X-ray structure of potato epoxide hydrolase sheds light on substrate specificity in plant enzymes

SHERRY L. MOWBRAY,[1] LISA T. ELFSTRÖM,[2] KERSTIN M. AHLGREN,[1,4]
C. EVALENA ANDERSSON,[3] AND MIKAEL WIDERSTEN[2]

[1]Department of Molecular Biology, Swedish University of Agricultural Sciences, SE-751 24 Uppsala, Sweden
[2]Department of Biochemistry and Organic Chemistry, Uppsala University, SE-751 23 Uppsala, Sweden
[3]Department of Cell and Molecular Biology, Uppsala University, SE-751 24 Uppsala, Sweden

(RECEIVED August 18, 2005; FINAL REVISION April 4, 2006; ACCEPTED April 4, 2006)

Abstract

Epoxide hydrolases catalyze the conversion of epoxides to diols. The known functions of such enzymes include detoxification of xenobiotics, drug metabolism, synthesis of signaling compounds, and intermediary metabolism. In plants, epoxide hydrolases are thought to participate in general defense systems. In the present study, we report the first structure of a plant epoxide hydrolase, one of the four homologous enzymes found in potato. The structure was solved by molecular replacement and refined to a resolution of 1.95 Å. Analysis of the structure allows a better understanding of the observed substrate specificities and activity. Further, comparisons with mammalian and fungal epoxide hydrolase structures reported earlier show the basis of differing substrate specificities in the various epoxide hydrolase subfamilies. Most plant enzymes, like the potato epoxide hydrolase, are expected to be monomers with a preference for substrates with long lipid-like substituents of the epoxide ring. The significance of these results in the context of biological roles and industrial applications is discussed.

Keywords: X-ray crystallography; epoxide hydrolase; active site; *trans*-stilbene oxide; substrate specificity

Epoxide hydrolases (EHs) catalyze the conversion of epoxides to their corresponding diols. The functions of mammalian EHs are diverse, including detoxification of xenobiotics and drug metabolism (Morisseau and Hammock 2005; Newman et al. 2005). Plant EHs have been suggested to play a part in general defense systems by contributing to the synthesis of cutin monomers and anti-fungal chemicals, as well as by participating in stress responses. Epoxide-containing signaling substances such as leukotrienes, epoxyeicosatrienoic acids, and insect juvenile hormone are substrates for specialized EHs. Roles in catabolic pathways of intermediary metabolism in prokaryotes have also been described.

Most EHs belong to the versatile family of α/β-hydrolase fold enzymes (Heikinheimo et al. 1999); these enzymes (EC 3.3.2.3) produce vicinal diols as products. Four EH structures from this superfamily have been reported to date, representing *Agrobacterium radiobacter* EH (ArEH) (Nardini et al. 1999), the soluble EHs from mouse (MmsEH) (Argiriadi et al. 1999, 2000) and human (HssEH) (Gomez et al. 2004), and an enzyme from the fungus *Aspergillus niger* (AnEH) (Zou et al. 2000). The α/β EH fold includes a core domain with a central eight-stranded β-sheet flanked by α-helices, as well as a mainly helical domain that forms a lid over the core, so forming the active

[4]Present address: Department of Medical Sciences, Uppsala University, University Hospital, SE-751 85 Uppsala, Sweden.
Reprint requests to: Sherry Mowbray, Department of Molecular Biology, Swedish University of Agricultural Sciences, Box 590, Biomedical Center, SE-751 24 Uppsala, Sweden; e-mail: mowbray@xray.bmc.uu.se; fax: +46-18-53-69-71.

Abbreviations: EH, epoxide hydrolase; StEH1, epoxide hydrolase 1 from *Solanum tuberosum*; ArEH, epoxide hydrolase from *Agrobacterium radiobacter* AD1; MmsEH, soluble epoxide hydrolase from *Mus musculus*; HssEH, soluble epoxide hydrolase from *Homo sapiens*; AnEH, epoxide hydrolase from *Aspergillus niger*; PEG, polyethylene glycol; RMS, root-mean-square; TSO, *trans*-stilbene oxide.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051792106.

site. Three residues of the α/β domain form a catalytic triad, while two tyrosines of the lid assist in opening the epoxide ring. The distinct catalytic behaviors of mammalian, fungal, and bacterial EHs have been studied extensively.

Potato (*Solanum tuberosum*) expresses four similar EH enzymes, which are regulated in response to environmental and developmental signals (Stapleton et al. 1994). A number of differences between the catalytic properties of one of these (here designated StEH1) and the mammalian EHs were noted earlier (Morisseau et al. 2000). However, these studies, like most others with EHs, focused primarily on racemic mixtures of substrates, making the results difficult to interpret. We recently investigated the catalytic properties of StEH1 with two pure enantiomeric forms of *trans*-stilbene oxide (TSO) (Fig. 1A), showing that this enzyme in fact displays distinct substrate preferences in both binding and catalysis (Elfström and Widersten 2005). Although StEH1, like other plant enzymes, is structurally most similar to the mammalian soluble EHs, the relationships are too distant for those structures to provide insights into how the differences in substrate specificity arise. The lid domains, in particular, which define much of the binding site's shape and chemical character, are very dissimilar; indeed, many published EH sequence alignments appear to contain errors/uncertainties in the lid regions.

In the present study we describe the structure of StEH1, the first plant EH to be reported. The results show the structural basis for the properties of the potato enzymes and allow a greatly improved understanding of other plant EHs. Comparisons to the mammalian and fungal enzyme structures indicate the origins of the observed differences in substrate specificities among the different classes. Such insights are useful both in interpreting existing biological data and in guiding the choice of enzymes for use in applications such as the synthesis of fine chemicals (Archelas and Furstoss 2001).

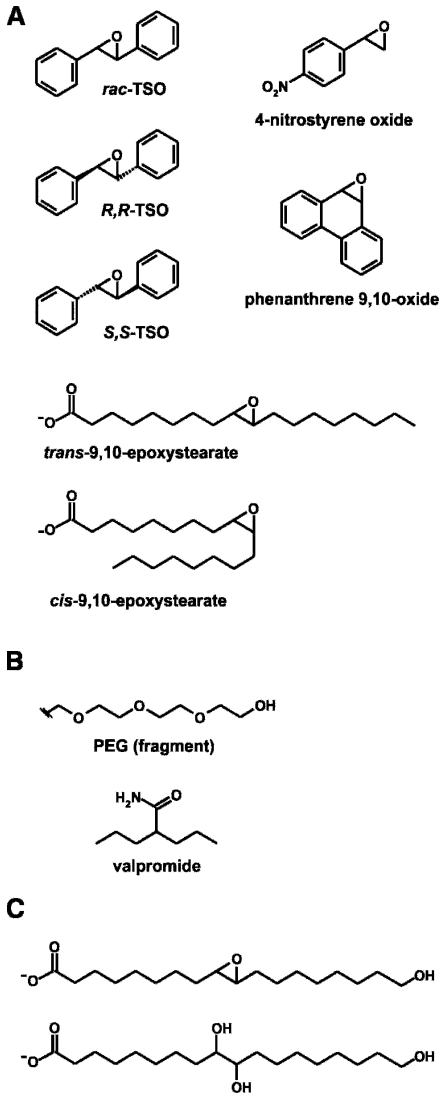

Figure 1. StEH1 ligands. (*A*) Substrates used in in vitro assays as discussed in the text. (*B*) Inhibitors valpromide and PEG. (*C*) Examples of cutin monomer precursor molecules thought to be of significance in vivo (Walton and Kolattukudy 1972).

Results

Overall structure

The structure of StEH1 was solved by molecular replacement and refined to a resolution of 1.95 Å. Statistics for the experimental data and final refined model are summarized in Table 1. Two molecules are found in the asymmetric unit. The final map shows continuous electron density for residues 2–321 of the A molecule and 3–321 of the B molecule; thus, the structure includes all but the very first residue(s) of the native protein. As seen for earlier EH structures, the fold (Fig. 2) consists of two distinct regions: the α/β-hydrolase domain (composed of two segments, residues ~1–140 and 238–321) and the largely helical lid domain (residues ~141–237).

The two molecules of the asymmetric unit are very similar, with an RMS distance of 0.4 Å when all Cα atoms are compared; only residues 11, 94–95, 145–151, 204–210, and 270–272 lie outside of the 0.7 Å cutoff used in an improved structural alignment. Although these differences are well supported by electron density, they are almost certainly related to crystal packing effects. However, movements near residues 150 and 270 are at the active site and are associated with differences in the ligands bound (see below). The two molecules do not form a dimer, nor does a dimer interaction arise through crystallographic symmetry.

Active site

The active site of StEH1 can be identified with confidence, both by analogy with earlier EH structures and

Table 1. *Data collection and refinement statistics*

Data collection statistics	
Data collection beamline/detector	ESRF ID14:2/ADSC scanner
Cell axial lengths (Å)	55.9, 98.5, 122.0
Resolution range (Å)	31.6 -1.95 (2.06 - 1.95)
No. of reflections measured	250,767
No. of unique reflections	49,514
Average multiplicity	5.1 (5.0)
Completeness (%)	99.2 (98.6)
R_{meas} (%)	6.3 (47.1)
$< <I>/\sigma<I> >$	20.9 (3.1)
Refinement statistics	
Resolution range (Å)	31.6 - 1.95
No. of reflections used in working set	47,002
No. of reflections for Rfree calculation	2295
R-value, Rfree (%)	20.3, 24.0
No. of nonhydrogen atoms	5547
No. of solvent waters	405
Mean B-factor, protein atoms, A and B molecules (Å^2)	32.0, 32.6
Mean B-factor, solvent atoms (Å^2)	36.3
Mean B-factor, tetraethylene glycol atoms, A molecule (Å^2)	56.9
Mean B-factor, valpromide atoms, B molecule (Å^2)	31.2
Mean B-factor, ethylene glycol atoms, B molecule (Å^2)	46.8
Ramachandran plot outliers (%)[a]	1.7
RMS deviation from ideal bond length (Å)[b]	0.008
RMS deviation from ideal bond angle (°)	1.06

The space group is $P2_12_12_1$. Information in parentheses refers to the highest resolution shell.
[a] Calculated using a strict-boundary Ramachandran plot (Kleywegt and Jones 1996).
[b] Using the parameters of Engh and Huber (1991).

based on mutagenesis studies (Elfström and Widersten 2005). Asp105 is the catalytic nucleophile, and His300/Asp265 comprise a general base-charge relay pair. Two tyrosine residues from the lid (Tyr154 and Tyr235) are expected to assist ring opening by hydrogen bonding to the oxygen of the substrate's epoxide ring (Elfström and Widersten 2006).

The competitive inhibitor valpromide (Fig. 1B, with a measured K_I of 0.8 mM) was included in the crystallization solutions at a concentration of 5 mM. The clearest electron density in the active site of the A molecule is, however, much more compatible with a molecule of PEG (polyethylene glycol; cf. Figs. 1B and 3A). Subsequent kinetic studies indicated that PEG 10,000 is indeed a competitive inhibitor, with a K_I of 0.4 mM. Since PEG was present at 30 mM in the mother liquor, its binding in the active site was reasonable; this ligand was therefore modeled as tetraethylene glycol. The terminal oxygen of the PEG lies close to the lid tyrosines, while the remainder of the visible density meanders out to the surface of the protein through a largely nonpolar binding cavity. In the B molecule, good electron density is observed for valpromide, together with a smaller segment of PEG; in this case, the PEG density was modeled as ethylene glycol (Fig. 3B). The amide nitrogen of valpromide interacts with the side chain of Asp105 (N–O distance 3.1 Å). The inhibitor's amide oxygen hydrogen bonds to a well-ordered water molecule (Fig. 3B), which in turn interacts with the hydroxyl groups of the two lid tyrosines. Introduction of this water forces a slight shift in the position of the PEG compared with that observed in the A molecule. The electron density suggests that both valpromide and the water molecule are in fact present in the A site as well, but at partial occupancy (Fig. 3A). The somewhat higher temperature factors associated with the tetraethylene glycol ligand in this site (Table 1) are largely attributable to this mixture of binding modes. When valpromide/water are absent, PEG can interact directly with the lid tyrosines; when valpromide/water are present, the PEG must move slightly away. In soaks of crystals in the absence of valpromide, the electron density for the valpromides was lost, but that for PEG remained. Changes in main-chain conformation are associated with the different ligands in the two sites; alternate side-chain conformations, e.g., at Leu266 and Ile270, also illustrate how the protein can adapt to various compounds in its active site.

Predicted modes of substrate binding

The water molecule (shown in Fig. 3B) that hydrogen bonds to the lid tyrosines indicates the expected position of the oxygen in the epoxide ring of bound substrate. Further, the carbons of the epoxide ring must lie near the

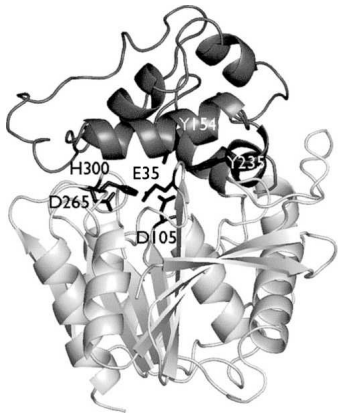

Figure 2. Overall structure of StEH1. The structure of StEH1 is illustrated as a ribbon cartoon. The α/β-hydrolase domain is colored light gray, and the lid is dark gray. Catalytic residues are shown as stick representations in black with labels.

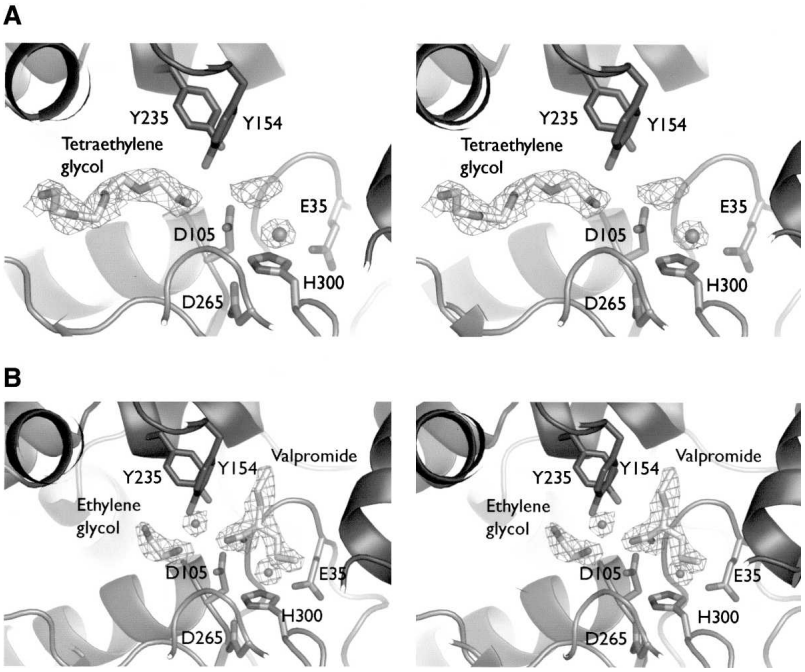

Figure 3. Active sites. Electron densities are shown for bound ligands in the A (*A*) and B (*B*) molecules of the asymmetric unit, using SIGMAA-weighted (Read 1986) $|2F_{obs}-F_{calc}|$ maps contoured at 1σ. Residues expected to participate in StEH1 catalysis are labeled. The putative catalytic water lies between His300 of the charge relay and a nearby acidic residue (Glu35). Additional density is modeled as tetraethylene glycol in the A molecule, and as valpromide and ethylene glycol in the B molecule. Water number 247, with a temperature factor (29 $Å^2$) similar to that of the protein, interacts with both valpromide and the lid tyrosines of the B molecule, and marks the expected position of the epoxide ring's oxygen atom.

attacking carboxylate of Asp105 for catalysis to take place. Such considerations, taken together with the shape of the active site, restrict the possible modes of binding substrates, as illustrated for two enantiomers of TSO in Figure 4. Computer docking with AutoDock3 also suggested a tight clustering of similar poses in both molecules, with <1 Å RMS distance between all atoms in the highest-scoring poses. *S,S*-TSO consistently docked ~1 Å deeper inside the protein. The results did not differ significantly with changes in starting position of the ligand, docking parameters or restraints on bond rotations, or indeed when using the slightly different structures of the A and B molecules. The docking methods are further validated by the fact that valpromide was predicted to bind at a position identical to that observed in the actual crystal structure.

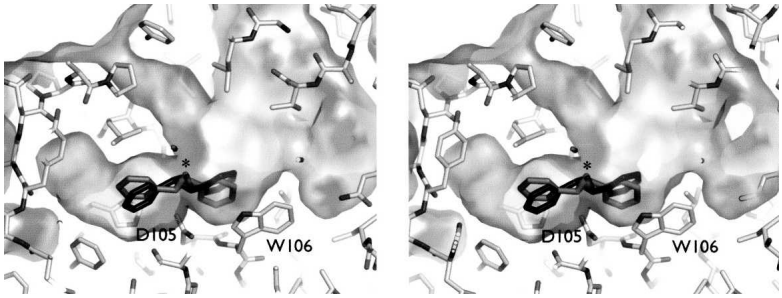

Figure 4. Docked substrates and shape of the StEH1 active site. The expected modes of binding of *S,S*-TSO (light gray) and *R,R*-TSO (black) are compared, with a surface showing the volume of the active site (a 1.2 Å probe was used to calculate accessible surface area). The epoxide oxygen is marked with a star. The "inside" of the binding site is shown at *left*, with the nonpolar branch coming toward the viewer, and the nonpolar one going away. The wider and less polar of the "outside" branches of the active-site cavity extends to the *right*, while a "pipe" of bound waters points upward from the proposed position of the epoxide ring. Trp106 (the likely source of the observed quenching of fluorescence signal when TSO is bound) is labeled, as is the nucleophile.

Table 2. *Comparison of StEH1 with related α/β hydrolase-fold EHs*

Protein (Source)	No. of residues	No. of atoms within 3.5 Å cut-off (% id. of matched)	RMS dist. to StEH1 (Å)	PDB code (Reference)	Equiv. to Asp105 (nucleophile)	Equiv. to His300 (general base)	Equiv. to Asp265 (charge-relay)
HssEH	302[a]	261 (32.2)	1.34	1VJ5 (Gomez et al. 2004)	Asp333	His523	Asp495
MmsEH	310[b]	263 (31.6)	1.40	1EK2 (Argiriadi et al. 2000)	Asp333	His523	Asp495
ArEH	282	228 (30.3)	1.30	1EHY (Nardini et al. 1999)	Asp107	His275	Asp246
AnEH	385	202 (19.2)	1.73	1QO7 (Zou et al. 2000)	Asp192	His374	Asp348

The comparisons reflect the A molecules of the respective asymmetric units.
[a] Using only domain 2, residues 245–547.
[b] Using only domain 2, residues 235–544. Note that residue numbering here follows that of the human enzyme.

Structural and sequence alignments

Structural comparisons are summarized in Table 2. As expected from earlier comparisons based on sequences alone, StEH1 is most closely related to the mammalian sEHs and is less similar to the bacterial enzyme. It is even more distantly related to the fungal enzyme that represents the subfamily containing the human microsomal EH. A structure-based sequence alignment is shown in Figure 5. Not unexpectedly, similarities are greatest in the α/β-hydrolase domain. Despite the structural conservation

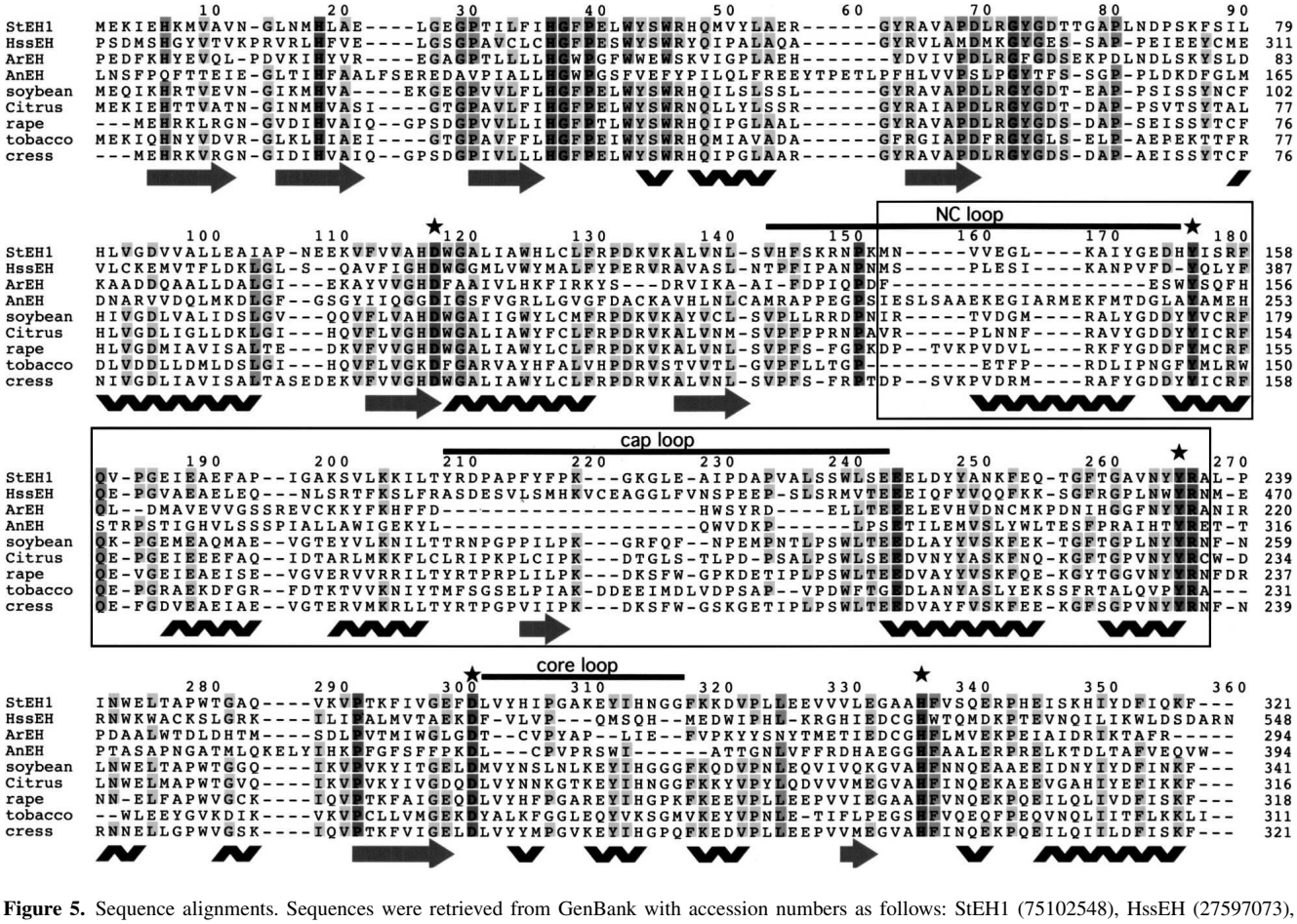

Figure 5. Sequence alignments. Sequences were retrieved from GenBank with accession numbers as follows: StEH1 (75102548), HssEH (27597073), ArEH (2292731), AnEH (6165234), *Glycine max* (soybean, 1304226), *Citrus jambhiri* (rough lemon, 46090792), *Brassica napus* (rape, 22208299), *Nicotiana tabacum* (tobacco, 1354848), and *Arabidopsis thaliana* (cress, 12642902). A structure-based alignment was followed by alignment of additional sequences; shading indicates the degree of sequence conservation. Residues of the lid region are shown boxed, and the NC- and cap loops as defined by Barth et al. (2004), as well as the core loop as defined here, are labeled. Secondary structure elements of StEH1 were assigned using DSSP (Kabsch and Sander 1983), and catalytic residues are marked with stars. The first 233 (and last seven) residues of HssEH, as well as the first five of ArEH, 78 of AnEH, and 25 of the putative transit peptide of the soybean enzyme (Arahira et al. 2000), are omitted for clarity.

of most helical units of the lid, sequence similarity is much weaker in this region. Only the two active-site tyrosines are truly conserved within the lid; other apparently conserved residues do not have a consistent structural explanation.

Superposition based on the Cα positions of the five catalytic residues from StEH1, HssEH, and AnEH clearly shows differences in the character of their active sites (Fig. 6). The site of StEH1 is constricted near the catalytic residues, but broader to either side (Figs. 4, 6A). Two narrow volumes lie on the ''inside'' of the active site, i.e., on the side of the catalytic residues at left in the two figures. These are partly filled by the two aliphatic tails of valpromide in the B molecule. One branch is lined almost exclusively with phenylalanine side chains and is completely enclosed, while the other has more polar components and eventually leads out to the solvent. Two branches are also observed on the ''outside'' of the StEH1 active site (at right in the two figures). The broader of these points directly toward the solvent and is similar in all three enzymes; this branch is lined primarily with nonpolar residues and contains bound PEG in the StEH1 structure. The second branch is much narrower but also leads eventually to the exterior of the protein; this tunnel is lined with both polar and nonpolar groups and is filled by water molecules in StEH1 (Fig. 4). The site of HssEH (Fig. 6B) is broader near the catalytic residues and much broader on the ''inside''; the two narrow tunnels of StEH1 are merged into a single large cavity in HssEH. Variations

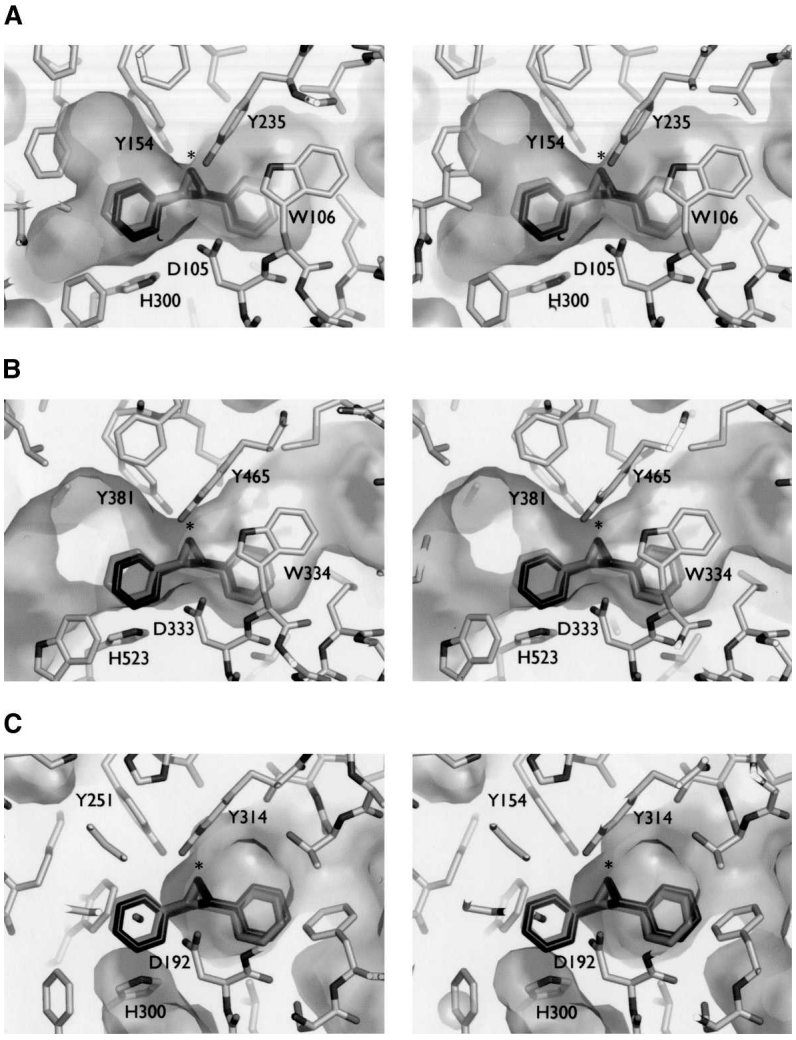

Figure 6. Comparisons of shape of StEH1, HssEH, and AnEH active sites. (*A*) The active-site volume of StEH1 is portrayed from a different view from that used in Figure 4. At *left*, the nonpolar branch of the ''inside'' of the binding site extends upward, while the more polar branch points downward. Of the two ''outside'' branches of the active site, the less-polar one is found at *right*, and the ''pipe'' of waters lies behind. Superpositions based on the catalytic residues were used to generate equivalent views of HssEH and AnEH in *B* and *C*, respectively. *R, R*- and *S, S*-TSO, docked and shaded as in Figure 4, are shown to assist in the comparisons; the epoxide oxygen is again marked with an asterisk.

in the structure of the lid domains underlie these differences between the two proteins. Changes near residue 150 of StEH1 (position ~170 in the alignment in Fig. 5) explain the differences in the ''outside'' of the active sites. Differences near StEH1's residues 266–281 (positions 301–316 in Fig. 5) are also noteworthy for their contributions to the outside regions of the site. Interestingly, both segments include regions of StEH1 that adapt slightly when different ligands are bound. Changes near residues 183–202 of StEH1 (positions ~208–231 in Fig. 5) open up the ''inside'' of HssEH's active site. More extensive changes in AnEH not only block the narrower of the two outside branches of the site but close the entire volume on the inside (Fig. 6C).

The observation that StEH1 is a monomer in the crystal was consistent with our results from size-exclusion chromatography. The value obtained for the apparent molecular mass was 39 kDa, in good agreement with that predicted from the primary structure including the His-tag (37.1 kDa).

Discussion

The α/β-EH reaction is believed to occur in two steps (Lacourciere and Armstrong 1993). First, a *trans*-specific attack of StEH1's Asp105 on the oxirane ring will result in the formation of a covalent enzyme-substrate ester intermediate. The lid tyrosines (Tyr154 and Tyr235) will assist in the opening of the ring through hydrogen bonding (Elfström and Widersten 2006). In the second step of the reaction, the alkyl-enzyme intermediate will be hydrolyzed by the attack of a water molecule that has been activated through proton extraction using the His300–Asp265 charge relay. A crystallographic water molecule (waters 21 and 49 in molecules A and B, respectively) positioned by the side chains of His300 and Glu35, as well as the backbone carbonyl oxygen of Phe33, is ideally situated to take on this role in the second step (Fig. 3). The water's main-chain interaction is provided by an aromatic-proline unit that is linked by a *cis* peptide bond; the *cis* conformation is stabilized by planar stacking of the two side chains and by the structure in the surrounding protein. The His-Gly-(aromatic)-Pro sequence (positions 36–39 in Fig. 5) is one of the most highly conserved motifs among EHs generally because of its crucial roles in shaping the active site; without the *cis* peptide bond, the catalytic water will not be correctly placed.

Comparison of kinetic parameters shows that the *S,S*-TSO is hydrolyzed at a ~10-fold slower rate than the *R,R* enantiomer (Elfström and Widersten 2005). In the first step of the EH reaction, binding of *S,S*-TSO (K_S) is at least threefold tighter than that of *R,R*-TSO, whereas the alkylation rate of *S,S* is more than fivefold slower (Elfström and Widersten 2005). Although there is at present no structure representing an α/β EH bound to any substrate, the available data indicate how they must be placed in the active site. As in previous epoxide hydrolase structures (e.g., Zou et al. 2000), a crystallographic water (water 247 in the present structure) marks the expected position of the oxygen of the epoxide ring. Furthermore, the carbons of the ring must be situated near the appropriate catalytic residues and water. Combined with considerations of the shape of the StEH1 active site, it is possible to envision how some substrates will bind, as illustrated for TSO in Figure 4. Computer docking studies confirm that such models provide a good working hypothesis. It is clear that the two enantiomers cannot be docked in the same way. The plane of the epoxide ring is expected to differ by ~90°, placing the ring carbons differently with respect to the nucleophile and the lid tyrosines. The kinetic data indicate that the arrangement is less optimal for hydrolysis of *S,S*-TSO. The docking experiments further suggest that *S,S*-TSO binds more deeply within the active site. Better binding for this substrate could arise, for example, from good stacking interactions with Trp106 and nonpolar interactions with nearby Leu109, which are not found for the *R,R*-TSO substrate. This hypothesis is supported by the observation that *S,S*-TSO causes more quenching of a tryptophan fluorescence signal than *R,R*-TSO (Elfström and Widersten 2005). The outermost of the two ring carbons lies closest to the nucleophile in the docked *S,S* enantiomer; the opposite is observed for *R,R*-TSO. This would be expected to give rise to differences in regioselectivity in the attack on the ring carbons, a question that should be addressed in future experiments.

Racemic 4-nitrostyrene oxide has a k_{cat} comparable to that of *rac*-TSO, but a substantially larger K_m. This suggests that the epoxide ring of 4-nitrostyrene oxide docks in the same place, but either the loss of a ring or the introduction of the nitro group weakens binding or decreases the alkylation rate. Inspection of the StEH1 active site suggests that both effects will have an impact on binding. However, since this substrate is asymmetric and the kinetic data were obtained for a racemic mixture, further speculation is probably unwise.

StEH1 shows no activity with phenanthrene 9,10-oxide (Elfström and Widersten 2005). Bulky and sterically constrained compounds of this type (Fig. 1A) cannot be placed in the StEH1 active site in an orientation consistent with catalysis; clashes between the fused-ring system of the epoxide and protein side chains inevitably result.

The biologically relevant substrates of StEH1 and similar plant enzymes are thought to be precursors of cutin (Fig. 1C; Walton and Kolattukudy 1972). Like other plant enzymes, StEH1 has a relatively high activity on epoxystearate, thereby producing a cutin precursor, 9,10-dihydroxystearate (catalytic efficiency 150 nmol min^{-1} mg^{-1}). The longer PEG molecule observed in the A molecule of the StEH1

structure provides a good starting point for understanding how substrates with long linear substituents might bind. Further, an analysis of the active site's shape (Figs. 4, 6A) suggests that two long tail groups could be accommodated on each substituted position of the epoxide ring; the arrangement of the observed cavities within the protein further suggests that they could be used to process *cis* and *trans* enantiomers of epoxystearate, in accordance with the observed catalytic properties. Three other EH sequences have been identified in potato (Stapleton et al. 1994), which are 90%–95% identical to StEH1 at the amino acid sequence level. Our analysis suggests that all will have properties similar to StEH1, based on the strong similarities in the lid regions. Little is known about how the expression patterns of the four different potato EHs compare, although it seems possible that such differences could explain the existence of four genes. Similarities in the lid sequences of most other plant enzymes (Fig. 5) suggest that their specificities will also resemble those of StEH1.

Valpromide is bound in the active site of StEH1 in a manner consistent with the observed competitive inhibition. However, because PEG was also present at inhibitory concentrations in the crystallization conditions, this might not be the only relevant mode for binding valpromide in solution. Computer docking studies suggest that the observed position is indeed the most likely one in the case of StEH1, even in the absence of PEG. Identification of crystallization conditions that lack PEG will doubtless assist future efforts to obtain complexes with substrates and inhibitors.

A recent study (Barth et al. 2004) investigated correlations between the sizes of two loops in EHs and their substrate specificity. The conclusions can be re-evaluated using improved alignments based on the structure presented here. One of the loops (the so-called ''NC-loop'') connects the N-terminal segment of the core domain to the lid, and includes positions 143–175 in Figure 5. The other (the ''cap loop'') comprises positions 208–242 of the same alignment. The NC-loop in StEH1 is 24 residues in length (residues 130–153), which classifies it as medium-sized. The investigators' prediction that an NC-loop in excess of 20 residues would contain a helix is correct in this case; further, predicted mobility of this loop is consistent with the observed differences between the A and B molecules of the StEH1 structure. We note that the NC-loop is most important in determining the character of the ''outside'' regions of an EH active site. The cap loop of StEH1 corresponds to that mentioned earlier with regard to the inside regions of the active site. This loop is 31 residues long in StEH1 (residues 183–213), which classifies it as long; comparison of the A and B molecules of StEH1 again suggests mobility. Mammalian EHs show a similar pattern, with sizes of 23 and 34–35 for the two loops. It was stated (Barth et al. 2004) that this combination of loop sizes underlies the enzymes' preferences for elongated epoxides such as fatty acid epoxides or TSO. However, the investigators could not have predicted that both loops, despite their similar length, would differ dramatically in structure between the mammalian and potato enzymes, with clear effects on the shape of their active sites (Fig. 6). The role of StEH1's residues 266–281 (designated as the ''core'' loop because of its location in the α/β-hydrolase core rather than the lid; positions 301–316 in Fig. 5) in forming an adaptable loop that contributes to the outer portions of the site was also not predicted. The fact that this loop tends to be longer in plant EH sequences is additional useful information. The addition of the plant enzyme structure thus provides a firmer basis for future comparisons of this type.

It is clear that changes in the lids underlie most of the differences in specificity among EHs. Very broad active sites (as illustrated for HssEH in Fig. 6B) allow a wide variety of substrates to be bound, consistent with roles in detoxification and drug metabolism. As the plant enzymes are more likely to act within specific niches in metabolism, a higher degree of enantioselectivity may be a more important consideration, and so the shape of the active-site cavity must be more restricted. The effective loss of the ''inside'' end of the binding site in AnEH (Fig. 6C) gives rise to its marked preference for monosubstituted epoxide substrates; this also must be correlated with the preferred biological substrates. A better understanding of the reasons for specificity in the different EHs will allow a more effective exploration of their properties in synthetic applications. Although some protein flexibility is clearly allowed, the intrinsic shape of the different active sites indicates which substrates can reasonably be expected to bind to each.

The ability to recognize monomers versus dimers from amino acid sequences alone would be advantageous, since monomers are generally more useful for applications. Potato and *Arabidopsis* EHs have been reported to be monomers (Morisseau et al. 2000), as has the enzyme from rape (Bellevik et al. 2002). Soybean sEH was originally thought to be a dimer (Blée and Schuber 1992), but more recent studies suggest that it is also a monomer (Arahira et al. 2000). AnEH is a dimer in which the subunits make contact via the relatively flat surfaces of their lids; the dimer is further stabilized by an extension of the protein at the N-terminal end (Zou et al. 2000). ArEH has been reported to be a monomer (Nardini et al. 1999), although a re-evaluation of the crystal structure suggested that dimeric forms might also exist in solution (Zou et al. 2000); this protein, however, lacks a stabilizing N-terminal extension. Mammalian sEHs are dimers but of a very different construction; an N-terminal phosphatase domain provides the necessary interactions. It is shown here that StEH1 cannot form a dimer like that of AnEH because of two insertions in its lid (near residues 143–153 and 184–205; positions 161–175 and 209–234 in Fig. 5).

Clearly, the NC and cap loops are intimately involved in the monomer/dimer issue as well as that of substrate specificity. StEH1 also lacks an N-terminal extension or a domain that would provide an alternate means of dimerization. Based on similar characteristics of their lid and N-terminal sequences (Fig. 5), we predict that most plant enzymes will be monomers.

Materials and methods

Protein purification

The construct containing a C-terminal 5-His tag was described and protein prepared, as reported elsewhere (Elfström and Widersten 2005). The sequence of StEH1 is identical to that of GenBank (Benson et al. 2005) accession number 75102548 with the exception of a Glu to Lys replacement at residue 2 that results from a point mutation (G to A) in the corresponding codon.

Determination of the oligomeric state of StEH1

The native molecular mass of StEH1 was determined by size-exclusion chromatography using a Sephacryl S-200 column (Ø 16 mm, length 60 cm) equilibrated with 0.1 M sodium phosphate (pH 7.5). A mixture of StEH1 (0.2 mg), together with equine cytochrome c (0.3 mg; M_R = 11.7 kDa), human glutathione transferase M2–2 (0.2 mg; M_R = 51.2 kDa), and porcine lactate dehydrogenase (0.1 mg; M_R = 146 kDa), was eluted at a flow rate of 0.2 mL/min. Eluted proteins were identified by OD_{408} (cytochrome *c*) or enzymatic assays for pyruvate reduction (lactate dehydrogenase), glutathione aryl transferase activity (glutathione transferase M2-2), or hydrolysis of *R,R-trans*-stilbene oxide (StEH1).

Enzyme assays and inhibition measurements

EH activity was measured in 0.1 M sodium phosphate with 1% (v/v) acetonitrile (pH 6.8), at 30°C. The initial rates of hydrolysis of *R,R*-TSO (Fig. 1A) were recorded as described earlier (Elfström and Widersten 2005). Steady-state parameters k_{cat} and K_m were obtained after curve fitting the Michaelis-Menten equation to the experimental data using SIMFIT (http://www.simfit.man.ac.uk/). K_I values for PEG 10,000 and valpromide were obtained after determination of values for apparent k_{cat} and K_m in the hydrolysis of *R,R*-TSO, in the presence of 5 mM PEG 10,000 or 2 mM valpromide, respectively.

Crystallization, data collection, and structure solution

Crystals were grown at room temperature using the sitting-drop vapor diffusion method. Then, 1-μL drops of protein solution (7.2 mg/mL, i.e., 0.2 mM, in 30 mM Tris-HCl at pH 7.4, 5 mM valpromide) were mixed with 1-μL drops of reservoir solution (containing 90 mM Na-HEPES at pH 7.5, 25% PEG 10,000). Crystals ($\sim 0.05 \times 0.1 \times 0.4\ \mu m^3$) appeared within 3 d. Prior to flash cooling at 100 K in liquid nitrogen, the crystals were soaked in a cryoprotectant solution comprised of 45 mM Na-HEPES (pH 7.5), 30% PEG 10,000, and 5 mM valpromide. X-ray data were collected at beamline ID14–2 at the European Synchrotron Radiation Facility, (ESRF), Grenoble, and processed using MOSFLM (Leslie 1999) and SCALA (Evans 1993) as implemented in the CCP4 interface (CCP4 1994; Potterton et al. 2003). Data statistics are presented in Table 1.

Analysis of the cell contents indicated that there were two molecules in the asymmetric unit, with 45% solvent and a V_M of 2.2 (Matthews 1968). Molecular replacement with MOLREP (Vagin and Teplyakov 1997) using a HssEH structure (PDB entry [Berman et al. 2000] 1VJ5 [Gomez et al. 2004]) as a search model gave two solutions. Auto-building with ARP/wARP (Perrakis et al. 1999) failed, but the improved maps so generated could be averaged in the program O (Jones et al. 1991) and used to build a rough model for ~80% of the residues of the A subunit of the potato EH. Noncrystallographic symmetry generated the second molecule of the asymmetric unit. Using this model to calculate starting phases, ARP/wARP successfully built 626 of 656 residues in the asymmetric unit, with 97% connectivity. Rebuilding in O placed the remaining residues (most notably the *cis* peptide of each molecule). After refinement in REFMAC5 (Murshudov et al. 1997) with a matrix weight of 0.05, rebuilding as needed, and placement of ligands (valpromide and PEG) and waters, the final model described in Table 1 was obtained. Back-soaks in the absence of valpromide were carried out for 1 min in cryo solution containing 1% acetonitrile.

Other methods

Similar sequences were identified using BLAST (Altschul et al. 1997), and structures were compared using O and LSQMAN (Kleywegt and Jones 1997). Structure-based sequence alignments, followed by alignments of additional sequences, were carried out using Indonesia (D. Madsen, G. Kleywegt, and P. Johansson, unpubl. [http://xray.bmc.uu.se/~dennis/]). In silico docking experiments were carried out with AutoDock3.0.3 in the AutoDockTools interface (Morris et al. 1998) using the genetic algorithm protocols with a wide variety of search parameters. Figures were prepared with the programs PyMOL (DeLano Scientific [http://www.pymol.org]), Canvas (ACD Systems, Inc.), and ISIS Draw (Molecular Design, Ltd.).

Deposition of data

Atomic coordinates and structure factor data have been deposited in the PDB with entry code 2CJP.

Acknowledgments

This work was supported by grants from the Swedish Research Council (VR) to S.L.M. and the Carl Trygger Foundation to M.W. We thank Hugo G. de Terán for invaluable assistance with the computer docking experiments.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Arahira, M., Nong, V.H., Udaka, K., and Fukazawa, C. 2000. Purification, molecular cloning and ethylene-inducible expression of a soluble-type epoxide hydrolase from soybean (Glycine max [L.] Merr.). *Eur. J. Biochem.* **267:** 2649–2657.

Archelas, A. and Furstoss, R. 2001. Synthetic applications of epoxide hydrolases. *Curr. Opin. Chem. Biol.* **5:** 112–119.

Argiriadi, M.A., Morisseau, C., Hammock, B.D., and Christianson, D.W. 1999. Detoxification of environmental mutagens and carcinogens: Structure, mechanism, and evolution of liver epoxide hydrolase. *Proc. Natl. Acad. Sci.* **96:** 10637–10642.

Argiriadi, M.A., Morisseau, C., Goodrow, M.H., Dowdy, D.L., Hammock, B.D., and Christianson, D.W. 2000. Binding of alkylurea inhibitors to epoxide hydrolase implicates active site tyrosines in substrate activation. *J. Biol. Chem.* **275:** 15265–15270.

Barth, S., Fischer, M., Schmid, R.D., and Pleiss, J. 2004. Sequence and structure of epoxide hydrolases: A systematic analysis. *Proteins* **55:** 846–855.

Bellevik, S., Zhang, J., and Meijer, J. 2002. *Brassica napus* soluble epoxide hydrolase (BNSEH1). *Eur. J. Biochem.* **269:** 5295–5302.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Wheeler, D.L. 2005. GenBank. *Nucleic Acids Res.* **33:** D34–D38.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The protein data bank. *Nucleic Acids Res.* **28:** 235–242.

Blée, E. and Schuber, F. 1992. Occurrence of fatty acid epoxide hydrolases in soybean (*Glycine max*). Purification and characterization of the soluble form. *Biochem. J.* **282:** 711–714.

Collaborative Computing Project, Number 4 (CCP4). 1994. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50:** 760–763.

Elfström, L.T. and Widersten, M. 2005. Catalysis of potato epoxide hydrolase, StEH1. *Biochem. J.* **290:** 633–640.

———. 2006. Implications for an ionized alkyl-enzyme intermediate during StEH1-catalyzed *trans*-stilbene oxide hydrolysis. *Biochemistry* **45:** 205–212.

Engh, R.A. and Huber, R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47:** 392–400.

Evans, P.R. 1993. Data reduction. In *Proceedings of CCP4 Study Weekend on Data Collection & Processing* (eds. L. Sawyer et al.), pp. 114–122. Daresbury Laboratory, Warrington, UK.

Gomez, G.A., Morisseau, C., Hammock, B.D., and Christianson, D.W. 2004. Structure of human epoxide hydrolase reveals mechanistic inferences on bifunctional catalysis in epoxide and phosphate ester hydrolysis. *Biochemistry* **43:** 4716–4723.

Heikinheimo, P., Goldman, A., Jeffries, C., and Ollis, D.L. 1999. Of barn owls and bankers: A lush variety of α/β hydrolases. *Structure* **7:** 141–146.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47:** 110–119.

Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22:** 2577–2637.

Kleywegt, G.J. and Jones, T.A. 1996. Phi/Psi-cology: Ramachandran revisited. *Structure* **4:** 1395–1400.

———. 1997. Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277:** 525–545.

Lacourciere, G.M. and Armstrong, R.N. 1993. The catalytic mechanism of microsomal epoxide hydrolase involves an ester intermediate. *J. Am. Chem. Soc.* **115:** 10466–10467.

Leslie, A.G. 1999. Integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **55:** 1696–1702.

Matthews, B.W. 1968. Solvent content of protein crystals. *J. Mol. Biol.* **33:** 491–497.

Morisseau, C. and Hammock, B.D. 2005. Epoxide hydrolases: Mechanisms, inhibitor designs, and biological roles. *Annu. Rev. Pharmacol. Toxicol.* **45:** 311–333.

Morisseau, C., Beetham, J.K., Pinot, F., Debernard, S., Newman, J.W., and Hammock, B.D. 2000. Cress and potato soluble epoxide hydrolases: Purification, biochemical characterization, and comparison to mammalian enzymes. *Arch. Biochem. Biophys.* **378:** 321–332.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19:** 1639–1662.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53:** 240–255.

Nardini, M., Ridder, I.S., Rozeboom, H.J., Kalk, K.H., Rink, R., Janssen, D.B., and Dijkstra, B.W. 1999. The X-ray structure of epoxide hydrolase from *Agrobacterium radiobacter* AD1. An enzyme to detoxify harmful epoxides. *J. Biol. Chem.* **274:** 14579–14586.

Newman, J.W., Morisseau, C., and Hammock, B.D. 2005. Epoxide hydrolases: Their roles and interactions with lipid metabolism. *Prog. Lipid Res.* **44:** 1–51.

Perrakis, A., Morris, R., and Lamzin, V.S. 1999. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6:** 458–463.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. 2003. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59:** 1131–1137.

Read, R.J. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42:** 140–149.

Stapleton, A., Beetham, J.K., Pinot, F., Garbarino, J.E., Rockhold, D.R., Friedman, M., Hammock, B.D., and Belknap, W.R. 1994. Cloning and expression of soluble epoxide hydrolase from potato. *Plant J.* **6:** 251–258.

Vagin, A. and Teplyakov, A. 1997. MOLREP: An automated program for molecular replacement. *J. Appl. Crystallogr.* **30:** 1022–1025.

Walton, T.J. and Kolattukudy, P.E. 1972. Determination of the structures of cutin monomers by a novel depolymerization procedure and combined gas chromatography and mass spectrometry. *Biochemistry* **11:** 1885–1896.

Zou, J., Hallberg, B.M., Bergfors, T., Oesch, F., Arand, M., Mowbray, S.L., and Jones, T.A. 2000. Structure of *Aspergillus niger* epoxide hydrolase at 1.8 Å resolution: Implications for the structure and function of the mammalian microsomal class of epoxide hydrolases. *Struct. Fold. Des.* **8:** 111–122.