

Molecular Characterization of the *sor* Gene, Which Encodes the Sulfur Oxygenase/Reductase of the Thermoacidophilic Archaeum *Desulfurolobus ambivalens*

ARNULF KLETZIN

Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, W-8033 Martinsried, Germany

Received 30 March 1992/Accepted 8 July 1992

A 5.8-kbp *Hind*III fragment containing the *sor* gene which encodes the aerobically induced sulfur oxygenase/reductase of the thermoacidophilic, chemolithoautotrophic, and facultatively anaerobic archaeum *Desulfurolobus ambivalens*, was cloned in pUC18 by using an oligonucleotide derived from the N-terminal amino acid sequence for identification (pSOR-1/17). The native enzyme is a 550,000-molecular-weight oligomer composed of single 40,000-molecular-weight subunits; this oligomer is capable of the simultaneous oxidation and reduction of sulfur (A. Kletzin, J. Bacteriol. 171:1638-1643, 1989). From the fragment, 3,025 bp that contained the entire *sor* gene were sequenced. The *sor* gene encoded a protein with 309 amino acid residues (molecular weight, 35,317). The transcript length was determined by Northern RNA hybridization to be 960 to 1,020 nucleotides, and the transcriptional start site was mapped by primer extension analysis. The transcript of the *sor* gene in aerobically grown cells was amplified 38- to 42-fold relative to that in anaerobically grown cells. An initial transcriptional characterization of three neighboring genes of unknown function is also reported.

Depending on the culture conditions, the extremely thermophilic, facultatively anaerobic, chemolithoautotrophic, and acidophilic organisms *Desulfurolobus ambivalens*, *Acidianus infernus*, and *A. brierleyi* can grow either by oxidation of sulfur with oxygen, producing sulfuric acid, or by anaerobic reduction of sulfur with hydrogen, producing H₂S (23, 33). Recently, it was found that they can also grow by oxidation of hydrogen with oxygen (25a). No other organisms known at present show this combination of metabolic pathways (26). The oxidation of sulfur is a characteristic pathway for energy conversion utilized by species of the genus *Sulfolobus* and its relatives (25, 26) among the thermophilic, sulfur-dependent archaeobacteria (*Archaea*). It is also common among the thiobacilli within the bacterial domain (17). Chemolithoautotrophic growth by reduction of sulfur with hydrogen is known only for species of the order *Thermoproteales* and the two above-mentioned genera of the order *Sulfolobales* (5, 26) but has not been found within the domain *Bacteria* (eubacteria; for the definition of the domains *Archaea* and *Bacteria*, see reference 32). The switch between the two pathways of energy conversion in both of the two distantly related *Acidianus* spp. is accompanied by a dramatic change in the composition of the total soluble protein (10). In addition, a major change in the composition of the quinones and lipids in *D. ambivalens*, an organism which is closely related to *A. infernus* (24), has been observed (28).

Little is known about archaeal sulfur metabolism. Two enzymes from the sulfur oxidation pathway of the *Sulfolobales* have been described: a sulfur oxygenase from an *A. brierleyi* isolate (*Sulfolobus brierleyi* [4]) and a sulfur oxygenase/reductase (SOR) from *D. ambivalens* have been purified (9). Both enzymes were similar in size and structure. In contrast to the sulfur oxygenase from *A. brierleyi*, the SOR simultaneously oxidizes and reduces sulfur in the presence of oxygen with sulfite, thiosulfate, and H₂S as products. Thiosulfate was formed by a nonenzymatic reaction of sulfite with sulfur. The inseparable sulfur-oxidizing and -reducing

activities were due to a soluble, cytoplasmic 550-kDa enzyme composed of a single subunit with an apparent molecular mass of 40 kDa. The SOR was aerobically induced. It was not detected in anaerobically growing, sulfur-reducing cells (9).

The cloning and sequencing of a gene encoding the subunit of the SOR as a tool for the investigation of its oxygen-dependent expression is reported here. It is also the first sequence of an archaeal sulfur-metabolizing enzyme. Indeed, nothing on enzymes involved in dissimilatory sulfur oxidation in any of the primary domains has been published. The only sequences available for sulfur-metabolizing enzymes are from the polysulfide reductase of *Wolinella succinogenes* (11), bovine liver rhodanese (21), and the phosphoadenylsulfate reductase pathways of assimilatory sulfate reduction of *Escherichia coli* (7, 12), *Salmonella typhimurium* (14, 15), and *Desulfovibrio vulgaris* (27).

MATERIALS AND METHODS

Materials, cells, and DNA and RNA purification procedures. All materials were of the highest quality commercially available. Molecular biology enzymes were obtained from Boehringer GmbH, Mannheim, Germany. *D. ambivalens* DSM 3772 (German Collection of Microorganisms, Braunschweig, Germany) was grown aerobically and anaerobically by published procedures (34). DNA was prepared from anaerobically grown cells by standard procedures (22). RNA was prepared from aerobically and anaerobically grown log-phase cells by the acidic guanidinium thiocyanate method (1).

Enzyme preparation and amino acid sequencing. The SOR was purified by published procedures (9). The N-terminal amino acid sequence was determined by F. Lottspeich (Martinsried) with a 477A gas-phase sequencer (Applied Biosystems) as recommended by the manufacturer.

Oligodeoxynucleotide probes. The oligonucleotide probes Sox1 3 (AAAACNTTTTGAATGTTTGC, derived from the

N-terminal amino acid sequence, bp 904 through 923), SorT (CTACTGAGGCAAACATT, reverse complement of bp 915 through 931), Sor2T (AGCTGTTTAGCAAATGG, reverse complement of bp 2030 through 2046), Sor3T (CATT TCAATTAGTTTTT, bp 551 through 567), Sor4T (ATAG GCTGGGCGCACAT, reverse complement of bp 2902 through 2918), Seq48 (AAATAGAATATAAATAT, reverse complement of bp 736 through 752), Seq49 (GTACAAAAG GAGAAAAA, reverse complement of bp 1791 through 1807), Seq50 (GGTATTTTGATAGGCTT, bp 2231 through 2247), T3, and T7 (Stratagene) were synthesized on a DNA synthesizer 381A (Applied Biosystems) as recommended by the manufacturer. Labeling with T4 polynucleotide kinase was done by standard procedures (22), except that the reaction was incubated for 30 min at room temperature.

Cloning, subcloning, and sequencing. A 5.8-kbp *Hind*III fragment was cloned into pUC18, yielding plasmid pSOR-1/17 (host strain, *E. coli* JM83; Amersham), after digestion of genomic DNA, Southern hybridization by standard procedures (22), and identification of a fragment with the ³²P-labeled oligonucleotide Sox1 3. Appropriate restriction fragments were subcloned into the vectors M13mp10, M13tg131, M13mp18, and M13mp19 (host strain, *E. coli* DH5 α ; Amersham) and the four pBluescript II vectors (host strain, *E. coli* XL1-Blue; Stratagene), allowing the sequencing of both strands. All sequencing was carried out by the dideoxy-chain termination method with a Sequenase kit (U.S. Biochemical Corp.) as recommended by the manufacturer. The sequences from clones in M13 vectors were primed with the universal primer (-40; U.S. Biochemical); the sequences in pBluescript vectors were primed with the T3 or T7 oligonucleotide. Sequencing gels were cast by using Sequagel solutions (National Diagnostics) and run in Maxiphor chambers (Pharmacia LKB).

Primer extension reactions, Northern RNA hybridization. The *in vivo* transcription initiation of the *sor* gene was mapped by primer extension analysis by published procedures (8) with the oligonucleotide SorT as a primer. The oligonucleotides Sor2T, Sor3T, and Sor4T were used to investigate the transcription initiation of the open reading frames pSOR-1/17-ORF2, 3, and 4. The separation of total RNA in glyoxal gels followed by Northern hybridization with the labeled fragments *sor*-*Sca*I (*sor* gene, bp 1061 through 1652), *Nsi*I-*Pst*I (pSOR-1/17-ORF2, bp 1975 through 2605), *Eco*RI-*Kpn*I (pSOR-1/17-ORF3, bp 1 through 484), and *Nde*I-*Hind*III (pSOR-1/17-ORF4, bp 2794 through 3025) were carried out by standard procedures (22).

Sequence comparison. The searches in the Mipsx protein data bank (Martinsried, Germany) were carried out with the program PIRFASTA (National Biomedical Research Foundation [16]). Similar sequences were subjected to sequence comparison with the GAP programs (University of Wisconsin Genetics Computer Group [3]) for pairwise alignments and determination of pairwise identities and to RDF2 analysis (National Biomedical Research Foundation [16]) for evaluation of the significance of the results.

Nucleotide sequence accession number. The nucleotide sequence shown in Fig. 2 was submitted to the EMBL data library under accession number X56616.

RESULTS AND DISCUSSION

Cloning and sequencing. The sequence of the first 28 N-terminal amino acid residues of the subunit of the SOR was kindly determined by F. Lottspeich, Martinsried, by analysis in a gas-phase sequencer (13a) (see Fig. 2). The

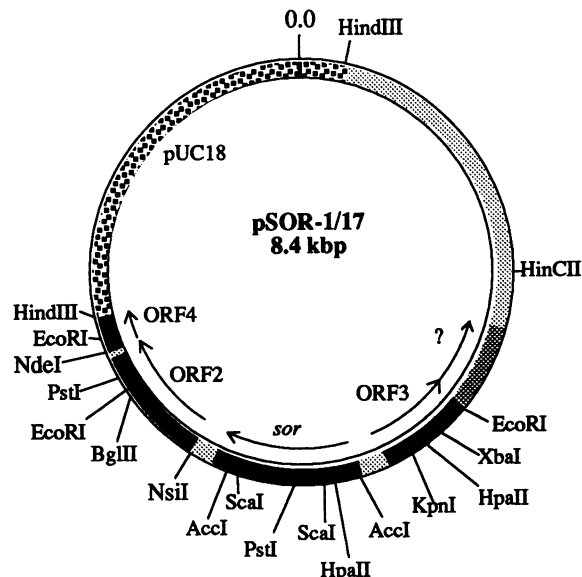


FIG. 1. Plasmid map of pSOR-1/17 with important restriction sites and the orientation of the reading frames. The question mark (?) indicates the unsequenced end of pSOR-1/17-ORF3, and 0.0 indicates the zero position of the pUC18 plasmid.

derived oligonucleotide Sox1 3 was used to identify and clone a 5.8-kbp *Hind*III fragment of genomic DNA of *D. ambivalens* into pUC18, yielding plasmid pSOR-1/17 (Fig. 1). A 3,025-bp region of this fragment, beginning at an *Eco*RI site 852 bp upstream of the *sor* gene (position 1) and ending at the downstream *Hind*III site, was sequenced (Fig. 1).

The fragment contains the entire genetic information for the SOR subunit, another open reading frame (pSOR-1/17-ORF2, Fig. 1), and two other incomplete open reading frames extending beyond the sequenced region (pSOR-1/17-ORF3 and 4; Fig. 1). The nucleotide sequence and the derived amino acid sequences are shown in Fig. 2. The N-terminal amino acids of the *sor* gene exactly match the amino acid sequence obtained on the protein level, showing that the identification of the clone was correct (Fig. 2).

The gene for the subunit of the SOR contained 309 codons, corresponding to a protein with a molecular weight of 35,317. The stoichiometrical ratio calculated from the sequence and the molecular weight of the native enzyme (550,000, determined with gel filtration chromatography [9]) was approximately 16 subunits per molecule of holoenzyme. The calculated molecular weight of 35,317 for the subunit was approximately 4,000 to 5,000 molecular weight units below the apparent molecular weight (40,000 [9]) determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis; this molecular weight is tolerable with respect to the error range of the method. The net charge of the subunit calculated from the amino acid composition was -4, and the isoelectric point was 6.52. Analysis of the sequence with regard to hydropathy (13) and predictions of its secondary structure (2, 6) did not yield features that were remarkable or could be interpreted without further structural information (data not shown). No similar sequences were found in a search of the Mipsx protein data bank (Martinsried, Germany).

Only two features of the amino acid composition were remarkable. The SOR contains 7 tryptophan residues (2.2

GAATTCCTATACTTTCTAGTTCTCTGGATCTTTCTCGTAAATTTTCACGAATTGATTAATAATCTTTATCCTTTA 75
 ?← I G I S E L E R P D K E Y I K V F Q N F D K D K
 CTATTTCTTCGATCTTTTCTCATAAGGTTCAAGTTCTTTAGGTGAGAGAATCCTTTACTTATTAATTCAGCGA 150
 V I E E I K K E Y P E L E K P S L V G K S I L E A
 CCAAGAATAGTAAATTAACCTTATATCGTAAGCTGGAAAATAAAAAGTTCTAATAAATATTTCTATTATTTT 225
 V L F L L N F R I D Y A P F N F L E L L Y K K N N
 CGACGTCATTCTTCAATTTTCTAGGTTAACTATGTCAGAAATTAGGGACCAGCTAGCTATTCTATCTAGTT 300
 E V D N K D I I G L N V I D L I L S W S A I R D L
 CATTACCTTTAAGTAAAGCTATATATTTCTCCTTTCTTTTTCACGCTTTCTATATCTATTACACCGGAATTTA 375
 E N G K L L A I Y K E K E K K V S E I D I V G S N
 CTAGGCTTTTACAGTAGCCAAGCCTCAGAATCGTAAAGGAGAGATAAGAAATCGTATTTTTTACGCTTTTAA 450
 V L D K V L L W A E S D Y L L S L F D Y K K V D K
 CAGTTATGATTCCTTTATTAATCAATTGCGGTACCATTCCAAGCATGTAGCTTTACTGTATCGTCATTGTAAAC 525
 V T I I G K N I L Q P V M R C T W A H L K V T D D N Y
 TGAGTAAATCGAGGAATTTATCCTTCAATTTAGTTTCTTTAGTACAATATTGTTCTGAAATTAGCCAAT 600
 S L L D L F K D K M E I L K E K T V I N N Q I L W
 CTATTTTGAACCAACCATCTAATCTCGTAAGGGTTTGGGAAGATAATAGTGCTAAATAATCGACCATAATCT 675
 D I K F W G D L R D Y P N Q S S L L A L Y D V M
 AATCTACTTTTATTCAGGAATAAATCTATTCTTTTGTAGAGAAAAAATACCAAAAATATTATATTCTAT 750
 ← pSOR-1/17-ORF3
 TTTACATATAGGTTAATGTTAAAGCTAAGTTATACTTTATCTAGTTTTTATGAAAGAAAGATATAAAAAG 825
 TAGACAGAAATGTATATTTGACCAAAAATGCCGAAACCATACGTTGCTATAAACATGGCAGAATTAAAGAATGAA 900
 sor → M P K P Y V A I N M A E L K N E
 amino acid sequencing P K P Y V A I N M A E L K N E
 CCTAAAACCTTTGAAATGTTTGCCTCAGTAGGACCGAAGGCTGCATGGTAACAGCAAGGCATCCGGGCTTTGTT 975
 P K T F E M F A S V G P K V C M V T A R H P G F V
 P K T F E M F A (S) V G P K (V) ? M (V) ?
 GGTTTTCAAACCATATACAAATAGGAATTTTGCCTTCGAAACAGATACGGCGGAGCTAAAATGGACATGACT 1050
 A G F Q N H I Q I G I L P F G N R Y G G A K M D M T
 GAGAAAGTAGTACTGTCAGAGTTTACAGTACACCTTCTGGAAAGATTGGAAGACCATGAAAGAAATGCACAGG 1125
 K E S S T V R V L Q Y T F W K D W K D H E E M H R
 CAAAATCGGAGTTACTTATTCAGGCTATGCTATTCATGCGCTTCAAAATGATATGGGGACCTGGGAGCCAAT 1200
 Q N W S Y L F R L C Y S C A S Q M I W G P W E P I
 TATGAAATAATCTACGCAAACATGCCATAAACACTGAAATGACCGACTTACTGCAGTTGTAGAAAGAAGTTC 1275
 Y E I I Y A N M P I N T E M T D F T A V V G K K F
 GCAGAAGGAAAGCCTTTAGATATTCAGTTATTTTCAACCATATGGAAAGAGAGTTGTTGCCTTTGCAGAGCAC 1350
 A E G K P L D I P V I S Q P Y G K R V V A F A E H
 TCAAGTAAATCCAGCAAGAGAAGCAATTTGAGGACGCAATAGTTAGGACTTTAGAAATGTTAAGAAAGCTCCT 1425
 S V I P G K E K Q F E D A I V R T L E M L K K A P
 GGCTTCTTAGGTGCAATGGTATTAAGGAAATGAGGTTTCCGGAATTGGAAGCATGCAATTCGGTGCCAGGGA 1500
 G F L G A M V L E I G V S C A I G S M Q F G A K G
 TTCCATCAAGTCTTAGAGAACCCTGGATCACTTGAGCCAGATCCAAATAATGTAATGTATTAGTCCAGAGCA 1575
 F H Q V L E N P G S L E P D P N N V M Y S V P E A
 AAGAATACTCCACAACAATACATAGTTTCATGTAGAATGGGCAAATAGTGTCTTTAATGTTGAAATGGGTAGA 1650
 K N T P Q Q Y I V H V E W A N T D A L M F G M G R
 GTACTATTATATCTGAGTAAAGCAAGTACAGCAGCAAGTTTACACATTAGTATACGGACCTTACATTAGA 1725
 V L L Y P E L R Q V H D E V L D T L V Y G P Y I R
 ATATTAATCCAATGATGGAAGGCATTTCTGGAGAGAAATTTAAACGAAACAATAATTTTTTATTTTTTCTCCT 1800
 I L N P M M E G T F W R E Y A L N E Q *
 TTTGTACTTCTTCTCATCTATTATGTAATATTTGGAAGAGTGCCTATTATGAGATAATTATATAATTGAAC 1875
 AGTATCGCTTAGGTTGTGGTAAAGTTATATTATATTGTACCTTATGTGTAATTAATTCATTATTTATACATT 1950
 TGTTAAGGATGGGTTATATTATGCATCTCCCAGTTCTTTTAGGCTTTAAGGTTCTAATAGGTGGAATAATAT 2025
 pSOR-1/17-ORF2 → M F M A L R F L I G G I I L
 ATTACCATTGCTAAACAGCTAACATTGAATAGGATATTTTCTTCTCAATTTTTACTACTTTAAGTACAT 2100
 L P F A L T L T L N R D I F L L S I F T T L S T S
 TTTTTGGGCTTACGGQTTACTTTACGTTGAACCGTCTGAATCAGCAGTATTAAGTTATACAAATGCAATTTGC 2175
 F W A Y G L L Y V E P S E S A V L S Y T M P L I A
 CATTCTTTATCGACALTAATACTTAGAGAAAAGACTAAGACTAAGTATAGGTATTTGATAGGCTTCAG 2250
 I P L S T L I L R E K T T K T E V I G I L I G F S
 CGGTGTTGTAATTTATCTCTAAATCTAGGTTTACTTCTCATTAAATAGGTATAGTATTAACGTAATAAATG 2325
 G V V I Y S L N L G I Y F S L I G I V L T I N A C
 TTTCTTTTGGGCATTATTTACAGTCTATTTTAGAAAACAAAGGGGTTTGTAGTACTTCAAGTAAATGCTGTACA 2400
 F F W A L F T V Y F R K L R G F D A T S V N A V Q
 ATTACTATGGGTTCTTTAATCTTTTTACGCTATCTCTATTCATTTACTTTAATATTCGATAAATTTCTCCT 2475
 L L L G S L I F F T L S P I Q F Y F K Y S I N F L
 AGTAGATCTATTATACGTTTTCAGTACTTGGAGGAGGATTTCTTTTTACTTATGGAATTCATGCTAAAACTGA 2550
 V D L L Y V S V L G G G I S F Y L W N S M L K T E
 GAGAGTACAAAGGTTACTGTCCTTAGTTTTTTCGGTCCCTGCGGTAAGTACTGCAGTTGATGAGCTTAGAGGAGT 2625
 R V S K V T V L S F S V P A V S T A V D E L R G V
 TAATGTAACATTGGTATGATGAAGGGATAGGTTGTTGTTTCTTGAATTTAATATCTAGGTTAGAGAAGAA 2700
 N V N I G M I E G I G V M F L G I L I S R L E K K
 GATTAACAATCTAATAATAAATGGAAGATTTTCATGTATAGCGAATTTTATCTCAACTATTTTTAAATACTAG 2775
 I N K S N I I N G R F H V *
 TGCAAGAAAGAAATAGCATAGGTTATCCTTACGAAATCCATACCTCAAGGAGGTTATCCCACTCAAGGT 2850
 pSOR-1/17-ORF4 → M S Y P Y G N P Y Y P Q G Y P T Q G
 TATCCGCAAGGAAATCCGTTACTATCAACAGAATTTCTAATTTTTCTTTCATTATGTGCGCCAGCCTATAGGCTTA 2925
 Y P Q G N N P Y Y Q N S N F S F I M C A Q P I G L
 GGAGAAAACAGATGATACCAATAAATCCGATGACCTACAAATATGTTGCACAACAAGCAAGTAATGTAC 3000
 G G K Q M M I P I N H P I D L Q Y V A Q Q A V M Y
 TTAATGGGTCAGGATTTCAAGCTT 3025
 L M G Q G F Q A → ?

FIG. 2. Nucleotide and derived amino acid sequences of pSOR-1/17. The N-terminal amino acid sequence was determined by F. Lottspeich (Martinsried) with a gas-phase analyzer. Numbering begins at the left *EcoRI* site (Fig. 1). Transcriptional start sites are indicated in boldface type, and box A regions derived from the start sites are underlined. The pSOR-1/17-ORF3 sequence is in the opposite direction.

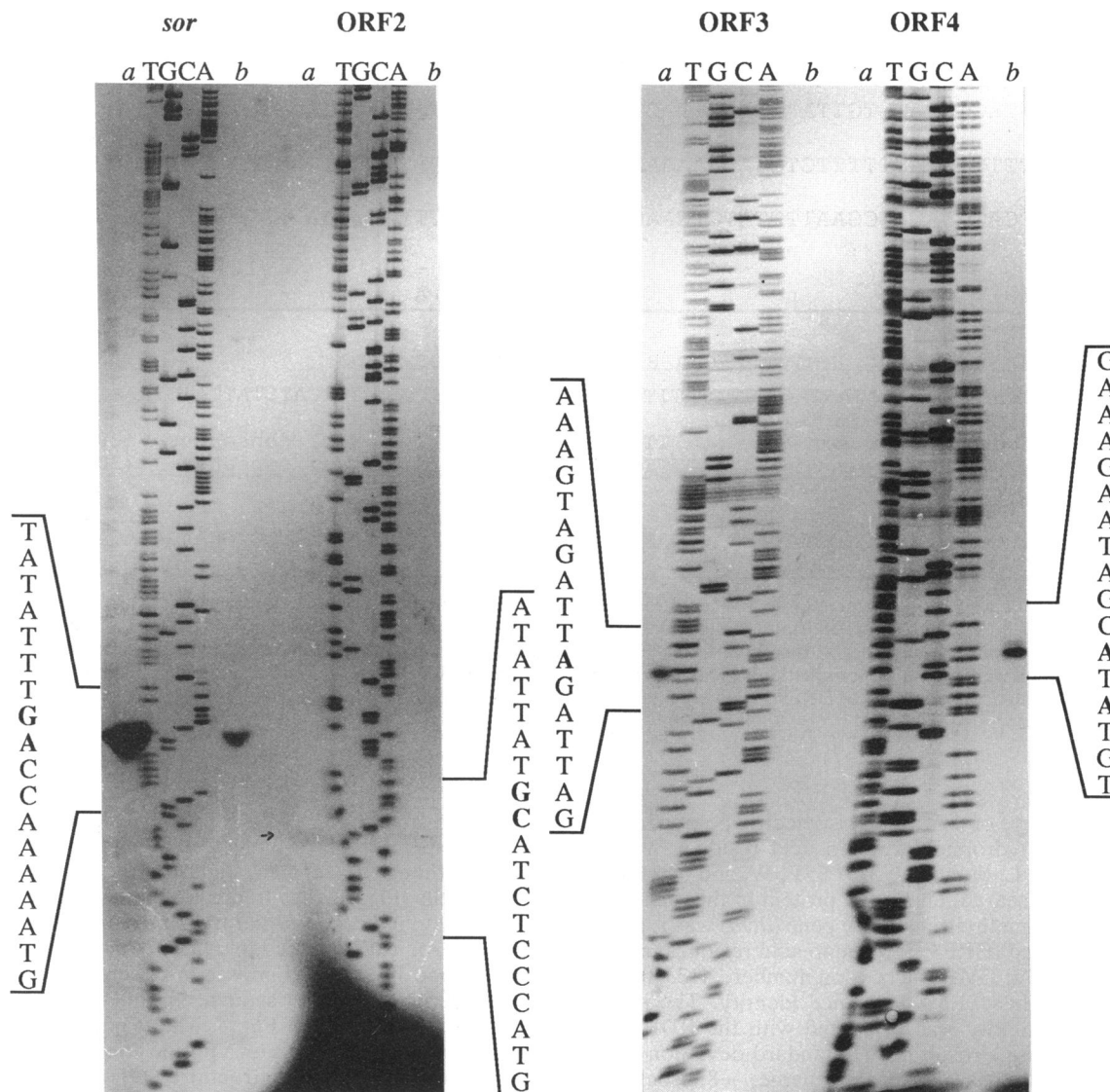


FIG. 3. Primer extension mapping of the pSOR-1/17 transcripts. Samples of 10 μ g of RNA were used in each lane. Lanes: *a*, primer extension signals of RNA of aerobically grown cells; T, G, C, and A, nucleotide sequences obtained with the same oligonucleotide as the primer extension signal (the sequence of the opposite DNA strand is given next to the autoradiogram); *b*, primer extension signals of RNA of anaerobically grown cells. The transcriptional start sites are indicated in boldface type.

mol%) per subunit; the functional importance of the high content of this usually rare amino acid remains unknown. Second, three cysteine residues per subunit were found. These could be of importance for the function of the enzyme. The SOR simultaneously oxidized sulfur to sulfite and reduced sulfur to H_2S in the presence of oxygen. It was inactive under a hydrogen atmosphere (9). The exact reaction mechanism, however, remained obscure. I proposed previously that the oxidation and reduction reactions were due to a sulfur oxygenase reaction, possibly linked to a sulfur disproportionation step with sulfur as an electron donor (9). No cofactor is necessary for the activity; no enzyme-bound cofactor has been found in active preparations. The activity was strongly inhibited by SH-modifying or -binding reagents (9). I presumed that sulfur bound to the enzyme via an SH group was the actual substrate, as previously demonstrated for bovine liver rhodanese (29).

Three cysteine residues per subunit are found in the sequence at two locations within the motifs KVCMT (amino acids 29 through 34) and RLCYSCAS (amino acid 99 through 106). The latter motif is conserved in the large subunit of nickel-containing hydrogenases ($RxC_{xx}C$ [18]), and it occurs frequently in various combinations in chain A of the assimilatory sulfite reductase from *S. typhimurium* (7), e.g., SRCISCGRC (amino acids 229 through 337) or TDCYSAA (amino acids 152 through 158). This protein, however, is a cysteine-rich iron-sulfur protein with an entirely different structure (14, 15). The similarities suggest the participation of an iron-sulfur center in the reaction catalyzed by the SOR. No similar structures could be detected in the amino acid sequences of other proteins that interact with sulfur derivatives.

The protein derived from the reading frame pSOR-1/17-ORF2 has 253 amino acid residues, a molecular weight of

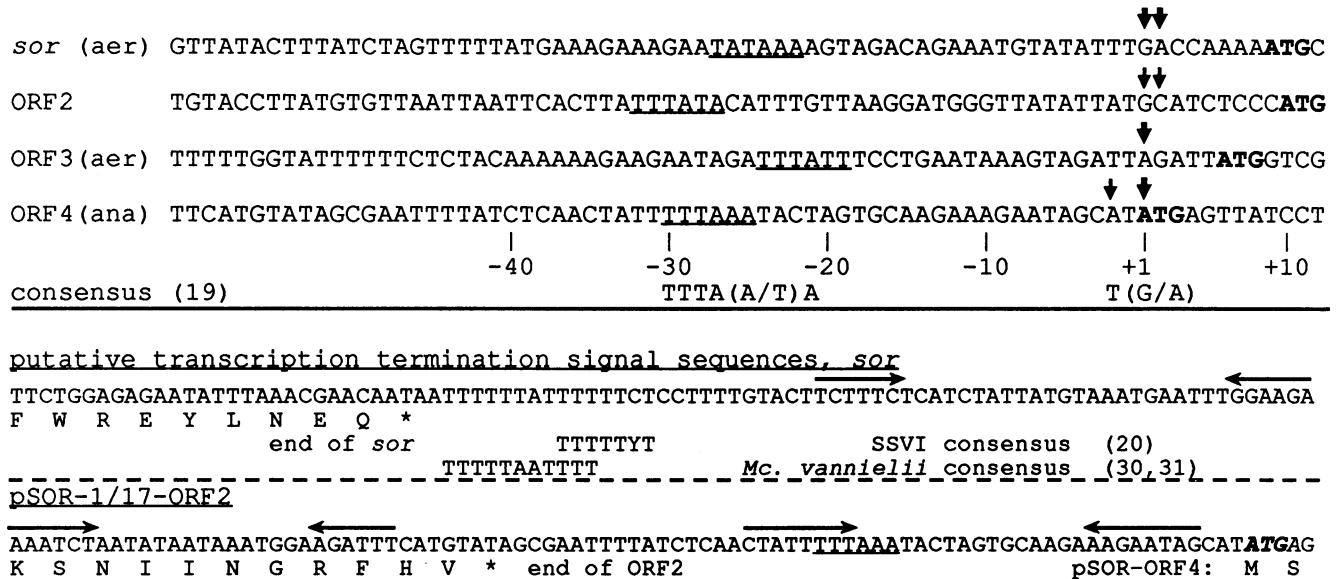


FIG. 4. Promoter and putative terminator sequences. Promoter sequences are derived from the transcriptional start sites (Fig. 2). Box A regions are underlined. Vertical arrows indicate strong (thick arrows) and weak (thin arrows) transcriptional start sites. The start of reading frame is in boldface type. The numbers within parentheses indicate references. Aerobically (aer) and anaerobically (ana) induced transcripts are indicated. Putative transcription termination signals for the *sor* gene and pSOR-1/17-ORF2 (ORF2) and published consensus sequences are indicated. Translational start site of ORF3 is indicated in boldface type, and the box A region is underlined. Horizontal arrows indicate possible stem and loop structures.

28,282, a net charge of +7, and an isoelectric point of 10.49. The content of hydrophobic amino acid residues was remarkably high (I, L, M, and V, 37.9%; F, W, and Y, 14.2%). In the data bank search, an *E. coli* protein in the vicinity of the deoxycytidyl methyltransferase gene (*dcm*) was found to be distantly related (ORF4; 306 amino acid residues; molecular weight, 32,172; EMBL accession number, X13330 [6a]). Both sequences had 21.5% sequence identity. The significance of the similarity was evaluated with the RDF2 program (16), yielding a result of 6.1 standard deviation units. The *E. coli* protein is also basic (charge, +8; isoelectric point, 10.64) and contains highly hydrophobic domains. The structural and similarity data suggest a similar physiological function and a distant phylogenetic relationship. No sequences similar to the reading frames pSOR-1/17-ORF3 and 4 were found. The physiological functions of pSOR-1/17-ORF2, 3, and 4 remain unknown.

Transcription. Northern hybridization of *D. ambivalens* RNA with a ³²P-labeled *Sca*I restriction fragment specific for the *sor* gene (Fig. 1) resulted in the detection of a 960- to 1,020-nucleotide transcript in RNA from aerobically but not of anaerobically grown cells (data not shown). The transcriptional start site was mapped, and the transcription induction was measured by primer extension analysis (Fig. 3). With this method, *sor* gene transcripts were detectable in RNA from aerobically and anaerobically grown cells. The densitometric comparison of signal strength showed that the signal of RNA from aerobically grown cells was 38- to 42-fold stronger than that from anaerobically grown cells (Fig. 3). In spite of the relatively high level of transcription, compared with those of the other inducible transcripts mapped (see below) (Fig. 3), no enzyme activity was found in anaerobically grown cells (9). The reason for this remains unknown.

The *sor* gene and the three other open reading frames were transcribed separately. The primer extension signal of the

transcript of pSOR-1/17-ORF2 was weak. It was found in RNA from cells grown under aerobic and anaerobic conditions, although it seemed slightly stronger in aerobically grown cells (Fig. 3). The transcript of pSOR-1/17-ORF3 was aerobically induced, and the transcript of pSOR-1/17-ORF4 was anaerobically induced. No trace of a signal was obtained in the opposite mode of growth in either case (Fig. 3).

The derived promoter sequences from the transcript mapping are summarized in Fig. 4. The major site of transcription initiation was always on a purine after a pyrimidine. For all promoters, box A-like structures (19) could be found. The only feature shared by the two aerobically induced transcription units, the *sor* gene and pSOR-1/17-ORF3, was a purine-rich stretch immediately preceding box A (Fig. 4).

The transcription termination sites of the two complete open reading frames were not precisely mapped. From the experimentally determined transcript length of 960 to 1,020 nucleotides for the *sor* gene (see above), a termination 30 to 90 nucleotides downstream of the reading frame would be expected. A comparison of the pSOR-1/17 sequences at the end of the *sor* gene with known termination signal sequences showed high similarities with two different, independently determined terminator structures of the virus SSVI (20) and stable RNA genes from *Methanococcus vannielii* (Fig. 4) (30, 31). For these reasons I concluded that there is a strong terminator at this position (Fig. 4). The calculated transcript length would be ≈960 nucleotides, matching the experimentally determined length. The termination sequence of pSOR-1/17-ORF2 could not be aligned as easily to known signal sequences. Two rather weak stem-and-loop structures and a weak similarity to the two above-mentioned consensus sequences were found (Fig. 4).

This first sequence of a sulfur-oxidizing enzyme should contribute to the understanding of sulfur metabolism in the *Sulfolobales*. With respect to the metabolic switch between

aerobic and anaerobic growth in *D. ambivalens*, the basis for the investigation of inducible promoters is laid.

ACKNOWLEDGMENTS

I thank W. Zillig for continuing support during the entire work. Further thanks are due to F. Lottspeich for the determination of the N-terminal amino acid sequence and to Manfred Reitmeier for assistance in the cloning of the gene.

REFERENCES

- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156-159.
- Chou, P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**:45-147.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of programs for sequence analysis. *Nucleic Acids Res.* **12**:387-395.
- Emmel, T., W. Sand, W. A. König, and E. Bock. 1986. Evidence for the existence of a sulfur oxygenase in *Sulfolobus brierleyi*. *J. Gen. Microbiol.* **132**:3415-3420.
- Fischer, F., W. Zillig, K. O. Stetter, and G. Schreiber. 1983. Chemolithoautotrophic metabolism of anaerobic extremely thermophilic archaeobacteria. *Nature (London)* **301**:511-513.
- Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting secondary structure of globular proteins. *J. Mol. Biol.* **120**:97-120.
- Hanck, T. 1988. Unpublished data.
- Huang, C. J., and E. L. Barrett. 1991. Sequence analysis and expression of the *Salmonella typhimurium* *asr* operon encoding production of hydrogen sulfide from sulfite. *J. Bacteriol.* **173**:1544-1553.
- Hüdepohl, U., W.-D. Reiter, and W. Zillig. 1990. *In vitro* transcription of two rRNA genes of the archaeobacterium *Sulfolobus* sp. B12 indicates a factor requirement for specific initiation. *Proc. Natl. Acad. Sci. USA* **87**:5851-5855.
- Kletzin, A. 1989. Coupled enzymatic production of sulfite, thiosulfate, and hydrogen-sulfide from sulfur: purification and properties of a sulfur oxygenase reductase from the facultatively anaerobic archaeobacterium *Desulfurolobus ambivalens*. *J. Bacteriol.* **171**:1638-1643.
- Kletzin, A., A. Segerer, and F. Klink. 1989. Changes in protein composition of facultatively aerobic sulfur-dependent archaeobacteria depending on growth conditions. *Arch. Microbiol.* **151**:282-284.
- Krafft, T., M. Bokranz, O. Klimmek, I. Schröder, F. Fahrenholz, E. Kojro, and A. Krüger. 1992. Cloning and nucleotide-sequence of the *psrA* gene of *Wolinella succinogenes* polysulfide reductase. *Eur. J. Biochem.* **206**:503-510.
- Krone, F. A., G. Westphal, and J. D. Schwenn. 1991. Characterization of the gene *cysH* and of its product phospho-adenylsulfate reductase from *Escherichia coli*. *Mol. Gen. Genet.* **225**:314-319.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105-132.
- Lottspeich, F. Personal communication.
- Ostrowski, J., M. J. Barber, D. C. Rueger, B. E. Miller, L. M. Siegel, and N. M. Kredich. 1989. Characterization of the flavo-protein moieties of NADPH-sulfite reductase from *Salmonella typhimurium* and *Escherichia coli*. Physicochemical and catalytic properties, amino acid sequence deduced from DNA sequence of *cysJ* and comparison with NADPH-cytochrome P-450 reductase. *J. Biol. Chem.* **264**:15726-15737.
- Ostrowski, J., J.-Y. Wu, D. C. Rueger, B. E. Miller, L. M. Siegel, and N. M. Kredich. 1989. Characterization of the *cysJIIH* regions of *Salmonella typhimurium* and *Escherichia coli* B: DNA sequences of *cysI* and *cysH* and a model for the siroheme-Fe4S4 active center of sulfite reductase hemoprotein based on amino acid homology with spinach nitrite. *J. Biol. Chem.* **264**:15726-15737.
- Pearson, W. R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* **183**:63-99.
- Pronk, J. T., R. Meulenberg, W. Hazeu, P. Bos, and J. G. Kuenen. 1990. Oxidation of reduced inorganic sulfur-compounds by acidophilic thiobacilli. *FEMS Microbiol. Rev.* **75**:293-306.
- Przybyla, A. E., J. Robbins, N. Menon, and H. D. Peck. 1992. Structure-function relationships among the nickel-containing hydrogenases. *FEMS Microbiol. Rev.* **88**:109-136.
- Reiter, W.-D., P. Palm, and W. Zillig. 1988. Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic *pol II* promoters. *Nucleic Acids Res.* **16**:1-19.
- Reiter, W.-D., P. Palm, and W. Zillig. 1988. Transcription termination in the archaeobacterium *Sulfolobus*: signal structures and linkage to transcription initiation. *Nucleic Acids Res.* **16**:2445-2459.
- Russell, J., L. Weng, P. S. Keim, and R. L. Henrikson. 1978. The covalent structure of bovine liver rhodanese. *J. Biol. Chem.* **253**:8102-8108.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Segerer, A., K. O. Stetter, and F. Klink. 1985. Two contrary modes of chemolithotrophy in the same archaeobacterium. *Nature (London)* **313**:787-789.
- Segerer, A. H., A. Trincone, M. Gahrtz, and K. O. Stetter. 1991. *Stygiolobus azoricus* gen. nov., sp. nov. represents a novel genus of anaerobic, extremely thermoacidophilic archaeobacteria of the order *Sulfolobales*. *Int. J. System. Bacteriol.* **41**:495-501.
- Shivvers, D. G., and T. D. Brock. 1973. Oxidation of elemental sulfur by *Sulfolobus acidocaldarius*. *J. Bacteriol.* **114**:706-710.
- Stetter, K. O. In T. Thanh Van (ed.), *Frontiers of life. IIIrd Rencontres de Blois*, in press. Editions Frontieres, Gif-sur-Yvette Cedex, France.
- Stetter, K. O., G. Fiala, G. Huber, R. Huber, and A. Segerer. 1990. Hyperthermophilic microorganisms. *FEMS Microbiol. Rev.* **75**:117-124.
- Tan, J., L. R. Helms, R. P. Swenson, and J. A. Cowan. 1991. Primary structure of the assimilatory-type sulfite reductase from *Desulfovibrio vulgaris* (Hildenborough)—cloning and nucleotide sequence of the reductase gene. *Biochemistry* **30**:9900-9907.
- Trincone, A., V. Lanzotti, B. Nicolaus, W. Zillig, M. DeRosa, and A. Gambacorta. 1989. Comparative lipid-composition of aerobically and anaerobically grown *Desulfurolobus ambivalens*, an autotrophic thermophilic archaeobacterium. *J. Gen. Microbiol.* **135**:2751-2757.
- Weng, L., R. L. Henrikson, and J. Westley. 1978. Active site cysteinyl and arginyl residues of rhodanese. *J. Biol. Chem.* **253**:8109-8119.
- Wich, G., H. Hummel, M. Jarsch, U. Bar, and A. Böck. 1986. Transcription signals for stable RNA genes in *Methanococcus*. *Nucleic Acids Res.* **14**:2459-2479.
- Wich, G., L. Siebold, and A. Böck. 1986. Genes for tRNA and their putative expression signals in methanogens. *Syst. Appl. Microbiol.* **7**:18-25.
- Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Toward a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eukarya*. *Proc. Acad. Natl. Sci. USA* **87**:4576-4579.
- Zillig, W., S. Yeats, I. Holz, A. Böck, F. Gropp, M. Rettenberger, and S. Lutz. 1985. Plasmid-related anaerobic autotrophy of the novel archaeobacterium *Sulfolobus ambivalens*. *Nature (London)* **313**:789-791.
- Zillig, W., S. Yeats, I. Holz, A. Böck, M. Rettenberger, F. Gropp, and G. Simon. 1986. *Desulfurolobus ambivalens* gen. nov., sp. nov., an autotrophic archaeobacterium facultatively oxidizing and reducing sulfur. *Syst. Appl. Microbiol.* **8**:197-203.