

Structure and evolution of the serum paraoxonase family of detoxifying and anti-atherosclerotic enzymes

Michal Harel^{1,5}, Amir Aharoni^{3,5}, Leonid Gaidukov³, Boris Brumshtein¹, Olga Khersonsky³, Ran Meged¹, Hay Dvir^{1,2}, Raimond B G Ravelli⁴, Andrew McCarthy⁴, Lilly Toker², Israel Silman², Joel L Sussman¹ & Dan S Tawfik³

Members of the serum paraoxonase (PON) family have been identified in mammals and other vertebrates, and in invertebrates. PONs exhibit a wide range of physiologically important hydrolytic activities, including drug metabolism and detoxification of nerve agents. PON1 and PON3 reside on high-density lipoprotein (HDL, 'good cholesterol') and are involved in the prevention of atherosclerosis. We describe the first crystal structure of a PON family member, a variant of PON1 obtained by directed evolution, at a resolution of 2.2 Å. PON1 is a six-bladed β -propeller with a unique active site lid that is also involved in HDL binding. The three-dimensional structure and directed evolution studies permit a detailed description of PON1's active site and catalytic mechanism, which are reminiscent of secreted phospholipase A2, and of the routes by which PON family members diverged toward different substrate and reaction selectivities.

Serum paraoxonase (PON1) is a mammalian enzyme that catalyzes the hydrolysis and thereby the inactivation of various organophosphates including the nerve agents sarin and soman¹. In recent years PON1 and its two known Q and R isoenzymes have been found to play important roles in drug metabolism and in the prevention of atherosclerosis^{1,2}. PON1 is the best-studied member of a family of mammalian enzymes that includes PON2 and PON3, which share ~60% sequence identity with PON1. PON1 and PON3 reside in the cholesterol-carrying particles HDL (good cholesterol), whereas PON2 is found in many tissues. Polymorphism of the PON1 gene affects the blood levels of PON1 and its catalytic efficacy; both factors strongly affect an individual's susceptibility to atherosclerosis and to pollutants and insecticides¹. Mice lacking PON1 are highly susceptible to atherosclerosis and organophosphate poisoning³. *In vitro* assays show that PON1 and PON3 inhibit lipid oxidation in low-density lipoprotein (LDL, 'bad cholesterol'), thus reducing levels of oxidized lipids that are involved in the initiation of atherosclerosis^{4,5}. Because atherosclerosis is the underlying cause of 50% of mortality in Western societies, and organophosphates present an environmental risk as well as a terrorist threat, PONs have become the subject of intensive research.

Despite many efforts, the structure and mechanism of action of PONs have remained unclear. The name, paraoxonase, is purely historical, as the PON family is a hydrolase family with one of the broadest specificities known. PON1 is a proficient esterase toward several synthetic substrates, whereas PON2 and PON3 exhibit high lactonase activity. But the paraoxonase activity of PON1 is rather weak, and PON2 and PON3 exhibit almost no paraoxonase activity¹. However,

all of these activities toward man-made chemicals are promiscuous activities of PONs rather than their primary functions. A variety of physiological roles have been proposed for PONs, including phospholipase A2 action (in hydrolysis of platelet-activating factor (PAF)⁶ and of oxidized lipids⁷), and hydrolysis and inactivation of homocysteine thiolactone, a risk factor for atherosclerotic vascular disease⁸. The anti-atherosclerotic activity of PONs is closely linked to their localization on HDL particles. HDL has two key roles: mediation of cholesterol efflux, such as from macrophage foam cells in atherosclerotic lesions, and limitation of lipid oxidation in LDL⁹. PONs have been implicated in both^{2,10,11}. It has been suggested that the hydrophobic N terminus of PON1 mediates its anchoring to HDL¹², but the mode of binding of PONs and other HDL-associated proteins to HDL is still unknown.

It has also been unclear how PONs exhibit such a broad spectrum of activities and whether these activities are mediated by a single active site—and if so, what dictates the substrate selectivities, and how did PONs diverge toward their individual activities. The three-dimensional structure of PONs and their catalytic mechanism were thus far unknown. Structural and functional characterization of the PONs and their engineering were hindered by lack of an ample source of recombinant protein. We recently described the directed evolution of PON1 and PON3 variants that express in a soluble and active form in *Escherichia coli* and have enzymatic properties identical to those reported for PONs purified from sera¹³. We report the crystal structure of a recombinant PON1 variant derived from rabbit PON1, which is highly similar to human PON1. By combining directed evolution,

Departments of ¹Structural Biology, ²Neurobiology and ³Biological Chemistry, The Weizmann Institute of Science, Rehovot 76 100, Israel. ⁴European Molecular Biology Laboratory, Grenoble outstation, 38042 Grenoble Cedex 9, France. ⁵These authors contributed equally to this work. Correspondence should be addressed to J.L.S. (joel.sussman@weizmann.ac.il) or D.S.T. (tawfik@weizmann.ac.il).

Published online 18 April 2004; doi:10.1038/nsmb767

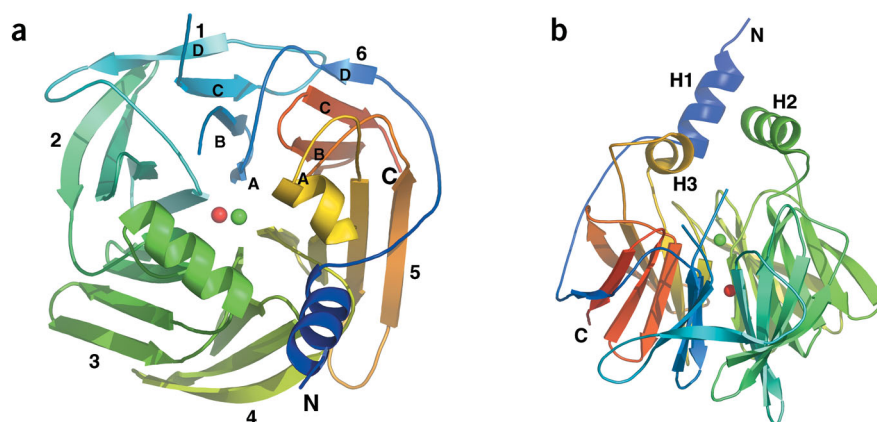


Figure 1 Overall structure of PON1. (a) View of the six-bladed β -propeller from above. The top of the propeller is, by convention, the face carrying the loops connecting the outer β -strand of each blade (strand D) with the inner strand (A) of the next blade¹⁹. Shown are the N and C termini, and the two calcium atoms in the central tunnel of the propeller (Ca1, green; Ca2, red). (b) A side view of the propeller, including the three helices at the top of the propeller (H1–H3). N-terminal residues 1–15 and a surface loop connecting strands 1B and 1C (residues 72–79) are not visible in the structure.

site-directed mutagenesis and kinetic studies, we provide a comprehensive description of the overall architecture of PON1, details of its active site, catalytic mechanism and HDL-binding mode, and the evolution of members of the PON family. We show how directed evolution in the laboratory follows in the footsteps of natural evolution by swiftly providing new PON variants with predetermined catalytic specificities.

RESULTS

Crystallization and structure determination

Previous attempts to determine the structure of PON1 relied on limited amounts of serum-purified proteins and led to the crystallization of a protein that copurified with it¹⁴. Human PON1 is rather unstable, and tends to aggregate in the absence of detergents¹⁵. These factors led us to identify PON variants for bacterial expression and increased solubility by directed evolution¹³. Family shuffling of four PON1 genes (human, rabbit, mouse and rat) and screening for esterolytic activity led to recombinant PON1 variants (rePON1s) that express in *E. coli*. These variants diverged from wild-type rabbit PON1 by 14–31 amino acids from other PON1 genes. The rePON1 variants had enzymatic properties essentially identical to those of wild-type PON1 (ref. 13) and similar biological activities in inhibiting LDL oxidation and mediating macrophage cholesterol efflux (M. Aviram, The Technion Faculty of Medicine, personal communication). Variants from the first round of evolution (such as G1A5 and G1C4 (ref. 13)) aggregated, and none crystallized. The second-generation variants (obtained by shuffling of the first-generation variants and screening for the highest expression levels) did not aggregate, and at least one (G2E6) gave stable and well-diffracting crystals.

RePON1-G2E6 has 91% identity to wild-type rabbit PON1, with the vast majority of variations deriving from human, mouse or rat wild-type PON1. Rabbit and human PON1s are also very similar in sequence (86%) and function¹⁶. Sequence variations between rePON1-G2E6 and rabbit and human PON1 are in regions that do not affect their active sites and overall structures (for detailed analysis, see **Supplementary Note** online). The refined crystal structure of rePON1 at a resolution of 2.2 Å contains one molecule per asymmetric unit. It was solved by single isomorphous replacement and anomalous scattering from data collected on crystals of the native protein and the

selenomethionine (SeMet) protein at a resolution of 2.6 Å (see Methods). The structure shows all residues except N-terminal residues 1–15 and a surface loop (residues 72–79). Two calcium atoms, a phosphate ion and 115 water molecules are also seen.

The overall architecture of PON1

PON1 is a six-bladed β -propeller, and each blade contains four strands (Fig. 1). The ‘velcro’ closure characteristic of this fold¹⁷ is complemented by a disulfide bridge between Cys42 (strand 6D) and Cys353 (strand 6C). This covalent closure of the N and C termini is rarely seen in β -propellers with more than four blades, but is conserved throughout the PON family.

Two calcium ions, 7.4 Å apart, are seen in the central tunnel of the propeller, one at the top (Ca1) and one in the central section (Ca2). Ca2 is most probably a ‘structural calcium’ whose dissociation leads to irreversible

denaturation¹⁸ (**Supplementary Fig. 1** online). Ca1 is assigned as the ‘catalytic calcium’¹⁸. It seems to interact with five protein residues (the side chain oxygens of Asn224, Asn270, Asn168, Asp269 and Glu53), 2.2–2.5 Å away. Two other potential ligands are a water molecule and one of the oxygens of a phosphate ion. The two calcium ions exhibit markedly different affinities¹⁶. Ca1’s ligation is more extensive than Ca2’s. However, two of Ca1’s ligating residues (Asn224 and Asp269) exhibit distorted dihedral angles. This and the higher solvent accessibility of Ca1 indicate that Ca2 is the higher-affinity calcium.

PON1’s structure resembles the *Loligo vulgaris* DFPase¹⁹. Both are six-bladed propellers with two calcium atoms in their central tunnel. They also share functional homology, as both have phosphotriesterase (PTE) activity, although PON1 is primarily an esterase or lactonase. However, there is no clear sequence similarity between them (BLAST *E*-score $\gg 3.6$), although more sensitive algorithms indicate weak but significant similarity¹⁴. This is not surprising: low sequence similarity is a distinct characteristic of β -propellers¹⁷. DFPase, for example, has no sequence similarity to other six-blade β -propellers¹⁹. Closer inspection reveals that PON1 and DFPase differ markedly in their overall architecture, active site structure and mechanism. Most distinctly, PON1 has a unique addition in the form of an active site canopy defined by helices H2 and H3 and the loops connecting them to the β -propeller scaffold. This addition provides PON with an uncharacteristically closed active site, since most β -propellers, including DFPase, have uncovered active sites defined only by loops that connect the β -strands. This addition seems to have a critical role in PON1’s function, in defining the active site architecture and sequestering it from solvent, and in anchoring PON1 to the HDL particle.

Detergent-solubilized PON1 forms dimers and higher oligomers¹⁵, but in the crystal there is only one molecule per asymmetric unit, and very few contacts between symmetry-related molecules. Crystallization may favor a monomeric form, but it seems more likely that oligomerization of PON1 is a consequence of its anchoring to detergent micelles in a mode similar to its anchoring to HDL. PON1 expressed in animal cells is glycosylated¹. Glycosylation is not essential for the hydrolytic activities of PONs^{13,20} but may be important in increasing their solubility and stability, or in preventing nonspecific binding to cell membranes (as proposed for other HDL-associated enzymes²¹). There are four potential *N*-glycosylation sites on PON1

(NX(S/T) sites). Two (Asn227 and Asn270) are in the central tunnel of the propeller and are largely inaccessible to solvent. Asn253 and Asn324 are located on surface loops, and are most probably (as previously proposed²⁰) PON1's glycosylation sites.

Directed evolution of substrate specificity

Site-directed mutagenesis is routinely used to identify active site residues. This approach suffers, however, from a well-recognized drawback: loss of activity does not necessarily indicate direct involvement of a particular amino acid in the protein's function because mutations often disrupt the overall structure. Indeed, whereas certain residues identified by site-directed mutagenesis as essential for PON1's activity^{20,22,23} are related to its active site, others are not (such as Trp281). In contrast, mutations identified after directed evolution toward a modified function are inevitably relevant to activity, and involve residues located within or near the active site.

Gene libraries of rePON1 were prepared by random mutagenesis and cloned in bacterial colonies as described¹³. Colonies on agar plates were screened with several different substrates representing each of the substrate-reaction types catalyzed by PONs: PTE, lactonase and esterase. For the latter, both acetate and octanoate esters were used, representing, respectively, 'short chain' esters toward which PON1 exhibits high activity, and 'long chain' esters, typical substrates of lipases for which PONs exhibit lower activity. The best clones identified from a screen of 10^3 – 10^4 library clones were shuffled and the resulting libraries screened for the same activity. Two to four rounds of selection were done, after which the sequences and catalytic activities of the selected variants were determined (Table 1).

The newly evolved variants clearly define a set of amino acids whose alteration markedly shifts PON1's reactivity and substrate selectivity (Table 1). These shifts involve not only an increase of 16–46-fold in activity toward the substrate for which each particular variant was evolved, but also a marked decrease in activity on substrates that had not been selected for (6–167-fold). Overall, shifts in substrate selectivity of up to 4,600-fold were observed (such as variant 7HY, which exhibits, relative to wild-type PON1, 46-fold higher long-chain esterase activity and 100-fold lower PTE activity). Some of the new variants, all derived from PON1, represent substrate and reaction selectivities closer to PON2 or PON3. For example, variants 2AC and 1HT exhibit ~20-fold higher esterase and lactonase activity relative to PON1, and 530-fold weaker PTE activity. The positions identified by directed evolution all appear in the same region at the top of the β -propeller, clearly delineating the entrance to and walls of PON1's active site (Fig. 2). The amino acids identified by the directed evolution also provide insights into how the substrate selectivity of the PON family members evolved in nature (see below).

The catalytic mechanism

At the very bottom of the active site cavity lies the upper calcium (Ca1), and a phosphate ion that was present in the mother liquor (Fig. 2b; Supplementary Fig. 2 online). One of the phosphate's oxygens is only 2.2 Å from Ca1. This phosphate ion may be bound in a mode similar to that of the intermediates in the hydrolytic reactions

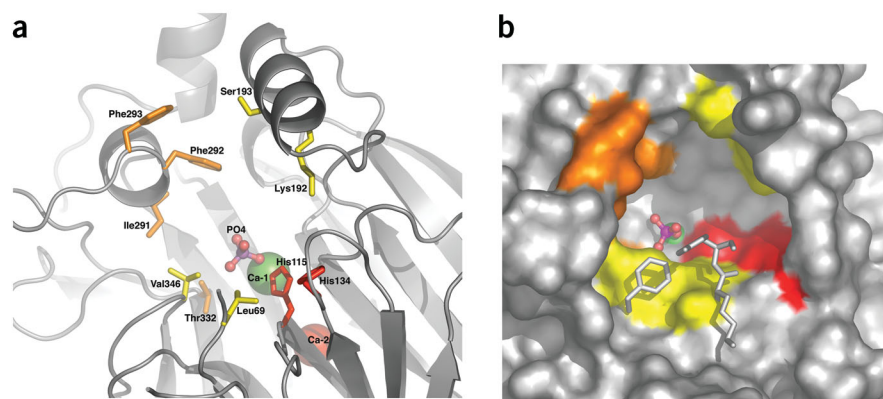


Figure 2 PON1's active site viewed from above the propeller. **(a)** Central tunnel of the propeller with the two calcium atoms, and the side chains of the residues found to be mutated in the newly evolved PON1 variants for esterase and lactonase (orange) or for phosphotriesterase activity (yellow), including the R192Q human polymorphism (in the rePON1-G2E6 variant, this position is a lysine). The putative catalytic His-His dyad is red (see text and Fig. 4). **(b)** A surface view of the active site. Lys70, Tyr71 and Phe347 are shown as sticks to permit a better view of the active site. At the deepest point of the cavity lies the upper calcium atom (Ca1, green) to which a phosphate ion is bound.

catalyzed by PON. One of its negatively charged oxygens, the one nearest to Ca1, may mimic the oxyanionic moiety of these intermediates, which is stabilized by the positively charged calcium. This type of 'oxyanion hole' is present in secreted phospholipase A2 (PLA2)²⁴, and has also been suggested to occur in DFPase¹⁹. Two other phosphate oxygens may be mimicking the attacking hydroxyl ion and the oxygen of the alkoxy or phenoxy leaving groups of ester and lactone substrates.

To help elucidate PON1's mechanism, we determined its pH-rate profile with two typical substrates: an ester (2-naphthyl acetate, 2NA) and a phosphotriester (paraoxon). Both profiles exhibit a bell-shaped curve (Fig. 3). The minor basic shoulder fits an apparent pK_a of 9.8 (paraoxon) or 9.0 (2NA), probably reflecting the deprotonation of a basic side chain that affects the active site but is not directly involved in catalysis. The fully pronounced acidic shoulder, with an apparent pK_a of ~7.1, may be ascribed to a histidine imidazole involved in a base-catalyzed, rate-determining step. In hydrolytic enzymes, histidine often serves as a base, deprotonating a water molecule and generating the attacking hydroxide ion that produces hydrolysis. In secreted PLA2, the attacking hydroxide is generated by a His-Asp dyad, in which the imidazole acts as a base to deprotonate a water molecule, and the aspartate carboxylate increases the imidazole's basicity via a proton-shuttle mechanism. The closest histidine nitrogen in PLA2 is 6.3 Å from the catalytic calcium, and two water molecules are involved: one attacks the substrate (after deprotonation) and another 'catalytic water' mediates between the attacking water and the histidine base²⁴. The DFPase active site also contains a His-Glu dyad with a histidine nitrogen 7.2 Å from the catalytic calcium¹⁹.

Did PON1 adopt the same mechanism, namely a His-Glu/Asp dyad acting as base on a cascade of two water molecules? PON1's active site contains such a dyad (Asp183-His184). However, His184 is ~11 Å from Ca1, and the H184N mutant of human PON1 is active²⁰. Another putative dyad is His285-Asp269. Yet Asp269 ligates Ca1, and His285 is ~8 Å from Ca1 and ~5 Å from the nearest phosphate oxygen. We identified, however, a His-His dyad near both Ca1 and the phosphate ion (Fig. 4). We hypothesize that His115 (the closer nitrogen of which is only 4.1 Å from Ca1) acts as a general base to deprotonate a single water molecule and generate the attacking hydroxide, whereas His134 acts in a proton shuttle mechanism to increase His115's

basicity. Notably, His115 adopts distorted dihedral angles, a phenomenon observed in catalytic residues of many enzymes. In support of the postulated mechanism are the H115Q mutation, which caused a marked decrease ($\sim 2 \times 10^4$ -fold) in activity, and the H134Q mutation, which produced a milder, yet substantial, decrease (6–150-fold; **Supplementary Table 1** online).

Notably, Cys284, which is conserved in all PONs, has been proposed to serve in alternative functions of PON1 related to atherosclerosis²⁵. Cys284 is part of a highly conserved stretch that includes active site His285, and is packed against four highly conserved residues from the adjacent strands (Leu267, Val268, Glu303 and Leu305; **Fig. 5**). Because it is buried it is unlikely to have a functional role. Its mutation, however, is likely to destabilize the core structure, thereby affecting function indirectly. Indeed, we found that Cys284 mutants of rePON1 are poorly expressed and relatively unstable (**Supplementary Table 1** online).

The PON family members

Initially identified in mammals, PON and PON-related genes have now been found in fowl, zebrafish and even in invertebrates such as *Caenorhabditis elegans*¹. The more closely related PON genes are divided into three classes or subfamilies: PON1, PON2 and PON3. Polymorphic forms of PON are also very common: over 200 single nucleotide polymorphisms (SNPs) have been identified in the human PON1 gene alone¹. The three-dimensional structure of PON1 provides insights into how the substrate and reaction selectivity of different PONs is determined, and how they diverged toward different activities. Further insights are provided into the possible effects of various SNPs on PON's activity and stability.

The residues maintaining the hydrophobic core of the β -propeller, its central tunnel, the two calciums and the 'velcro' closure are conserved throughout the family (**Fig. 5**). The key elements of the catalytic site are also highly conserved; these include the residues that ligate Ca1, neighboring residues that form hydrogen bonds with the latter, and the catalytic histidines (**Fig. 5**). Thus, the PON subfamilies diverged while maintaining their overall active site structure and catalytic machinery. However, their substrate and reaction selectivities

Table 1 New PON1 variants generated by directed evolution

Variant	Phosphotriesterase activity ^a	Lactonase activity ^a	Esterase activity ^a		
			Short-chain ester	Long-chain esters	
rePON1 (wild-type-like activity)	3.5×10^3	1.4×10^2	3.0×10^4	1.7×10^2	
Variants with 'specialized' substrate selectivities obtained by directed evolution					
No.	Mutations ^b				
7PC	V346A	1.3×10^4 (3.7) ^b	5.0×10^1 (0.36)	1.4×10^3 (0.05)	1.4×10^1 (0.08)
4PC	L69V, S193P, V346A	5.7×10^4 (16.3)	0.9 (0.006)	4.4×10^2 (0.015)	ND
1HT	I291L, T332A, G339E	6.0×10^2 (0.17)	3.0×10^3 (25.9)	8.6×10^3 (0.3)	ND
2AC	F292S, V346M, V30A, E249K	1.1×10^2 (0.03)	6.4 (0.04)	6.0×10^5 (20)	7.0×10^2 (4.1)
7HY	F292V, Y293D, I109M	4.1×10^1 (0.01)	4.1 (0.03)	1.2×10^5 (4.0)	8.0×10^3 (47)
4HY	I74L, F292L, K84Q, I343M	5.3×10^1 (0.015)	5.9 (0.04)	5.2×10^4 (1.7)	6.5×10^3 (38)

^aActivities were determined at 0.1 mM substrate, where substrate concentration is well under K_m (k_{cat}/K_m conditions) and are expressed as μmol product released per min per mg enzyme. In parentheses are the activities of the new variants relative to wild-type PON1's activity on the same substrate. ^bMutations are given in relation to the sequence of the wild-type-like variant G3C9. In bold are positions found to be mutated in all the highest-activity variants for a given substrate. Typically, the same mutations could be individually identified in the sequence of selected variants from the first and second rounds of evolution, and appear together in the third-generation variants. Mutations that appear in only one selected variant, but not in others selected for the same substrate, and/or do not appear in the first and second round of evolution, are noted in regular print. ND, not detected.

changed markedly. Our results show how PONs readily adopt new selectivities. PON1 variants evolved in the laboratory show patterns reminiscent of PON2 or PON3 (such as variants 1HT and 2AC, with high lactonase and esterase activity and low PTE activity). Our results define a set of 16 residues that make up the walls and perimeter of the PON active site, and thus govern substrate selectivity (**Table 2**). Variants exhibiting patterns of activities that have not yet been identified in natural PONs carry mutations at these positions, but to amino acids other than those observed in wild-type PONs (such as 4PC, which has higher phosphotriesterase PTE activity than wild-type PON1 and markedly lower esterase activity; **Table 1**). At some stage in evolution, changes in the selectivity-determining residues (**Table 2**) led to divergence of the individual PON subfamilies, each of which is highly conserved with respect to these residues. We presume that each of the subfamilies evolved, and is evolutionarily preserved, for a different substrate of key physiological importance, yet the identities of these substrates remain obscure.

There are also residues outside the active site that vary from one subfamily to another, yet are conserved within each family. An example is Asn253, one of the two presumed glycosylation sites of PON1. This site is abolished in PON2 and PON3 owing to a change in the third position of the NX(S/T) consensus sequence from threonine to asparagine or aspartate (**Fig. 5**). Other residues or clusters of residues seem to be specific for each subfamily (such as in the region of positions 20–50). These may be linked to nonhydrolytic roles of PONs and to their localization (for example, PON1 and PON3 are exclusively localized in the liver and in HDL, whereas PON2 is found in many tissues).

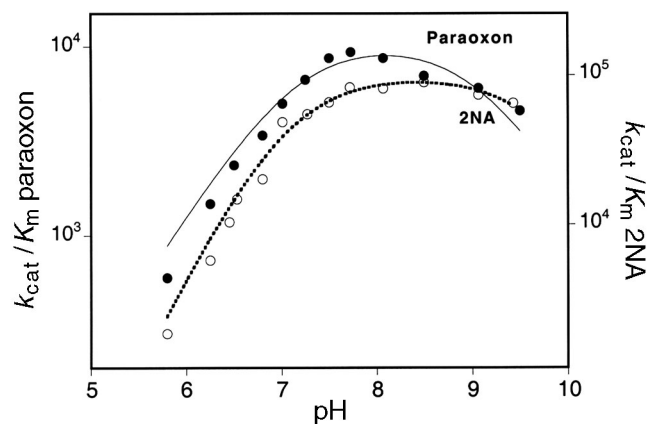


Figure 3 pH-rate profiles of rePON1-G2E6 with 2-naphthyl acetate (2NA, \circ) and paraoxon (\bullet). The data fit a bell-shaped model with the following values for paraoxon: $(k_{cat}/K_m)^{\text{max}} = 7,016 \text{ M}^{-1} \text{ s}^{-1}$, $\text{p}K_a^1 = 7.06$ and $\text{p}K_a^2 = 9.78$; and for 2NA: $(k_{cat}/K_m)^{\text{max}} = 1.67 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $\text{p}K_a^1 = 7.15$ and $\text{p}K_a^2 = 9.03$.

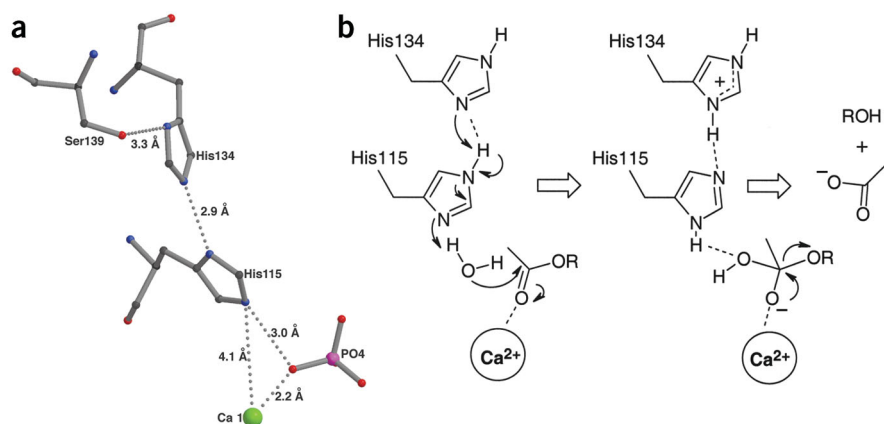


Figure 4 The postulated catalytic site and mechanism of PON1. (a) The catalytic site includes the upper calcium atom (Ca1), the phosphate ion at the bottom of the active site and the postulated His-His dyad. (b) Schematic representation of the proposed mechanism of action of PON1 on ester substrates such as phenyl and 2-naphthylacetate. The first step involves deprotonation of a water molecule by the His-His dyad to generate an hydroxide anion that attacks the ester carbonyl, producing an oxyanionic tetrahedral intermediate. This intermediate breaks down (second step) to an acetate ion and either phenol or 2-naphthol.

Some remote PON family members in bacteria and fungi exhibit functional and presumably structural resemblance to mammalian PONs²⁶. Nature has also recruited the PON scaffold for completely different tasks. For example, MEC-6 (a *C. elegans* protein²⁷) maintains the key structural elements of PON1 (most notably, the Cys42-Cys353 disulfide bridge and two of the three residues that ligate the structural calcium). The hydrolytic site, however, including the residues that ligate Ca1 and the His-His dyad, were mutated away.

Anchoring of PON1 to HDL

PON1 and PON3 are synthesized in the liver and secreted into the blood, where they specifically associate with HDL. HDL mediates reverse transport of cholesterol from peripheral cells and limits LDL oxidation by the action of HDL-associated enzymes such as platelet-activating factor acetylhydrolase (PAF-AH), PON1 and PON3 (refs. 9,28). HDL is a particle of diameter ~10 nm composed primarily of membrane components (phospholipids, cholesterol and cholesterol esters), and apolipoprotein A-I (apoA-I)²⁹, the amphipathic helices of which are thought to wrap around the particle's membrane-like bilayer like a belt³⁰. Several other proteins are associated with HDL²⁸, including lecithin-cholesterol acyl transferase (LCAT)²¹, but their mode of binding to HDL is still under investigation. PON1 is the first HDL-associated protein whose three-dimensional structure has been solved.

PON1 retains its hydrophobic N terminus, which resembles a signal peptide and is thought to be involved in anchoring of PON1 to HDL¹². Most of the N terminus is disordered and invisible in the crystal structure, yet its hydrophilic part, which extends beyond the signal peptide (residues 19–28), adopts a helical structure (H1). The entire sequence of the N terminus is compatible with a transmembrane helix, yet following a secondary structure prediction, we modeled only residues 7–18 as part of H1 (Fig. 6). Helix H2, adjacent to H1, has a clear amphipathic character. Unexpectedly, however, its hydrophobic face points toward the solvent, as do

several residues from the two loops that connect H2 to the propeller scaffold. Helices H1 and H2 form, therefore, two adjacent hydrophobic patches that provide a potential membrane-binding surface (Fig. 6a). The interface with HDL was further defined by a characteristic 'aromatic belt' rich in tryptophan and tyrosine side chains, and by a lysine side chain on H1 (ref. 31). Notably, the glycosylation sites point away from the interface (Fig. 6b).

DISCUSSION

This study complements earlier examples of the application of directed evolution to protein crystallization³². Directed evolution was also applied toward the identification of PON1's active site, and provided key insights into how the substrate selectivity of the PON family members evolved in nature.

The crystal structure reveals the overall fold of the PON family and the details of PON1's structure and can be used to propose a catalytic mechanism. Such a mechanism, based on a His-His dyad, has not been described before, although its key elements are reminiscent of secreted PLA2. Future research should allow verification and refinement of PON1's mechanism, particularly regarding its wide substrate range. Catalysis of both C-O and P-O hydrolyses at one site is unusual but not unprecedented^{33,34}. The structure, the directed evolution results,

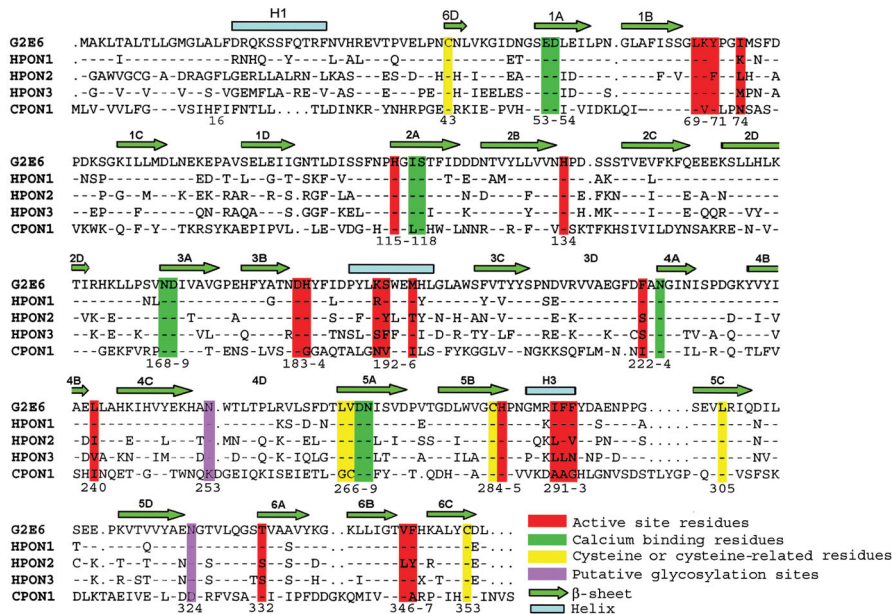


Figure 5 Sequence alignment of representative members of the PON family. Shown are human PON1, PON2 and PON3 (with an 'H' prefix), *C. elegans* PON1 (CPON1), and rePON1 variant G2E6, the structure of which is reported here.

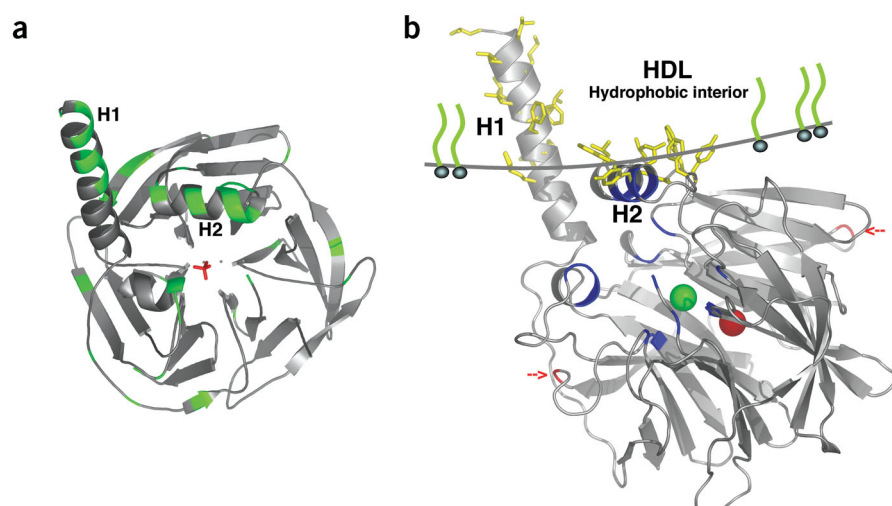


Figure 6 Proposed model for anchoring of PON1 to the surface of HDL. **(a)** Tertiary structure cartoon of rePON1 showing its exposed hydrophobic surfaces. N-terminal residues 7–18, missing in the crystal structure and predicted to be helical, were modeled as part of H1. Denoted are all the hydrophobic residues (leucine, phenylalanine, proline, isoleucine, tyrosine, tryptophan and valine) appearing with accessible surface area $\geq 20 \text{ \AA}^2$. **(b)** Hydrophobic residues proposed to be involved in HDL anchoring (side chains yellow). The line defined by the side chains of Tyr185, Phe 186, Tyr190, Trp194, Trp202 (helix H2 and the adjacent loops) and Lys21 (helix H1) models the putative interface between HDL's hydrophobic interior and the exterior aqueous phase. The hydrophobic side chains of leucine and phenylalanine residues of H1 are primarily within the apolar region³¹. The active site and the selectivity-determining residues (**Table 2**) are blue, and the proposed glycosylation sites (Asn253 and Asn324) are red.

the pH-rate profiles and previous biochemical data^{1,13} show that both these activities take place at the same site. At this stage, however, a possibility remains that certain PON1 activities (for example, as a homocysteine thiolactonase⁸) make use of a different subset of residues of this site, including His285, whose side chain also points toward the center of the cavity and to the phosphate ion. In addition, nucleophilic catalysis by His115 cannot yet be ruled out, although there is currently no evidence to support catalysis via an acyl- or phosphoryl-enzyme intermediate.

The three-dimensional structure does provide a hint regarding the origins of PON1's notably wide substrate range. Hydrophobicity is common to almost all of PON1's effective substrates. The hydrophobicity and depth of PON1's active site explain this preference, and account for the fact that PON1's substrates, whether poor or effective, have K_m values in the millimolar range, but markedly different k_{cat} values^{1,13}. PON1's multispecificity is, therefore, driven primarily by non-specific hydrophobic forces, as has been observed for other enzymes with deep hydrophobic active sites, such as acetylcholinesterase³⁵. We postulate that poor and effective substrates bind at the active site with similar affinity; yet the mode of binding differs, as the poor substrates are inadequately positioned relative to Ca1 and to the catalytic base. The mutations observed after the directed evolution process probably reshape the active site walls and perimeter, thereby improving the positioning of some substrates (and of their respective catalytic intermediates and transition states) and worsening that of others. Reshaping of the active site walls also drives the divergence of various PON subfamilies in nature.

Polymorphism of the PON gene and its effects on susceptibility to organophosphate poisoning and to atherosclerosis are the subject of intensive research. The two most common PON1 forms are R192Q and L55M. The structure reveals that Lys192 is part of the active site wall. In human PON1, this position is normally arginine, but a com-

monly observed polymorphism (R192Q) causes about a ten-fold decrease in paraoxonase activity and higher susceptibility to atherosclerosis¹. Given the marked effects that changes in other active site residues have on its substrate selectivity (**Table 1**), the human R192Q variant may indeed have substantially reduced activities toward PON1's physiological substrates, causing increased susceptibility to atherosclerosis. The L55M mutation may considerably affect PON1's stability and thereby account for the lower enzymatic activity³⁶. This is due to the key role of Leu55 in packing the propeller's central tunnel, and of its neighboring residues (Glu53 and Asp54), which ligate both Ca1 and Ca2. Future work will provide a more detailed characterization of these mutations and their *in vivo* activities.

Several features of PON1 are reminiscent of lipases, and of secreted PLA2 in particular. The postulated model of HDL anchoring supports a mode of interfacial activation whereby HDL anchoring modifies PON1's active site. The top of the active site is partly covered by the backbones of Lys70 and Tyr71. These residues form the end of a loop that is disordered and hence not seen in the structure (residues 72–79). This loop may be part of an

active site lid, as we have also observed mutations in residue 74 that led to changes in substrate selectivity (**Table 1**). Interfacial activation, which leads to markedly higher activity toward lipid substrates, was seen in both lipases, including secreted PLA2 (ref. 24), and LCAT²¹. It has also been suggested that, like LCAT, PON1 interacts specifically with apoA-I, and that these interactions modulate its enzymatic

Table 2 The selectivity-determining residues of PONs

Subfamily Position	PON1	PON2	PON3	Newly evolved PON1s Residue	Selectivity
69	L	L	L	V/I	PTE ^a
74 ^b	I	L	M	L/M	Lactonase/esterase
75 ^b	K/M	K/H	P		
76 ^b	S	S	N/A		
78 ^b	N/D	A	A		
190	Y	F/I	L/W/F		
192	K/R	K/M	S/A/V		Human R/Q SNP
193	S	Y/F	F/L	P	PTE ^a
196	M	M/T	M		
222	F	S	S		
240	L	I	V		
291	I	L/V	L	L	Lactonase
292	F	F/Y	L	L/V/S	Esterase
293	F/Y	V/Y/I	N/I	D	Lipase-like ^c
332	T	S	S/T	A	Lactonase
346	V	L/V	I/V	A	PTE ^a

^aPTE, phosphotriesterase. The PTE activity of these mutants is much higher than that of wild-type PON1, which is the best PTE among all PONs (**Table 1**). ^bResidues 74–79 belong to the selectivity-determining residues, which differ between the PON subfamilies (**Fig. 5**) but are conserved within them. These residues are part of a mobile loop (residues 72–79) containing residue 74, which is not seen in the electron density but is presumably part of the active site. ^cLipase-like activity refers to esters of long-chain carboxylic acids (**Table 1**).

Table 3 Data collection and refinement statistics

	Native	SeMet protein
Data collection		
Wavelength (Å)	0.9796	0.9794
Space group	<i>P</i> ₄ ₃ ₂ ₁ ₂	<i>P</i> ₄ ₃ ₂ ₁ ₂
Unit cell dimensions (Å)		
<i>a</i>	98.44	98.49
<i>c</i>	139.17	139.56
Resolution range (Å)	20–2.2	30–2.6
Number of unique reflections	35,312	39,473
Completeness (%) ^a	99.7 (97.9)	97.8 (97.7)
<i>I</i> / $\sigma(I)$ ^a	12.7 (2.7)	13.7 (4.2)
<i>R</i> _{sym} (<i>I</i>) (%) ^a	8.6 (66.1)	10.4 (51.0)
Refinement and model statistics		
Resolution range (Å)		20–2.2
Number of reflections		33,505
<i>R</i> _{work} (%) ^b		18.5
<i>R</i> _{free} (%) ^c		21.7
Average <i>B</i> -factors (Å ²)		34.76
R.m.s. deviation from ideal values		
Bond length (Å)		0.028
Bond angle (°)		2.02
Dihedral angles (°)		28.7
Improper torsion angles (°)		2.06
Estimated coordinate error		
Low-resolution cutoff (Å)		5.0
ESD from Luzzati plot (Å) ^d		0.32
ESD from SIGMAA (Å) ^d		0.34
Ramachandran outliers (%) ^e		3.9

^aValues in parentheses are for the highest-resolution shell. ^b $R_{\text{work}} = \sum |F_o| - |F_c| / \sum |F_o|$ where F_o denotes the observed structure factor amplitude, and the F_c , the structure factor amplitude calculated from the model. ^c R_{free} is as for R_{work} , but calculated with 10% of randomly chosen reflections omitted from the refinement. ^dESD, estimated standard deviation. ^eRamachandran plot outliers are all glycines except for His115, Asn224, Asp269 and His348.

activity³⁷. This hypothesis is supported by the marked proximity of PON1's HDL-anchoring region to its active site. In fact, the proposed HDL-anchoring region is part of an active site lid, and several of the selectivity-determining residues (Table 2) are on amphipathic helix H2 and on the loops linked to H1 (Fig. 6).

The mechanistic and functional resemblances to PLA2 revealed here support the hypothesis that PON1's anti-atherosclerotic activity is mediated by its ability to hydrolyze certain lipids⁶, including oxidized lipids in LDL⁷ and on macrophages¹¹. The structure of PON1 will direct further research aimed at elucidating PON1's mode of binding to HDL and its effect on the enzyme's activities, as well as the precise physiological roles of these activities. The directed evolution results demonstrate the PON family's notable ability to evolve, and introduce the possibility of enhancing the activity of PONs in detoxifying and decontaminating toxic organophosphates and retarding atherosclerosis.

METHODS

Structure determination. The structure was determined for rePON1 variant G2E6 (see Fig. 5 for sequence). rePON1-G2E6 was expressed and purified in fusion with thioredoxin (via a linker containing a His₆-tag)¹³. The thioredoxin tag was spontaneously cleaved after incubation at 25 °C for 10 d, leaving PON1's intact sequence plus five amino acids (DDDKA) from the linker peptide (see Supplementary Methods online). The cleaved PON1 was purified by gel filtration and concentrated to 10 mg ml⁻¹. The SeMet rePON1-G2E6 was expressed in *E. coli* B834(DE3) cells. The protein was purified and crystallized

as native rePON1, but from a different mother liquor. The rePON1 variants were crystallized using the Douglas Instruments robot IMPAX 1–15 microbatch method. Data were collected on beamline ID14–4 at the European Synchrotron Radiation Facility. Three data sets were collected for the SeMet protein crystal and one data set for the native crystal, both at 100 K (Table 3). Details of the crystallization and refinement protocols are provided in Supplementary Methods online. Figures were created with PyMol (<http://pymol.sourceforge.net/>) and MolScript (Fig. 4; <http://www.avatar.se/molscript>). Accessible surface area was calculated using AREAIMOL in the CCP4 package³⁸.

Directed evolution. rePON1 variant G3C9 was used as the starting point for directed evolution. Its sequence is almost identical to those of wild-type rabbit PON1 and rePON1 variant G2E6, whose three-dimensional structure was determined (see Supplementary Note online), and has the same enzymatic parameters. Libraries were generated and cloned as described¹³. PTE activity was screened and subsequently quantified with the fluorogenic organophosphate substrate 7-*O*-diethylphosphoryl-3-cyano-7-hydroxycoumarin as described¹³. Thioloactonase activity was initially screened using 2NA on agar plates. Positive clones were picked from replica plates and grown in 96-well plates as described¹³. The crude cell lysates and purified proteins were assayed for hydrolysis of γ -butyrolactone using 5,5'-dithio-bis-2-nitrobenzoic acid for detection of product by absorbance at 412 nm. Esterase activity was screened using the fluorogenic substrate 7-acetoxycoumarin (for short-chain esterase activity), or with 2-naphthyl octanoate (for long-chain esterase activity) using fast red for the detection of 2-naphthol¹³. Colonies exhibiting the highest activity were grown in 96-well plates, and the crude cell lysate was assayed spectrophotometrically at 365 nm for hydrolysis of 7-acetoxycoumarin and at 320 nm for 2-naphthyl laurate. The activity of the variants described above was determined with the same substrates and assays.

Enzyme kinetics. k_{cat} and K_m values were determined for rePON1-G2E6 with 2NA and paraoxon¹³ at pH 5.6–9.5. Buffers used were MES (pH 5.6–6.5) and bis-trispropane (pH 6.5–9.4) at 0.1 M, plus 1 mM CaCl₂; the ionic strength was adjusted to 0.2 M with NaCl. Kinetic parameters were obtained from three to five independent measurements averaged with standard deviations of 2–23%. k_{cat} / K_m values for each pH value ($(k_{\text{cat}} / K_m)^{\text{H}}$) were fitted to a 'bell-shaped' model using the equation $(k_{\text{cat}} / K_m)^{\text{H}} = (k_{\text{cat}} / K_m)^{\text{max}} / [(10^{-\text{pH}} / 10^{-\text{pKa1}}) + (10^{-\text{pKa2}} / 10^{-\text{pH}}) + 1]$; where $(k_{\text{cat}} / K_m)^{\text{max}}$ is the pH-independent (or plateau value) of k_{cat} / K_m , and pKa^1 and pKa^2 are the apparent pKa values for the acidic and basic groups, respectively.

Coordinates. The coordinates of rePON1 have been deposited in the Protein Data Bank (accession code 1V04).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We acknowledge financial support by the Minerva Foundation and the Israel Science Foundation to D.S.T. and US army Medical Research & Materiel Command to I.S. and J.L.S. The structure was determined in collaboration with the Israel Structural Proteomics Center, supported by the Israel Ministry of Science & Technology, the European Commission Structural Proteomics Project (SPINE) and the Divadol Foundation. J.L.S. is the Morton and Gladys Pickman Professor of Structural Biology.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 23 February; accepted 5 April 2004

Published online at <http://www.nature.com/natstructmolbiol/>

1. Draganov, D.I. & La Du, B.N. Pharmacogenetics of paraoxonases: a brief review. *Naunyn Schmiedeberg's Arch. Pharmacol.* **369**, 78–88 (2004).
2. Lusis, A.J. Atherosclerosis. *Nature* **407**, 233–241 (2000).
3. Shih, D.M. *et al.* Mice lacking serum paraoxonase are susceptible to organophosphate toxicity and atherosclerosis. *Nature* **394**, 284–287 (1998).
4. Mackness, M.I., Arrol, S. & Durrington, P.N. Paraoxonase prevents accumulation of lipoperoxides in low-density-lipoprotein. *FEBS Lett.* **286**, 152–154 (1991).
5. Reddy, S.T. *et al.* Human paraoxonase-3 is an HDL-associated enzyme with biological

- activity similar to paraoxonase-1 protein but is not regulated by oxidized lipids. *Arterioscler. Thromb. Vasc. Biol.* **21**, 542–547 (2001).
6. Rodrigo, L., Mackness, B., Durrington, P.N., Hernandez, A. & Mackness, M.I. Hydrolysis of platelet-activating factor by human serum paraoxonase. *Biochem. J.* **354**, 1–7 (2001).
 7. Ahmed, Z. *et al.* Apolipoprotein A-I promotes the formation of phosphatidylcholine core aldehydes that are hydrolyzed by paraoxonase (PON-1) during high density lipoprotein oxidation with a peroxyxynitrite donor. *J. Biol. Chem.* **276**, 24473–24481 (2001).
 8. Jakubowski, H. Calcium-dependent human serum homocysteine thiolactone hydrolase. A protective mechanism against protein N-homocysteinylation. *J. Biol. Chem.* **275**, 3957–3962 (2000).
 9. Lund-Katz, S., Liu, L.J., Thuahnai, S.T. & Phillips, M.C. High density lipoprotein structure. *Front. Biosci.* **8**, D1044–D1054 (2003).
 10. Fuhrman, B., Volkova, N. & Aviram, M. Oxidative stress increases the expression of the CD36 scavenger receptor and the cellular uptake of oxidized low-density lipoprotein in macrophages from atherosclerotic mice: protective role of antioxidants and of paraoxonase. *Atherosclerosis* **161**, 307–316 (2002).
 11. Rozenberg, O., Shih, D.M. & Aviram, M. Human serum paraoxonase 1 decreases macrophage cholesterol biosynthesis: possible role for its phospholipase-A(2)-like activity and lysophosphatidylcholine formation. *Arterioscler. Thromb. Vasc. Biol.* **23**, 461–467 (2003).
 12. Sorenson, R.C. *et al.* Human serum paraoxonase/arylesterase's retained hydrophobic N-terminal leader sequence associates with HDLs by binding phospholipids: apolipoprotein A-I stabilizes activity. *Arterioscler. Thromb. Vasc. Biol.* **19**, 2214–2225 (1999).
 13. Aharoni, A. *et al.* Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization. *Proc. Natl. Acad. Sci. USA* **101**, 482–487 (2004).
 14. Fokine, A. *et al.* Direct phasing at low resolution of a protein copurified with human paraoxonase (PON1). *Acta Crystallogr. D.* **59**, 2083–2087 (2003).
 15. Josse, D. *et al.* Oligomeric states of the detergent-solubilized human serum paraoxonase (PON1). *J. Biol. Chem.* **277**, 33386–33397 (2002).
 16. Kuo, C.L. & La Du, B.N. Comparison of purified human and rabbit serum paraoxonases. *Drug Metab. Dispos.* **23**, 935–944 (1995).
 17. Jawad, Z. & Paoli, M. Novel sequences propel familiar folds. *Structure* **10**, 447–454 (2002).
 18. Kuo, C.L. & La Du, B.N. Calcium binding by human and rabbit serum paraoxonases. Structural stability and enzymatic activity. *Drug Metab. Dispos.* **26**, 653–60 (1998).
 19. Scharff, E.I., Koepke, J., Fritsch, G., Lucke, C. & Ruterjans, H. Crystal structure of diisopropylfluorophosphatase from *Loligo vulgaris*. *Structure* **9**, 493–502 (2001).
 20. Josse, D. *et al.* Identification of residues essential for human paraoxonase (PON1) arylesterase/organophosphatase activities. *Biochemistry* **38**, 2816–2825 (1999).
 21. Jonas, A. Lecithin cholesterol acyltransferase. *Biochim. Biophys. Acta* **1529**, 245–256 (2000).
 22. Josse, D. *et al.* The active site of human paraoxonase (PON1). *J. Appl. Toxicol.* **21**, S7–S11 (2001).
 23. Josse, D., Xie, W.H., Masson, P., Schopfer, L.M. & Lockridge, O. Tryptophan residue(s) as major components of the human serum paraoxonase active site. *Chem. Biol. Interact.* **120**, 79–84 (1999).
 24. Sekar, K. *et al.* Phospholipase A2 engineering. Structural and functional roles of the highly conserved active site residue aspartate-99. *Biochemistry* **36**, 3104–3114 (1997).
 25. Aviram, M. *et al.* Paraoxonase active site required for protection against LDL oxidation involves its free sulfhydryl group and is different from that required for its arylesterase/paraoxonase activities: selective action of human paraoxonase allozymes Q and R. *Arterioscler. Thromb. Vasc. Biol.* **18**, 1617–1624 (1998).
 26. Kobayashi, M., Shinohara, M., Sakoh, C., Kataoka, M. & Shimizu, S. Lactone-ring-cleaving enzyme: genetic analysis, novel RNA editing, and evolutionary implications. *Proc. Natl. Acad. Sci. USA* **95**, 12787–12792 (1998).
 27. Chelur, D.S. *et al.* The mechanosensory protein MEC-6 is a subunit of the *C. elegans* touch-cell degenerin channel. *Nature* **420**, 669–73 (2002).
 28. Navab, M. *et al.* High density associated enzymes: their role in vascular biology. *Curr. Opin. Lipidol.* **9**, 449–456 (1998).
 29. Borhani, D.W., Rogers, D.P., Engler, J.A. & Brouillette, C.G. Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc. Natl. Acad. Sci. USA* **94**, 12291–12296 (1997).
 30. Segrest, J.P., Harvey, S.C. & Zannis, V. Detailed molecular model of apolipoprotein A-I on the surface of high-density lipoproteins and its functional implications. *Trends Cardiovasc. Med.* **10**, 246–252 (2000).
 31. Killian, J.A. & von Heijne, G. How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429–434 (2000).
 32. Waldo, G.S. Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.* **7**, 33–38 (2003).
 33. Bencharit, S., Morton, C.L., Xue, Y., Potter, P.M. & Redinbo, M.R. Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. *Nat. Struct. Biol.* **10**, 349–356 (2003).
 34. Millard, C.B., Lockridge, O. & Broomfield, C.A. Organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase: synergy results in a somanase. *Biochemistry* **37**, 237–247 (1998).
 35. Greenblatt, H.M., Dvir, H., Silman, I. & Sussman, J.L. Acetylcholinesterase: a multifaceted target for structure-based drug design of anticholinesterase agents for the treatment of Alzheimer's disease. *J. Mol. Neurosci.* **20**, 369–383 (2003).
 36. Leviev, I., Deakin, S. & James, R.W. Decreased stability of the M54 isoform of paraoxonase as a contributory factor to variations in human serum paraoxonase concentrations. *J. Lipid Res.* **42**, 528–535 (2001).
 37. Oda, M.N., Bielicki, J.K., Berger, T. & Forte, T.M. Cysteine substitutions in apolipoprotein A-I primary structure modulate paraoxonase activity. *Biochemistry* **40**, 1710–1718 (2001).
 38. Lee, B. & Richards, F.M. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–389 (1971).