Structure of Human Estrone Sulfatase Suggests Functional Roles of Membrane Association*

Estrone sulfatase (ES; 562 amino acids), one of the key enzymes responsible for maintaining high levels of estrogens in breast tumor cells, is associated with the membrane of the endoplasmic reticulum (ER). The structure of ES, purified from the microsomal fraction of human placentas, has been determined at 2.60-Å resolution by x-ray crystallography. This structure shows a domain consisting of two antiparallel α-helices that protrude from the roughly spherical molecule, thereby giving the molecule a “mushroom-like” shape. These highly hydrophobic helices, each about 40 Å long, are capable of traversing the membrane, thus presumably anchoring the functional domain on the membrane surface facing the ER lumen. The location of the transmembrane domain is such that the opening to the active site, buried deep in a cavity of the “gill” of the “mushroom,” rests near the membrane surface, thereby suggesting a role of the lipid bilayer in catalysis. This simple architecture could be a prototype utilized by the ER membrane in dictating the form and the function of ER-resident enzymes.

Human estrone (E1)/DHEA sulfatase (ES), along with cytochrome P450 aromatase and 17β-hydroxysteroid dehydrogenase 1 (17HSD1), is responsible for maintaining high levels of the active estrogen, 17β-estradiol (E2), in tumor cells. ES catalyzes the hydrolysis of E1 sulfate, which is subsequently reduced to E2 by 17HSD1 (see Refs. 1 and 2 and references therein). The presence of ES in breast carcinomas and ES-dependent proliferation of breast cancer cells have been demonstrated (3, 4). Germ line mutations and inactive enzyme have also been associated with X-linked ichthyosis, a disease related to scaling of the skin (5, 6).

ES is found in the microsomal fraction of human placenta (7, 8). The membrane-bound enzyme is distributed in the rough endoplasmic reticulum (ER) including the perinuclear cisternae, the Golgi cisternae, the endocytic structures, and the plasma membrane (9). Localization of ES in the smooth and rough ER was demonstrated by immunohistochemical labeling (10). ES is synthesized in several tissues, including human placenta, skin fibroblasts, breasts, and fallopian tubes (7–16). It has been proposed that intracrine biosynthesis by P450arom, ES, and 17HSD1 in the breast accounts for most of the estrogens in postmenopausal women and, hence, could be a major factor in hormonal breast cancers (3). Selective estrogen enzyme modulators that inhibit these enzymes have shown promise as antiproliferative agents in hormonal breast carcinoma (3, 17–19). Rational design of specific ligands requires detailed understanding of the molecular structure of the active site.

The cytoplasmic side of the rough ER is the site for biosynthesis for membrane and secreted proteins, which are then transported to the lumen side across the ER membrane by translocation signal peptides at the N terminus. Many ER-resident steroidogenic enzymes are known to possess strong ER membrane association that presumably plays roles in their functionality. To the best of our knowledge, no structure of an ER-resident protein showing a trans-membrane association is available to date. Here, we report the crystal structure of the full-length, active ES purified from human placenta, as determined by x-ray crystallography at 2.60-Å resolution. The proximity of the putative membrane-spanning domain and lipid-associated regions to the active site suggests a direct influence of the lipid bilayer on the enzyme activity, thereby providing the first structural evidence of functional significance for membrane association in this class of microsomal enzymes.

EXPERIMENTAL PROCEDURES

Purification and Crystallization—Details of purification and crystallization of the enzyme have been previously described (8). Briefly, the full-length enzyme in an active form was purified from human placenta. Diffraction quality crystals were obtained from 45–50% methylpentanediol in 100 mM Tris-HCl, pH 8.5, containing 0.10–0.20 M ammonium phosphate (monobasic) and 0.3% n-octyl-β-D-glucopyranoside (BOG). The space group is P21 with unit cell dimensions of a = b = 116.98 Å, c = 102.66 Å, and α = β = 90°, γ = 120°. There are six symmetry-related positions, and the unit cell volume is 1.2 × 106 Å3. With a molecular mass of roughly 65 kDa, the specific volume for 1 molecule per asymmetric unit is 3.1 Å3/Da. Assuming a specific volume of 0.746 cm3/g for proteins, the solvent content of the unit cell is 60.3%.

Data Collection.—The native ES data was collected at the F-1 beam line of the Cornell High Energy Synchrotron Source (Ithaca, NY). The station was equipped with a Quantum-210 CCD detector and an Oxford Cryostream cryostat. The crystal to detector distance was 250 mm, and

* This work was supported in part by National Institute of Health Grant GM62794 (to D. G.). The macromolecular diffraction facility at the Cornell High Energy Synchrotron Source is supported by National Institutes of Health Grant RR-01646. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Francisco G. Hernandez-Guzman‡§, Tadayoshi Higashiyama‡, Walter Pangborn‡, Yoshio Osawa‡, and Debashis Ghosh‡§

From the ‡Department of Structural Biology, Hauptman-Woodward Medical Research Institute, Buffalo, New York 14203 and the §Department of Molecular and Cellular Biophysics, Roswell Park Cancer Institute, Buffalo, New York 14263
the wavelength used was 0.948 Å. The diffraction data were collected in 110 frames, each of 1° oscillation and 20-s exposure. The data was processed and scaled using the programs Data Processing Suite (20) and MOSFLM (21). A total of 150,699 observations were measured for 24,534 unique reflections to a 2.60 Å resolution with an overall Rmerge value of 0.068 on intensities. The data overall was 97.4% complete, and the completeness and average intensity to S.D. ratio in the last shell (2.74–2.60 Å) were 96.2% and 3.1, respectively. Data for the heavy atom derivative crystals were gathered under cryogenic conditions at the in-house R axis-IV image plate detector receiving x-rays (Cu Kα) from a rotating anode source and processed with DENOZO and SCALEPACK (22). Table I provides a summary of the data collection statistics.

Structure Solution and Refinement—The initial phasing of the data was carried out by the single isomorphous replacement and anomalous scattering (SIRAS) method. Two derivative data sets, each having the same single major mercury-binding site and prepared by soaking the crystals with HgCl₂, were used for the purpose. The software package SOLVE (23) was used to determine and refine the position of the heavy atom. The program package PHASES (24) was used for phase combination, solvent flattening, and the final phase calculation. A summary of phase calculation, combination, and solvent flattening is provided in Table II. Interpretation of the electron density map and model building was carried out using CHAIN (25). A molecular replacement solution (using MOLREP from the CCP4 suite (26)) with the structure of arylsulfatase A (ASA; Protein Data Bank code 1AUK) as the search model aided the chain tracing in a few weakly defined regions of similarity. The identification of the molecular replacement solution was possible by cross-phasing the (mercury minus native) difference Fourier synthesis with phases from the correct molecular replacement solution. A single peak, corresponding to the mercury-binding site and having a height 4 times that of the next background peak, confirmed the correctness with phases from the correct molecular replacement solution. A summary of x-ray diffraction data for native and derivative crystals is shown in Fig. 1a. The tertiary structure consists of two domains: a globular (55 × 60 × 70 Å), polar domain containing the catalytic site and the putative transmembrane domain consisting of two antiparallel hydrophobic α-helices. The major polar domain consists of two subdomains with the α/β sandwich fold (Fig. 2b). Subdomain 1 (SD1), wound around a central 11-stranded (strands 1, 2, 4–11, and 17) mixed β-sheet flanked by 13 α-helices/helical turns (helices 1–7 and 10–15), contains the catalytic core. Subdomain 2 (SD2), consisting of roughly 110 C-terminal residues and wound around a four-stranded antiparallel β-sheet (strands 13–16) flanked by α-helix 16, packs against turn and loop regions of the β-sheet of subdomain 1. The two putative transmembrane helices protrude on one side of the nearly spherical polar domain, thereby giving the overall molecule a “mushroom-like” shape. The polar domain resembles in shape, size, and fold the two known structures of the soluble forms of human sulfatases, arylsulfatases A (ASA) (29) and B (ASB) (30). A comparison of three sequences is shown in Fig. 3a. The sequence identities of ES were 32 and 24% with ASA and ASB, respectively (Protein Data Bank codes 1AUK and 1FSU, respectively). The similarity of the tertiary structures is evident when three structures are superimposed (Fig. 3b). The root mean-squared deviations for 151 carbon atoms of identical/conserved residues were 0.90 and 1.02 Å for ES:1AUK and ES:1FSU, respectively. However, in addition to the insertion of the transmembrane domain in ES, there are significant additions, deletions, and substitutions in ES, especially in the C-terminal SDS, which are critical to its membrane association and anchoring of the catalytic domain to the membrane surface.

### RESULTS

Quality of the Model—Fig. 1, a and b, show representative experimental and refined electron density maps, respectively, for the putative transmembrane helices. The atomic model for the region is shown in Fig. 1c. The overall secondary and tertiary structures of the molecule are shown in Fig. 2a and the folding topology in Fig. 2b. The residues of the full-length functional ES are numbered from 22 to 583. The N-terminal histidine and residues 477–480 in a weakly defined loop could not be modeled. In addition, the C-terminal residues 576–583 did not have any appreciable electron density. The backbone conformation of one residue, Ala389, was in the disallowed range despite its well defined electron density. More than 98% of all nonproline and nonglycine residues were in the allowed ranges of backbone conformations defined by Ramachandran plot. The crystallographic R-factor for all reflections was 0.25, and Rfree was 0.30. The average temperature factor (B) for the protein molecule was 58.2 Å², in good agreement with an overall temperature factor of 62.5 Å² calculated from the Wilson plot. The average R, Rfree, and B values calculated for 31 membrane protein structures (28), having at least 2500 atoms in the asymmetric unit and determined at a resolution between 2.2 and 3.3 Å are 0.24, 0.29, and 56 Å², respectively. Thus, the overall quality of the ES model is typical of membrane protein structures of similar size determined at a medium resolution range.

### Structure of Human Estrone Sulfatase

#### Table I

| Data set | Source        | Wavelength | Soaking conditions | No. of crystals used | Resolution (Å) | No. of unique reflections | Overall R (last 0.1 Å shell) | Overall data completeness (last 0.1 Å shell) | Rsym | Rfree |
|----------|---------------|------------|--------------------|----------------------|----------------|--------------------------|-------------------------------|-----------------------------------------------|------|-------|
| Native   | CHESS         | 0.948      |                    | 1                    | 30-2.60        | 24,534                   | 7.8 (3.1)                     | 97.4 (96.2)                     | 0.068 |       |
| HA7: HgCl₂ | In house R-AXIS-IV | 1.5418     | 0.1 mm 6 days      | 1                    | 99-3.80        | 8265                     | 28.3 (15.2)                   | 99.1 (97.8)                     | 0.057 | 0.101 |
| HA12: HgCl₂ | In house R-AXIS-IV | 1.5418     | 0.5 mm 7 days      | 1                    | 99-2.90        | 18,358                   | 25.6 (4.2)                    | 98.4 (92.3)                     | 0.063 | 0.099 |

#### Description of the Overall Structure

The three-dimensional structure of the full-length ES from human placenta as determined by x-ray crystallography is shown in Fig. 2a. The tertiary structure consists of two domains: a globular (55 × 60 × 70 Å), polar domain containing the catalytic site and the putative transmembrane domain consisting of two antiparallel hydrophobic α-helices. The major polar domain consists of two subdomains with the α/β sandwich fold (Fig. 2b). Subdomain 1 (SD1), wound around a central 11-stranded (strands 1, 2, 4–11, and 17) mixed β-sheet flanked by 13 α-helices/helical turns (helices 1–7 and 10–15), contains the catalytic core. Subdomain 2 (SD2), consisting of roughly 110 C-terminal residues and wound around a four-stranded antiparallel β-sheet (strands 13–16) flanked by α-helix 16, packs against turn and loop regions of the β-sheet of subdomain 1. The two putative transmembrane helices protrude on one side of the nearly spherical polar domain, thereby giving the overall molecule a “mushroom-like” shape.

The polar domain resembles in shape, size, and fold the two known structures of the soluble forms of human sulfatases, arylsulfatases A (ASA) (29) and B (ASB) (30). A comparison of three sequences is shown in Fig. 3a. The sequence identities of ES were 32 and 24% with ASA and ASB, respectively (Protein Data Bank codes 1AUK and 1FSU, respectively). The similarity of the tertiary structures is evident when three structures are superimposed (Fig. 3b). The root mean-squared deviations for 151 carbon atoms of identical/conserved residues were 0.90 and 1.02 Å for ES:1AUK and ES:1FSU, respectively. However, in addition to the insertion of the transmembrane domain in ES, there are significant additions, deletions, and substitutions in ES, especially in the C-terminal SDS, which are critical to its membrane association and anchoring of the catalytic domain to the membrane surface.

### The Active Site Architecture

The location of the transmembrane domain is such that the opening to the active site, buried deep in a cavity in the “gill” of the “mushroom,” rests near the membrane surface. A close-up view of the catalytic site is shown in Fig. 4a. The catalytic amino acid hydroxylformylglyc-
Structure of Human Estrone Sulfatase

TABLE II
Phasing results for two HgCl₂ derivative data sets (HA7 and HA12)

| Data set | Resolution range | No. of reflections phased | R-centric | Phase combination results | Solvent flattening results |
|----------|------------------|---------------------------|-----------|---------------------------|---------------------------|
|          | Å                |                           |           | FOM<sup>ª</sup> | Phasing power<sup>ª</sup> | Phasing power SAS | Highest resolution | Reflections | Combined FOM | Correlation coefficient | R<sub>free</sub> |
| HA7      | 99-3.8           | 7723                      | 0.59      | 0.516                    | 2.70                     | 2.56          | 2.90Å         | 16,938      | 0.774        | 0.849                  | 0.286          |
| HA12     | 15-2.9           | 16,995                    | 0.58      | 0.482                    | 2.60                     | 1.92          |                           |             |              |                        |                |

<sup>a</sup> SIR, single isomorphous replacement.
<sup>b</sup> SAS, single anomalous scattering.
<sup>c</sup> R-centric = Σ/F<sub>obs</sub> − F<sub>calc</sub> for centric reflections only, where F<sub>calc</sub> and F<sub>obs</sub> refer to the heavy atom derivative and parent measured structure factors, respectively.
<sup>d</sup> FOM, figure of merit = (cos Ω<sub>α</sub>), mean value of the cosine of the error in phase angle α of a reflection determined from the calculated probability function P(α) for the reflection having the phase angle α.

Phasing power = mean |F<sub>H.A.(calc)</sub> − F<sub>obs</sub>| of all reflections (h). Lack of closure = |F<sub>calc</sub> + F<sub>H.A.(calc)</sub> − F<sub>obs</sub> with F<sub>H.A.(calc)</sub> defined as the calculated heavy atom structure factor amplitude.

<sup>e</sup> R<sub>ave</sub> = Σ |F<sub>calc</sub> − F<sub>obs</sub>|/Σ |F<sub>calc</sub>|, where F<sub>calc</sub> is the observed structure factor amplitudes and F<sub>obs</sub> is the calculated structure factor amplitudes obtained from the map inversion.

<sup>f</sup> R<sub>free</sub> = (No. of reflections used, % complete) 0.254(23,350,92%)

<sup>g</sup> rmsδ bonds 0.0072
<sup>h</sup> rmsδ angles 1.554
<sup>i</sup> Average isotropic B values

<sup>j</sup> Main chain 57.9
<sup>k</sup> Side chain 58.6
<sup>l</sup> Whole chain 58.2
<sup>m</sup> Overall solvent 47.5

Ramachandran plot statistics
<sup>n</sup>% residues of 468 nonglycine and nonproline in most favored (number of residues) 75.0%(351)
<sup>o</sup>% residues of 468 nonglycine and nonproline in additional allowed region ns (number of residues) 23.2%(109)
<sup>p</sup>% in disallowed (number of residues) 0.2%(1)

<sup>ª</sup> R-factor = Σ |F<sub>calc</sub> − F<sub>obs</sub>|/Σ |F<sub>calc</sub>|
<sup>b</sup> R<sub>ave</sub> same calculation as R-factor except that a 5% subset of randomly selected reflections was used.
<sup>c</sup> rmsδ, root mean square deviation.

Table III
Summary of refinement results

| Parameter | Value |
|-----------|-------|
| Protein atoms in the model (1 polypeptide chain) | 4339 |
| No. of NAG atoms (2 molecules in 2 sites) | 28 |
| No. of BOG atoms (2 molecules) | 40 |
| No. of ions modeled | 1 |
| Solvent oxygens | 130 |
| Total number of atoms | 4539 |
| Resolution range (Å) | 60.0-2.60 |
| Unique data used (P > 0) (% complete) | 24,539/96.7 |
| Crystallographic R (R-factor<sup>ª</sup>) | 0.254(23,350,92%) |
| R<sub>ave</sub> (No. of reflections used, % complete) | 0.30(1,183,4.7%) |
| Average isotropic B values | |
| Main chain | 57.9 |
| Side chain | 58.6 |
| Whole chain | 58.2 |
| Overall solvent | 47.5 |

Ramachandran plot statistics
| % residues of 468 nonglycine and nonproline in most favored (number of residues) | 75.0%(351) |
| % residues of 468 nonglycine and nonproline in additional allowed region ns (number of residues) | 23.2%(109) |
| % in disallowed (number of residues) | 0.2%(1) |

<sup>ª</sup> R-factor = Σ |F<sub>calc</sub> − F<sub>obs</sub>|/Σ |F<sub>calc</sub>|
<sup>b</sup> R<sub>ave</sub> same calculation as R-factor except that a 5% subset of randomly selected reflections was used.
<sup>c</sup> rmsδ, root mean square deviation.

Estrone (FG)₃, created by posttranslational modification of Cys⁷⁵ (31), was found to be covalently linked to a sulfatate moiety (i.e. as a sulfatate ester of FG (FGS)). A large spherical electron density at the center of the catalytic site near the FGS₇⁵ side chain was interpreted to be the cation required for the ES activity. The refinement was carried out with three possible bivalent cations, a Mg²⁺, Ca²⁺, or Mn²⁺, at the metal binding site. Refinement with a Ca²⁺ ion at the site resulted in a temperature factor of 41.4 Å², which correlated very well with the average temperature factor of 43 Å² of eight coordinating oxygen atoms from side chains and exhibited no residual electron density (positive or negative) at the metal center. A Mg²⁺ and a Mn²⁺ at the site, on the other hand, yielded temperature factors of 23.6 and 51.0 Å², respectively. Although neither can be completely ruled out, Ca²⁺, in all likelihood, is the cation present in the crystal structure of ES. ASB (Protein Data Bank code 1FSU) also has a Ca²⁺ at the active site (30).

Whereas the oxygen atoms of the Asp³⁵, Asp³⁶, Asp³⁴², Gln³⁴¹, and FGS₇⁵ side chains serve as ligands for the bivalent cation (Ca²⁺...O distances range between 2.2 and 2.8 Å), Lys¹³⁴, Lys³⁴³, and Arg⁷⁹ are involved in neutralization of negative charges of the carboxylic moieties. The positively charged amino groups of the Lys¹³⁴ and Lys³⁶⁸ side chains are also within contact distance (2.7–3.1 Å) of the sulfate oxygens of FGS₇⁵ (Fig. 4a). In addition, two sulfate oxygens (O₉₂ and O₂) are within coordination distances of Ca²⁺ (2.6–2.7 Å). Several histidine residues in the immediate vicinity may play important roles in catalysis as well. The imidazole ring of His¹³⁶ is within a hydrogen-bonding distance (2.6 Å) of the hydroxyl of FGS, and His²⁹⁰ Nε₂ is 2.6 Å away from the sulfate oxygen (O₂) of FGS. Also, the His³⁸⁶ side chain is linked to Lys³⁶⁸ and the Thr²⁹¹ side chain via a bridging water molecule. The main chain NH groups of FGS₇⁵ and Thr²⁹¹ point toward the sulfate-binding cavity and may thus be responsible for the formation of an oxyanion hole. A few additional solvent molecules, the presence of which may be important for the completion of the hydrolysis, are located inside the substrate-binding cavity.

The catalytic end of the active site in ES is highly homologous to those in ASA and ASB. Nine of the 10 catalytically important residues, namely Asp³⁵, Asp³⁶, FGS₇⁵, Arg⁷⁹, Lys¹³⁴, His¹³⁶, His²⁹⁰, Asp³⁴², and Lys³⁶⁸, are identical in all three enzymes. When these nine α-carbon atom positions were superimposed by least-squares minimization, the root mean-squared deviation was 0.39 Å between ES and either ASA or ASB. The 10th residue is Glu³⁴³, a ligand to the cation, which
is an asparagine in both ASA and ASB. However, the substrate binding end of the catalytic cavity has a considerably different architecture in ES. In Fig. 4b, a E1 sulfate molecule is placed in the active site with its sulfate moiety superimposed with the crystallographically observed sulfate of FGS75. The resulting model of the covalently linked substrate mimics the first step of the proposed mechanism (see “Discussion”). Residues Leu74, Arg98, Thr99, Val101, Leu103, Leu167, Val177, Phe178, Thr180, Gly181, Thr484, His485, Val486, and Phe488 surround the steroid backbone, yielding mostly hydrophobic contacts and could therefore be involved in substrate recognition. Many of the residues from the membrane-associated regions described above line the groove leading to the active site.

Transmembrane Domain—Two antiparallel helices 8 and 9, each about 40 Å long and situated between residues 179 and 241, protrude out of the catalytic polar domain, shown in Fig. 1c. Consisting primarily of hydrophobic residues, the helices probably insert into the lipid bilayer. Side chains of Arg184 and Lys183, the only two charged residues on the helices, are in the proximity to interact with surface phosphate moieties. The presence of Pro189 nearby causing a bend in helix 8 could help these side chains to orient toward the surface polar groups. Other polar residues in the domain are Gin121, Thr126, Thr129, and Asn204 on helix 8, a free sulfahydryl group Cys205 on the last turn of the helix, His210 on the short β-stretch connecting the two helices, and Ser218 and Thr229 on helix 9. Another proline residue, Pro212, is located at the N terminus of helix 9. The two helices interdigitate tightly with each other through hydrophobic interactions among leucine, isoleucine, valine, alanine, and phenylalanine side chains, burying 1220 Å² of accessible surface between them, about one-fifth of the combined surface area of two helices. In analogy with previously described packing topology of parallel dimeric transmembrane helices (32, 33), antiparallel helices 8 and 9 have a right-handed crossover with a crossing angle of about −21°, similar to −40° observed previously (33). Near the opening to the active site cleft, a constellation of large hydrophobic side chains, Phe178, Phe182, Phe187, Phe230, Phe233, Tyr236, and Phe237 from the transmembrane domain and Phe104, Tyr193 (from a helical turn 490–493), Trp155, Phe553, and Leu556 (from the one-turn α-helix 17; residues 550–553), and Trp555 of the catalytic domain line the surface of a hydrophobic “tunnel” leading to the active site (Fig. 1c). The Arg200 and Thr203 side chains position themselves as gatekeepers to the “tunnel” (Fig. 4b). Further above, shown at the top of Fig. 1c, the disulfide bridge between Cys170 and Cys242 at the beginning and the end of the transmembrane helices, respectively, forms a “zipper lock” that probably helps stabilize the helix-helix packing.
Three two-helix bundles from three symmetry-related molecules pack about the crystallographic 3-fold axis, concealing one side of the hydrophobic surface between residues 195–208 of helix 8 and residues 216–232 of helix 9. Residues from the section 198–208 of helix 8 also interact with Trp555, Pro557, and Trp558 side chains from the C-terminal loop of a neighboring molecule. This intermolecular hydrophobic interaction, which conceals about 890 Å² of accessible surface of each helix pair may in part be responsible for the observed dynamic stability of the helices in the crystal structure. One ordered BOG molecule (shown in Fig. 4c) was located on the other side of each helix pair, partially shielding them from exposure to the solvent. The O3 atom of the pyranose ring of BOG is situated at a hydrogen bonding distance from the hydroxyl of the Ser218 side chain, whereas the hydrophobic tail rests at a van der Waals contact distance from the Leu222, Ile226, and Pro189 side chains.

**DISCUSSION**

**The Catalytic Mechanism**—As described under “Results,” 9 of 10 catalytically important residues are identical in all three mammalian sulfatases. The 10th residue is Gln343, as opposed to asparagines in ASA and ASB. Therefore, the catalytic mechanism by which E1 sulfate and DHEA sulfate is hydrolyzed by ES into sulfate and free steroids, in all likelihood, is quite similar to those proposed for ASA and ASB (29, 30). The cation found in ASA was a Mg²⁺/H⁺; however, in ASB, as in ES, the metal density was assigned to a Ca²⁺. Interestingly, the catalytic formylglycine in ASB was also found as a sulfate derivative (30). Our results clearly demonstrate that the catalytic formylglycine 75 is a hydroxylformylglycine and that it is covalently linked to a sulfate moiety. It is thus likely that the catalytic residue as a sulfated hydroxylformylglycine is the resting state of the enzyme, as has been suggested for ASB (30). It is the last intermediate step of the four-step mechanism described for all sulfatases as well as for alkaline phosphatase (34). According to the scheme described for ES in Fig. 5, therefore, the HSO₄⁻ moiety is released in the first step in a nucleophilic attack by a water molecule. This reaction also serves as the process of activation of the enzyme to prepare it for the next cycle of catalysis. In analogy with the mechanism described for other sulfatases, several positively charged side chains, Lys134, Lys368, and Arg79, as well as His136 and His290 participate in catalysis in addition to their role in charge neutralization inside the active site cavity. Nucleophilic attack on the sulfur atom by one of the hydroxyls of the hydroxylformylglycine follows its activation by Ca²⁺, whereas the other hydroxyl is deprotonated by His136. This causes covalent linkage of the sulfate moiety with the formylglycine side chain and release of the free E1.

**Evidence for Transmembrane Association**—A pair of hydrophobic helices each consisting of 14–23 amino acids has been shown to form a right-handed transmembrane dimer (32, 33). Helices 8 and 9 in ES, each containing about 25–30 residues, roughly 40 Å long (Fig. 1c) and situated between residues 179 and 241 thus presumably traverse the membrane and anchor the functional domain to the membrane surface facing the ER lumen. The location of two charged side chains, Lys183 and
Arg184, near the polar surface on the lipid bilayer is also consistent with the proposed transmembrane insertion of a two-helix bundle. The presence of lysine residues at interfacial polar locations of transmembrane helices has been previously observed (35). Due to the presence of several disulfide groups and four glycosylation sites, it is highly unlikely that the polar catalytic domain of ES is situated in the reducing environment of the cytoplasm. Therefore, the polar catalytic domain rests on the lumen side of the lipid bilayer.

In addition to helices 8 and 9, two other hydrophobic regions from SD2 could possibly have membrane association, as suggested by their location, hydrophobicity, and high thermal mo-
tion. High B-factors could be a direct consequence of their removal from the lipid bilayer and exposure of hydrophobic residues to a more polar environment. In general, SD2 has a significantly higher average isotropic temperature factor (91.7 Å²) than SD1 (51.9 Å²). Two polypeptide segments, Phe 548–Leu 568 (including two one-turn helices, 17 (residues 550–553) and 18 (residues 557–559)) that is spatially proximal to the active site groove and Phe 468–Pro 500 (including one helical turn 490–493) that consists of several hydrophobic side chains and displays high thermal motion, presumably associate with the lipid bilayer (Figs. 2a and 6a). Three tryptophan side chains, Trp 555, Trp 558, and Trp 560, from the first segment position themselves at the lipid bilayer interface (Fig. 1c). The existence of tryptophan side chains at the polar face of the bilayer near the lipid carbonyl groups has been noted (35). In addition, the SD1 side chains Arg 98 and Lys 353 that are from two loops that approach the lipid bilayer may also interact with the phosphate moieties (Fig. 6b). Interestingly, each of these residues belongs to one of two regions having 4–7-residue peptide insertions when compared with the structures of the soluble ASA and ASB (Fig. 3a). The Arg 98 side chain, in particular, as well as Thr 99, may have contacts with the 17α-hydroxyl (or 17-keto) end of the substrate (Fig. 4b) and therefore may play significant roles in substrate recognition.

**Passage to and from the Active Site**—The proposed transmembrane architecture of ES and the relative location of the active site suggest a number of possibilities for the passage of a steroidal substrate to and from the active site. Fig. 6b highlights three flexible loops that may conceivably open to allow a
above could also serve as the “left swing door” to and from the active site.

It is, however, possible that the substrate and the product pass through the ER membrane to enter or leave the active site. Similar mechanisms have been proposed for transport of substrates for the membrane-associating enzymes fatty acid amide hydrolase (36), squalene cyclase (37), and prostaglandin H₂ synthase (38), all of which use lipophilic substrates and have an active site close to the lipid bilayer surface. Although the steroid is highly hydrophobic, the polar sulfate moiety may require shielding, either by counterions or solvation, to cross the membrane. It is known that the bilayer interface is hydrated and contains polar molecules (39). Transmembrane helices 8 and 9 pack tightly to each other from one end of the bilayer to the other, creating a large opening to the “gill” of the “mushroom” that is lined with several aromatic residues (Phe₁⁰⁴, Phe₁⁷₈, Phe₁₈₂, Phe₃₃₃, Trp₅₅₀, and Phe₁₆₅₃) and other hydrophobic side chains (Leu₁₅₈, Leu₂₉₉, and Leu₅₅₄) (Figs. 1c and 4b). In this way, a smooth hydrophobic surface traversing the bilayer is created on the two-helix bundle, shown in Fig. 6a, that contains only two charged residues (Lys₁₅₂ and Arg₁₆₄). This hydrophobic “tunnel” connects the active site to the interior of the lipid bilayer, as depicted in Fig. 6a.

The ER membrane is the site for biosynthesis of transmembrane as well as soluble proteins. Some of the most important steroidogenic enzymes, such as cytochrome P450 aromatase and 17HSD1, are known to be bound to the ER membrane. The ES molecule may, in fact, be embedded in the ER membrane at a translocator pore or a receptor that facilitates the passage of sulfated steroids directly into the active site at the lumen face of the membrane. Hydrolyzed steroids are then released into the lumen, possibly via one of the “swing doors” described above, for distribution to targeted organelles such as the nucleus and the mitochondria for use in hormone-dependent signaling. Another plausible scenario is that sulfated steroids enter the ES active site from the lumen side through the “swing doors,” and the continuum of the lipid bilayer of the ER, acting as a solvent for highly hydrophobic steroidal molecules, could be involved in dispersing products of ES-catalyzed hydrolysis to other organelles. Steroid carrier proteins, such as sex hormone-binding globulin, could also have a role in this process.

Conclusion—The ES crystal structure elucidates the molecular architectures of the active site and putative transmembrane regions. This structure provides clues to the mode of action, but it also raises questions regarding functional implications of the membrane association. These questions are relevant not only to ES but also perhaps to other ER membrane-bound and activated enzymes. Future work with site-directed mutants of the recombinant enzyme, both in transfected cells and in purified solutions, should address some of these issues.

Acknowledgment—We thank Dr. Charles M. Weeks for critically reading the manuscript and making valuable suggestions.

REFERENCES
1. Reed, M. J., and Purohit, A. (1997) Endocrin. Rev. 18, 701–715
2. Ghosh, D., and Vihko, P. (2002) Chem.-Biol. Interact. 130–132, 657–650
3. Pasqualini, J. R., and Chetrite, G. S. (2002) Breast Cancer: Prognosis, Treatment and Prevention (Pasqualini, J. R., ed) pp. 187–249, Marcel Dekker, New York, NY
4. Billiach, A., Nusbaumer, P., and Lehr, P. (2000) J. Steroid Biochem. Mol. Biol. 73, 225–235
5. Hernandez-Martín, A., Gonzalez-Sarmiento, R., and De Unamuno, P. (1999) Br. J. Dermatol. 141, 612–627
6. Alperin, E. S., and Shapiro, L. J. (1979) J. Biol. Chem. 254, 20756–20763
7. Vaccaro, A. M., Salvioli, R., Muscillo, M., and Renola, L. (1987) Enzyme 37, 115–126
8. Hernandez-Guzman, F. G., Higashiya, T., Osawa, Y., and Ghosh, D. (2001) J. Steroid Biochem. Mol. Biol. 78, 441–450
9. Stein, C., Hille, A., Seidel, J., Rijndijk, S., Waheed, A., Schmidt, B., Gueze, H., and von Figura, K. (1989) J. Biol. Chem. 264, 13865–13872
10. Noel, H., Plante, L., Bleau, G., Chapdelaine, A., and Roberts, K. D. (1983) J. Steroid Biochem. 19, 1591–1598
11. Dibbelt, L., and Kuss, E. (1986) *Biol. Chem. Hoppe-Seyler* **367**, 1223–1229
12. Burns, G. R. J. (1983) *Biochim. Biophys. Acta* **759**, 199–204
13. Van der Loos, C. M., Van Breda, A. J., Van den Berg, F. M., Walboomers, J. M. M., and Jobis, A. C. (1985) *J. Inherited Metab. Dis.* **8**, 97–103
14. Suzuki, T., Hirato, K., Yanaihara, T., Kido, M., Satoh, T., Hoshino, M., and Yanaihara, N. (1992) *Endocrinol. Jpn.* **39**, 93–101
15. Pursuit, A., Potter, B. V. L., Parker, M. G., and Reed, M. J. (1998) *Chem.-Biol. Interact.* **109**, 183–193
16. Van der Loos, C. M., Van Breda, A. J., Van den Berg, F. M., Walboomers, J. M. M., and Jobsis, A. C. (1983) *J. Inherited Metab. Dis.* **7**, 97–103
17. Suzuki, T., Hirato, K., Yanaihara, T., Kodofuku, T., Sato, T., Hoshino, M., and Yanaihara, N. (1992) *Endocrinol. Jpn.* **39**, 93–101
18. Purohit, A., Potter, B. V. L., Parker, M. G., and Reed, M. J. (1998) *Chem.-Biol. Interact.* **109**, 183–193
19. Yanaihara, A., Yanaihara, T., Toma, Y., Shimizu, Y., Saito, H., Okai, T., Higashiyama, T., and Osawa, Y. (2001) *Steroids* **66**, 87–91
20. Reed, M. J., Purohit, A., Woo, L. W. L., and Potter, B. V. L. (1996) *J. Endocrinol. Endocr. Rel. Cancer* **3**, 9–23
21. Leslie, A. G. W. (1990) *Crystallographic Computing*, Oxford University Press, Oxford, UK
22. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
Structure of Human Estrone Sulfatase Suggests Functional Roles of Membrane Association
Francisco G. Hernandez-Guzman, Tadayoshi Higashiyama, Walter Pangborn, Yoshio Osawa and Debasis Ghosh

J. Biol. Chem. 2003, 278:22989-22997.
doi: 10.1074/jbc.M211497200 originally published online March 25, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M211497200

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

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

This article cites 38 references, 8 of which can be accessed free at
http://www.jbc.org/content/278/25/22989.full.html#ref-list-1