A new type of bacterial sulfatase reveals a novel maturation pathway in prokaryotes

Olivier O. Berteau, Alain A. Guillot, Alhosna Benjdia, Sylvie S. Rabot

▶ **To cite this version:**

Olivier O. Berteau, Alain A. Guillot, Alhosna Benjdia, Sylvie S. Rabot. A new type of bacterial sulfatase reveals a novel maturation pathway in prokaryotes. Journal of Biological Chemistry, 2006, 281 (32), pp.22464-22470. 10.1074/jbc.M602504200 . hal-02656640

HAL Id: hal-02656640
https://hal.inrae.fr/hal-02656640

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 32, pp. 22464–22470, August 11, 2006
© 2006 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

A New Type of Bacterial Sulfatase Reveals a Novel Maturation Pathway in Prokaryotes*

Received for publication, March 16, 2006, and in revised form, June 8, 2006 Published, JBC Papers in Press, June 9, 2006, DOI 10.1074/jbc.M602504200

Olivier Berteau[‡1], **Alain Guillot**[§], **Alhosna Benjdia**[‡], **and Sylvie Rabot**[‡]

From the [‡]*Unité d'Ecologie et Physiologie du Système Digestif and the* [§]*Unité Biochimie Bactérienne, Institut National de la Recherche Agronomique, 78352 Jouy-en-Josas, France*

Sulfatases are a highly conserved family of enzymes found in all three domains of life. To be active, sulfatases undergo a unique post-translational modification leading to the conversion of either a critical cysteine ("Cys-type" sulfatases) or a serine ("Ser-type" sulfatases) into a C*α*-formylglycine (FGly). This conversion depends on a strictly conserved sequence called "sulfatase signature" (C/S)*X*P*X*R. In a search for new enzymes from the human microbiota, we identified the first sulfatase from Firmicutes. Matrix-assisted laser desorption ionization time-of-flight analysis revealed that this enzyme undergoes conversion of its critical cysteine residue into FGly, even though it has a modified (C/S)*X*A*X*R sulfatase signature. Examination of the bacterial and archaeal genomes sequenced to date has identified many genes bearing this new motif, suggesting that the definition of the sulfatase signature should be expanded. Furthermore, we have also identified a new Cys-type sulfatase-maturating enzyme that catalyzes the conversion of cysteine into FGly, in anaerobic conditions, whereas the only enzyme reported so far to be able to catalyze this reaction is oxygen-dependent. The new enzyme belongs to the radical *S*-adenosyl-L-methionine enzyme superfamily and is related to the Ser-type sulfatase-maturating enzymes. This finding leads to the definition of a new enzyme family of sulfatase-maturating enzymes that we have named anSME (anaerobic sulfatase-maturating enzyme). This family includes enzymes able to maturate Cys-type as well as Ser-type sulfatases in anaerobic conditions. In conclusion, our results lead to a new scheme for the biochemistry of sulfatases maturation and suggest that the number of genes and bacterial species encoding sulfatase enzymes is currently underestimated.

Sulfatases are widespread enzymes found from prokaryotes to eukaryotes. They are involved in various metabolic processes, ranging from sulfate starvation response in bacteria to hormone biosynthesis and the modulation of developmental cell signaling in mammals (1). In humans, their biological relevance is particularly underlined by their involvement in several inherited diseases such as mucopolysaccharidoses (2), metachromatic leukodystrophy (3), X-linked ichthyosis, chondrodysplasia punctata (4), and the rare multiple sulfatase deficiency syndrome (1, 5).

Sulfatases act on a broad diversity of substrates, which leads to their classification by the IUBMB into 17 classes (from EC 3.1.6.1 to EC 3.1.6.18). Despite this apparent heterogeneity, the primary and tertiary structures of sulfatases are highly conserved (6–9). Probably, the most striking feature of sulfatases is that they undergo a unique co- or post-translational modification that produces a C*α*-formylglycine (FGly)[2] residue in their active site (10, 11). This residue originates from the conversion of a serine (in prokaryotes) or a cysteine (in prokaryotes and eukaryotes), thus defining two classes of sulfatases, the "Ser-type" and the "Cys-type" sulfatases (10, 11).

These unique modifications are mediated by one of two different enzymes, formylglycine-generating enzyme (FGE) (12, 13) and AtsB (14), responsible for the conversion of cysteine or serine to FGly, respectively.

Both systems are highly divergent but recognize the same consensus motif "(C/S)*X*P*X*R" regarded as the "sulfatase signature" (15–17). This consensus sequence is conserved across all known members of the sulfatase family and has been described as being essential for the conversion of cysteine or serine to FGly and to the proper conformation of the active site of sulfatases (15–17).

In a search for enzymes from human microbiota that are able to hydrolyze nutritionally relevant sulfated compounds, namely glucosinolates (*β*-thioglucoside-*N*-hydroxysulfates) (18), we investigated the biochemical activities of *Clostridium perfringens*. This led us to identify the first sulfatase from Firmicutes. This sulfatase, a Cys-type, is the first one described that lacks the canonical sulfatase signature sequence. Furthermore, while searching for enzymes responsible for its maturation *in vivo*, we could not identify any FGE-type enzyme, the only Cys-type maturating enzyme identified to date. Instead, we discovered the *C. perfringens* genome encodes a protein related to the Ser-type sulfatase-maturating enzyme, AtsB. We demonstrate that this enzyme activates the newly discovered sulfatase, making it the first enzyme described as being able to convert cysteine to FGly in anaerobic bacteria.

EXPERIMENTAL PROCEDURES

Chemicals—p-Nitrophenyl sulfate was purchased from Sigma. Enzymes, oligonucleotides, and culture media were pur-

* This work was supported by a Ph.D. Grant from Région Ile-de-France (to A. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed. E-mail: Olivier.Berteau@jouy.inra.fr.

[2] The abbreviations used are: FGly, C*α*-formylglycine; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; FGE, formylglycine-generating enzyme; DNPH, 2,4-dinitrophenylhydrazone acid matrix; CHCA, *α*-cyano-4-hydroxycinnamic acid matrix; PNP-S, *p*-nitrophenyl sulfate.

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

chased from VWR Scientific. Other chemicals and reagents were obtained from commercial sources and were of analytical grade.

Bacterial Strains, Plasmids, and DNA Manipulations—The *C. perfringens* strain used in this study was the ATCC 13124 strain. *Escherichia coli* DH5α was used for routine DNA manipulations. *E. coli* BL21 (DE3) (Stratagene) was used for *C. perfringens* sulfatase and maturation enzyme overexpression. The pET-28(a) plasmid (Novagen) used to clone the *cpe0231* gene was engineered to bear the ampicillin resistance gene. The *cpe0635* gene was cloned in the pRSF plasmid (Novagen). T4 DNA ligase was from Promega. The plasmid DNA purification kit and QIAprep spin were from Qiagen. DNA fragments were extracted from agarose gel and purified with Wizard SV gel and PCR clean up system kit (Promega). DNA sequencing was performed by VWR Scientific.

Preparation of C. perfringens Protein Extracts—*C. perfringens* ATCC 13124 was grown overnight in an anaerobic chamber (Bactron, Cornelius, OR) in TYH broth (tryptone, yeast extract, and hemin broth) supplemented with or without 50 mM ammonium sulfate. At the end of the culture, the cells were harvested by centrifugation (4,000 × *g* at 4 °C for 20 min) and disrupted by sonication (1 min). After centrifugation (10,000 × *g* at 4 °C for 20 min), the resulting supernatant constituting the protein extract was stored at −20 °C.

Enzyme Assays—Sulfatase activity was assayed at 30 °C for 10 min using 50 mM *p*-nitrophenyl sulfate in 100 mM Tris buffer, pH 7.15. The *p*-nitrophenol released was measured spectrophotometrically at 405 nm (ϵ = 9 000 $mol^{-1}\cdot cm^{-1}$ at pH 7.15).

Cloning and Construction of the pET-6His-CPE0231 Overexpressing Plasmid—*C. perfringens* ATCC 13124 was grown anaerobically in BHI medium, pH 7.0, and the cells were harvested to extract the genomic DNA using the Wizard genomic kit from Promega. The *cpe0231* gene encoding the putative sulfatase was amplified by a PCR-based method using genomic DNA as a template. The following primers were used: 5′-catATGaagccaaatattgtgttaatcatggtt-3′ (NdeI site underlined and ATG codon in uppercase) hybridized to the noncoding strand at the 5′ terminus of the gene and 5′-ctcgagttattatcttatatgttttaaagtgcttac-3′ (XhoI site underlined) hybridized to the coding strand. PCR was run as follows. Genomic DNA (1 μg) in the presence of the primers (0.5 μM each) was mixed with the Hot Start kit (Promega), and 30 cycles of PCR were performed (1 min at 95 °C, 1 min at 50 °C, and 2 min at 72 °C), followed by a final 10-min elongation step at 72 °C. The PCR product was digested with NdeI and XhoI and then ligated with T4 DNA ligase into pET28(a) plasmid digested previously with the same restriction enzymes. The entire sequence of the cloned gene was sequenced to ensure that no errors were introduced during the PCR. The plasmid was named pET-6His-CPE0231.

Cloning and Construction of the pRSF-CPE0635 Overexpressing Plasmid—In a similar manner, the *cpe0635* gene was cloned with the following modifications. The primers used were 5′-catATGccaccattaagtttgcttattaagcca-3′ (NdeI site underlined, ATG codon in uppercase) hybridized to the noncoding strand at the 5′ terminus of the gene and 5′-ctcgagttattatttaatattgttgcaacatttat-3′ (XhoI site underlined) hybridized to the coding strand. The PCR product was digested with NdeI and XhoI and then ligated with T4 DNA ligase into pRSF plasmid digested previously with the same restriction enzymes. The entire sequence of the cloned gene was sequenced to ensure that no errors were introduced during the PCR, and the plasmid was named pRSF-CPE0635.

Protein Expression—*E. coli* BL21(DE3) was transformed with pET-6His-CP0231 and then grown overnight at 37 °C in LB medium (100 ml) supplemented with ampicillin (100 $\mu g\cdot ml^{-1}$). The overnight culture was then used to inoculate fresh LB medium (1 liter) supplemented with the same antibiotic, and bacterial growth proceeded at 37 °C until the A_{600} reached 0.6. The cells were induced by adding 500 μM of isopropyl 1-thio-β-D-galactopyranoside and were collected after 3 h of culture by centrifugation at 4,000 × *g* at 4 °C for 30 min. After resuspension in Tris buffer (0.1 M, pH 7.0), the cells were disrupted by sonication and centrifuged at 220,000 × *g* at 4 °C for 1 h. The solution was then loaded onto a nickel-nitrilotriacetic acid-Sepharose column equilibrated with 0.1 M Tris buffer, pH 7.0. The column was washed extensively with the same buffer. Some of the adsorbed proteins were eluted by a washing step with 30 and 100 mM imidazole, and the overexpressed protein was eluted by applying 500 mM imidazole. Fractions containing the protein were immediately concentrated in Ultrafree cells (Millipore) with a molecular cut-off of 10 kDa.

Protein Co-expression—*E. coli* BL21(DE3) cells transformed previously with the plasmid pET-6His-CP0231 were made chemically competent and then further transformed with the plasmid pRSF-CPE0635. Transformed bacteria were selected on LB medium containing ampicillin and kanamycin, and protein expression was induced with 500 μM isopropyl 1-thio-β-D-galactopyranoside.

MALDI-TOF Analysis—Samples were prepared as follows. The overexpressed protein (50 $pmol\cdot\mu l^{-1}$) was digested overnight with trypsin (20 $ng\cdot\mu l^{-1}$) in ammonium carbonate buffer, pH 8.0, at 37 °C. Then 5 μl of the solution was further hydrolyzed with 5 μl of CNBr (20 $mg\cdot ml^{-1}$ in 0.2 M HCl) in the dark at 45 °C for 4 h.

The α-cyano-4-hydroxycinnamic acid matrix (CHCA) was prepared at 4 $mg\cdot ml^{-1}$ in 0.15% trifluoroacetic acid, 50% acetonitrile. The 2,4-dinitrophenylhydrazone acid matrix (DNPH) was prepared at 1.3 $mg\cdot ml^{-1}$ in 0.5% trifluoroacetic acid, 50% acetonitrile. Equal volumes (1 μl) of matrix and sample were spotted onto the MALDI-TOF target plate. MALDI-TOF analysis was then performed on a Voyager DE STR Instrument (Applied Biosystems, Framingham, CA). Spectra were acquired in the reflector mode with 20-kV accelerating voltage, 62% grid voltage, and a 120-ns delay.

Computational Analysis—The 16 S rDNA sequences from each bacterial strain were aligned with each other using ClustalW (19). The phylogenetic tree was drawn by using the neighbor-joining method (20) with the Kimura two-parameter calculation model.

RESULTS

Sulfatase Activity of C. perfringens—Protein extract from *C. perfringens* was assayed for sulfatase activity using the

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

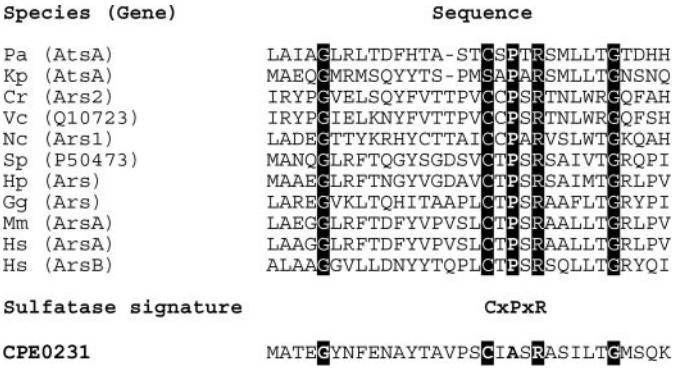

FIGURE 1. **Sequence alignment of several authentic sulfatases representative of bacteria (*Pa*, *Pseudomonas aeruginosa* (36); *Kp*, *K. pneumoniae* (14)), protists (*Cr*, *Chlamydomonas reinhardtii* (38); *Vc*, *Volvox carteri* (39)), fungi (*Nc*, *Neurospora crassa* (40)), lower eukaryotes (*Sp*, *Strongylocentrotus purpuratus* (41); *Hp*, *Hemicentrotus pulcherrimus* (42)), birds (*Gg*, *Gallus gallus* (43)), mammals (*Mm*, *Mus musculus* (44); *Hs*, *Homo sapiens* (45, 46)), and the cloned sulfatase from *C. perfringens* (CPE0231).** Gene names or the protein accession numbers, when no name is available, are given in *parentheses*.

synthetic substrate, *p*-nitrophenyl sulfate (PNP-S). At neutral pH, the extract exhibited a weak sulfatase activity of 0.25 nmol·min^{-1}·mg^{-1}, which almost vanished when we grew *C. perfringens* in presence of 50 mM ammonium sulfate. This result suggested that *C. perfringens* possesses at least one sulfatase enzyme regulated by the sulfate content of the medium.

A search for genes annotated as sulfatases in the *C. perfringens* genome did not show any obvious candidates (15), whereas a blast search with sulfatase sequences led to the identification of one gene, *cpe0231*, annotated as a "probable phosphonate monoester hydrolase" (21, 22). Despite its annotation, the sequence of the putative protein coded by the *cpe0231* gene shared significant identity (from 20 to 24%) with other known bacterial sulfatases. Nevertheless, although all the sequenced sulfatases studied to date share a consensus motif (C/S)*X*P*X*R, the protein encoded by the *cpe0231* gene was predicted to be devoid of the strictly conserved proline (Fig. 1).

Cloning and Sulfatase Activity of the cpe0231 Gene—The *cpe0231* gene was cloned and overexpressed in *E. coli* BL21 (DE3). Following purification, the recombinant protein was analyzed by gel electrophoresis under denaturing conditions. The protein was essentially pure and migrated in accordance with its predicted molecular mass of 58 kDa (data not shown). The pure protein was then investigated for its sulfatase activity, using PNP-S as a substrate. Analysis revealed that it efficiently hydrolyzed PNP-S with a K_m value of 38 mM.

Cations, usually calcium, have been described as promoting the activity of sulfatases (6–9). Therefore we checked if calcium is able to increase the activity of the cloned enzyme. Indeed, the sulfatase activity increased when we added up to 10 mM of $CaCl_2$, leading to a slight K_m decrease from 38 to 30 mM but a doubling of its maximum velocity (V_m = 2.75 μmol·min^{-1}). Thus, in optimal conditions, the specific activity of the purified enzyme toward PNP-S was determined to be 0.08 μmol·min^{-1}·mg^{-1}.

MALDI-TOF Analysis of the Protein Encoded by the cpe0231 Gene—Sequence analysis of the protein encoded by the *cpe0231* gene indicated that Cys-51 (Cys-67 with the addition

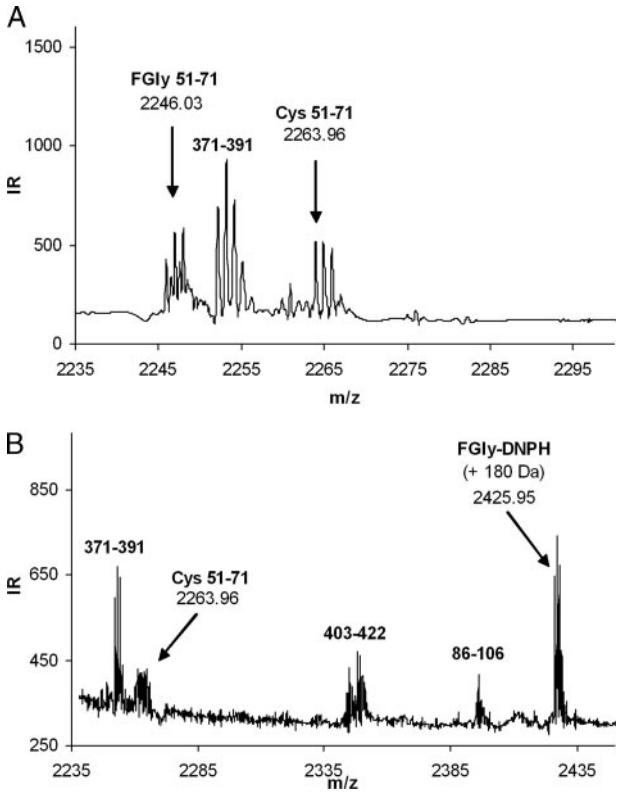

FIGURE 2. **MALDI-TOF analysis of the peptides obtained after CNBr/trypsin digestion of *C. perfringens* sulfatase (CPE0231) using a CHCA matrix (*A*) or a DNPH matrix (*B*).** *Numbers* indicate the amino acid residues. FGly-51–71 points at the peptide bearing the Cα-formylglycine residue; Cys-51–71 points at the unmodified peptide bearing cysteine. FGly-DNPH is the FGly-DNP hydrazone derivative.

of the Tag in the recombinant enzyme) is likely to be the key residue involved in sulfate hydrolysis and thus to undergo the conversion into FGly. To provide evidence of this conversion, we applied MALDI-TOF analysis to the peptides released from a trypsin/CNBr digestion. We clearly obtained two relevant peptides, labeled Cys-51–71 and FGly-51–71 (Fig. 2*A*). The former has a molecular mass of 2263.96 Da and thus corresponds to the peptide ^{51}ATEGYNFENAYTAVPSCIASR^{71} (theoretical molecular mass of 2264.03 Da) containing the critical cysteine Cys-67; the latter has a molecular mass of 2246.03 Da, *i.e.* 18 Da less, as expected from the conversion of the cysteine residue into FGly (Fig. 2*A*). To ascertain the nature of the modification undergone by the peptide 51–71, we performed MALDI-TOF analysis using DNPH as a matrix. Indeed, FGly-containing peptides specifically react with DNPH to form a hydrazone derivative with a mass increment of 180 Da (23). With the DNPH matrix a new peptide with the expected molecular mass of 2425.95 Da appeared (Fig. 2*B*). Nevertheless, in *E. coli*, the overexpressed sulfatase was only partially maturated, as we were able to detect a large amount of the nonmaturated peptide Cys-51–71 (Fig. 2).

Thus, the protein encoded by the *cpe0231* gene, as with all the sulfatases studied to date, is endowed with a conversion of its critical residue into FGly. However, this sulfatase possesses the sequence "FGly*X*A*X*R" at its active site, whereas the other known sulfatases, regardless of their origin (*i.e.* prokaryotic or

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

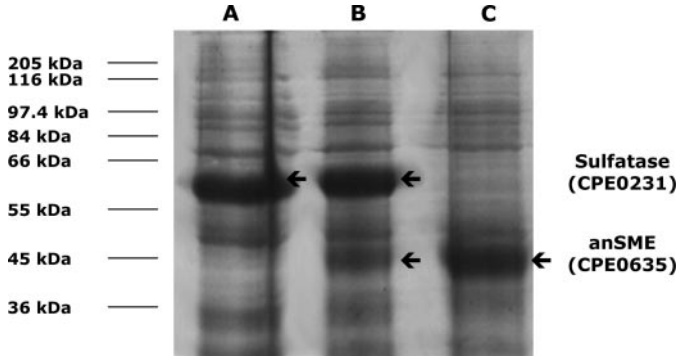

FIGURE 3. **Gel electrophoresis analysis of *C. perfringens* sulfatase (CPE0231) (*lane A*), the sulfatase co-expressed with the protein encoded by the *cpe0635* gene (named anSME) (*lane B*), and anSME (*lane C*) expressed in *E. coli* cells.**

eukaryotic) or their type (Ser-type or Cys-type), bear the sequence FGly*X*P*X*R.

Cloning and Sulfatase Maturation Activity of the cpe0635 Gene—C. perfringens does not encode any FGE-type enzyme, the only Cys-type maturation enzyme identified to date. Nevertheless, an examination of *C. perfringens* genome revealed the existence of the gene *cpe0635*, located almost 500 kbp downstream of the sulfatase gene. This gene codes for a protein that shares homology (48%) with AtsB, the maturation enzyme of Ser-type sulfatases in *Klebsiella pneumoniae* (14). AtsB has been described to be strictly dependent on the presence of an N-terminal signal peptide and on a critical serine residue in the sulfatase active site (24), both features being absent with *C. perfringens* sulfatase.

We cloned CPE0635 into a plasmid allowing co-expression with the *C. perfringens* sulfatase gene in *E. coli* BL21 (DE3). As shown in Fig. 3 (*lane B*), we obtain an efficient co-expression of both proteins. As only the sulfatase bears a His tag, it was easily purified and further analyzed to check if co-expression with the *cpe0635* gene enhanced the sulfatase maturation.

MALDI-TOF analysis revealed that when co-expressed with the *cpe0635* gene, the sulfatase was fully mature because no immature Cys-51–71 peptide could be detected (Fig. 4*A*). Furthermore, analysis with the DNPH matrix confirmed the nature and the efficiency of the maturation, as the signal originating from the peptide "FGly-DNPH" dominates the MALDI-TOF spectrum (Fig. 4*B*). As a consequence of a more efficient maturation, the specific activity of *C. perfringens* sulfatase increased more than 6-fold, rising from 0.08 to 0.53 μmol·min^{-1}·mg^{-1}.

"CXAXR" Sulfatases in Other Bacteria—We searched to see if the *C. perfringens* sulfatase motif C*X*A*X*R exists in other potential microbial sulfatases. For this purpose we performed a blast search, using the following motif (CS)*X*A*X*R, in all the bacterial genomes sequenced thus far. Only sequences that have significant homologies to authentic sulfatases were further considered. We found several genes encoding potential sulfatases lacking the consensus proline at position 3 of the sulfatase signature. These genes were found in 22 bacterial species corresponding to the major bacterial phyla (Fig. 5): Proteobacteria (*Bordetella*, *Bradyrhizobium, Colwellia, Mesorhizobium, Novosphingobium, Pseudoalteromonas, Pseudomonas, Ralstonia, Shewanella, Silicibacter, Sinorhizobium,* and *Vibrio*),

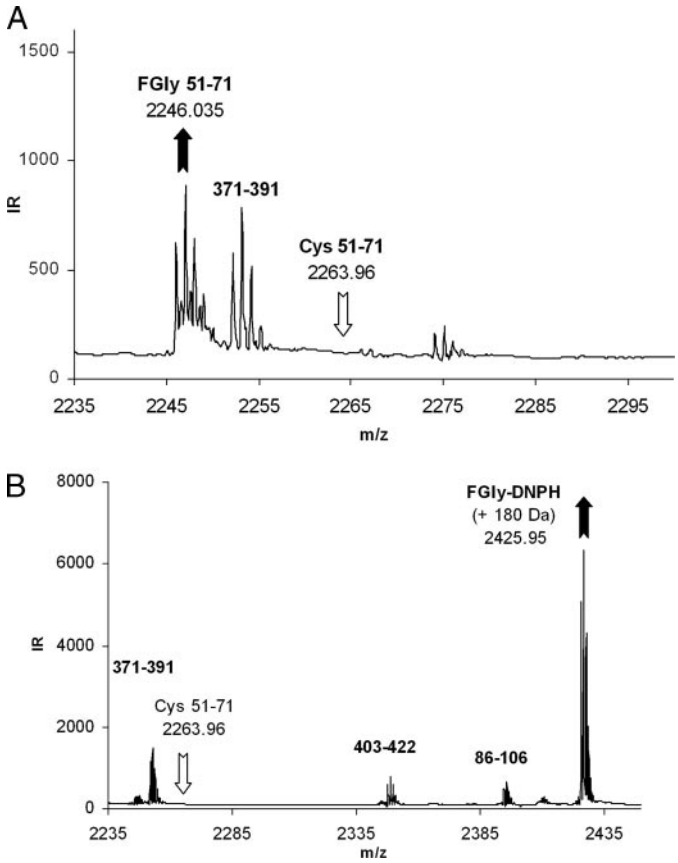

FIGURE 4. **MALDI-TOF analysis of the peptides obtained after CNBr/trypsin digestion of *C. perfringens* sulfatase (CPE0231) co-expressed with the *cpe0635* gene using a CHCA matrix (*A*) or a DNPH matrix (*B*).** *Filled* and *open arrows* indicate increase and decrease of peptide content. *Numbers* indicate the amino acid residues. FGly-51–71 points at the peptide bearing the Cα-formylglycine residue. Cys-51–71 points at the unmodified peptide bearing cysteine. FGly-DNPH is the FGly-DNP hydrazone derivative.

Planctomycetes (*Rhodopirellula*), Cyanobacteria (*Crocosphaera*), Bacteroidetes (*Bacteroides*), and in two new phyla regarded until recently as deprived of sulfatases (17). These phyla are the large phylum of the Firmicutes and the recently described phylum of the Acidobacteria (25). We also identified potential sulfatases with the C*X*A*X*R motif in one archaeal genus, *Methanosarcina*.

Examination of the genomes encoding sulfatases with the C*X*A*X*R motif revealed various situations. Some genomes, as those of *C. perfringens* or *Methanosarcina mazei*, only encode sulfatases with the C*X*A*X*R motif, whereas others, as those of *Rhodopirellula baltica* or *Solibacter usitatus*, encode both types of cysteine sulfatases, *i.e.* with either the "C*X*P*X*R" or the C*X*A*X*R motif. Finally, we also found, although less frequently, genomes encoding potential serine sulfatases with an "S*X*A*X*R" motif, as exemplified by *Vibrio parahemeolyticus*, *Vibrio vulnificus,* and *Bacteroides thetaiotaomicron*.

DISCUSSION

In a search for enzymes from *C. perfringens* able to hydrolyze sulfated compounds, namely glucosinolates (18, 26), we found that a protein extract of *C. perfringens* exhibits a sulfatase activity toward the synthetic substrate PNP-S.

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

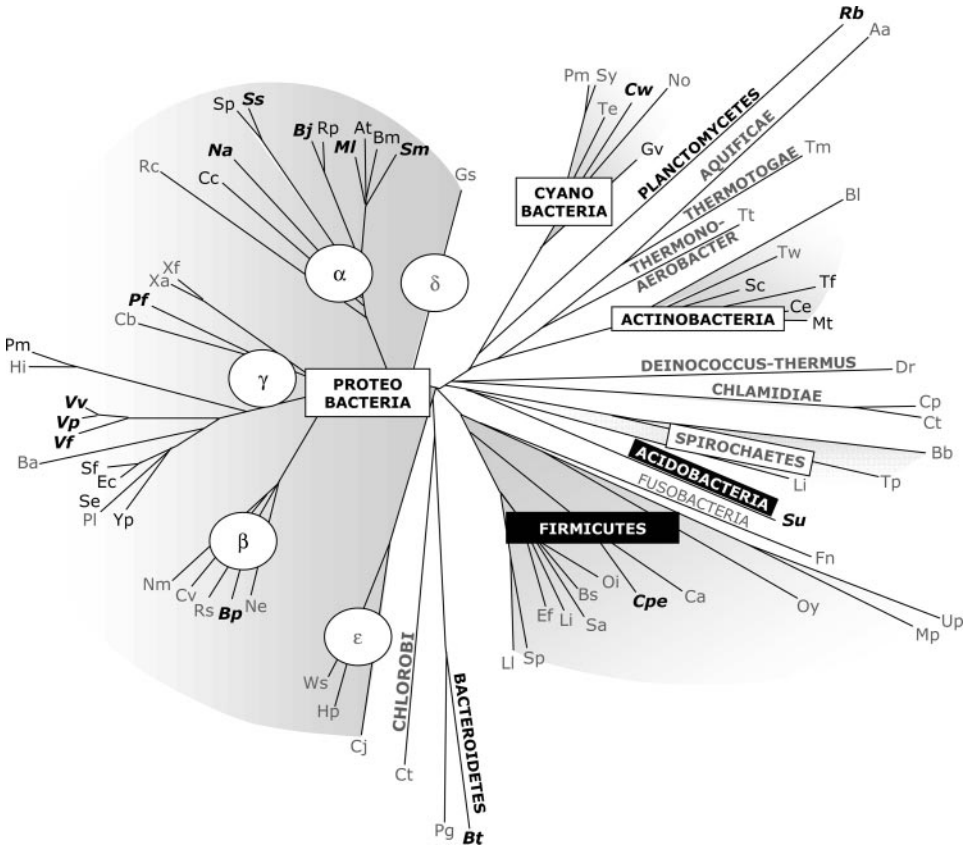

FIGURE 5. **Distribution of potential sulfatases among representative sequenced bacteria.** The phyla described previously as deprived of sulfatase are shown in *black*. Species bearing only sulfatases with the consensus motif (C/S)*XPX*R are in *black letters,* and species bearing sulfatases with the (C/S)*XAX*R motif are in *boldface italic black letters.* The species and phyla deprived of sulfatase are in *gray letters.* Abbreviations: *Aa*, *Aquifex aeolicus*; *At*, *Agrobacterium tumefaciens*; *Ba*, *Buchnera aphidicola*; *Bb*, *Borrelia burgdorferi*; *Bj*, *Bradyrhizobium japonicum*; *Bl*, *Bifidobacterium longum*; *Bm*, *Brucella melitensis*; *Bs*, *Bacillus subtilis*; *Bp*, *Bordetella parapertussis*; *Bt*, *Bacteroides thetaiotaomicron*; *Ca*, *Clostridium acetobutylicum*; *Cb*, *Coxiella burnetii*; *Ce*, *Corynebacterium efficiens*; *Cj*, *Campylobacter jejuni*; *Cc*, *Caulobacter crescentus*; *Cp*, *Chlamydophila pneumoniae*; *Cpe*, *C. perfringens*; *Ct* (Chlamidiae), *Chlamydia trachomatis*; *Ct* (Chlorobi), *Chlorobium tepidum*; *Cv*, *Chromobacterium violaceum*; *Cw*, *Crocosphaera watsonii*; *Dr*, *Deinococcus radiodurans*; *Ec*, *E. coli*; *Ef*, *Enterococcus faecalis*; *Fn*, *Fusobacterium nucleatum*; *Gs*, *Geobacter sulfurreducens*; *Gv*, *Gloeobacter violaceus*; *Hi*, *Haemophilus influenzae*; *Hp*, *Helicobacter pylori*; *Ll*, *Lactococcus lactis*; *Li* (Spirochaetes), *Leptospira interrogans*; *Li* (Firmicutes), *Listeria innocua*; *Ml*, *Mesorhizobium loti*; *Mp*, *Mycoplasma pneumoniae*; *Mt*, *Mycobacterium tuberculosis*; *Na*, *Novosphingobium aromaticivorans*; *Ne*, *Nitrosomonas europaea*; *Nm*, *Neisseria meningitidis*; *No*, *Nostoc* sp.; *Oi*, *Oceanobacillus iheyensis*; *Oy*, *Onion yellows phytoplasma*; *Pf*, *Pseudomonas fluorescens*; *Pg*, *Porphyromonas gingivalis*; *Pl*, *Photorhabdus luminescens*; *Pm* (Proteobacteria), *Pasteurella multocida*; *Pm* (Cyanobacteria), *Prochlorococcus marinus*; *Rc*, *Rickettsia conorii*; *Rb*, *Rhodopirellula baltica*; *Rp*, *Rhodopseudomonas palustris*; *Rs*, *Ralstonia solanacearum*; *Sa*, *Staphylococcus aureus*; *Sc*, *Streptomyces coelicolor*; *Se*, *Salmonella enterica*; *Sf*, *Shigella flexneri*; *Sh*, *Shewanella* sp; *Sm*, *Sinorhizobium meliloti*; *Su*, *Solibacter usitatus*; *Sp* (Firmicutes), *Streptococcus pneumoniae*; *Sp* (Proteobacteria), *Silicibacter pomeroyi*; *Ss*, *Silicibacter* sp; *Sy*, *Synechococcus* sp.; *Te*, *Thermosynechococcus elongates*; *Tf*, *Thermobifida fusca*; *Tm*, *Thermotoga maritima*; *Tp*, *Treponema pallidum*; *Tt*, *Thermoanaerobacter tengcongensis*; *Tw*, *Tropheryma whipplei*; *Up*, *Ureaplasma parvum*; *Vf*, *Vibrio fischeri*; *Vp*, *V. parahemeolyticus*; *Vv*, *Vibrio vulnificus*; *Ws*, *Wolinella succinogenes*; *Xa*, *Xanthomonas axonopodis*; *Xf*, *Xylella fastidiosa*; *Yp*, *Yersinia pestis*.

Analysis of the genome of *C. perfringens* showed no genes annotated as "sulfatase" (22), and we found no protein with the so-called sulfatase signature (C/S)*XPX*R (17, 27, 28). This result was expected as *Clostridia* and the Firmicutes in general are regarded as devoid of sulfatases (17). Nevertheless, we identified the *cpe0231* gene as a potential sulfatase-coding gene.

This gene encodes a protein that possesses a C*XAX*R motif instead of the consensus motif (C/S)*XPX*R. Previous studies have clearly established that this substitution (alanine for proline) prevents sulfatases maturation (15, 16) leading to the current definition of the sulfatase signature (17, 27).

Nevertheless, the purified protein exhibited sulfatase activity and shares many common features with sulfatases described previously. Its activity is increased by calcium, a co-factor found in most of the active sites of sulfatases studied thus far (6–9), and, above all, it undergoes the post-translational modification, unique to sulfatase enzymes (10, 15), namely the conversion of the cysteine residue present in the new motif C*XAX*R into a FGly. The sulfatase from *C. perfringens* is thus an authentic sulfatase bearing the motif FGly*XAX*R in its active site and should be classified in the IUBMB subclass of arylsulfatases (EC 3.1.6.1).

These results rule out the current paradigm defining the sulfatase signature (17, 27) and demonstrate that the definition of the sulfatase signature should be expanded to include the novel motif C*XAX*R. Our results also indicate that the number of genes encoding sulfatases and of bacterial species bearing sulfatase enzymes is currently underestimated. This has biochemical significance because recent findings highlight the involvement of sulfatases in the pathogenic properties of bacteria (29, 30).

Our study also brings new data in the general scheme of prokaryote sulfatase maturation. Until now it was considered that the maturation of sulfatases is dependent on two enzymatic systems, according to the nature of the amino acid residue encoded in the active site. The first, which has attracted considerable attention, is FGE, an oxygen-dependent oxidoreductase found from prokaryotes to eukaryotes (12, 13, 31). This enzyme is the only one described to date able to convert cysteine into FGly (12, 28, 32, 33).

The other system, which accounts for the maturation of Ser-type sulfatases, is called AtsB and probably belongs to the radical *S*-adenosyl-L-methionine enzyme superfamily (34). AtsB has been reported previously to be specific of sulfatases bearing the S*XPX*R motif and to be strictly dependent on the presence of a signal peptide (24). Furthermore, yeast two-hybrid assays have shown that the presence of cysteine instead of a serine in the active site disrupted AtsB/Sulfatase interactions (24, 35) preventing further maturation of the enzyme.

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

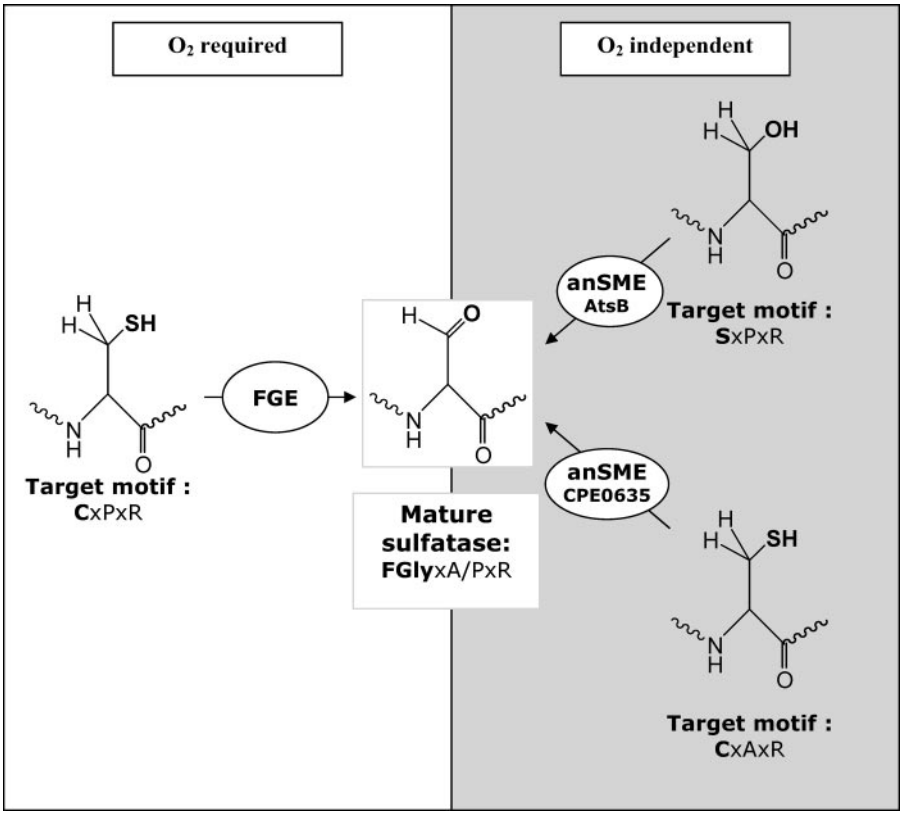

FIGURE 6. **Maturation pathway of sulfatases among prokaryotes.** As shown, the newly discovered system allows the maturation of Cys-type sulfatases in anaerobic conditions.

In *C. perfringens* we were unable to identify any FGE-related protein able to account for the physiological maturation of the cloned Cys-type sulfatase. Surprisingly, we identified a gene encoding a protein showing high homology with AtsB, the enzyme responsible for the maturation of Ser-type sulfatases. AtsB and the enzyme found in *C. perfringens* not only have high sequence identity but also bear the CX_3CX_2C motif, which is a signature for radical *S*-adenosyl-L-methionine enzymes (34).

Co-expression of both genes, encoding the sulfatase and the putative maturation enzyme, led to a fully maturated sulfatase. Thus, *in vivo*, *C. perfringens* sulfatase is not modified by an FGE-type enzyme but by an enzyme closer to the Ser-type sulfatase maturation machinery.

The Cys-type maturating enzyme identified in *C. perfringens* strongly differs from the other Cys-type maturating enzyme (FGE) because it is able to catalyze the conversion of cysteine into FGly under strict anaerobic conditions. Indeed, FGE has been reported recently to be strictly dependent on molecular oxygen (28, 33), which raised the question of how Cys-type sulfatases can be maturated in anaerobic or facultative aerobic bacteria?

The new maturation enzyme identified in *C. perfringens*, a strictly anaerobic microorganism, answers this question, at least in part, and provides evidence that Cys-type sulfatases can be maturated by oxygen-independent oxidoreductases. Supporting this conclusion, a search among sequenced bacterial and archaeal genomes showed that strict anaerobes never bear FGE-related genes but rather AtsB-related genes like many facultative anaerobes such as *E. coli*.

Until now it was unclear why *E. coli*, while being deprived of any FGE-related enzyme, is able to mature Cys-type sulfatases (36, 37). This led to the hypothesis of the existence of a yet unidentified system responsible for the maturation of Cys-type sulfatases among prokaryotes (11, 14).

Our findings, in agreement with the recent bio-computing study of Ballabio and co-workers (17), demonstrate that bacteria such as *C. perfringens*, which possess only enzymes related to AtsB, are able to maturate Cys-type sulfatases. Thus, it is tempting to hypothesize that a similar pathway accounts for the maturation of a Cys-type sulfatase in *E. coli*, which possesses two AtsB-related genes.

We thus propose a new scheme for sulfatase maturation among prokaryotes (Fig. 6). In this new scheme, sulfatase maturation is based on oxygen-dependent (FGE) *versus* oxygen-independent (AtsB, CPE0635) machineries, rather than amino acid specific enzymes as currently stated (17, 27, 32).

We propose that AtsB and CPE0635 belong to the same family that we suggest to name anSME (*i.e.* anaerobic sulfatase maturation enzyme). In our proposed sulfatase maturation pathway, anSMEs are able to modify both types of sulfatases (*i.e.* Cys- and Ser-types), and FGE is strictly dedicated to the maturation of Cys-type sulfatases in aerobic organisms.

Acknowledgment—We thank Jon K. Rubach for critical reading of this manuscript.

REFERENCES

1. Parenti, G., Meroni, G., and Ballabio, A. (1997) *Curr. Opin. Genet. Dev.* **7,** 386–391
2. Dorfman, A., and Matalon, R. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73,** 630–637
3. Gieselmann, V., Matzner, U., Hess, B., Lullmann-Rauch, R., Coenen, R., Hartmann, D., D'Hooge, R., DeDeyn, P., and Nagels, G. (1998) *J. Inherit. Metab. Dis.* **21,** 564–574
4. Franco, B., Meroni, G., Parenti, G., Levilliers, J., Bernard, L., Gebbia, M., Cox, L., Maroteaux, P., Sheffield, L., Rappold, G. A., Andria, G., Petit, C., and Ballabio, A. (1995) *Cell* **81,** 15–25
5. Diez-Roux, G., and Ballabio, A. (2005) *Annu. Rev. Genomics Hum. Genet.* **6,** 355–379
6. Bond, C. S., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J., and Guss, J. M. (1997) *Structure* (*Lond.*) **5,** 277–289
7. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., and Saenger, W. (1998) *Biochemistry* **37,** 3654–3664
8. Hernandez-Guzman, F. G., Higashiyama, T., Pangborn, W., Osawa, Y., and Ghosh, D. (2003) *J. Biol. Chem.* **278,** 22989–22997
9. Boltes, I., Czapinska, H., Kahnert, A., von Bulow, R., Dierks, T., Schmidt,

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010

B., von Figura, K., Kertesz, M. A., and Uson, I. (2001) *Structure* (*Lond.*) **9,** 483–491
10. Schmidt, B., Selmer, T., Ingendoh, A., and Vonfigura, K. (1995) *Cell* **82,** 271–278
11. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M. A., and von Figura, K. (1998) *J. Biol. Chem.* **273,** 25560–25564
12. Cosma, M. P., Pepe, S., Annunziata, I., Newbold, R. F., Grompe, M., Parenti, G., and Ballabio, A. (2003) *Cell* **113,** 445–456
13. Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J. H., Preusser, A., Mariappan, M., and von Figura, K. (2003) *Cell* **113,** 435–444
14. Szameit, C., Miech, C., Balleininger, M., Schmidt, B., von Figura, K., and Dierks, T. (1999) *J. Biol. Chem.* **274,** 15375–15381
15. Dierks, T., Schmidt, B., and VonFigura, K. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11963–11968
16. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) *EMBO J.* **18,** 2084–2091
17. Sardiello, M., Annunziata, I., Roma, G., and Ballabio, A. (2005) *Hum. Mol. Genet.* **14,** 3203–3217
18. Fahey, J. W., Zalcmann, A. T., and Talalay, P. (2001) *Phytochemistry* **56,** 5–51
19. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
20. Saitou, N., and Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425
21. Dotson, S. B., Smith, C. E., Ling, C. S., Barry, G. F., and Kishore, G. M. (1996) *J. Biol. Chem.* **271,** 25754–25761
22. Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S., and Hayashi, H. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 996–1001
23. Peng, J., Schmidt, B., von Figura, K., and Dierks, T. (2003) *J. Mass Spectrom.* **38,** 80–86
24. Marquordt, C., Fang, Q., Will, E., Peng, J., von Figura, K., and Dierks, T. (2003) *J. Biol. Chem.* **278,** 2212–2218
25. Quaiser, A., Ochsenreiter, T., Lanz, C., Schuster, S. C., Treusch, A. H., Eck, J., and Schleper, C. (2003) *Mol. Microbiol.* **50,** 563–575
26. Leoni, O., Iori, R., Haddoum, T., Marlier, M., Wathetet, J. P., Rollin, P., and Palmieri, S. (1998) *Ind. Crops Prod.* **7,** 335–343
27. Hanson, S. R., Best, M. D., and Wong, C. H. (2004) *Angew. Chem. Int. Ed. Engl.* **43,** 5736–5763
28. Roeser, D., Preusser-Kunze, A., Schmidt, B., Gasow, K., Wittmann, J. G., Dierks, T., von Figura, K., and Rudolph, M. G. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 81–86
29. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem. Biol.* **9,** 767–776
30. Hoffman, J. A., Badger, J. L., Zhang, Y., Huang, S. H., and Kim, K. S. (2000) *Infect. Immun.* **68,** 5062–5067
31. Landgrebe, J., Dierks, T., Schmidt, B., and von Figura, K. (2003) *Gene* (*Amst.*) **316,** 47–56
32. Preusser-Kunze, A., Mariappan, M., Schmidt, B., Gande, S. L., Mutenda, K., Wenzel, D., von Figura, K., and Dierks, T. (2005) *J. Biol. Chem.* **280,** 14900–14910
33. Dierks, T., Dickmanns, A., Preusser-Kunze, A., Schmidt, B., Mariappan, M., von Figura, K., Ficner, R., and Rudolph, M. G. (2005) *Cell* **121,** 541–552
34. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., and Miller, N. E. (2001) *Nucleic Acids Res.* **29,** 1097–1106
35. Fang, Q., Peng, J., and Dierks, T. (2004) *J. Biol. Chem.* **279,** 14570–14578
36. Beil, S., Kehrli, H., James, P., Staudenmann, W., Cook, A. M., Leisinger, T., and Kertesz, M. A. (1995) *Eur. J. Biochem.* **229,** 385–394
37. Myette, J. R., Shriver, Z., Claycamp, C., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2003) *J. Biol. Chem.* **278,** 12157–12166
38. Ravina, C. G., Chang, C.-I., Tsakraklides, G. P., McDermott, J. P., Vega, J. M., Leustek, T., Gotor, C., and Davies, J. P. (2002) *Plant Physiol.* **130,** 2076–2084
39. Hallmann, A., and Sumper, M. (1994) *Eur. J. Biochem.* **221,** 143–150
40. Paietta, J. V. (1989) *Mol. Cell. Biol.* **9,** 3630–3637
41. Yang, Q., Angerer, L. M., and Angerer, R. C. (1989) *Dev. Biol.* **135,** 53–65
42. Morokuma, J., Akasaka, K., Mitsunaga-Nakatsubo, K., and Shimada, H. (1997) *Dev. Growth Differ.* **39,** 469–476
43. Ferrante, P., Messali, S., Ballabio, A., and Meroni, G. (2004) *Gene* (*Amst.*) **336,** 155–161
44. Kreysing, J., Polten, A., Hess, B., von Figura, K., Menz, K., Steiner, F., and Gieselmann, V. (1994) *Genomics* **19,** 249–256
45. Collins, J. E., Wright, C. L., Edwards, C. A., Davis, M. P., Grinham, J. A., Cole, C. G., Goward, M. E., Aguado, B., Mallya, M., Mokrab, Y., Huckle, E. J., Beare, D. M., and Dunham, I. (2004) *Genome Biol.* **5,** 84
46. Modaressi, S., Rupp, K., von Figura, K., and Peters, C. (1993) *Biol. Chem. Hoppe-Seyler* **374,** 327–335

Downloaded from www.jbc.org at INRA Institut National de la Recherche Agronomique, on November 8, 2010