Arylsulfatase from *Klebsiella pneumoniae* Carries a Formylglycine Generated from a Serine*

(Received for publication, December 2, 1997, and in revised form, January 6, 1998)

Claudia Miech, Thomas Dierks, Thorsten Selmer, Kurt von Figura, and Bernhard Schmidt‡

*From the Georg-August-Universität, Abt. Biochemie II, Gosslerstraße 12d, D-37073 Göttingen, Germany*

Eukaryotic sulfatases share an unusual posttranslational protein modification, which converts a cysteine into α-formylglycine. The α-formylglycine is essential for the catalytic activity. *Klebsiella pneumoniae* expresses an inducible arylsulfatase for which the DNA predicts a serine at the position occupied by the α-formylglycine residue in eukaryotic sulfatases. Structural analysis showed that the majority of the arylsulfatase polypeptides from *K. pneumoniae* carries the α-formylglycine, whereas the remaining arylsulfatase polypeptides contain the predicted serine residue. This demonstrates the evolutionary conservation between prokaryotes and eukaryotes of this novel protein modification that so far has been found only in sulfatases. α-Formylglycine in *Klebsiella* is generated from a serine and not from a cysteine as in eukaryotes.

Eukaryotic sulfatases share a cysteine residue, which is posttranslationally converted into α-formylglycine (FGly; 2-amino-3-oxopropanoic acid). This novel posttranslational modification occurs in the endoplasmic reticulum and is directed by a linear sequence surrounding the cysteine to be modified (1–3). The FGly residue is critical for the catalytic activity of sulfatases. Crystallographic analysis of two sulfatases has shown that the FGly residue is part of the catalytic site (4, 5). The aldehyde is likely to be hydrated and to serve as an acceptor for sulfate during catalysis (5).2

In prokaryotes expression of sulfatases is generally controlled by the sulfur or the carbon content of their environment. Under appropriate conditions such as the absence of SO4\(^{2-}\) and the presence of alkyl sulfates, the sulfatases are expressed in the periplasmic space (for review see Ref. 7). The sequences of five prokaryotic sulfatases have been reported (8–12), and they share sequence homology with eukaryotic sulfatases to a similar extent as the members of the eukaryotic sulfatase family among each other. Surprisingly, in the genomes encoding the sulfatases from *Klebsiella pneumoniae* (8) and from *Escherichia coli* (9), a serine residue is predicted at a position where all other known sulfatase DNAs predict a cysteine that is known to be converted into FGly in eukaryotes. To examine whether the FGly residue is found also in prokaryotic sulfatases and whether it can be generated also from a serine residue, we purified arylsulfatase from *K. pneumoniae* and examined the protein for the presence of a FGly residue.

**EXPERIMENTAL PROCEDURES**

**Materials**—*K. pneumoniae* strain DSM 681 was obtained from DSM GmbH, Braunschweig. \([^{3}H]NaBH_4\) was from Amersham Buchler.

**Sulfatase Production and Purification**—Bacterial culture and purification of arylsulfatase were performed as described by Okamura et al. (13) with some modifications: *K. pneumoniae* DSM 681 from blood agar plates was grown overnight in aliquots of 0.4 liters of medium containing methionine as sulfur source. At an *A*\(_{595}\) of 1.5–2.5 cells were sedi-mented, washed, and disrupted for 20 min in aliquots of 30 ml using a Sonicator W 220-F, Heat Systems-Ultrasonics, Inc. For analysis of arylsulfatase activity 200 μl of sample were incubated with 200 μl of 20 mM p-nitrocatecholsulfate in 10 mM Tris/HC1, 150 mM NaCl, pH 7.4, for 10 min at 37 °C. After addition of 1 ml of 1 N NaOH, absorbance at 515 nm was measured. The initial ammonium sulfate precipitation (13) was omitted, and Sephadex G-100 was substituted by Superdex 75 (120-ml volume, Pharmacia Biotech Inc.) equilibrated with 20 mM Tris/HC1, 150 mM NaCl, pH 7.2 on a fast protein liquid chromatograph. Arylsulfatase activity eluted after 66 ml. After concentration in an ultra thimble and dialysis against 20 mM Tris/HC1, pH 7.4 arylyglycosidase was loaded on a MonoQ fast protein liquid chromatography column replacing the DEAE-Sephadex A-25 column (13). By this procedure arylsulfatase was purified about 224-fold to a specific activity of 123 units/mg protein yielding about 1 mg of enzyme from 30 liters of culture with a recovery of 7%. For further analysis the sulfatase was desalted by RP-HPLC on a SMART system (Pharmacia) using an Aquapore Butyl 7 micron (220 × 2.1 mm) column (Applied Biosystems) equilibrated with 0.1% trifluoroacetic acid/H\(_2\)O. Arylsulfatase A was eluted by an acetonitrile gradient from 0 to 90% in 36 min.

**Reduction with \([^{3}H]NaBH_4\)**—30 μg of arylsulfatase were lyophilized and solubilized in 4 x guanidinium hydrochloride, 25 mM Tris/HC1, 10 mM EDTA, pH 9. Reduction with \([^{3}H]NaBH_4\), desalting, tryptic digestion, and purification of the peptides on RP-HPLC were performed as described, omitting reducing and with 20 mM l-cysteine as sulfur source. Before digestion with trypsin an aliquot of the \(^{3}H\)-labeled arylsulfatase was analyzed by SDS-PAGE using high Tris gels containing 10% acrylamide, 0.13% bis-acrylamide (14), followed by Coomassie Blue staining and phosphoimaging. Fractions from RP-HPLC were analyzed by liquid scintillation counting, mass spectrometry, amino acid sequencing, and radiosequencing (1, 3).

**Proteolytic In-gel Digestion**—180 μg of arylsulfatase were lyophilized and subjected to SDS-PAGE (see above). After staining with Coomassie Blue (0.195% Coomassie R-250, 0.005% Coomassie G-250, 0.5% acetic acid, 20% methanol) for 20 min and destaining with 30% methanol, the gel slice containing arylsulfatase was excised and cut into small pieces. The gel pieces were washed twice with 0.5 ml of 50% acetonitrile, 50% 100 mM ammonium carbonate (buffer A) and dried on air. 3.6 μg of trypsin in 100 μl of buffer A were added. After 30 min of incubation at room temperature buffer A was added to cover the gel pieces. Following incubation for 16 h at 37 °C, peptides were extracted with 0.2 ml of 50% acetonitrile, 50% trifluoroacetic acid, 0.2 ml of 50% acetonitrile/H\(_2\)O, and 0.2 ml of 75% acetonitrile/H\(_2\)O each at 60 °C for 30 min. The supernatants were pooled, concentrated to 20 μl, filled up to 100 μl with 0.1% trifluoroacetic acid/H\(_2\)O and subjected to RP-HPLC (see above). Fractions containing nonmodified and modified peptide 2, as identified by mass spectrometry (MALDI III, Shimadzu), were pooled, concentrated to 50% volume, and applied to a μPak C2/C18 column (Phar-macia) equilibrated with 0.1% trifluoroacetic acid/H\(_2\)O. The peptides

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

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc.

Printed in U.S.A.
were eluted by an acetonitrile gradient (0.5%/min) at a flow rate of 0.3 m/min. Fractions were analyzed by mass spectrometry using indole-2-carboxylic acid or p-nitroaniline as matrices (1, 3). Fractions containing modified or nonmodified peptide 2, respectively, were pooled and subjected to amino acid sequencing (1).

RESULTS AND DISCUSSION

*K. pneumoniae* was grown in the presence of 1 mM methionine to induce the expression of arylsulfatase (8). The arylsulfatase was purified from the periplasmic fraction to a specific activity of 123 units/mg. The final preparation contained one major polypeptide (Fig. 1, lane 1), which was identified as arylsulfatase by protein sequencing. Among the contaminating proteins, which became visible upon overloaded the gel, the xylene binding protein (15) represented the major species accounting for up to 5% of total protein (data not shown).

In the arylsulfatase from *K. pneumoniae*, serine 72 is equivalent to the cysteine residue that in eukaryotic sulfatases is converted into FGly. To examine whether serine 72 is converted to FGly, the arylsulfatase was incubated with NaB[3H]H₄ in the presence of 4 M guanidinium chloride. This would reduce the aldehyde group of FGly and generate a [3H]serine residue (1, 3). An aliquot of the reduced arylsulfatase was analyzed by SDS-PAGE followed by Coomassie Blue staining and phosphoimaging. Phosphoimaging showed that the arylsulfatase had incorporated [3H]radioactivity with residue 72 strongly suggests that serine 72 in the DNA sequence predicts a serine. The association of the radioactivity at this position corresponds to residue 72, for which the radioactivity (shaded area) is indicated. The absorbance and the radioactivity (shaded area) are shown. Fraction 36 containing the peak of radioactivity associated with tryptic peptide 2 is indicated. B, mass spectrometry of fraction 36. The material was embedded in indole-2-carboxylic acid. C, fraction 36 was subjected to amino acid sequencing. The radioactivity released in each cycle is given. The sequence indicated on the abscissa corresponds to that of tryptic peptide 2 (residues 63–76) of *Klebsiella* arylsulfatase.

When sulfatases are expressed at high level in mammalian cells, a fraction of the cysteine can escape conversion to FGly, presumably due to saturation of the oxidizing machinery (1). Because the *Klebsiella* arylsulfatase had been reduced prior to tryptic digestion, it was not possible to determine whether a fraction of the arylsulfatase had retained serine 72 predicted by the DNA. Therefore the experiment was repeated with the omission of the reduction of the arylsulfatase. In addition the tryptic digestion was performed after separation by SDS-PAGE to remove the remaining contaminants. The tryptic peptides were separated by RP-HPLC. Fractions containing masses predicted for the nonmodified serine 72-containing form of peptide 2 (1500 Da) or for the modified FGly 72-containing form of peptide 2 (1588 Da, designated as peptide 2*) were identified by mass spectrometry and subjected to rechromatography (Fig. 3, upper panel). A [3H]-labeled standard of peptide 2 (obtained as described above) eluted in fractions 31–34 as indicated in Fig. 3 (upper panel) by a horizontal bar. The material in these fractions had a mass of 1589–1590 Da, and sequencing of the pooled fractions 31–34 yielded the full-length sequence of peptide 2 (Fig. 3, lower panel).

The peptide(s) in fractions 26–30 had a mass of 1587–1589 Da. To determine whether this material corresponds to the modified peptide 2*, the mass was additionally determined.
Formylglycine in Arylsulfatase Generated from Serine

CONCLUSIONS

The present data clearly demonstrate that the arylsulfatase from K. pneumoniae contains an FGly residue at position 72, where its DNA predicts a serine residue. About 40% of the arylsulfatase polypeptides, which were synthesized under conditions of strong induction, still contained the predicted serine 72, suggesting that the capacity of the machinery generating the FGly residue is saturable. FGly 72 in the Klebsiella sulfatase occupies a position that is homologous to that of the FGly residue in eukaryotic sulfatases. Thus, the FGly residue, which is critical for sulfate ester cleavage, is conserved between prokaryotes and eukaryotes. The pathways responsible for this unusual protein modification, however, may be different. If in eukaryotic sulfatases the cysteine is replaced by a serine, catalytically inactive sulfatases are formed (6) in which the serine residue is not converted into FGly. Whether the prokaryotic machinery converting serine into FGly can also catalyze the conversion of cysteine to FGly remains to be determined. At least one prokaryotic sulfatase is known in Pseudomonas aeruginosa for which the DNA sequence predicts a cysteine at the position occupied by FGly in sulfatases from eukaryotes and K. pneumoniae (10).

The ability of K. pneumoniae to convert a serine into FGly in arylsulfatase should facilitate genetic approaches to identify components required for this novel protein modification. This in turn may help to explain the genetic defect in multiple sulfatase deficiency. In this human disease catalytically inactive sulfatases are synthesized due to a failure to generate the FGly residues in sulfatases (1).

Acknowledgments—We thank Katja Unthan-Hermeling, Klaus Neifer, and Michael Recksiek for experimental assistance, Dr. H. Eiffert (Institute of Medical Microbiology, University Göttingen) for help in growing Klebsiella, and Dr. U. Pleßmann (Max-Planck-Institute for Biophysical Chemistry Göttingen) for advice in proteolytic in-gel digestion.

REFERENCES

1. Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995) Cell 82, 271–278
2. Dierks, T., Schmidt, B., and von Figura, K. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 11063–11068
3. Selmer, T., Hallmann, A., Schmidt, B., Sumper, M., and von Figura, K. (1996) Eur. J. Biochem. 238, 341–345
4. Bond, C. S., Clemments, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J., and Gass, J. M. (1997) Structure (London) 5, 275–289
5. Lukatel, G., Krauf, N. Theis, K., Selmer, T., Gieselmann, V., von Figura, K., and Saenger, W. (1998) Biochemistry, in press
6. Brooks, D. A., Robertson, D. A., Bindless, C., Litijens, T., Anson, D. S., Peters, C., Morris, C. P., and Hopwood, J. J. (1995) Biochem. J. 307, 457–463
7. Fitzgerald, J. W. (1976) Bacteriol. Rev. 40, 698–721
8. Murooka, Y., Ishibashi, K., Yasumoto, M., Sasaki, M., Sugino, H., Azakami, H., and Yamashita, M. (1990) J. Bacteriol. 172, 2131–2140
9. Daniela, K. L., Plunkett, G. III, Burland, V., and Blattner, F. R. (1992) Science 257, 771–778
10. Beil, S., Kehrli, H., James, P., Stendemann, W., Cook, A. M., Leisinger, T., and Kortes, M. A. (1995) Eur. J. Biochem. 229, 385–394
11. Burland, V., Plunkett, G. III, Daniels, D. L., and Blattner, F. R. (1993) Genomics 16, 551–561
12. Philipp, W. J., Poole, S., Eigelmeier, K., Pascopella L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B. R., Jacobs, W. R., Jr., and Cole, S. T. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 732–737
13. Okamura, H., Yamada, T., Murooka, Y., and Harada, T. (1976) Agr. Biol. Chem. 40, 1071–1076
14. 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
15. Sofia, H. J., Burland, V., Daniels, D. L., and Plunkett, G. III (1994) Nucleic Acids Res. 22, 2576–2586
Arylsulfatase from *Klebsiella pneumoniae* Carries a Formylglycine Generated from a Serine
Claudia Miech, Thomas Dierks, Thorsten Selmer, Kurt von Figura and Bernhard Schmidt

*J. Biol. Chem.* 1998, 273:4835-4837.  
doi: 10.1074/jbc.273.9.4835

Access the most updated version of this article at http://www.jbc.org/content/273/9/4835

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 14 references, 6 of which can be accessed free at http://www.jbc.org/content/273/9/4835.full.html#ref-list-1