Proposed active site domain in estrogen sulfotransferase as determined by mutational analysis

WILLIAM J. DRISCOLL, KATSUTOSHI KOMATSU*, AND CHARLES A. STROTT†

Section on Steroid Regulation, Endocrinology and Reproduction Research Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-4510

Communicated by Seymour Lieberman, St. Luke's Roosevelt Hospital Center, New York, NY, September 28, 1995 (received for review July 20, 1995)

ABSTRACT Point mutations were selectively introduced into a cDNA for guinea pig estrogen sulfotransferase (gpEST); each construct was then expressed in Chinese hamster ovary K1 cells. The molecular site chosen for study is a conserved GXXGXXK sequence that resembles the P-loop-type nucleotide-binding motif for ATP- and GDP-binding proteins and is located near the C terminus of all steroid and phenol(aryl) sulfotransferases for which the primary structures are known. Preliminary experiments demonstrated that the GXXGXXK motif is essential for binding the activated sulfonate donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS). The present study was undertaken to ascertain the relative importance of each individual residue of the motif. While the mutation of a single motif residue had little effect on the interaction between gpEST and PAPS as determined by kinetic analysis and photoaffinity labeling, the mutation of any two residues in concert resulted in an approximate 10-fold increase in the Km for PAPS and reduced photoaffinity labeling. The mutation of all three motif residues resulted in an inactive enzyme and complete loss of photoaffinity labeling. Interestingly, several mutants also displayed a striking effect on the Km for the steroid substrate; double mutants, again, demonstrated greater perturbations (8- to 28-fold increase) than did single mutants. Unexpectedly, whereas the mutation of nonmotif residues had a negligible effect on the Km for PAPS, a marked increase in the Km for the estrogen substrate (>30-fold) was noted. On the basis of these findings, it is concluded that the sequence GISGDWKFN within the C-terminal domain of gpEST represents a critical component of the active site.

The biotransformation of compounds by sulfonation occurs widely, effecting a marked change in the physicochemical properties of the sulfonated products. Sulfonation of drugs and xenobiotics functions primarily to inactivate and clear these generally hydrophobic compounds, although there are examples where the active form of a drug is sulfonated (1). Likewise, the sulfonation of endogenous substances is a fundamental metabolic mechanism that covers an extended range of molecules of diverse size, structure, chemistry, and function. Membrane and secretory proteins are post-translationally modified by sulfonation. Macromolecules such as glycosaminoglycans and proteoglycans, components of cell surface and connective tissue structures, are subjected to modification by the addition of sulfonate groups. Importantly, a multitude of low molecular weight compounds, including neurotransmitters and hormones—e.g., catecholamines, isothiocyanates, and steroids—are modified by sulfonylation.

Inherently, the enzymes that carry out the transfer of a sulfonate group (SO₃⁻), termed sulfotransferases, must interact with two substrates, namely, the sulfonate donor and acceptor molecules. In mammals, the universal activated sulfonate donor is 3′-phosphoadenosine 5′-phosphosulfate (PAPS) (1). We recently published a preliminary report providing evidence that a GXXGXXK sequence, related to the P-loop motif found in ATP- and GTP-binding proteins (2), is involved in binding the cofactor PAPS by estrogen sulfotransferase (EST; EC 2.8.2.4) from guinea pig (gpEST) (3). This motif is positioned within a highly conserved region located near the C termini of all steroid and phenol(aryl) sulfotransferases for which primary structures are known (3). Interestingly, the first glycine in the GXXGXXK motif is located at the 5′-splice site for the 3′-terminal exon of gpEST (exon 8), a pattern consistent for all steroid sulfotransferase gene structures cloned to date (4).

To examine the GXXGXXK motif in greater detail, point mutations that replaced the conserved glycines (G260 and G263) and lysine (K266) with alanines were introduced singly or in combination into a gpEST cDNA that had been cloned in our laboratory and extensively characterized (5, 6). Evidence is presented indicating that, in addition to the critical role played by the GXXGXXK motif in binding the sulfonate donor, PAPS, there are residues, specifically glycine-260 and asparagine-267, that also have profound effects on the interaction between enzyme and estrogen substrate. These results suggest that this conserved region of gpEST, encoded by the 3′-terminal exon, interacts with both cofactor and steroid substrate and thus constitutes, at least in part, the active site where sulfonate is transferred from donor to recipient.

MATERIALS AND METHODS

Mutant cDNA Constructs. Wild-type EST cDNA in pSV/SPORT-1 (GIBCO/BRL) (3) was used as template to generate mutant constructs by the recombinant circle polymerase chain reaction (RCPCR) technique (7). Synthetic oligonucleotide primers were designed to introduce point mutations into the PCR products that would result in the substitution of alanine at the desired position in the expressed proteins. Appropriate mutations in each construct were verified by dideoxy nucleotide chain-termination sequencing (Sequenase version 2.0, United States Biochemical), and accurate vector replication by PCR was confirmed by reversal of the inactive G₁+G₂+K→A mutant back to active wild type by using the identical RCPCR protocol.

Cell Culture and Transfection. Chinese hamster ovary K1 (CHO-K1) cells were grown in 75-cm² culture flasks and, at approximately 70% confluency, were transfected with 15 μg of plasmid DNA by using 100 μl of Lipofectamine (GIBCO/BRL) in accordance with the manufacturer's instructions.

Abbreviations: PAPS, 3′-phosphoadenosine 5′-phosphosulfate; EST, estrogen sulfotransferase; gpEST, guinea pig EST; RCPCR, recombinant circle polymerase chain reaction; estrone, 3-hydroxy-1,3,5(10)-estratrien-17-one.

*To whom reprint requests should be addressed: National Institutes of Health, Building 49, Room 6A36, Bethesda, MD 20892-4510.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
BRL) in a total volume of 11 ml of serum-free medium according to the supplier's instructions. Cells were harvested 48 hr after transfection, and cytosol was prepared as described previously (5). Protein concentrations were estimated by the method of Lowry et al. (8), using bovine serum albumin as the standard.

**Immunoblotting.** Cytosolic proteins (10–40 µg) were resolved by SDS/PAGE in a 15% polyacrylamide gel as described by Laemmli (9) and transferred to nitrocellulose membranes (Schleicher & Schuell) (10). Standard immunoblotting was performed as previously described, using a specific EST polyclonal antibody at a dilution of 1:4000 (11).

**EST Assay.** Assays for EST activity were performed in 250 µl of assay buffer (100 mM Tris-HCl, pH 7.7/5 mM magnesium acetate) containing various concentrations of [3H]estrone (43.5 Ci/mmol; New England Nuclear; 1 Ci = 37 GBq) and PAPS (Sigma) as described previously (12). Concentrations of each substrate were varied to accommodate the specific mutant construct being analyzed and the kinetic parameter being measured. Data points represent the average of duplicate determinations. The program ENZFIT (Elsevier Biosoft) was used to calculate kinetic constants and plot the data.

**Photoaffinity Labeling.** Cytosols were diluted in assay buffer and incubated on ice with ~1 µCi of [35S]PAPS (2.9 µM) (New England Nuclear) in a total volume of 150 µl for 30 min. Samples were kept on ice and irradiated with a UVS*12 Mineralight (Ultra-violet Products) for 15 min at a distance of <1 cm. As controls, additional aliquots of wild-type cytosol either were not irradiated or were incubated as described above in the presence of 290 µM additional nonradioactive PAPS. Samples were denatured with an equal volume of SDS/PAGE sample diluent, boiled for 5 min, and loaded onto an SDS/15% polyacrylamide gel (20 µg of protein per lane). Resolved proteins were blotted to nitrocellulose and the dried membrane was exposed to x-ray film for 6–20 days at ~80°C.

**RESULTS AND DISCUSSION**

Nine mutant gpEST constructs were generated from the wild-type cDNA by the RCP PCR technique and are detailed in Table 1. In the notation used, G1 and G2 refer to the first and second glycines of the GXXGXXK motif, respectively. The mutant constructs consist of the GXXGXXK motif residues (G260, G263, and K266) singly mutated to alanine as well as combinatorial mutations of the motif residues. Additionally, as “controls,” nonmotif residues S262, D264, and N267 were also mutated to alanine as indicated in Table 1. The correctness of each mutation was verified by DNA sequencing of the appropriate region of the cDNA (data not presented).

Expression of gpEST protein by the mutant cDNA constructs was verified by immunoblot analysis employing a highly specific anti-gpEST polyclonal antibody (11). It is evident, as demonstrated in Fig. 1, that all nine mutant constructs expressed the 34-kDa gpEST protein. The failure of immunoblot analysis to detect any immunoreactive protein in the cytosol of nontransfected CHO-K1 cells (Fig. 1, lane 10) is in keeping with the lack of detectable endogenous EST activity by these cells (6). The slightly faster migration in SDS/PAGE of the G1+G2+K->A construct (Fig. 1, lane 8) is not clearly understood, but it was reproducible and could conceivably represent an altered conformation compared with the other constructs. Although proteolytic truncation cannot be definitively ruled out, the fact that the S+D+N->A triple mutation was not similarly affected suggests that this is probably not the explanation. Additionally, the G1+G2+K->A mutant cDNA construct was sequenced in its entirety and found to be complete.

Kinetic analyses were performed to detect perturbations in the interaction between PAPS and the mutant gpEST constructs compared to the wild type (Fig. 2 and Table 2). Because K_m reflects a complex ratio of individual rate constants, the measurement is used here only to reflect changes in the overall catalytic efficiency of the expressed enzyme. No attempt has yet been made to determine whether ligand affinity or kcat or both might be affected by specific mutations. The data presented in Fig. 2 demonstrate that all mutant constructs except the triple motif mutant G1+G2+K->A were expressed as active enzymes. In repeated trials, this construct failed to demonstrate detectable EST activity, although immunoreactive gpEST protein was adequately expressed as noted above (cf. Fig. 1). The mutation of any single GXXGXXK motif residue failed to cause an alteration in the K_m for PAPS compared with the wild-type enzyme. For the G1->A and G2->A mutant constructs, a negligible effect on PAPS K_m can be understood on the basis of these being relatively conservative structural changes; however, it was surprising that the K->A mutation did not affect the PAPS K_m, as it might be predicted that the positively charged amino acid would be instrumental in interacting with the negatively charged cofactor. Each doubly mutated construct involving motif residues produced enzymes exhibiting K_m values for PAPS that were approximately 10-fold higher than wild type, thus indicating that structural perturbations affecting the catalytic efficiency of gpEST had occurred. On the other hand, the K_m for PAPS was unchanged from wild type for either the double (S+D->A) or triple (S+D+N->A) “control” mutants, demonstrating the singular importance of motif residues specifically.

Table 1. EST mutant constructs

<table>
<thead>
<tr>
<th>Construct</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>G I S G D W K N</td>
</tr>
<tr>
<td>G1-&gt;A</td>
<td>A I S G D W K N</td>
</tr>
<tr>
<td>G2-&gt;A</td>
<td>G I S A D W K N</td>
</tr>
<tr>
<td>K-&gt;A</td>
<td>G I S G D W A N</td>
</tr>
<tr>
<td>G1+G2-&gt;A</td>
<td>A I S A D W K N</td>
</tr>
<tr>
<td>G1+K-&gt;A</td>
<td>A I S G D W A N</td>
</tr>
<tr>
<td>G2+K-&gt;A</td>
<td>G I S A D W A N</td>
</tr>
<tr>
<td>G1+G2+K-&gt;A</td>
<td>A I S A D W A N</td>
</tr>
<tr>
<td>S+D-&gt;A</td>
<td>G I A G A W K N</td>
</tr>
<tr>
<td>S+D+N-&gt;A</td>
<td>G I A G A W K A</td>
</tr>
</tbody>
</table>

Amino acid sequence is given for residues 260–267 of gpEST. The residues mutated from the wild type are underlined.
residues within the region of the GXXGXKK motif fail to affect the $K_m$ for PAPS.

To gain some insight into why the substitution of alanine for lysine-266 might have no effect on the $K_m$ for PAPS, we examined the nucleotide-binding site for the human p21 ras protein (13, 14). This GTP/GDP-binding protein has a P-loop GXXGXKK motif located near its N terminus and was used as a model structure because there are no crystallographic data yet available for any sulfotransferase. Computer-assisted examination of p21 ras suggests that the positive charge on arginine-258 in gpEST could be positioned so as to effectively compensate for the loss of lysine-266 when it is mutated to alanine, thus leaving the $K_m$ for PAPS unaffected. In contrast to this study, when the conserved P-loop lysine in p21 ras was mutated, the affinity for guanine nucleotides was significantly compromised (15); however, and most importantly, the residue in p21 ras that would correspond to arginine-258 in gpEST is a valine, which would not be able to provide ionic or hydrogen bonding potential.

The interaction between PAPS and gpEST was also evaluated by photoaffinity labeling using [35S]PAPS (Fig. 3). It has been previously shown that [35S]PAPS is an efficient photoaffinity ligand for sulfotransferases (16). CHO-K1 cells expressing wild-type gpEST demonstrated radioactive labeling of an appropriate 34-kDa protein with [35S]PAPS (Fig. 3, lane 1). Labeling was dependent on UV irradiation (Fig. 3, lane 2) and was inhibited by the presence of a 100-fold excess of nonradioactive PAPS (Fig. 3, lane 3), thus demonstrating specificity for the cofactor. Furthermore, nontransfected CHO-K1 cells failed to demonstrate any labeling of cytosolic proteins (Fig. 3, lane 4). It is evident that the mutant constructs exhibited variable reactivity for photoaffinity labeling and were less efficiently labeled than wild type, a finding consistent for multiple experimental trials. Cytosols prepared from cells transfected with constructs containing singly mutated GXXGXKK motif residues or "control" mutant constructs involving nonmotif residues (Fig. 3, lanes 6–9 and 13) exhibited labeling of a 34-kDa protein. More specifically, constructs containing nonmotif residue mutations and the K→A substitution mutant were more efficiently labeled than either of the single-motif mutants involving glycine residues. Of the three double-motif mutants, only the G2+K→A construct demonstrated significant labeling (Fig. 3, lane 12), while cytols prepared from cells transfected with double-mutant constructs containing G1→A substitutions demonstrated only negligible

Collectively, these results indicate that the GXXGXKK motif is an important element for PAPS/gpEST interaction, albeit there appears to be some structural tolerance in that the substitution of alanine for any single residue of the motif has little affect on the $K_m$ for PAPS. It is only with double mutations involving motif residues that a significant kinetic effect is evident, and, most importantly, the triple motif mutant, G1+G2+K→A, results in a completely inactive protein. The structural significance of the GXXGXKK motif in PAPS/gpEST interaction gains additional support from the fact that double and triple mutations involving nonmotif

---

**Table 2. $K_m$ values for PAPS**

<table>
<thead>
<tr>
<th>Construct</th>
<th>$K_m$, $\mu$M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type*</td>
<td>2.5 ± 0.2 (5)</td>
</tr>
<tr>
<td>G1→A</td>
<td>2.5 ± 0.2 (3)</td>
</tr>
<tr>
<td>G2→A*</td>
<td>2.6 ± 0.3 (3)</td>
</tr>
<tr>
<td>K→A*</td>
<td>2.6 ± 0.1 (3)</td>
</tr>
<tr>
<td>S→D→A*</td>
<td>2.6 ± 0.2 (3)</td>
</tr>
<tr>
<td>S→D→N→A‡</td>
<td>2.5 ± 0.2 (4)</td>
</tr>
<tr>
<td>G1+G2→A‡</td>
<td>28.4 ± 1.3 (3)</td>
</tr>
<tr>
<td>G1+K→A‡</td>
<td>34.6 ± 0.6 (3)</td>
</tr>
<tr>
<td>G2+K→A*</td>
<td>27.3 ± 1.3 (4)</td>
</tr>
<tr>
<td>G1+G2+K→A‡</td>
<td>ND</td>
</tr>
</tbody>
</table>

Constructs are described in Table 1. Results are given as mean ± SEM for the number of trials shown in parentheses. ND, not determinable.

* Assays were performed at 2 $\mu$M estrone.
‡ Assays were performed at 5 $\mu$M estrone.
‡ Assays were performed at 10 $\mu$M estrone.

---

**Fig. 2.** Activity analysis for EST cDNA constructs expressed in CHO-K1 cells. Cytosols prepared from CHO-K1 cells transfected with wild type and the nine EST cDNA mutant constructs as indicated were assayed for EST activity at various concentrations of PAPS. (A) Constructs assayed at 0–10 $\mu$M PAPS. (B) Constructs assayed at 0–20 (S+D+N→A) or 0–50 $\mu$M PAPS. Data points represent averaged duplicate determinations. Graphical curve fitting was performed by using the EZFITTER program.

**Fig. 3.** Photoaffinity labeling of CHO-K1 cell cytosols with [35S]PAPS. CHO-K1 cellular cytosols were prepared and photoaffinity labeled with [35S]PAPS. Lane 1, transfection with wild-type EST cDNA; lane 2, wild-type transfection without ultraviolet irradiation; lane 3, wild-type transfection in the presence of 100-fold molar excess of nonradioactive PAPS; lane 4, nontransfected CHO-K1 cell cytosol; lanes 5–13, transfection with the EST cDNA mutant constructs as indicated in the figure. Total cytosolic protein loaded was 20 $\mu$g per lane.
glycine residues, in particular G1, significantly alter the environment of the PAPS binding site so that photoaffinity labeling of the protein is much less efficient. Additionally, the fact that single-residue and "control" mutations profoundly reduce photoaffinity labeling compared with wild type suggests that these mutations, while not affecting Km measurements, caused an alteration in the spatial relationship between the ligand and the amino acid residue(s) that is covalently modified by the photoaffinity labeling. It should be noted that these studies do not attempt to identify which amino acid(s) in the protein is modified by this procedure (although it is of considerable interest and a subject for future investigation), and we do not mean to suggest that a residue within or adjacent to the motif is in fact the site of photoaffinity labeling. It is quite possible, due to protein folding and the inherent three-dimensional nature of a binding pocket, that a residue far removed in linear sequence from the GXXGXXK motif could serve as the site for covalent attachment. In fact, Zheng and colleagues (17) found that chemical affinity labeling of aryl sulfotransferase IV (EC 2.8.2.9) by an ATP analog resulted in the covalent attachment of the ligand to lysine-65 and cysteine-66, which are located in the N-terminal portion of the protein. While it can be reasonably concluded from the data presented that these residues are likely in proximity to the PAPS-binding site, this does not address their functional significance and, additionally, the authors clearly note that the sequence surrounding these residues as well as the residues themselves are not highly conserved among sulfotransferases. Studies such as those presented here will be necessary to determine whether this domain possesses structural importance for catalytic activity or simply served as a favorable site for covalent attachment of the nucleotide analog that was used. Importantly, as previously noted for all steroid and phenol(aryl) sulfotransferases, aryl sulfotransferase IV contains the conserved GXXGXXK motif in its C-terminal domain between residues 246 and 252.

Kinetic analyses were also performed to determine the Km values for the steroid substrate, estrone, using cytosols from CHO-K1 cells expressing the various cDNA constructs (Table 3). Because we initially believed that the GXXGXXK motif was specifically involved with PAPS binding, we anticipated that the mutant constructs would express negligible kinetic effects for the steroid substrate, as the binding domain for this molecule was expected to reside at a unique site. Unexpectedly, several mutant constructs dramatically increased the Km for estrone in a pattern that was pointedly different from that found for PAPS. Specifically, while no single-mutant construct significantly altered the Km for PAPS, the G1→A mutant resulted in an impressive 13-fold increase in the Km for the steroid substrate (Table 3). The other single-mutant constructs, G2→A and K→A, only modestly increased the Km for estrone. Additionally, whereas the double-motif mutant not involving the G1 residue resulted in an 8-fold increase in the Km for estrone, the two double mutants that did involve G1→A substitutions produced a >28-fold increase in the Km for the estrogen substrate. Furthermore, it should be recalled that photoaffinity labeling experiments also implicated the G1 residue as structurally significant. These results indicate that the G1 residue of the GXXGXXK motif plays an especially important role in the interaction between the steroid substrate and the enzyme.

Another unexpected finding that arose from these kinetic analyses was that the mutation of nonmotif "control" residues led to significantly elevated Km values for estrone. Whereas neither the double S+D→A nor the triple S+D+N→A mutant caused an alteration in the Km for PAPS, both constructs increased the Km for the steroid substrate (Table 3). The S+D→A construct led to a modest but significant ≈2.5-fold increase in the Km for estrone, whereas the S+D+N→A construct resulted in a pronounced increase of greater than 30-fold. The difference between these two constructs, of course, is the additional substitution of alanine for asparagine-267 in the latter construct, thus suggesting a critical role for this residue in the catalytic efficiency of gpEST. Mechanistically, it is possible that this asparagine could interact with the 3-hydroxyl on the steroid A ring through hydrogen bonding. Additionally, it is interesting that this asparagine is conserved for all steroid sulfotransferases, whereas threonine is found at this position in phenol(aryl) sulfotransferases (4). We previously reported the PAPS Km for the S+D+N→A construct to be approximately 3-fold greater than wild type (3). That analysis, however, was performed at a limiting concentration of estrone, before it was realized that this mutation significantly elevated the Km for estrone, and has been corrected in this study.

In summary, evidence is presented indicating that G260, G263, and K266, amino acid residues that constitute the GXXGXXK P-loop motif, play a critical role in the interaction between gpEST and the universal sulfonate donor, PAPS. Additionally and importantly, there are residues, specifically G260 and N267, that have profound effects on the ability of the enzyme to efficiently sulfonate the estrogen substrate. Thus, on the basis of these findings, it is proposed that the sequence GISGDWK, encoded by the 5′ end of the 3′-terminal exon ( exon 8) represents an essential component of the active site for gpEST and that, insomuch as the GXXGXXK motif is universally conserved, this area will also prove to be catalytically important for all steroid and phenol(aryl) sulfotransferases.

We thank Dr. Byungkook Lee of the Division of Cancer Biology, National Cancer Institute, for conducting the p21 ras computer modeling.