

# DNA Microarray Analysis of the Hyperthermophilic Archaeon *Pyrococcus furiosus*: Evidence for a New Type of Sulfur-Reducing Enzyme Complex

GERRIT J. SCHUT,<sup>1</sup> JIZHONG ZHOU,<sup>2</sup> AND MICHAEL W. W. ADAMS<sup>1\*</sup>

*Department of Biochemistry and Molecular Biology and Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602,<sup>1</sup> and Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831<sup>2</sup>*

Received 4 May 2001/Accepted 21 September 2001

DNA microarrays were constructed by using 271 open reading frame (ORFs) from the genome of the archaeon *Pyrococcus furiosus*. They were used to investigate the effects of elemental sulfur (S<sup>0</sup>) on the levels of gene expression in cells grown at 95°C with maltose as the carbon source. The ORFs included those that are proposed to encode proteins mainly involved in the pathways of sugar and peptide catabolism, in the metabolism of metals, and in the biosynthesis of various cofactors, amino acids, and nucleotides. The expression of 21 ORFs decreased by more than fivefold when cells were grown with S<sup>0</sup> and, of these, 18 encode subunits associated with three different hydrogenase systems. The remaining three ORFs encode homologs of ornithine carbamoyltransferase and HypF, both of which appear to be involved in hydrogenase biosynthesis, as well as a conserved hypothetical protein. The expression of two previously uncharacterized ORFs increased by more than 25-fold when cells were grown with S<sup>0</sup>. Their products, termed SipA and SipB (for sulfur-induced proteins), are proposed to be part of a novel S<sup>0</sup>-reducing, membrane-associated, iron-sulfur cluster-containing complex. Two other previously uncharacterized ORFs encoding a putative flavoprotein and a second FeS protein were upregulated more than sixfold in S<sup>0</sup>-grown cells, and these are also thought to be involved in S<sup>0</sup> reduction. Four ORFs that encode homologs of proteins involved in amino acid metabolism were similarly upregulated in S<sup>0</sup>-grown cells, a finding consistent with the fact that growth on peptides is a S<sup>0</sup>-dependent process. An ORF encoding a homolog of the eukaryotic rRNA processing protein, fibrillarin, was also upregulated sixfold in the presence of S<sup>0</sup>, although the reason for this is as yet unknown. Of the 20 S<sup>0</sup>-independent ORFs that are the most highly expressed (at more than 20 times the detection limit), 12 of them represent enzymes purified from *P. furiosus*, but none of the products of the 34 S<sup>0</sup>-independent ORFs that are not expressed above the detection limit have been characterized. These results represent the first derived from the application of DNA microarrays to either an archaeon or a hyperthermophile.

Hyperthermophiles are microorganisms that grow optimally at temperatures of 80°C and higher, and most are classified as archaea (65). They are a rather diverse group with respect to their metabolic capabilities, but many of them utilize peptides as a carbon source and reduce elemental sulfur (S<sup>0</sup>) to H<sub>2</sub>S (66). Of these, the majority are obligately proteolytic and show little if any growth unless S<sup>0</sup> is added to the growth medium. The exceptions include some species of *Pyrococcus* that metabolize poly- and oligosaccharides, as well as peptides (3, 4, 15). For example, *Pyrococcus furiosus* grows on a disaccharide (maltose) to high cell densities in the absence of S<sup>0</sup> and produces H<sub>2</sub> as an end product rather than H<sub>2</sub>S. On the other hand, *P. furiosus* requires S<sup>0</sup> in the growth medium when it utilizes peptides as a carbon source (1). A comparison of the activities of some key metabolic enzymes in cells grown either on peptides (with S<sup>0</sup>) or on maltose (with or without S<sup>0</sup>) provided the first evidence for a highly regulated fermentative-based metabolism in *P. furiosus*, with S<sup>0</sup> or its metabolites playing a key regulatory role (1). The molecular mechanism by which S<sup>0</sup> achieves this control is not known.

The pathways by which *P. furiosus* metabolizes peptides and carbohydrates are reasonably well established. The organism contains a modified Embden-Meyerhof pathway in which hexokinase and phosphofructokinase are ADP rather than ATP dependent (29, 68). In addition, the expected glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase are replaced by a single phosphate-independent enzyme, glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR [42]), which is regulated at the transcriptional level (69). The pyruvate produced by this pathway is converted to the end product acetate in two steps involving pyruvate ferredoxin oxidoreductase (POR [5]) and acetyl-coenzyme A (CoA) synthetases I and II (40, 61), with the concomitant production of ATP from ADP and phosphate. The oxidative steps in the pathway of acetate production from glucose therefore involve two ferredoxin-dependent enzymes (GAPOR and POR), and a similar mechanism is present during peptide catabolism by *P. furiosus*. The organism contains, in addition to POR, three other, ferredoxin-dependent, 2-keto acid oxidoreductases, and these convert transaminated amino acids into their corresponding CoA derivatives (5, 19, 39), which are then transferred to their corresponding organic acids by the two acetyl-CoA synthetases with concomitant ATP synthesis (40).

How *P. furiosus* and related species couple the oxidation of the reduced ferredoxin generated by the catabolic pathways to

\* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Life Sciences Bldg., University of Georgia, Athens, GA 30602-7229. Phone: (706) 546-2060. Fax: (706) 542-0229. E-mail: adams@bmb.uga.edu.

the reduction of  $S^{\circ}$  is not known. The organism contains three different hydrogenases: two cytoplasmic, NAD(P)H-dependent enzymes, and one as part of a membrane-bound complex (9, 38, 59). The cytoplasmic hydrogenases were thought to play a role in  $S^{\circ}$  reduction since they can reduce  $S^{\circ}$  in vitro (37), but this appears not to be the case as the total hydrogenase activity of  $S^{\circ}$ -grown cells is dramatically lower than that of cells grown without  $S^{\circ}$  (1). The mechanism by which  $S^{\circ}$  affects hydrogenase activity is not understood. Moreover, the nature of the enzyme system that reduces  $S^{\circ}$  remains unknown. In mesophilic organisms, the only  $S^{\circ}$ -reducing enzyme that has been well-characterized is the trimeric, membrane-bound, molybdenum (Mo)-containing, polysulfide reductase of the mesophilic bacterium, *Wolinella succinogenes* (32), but the genome of *P. furiosus* (54) does not contain ORFs that would encode a homolog of the mesophilic enzyme. Moreover, *P. furiosus* is not known to utilize Mo but contains at least three enzymes that utilize the analogous element tungsten (W [27]), perhaps suggesting a role for W in  $S^{\circ}$  reduction.

The ability of *P. furiosus* to grow well both with and without  $S^{\circ}$ , together with the availability of its complete genome sequence, provide an opportunity to investigate  $S^{\circ}$  metabolism by expression analysis by using DNA microarrays. In fact, the genus *Pyrococcus* is unique in that complete genome sequences are available from three species; *P. horikoshii*, *P. abyssi*, and *P. furiosus* (28, 54; D. Prieur, P. Forterre, J.-C. Thierry, and J. Dietrich [http://www.genoscope.cns.fr/Pab/]). While DNA microarrays (60) have revolutionized functional genomics in eukaryotic systems (e.g., references 14, 56, and 63), there have been far fewer applications with prokaryotes. Studies initially focused on pathogens (11, 17, 58, 67, 74) and more recently have included *Escherichia coli* (2, 30, 53, 73, 79), *Bacillus subtilis* (76, 78) and a cyanobacterium (21). As yet there have been no reports of using this technique with either a hyperthermophile or an archaeon. The genome of *P. furiosus* contains ca. 2,200 open reading frames (ORFs) with >50% of unknown function. As a prelude to a complete genomic analysis, we focus here on 271 ORFs that encode proteins that are proposed to be involved in the primary metabolic pathways, energy conservation, and metal metabolism. The results indicate that  $S^{\circ}$  or its metabolites play a major regulatory role at the transcriptional level and that  $S^{\circ}$  reduction appears to be accomplished by a new type of enzyme system involving the products of previously uncharacterized ORFs.

#### MATERIALS AND METHODS

**Primer design and PCR.** Gene sequences were obtained from the *P. furiosus* genome site (<http://comb5-156.umbi.umd.edu/>) (54). Of the 275 ORFs that were examined in this study, 87 of them are listed in Tables 1 to 4. A complete list of the 275 ORFs is available at <http://adams.bmb.uga.edu/pubs/sup228.pdf>. Primers were designed for the ORFs by using the Primer 3 program (Whitehead Institute, MIT) and were purchased from Stanford University and from MWG Biotech (High Point, N.C.). For the initial round of PCR the primer pairs were designed to give products corresponding to the complete ORFs, and this yielded products for 249 of them as verified by gel electrophoresis. The remaining 42 primer pairs were designed to yield a product of ca. 1 kb (unless the target ORF was smaller). Of these only four ORFs did not yield suitable PCR products, and these were not pursued further. PCR products were purified with a Strataprep 96 PCR purification kit (Stratagene, La Jolla, Calif.) and eluted in 33% (vol/vol) dimethyl sulfoxide. They were spotted in duplicate onto aminosilane coated slides (Sigma, St. Louis, Mo.) in four subarrays by using a robotic slide printer (Omnigridd; Genemachines, San Carlos, Calif.). The slides were processed as previously described (8).

**Growth of *P. furiosus* and RNA preparation.** *P. furiosus* was grown in batch mode in a 20-liter custom fermentor at 95°C with maltose as the primary carbon source and was harvested in mid-log phase (1). The only variable was the presence or absence of  $S^{\circ}$ . The cells that were used to prepare RNA for the microarray analyses were from experiments that have been described previously (1). In that case, activity assays were determined for more than 20 enzymes involved in the primary metabolic pathways by using cytoplasmic and membrane fractions, and these results are referred to below. Samples (2,000 ml) of the same cultures were cooled on ice, and total RNA was extracted by using acid-phenol extraction (71). No significant contamination of the RNA with genomic DNA could be detected by Northern blot hybridization or preliminary microarray experiments (data not shown), and so a DNase treatment was not used.

**Preparation of cDNA and hybridization conditions.** Fluorescently labeled cDNA was prepared with the ARES DNA Labeling Kit (Molecular Probes, Eugene, Oreg.). In brief, 15  $\mu$ g of total RNA was reversed transcribed in a total volume of 20  $\mu$ l by using Stratascript RT (Stratagene, La Jolla, Calif.) in the presence of 1 mM dATP, 1 mM CTP, 1 mM GTP, 0.3 mM dTTP, 0.5 mM aminoallyl dUTP, and 1  $\mu$ g of random 9-mers (Stratagene) according to the manufacturer's instructions. After a 1.5 h of incubation at 42°C, RNA was destroyed by the addition of 0.1 N NaOH, followed by a 10-min incubation at 70°C. After neutralization with 0.1 N HCl, amine-modified cDNA was purified by using a QIAquick PCR Purification Kit (Qiagen, Valencia, Calif.), except that the wash buffer was replaced with 75% (vol/vol) ethanol, and the cDNA was eluted with 45  $\mu$ l of distilled water and dried under vacuum. The amine modified cDNA was labeled with Alexa 488 or Alexa 594 dyes (Molecular Probes) according to the manufacturer's instructions. Alexa 546 was used as third dye for the triple-labeling experiments. The labeled cDNA was purified with the Qiagen kit as described above and was dried under vacuum. Differentially labeled (Alexa 488 and Alexa 594) cDNA derived from *P. furiosus* cells grown in the presence or absence of  $S^{\circ}$  was pooled and hybridized to the microarrays. The labeled cDNA was dissolved in 20  $\mu$ l of hybridization buffer (Sigma), and hybridization was performed under a coverslip at 65°C in a humidity chamber (Arrayit, Sunnyvale, Calif.) for between 10 and 15 h. The slides were then washed twice for 5 min in each of 2 $\times$  SSC (1 $\times$  SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.1% sodium dodecyl sulfate (SDS), and 0.2 $\times$  SSC-0.1% SDS and then rinsed in distilled water and blown dry with compressed air. The intensities of the Alexa 488 and Alexa 594 dyes were measured by using a Scan Array 5000 spectrometer (Packard, Meriden, Conn.) with the appropriate laser and filter settings and analyzed by using Quantarray (Packard).

**Data analysis.** To compute signal intensities for each ORF, the areas surrounding each spot were taken as the background, and 17 control spots, where only spotting buffer was printed, were included on the 384-spot array. The control spots were corrected for local background, and the average was used to correct the signals from each ORF after their background correction. The relative amounts of the transcripts (with or without  $S^{\circ}$ ) are presented in a linear fashion by converting all ratios to a  $\log_2$  function. For the two different growth conditions tested, a negative value of less than  $-2.3$  represents a >5-fold down-regulation of expression from a given ORF by  $S^{\circ}$ , and this corresponds to a green color in a false color overlay. Conversely, a positive value greater than 2.3 represents a >5-fold upregulation and corresponds to a red color in a false color overlay. ORFs that display  $\log_2$  values between approximately  $-1$  and 1 are minimally affected by  $S^{\circ}$  and show a yellow color on the overlay. The detection limit of fluorescent signals was set arbitrary to 1,000 intensity units (see Fig. 1) and such spots are not visible on the false overlay. Only ORFs that display intensities more than twice the detection limit are considered valid. Conversion of intensity ratios enables the data to be presented as a  $\log_2$  value with a standard deviation (SD). This represents an average of six hybridization experiments with cDNA derived from five different cultures of *P. furiosus*: two grown with  $S^{\circ}$  and three grown without it. In preliminary experiments we did not observe any significance difference in the quality of the data between different Alexa dyes (Molecular Probes) nor when the Alexa dyes were replaced with CY3 and CY5 (Amersham, Piscataway, N.J.), dyes that are more usually used as fluorescent labels (12). The advantage of the Alexa dyes is that labeling with three dyes is possible with the lasers and filters available in the Scanarray 5000 (Packard). Statistical significance of the observed fluorescence signal ratios was obtained by a paired *t* test analysis by using the statistical software JMP4 (SAS Institute, Cary, N.C.).

#### RESULTS AND DISCUSSION

**Selection of ORFs.** Of the ca. 2,200 ORFs in the *P. furiosus* genome (54), 271 were selected for microarray analysis. The

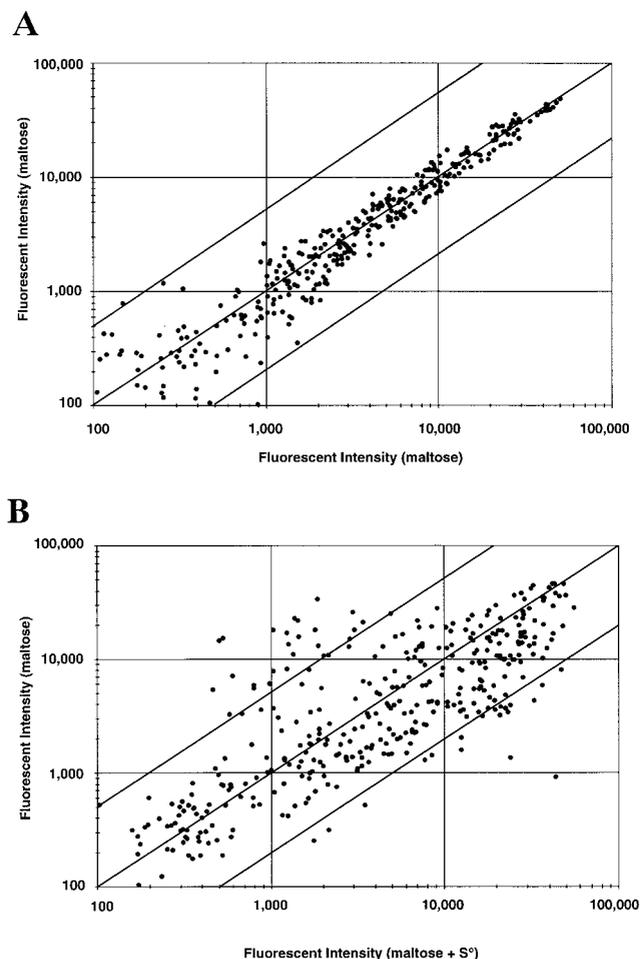


FIG. 1. Fluorescence intensities of DNA microarrays. (A) cDNA versus cDNA derived from the same cultures of cells grown on maltose (no  $S^{\circ}$ ). (B) cDNA versus cDNA derived from cells grown on maltose with or without  $S^{\circ}$ . The upper and lower diagonal lines indicate fivefold changes in the signal intensities. See the text for details.

targets were putative genes involved in the primary metabolic pathways of carbon, nitrogen, and energy metabolism, and those that are, or might be, related to the metabolism of  $S^{\circ}$  and  $H_2$ . They included ORFs that are proposed to encode various enzymes and proteins involved in the pathways of sugar and peptide catabolism; the utilization of metals (such as Fe, Ni, W, and Mo); and the biosynthesis of cofactors, amino acids, lipids, polyamines, ribosomes, and nucleotides, together with putative chaperonins, ATPases, and transcriptional regulators. Unless otherwise indicated, each ORF is referred to by its end nucleotide number and by its current annotation, which is based on homology searches (depicted in brackets [see Table 1 and reference 54]). It is specifically noted in Tables 1 to 4 where there are experimental data derived by using *P. furiosus* to support the ORF assignment (given without brackets and with a reference, see Table 1).

**Experimental protocols and data analysis.** The efficacy of the microarray experiment with mRNA derived from *P. furiosus* cells is shown in Fig. 1. Figure 1A shows data with differentially labeled cDNA samples of RNA prepared from the

same batch of cells, grown with maltose in the absence of  $S^{\circ}$ , and hybridized to the same microarray. The intensity of the fluorescence signals varied over a range of more than  $10^3$ , a finding consistent with what has been reported with other organisms (26, 34, 46, 78). Those ORFs with intensities of below ca. 2,000 (Fig. 1A) are considered below the detection limit, and it cannot be concluded that these ORFs are expressed to a significant degree. From Fig. 1A it is clear that ORFs that give rise to low-intensity signals show a high deviation because of background fluorescence, whereas the data derived from ORFs that give high signal intensities lie near or on the diagonal.

As shown in Fig. 1B,  $S^{\circ}$  and/or its metabolites clearly have a dramatic effect on the expression of many of the 271 ORFs that were examined. It was previously shown that the presence of  $S^{\circ}$  in the growth medium affects the activities of several metabolic enzymes (1). The microarray data indicate that this regulation is primarily at the transcriptional level. We focus below on ORFs whose expression appears to be strongly regulated by  $S^{\circ}$ , where the signal intensity changes by at least fivefold (shown by the upper and lower diagonal lines in Fig. 1B). The standard deviations ( $\log_2$  intensity ratios) for almost all of these ORFs are  $<2$ , and they are differentially regulated with at least a 95% confidence level as determined by using a paired *t* test analysis ( $P > 0.95$ ). However, there are three exceptions. These are ORFs 1419794 (formate dehydrogenase  $\alpha$ -chain), 156299 (3-isopropyl malate dehydrogenase), and 438413 (phosphoribosylglycinamide formyltransferase) whose  $\log_2$  intensities showed deviations of  $>2$  ( $P < 0.90$ ). For these ORFs a large variation in the ratios was obtained by using the same growth condition (maltose without  $S^{\circ}$ ). This was verified when cDNA samples from three independently grown cultures were separately labeled with the Alexa 488, Alexa 546, or Alexa 594 dyes, and all three cDNAs were hybridized to the same slide. Only these three (of 271) ORFs showed distinct false colors, indicating differential expression in one of the three samples. The reasons for this are not clear, but these ORFs appear not to be regulated by  $S^{\circ}$  and are not considered further.

**ORFs strongly downregulated by  $S^{\circ}$ .** The expression of 21 of the 271 ORFs examined decreased by more than a factor of 5 when RNA was derived from cells grown with  $S^{\circ}$  (Fig. 1B), and these are listed in Table 1. Of these, 18 ORFs encode subunits associated with the three different hydrogenase systems that have been characterized from *P. furiosus*. These include those encoding six of the eight subunits of the two cytoplasmic hydrogenases (I and II [38, 50]), whose expression decreases between 6- and 14-fold in the presence of  $S^{\circ}$  (Table 1). Expression of the other two ORFs (encoding the  $\alpha$ - and  $\gamma$ -subunits of hydrogenase I) decrease by 2.2- to 4.8-fold. These data are in accord with the report that the presence of  $S^{\circ}$  in the medium decreased the total hydrogenase activity of the cytoplasmic fraction by ca. 16-fold (1). Since the specific activity of hydrogenase I is ca. 10-fold higher than that of hydrogenase II in *in vitro* assays, it was not possible to determine from the activity analyses if hydrogenase II activity was regulated by  $S^{\circ}$ . However, the microarray data clearly show that both hydrogenases are strongly regulated. Moreover,  $S^{\circ}$  regulates the expression of the genes encoding the cytoplasmic hydrogenases in a negative fashion, rather than  $S^{\circ}$  (or its metabolites) having some inhibitory effect on activity. Obviously, these hydrogenases are

TABLE 1. ORFs whose expression is dramatically (&gt;5-fold) downregulated by S°

ORF <sup>a</sup>	Description <sup>b</sup>	Mean intensity ratio (log <sub>2</sub> ± SD) <sup>c</sup>	Change in expression (fold) <sup>c</sup>
577932	[Hydrogenase expression and/or formation regulatory protein, <i>hypF</i> ]	5.28 ± 0.89	39.0
1337916	Membrane-bound hydrogenase ORF1, <i>mbh1</i> <sup>d</sup>	4.14 ± 0.80	17.6
1251888	Hydrogenase II gamma, <i>hyhG2</i> (38)	3.78 ± 1.47	13.7
1338167	Membrane-bound hydrogenase ORF2, <i>mbh2</i> <sup>d</sup>	3.77 ± 1.29	13.6
1339081	Membrane-bound hydrogenase ORF5, <i>mbh5</i> <sup>d</sup>	3.68 ± 1.30	12.8
1339520	Membrane-bound hydrogenase ORF6, <i>mbh6</i> <sup>d</sup>	3.65 ± 0.77	12.6
1253842	Hydrogenase II alpha, <i>hyhL2</i> (38)	3.57 ± 1.45	11.8
1341399	Membrane-bound hydrogenase ORF8 (like <i>cooM</i> and <i>mbh8</i> ) <sup>d</sup>	3.49 ± 0.79	11.2
1338538	Membrane-bound hydrogenase ORF3, <i>mbh3</i> <sup>d</sup>	3.31 ± 1.66	9.9
1252601	Hydrogenase II delta, <i>hyhS2</i> (38)	3.25 ± 1.43	9.5
1342770	Membrane-bound hydrogenase ORF11, <i>mbh11</i> <sup>d</sup>	3.06 ± 1.22	8.3
866528	Hydrogenase I delta, <i>hyhS1</i> (37)	2.95 ± 0.66	7.7
1345018	Membrane-bound hydrogenase ORF13 (like <i>hydC</i> , <i>cooK</i> , <i>echB</i> , and <i>mbh13</i> ) <sup>d</sup>	2.77 ± 0.85	6.8
1345434	Membrane-bound hydrogenase ORF14 (like <i>hycF</i> , <i>echF</i> , <i>cooX</i> , and <i>mbh14</i> ) <sup>d</sup>	2.73 ± 1.40	6.6
1344050	Membrane-bound hydrogenase ORF12, catalytic NiFe subunit, <i>mbh12</i> (59)	2.68 ± 0.59	6.4
864857	Hydrogenase I beta, <i>hyhB1</i> (37)	2.66 ± 1.08	6.3
1251025	Hydrogenase II beta, <i>hyhB2</i> (38)	2.65 ± 1.04	6.3
1342256	Membrane-bound hydrogenase ORF10, small subunit homolog, <i>mbh10</i> <sup>d</sup>	2.59 ± 0.87	6.0
51760	[Conserved hypothetical protein]	2.59 ± 1.14	6.0
1338785	Membrane-bound hydrogenase ORF 4, <i>mbh4</i> <sup>d</sup>	2.54 ± 2.16	5.8
615154	OTCase, <i>argF</i> (36)	2.43 ± 0.49	5.4

<sup>a</sup> ORF designation is the end nucleotide number (54).

<sup>b</sup> The ORF description is derived either from annotation by homology (given within brackets) or from the indicated reference where there is experimental data to support the ORF assignment specifically in *P. furiosus* (given without brackets).

<sup>c</sup> The intensity ratio is expressed as a log<sub>2</sub> value so that the standard deviation can be given. For case of comparison between ORFs, the apparent change in the expression level of a given ORF is also indicated.

<sup>d</sup> This work (see text for details).

unlikely to play a role in S° reduction, as originally suggested (37).

It was previously shown that the activity of the membrane-bound hydrogenase of *P. furiosus* decreases by almost 30-fold when cells are grown in the presence of S° (1). After solubilization and purification from cells grown in the absence of S°, the enzyme contained two subunits (Mbh11 and Mbh12) and, based on their N-terminal sequences, the genes encoding them were proposed to be part of a 14-ORF operon (Mbh1 to -14 [59]). All but one (Mbh7) of the 14 ORFs were included on the DNA microarray. As shown in Table 1, the expression of all but one of them decreased by more than 5-fold in the presence of S° (the remainder, Mbh9, decreased by 4.2-fold). These data confirm that the 14 ORFs do indeed constitute an operon and that its products are involved in the metabolism of H<sub>2</sub>. The average fluorescence intensities associated with the 13 ORFs of the membrane-bound hydrogenase operon are shown in Fig. 2, along with the values for the eight subunits that encode the two cytoplasmic hydrogenases. The ORFs are depicted according to their order in their respective operons. All show a similar pattern in that they are downregulated to close to the detection limit (see Fig. 1B) in the presence of S°, a finding consistent with barely detectable hydrogenase activities in cell extracts (1). The apparent exception is the ORF encoding the α-subunit of hydrogenase I. Although this ORF is expressed at about five times the detection limit, the results show that it is regulated by S° as determined by the paired *t* test analysis (*P* > 0.98). However, the 3' end of this ORF (ORF 867811) overlaps by two nucleotides with the 3' end of an ORF (ORF 867808) on the opposing strand. If the untranslated 3' end of ORF 867811 mRNA extends far enough into the complementary strand of *hyhA*, this could give rise to a "false signal" due to

cross-hybridization. ORF 867811 encodes a putative ABC transporter (ATP binding protein) and was not included on the microarray. The interference of overlapping mRNAs in microarray analyses has been reported (30), although it has not been systematically analyzed with any genome. With the 18 hydrogenase-related ORFs (Fig. 2), there is considerable variation in the absolute signal intensities, but this appears to arise to some extent from the differences in ORF sizes, several of which are <500 bp.

Hence, of the 21 ORFs (from the total 271 ORFs) that are strongly downregulated (>5-fold) in the presence of S°, 18 represent structural genes for the three hydrogenases. Of the other three S°-responsive ORFs, one (ORF 577932) showed an almost 40-fold decrease in expression, the largest of any of the downregulated ORFs (Table 1). ORF 577932 is annotated as *hypF* and shows 37% sequence identity to its closest homolog, HypF, from the aerobic bacterium *Ralstonia eutrophica* (75). HypF is thought to be involved in biosynthesis of the metal site of [NiFe]-hydrogenases. There is only one gene of this type in *P. furiosus* so it appears that, as in *E. coli* (25), one HypF protein is involved in the maturation of three different hydrogenases (two cytoplasmic and one membrane bound). One would expect HypFA, as a processing enzyme, to be present at low intracellular concentrations, and the reason why its gene appears to be so strongly regulated by S° is not clear at this point. *P. furiosus* contains homologs of several other genes that are involved in hydrogenase maturation and metal insertion in mesophilic bacteria. These include *hycI* (ORF 637303) and *hypACDE* (ORFs 636092, 566869, 567973, and 626147; a version of *hypB* appears not to be present in the *P. furiosus* genome). These ORFs were also included on the microarray. The expression of the *hyp* genes, which are all expressed at

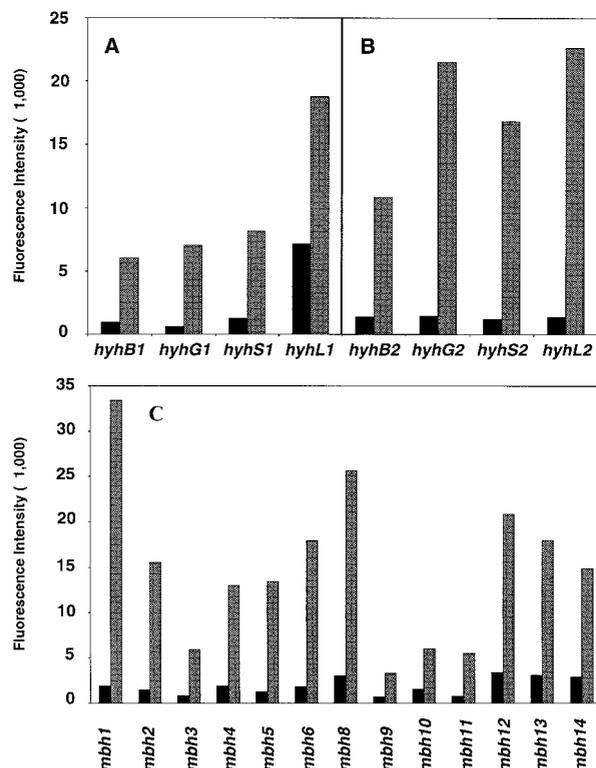


FIG. 2. Expression of ORFs encoding hydrogenase-related subunits with or without  $S^{\circ}$ . The ORFs encoding the subunits of the cytoplasmic hydrogenase I (A), cytoplasmic hydrogenase II (B), and the membrane-bound hydrogenase (C) are arranged according to their positions in their respective operons. These are plotted against signal intensities for cDNA obtained from cells grown with (solid bars) or without (shaded bars)  $S^{\circ}$ . The results with all subunits have a confidence level of at least 96% ( $P > 0.98$  for the cytoplasmic hydrogenases and  $P > 0.96$  for the membrane-bound enzyme).

significant levels (1.5 to 9.6 times the detection limit; see Fig. 1B) decreased by two- to threefold when  $S^{\circ}$  was added, but the expression of *hycI* was below the detection limit under both conditions.

One of the two remaining ORFs that are strongly downregulated by  $S^{\circ}$  (Table 1) may also be related to hydrogenase

biosynthesis. This is ORF 615154, which encodes ornithine carbamoyltransferase (OTCase). This enzyme has been purified from *P. furiosus*, its gene has been cloned and expressed, and its crystal structure is available (36, 70). OTCase catalyzes the transfer of a carbamoyl group from carbamoyl phosphate to ornithine, generating citrulline in the arginine biosynthetic pathway. The potential link between OTCase and the hydrogenases stems from the recent proposal that the source of the C1 (CO and CN) ligands to the metals at the active site of hydrogenases is the carbamoyl group of carbamoyl phosphate (48). In view of the regulation of both OTCase and hydrogenase expression by  $S^{\circ}$  in *P. furiosus*, we speculate that citrulline is the precursor of the carbamoyl group for the synthesis of the active sites of the three hydrogenases in this organism. Citrulline may well be preferred to carbamoyl phosphate because of the high thermal lability of the latter compound (35). One would expect citrulline also to be needed for arginine biosynthesis even in the presence of  $S^{\circ}$ , and this appears to be the case, since OTCase is expressed at close to three times the detection level in  $S^{\circ}$ -grown cells.

The remaining ORF that is strongly downregulated by  $S^{\circ}$  is ORF 51760, and this is annotated as a conserved hypothetical protein (Table 1). It is adjacent to the gene (ORF 49183) encoding PEP synthetase, a gluconeogenic enzyme recently purified from *P. furiosus* (24; see also below), and an ORF (ORF 53135) encoding a conserved hypothetical protein, neither of which showed significant (<2-fold) changes in expression according to the microarray data. The function of the  $S^{\circ}$ -responsive ORF 51760 is therefore unclear at present.

**ORFs strongly upregulated by  $S^{\circ}$ .** Of the 271 ORFs examined, the expression of 12 of them was strongly upregulated (>5-fold) when  $S^{\circ}$  was present in the growth medium (Table 2). In contrast to the ORFs that were downregulated by  $S^{\circ}$ , all of the upregulated ORFs represent putative proteins since none of their products have been characterized from *P. furiosus*. Moreover, the two most strongly regulated ORFs, 1871822 and 1872873, which are upregulated more than 25-fold, are next to each other on the genome. The products of these two ORFs show no sequence similarity to any known protein. Nevertheless, they will be referred to as SipA (1871822) and SipB (1872873) for "sulfur-induced protein" since expression of their genes clearly responds to  $S^{\circ}$  or its metabolites. Both genes

TABLE 2. ORFs whose expression is dramatically (>5-fold) upregulated by  $S^{\circ}$

ORF <sup>a</sup>	Description <sup>a</sup>	Mean intensity ratio (log <sub>2</sub> ± SD) <sup>a</sup>	Change in expression (fold) <sup>a</sup>
1871822	[Conserved hypothetical protein, <i>sipA</i> ] <sup>b</sup>	5.94 ± 1.43	61.4
1872873	[Putative polyferredoxin, <i>sipB</i> ] <sup>b</sup>	4.65 ± 1.90	25.1
1487371	[Tryptophan synthase, subunit beta, <i>trpB-1</i> ]	2.98 ± 0.69	7.9
1805557	[Conserved hypothetical protein]	2.93 ± 1.29	7.6
1008251	[Aspartokinase II alpha subunit]	2.92 ± 0.89	7.6
1131551	[NADH oxidase, <i>noxA-2</i> ]	2.88 ± 0.39	7.4
1825269	[Thermosome, single subunit]	2.83 ± 1.69	7.1
900019	[Acetolactate synthase]	2.78 ± 1.02	6.9
65527	[Fibrillar-like pre-rRNA processing protein]	2.70 ± 0.80	6.5
204761	[Oligopeptide transport system permease protein]	2.68 ± 0.45	6.4
102519	[Glutaredoxin-like protein]	2.61 ± 1.72	6.1
1352206	[NADH dehydrogenase subunit]	2.59 ± 0.59	6.0

<sup>a</sup> See Table 1 for details.

<sup>b</sup> See text for details.

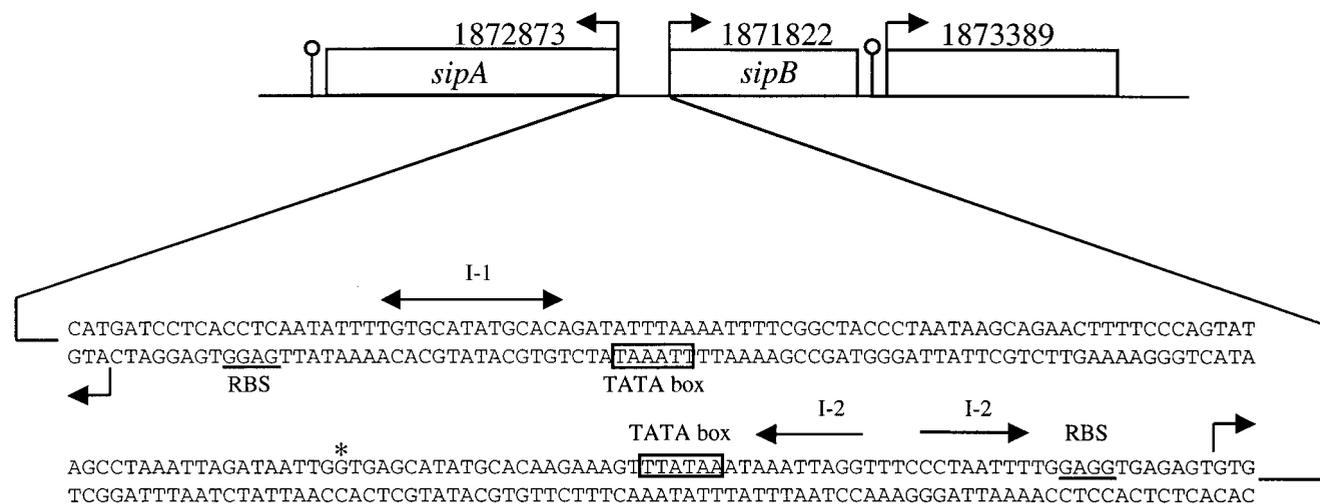


FIG. 3. Genome organization of *sipA* and *sipB*. The back to back ORFs coding for SipA (179 amino acids) and SipB (115 amino acids) are shown with the indicated directions for transcription. Putative transcription stop sites are indicated with a “ball on a stick,” and putative translational start sites are indicated with right-angled arrows. The putative start sites are preceded with possible ribosomal binding sites (underlined) and TATA boxes (boxed). The start site for *sipB* shown here is not the same as that given in the genome annotation (which is indicated with an asterisk [54]). Both *sipA* and *sipB* contain an inverted repeat (I-1 and I-2) directly next to the proposed TATA boxes. The position of the ORF (1873389) encoding the SipB extension (142 amino acids) is indicated.

are also very tightly regulated, since there appears to be little or no expression from them when cells are grown in the absence of  $S^{\circ}$  (their signal intensities are less than the detection limit).

What is now termed *sipA* was included in the list of ORFs for microarray analysis because its product had been previously identified (from the N-terminal sequence) after gel electrophoresis of *P. furiosus* extracts. The protein was present in washed membrane preparations of cells grown with  $S^{\circ}$  but was not detected in cells grown without  $S^{\circ}$  (23). SipA is a 19-kDa protein that contains a single Cys residue. Although it is strongly associated with the membrane, sequence analyses do not reveal any transmembrane helix motifs (23). What we now term *sipB* is currently annotated as a “putative polyferredoxin,” and it contains Cys motifs that could coordinate two [4Fe-4S] clusters. However, the closest homolog (25% sequence identity) to this putative 15-kDa protein is actually the  $\delta$ -subunit of the *P. furiosus* enzyme 2-ketoisovalerate ferredoxin oxidoreductase (VOR; see below), a protein that also contains two [4Fe-4S] centers (see reference 41). Interestingly, although *sipA* and *sipB* are side by side, they are transcribed in opposite directions. Figure 3 shows a detailed analysis of putative transcriptional and ribosomal binding sites for *sipA* and *sipB*, and these strongly suggest that both genes are expressed from a shared promoter domain, as previously shown for the genes encoding  $\beta$ -glucosidase and alcohol dehydrogenase of *P. furiosus* (72). Homologs of *sipA* and *sipB* are present in the genomes of the  $S^{\circ}$ -reducing *Pyrococcus* species, *P. abyssi* (PAB1692 and PAB0578), and *P. horikoshii* (PH1227 and PH0982), but their genes are not arranged back to back such since they are in *P. furiosus* and SipB is much larger (by  $\sim$ 15 kDa) in these other two species. In *P. furiosus*, this SipB 15-kDa extension is encoded by a separate ORF (1873389) that is adjacent to *sipB* (see Fig. 3). The product of ORF 1873389 shows no significant sequence similarity to any characterized

protein. This ORF was included in the microarray analyses, but it is not expressed above the detection limit when cells were grown with or without  $S^{\circ}$  (Table 4). This difference in genome arrangement among the *Pyrococcus* species might be related to the ability of *P. furiosus*, but not *P. abyssi* and *P. horikoshii*, to grow in the absence of  $S^{\circ}$ .

Considering the tight regulation of *sipA* and *sipB* by  $S^{\circ}$  and the likely electron transfer ability of SipB, it seems reasonable to suggest that SipA and SipB form part of a membrane-associated complex that directly reduces  $S^{\circ}$  to  $H_2S$ . Two of the other ORFs listed in Table 2 that are strongly upregulated by  $S^{\circ}$ , ORFs 1131551 and 102519, may also be directly involved in  $S^{\circ}$  reduction since both show sequence similarity to the disulfide oxidoreductase family of enzymes. ORF 1131551 is annotated as NADH oxidase (*noxA-2*) and would encode a 49-kDa cytoplasmic protein that shows significant sequence similarity to NADH peroxidase, a flavoprotein with an active site cysteine (64). ORF 102519 is predicted to encode a 26-kDa cytoplasmic protein that contains four Cys residues (as two -CxxC-motifs). Both ORFs are expressed at near the detection limit in the absence of  $S^{\circ}$  but increase to the “highly expressed” category ( $>10$ -fold the detection limit; see below) in  $S^{\circ}$ -grown cells, and neither appears to be part of an operon. Homologs of both ORFs are also present in *P. abyssi* (PAB0936 and PAB2245) and *P. horikoshii* (PH0572 and PH0178).

Two of the other  $S^{\circ}$ -induced proteins that might be related to  $S^{\circ}$  metabolism are ORF 1805557, which is annotated as a conserved hypothetical protein, and ORF 1352727, which is annotated as an NADH dehydrogenase (Table 2). The expression of ORF 1805557 changes from below the detection limit to twice the detection limit when cells are grown with  $S^{\circ}$ . This ORF would encode a 27-kDa cytoplasmic protein where the N-terminal part has high sequence similarity to ATP/GTP-binding proteins (13), while the C-terminal half contains six Cys residues, including a CxxCxxx motif. ORF 1352727 is

part of a putative operon containing 13 ORFs, most of which have homology to membrane-bound NADH dehydrogenases. Hagen and coworkers suggested that this operon encodes a (fourth) hydrogenase system in *P. furiosus* (62), but this seems unlikely in view of the upregulation of ORF 1352727 by S° and the corresponding dramatic decrease in membrane-associated hydrogenase activity (1). Moreover, two other ORFs (1350388 and 1359534) in this putative operon were included in the microarray analysis, and these are also upregulated by S° (three- and fourfold [data not shown]). It is tempting to postulate that ORFs 1805557 and 1352727 (and possibly other ORFs in the putative operon) are involved in nucleotide binding and electron transfer during S° reduction, although further evidence is obviously required to substantiate this.

Four of the remaining six ORFs that are upregulated more than fivefold by S° (Table 2) appear to encode proteins that are involved in amino acid metabolism. These are annotated as an oligopeptide permease (ORF 204761), acetolactate synthase (which is involved in the biosynthesis of branched chain amino acids [ORF 900019]), and as subunits of aspartokinase (ORF 1008251) and tryptophan synthase (ORF 1487371). Although all are putative proteins at present, these data support the relationship between S° and amino acid metabolism that was recently established by growth studies of *P. furiosus*, in which it was shown that peptides can only serve as the primary carbon source if S° is present (1). All four of these ORFs are expressed at significant levels in the absence of S° (more than twice the detection limit), but all four fall into the highly expressed category (>10-fold the detection limit [see below]) in S°-grown cells. Of the other two upregulated ORFs in Table 2, one is annotated as the thermosome (ORF 1825269). This protein is highly conserved in archaea and is a type II chaperonin, a class that also includes the eukaryotic TCP-1(CCT) proteins (7). The thermosome is expressed at about twice the detection limit in the absence of S°, although why its expression should increase ~7-fold in the presence of S° is not clear.

The other S°-responsive ORF in Table 2 is annotated as a fibrillar-like protein (ORF 62257). It shows more than 50% sequence identity to a key eukaryotic nucleolar protein that associates with small nucleolar RNAs (snoRNAs) directing 2'-O-ribose methylation of the rRNA. Homologs of fibrillar and some of its associated proteins, such as Nop56/58, as well as snoRNAs, have only recently been identified in archaea, although their function in these organisms is not clear (16, 33, 44, 45, 47). In *P. furiosus*, the ORF encoding the fibrillar homolog is directly adjacent to a homolog of Nop56/58 (ORF 66216). The microarray data indicate that these ORFs are expressed at twice and six times the detection limit, respectively, in the absence of S°, but their expression increases a further six- and threefold, respectively, when S° is present. Indeed, the microarray experiment included two copies of the ORF encoding the fibrillar homolog, and both gave remarkably similar results with RNA from cells grown with (an intensity of  $22,237 \pm 1,074$ , see Fig. 1B) and without S° ( $3,590 \pm 185$ ). *P. furiosus* also contains five homologs of a family of RNA m<sup>5</sup>C methyl transferases (52), and four were included in the microarray analysis (ORFs 1191037, 674979, 179810, and 1197174). All of them are annotated as a nucleolar NOL1-NOP2-sun family protein, but none appear to be regulated by S°. Thus, it is not obvious why S° should cause the genes

encoding homologs of eukaryotic nucleolar rRNA processing proteins to be expressed at such high levels. Fibrillar is also found in archaea that do not metabolize S°, such as methanogens (20), so this protein must have a more general role, albeit one that is associated with S° metabolism, at least in *P. furiosus*.

**Sulfur-independent ORFs.** In the preceding discussion we have considered all 33 of the 271 ORFs that show a >5-fold change in expression when the cells were grown in the presence of S°. In addition to these 33, there are 84 ORFs that showed between a two- to fivefold response to S° (data not shown). At this point further comment on the nature and function of these ORFs seems premature since there are no obvious trends and corroborating experimental data are required. That leaves 154 of the 271 ORFs remaining that show little if any response to S° (<2-fold), and these can be separated into three groups: (i) 20 ORFs that appear to be the most highly expressed, where the transcripts are at >10 times the detection level (> 20,000 intensity units [see Fig. 1B]); (ii) 34 ORFs that appear to be poorly expressed, if at all, where the transcripts are below the detection level (<2,000 intensity units); and (iii) 100 ORFs that appear to be moderately expressed (2,000 to 20,000 intensity units). It should be noted that the relative signal intensities do not necessarily correlate with the degrees of expression of different ORFs because intensity depends on many factors, including PCR product size and efficiencies of labeling and hybridization. However, the results and discussion below show that there may be some merit in this general assumption.

As discussed above, it was anticipated that S° reduction by *P. furiosus* would involve Mo- or W-containing proteins, analogous to the situation in mesophilic S° reducers. For example, the polysulfide (sulfur) reductase of the mesophilic bacterium, *W. succinogenes*, is a molybdoenzyme (32). Consequently, the microarray for *P. furiosus* included numerous ORFs involved in the metabolism of these metals, as well as putative molybdo- and tungstoenzymes. In fact, the genome contains two ORFs (1175771 and 1419794) that show sequence similarity to members of the ubiquitous molybdenum-containing family of enzymes (22). However, both ORFs are expressed either below or close to the detection limit and show no response to the presence of S° (data not shown). In addition, *P. furiosus* appears to contain five distinct tungstoenzymes. Three of them have been purified and characterized (abbreviated AOR, GAPOR, and FOR and encoded by ORFs 356987, 478142, and 1145403, respectively), and two are putative (WOR4 and WOR5, encoded by ORFs 1812948 and 1385199, respectively [see reference 57]). FOR, GAPOR, and AOR are in the highly expressed category (>10-fold detection limit [see Table 3]), while WOR4 and WOR5 are moderately expressed and are only slightly upregulated in the presence of S° (data not shown). Similarly, *P. furiosus* contains several ORFs that would encode homologs of proteins involved in the biosynthesis of pterin (designated *moa*, *mob*, and *moe* [see reference 51]), the cofactor that binds the Mo and W in molybdo- and tungstoenzymes. None of these ORFs showed any significant changes in expression when cells were grown with S°, with the exception of ORF 1805557 (discussed above). In light of these data, it is concluded that S° reduction in *P. furiosus* does not involve Mo-, W-, or pterin-containing proteins—or at least those that can be identified by sequence similarity to known proteins of this type.

Some comment is also necessary on the nature of the pro-

TABLE 3. Highly expressed S<sup>o</sup>-independent ORFs

ORF <sup>a</sup>	Description <sup>b</sup>
49183.....	Phosphoenolpyruvate synthetase, <i>ppsA</i> (24)
143318.....	[Conserved hypothetical protein]
232621.....	Enolase (2-phosphoglycerate dehydratase) (49)
236793.....	[Hexulose-6-phosphate synthase]
358419.....	Aldehyde ferredoxin oxidoreductase, <i>aor</i> (43)
478142.....	GAPOR, <i>gor</i> (42)
683389.....	[Methylmalonyl-CoA decarboxylase, subunit alpha, <i>mmdA</i> ]
720985.....	[Alkyl hydroperoxide reductase subunit C]
925374.....	POR beta, <i>porB</i> (6)
926380.....	POR alpha, <i>porA</i> (6)
927947.....	VOR subunit beta, <i>vorB</i> (19)
928888.....	VOR subunit alpha, <i>vorA</i> (19)
1145403.....	Formaldehyde ferredoxin oxidoreductase, <i>for</i> (57)
1208774.....	[LSU ribosomal protein L10]
1210188.....	Superoxide reductase, <i>sor</i> (77)
1210814.....	[Rubrerythrin]
1425928.....	[Ethylene-inducible protein homolog]
1493675.....	Glutamate dehydrogenase, <i>gdh</i> (55)
1599202.....	Intracellular protease, <i>pfpI</i> (18)
1619038.....	[Putative trehalose synthase]

<sup>a</sup> ORFs displaying average fluorescent intensities >20,000 (see Fig. 1B).

<sup>b</sup> See Table 1 for details.

teins encoded by the ORFs that are designated as either poorly or highly expressed. For example, the ORF with the highest intensity on the microarray (~23 times the detection limit) encodes phosphoenolpyruvate synthase (*ppsA*). This enzyme was recently purified from *P. furiosus* and was shown to be present at extremely high cellular concentrations (24), a finding consistent with the highly expressed connotation. In fact, of the 20 ORFs that show the highest intensities on the microarray (>10 times the detection limit; see Table 3), 12 of them encode proteins that have been purified from *P. furiosus*. The biochemical data therefore supports the notion that, in general, these ORFs encode proteins that are indeed present at relatively high cellular concentrations and that signal intensity is a reasonable measure of this. Further evidence for this qualitative relationship comes from the 34 ORFs that are in the poorly expressed category, since none of the proteins that they are proposed to encode have been characterized from *P. furiosus*, see Table 4 (although one, encoding a putative phosphoglycerate kinase, has been cloned and expressed from the related organism, *P. woesei* [10]). Of the 100 (of 271) ORFs that are in the moderately expressed category, the products of 12 of them have been characterized (data not shown).

Four of the highly expressed ORFs listed in Table 3 encode the two  $\alpha$ -subunits and the two  $\beta$ -subunits of two closely related, heterotetrameric enzymes, POR (6) and VOR (19). Both of these enzymes have been purified from *P. furiosus*. The genes encoding them form three adjacent operons, in which the third is a common gene encoding their  $\gamma$  subunits (31). The gene arrangement and their transcript intensities in cells with or without S<sup>o</sup> are shown in Fig. 4. Although the twofold difference in expression of *porG* is statistically significant ( $P > 0.95$ ), S<sup>o</sup> does not dramatically affect the expression of POR and, accordingly, the specific activity of POR is comparable in the two cell types (1). Moreover, while the apparent amounts of the mRNA species encoding the two  $\alpha$ -subunits and the two  $\beta$ -subunits are consistent with the high cellular concentrations

TABLE 4. Poorly expressed S<sup>o</sup>-independent ORFs<sup>a</sup>

ORF	Description <sup>b</sup>
53135.....	[Conserved hypothetical protein]
349245.....	[Flagellum-related protein D, putative]
350457.....	[Conserved hypothetical protein]
350944.....	[Flagellin B2 precursor]
351748.....	[Flagellin B2 precursor]
373060.....	[ABC transporter, OppBC family]
379218.....	[ $\beta$ -Galactosidase precursor]
488057.....	[DNA mismatch repair protein, <i>MutS</i> ]
562899.....	[Molybdopterin converting factor, subunit 1, <i>moaD</i> ]
575520.....	[NADH oxidase, <i>noxA-4</i> /nitrite reductase]
637268.....	[Molybdopterin-guanine dinucleotide biosynthesis protein, <i>mobA</i> ]
637303.....	[Hydrogenase maturation protease, <i>hycI</i> ]
722839.....	[Conserved hypothetical protein]
738742.....	[Conserved hypothetical protein]
743892.....	[Putative proline depeptidase]
834030.....	[Conserved hypothetical protein]
838710.....	[Conserved hypothetical protein]
880926.....	[Ferric enterobactin transport ATP-binding protein homolog]
881669.....	[Iron(III) ABC transporter, permease protein, <i>hemU-1</i> ]
882732.....	[Iron(III) ABC transporter ATP-binding protein, <i>hemV-2</i> ]
962748.....	[Alkaline phosphatase IV precursor]
1012695.....	[Phosphoglycerate kinase]
1138558.....	[Transcriptional regulator (FurR family)]
1158892.....	[Dissimilatory sulfate adenylyl transferase]
1165967.....	[4-Aminobutyrate aminotransferase]
1197174.....	[Putative nucleolar protein II, Nol1-Nop2-sun family]
1208389.....	[Conserved hypothetical protein]
1417217.....	[Sugar-binding transport ATP-binding protein]
1647980.....	[2-keto-acid:ferredoxin oxidoreductase subunit alpha]
1668574.....	[Sarcosine oxidase, alpha subunit, <i>SoxA</i> ]
1669077.....	[Putative polyferredoxin, <i>muhB</i> ]
1711295.....	[Molybdenum cofactor biosynthesis protein, <i>moaC</i> ]
1873595.....	[Nitrogen reductase, N terminus]
1873914.....	[Ferredoxin-family protein]

<sup>a</sup> ORFs displaying average fluorescent intensities below 2,000 (see Fig. 1B).

<sup>b</sup> See Table 1 for details.

of the enzymes (1, 6, 19), this is not true for the other subunits and particular for the ORF encoding the  $\delta$ -subunit of VOR, which is barely above the detection limit. How these three operons are regulated such that they produce two functional enzymes at high cellular concentrations and with a common subunit is not clear from the transcript levels measured by using the DNA microarray. Thus, while there may be some overall relationship between transcript intensities on the microarray slide and cellular concentration, that encoding the  $\delta$ -subunit of VOR is a notable exception. On the other hand, the biochemical and electrophoretic data strongly support the dramatic downregulation of the expression of the three hydrogenases by S<sup>o</sup>, and the dramatic upregulation of at least one of the "conserved hypothetical" ORFs (*sipA*) that are now proposed to form a new type of S<sup>o</sup>-responsive complex. Further analyses of this type are now required to determine the nature of this complex, whether it is directly involved in S<sup>o</sup> reduction, and the nature of the effector molecule that mediates the S<sup>o</sup> response.

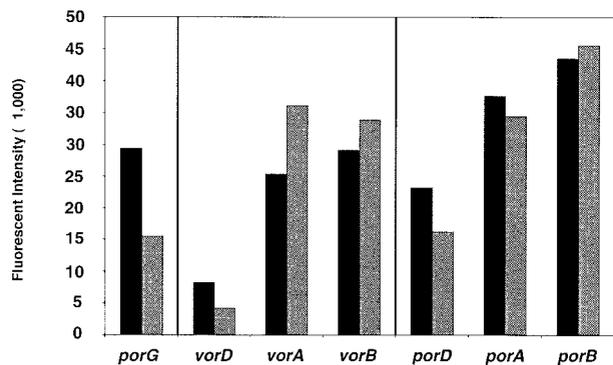


FIG. 4. Expression of the ORFs encoding POR and VOR with or without S<sup>0</sup>. The genes (A, B, and D) encoding the  $\alpha$ -,  $\beta$ -, and  $\delta$ -subunits of POR and of VOR are arranged according to their positions in their respective operons, along with the gene (G) encoding the  $\gamma$ -subunit that is shared by the two enzymes (31). For each ORF, the signal intensities are indicated for cDNA obtained from cells grown with (solid bars) or without (shaded bars) S<sup>0</sup>.

#### ACKNOWLEDGMENTS

This research was funded by grants to M.W.W.A. from the National Institutes of Health (GM 60329), the National Science Foundation (MCB 9904624, MCB 9809060, and BES-0004257), and the Department of Energy (FG05-95ER20175 and contract 992732401 with Argonne National Laboratory) and to J.Z. from the Department of Energy under the Microbial Genome Program and the Natural and Accelerated Bioremediation Research Program of the Office of Biological and Environmental Research. Oak Ridge National Laboratory is managed by the University of Tennessee-Battelle LLC for the Department of Energy under contract DE-AC05-00OR22725.

We thank Gary Li, Scott Lee, and Marc Sudman for their help in preparing the microarray slides; F. Robb and R. Weiss for making the *P. furiosus* sequence available prior to publication; and Frank E. Jenney, Jr., Angeli Lal Menon, James F. Holden, Eleanor Green, and Rajat Sapra for many helpful discussions.

#### REFERENCES

- Adams, M. W. W., J. F. Holden, A. L. Menon, G. J. Schut, A. M. Grunden, C. Hou, A. M. Hutchins, F. E. Jenney, Jr., C. Kim, K. Ma, G. Pan, R. Roy, R. Sapra, S. V. Story, and M. F. Verhagen. 2001. Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **183**:716–724.
- Arfin, S. M., A. D. Long, E. T. Ito, L. Toller, M. M. Riehle, E. S. Paegle, and G. W. Hatfield. 2000. Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J. Biol. Chem.* **275**:29672–29684.
- Barbier, G., A. Godfroy, J. R. Meunier, J. Querellou, M. A. Cambon, F. Lesongeur, P. A. Grimont, and G. Raguene. 1999. *Pyrococcus glycovorans* sp. nov., a hyperthermophilic archaeon isolated from the East Pacific Rise. *Int. J. Syst. Bacteriol.* **49**(Pt. 4):1829–1837.
- Blamey, J., M. Chiong, C. Lopez, and E. Smith. 1999. Optimization of the growth conditions of the extremely thermophilic microorganisms *Thermococcus celer* and *Pyrococcus woesei*. *J. Microbiol. Methods* **38**:169–175.
- Blamey, J. M., and M. W. W. Adams. 1994. Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry* **33**:1000–1007.
- Blamey, J. M., and M. W. W. Adams. 1993. Purification and characterization of pyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochim. Biophys. Acta* **1161**:19–27.
- Bosch, G., W. Baumeister, and L. O. Essen. 2000. Crystal structure of the beta-apical domain of the thermosome reveals structural plasticity in the protrusion region. *J. Mol. Biol.* **301**:19–25.
- Brown, P. O., and D. Botstein. 1999. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21**:33–37.
- Bryant, F. O., and M. W. W. Adams. 1989. Characterization of hydrogenase from the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *J. Biol. Chem.* **264**:5070–5079.
- Crowhurst, G., J. McHarg, and J. A. Littlechild. 2001. Phosphoglycerate kinases from bacteria and archaea. *Methods Enzymol.* **331**:90–104.
- de Saizieu, A., U. Certa, J. Warrington, C. Gray, W. Keck, and J. Mous. 1998.

- Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays. *Nat. Biotechnol.* **16**:45–48.
- Duggan, D. J., M. Bittner, Y. Chen, P. Meltzer, and J. M. Trent. 1999. Expression profiling using cDNA microarrays. *Nat. Genet.* **21**:10–14.
- Eaves, D. J., T. Palmer, and D. H. Boxer. 1997. The product of the molybdenum cofactor gene *mobB* of *Escherichia coli* is a GTP-binding protein. *Eur. J. Biochem.* **246**:690–697.
- Edman, C. F., P. Mehta, R. Press, C. A. Spargo, G. T. Walker, and M. Nerenberg. 2000. Pathogen analysis and genetic predisposition testing using microelectronic arrays and isothermal amplification. *J. Investig. Med.* **48**:93–101.
- Fiala, G., and K. O. Stetter. 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Arch. Microbiol.* **145**:56–61.
- Gaspin, C., J. Cavaille, G. Erauso, and J. P. Bachellerie. 2000. Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes. *J. Mol. Biol.* **297**:895–906.
- Gingeras, T. R., G. Ghandour, E. Wang, A. Berno, P. M. Small, F. Drobniowski, D. Alland, E. Desmond, M. Holodniy, and J. Drenkow. 1998. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res.* **8**:435–448.
- Halio, S. B., I. I. Blumentals, S. A. Short, B. M. Merrill, and R. M. Kelly. 1996. Sequence, expression in *Escherichia coli*, and analysis of the gene encoding a novel intracellular protease (PfpI) from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **178**:2605–2612.
- Heider, J., X. Mai, and M. W. W. Adams. 1996. Characterization of 2-ketoisovalerate ferredoxin oxidoreductase, a new and reversible coenzyme A-dependent enzyme involved in peptide fermentation by hyperthermophilic archaea. *J. Bacteriol.* **178**:780–787.
- Hickey, A. J., A. J. Macario, and E. Conway de Macario. 2000. Identification of genes in the genome of the archaeon *Methanosarcina mazei* that code for homologs of nuclear eukaryotic molecules involved in RNA processing. *Gene* **253**:77–85.
- Hihara, Y., A. Kamei, M. Kanehisa, A. Kaplan, and M. Ikeuchi. 2001. DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell* **13**:793–806.
- Hille, R. 1996. The mononuclear molybdenum enzymes. *Chem. Rev.* **96**:2757–2816.
- Holden, J. F., F. L. Poole, S. L. Tollaksen, C. S. Giometti, H. Lim, J. R. Yates, and M. W. W. Adams. 2001. Identification of membrane proteins in the hyperthermophilic archaeon *Pyrococcus furiosus* using proteomics and prediction programs. *Comp. Funct. Genomics* **2**:275–288.
- Hutchins, A. M., J. F. Holden, and M. W. W. Adams. 2001. Phosphoenolpyruvate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **183**:709–715.
- Jacobi, A., R. Rossmann, and A. Bock. 1992. The *hyp* operon gene products are required for the maturation of catalytically active hydrogenase isoenzymes in *Escherichia coli*. *Arch. Microbiol.* **158**:444–451.
- Jia, M. H., R. A. Larossa, J. M. Lee, A. Rafalski, E. Deroose, G. Gonye, and Z. Xue. 2000. Global expression profiling of yeast treated with an inhibitor of amino acid biosynthesis, sulfometuron methyl. *Physiol. Genomics* **3**:83–92.
- Johnson, M. K., D. C. Rees, and M. W. W. Adams. 1996. Tungstoenzymes. *Chem. Rev.* **96**:2817–2839.
- Kawarabayashi, Y., M. Sawada, H. Horikawa, Y. Haikawa, Y. Hino, S. Yamamoto, M. Sekine, S. Baba, H. Kosugi, A. Hosoyama, Y. Nagai, M. Sakai, K. Ogura, R. Otsuka, H. Nakazawa, M. Takamiya, Y. Ohfuku, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, K. Aoki, and H. Kikuchi. 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3 (supplement). *DNA Res.* **5**:147–155.
- Kengen, S. W., J. E. Tuininga, F. A. de Bok, A. J. Stams, and W. M. de Vos. 1995. Purification and characterization of a novel ADP-dependent glucokinase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **270**:30453–30457.
- Khodursky, A. B., B. J. Peter, N. R. Cozzarelli, D. Botstein, P. O. Brown, and C. Yanofsky. 2000. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**:12170–12175.
- Kletzin, A., and M. W. W. Adams. 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. *J. Bacteriol.* **178**:248–257.
- Krafft, T., R. Gross, and A. Kroger. 1995. The function of *Wolinella succinogenes* *psr* genes in electron transport with polysulphide as the terminal electron acceptor. *Eur. J. Biochem.* **230**:601–606.
- Lafontaine, D. L., and D. Tollervey. 1998. Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. *Trends Biochem. Sci.* **23**:383–388.
- Lee, H., G. H. Greeley, and E. W. Englander. 2001. Age-associated changes in gene expression patterns in the duodenum and colon of rats. *Mech. Ageing Dev.* **122**:355–371.
- Legrain, C., V. Stalon, J. P. Noullez, A. Mercenier, J. P. Simon, K. Broman,

- and J. M. Wiame. 1977. Structure and function of ornithine carbamoyltransferases. *Eur. J. Biochem.* **80**:401–409.
36. Legrain, C., V. Villeret, M. Roovers, D. Gigot, O. Dideberg, A. Pierard, and N. Glandsdorff. 1997. Biochemical characterisation of ornithine carbamoyltransferase from *Pyrococcus furiosus*. *Eur. J. Biochem.* **247**:1046–1055.
  37. Ma, K., R. N. Schicho, R. M. Kelly, and M. W. W. Adams. 1993. Hydrogenase of the hyperthermophile *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: evidence for a sulfur-reducing hydrogenase ancestor. *Proc. Natl. Acad. Sci. USA* **90**:5341–5344.
  38. Ma, K., R. Weiss, and M. W. W. Adams. 2000. Characterization of hydrogenase II from the hyperthermophilic archaeon *Pyrococcus furiosus* and assessment of its role in sulfur reduction. *J. Bacteriol.* **182**:1864–1871.
  39. Mai, X., and M. W. W. Adams. 1994. Indolepyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*. A new enzyme involved in peptide fermentation. *J. Biol. Chem.* **269**:16726–16732.
  40. Mai, X. H., and M. W. W. Adams. 1996. Purification and characterization of two reversible and ADP-dependent acetyl coenzyme-A synthetases from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **178**:5897–5903.
  41. Menon, A. L., H. Hendrix, A. Hutchins, M. F. J. M. Verhagen, and M. W. W. Adams. 1998. The delta-subunit of pyruvate ferredoxin oxidoreductase from *Pyrococcus furiosus* is a redox-active, iron-sulfur protein: evidence for an ancestral relationship with 8Fe-type ferredoxins. *Biochemistry* **37**:12838–12846.
  42. Mukund, S., and M. W. W. Adams. 1995. Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **270**:8389–8392.
  43. Mukund, S., and M. W. W. Adams. 1991. The novel tungsten-iron-sulfur protein of the hyperthermophilic archaeobacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase. Evidence for its participation in a unique glycolytic pathway. *J. Biol. Chem.* **266**:14208–14216.
  44. Newman, D. R., J. F. Kuhn, G. M. Shanab, and E. S. Maxwell. 2000. Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA* **6**:861–879.
  45. Noon, K. R., E. Bruenger, and J. A. McCloskey. 1998. Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*. *J. Bacteriol.* **180**:2883–2888.
  46. Oh, M. K., and J. C. Liao. 2000. Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnol. Prog.* **16**:278–286.
  47. Omer, A. D., T. M. Lowe, A. G. Russell, H. Ebhardt, S. R. Eddy, and P. P. Dennis. 2000. Homologs of small nucleolar RNAs in *Archaea*. *Science* **288**:517–522.
  48. Paschos, A., R. S. Glass, and A. Bock. 2001. Carbamoyl phosphate requirement for synthesis of the active center of [NiFe]-hydrogenases. *FEBS Lett.* **488**:9–12.
  49. Peak, M. J., J. G. Peak, F. J. Stevens, J. Blamey, X. Mai, Z. H. Zhou, and M. W. W. Adams. 1994. The hyperthermophilic glycolytic enzyme enolase in the archaeon *Pyrococcus furiosus*: comparison with mesophilic enolases. *Arch. Biochem. Biophys.* **313**:280–286.
  50. Pedroni, P., A. Della Volpe, G. Galli, G. M. Mura, C. Pratesi, and G. Grandi. 1995. Characterization of the locus encoding the [Ni-Fe] sulfhydrogenase from the archaeon *Pyrococcus furiosus*: evidence for a relationship to bacterial sulfite reductases. *Microbiology* **141**:449–458.
  51. Rajagopalan, K. V. 1997. Biosynthesis and processing of the molybdenum cofactors. *Biochem. Soc. Trans.* **25**:757–761.
  52. Reid, R., P. J. Greene, and D. V. Santi. 1999. Exposition of a family of RNA m(5)C methyltransferases from searching genomic and proteomic sequences. *Nucleic Acids Res.* **27**:3138–3145.
  53. Richmond, C. S., J. D. Glasner, R. Mau, H. Jin, and F. R. Blattner. 1999. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **27**:3821–3835.
  54. Robb, F. T., D. L. Maeder, J. R. Brown, J. DiRuggiero, M. D. Stump, R. K. Yeh, R. B. Weiss, and D. M. Dunn. 2001. Genomic sequence of hyperthermophile *Pyrococcus furiosus*: implications for physiology and enzymology. *Methods Enzymol.* **330**:134–157.
  55. Robb, F. T., J. B. Park, and M. W. W. Adams. 1992. Characterization of an extremely thermostable glutamate dehydrogenase: a key enzyme in the primary metabolism of the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Biochim. Biophys. Acta* **1120**:267–272.
  56. Ross, D. T., U. Scherf, M. B. Eisen, C. M. Perou, C. Rees, P. Spellman, V. Iyer, S. S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J. C. Lee, D. Lashkari, D. Shalon, T. G. Myers, J. N. Weinstein, D. Botstein, and P. O. Brown. 2000. Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **24**:227–235.
  57. Roy, R., A. L. Menon, and M. W. W. Adams. 2001. Aldehyde oxidoreductases from *Pyrococcus furiosus*. *Methods Enzymol.* **331**:132–144.
  58. Salama, N., K. Guillemin, T. K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow. 2000. A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc. Natl. Acad. Sci. USA* **97**:14668–14673.
  59. Sapra, R., M. F. Verhagen, and M. W. W. Adams. 2000. Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **182**:3423–3428.
  60. Schena, M., D. Shalon, R. W. Davis, and P. O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**:467–470.
  61. Schönheit, P., and T. Schäfer. 1995. Metabolism of hyperthermophiles. *World J. Microbiol. Biotech.* **11**:26–57.
  62. Silva, P. J., E. C. van den Ban, H. Wassink, H. Haaker, B. de Castro, F. T. Robb, and W. R. Hagen. 2000. Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. *Eur. J. Biochem.* **267**:6541–6551.
  63. Staudt, L. M., and P. O. Brown. 2000. Genomic views of the immune system. *Annu. Rev. Immunol.* **18**:829–859.
  64. Stehle, T., S. A. Ahmed, A. Claiborne, and G. E. Schulz. 1991. Structure of NADH peroxidase from *Streptococcus faecalis* 10C1 refined at 2.16 Å resolution. *J. Mol. Biol.* **221**:1325–1344.
  65. Stetter, K. O. 1999. Extremophiles and their adaptation to hot environments. *FEBS Lett.* **452**:22–25.
  66. Stetter, K. O. 1996. Hyperthermophilic prokaryotes. *Fems Microbiol. Rev.* **18**:149–158.
  67. Talaat, A. M., P. Hunter, and S. A. Johnston. 2000. Genome-directed primers for selective labeling of bacterial transcripts for DNA microarray analysis. *Nat. Biotechnol.* **18**:679–682.
  68. Tuininga, J. E., C. H. Verhees, J. van der Oost, S. W. Kengen, A. J. Stams, and W. M. de Vos. 1999. Molecular and biochemical characterization of the ADP-dependent phosphofructokinase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **274**:21023–21028.
  69. van der Oost, J., G. Schut, S. W. Kengen, W. R. Hagen, M. Thomm, and W. M. de Vos. 1998. The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. *J. Biol. Chem.* **273**:28149–28154.
  70. Villeret, V., B. Clantin, C. Tricot, C. Legrain, M. Roovers, V. Stalon, N. Glandsdorff, and J. Van Beeumen. 1998. The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extremely high temperatures. *Proc. Natl. Acad. Sci. USA* **95**:2801–2806.
  71. Voorhorst, W. G., R. I. Eggen, E. J. Luesink, and W. M. de Vos. 1995. Characterization of the *celB* gene coding for β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *J. Bacteriol.* **177**:7105–7111.
  72. Voorhorst, W. G., Y. Gueguen, A. C. Geerling, G. Schut, I. Dahlke, M. Thomm, J. van der Oost, and W. M. de Vos. 1999. Transcriptional regulation in the hyperthermophilic archaeon *Pyrococcus furiosus*: coordinated expression of divergently oriented genes in response to beta-linked glucose polymers. *J. Bacteriol.* **181**:3777–3783.
  73. Wei, Y., J. M. Lee, C. Richmond, F. R. Blattner, J. A. Rafalski, and R. A. LaRossa. 2001. High-density microarray-mediated gene expression profiling of *Escherichia coli*. *J. Bacteriol.* **183**:545–556.
  74. Wilson, M., J. DeRisi, H. H. Kristensen, P. Imboden, S. Rane, P. O. Brown, and G. K. Schoolnik. 1999. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc. Natl. Acad. Sci. USA* **96**:12833–12838.
  75. Wolf, I., T. Buhrke, J. Dervede, A. Pohlmann, and B. Friedrich. 1998. Duplication of hyp genes involved in maturation of [NiFe] hydrogenases in *Alcaligenes eutrophus* H16. *Arch. Microbiol.* **170**:451–459.
  76. Ye, R. W., W. Tao, L. Bedzyk, T. Young, M. Chen, and L. Li. 2000. Global gene expression profiles of *Bacillus subtilis* grown under anaerobic conditions. *J. Bacteriol.* **182**:4458–4465.
  77. Yeh, A. P., Y. Hu, F. E. Jenney, Jr., M. W. W. Adams, and D. C. Rees. 2000. Structures of the superoxide reductase from *Pyrococcus furiosus* in the oxidized and reduced states. *Biochemistry* **39**:2499–2508.
  78. Yoshida, K., K. Kobayashi, Y. Miwa, C. M. Kang, M. Matsunaga, H. Yamaguchi, S. Tojo, M. Yamamoto, R. Nishi, N. Ogasawara, T. Nakayama, and Y. Fujita. 2001. Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res.* **29**:683–692.
  79. Zimmer, D. P., E. Soupene, H. L. Lee, V. F. Wendisch, A. B. Khodursky, B. J. Peter, R. A. Bender, and S. Kustu. 2000. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. *Proc. Natl. Acad. Sci. USA* **97**:14674–14679.