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

Jinyu Zou[1], B Martin Hallberg[1], Terese Bergfors[1], Franz Oesch[2], Michael Arand[2], Sherry L Mowbray[3] and T Alwyn Jones[1]*

Background: Epoxide hydrolases have important roles in the defense of cells against potentially harmful epoxides. Conversion of epoxides into less toxic and more easily excreted diols is a universally successful strategy. A number of microorganisms employ the same chemistry to process epoxides for use as carbon sources.

Results: The X-ray structure of the epoxide hydrolase from *Aspergillus niger* was determined at 3.5 Å resolution using the multiwavelength anomalous dispersion (MAD) method, and then refined at 1.8 Å resolution. There is a dimer consisting of two 44 kDa subunits in the asymmetric unit. Each subunit consists of an α/β hydrolase fold, and a primarily helical lid over the active site. The dimer interface includes lid–lid interactions as well as contributions from an N-terminal meander. The active site contains a classical catalytic triad, and two tyrosines and a glutamic acid residue that are likely to assist in catalysis.

Conclusions: The *Aspergillus* enzyme provides the first structure of an epoxide hydrolase with strong relationships to the most important enzyme of human epoxide metabolism, the microsomal epoxide hydrolase. Differences in active-site residues, especially in components that assist in epoxide ring

metabolic enzyme. The N-terminal domain that is characteristic of microsomal epoxide hydrolases corresponds to a meander that is critical for dimer formation in the *Aspergillus* enzyme.

Introduction

Epoxides are frequent intermediates in the metabolism of lipophilic xenobiotics. Because of their high ring tension and polarized C–O bonds, they are often electrophilically reactive, forming adducts with cellular DNA and so producing mutations that can cause cancer. Epoxide hydrolases (EH; EC 3.3.2.3) convert epoxides to the more water-soluble and usually less toxic diols, and are therefore key enzymes in the defense of the human organism against the hazardous properties of xenobiotic compounds [1].

The epoxide hydrolase implicated in most metabolism of xenobiotics in mammals is the membrane-bound microsomal epoxide hydrolase (mEH; [2]). It is particularly abundant in human liver, the major site of foreign-compound metabolism. A second EH that is also capable of hydrolyzing xenobiotic epoxides is the soluble epoxide hydrolase (sEH; [3]). Although direct sequence comparison of the two mammalian EHs did not at first suggest a structural relationship [4], the low yet significant sequence similarity of each enzyme to the bacterial haloalkane dehalogenase (HAD; [5]) indicated a common origin [6].

This relationship gave many clues about the chemistry to be expected for the EH reaction. HAD belongs to the class of α/β hydrolase fold enzymes [7,8] that, despite a generally low sequence similarity, are thought to be of common phylogenetic origin. The catalytic triad of HAD is composed of Asp124–Asp260–His289, with Asp124 being the catalytic nucleophile needed to form an ester intermediate with the substrate dichloroethane [9,10]. This ester intermediate differs from that of the classical serine hydrolases, such as acetylcholine esterase, in that the acid component is derived from the enzyme rather than from the substrate. The product release is afforded by hydrolytic cleavage of the covalent intermediate through attack of a water molecule activated by the His289–Asp260 charge-relay system [11]. Sequence comparisons suggest a conservation of this catalytic triad in both mammalian EHs [6,12,13], which has more recently been substantiated using biochemical characterization of the proteins [14–18].

Over the past several years, a growing number of EHs have been identified in microorganisms, largely motivated by their potential as enantioselective biocatalysts for the chiral synthesis of fine chemicals, for example drug precursors. At present, all of these can be classified as α/β hydrolase fold enzymes, with only a single exception [19]. The solution of the structure of the EH from *Agrobacterium radiobacter* [20] gave the first picture of an EH structure, and many insights into the sEH family. We have recently reported the cloning and molecular analysis of an epoxide hydrolase from the fungus *Aspergillus niger* (AnEH; [21]). This enzyme is especially interesting for two reasons. First, it is a particularly useful biocatalyst because it shows exceptionally fast substrate turnover as compared with other EHs, while at the same time being highly enantioselective towards industrially important epoxides [22]. Second, and of broader interest, it is the first known soluble member of the family of mEHs. Structural information about this branch of the EH enzymes has previously been unobtainable, largely because of the difficulties in crystallizing membrane proteins. AnEH thus offers an excellent opportunity to study some basic features of the enzymatic reaction of mEHs that have previously been hard to address. One important question is the nature of the proton donor that, through interactions with the epoxide ring oxygen atom, is expected to assist in enzymatic attack. A second puzzle centers on the reasons why *trans*-substituted epoxides are generally not hydrolyzed by the mEHs, despite the otherwise promiscuous substrate selectivity of these enzymes. Third, the function of the long N-terminal extension (~90 residues) preceding the α/β hydrolase fold in the mEH family has remained obscure. A possible role in molecular interactions with other drug-metabolizing enzymes (for example cytochromes P450, UDP-glucuronosyl transferases) for efficient substrate channeling is conceivable but has not, as yet, been tested.

In the present paper, we describe the three-dimensional structure of AnEH determined at a resolution of 1.8 Å, the first structure of the mEH family to be solved.

Results and discussion

The three-dimensional crystal structure was determined at 3.5 Å resolution by the multiwavelength anomalous diffraction (MAD) method using crystals of a seleno-methionine (SeMet) substituted form of the enzyme. The model was then refined at 1.8 Å resolution to a crystallographic R factor of 22.5% and an R_{free} [23] of 23.3%. There are two subunits of epoxide hydrolase in the asymmetric unit, with a dimer formed around a nearly perfect (179.2°) non-crystallographic twofold axis. The current model includes residues 3–396 (out of 398), except for the partially disordered residues from 320–328. Because of the expression construct used, the sequence KA (in single-letter amino acid code) is present at residues 3–4 instead of the AP of

the wild-type protein (see the Materials and methods section). A total of 374 water molecules are included. Some statistics for the X-ray data collection and refinement are shown in Tables 1 and 2.

Overall structure and fold of the subunit

The epoxide hydrolase subunit can be divided into three parts: a core α/β domain that contains the active-site residues, a lid protruding from this domain (that caps the

Table 1

Data collection statistics.

	Native	Se-AnEH Peak	Se-AnEH Inflection	Se-AnEH Remote
Wavelength (Å)	0.937	0.9794	0.9796	0.9312
Resolution (Å)	1.8–50	2.7–30	2.7–30	2.7–30
No. observations	263,931	161,798	163,833	163,296
No. unique reflections	74,416	22,522	22,587	22,546
Completeness (%)	99.8 (99.9)	99.5 (98.5)	99.5 (98.4)	99.6 (99.6)
R_{merge}* (%)	7.3 (32.1)	5.1 (12)	5.8 (17)	6.1 (10.4)
Average I/σ(I)[†]	15 (5)	11 (5)	10 (4)	13 (6)
Anomalous differences (%)[†‡]	–	5.6	4.1	4.8
Dispersive differences to remote wavelength (%)[†‡]	–	4.8	4.9	–
Refined scattering factors f'/f'' (e⁻)[†‡]	–	–8.1/7.3	–8.5/3.5	–1.7/3.3
FOM[‡]	–	–	–	0.81
FOM after DM[‡]	–	–	–	0.84
FOM after twofold averaging in DM[‡]	–	–	–	0.88

Both the native and Se-AnEH crystallized in space group P2₁. The unit-cell dimensions were a = 62.7 Å, b = 89.3 Å, c = 75.8 Å and β = 105° for the native protein, and a = 62.8 Å, b = 89.8 Å, c = 75.7 Å and β = 105.4° for the Se-AnEH protein. Statistics for the highest resolution shells, 1.86–1.80 Å for the native protein and 2.8–2.7 Å for Se-AnEH, are given in parentheses. *As taken from SCALEPACK [49]. †As taken from SOLVE [47]. ‡Valid for data truncated to 3.5 Å resolution. FOM, figure of merit.

Table 2

Refinement statistics.

Resolution range (Å)	20.0–1.8
No. of reflections	74,309
No. of reflections used for R_{free} calculation	1524
R value, R_{free} value (%)	22.5, 23.3
No. of non-hydrogen protein atoms	6128
No. of solvent waters	374
Mean B factor, protein mainchain atoms (Å²)	22.3
Mean B factor, protein sidechain atoms (Å²)	23.2
Average B factor, solvent atoms (Å²)	30.4
Ramachandran plot outliers (%)*	2.5
Rms deviation bond length (Å)	0.006
Rms deviation bond angle (°)	1.3
Rms difference NCS, on Cα atoms (Å)	0.020
Rms difference NCS, on all atoms (Å)	0.023
Rms difference NCS ΔB, on Cα atoms (Å²)	0.35
Rms difference NCS ΔB, on all atoms (Å²)	0.38

*A stringent-boundary Ramachandran plot was used [60].

active site), and a long N-terminal meander that in turn caps the lid (Figures 1a,b).

The α/β fold follows the classic design of an α/β hydrolase [8], consisting of a twisted eight-stranded β sheet flanked on both faces by α helices. As in the canonical fold, all strands are parallel, except the second, which is antiparallel (Figure 1b). The two Ramachandran outliers are common features of the α/β hydrolase family. The first involves the mainchain conformation of the active-site nucleophile (in this case Asp192). The second lies in the insert at the C-terminal end of β4 (normally a glycine residue in the closest family members, but in this case Thr153). This insert has a highly conserved sequence and has been used as one of the key motifs for epoxide hydrolase identification in genomic sequence searches (M Arand, unpublished observations). Its function appears to be to create a niche to support one of the helices

Figure 1

Structure of AnEH. **(a)** Cα trace, with colors going through the rainbow from red at the N terminus to blue at the C terminus. The sidechains of the catalytic triad (Asp192, His374 and Asp348), Tyr251, Tyr314, and Glu123 are shown in red. The approximate placement of residues 320–328 (omitted from the final model) is indicated by dotted lines. **(b)** Topology diagram colored using the same rainbow scheme. Secondary-structure descriptors for the α/β domain follow earlier nomenclature [7], whereas the N-terminal meander and lid domains use separate systems, prefixed with N and L, respectively. Residues of the catalytic triad are indicated by red stars, and the two active-site tyrosine residues of the lid by magenta stars.
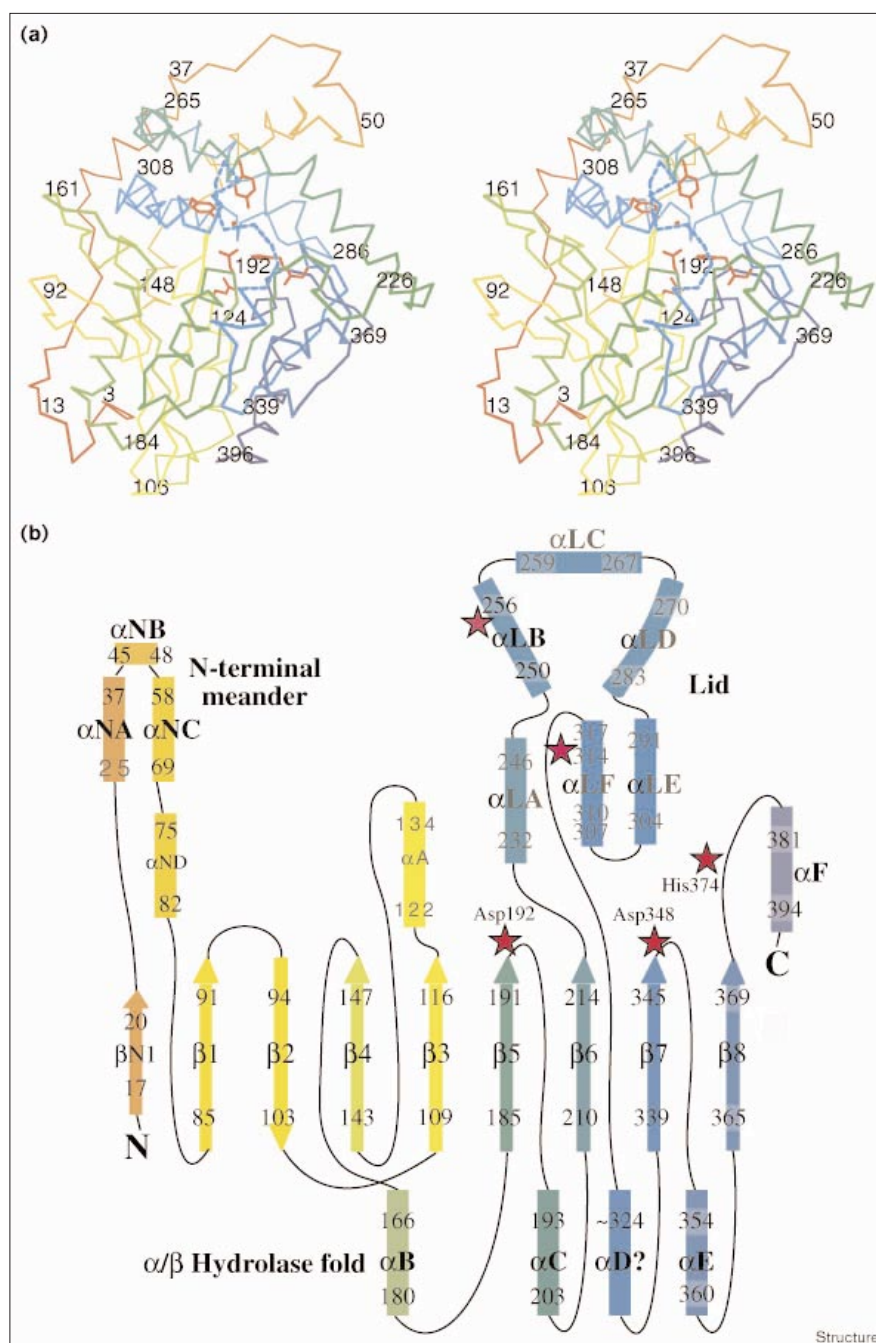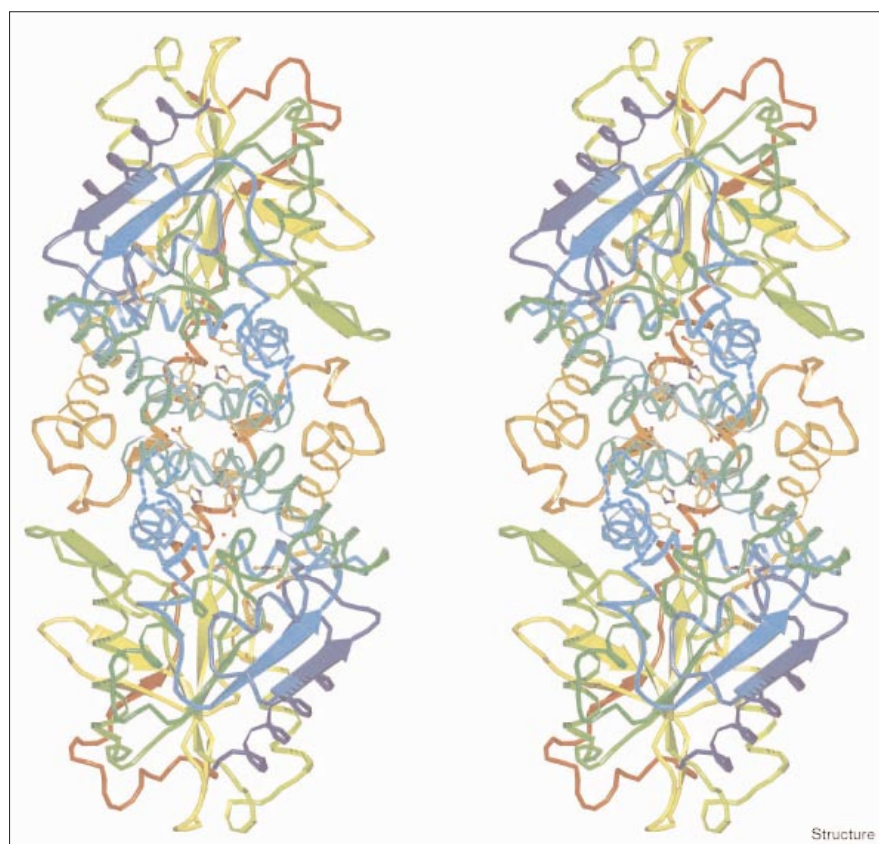
Figure 2

Stereo ribbon drawing showing the interface and symmetry of the AnEH dimer interaction. Both subunits are colored in the rainbow scheme used in Figure 1. Sidechains of the dimer interface are shown, as well as a series of interactions (Trp276, His255, and Tyr314) that leads from the dimer interface to the active site.

making up the active-site lid. Strand β6 ends at residue 214 and the chain then makes a sharp change in direction. The space next to the abutting helix αC is poorly packed with sidechains, creating a solvent-accessible tunnel that reaches to the catalytic nucleophile at residue 192. The α/β hydrolase domain continues from residue 329. Most of the intervening region is made up by the lid.

In AnEH, the lid is composed of six α helices corresponding to residues 231–319. The first helix leads up to the active site, then αLB (where the prefix L denotes the lid domain) bridges it. Helices αLB, αLC and αLD together form a nearly triangular arrangement; αLD is somewhat curved, with a bend near Trp276. Another change in direction brings αLE and αLF back, such that these two helices are wedged between the triangle and the α/β hydrolase domain. Helix αLF is distorted near Tyr314, whose sidechain points into the active site. The residues (320–328) that link the lid to αD of the α/β hydrolase fold are poorly defined in our electron-density map.

The N-terminal region of AnEH (residues 13–85) forms an unusual curved meander approximately 60 Å in length. At the beginning and end of the meander, hydrophobic sidechains pack onto the α/β hydrolase core. More central

segments interact first with the helices of the lid, then with the lid of the twofold-related molecule (Figures 1a,2).

Two *cis*-prolines are found in each subunit, both at critical parts of the fold. Pro118 is located at the end of β3 in a reverse turn adjacent to the active-site nucleophile; the mainchain carbonyl oxygen of the preceding residue, Trp117, forms part of the oxyanion hole. Pro158 is part of the lid support structure at the end of β4. Both regions are highly conserved in the sequences making up this subfamily of the epoxide hydrolases, and in many other α/β hydrolases.

The dimer interaction
Gel filtration chromatography studies suggest that AnEH is a dimer in solution (M Arand *et al.*, unpublished observations). In the crystal structure, a dimer is formed about a nearly perfect noncrystallographic twofold axis (Figure 2) with a total loss of 5500 Å^2 of accessible surface area. The dimer interface results primarily from the interaction of the lids, and the tips of the N-terminal meanders. The triangular arrangement of lid helices αLB, αLC and αLD is centered on the indole ring of Trp276, which in turn packs against the twofold related shape in the dimer interface. Although most of the sidechains in the interface are

Figure 3
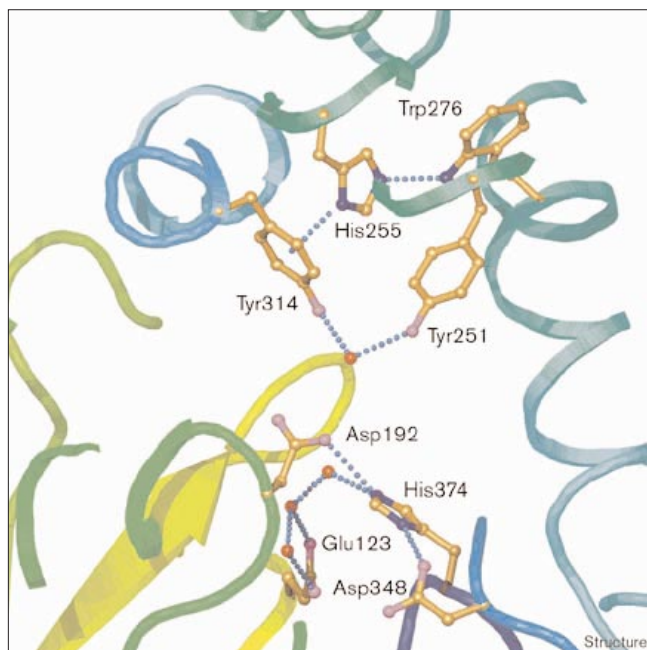
Catalytic site residues (Asp192, His374 and Asp348), with the most important hydrogen-bonding interactions displayed. The water molecule (red) interacting with Tyr251 and Tyr314 is believed to indicate the position of the oxirane ring oxygen. Additional water molecules near the charge-relay residues (His374 and Asp348) and Glu123 are thought to provide a reservoir of water close to the active site. The series of interactions leading via the rings of Trp276 and His255 to the active-site residue Tyr314 is also shown. Colors for the backbone use the same rainbow scheme found in earlier figures.

hydrophobic, there are also many salt links and other polar interactions. These include the buried but hydrated Glu254, and a series of interactions leading via the rings of Trp276, His255 and Tyr314 to the active site (Figure 3). The tip of each N-terminal meander also interacts with the lid helices of both monomers so that αLB, αLC and αLD are almost completely buried within the dimer interface. Consequently, this region has the lowest crystallographic temperature factors in the entire structure.

The active site
There are two active sites per dimer; each is composed of residues drawn from a single subunit. The residues implicated directly in catalysis by analogy to other enzymes [21] are the catalytic triad of Asp192, His374 and Asp348 (Figure 3).

The reaction is believed to occur in two steps ([24]; Figure 4). First, a *trans*-specific attack of Asp192 on the oxirane ring results in the formation of a covalent enzyme–substrate ester intermediate. This is thought to be the rate-determining step for AnEH [21]. It is believed that a proton-donating group from the protein assists

epoxide ring opening, in a classic push–pull mechanism. In the second step, the intermediate is hydrolysed by the attack of a water molecule that has been activated through proton extraction using the His374–Asp348 charge relay.

Asp192 is located in the cross-over connection between β5 and αC. The mainchain conformation is strained at this point, and the residue is one of the two Ramachandran outliers in the structure. The local sequence (GGDIGS, residues 190–195) is rich in glycines, and agrees well with the generic sm–X–nu–X–sm–sm motif of the α/β fold hydrolases, where sm is a small residue, nu is the catalytic nucleophile and X is any residue. The lack of a sidechain at 190 allows the close packing of β5 and αC. Similarly, a glycine at 194 allows the close proximity of β5 and β3, whereas a larger sidechain at residue 191 would affect access to the active site. The other members of the triad are positioned as expected: His374 lies on the loop between β8 and αF, and Asp348 on the loop between β7 and αE.

Figure 3 shows the hydrogen-bonding network near the active site. Although there is a hydrogen bond between His374 and Asp348, the interaction of Oδ1 of Asp192 with the imadazole ring is more distant, and has poor hydrogen-bonding geometry. Oδ2 of Asp192, on the other hand, forms good hydrogen bonds with the mainchain amide nitrogens of residues 117 and 193. These hydrogen bonds constitute the classical α/β hydrolase oxyanion hole interactions [9]. The only other nearby contact for Asp192 in the apo structure is to a water molecule (3.3 Å separation). This water molecule in turn interacts with the hydroxyl groups of two tyrosine residues, 251 and 314, whose sidechains hang down from αLB and αLF of the lid (Figures 1a,3). These tyrosines are perfectly situated to act as proton donors for the epoxide ring oxygen, and so to assist in ring opening in the first step of the reaction. The Trp276, His255, Tyr314 series of interactions that begins near the subunit–subunit interface might well affect the chemical properties of the active-site tyrosine pair in a manner that is important in catalysis. Such supporting interactions could, for instance, reduce the pK_a of Tyr314 from the normal value of ~10 by delocalising the negative charge of the deprotonated species.

It was found that mutation of a charge-relay residue in rat mEH (Glu404→Gln) left some residual activity, and on this basis it was suggested that Glu376 could contribute to water activation during the second step of catalysis [1,17]. Glu376 corresponds to Glu316 in AnEH. This residue is not, in fact, located in the active site, but is instead aimed in the opposite direction, forming a hydrogen bond with the mainchain nitrogen of Leu166. As such, it is not in the appropriate location for a catalytic role, although it might well have an effect on catalysis for indirect structural reasons. Very recent work has shown that expression of the Glu376Gln mutant protein in yeast

Figure 4
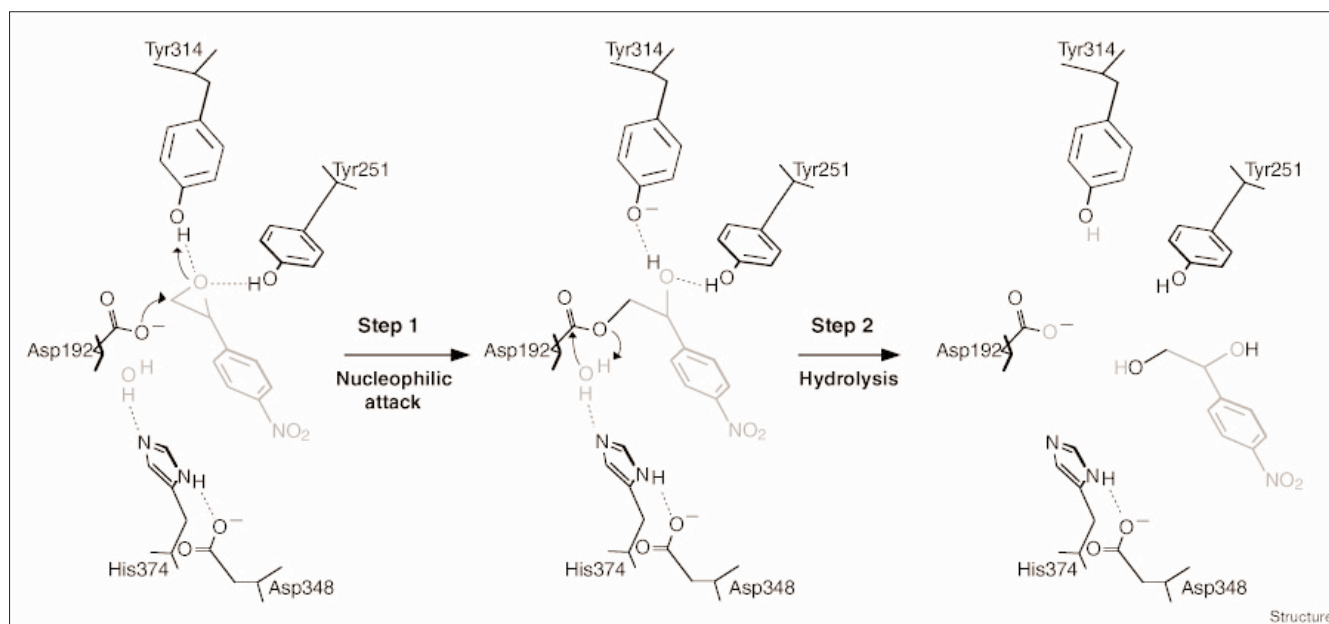
Mechanism of 4-nitrostyrene oxide hydrolysis by AnEH; residues are shown as close as possible to their actual relative orientation and position. Initial binding and positioning of the substrate in the active center of the enzyme is thought to be directed by hydrogen bonding between the epoxide ring oxygen and the hydroxyl groups of Tyr251 and Tyr314. In the first step of the enzymatic reaction, the carboxylate sidechain of the catalytic nucleophile Asp192 attacks the unsubstituted carbon atom of the oxirane ring, forming an ester bond and opening the ring. The process is facilitated through simultaneous proton donation by one of the tyrosine residues (the series of hydrogen bonds between the π system of Tyr314 and His255/Trp276 discussed in the main text favors Tyr314, as it is expected to result in a lower pK_a). The two parts together comprise a classic push–pull mechanism. The second step of the reaction consists of the hydrolysis of the ester intermediate via the attack of a water molecule. This water is activated through proton abstraction by the His374/Asp348 charge-relay system, possibly further assisted by Glu123 (not shown).

(instead of in bacteria) leads to a fully active mEH (M Arand *et al.*, unpublished observations).

In contrast, Glu123, which is well conserved in this subfamily of EHs, could be involved in water activation. This residue protrudes into the active-site region and is not involved in hydrogen-bonding interactions with other protein atoms. The carboxylate group is hydrated, and is bridged to the imadazole ring of His374 via two water molecules (Figure 3). The water molecule closest to the histidine is ideally positioned for the hydrolysis step of the reaction, and is only 4.2 Å from the carboxylate group of Glu123.

A previous report that one or more cysteine residues are important in activity [22] appears to be attributable to the presence of Cys350 in the wall of the active site.

Full access to the active site is prevented by the α-helical lid. The known EHs have characteristic enantio- and regio-selectivities, which can presumably be explained in terms of the shape and character of the active-site cavity. AnEH preferentially hydrolyzes the *R* enantiomer of mono-substituted oxirane substrates [22], attacking at the least hindered carbon of the ring. Retention of configuration in the product results, for both *R* and *S* enantiomers. We have modeled the binding of both enantiomers of the substrate 4-nitrostyrene oxide into the active site, and conclude that the selectivity is because of the different direction of attack by Asp192 on the two substrates. If the epoxide ring oxygen of the *R* isomer is placed at the site of the water interacting with the tyrosines, we can point the least hindered carbon towards Asp192 while placing the aromatic ring in the opening to the active site (Figure 5), in a seemingly optimal arrangement for catalysis. The aromatic ring of *p*-nitrostyrene of this substrate should make favorable perpendicular stacking interactions to Phe196 and Phe244 [25]. Similarly, the *S* isomer can be placed in a position for SN2 attack by Asp192, and for hydrogen bonding by the two tyrosines. The plane of the epoxide ring is, however, aligned quite differently with respect to the catalytic nucleophile, and there will be much more crowding for substituents to the C7 atom (Figure 5). This should explain the reduced binding of the *S* enantiomer of 4-nitrostyrene oxide [26]. Substitution at C8 destroys activity for both enantiomers, and is probably a result of close contacts between the substituent and Asp192, as well as Trp284 and Leu349.

Similarity to other epoxide hydrolases

Sequence similarity between mEH and sEH could not be detected directly, but was discovered through their common similarity to sequences of the α/β hydrolase fold family [6]. The structure of the EH from *Agrobacterium radiobacter* was reported recently [20], the first EH structure to be published. This enzyme represents the subfamily of sEH proteins, as AnEH exemplifies the mEH branch. Table 3 gives some comparative information for the two EH structures; they are superimposed in Figure 6. There is 19% sequence identity in equivalent segments of the bacterial protein and AnEH. Two residues of the catalytic triad, those equivalent to Asp192 and His374, are placed in an equivalent manner in a similar structural setting. Surprisingly, the authors state that the third member of the triad (Asp246) is forced out of the active site by non-productive binding of Gln134, which they attribute to crystal-packing effects.

The two hydrogen-bond donors of AnEH that are believed to assist in ring opening, Tyr251 and Tyr314, are conserved as Tyr152 and Tyr215 in the bacterial enzyme. Tyr→Phe mutations of both these donors together dramatically impair performance of the enzyme (referred to as unpublished work in [20]), supporting their role as important players in catalysis. The hydrogen-bonding interactions observed to support Tyr314 in AnEH (Figure 3) are also conserved in the bacterial protein as His156 (equivalent to His255) and Tyr175 (equivalent to Trp276). Other differences in the active-site region of the enzymes are presumably responsible for their respective substrate specificities (for example, Leu193 of AnEH is changed to Phe108 in the bacterial protein).

Although the authors state that the bacterial enzyme is monomeric [20], there are four molecules of the bacterial EH in the reported asymmetric unit. These, in fact, represent two pairs of dimers that are very similar to the dimer observed for AnEH. We calculate the total buried surface area for each dimer in the bacterial EH crystal to be 1200 Å^2. This suggests that a dimer is present in solution, perhaps in a reversible equilibrium with the monomer. As with AnEH, the dimer interface is mostly hydrophobic in character (for example, the Trp151/Phe155/Leu158/Val164/Tyr175 cluster), although the residues involved are not well conserved between the two proteins. The smaller size of the contact surface in the bacterial enzyme is largely because of its lack of the N-terminal meander. Some other epoxide hydrolases have been reported to exist as dimers in solution, including the human sEH [27] as well as the soybean enzyme [28].

The structure of the murine sEH has been published very recently [29]. As coordinates are not yet available, and the resolution of the sEH structure is also lower than for

Figure 5
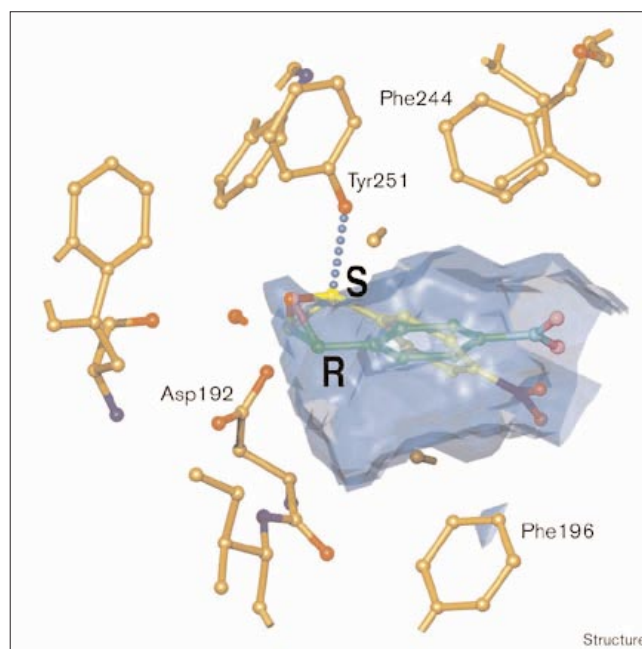
Substrate 4-nitrostyrene oxide docked into the active site of the apo enzyme, shown with a solvent accessible surface. The site is complementary to the *R* enantiomer (green carbon atoms), and a poorer fit for the *S* enantiomer (yellow carbon atoms). The close contact (< 3 Å) expected between C7 of the *S* isomer and the hydroxyl group of Tyr251 is shown.

AnEH, our comparisons must at present remain qualitative. The catalytic acid of sEH is Asp333, and is supported by nearby charge-relay residues (His523 and Asp495). Tyrosine residues 381 and 465 appear to take on the positions of 251 and 314 of AnEH, although the authors considered only the significance of Tyr381 in catalysis. All of the available evidence points to a catalytic mechanism similar to that used by AnEH.

The approximately 30% sequence identity of AnEH to the human mEH allowed us to build a model of the latter enzyme, based on our coordinates. The residues of the catalytic triad [18] can be placed with great confidence. The nucleophile, Asp226, and His431 of the charge relay replace the equivalent residues of AnEH. The substitution of Asp348 with Glu404 in the mEH charge-relay system might slow down the hydrolysis step and thus contribute to the fact that this step is rate-limiting for the mammalian enzyme [1,21]. Indeed, a replacement of Glu404 with an aspartic acid leads to a strongly increased turnover rate, with altered kinetic characteristics that support such an interpretation [18]. Although sequence relationships are weaker in the lid, Tyr299 and Tyr374 appear in equivalent positions to the two tyrosines of the AnEH active site. The residues that support these two tyrosines in AnEH are mutated in mEH: His255 is

Table 3

Comparison to other known structures.

Protein (Source)	No. residues	No. atoms within a 3.8 Å cut-off (% identical of matched)	Rms difference to AnEH (Å)	PDB code and reference	Nucleophile	Charge-relay residues	
					Equiv. to Asp192	Equiv. to His374	Equiv. to Asp348
Bacterial EH (*Agrobacterium radiobacter*)	282	224 (18.8)	1.72	1EHY [20]	Asp107	His275	Asp246?
Bromoperoxidase A2 (*Streptomyces aureofaciens*)	277	193 (13.5)	1.95	1BRO [61]	Ser98	His257	Asp228
Chloroperoxidase F (*Pseudomonas fluorescens*)	273	194 (16.5)	1.97	1A8S [62]	Ser94	His253	Asp224
Chloroperoxidase L (*Streptomyces lividans*)	275	192 (16.9)	1.99	1A88 [62]	Ser96	His255	Asp226
Bromoperoxidase A1 (*Streptomyces aureofaciens*)	274	182 (12.7)	1.88	1A8Q [62]	Ser94	His252	Asp223
Haloalkane dehalogenase (*Xanthobacter autotrophicus*)	310	167 (16.2)	2.13	2HAD [9]	Asp124	His289	Asp260
Hydroxynitrile lyase (*Hevea brasiliensis*)	256	149 (10.7)	1.85	1YAS [63]	Ser80	His235	Asp207
Proline iminopeptidase (*Xanthomona campestis*)	313	155 (17.4)	1.95	1AZW [64]	Ser110	His294	Asp266
Lipase (*Candida antarctica*)	317	103 (12.6)	2.11	1TCB [65]	Ser105	His224	Asp187

The putative third member of the charge relay of the bacterial EH is indicated with a question mark, although the currently available structure does not place this residue in the appropriate position. Equiv., equivalent.

replaced by Gln303, and Trp226 by Tyr324. Although details of the mEH structure are impossible to predict, mutations such as Phe196→Leu230, Cys350→Phe406, and Phe244→Ser292 appear likely to alter the shape and character of the active-site cavity and help rationalize why mEH is able to hydrolyse a broad range of monosubstituted and *cis*-disubstituted epoxides, whereas AnEH can hydrolyze only monosubstituted epoxides efficiently.

Figure 6
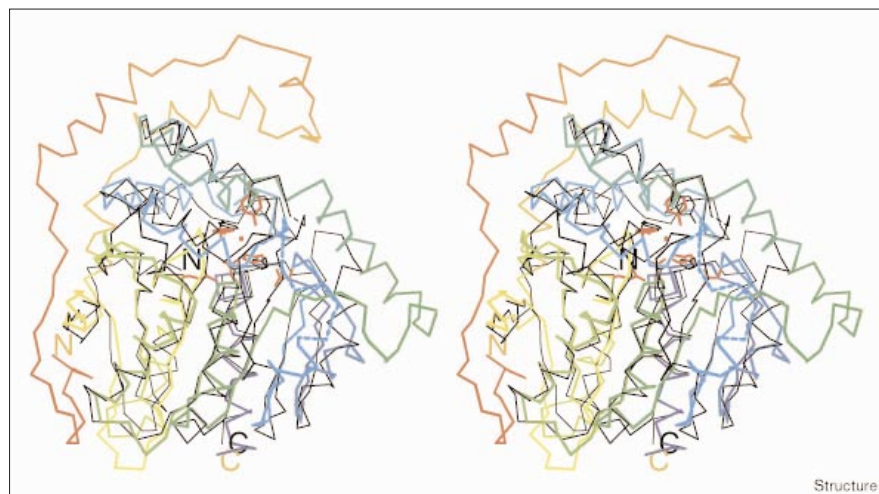
Superposition of the agrobacterial EH (black) on AnEH (rainbow coloring), showing similarity of the active-site region and lid. The known residues of catalysis are included; dashed lines indicate the location of the disordered regions of the two proteins.

Figure 7

Structural alignments of the proteins listed in Table 3, showing the similarity of the α/β domains and the different construction of the lids. AnEH is shown in black, the bacterial EH in red, bromoperoxidase A2 in magenta, chloroperoxidase F in gold, chloroperoxidase L in green, bromoperoxidase A1 in cyan, haloalkane dehalogenase in yellow, hydroxynitrile lyase in blue, proline iminopeptidase in maroon, and the lipase in lime green.

Structure
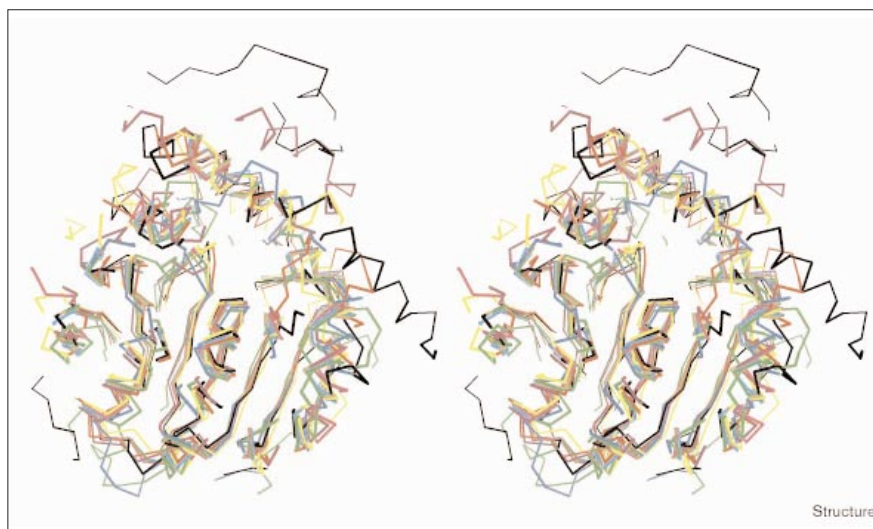
The N-terminal meander of AnEH is preserved intact in mEH, and augmented with a membrane anchor. The extension is characteristic of the subfamily, and its importance in dimer formation in AnEH suggests that mEH will also be dimeric. The low level of the sequence identity, however, makes absolute statements impossible. Because mEH is membrane-bound, this aspect of its structure has so far remained unexplored. The presence of two membrane anchors in an mEH dimer could, however, have important consequences for function, for example, by pre-orienting the enzyme for binding of either substrates or functional partners.

Relationship to other α/β hydrolase structures

Sequence relationships of EHs to other enzymes in the α/β hydrolase family had suggested a structural relationship [6,30] that is confirmed by the AnEH structure. Table 3 shows a comparison of the structure of AnEH to the most similar models currently available in the Protein Data Bank [31].

Greatest similarity among the group members is maintained in the region of the α/β hydrolase fold (Figure 7). This portion bears the nucleophile, either serine or aspartic acid, as required by the chemistry of the reaction. In the EHs, for example, attack by a serine nucleophile (like that found in bromoperoxidases) would result in an overly stable reaction intermediate, and therefore aspartate is used instead. Each of the proteins listed here utilizes a histidine/aspartate charge relay, although the aspartate might be replaced by a glutamate in other family members. The bacterial EH appears to be the only case of an EH/HAD family member for which the catalytic triad is reported to be incomplete [20].

Increasing differences in lid design explain most of the progressive decrease in the number of equivalent residues as the structures diverge. Although each enzyme retains some sort of lid structure, the basic requirements can clearly be met in a number of ways that can be tailored to suit the individual reactions. Only the bacterial EH retains helices in positions equivalent to all six of the lid helices of AnEH. With the exception of the *Candida antarctica* lipase (which lacks a lid), the others possess helices equivalent to αLD, αLE, and (except the hydroxynitrile lyase) αLF. Lid groups that correspond to the active-site tyrosines of AnEH are frequent features, including for example, Trp127 and Trp175 in the haloalkane dehalogenase.

Biological implications

The structure of *Aspergillus niger* epoxide hydrolase (AnEH) is representative of the family of microsomal epoxide hydrolases (mEHs) and thus gives insights into several previously unresolved issues. The shape of the substrate-binding site, with the catalytic nucleophile waiting for the substrate at the end of a relatively narrow hydrophobic tunnel, offers a perfect explanation for the observed inability of this class of EHs to hydrolyze the bulky *trans*-substituted epoxides. In addition, two tyrosine residues are observed in the active site that can bind the ring oxygen of the epoxide and assist catalysis by subsequent proton donation. They can also serve to bind fatty acid amides, which are known inhibitors of mammalian mEHs [32], via an interaction of their carbonyl moiety.

Possible physical interactions of mEHs with other enzymes of the xenobiotic metabolism cascade have long been postulated [33,34]. The N-terminal portion of the protein that precedes the canonical core α/β hydrolase

fold of the family of epoxide hydrolases and haloalkane dehalogenases is unique to the mEHs, and could thus be suspected of mediating such an interaction. Based on the three-dimensional structure, however, it can no longer be favored in this role because of its distance from the substrate entry site. On the other hand, a poorly ordered loop immediately adjacent to this entry (residues 320–328 in the current model), offers a tempting possibility for reversible contacts with, for example, epoxide-generating cytochrome P450-dependent monooxygenases. Suggestions that mEH might be a dimer offer additional intriguing possibilities for the design of biochemical experiments.

Recently, mammalian mEH has been reported to be a membrane transporter for bile acids, although this issue is presently a matter of debate [35–37]. Based on the structure of AnEH, we suggest that mammalian mEH is unlikely to act as a membrane transporter. There is no indication of a structural element other than the mEH N-terminal membrane anchor (not present in AnEH) that could span a lipid membrane.

Finally, the three-dimensional structure of AnEH provides a basis for the rational design of mutants of the enzyme with altered substrate preference and enhanced enantioselectivity, further increasing its value as a biocatalyst in the production of fine chemicals.

Materials and methods
Purification and crystallization
AnEH from *Aspergillus niger* strain LCP521 (EMBL accession number AJ238460) was over-expressed and purified in a manner similar to that described previously [21]. The expression construct was modified to enhance translation efficiency in *Escherichia coli* by improving the binding to the 16S ribosomal RNA [38]. A stretch of 15 nucleotides was inserted into the 5′-end of the AnEH coding region resulting in the sequence *ATG* TCC GCT Ctc atg aat cac aaa gCG TTC (start codon italicized; insert shown in lower case letters). This altered the N-terminal sequence of the protein from MSA*PF* to MSA*LMNHKA*F. Native enzyme was expressed in the *E. coli* strain BL21(DE3): a 3 l culture in terrific broth [39] was supplemented with 100 mg ampicillin per litre, and grown at 12% oxygen saturation, in a fermentor (30°C). T7 RNA polymerase was induced with 0.4 mM IPTG (isopropyl-β-D-thiogalactose) at $OD_{600} = 15$, and the cells harvested 2.5–3 h later. Purification of the native protein was as previously described.

For the MAD experiments, the enzyme was expressed in the met⁻ *E. coli* strain B834(DE3), using M9 minimal medium supplemented with amino acids (5 mg/l each of tryptophan and tyrosine, 40 mg/l of selenomethionine, and 50 mg/l each of the other 17 amino acids), 4 g/l glucose and 100 mg/l ampicillin. Because of the reduced growth of the bacteria under these conditions, IPTG was added at an OD_{600} of 1.0. Purification of Se-AnEH was as described above, except that 1 mM DTT was included in all buffers. Because of its lower solubility, the SeMet–protein had to be eluted from the Phenyl Sepharose column (the second chromatographic step in the purification) with 50% glycerol in 5 mM Tris-HCl pH 8, 0.5 mM EDTA, 1 mM DTT. The incorporation of SeMet was confirmed by a total amino acid analysis (using β-mercaptosulfonic acid hydrolysis to avoid methionine degradation).

Crystallization trials were carried out using the sitting-drop vapour diffusion method [40] in microbridges [41] at 277K. Two commercial kits (Hampton I, Wizard I), based on the sparse matrix designs of Jancarik and Kim [42]; Hampton Research, California) and Safarty and Hol (described in [43]; Emerald BioStructures, Bainbridge Island, Washington), respectively, were used for screening. Crystals grew readily under a variety of conditions. The best-diffracting crystals were obtained by mixing 2 µl of the protein solution (17 mg/ml in a buffer of 10 mM Tris-HCl pH 7.4, 1 mM EDTA, 20 mM NaCl and 0.02% sodium azide) with 2 µl of the reservoir (20% PEG 6000, 0.1 M MES pH 6.0 and 0.1 M unbuffered sodium acetate). Crystals (0.3 × 0.4 × 0.02 mm) grew within one week. The Se-AnEH was concentrated to 30 mg/ml in l0 mM Tris pH 7.4, 1 mM EDTA, 20 mM NaCl, 1 mM DTT, excluding azide, for crystallization trials. Successful conditions were similar to those for the native protein, except that the reservoir contained 10% rather than 20% PEG 6000. β-Mercaptoethanol (0.1%) was initially included in the reservoirs to prevent oxidation of the selenium, in addition to the DTT present in the buffer. This was found to be unnecessary and therefore omitted in later experiments. The drops were streak-seeded [44] with native crystals; a single grain of sand was also added to some drops [45]. Addition of the sand nucleant resulted in thicker crystals with identical morphology.

Data collection, structure solution and refinement
Native crystals were flash-frozen in liquid nitrogen after a 1 min soak in a cryoprotectant solution (30% PEG 6000, 0.1 M MES pH 6.0, 0.1 M sodium acetate). They diffracted to 2.5 Å on a home X-ray source, and to 1.8 Å at the ESRF (Grenoble) ID14 EH4 beam line. The Se-substituted crystals were cryoprotected in the same manner, and a three-wavelength MAD dataset collected on the same ESRF beamline. An X-ray absorption spectrum was collected near the Se K-absorption edge for Se-AnEH by measuring the fluorescent signal perpendicular to the beam during an energy scan. Data were then collected to 2.7 Å resolution at wavelengths corresponding to the inflection point of the spectra (12657 eV), the peak of the spectra (12659 eV), and a remote wavelength at the peak brilliance of the beamline at 13315 eV. The data were acquired on an ADSC Quantum-4 charged coupled device (CCD) detector using 1° oscillations. No measures were taken to record the Bijvoet pairs close in time or to reduce absorption effects by crystal alignment or reverse-beam procedures. The MAD data collection, including change of wavelengths, was completed in 2 h. All images were integrated and scaled using the HKL suite of programs [46]. Statistics relating to crystal quality and data collection are summarized in Table 1.

The sites of 15 of the possible 22 selenium atoms were located, refined and used for phasing in the program package SOLVE [47]. Inspection of a solvent-flattened map generated by DM [48,49] phased to 3.5 Å resolution clearly showed a closely entwined dimer with good solvent separation. An approximate NCS operator was determined by building a number of short equivalent α-helical fragments in O [50]. This operator was improved using IMP [51] and the map cyclically averaged using DM. The first model was built in O using skeletons, mainchain and sidechain databases [52]; numbering was according to the sequence of the wild-type protein, with an asparagine as residue 1 instead of the wild-type methionine (see above). Refinement was carried out with the high-resolution native dataset. CNS [53] was used throughout the refinement with force field parameters derived from the small-molecule database [54]. NCS restraints were relaxed as the refinement progressed. After the second refinement-rebuilding macro-cycle, waters were added using wARP [55]. Only water molecules satisfying the NCS were included in subsequent rounds of refinement. Statistics for the final model are shown in Table 2.

Surface accessibility calculations were carried out using the algorithm of Lee and Richards [56]. Coordinates of similar proteins were obtained from the PDB [31]. Similar proteins were identified using the bromoperoxidase A2 structure as an intermediate step, that is, defined as proteins with Z scores greater than 5.5 to 1BRO in the RCSB PDB

(http://www.rcsb.org/pdb/) according to Combinatorial Extension [57]. These structures were then compared with AnEH using the *brute* and *imp* options of the program LSQMAN [58,59]. Where more than one molecule is present in the asymmetric unit, the A chain was used for the analysis. Figures were prepared with the programs O and OPLOT [50], POV-Ray version 3.1 (http://www.povray.org), and Canvas 6 (Deneba Systems, Inc.).

Accession numbers
Coordinates and structure factors have been deposited at the Protein Data Bank with tracking code 1QO7.

Acknowledgements

References

1. Armstrong, R.N. (1999). Kinetic and chemical mechanism of epoxide hydrolase. *Drug Metab. Rev.* **31**, 71-86.
2. Oesch, F. (1973). Mammalian epoxide hydrases: inducible enzymes catalysing the inactivation of carcinogenic and cytotoxic metabolites derived from aromatic and olefinic compounds. *Xenobiotica* **3**, 305-340.
3. Ota, K. & Hammock, B.D. (1980). Cytosolic and microsomal epoxide hydrolases: differential properties in mammalian liver. *Science* **207**, 1479-1481.
4. Knehr, M., Thomas, H., Arand, M., Gebel, T., Zeller, H.D. & Oesch, F. (1993). Isolation and characterization of a cDNA encoding rat liver cytosolic epoxide hydrolase and its functional expression in *Escherichia coli*. *J. Biol. Chem.* **268**, 17623-17627.
5. Janssen, D.B., Fries, F., van der Ploeg, J., Kazemier, B., Terpstra, P. & Witholt, B. (1989). Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhlA* gene. *J. Bacteriol.* **171**, 6791-6799.
6. Arand, M., Grant, D.F., Beetham, J.K., Friedberg, T., Oesch, F. & Hammock, B.D. (1994). Sequence similarity of mammalian epoxide hydrolases to the bacterial haloalkane dehalogenase and other related proteins. *FEBS Lett.* **338**, 251-256.
7. Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M. *et al.* & Goldman, A. (1992). The α/β hydrolase fold. *Protein Eng.* **5**, 197-211.
8. Heikinheimo, P., Goldman, A., Jeffries, C. & Ollis, D.L. (1999). Of barn owls and bankers: a lush variety of α/β hydrolases. *Structure* **7**, 141-146.
9. Franken, S.M., Rozeboom, H.J., Kalk, K.H. & Dijkstra, B.W. (1991). Crystal structure of haloalkane dehalogenase: an enzyme to detoxify haolgenated alkanes. *EMBO J.* **10**, 1297-1302.
10. Verschueren, K.H.G., Seljée, F., Rozeboom, H.J., Kalk, K.H. & Dijkstra, B.W. (1993). Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* **363**, 693-698.
11. Tsukada, H. & Blow, D.M. (1985). Structure of α-chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.* **184**, 703-711.
12. Lacourciere, G.M. & Armstrong, R.N. (1994). Microsomal and soluble epoxide hydrolases are members of the same family of C–X bond hydrolase enzymes. *Chem. Res. Toxicol.* **7**, 121-124.
13. Pries, F., *et al.,* & Janssen, D.B. (1994). Site-directed mutagenesis and oxygen isotope incorporation studies of the nucleophilic aspartate of haloalkane dehalogenase. *Biochemistry* **33**, 1242-1247.
14. Pinot, F., *et al.,* & Hammock, B.D. (1995). Molecular and biochemical evidence for the involvement of the Asp333–His523 pair in catalytic mechanism of soluble epoxide hydrolase. *J. Biol. Chem.* **270**, 7968-7974.
15. Arand, M., Wagner, H. & Oesch, F. (1996). Asp333, Asp495, and His523 form the catalytic triad of rat soluble epoxide hydrolase. *J. Biol. Chem.* **271**, 4223-4229.
16. Laughlin, L.T., Tzeng, H.F., Lin, S. & Armstrong, R.N. (1998). Mechanism of microsomal epoxide hydrolase. Semifunctional site-specific mutants affecting the alkylation half-reaction. *Biochemistry* **37**, 2897-2904.
17. Tzeng, H.F., Laughlin, L.T. & Armstrong, R.N. (1998). Semifunctional site-specific mutants affecting the hydrolytic half-reaction of microsomal epoxide hydrolase. *Biochemistry* **37**, 2905-2911.
18. Arand, M., *et al.,* & Oesch, F. (1999). Catalytic triad of microsomal epoxide hydrolase: replacement of Glu404 with Asp leads to a strongly increased turnover rate. *Biochem. J.* **337**, 37-43.
19. Barbirato, F., Verdoes, J.C., de Bont, J.A. & van der Werf, M.J. (1998). The *Rhodococcus erythropolis* DCL14 limonene-1,2-epoxide hydrolase gene encodes an enzyme belonging to a novel class of epoxide hydrolases. *FEBS Lett.* **438**, 293-296.
20. Nardini, M., *et al.,* & 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.
21. Arand, M., *et al.,* & Oesch, F. (1999). Cloning and molecular characterization of a soluble epoxide hydrolase from *Aspergillus niger* that is related to mammalian microsomal epoxide hydrolase. *Biochem. J.* **344**, 273-280.
22. Morisseau, C., Archelas, A., Guitton, C., Faucher, D., Furstoss, R. & Baratti, J.C. (1999). Purification and characterization of a highly enantioselective epoxide hydrolase from *Aspergillus niger. Eur. J. Biochem.* **263**, 386-395.
23. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
24. Lacourciere, G.M. & Armstrong, R.N. (1993). The catalytic mechanism of microsomal epoxide hydrolase involves an ester intermediate. *J. Am. Chem. Soc.* **115**, 10466-10467.
25. McGaughey, G.B., Gagne, M. & Rappe, A.K. (1998). pi-Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* **273**, 15458-15463.
26. Pedragosa-Moreau, S., Archelas, A. & Furstoss, R. (1996). Microbiological transformations. 32. Use of epoxide hydrolase mediated biohydrolysis as a way to enantiopure epoxides and vicinal diols: application to substituted styrene oxide derivatives. *Tetrahedron* **52**, 4593-4606.
27. Wang, P., Meijer, J. & Guengerich, F.P. (1982). Purification of human liver cytosolic epoxide hydrolase and comparison to the microsomal enzyme. *Biochemistry* **21**, 5769-5776.
28. Blée, E. & 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.
29. Argiriadi, M.A., Morisseau, C., Hammock, B.D. & Christianson, D.W. (1999). Detoxification of environmental mutagens and carcinogens: structure, mechanism, and evolution of liver epoxide hydrolase. *Proc. Natl Acad. Sci. USA* **96**, 10637-10642.
30. Puente, X.S. & Lopez-Otin, C. (1997). The PLEES proteins: a family of structurally related enzymes widely distributed from bacteria to humans. *Biochem. J.* **322**, 947-949.
31. Bernstein, F.C., *et al.,* & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.
32. Pacifici, G.M., Franchi, M., Bencini, C. & Rane, A. (1986). Valpromide inhibits human epoxide hydrolase. *Brit. J. Clin. Pharmacol.* **22**, 269-274.
33. Oesch, F. & Daly, J. (1972). Conversion of naphthalene to trans-naphthalene dihydrodiol: evidence for the presence of a coupled aryl monooxygenase-epoxide hydrase system in hepatic microsomes. *Biochem. Biophys. Res. Commun.* **46**, 1713-1720.
34. Etter, H.U., Richter, C., Ohta, Y., Winterhalter, K.H., Sasabe, H. & Kawato, S. (1991). Rotation and interaction with epoxide hydrolase of cytochrome P-450 in proteoliposomes. *J. Biol. Chem.* **266**, 18600-18605.
35. Alves, C., von Dippe, P., Amoui, M. & Levy, D. (1993). Bile acid transport into hepatocyte smooth endoplasmic reticulum vesicles is mediated by microsomal epoxide hydrolase, a membrane protein exhibiting two distinct topological orientations. *J. Biol. Chem.* **268**, 20148-20155.
36. Honscha, W., Platte, H.D., Oesch, F. & Friedberg, T. (1995). Relationship between the microsomal epoxide hydrolase and the hepatocellular transport of bile acids and xenobiotics. *Biochem. J.* **311**, 975-979.
37. von Dippe, P., Amoui, M., Stellwagen, R.H. & Levy, D. (1996). The functional expression of sodium-dependent bile acid transport in Madin-Darby canine kidney cells transfected with the cDNA for microsomal epoxide hydrolase. *J. Biol. Chem.* **271**, 18176-18180.
38. Sprengart, M.L., Fuchs, E. & Porter, A.G. (1996). The downstream box: an efficient and independent translation initiation signal in *Escherichia coli. EMBO J.* **15**, 665-674.
39. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular cloning: a*

laboratory manual. 2nd ed. Vol. 1–3. Cold Spring Harbour Laboratory Press, Cold Spring Harbour.

40.   McPherson, A.J. (1982). *Preparation and Analysis of Protein Crystals*. John Wiley and Sons, New York.

41.   Harlos, K. (1992). Micro-bridges for sitting-drop crystallizations. *J Appl. Crystallogr.* **25**, 536-538.

42.   Jancarik, J. & Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409-411.

43.   Bergfors, T. (1998). *Crystallization of Proteins: Techniques, Strategies, and Tips. A Laboratory Manual*. IUL Biotechnology Series, (Bergfors, T.M. ed) La Jolla, CA: International University Line.

44.   Stura, E. & Wilson, I. (1992). Seeding techniques. In *Crystallization of Nucleic Acids and Proteins*. (Ducruix, A. & Giegé, R. eds). IRL Press, Oxford.

45.   Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665-1672.

46.   Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

47.   Terwilliger, T.C. & Berendzen, J. (1999). Automated structure solution for MIR and MAD. *Acta Crystallogr. D* **55**, 849-861.

48.   Cowtan, K. (1994). An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallogr.* **31**, 34-38.

49.   Collaborative Computing Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763.

50.   Jones, T.A. & Kjeldgaard, M.O. (1997). Electron-density map interpretation. *Methods Enzymol.* **277**, 173-208.

51.   Kleywegt, G.J. & Jones, T.A. (1994). Halloween... Masks and Bones, In *From First Map to Final Model*. (Bailey, S., Hubbard, R. & Waller, D., eds.), pp. 59-66, SERC Daresbury Laboratory, Warrington, UK.

52.   Jones, T.A., Zou, J.-Y., Cowan, S.W. & 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.

53.   Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., *et al.* & Warren, G.L. (1998). Crystallography and NMR system (CNS): a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905-921.

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

55.   Perrakis, A., Sixma, T.K., Wilson, K.S. & Lamzin, V.S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr. D* **53**, 448-455.

56.   Lee, B. & Richards, F.M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379-400.

57.   Shindyalov, I.N. & Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11**, 739-747.

58.   Kleywegt, G.J. (1996). Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D* **52**, 842-857.

59.   Kleywegt, G.J. & Jones, T.A. (1997). Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277**, 525-545.

60.   Kleywegt, G.J. & Jones, T.A. (1996). Phi/Psi-cology: Ramachandran revisited. *Structure* **4**, 1395-1400.

61.   Hecht, H.J., Sobek, H., Haag, T., Pfeifer, O. & van Pée, K.-H. (1994). The metal-ion-free oxidoreductase from *Streptomyces aureofaciens* has an α/β hydrolase fold. *Nat. Struct. Biol.* **1**, 532-537.

62.   Hofmann, B., Tolzer, S., Pelletier, I., Altenbuchner, J., van Pee, K.H. & Hecht, H.J. (1998). Structural investigation of the cofactor-free chloroperoxidases. *J. Mol. Biol.* **279**, 889-900.

63.   Wagner, U.G., Hasslacher, M., Griengl, H., Schwab, H. & Kratky, C. (1996). Mechanism of cyanogenesis: the crystal structure of hydroxynitrile lyase from *Hevea brasiliensis. Structure* **4**, 811-822.

64.   Medrano, F.J., Alonso, J., Garcia, J.L., Romero, A., Bode, W. & Gomis-Ruth, F.X. (1998). Structure of proline iminopeptidase from *Xanthomonas campestris* pv. citri: a prototype for the prolyl oligopeptidase family. *EMBO J.* **17**, 1-9.

65.   Uppenberg, J., Hansen, M.T., Patkar, S. & Jones, T.A. (1994). The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica. Structure* **2**, 293-308.