

Divergence and Convergence in Enzyme Evolution: Parallel Evolution of Paraoxonases from Quorum-quenching Lactonases*

Published, JBC Papers in Press, November 8, 2011, DOI 10.1074/jbc.R111.257329

Mikael Elias¹ and Dan S. Tawfik²

From the Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel

We discuss the basic features of divergent *versus* convergent evolution and of the common scenario of parallel evolution. The example of quorum-quenching lactonases is subsequently described. Three different quorum-quenching lactonase families are known, and they belong to three different superfamilies. Their key active-site architectures have converged and are strikingly similar. Curiously, a promiscuous organophosphate hydrolase activity is observed in all three families. We describe the structural and mechanistic features that underline this converged promiscuity and how this promiscuity drove the parallel divergence of organophosphate hydrolases within these lactonase families by either natural or laboratory evolution.

Common Descent

Darwin's theory of evolution is associated mostly with the idea of natural selection, as manifested in the title of his book, *On the Origins of Species by Means of Natural Selection*. However, the more fundamental aspect of Darwinian theory is in fact "common descent," the notion that all extant species diverged from one common ancestor (1). The crux of evolutionary analyses is therefore understanding the routes and mechanisms by which different species diverged from one common ancestor. The ultimate goal is a detailed description of the tree of life, starting from the earliest life forms to the last universal common ancestor (LUCA)³ that appeared over 3.5 billion years ago and subsequently to all extant species. The same principle applies to proteins. We assume that all extant proteins diverged from a rather small set of proteins that existed in the LUCA. The minimal set of the LUCA proteins can be inferred from phylogenetic trees based on the existing protein families and superfamilies. Proteins found in all three kingdoms (Archaea, Eukarya, and Bacteria) and in nearly all organisms are presumed to date back to the progenitor from which these kingdoms have diverged, *i.e.* to the LUCA (2). The LUCA protein set

can also be inferred by identifying the most ancient functions (3). Although small (several hundred, possibly ~1000) (4), the LUCA protein set presumably included the progenitors of the major protein superfamilies that are presently seen in all three kingdoms (Refs. 5–7 and references therein).

The proteins of the pre-LUCA era remain a complete mystery. We assume that there were many fewer proteins than in the LUCA set and that these proteins may have been preceded by short polypeptides within an RNA world (8). Thus, we assume that all existing proteins, be they as different as they are in sequence, structure, and function, are related to one another via a small number of ancestral peptides. The identity of these ancestral peptides is unknown. However, certain sequence motifs, such as the P-loop motif, are present in a large number of different protein superfamilies, suggesting that they were already present in the early protein ancestors (9, 10).

Divergent, Convergent, or Perhaps Parallel Evolution?

Divergent evolution can be traced through sequence and/or fold similarities (see definition in Fig. 1A). Members of the same enzyme family are related by sequence, primarily by conserved motifs that encode key active-site residues (see definition in Fig. 1A). However, sequence diverges far more rapidly than structure, and hence, the assignment of the same fold and active-site chemistry serves as the basis for classifying enzyme superfamilies (11, 12). All members of a given superfamily (see definition in Fig. 1A) are related by divergent evolution, albeit through mostly uncharacterized nodes (Fig. 1B). For several superfamilies, putative routes of divergence have been described, from the LUCA progenitors to the current, highly diverse superfamily (for examples, see Refs. 13–17). However, due to the loss of sequence identity, the ancestor(s) and the precise tree of divergence (the lineages, ancestral nodes, and sequences of the latter) of these superfamilies (let alone of other) cannot be inferred, at least not by the present sequence alignment-based methodologies. It has been speculated that superfamily ancestors exhibited broad substrate specificity and could possibly catalyze a range of reactions (18). However, the current reaction and substrate diversity of superfamilies is so large such that the inference of the ancestral function(s) is a challenging task.

Beyond the superfamily level, common ancestry is much harder to assign. Thus, enzymes belonging to different folds, or even superfamilies, but sharing the same substrate and reaction and strikingly similar active-site arrangements are commonly described as the outcome of convergent evolution (namely the acquisition of the same trait in unrelated lineages) (see definition in Fig. 1A). However, if one assumes one or few common ancestors, then all organisms and all their molecular components, including enzymes, share a common ancestor and are therefore related by divergent evolution. Parallel evolution, *i.e.* the acquisition of the same trait in species descending from the same ancestor but connected by non-continuous lineages (see definition in Fig. 1A), might be a more adequate term for many of these so-called cases of convergent evolution. Thus, although the terms convergent and divergent evolution seem absolute,

* This work was supported in part by the Israel Science Foundation. This is the second article in the Thematic Minireview Series on Enzyme Evolution in the Post-genomic Era.

¹ Supported by Marie Curie Intra-European Fellowship 252836.

² Nella and Leon Benozio Professor of Biochemistry. To whom correspondence should be addressed. E-mail: tawfik@weizmann.ac.il.

³ The abbreviations used are: LUCA, last universal common ancestor; QQL, quorum-quenching lactonase; OPH, organophosphate hydrolase; HSL, *N*-acylhomoserine lactone; PLL, phosphotriesterase-like lactonase; PTE, phosphotriesterase; SsoPox, *S. solfataricus* Pox; PON, serum paraoxonase.

A: Definitions of key terms

Family: A group of proteins sharing the same three-dimensional structure and function. A typical family includes orthologues (proteins from different organisms, sharing the same function and clear sequence identity) as well as close paralogues or isozymes (proteins within the same organism that exhibit closely related functions).

Superfamily: A large group of proteins sharing the same fold (topology of secondary-structure elements), and for most members, the same active site architecture. A typical super-family is comprised of dozens of families, each exhibiting a different function. These functions typically share a common element (*e.g.* a key chemical step in enzyme catalysis) and also the active-site residues executing this element.

Divergent evolution: The evolution of two or more distinctly different protein families from one common ancestor. The diverged families typically belong to the same superfamily, share common structural and functional motifs, and exhibit detectable sequence identity.

Convergent evolution: The emergence of similar functional and structural properties within protein families that did not diverge from the same ancestor. Typically, these families belong to different superfamilies and do not exhibit sequence identity (not even at the level of short sequence motifs).

Parallel evolution: The appearance of similar functional and structural properties within protein families that evolved in separate lineages connected by a distant, ancient ancestor. These families may belong to different superfamilies or can also be part of the same superfamily. They share common structural, functional, but show no detectable sequence identity.

Promiscuous activity: enzymatic activities with substrates and reactions other than the native ones (the substrates and reaction for which the enzyme evolved). Promiscuous activities are purely accidental, and emerged under no selection or physiological relevance.

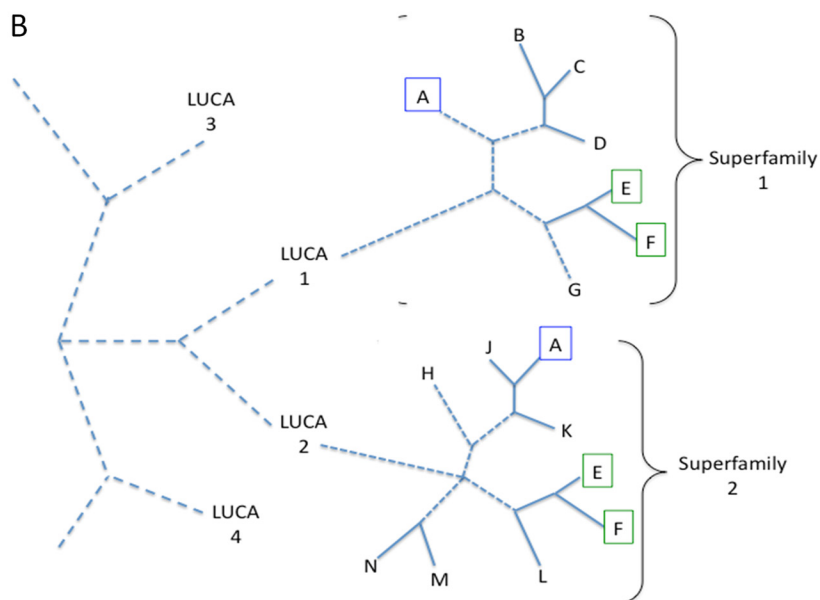


FIGURE 1. **Definition of key terms and schematic tree representation of protein universe.** *A*, definition of key terms used in this minireview. *B*, protein families (including certain orthologs and close paralogues) are represented as *A, B*, etc., and superfamilies, as *1, 2*, etc. The routes of divergence of closely related proteins can be readily inferred by sequence resemblance and are represented as *continuous lines*. However, most families within a given superfamily are so remote in sequence that the routes of divergence remain hypothetical (depicted by *dashed lines*). The progenitors of many of the contemporary protein superfamilies were present in the LUCA. The LUCA proteins diverged from unknown ancestors and by unknown routes, as indicated by *thick dashed lines*. Cases of convergent evolution are represented by family *A*, *i.e.* enzymes with activity *A* that appeared in parallel in both superfamilies *1* and *2*. If, however, the event could be traced back to the node connecting the LUCA ancestors of these two superfamilies, this would be a case of parallel evolution. In some cases, the same family (*e.g.* OPH, depicted as *F*) has diverged in parallel, within two or more superfamilies, from the same ancestral family (QQL, depicted as *E*).

they are not: certain cases of convergent evolution actually concern cases of parallel evolution related by ancestors that are too ancient to be tracked down (Fig. 1*B*).

Two notable examples of convergent enzyme evolution are glycosidases and proteinases (other examples and general aspects of convergent and divergent evolution are discussed in the accompanying minireview by Galperin and Koonin (78)).

Over 100 different families of *O*-glycosyl hydrolases share essentially the same active-site architecture: two juxtapositioned acidic residues (Asp-Glu dyad) at a distance of $\sim 5 \text{ \AA}$ (retaining hydrolases) or $\sim 10 \text{ \AA}$ (invertin hydrolases). This architecture is found within numerous different folds (19). Many of these folds are shared by enzymes exhibiting different activities. For example, glycosyl hydrolase families are com-

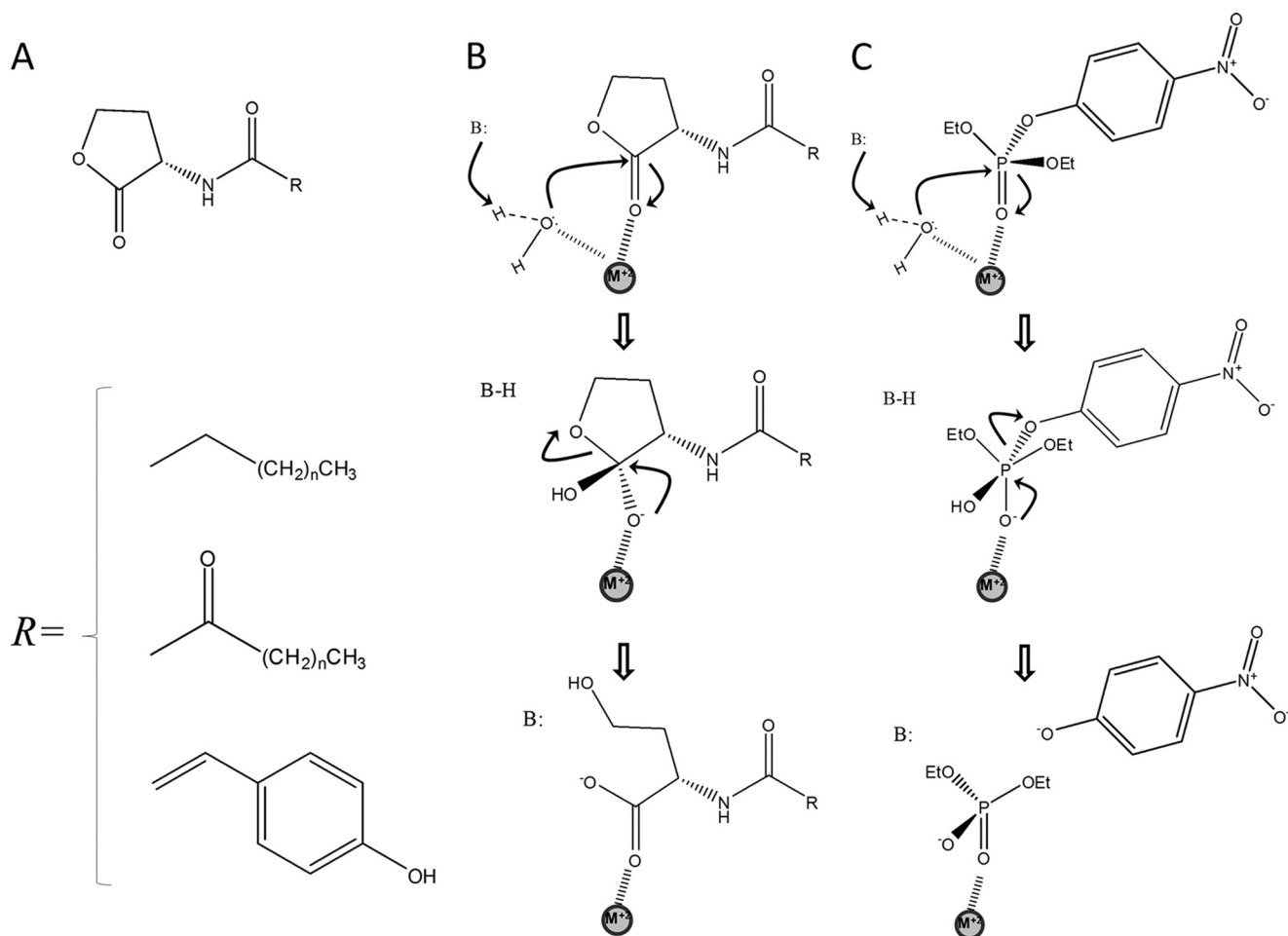


FIGURE 2. **HSLs, their hydrolysis by QQLs, and promiscuous paraoxonase activity of QQLs.** A, structure of HSLs. B, schematic representation of the common catalytic features of the three QQL families described here. The lactone substrate binds to the metal cation via its carbonyl oxygen, thus making the carbonyl carbon more electrophilic. The attacking water molecule is deprotonated either by the active-site metal or by an amino acid side chain acting as a base. The resulting tetrahedral intermediate is subsequently broken (with protonation of the alkoxide leaving group) to give the hydrolyzed product. C, catalytic features of the promiscuous paraoxonase activity of QQLs. Binding of the substrate phosphoryl oxygen and formation and stabilization of the pentacovalent intermediate make use of the same active-site features that mediate the lactonase activity.

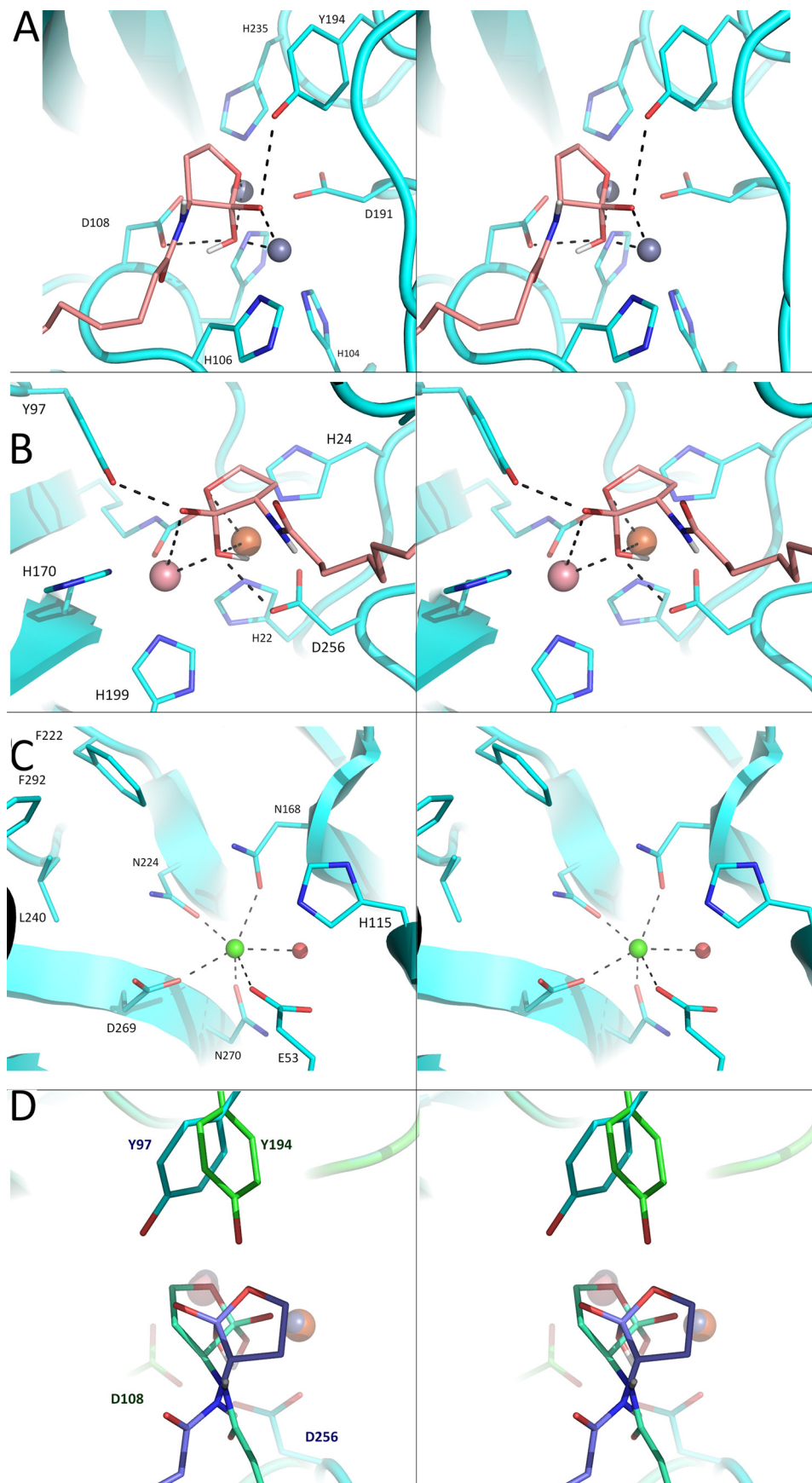
monly found with β/α -barrels (TIM) or β -propeller folds. However, these folds encompass numerous other activities, essentially from all six EC classes. Did these different families diverge from an ancestral protein with an Asp-Glu dyad that later adopted different folds, or did these families evolve in parallel while converging into the same active-site arrangement? The divergence routes within and between superfamilies remain largely unknown, and so are the answers to the above questions.

Another interesting case regards membrane proteinases. Three major active-site arrangements have long been known in ordinary (soluble) proteinases, namely serine/cysteine and aspartic proteinases and metalloproteinases. These three classes have also been identified among membrane proteinases. The active-site architectures of membrane proteinases (e.g. a Ser-His dyad and an oxyanion hole) are highly similar to those of the respective soluble proteinases (20). However, the structure of membrane proteins is fundamentally different. Soluble and membrane proteins may have diverged from a common ancestor, possible via an inside-out rearrangement (21, 22). Did an ancient progenitor of serine/cysteine proteinases diverge

into soluble and membrane forms, or did the soluble and membrane proteinase families evolve in parallel? Although distinguishing between these two scenarios is, at present, impossible, discoveries of new proteinase families may reveal a bridging scenario between the currently unrelated soluble and membrane proteinases.

Quorum-quenching Lactonases

The emergence of quorum-quenching lactonases (QQLs) within different superfamilies constitutes another example of convergent/parallel evolution, and so does the intriguing case of parallel divergence of organophosphate hydrolases (OPHs) from these lactonases. Bacterial quorum sensing is mediated by a wide range of molecules, from small ligands to proteins. *N*-Acylhomoserine lactones (HSLs) were the first class of quorum-sensing molecules to be discovered (Fig. 2A) (23). Their role in mediating the pathogenicity of *Pseudomonas aeruginosa* constitutes a well established experimental model (24). In essence, the binding of a given HSL to a transcription factor (LuxR) turns on the expression of a set of genes that relate to the virulent state. Specificity is determined by the *N*-acyl group,



alkyl chains with varying length and substituents (e.g. 3-oxo group), or *p*-coumaroyl (25, 26). The quorum signal is quenched by lactonases that open the lactone ring or by deacylases that cleave the *N*-acylamide bond (27). QQLs therefore possess a therapeutic potential in preventing biofilm formation and virulence (24).

QQLs are found in Bacteria, Archaea, and Eukarya, including mammals. In some bacterial strains, a QQL is present alongside an HSL-based quorum-sensing machinery. However, in many cases, a QQL is found with no other HSL components (e.g. in *Mycobacterium tuberculosis* (28)). The role of these QQLs might be to intercept the quorum signal of other bacteria or to utilize HSLs as a carbon and nitrogen source.

QQLs exhibit different specificities with respect to short or long *N*-acyl chains and/or for the presence of a 3-oxo group, but some of these enzymes also hydrolyze lactones other than HSLs. As exemplified below, lactonase families appear to have diverged such that some members specialize in hydrolyzing HSLs, whereas other members hydrolyze other lactones (either γ -lactones, as are HSLs, or δ -lactones).

At present, three lactonase families that include QLLs are known in detail. These belong to three different superfamilies, as described below. However, QQLs that belong to other superfamilies, such as AiiM from *Microbacterium testaceum*, which seems to belong to the α/β -hydrolase superfamily (29), have also been identified.

AiiA: Metallo- β -lactamase-like Lactonase

The first QQL to be discovered, AiiA, was found in *Bacillus thuringiensis*. Its structure and mechanism of catalysis have been studied in detail (30–33). AiiA and the related AiiB from *Agrobacterium tumefaciens* (~23% sequence identity) (34) belong to the metallo- β -lactamase superfamily. This superfamily includes enzymes with many different hydrolytic activities (cleavage of C–O/N/S bonds, as well as S–O and P–O), as well as non-hydrolytic activities (35, 36).

The active site of AiiA is composed of a bimetallic center (two zinc cations), coordinated by five histidines and two aspartates (Fig. 3A). The lactone substrate binds the bimetallic center via the two oxygen atoms of its ester group. The carbonyl oxygen also interacts with Tyr-194. The *N*-acyl chain resides in a hydrophobic crevice. A water molecule bridging the two metals constitutes the putative nucleophile that attacks the *sp*² carbonyl carbon of the lactone, thus forming a tetrahedral intermediate (Fig. 2B). This activated intermediate breaks down to yield the ring-open form of carboxylate and alcohol. The metal-

ligating residue Asp-108 is thought to act as a base that activates the bridging water for the attack and subsequently as an acid that protonates the alkoxide leaving group (32, 33).

Phosphotriesterase-like Lactonases

The second known QQL family is named phosphotriesterase-like lactonases (PLLs). It belongs to the amidohydrolase superfamily, a highly diverse superfamily that, as is the case with the metallo- β -lactamase superfamily, includes many different hydrolytic and other activities (37). However, unlike the metallo- β -lactamases, amidohydrolases do not possess a unique fold. The (β/α)₈-barrel (or TIM barrel) fold adopted by all amidohydrolases is the most common fold among enzymes (38) and is shared by many other superfamilies. Whether these different superfamilies diverged from a common ancestor that already possessed a (β/α)₈-barrel fold or whether they converged to the same fold remains unclear (39).

The PLL family was discovered in an attempt to decipher the evolutionary origins of bacterial phosphotriesterases (PTEs; discussed in detail below). Genes showing 26–35% amino acid identity to PTEs turned out to encode lactonases and have thus been named PLLs (28). Indeed, the first family member was biochemically characterized as paraoxonase (40) but was later found to be a lactonase with promiscuous paraoxonase activity (28). Additional members of the PLL family have been discovered (41–43), including members whose substrates are lactones other than HSLs (44). This shift in specificity is not unique though. Divergence of family members that have specialized in hydrolyzing lactones other than HSLs has also occurred within mammalian serum paraoxonases (described below).

SsoPox is a hyperthermophilic PLL from the archaeon *Sulfolobus solfataricus*. Its structure was determined in the free form and also in complex with an HSL analog (45, 46). The active site of *SsoPox* shows striking resemblance to that of AiiA (Fig. 3B). The two metal cations are chelated by four histidines, one aspartic acid, and one carboxylated lysine. A water molecule that bridges the two metal cations comprises the putative nucleophile, and the carbonyl oxygen also interacts with the Tyr-97 hydroxyl group. Finally, Asp-256 in *SsoPox* adopts a very similar position to Asp-108 in AiiA, suggesting that this Asp plays a similar role in catalysis in both lactonase families.

Serum Paraoxonases

In contrast to the first two families, which are bacterial, the third family is primarily mammalian, with few other vertebrate

FIGURE 3. Stereo representations of active-site models of representative lactonases from three known QQL families. A, docking model of AiiA (Protein Data Bank code 2A7M), representing the metallo- β -lactamase QQL family, with the tetrahedral reaction intermediate of *N*-dodecanoylhomoserine lactone. The active-site zinc cations are shown as gray spheres. The carbonyl oxygen interacts with one of the zincs, as well as with the hydroxyl of Tyr-104. The intermediate's hydroxyl interacts with the other zinc cation and with the metal-ligating residue Asp-108, which acts as a base (to generate the attacking hydroxide) and subsequently as an acid (to protonate the alkoxide leaving group). B, same model for *SsoPox* (Protein Data Bank code 2VC5), representing the PLL family. Tyr-97 and Asp-256 play equivalent roles to Tyr-104 and Asp-108 in AiiA. The spheres represent an iron (orange) and a cobalt (pink) cation, although this enzyme is active with other transition metals. C, active site of PON1 (Protein Data Bank code 1V04), representing the PON family. Shown are the catalytic calcium cation (green sphere) and His-115, which is thought to act as a base in generating the attacking hydroxide. Reasonably accurate docking models of the bound lactone could not be generated because the only available PON structure is without a ligand, at pH 4.5 when the enzyme is inactive, and with an active-site loop missing. D, AiiA and *SsoPox* active sites are mirror images of one another. The AiiA (green) and *SsoPox* (blue) active-site models were manually aligned based on their bimetallic centers and their bridging water molecules. Docking models were generated with AutoDock 4.0 (77). The docked ligands were generated using J.Ligand and DockingServer. A single negative charge was attributed to the intermediate's carbonyl oxygen atom and +2 for the metal cations. The remaining charges for both the intermediate and the enzyme residues were attributed using the Gasteiger method. The docking calculations were performed using a Lamarckian genetic algorithm. The images were generated with PyMOL.

representatives (47). As is the case with PLLs, the identification of serum paraoxonases (PONs) and the family's name relate to their paraoxon-hydrolyzing activity. Although the earliest identification might have been of a lactonase (48), it was only much later realized that PONs are in fact lactonases (49, 50) that also exhibit some quorum-quenching activity (51). The paraoxonase activity exhibited by PONs turned out to be a promiscuous coincidental activity (50, 52). Examination of all three PON families revealed that PON2 (an isozyme distributed in all human tissues) may indeed specialize in hydrolyzing HSLs with long acyl chains (53, 54). The remaining two isozymes, PON1 and PON3, found in the liver and blood, show lower activity toward HSLs. Their primary substrates seem to be δ -lactones and/or γ -lactones with lipophilic side chains (49, 50, 55).

PONs adopt a six-bladed β -propeller fold and a central tunnel with two calcium cations, one playing a structural role and the other serving in catalysis (56). The catalytic calcium is coordinated by three asparagines, one glutamate, and one aspartate (Fig. 3C). As is the case with the other two families, the PON active-site metal cation aligns the lactone molecule via its carbonyl oxygen and stabilizes the oxyanion tetrahedral intermediate that is formed upon hydroxide attack (Fig. 2B). The latter is thought to be generated via deprotonation of a water molecule by His-115 (assisted by His-134 in a dyad format) (57).

Convergence of Lactonase Mechanistic and Structural Features

The three lactonase families described above share no sequence or fold similarity, but their catalytic chemistries are similar (Figs. 2 and 3). In particular, AiiA and PLLs constitute a remarkable example of convergent or parallel evolution. These two families not only share the same chemistry but also exhibit essentially identical active-site architectures. Indeed, the models for the bound reaction intermediates for the representatives of these two families, *SsoPox* and AiiA, show highly similar interaction networks, including the interactions with the bimetallic center, the metal-ligating Asp residue that acts as a base/acid, and the tyrosine hydroxyl group (Fig. 3, A and B). In fact, the model suggests that these enzyme active-sites are mirror images of one another (Fig. 3D).

Convergence of Promiscuous Paraoxonase Activity

That the native activity of QQLs evolved in parallel within three different superfamilies (and possibly more) is obvious. The different QQL families diverged from different hydrolases that were available as starting points in different organisms in which hydrolyzing HSLs turned out to be advantageous. The ancestral forms of these lactonase families remain unknown. However, at least for AiiA and PLLs, the superfamily context suggests that these were hydrolases and, most likely, C–O or C–NH hydrolases. Convergence of the active-site features may therefore stem from similarities in the active-site architectures of the starting points and from having the same function, namely hydrolysis of HSLs.

In addition to their native function (HSL hydrolysis in the case of QQLs), enzymes exhibit promiscuous functions that have never been selected for (18, 58). These side activities result from the fact that no protein is likely to exhibit absolute speci-

ficity, and substrates other than the native one may bind and get transformed with low efficiency (18). By definition (see Fig. 1A), promiscuous functions are functions that have not been selected for and are therefore coincidental (18). A curious observation regarding QQLs and the related lactonases is that, irrespective of the fold and specific active-site arrangement, members from all three families described above have been found to exhibit promiscuous paraoxonase activity (and OPH activity in general). In fact, as described above, two of the three known families, PLLs and PONs, were initially identified as paraoxonases and were only later defined as lactonases. Representing the third family, AiiA may also exhibit promiscuous paraoxonase activity.⁴ Furthermore, numerous hydrolases (serine hydrolases in particular) react with organophosphates such as paraoxon. However, the reaction occurs with no or extremely slow turnover primarily because the phosphoryl-enzyme intermediate is not further hydrolyzed. In contrast, the lactonases described here hydrolyze organophosphates with turnover.

The paraoxonase activity is unlikely to have emerged under selection. Paraoxon is a newly introduced, man-made substance present in distinct environments only. Phosphonodiester, the closest natural compounds to paraoxon (a phosphotriester), were thus far found only in specific marine environments (59). It is therefore unlikely that the paraoxonase activity was selected for in diverse organisms ranging from archaea to humans. Furthermore, the magnitude of paraoxonase activity dramatically varies from one family member to another as expected for a coincidental activity. The turnover numbers (k_{cat}) with paraoxon vary from 0.5 up to $10^4 \text{ M}^{-1} \text{ s}^{-1}$ for bacterial PLLs (28) and from non-detectable up to $10^4 \text{ M}^{-1} \text{ s}^{-1}$ for PONs (49, 50).

Why Do Lactonases Exhibit Promiscuous Paraoxonase Activity?

The convergence of the promiscuous paraoxonase activity seems to stem from a fortuitous overlap between the substrates and intermediates for the native promiscuous reactions. Lactones have an sp^2 carbonyl group that becomes tetrahedral (sp^3) upon nucleophilic attack by a hydroxide ion or by an active-site catalytic side chain. Paraoxon possesses a tetrahedral configuration at the ground state that becomes bipyramidal upon attack (Fig. 2). How do these different substrates and intermediates overlap?

Crystal structures of members from all three families described above are available, some with substrate analogs. There are no structures, however, of the same enzyme with both a lactone and organophosphate analog. We therefore generated docking models of HSL and paraoxon and of their respective reaction intermediates in the active sites of AiiA and *SsoPox*. As observed in the structural models of HSL and its reaction intermediate (Fig. 3, A and B), a very similar binding mode is observed for paraoxon in both enzymes. Specifically, in both enzymes, superposition of the reaction intermediates for the lactone and paraoxon results in the following: (i) the ethyl ester substituent of paraoxon transition state overlaps the ester group of the lactone, and (ii) the phenoxy

⁴ H.-S. Kim, personal communication.

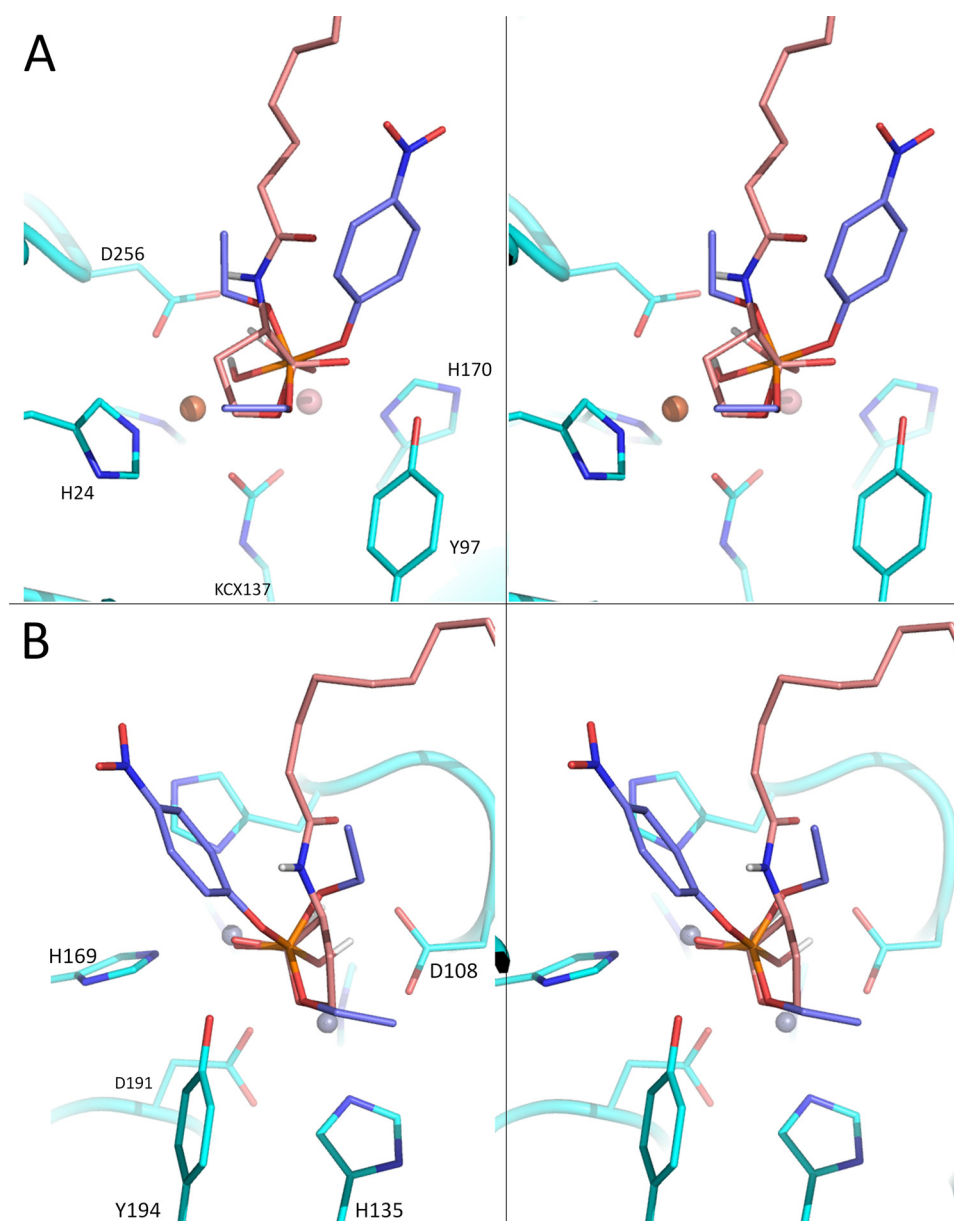


FIGURE 4. **Stereo view of superposition of lactonase and paraoxonase reaction intermediates.** Presented are docking models of SsoPox (Protein Data Bank code 2VC5; A) and AiiA (code 2A7M; B) with the reaction intermediates that form upon a nucleophilic attack by hydroxide. The structures of the reaction intermediates for *N*-acylhomoserine (carbon atoms in pink) and paraoxon (carbon atoms in blue) are aligned. Note that the intermediate for lactone hydrolysis is tetrahedral, and that for paraoxon is bipyramidal.

leaving group of paraoxon occupies a similar position as the lactone oxyanion (Fig. 4, A and B).

These active-site models, which we hope will be experimentally validated in the future, suggest that the appearance of the promiscuous paraoxonase activity in three evolutionarily independent lactonase families is not a coincidence. Rather, it is an outcome of overlaps in specific substrate, reaction, and active-site features. These overlaps go beyond the obvious, *e.g.* the attack by hydroxide or stabilization of the oxyanion intermediates of lactones and paraoxon. The lactone structure, as opposed to ordinary (non-cyclic) esters, appears to maximize this overlap, as most esterases do not exhibit promiscuous paraoxonase activity with multiple turnovers (28).

Parallel Evolution of Paraoxonases

Latent promiscuous functions provide the starting points for the evolution of new protein functions (Refs. 18 and 58; see also the accompanying minireview by Copley (79)). Indeed, the paraoxonase promiscuity of lactonases seems to have been exploited by nature as a starting point for the evolution of new enzymes that specialize as paraoxonases. These paraoxonases seem to have diverged only few decades after the introduction of parathion, the first widely used organophosphate pesticide. The product of the spontaneous oxidation of parathion, paraoxon, is a potent inhibitor of insect acetylcholinesterase. By the 1950s, parathion was routinely used in many countries. Microorganisms that metabolize parathion and/or paraoxon have been reported as early as the mid-1960s (60). Their enzymes

degrade organophosphates, presumably to supplement shortages in inorganic phosphate (61).

By 1981, the first bacterial paraoxonase, called PTE or organophosphate-degrading enzyme, was identified in *Pseudomonas diminuta* (62) on a plasmid that comprises part of a transposable element (63, 64). Closely related enzymes ($\geq 95\%$ identity) were found in other bacteria (65, 66). Paraoxon (or methyl paraoxon (66)) is by far the best substrate for these enzymes, with rates approaching the diffusion limit ($k_{\text{cat}}/K_m \geq 4 \times 10^7$) (67). Parathion is also hydrolyzed by PTE with very similar kinetic parameters.⁵ How could proficient and specialized paraoxonases such as PTE evolve in a matter of a few decades?

The ancestor of PTE must have exhibited promiscuous paraoxonase activity that served as the starting point for a specialized paraoxonase. Along the same vein, promiscuous activities of PTE may comprise vestiges of the native activity of its progenitor. Among other promiscuous hydrolytic activities, PTE was found to exhibit relatively high lactonase activity (68). This suggested that PTE could have diverged from a lactonase. Indeed, genes exhibiting $\sim 30\%$ sequence identity to PTE and that were annotated as “putative PTEs” turned out to be members of a newly identified family of lactonases, which we named PLLs. All key active-site residues are shared between PTE and PLL (28), and all currently known PLLs exhibit promiscuous paraoxonase activity (28, 41–43). It was therefore assumed that a yet unknown PLL gave rise to PTE via a series of sequence modifications, including insertion and deletions in the active-site loops (28).

A different bacterial strain (*Pseudomonas* sp. WBC-3) that can grow on methyl parathion as a sole carbon and nitrogen source was isolated near a pesticide-manufacturing site in China. Its PTE seems to have evolved toward degradation of methyl parathion ($k_{\text{cat}}/K_m \sim 10^6 \text{ M}^{-1} \text{ s}^{-1}$). Methyl parathion hydrolases with $>90\%$ amino acid identity were isolated from similar sites. The sequence and structure indicate that this family of paraoxonases belongs to the metallo- β -lactamase superfamily (69). Our analysis indicates that AiiA (Protein Data Bank code 3DHA; 24.1% sequence identity) and AiiB (code 2R2D; 16.4% sequence identity) are among the best hits in a search for related structures and sequences in the Molecular Modeling Database (70). We surmise that this homology is no coincidence and that the parathion hydrolases probably diverged from a QQL ancestor that resembles AiiA in structure, function, and sequence. One might also speculate that other enzymes whose structures are closely similar to AiiA but whose functions are unknown (e.g. Protein Data Bank codes 3ESH, 2ZWR, and 3KLD) are either lactonases or OPHs.

In the case of the third family, the mammalian PONs, no specialized paraoxonase has been thus far identified. However, laboratory evolution experiments have unambiguously shown that at least two family members, PON1 and PON3, could be readily evolved into a highly proficient OPH (71, 72). Notably, this functional transition is mediated by mutations in His-115, a residue that is thought to act as the catalytic base (Fig. 3C). Mutations of

His-115 dramatically reduce the lactonase activity while improving the OPH activity (57). The laboratory experiments therefore parallel the two natural examples in diverging a specialized OPH from a lactonase that belongs to yet another superfamily.

Concluding Remarks

Through structural and functional similarities, the fingerprints of divergent evolution are evident in the two cases of recently evolved paraoxonases described above. Because the divergence of paraoxonases is a relatively recent event, one also expects to find lactonases that are highly similar in sequence. However, the currently known lactonases and the related paraoxonases exhibit $<25\%$ sequence identity. This is not surprising, however. The available sampling of protein sequences is extremely small and highly biased, thus resulting in members of the same family being related by as little as 25% identity (e.g. AiiA and AiiB). The likelihood of discovering the actual ancestors of the paraoxonases that evolved only 2 decades ago is therefore slim, not to mention the identification of divergence events that occurred hundreds of millions of years ago.

The prediction that the PTE ancestor is a QQL was also based on considerable luck and somewhat prepared minds. Nonetheless, the functional, structural, and laboratory evolutionary linkages between these two seemingly unrelated classes of enzymes, lactonases and paraoxonases, suggest that the parallel evolution of paraoxonases in two different superfamilies is more than a coincidence. It therefore appears that the identification of patterns of overlapping activities, both native and promiscuous, between different families within the same superfamily constitutes a powerful means of deciphering evolutionary relationships (12, 68, 73–76). As discussed here, these patterns relate to specific overlaps between transition states and/or intermediates of different reactions (see also Refs. 12 and 73–76). These overlaps are not random, as suggested by the fact that they are observed in three different folds and active-site architectures.

Further exploration of patterns of overlapping activities, alongside a systematic mapping of sequences, functions, and structures within superfamilies (see also the accompanying minireview by Gerlt *et al.* (80)), may enable the identification of more cases of divergent evolution. The ultimate goal is the generation of a unified tree that describes the divergence of all protein families and superfamilies from a few early progenitors. How far-reaching (or perhaps far-fetched) is the task of unraveling the complete history of protein evolution based on the anecdotal and biased information we currently have? This is best described by Charles Darwin’s analogy: “I look at the geological record as a history of the world imperfectly kept and written in a changing dialect. Of this history we possess the last volume alone, relating only to two or three countries. Of this volume, only here and there a short chapter has been preserved, and of each page, only here and there a few lines.”

Acknowledgments—We thank Livnat Afriat and Colin Jackson for valuable input. While writing this minireview, D. S. T. was on sabbatical at the Université Joseph Fourier and the Institut de Biologie Structurale (Grenoble, France).

⁵ L. Afriat, unpublished data.

REFERENCES

1. Theobald, D. L. (2010) *Nature* **465**, 219–222
2. Koonin, E. V., and Galperin, M. Y. (2002) *Sequence-Evolution-Function: Computational Approaches in Comparative Genomics*, 1st Ed., Springer, New York
3. Kim, K. M., and Caetano-Anollés, G. (2010) *Mol. Biol. Evol.* **27**, 1710–1733
4. Ouzounis, C. A., Kunin, V., Darzentas, N., and Goldovsky, L. (2006) *Res. Microbiol.* **157**, 57–68
5. Koonin, E. V. (2003) *Nat. Rev. Microbiol.* **1**, 127–136
6. Bolton, D. P., and Marsh, J. (1984) *J. Physiol.* **351**, 451–459
7. Tuller, T., Birin, H., Gophna, U., Kupiec, M., and Ruppín, E. (2010) *Genome Res.* **20**, 122–132
8. Kun, Á., Pongor, S., Jordán, F., and Szathmáry, E. (2008) in *The Codes of Life* (Barbieri, M., and Hoffmeyer, J. eds) pp. 39–58, Springer-Verlag, Amsterdam
9. Lupas, A. N., Ponting, C. P., and Russell, R. B. (2001) *J. Struct. Biol.* **134**, 191–203
10. Trifonov, E. N., and Frenkel, Z. M. (2009) *Curr. Opin. Struct. Biol.* **19**, 335–340
11. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) *Science* **267**, 1159–1161
12. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) *Curr. Opin. Chem. Biol.* **10**, 492–497
13. Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) *J. Mol. Biol.* **317**, 41–72
14. Aravind, L., Anantharaman, V., and Koonin, E. V. (2002) *Proteins* **48**, 1–14
15. Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004) *J. Struct. Biol.* **146**, 11–31
16. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) *J. Mol. Biol.* **361**, 1003–1034
17. Iyer, L. M., Abhiman, S., de Souza, R. F., and Aravind, L. (2010) *Nucleic Acids Res.* **38**, 5261–5279
18. Khersonsky, O., and Tawfik, D. S. (2010) *Annu. Rev. Biochem.* **79**, 471–505
19. Davies, G., and Henrissat, B. (1995) *Structure* **3**, 853–859
20. Erez, E., Fass, D., and Bibi, E. (2009) *Nature* **459**, 371–378
21. Engelman, D. M., and Zaccari, G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5894–5898
22. Blobel, G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1496–1500
23. Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994) *J. Bacteriol.* **176**, 269–275
24. Estin, M. L., Stoltz, D. A., and Zabner, J. (2010) *Adv. Exp. Med. Biol.* **660**, 183–193
25. Jayaraman, A., and Wood, T. K. (2008) *Annu. Rev. Biomed. Eng.* **10**, 145–167
26. Schaefer, A. L., Greenberg, E. P., Oliver, C. M., Oda, Y., Huang, J. J., Bittan-Banin, G., Peres, C. M., Schmidt, S., Juhaszova, K., Sufrin, J. R., and Harwood, C. S. (2008) *Nature* **454**, 595–599
27. Bokhove, M., Nadal Jimenez, P., Quax, W. J., and Dijkstra, B. W. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107**, 686–691
28. Afriat, L., Roodveldt, C., Manco, G., and Tawfik, D. S. (2006) *Biochemistry* **45**, 13677–13686
29. Wang, W. Z., Morohoshi, T., Ikenoya, M., Someya, N., and Ikeda, T. (2010) *Appl. Environ. Microbiol.* **76**, 2524–2530
30. Liu, D., Lepore, B. W., Petsko, G. A., Thomas, P. W., Stone, E. M., Fast, W., and Ringe, D. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11882–11887
31. Kim, M. H., Choi, W. C., Kang, H. O., Lee, J. S., Kang, B. S., Kim, K. J., Derewenda, Z. S., Oh, T. K., Lee, C. H., and Lee, J. K. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17606–17611
32. Liu, D., Momb, J., Thomas, P. W., Moulin, A., Petsko, G. A., Fast, W., and Ringe, D. (2008) *Biochemistry* **47**, 7706–7714
33. Momb, J., Wang, C., Liu, D., Thomas, P. W., Petsko, G. A., Guo, H., Ringe, D., and Fast, W. (2008) *Biochemistry* **47**, 7715–7725
34. Liu, D., Thomas, P. W., Momb, J., Hoang, Q. Q., Petsko, G. A., Ringe, D., and Fast, W. (2007) *Biochemistry* **46**, 11789–11799
35. Daiyasu, H., Osaka, K., Ishino, Y., and Toh, H. (2001) *FEBS Lett.* **503**, 1–6
36. Bebrone, C. (2007) *Biochem. Pharmacol.* **74**, 1686–1701
37. Seibert, C. M., and Raushel, F. M. (2005) *Biochemistry* **44**, 6383–6391
38. Wierenga, R. K. (2001) *FEBS Lett.* **492**, 193–198
39. Gerlt, J. A., and Raushel, F. M. (2003) *Curr. Opin. Chem. Biol.* **7**, 252–264
40. Merone, L., Mandrich, L., Rossi, M., and Manco, G. (2005) *Extremophiles* **9**, 297–305
41. Chow, J. Y., Wu, L., and Yew, W. S. (2009) *Biochemistry* **48**, 4344–4353
42. Chow, J. Y., Xue, B., Lee, K. H., Tung, A., Wu, L., Robinson, R. C., and Yew, W. S. (2010) *J. Biol. Chem.* **285**, 40911–40920
43. Hawwa, R., Larsen, S. D., Ratia, K., and Mesecar, A. D. (2009) *J. Mol. Biol.* **393**, 36–57
44. Xiang, D. F., Kolb, P., Fedorov, A. A., Meier, M. M., Fedorov, L. V., Nguyen, T. T., Sterner, R., Almo, S. C., Shoichet, B. K., and Raushel, F. M. (2009) *Biochemistry* **48**, 2237–2247
45. Del Vecchio, P., Elias, M., Merone, L., Graziano, G., Dupuy, J., Mandrich, L., Carullo, P., Fournier, B., Rochu, D., Rossi, M., Masson, P., Chabriere, E., and Manco, G. (2009) *Extremophiles* **13**, 461–470
46. Elias, M., Dupuy, J., Merone, L., Mandrich, L., Porzio, E., Moniot, S., Rochu, D., Lecomte, C., Rossi, M., Masson, P., Manco, G., and Chabriere, E. (2008) *J. Mol. Biol.* **379**, 1017–1028
47. Draganov, D. I., and La Du, B. N. (2004) *Naunyn-Schmiedeberg Arch. Pharmacol.* **369**, 78–88
48. Roth, R. H., Levy, R., and Giarmar, N. J. (1967) *Biochem. Pharmacol.* **16**, 596–598
49. Draganov, D. I., Teiber, J. F., Speelman, A., Osawa, Y., Sunahara, R., and La Du, B. N. (2005) *J. Lipid Res.* **46**, 1239–1247
50. Khersonsky, O., and Tawfik, D. S. (2005) *Biochemistry* **44**, 6371–6382
51. Yang, F., Wang, L. H., Wang, J., Dong, Y. H., Hu, J. Y., and Zhang, L. H. (2005) *FEBS Lett.* **579**, 3713–3717
52. Draganov, D. I. (2010) *Chem. Biol. Interact.* **187**, 370–372
53. Teiber, J. F., and Draganov, D. I. (2011) *Methods Mol. Biol.* **692**, 291–298
54. Teiber, J. F., Horke, S., Haines, D. C., Chowdhary, P. K., Xiao, J., Kramer, G. L., Haley, R. W., and Draganov, D. I. (2008) *Infect. Immun.* **76**, 2512–2519
55. Connelly, P. W., Picardo, C. M., Potter, P. M., Teiber, J. F., Maguire, G. F., and Ng, D. S. (2011) *Biochim. Biophys. Acta* **1811**, 39–45
56. Harel, M., Aharoni, A., Gaidukov, L., Brumshtein, B., Khersonsky, O., Meged, R., Dvir, H., Ravelli, R. B., McCarthy, A., Toker, L., Silman, I., Sussman, J. L., and Tawfik, D. S. (2004) *Nat. Struct. Mol. Biol.* **11**, 412–419
57. Khersonsky, O., and Tawfik, D. S. (2006) *J. Biol. Chem.* **281**, 7649–7656
58. O'Brien, P. J., and Herschlag, D. (1999) *Chem. Biol.* **6**, R91–R105
59. Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B., and Webb, E. A. (2006) *Nature* **439**, 68–71
60. Munnecke, D. M., and Hsieh, D. P. (1974) *Appl. Microbiol.* **28**, 212–217
61. Yang, K., and Metcalf, W. W. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7919–7924
62. Serdar, C. M., Gibson, D. T., Munnecke, D. M., and Lancaster, J. H. (1982) *Appl. Environ. Microbiol.* **44**, 246–249
63. Horne, I., Qiu, X., Russell, R. J., and Oakeshott, J. G. (2003) *FEMS Microbiol. Lett.* **222**, 1–8
64. Siddavattam, D., Khajamohiddin, S., Manavathi, B., Pakala, S. B., and Merrick, M. (2003) *Appl. Environ. Microbiol.* **69**, 2533–2539
65. Mulbry, W., Ahrens, E., and Karns, J. (1998) *Pesticide Sci.* **52**, 268–274
66. Yang, H., Carr, P. D., McLoughlin, S. Y., Liu, J. W., Horne, I., Qiu, X., Jeffries, C. M., Russell, R. J., Oakeshott, J. G., and Ollis, D. L. (2003) *Protein Eng.* **16**, 135–145
67. Omburo, G. A., Kuo, J. M., Mullins, L. S., and Raushel, F. M. (1992) *J. Biol. Chem.* **267**, 13278–13283
68. Roodveldt, C., and Tawfik, D. S. (2005) *Biochemistry* **44**, 12728–12736
69. Dong, Y. J., Bartlam, M., Sun, L., Zhou, Y. F., Zhang, Z. P., Zhang, C. G., Rao, Z., and Zhang, X. E. (2005) *J. Mol. Biol.* **353**, 655–663
70. Wang, Y., Adress, K. J., Chen, J., Geer, L. Y., He, J., He, S., Lu, S., Madej, T., Marchler-Bauer, A., Thiessen, P. A., Zhang, N., and Bryant, S. H. (2007) *Nucleic Acids Res.* **35**, D298–D300
71. Aharoni, A., Gaidukov, L., Yagur, S., Toker, L., Silman, I., and Tawfik, D. S. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 482–487
72. Gupta, R. D., Goldsmith, M., Ashani, Y., Simo, Y., Mullokandov, G., Bar, H., Ben-David, M., Leader, H., Margalit, R., Silman, I., Sussman, J. L., and

MINIREVIEW: *Parallel Evolution of Paraoxonases and Lactonases*

- Tawfik, D. S. (2011) *Nat. Chem. Biol.* **7**, 120–125
73. Schmidt, D. M., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) *Biochemistry* **42**, 8387–8393
74. Taylor Ringia, E. A., Garrett, J. B., Thoden, J. B., Holden, H. M., Rayment, I., and Gerlt, J. A. (2004) *Biochemistry* **43**, 224–229
75. Poelarends, G. J., Serrano, H., Johnson, W. H., Jr., Hoffman, D. W., and Whitman, C. P. (2004) *J. Am. Chem. Soc.* **126**, 15658–15659
76. Wang, S. C., Johnson, W. H., Jr., and Whitman, C. P. (2003) *J. Am. Chem. Soc.* **125**, 14282–14283
77. Huey, R., Morris, G. M., Olson, A. J., and Goodsell, D. S. (2007) *J. Comput. Chem.* **28**, 1145–1152
78. Galperin, M. Y., and Koonin, E. V. (2012) *J. Biol. Chem.* **287**, 21–28
79. Copley, S. D. (2012) *J. Biol. Chem.* **287**, 3–10
80. Gerlt, J. A., Babbitt, P. C., Jacobson, M. P., and Almo, S. C. (2012) *J. Biol. Chem.* **287**, 29–34