doi:10.1006/jmbi.2000.4297 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **305**, 269–277

JMB

Crystal Structure of an Enzyme-Substrate Complex Provides Insight into the Interaction between Human Arylsulfatase A and its Substrates During Catalysis

Rixa von Bülow[1,2], Bernhard Schmidt[2], Thomas Dierks[2] Kurt von Figura[2] and Isabel Usón[1]*

[1]*Lehrstuhl für Strukturchemie Institut für Anorganische Chemie, Universität Göttingen Tammannstrasse 4 37077 Göttingen, Germany*

[2]*Institut für Biochemie und Molekulare Zellbiologie, Abt. Biochemie II, Universität Göttingen, Heinrich Düker Weg 12, 37073 Göttingen Germany*

**Corresponding author*

Arylsulfatase A (ASA) belongs to the sulfatase family whose members carry a C^{α}-formylglycine that is post-translationally generated by oxidation of a conserved cysteine or serine residue. The crystal structures of two arylsulfatases, ASA and ASB, and kinetic studies on ASA mutants led to different proposals for the catalytic mechanism in the hydrolysis of sulfate esters.

The structures of two ASA mutants that lack the functional C^{α}-formylglycine residue 69, in complex with a synthetic substrate, have been determined in order to unravel the reaction mechanism. The crystal structure of the inactive mutant C69A-ASA in complex with *p*-nitrocatechol sulfate (pNCS) mimics a reaction intermediate during sulfate ester hydrolysis by the active enzyme, without the covalent bond to the key side-chain FGly69. The structure shows that the side-chains of lysine 123, lysine 302, serine 150, histidine 229, the main-chain of the key residue 69 and the divalent cation in the active center are involved in sulfate binding. It is proposed that histidine 229 protonates the leaving alcoholate after hydrolysis.

C69S-ASA is able to bind covalently to the substrate and hydrolyze it, but is unable to release the resulting sulfate. Nevertheless, the resulting sulfation is low. The structure of C69S-ASA shows the serine side-chain in a single conformation, turned away from the position a substrate occupies in the complex. This suggests that the double conformation observed in the structure of wild-type ASA is more likely to correspond to a formylglycine hydrate than to a twofold disordered aldehyde oxo group, and accounts for the relative inertness of the C69S-ASA mutant. In the C69S-ASA-pNCS complex, the substrate occupies the same position as in the C69A-ASA-pNCS complex, which corresponds to the non-covalently bonded substrate. Based on the structural data, a detailed mechanism for sulfate ester cleavage is proposed, involving an aldehyde hydrate as the functional group.

Keywords: crystal structure; arylsulfatase-A complex; catalytic mechanism; formylglycine; multiple sulfatase deficiency (MSD)

Introduction

Mammalian sulfatase enzymes catalyze the hydrolysis of sulfate ester bonds in the pathways for catabolism of glycosaminoglycans and glycolipids and in the synthesis of steroid hormones.[1,2] The active sites of all studied sulfatases carry a C^{α}-formylglycine (FGly) that is post-translationally generated by oxidation of a conserved cysteine or, in some bacterial sulfatases, of a serine residue.[3–6] The sequence identity between sulfatases, ranging

Abbreviations used: ASA, cerebroside-3-sulfate 3-sulfohydrolase (arylsulfatase A); ASB, *N*-acetylgalactosamine 4-sulfatase (arylsulfatase B); FGly, C^{α}-formylglycine; pNCS, *p*-nitrocatechol sulfate; EDTA, ethylenediamine tetra acetic acid.

E-mail address of the corresponding author: uson@shelx.uni-ac.gwdg.de

0022-2836/01/020269–9 $35.00/0

from 20 to 60 %, with the N-terminal third of the polypeptide chain and residues around the active site being especially conserved, suggests that members of this family have a common origin and share structural fold and active site. Human sulfatases possess, nevertheless, a high substrate specificity and eight distinct genetic disorders have so far been described in humans associated to the malfunction of single sulfatases.[7–10] The lack of the post-translational modification of the cysteine residue to C^{α}-formylglycine (FGly) in the active site of all sulfatases leads to multiple sulfatase deficiency (MSD), a rare autosomal recessive disease.[8,3] Ten human sulfatases have been characterized to date, six in the lysosomes. Sulfatases that are able to hydrolyze *in vitro*, in addition to their natural substrates, sulfate ester bonds in synthetic arylsulfates are called arylsulfatases.

The crystal structures of two human sulfatases have been determined, cerebroside-3-sulfate 3-sulfohydrolase (arylsulfatase A, ASA) at 2.1 Å[11] and *N*-acetylgalactosamine-4-sulfate 4-sulfohydrolase (arylsulfatase B, ASB) at 2.5 Å,[12] both sharing structural similarities in their folds and active sites with alkaline phosphatase.[13] ASA and ASB are lysosomal arylsulfatases, showing optimal catalytic activity at pH values around 5.[14,15] The reaction mechanisms proposed for sulfate ester hydrolysis on the basis of these structural data share the formation of a covalently sulfated enzyme intermediate, but entail some differences involving formation and cleavage of this intermediate and the nature of the active side-chain initiating reaction. An aldehyde hydrate has been proposed in the case of ASA *versus* the unmodified aldehyde for ASB. The mechanism derived from the aldehyde hydrate would be a nucleophilic attack on the sulfate sulfur atom, involving a pentacoordinated sulfur intermediate from which the alcohol would be eliminated. On the other hand, the catalytic mechanism, as proposed for ASB, could start through an electrophilic reaction of the aldehyde C^{β} on a sulfate oxygen atom, followed by the participation of an unspecified nucleophile to perform hydrolysis of the ester bond. Furthermore, the mechanism proposed for ASA would cause an inversion of the sulfur configuration, while that proposed for ASB is likely to proceed with retention. Inversion of the sulfur configuration has been implied in the case of a lower eukaryotic arylsulfatase but there are no data concerning ASA, ASB or closer related sulfatases.[16]

Analysis of the interaction of wild-type ASA with its substrate and of a covalently sulfated ASA intermediate is hampered by the fact that ASA degrades its substrate too fast to allow trapping of actual intermediate complexes, even at low temperature,[17] and that ASA crystals could not be frozen without dramatic loss of resolution. An indirect approach was therefore chosen here by using the two ASA mutants, C69A and C69S, both of which lack the FGly residue essential for catalysis. The C69A mutant is inactive because the substitution of the FGly side-chain by the methyl group of alanine excludes the formation of a covalent bond between the enzyme and the substrate sulfur atom. As the geometry and polarity of the environment of the active-site residue 69 should be preserved, analysis of the C69A-ASA substrate complex should show the location and interactions in the reaction intermediate with the exception of the covalent bonding.

For the C69S mutant it is known that the hydroxyl group of serine 69 carries out a nucleophilic attack onto the substrate sulfur, which results in an enzyme intermediate covalently sulfated at serine 69. The subsequent release of the sulfate from the enzyme is blocked. The efficiency of the enzyme sulfation, however, was low, and at best 20 % of the C69S-ASA was recovered as sulfated product.[18] The sulfated form of C69S was expected to provide insight into the structure of the sulfated intermediate of ASA.

Here, we report on the crystal structures of the ASA mutants C69A and C69S soaked with buffer containing *p*-nitrocatechol sulfate (pNCS), its most commonly used synthetic substrate. The natural substrate, cerebroside 3-sulfate, cannot be degraded in the absence of the activator protein saposin B, as it is otherwise insoluble. While a covalently sulfated intermediate could not be seen in the C69S-ASA crystal, we were able to determine the interactions and non-bonded distances between ASA and pNCS. Based on the structure of the complex between C69A-ASA and pNCS, a detailed reaction mechanism for sulfate ester cleavage is presented.

Results

Crystal structure of C69A-ASA in complex with *p*-nitrocatechol sulfate

Crystals of this mutant were obtained under the same conditions as those of the wild-type, yielding the same space group and unit cell.[11] Keeping a reduced variability in the experimental conditions was desirable to compare both structures, and in addition the pH of 5.3 is relevant to the lysosomal environment, ASA displaying optimal activity at a pH range around 5.[14] The substrate was incorporated by soaking the crystals in mother-liquor solution containing additionally 20 mM pNCS for a period of at least two hours prior to data collection to a resolution of 2.35 Å. The structure was isomorphous to that of the wild-type. Both final models lack the same residues, corresponding to disordered regions (the loop comprising Gly444 to Ala447 and the C-terminal Asp504 to Ala507 besides a few side-chains on the protein surface). After refinement, the water molecules modeled in the wild-type structure were incorporated if they lay in positive difference electron density, and discarded from the model if their *B*-values refined to more than 80 $Å^2$. The residues forming the active site and the water molecules located in it in the

wild-type were eliminated from the model to calculate a sigmaA-weighted omit map,[19] which shows the electron density corresponding to a pNCS molecule, with the sulfate group well defined and the aryl group discretely disordered over two sites, with relative occupancies of 60 and 40 %, respectively. The pNCS group was included in the refinement with the help of restraints derived from high-resolution structural data for this substrate.[20] The electron density at residue 69 agrees well with the presence of a methyl group, and is missing in the sites where the two oxygen atoms of a disordered formylglycine residue or its hydrate would be located, as determined in the wild-type and the structure of P426L-ASA.[21] The coordinates of the remaining residues constituting the active-site pocket are identical to those of the wild-type within the experimental uncertainty.

Figure 1 shows a view of the structure of the protein, highlighting the location of substrate and side-chains in the active center within the whole protein frame. Figure 2(a) shows a stereo view of the difference omit map at a contour level of 1σ, and the final model in the area of the active site. Both the use of omit maps and refinement against a maximum likelihood target were employed to reduce model bias, severe at this resolution with an almost identical starting model.[22] Indeed, with the exception of residues E198, A199 and C493, deviations between C^{α} atoms in both structures are less than 0.5 Å, the average deviation being 0.19 Å, that is, identical within the experimental error margin. Table 1 summarizes the interactions and non-bonded distances between protein atoms and substrate as determined in this structure. The sulfate group in the substrate is held close to the side-chain of Ala69 by the positively charged residues His229, Lys302 and the Mg^{2+} cation. Further interactions to the side-chains of Lys123, Ser150 and the main-chain of Ala69 stabilize this position (Figure 2(b)).

Crystal structure of C69S-ASA and its complex with *p*-nitrocatechol sulfate

The structure of the ASA mutant carrying a serine residue in place of cysteine was determined at 2.4 Å. In the case of eukaryotic sulfatases, serine does not get post-translationally modified to the FGly residue.[23,18] In the wild-type crystal structure, the FGly side-chain was interpreted as being disordered between two discrete sites, with possible participation of its hydrated form, that is, with a geminal diol on the C^{β} carbon atom[11] (Figure 2(c)). One of the hydroxyl groups is coordinating the magnesium cation present in the active site and is directed towards the opening of the active-site cleft, while the other hydroxyl group is establishing hydrogen bonds with the side-chains of His125 and Arg73, and thus turned away from the site a substrate would occupy (Figure 2(c)). It is in this second orientation that the serine side-chain in the C69S mutant is exclusively found in the crystal structure (Figure 3).

Table 1. Polar interactions between ASA and pNCS (distances in Å)

S150 O^{γ}-O2	3.3	A69 C^{β}-O1	3.5
H229 $N^{\varepsilon 2}$-O1	2.9	A69 C^{β}-O2	3.4
H229 $N^{\varepsilon 2}$-O4	2.4	A69 C^{β}-O3	3.3
K302 N^{η}-O4	2.8	A69 N-O3	3.3
W-O2	2.5	K123 N^{η}-O2	3.3
W-O4	3.0	K123 N^{η}-O1	3.2
Mg^{2+}-O1	2.7		

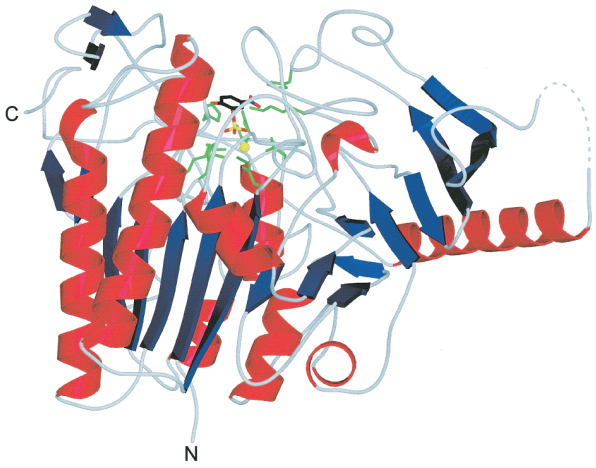

Figure 1. Structure of ASA-C69A mutant in complex with pNCS. The side-chains of the residues relevant for catalysis are depicted together with the major component of the disordered pNCS molecule. The second site of the aryl moiety is omitted for clarity.

Crystals of this mutant were grown under the same conditions as the wild-type, thus minimizing variability, and as the structure was solved by replacement with the wild-type coordinates, differences would be more likely to be downweighted due to model bias than to be artifacts. In fact, the mean C^{α} deviation between the structures of wild-type and C69S-ASA is 0.16 Å.

Crystals of this mutant were also soaked in mother liquor containing 20 mM pNCS for a period of at least two hours before data collection to a resolution of 2.65 Å. The substrate-binding pocket shows residual electron density after modeling the water molecules, but not distinct enough to be clearly interpretable as the whole substrate. Either the substrate lies in the active site too inhomogeneously to be modeled, or substrate diffusion into the active cavity of the crystal was hindered and too slow to reach completeness. The better-defined density corresponds, nevertheless, to the same position that the sulfate group occupies in the structure of C69A-pNCS. It is not possible to resolve the disorder of the aryl group, or to detect a covalently sulfated minor component if present in the crystal. Figure 3 shows a stereo view of the C69S-ASA structure in the area of the active-site

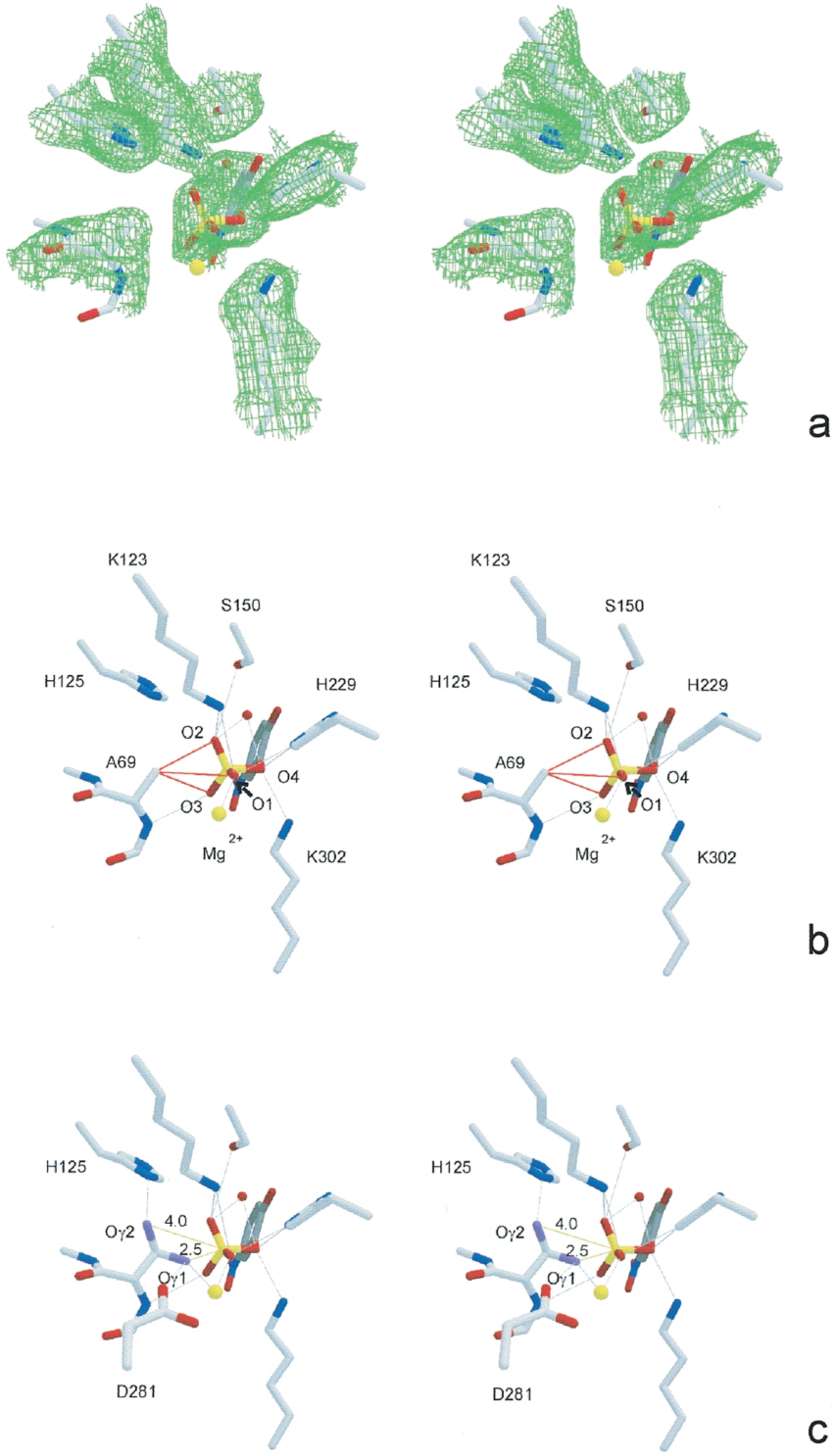

Figure 2. Active-site region in the C69A-pNCS structure showing stereo views of: (a) The protein residues and substrate in the final model and a difference electron density omit map contoured at a 1σ level. The electron density around the magnesium cation has been omitted for clarity, but its position is shown. (b) Intermolecular contacts between protein residues and substrate in the final model. The distances in Å between interacting atoms are listed in Table 1. (c) Model with the superimposed coordinates of the FGly side-chain, as determined in the wild-type, corresponding either to a disordered aldehyde or the geminal diol of an aldehyde hydrate. The side-chain of Arg73, which forms an hydrogen bond to $O^{\gamma 2}$, has been omitted for clarity.

pocket and a sigmaA weighted electron density map[19] of the C69S-pNCS soaked crystal.

Discussion

Substrate binding

The structure of the C69A-ASA pNCS complex (Figure 2) shows the substrate located at the bottom of a very narrow cleft that constitutes the active site. The sulfate tetrahedron is situated with the face opposing the ester oxygen (O4) vertex orientated towards the side-chain of residue 69. In the wild-type structure this would be the optimal orientation for a nucleophilic attack of the FGly side-chain onto the sulfate sulfur (Figure 2(c)).

The active center regions in ASA, ASB and alkaline phosphatase show remarkable structural

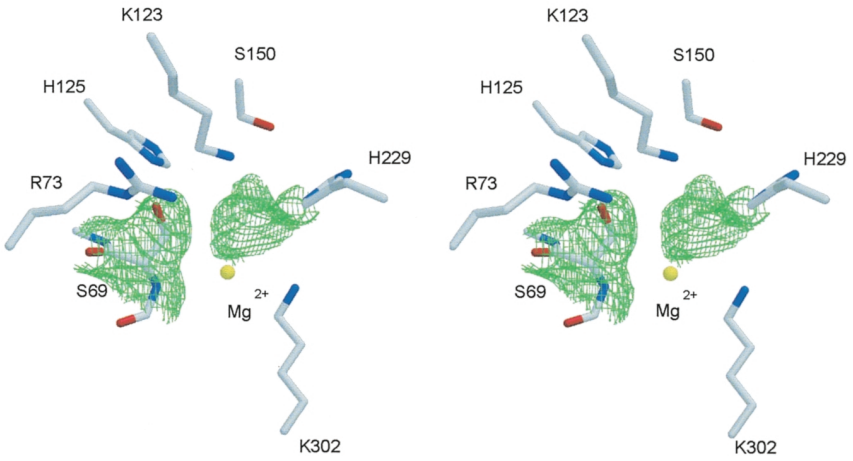

Figure 3. Stereo view of the C69S model in the area of the active site pocket and sigmaA weighted electron density map[19] of the C69S-pNCS-soaked crystal contoured at a 1σ level around the substrate and serine 69.

homology. Table 2 lists the relevant amino acid residues and metal atoms occupying nearly equivalent positions in all three enzymes, which should therefore fulfill parallel roles during catalysis. The sulfatases contain metal cations that have been modeled as Mg and Ca, respectively, and are coordinated by three aspartate side-chains, an asparagine residue and the key FGly. The corresponding site in alkaline phosphatase is occupied by Zn(2), coordinated by the hydroxyl group in the catalytic serine, two aspartate residues and a histidine residue. Not surprisingly, as the geometry of phosphate and sulfate groups are comparable, the presence of inorganic phosphate inhibits catalysis in sulfatases,[24] and a very low phosphatase activity can be measured for these enzymes.[25,12] Nevertheless, the nature of the interaction between substrate and enzyme, together with the narrow topology of the active center located at the bottom of a cavity, must be the key to the high specificity shown *in vivo* in an environment where phosphorylated substrates are predominant. A contribution of the hydroxyl groups on the sugars bonded to synthetic and natural substrates to this recognition cannot be ruled out.

The interactions of pNCS with the inert mutant C69A-ASA, lacking the functional FGly group essential for catalysis, reveal the nature of this recognition. The sulfate group, common to the target substrates, binds the enzyme through strong hydrogen bonds to the positively charged side-chains (at lysosomal pH) of His229 and Lys302. Weaker bonds, as indicated by the significantly longer distances, are formed to the side-chains of Lys123 and Ser150 as well as to the main-chain nitrogen atom in residue 69 and the divalent magnesium cation present in the active center (Figure 2(b), Table 1). Thus, the polarity is better tailored to sulfate than phosphate, given the difference in p*K*a values (phosphate would be expected to be protonated at lysosomal pH, ~5). The protonated lysine and histidine side-chains and the main-chain N-H being H-donors rather than acceptors would not be suited to a protonated substrate.

The wild-type enzyme has to interact additionally through the side-chain of the FGly69 to form a covalent bond with the substrate. A superposition of the wild-type coordinates onto the model of the C69A-ASA complex, to see the relative position of the hydroxyl or disordered oxo groups to the substrate, can be performed to illustrate this interaction during catalysis (Figure 2(c)).

The *p*-nitrocatechol moiety also interacts with the enzyme, albeit through less specific contacts, as evidenced by its disordered location among two different sites found in the crystal structure. Indeed, most natural substrates carry considerably bulkier groups. Still, the interactions involving the hydroxyl groups in pNCS should be qualitatively comparable to those taking place with the sugar residue in the natural substrate cerebrosid 3-sulfate. Involved in this binding are the charged side-chains of Arg288 and Asp152. No hydrophobic interaction with the aromatic ring is found, as expected, since the natural substrate lacks one. Other potential contacts between protein and the sugar moiety could be established to the side-chains of His405, Glu155 and Asp173. Further in the pocket, hydrophobic interactions to Leu68, Val91 and Val93 could stabilize the positioning of

Table 2. Metal cations and amino acid residues occupying equivalent positions in the structures of ASA, ASB and alkaline phosphatase, and thus expected to fulfil parallel roles during catalysis

ASA	ASB	AP
FGly69	FGly91	Ser102
Mg^{2+}	Ca^{2+}	$Zn2^{2+}$
Lys123	Lys145	$Arg166N^{\eta 1}$
His229	His242	$Arg166N^{\eta 2}$
Lys302	Lys318	$Zn1^{2+}$

the hydrophobic chains in the physiological substrate.

Catalytic mechanism of ASA

Based on the crystal structures of wild-type ASA and ASB two different mechanisms have been proposed for the hydrolysis of sulfate ester bonds, as described above. Neither structure can in itself characterize unequivocally a single general model for the sulfate ester cleavage or exclude the alternative mechanism, so that additional kinetic and structural data of reaction intermediates are required to resolve the question. Kinetic studies on the hydrolysis of radioactively labeled pNCS through C69S-ASA showed this mutant to be able to bind and hydrolyze the substrate leading to an enzyme incorporating covalently bonded sulfate to the side-chain of serine 69, though only to a 20 % extent, in contrast to no remaining covalent sulfation after catalysis in wild-type ASA.[18] Within the model mechanism proposed for ASA, the serine side-chain would still be able to carry out a nucleophilic attack on the sulfate sulfur through its single hydroxyl group, and hydrolysis of the substrate ester could also proceed, but lacking the geminal hxdroxyl group, the resulting covalently bonded sulfate could not be cleaved from the enzyme. As the serine C^{β} is saturated, this mutant could not initiate reaction at all through an electrophilic attack of this carbon atom on a sulfate oxygen atom, as proposed for ASB. The initial state in the catalytic mechanism proposed for ASA is expected to be a C^{α}-formylglycine hydrate although this cannot be determined from the crystal structure of the wild-type[11] or the P426L mutant[21]. Given the modest resolution of these structures, it is not possible to discriminate between the geminal diol and a disordered aldehyde side-chain. Nevertheless, the absence of a detectable second conformation of the serine side-chain in both the C69S-ASA free and pNCS-soaked structures makes a disordered aldehyde side-chain less likely as there is clearly an energetically favored conformation for the orientation of the serine hydroxyl group (Figure 3). The structure of C69S-ASA shows the hydroxyl group in the side-chain of residue 69 to be fixed between the side-chains of Arg73 ($73N^{\varepsilon}$-$69O^{\gamma}$: 2.85 Å) and His125 ($125N^{\delta1}$-$69O^{\gamma}$: 2.93 Å), occupying the site of $O^{\gamma2}$ in the wild-type, situated more than 4 Å away from the sulfur center in the complex. This would account for the inertia in substrate binding.

Figure 4 shows a scheme of the proposed catalytic mechanism for the cleavage of sulfate ester bonds by ASA. In the hypothetic geminal diol, one of the oxygen atoms, $O^{\gamma1}$, would be positioned at 2.5 Å from the sulfate sulfur atom, close enough to start a nucleophilic attack on the sulfur atom. The orientation of the sulfate group would also be optimal for such a reaction. The sulfate group is recognized by the charged side-chains of His229 and Lys302 interacting with the ester oxygen atom (O4) of the substrate. The closest distance of a sulfate oxygen (O1) to the magnesium divalent cation is 2.7 Å. Additional contacts to the side-chains of Lys123, Ser150 and the main-chain nitrogen atom of Ala69 all contribute to sulfate coordination, and to electron density withdrawal from the sulfate oxygen atoms, leading to an increased electrophilicity of the sulfur center. On the other hand, the

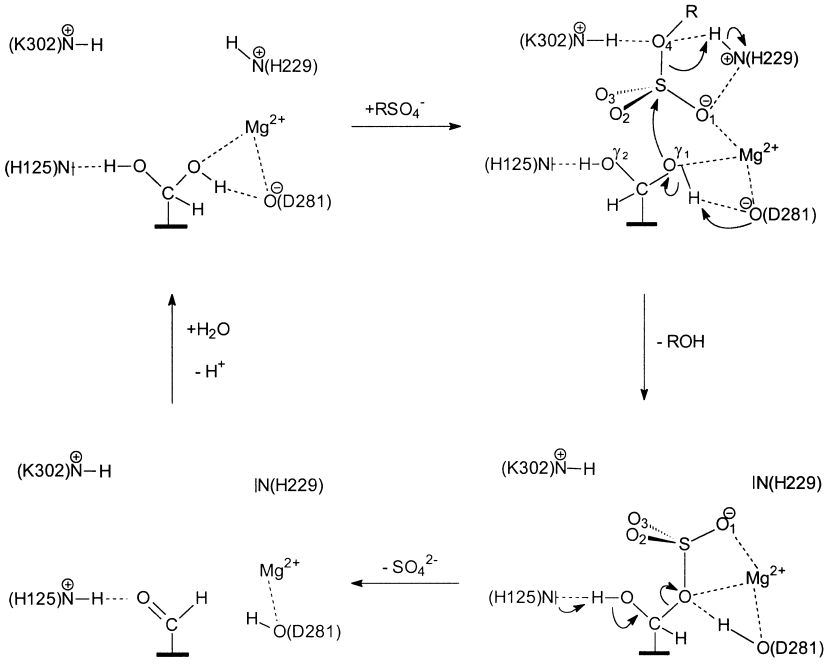

Figure 4. Scheme of the proposed catalytic mechanism. The ASA-C69A/pNCS complex represents the state of the first intermediate step (upper right), albeit lacking the covalent bond to the side-chain of residue 69.

nucleophilicity of $O^{\gamma 1}$ in the postulated aldehyde hydrate should be enhanced both by its coordination to the magnesium cation and the possibility of a proton transfer to the carboxyl group of Asp281, which would be stabilized in turn by the divalent cation.

This constellation is proposed to induce a S_N2 substitution reaction, with a pentacoordinated sulfur intermediate and resulting in the heterolytic cleavage of the S-O ester bond. This mechanism is consistent with the inversion of the tetrahedral sulfate environment implied in the case of a sulfatase from *Aspergillus orizae*,[16] but there are no data available on the resulting configuration for sulfatases from higher eukaryotic organisms. The departing alcoholate can accept a proton either from the side-chains of His229 or Lys302 or else from a neighbor water (WAT154) molecule located 3.0 Å away from the ester oxygen in the crystal structure. His229 is the most likely candidate to donate this proton, as its ε-amino group is the closest contact, with a distance of 2.4 Å. Incidentally, the analogous position to that of the ζ-amino group of Lys302 in the structure of alkaline phosphatase is occupied by a zinc cation and the equivalent site to that of His229 is occupied by an arginine residue. The former is more likely to be involved in anchoring and the latter in protonating the alcoholate.[26]

After cleavage, the interactions from the protein to the now free alcohol are not strong enough to retain it in the catalytic pocket, which could be due both to deprotonation of the His229 side-chain and to the cleavage from the tightly bonded sulfate, so it is free to diffuse out. The sulfate should now be covalently bonded to the enzyme and additionally coordinated to the magnesium cation through two of its oxygen atoms.

Cleavage of the sulfate group should take place to regenerate the aldehyde. This requires the deprotonation of the alcohol $O^{\gamma 2}$ group in the side-chain of residue 69. His125 is the most suited candidate to receive this proton. As the active cleft has a narrow shape, the sulfate group would be expected to diffuse out before water coming into the pocket could regenerate the aldehyde hydrate, stabilized by hydrogen bonds to His125, Arg73 and the metal center.

Conclusion

The crystal structure of the inactive mutant C69A-ASA soaked in pNCS mimics a reaction intermediate in the mechanism of sulfate ester hydrolysis catalyzed by this enzyme. In the absence of the covalent bond to the key side-chain of FGly69 the reaction cannot proceed, and the structure shows how the sulfate group in the substrate is non-covalently bonded to the enzyme through lysine residues 123 and 302, serine 150, histidine 229, the main-chain of the key residue 69 and the divalent cation in the active center. The ester oxygen atom interacts with the side-chains of lysine 302 and a histidine 229. The latter establishes the closest contact and is the most likely candidate to protonate the alcoholate being eliminated during catalysis.

The structure of C69S-ASA shows the serine 69 side-chain in a single conformation, suggesting that the double conformation observed in the structure of wild-type ASA and P426L-ASA is more likely to correspond to an aldehyde hydrate than to a twofold disordered aldehyde oxo group. The C69S-ASA mutant had been proven in previous experiments to be able to bind covalently to the substrate and hydrolyze it, but release of the product is hindered. The orientation of the serine side-chain in the complex would account for its relative inertness.

Materials and Methods

Expression and purification

The plasmid carrying mutated cDNA for the C69A-ASA mutant was obtained as described.[27] In the presence of pGK-hygro as selection marker the cDNA was stably transfected into mouse embryonic fibroblasts deficient for both mannose 6-phosphate receptors (mpr^- MEF cells[28]). Cells expressing C69S-ASA were obtained earlier.[18] After selection and expansion of the clones stably expressing the mutants, sulfatases were purified from cell secretions by affinity chromatography and dialyzed against 10 mM Tris-HCl, 150 mM NaCl (pH 7.4). Purity of ASA was measured by using SDS-PAGE gel-electrophoresis and the concentration was estimated by measuring the absorbance at 280 nm.

Crystallization, data collection and processing

Crystallization of the ASA-mutants was achieved by the hanging drop vapor diffusion method[29] at $20 \pm 2\,^{\circ}C$. The hanging drop was prepared by mixing equal volumes of the protein solution in the above described buffer and the reservoir solution containing 100 mM sodium acetate buffer (pH 5.3-5.4) and 11-13% (w/v) polyethylene glycol 6000. Single crystals suitable for X-ray analysis showed a tetragonal shape with maximum size of 0.6 mm × 0.5 mm × 0.5 mm. For complex formation, crystals were soaked in the mother liquor containing 20 mM *p*-nitrocatechol sulfate (Sigma), and were left for two hours at $4\,^{\circ}C$. Freezing the crystals for data collection was attempted using cryobuffers containing different salts, PEGs, alcohols, sugars and oils, however, the crystals always broke and diffracted to about 1 Å poorer resolution than the non-frozen crystals. Also, crystals that grew out of cryobuffer diffracted to about 1 Å poorer (>3 Å) at room temperature.

All crystals were mounted in sealed glass capillaries. All X-ray diffraction data were collected at room temperature using MAR Research image plate detectors at EMBL c/o DESY, Hamburg, on the X11 synchrotron beamline with an X-ray wavelength of 0.91 Å. The crystals belonged to the tetragonal space group *I*422. Data of the C69A/pNCS complex were collected from a single crystal. The crystals of the mutant C69S suffered from radiation damage and therefore complete data were obtained using two crystals. Data were processed with the programs DENZO and SCALEPACK.[30] Stat-

Table 3. Statistics of the data collection and structure refinement

	ASA-C69A/pNCS	ASA-C69S	ASA-C69S/pNCS
Resolution (Å)	2.35	2.40	2.65
R_{sym} (%)[a,b]	8.5 (33.7)	7.5 (43.9)	7.7 (36.4)
Reflections (unique)	35,243	33,144	24,966
Completeness (%)[a]	99.3 (100.0)	99.2 (98.3)	99.8 (100.0)
$I > 3\sigma(I)$ (%)[a]	75.4 (43.2)	73.9 (43.1)	69.3 (31.9)
Protein atoms	3554	3540	3572
Solvent atoms	167	171	173
r.m.s. bond length (Å)	0.014	0.017	0.019
r.m.s. bond angle (deg.)	3.0	4.0	3.2
R_{work}-factor (%), (R_{free}-factor) (%)	19.1 (23.6)	19.3 (24.3)	17.3 (24.3)

[a] Outer resolution shell (0.1 Å).
[b] $R_{sym} = (\Sigma|I_{hkl} - \langle I_{hkl}\rangle|)/\Sigma(I_{hkl})$.

istics on data quality and refinement are summarized in Table 3.

Structure solution and refinement

The structures were solved by using the atomic coordinates of the wild-type of human ASA (identification code: 1AUK, Protein Data Bank of the Brookhaven National Laboratory).[11] The residue in position 69 was first treated as glycine and all water molecules were removed. Solvent water molecules and the cation were included from the structure of the wild-type after examination of the difference electron density. Additional water molecules were included if they had suitable stereochemistry and individual *B*-factors below 80 Å. The structures were refined with the program REFMAC[31] using maximum likelihood as target. An R_{free} factor[32] was calculated based on 5% of the data chosen at random but common for all three structures to avoid overfitting.

Graphical modeling was performed using the program XtalView.[33] Omit maps were used leaving out the amino acids, solvent and substrate in the region of the active site. Figures were drawn with MOLSCRIPT,[34] Bobscript[35] and RASTER3D.[36]

The final models comprise 3564 (C69A/pNCS), 3540 (C69S) and 3572 (C69S/pNCS) protein atoms, and 167 (C69A/pNCS), 171 (C69S) and 173 (C69S/pNCS) water molecules. A magnesium ion and 28 atoms of sugar residues are found in all structures. The structure of the complexed C69A-ASA includes 15 atoms of the substrate. The final models were refined to *R*-factors of 19.4% (C69A/pNCS), 19.6% (C69S) and 17.4% (C69S/pNCS). Data statistics and final refinement figures are summarized in Table 3. The program PROCHECK[37] was used to check the stereochemical and geometrical outliers in the final structure. All main-chain angles were within allowed regions of the Ramachandran plot[38] except for those of Asn184 which carries the oligosaccharide chain.

RCSB Protein Data Bank accession numbers

The coordinates and structure factors have been deposited in the RCSB Protein Data Bank with the entry codes 1e2s and r1e2ssf for C69A-ASA/pNCS, 1e1z and r1e1zsf for C69S-ASA and 1e3c and r1e3csf for C69S-ASA/pNCS.

Acknowledgments

We thank Garib Murshudov for support and advice with REFMAC under Linux, George M. Sheldrick for discussion and support. We thank Katja Unthan-Hermeling for technical assistance and Ehmke Pohl for critically reading the manuscript. Data collection at the synchrotron beamline X11 in Hamburg at the EMBL outstation and assistance from Victor Lamzin and Thomas R. Schneider during data collection is gratefully acknowledged. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

1. Parenti, G., Meroni, G. & Ballabio, A. (1997). The sulfatase gene family. *Curr. Opin. Genet. Dev.* **7**, 386-391.
2. von Figura, K., Schmidt, B., Selmer, T. & Dierks, T. (1998). A novel protein modification generating an aldehyde group in sulfatases: its role in catalysis and disease. *BioEssays,* **20**, 505-510.
3. Schmidt, B., Selmer, T., Ingendoh, A. & von Figura, K. (1995). A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. *Cell,* **82**, 271-278.
4. Selmer, T., Hallmann, A., Schmidt, B., Sumper, M. & von Figura, K. (1996). The evolutionary conservation of a novel protein modification, the conversion of cysteine to serinesemialdehyde in arylsulfatase from *Volvox carteri*. *Eur. J. Biochem.* **238**, 341-345.
5. Miech, C., Dierks, T., Selmer, T., von Figura, K. & Schmidt, B. (1998). Arylsulfatase from Klebsiella pneumoniae carries a formylglycine generated from a serine. *J. Biol. Chem.* **273**, 4835-4837.
6. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M. A. & von Figura, K. (1998). Post-translational formation of formylglycine in prokariotic sulfatases by modification of either cysteine or serine. *J. Biol. Chem.* **273**, 25560-25564.
7. Neufeld, E. F. & Muenzer, J. (1995). The mucopolysaccharidoses. In *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds), pp. 2465-2494, McGraw-Hill, New York.

8. Kolodny, E. H. & Fluharty, A. L. (1995). Metachromatic leukodystrophy and multiple sulfatase deficiency: sulfatide lipidosis. In *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds), pp. 2693-2741, McGraw-Hill, New York.
9. Ballabio, A. & Shapiro, L. (1995). Steroid sulfatase deficiency and X-linked ichthyosis. In *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds), pp. 2999-3022, McGraw-Hill, New York.
10. 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. & Ballabio, A. (1995). A cluster of sulfatase genes on Xp22. 3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embriopathy. *Cell,* **81**, 15-25.
11. Lukatela, G., Krauss, N., Threis, K., Selmer, T., Gieselmann, V., von Figura, K. & Saenger, W. (1998). Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. *Biochemistry,* **37**, 3654-3664.
12. Bond, S. C., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J. & Guss, J. M. (1997). Structure of a human lysosomal sulfatase. *Structure,* **5**, 277-289.
13. Sowadski, J. M., Handschumacher, M. D., Krishna Murty, H. M., Foster, B. A. & Wyckoff, H. W. (1985). Refined structure of alkaline phosphatase from *Escherichia coli* at 2.8 Å resolution. *J. Mol. Biol.* **186**, 417-433.
14. Waldow, A., Schmidt, B., Dierks, T., von Bülow, R. & von Figura, K. (1999). Amino acid residues forming the active site of arylsulfatase A. *J. Biol. Chem.* **274**, 12284-12288.
15. McGovern, M. M., Vine, D. T., Haskins, M. E. & Desnick, R. J. (1982). Purification and properties of feline and human arylsulfatase B isozymes. Evidence for feline homodimeric and human monomeric structures. *J. Biol. Chem.* **257**, 21605-12610.
16. Chai, C. L., Loughlin, W. A. & Lowe, G. (1992). The stereochemical course of sulphuryl transfer catalysed by the arylsulphatase II from *Aspergillus oryzae*. *Biochem. J.* **287**, 805-812.
17. Waheed, A. & van Etten, R. L. (1980). The structural basis of the anomalous kinetics of rabbit liver arylsulfatase A. *Arch. Biochem. Biophys.* **203**, 11-24.
18. Recksiek, M., Selmer, T., Dierks, T., Schmidt, B. & von Figura, K. (1998). Sulfatases, trapping of the sulfated enzyme intermediate by substituting the active site formylglycine. *J. Biol. Chem.* **273**, 6096-6103.
19. Read, R. J. (1986). Improved coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A,* **42**, 120-129.
20. von Bülow, R. & Usón, I. (2000). The dipotassium salt of *p*-nitrocatecholsulfatase. *Acta Crystallog. sect. C,* **56**, 152-153.
21. von Bülow, R. (1999). Strukturanalyse zum katalysemechanismus und zur Stabilität der Arylsulfatase A. PhD thesis, pp. 85-108. University of Göttingen, Germany.
22. Kleywegt, G. (2000). Validation of protein crystal structures. *Acta Crystallog. sect. D,* **56**, 249-265.
23. Dierks, T., Schmidt, B. & von Figura, K. (1997). Conversion of cystein to formylglycine: a protein modification in the endoplasmatic reticulum. *Proc. Natl Acad. Sci. USA,* **94**, 11963-11968.
24. Lee, G. D. & van Etten, R. L. (1975). Purification and properties of homogeneous arylsulfatase A from rabbit liver. *Arch. Biochem. Biophys.* **166**, 280-294.
25. Uchida, T., Egami, F. & Roy, A. B. (1981). 3′,5′-cyclic nucleotide phosphodiesterase activity of the sulfatase of ox liver. *Biochim. Biophys. Acta,* **657**, 356-363.
26. Kim, E. E. & Wyckoff, H. W. (1991). Reaction mechanism of alkaline phosphatase based on crystal structures. *J. Mol. Biol.* **218**, 449-464.
27. Knaust, A., Schmidt, B., Dierks, T., von Bülow, R. & von Figura, K. (1998). Residues critical for formylglycine formation and/or catalytic activity of arylsulfatase A. *Biochemistry,* **37**, 13941-13946.
28. Kasper, D., Dittmer, F., von Figura, K. & Pohlmann, R. (1996). Neither type of mannose 6-phosphate receptor is sufficient for targeting of lysosomal enzymes along intracellular routes. *J. Cell Biol.* **134**, 615-623.
29. McPherson, A. (1998). Crystallisation of biological macromolecules. *J. Crystal Growth,* **122**, 161-167.
30. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
31. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallorg. sect. D,* **53**, 240-255.
32. Brünger, A. T. (1993). Assessment of phase accuracy by cross validation: the free *R* value. Methods and applications. *Acta Crystallog. sect. D,* **49**, 24-36.
33. McRee, D. (1999). XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165.
34. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
35. Esnouf, R. M. (1997). BOBSCRIPT. *J. Mol. Graph.* **15**, 132-134.
36. Merrit, E. A. & Bacon, D. J. (1997). RASTER3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, M. T. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crytallog.* **26**, 283-291.
38. Ramachandran, G. N. & Sassiekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **28**, 283-437.

Edited by R. Huber

(Received 11 August 2000; received in revised form 26 October 2000; accepted 26 October 2000)