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[†,‡]

G. Lukatela,[§,||] N. Krauss,[§] K. Theis,[§] T. Selmer,[||] V. Gieselmann,[||,⊥] K. von Figura,*[,||] and W. Saenger*[,§]

Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-14195 Berlin, Germany, Institut für Biochemie 2, Georg August Universität Göttingen, Gosslerstrasse 12d, D-37073 Göttingen, Germany, and Biochemisches Institut, Christian Albrechts Universität, Kiel, Olshausenstrasse 40, D-24118 Kiel, Germany

Received June 23, 1997; Revised Manuscript Received October 15, 1997

ABSTRACT: Human lysosomal arylsulfatase A (ASA) is a prototype member of the sulfatase family. These enzymes require the posttranslational oxidation of the $-CH_2SH$ group of a conserved cysteine to an aldehyde, yielding a formylglycine. Without this modification sulfatases are catalytically inactive, as revealed by a lysosomal storage disorder known as multiple sulfatase deficiency. The 2.1 Å resolution X-ray crystal structure shows an ASA homooctamer composed of a tetramer of dimers, $(\alpha_2)_4$. The α/β fold of the monomer has significant structural analogy to another hydrolytic enzyme, the alkaline phosphatase, and superposition of these two structures shows that the active centers are located in largely identical positions. The functionally essential formylglycine is located in a positively charged pocket and acts as ligand to an octahedrally coordinated metal ion interpreted as Mg^{2+}. The electron density at the formylglycine suggests the presence of a 2-fold disordered aldehyde group with the possible contribution of an aldehyde hydrate, $-CH(OH)_2$, with *gem*-hydroxyl groups. In the proposed catalytic mechanism, the aldehyde accepts a water molecule to form a hydrate. One of the two hydroxyl groups hydrolyzes the substrate sulfate ester via a transesterification step, resulting in a covalent intermediate. The second hydroxyl serves to eliminate sulfate under inversion of configuration through C−O cleavage and reformation of the aldehyde. This study provides the structural basis for understanding a novel mechanism of ester hydrolysis and explains the functional importance of the unusually modified amino acid.

Sulfatases are an evolutionarily highly conserved gene family (Franco et al., 1995; Schmidt et al., 1995). They hydrolyze sulfate ester bonds in a wide variety of structurally different compounds ranging from complex glucosaminoglucans and glycolipids to sulfated hydroxysteroids and amino acids. The physiological importance of these enzymes is illustrated by seven distinct human disorders, each caused by the deficiency of a specific sulfatase [for reviews see Ballabio and Shapiro (1995) and Neufeld and Muenzer (1995)]. In a rare disorder called multiple sulfatase deficiency (MSD), all sulfatases are catalytically inactive (Kolodny and Fluharty, 1995) because they are defective in a posttranslational modification reaction common to all sulfatases. In this reaction, a cysteine conserved among all eukaryotic sulfatases is oxidized at Cβ, yielding L-2-amino-3-oxopropionic acid (L-C_α-formylglycine, FGly), with an aldehyde as the functional group of the side chain (Schmidt et al., 1995; Selmer et al., 1996).

Among the sulfatase family, lysosomal arylsulfatase A (ASA; EC 3.1.6.8) has been most extensively studied. In humans, it is synthesized as a preprotein with 507 amino acids. After translocation into the endoplasmatic reticulum, a signal peptide of 18 amino acids is cleaved, yielding a mature protein of 489 amino acids with a calculated mass of 51 908 Da. It shares 47−59% sequence similarity with the other eight known human sulfatases (Franco et al., 1995). Three mostly high-mannose-type oligosaccharide side chains are attached at Asn158, Asn184, and Asn350, of which the first and third are phosphorylated (Sommerlade et al., 1994). In acidic milieu (pH ≈ 5) characteristic for lysosomes, human ASA exists as an octamer, which at neutral pH dissociates into dimers (T. Selmer, unpublished).

The major physiological substrate of human ASA is a sphingolipid sulfate ester, cerebroside 3-sulfate, a major constituent of the myelin sheet. *In vivo*, the 3-sulfate group can be hydrolyzed by ASA only if cerebroside 3-sulfate is complexed with the small activator protein saposin B solubilizing the hydrophobic substrate. *In vitro*, detergents can substitute for the activator protein. ASA was shown to be relatively unspecific, cleaving a wide variety of synthetic sulfate esters with a pH optimum between 5 and 6 (Nicholls & Roy, 1971; Roy, 1976; Waheed & van Etten, 1980a,b). With synthetic aromatic substrates much higher rates of hydrolysis are observed than with physiological substrates. This is associated with "anomalous kinetics" characterized by a partial inactivation due to the covalent incorporation

[†] This work was supported by the Deutsche Forschungsgemeinschaft, by Studienstiftung des Deutschen Volkes (K.T.), and by Fonds der Chemischen Industrie.

[‡] PDB file name: 1auk.

* Corresponding authors. K.v.F.: tel, +49-551-395948; fax, 49-551-395979; e-mail, kfigura@uni-bc2.gwdg.de. W.S.: tel, +49-30-8383412; fax, +49-30-8386702; e-mail, saenger@chemie.fu-berlin.de.

[§] Freie Universität Berlin.

[||] Georg August Universität Göttingen.

[⊥] Christian Albrechts Universität Kiel.

S0006-2960(97)01492-X CCC: $15.00
Published on Web 02/28/1998

Table 1: X-ray Diffraction Data and Phasing Statistics

(a) X-ray Diffraction Data

data set	source[a]	wavelength (Å)	resolution (Å)	R_{sym}[b]	refl with $I > 3\sigma(I)$ (%)	completeness (%)	no. of sites	R_{iso}[c]	phasing power[d]	R_{Cullis}[e]	anomalous data
native 1[f]	A	0.93	30−2.6	0.063 (0.276)[k]	80.3 (56.0)	96.1 (96.7)					
native 2[g]	B	0.92	30−2.1	0.051 (0.379)	78.9 (48.5)	99.4 (99.2)					
EMTS[h]	A	0.93	20−3.0	0.063 (0.170)	84.3 (65.7)	97.4 (99.0)	2	0.148 (0.178)	2.10	0.60	yes
$Pbac_2$[i]	C	1.5418	20−3.4	0.092 (0.232)	76.8 (55.4)	90.4 (85.3)	2	0.127 (0.180)	1.06	0.70	no
$K[Au(CN)_2]$	C	1.5418	20−3.4	0.098 (0.253)	76.7 (52.1)	90.3 (80.3)	1	0.116 (0.193)	1.02	0.81	no
Na_2WO_4	C	1.5418	20−6.0	0.053 (0.072)	92.3 (89.5)	96.3 (94.9)	1	0.079 (0.067)	0.62	0.93	no
UO_2ac_2[i]	C	1.5418	20−5.0	0.046 (0.058)	90.4 (87.9)	96.3 (95.2)	1	0.075 (0.073)	0.74	0.84	no

(b) Mean Figure of Merit as a Function of Resolution

resolution (Å)	11.71	8.28	6.40	5.22	4.40	3.81	3.36	3.00	total
no. of reflections	228	504	923	1452	2080	2838	3722	4692	16439
mean figure of merit[j]	0.806	0.851	0.821	0.756	0.659	0.579	0.478	0.343	0.539

[a] A: beam line X31 at EMBL outstation, DESY, Hamburg, Germany; MAR Research imaging plate detector (180 mm diameter). B: station 9.5 at SRS, Daresbury, U.K.; MAR Research imaging plate detector (300 mm diameter). C: FX 571 rotating anode generator (Enraf Nonius); FAST area detector (Enraf Nonius). [b] $R_{sym} = \sum_h\sum_i|I_i(h) - \langle I(h)\rangle|/\sum_h\sum_i I_i(h)$, where $I_i(h)$ are individual intensity measurements of a reflection and $\langle I(h)\rangle$ is the mean intensity of this reflection. [c] $R_{iso} = \sum_h|F_{PH}(h) - F_P(h)|/\sum F_P(h)$, where $F_{PH}(h)$ and $F_P(h)$ are derivative and native structure factor amplitudes. [d] Phasing power = mean ($F_{H(calc)}(h)$/lack of closure(h)), where lack of closure(h) is $|\mathbf{F}_P + \mathbf{F}_{H(calc)}| - F_{PH}$ with $\mathbf{F}_P$ being the native structure factor and $\mathbf{F}_{H(calc)}$ and $F_{H(calc)}$ being the calculated heavy atom structure factor and structure factor amplitude, respectively. [e] $R_{Cullis} = \sum_h$lack of closure(h)/$\sum_h|F_{PH} - F_P|$ for centric reflection only. [f] Data set used in MIRAS phasing and initial model building and refinement. [g] Data set used in final model refinement. [h] EMTS: sodium 2-ethylmercurithiosalicylate. [i] ac: acetate. [j] Figure of merit = $\langle\cos\Delta\alpha_i\rangle$, mean value of the cosine of the error in phase angle α of a reflection determined from the calculated probability function $P(\alpha_i)$ for the reflection having a phase angle α_i. [k] Parameters in parentheses belong to the highest resolution shell.

of SO_4^{2-} into ASA. The obtained turnover-modified ASA is partially reactivated in the presence of SO_4^{2-} and substrate (Waheed & van Etten, 1979; Prosser & Roy, 1980).

Here we report the three-dimensional structure of human lysosomal ASA based on X-ray diffraction data to 2.1 Å resolution and propose a mechanism for the hydrolysis of sulfate esters that explains the critical function of the aldehyde in position 69. Unexpectedly, ASA contains a magnesium ion in its active site and is structurally related to alkaline phosphatase, a typical member of a family of hydrolytic enzymes containing one or more zinc ions in the active site (Vallee & Auld, 1993).

EXPERIMENTAL PROCEDURES

Crystallization. Catalytically active human ASA was expressed in BHK cells cotransfected with the cDNAs of human ASA and human M_r 46 000 mannose 6-phosphate receptor (Chao et al., 1990). Amounts of 3−5 mg of ASA/L of medium could be obtained in 3−4 days. ASA was purified to homogeneity in a one-step immunoaffinity procedure (Schmidt et al., 1995) and concentrated in an ultra thimble UH 100/10 (Merck) up to ∼10 mg/mL using a buffer containing 10 mM Tris-HCl (pH 7.4) and 150 mM NaCl. Single crystals were grown by vapor diffusion in hanging drops at 18 °C. The protein solution (5 μL) was mixed with 5 μL of reservoir solution prepared with 100 mM sodium acetate buffer, pH 5.0−5.4, containing 10−13% polyethylene glycol 6000. To avoid intergrowth of crystals and to increase the amounts of single crystals, 1 μL of 1.25% octyl β-glucopyranoside in 10 mM Tris-HCl (pH 7.4) and 150 mM NaCl was added to the crystallization mixture (McPherson et al., 1986).

Data Collection. Crystals suitable for X-ray analysis grew within 1−2 weeks as bipyramids with a maximum size of 0.5 mm × 0.5 mm × 0.6 mm. They belong to the tetragonal space group $I422$. The unit cell with parameters a =131.7 Å and c = 192.0 Å contains one monomer per asymmetric unit, corresponding to a solvent content of 63% and a Matthews coefficient of V_m = 3.3 Å^3/Da (Matthews, 1968). X-ray diffraction data were collected using MAR Research image plate detectors on synchrotrons at 4 °C or a FAST detector (Enraf Nonius) on a rotating anode X-ray generator at room temperature; see Table 1. Crystal decay was not critical under these conditions so that each data set could be collected using only one crystal. In the case of the high-resolution native data set the radiation damage was compensated by irradiating the crystal at two different positions.

Phasing and Refinement. Heavy atom derivatives were prepared by soaking crystals at 18 °C in 1−2 mM solutions of the respective compounds in the same buffer which was used as reservoir solution in crystallization. The data were evaluated using DENZO (Otwinowski, 1993) or MADNES (Messerschmidt & Pflugrath, 1987). Heavy atoms were located from difference Patterson maps, and cross-phased difference Fourier syntheses were used to confirm their coordinates and to determine correlations of origins between derivatives. MIRAS (multiple isomorphous replacement including anomalous scattering) phases were calculated with MLPHARE (CCP4, 1994), and electron density maps were calculated at 3.0 Å resolution and modified by solvent flattening and histogram matching using the program DM (CCP4, 1994). The modified MIRAS electron density map was interpreted with O (Jones et al., 1991). During model building, combination of phases of the partial models with MIRAS phases using SIGMAA (CCP4, 1994) helped to improve the electron density. The obtained models were refined with X-PLOR (Brünger, 1992) using simulated annealing at T = 4000 K. Phases were extended in several steps to 2.1 Å resolution by alternate cycles of simulated annealing and manual rebuilding based on simulated annealing omit maps, and the obtained model was refined with the program REFMAC (CCP4, 1994). Water molecules were located from difference Fourier peaks and form at least one hydrogen bond to protein functional groups or to other water

molecules. Only water molecules complying with the above criteria and refined temperature factors $B < 80$ Å^2 were included in the atom list. The current model contains amino acid residues from Arg19 to Gly443 and Ala448 to Pro503, 180 solvent molecules, one magnesium ion, and the first two *N*-acetylglucosamine residues of the glycosylated Asn184. This is the only amino acid which, according to the Ramachandran plot, exhibits backbone ϕ, ψ torsion angles in a strained, not sterically allowed region. Four residues are in marginally strained, but still "allowed" regions; these are cysteines 172, 493, and 502 and Met289. The *R*-value of 23.2% for 49 794 reflections between 30 and 2.1 Å mirrors the emphasis on ideal standard geometry during crystallographic refinement (*R*-free of 27.3% for 2463 reflections). The root mean square deviations from ideal geometry of the bond lengths and angles are 0.013 Å and 1.8°, respectively. The atomic coordinates have been deposited with the Brookhaven Protein Data Bank.

RESULTS

Description of the Overall Structure. The electron density of ASA at 2.1 Å resolution enabled the location of the polypeptide chain starting with Arg19 (first amino acid after cleavage of the signal peptide) up to Pro503; the last four C-terminal amino acids, Asp504 to Ala507 and the tetrapeptide Gly444 through Ala447 could not be identified. The overall shape of the globular ASA monomer is that of a hat with dimensions 70 Å × 45 Å at the base and 50 Å height (Figure 1A,B). Two ASA monomers related by a crystallographic 2-fold axis form the homodimer with contacts between loop structures arranged on a "flat" part on one of the sides of the hat-shaped monomer. The dimer is best described as almost cylindrical with dimensions 65 Å in diameter and 90 Å in length (Figure 1C).

The secondary structure of ASA is of the mixed α/β type with 26% α-helix and 3_{10} helix, 16% β-sheet, and 46% β- and 3_{10}-turns; only 12% of the polypeptide chain cannot be classified in terms of secondary structure. The core of the enzyme consists of two β-pleated sheets (Figure 1D). The minor β-sheet is formed by four antiparallel β-strands, β15 to β18, and the major one is composed of ten β-strands. They are in parallel orientation with the exception of strands β8 and β12, which are antiparallel to the others. At the C-termini of strands β13 and β2, the two short strands, β14 and β3, form a separate parallel β-sheet; between the antiparallel strands β5 and β8, a twisted hairpin encompassing strands β6 and β7 is extruded and stabilized by two disulfide bridges. This hairpin is at the periphery of the molecule, close to the C-terminus (Figure 1B).

The major β-pleated sheet is sandwiched between helices αA, αD, and αE on one side and helices αB, αC, αG, and αH on the other; the short helix αF connects antiparallel β11, β12 (Figure 1A,B). The major and minor β-sheets are in hydrogen bonding contact through several buried water molecules and covalently linked by disulfide Cys300−Cys414.

At the C-terminus, a loop including the long helix αI is formed by nearly 80 amino acids and stretches over the whole ASA monomer, from β18 of the minor β-sheet to β9 of the major β-sheet (Figure 1B). Just N-terminal of the long α-helix I is the position of the tetrapeptide (residues 444−447) which could not be modeled. The electron density in this region is weak and suggests that the polypeptide segment is too flexible to be located. This is also the region where ASA is susceptible to proteolysis (Fujii et al., 1992).

Human ASA has a high content of 49 Pro residues per monomer, seven of which are of the *cis* type (residues 89, 164, 191, 231, 425, 495, and 497). Most of the prolines, including those in the *cis* conformation, occur in loop regions and are not conserved. There is an additional *cis*-peptide, His227−Thr228, which together with *cis*-Pro231 is located at the active site and important for its geometry. In three of the human sulfatases, the residue corresponding to Thr228 is a proline, which favors formation of a *cis*-peptide.

Of the 15 cysteines in ASA, 12 form disulfide bonds. One of the disulfides connects the two β-sheets, two stabilize the hairpin extrusion formed by strands β6 and β7, and three knot the C-terminal 20 amino acids; this cluster of 6 cysteines is unique to ASA. The remaining 3 cysteines are single: Cys38 and Cys294 are located in the homodimer interface and shielded from solvent (see below), Cys69 at the N-terminus of helix αB is the residue posttranslationally converted into C_α-formylglycine (FGly). As the cysteines occur in nonconserved loop regions, none of the bridged cysteine pairs is conserved in other sulfatases.

ASA Dimers Form Octamers at Low pH. The present crystals were grown at acidic pH, 5.0−5.4; they contain ASA homooctamers, as expected from gel filtration experiments (T. Selmer, unpublished). The octamers are composed of dimers, $(\alpha_2)_4$, the eight monomers being related by the symmetry elements of point group 422 (D_4) (Figure 2A). There are two different 2-fold symmetrical monomer−monomer contacts, one extended, 1660 Å^2, forming the dimer and the other smaller, 900 Å^2, between the dimers. The contacts involve not only different segments of the polypeptide chain but also numerous water molecules, each in symmetry equivalent duplicate.

The dimer forming contact between monomers comprises strand β13 at one corner of the major β-sheet and four loops, two containing free cysteines which are shielded from solvent: loop 32−44 (free Cys38); loop 286−300 (free Cys294) covalently linked by disulfide Cys300−Cys414 to loop 406−414; and loop 430−439. There are only 2 × 3 direct hydrogen bonds between the two monomers (see caption to Figure 2), augmented by hydrogen bonds mediated by 2 × 25 water molecules trapped in the contact interface (Figure 2B). In spite of these few direct monomer−monomer interactions, the cohesion of the monomers in the dimer is so strong that in the previous literature, the 120 kDa ASA species occurring at neutral pH was described as being a monomer (Nicholls & Roy, 1971; Waheed & van Etten, 1985).

The most prominent feature of the dimer−dimer contact is the long α-helix I, supported by two segments of the minor β-sheet, loop β15−β16 and strand β18. Two ASA molecules are so arranged that the helices αI are nearly antiparallel at an angle of 156° and interdigitate through aliphatic side chains spaced 3 to 4 residues apart along helix αI. This indicates that the octamer is stabilized by hydrophobic interactions, as has been very early suggested for ASA from ox liver (Nichol & Roy, 1966). In addition, there are 2 × 3 direct monomer−monomer hydrogen bonds and several others mediated by the 2 × 9 trapped water molecules

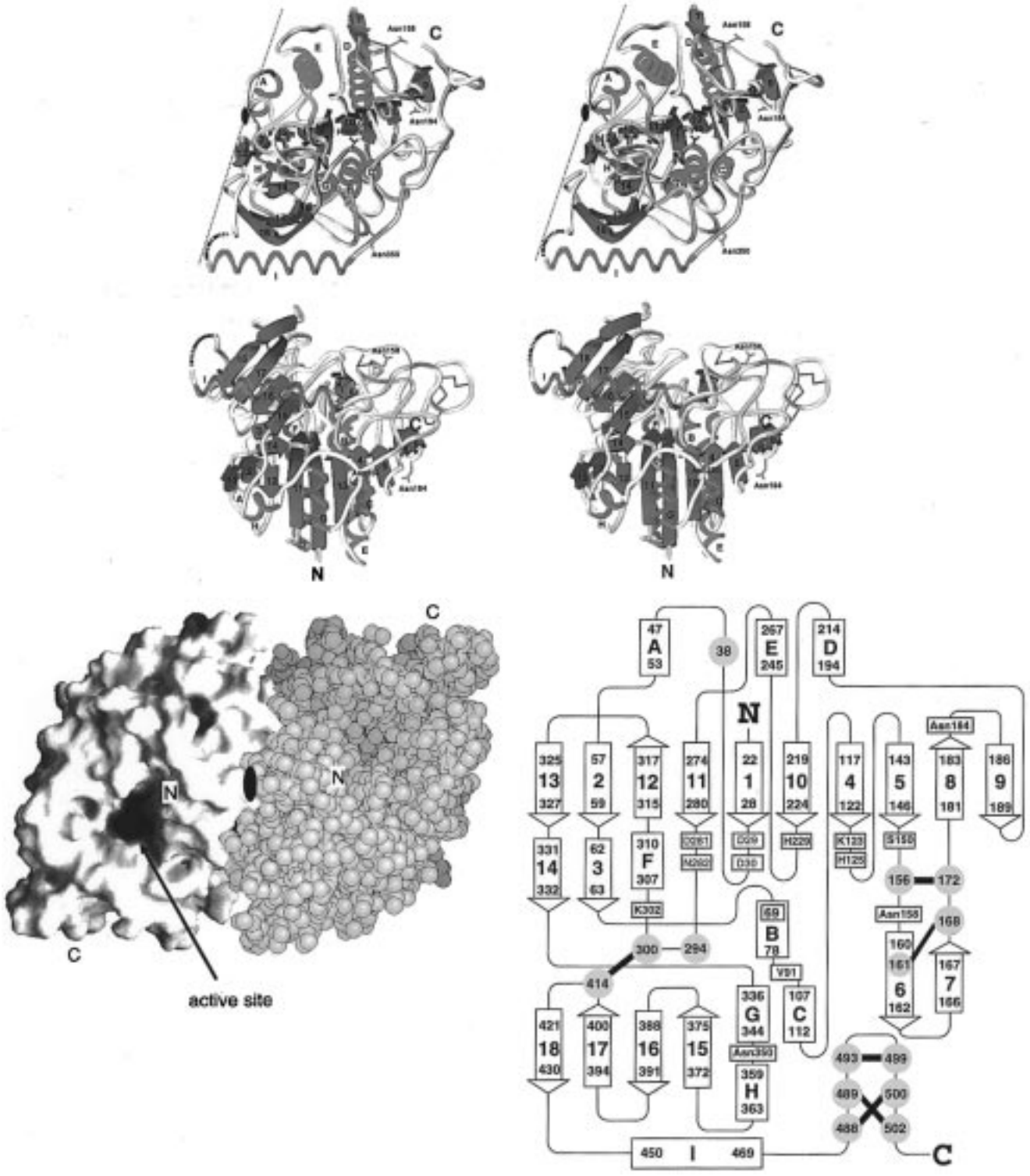

FIGURE 1: Schematic stereo representation of the structure of ASA in two views looking on the base of the hat-shaped monomer (A, top) and, after 90° rotation about the horizontal, on the side (B, middle). N- and C-termini are indicated; α-helices (red spirals) and β-strands (blue arrows) are labeled A−I and 1−18, respectively. Glycosylation sites Asn158, Asn184, and Asn350 are in red, and cysteines are in green; the modified Cys69 is shown as a hydrate with the *gem*-diol group, $-CH(OH)_2$, in blue and labeled FGly69 (hidden in panel B). The break of the polypeptide chain comprising segment Gly444−Ala447 is dashed. The left side of the molecule in (A) marked by a dashed line is the dimer interface, and the 2-fold rotation axis perpendicular to the paper plane indicated by (a black oval) relates the two monomers in the dimer (see also panel C). Drawn with SETOR (Evans, 1993). (C, bottom left) ASA dimer. One subunit is represented as a van der Waals surface in yellow in the same orientation as in (A); the other is represented with an electrostatic potential surface calculated with His, Lys, and Arg charged positive and Asp and Glu charged negative. Calculated and drawn with GRASP (Nicholls, 1992). Negative potentials are in red and positive in blue; the deep blue patch is at the active site. The subunits are related by a crystallographic 2-fold axis (indicated by a black oval) perpendicular to the paper plane. (D, bottom right) Secondary structure of ASA. α-Helices are shown as columns and β-strands as arrows numbered as in panels A and B with N- and C-terminal amino acids in each of these structure elements numbered; cysteines are in yellow with disulfide bridges shows as thick black lines; amino acids in the active site are in red boxes (green numbering coordinating to Mg^{2+} and black numbering to binding to sulfate oxygens; V91 is boxed because it is located in the active site and probably protects FGly69 from solvent); glycosylation sites are boxed in blue.

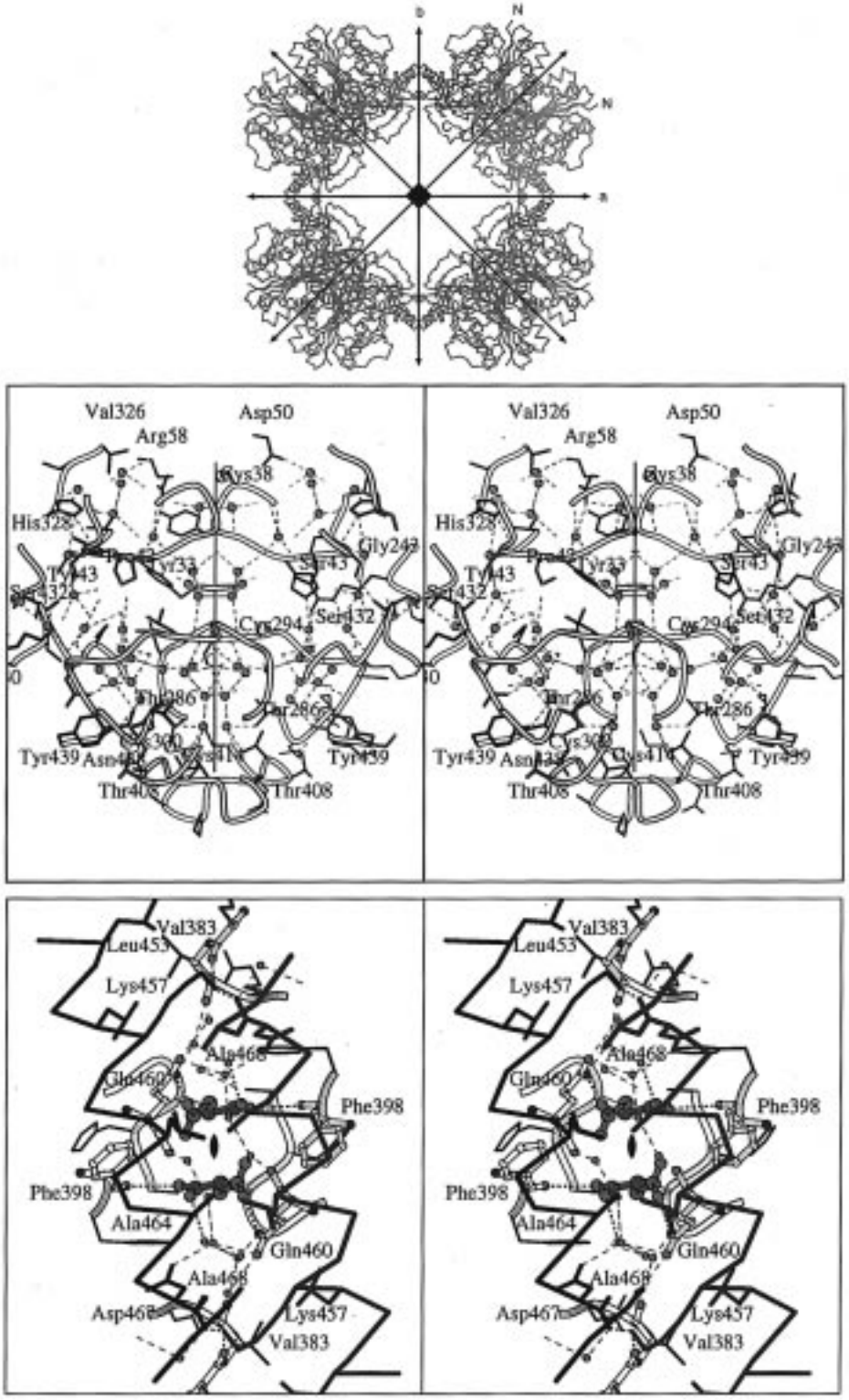

FIGURE 2: (A, top) ASA octamer formed by four dimers with crystallographic symmetry axes indicated. In one of the dimers, the monomers are indicated by different colors. Drawn with O (Jones, 1991). (B, middle) Monomer−monomer contact interface in the dimer [generated by the diagonal (a + b) in panel A]. The ASA molecule toward the viewer is drawn in blue, water molecules are shown as red spheres, cysteine sulfur is in yellow, and the crystallographic 2-fold axis relating the two monomers is drawn as a black line. There are 2 × 3 direct monomer−monomer hydrogen bonds (dashed black lines): His328NϵH···OPro42, 3.4 Å; Ser43Oγ···OSer432, 2.6 Å; and Tyr439-OηH···OγThr408, 2.9 Å. The 2 × 25 water molecules in the interface mediate hydrogen bonds between the ASA monomers (dashed red lines). Drawn with MOLSCRIPT (Kraulis, 1993). (C, bottom) Dimer−dimer contact interface (generated by axis a or b in panel A), looking along the 2-fold symmetry axis indicated by a black oval. The amino acids involved in the interface contact are drawn as ball and sticks; one monomer is in blue and the other is in black. The cohesion between the dimers is due to hydrophobic interactions between helices αI formed by side chains indicated with thick lines: Leu453; Leu456; the aliphatic side chain of Lys457; Leu461; Ala464; Ala468; supported by several hydrogen bonds mediated by 2 × 9 water molecules and 2 × 2 direct monomer−monomer hydrogen bonds: Lys457Nξ donates hydrogen bonds to Glu382O, 3.1 Å, and Asp467O, 3.0 Å. The side chain of Glu424 is 2-fold disordered; the position forming the intermolecular hydrogen bond Glu424OϵH···OPhe398 and stabilizing the octamer is drawn in green, and the other forming the intramolecular hydrogen bond is in purple. Drawn with MOLSCRIPT (Kraulis, 1993).

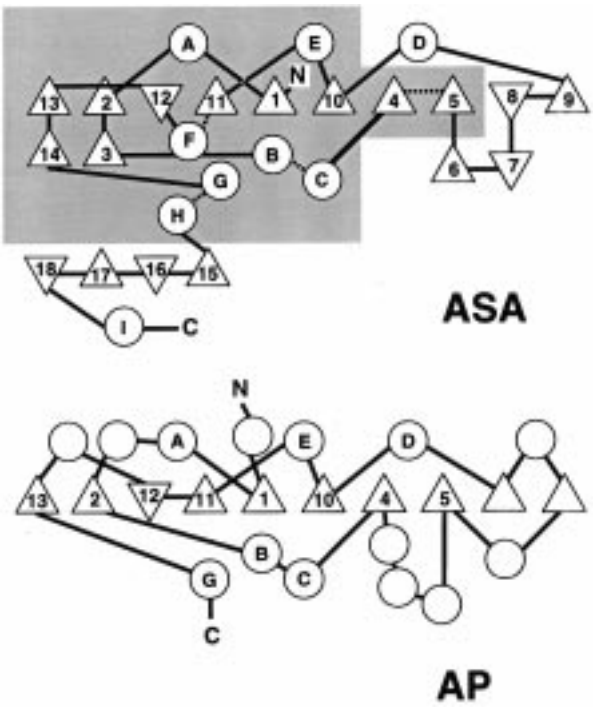

FIGURE 3: Comparison of the folds of ASA and *E. coli* alkaline phosphatase (AP). α-Helices and β-strands are indicated as circles and triangles, respectively, and numbered as in Figure 1A,B. △ and ▽ represent β-strands pointing toward and away from the viewer, respectively, according to Figure 1A. In ASA, the symbols within the shaded area show structural elements that are highly conserved among the human sulfatases; dashed lines indicate loop regions of variable length in different human sulfatases.

(Figure 2C). The association between the dimers has been shown to occur even across ASA isolated from different species (Waheed & van Etten, 1985), suggesting that the interface structure is conserved.

ASA and Alkaline Phosphatase from Escherichia coli Are Structurally Related. A systematic comparison of secondary structure elements between ASA and all proteins in the Brookhaven Protein Data Bank using the program DALI (Holm & Sander, 1993) gave one single hit with a *z*-score of 12.1, indicating that alkaline phosphatase (AP) from *E. coli* and human ASA share a high degree of structural similarity (Figure 3). Besides AP the program DALI indicates no other closely related structure as expressed by a *z*-score below 3.5 for the remaining entries in the Data Bank.

The main structural feature of AP is a β-pleated sheet with 10 strands sandwiched between α-helices (Sowadski et al., 1985; Kim & Wyckoff, 1991). Except for the two β-strands 8, 9 on one corner of the major β-pleated sheet in ASA, all the connections between the strands have the same orientation, and α-helices A−E have comparable lengths and positions in ASA and AP (Figure 3). Superposition of the two structures gave 13 segments of contiguous Cα atoms (altogether 166 Cα atoms in the region 21−347 in ASA and 43−449 in AP) with a root mean square deviation of 1.9 Å. Like ASA, AP occurs as a homodimer, and the interfaces between the two monomers are in comparable positions. AP belongs to a family of hydrolytic enzymes cleaving P−O, C−O, and C−N bonds in various esters and amides (Vallee & Auld, 1993). They all carry one or more metal ions (Zn, Co, Mg) at the C-terminal side of a parallel β-sheet that are directly involved in catalysis. The structural relationship of ASA to this enzyme family came as a surprise because the amino acid sequence similarity is low, and available data did not suggest a metal-dependent catalysis by ASA.

The Metal Binding Site in ASA. In the electron density maps, a strong difference density peak persisted that was not connected to the polypeptide chain. There are three indications that this peak is located at a binding site for a divalent cation: (I) Three negatively charged side chains (Asp29, Asp30, Asp281) are in close vicinity to coordinate the cation by four of the carboxylate oxygens. (II) In a Pb^{2+} derivative used for solving the phase problem, the difference Fourier map showed a weak density at this site, which we interpret as a partial exchange of the native metal ion against Pb^{2+}. (III) In the related alkaline phosphatase, a Zn^{2+} cation is in an identical position. In addition to the three aspartates, Asn282Oδ and the side chain of FGly69 coordinate the metal ion, completing a distorted octahedral coordination sphere (Figure 4). The nature of the ligands (all oxygen), the metal−oxygen distances (given in the caption to Figure 4), the electron density, and the temperature factor ($B = 24.2$ Å^2 within the range of those of the ligand binding oxygen atoms, $B = 19.8-35.8$ Å^2) identify this metal ion as Mg^{2+}. Attempts to refine this metal ion as Ca^{2+}, Cu^{2+}, or Zn^{2+} led to unreasonable *B*-factors, and atom absorption spectroscopy also identified this cation as Mg^{2+} (H. Engert, unpublished). All the amino acids coordinating to Mg^{2+} are conserved among the sulfatases except for Asn282, which is replaced in some of the human sulfatases by the homologous Gln or by His, which can coordinate Mg^{2+} with an unprotonated imidazole nitrogen.

The Active Site of ASA. Two earlier observations point to a critical role of FGly69 for catalysis of ASA. Genetic deficiency of this residue is associated with enzymatic inactivity (Schmidt et al., 1995), and aldehyde reagents such as sulfite, hydrazine, hydroxylamine, and cyanide act as inhibitors for ASA (Nicholls & Roy, 1971; Roy, 1976; Lee & van Etten, 1975). The crystal structure reveals that FGly lies at the bottom of a cavity which, according to the electrostatic potential calculation shown in Figure 1C, is positively charged. This supports the view that the cavity harboring FGly69 is indeed the active site binding the negatively charged sulfate group of the substrate.

The active site cavity is located at the C-terminal end of the major β-sheet and is lined mainly with charged amino acids. These charged groups are found in loops at the C-termini of β11 (Asp281, Asn282), β1 (Asp29, Asp30), β10 (His229), β4 (Lys123, His125), and β5 (Ser150) and at the N-termini of helices αF (Lys302) and αB (Arg 73) (Figure 1A,D).

It is not clear from previous studies whether FGly69 occurs as an aldehyde or an aldehyde derivative. To avoid bias in the interpretation of the electron density at FGly69, this amino acid was treated as glycine during structure refinement. The $|F_o - F_c|$ difference electron density map in Figure 4A shows V-shaped density with one short and one long arm, the branch point being centered at Cβ of FGly69. We interpret this electron density as a 2-fold disordered aldehyde group, with possible contribution from aldehyde hydrate, $-CH(OH)_2$ (Baker & Engel, 1992). One of the two oxygen positions is coordinated to Mg^{2+}, and the other is hydrogen bonded to His125Nδ at 2.7Å.

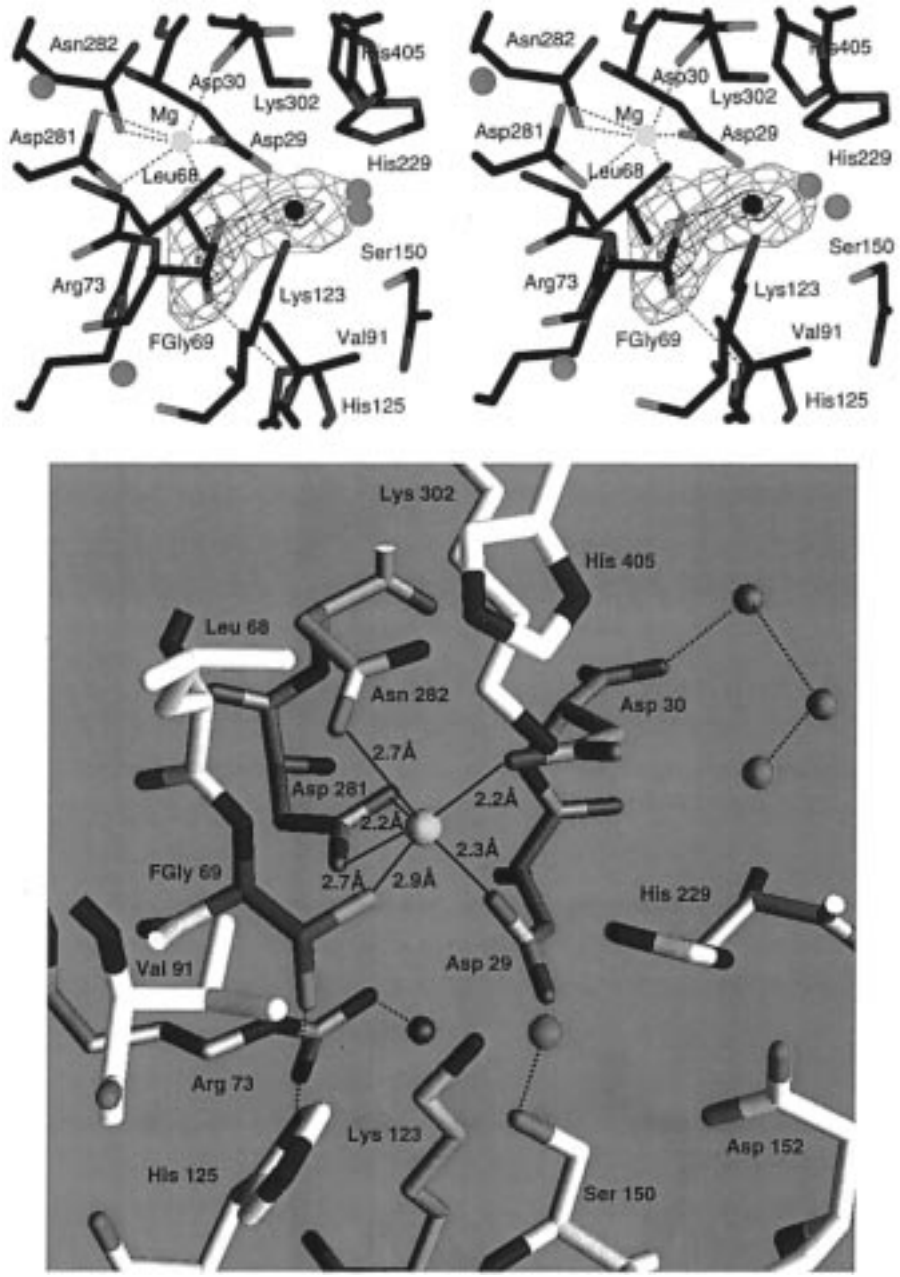

FIGURE 4: (A, top) Stereoview showing a section of the $|F_o - F_c|$ difference electron density map at the active site (green contours, 1.7σ level, and blue contours, 10σ level) in the vicinity of FGly69, calculated with glycine in position 69. The aldehyde group of FGly69 is seen 2-fold disordered with possible contribution from the aldehyde hydrate *gem*-diol group, $-CH(OH)_2$. The black sphere in the long arm of the difference electron density is likely to represent chloride ion occupied 30%; for sterical reasons it is only there if the OH group or aldehyde oxygen coordinated to Mg^{2+} (yellow sphere) is not present; water molecules are shown as red spheres. Drawn with O (Jones, 1991). (B, bottom) Active site of ASA (in a view similar but not identical to panel A) with hydrogen bonding (<3.5 Å dotted lines) and Mg^{2+} coordination interactions (solid lines). Mg^{2+} (yellow) is surrounded in distorted octahedral geometry by Asp29Oδ and Asn282Oϑ in apical positions and by Asp30Oϑ, Asp281Oϑ1, Oϑ2 and FGly69 Oγ in equatorial positions. Red spheres indicate water molecules. Both conformations of the side chain of the disordered FGly69 (drawn in green), which could also be partially present as the *gem*-diol group of an aldehyde hydrate, are shown. Note that one of the oxygen sites is coordinated to Mg^{2+} and the other hydrogen bonded to His125Nδ. Drawn with MOLSCRIPT (Kraulis, 1993).

The additional electron density in the long arm (see Figure 4A) could be due to a chloride ion. This site is only partially occupied (30%), presumably because one of the conformations of the aldehyde (Oγ pointing toward Mg^{2+}) and/or the presence of an aldehyde hydrate prevent full occupation due to a short O···Cl distance. The positively charged groups of Lys123, His229, and Lys302 are in close vicinity to Cl^- with His229Nϵ and Lys302Nξ at 3.6 and 3.4 Å, respectively; the distances to other functional groups are >4 Å. The contacts of Cl^- to His229 and Lys302 are longer than the ~3.2 Å expected for N−H···Cl^- hydrogen bonds (Jeffrey & Saenger, 1991) but shorter than the calculated van der Waals distances, ~4 Å (Weast et al., 1962). Since this is the binding site for the sulfate group of the substrate, Cl^- appears to be a compromise to fill the site in the absence of the substrate.

If the sulfate group of a substrate RO−SO_3^- is modeled in the site occupied by Cl^- and hydrogen bonded to the above mentioned positively charged amino acids, it is close to the functional group of FGly69 and effectively closes the narrow

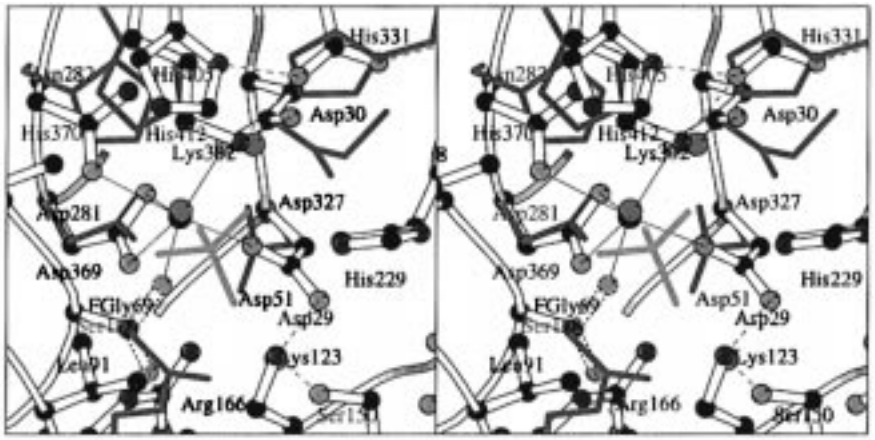

FIGURE 5: Stereoview of the local superposition of the active centers of alkaline phosphatase AP (green) and ASA; differences between Cα positions of Asp29, FGly69, and Asp281 in ASA and Asp51, Ser102, and Asp369 in AP, respectively, were minimized. Metal ions in AP are drawn as small green spheres. Ser102 (AP) coincides with FGly69 (ASA), Zn_1^{2+} (AP) with Mg^{2+} (ASA), Zn_2^{2+} (AP) with Lys302Nξ (ASA), and Mg^{2+} (AP) with Lys123Nξ (ASA). The phosphate bound to AP drawn as a red tetrapod is coordinated to Zn_1^{2+}, Zn_2^{2+}, and Arg166Nε,Nη. Drawn with MOLSCRIPT (Kraulis, 1993).

active site to the extent that water cannot enter. This modeling suggests that the moiety R of a substrate RO−SO_3^- is located on the surface of ASA, with the hydrophobic tail forming a complex with the activator protein saposin. There is no obvious pocket to harbor R, in agreement with the hydrolysis of a number of structurally unrelated synthetic substrates by ASA.

Among the nine human sulfatases with known sequence, the aspartates (29, 30, and 281) coordinating to magnesium and the positively charged amino acids Arg 73, Lys123, and Lys302 as well as His125 and His229 are strictly conserved, suggesting structural and electrostatic similarity of the active sites of these enzymes.

DISCUSSION

Proposal for the Catalytic Mechanism of ASA. The reaction mechanisms of ASA and AP are likely to share some properties. This assumption is based on the high degree of structural similarity of the two enzymes at and around the active site (see below) and the observation that ox liver ASA displays phosphodiesterase activity in that it hydrolyzes adenosine and guanosine 3′,5′-cyclic phosphates to the corresponding nucleoside 5′-monophosphates (Uchida et al., 1981). The analogy is further strengthened by infrared spectroscopy using $H_2{}^{18}O$ (Spencer, 1958, 1962) and by ^{17}O NMR (Marker & Roy, 1983), indicating that ASA hydrolyzes the RO−SO_3^- bond comparable to the hydrolysis of RO−PO_3^{2-} by AP (Coleman, 1992).

The global superposition of ASA and of AP in the AP−phosphate complex (Kim & Wyckoff, 1991) described above was further refined by local superpositioning of the active site Cα groups of Asp29, FGly69, and Asp281 in ASA and their equivalents in AP (Figure 5). In AP, the ester oxygen of substrate RO−PO_3^{2-} has been proposed to be coordinated by the zinc cation Zn_1^{2+}, and the three non-ester oxygens are coordinated by a second Zn_2^{2+} and in bidentate mode by Arg166Nε,Nη. In ASA, Lys302Nξ is in a similar position to Zn_1^{2+}; Mg^{2+} replaces Zn_2^{2+} in virtually the same site, and the side chains of the pair Lys123Nξ and His229 may substitute for Arg166Nε,Nη. A noteworthy detail in ASA is His227 forming an energetically unfavorable *cis*-peptide bond to Thr228 and the occurrence of *cis*-Pro231. Evidently, these *cis* conformations are required to allow the participation of His229 in substrate binding or catalysis.

Since the two catalytically essential zinc cations in AP are in the same positions as Lys302Nξ and Mg^{2+} in ASA, we suggest as a working hypothesis the mechanistic scheme shown in Figure 6. The reaction is initiated (I) with the addition of water to the aldehyde function of FGly69 to form the diol and by binding of the substrate RO−SO_3^- to the active site; in analogy to Zn_1^{2+} in AP, Lys302Nξ may hydrogen bond to the ester oxygen, and the three non-ester oxygens hydrogen bond or form salt bridges to the side chains of Lys123, Ser150, and His229.

The key feature of the proposed mechanism of ASA is the participation of FGly69 as an aldehyde hydrate whose *gem*-hydroxyl groups serve functions reminiscent of Ser102Oγ*H* and a water molecule in AP. The hydrated aldehyde, however, may be considered as a "mild acid" (Bell & Onwood, 1962) and is a much more potent functional group than Ser102Oγ*H* because the *gem*-hydroxyl groups have a lower pK_A (13.6 in acetaldehyde hydrate) compared to a primary hydroxyl group with a pK_A nearly 5 units higher (18.0 in ethanol; McEwen, 1936). As illustrated in Figure 6 (step II), one of the aldehyde hydrate hydroxyl groups attacks the sulfur of the substrate to form the covalent intermediate E−$CH(OH)OSO_3H$. In a subsequent step (step III), the deprotonated His125 activates the second hydroxyl of the aldehyde hydrate group to induce the elimination of the SO_4^{2-} anion, and the aldehyde is regenerated. The involvement of His125 imidazole agrees with the observed optimum for ASA catalysis, pH 5−6. This step contrasts the hydrolysis of the covalent phosphoserine intermediate E−OPO_3^{2-} in AP where water attacks phosphorus from the solvent side.

Common to the reaction mechanisms in ASA and AP is the transesterfication step resulting in a protein−phosphate (AP) or protein−sulfate (ASA) ester intermediate. The water required in both reactions for hydrolysis is coordinated and activated by a Zn^{2+} cation in AP, whereas it is activated in a "covalent form" by formation of the aldehyde hydrate in ASA. As a consequence, cleavage of the ester bond is different in both mechanisms: in AP it is initiated by the nucleophilic attack of the activated water (cleavage of a P−O

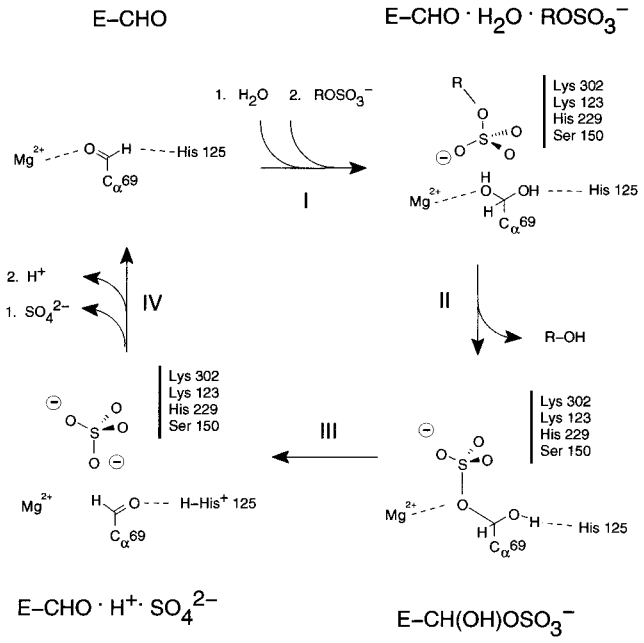

FIGURE 6: Catalytic mechanism of ASA proposed in analogy to that deduced for AP (Sowadski et al., 1985; Kim & Wyckoff, 1991). Formylglycine 69 (FGly69) takes up a water to form the hydrate, $-CH(OH)_2$; Lys123, Ser150, His229, and Lys302 are in close proximity of the sulfate group of the substrate when it enters the active site (I). In $E-CH(OH)_2 \cdot RO-SO_3H$, the sulfate is coordinated to these side chains; Mg^{2+} activates one of the hydroxyl groups of the aldehyde hydrate for nucleophilic attack at sulfur to form the covalent intermediate $E-CH(OH)OSO_3H$ (II). The side chain of His125 hydrogen bonds and deprotonates the second OH of FGly69 $-CH(OH)_2$; elimination of HSO_4^- from the intermediate takes place and re-forms the FGly69 aldehyde (III). Note that this mechanism requires the configuration of the sulfate group to be inverted whereas that of the phosphate group in AP catalysis is retained. The transesterification could occur in a dissociative (via SO_3) or associative (via pentacovalent sulfur intermediate) mechanism, as discussed for phosphate esters (Coleman, 1992), and our scheme deliberately leaves open this issue.

bond), whereas in ASA it results from elimination of sulfate (cleavage of a C−O bond instead of an S−O bond) and regeneration of the aldehyde. The catalytic cycle is closed by the release of sulfate and the deprotonation of His125 (Figure 6, step IV).

The proposed catalytic mechanism of ASA is supported by the recent observation that an ASA variant in which FGly69 is replaced by Ser69 hydrolyzes *p*-nitrocatechol sulfate forming the covalent reaction intermediate $E-CH_2-OSO_3H$. The latter, however, is not processed further (T. Selmer, unpublished). This is in agreement with the proposed mechanism, as the hydroxyl group required for elimination of the sulfate is missing in the reaction intermediate of this variant. At the same time the sulfurylation of Ser69 in the variant rules out the formal possibility that the reaction with wild-type ASA is initiated by nucleophilic attack of a non-ester sulfate oxygen on the carbonyl C atom of FGly69.

The different catalytic mechanisms proposed for AP and ASA have distinct effects on the configuration of the released phosphate and sulfate, respectively. The phosphate ester is hydrolyzed by AP in two steps, each involving inversion of configuration at phosphorus, yielding a net retention of configuration. In contrast, the configuration of the sulfate is inverted during hydrolysis by ASA, because the nucleophilic attack and the elimination reaction at the sulfate group occur from the same side.

The proximity of Mg^{2+} to FGly69 and to the sulfate group of the substrate and its location at the site occupied in AP by the Zn_2^{2+} cation suggest that Mg^{2+} is involved in one or several steps of the catalytic mechanism as proposed in Figure 6. This assumption, however, has to be taken with caution. Incubation of ASA with EDTA for 2 days does not affect catalytic activity (T. Selmer, unpublished), suggesting that Mg^{2+} is either not removed from its binding site or not essential for ASA activity.

Glycosylation Sites. As expected, the three N-glycosylation sites in ASA, Asn158, Asn184, and Asn350, are all located at the periphery of the molecule (Figure 1A,B). The sugar moieties of the attached oligosaccharides are disordered and could not be identified in the electron density except for the first two *N*-acetylglucosamines covalently linked to Asn184. This is the only amino acid with ϕ,ψ torsion angles not in allowed regions of the Ramachandran plot, suggesting that the two *N*-acetylglucosamines are in close contact with the surface of ASA and exert some strain on this section of the polypeptide chain. Unlike oligosaccharides attached at Asn158 and Asn350, the one at Asn184 is not phosphorylated (Sommerlade et al., 1994). The close contact of the latter with the surface of ASA may impair its accessibility for the phosphotransferase.

Regulation of Dimer−Octamer Association. The pH-dependent dimer−octamer equilibrium had been proposed to be regulated with deprotonation−protonation of carboxylate or imidazole side chains (Nichol & Roy, 1966; Nicholls & Roy, 1971). The only such amino acid at the dimer−dimer interface is Glu424 located on β18 (Figure 2C). This is where, due to *cis*-Pro425, strand β18 forms a β-bulge and bends toward helix αI (Figure 1A), pushing Glu424 close to the 2-fold rotation axis relating two dimers in the octamer.

The two symmetry-related carboxylates of Glu424 are only 5 Å apart and are found 2-fold disordered in the electron density map. One of the carboxylate positions is likely to represent the protonated state because this would allow hydrogen bonding to the peptide O of Phe398 at 3.2 Å in the symmetry-related dimer, thereby stabilizing the octamer (green dashed line in Figure 2C). The other position forms an intramolecular hydrogen bond to Gln460Nϵ at 3.0 Å (purple dashed line in Figure 2C).

Both Glu424 and 2 × 7 water molecules are enclosed within a hydrophobic cavity formed by *cis*-Pro425, Pro426, Phe398, Phe399, and Ala422 and shielded from solvent. We can envisage that Glu424 exerts its switch function as follows: at pH 5, Glu424 carboxylate is protonated, forming the intermolecular hydrogen bond to Phe398O and stabilizing the octamer. At pH 7, Glu424 is deprotonated; the hydrogen bond to Phe398O is no longer possible, and the octamer is destabilized by electrostatic repulsion between the negatively charged Glu424 side chains. In the dissociated form, the side chain changes conformation and Glu424 carboxylate forms an intramolecular hydrogen bond to Gln460Nϵ. According to the amino acid sequence (Franco et al., 1995), arylsulfatases C and E may also form dimer−octamer equilibration regulated by pH. They carry Asp instead of Glu424 in ASA and have identical amino acids in positions corresponding to *cis*-Pro425, Pro426, and Gln460.

Comparison to Other Human Sulfatases. The regions with highest sequence similarity among human sulfatases correspond in ASA to residues 19−146 (β-strands 1−5 and α-helices A−C) and residues 219−363 (β-strands 10−14 and α-helix E−H) (Figure 1D). This indicates that most of the central major β-sheet and its decorating α-helices are structurally conserved among sulfatases. The connecting sequence represented in ASA by β-strands 6−9 and helix D lacks sequence and presumably also structural similarity between the sulfatases. Interestingly, this is the region where a hydrophobic insertion is found in steroidsulfatase (arylsulfatase C), which has been shown to anchor this enzyme in the membrane in such a way that the N- and C-terminal portions of the polypeptide are both exposed at the lumenal side of the membrane (Stein et al., 1989). Steroidsulfatase shares this hydrophobic insertion with arylsulfatases D and E, which have not yet been characterized biochemically. According to their amino acid sequence, they are more closely related to each other and to steroidsulfatase than to other sulfatases (Franco et al., 1995).

Comparison with the Structure of Arylsulfatase B. Very recently the crystal structure of a related sulfatase, the human lysosomal arylsulfatase B, specific for hydrolysis of *N*-acetylgalactosamine 4-sulfate was published (Bond et al., 1997). The enzyme was overexpressed in CHO cells and crystallized under similar conditions as ASA. The crystals diffracting to 2.5 Å resolution contained a mixture of 20% precursor and 80% mature forms of arylsulfatase B, with the former a single polypeptide chain of 522 amino acids and the latter cleaved at positions 423 and 465, the three fragments held together by disulfide bonds.

In contrast to ASA, arylsulfatase B occurs only as monomer in solution and in the solid state. The secondary structure and the tertiary fold of ASA and arylsulfatase B are virtually identical except for the loops, some of which have different lengths due to insertions/deletions in the amino acid sequence, and for the number and positions of cysteines and glycosylation sites. The extruded hairpin $\beta6-\beta7$ lies flat on the surface of the protein in ASA but sticks out into solution in arylsulfatase B. If this hairpin would serve as a recognition element for targeting to the lysosome, as proposed for arylsulfatase B (Bond et al., 1997), the hairpin would have to adopt in ASA a conformation different from that in the crystal.

The active site geometries of the two enzymes are identical, as are the strictly conserved functional amino acids. This correspondence even extends to the *cis*-peptide preceding His229 (ASA) and His 242 (arylsulfatase B). The exception is that in arylsulfatase B the electron density at the modified cysteine was interpreted as a sulfate ester of the hydrated FGly, $-CH(OH)OSO_3^-$. Interestingly, this sulfate ester would correspond to an intermediate postulated in the reaction scheme formulated for ASA. In arylsulfatase B, the active site coordinates a metal ion identified as Ca^{2+} by the same procedures and reasonings as the Mg^{2+} in ASA. The metal ions in both enzymes are located at corresponding positions with almost identical local environment. The only difference between the metal coordination sphere in the two structural models is that in arylsulfatase B the Ca^{2+} ion is additionally coordinated through an oxygen atom of the sulfate group. In both structures it is not clear whether the metal ions identified are the *in vivo* ions as these could have been replaced during production and purification of the proteins. As in ASA, there is no significant effect on enzyme activity if arylsulfatase B is incubated with EDTA.

An enzymatic mechanism has been proposed for arylsulfatase B which differs from that illustrated in Figure 6 for ASA. Arylsulfatase B sulfate was assumed to be a "resting" form, the sulfate group being released by hydrolysis or in equilibrium with the free aldehyde. If the substrate *N*-acetylgalactosamine 4-sulfate enters the active site, its sulfate binds covalently to FGly (residue 91 in ASB). A nucleophile (water, a residue of the protein, or an OH group of the substrate) is supposed to attack the sulfur atom, forming a pentacoordinated intermediate which, after breaking the S−O (galactose) bond, results in arylsulfatase B sulfate. This mechanism is analogous to that derived for alkaline phosphatase. In contrast to that proposed for ASA, it requires inversion of configuration of the sulfate group.

Conclusions. The three-dimensional structure of ASA reveals that this enzyme (and related sulfatases) belongs to the divalent metal ion containing hydrolases represented by AP. Both types of enzymes act on similar substrates. In AP, the active site carries two Zn^{2+} and Ser102Oγ; the corresponding moieties in ASA are Mg^{2+}, Lys302Nξ, and the aldehyde function of FGly69. A catalytic mechanism is proposed for ASA where the aldehyde group of FGly69 activates a water molecule by forming a hydrate, $-CH(OH)_2$. Subsequently, one of the oxygens of the resulting *gem*-hydroxyl groups attacks the sulfur atom of the substrate $ROSO_3^-$. In a final step, this oxygen leaves as part of SO_4^{2-}, and the other oxygen re-forms the carbonyl bond of the aldehyde. The proposed mechanism is associated with inversion of configuration of the sulfate group, distinct to hydrolysis of phosphate esters by AP, where configuration is retained. This difference can serve as a tool to experimentally probe the scheme depicted in Figure 6 if substrates with chiral sulfate groups (Nicolaides & Radom, 1995) are utilized in the reaction with ASA.

ACKNOWLEDGMENT

We thank D. Röleke and Dr. K. Gessler for help with data collection, Dr. B. Schmidt, Göttingen, for critically reading the manuscript, and Dr. G. Schluckebier for help with structure comparisons. Data collection using synchrotrons at Orsay (LURE), Daresbury (SRS), and Hamburg (EMBL outstation and MPG beamlines at DESY) is gratefully acknowledged. N.K. is on leave from Max-Delbrück-Centrum für Molekulare Medizin, Forschungsgruppe Kristallographie, Robert-Rössle-Strasse 10, D-13122 Berlin.

REFERENCES

Baker, A. D., & Engel, R. (1992) *Organic Chemistry*, p 81, West Publishing Co., St. Paul, MN.

Ballabio, A., & Shapiro, L. J. (1995) Steroid sulfatase deficiency and X-linked ichthyosis, in *The metabolic and molecular bases of inherited disease* (Scriver, C. R., Beaudet, A. L., and Sly, W. S., Eds.) pp 2999−3022, McGraw-Hill, New York.

Bell, R. P., & Onwood, D. N. (1962) *Trans. Faraday Soc. 58*, 1557−1561.

Bond, C. S., Clements, P. R., & Ashby, S. R. (1997) *Structure 5*, 277−289.

Brünger, A. T. (1992) *X-PLOR Version 3.1*, Yale University Press, New Haven and London.

CCP4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.

Chao, H. H. J., Waheed, A., & Pohlmann, R (1990) *EMBO J. 9*, 3507−3513.
Coleman, J. E. (1992) *Structure 21*, 441−483.
Evans, S. V. (1993) SETOR: Hardware lighted three-dimensional solid model representation of macromolecules, *J. Mol. Graphics 11*, 134−138.
Franco, B., Meroni, G., & Parenti, G. (1995) *Cell 81*, 15−25.
Fujii, T., Kobayashi, T., & Honke, K (1992) *Biochim. Biophys. Acta 1122*, 93−98.
Holm, L., & Sander, L. (1993) *DALI version 2.0, J. Mol. Biol. 233*, 123−138.
Jeffrey, G. A., & Saenger, W. (1991) *Hydrogen Bonding in Biological Structures*, Springer Verlag, Heidelberg and New York.
Jones, T. A., Zou, J.-Y., & Cowan, S. W. (1991) *Acta Crystallogr. A47*, 110−119.
Kim, E. E., & Wyckoff, H. W. (1991) *J. Mol. Biol. 218*, 449−464.
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., Eds.) pp 2693−2741, McGraw-Hill, New York.
Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
Lee, G. D., & van Etten, R. L. (1975) *Arch. Biochem. Biophys. 166*, 280−294.
Marker, A., & Roy, A. B. (1983) *Biochim. Biophys. Acta 742*, 446−451.
Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497
McEwen, W. K. (1936) *J. Am. Chem. Soc. 58*, 1124−1129.
McPherson, A., Koszelak, S., & Axelrod, H (1986) *J. Crystal Growth 76*, 547−553.
Messerschmidt, A., & Pflugrath, J. W. (1987) *J. Appl. Crystallogr. 20*, 306−315.
Neufeld, E. F., & Muenzer, J. (1995) The mucopolysaccharidoses, in *The metabolic and molecular bases of interited disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D. Eds.) pp 2645−2694, McGraw-Hill, New York.
Nichol, L. W., & Roy, A. B. (1966) *Biochemistry 5*, 1379−1388.
Nicholls, A. (1992) *GRASP*, Columbia University, 630 Westth St., New York, NY 10032.
Nicholls, R. C., & Roy, A. B. (1971) Arylsulfatases, in *The Enzymes, 3rd Ed.*, Vol. 5, pp 21−41, Academic Press, New York.
Nicolaides, A., & Radom, L. (1995) *J. Chem. Soc., Chem. Commun.*, 701−702.
Otwinowski, Z. (1993) Oscillation Data Reduction Program, in *Proceedings of the CCP4 Study Weekend: "Data Collection and Proceeding"* (Sawyer, L., Isaacs, N., & Bailey, S., Eds.) pp 56−62, SERC Laboratory, Daresbury.
Prosser, C. I., & Roy, A. B. (1980) *Biochim. Biophys. Acta 613*, 130−139.
Roy, A. B. (1976) *J. Exp. Biol. Med. Sci. 54*, 111−135.
Schmidt, B., Selmer, T., & Ingendoh, A.(1995) *Cell 82*, 271−278.
Selmer, T., Hallmann, A., & Schmidt, B. (1996) *Eur. J. Biochem. 238*, 341−345.
Sommerlade, H.-J., Selmer, T., & Ingendoh, A. (1994) *J. Biol. Chem. 269*, 20977−20981.
Sowadski, J. M., Handschumacher, M. D., & Krishna Murthy, H. M. (1985) *J. Mol. Biol. 186*, 417−433.
Spencer, B. (1958) *Biochem. J. 69*,155−159.
Spencer, B. (1962) *Biochem. J. 73*, 442−447.
Stein, C, Hille, A., & Seidel, J. (1989) *J. Biol. Chem. 264*, 13865−13872
Uchida, T., Egami, F., & Roy, A. B. (1981) *Biochim. Biophys. Acta 657*, 356−363.
Vallee, B. L., & Auld, D. S. (1993) *Biochemistry 32*, 6493−6501.
Waheed, A., & van Etten, R. L. (1979) *Arch. Biochem. Biophys. 195*, 248−251.
Waheed, A., & van Etten, R. L. (1980a) *Arch. Biochem. Biophys. 203*, 11−24.
Waheed, A., & van Etten, R. L. (1980b) *Biochim. Biophys. Acta 614*, 92−101.
Waheed, A., & van Etten, R. L. (1985) *Int. J. Pept. Protein Res. 26*, 362−372.
Weast, R. C., Selby, S. M., & Hodgman, C. D. (1962) *Handbook of Chemistry and Physics*, p D90, The Chemical Rubber Co., Cleveland, OH.

BI9714924