(1)
Glu-Ala-Val-Glu-Pro-Met-Trp(Ser,Arg)-Glu-Glu-Leu-Arg

(2)
Thr-Pro-Ser-Gly-Leu-Leu-Thr-Pro-Asp-Val-Leu-Asp-Ala-Val-Ser-Arg

The second probe (P2, 17-mer), with a redundancy of 64 (taking account of the preferred codon usage), was derived from the sequence Glu-Pro-Met-Trp(Ser,Arg)-Glu-Glu in peptide (1). The sequencing results for (1) were ambiguous and could not discount the possibility of Arg or Ser as replacements for the Trp residue. Also, the residue identified as serine (underlined) in sequence (2) was found to be cysteine from the corresponding nucleotide sequence.

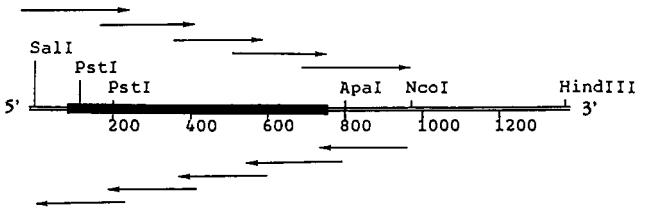


FIGURE 2: Restriction map of the 1.4 kb insert containing the gene encoding the *D. vulgaris* assimilatory-type sulfite reductase (from left to right, 5′ to 3′, heavy line) and sequencing strategy.

```
ATA GAT ACG CAA GAA CAG GAG ACC ATG ATG AGT GAC GAA CCT AAG GGC GCG ATA CTG CAG     60
                            Met Ser Asp Glu Pro Lys Gly Ala Ile Leu Gln

CGC GAC AAG CTC ACC TAC GCC ATC GTA CCC CGC ACC CCG TGC GGT CTT CTC ACC CCC GAC    120
Arg Asp Lys Leu Thr Tyr Ala Ile Val Pro Arg Thr Pro Cys Gly Leu Leu Thr Pro Asp

GTG CTT GAT GCC GTG TCG CGG GTC TGC AGG AAG TAC GAA GTG CCG ATC ATC AAG ATC ACA    180
Val Leu Asp Ala Val Ser Arg Val Cys Arg Lys Tyr Glu Val Pro Ile Ile Lys Ile Thr

TCC GGG CAG CGT CTG GCC CTT GTC GGC ATG AAG AAG GAA GCC GTC GAA CCC ATG TGG GAA    240
Ser Gly Gln Arg Leu Ala Leu Val Gly Met Lys Lys Glu Ala Val Glu Pro Met Trp Glu

GAG CTT CGC CTT GAC GTC GGA CGT GCC GTA GAA CTT TGC GTG CAC TAT GTG CAG GCG TGC    300
Glu Leu Arg Leu Asp Val Gly Arg Ala Val Glu Leu Cys Val His Tyr Val Gln Ala Cys

CCC GGT ACG GCG GTC TGC CGG TTC GGC CTT CAG GAT TCC CTC GGC ATC GGC GTG GCC ATA    360
Pro Gly Thr Ala Val Cys Arg Phe Gly Leu Gln Asp Ser Leu Gly Ile Gly Val Ala Ile

GAA GAG GAA TAT GTC GGG CAC GAC TTC CCA GCC AAG GTC AAG TTC GGC ATT TCG GGC TGT    420
Glu Glu Glu Tyr Val Gly His Asp Phe Pro Ala Lys Val Lys Phe Gly Ile Ser Gly Cys

CCC TTC TGC TGT GGC GAA ACG TAC CTG CGA GAT GTA GGT CTC GTG GGC ACC AAG AAG GGC    480
Pro Phe Cys Cys Gly Glu Thr Tyr Leu Arg Asp Val Gly Leu Val Gly Thr Lys Lys Gly

TGG ACG CTC ATC GTG GGT GGC AAT TCC GGC GGG CAT CCG CGC ATC GGC GAT GTG CTT GCA    540
Trp Thr Leu Ile Val Gly Gly Asn Ser Gly Gly His Pro Arg Ile Gly Asp Val Leu Ala

GAA GAG CTT TCC ACT GAT GAG GCC AAG GGG CTG ATA CGT AAG TTC ATG GAG TTC TAT CGC    600
Glu Glu Leu Ser Thr Asp Glu Ala Lys Gly Leu Ile Arg Lys Phe Met Glu Phe Tyr Arg

GAT AAC TCG GGT AAA CGG CTG CGT GTC TCC AAG TTC GTC GAG AAG ACG GGC ATC GAG GCC    660
Asp Asn Ser Gly Lys Arg Leu Arg Val Ser Lys Phe Val Glu Lys Thr Gly Ile Glu Ala

ATA CGG CAG GCG GTA CTC GGT TAG AAT CAG GGG CTG CGG CCC CCT TCC TTT TTG ATG CCC    720
Ile Arg Gln Ala Val Leu Gly Non

CCT CGT GGG AGG GCC CC                                                           737
```

FIGURE 3: Nucleotide and deduced amino acid sequence of *D. vulgaris* assimilatory-type sulfite reductase. The Shine–Dalgarno sequence (—) and two regions of dyad symmetry possibly involved in transcription termination (— —) are indicated.

Both radiolabeled mixed probes hybridized strongly and consistently to a 1.4 kb fragment in Southern blots of *Hin*dIII/*Sal*I-digested genomic DNA (data not shown). A subgenomic DNA library in pBS(+)KS was prepared that contained size-selected *Hin*dIII/*Sal*I fragments (0.9–2.0 kb) corresponding to the hybridizing region. Approximately 3000 recombinant clones were obtained. After colony hybridization, six positive clones were identified and rescreened. Of the six, two recombinant plasmids were found to contain 1.4 kb *Hin*dIII/*Sal*I inserts that hybridized strongly and specifically to both probes. Restriction mapping and subsequent Southern blot analysis showed that the entire gene and flanking regions were included in the insert.

*Nucleotide Sequence Analysis.* The partial restriction map and the sequencing strategy used to acquire the complete nucleotide sequence of the structural gene and flanking sequences are shown in Figure 2. The nucleotide sequence of the *D. vulgaris* sulfite reductase gene, as well as part of the 5′- and 3′-flanking regions, was established by using the data derived from both strands of DNA and is shown in Figure 3.

*(1) Shine–Dalgarno Sequence.* The initiation codon (ATG) at the beginning of the gene sequence is preceded by -AGGAGAC- (Figure 3), which is very similar to the concensus

Table I: Codon Usage in *D. vulgaris* (Hildenborough)[a]

| A | B | C | D | | | A | B | C | D | | | A | B | C | D | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 2 | UUU | Phe | 0 | 1 | 1 | 0 | UCU | Ser | 3 | 0 | 4 | 0 | UAU | Tyr |
| 7 | 6 | 18 | 3 | UUC | Phe | 5 | 4 | 9 | 5 | UCC | Ser | 3 | 5 | 18 | 3 | UAC | Tyr |
| 0 | 0 | 0 | 0 | UUA | Leu | 0 | 0 | 0 | 0 | UCA | Ser | 0 | 0 | 0 | 0 | UAA | Ter |
| 0 | 0 | 2 | 0 | UUG | Leu | 3 | 1 | 7 | 0 | UCG | Ser | 1 | 1 | 2 | 1 | UAG | Ter |
| 9 | 4 | 4 | 1 | CUU | Leu | 1 | 0 | 2 | 1 | CCU | Pro | 1 | 1 | 1 | 2 | CAU | His |
| 6 | 7 | 15 | 2 | CUC | Leu | 5 | 3 | 22 | 3 | CCC | Pro | 2 | 0 | 12 | 7 | CAC | His |
| 0 | 0 | 0 | 0 | CUA | Leu | 1 | 0 | 0 | 0 | CCA | Pro | 0 | 0 | 1 | 0 | CAA | Gln |
| 5 | 1 | 13 | 3 | CUG | Leu | 3 | 0 | 8 | 1 | CCG | Pro | 5 | 3 | 15 | 1 | CAG | Gln |
| 1 | 1 | 2 | 0 | AUU | Ile | 1 | 0 | 1 | 0 | ACU | Thr | 1 | 0 | 0 | 0 | AAU | Asn |
| 8 | 8 | 21 | 0 | AUC | Ile | 4 | 3 | 26 | 5 | ACC | Thr | 1 | 2 | 8 | 3 | AAC | Asn |
| 4 | 0 | 2 | 0 | AUA | Ile | 1 | 1 | 0 | 0 | ACA | Thr | 1 | 1 | 2 | 1 | AAA | Lys |
| 3 | 1 | 22 | 4 | AUG | Met | 4 | 3 | 8 | 0 | ACG | Thr | 14 | 3 | 47 | 20 | AAG | Lys |
| 0 | 1 | 2 | 3 | GUU | Val | 0 | 2 | 4 | 1 | GCU | Ala | 6 | 3 | 4 | 1 | GAU | Asp |
| 9 | 6 | 23 | 1 | GUC | Val | 9 | 10 | 32 | 11 | GCC | Ala | 5 | 14 | 25 | 8 | GAC | Asp |
| 4 | 0 | 0 | 2 | GUA | Val | 1 | 2 | 10 | 0 | GCA | Ala | 10 | 9 | 17 | 5 | GAA | Glu |
| 9 | 2 | 9 | 7 | GUG | Val | 4 | 3 | 9 | 2 | GCG | Ala | 7 | 4 | 14 | 0 | GAG | Glu |

| A | B | C | D | | |
|---|---|---|---|---|---|
| 2 | 0 | 2 | 1 | UGU | Cys |
| 6 | 4 | 17 | 8 | UGC | Cys |
| 0 | 0 | 0 | 0 | UGA | Ter |
| 2 | 2 | 6 | 1 | UGG | Trp |
| 4 | 1 | 5 | 0 | CGU | Arg |
| 5 | 3 | 13 | 1 | CGC | Arg |
| 1 | 0 | 0 | 0 | CGA | Arg |
| 4 | 2 | 3 | 0 | CGG | Arg |
| 1 | 1 | 0 | 0 | AGU | Ser |
| 0 | 1 | 8 | 1 | AGC | Ser |
| 0 | 0 | 0 | 0 | AGA | Arg |
| 1 | 1 | 0 | 1 | AGG | Arg |
| 6 | 5 | 9 | 4 | GGU | Gly |
| 15 | 9 | 35 | 6 | GGC | Gly |
| 1 | 1 | 1 | 0 | GGA | Gly |
| 3 | 3 | 4 | 0 | GGG | Gly |

[a] Column A is sulfite reductase (this work), column B is flavodoxin, column C is hydrogenase, subunits I and II, and column D is cytochrome $c_3$.

Table II: Amino Acid Composition of Assimilatory-Type Sulfite Reductase from *D. vulgaris* (Hildenborough)

| amino acid | no. of residues per protein | | amino acid | no. of residues per protein | |
|---|---|---|---|---|---|
| | deduced | analyzed | | deduced | analyzed |
| Asp (D) | 11 | 16 (Asp + Asn) | Cys (C) | 8 | 7 |
| Asn (N) | 2 | | Met (M) | 3 | 2 |
| Glu (E) | 17 | 23 (Glu + Gln) | Ser (S) | 9 | 10 |
| Gln (Q) | 5 | | Thr (T) | 10 | 10 |
| Lys (K) | 15 | 16 | Pro (P) | 10 | 10 |
| Arg (R) | 15 | 14 | Gly (G) | 25 | 27 |
| Trp (W) | 2 | –[a] | Ala (A) | 14 | 16 |
| Tyr (Y) | 6 | 5 | Ile (I) | 13 | 13 |
| Phe (F) | 7 | 7 | Leu (L) | 20 | 22 |
| His (H) | 3 | 2 | Val (V) | 22 | 22 |

[a] Not determined.

sequence for Shine–Dalgarno sites in *E. coli* (AGGAGGT; Gren, 1984).

(*2*) *Termination Site.* A stem and loop structure that is consistent with known RNA polymerase termination signals (Adhya & Gottesman, 1978) was found in the region downstream from the TAG stop codon (indicated by arrows). This site is GC rich and is followed by a short stretch of T's.

(*3*) *Coding Region.* The assimilatory-type sulfite reductase from *D. vulgaris* is encoded by an open reading frame of 651 bp. The nucleotide and amino acid sequences of the protein derived from this open reading frame are shown in Figure 3. Codon usage for the reductase gene is summarized in Table I together with the codon usage for flavodoxin, two subunits of the periplasmic hydrogenase, and for cytochrome $c_3$, which had been previously reported for this organism (Krey et al., 1988; Voordouw et al., 1985). Clear preferences can be observed in the selection of certain codons for most of the amino acids in these four genes; in particular, the use of either G or C in the wobble position. Within experimental error, the deduced and experimentally determined amino acid compositions given in Table II are similar. The low isoelectric point (pI 4.6) most likely results from the large number of carboxylate substituents associated with the siroheme chromophore since the number of positively and negatively charged amino acid residues are similar. The reductase gene encodes a translation product of 217 amino acids, although the initial methionine is absent from the mature protein. Previously, the molecular weight of the apoprotein was estimated as 27 000 by sedimentation equilibrium (Lee et al., 1973); however, the value of 24 000 determined from the primary sequence is consistent with that reported from disc gel electrophoresis (Lee et al., 1973) and repeated by us, employing SDS–PAGE. Since the molecular weight of the apoprotein deduced from the primary structure is similar to that found from SDS–PAGE experiments, we conclude that the DNA sequence is free from frame-shift errors that might alter the translational reading frame or lead to an incorrect translation stop codon. This is further supported by the identification of the termination signal flanking the 3′-end of the nucleotide sequence.

*Computer-Assisted Comparison of Primary Sequences.* The primary sequences of the *E. coli* and *S. typhimurium* assimilatory sulfite reductases are very similar (differing in only 34 noncritical residues out of 570; Ostrowski et al., 1989), and so each may be referred to interchangeably in the following discussion. The amino acid sequence of the *D. vulgaris* enzyme was compared with those from *S. typhimurium* sulfite reductase and spinach nitrite reductase by use of the dot matrix method (Figure 4). This is a particularly valuable procedure for determining regions of sequence homology in proteins differing greatly in size. By use of this type of analysis, the *D. vulgaris* enzyme appears to be most highly homologous with the C-terminal halves of *S. typhimurium* sulfite reductase and spinach nitrite reductase, although in both cases the homology is weak. For each of these enzymes, there is a high degree of homology around one pair of cysteines in the C-terminus that may form part of a $[Fe_4S_4]$ cluster binding domain (Ostrowski et al., 1989).

The amino acid sequence of *D. vulgaris* sulfite reductase was aligned with those of *S. typhimurium* sulfite reductase and spinach nitrite reductase by using the program BESTFIT. The low homology between the former and either the N- or C-terminal domains of the larger proteins was also reflected by this comparative sequence alignment. Typically, similar fits could be obtained with either the N- or C-terminal domains, depending on the parameters used, although the best fit was always to the C-terminus. For example, with gap and length weights of 7.0 and 0.5, respectively, alignment of the *D. vulgaris* sequence to that of *S. typhimurium* sulfite reductase was optimized for the 362–557 domain of the latter, giving one gap in the C-terminal sequence of this enzyme with 15% identity and 40% similarity with the corresponding *D. vulgaris* sequence. For spinach nitrite reductase (gap and length weights of 4.0 and 0.3, respectively) an optimal alignment from residues 350–562, with four gaps in the sequence, gave a 20% identity and 47% similarity. The N-terminal alignments tend to have a larger number of gaps in the sequence, which is also reflected in the dot matrix comparison in Figure 4. The previously reported sequence alignment between *S. typhimurium* sulfite reductase and spinach nitrite reductase (Ostrowski et al., 1989), which are enzymes of comparable size (570 and 562 amino acids, respectively), also displayed a low homology (25% identity, 58% similarity). It is clear that there is very little overall sequence homology when comparing any of these enzymes, although this is not unexpected since these proteins derive from quite distinct sources (an anaerobic bacterium, an aerobic bacterium, and a plant source). Moreover, the *E. coli* and *S. typhimurium* enzymes are derived from multisubunit complexes (Ostrowski et al., 1989; Siegel et al., 1974a,b; Siegel, 1975) and may require distinct structural elements and charge distributions to better define essential protein–protein contacts. We would, however, expect more uniformity among the three smaller assimilatory sulfite reductases isolated from the sulfate-reducing strains noted earlier. For these reasons, we restrict our attention to those parts of the sequence that might possess a higher degree of sequence homology, viz., the binding domains for the prosthetic centers.

*Ligands Binding to the Prosthetic Centers.* The cysteine content of each enzyme is 8, 6 (4), 6 (4), and 12 (7) for the

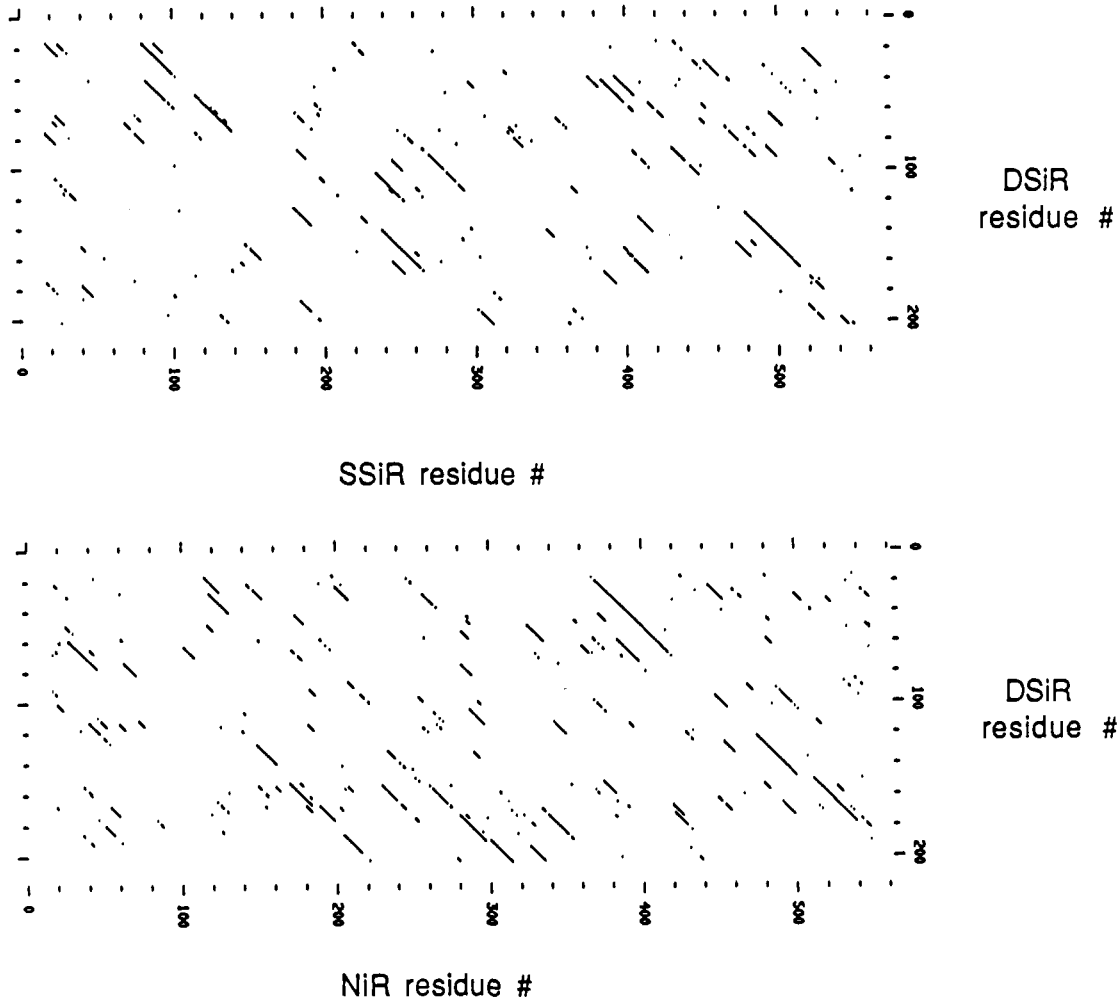


FIGURE 4: Dot matrix comparison of the amino acid sequences of *S. typhimurium* sulfite reductase (SSiR) and spinach nitrite reductase (NiR) to that of *D. vulgaris* sulfite reductase (DSiR). Sequence homologies were established by using the program COMPARE and then displayed by use of DOTPLOT. The procedure employed a window of 30 residues with a stringency of 12. The numbers along the axes correspond to the amino acid residues in each enzyme. The sequences of SSiR and NiR were from Ostrowski et al. (1989) and Back et al. (1988). Diagonal lines indicate regions of possible homology according to criteria determined by the window size ($x$-axis) and stringency ($y$-axis) of the analysis. A dot indicates a match of $y$ out of $x$ amino acids. If diagonal lines are offset, then this is indicative of a possible gap in the sequence alignment of the two proteins.

*D. vulgaris*, *E. coli*, and *S. typhimurium* sulfite reductases and spinach nitrite reductase, respectively. The parenthetical values for the larger enzymes refer to the number of cysteines in the C-terminal halves of the sequences that are most likely to be involved in cluster coordination. Figure 5 shows the local alignments around the two cysteine pairs previously proposed to bind the [$Fe_4S_4$] cluster. The sequence Gly-Cys-Pro-Phe-Cys-Cys-Gly- at position 129–135 of the *D. vulgaris* enzyme shows a high degree of homology with the equivalent sequences for the *S. typhimurium* and *E. coli* enzymes (residues 478–484) or spinach nitrite reductase (residues 513–519), which show Gly-Cys-Pro-Asn-Gly-Cys-Gly and Gly-Cys-Pro-Asn-Ser-Cys-Gly, respectively. The Cys-$(X)_3$-Cys motif is therefore very likely to form an integral part of the cluster binding domain.

The second set of putative Cys binding pairs Cys-$(X)_5$-Cys (Figure 5) lie 33, 38, and 34 residues away from the Cys-$(X)_3$-Cys sequence in the *D. vulgaris*, *S. typhimurium*, and spinach enzymes, respectively. The Cys-$(X)_5$-Cys sequences show no homology other than the Ala residue before the first Cys. The possibility of a common site of helical contact in this region of the enzyme that was previously postulated due to the identical amino acid residues (Ala and Glu) at positions 3 and 6 after the second Cys in the *S. typhimurium* sulfite reductase and spinach nitrite reductase (Ostrowski et al., 1989)

```
DSiR
        82 - L C V H Y V Q A C P G T A V C R F G L Q D - 112
       128 - S G C P F C C G E - 136

SSiR
       432 - M A C V S F P T C P L A M A E - 446
       477 - T G C P N G C G R - 485

NiR
       471 - V A C T G S Q F C G Q A I I E - 485
       512 - T G C P N S C G Q V Q V A D I G
                 F M G C M T R D E N G K P C E - 542
```

FIGURE 5: Comparison of the amino acid sequences for putative cluster binding domains in *D. vulgaris* (DSiR) and *S. typhimurium* (SSiR) (also *E. coli*) sulfite reductase and spinach nitrite reductase (NiR). Cysteine residues are underlined.

may be valid; the *D. vulgaris* enzyme possesses Gly and Asp residues at these positions, respectively. However, the absence of Pro in the $(X)_5$ region of the nitrite reductase makes a strict structural analogy among these three coordination domains tenuous. The *D. vulgaris* enzyme has an additional cysteine before the Cys-$(X)_5$-Cys sequence that gives rise to a possible

Cys-$(X)_6$-Cys binding domain. At this moment it is not possible to reliably predict which of these two pairs, *if either*, is likely to coordinate to the cluster. A Cys-$(X)_{14}$-Cys motif from position 24 to 39 (Figure 3) could also be considered since Cys pairs with larger intervening sequences commonly arise in low potential ferredoxins.

Huynh et al. (1984) have previously quantitated the amount of labile sulfide by chemical analysis and suggested the possibility of a bridging sulfide between the siroheme and cluster chromophores. This stands in contrast to the more thoroughly characterized *E. coli* enzyme where a bridging cysteine has been proposed (Cline et al., 1986; Han et al., 1989). The apparent lack of sequence homology between these various enzymes and the quite distinct origins of their host cells should lend a note of caution to any rigorous assignment of the bridging ligand at this time. While it is obvious that common sequences do exist in all of these enzymes, there are sufficient differences in the order and/or identity of the intervening residues between putative pairs of coordinating cysteines to make any rigorous assessment premature. The *E. coli* enzyme is currently the only sulfite reductase where there is strong evidence for the identity of the bridging ligand. Such a ligand is likely to be subject to very rigorous steric and orientational limitations. The many variations in residues adjacent to possible bridging cysteines lead one to conclude that both Cys and inorganic sulfide must be considered as potential ligands, depending on the enzyme. Further work is required.

Recently, Pierik and Hagen (1991) have challenged the coupled siroheme-$[Fe_4S_4]$ complex as a general active site model for sulfite, and a number of nitrite, reductases. This work was based primarily on EPR data from the $\alpha_2\beta_2$ dissimilatory sulfite reductase (desulfoviridin) from *D. vulgaris*. The $^1H$ NMR results from our laboratory on the assimilatory enzyme, and in particular the observed Curie temperature dependence of the chemical shifts, are very indicative of coupled prosthetic centers.

We have considered the bridging ligand to the siroheme above. Of the residues most likely to act as the sixth ligand to the siroheme (His, Met, Cys, and possibly Tyr) there are three His, two Met, eight Cys (at least three of which must be involved in coordination to the cluster), and six Tyr. $^1H$ NMR studies from our laboratory have suggested histidine as a sixth axial ligand to the low-spin siroheme in the *D. vulgaris* enzyme (Cowan & Sola, 1990), which is supported by careful analysis of published EPR spectra (Tan & Cowan, 1991).

The lack of similarity in molecular weight and distinct coordination numbers of the prosthetic siroheme, in comparison to previously characterized assimilatory enzymes (Murphy et al., 1974; Siegel et al., 1982; Crowe et al., 1983; Cammack et al., 1978), suggests that these low molecular weight sulfite reductases fall into a distinct class.

## Acknowledgments

We thank Mesut Eren for valuable advice, Jane Tolley in the Ohio State Biochemical Instrument Center for carrying out the amino acid analyses and synthesis of oligonucleotides, and L.-Y. Hsu for guidance on computer-assisted sequence analysis.

## References

Adhya, S., & Gottesman, M. (1978) *Annu. Rev. Biochem. 47*, 967–996.
Barber, J. (1984) *Trends Biochem. Sci. 9*, 79–80.
Bartsch, R. G. (1971) *Methods Enzymol. 23*, 644–649.
Bartsch, R. G. (1978) *Methods Enzymol. 53*, 329–340.
Cammack, R., Hucklesby, D., & Hewitt, E. J. (1978) *Biochem. J. 171*, 519–526.
Christner, J. A., Janick, P. A., Siegel, L. M., & Munck, E. (1983) *J. Biol. Chem. 258*, 11157–11164.
Christner, J. A., Munck, E., Kent, T. A., Janick, P. A., Salerno, J.-C., & Siegel, L. M. (1984) *J. Am. Chem. Soc. 106*, 6786–6794.
Cline, J. F., Janick, P. A., Siegel, L. M., & Hoffman, B. M. (1986) *Biochemistry 25*, 4647–4654.
Cowan, J. A., & Sola, M. (1990) *Inorg. Chem. 29*, 2176–2179.
Cowan, J. A., Upmacis, R. K., Beratan, D. N., Onuchic, J. N., & Gray, H. B. (1989) *Ann. N.Y. Acad. Sci. 550*, 68–84.
Crowe, B. A., Owen, P., Patil, D. S., & Cammack, R. (1983a) *Eur. J. Biochem. 137*, 191–196.
Crowe, B. A., Owen, P., & Cammack, R. (1983b) *Eur. J. Biochem. 137*, 185–190.
Gren, E. J. (1984) *Biochemie 66*, 1–29.
Han, S., Madden, J. F., Siegel, L. M., & Spiro, T. G. (1989) *Biochemistry 28*, 5477–5485.
Hatefi, Y., Ragan, C. I., & Galante, Y. M. (1985) in *The Enzymes of Biological Membranes 4*, pp 1–70, Plenum Press, NY.
Helms, L., Krey, G. D., & Swenson, R. P. (1990) *Biochem. Biophys. Res. Commun. 168*, 809–817.
Huynh, B. H., Kang, L., DerVartanian, D. V., Peck, H. D., & Le Gall, J. (1984) *J. Biol. Chem. 259*, 15373–15376.
Janick, P. A., & Siegel, L. M. (1983a) *Biochemistry 22*, 504–515.
Janick, P. A., & Siegel, L. M. (1983b) *Biochemistry 21*, 3538–3547.
Janick, P. A., Rueger, D. C., Krueger, R. J., Barber, M. J., & Siegel, L. M. (1983) *Biochemistry 22*, 396–408.
Krey, G. D., Vanin, E. F., & Swenson, R. P. (1988) *J. Biol. Chem. 263*, 15436–15443.
Krueger, R., & Siegel, L. M. (1982) *Biochemistry 27*, 2905–2909.
Lancaster, J. R., Vega, J. M., Kamin, H., Orme-Johnson, N. R., Orme-Johnson, W. H., Krueger, R. J., & Siegel, L. M. (1979) *J. Biol. Chem. 254*, 1268–1272.
Lee, J. P., Le Gall, J., & Peck, H. D. (1973) *J. Bacteriol. 115*, 529–542.
Le Gall, J., & Postgate, J. R. (1973) *Adv. Microbiol. Physiol. 10*, 81–133.
Madden, J. F., Han, S., Siegel, L. M., & Spiro, T. G. (1989) *Biochemistry 28*, 5471–5477.
McRee, D. E., Richardson, D. C., Richardson, J. S., & Siegel, L. M. (1986) *J. Biol. Chem. 261*, 10277–10281.
Mortenson, L. E., & Thorneley, R. N. F. (1979) *Annu. Rev. Biochem. 48*, 387–418.
Moura, I., Lina, A. R., Moura, J. J. G., Xavier, A. V., Fauque, G., Peck, H. D., & Le Gall, J. (1986) *Biochem. Biophys. Res. Commun. 141*, 1032–1041.
Moura, I., Le Gall, J., Lino, A. R., Peck, H. D., Fauque, G., Xavier, A. V., DerVartanian, D. V. G., Moura, J. J. G., & Huynh, B. H. (1988) *J. Am. Chem. Soc. 110*, 1075–1082.
Murphy, M. J., Siegel, L. M., Kamin, H., & Rosenthal, D. (1973) *J. Biol. Chem. 248*, 2801–14.
Murphy, M. J., Siegel, L. M., Tove, S. R., & Kamin, H. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 612–616.
Ostrowski, J., Wu, J.-Y., Ruegers, D. C., Miller, B. E., Siegel, L. M., & Kredich, N. M. (1989) *J. Biol. Chem. 264*, 15726–15737.
Papa, S. (1983) in *Membranes and Transport*, pp 363–368, Plenum Press, NY.

Peck, H. D., & Le Gall, J. (1982) *Phil. Trans. R. Soc. London B298*, 443–466.
Pierik, A. J., & Hagen, W. R. (1991) *Eur. J. Biochem. 195*, 505–516.
Postgate, J. R. (1984) in *The Sulphate-Reducing Bacteria*, p 86, Cambridge University Press, New York.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning*, Cold Spring Harbor Laboratory Press.
Siegel, L. M. (1975) in *Metabolism of Sulfur Compounds* (Greenberg, D. M., Ed.) pp 217–286, Academic Press, NY.
Siegel, L. M., & Davis, P. S. (1974) *J. Biol. Chem. 249*, 1587–98.
Siegel, L. M., Murphy, M. J., & Kamin, H. (1973) *J. Biol. Chem. 248*, 251–264.
Siegel, L. M., Davis, P. S., & Kamin, H. (1974) *J. Biol. Chem. 249*, 1572–1586.
Siegel, L. M., Rueger, D. C., Barber, M. J., Krueger, R. J., Orme-Johnson, N. R., & Orme-Johnson, W. H. (1982) *J. Biol. Chem. 257*, 6343–6350.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503.
Tan, J., & Cowan, J. A. (1991) *Biochemistry 31*, 8910–8917.
Vega, J. M., & Kamin, H. (1977) *J. Biol. Chem. 252*, 896–909.
Wilkerson, J. O., Janick, P. A., & Siegel, L. M. (1983) *Biochemistry 22*, 5048.
Witt, S. N., Blair, D. F., & Chan, S. E. (1986) *J. Biol. Chem. 261*, 8104–8107.
Wood, W. I., Gitschier, J., Lasky, L. A., & Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1585–1588.
Young, L. J., & Siegel, L. M. (1988) *Biochemistry 27*, 2790–2800.

# Hydrogen Exchange in Thermally Denatured Ribonuclease A†

Andrew D. Robertson‡ and Robert L. Baldwin*
*Biochemistry Department, Stanford University School of Medicine, Stanford, California 94305*
*Received May 1, 1991; Revised Manuscript Received August 6, 1991*

ABSTRACT: Hydrogen exchange has been used to test for the presence of nonrandom structure in thermally denatured ribonuclease A (RNase A). Quenched-flow methods and 2D $^1$H NMR spectroscopy were used to measure exchange rates for 36 backbone amide protons (NHs) at 65 °C and at pH* (uncorrected pH measured in $D_2O$) values ranging from 1.5 to 3.8. The results show that exchange is approximately that predicted for a disordered polypeptide [Molday, R. S., Englander, S. W., & Kallen, R. G. (1972) *Biochemistry 11*, 150–158]; we thus are unable to detect any stable hydrogen-bonded structure in thermally denatured RNase A. Two observations suggest, however, that the predicted rates should be viewed with some caution. First, we discovered that one of the approximations made by Molday et al. (1972), that exchange for valine NHs is similar to that for alanine NHs, had to be modified; the exchange rates for valine NHs are about 4-fold slower. Second, the pH minima for exchange tend to fall at lower pH values than predicted, by as much as 0.45 pH units. These results are in accord with those of Roder and co-workers for bovine pancreatic trypsin inhibitor [see Table I in Roder, H., Wagner, G., & Wüthrich, K. (1985) *Biochemistry 24*, 7407–7411]. The origin of the disagreement between predicted and observed pH minima is unknown but may be the high net positive charge on these proteins at low pH. In common with some other thermally unfolded proteins, heat-denatured ribonuclease A shows a significant circular dichroism spectrum in the far-ultraviolet region [Labhardt, A. M. (1982) *J. Mol. Biol. 157*, 331–355]. This spectrum is reduced to a low level by adding guanidine hydrochloride. The nature of the residual structure responsible for this spectrum is not known. Our results show that it is not stable helix formation by the three $\alpha$-helices of native ribonuclease A, which would give measurable protection against amide proton exchange.

One of the primary goals of modern protein chemistry is to understand the structural basis for the thermodynamic stability of proteins. For most proteins, only two states, the native and denatured states, are populated to a significant extent at equilibrium and it is the modest free energy difference between these states that determines the stability of a protein (Becktel & Schellman, 1987). X-ray crystallography and, more recently, NMR[1] spectroscopy provide high-resolution information about the structure of native proteins, which has guided much of our thinking about the energetics of protein folding (Alber, 1989). Such detailed knowledge is lacking for the denatured state; until recently, structural information for the denatured state has been derived almost exclusively from studies of its macroscopic properties. The absence of more

† This work was supported by the National Institutes of Health (GM 19988). A.D.R. is a fellow of the Damon Runyon-Walter Winchell Cancer Fund (DRG-970).
‡ Present address: Department of Biochemistry, College of Medicine, University of Iowa, Iowa City, IA 52242

[1] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; CD, circular dichroism; COSY, two-dimensional chemical shift correlation spectroscopy; FID, free induction decay; FMOC, 9-fluorenylmethoxycarbonyl; FPLC, fast protein liquid chromatography; GuHCl, guanidine hydrochloride; NMR, nuclear magnetic resonance; PDLA, poly(D,L-alanine); pH*, glass electrode pH reading of $D_2O$ solutions without correction for isotope effects; $pK_{D_2O}$, $pK$ of $D_2O$; RELAY, relayed coherence transfer spectroscopy; RNase A, bovine pancreatic ribonuclease A (EC 3.1.27); [$^2$H]TSP, sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]-propionate.

0006-2960/91/0430-9907$02.50/0 