Conversion of cysteine to formylglycine in eukaryotic sulfatases occurs by a common mechanism in the endoplasmic reticulum

Thomas Dierks1,*, Maria Rita Lecca2, Bernhard Schmidt, Kurt von Figura

Institut für Biochemie und Molekulare Zellbiologie, Abt. Biochemie II, Universität Göttingen, Gosslerstr. 12d, 37073 Göttingen, Germany

Received 18 December 1997

Abstract Sulfatases undergo an unusual protein modification leading to conversion of a specific cysteine residue into α-formylglycine. This conversion is essential for catalytic activity. In arylsulfatase A the α-formylglycine is generated inside the endoplasmic reticulum at a late stage of protein translocation. Using in vitro translation in the presence of transport-competent microsomes we found that arylsulfatase B is also modified in a similar way by the formylglycine-generating machinery. Modification depended on protein transport and on the correct position of the relevant cysteine. Arylsulfatase A and B did not compete for modification, as became apparent in co-expression experiments. This could argue for an association of the modification machinery with the protein translocation apparatus.

© 1998 Federation of European Biochemical Societies.

Key words: Sulfatase; Endoplasmic reticulum; Protein modification; Protein transport; Multiple sulfatase deficiency

1. Introduction

In eukaryotic sulfatases a α-formylglycine (FGly, 2-amino-3-oxypropanoic acid) is found at a position where the genes encode a cysteine [1,2]. Deficiency of the FGly is associated with catalytic inactivity of the sulfatases, as is found in multiple sulfatase deficiency, a rare inherited human lysosomal storage disorder [1,3,4]. The FGly residue is part of the catalytic site, as has been shown by crystallographic analysis of two lysosomal sulfatases, arylsulfatase A (ASA) and arylsulfatase B (ASB) [5,6]. The aldehyde group of the FGly residue, most likely in its hydrated form [5,7], serves as an acceptor for sulfate during sulfate ester cleavage [5–7]. In acting as a general dap the FGly hydrate allows for efficient ester hydrolysis at the acidic pH of lysosomes. The catalytic mechanism involves trans-esterification of the sulfate group from the substrate to the first hydroxyl, from where it is eliminated due to the presence of the second hydroxyl [5,7].

The FGly residue in ASA is generated by protein modification at a late stage of co-translational protein translocation into the endoplasmic reticulum (ER), as could be shown by in vitro synthesis and translocation of nascent ASA polypeptides into canine pancreas microsomes [8]. The modification is directed by a linear sequence of 16 residues surrounding the cysteine to be modified [8]. This sequence includes an X-P-S-R and an L/M-T-G-R/K tetrapeptide, which are located C-terminal of the cysteine and which are conserved among all human sulfatases [2,9].

Since in multiple sulfatase deficiency the activity of all tested sulfatases was severely decreased, it is likely that the microsomal modifying machinery acts on all newly synthesized sulfatase polypeptides. So far nothing is known about the modifying machinery and its components, which are supposed to catalyze a two-step reaction involving an oxidation of the thiol to a thioaldehyde group and a hydrolytic release of H2S [1]. To establish whether all sulfatases are modified in the ER and to qualify the in vitro translation/translocation system as a biochemical means to characterize the modifying machinery we determined in vitro the modification of ASB and compared it to that of ASA.

2. Materials and methods

2.1. Site-directed mutagenesis

Mutagenesis of a cDNA coding for human ASB [10] was carried out by the QuikChange method (Stratagene) using complementary primers coding sequence GTGTCCTGGAACATGATGACTACCG-CAG that substituted the Asn codon 84 by a methionine codon. Inversion of codon 91 (Cys) and 92 (Thr) was achieved by PCR using a non-coding primer (CTGTAGCCGCGCCAGTGAGCAGCTGG-CTCCGGAGCGGCGTACCGGCGTC) covering the near- by Xcm site. The PCR product was subcloned as a HindIII/XcmI fragment replacing the corresponding fragment of the template DNA pRL2 (see below). No Pfu- or Taq-polymerase errors were detected upon sequencing of the entire coding sequences.

2.2. Protein expression and purification

The cDNAs of wild-type ASB and ASB-N84M, respectively, were cloned as BamHI/EcoRI fragments into the pMPSVHE vector [11] downstream of the myeloproliferative sarcoma virus promoter. The resulting plasmid and PGK-hygro as selection marker were used for stable transfection of mouse embryonic fibroblasts deficient in both mannose 6-phosphate receptors, as described [12]. The expressed ASB protein was purified from the secretions of the cells by affinity chromatography [13]. The specific activity of both wild-type and mutant enzyme was similar (about 100 U/mg). Synthesis of the ASB proteins in an immature form (64 kDa), that due to constitutive secretion from the mannose 6-phosphate receptor-deficient cells circumvent lysosomal processing [14], did not affect the catalytic activity, since mature ASB expressed in BHK cells had the same sulfatase activity. The expressed proteins carried exclusively FGly and no cysteine in position 91, as was verified by mass spectrometry (see below). Expression and purification of ASA-F59M protein was described earlier [8].

1 Corresponding author. Fax: (49) (551) 395979. E-mail: dierks@ukb2-00.uni-bc.gwdg.de
2 The first two authors contributed equally to this work.

Abbreviations: ASA, arylsulfatase A (cerebroside-3-sulfate 3-sulfohydrolase, EC 3.1.6.6); ASB, arylsulfatase B (N-acetylgalactosamine-4-sulfate 4-sulfohydrolase, EC 3.1.6.9); DNP, 2,4-dinitrophenyl; ER, endoplasmic reticulum; FGly, Cα-formylglycine; P3, cysteine-containing form of tryptic peptide 3 of ASB; P5*, FGly-containing form of tryptic peptide 3 of ASB; P3C, C-terminal fragment of P3 generated by endopeptidase AspN; P3*C, C-terminal fragment of P3* generated by endopeptidase AspN; RP-HPLC, reversed phase-high performance liquid chromatography.
2.3. In vitro synthesis of ASB and ASA derivatives

The PCR coding for residues 39–134 of ASB-N84M was amplified by using two different primers (AATGGCGGCTCCGCAAGC- CGGGGGCAAGCAGGCCGCCG). The PCR fragment was cloned into an expression vector (pTD3) for in vitro expression of the resulting plasmid. The resulting mASB (ASA-F59M) was synthesized in vitro in the presence of [35S]methionine and dog pancreatic microsomes. The translation product imported into the primary translation product and becomes converted into FGly upon translocation into microsomes [8]. In order to analyze the presence or absence of FGly-91 in ASB synthesized in vitro in the absence or presence of microsomes, we expressed an N-terminal fragment of ASB-N84M comprising residues 39–134 of ASB-N84M fused to the signal peptide of preprolactin. The exchange of the authentic signal peptide (residues 1–38) for that of preprolactin should ensure efficient translocation into the microsomes. To allow for incorporation of a [35S]methionine-label during translation into the ASB

Fig. 1. In vitro modification of ASB in the endoplasmic reticulum.
A: A fusion protein consisting of the signal peptide of preprolactin and the N-terminal residues 39–134 of mature ASB-N84M was synthesized in vitro in the presence of [35S]methionine and dog pancreatic microsomes. The translation products were analyzed using either purified ASB-N84M or ASA-F59M protein as a carrier (see below). Purification of translation products imported by the microsomes using differential centrifugation and protease K digestion was described earlier [8]. Aliquots (2 μl) were analyzed by SDS-PAGE on high-Tris gels [8] and phosphorimaging (Figs. 1, 3 and 4). The remaining 48 μl (98 μl in the co-expression sample) were used for peptide analysis.

2.4. Peptide analysis

Purified ASB (ASA) in vitro translation/translocation products were mixed with 30–40 μg of unlabeled ASB-N84M (ASA-F59M) carrier protein and subjected to reductive carboxymethylation and generation of peptides by trypsin or endoproteinase AspN, as described [1,8]. Separation of peptides by RP-HPLC, mass spectrometry and sequencing of peptides were also described earlier [1]. The protocols for reaction of peptides with 2,4-dinitrophenylhydrazine and for analysis of the peptide hydrazones are given in ref. [8].

3. Results

3.1. Modification of Cys-91 of ASB in the endoplasmic reticulum

The FGly residue in human ASB is found in position 91 [1]. This position is equivalent to residue 69 of ASA, where a cysteine that is encoded by the ASA gene is incorporated into the primary translation product and becomes converted into FGly upon translocation into microsomes [8]. In order to analyze the presence or absence of FGly-91 in ASB synthesized in vitro in the absence or presence of microsomes, we expressed an N-terminal fragment of ASB-N84M comprising residues 39–134 of ASB-N84M fused to the signal peptide of preprolactin. The exchange of the authentic signal peptide (residues 1–38) for that of preprolactin should ensure efficient translocation into the microsomes. To allow for incorporation of a [35S]methionine-label during translation into the ASB
(39–134) fragment, which lacks methionines, Asn-84 was replaced by methionine. Expression of a full-length ASB-N84M mutant in eukaryotic cells yielded a catalytically active protein (see Section 2) indicating that substitution of Asn-84 by methionine does not interfere with generation of FGly-91. This could also be demonstrated by structural analysis of the ASB-N84M protein (see below).

A cDNA encoding the preprolactin-ASB derivative described above was subjected to coupled in vitro transcription and translation in the presence of [35S]methionine and transport-competent microsomes. The translation product imported into and processed by the microsomes was purified (mASB, Fig. 1A) and mixed with purified ASB-N84M protein serving as carrier. This mixture was subjected to reductive carboxymethylation of cysteines and tryptic digestion. Residue 91 is part of the tryptic peptide 3 of ASB comprising residues 69–95 (Fig. 1A). After separation of the tryptic peptides by RP-HPLC one major 35S-labeled peak was recovered (Fig. 1B). A part of the radioactivity coeluted with the FGly-91 containing peptide 3 (designated P3*) of the ASB-N84M carrier (Fig. 1B). The latter was identified by mass spectrometry (2886 Da) and amino acid sequencing of the respective fractions. The majority of the radioactivity eluted at a 1.0% higher acetonitrile concentration, where the joined peptides 6 plus 7 (which lack methionines), but no peptide 3 of the ASB-N84M carrier protein could be identified by sequencing and mass spectrometry. Radiosequencing of this material revealed, however, that a methionine was present in position 16, as expected for peptide 3 of the in vitro translation product. From experiments with ASA it is known that the peptide 2 carrying the cysteine to be modified elutes at a 1.5% higher acetonitrile concentration than the FGly-containing form of this peptide [1,8]. It was therefore likely that the majority of the radiolabeled peptide of ASB-N84M represented the Cy35S-labeled peak was recovered after HPLC of trypsin- and endoproteinase AspN-generated peptides (see Fig. 1C). In addition, the fraction of [35S]P3*C that was converted into the corresponding DNP-hydrazone (see Fig. 2) is given. Lane 1 shows the results obtained after analysis of the precursor form of the same ASB construct synthesized in the absence of microsomes. The apparent modification represents background radioactivity coeluting with unlabeled P3*C of the carrier protein, since it did not react with DNP-hydrazone (see text). The same holds true for the translation/translocation product shown in lane 3, which was synthesized in the presence of microsomes but carried the relevant cysteine in position 92 due to an inversion of codons 91 (Cys) and 92 (Thr) in the cDNA.

91 containing form of peptide 3 (designated P3), which is absent in wild-type ASB [1] and in the ASB-N84M carrier (data not shown).

Due to the high hydrophobicity of the large peptide 3 RP-HPLC did not lead to adequate separation of P3* and P3 to allow quantification of modification. Therefore the fractions containing the peak of radioactivity were pooled (see Fig. 1B) and digested with endoproteinase AspN. This generates the 13-mer peptide 3C [1] comprising residues 83–95 of ASB-N84M. By RP-HPLC the 35S-labeled peptide 3C could be resolved in two forms (Fig. 1C). About 20% of the radioactivity represented [35S]P3*C, as identified by its coelution with carrier P3*C (mass: 1557 Da) and by radiosequencing (Fig. 1D). About 80% of the radioactivity did not coelute with a carrier peptide, but by radiosequencing could be identified as [35S]P3*C (Fig. 1E).

To examine for the presence of an aldehyde function [35S]P3*C and [35S]P3*C were subjected to reaction with dimethylphenyldrazine (DNP-hydrazone) [8]. Only [35S]P3*C gave rise to hydrazone formation which could be identified after separation from non-reacted [35S]P3*C by RP-HPLC (Figs. 2 and 3, lane 2). The [35S]P3*C-DNP-hydrazone coeluted with the unlabeled P3*C-DNP-hydrazone of the carrier, which had the predicted mass of 1737 Da, i.e. 180 Da more than P3*C. It should be noted that hydrazone formation of most FGly-containing peptides is only partial. Quantitative hydrazone formation was found only for smaller and more hydrophilic peptides [8].

Taken together these data demonstrate that about 20% of the ASB-N84M fragment synthesized in vitro carried a FGly residue, when translation had been coupled to translocation into microsomal membranes. In the absence of microsomes the translation product representing the precursor form of the preprolactin-ASB-N84M construct carried no aldehyde

---

**Fig. 2.** Presence of an aldehyde group in peptide 3C* after in vitro translocation of ASB into microsomes. A: The [35S]methionine-labeled peptide 3C* coeluting with unlabeled P3*C (see Fig. 1C and D) was tested for the presence of an aldehyde group by reaction with DNP-hydrazone. The incubation mixture was subjected to RP-HPLC in order to separate the DNP-hydrazone derivative from the parent peptide and the reagent. The positions of P3*C and P3*C-DNP-hydrazone are indicated, as identified by mass spectrometry. The radioactivity profile (see histogram) shows that [35S]P3*C was also converted into the hydrazone with an efficiency of approximately 40%. B: Reaction of [35S]P3*C (see Fig. 1C and E) did not give rise to DNP-hydrazone formation.

**Fig. 3.** Conversion of cysteine into FGly in ASB depends on import into microsomes and on the correct position of the cysteine. The in vitro translation/translocation product shown in lane 2, representing the ASB construct shown in Fig. 1A, was analyzed for cysteine modification as described in Figs. 1 and 2. The values given for modification represent the percentage of [35S]P3*C of total [35S]P3*C recovered after HPLC of trypsin- and endoproteinase AspN-generated peptides (see Fig. 1C). In addition, the fraction of [35S]P3*C that was converted into the corresponding DNP-hydrazone (see Fig. 2) is given. Lane 1 shows the results obtained after analysis of the precursor form of the same ASB construct synthesized in the absence of microsomes. The apparent modification represents background radioactivity coeluting with unlabeled P3*C of the carrier protein, since it did not react with DNP-hydrazone (see text). The same holds true for the translation/translocation product shown in lane 3, which was synthesized in the presence of microsomes but carried the relevant cysteine in position 92 due to an inversion of codons 91 (Cys) and 92 (Thr) in the cDNA.
giving a positive reaction with DNP-hydrazine was quantitated (see column with the tryptic P2* of the carrier ASA-F59M protein and described in Fig. 3. For ASA the radioactivity coeluting from the RP-gradient. The modification efficiency of ASB was calculated as depicted in the presence of the FGly-containing peptides P3*C (ASB) or P2* (ASA), respectively. The import products purified from the 100-μl coexpression sample were analyzed in duplicate. One half was mixed with ASB-N84M carrier protein, the other half with ASA-F59M carrier protein. Analysis of the [35S]labeled tryptic peptides in a single RP-HPLC run was possible, since the tryptic peptides 2 of ASA and 3 of ASB elute at well separated positions in the acetonitrile gradient. The modification efficiency of ASB was calculated as described in Fig. 3. For ASA the radioactivity coeluting from the RP-column with the tryptic P2* of the carrier ASA-F59M protein and giving a positive reaction with DNP-hydrazine was quantitated (see [8]).

The specificity of the modifying machinery was tested by inverting the position of the cysteine to be modified and the position of its C-terminal neighbor (threonine 92). After translation and import of this ASB-mutant into microsomes peptide analysis revealed that only some background radioactivity (6% of P3C-associated radioactivity) coeluted with P3*C of the carrier protein (Fig. 3, lane 3). This radioactive material did not react with DNP-hydrazine.

3.2. Simultaneous modification of ASA and ASB

The relative efficiencies of in vitro modification of ASB and ASA, quantitated as percentage of ASB-[35S]P3*C or ASA-[35S]P2*, respectively, of total [35S]P3C or [35S]P2, respectively, were very similar when assayed in parallel (Fig. 4, lanes 1 and 2). We wanted to know whether the modification efficiency of ASA and/or ASB is affected when the two sulfatases are translated and translocated simultaneously. When co-expressed in the presence of microsomes, ASA and ASB were translated and translocated with the same efficiency, as compared to the single expression controls (Fig. 4, lane 3; it should be noted that the aliquot subjected to SDS-PAGE in lane 3 is only 50% of that applied in lanes 1 and 2). Analysis of the tryptic peptides of these translocation products, which in the case of ASB had to be digested additionally with endoprotease AspN (see Fig. 1), clearly showed a similar extent of modification in the co-expressed sulfatases as in the singly expressed ASA or ASB. If changed at all, a slight increase of relative modification was observed for both ASA and ASB (Fig. 4). Calculation of total radioactivity recovered in the HPLC fractions associated with the modified peptides of ASA and ASB revealed that similar amounts of molecules had been modified per equivalent of microsomes [16] in all three samples shown in Fig. 4.

4. Discussion

ASB is subjected to FGly formation in the ER by a mechanism that shares all characteristics observed earlier for ASA [8]. FGly formation depended on protein import into the ER and on the presence of a cysteine in the primary translation product located at the correct position within a sequence that is highly conserved among eukaryotic sulfatases and obviously determines this novel protein modification (see Section 1). This supports the notion that most likely all eukaryotic sulfatases are subjected to this modification by a common modifying machinery located in the ER.

This machinery obviously is saturable under in vitro conditions [1,8]. For ASA it was shown in vitro that reducing the expression to about one-twentieth of the level used in the present and in an earlier study [8] doubles the relative modification efficiency from about 20% to about 40% [8]. Using the high expression conditions we observed a similar modification efficiency of about 20% for both ASA and ASB. This efficiency did not drop when ASA and ASB were translocated simultaneously into microsomes. This may indicate that the modifying machinery has a similar affinity for ASA and ASB. However, several other observations are inconsistent with this conclusion. In multiple sulfatase deficiency usually low residual activities of the various sulfatases are detectable. This is attributed to a residual activity of the modifying machinery. Characteristically the residual activity of ASB is the highest among all sulfatases, e.g. 2-4 times higher than that of ASA [17,18]. This suggests that in vivo modification of ASB is more efficient than that of other sulfatases. Accordingly, in recombinant ASB Cys-91 is quantitatively modified to FGly, while in recombinant ASA expressed under similar conditions 10-40% of the Cys-69 escape modification [1]. In contrast to these in vivo data the in vitro data presented in this study would argue for a similar affinity of the modifying machinery for ASB and ASA. The similar and overall rather low modification efficiencies observed in the in vitro system, however, may in fact result from limitation at a certain step of modification that is not limiting in vivo. If the translocation and modifying machineries are coupled at the luminal side of the ER membrane, limited modification may merely reflect an excess of translocation over modification under in vitro conditions. Until now attempts to uncouple translocation and modification, e.g. by studying modification of a sulfatase precursor polypeptide in a microsomal detergent extract, have proven unsuccessful (not shown). Obviously, the translocation and modification machineries have to be reconstituted in order to develop our in vitro system into an
assay that is suitable to biochemically identify the modifying enzyme(s).

Acknowledgements: We thank Meike Pauly-Evers and Christoph Peters (Freiburg) for providing the affinity column for ASB purification. The technical assistance of Petra Schlotterhose is gratefully acknowledged. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References