Assembly of Iron-Sulfur Clusters

IDENTIFICATION OF AN iscSUA-hscBA-fdx GENE CLUSTER FROM AZOTOBACTER VINELANDII*

(Received for publication, July 3, 1997, and in revised form, February 2, 1998)

Limin Zheng†, Valerie L. Cash‡, Dennis H. Flint§, and Dennis R. Dean∥

From the †Department of Biochemistry, Virginia Tech, Blacksburg, Virginia 24061-0346 and §E. I. du Pont de Nemours and Co., Central Research and Development, Experimental Station, Wilmington, Delaware 19890-0328

An enzyme having the same l-cysteine desulfurization activity previously described for the NifS protein was purified from a strain of Azotobacter vinelandii deleted for the nifS gene. This protein was designated IscS to indicate its proposed role in iron-sulfur cluster assembly. Like NifS, IscS is a pyridoxal-phosphate containing homodimer. Information gained from microsequencing of oligopeptides obtained by tryptic digestion of purified IscS was used to design a strategy for isolation and DNA sequence analysis of a 7,886-base pair A. vinelandii genomic segment that includes the iscS gene. The iscS gene is contained within a gene cluster that includes homologs to nifU and another gene contained within the major nif cluster of A. vinelandii previously designated orf6. These genes have been designated iscU and iscA, respectively. Information available from complete genome sequences of Escherichia coli and Hemophilus influenzae reveals that they also encode iscSUA gene clusters. A wide conservation of iscSUA genes in nature and evidence that NifU and NifS participate in the mobilization of iron and sulfur for nitrogen-specific iron-sulfur cluster formation suggest that the products of the iscSUA genes could play a general role in the formation or repair of iron-sulfur clusters. The proposal that IscS is involved in mobilization of sulfur for iron-sulfur cluster formation in A. vinelandii is supported by the presence of a cysE-like homolog in another gene cluster located immediately upstream from the one containing the iscSUA genes. O-Acetylserine synthase is the product of the cysE gene, and it catalyzes the rate-limiting step in cysteine biosynthesis. A similar cysE-like gene is also located within the nif gene cluster of A. vinelandii. The likely role of such cysE-like gene products is to increase the cysteine pool needed for iron-sulfur cluster formation. Another feature of the iscSUA gene cluster region from A. vinelandii is that E. coli genes previously designated as hscB, hscA, and fdx are located immediately downstream from, and are probably co-transcribed with, the iscSUA genes. The hscB, hscA, and fdx genes are also located adjacent to the iscSUA genes in both E. coli and H. influenzae. The E. coli hscA and hscB gene products have previously been shown to bear primary sequence identity when respectively compared with the dnaK and dnaJ gene products and have been proposed to be members of a heat-shock-cognate molecular chaperone system of unknown function. The close proximity and apparent co-expression of iscSUA and hscBA in A. vinelandii indicate that the proposed chaperone function of the hscBA gene products could be related to the maturation of iron-sulfur cluster-containing proteins. Attempts to place non-polar insertion mutations within either A. vinelandii iscS or hscA revealed that such mutations could not be stably maintained in the absence of the corresponding wild-type allele. These results reveal a very strong selective pressure against the maintenance of A. vinelandii iscS or hscA knock-out mutations and suggest that such mutations are either lethal or highly deleterious. In contrast to iscS or hscA, a strain having a polar insertion mutation within the cysE-like gene was readily isolated and could be stably maintained. These results show that the cysE-like gene located upstream from iscS is not essential for cell growth and that the cysE-like gene and the iscSUA-hscBA-fdx genes are contained within separate transcription units.

Iron-sulfur clusters ([Fe-S] clusters) are found in numerous proteins that have important redox, catalytic, or regulatory properties, yet the mechanism(s) by which such clusters are formed or repaired in vivo is not known. We have previously characterized the nifU and nifS gene products (NifU and NifS) that are proposed to have specific roles in the formation or repair of the [Fe-S] cores of metalloglutathione clusters contained within the catalytic components of nitrogenase (1, 2). NifS is a pyridoxal phosphate (PLP)1-dependent L-cysteine desulfurase (1), and it is able to catalyze the in vitro reconstitution of an apo-form of the nitrogenase Fe protein whose [4Fe-4S] cluster has been removed by chelation (3). The active species in this reaction is an enzyme-bound persulfide that is formed through nucleophilic attack by an active site cysteine on the PLP-substrate cysteine adduct (4). Although the specific role of NifU is not yet known, we have suggested that it might function either to deliver the iron necessary for [Fe-S] cluster formation or to provide an intermediate site for [Fe-S] cluster assembly (2, 5). Because the turnover rate for nitrogenase is so slow (6), cells that depend on nitrogen fixation as their sole source of metabolizable nitrogen must accumulate large amounts of the nitrogenase polypeptides. For example, the nitrogenase Fe protein and MoFe protein comprise approximately 5–10% of the total soluble protein fraction in nitrogen fixing Azotobacter vinelandii cells (6). Because each Fe protein contains four iron and four
inorganic sulfur atoms, and each MoFe protein αβ-unit contains 15 iron and 16 inorganic sulfur atoms (7), this situation must place a great demand on the physiological mobilization of the iron and sulfur necessary for the maturation of the nitrogenase catalytic components. Thus, we have suggested the possibility that the reactions catalyzed by NifU and NifS might represent a specialized way to boost the mobilization of iron and sulfur necessary for nitrogenase maturation and that there might also be NifU- and NifS-like “housekeeping” counterparts involved in the formation or repair of [Fe-S] clusters present in other iron-sulfur proteins (1, 5). Several lines of evidence support these ideas. First, a requirement for a boost in the mobilization of sulfur for [Fe-S] cluster assembly under nitrogen-fixing conditions is indicated by the presence of a nif-specific cysE-like homolog in *A. vinelandii* (8, 9) The cysE gene product encodes an O-acetylserylserine synthase that catalyzes the rate-limiting step in cysteine biosynthesis (10). Thus, an increase in cysteine formation when *A. vinelandii* is grown under nitrogen-fixing conditions probably occurs in response to a demand for NifS activity which utilizes t-cysteine as substrate in the mobilization of sulfur for nitrogenase [Fe-S] cluster formation. Second, we have purified a NifS-like protein from the non-nitrogen-fixing organism *Escherichia coli* and have shown that this protein is also able to activate sulfur for [Fe-S] cluster formation *in vitro* (11). Third, genome sequencing projects have revealed the presence of NifU- and NifS-like homologs in a variety of other non-nitrogen-fixing organisms, including *E. coli* (12), *Hemophilus influenzae* (13), *Saccharomyces cerevisiae* (14), and human (15). In the present work we have identified and sequenced a genomic region from *A. vinelandii* that encodes genes that are homologous to nifS, nifU, as well as other genes whose products could have roles in the maturation of [Fe-S] cluster-containing proteins.

**MATERIALS AND METHODS**

**Chemicals, Assays, and Media—**All chemicals were obtained from Sigma unless otherwise noted. Assays for cysteine desulfurase activity (NifS activity) were conducted by following the production of alanine, sulfur, and S2− from t-cysteine as described previously (16). For assay of IscS purified from *A. vinelandii*, PLP was added to a concentration of 10 μM. Protein was measured by the Bradford method (17).

**Protein Purification—**For purification of IscS from *A. vinelandii* strain DJ116 (∆nifS) (19), approximately 1 kg of cell paste of strain DJ116 was mixed with 2 volumes of buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl2, and 1 mM dithiothreitol (DTT)) and ruptured in a microfluidizer. Insoluble cell debris was removed by centrifugation. A solution of 10 mg/ml proteaminate was slowly added to the supernatant to give a final protamine sulfate concentration that was 5% by weight of the total protein in the sample. Precipitate was removed by centrifugation, and the soluble fraction was applied to a 5-liter Q-Sepharose column and eluted with the above buffer that contained a linear gradient of increasing KCl. Fractions were assayed for t-cysteine desulfurase (IscS) activity, and fractions from the large peak of activity were pooled. After bringing the pooled fraction to 10% w/v in (NH4)2SO4, it was applied to a 3.5-liter phenyl-Sepharose column and eluted with a decreasing (NH4)2SO4 gradient. Fractions exhibiting IscS activity were pooled, dialyzed, and loaded onto a 400-ml Sigma Green 10 column and subsequently eluted with an increasing KCl gradient. The active fractions were pooled and loaded onto a Superdex 25/600 column and eluted with buffer. Active fractions from the Superdex column were pooled and chromatographed on a 20-ml Mono-Q column using a linear increasing KCl gradient. The active fractions were pooled, and the protein was concentrated to about 10 mg/ml and pelleted in liquid N2 until used.

Q-Sepharose and phenyl-Sepharose columns were run at room temperature using an Amersham Pharmacia Biotech BioPilot system and the other columns were run at 4 °C using an Amersham Pharmacia Biotech FPLC. In addition to normal activity assays, purification of IscS was monitored by non-denaturing gel electrophoresis using the lead stain method (20) to identify the band that exhibits IscS activity (i.e. release of L-cysteine) and by monitoring gel electrophoresis of the lead-stained bands.

Purification of recombinant IscS produced in *E. coli* was performed as described previously for NifS (1), except that the phenyl-Sepharose step was replaced by Sephacryl S-300 HR chromatography. The very high level of accumulation of IscS expressed in *E. coli* using a hybrid T7–7 plasmid construct (described below) was about the same as for NifS expressed in *E. coli*. IscS was recombinantly produced in the same way (1). Because of the abundance of IscS produced in this way, its purification could be monitored by the yellow color of its PLP cofactor and by denaturing SDS-polyacrylamide gel electrophoresis of sample fractions.

**Peptide Microsequence Analysis—**Tryptic peptides of IscS purified from *A. vinelandii* were generated by digestion of a reduced and alkylated (by iodoacetamide) sample with trypsin. The oligopeptides were separated on a Vydac C-18 column using a Waters high pressure liquid chromatograph. To identify the reactive cysteine residue, the protein was treated with 1.2 eq of N-iodoacetyl-N′-((5-sulfol-1-naphthyl)ethylendiamine) (1-AEDANS) before alkylation by iodoacetamide. Oligopeptides were separated by high pressure liquid chromatography as above, and the fluorescein oligopeptide fraction was identified by illumination with a blue lamp. Microsequence analysis was performed on a Peron-Elmer 5800 gas-phase sequencer equipped with a Beckman System Gold 125S pump module, Beckman PTH Microcolumn, and a 166 UV detector module. The molecular weight of several of the isolated oligopeptide fragments was determined by infusion electrospray into a Micromass Trio-2000 quadrupole mass spectrometer.

**DNA Biochemistry—**Oligonucleotides for polymerase chain reaction (PCR) amplification of a segment of the *iscE* gene from *A. vinelandii* DJ116 genomic DNA were designed based on the sequence of tryptic peptide fragments of IscS as described above and from the known codon usage bias for *A. vinelandii* genes (8). The PCR primers were used 5′-GCNACCCACCCGC/TGNTGAt/cCc/CC-3′ (forward primer) and 5′-GAANGc/cTCNCCCATCNCAGGT-3′ (reverse primer). “N” refers to any nucleotide, and lowercase letters in parentheses indicate a degeneracy for the two nucleotides indicated at that position. A *Stratagene* Robocycler Gradient 40 thermocycler was used to amplify a 0.72-kb segment of DNA. The 10-μl PCR reaction mixture contained 2.5 units of *Taq* polymerase (Perkin-Elmer), 0.2 mM MgCl2, 0.5 mM of each dNTP, and 0.8 μl of each primer. Cycling temperatures were 94 °C for 1.5 min, 65 °C for 2.5 min, and 72 °C for 3 min. The 0.72-kb DNA product was purified on a gDNA gel, and the 0.72-kb oligonucleotide band was excised and ligated in the cloning vector pUC118 (22) that had been digested with restriction enzyme *Hin*dIII. This hybrid plasmid is designated pDB918. The nucleotide sequence of the cloned fragment was determined by the Sanger dideoxy method (23). A genomic library of DJ116 DNA was prepared by partially digested chromosomal DNA with *XhoI*, and the sticky ends of the fragments were filled with T4 polymerase (Life Technologies, Inc.) in the presence of 1 mM dCTP and dTTP. The λ-DASH-II cloning vector DNA obtained from Stratagene was digested with *BamHI*, and its sticky ends were partially filled by incubation with *T4* DNA polymerase in the presence of 1 mM dATP and dGTP. A mixture of ligated vector and target DNA was packaged into λ phage using the Stratagene Gigapack II system using the conditions recommended by the supplier. The library had a titer of 5.0 × 1010 plaques/ml lysate, and the average insert size was about 20 kb. The phage were plated on LB medium using XL1-blue MRA cells (Stratagene) as recipient, and DNA was lifted from plaques using Amersham Pharmacia Biotech Hybond-N+ membranes. DNA was denatured by placing the membranes on top of a 0.5 N NaOH-saturated Whatman paper for 5 min and then washed in 75 mM sodium citrate, 750 mM NaCl, pH 7.0, for 1 min. The library was screened using the cloned 0.72-kb PCR-amplified DNA fragment as probe. Screening was done using an Amersham Pharmacia Biotech ECL kit following procedures recommended by the supplier. Prehybridization was done for 1 h at 42 °C, and hybridization was done for 4 h at the same temperature. Following the washing step and the color development step, the membrane was exposed to x-ray film over a period of 2 weeks. Five positive plaques were identified, and a second screening was performed using phage that were diluted from four positives detected in the original screening procedure. DNA was prepared from eight positive individual plaques identified in the second screen for further verification. Restriction enzyme mapping of DNA prepared from these samples revealed that three different overlapping genomic regions had been cloned. Southern analysis using the 0.72-kb

---

2 This information is available at the Yeast MIPS Web page (http://www.mips.biochem.mpg.de/mips/yeast/).

3 T. Rouault, personal communication.
fragment showed hybridization to a 4.0-kb EcoRI restriction enzyme fragment. This fragment could not be stably cloned into pUC119, but it was found that a 2.5-kb EcoRI-SstI fragment, which also contained iscS and flanking regions, could be stably cloned. This fragment was subsequently ligated into EcoRI-SstI-digested pUC119 and the resultant plasmid designated pDB933. A variety of other genomic fragments that are also contained within this genomic region were also cloned into pUC119 and sequenced. Among the most significant of these fragments are the following: a 2.5-kb EcoRI fragment (pDB948); a 1.5-kb SstI-EcoRI fragment (pDB932); a 1.0-kb EcoRI fragment (pDB946); and a 5.0-kb EcoRI fragment (pDB944). The nucleotide sequence of various DNA fragments was determined by the Sanger dideoxy method (23).

The strategy and methods for placing the iscS gene under the control of the T7 transcription and translation control elements using the pt7 vector (24) were done as described previously for the nifS gene (1). Primers used for PCR amplification of the iscS gene were 5'-CATGAAGTTGCCGATTTAT-3' (forward primer) and 5'-CATG-9 (reverse primer).

**Mutagenesis of iscS, hscA, and the cysE-like Gene**—For mutagenesis of iscS, pDB932 DNA was digested with PstI and ligated to remove an approximately 900-base pair fragment that extends from the PstI site in iscU to the rightward PstI site contained within the polynuclein region of the original pUC119 cloning vector. The resultant plasmid was designated pDB941. Plasmid pDB941 contains three SstI restriction enzyme sites located within the iscS coding sequence. These fragments were excised from pDB941 by digestion with SstI and replaced by ligating the sample with a purified 1.4-kb SstI fragment that contains a kanamycin resistance (KmR) cartridge inserted from the vector pUC4-KAPA (purchased from Amersham Pharmacia Biotech). This plasmid was designated pDB952, and it contains a KmR gene cartridge whose direction of transcription is the same as the iscS transcription. We have previously found that, when in this orientation, the KmR insertion cartridge used in the present work is not polar upon expression of downstream A. vinelandii genes, but when present in the opposite direction, it is polar upon expression of downstream genes (8, 19). A similar strategy to the one used for construction of pDB952 was also used for constructing plasmids that have the same KmR cartridge inserted into either the cysE-like gene or the hscA gene. The parental plasmid for cysE-like gene mutagenesis was pDB948. Mutagenized plasmids that carry the KmR cartridge inserted within the cysE-like coding sequence either in the same or the reverse orientation as cysE-like coding sequence was designated pDB1008 and pDB1007, respectively. The parental plasmid used for hscA mutagenesis was pDB932. Mutagenized plasmids that carry the KmR cartridge inserted into the hscA-coding sequence either in the same or the reverse orientation as hscA transcription were designated pDB1109 and pDB1108, respectively. The respective insertion mutations contained within the various plasmid constructs were transferred to the A. vinelandii chromosome through double-reciprocal recombination events that occurred during transformation (25) using either wild-type or strain DJ116 (ΔnifS) as host and plasmid DNA as the donor. To ensure that incorporation of the KmR cartridge into the A. vinelandii chromosome occurred only via double-reciprocal recombination events, plasmids DNAs were linearized by restriction enzyme digestion prior to transformation. For those transformants chosen for further study, confirmation that double-reciprocal recombination did occur was demonstrated by showing that the resultant strain was AmpR and KmR. Because ColEI-derived plasmids are incapable of autonomous replication in A. vinelandii, any transformant arising from a single Campbell-like recombination event must have both KmR and AmpR phenotypes (8, 19).

### RESULTS

**Purification of a Protein Having a NifS-like Activity from A. vinelandii**—A protein that catalyzes the same enzymatic activity as NifS was isolated from extracts prepared from an A. vinelandii strain (DJ116) deleted for the nifS gene. We have designated the NifS-like protein from A. vinelandii "IcsS" to indicate its proposed role in iron-sulfur-cluster formation, and that nomenclature will be used hereafter. Crude extracts were sequentially processed by proteamine sulfate treatment and column chromatography using Q-Sepharose, phenyl-Sepharose, Sigma Green-19, Superdex, and Mono-Q columns, respectively. Fractions having IcsS activity were monitored by their ability to release sulfur or S2- and alanine from L-cysteine. Fractionation of the extract by Q-Sepharose chromatography yielded a minor peak and a large major peak of IcsS activity. The fractions containing the major peak were pooled and further purified, whereas the minor peak fraction was not processed further. When the material in the large activity peak was purified using the columns listed above, a single protein whose purity was judged to be greater than 90%, by native and denaturing gel electrophoresis, was obtained. Data from a typical isolation are shown in Table I.

**Characterization of IcsS—**IcsS isolated from A. vinelandii extracts is yellow, and the optical spectrum indicates the presence of PLP (Fig. 1). When 1 eq of L-cysteine was added to the enzyme, the visible region of the optical spectrum changed as shown in Fig. 1. These spectra are similar to those exhibited by the A. vinelandii NifS protein (1) and the E. coli NifS-like protein (11). When DTT was present in the assay, IcsS catalyzed the desulfurization of L-cysteine to yield both alanine and S2- in approximately equal amounts. In the absence of DTT, the enzyme catalyzed the release of both sulfur and S2- from L-cysteine, the sum of which approximately equaled the amount of alanine formed. Apparently, some of the sulfur released from L-cysteine was reduced to S2- in the absence of DTT by the excess of substrate cysteine present in the reaction. Treatment of IcsS with 1.2 eq of the alkylating reagent I-AEDANS resulted in the complete inhibition of its L-cysteine desulfurization activity. These results are the same as found for the A. vinelandii NifS protein (1) and the NifS-like protein purified from E. coli (11).

**Microsequencing Experiments**—For peptide microsequencing experiments a highly purified form of IcsS was obtained by further processing a sample prepared as described above by first removing DTT from the buffer, binding of the sample to a thiopropyl-Sepharose column, followed by its elution with 5 mM DTT. The protein prepared in this way was subjected to automated N-terminal sequence analysis, which gave the sequence KLPYILDSATTPVDPRTAQKMF. This sequence matches the polypeptide sequence deduced from the gene sequence determined in experiments described below, except that the termi-

### Table I

| Fraction       | Volume | Protein | Activity<sup>a</sup> | Specific activity<sup>b</sup> | Yield | Purification |
|----------------|--------|---------|----------------------|--------------------------------|-------|--------------|
| Crude extract  | 1700   | 82      | NA<sup>c</sup>       | 1                              | 1     | 670          |
| Protamine sulfate | 1700 | 76      | 7.7                  | 1.9                            | 100   | 19           |
| DEAE-Sepharose | 2200   | 4       | 5.5                  | 7.0                            | 75    | 70           |
| Phenyl-Sepharose | 900  | 0.81    | 4.7                  | 16.2                           | 51    | 160          |
| Sigma Green 19 | 200    | 0.29    | 2.5                  | 56.8                           | 32    | 570          |
| Superdex 35/600 | 33    | 0.044   | 2.2                  | 67.6                           | 29    | 670          |
| Mono Q        | 6      | 0.033   |                      |                                |       |              |

<sup>a</sup> Activity is expressed as nanomoles of product per min of assay.  
<sup>b</sup> Specific activity is expressed as nanomoles of product per min/mg of protein.  
<sup>c</sup> Not assayed.
nal methionine is not present in the isolated protein. Following tryptic digestion of purified IscS, the resulting oligopeptides were separated by high pressure liquid chromatography using a C-4 column. Primary sequences were determined for five of the major oligopeptides separated by this method, and the sequence of four of these was corroborated by mass spectrometry. The sequences of these peptides were all readily aligned with portions of the primary sequence of the A. vinelandii NifS protein and other NiFS-like proteins known to share considerable sequence identity with NiFS (Fig. 2). The five oligopeptide sequences determined by microsequencing show a nearly perfect match to the corresponding polypeptide sequences deduced from the DNA sequence, and for those sequences that were confirmed by mass spectrometry, there is a perfect match (Fig. 2).

A tryptic digest was also prepared from a sample of IscS that had been alkylated by prior treatment with 1.2 eq of I-AEDANS. When the oligopeptides from this sample were separated by high pressure liquid chromatography as above, one fraction also exhibited fluorescence. The oligopeptide contained within the fluorescent fraction was sequenced and found to be identical to a previously sequenced peptide shown in Fig. 2, with the exception that no cysteine was found at the residue 10 position of that peptide. This result indicates that the cysteine was modified by I-AEDANS. The alignments in Fig. 2 show that this IscS cysteine residue (cysteine 328) corresponds to cysteine 325 in the NiFS sequence. The NiFS cysteine 325 residue was previously shown by alkylation experiments and amino acid substitution experiments to be the active site cysteine upon which the persulfide is formed (4).

Purification and Characterization of Recombinantly Produced IscS—As described below and under “Materials and Methods,” primary sequence information obtained from microsequencing of IscS tryptic peptides provided the information necessary to design a strategy for isolation of a genomic DNA fragment that encodes IscS. The cloned iscS gene was then used to construct a hybrid plasmid for which the expression of iscS is placed under the control of the T7 transcriptional and translational control elements. By using this construct, it was possible to produce A. vinelandii IscS at high levels in E. coli. IscS produced in this way was purified as described under “Materials and Methods.” Recombinantly produced A. vinelandii IscS exhibited the same properties as described for IscS purified from A. vinelandii. Gel exclusion chromatography showed that recombinant IscS is a homodimer that exhibits a native $M_r$ of approximately 104,000. When assayed in the presence of DTT, the maximum activity obtained for IscS (124 nmol of product/min/mg) compared favorably with the maximum activity of purified NiFS assayed under the same conditions (168 nmol of product/min/mg). However, a true maximum activity for IscS could not be obtained nor could kinetic parameters be determined, because L-cysteine concentrations greater than 0.2 mM inhibit activity, and product detection limitations prevent assay of IscS at L-cysteine concentrations lower than 0.2 mM. This pattern of inhibition of IscS by L-cysteine was not observed for NiFS. It is noted that the specific activity of purified recombinant IscS, when assayed under the conditions reported in Table I (2.5 mM L-cysteine), gives approximately the same specific activity as IscS purified from A. vinelandii.

Isolation of the iscSU Genetic Region—Based on the oligopeptide sequences obtained in the microsequencing experiments, and the known codon bias for A. vinelandii, oligonucleotide primers were designed for PCR amplification of a segment of the iscS gene from genomic A. vinelandii DNA prepared from strain DJ116. The authenticity of this fragment was confirmed.
by DNA sequence analysis, and the fragment was then used to probe an *A. vinelandii* genomic library prepared using the λ-DASH-II cloning vector as host. A 4.0-kb EcoRI restriction enzyme fragment located within a hybrid λ phage exhibited hybridization to the PCR-amplified *iscS* fragment. A 2.5-kb EcoRI-SstI subfragment of the 4.0-kb EcoRI fragment was subsequently subcloned into pUC119, and its nucleotide sequence was determined. Analysis of the nucleotide sequence of this fragment revealed three complete coding sequences (Fig. 3), and these coding sequences respectively exhibited sequence identity when compared with NifS, NifU, and the deduced primary sequence of another gene product previously designated Orf6 (19). This gene is encoded within the major *nif* cluster of *A. vinelandii* (8) and is located adjacent to *nifU* (indicated as “iscA” in Fig. 3). We have provisionally designated the *nifS* homolog *iscS*, the *nifU* homolog *iscU*, and the *orf6* homolog as *iscA* to indicate their potential roles in iron-sulfur-cluster assembly. Although, the *iscSUA* genes are clustered, they are organized in an opposite fashion when compared with *iscSUA* homologs that are present in *E. coli* (12, 13) (Fig. 3). Furthermore, a general conservation in the relative genomic positions of these clusters in *A. vinelandii*, *E. coli*, and *H. influenzae* is indicated by interspecies conservation in sequences located both upstream and downstream of the respective *isc* gene cluster region.

The deduced primary sequence of the *E. coli* gene we have designated *iscS* corresponds to the NiFS-like protein previously isolated from *E. coli* (11). This conclusion is based on comparison of the deduced primary sequence from the gene sequence to oligopeptide sequences obtained from microsequencing of tryptic digests of the isolated protein.

A search of the data banks revealed that proteins homologous to *iscSUA* are widely distributed in nature. For example, apparent counterparts to *iscS*, *iscU*, or *iscA* can be found in *E. coli* (12), *H. influenzae* (13), *Streptococcus pyogenes*, *Neisseria gonorrhoeae*, *Synechocystis*, *Archaeoglobus fulgidus* (26), *S. cerevisiae* (27), and human (15). In some cases, not all *isc* genes have been identified in the same organism. For example, no *iscU* gene is apparent in the *Synechocystis* genome, although this organism contains homologs to both *iscS* and *iscA*. Also, some organisms have more than one copy of a particular *isc* homolog. For example, there are two different genes contained within the *E. coli* genome that are homologous to *iscA*, and in *S. pyogenes*, there are at least two copies of *iscS* homologs. Even though homologs to the *isc* genes appear to be widespread throughout nature, it is clear that their respective functions cannot represent a universal mechanism for [Fe-S] cluster assembly as previously suggested (1). The reason for this is that no *isc* homologs are present in *Methanococcus jannaschii* even though this organism produces many [Fe-S] cluster-containing proteins (28). However, at least one Archeon, *A. fulgidus*, does contain genes corresponding to both *iscS* and *iscU* (two copies each, see Ref. 26).

### Primary Sequence Comparisons of *iscSUA* Homologs—The interspecies primary sequences among proposed *IscS* (Fig. 2), *IscU* (Fig. 4), and *IscA* (Fig. 5) proteins from *A. vinelandii*, *H. influenzae*, and *E. coli* are aligned with each other and with the *A. vinelandii* NiFS, NiFU, and *nif*-specific *IscA* homologs, respectively. In the case of *IscS* from *A. vinelandii* and *E. coli*, these proteins are legitimately assigned as having an *IscS* function because they have been purified and characterized. In other cases, *IscS* function can only be inferred from primary sequence comparisons. The criteria that we have used to tentatively assign *IscS* function to a particular protein sequence is as follows: (i) there must be strong primary sequence conservation throughout when compared with the *IscS* proteins having established cysteine desulfurase activity; (ii) the active site PLP-binding lysine residue corresponding to lysine 202 in the NiFS sequence must be conserved; and (iii) the active site cysteine residue corresponding to cysteine 325 in the NiFS sequence must be conserved in the same relative position within

---

5 Deposited by B. A. Roe, S. P. Lin, L. Song, X. Yuan, S. Clifton, and D. W. Dyer, and available from the Gonococcal Genome Sequencing Project Home page (http://www.genome.ou.edu/gono.html).

6 Available from the Cyanobase Genome Data Base Web page (http://www.kazusa.or.jp/cyano/cyano.orig.html).

7 T. Rouault, personal communication.
its primary sequence. Data base searches reveal that a number of putative proteins have been designated "NifS-like" based only on significant sequence similarity surrounding and including the PLP-binding active site lysine corresponding to lysine 202 from NifS. The primary sequences of several of these proteins, however, do not show conservation of the NifS active site cysteine 325 residue, so it is possible that they represent PLP-dependent proteins that have functions unrelated to the NifS-IscS family. This possibility has been proven for at least one of the NifS-like protein from *E. coli*, which has been purified and shown to have cysteine sulfinate desulfinase activity rather than cysteine desulfurase activity (29). A complete comparison of NifS-like primary sequences currently available in the data bank has been published by Mihara *et al.* (29).

Placement of proteins in the IscU or IscA families can be even more tenuous because there are no functional assays for these proteins. We place proteins in the IscU family based on their strong primary sequence conservation throughout when compared with IscU from *A. vinelandii* and if the three NifU cysteines (residues 35, 62, and 106), corresponding to those targeted as potential iron-binding or cluster assembly sites within NifU (2), are conserved in highly homologous segments of the primary sequence. We have recently found that these three cysteines within NifU comprise an iron-binding site that is separate and independent from the [2Fe-2S] cluster coordinating ligands provided by NifU cysteine residues 137, 139, 172, and 175. It should be noted that IscU proteins are considerably truncated when compared with NifU and do not contain residues corresponding to the [2Fe-2S]-coordinating cysteines found in NifU (Figs. 3 and 4).

In the case of IscA, assignment of homologous proteins is made on the basis of their primary sequence conservation throughout when compared with the *A. vinelandii* nif-specific IscA protein and if there is conservation of three cysteines corresponding to residues 35, 99, and 101 within the nif-specific IscA sequence (Fig. 5). Although no function for the IscA family of proteins is yet known, the conservation of the cysteine residues noted above makes them candidates to serve as follows: (i) an iron carrier, (ii) a sulfur carrier, or (iii) a site for [Fe-S] cluster intermediate assembly.

Homologs to the *E. coli* hscBA and fdx Genes Are Located Downstream from the *iscSUA* Gene Cluster—After submission of the sequence of the *iscSUA* gene cluster to the data bank, our attention was brought to the fact that, in *E. coli*, genes designated as hscB, hscA, and fdx are located immediately downstream from the proposed *iscSUA* gene cluster. The *hscB* and *hscA* genes encode proteins that bear a high degree of sequence identity when compared with the molecular chaperone proteins encoded by *dnaJ* and *dnaK*, respectively (30, 31). The location of the *hscBA* genes immediately downstream from *iscSUA* in *E. coli* has been published by Mihara *et al.* (29).
coli raises the intriguing possibility that these heat-shock cognate proteins, whose functions are not yet known, might serve as molecular chaperones that specifically assist in the maturation of [Fe-S] proteins. Subsequent sequence analysis of the A. vinelandii genome in the region immediately downstream from iscSUA revealed the presence of homologs to the hscBA genes, a gene that encodes a homolog for a [2Fe-2S] cluster-containing ferredoxin (32), and another putative gene that has not been characterized (Fig. 3). In contrast to E. coli, where there is a relatively large separation between iscA and hscB (96 nucleotides), only 16 nucleotides separate the termination codon of iscA and the initiation codon of hscB. This spatial organization indicates that the A. vinelandii iscSUA and hscBA genes (as well as genes indicated as foxs and orfβ in Fig. 3) are very likely to be co-transcribed. Inspection of the nucleotide sequence immediately following the gene indicated as orfβ in Fig. 3 indicates the likely presence of a rho-independent transcriptional termination signal. Also, there are 96 nucleotides separating orfβ and the next gene, which encodes nucleotide dikinase. The ndk gene is also found at the same relative position on the E. coli genome and is separated from the orfβ gene by 91 nucleotides (Fig. 3).

Identification of a cysE-like Gene That Is Located Upstream from the iscSUA Gene Cluster—Prompted by the identification of the hscBA genes and the potential involvement of their products in maturation of [Fe-S] proteins, the DNA sequence of the A. vinelandii genomic segment located immediately upstream from the iscS gene was also determined. The result of this analysis revealed that another putative gene (indicated as orf2 in Fig. 3) precedes iscS and is separated from it by only 28 nucleotides. A homolog to this gene is also present in the A. vinelandii genome, where there is a relatively large separation between orf2 and orf1 by 135 nucleotides. This spatial organization indicates that orf1 and cysE2 are missing the first 75 amino acids when compared with the bona fide cysE gene product from E. coli (Fig. 6).

Genetic Analysis of the isc Gene Cluster Region—In order to attempt inactivation of the chromosomally encoded iscS gene, a plasmid (pDB952) was constructed that has a large portion of the iscS gene deleted and the deleted segment replaced by a KmR gene cartridge. The deletion and insertion within pDB952 was incorporated into the genomes of A. vinelandii strains DJ116 (∆nisS) by transformation followed by selection for KmR and scoring for AmpS. The strain constructed in this way exhibited a normal growth rate, initially indicating that no obvious phenotype is associated with the inactivation of the iscS gene. However, when these cells were cultured in the absence of kanamycin, the KmR phenotype was rapidly lost. For example, after growing these cells for approximately 15 generations without selective pressure, no KmR cells could be recovered. In another experiment, single KmR colonies were picked and diluted in liquid medium and then spread on Petri plates supplemented with or without added kanamycin. The ratio of KmR to KmS cells was found to be approximately 1 to 5 indicating that, even under kanamycin-selective pressure, cells harboring only the wild-type iscS allele rapidly segregate from those harboring the KmR cartridge. A homologous excision event cannot account for the rapid loss of the KmR marker because the original integration of the KmR cartridge occurred through double-reciprocal recombination. These results show that there is strong selective pressure to maintain the wild-type iscS allele and that both the wild-type and the inactivated iscS alleles are maintained in the same cells when placed under the appropriate selection pressure. There have been two previous reports of the same phenomenon occurring in A. vinelandii where strong selective pressure was used to maintain both a

---

9 L. Vickery, personal communication.
gene that was interrupted by an antibiotic resistance marker and the corresponding wild-type allele (33, 34). That such a phenomenon can occur in A. vinelandii has been attributed to its proposed ability to maintain multiple identical chromosomess (34, 35). It is also noted that there are many cases of insertional activation of non-essential genes in A. vinelandii which rapidly undergo homogenization and remain stable in the absence of antibiotic-selective pressure (8, 19).

In the present work the ability of an A. vinelandii cell population to maintain both wild-type and insertional inactivated iscsS alleles was determined by a series of PCR experiments using DNA isolated from a KmR isolate obtained after transforming DJ116 with pDB952 DNA. This isolate was maintained under KmR-selective pressure. DNA isolated from the parental DJ116 cells served as the control, and the results of these experiments are shown in Fig. 7. In the first experiment, PCR primers were used to amplify a fragment of DNA corresponding to that located within (reverse primer), and just outside (forward primer), the iscsS-coding sequence deleted from pDB952. If the wild-type allele persists in the transformed strain, then an identical 0.72-kb product should accumulate from PCR amplification when using either wild-type or transformed cell DNA. This result is shown in lanes 1 and 2 of Fig. 7. In the second experiment, PCR primers were used to amplify an internal segment of the KmR gene cartridge. The results of these experiments show that PCR amplification using the parental DJ116 strain genomic DNA does not result in any product accumulation but that PCR amplification using transformed cell genomic DNA results in accumulation of the predicted 0.91-kb KmR gene fragment (lanes 3 and 4 in Fig. 7). A final DNA amplification experiment was performed to determine if the KmR gene cartridge present in DNA from the transformed cells is correctly located within the predicted genomic position. This was accomplished by using a forward PCR primer that corresponds to a sequence located within the KmR cartridge and a reverse primer that corresponds to a sequence located downstream from the iscsS region deleted in pDB952. If the KmR gene cartridge present in a population of the transformed cell genomic DNA is inserted within the iscsS gene, and in the same transcriptional orientation as iscsS gene transcription, then a PCR product of 1.29-kb is predicted from amplification of transformed cell DNA. Also, because there is no KmR cartridge insert within DJ116 DNA, PCR amplification using DJ116 DNA should result in no product accumulation. These predictions were also confirmed as shown in lanes 5 and 6 in Fig. 7. In summary, the results of these experiments show that a non-polar insertion mutation within iscsS can only be maintained if a wild-type iscsS allele is also present. These results, and the rapid loss of the insertional inactivated allele under non-selective conditions, demonstrate that inactivation of iscsS is either lethal or is highly deleterious. In a series of experiments similar to those described above, both polar and non-polar insertion mutations (carried by plasmids pDB1005 and pDB1009, respectively) were also placed within the hscA coding sequence (Fig. 3). As in the case of the iscsS non-polar insertion mutation, neither a polar nor a non-polar insertion mutation within hscA could be stably maintained in the absence of the wild-type hscA allele. The same PCR approach described above was also used to analyze DNA isolated from A. vinelandii cells that were transformed with pDB1005 DNA and cultured under KmR-selective pressure. This analysis revealed that DNA from such cells also maintained both the wild-type and insertional inactivated hscA alleles. Thus, like iscsS, inactivation of hscA is either lethal or highly deleterious. These results contrast with similar experiments with E. coli (30, 31) where it was shown that hscA is not essential for cell growth.

Insertional inactivation of cysE2 was accomplished by transforming wild-type A. vinelandii cells with plasmids that contain either polar (pDB1007) or non-polar (pDB1008) KmR insertion mutations within the cloned cysE2 gene (Fig. 3). In contrast to the iscsS and hscA genes, an insertion mutation within cysE2 could be stably maintained with no segregation of KmR cells even when cultured for extended periods under non-selective conditions. Thus, cysE2 gene expression is not essential for cell growth. Also, because a polar mutation within cysE2 can be stably maintained, the expression of the downstream gene cluster that contains iscSU and hscBA must be able to occur in the absence of cysE2 expression.
conditions of non-selective antibiotic pressure, providing that the cells are cultured under diazotrophic growth conditions. However, it was found that an iscS Km<sup>+</sup> insertion mutation could not be stably maintained even when the transformants were cultured under diazotrophic growth conditions. Thus, either NifS or IscS activity is not significantly interchangeable or the amount of NifS needed to replace IscS function is not accumulated under the experimental conditions used. Evidence supporting this latter possibility was recently reported (35). In these experiments it was shown that heterologous co-expression of nifS from <i>A. vinelandii</i> and glutamine phosphoribosylpyrophosphate amidotransferase from <i>Bacillus subtilis</i> in <i>E. coli</i> boosted the amount of the mature [4Fe-4S] cluster-containing <i>B. subtilis</i> enzyme accumulated.

**DISCUSSION**

The IscS protein from <i>A. vinelandii</i> was purified based on its ability to catalyze the release of sulfur from l-cysteine. IscS exhibits strong primary sequence identity when compared with NifS and catalyzes the same reaction as NifS. Isolation and sequence analysis of a genomic fragment containing iscS revealed that it is clustered, and probably co-transcribed, with seven other genes. There is currently good evidence that both NifS and IscS activity is not significantly interchangeable or that both NifU and IscS are involved in the formation or repair of [Fe-S] clusters needed for activation of nitrogenase components (1, 2).

By analogy to the nif system we suggest that IscS, IscU, and iscC are likely to have housekeeping functions related to the assembly or repair of [Fe-S]-containing proteins other than nitrogenase (18). The location and apparent co-expression of iscSU and iscBA genes indicate the possible presence of a macromolecular system that functions in the proper folding of [Fe-S]-containing proteins, as well as participating in the formation and insertion of their corresponding clusters. The presence of the <i>cysE2</i> gene in the transcription unit located upstream from <i>iscS</i> in <i>A. vinelandii</i> also suggests the possibility of a mechanism for the specific targeting of l-cysteine for [Fe-S] cluster formation.

**Acknowledgments**—We thank those investigators that have made genome sequence analyses available prior to publication. We also acknowledge the expert technical assistance of Christina OKernick and Laura Taylor. We are particularly grateful to professor Larry Vickery for pointing out the hscBA genes and suggesting the potential role of their products in maturation of [Fe-S] cluster-containing proteins.

**REFERENCES**

1. Zheng, L., White, R. H., Cash, V. L., Jack, R. F., and Dean, D. R. (1993) <i>J. Biol. Chem.</i> 268, 18723–18726
2. Zheng, L., White, R. H., Cash, V. L., and Dean, D. R. (1994) <i>Biochemistry</i> 33, 13445–13463
3. Zheng, L., White, R. H., Cash, V. L., Jack, R. F., and Dean, D. R. (1993) <i>J. Biol. Chem.</i> 268, 18723–18726
4. Zheng, L., White, R. H., Cash, V. L., and Dean, D. R. (1994) <i>Biochemistry</i> 33, 4714–4720
5. Dean, D. R., Bolin, J. T., and Zheng, L. (1993) <i>J. Bacteriol.</i> 175, 6737–6744
6. Shah, V. K., and Brill, W. J. (1973) <i>Biochem. Biophys. Acta</i> 305, 445–454
7. Howard, J. B., and Rees, D. C. (1996) <i>Chem. Rev.</i> 96, 2965–2982
8. Jacobson, M. R., Brigue, K. E., Bennett, L. T., Setterquist, R. A., Wilson, M. S., Cash, V. L., Beynon, J., Newton, W. E., and Dean, D. R. (1989) <i>J. Bacteriol.</i> 171, 1017–1027
9. Evans, D. J., Jones, R., Woodley, P. R., Wilborn, J. R., and Rosbon, R. L. (1991) <i>J. Bacteriol.</i> 173, 5457–5469
10. Denk, D., and Boek, A. J. (1987) <i>J. Gen. Microbiol.</i> 132, 515–525
11. Flint, D. H. (1996) <i>J. Biol. Chem.</i> 271, 16068–16074
12. Blattner, F. R., Plunkett, G., III, Blisch, A. C., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glaser, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) <i>Science</i> 277, 1453–1462
13. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., K Earlavaz, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L.-I., Glinek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D., Sondek, D. M., Brandon, R. C., Fine, L. D., Frischmann, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) <i>Science</i> 269, 496–512
14. Oliver, S. G., der Aart, Q. J., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., Alexandraki, D., Antoine, G., Anwar, R., Ballesta, J. P., Benit, P., et al. (1992) <i>Nature</i> 360, 38–46
15. Hwang, D. M., Dempsey, A., Tan, K. T., and Liew, C. C. (1996) <i>J. Mol. Biol.</i> 260, 780–784
16. Flint, D. H., Tuminello, J. F., and Miller, T. J. (1996) <i>J. Biol. Chem.</i> 271, 16053–16067
17. Bradford, M. M. (1976) <i>Anal. Biochem.</i> 72, 248–254
18. Strandberg, G. W., and Wilson, P. W. (1968) <i>Curr. Microbiol.</i> 14, 25–31
19. Jacobson, M. R., Cash, V. L., Weiss, M. C., Laird, N. F., Newton, W. E., and Dean, D. R. (1989) <i>Mol. Gen. Genet.</i> 219, 49–57
20. Kredich, N. M., Keenan, B. S., and Poette, L. J. (1972) <i>J. Biol. Chem.</i> 247, 7157–7162
21. Laemmli, U. K. (1970) <i>Nature</i> 227, 680–685
22. Vieira, J., and Messing, J. (1987) <i>Methods Enzymol.</i> 153, 3–11
23. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) <i>Proc. Natl. Acad. Sci. U. S. A.</i> 74, 5463–5467
24. Taber, S., and Richardson, C. C. (1985) <i>Proc. Natl. Acad. Sci. U. S. A.</i> 82, 1074–1078
25. Page, W. J., and von Tiggesstrom, M. (1979) <i>J. Bacteriol.</i> 139, 1058–1061
26. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Geoghagen, N. S. M., Collado-Vides, J. P., Benit, P., et al. (1997) <i>Nature</i> 387, 33–46