THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 4, Issue of January 24, pp. 2212–2218, 2003
Printed in U.S.A.
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.

Posttranslational Modification of Serine to Formylglycine in Bacterial Sulfatases

RECOGNITION OF THE MODIFICATION MOTIF BY THE IRON-SULFUR PROTEIN AtsB*

Received for publication, September 16, 2002, and in revised form, October 28, 2002
Published, JBC Papers in Press, November 4, 2002, DOI 10.1074/jbc.M209435200

Claudia Marquordt, Qinghua Fang, Elke Will‡, Jianhe Peng, Kurt von Figura, and Thomas Dierks§

From the Institut für Biochemie und Molekulare Zellbiologie, Abt. Biochemie II, Universität Göttingen, Heinrich-Düker-Weg 12, 37073 Göttingen, Germany and ‡Abteilung für Molekulare Genetik, Max-Planck-Institut für Biophysikalische Chemie, 37070 Göttingen, Germany

Cα-formylglycine is the catalytic residue of sulfatases. Formylglycine is generated by posttranslational modification of a cysteine (pro- and eukaryotes) or serine (prokaryotes) located in a conserved (C/S)*X*P*X*R motif. The modifying enzymes are unknown. AtsB, an iron-sulfur protein, is strictly required for modification of Ser[72] in the periplasmic sulfatase AtsA of *Klebsiella pneumoniae*. Here we show (i) that AtsB is a cytosolic protein acting on newly synthesized serine-type sulfatases, (ii) that AtsB-mediated FGly formation is dependent on AtsA's signal peptide, and (iii) that the cytosolic cysteine-type sulfatase of *Pseudomonas aeruginosa* can be converted into a substrate of AtsB if the cysteine is substituted by serine and a signal peptide is added. Thus, formylglycine formation in serine-type sulfatases depends both on AtsB and on the presence of a signal peptide, and AtsB can act on sulfatases of other species. AtsB physically interacts with AtsA in a Ser[72]-dependent manner, as shown in yeast two-hybrid and GST pull-down experiments. This strongly suggests that AtsB is the serine-modifying enzyme and that AtsB relies on a cytosolic function of the sulfatase's signal peptide.

Almost all sulfatases that have been described are members of an evolutionary conserved protein family showing extensive homology among enzymes of prokaryotic, lower eukaryotic, and mammalian origin (1–3). The three-dimensional fold and, in particular, active site region of human and bacterial sulfatases are strikingly similar (4–6). The catalytic residue is a Cα-formylglycine (FGly).[1] Its formyl group is hydrated, leading to two geminal hydroxyls at the α-carbon that both are required for catalysis (6, 7). During sulfate ester cleavage, one of the hydroxyls undergoes covalent sulfation with consecutive desulfation induced by the second hydroxyl (6–8). The importance of this novel catalytic mechanism is reflected by the fact that failure to generate the FGly residue leads to synthesis of catalytically inactive sulfatase polypeptides, as is observed in multiple sulfatase deficiency, a rare but fatal human lysosomal storage disorder (9–11).

In eukaryotes, FGly is generated in the endoplasmic reticulum by oxidation of a conserved cysteine residue (12, 13). This late cotranslational or early posttranslational protein modification is directed by a short linear sequence motif comprising a proline and an arginine as the key residues in +2- and +4-positions (C*X*P*X*R) and, in addition, an adjacent auxiliary element (LTG; +8 to +10) (2). Replacing the key cysteine by serine or any other amino acid abolishes FGly formation completely (2, 7, 14). The motif has to be accessible to the modifying machinery prior to folding of the nascent polypeptide into its native structure (2, 12). To date, none of the components or cofactors of this machinery have been identified in eukaryotes. These components are comprised among the luminal contents of the endoplasmic reticulum, which *in vitro* mediate FGly modification independent of protein translocation and independent of a signal peptide in the sulfatase substrate (15).

Due to the clear conservation of the FGly modification motif, most of the sulfatases encoded in various eubacterial genomes are predicted also to undergo FGly modification by oxidation of a cysteine. This was shown experimentally for the arylsulfatase of *Pseudomonas aeruginosa* (PAS), a member of the cysteine-type sulfatases. Even after strong overexpression in *Escherichia coli*, this cytosolic sulfatase was quantitatively converted to the active FGly-bearing enzyme (16). Hence, the *E. coli* cytosol contains the modifying machinery. This machinery is expressed even under excessive supply with inorganic sulfate. Thus, expression of *E. coli*'s cysteine-modifying system is independent of the sulfur status of the cells, in contrast to expression of the sulfatase structural genes, as studied in *P. aeruginosa* and in *Klebsiella pneumoniae* (17–19).

The other well characterized bacterial sulfatase, the arylsulfatase AtsA *of K. pneumoniae*, is a serine-type sulfatase that carries an FGly residue generated by oxidation of a serine rather than a cysteine (20). Generation of FGly (*i.e.* serine semialdehyde) from serine most likely is a one-step oxidation process. In contrast to the cytosolic cysteine-type sulfatases, serine-type sulfatases are located in the periplasm (18, 21, 22). The key FGly motif (S*X*P*X*R) and also the auxiliary downstream element (LTG) are also conserved in serine-type sulfatases (2).

Despite these similarities, bacteria have two different pathways for FGly generation from cysteine and serine, respectively. This is indicated by two observations. First, substitution of the cysteine to be modified in PAS by serine totally blocks FGly formation (16). Second, expression of active, FGly-con-

* This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. 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.

§ To whom correspondence should be addressed. Tel.: 49-551-3919706; Fax: 49-551-395979; E-mail: tdierks@gwdg.de.

[1] The abbreviations used are: FGly, Cα-formylglycine; GST, glutathione *S*-transferase; PAS, *P. aeruginosa* sulfatase; P2/P2*, tryptic peptide 2 of AtsA containing Ser[72]/FGly[72]; UTR, untranslated region; NTA, nitrilotriacetic acid; HA, hemagglutinin; HPLC, high pressure liquid chromatography; PBS, phosphate-buffered saline; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside.

This paper is available on line at http://www.jbc.org

taining AtsA in *E. coli* essentially requires coexpression of the *K. pneumoniae atsB* gene (21), whereas the genomic background of *E. coli* is sufficient for expression of active and modified PAS (16).

The *atsB* gene is located on the same operon as the structural *atsA* gene. AtsB acts in *trans* on AtsA, since it was fully functional when both genes were co-expressed from two different plasmids (21). Despite the presence of two serine-type sulfatase operons in *E. coli*, each consisting of a sulfatase gene and an *atsB* homolog, this species has not been found to express endogenous sulfatases. The inability of the chromosomal *atsB* homologs of *E. coli* to substitute for the *Klebsiella atsB* most likely is explained by repression of its sulfatase operons.

AtsB is predicted to be a 44-kDa iron-sulfur protein with three cysteine clusters that are conserved in all AtsB homologs (22). Iron-sulfur proteins are involved in redox reactions, but only recently a direct enzymatic oxidoreductase function has been assigned to this class of proteins (23–25). In this study, we addressed the following questions. Where in the cell does AtsB fulfill its function? Can it act on both cytosolic and secretory sulfatases? Does AtsB act specifically on the *Klebsiella* sulfatase or on serine-type sulfatases in general? And finally, does AtsB physically interact with the modification motif on the sulfatase polypeptide?

EXPERIMENTAL PROCEDURES

Site-directed Mutagenesis, Tagging, and 5′-UTR Optimization—*atsA*-S72C was generated by PCR mutagenesis of *atsA* (21) using a coding mutagenic primer (CGAGCAGGGAATGCGCATGAGCCAGTATTACACCTCGCCGATGTGCGCCCCGGC) that allowed subcloning of the PCR product via its *Bsm*I site. The generation of *pas*-C51S was described earlier (16). The 5′-UTR of the *pas* gene in pME4055 (16) was substituted by the short 5′-UTR of *atsA* (21) by adding an oligonucleotide (CCCAAGCTTGAACAGGAGAGTCAGTCGTG), with a *Hin*dIII site and an initiator GTG, 5′ of *pas*. The modified gene was cloned into pBluescriptII SK as a *Hin*dIII fragment downstream of the *lac* promotor.

The addition of an oligonucleotide coding for a C-terminal Arg-Gly-Ser-His_6 tag to *atsA* was described earlier (21). The same tag was fused to AtsB by adding the oligonucleotide GGGGGATATCATGCGCTAGTGATGGTGATGGTGATGCGATCCTCT 5′ of the *atsB* stop codon and subcloning of the obtained PCR product as a *Bsm*I/*Eco*RV fragment back into the *atsB* template vector.

Deletion/Addition of Signal Peptides—Deletion of AtsA's signal peptide was achieved by adding an oligonucleotide (CCCAAGCTTGAACAGGAGAGTCAGTCGTG) with a *Hin*dIII site and the indicated initiator GTG, 5′ of codon 21 of *atsA*, which encodes the first amino acid of mature AtsA. The PCR product was subcloned as a *Hin*dIII/*Xho*I fragment back into the *atsBA* template vector. For construction of PAS-C51S+SP, the AtsA signal peptide was added to its N terminus. Using overlapping extension PCR (26), we amplified the 5′-UTR and signal peptide codons of *atsA* (internal noncoding primer: GGGCGTTTGCTCGCGGCGTGCGCGCCACC) and, in a second PCR, the coding region of *pas* excluding the initiator ATG (internal coding primer: CACGCCGCGAGCAAACGCCCCAACTTCCTG). The two PCR products were fused by using them as templates in a third PCR reaction, due to hybridization of the overlapping complementary sequences introduced by the two internal primers. From the final PCR product, a 144-bp *Eco*RV fragment was subcloned into pBluescriptII SK-PAS-C51S, thereby replacing the corresponding part in the 5′ region of *pas*.

Protein Expression in E. coli, Subcellular Fractionation, and Protein Purification—*E. coli* DH5α was transformed with the following plasmids: pBluescriptII containing either *atsA*, *atsB*, *atsBA* (*atsA* and *atsB* with or without His tag codons, *atsA* with or without signal peptide codons), *pas*, or *pas*-C51S (with or without signal peptide codons). Coexpression of AtsB with PAS constructs was achieved in double transformants containing also the pBBR1MCS-*atsB* plasmid (*atsB* subcloned from *atsBA* as a *Kpn*I/*Hin*dIII fragment). The presence of the two plasmids was maintained in selective medium, containing ampicillin and chloramphenicol, and was routinely checked by PCR analysis. Growth conditions, preparation of periplasm, and purification of hexahistidine-tagged proteins were described earlier (21).

Generation of Antibodies—Using purified native PAS protein, provided by Dr. M. Kertesz (School of Biological Sciences, University of Manchester, UK) as antigen and specol as adjuvant, polyclonal antibodies were raised in rabbits injected with 400 μg (first injection) or 200 μg (two booster injections) of antigen. Anti-AtsB antibodies were generated similarly, using AtsB-His_6 protein as antigen. This was purified from inclusion bodies on Ni^{2+}-NTA-agarose (Qiagen) in the presence of 8 M urea according to the protocol of the manufacturer. Prior to rabbit injection, urea was removed by stepwise dialysis (4 M/2 M/1 M urea in PBS). The purity of AtsB was at least 95% (Fig. 1*B*). Antibodies were purified by preadsorption of antisera to immobilized *E. coli* protein. Anti-AtsB antibodies furthermore were affinity-purified by adsorption to SDS-PAGE-purified antigen that was blotted to and excised from a nitrocellulose membrane. Bound antibodies were eluted with 200 mM glycine (pH 2.8).

Protein and Peptide Analysis—Expressed AtsB, sulfatase, or fusion proteins were detected by Western blotting using anti-AtsA (21), anti-AtsB, anti-PAS (see above), anti-GST (Amersham Biosciences), anti-hexahistidine (Qiagen), or anti-HA 12C5A (Roche Molecular Biochemicals) as primary antibodies. ECL signals of corresponding secondary antibodies were detected by a LAS1000+ imaging system (Raytest) and quantitated by densitometry of digital images using Aida 3.10 software (Raytest). For SDS-PAGE, see Ref. 27. The activity of expressed sulfatases was determined in duplicate assays at saturating substrate concentration, as described earlier for AtsA (20) and PAS (16, 17).

The presence of FGly in AtsA was analyzed at the level of its tryptic peptides (see Refs. 20 and 21). During reversed-phase HPLC, P2 and P2* (see "Results"), as detected by mass spectrometry, were recovered in adjacent fractions. The amounts of these peptides were quantitated by sequencing on a Procise cLC protein sequencer (Applied Biosystems). The presence of FGly was verified by mass spectrometry on a matrix-assisted laser desorption ionization-time of flight Reflex III instrument (Bruker Daltonics), using a 337-nm nitrogen laser, with a 200-ns extraction delay. Spectra were obtained as averages of 100 laser shots. 10 mg/ml α-cyano-4-hydroxy-cinnamic acid (Bruker Daltonics) in 50% acetonitrile, 0.1% trifluoroacetic acid served as matrix. Samples were prepared by the drying droplet method, drying 0.5 μl each of sample and matrix solution on a stainless steel target. For FGly identification, 0.5 μl of 2,4-dinitrophenyl hydrazine (Fluka), saturated in 50% acetonitrile, 0.1% trifluoroacetic acid, was added to the dried sample/matrix mixture on the target.

Yeast Two-hybrid Experiments—Fragments of the *Klebsiella atsA* gene were cloned as PCR products into the pAS2 ("bait") vector in frame with the DNA sequence encoding the HA-tagged Gal4-binding domain (28). For PCR amplification of *atsA* fragments, plasmids encoding full-length *atsA* or its signal peptide-deleted version (see above) were used as templates. The forward primer (CCTGAAGGCCATGGAGGCCACAGGAGAGTCAGTCGTG) introduced the underlined *Sfi*I site 5′ of the *atsA* fragment, and the reverse primers (CGGGATCCGGAAGAACGATAGCCGTGGTGG or CGGGATCCTAGCGGTCGGTCAGCCGCAG) introduced the underlined *Bam*HI site 3′ of the wild type stop codon or of a stop codon inserted 3′ of codon 112 by the primer, respectively. The PCR products were cloned as *Sfi*I/*Bam*HI fragments into the pAS2 vector, yielding pAS2-AtsA-(1–112), pAS2-AtsA-(21–112), and pAS2-AtsA-(21–577) (see "Results"). The full-length *atsB* gene was cloned into the pACTII ("prey") vector in frame with the HA-tagged Gal4 activation domain (28) using a 3′ *Bam*HI site present in the multicloning sites of both originating (pBluescriptII KS) and receiving vector. At the 5′-end, a pBluescriptII KS *Kpn*I site, blunted with T4 polymerase, was ligated with a pACTII *Nco*I site, blunted with Klenow polymerase.

The yeast reporter strain Y190 was transformed with both one bait and one prey plasmid. As negative controls, transformations with empty pAS2 or pACTII vectors in combination with a hybrid construct were performed. Cotransformants were selected due to their tryptophan and leucine prototrophy conveyed by the two plasmids (see Ref. 28). Expression of the correct fusion proteins by these cotransformants was routinely controlled by Western blot analysis of cell lysates using anti-HA or anti-AtsB antibodies (not shown). The β-galactosidase activity induced in the case of reconstitution of the Gal4p transcription factor was detected by applying a filter assay with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) as a substrate (29); the permeabilized cells were incubated with X-gal for 6–8 h at 30 °C. The β-galactosidase activity was quantified in a fluid phase assay of cell lysates using *o*-nitrophenyl galactoside as a substrate (30) and calculated according to Ref. 31.

GST Pull-down Experiments—Codons 21–112 of *atsA* or *atsA-S72C* (see above) were amplified by PCR using primers that add a 5′ *Eco*RI site (GGAATTCTAACAGGAGAGTCAGTCGTG) and a 3′ *Sac*I site (AAGCTTGAGCTCTAGCGGTCGGTCAGCCGCAG). The PCR products were cloned as *Eco*RI/*Sac*I fragments into the pGEX-KG vector (32)

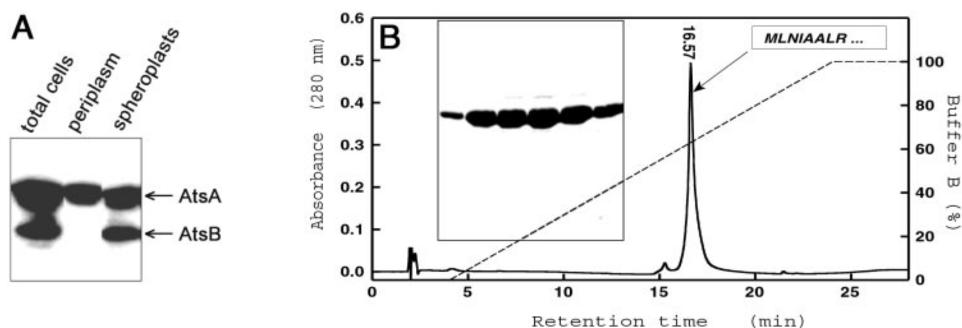


FIG. 1. **Subcellular localization of AtsA and AtsB.** *A*, *E. coli* cells expressing AtsA and AtsB were subjected to osmotic shock, and after sedimentation of spheroplasts, the periplasmic proteins were recovered from the supernatant. Equivalent amounts of periplasm and spheroplasts were analyzed by Western blotting and compared with a total cell lysate. AtsA and AtsB were simultaneously detected by the corresponding affinity-purified antibodies. *B*, hexahistidine-tagged AtsB was purified from *E. coli* coexpressing AtsA and AtsB-His$_6$ by Ni^{2+}-NTA-agarose chromatography (see Coomassie-stained gel of six consecutive eluate fractions, 0.5 column volumes each) and by HPLC on a C4 reversed-phase column (see chromatogram). The N-terminal sequence of the purified AtsB protein, eluting from the C4 column at 16.57 min retention time, is given. It was found to correspond to the N terminus encoded in the *atsB* gene.

in frame with its glutathione *S*-transferase encoding sequence. The fusion proteins, or GST only, were overexpressed in induced (0.2 mM isopropyl thiogalactoside), logarithmically growing *E. coli* DH5α. Bacteria were disrupted in a French press cell and treated with 5 M urea in PBS (pH 7.4) for 30 min at room temperature. The soluble material (75,000 × *g* supernatant) was subjected to dialysis against PBS to remove the urea. After dialysis and another centrifugation (75,000 × *g*), the supernatants or, as a control, PBS buffer were loaded on glutathione-agarose (incubation for 30 min at room temperature), which then was washed with 3 × 4 column volumes of PBS. The columns were then loaded at room temperature with the soluble fraction of an *E. coli* French press lysate (in PBS) containing expressed AtsB protein, and the flow-through was immediately collected without further incubation. After another three washing steps (as above), the columns were eluted twice with 1.5 column volumes of 20 mM glutathione in PBS (pH 8.0). The wash and eluate fractions as well as a final eluate, obtained by boiling the glutathione-agarose beads in SDS-PAGE sample buffer, were analyzed by Western blotting.

RESULTS

AtsB Is a Cytosolic Protein—The *Klebsiella* arylsulfatase AtsA, when purified from a total *K. pneumoniae* cell lysate, was found to be processed by signal peptidase and to carry FGly^{72} in 60% of sulfatase polypeptides (20). When hexahistidine-tagged AtsA, expressed in *E. coli* together with AtsB, was purified from the periplasm of the cells, FGly modification was observed for 48 ± 2% of polypeptides, whereas in the absence of AtsB no FGly was detected (21). To find out whether AtsB, like AtsA, has a periplasmic localization, as suggested earlier (18), we performed a subcellular fractionation of *E. coli* cells expressing both AtsA and AtsB. After osmotic shock of the cells AtsB was found exclusively in the spheroplast pellet and not in the supernatant containing the periplasmic proteins, among them AtsA (Fig. 1*A*). The appearance of AtsA also in the spheroplasts is attributed to incomplete disruption of the outer membrane. After two-step purification of AtsB, expressed in hexahistidine-tagged form, by Ni^{2+}-NTA-agarose chromatography and reversed-phase HPLC (Fig. 1*B*), the AtsB-His$_6$ protein was subjected to amino acid sequencing and found to have an intact N terminus (MLNIAALR). This excludes processing by the signal peptidase. In conclusion, AtsB is a cytosolic protein.

AtsB Coexpression Does Not Lead to Activation of PAS-C51S—The arylsulfatase PAS of *Pseudomonas aeruginosa* is a cysteine-type sulfatase that is quantitatively modified upon expression of its structural gene in *E. coli* (16). When the critical cysteine residue 51 was substituted by a serine (PAS-C51S), no FGly formation was observed (16). Since the sequence motifs (S*X*P*X*R and LTG; see Introduction) determining AtsB-dependent FGly formation in AtsA are contained also in PAS-C51S, the C51S form of PAS was expressed with or without AtsB. As shown in Fig. 2*A*, PAS-C51S was expressed as an

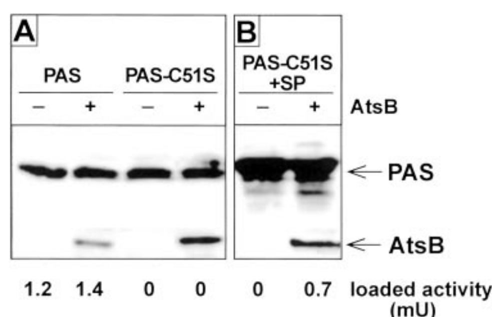


FIG. 2. **Expression of catalytically active PAS-C51S requires AtsB and the presence of a signal peptide.** The wild type and C51S forms of PAS were expressed in *E. coli* in the absence and presence of AtsB, as indicated. The Western blot shows the PAS and AtsB polypeptides, as recovered from the soluble fraction of total cell lysates. The sulfatase activities present in the samples loaded for SDS-PAGE are given below each lane. PAS and PAS-C51S were expressed as cytosolic (*A*) or as secretory proteins (*B*) (*i.e.* without or with the signal peptide of AtsA, respectively, engineered at the N terminus of PAS (PAS-C51S+SP)). PAS-C51S+SP was catalytically active when coexpressed with AtsB. Its specific activity, as calculated after densitometric quantification of PAS protein on the Western blot (see "Experimental Procedures"), corresponded to about 10% of wild type PAS activity.

inactive polypeptide in the absence and presence of AtsB. Control experiments demonstrated that expression of AtsB does not affect the expression of active wild type PAS (Fig. 2*A*). Therefore, we conclude that PAS, when converted to a serine-type sulfatase, is not a substrate for the AtsB-dependent FGly-generating machinery.

AtsB-dependent FGly Formation in AtsA Is Strongly Reduced after Removal of AtsA's Signal Peptide—One of the major differences between AtsA and PAS-C51S is the presence or absence of a signal peptide, respectively. Therefore, we deleted the signal peptide of AtsA and investigated whether a cytosolic version of AtsA (AtsAΔSP) is synthesized in active form when coexpressed with AtsB. It turned out that AtsAΔSP showed a very low, albeit significant, catalytic activity of 0.84 ± 0.14 units/mg (n = 5) (*i.e.* about 1% of wild type AtsA activity) (Fig. 3*A*). In the absence of AtsB, AtsAΔSP was expressed as a completely inactive protein. To examine for the presence of FGly, hexahistidine-tagged versions of wild type AtsA and AtsAΔSP were coexpressed with AtsB. The sulfatases were purified on Ni^{2+}-NTA-agarose and analyzed for FGly modification. For this purpose, tryptic peptides were generated and subjected to HPLC on a reversed-phase column, which allowed us to separate unmodified and modified peptide 2 (P2 and P2*) comprising serine or FGly at position 72, respectively. P2 and P2* eluted in adjacent fractions (20). By amino acid sequencing

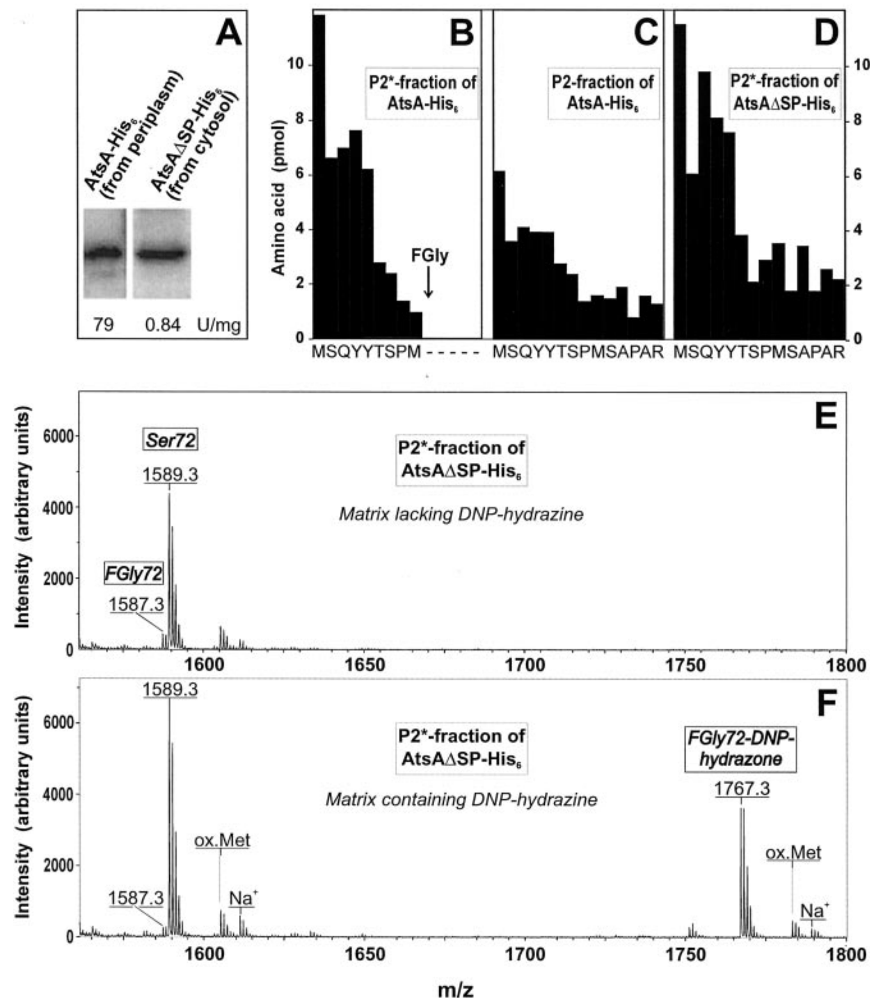


FIG. 3. **The signal peptide is required for efficient FGly modification of AtsA.** AtsB and AtsA-His$_6$ were coexpressed in *E. coli*, the latter as periplasmic wild type, or as cytosolic (*i.e.* signal peptide-deleted (AtsAΔSP-His$_6$)) proteins. The AtsA proteins, recovered from the periplasm (AtsA-His$_6$) or from spheroplasts (AtsAΔSP-His$_6$), were purified on Ni^{2+}-NTA-agarose, and their specific catalytic activities (units/mg of AtsA protein) were determined. *A*, the Western blot shows 0.1 and 0.2 μg of periplasmic and cytosolic AtsA-His$_6$, respectively, having a specific activity of 79 and 0.84 ± 0.14 units/mg (n = 5), respectively. *B–F*, the purified proteins were subjected to tryptic digestion, and their tryptic peptides were separated by reversed-phase HPLC. The modified and the unmodified forms of peptide 2 (P2* and P2, respectively) eluted in adjacent fractions (see Ref. 20), as detected by mass spectrometry. *B–C*, sequencing of the P2* and P2 fraction demonstrated that 56.2 ± 7.5% of the periplasmic wild type AtsA contained FGly. For P2*, Edman degradation is blocked at the position of the FGly (*B*, cycle 10) and reduced in the preceding cycle (20, 21). For FGly quantitation, only the amino acid yields in cycles 3–8 (QYYTSP) were considered, since the first two cycles showed some background. *D*, for the cytosolic AtsAΔSP-His$_6$, the HPLC fraction corresponding to P2* mainly contained unmodified P2, as evidenced by sequencing of the entire peptide showing a serine in position 72 and signals in the following cycles. *E*, mass spectrometry detected clearly low signals for P2* (1587 Da) and high signals for P2 (1589 Da). *F*, using 2,4-dinitrophenyl (*DNP*) hydrazine as a matrix for matrix-assisted laser desorption ionization mass spectrometry, P2*, but not P2, was converted into the corresponding hydrazone (1767 Da). Parts of P2, P2*, and the P2*-hydrazone derivative contained an oxidized methionine or were desorbed as Na^+ adducts, as indicated.

of these fractions (Fig. 3, *B* and *C*), we found that 56% of the wild type AtsA polypeptides carried the FGly.

In the cytosolic AtsAΔSP, traces of FGly could be detected. By mass spectrometry of the HPLC fraction that should contain P2*, we clearly could detect the FGly-containing peptide. As shown in Fig. 3*E*, some P2* (1587 Da) was detected that specifically reacted with 2,4-dinitrophenyl hydrazine, used as a matrix for matrix-assisted laser desorption ionization mass spectrometry, to the corresponding hydrazone (1767 Da; Fig. 3*F*). This hydrazone formation requires the presence of a formyl group (12). The majority of peptide in the analyzed HPLC fraction, however, corresponded to the Ser^{72}-containing P2 (1589 Da; Fig. 3*E*), indicating incomplete separation of P2* and P2 by HPLC. This was confirmed by amino acid sequencing. A serine residue was identified at the position of the expected FGly (Fig. 3*D*). Moreover, no clear decrease of sequencing efficiency at the position of the FGly and consecutive residues, an obligatory effect of FGly-containing peptides (Fig. 3, compare *B* and *D*) (9, 16, 20), was observed. As estimated from the yield of amino acids in the sequencing cycles before and after the position of the FGly in the P2* fraction (Fig. 3*D*) and, in comparison, in the P2-fraction (not shown), P2* contributes less than 10% to the P2/P2* mixture in the P2*-fraction and less than 2% to total P2/P2* of AtsAΔSP. In conclusion, expression of a signal peptide-deleted AtsA results in a sulfatase with a 100-fold lower specific enzymatic activity. This agrees with a similarly reduced FGly content.

AtsB-dependent Expression of Active Signal Peptide-containing PAS-C51S—To find out whether the presence of a signal peptide would allow for AtsB-dependent FGly formation in PAS-C51S, PAS-C51S+SP was constructed by fusing PAS-C51S to the signal peptide of AtsA. When expressed in the absence of AtsB, PAS-C51S+SP was catalytically inactive. However, when coexpressed with AtsB a significant sulfatase activity was measured (Fig. 2*B*), which corresponded to up to 10% of that of wild type PAS (Fig. 2*A*). It should be noted that

TABLE I
Analysis of β-galactosidase expression in various yeast two-hybrid cotransformants

Yeast two-hybrid interaction assays were performed using the indicated bait and prey hybrids (see "Experimental Procedures"). AtsB was expressed as full-length protein, whereas AtsA was lacking its signal peptide (amino acid residues 1–20) and/or its C-terminal residues 113–577. In addition, a S72C mutant of AtsA was investigated. To perform X-gal filter assays, the indicated cotransformants were grown on selection plates, transferred to a polyvinylidene difluoride membrane, permeabilized in liquid nitrogen, and then incubated on a filter soaked with X-gal. The development of no (−), weak (+), or strong (+++) blue color due to X-gal cleavage is given. Quantitation of β-galactosidase activity was performed in a fluid phase assay using *o*-nitrophenyl galactoside as a substrate (see "Experimental Procedures"). β-Galactosidase activity is expressed in units/A_{600} of cells employed for this assay.

Bait hybrid	Prey hybrid	Development of blue color in X-gal filter assay	β-Galactosidase activity
			units/$A_{600} \times 10^{-5}$
1. Gal4BD	Gal4AD-AtsB	−	2.4
2. Gal4BD-AtsA-(21–577)	Gal4AD-AtsB	−	2.7
3. Gal4BD-AtsA-(1–112)	Gal4AD-AtsB	−	1.9
4. Gal4BD-AtsA-(21–112)	Gal4AD-AtsB	+++	28.6
5. Gal4BD-AtsA-(21–112)-S72C	Gal4AD-AtsB	+	3.2
6. Gal4BD-AtsA-(21–112)	Gal4AD	−	9.9
7. Gal4BD-AtsA-(21–112)-S72C	Gal4AD	+	11.2

the signal peptide of most of the expressed PAS-C51S+SP remained unprocessed, as evidenced by a slightly lower electrophoretic mobility (Fig. 2*B*). In fact, less than 5% of PAS-C51S protein and less than 10% of arylsulfatase activity were recovered in the periplasm of cells coexpressing PAS-C51S+SP and AtsB (data not shown). Independent of this obvious translocation deficiency, the catalytic activity of PAS-C51S+SP in contrast to the inactivity of PAS-C51S supports the observation made for AtsAΔSP, namely that AtsB-dependent FGly formation in serine-type sulfatases requires the presence of a signal peptide. Furthermore, AtsB activity is not restricted to the sulfatase of its operon and can even act on a non-*Klebsiella* sulfatase that carries a serine-type FGly modification motif and a signal peptide.

AtsB Physically Interacts with the Serine-type FGly Modification Motif—As shown above, the signal peptide of a newly synthesized sulfatase polypeptide, determining its export to the periplasm, is important for AtsB-dependent conversion of serine to the active site FGly. Hence, the question arises as to whether the cytosolic AtsB makes contact to the sulfatase polypeptide and, if so, whether it interacts directly with the FGly motif or, more indirectly, via the signal peptide.

To address this question, we performed a yeast two-hybrid interaction assay using as "bait" the *N*-terminal 112 amino acids of AtsA with and without signal peptide (AtsA-(1–112) and AtsA-(21–112), respectively) or mature full-length AtsA (AtsA-(21–577)) and as "prey" full-length AtsB. Cell extracts of cotransformants were analyzed by Western blots showing correct expression of the AtsA and AtsB fused to HA-tagged Gal4-binding and Gal4 activation domain, respectively (not shown). The fusion proteins were detected by anti-HA antibodies and the Gal4AD-AtsB also by specific anti-AtsB antibodies. The cotransformants expressed the fusion proteins at similar levels. Only transformants coexpressing Gal4BD-AtsA-(21–112) and Gal4AD-AtsB showed induction of β-galactosidase expression, as determined both by a filter and a fluid phase assay for β-galactosidase activity (Table I, cotransformants 1–4). Thus, in the two-hybrid system, AtsA-AtsB interaction is observed only if the signal peptide and large parts of AtsA, located C-terminal of the modification motif, are deleted. The AtsA-(21–112)/AtsB interaction depended on AtsB, as the β-galactosidase activity was about 3-fold lower in the absence of AtsB (Table I, compare cotransformants 4 and 6).

To test for the dependence of the AtsA-(21–112)/AtsB interaction on the ^{72}SAPAR FGly modification motif, a ^{72}CAPAR (S72C) mutant was analyzed. A weak β-galactosidase activity was induced in AtsA-(21–112)-S72C/AtsB cotransformants (10% of that observed for AtsA-(21–112)/AtsB) (Table I, cotransformants 4 and 5). This residual β-galactosidase activity was not dependent on coexpression of AtsB (Table I, compare cotransformants 5 and 7). We noted that in the absence of AtsB in the prey, the expression of the Gal4BD-AtsA-(21–112) and Gal4BD-AtsA-(21–112)–S72C fusion proteins was considerably increased (not shown). This may contribute to the relatively high residual β-galactosidase activity observed in these cotransformants (Table I, cotransformants 6 and 7). Taken together, these results indicate that interaction of AtsA-(21–112) with AtsB depends on the serine-type FGly modification motif.

To substantiate the two-hybrid data, we performed biochemical *in vitro* interaction experiments using glutathione *S*-transferase (GST) pull-down assays. AtsA-(21–112) and its S72C mutant form were expressed in *E. coli* as C-terminal appendices of GST. The fusion proteins were recovered from inclusion bodies. After solubilization in 5 M urea, dialysis against PBS and centrifugation the soluble material was bound to glutathione-agarose columns. After washing, the soluble fraction of a lysate of *E. coli* expressing AtsB was applied. After further washing steps, the GSH-agarose column was eluted with glutathione. The eluate was analyzed on a Western blot using anti-GST and anti-AtsB antibodies (Fig. 4, *A* and *B*). GST-AtsA fusion proteins were recovered in the glutathione eluate as full-length proteins and, in part, as C-terminally truncated forms (Fig. 4*A*). Significant amounts of AtsB (about 600 ng of AtsB/μg of nontruncated GST-AtsA-(21–112) (*i.e.* 0.54 mol/mol)) were detected in the eluate of the GST-AtsA-(21–112) column, whereas only traces of AtsB were present in the eluate of the GST-AtsA-(21–112)-S72C column (about 0.04 mol of AtsB/mol of GST-AtsA-(21–112)-S72C; see Fig. 4*B*). No AtsB was detected in the eluates of columns loaded with GST or PBS (Fig. 4). In conclusion, AtsB firmly interacts with AtsA-(21–112) in a Ser^{72}-dependent manner.

DISCUSSION

AtsB Interacts Directly with the Serine-type FGly Modification Motif—AtsB is essential for FGly formation in the serine-type sulfatase AtsA from *K. pneumoniae* (21). Here we show that AtsB is a cytosolic protein (Fig. 1). As AtsA is located in the periplasmic space, a localization predicted for all known serine-type sulfatases, an interaction of AtsB and AtsA can only occur in the cytosol (*i.e.* prior to export of AtsA). In fact, serine modification is operative in the cytosol, since FGly formation in a signal peptide-deleted version of AtsA, although extremely inefficient (Fig. 3), still depended on AtsB. AtsB is not strictly specific for AtsA, with which it forms an operon at the genomic level. AtsB was able to activate an artificial serine-type sulfatase that was generated from the cytosolic cysteine-type sulfat-

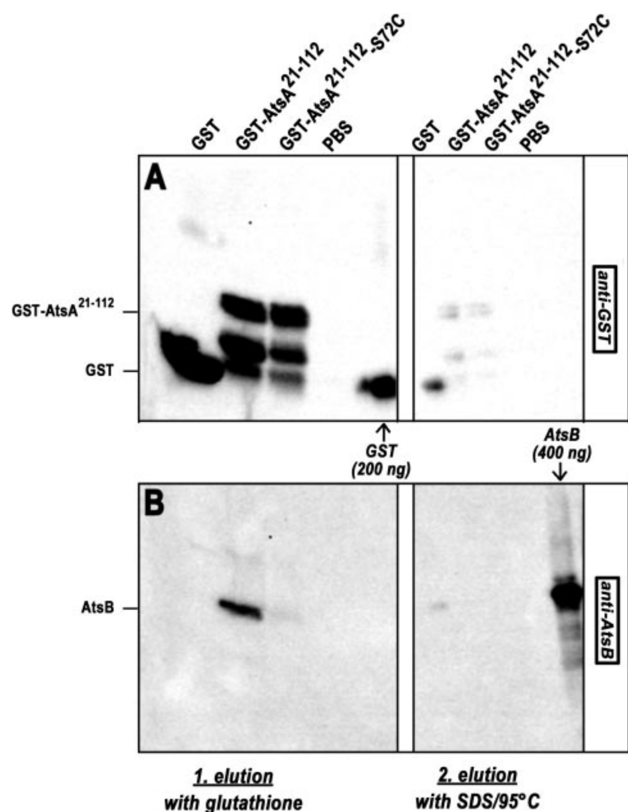
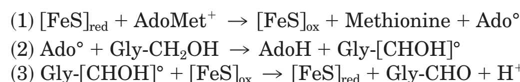


FIG. 4. **Binding of AtsB to immobilized GST-AtsA fusion proteins.** GST, GST-AtsA-(21–112) and GST-AtsA-(21–112)-S72C were solubilized in 5 M urea (see "Experimental Procedures"). After dialysis against PBS and centrifugation, the soluble material or, as a control, PBS buffer were loaded on glutathione-agarose, which then was washed thoroughly. All four columns were then loaded with the soluble fraction of an *E. coli* lysate (in PBS) containing AtsB protein. After another three washing steps, the columns were eluted with glutathione and then boiled with SDS-PAGE sample buffer. The eluates obtained with glutathione and SDS, were analyzed on a Western blot using goat anti-GST (*A*) or rabbit anti-AtsB antibodies (*B*). The positions of GST (27 kDa), GST-AtsA-(21–112) (39 kDa), and AtsB (44 kDa) are indicated. In addition to full-length GST-AtsA-(21–112) fusion proteins, the anti-GST antibody detected C-terminally truncated forms of about 33 and 29 kDa. 200 ng of purified GST (*A*) and 400 ng of purified AtsB (*B*) were used as standards, which allowed us to quantitate the detected Western blot signals. Controls demonstrated that the washings efficiently removed unbound GST fusion proteins and AtsB (not shown). Elution with glutathione was nearly quantitative, since only minute amounts of GST fusion proteins were extracted from the glutathione-agarose beads by SDS.

ase PAS of *P. aeruginosa* (Fig. 2). To convert PAS into a substrate of AtsB, the Cys^{51} residue had to be replaced by serine and an *N*-terminal signal peptide had to be added (see below).

More importantly, AtsB was shown to interact with an N- and C-terminally truncated AtsA polypeptide (AtsA-(21–112)) encompassing the serine-type FGly modification motif. This interaction depended on the presence of the serine to be modified, as shown both by yeast two-hybrid and GST pull-down experiments (Table I, Fig. 4). Hence, we conclude (i) that AtsB directly recognizes the critical Ser^{72} and (ii) that AtsA-(21–112), expressed as a C-terminal appendix of stably folded Gal4BD or GST domains, fulfills all structural requirements for proper association of its linear FGly modification motif with AtsB. This allows us to consider the possibility that AtsB itself is the oxidizing enzyme converting serine to FGly.

In line with this, enzymatic redox functions have been attributed to some AtsB-related FeS proteins. Among them are the anaerobic ribonucleotide reductase from *E. coli* (24) and the HemN protein catalyzing an oxygen-independent oxidation of coproporphyrinogen III in anaerobic heme biosynthesis (25). These and other related proteins generate radical species by reductive cleavage of *S*-adenosylmethionine through an unusual FeS center that is also present in AtsB (23, 33). A possible reaction sequence for AtsB-mediated FGly formation is outlined in Scheme 1. Transfer of an electron from the reduced FeS center to *S*-adenosylmethionine leads to its reductive cleavage (step 1). The generated deoxyadenosyl radical abstracts a hydrogen atom from the substrate (*i.e.* the Ser^{72} side chain, under formation of deoxyadenosine and a substrate radical) (step 2). The single electron of this radical is then accepted by the FeS center, leading to its rereduction, under formation of FGly (step 3).



$$(1)\ [FeS]_{red} + AdoMet^{+} \rightarrow [FeS]_{ox} + Methionine + Ado^{\circ}$$
$$(2)\ Ado^{\circ} + Gly\text{-}CH_2OH \rightarrow AdoH + Gly\text{-}[CHOH]^{\circ}$$
$$(3)\ Gly\text{-}[CHOH]^{\circ} + [FeS]_{ox} \rightarrow [FeS]_{red} + Gly\text{-}CHO + H^{+}$$

$AdoMet^{+}$ = *S*-adenosylmethionine
Ado° = 5′-deoxyadenosyl radical
AdoH = 5′-deoxyadenosine
Gly-CH_2OH = serine
Gly-CHO = FGly

SCHEME 1. **Proposed mechanism for AtsB-mediated FGly formation.**

According to this mechanism, FGly formation from serine is a single enzymatic reaction. Hence, AtsB may suffice for FGly formation in serine-type sulfatases, and additional components may merely serve auxiliary functions (see below). In eukaryotes, serine-type sulfatases are missing. It is therefore not surprising that only very weak homologs of AtsB are encoded in mammalian genomes, the best human homolog being viperin (34), an interferone-inducible antiviral protein that locates to the endoplasmic reticulum, the site where eukaryotic FGly formation (from cysteine) occurs. Viperin, however, shows only 26% identity with the N-terminal third of AtsB, and only one of AtsB's three predicted FeS centers is conserved.

AtsB-dependent FGly Formation Requires a Signal Peptide in the Sulfatase Substrate—In GST pull-down experiments, the interaction of AtsB with the serine-type FGly modification motif was independent of AtsA's signal peptide. Both in yeast two-hybrid (Table I) and in GST pull-down experiments (not shown) interaction was even impaired by the signal peptide. This is in contrast to the *in vivo* situation, where FGly formation in AtsA and in the serine-containing mutant of PAS was clearly dependent on the signal peptide. In the case of AtsA, deletion of the signal peptide reduced FGly formation about 100-fold (Fig. 3), and in the case of PAS-C51S, formation of catalytically active enzyme was totally abolished when a signal peptide was absent (Fig. 2). The apparent contradiction between the *in vitro* and the *in vivo* studies indicates that the presence of a serine-type FGly modification motif is sufficient for binding of AtsB but not for FGly formation. In line with this, we could not detect FGly in the GST-AtsA fusion protein after incubation for 1 h in the presence of AtsB (data not shown). The inhibitory effect of the signal peptide in the yeast two-hybrid experiments can be explained in several ways (*e.g.* by a signal peptide-induced misfolding or intracellular mislocalization of the Gal4BD-AtsA fusion protein).

To be subjected to AtsB-dependent FGly modification, newly synthesized AtsA obviously has to carry a signal peptide in addition to the serine-type FGly modification motif. This may be explained in one of three ways. The hydrophobic signal peptide may exert a direct intramolecular effect on the sulfatase's folding state and modification competence. Alternatively, AtsB binds directly, or indirectly via an adaptor protein, to the signal peptide, and this may be required for the FGly-

forming activity of AtsB. Furthermore, the signal peptide, after binding of an adaptor, may ferry newly synthesized AtsA to a site where AtsB is active. In the latter two cases, a chaperone or targeting factor of *E. coli*'s early secretory pathway such as trigger factor, SecB, or SecA may serve as the putative adaptor protein. A functional *in vitro* FGly-generating assay system reconstituted with cytosolic and/or membrane components of *E. coli* may help to resolve the mechanistic role of the signal peptide.

Acknowledgments—We thank Klaus Neifer for peptide and DNA sequencing and Rita Schmitz-Salue for assistance with the two-hybrid experiments. Furthermore we appreciate the gift of PAS protein from Michael Kertesz (Manchester, UK) used for immunization.

REFERENCES

1. von Figura, K., Schmidt, B., Selmer, T., and Dierks, T. (1998) *Bioessays* **20,** 505–510
2. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) *EMBO J.* **18,** 2084–2091
3. Selmer, T., Hallmann, A., Schmidt, B., Sumper, M., and von Figura, K. (1996) *Eur. J. Biochem.* **238,** 341–345
4. Bond, C. S., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J., and Guss, J. M. (1997) *Structure* **5,** 277–289
5. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., and Saenger, W. (1998) *Biochemistry* **37,** 3654–3664
6. Boltes, I., Czapinski, H., Kahnert, A., von Bülow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M. A., and Usón, I. (2001) *Structure* **9,** 483–491
7. Recksiek, M., Selmer, T., Dierks, T., Schmidt, B., and von Figura, K. (1998) *J. Biol. Chem.* **273,** 6096–6103
8. von Bülow, R., Schmidt, B., Dierks, T., von Figura, K., and Usón, I. (2001) *J. Mol. Biol.* **305,** 269–277
9. Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995) *Cell* **82,** 271–278
10. Hopwood, J. J., and Ballabio, A. (2001) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Valle, D., and Sly, W. S., eds) pp. 3725–3732, McGraw-Hill, Inc., New York
11. Dierks, T. (2001) in *Encyclopedia of Molecular Medicine* (Creighton, T. E., ed) pp. 974–976, John Wiley & Sons, Inc., New York
12. Dierks, T., Schmidt, B., and von Figura, K. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11963–11968
13. Dierks, T., Lecca, M. R., Schmidt, B., and von Figura, K. (1998) *FEBS Lett.* **423,** 61–65
14. Knaust, A., Schmidt, B., Dierks, T., von Bülow, R., and von Figura, K. (1998) *Biochemistry* **37,** 13941–13946
15. Fey, J., Balleininger, M., Borissenko, L. V., Schmidt, B., von Figura, K., and Dierks, T. (2001) *J. Biol. Chem.* **276,** 47021–47028
16. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M. A., and von Figura, K. (1998) *J. Biol. Chem.* **273,** 25560–25564
17. Beil, S., Kehrli, H., James, P., Staudenmann, W., Cook, A. M., Leisinger, T., and Kertesz, M. A. (1995) *Eur. J. Biochem.* **229,** 385–394
18. Murooka, Y., Ishibashi, K., Yasumoto, M., Sasaki, M., Sugino, H., Azakami, H., and Yamashita, M. (1990) *J. Bacteriol.* **172,** 2131–2140
19. Dodgson, K. S., White, G. F., and Fitzgerald, J. W. (1982) *Sulfatases of Microbial Origin*, pp. 133–154, CRC Press, Inc., Boca Raton, FL
20. Miech, C., Dierks, T., Selmer, T., von Figura, K., and Schmidt, B. (1998) *J. Biol. Chem.* **273,** 4835–4837
21. Szameit, C., Miech, C., Balleininger, M., Schmidt, B., von Figura, K., and Dierks, T. (1999) *J. Biol. Chem.* **274,** 15375–15381
22. Schirmer, A., and Kolter, R. (1998) *Chem. Biol.* **5,** R181-R186
23. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., and Miller, N. E. (2001) *Nucleic Acids Res.* **29,** 1097–1106
24. Ollagnier, S., Mulliez, E., Schmidt, P. P., Eliasson, R., Gaillard, J., Deronzier, C., Bergman, T., Graslund, A., Reichard, P., and Fontecave, M. (1997) *J. Biol. Chem.* **272,** 24216–24223
25. Layer, G., Verfurth, K., Mahlitz, E., and Jahn, D. (2002) *J. Biol. Chem.* **277,** 34136–34142
26. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 61–68
27. Dierks, T., Volkmer, J., Schlenstedt, G., Jung, C., Sandholzer, U., Zachmann, K., Schlotterhose, P., Neifer, K., Schmidt, B., and Zimmermann, R. (1996) *EMBO J.* **15,** 6931–6942
28. Tsukada, M., Will, E., and Gallwitz, D. (1999) *Mol. Biol. Cell* **10,** 63–75
29. Dalton, S., and Treisman, R. (1992) *Cell* **68,** 597–612
30. Guarente, L. (1983) *Methods Enzymol.* **101,** 181–191
31. Miller, J. H. (1972) *Experiments in Molecular Genetics*, pp. 352–355, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
32. Guan, K. L., and Dixon, J. E. (1991) *Anal. Biochem.* **192,** 262–267
33. Cheek, J., and Broderick, J. B. (2001) *J. Biol. Inorg. Chem.* **6,** 209–226
34. Chin, K. C., and Cresswell, P. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 15125–15130