Enzyme Evolution: An Epistatic Ratchet versus a Smooth Reversible Transition

Moshe Ben-David,[†,1] Misha Soskine,[‡,1] Artem Dubovetskyi,[1] Kesava-Phaneendra Cherukuri,[1] Orly Dym,[2] Joel L. Sussman,[3] Qinghua Liao,[4] Klaudia Szeler,[4] Shina Caroline Lynn Kamerlin,[*,4] and Dan S. Tawfik[*,1]

[1]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

[2]Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel

[3]Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel

[4]Department of Chemistry – BMC, Uppsala University, Uppsala, Sweden

[†]Present address: Ukko Ltd. HaMada 12, Rehovot, Israel

[‡]Present address: DNA Script, 29 Rue du Faubourg Saint-Jacques, Paris, France

***Corresponding authors:** E-mails: lynn.kamerlin@kemi.uu.se; dan.tawfik@weizmann.ac.il.

Associate editor: Miriam Barlow

Coordinates and structure factors of the 3m-OPH, 5m-OPH, and 5m-OPH-115H mutants have been deposited at the Protein Data Bank (accession numbers G682, 6H0A, and G6MU, respectively).

Abstract

Evolutionary trajectories are deemed largely irreversible. In a newly diverged protein, reversion of mutations that led to the functional switch typically results in loss of both the new and the ancestral functions. Nonetheless, evolutionary transitions where reversions are viable have also been described. The structural and mechanistic causes of reversion compatibility versus incompatibility therefore remain unclear. We examined two laboratory evolution trajectories of mammalian paraoxonase-1, a lactonase with promiscuous organophosphate hydrolase (OPH) activity. Both trajectories began with the same active-site mutant, His115Trp, which lost the native lactonase activity and acquired higher OPH activity. A neo-functionalization trajectory amplified the promiscuous OPH activity, whereas the re-functionalization trajectory restored the native activity, thus generating a new lactonase that lacks His115. The His115 revertants of these trajectories indicated opposite trends. Revertants of the neo-functionalization trajectory lost both the evolved OPH and the original lactonase activity. Revertants of the trajectory that restored the original lactonase function were, however, fully active. Crystal structures and molecular simulations show that in the newly diverged OPH, the reverted His115 and other catalytic residues are displaced, thus causing loss of both the original and the new activity. In contrast, in the re-functionalization trajectory, reversion compatibility of the original lactonase activity derives from mechanistic versatility whereby multiple residues can fulfill the same task. This versatility enables unique sequence-reversible compositions that are inaccessible when the active site was repurposed toward a new function.

Key words: **protein evolution, epistasis, evolutionary incompatibility, promiscuity, neo-functionalization.**

Introduction

The extent to which evolutionary processes are reversible is a fundamental aspect of organismal and molecular evolution (Bridgham et al. 2009). At the molecular level, it has been shown that although the phenotype may revert (e.g., a newly diverged receptor or enzyme may reacquire its ancestral activity), the genotype does not (Bridgham et al. 2009; Kaltenbach et al. 2015). Across long evolutionary times, irreversibility is manifested in an "epistatic ratchet" (Bridgham et al. 2009) or reversion incompatibility: When reverted in contemporary sequences, the mutation(s) that originally led the functional switch in the ancestor result in loss of function, typically of both the extant and the ancestral functions (Soylemez and Kondrashov 2012; Harms and Thornton 2013; Gong and Bloom 2014; Kaltenbach et al. 2015; Shah et al. 2015; Pareek et al. 2016; Starr and Thornton 2016; Cheema et al. 2017). Accordingly, reversion mutations have been shown to cause structural perturbances in the active site that in turn lead to loss of activity (Kaltenbach et al. 2015).

On the other hand, evolutionary transition states, that is, protein sequences that are compatible with both reverse and forward divergence have also been described (Wellner et al. 2013). In general, adaptive trajectories (e.g., acquisition of new enzyme functions) seem to be highly epistatic, and thereby more prone to irreversibility than neutral drifts (changes in sequence while maintaining the same function) (Bridgham et al. 2009; Afriat-Jurnou et al. 2012; Natarajan et al. 2013; Gong and Bloom 2014; Kaltenbach et al. 2015; Gupta and Adami 2016; Wu et al. 2018). However, non-epistatic neo-functionalization trajectories have also been described

Article

(Kaltenbach et al. 2018) and factors other than the acquisition of new biochemical function, such as expression levels, may also lead to high epistasis (Dasmeh et al. 2017). Overall, the mechanistic basis of reversion incompatibility versus compatibility remains largely unknown.

Here, we examined reversibility in two different laboratory evolution trajectories performed on the same enzyme: mammalian paraoxonase-1 (PON1), a calcium-dependent lactonase with promiscuous organophosphate hydrolase (OPH) activity (Draganov et al. 2005; Khersonsky and Tawfik 2005). Both trajectories began with the same mutation in an active-site residue, H115W. This mutation enhanced the promiscuous OPH activity while causing a >5,000-fold drop in the native lactonase activity (Ben-David et al. 2013). One evolutionary trajectory converted PON1 into a catalytically efficient and stereo-specific OPH for G-type nerve agent substrates (Gupta et al. 2011). The most active variant from this trajectory (Gupta et al. 2011) contained the H115W mutation *plus* four other active-site mutations and is dubbed hereafter, PON1-*5m-OPH*, or simply *5m-OPH*. In a separate experiment described here, we evolved the H115W mutant to regain the original lactonase activity. Following nine rounds of random mutagenesis and selection, we obtained an enzyme that lacked His115 yet largely regained the lactonase activity of wild-type PON1 (dubbed PON1-neo-lactonase, or neo-lactonase). The comparison of two different trajectories each toward a different substrate and reaction could be problematic (comparing apples to oranges), yet here, we compare two trajectories that begin not only with the same enzyme but also are initiated by the very same mutation.

We have subsequently explored the effect of reversion of position 115 to the wild-type state, that is, into histidine, along both trajectories. Reversion in the neo-functionalization trajectory that converted PON1 into an OPH resulted in loss of both the newly evolved OPH and the original lactonase activities. However, in the re-functionalization trajectory where the original lactonase activity was reacquired, the 115H revertants turned out to be fully active lactonases. We describe crystal structures of the evolved 5m-OPH variants and of it 115H revertant, and molecular simulations of the re-functionalized lactonases and their revertants. These shed light on the structural and mechanistic causes of reversion compatibility in the case of neo-functionalization versus incompatibility in the case of re-functionalization.

Results

The Directed Evolution of an OPH

The leading PON1 mutant from the OPH trajectory, 5m-OPH, contained H115W and four additional active-site mutations: L69G, H134R, F222S, and T332S. Selected for nerve agent hydrolysis, this mutant, originally dubbed 2D8, exhibited $>10^4$-fold increase with a cyclosarin surrogate used as the target (Gupta et al. 2011). Here, we monitored the OPH activity with the standard chromogenic OP substrate, paraoxon. The paraoxonase activity of 5m-OPH increased by ~44-fold compared with H115W, and ~100-fold compared with the wild-type enzyme, whereas its lactonase activity dropped down by $>10^3$-fold, indicating a new specialization as an OPH (table 1).

Crystal Structures of the Newly Evolved OPH Variants

To unravel the structural basis of PON1's shift from a lactonase into an OPH, we solved the crystal structure of 5m-OPH (at 2.1 Å resolution) and also of a triple mutant, L69S/H115W/F222S (PON1-3m-OPH), that represents the key intermediate of this neo-functionalization trajectory (2.4 Å; supplementary table 1, Supplementary Material online). We compared these structures to the available structure of wild-type PON1 (PDB ID: 3SRE; hereafter WT) and of PON1-H115W (PDB ID: 4HH0). The crystal structures were obtained with a crystallizable recombinant PON1 variant dubbed G2E6, whereas 5m-OPH was originally generated using an orthologous recombinant variant dubbed G3C9 (94.4% sequence identity to G2E6) (Aharoni et al. 2004). The activity levels and effects of mutations at these two backgrounds are

Table 1. Enzymatic Parameters of the Organophosphate Hydrolase PON1-5m-OPH Variants and Their Revertants.

Evolved Variants (115W)				Revertants (115H)			
Variant	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	**Variant**	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
Phosphotriesterase activity (paraoxon)							
WT-H115W	1.6 ± 0.1	7.3 ± 4	4,700 ± 246	WT-like PON1[a]	2.3 ± 0.3	5.0 ± 0.6	2,150 ± 301
3m-OPH	1.2 ± 0.08	41.3 ± 1.5	33,961 ± 2,272	3m-OPH	5.1 ± 0.07	0.26 ± 0.04	50 ± 9
5m-OPH	0.2 ± 0.01	43 ± 0.5	(2.09 ± 0.09) × 10^5	5m-OPH	1.0 ± 00.3	0.08 ± 0.01	78 ± 4
Lactonase activity (TBBL)							
WT-H115W	0.5 ± 0.014	0.13 ± 0.01	290 ± 4	WT-like PON1[a]	1.1 ± 0.1	189 ± 16	(1.69 ± 0.23) × 10^5
3m-OPH	n.a.	n.a.	n.a.	3m-OPH	n.a.	n.a.	n.a.
5m-OPH	>5[b]	n.d.	216 ± 11	5m-OPH	>5[b]	n.d.	29 ± 4

NOTE.—The wild-type-like starting point is denoted in the "Revertant" column, as it carries His at position 115, whereas the neo-functionalized 5m-OPH is in the "Evolved variants" column. Error ranges are derived from comparing >3 independent measurements. n.a., data not available; n.d., not determined. The enzymatic activity of 3m-OPH with TBBL is expected to be in the range of the WT-H115W and 5m-OPH mutants (i.e., significantly lower relative to the wild-type enzyme), due to the presence of the H115W mutation (Ben-David et al. 2013).

[a]3m-OPH and 5m-OPH were originally evolved at the background of recombinant PON1 variant G3C9 (variants 8C8 and 2D8, respectively [Gupta et al. 2011]). 3m-OPH contains only the active-site mutations of 8C8: H115W, L69S, and F222S (the surface mutations V97A and P135A were excluded due to their large distance from the active site [~18 Å]). Identical to 2D8, 5m-OPH contains five active-site mutations: H115W, L69G, H134R, F222S, and T332S. These were grafted at the background of an orthologous recombinant variant PON1-G2E6 (94.4% sequence identity to G3C9), which is amenable to crystallization and exhibits very similar kinetic parameters (Aharoni et al. 2004). WT-like PON1 therefore relates to PON1-G2E6, and all variants described in this table are based on G2E6.

[b]K_M and k_{cat} values could not be determined since saturation was not achieved during experiment (K_M is above substrate solubility).

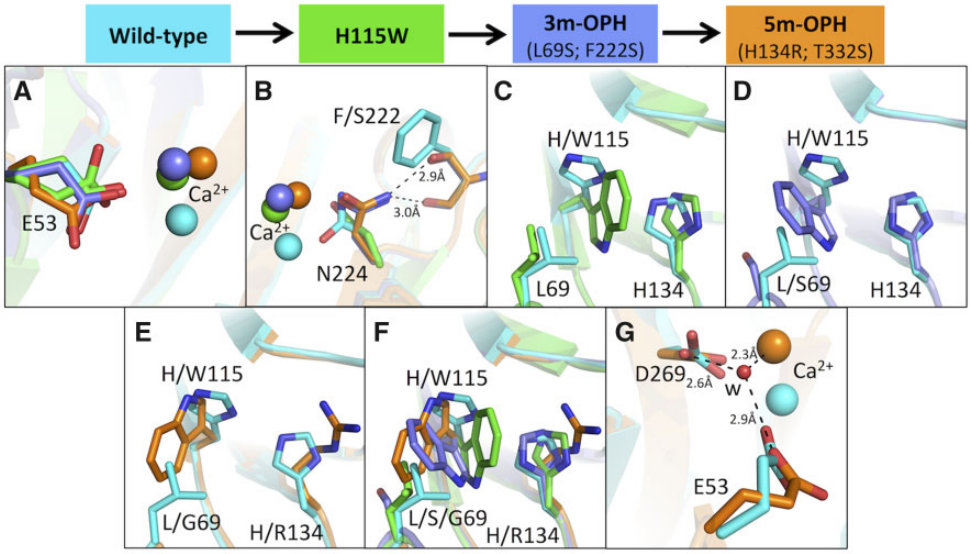

FIG. 1. PON1's neo-functionalization as an OPH. This figure illustrates the gradual change in PON1's active site from a lactonase (wild-type) to an organophosphate hydrolase (5m-OPH, the trajectory's final variant). Shown here are superimpositions of wild-type PON1 (PDB ID: 3SRE; *cyan*), the H115W single mutant (PDB ID: 4HHO; *green*), the triple-mutant intermediate 3m-OPH (*blue*), and the final penta-mutant 5m-OPH (*orange*). (*A*) Close-up of the catalytic calcium binding-site, highlighting the displacement of the calcium ion and of the side chains of the ligating E53. (*B*) The same, with a focus on the metal ligating side chain of N224 and the nearby residue 222, that changed from Phe in wild-type PON1 to Ser in 5m-OPH (dashed lines represent potential H-bonds; $d \approx 2.9$ Å). (*C*–*F*) Close-ups of active-site positions 69, 115, and 134, showing (*C*) H115W, (*D*) 3m-OPH, (*E*) 5m-OPH, and (*F*) the three mutants superimposed. (*G*) Upon the displacement of the catalytic calcium, a bridging water ("w"; *red*) had appeared in 5m-OPH that mediates the interaction of E53 with the catalytic calcium, as well as likely to mimic the attacking activated water.

highly similar (in the range of few-fold variations; see also table 1, footnote a). Nonetheless, for the sake of consistency, we grafted the five active-site mutations of 5m-OPH onto G2E6 and reanalyzed the enzymatic activities at this background (table 1).

The crystal structures of PON1-3m-OPH and PON1-5m-OPH indicated a significant reshaping of the active site. As seen in the H115W mutant that initiated the OPH trajectory (Ben-David et al. 2013), the catalytic calcium ion was displaced "upward" (away from the structural calcium which is buried in PON1's central tunnel) by 1.8 Å in H115W, and 2.4 Å in the further evolved 3m- and 5m-OPH (fig. 1*A* and supplementary fig. 1*A*–*D*, Supplementary Material online). Furthermore, in both these OPH mutants, the side chains of the calcium-ligating residues N224 and E53 adopted an orientation that differs from those observed in the WT and the H115W structures (fig. 1*A* and *B*). The new orientation of N224 favors the formation of H-bonds with both the side-chain and backbone oxygens of S222—one of the key OPH trajectory mutations (F222 in WT and H115W; fig. 1*B*). The rotamer of W115's side chain also differs compared with H115 in wild-type PON1 (fig. 1*C*–*F*). The side chain of E53 lost its direct contact with the catalytic calcium, and a bridging water had appeared that likely maintains this interaction (fig. 1*G*). Finally, PON1's flexible active-site loop (residues 70–81) is structured in 5m-OPH, with the backbone assuming almost the same orientation as in the WT when in complex with a lactone analog (PDB: 3SRG; supplementary fig. 1*E*, Supplementary Material online). However, the side chain of Y71 that is visible in the WT, and tightly H-bonded to D183, exhibited no density in 5m-OPH. The position of Y71 and its flexibility play a role in modulating solvent accessibility to PON1's active site, and in particular in relation to its OPH activity (Blaha-Nelson et al. 2017).

Overall, it appears that PON1's transition from a lactonase into an OPH involved a gradual shift in the position of the catalytic calcium and its ligating residues. The change in 115 from His to Trp triggered these transitions, but the additional mutations, foremost L69G (Ser in the intermediate PON1-3m-OPH mutant), and F222S, both play a role in shifting 115W's side chain and in inducing further changes in E53 and N224. The mechanistic implications of these structural changes are discussed below in conjunction with 5m-OPH's irreversibility.

Irreversibility of the OPH Trajectory

The $>10^3$-fold drop in 5m-OPH's lactonase activity was driven by the H115W mutation, whereas the four additional mutations in 5m-OPH induced only 10-fold further reduction in activity (table 1). However, reversion of 115 to His not only failed to restore 5m-OPH's lactonase activity but even decreased it by another 7-fold. The OPH activity also dropped 27-fold below the WT's level (table 1). In essence, 5m-OPH's revertant at position 115 (5m-OPH-115H) is a dead enzyme, both as a lactonase and as an OPH.

To unravel the mechanistic basis of this irreversibility, we have solved the crystal structure of the revertant, 5m-OPH-115H (at 2.7 Å; supplementary table 1, Supplementary Material online). Most conspicuously, in the revertant, H115 adopted a completely different rotamer compared with the WT (fig. 2*A* and *B* and supplementary fig. 1F, Supplementary Material online, in magenta; henceforth the

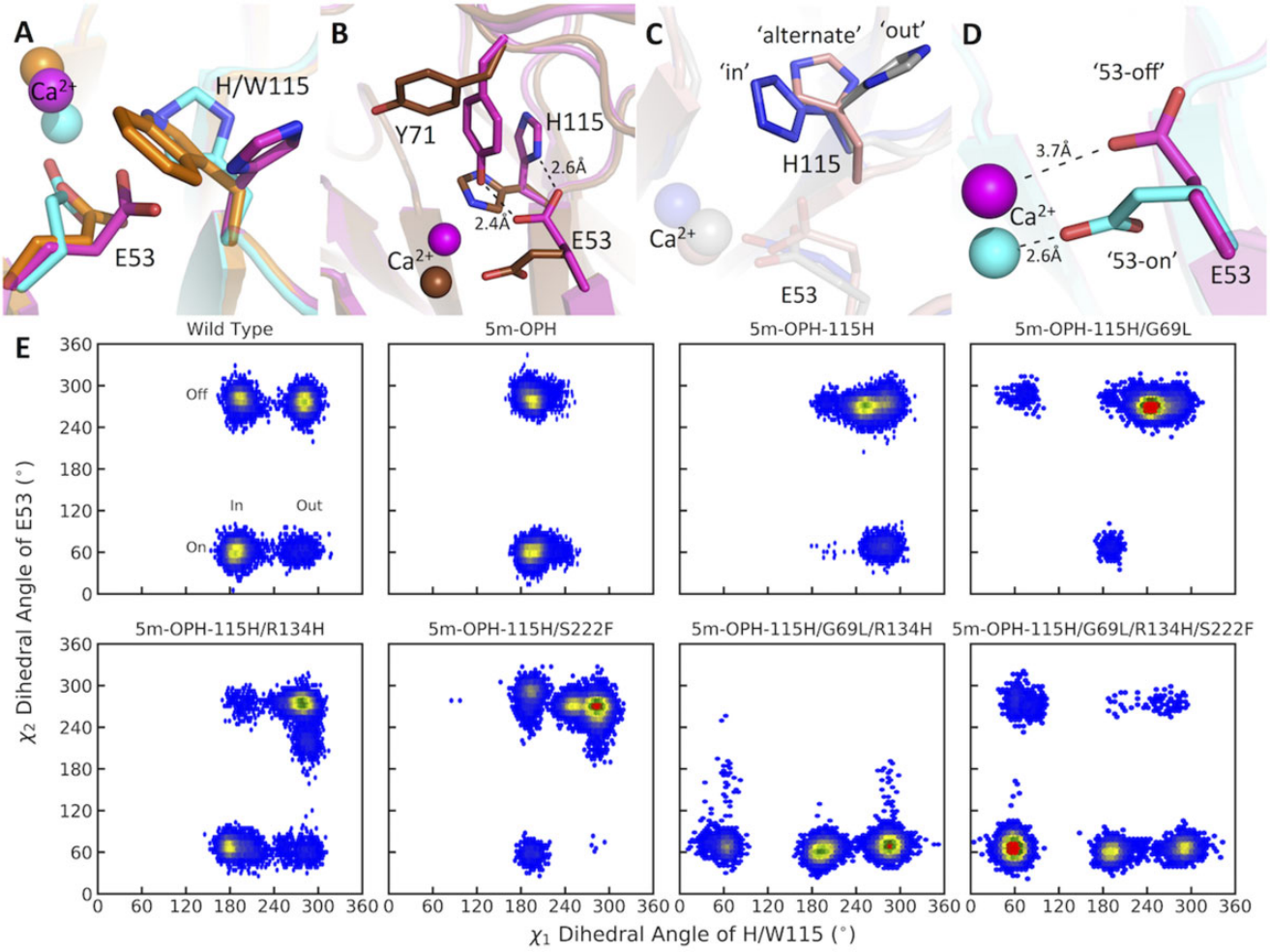

FIG. 2. Structural changes in response to the W115H reversion in PON1-5m-OPH. (*A*) Overlay of wild-type PON1 (*cyan*), 5m-OPH (*orange*), and its revertant, 5m-OPH-115H (*magenta*). (*B*) Overlay of 5m-OPH-115H (*magenta*) with wild-type PON1 in complex with a lactone analog (PDB: 3SRG; *brown*; ligand not shown). (*C*–*E*) Conformational space analysis of H115 and E53's side chains using HREX-MD (see Materials and Methods). (*C*) H115's major conformations: the "in" conformation observed in the wild-type crystal structure (sampling a χ_1-range of ~240–300°; *blue*); the "out" conformation observed in the penta-W115H revertant (sampling a χ_1-range of ~160–240°; *gray*); an alternate sparsely populated intermediary conformation that appears primarily in the 5m-OPH-W115H/G69L/R134H simulations (χ_1 values in the 30–90° range; *salmon*). (*D*) E53's major conformations: the "on" (as in wild-type; *cyan*) and "off" (as in 5m-OPH-115H; *magenta*) conformations and the parallel displacement of the catalytic metal ion. (*E*) Conformational space analysis of the χ_2-dihedral angle of the E53 side chain and the χ_1-dihedral angle of the H115 (W115 for 5m-OPH) side chain for: wild-type PON1 (PDB: 3SRG), the 5m-OPH mutant (the crystal structure described here), the reverted 5m-OPH-115H (the crystal structure described here), and five manually generated revertants: 5m-OPH-115H/G69L, 5m-OPH-115H/G69L/R134H, 5m-OPH-115H/S222F, 5m-OPH-115H/G69L/R134H, and 5m-OPH-115H/G69L/R134H/S222F.

"115-out" conformation, see fig. 2*C*). However, although the displacement of H115 may explain the failure to regain the lactonase activity, it does not explain the revertant's loss of OPH activity (given that, in wild-type PON1, the OPH activity was enhanced by mutating His115). The latter seems to be the outcome of other disruptions, most evidently the displacement of E53's side chain (fig. 2*A*–*C* and supplementary fig. 1*F*, Supplementary Material online). In the wild-type, E53's closer carboxylate oxygen is at a distance of 2.6 Å from the catalytic calcium ion, thus establishing a direct contact with the metal ion (henceforth, the on-calcium E53 rotamer, or "53-on" for brevity, fig. 2*D*). However, in H115W, 3m- and 5m-OPH, E53 is shifted away from the catalytic calcium (fig. 2*A* and *B*; henceforth the off-calcium, or "53-off" rotamer), although the contact with the calcium is maintained by a bridging water molecule (fig. 1*G*). However, in 5m-OPH, upon reversion of W115 to His, the distance between E53's carboxylate oxygen and the calcium is too large to establish a direct contact with the metal ion (3.7 Å) and the bridging molecule is also absent (fig. 2*D*). This "off" configuration of E53 seems to be stabilized by a close interaction with both the imidazole ring of H115 in its "out" conformation (E53-H115 2.6 Å) and Y71 (E53-Y71 2.4 Å; fig. 2*B*). Indeed, although the flexible active-site loop becomes structured in 5m-OPH, the side chain of one of its key residues, Y71, is disordered (i.e., lacks defined electron density). However, in the revertant structure (5m-OPH-115H), the side chain of Y71 is ordered and assumes a different rotamer orientation compared with the wild-type in complex with a lactone analog. Notably, this Y71 rotamer is within interacting distance from E53, thus

stabilizing the latter's off-calcium rotamer (fig. 2B). This observation is in-line with our previous studies where we have shown that Y71 can assume various conformations that might be linked to different functional states (Ben-David et al. 2012; Blaha-Nelson et al. 2017).

Overall, the crystal structures indicate that the wild-type-like conformation of these two residues, E53 and H115, is not restored in 5m-OPH, and this failure is the likely reason for the irreversibility of the 5m-OPH trajectory. This conclusion is supported by the molecular dynamics simulations detailed below.

The Mechanistic Basis of Irreversibility

The relocation of E53 represents a switch in the mechanism in catalysis of 5m-OPH. Owing to a high level of networking and versatility of its catalytic machinery, PON1 catalyzes both C–O and P–O hydrolysis: Various reaction steps can be handled by different residues, by sharing, or by completely shifting from one residue to another (Ben-David et al. 2012). Specifically, the task of activating the water for nucleophilic attack can shift from H115 in wild-type acting as lactonase, to E53 that seems to be the activating base in the wild-type when promiscuously acting as an OPH (and also to D269 as discussed in the next sections). This versatility also underlies PON1's evolvability: in H115W (Ben-David et al. 2012), and presumably in its more advanced OPH variants, E53 is repurposed to act as the activating base. Indeed, once specialized as OPH, a bridging water molecule appears that most likely mimics the attacking activated water (i.e., in 5m-OPH; fig. 1G). Upon reversion of 115 (5m-OPH-115H), a lose-lose situation arises: The His side chain fails to assume its wild-type position and cannot act as the activating base, hence the failure to regain the lactonase activity, and on the other hand, E53's dislocation and the loss of the bridging water diminishes its ability to act as the activating base in the OPH activity.

To better understand the conformational changes of E53 and H115, we performed Hamiltonian replica exchange-molecular dynamics (HREX-MD) simulations that systematically probed the rotamer space available for the side chains of residues 53 and 115 (fig. 2E; we have also attempted to validate the hypothesized mechanistic switches by empirical valence bond (EVB) simulations but failed to obtain reliable models; detailed in Materials and Methods). Interestingly, unlike the crystal structure, the simulated wild-type does not exhibit a single, discrete positioning of either E53 or H115. Alongside the dominant, active rotamer "115-in," the non-productive rotamer "115-out" was observed in ~40% of the simulations (fig. 2E; supplementary fig. 2A, Supplementary Material online). For E53, the "on" and "off" rotamers were also similarly represented, with a slight preference toward the "off" rotamer (fig. 2E; supplementary fig. 2B, Supplementary Material online). We note that in our simulations, "on" and "off" refer to both E53 and the calcium ion, which are either in a wild-type position ("on"), or with the E53 side chain swung out and the calcium displaced in position similar to that observed in the H115W mutant and the further evolved 3m- and 5m-OPH (fig. 1A) or in even more structurally disturbed positions (see supplementary figs. 2 and 3, Supplementary Material online). This displacement was proposed to lead to the divergence of PON1 from a lactonase to an OPH upon mutating H115 to Trp (Ben-David et al. 2013). The HREX-MD suggests that the potential for displacement of the catalytic calcium is already present in the wild-type enzyme.

Upon reversion of 115, in agreement with the complete loss of lactonase activity in 5m-OPH-115H, and as observed in the crystal structure (fig. 2A), the 115-"out" rotamer exclusively dominates and E53 is predominantly in its "off" rotamer (fig. 2A and E). Further, our simulations also indicate that the rotameric states of E53 and H115 are correlated in simulations of both the crystal structure of the wild-type and computationally generated structures of the reverted mutants; however, they are anticorrelated in simulations of 5m-OPH-115H ($P < 10^{-10}$ via chi-square tests; odds ratios also significant with $P < 10^{-10}$ via Fisher's exact test).

Reversibility in 5m-OPH Is Prevented by an Epistatic Ratchet

Which of the four other mutations in 5m-OPH is blocking reversion? Three mutations, L69G, H134R, and F222S, may have a major impact on 115's side-chain conformation, and thereby on the catalytic calcium and its ligating residues. The fourth mutation, T332S, appears in nearly all laboratory evolved PON1s irrespective of which activity was evolved and is largely neutral at the background of wild-type PON1 (Amitai et al. 2007). In contrast, positions 69 and 134 that are in direct contact with 115 are likely to be the major hurdle to reversion (fig. 1E and F). We therefore examined double-mutant revertants that included 115H and an additional reversion in either position 69 or position 134. These double mutants expressed at much lower levels compared with 5m-OPH-115H, and could not be purified, suggesting a strong destabilizing effect of the second mutation. Furthermore, in crude lysates, lactonase activity above background could not be detected (supplementary table 2, Supplementary Material online). Accordingly, in HREX-MD simulations of the 69 and 134 double revertants, the G69L reversion locked 115H in the "out" rotamer and E53 in the "off" rotamer (fig. 2E). The 115H-134H reversion showed some sampling of the "115-in" rotamer; however, the "115-out" and the "53-off" rotamer state dominated.

Even more strikingly, reversion of 115, 69, and 134 failed to reinstate activity. This triple revertant failed to purify and showed no detectable lactonase activity (supplementary table 2, upper panel, Supplementary Material online). Loss of activity was also observed with respect to the OPH activity. The evolved 5m-OPH is a priori ~50-fold more active compared with wild-type, as measured in crude cell lysates (supplementary table 2, Supplementary Material online). Reversion of 69, 134, and particularly of both, decreased the paraoxonase activity of 5m-OPH by >10-fold to a level that is only ~4-fold above wild-type (supplementary table 2, lower-left panel, Supplementary Material online). Further, at the background of the 115 reversion (5m-OPH-115H), reversion of either 69 or 134 completely inactivates the OPH activity.

Indeed, the HREX-MD simulations of this triple mutant indicated that only the "E53-on" rotamer was present, whereas for H115, alongside the wild-type in/out rotamers, a third "intermediate" rotamer appeared that is likely also noncatalytic. Finally, the HREX-MD simulations also indicated that reversion of 222 which was not tested experimentally (i.e., 5m-OPH-115H-S222F) results in a H115 side-chain rotamer distribution that is intermediary to reversion of 115 on its own and reversion of 115 combined with the R134H reversion (fig. 2*E*). Overall, it appears that PON1's catalytic cycle as an OPH (i.e., not necessarily the rate-determining chemical step, but other chemical steps and/or product release), demands a shuttling of E53 between its "on" and "off" rotamers and a concomitant shift in the calcium. However, in the revertants, one of the two rotamers tends to be favored. The failure to regain the lactonase activity is in agreement with the fact that neither of the rotamer spaces of the 53 or 115 revertants matches the rotamer space of the wild-type enzyme.

We note that simulating the simultaneous reversion of four key residues, that is, 5m-OPH-115H/G69L/R134H/F222S, which is equivalent to wild-type with 332T, and is fully functional (Amitai et al. 2007), leads to behavior that is intermediary to that of the wild-type enzyme and that of the 5m-OPH-115H revertant. That is, although our simulations still sample multiple rotameric states of both H115 and E53, H115 is now predominantly found in the "alternate" rotamer shown in fig. 2C, with E53 predominantly in its "on" rotamer (fig. 2*E* and supplementary figs. 2 and 3, Supplementary Material online). However, of all reverted mutants simulated here, 5m-OPH-115H/G69L/R134H/F222S is the only one in which the 115-"out" rotamer is no longer the dominant rotamer (note that, also here, the E53 and H115 rotamers are correlated). Overall, it appears that reverting positions 69, 134, and 222, simultaneously, is a prerequisite for enabling a transition away from H115-"out" toward ultimately H115-"in." Finally, we also performed HREX-MD simulations of 5m-OPH (fig. 2*E* and supplementary fig. 2, Supplementary Material online), where it can be seen that although W115 remains in a position analogous to the H115 "in" position (fig. 2), two distinct conformations of E53 are sampled in the simulations. Thus, the rotamer distribution of the revertant 5m-OPH-H115 differs not only from the wild-type enzyme (as expected, wild-type PON1 is primarily a lactonase) but also from its active counterpart 5m-OPH.

To conclude, it appears that reversion in 5m-OPH to the original lactonase activity is blocked by an "epistatic ratchet" (Bridgham et al. 2009) that involves multiple positions. In effect, 5m-OPH can only revert via simultaneous reversion of four out of its five mutations (i.e., W115H, R134H, G69L, and S222F). Thus, as far as we tested, a gradual, smooth reversion of 5m-OPH to give an even reasonably active lactonase is effectively impossible.

Directed Evolution of a Neo-lactonase

His115 makes an important contribution to PON1's lactonase activity. This residue is conserved throughout the serum paraoxonase phylogeny, including in PON1's paralogs PON2 and PON3, and even in distantly related bacterial PONs (Bar-Rogovsky et al. 2013). Accordingly, the H115W mutation decreases PON1's lactonase activity by ~5,000-fold (table 2). However, our previous study indicated that H115's role in activating the hydrolytic water is shared with E53 (Ben-David et al. 2012). Could PON1 evolve into a fully active lactonase in the absence of H115? To address this question, we applied random mutagenesis and screened for variants with higher lactonase activity. In a nutshell, PON1-H115W was subject to random mutagenesis incorporating ~2 mutations per gene. Approximately, 1,000 clones were screened using a chromogenic lactone substrate, 5-thiobutyl butyrolactone (TBBL), and the most active clones were isolated. Typically, ten variants displaying the highest lactonase activity were taken to the next round (≥1.2-fold lactonase activity relative to best variant from the previous round). The coding genes of these variants were used as a template for the next round of mutagenesis and screening. Nine such rounds of

Table 2. Enzymatic Parameters of the Neolactonase PON1 Variants and Their Revertants.

Evolved Variants (115W/G)				Revertants (115H)			
Variant	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)	**Variant**	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
Lactonase activity (TBBL)							
WT-H115W	1.40 ± 0.17	0.25 ± 0.02	178 ± 15	**WT-like PON1**[a]	0.31 ± 0.01	322 ± 3	$(1.04 \pm 0.05) \times 10^6$
9H8[b]	0.97 ± 0.13	8.7 ± 1.5	$(9.0 \pm 0.6) \times 10^3$	**9H8-115H**	0.20 ± 0.12	25 ± 6	$(1.3 \pm 0.8) \times 10^5$
9H10[c]	0.86 ± 0.03	23.4 ± 2.4	$(2.8 \pm 0.1) \times 10^4$	**9H10-115H**	0.03 ± 0.01	86 ± 6	$(3.1 \pm 0.1) \times 10^6$
Phosphotriesterase activity (paraoxon)							
WT-H115W	0.7 ± 0.2	21 ± 4	$(2.9 \pm 1.5) \times 10^4$	**WT-like PON1**[a]	0.8 ± 0.2	5.3 ± 0.2	7,320 ± 1116
9H8	ND[d]	ND	4 ± 2	**9H8-115H**	ND[d]	ND	8 ± 3
9H10	ND[d]	ND	32 ± 6	**9H10-115H**	ND[d]	ND	28 ± 4

NOTE.—The wild-type-like starting point is denoted in the "Revertant" column, as it carries His at position 115. Error ranges are derived from comparing >3 independent measurements.

[a]WT-like PON1 relates to PON1-G3C9, and all variants described in this table are based on G3C9. The kinetic parameters and the effect of mutations in G3C9 may vary by several folds compared with PON1 variant G2E6 used for crystallization (whose parameters are listed in table 1). Specifically, the H115W mutation increases G3C9's paraoxonase activity (k_{cat}/K_M) by ~4-fold and ~2-fold in G2E6; the lactonase activity is reduced in G2E6 by ~2×10^4-fold and by ~6×10^3-fold at the background of variant G3C9. However, these differences also relate in part to variability in protein preparations and assays.

[b]The 9H8 mutant contains the following mutations: **H115G**, K46E, K84E, N105S, T106A, H134R, K192R, M196T, F264L, F292S, and L355Q.

[c]The 9H10 mutant contains the following mutations: **H115G**, I74M, E93G, N105S, T106A, K162R, M196T, F264L, F292S, and K297M.

[d]K_M and k_{cat} values could not be determined since saturation was not achieved during experiment (K_M is above substrate solubility).

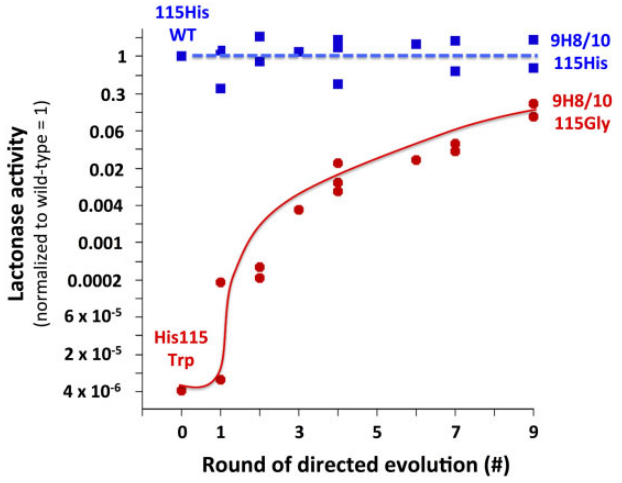

FIG. 3. PON1's re-functionalization trajectory. The trajectory variants are shown as red circles, starting from the H115W mutant used as the starting point (generation 0) to the ninth generation variants 9H8 and 9H10. The corresponding 115H revertants are shown as blue squares (round 0 therefore corresponds to wild-type). Plotted is the activity observed in crude cell lysates normalized to the activity of the recombinant WT-like PON1 variant G3C9 (sequences and raw data are listed in supplementary table 3, Supplementary Material online). The lines were drawn for visible clarity.

mutagenesis and screening were performed until the lactonase activity approached the level of wild-type PON1 (fig. 3).

Overall, up to 11 mutations were fixed in the evolved neolactonase variants (supplementary table 3, Supplementary Material online). Most mutations were remote and only a few occurred within the active site—foremost, a change of 115W to Cys, and ultimately to Gly occurred. The two most active variants isolated from the ninth round where 9H8 and 9H10. These were purified and assayed, indicating that the lactonase k_{cat}/K_M value increased by up to 160-fold compared with the starting point (the H115W mutant). The best variant, 9H10, exhibited a k_{cat}/K_M value of $2.8 \times 10^4\,s^{-1}\,M^{-1}$ (~37-fold lower than wild-type; table 2).

The Neo-lactonase Trajectory Is Entirely Reversible

In contrast to the OPH trajectory, reversion of 115 to His restored, or even increased activity, in all the intermediate variants as well as in the two most active variants, 9H8 and 9H10 (table 2 and supplementary table 3, Supplementary Material online). The H115 revertant of the latter exhibited a k_{cat}/K_M value for the lactonase activity that is even slightly higher than wild-type (table 2 and fig. 3).

The OPH activity that initially increased owing to the 115W mutation gradually declined along the gain of the lactonase activity and became nearly up to 7,000-fold lower in the ninth round variants 9H8 and 9H10. It remained similarly low upon reversion to 115H (table 2).

The Mechanistic Basis of Reversibility of the Neolactonase Trajectory

Why is reversion possible throughout the neolactonase trajectory? We attempted to crystallize the most active neolactonases and their revertant, alongside a few intermediates of this trajectory, but obtained no diffractable crystals (also when the trajectory mutations were grafted to the crystallizable G2E6 variant). However, in any case, structures of these neo-lactonases are unlikely to have revealed a fundamental difference compared with the published structure of PON1-H115W, and in particular, the structures of the revertants are likely to be highly similar to the wild-type structures. This is because lactonase trajectory mutants contained, apart from the 115 mutation, only one other active-site mutation—H134R, in a position that is peripheral to the active site (its closest atom is ~9 Å from the catalytic calcium) (supplementary fig. 4, Supplementary Material online). Three mutations are in a region that serves as the active-site mouth and also anchors PON1 to lipids (K192R, M196T, and F292S; the closest residue to the active site is F292 whose closest atom is ~10 Å from the catalytic calcium). The remaining mutations are in surface residues located far from the active site. Furthermore, with the exception of the initiating H115W mutation, the positions mutated along the neolactonase trajectory all adopted the same or a similar amino acid as in relatively close homologs of PON1, and/or mutation observed in laboratory drifts of PON1 (supplementary table 4, Supplementary Material online).

The neutral nature of the mutations may explain why the evolved neo-lactonases are reversion compatible, but two questions remain: 1) How do the evolved variants act as a proficient lactonase while lacking H115 that is important for the lactonase activity in wild-type. 2) How could a fully active neo-lactonase evolve with such subtle changes in the active site? We suspected that the divergence of these neo-lactonases reinforced an alternative, preexisting catalytic mechanism that is latent in wild-type PON1. The predominating lactonase mechanism in wild-type PON1 is presumed to involve H115 that, jointly with E53, aligns and activates the hydrolytic water (Ben-David et al. 2012). We suspected that E53 could take over completely thus compensating for the loss of H115 (along the above described mechanism of 5m-OPH). Alternatively D269, a calcium-ligating residue that was postulated to protonate the leaving group (Ben-David et al. 2012, 2015), could also act as a general base that deprotonates the nucleophilic water molecule, and could, in principle, also protonate the leaving group in the second step of the reaction.

To examine the above hypotheses, we performed EVB simulations (Warshel and Weiss 1980), starting from the crystal structure of wild-type PON1, and using computationally generated structural models of the trajectory's mutants, as described in the Materials and Methods section. The EVB approach is a hybrid QM/MM approach based on valence bond theory. It uses a classical force-field-based description of different reacting states, transferred to a quantum mechanical framework, in order to describe chemical reactivity. This approach has been successfully applied to describe the mechanisms and dynamics of a broad range of biomolecular systems and has been successfully applied to studies of PON1's mechanism (Ben-David et al. 2015; Blaha-Nelson et al. 2017). We considered two possible mechanisms: 1) The predominant

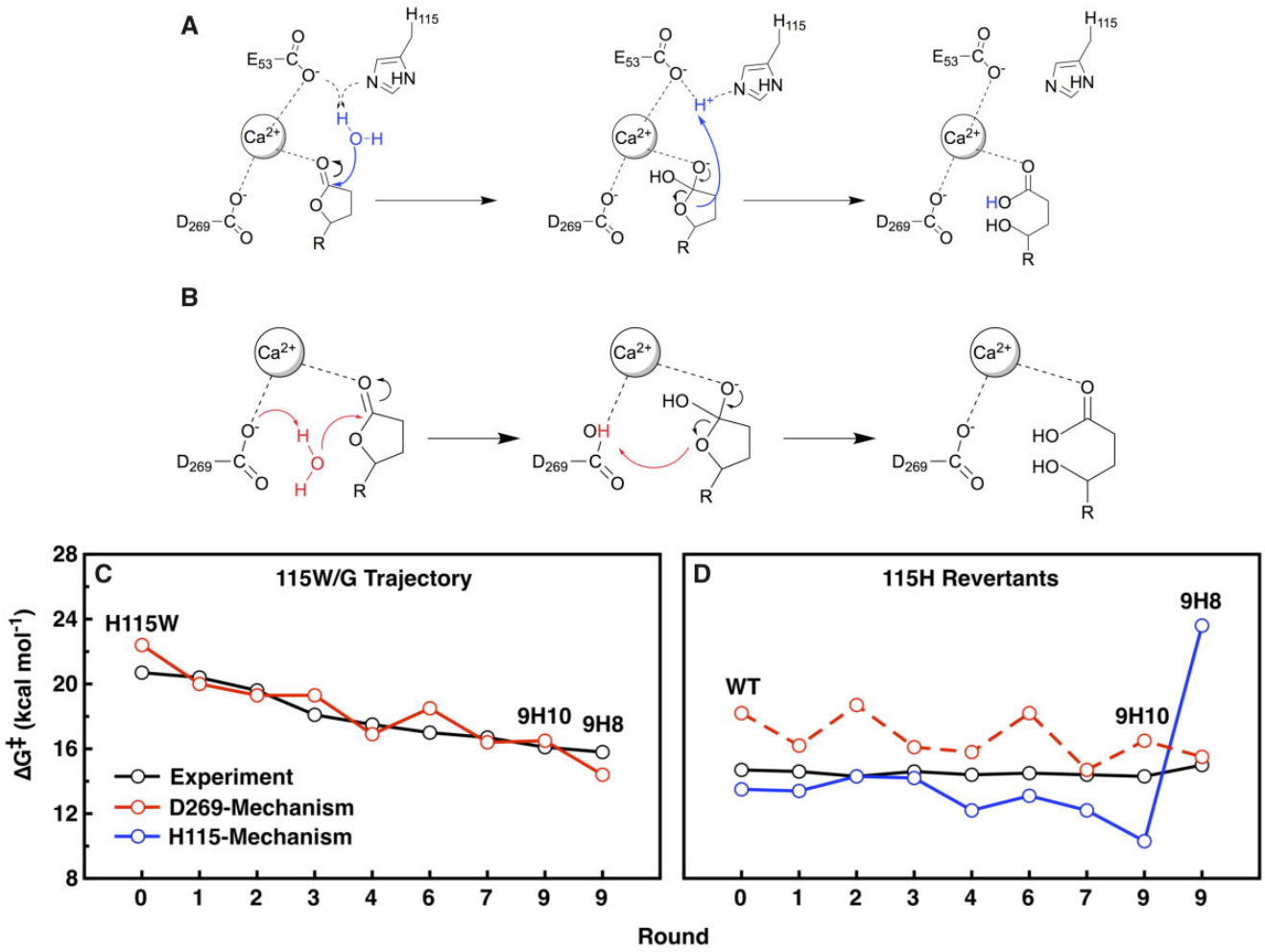

FIG. 4. The alternative lactonase mechanisms of PON1 and their contribution to the evolved neo-lactonase variants and their revertants. (*A*) The dominating lactonase mechanism in wild-type PON1 involves H115 that, jointly with the calcium-ligating E53, aligns and activates the hydrolytic water. The resulting tetrahedral oxyanion intermediate (and the corresponding transition states) is stabilized by the catalytic calcium (see also fig. 5, top). (*B*) Alternatively, another calcium-ligating residue, D269, acts as the activating base. Deprotonation of the water and the concomitant protonation of D269 weaken its interaction with the calcium and may thereby enable further stabilization of the negatively charged oxyanion intermediate (see also fig. 5, middle). When H115 is mutated, this alternative mechanism can replace the primary one (see also fig. 5, bottom) but can also act in the presence of H115. (*C*) Comparison between experimental (black) and calculated activation free energies for the trajectory variants (115W/G) with the alternative D269-based mechanism. (*D*) Comparison between the calculated activation energies for the corresponding revertants (115H) for both the D269-based mechanism (*red*) and the E53/H115-based mechanism (*blue*).

E53-H115-based mechanism (figs. 4*A* and 5, top panel) which is feasible in wild-type and in the 115H revertants (we also considered an alternative mechanism in which E53 acts as a general base on its own, but obtained an activation energy of 31.1 kcal mol^{-1} for wild-type PON1 indicating that this mechanism is irrelevant for the lactonase activity). 2) The D269-based mechanism which is, in principle, feasible for wild-type, the trajectory mutants (115W/G) and their 115H revertants (figs. 4*B* and 5, middle and bottom panels). Our calculations reproduced the experimental data with quantitative accuracy across the entire trajectory (fig. 4, supplementary tables 5 and 6, Supplementary Material online, with representative structures of key stationary points for TBBL hydrolysis by wild-type and mutant PON1 shown in fig. 5). They indicated that the D269-based mechanism is a viable backup to the E53/H115-based mechanism, and that this mechanism was also optimized along the evolution of a neolactonase (fig. 4*C*). However, as shown in figure 4*D* and supplementary tables 5 and 6, Supplementary Material online, upon reversion, the E53/H115-based mechanism took over the D269-based mechanism and could do so along the entire trajectory. That is, for almost all variants studied along this trajectory, the original E53/H115-based mechanism observed in the wild-type enzyme is between 6.0 and 1.8 kcal mol^{-1} lower in energy than the D269-based mechanism. Therefore, reversion of 115H leads PON1 to also revert to the wild-type mechanism, although as can be seen in supplementary table 3, Supplementary Material online, the D269-mechanism (which is the only mechanism available to the H115 mutants) becomes gradually more energetically favorable along the trajectory.

The only exception is the final round variant 9H8 where a complete swap had occurred to the D269-based mechanism. This variant uniquely includes the H134R mutation, and the

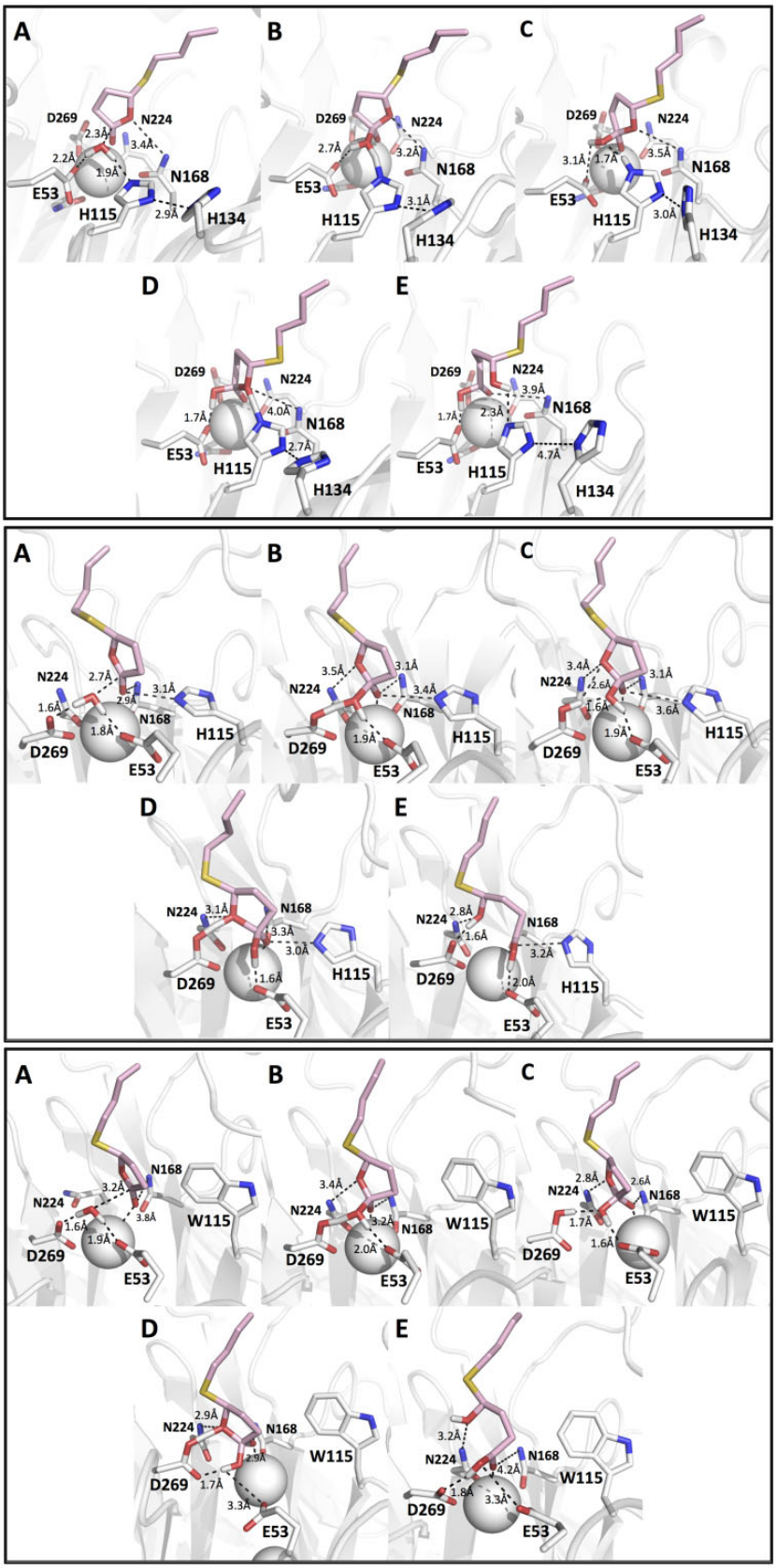

FIG. 5. The alternative lactonase mechanisms of PON1. Shown in each panel are representative snapshots from the EVB simulations depicting the key species along the catalytic cycle. (*A*) The enzyme–substrate complex (with TBBL as substrate), (*B*) the transition state for nucleophilic attack, (*C*) the tetrahedral oxyanionic intermediate, (*D*) the transition state for ring opening, and (*E*) the product (hydroxy acid) complex. Key distances are highlighted. (Top) Wild-type with H115 and E53 jointly acting as general base catalyst (schematically depicted in fig. 4*A*). (Middle) Wild-type with D269 acting as general base (fig. 4*B*). (Bottom) The H115W mutant with D269 acting as general base (fig. 4*B*).

simulations suggest that reversion led to H115 being pulled out of the active site into a noncatalytic "out" conformation (similar to the conformation shown in fig. 2C). However, crucially, although H115 does not contribute to 9H8's hydrolytic activity, it does not interfere either. It appears that in 9H8, the H115-mechanism becomes extremely energetically unfavorable, as shown in figure 4D and supplementary table 6, Supplementary Material online, resulting in the D269-mechanism becoming the preferred mechanism in its revertant. The loss of the H115 mechanism seems to relate to the H134R mutation that uniquely appears in 9H8 (supplementary table 3, Supplementary Material online), and in this mutation promoting the catalytically incompetent "out" rotamer of H115 (as in the 5m-OPH revertants; see fig. 2 and supplementary fig. 2, Supplementary Material online). Finally, as in our prior HREX-MD simulations (supplementary fig. 5, Supplementary Material online), we observed conformational displacement of the catalytic metal ion during our EVB simulations that appears to be correlated with conformational motion of E53 (supplementary fig. 6, Supplementary Material online). Therefore, repositioning of the metal ion may also play a role in the lactonase catalytic cycle.

Overall, the EVB simulations indicate a win-win scenario: The mutations gathered along the neolactonase trajectory amplified an alternative mechanism, as was the case in the OPH trajectory. However, in contrast to the OPH trajectory, at any point along the neolactonase trajectory, reversion to 115H could restore the original mechanism and/or did not disrupt the alternative D269 mechanism.

Discussion

The versatility and plasticity of PON1's catalytic machinery are key to its promiscuity and evolvability. These properties also account for the reversibility of one trajectory and irreversibility of the other. Stabilization of the oxyanionic transition states and intermediates by the calcium cation is the common denominator of all PON1's hydrolytic activities. However, the native lactonase activity, the promiscuous OPH activity, and other promiscuous activities such as hydrolysis of dihydrocoumarin, each use a different subset of catalytic residues. Specifically, the residues that activate the nucleophilic water molecule and those that protonate the leaving group exchange roles when catalyzing one activity or another. The calcium-ligating residues, E53 and D269, seem particularly versatile and so is H115. Neo-functionalization led to 5m-OPH—a variant of PON1 with high OPH activity and nearly no lactonase activity, where, as indicated by the crystal structures, a switch in mechanism had occurred. This switch is manifested mostly in the repositioning of E53, and the appearance of a bridging water molecule that most likely mimics the attacking activated water (fig. 1G). Upon reversion of 115 (i.e., in 5m-OPH-115H), E53 dislocates from the metal ion and cannot act as the activating base, and the newly introduced H115 side chain also assumes an inactive "out" conformation (fig. 2 and supplementary fig. 1F, Supplementary Material online). Consequently, both the evolved OPH and the original lactonase activities are lost. Similarly, profound structural perturbations have been seen upon reversions in an evolved bimetallo-enzyme, where the reversion mutations altered, for example, the distance between the catalytic metal ions (Kaltenbach et al. 2015). To have both E53 and H115 assume their active conformations, and thus regain the original lactonase activity, three additional mutations must be simultaneously reverted.

Reversibility in the newly diverged 5m-OPH is blocked because the newly evolved mechanism lacks a robust network of second-shell interactions that reinforce, and/or backup E53's new configuration, and its ligated water molecule, as is the case with PON1's native activity. Indeed, in our re-functionalization trajectory, two mechanisms, the original one (E53 and H115 activating the attacking water) and the alternative "backup" one (D269 as the activating base), coincide throughout this trajectory. Only in one of the last round variants, 9H8, did the alternative, D269, mechanism take over. However, although reintroduction of H115 did not revert the mechanism it did not disturb the new one either. It therefore follows that a preexisting mechanism, even if an alternative one, is more robust to reversion than a newly developed mechanism. Indeed, the new OPH activity and mechanism were instated via mutations in first-shell active-site residues and reconfiguration of key residues (primarily E53). These changes were profound and thus blocked reversibility. In contrast, the preexisting D269 mechanism could be developed via second- and even third-shell mutations and was thus amenable to reversibility. We note, however, that the alternative D269-based mechanism is not an inherent, evolved trait per se. Rather, we hypothesize that it is the outcome of optimizing PON1's catalytic efficiency, that is, of providing additional interactions that reinforce and optimize the primary mechanism.

Our results also confirm the notion that adaptive, neofunctionalization trajectories tend to be more highly epistatic, and irreversible, compared with "neutral drift"—the gradual accumulation of sequence exchanges while maintaining the same enzymatic function and mechanism. In fact, the reversibility of PON1's re-functionalization trajectory indicates a route by which active-site residues, that are generally highly conserved, may change by gradual, largely neutral drift. Active-site residues are the most conserved among all protein residues. Nonetheless, along long evolutionary times, even key active-site positions may change. However, the route followed here, namely, gradual compensation of a highly deleterious mutation in a key active-site position (H115W), is not a likely option for natural evolution. Reversion is by far the most likely outcome. Consider, however, the H115 revertants of the last rounds of our trajectory, that is, 9H8-, or 9H10-115H. These represent transition sequences that could be reached neutrally (i.e., while maintaining H115), because the revertants of each and every variant along the refunctionalization trajectory are nearly as active as wild-type PON1. Once these transition sequences have been reached, His115 can exchange to Gly with minimal loss of activity. Namely mutations that are largely neutral may enable the acceptance of a mutation that is otherwise highly deleterious (H115W/G in this case) (Wellner et al. 2013; Starr and Thornton 2016).

Exchanges in active-site residues are relatively rare. The laboratory re-functionalization trajectory analyzed here suggests that exchanges in key active-site residues that did occur in orthologous enzymes, occurred via transition sequences and enabling mutations. If such an exchange had occurred, it can be tracked down via an approach similar to the one applied here, namely re-functionalization of an enzyme in which the key catalytic residue was mutated. In phosphoglycerate kinase (PGK), for example, an exchange in a key catalytic residue occurred along its evolutionary history—position 219 diverges, and is Lys in all bacterial and eukaryote sequences, while mostly Ser, but never Lys, is found at this position in all archaeal PGKs. The K219S mutant of human PGK is largely inactive, yet regains activity by the accumulation of just two mutations in second-shell residues. Further, as seen in the refunctionalized PON1, these two mutations lead to a transition variant that is functional with both Ser and Lys at position 219, as well as with other amino acids seen in archaeal PGKs (Wellner et al. 2013).

Thus, given enough time, and some variability in selection levels, even the most conserved and functionally essential residues may change (Povolotskaya and Kondrashov 2010). In support that such changes may occur via neutral, or nearly neutral, enabling mutations we note that the positions mutated along the neolactonase trajectory all adopted the same or a similar amino acid as in relatively close homologs of PON1, and/or mutations observed in laboratory drifts of PON1 (supplementary table 4, Supplementary Material online).

Finally, we note that, although comparable in their lactonase activity, the revertants of the re-functionalization trajectory show nearly 1,000-fold lower paraoxonase activity compared with wild-type PON1. Thus, although the enabling mutations may be neutral with respect to the native lactonase activity that has been under selection, they do alter the active site with respect to alternative, promiscuous activities. This phenomenon is manifested in orders-of-magnitude variations in the level of promiscuous activities among orthologous enzymes that are rather similar in their native function (e.g., Afriat et al. 2006).

Materials and Methods

PON1 Mutants

Human serum PON1 is a well characterized enzyme (La Du et al. 1999). However, the human enzyme does not readily express in a soluble and functional form in *Escherichia coli*. We have generated PON1 variants by shuffling of several different mammalian PON1s including human PON1; these recombinant PON1s are well expressed and exhibit enzymatic parameters that are highly similar to human PON1 with an entire range of substrates (Aharoni et al. 2004), including with TBBL, a chromogenic lactone substrate that is routinely applied to assay PON1's lactonase activity (Khersonsky and Tawfik 2006), and with paraoxon that represents the OPH activity. One variant, dubbed G3C9, exhibited the highest stability and could also be expressed without a thioredoxin fusion tag (Aharoni et al. 2004) and has thus been the basis of all our directed evolution experiments. Albeit, this variant failed to crystallize, and hence all structural work has been done with G2E6 that exhibits very similar enzymatic parameters (Harel et al. 2004). These variants, G2E6 and G3C9, are nearly identical (94.4% amino acid identity). Nonetheless, to ensure that the enzymatic and structural analyses are consistent, we have grafted the active-site mutations of 5m-OPH that were identified by directed evolution at the background of G3C9 onto variant G2E6 (see also table 1). Mutants for both the G2E6 and G3C9 variants were generated by the Transfer-PCR (polymerase chain reaction) platform (Erijman et al. 2014) and verified by DNA sequencing.

Purification and Kinetics

Expression and purification of G2E6 mutants were performed as described (Aharoni et al. 2004) except that Ni-NTA chromatography comprised the last purification step. Protein purity was monitored by SDS-PAGE. Kinetic assays were performed in 50 mM Tris, 50 mM NaCl, 1 mM $CaCl_2$, pH 8.0, at room temperature. Rates with paraoxon (phosphotriesterase activity) and TBBL (lactonase activity) were determined as described (Aharoni et al. 2004; Khersonsky and Tawfik 2006), at substrate concentrations in a range from $0.3 \times K_M$ up to $(2–3) \times K_M$, and k_{cat}, K_M, and k_{cat}/K_M were obtained by fitting the data to the Michaelis–Menten model with Prism v. 7.04 (GraphPad Software). In cases in which substrate solubility was limiting, k_{cat}/K_M values were extracted from linear fits. Presented data are the mean of $\geq$2 independent experiments, and the error ranges represent the standard error of the mean.

Directed Evolution of the Neolactonase

We performed random mutagenesis by error-prone PCR of the G3C9-H115W mutant using mutazyme (Genemorph), at ~2 mutations/gene/round. The mutated genes were recloned into the overexpression pET32 plasmid (Novagen) and transformed into *E. coli* as previously described (Aharoni et al. 2004). Individual colonies were grown, lysed, and screened in 96-well plates using the chromogenic lactone TBBL, essentially as described (Amitai et al. 2007). At the starting point background of the H115W mutation, the activity of variants showing lactonase activity was too weak to be reliably detected in crude lysates; we thus performed the first round by screening for paraoxonase activity, thereby drifting to accumulate activity-modifying mutations (Amitai et al. 2007). The active variants isolated from this round were subsequently screened with TBBL. In each round, ~1,000 individual clones were screened, and ~10 positive variants were isolated (generally showing $\geq$1.2-fold TBBL hydrolyzing rate compared with the most active variant[s] from the previous round). These were regrown and assayed to confirm their activity, sequenced, and their coding genes served as a template for the next round of random mutagenesis and screening. Selected variants along the trajectory were purified and their kinetic parameters were determined as described above.

Crystallization, Data Collection, and Refinement

Crystallization trials were performed on concentrated solutions (10–20 mg/ml) of the 3m-OPH, 5m-OPH, and 5m-OPH-115H mutants at the background of the G2E6 PON1 variant, as previously described (Ben-David et al. 2012, 2013). Briefly, using Linbro plates (Hampton Research) and hanging drop vapor diffusion, crystals were formed by mixing equal volumes of protein with the precipitant buffer containing 20% PEG 3350, 0.2 M NaBr, 0.1 M Bis-tris propane, pH 6.5 (incubation at 19 °C). Complete data sets for 3m-OPH, 5m-OPH, and 5m-OPH-115H were collected on beamlines ID-14-1, ID-14-4, and ID-23-1, respectively, at the ESRF Synchrotron (Grenoble, France). The diffraction images were indexed, integrated, and scaled using the HKL2000 program package (Otwinowski and Minor 1997). Structure determination was carried out by molecular replacement (PHASER, CCP4 (McCoy et al. 2007)) using the published pH 6.5 structure (PDB: 3SRE). All steps of atomic refinement were carried out with FHENIX (Adams et al. 2010), the model was built into the current coordinates using the program Coot (Emsley and Cowtan 2004) and figures depicting structures were prepared using PyMOL (DeLano Scientific LLC, San Carlos, CA, http://www.pymol.org; last accessed December 28, 2019). Details of data collection and refinement statistics are displayed in supplementary table 1, Supplementary Material online. All three PON1 mutants did not display well-defined electron density for the first ~20 residues at the N-terminus. The 5m-OPH mutant showed higher resolution relative to the other two mutants and thus more details, such as one molecule of *n*-dodecyl-β-D-maltoside (only the sugar moiety), polyethylene glycol, and Bis-tris propane, were captured. Additionally, the 5m-OPH mutant showed unassigned electron density above the catalytic calcium that could correspond to either a phosphate ion, in a similar fashion to that seen in WT PON1 structures (PDB IDs: 3SRE and 1V04), or an alternative position of the catalytic calcium (Ben-David et al. 2013); however, none of these options was compatible with the model refinement, presumably due to lack of a well-defined electron density.

HREX-MD Simulations

Two sets of simulations were performed in this work: HREX-MD simulations (Bussi 2014) of substrate-free PON1, in order to explore the conformational space of the H115 side chain, as well as EVB simulations (Warshel and Weiss 1980) of the hydrolysis of TBBL by wild-type and mutant PON1 variants. All simulations were performed using the OPLS-AA force field (Jorgensen et al. 1996), with HREX-MD and EVB simulations performed in GROMACS v. 5.1.4 (van der Spoel et al. 2005; Abraham et al. 2015) and Q (Marelius et al. 1998), respectively. HREX-MD simulations were performed on wild-type PON1 (PDB ID: 3SRG [Ben-David et al. 2012]), the 5m-OPH mutant (PDB ID: 6H0A), the 5m-OPH-115H mutant (PDB ID: G6MU), as well as manually generated 5m-OPH-115H/G69L, 5m-OPH-115H/R134H, 5m-OPH-115H/F222S, 5m-OPH-111H/G69L/R134H, and 5m-OPH-115H/G69L/R134H/F222S mutants. All simulations were performed of PON1 in complex with a 98-lipid *n*-dodecyl-β-maltoside (DDM) micelle, with the micelle generated using the Micelle Maker web server. The micelle was generated by Micelle Maker (Krüger and Kamerlin 2017) using the standard Micelle Maker protocol and without postprocessing. The resulting structure was then described using the OPLS-AA force field in our HREX-MD simulations, for consistency with the protein. OPLS-AA parameters for DDM were taken from Damm et al. (1997) and Siu et al. (2012), to describe the carbohydrate and the hydrocarbon tail, respectively.

The manual G69L, R134H, and F222S reversions were inserted into the 5m-OPH-115H structure using PyMOL's Mutagenesis Wizard. The protein was manually attached to the micelle using PyMOL's Editing Mode, such that the micelle interacts with the surface of PON1 through helices H1 and H2 (see discussion in e.g., Harel et al. 2004). The final PON1-micelle complexes were then placed in an octahedral box of TIP3P water molecules (Jorgensen et al. 1983) extending at least 10 Å away from the surface of the complex in all directions. For both HREX-MD and EVB simulations, the two calcium ions in the structure were modeled by a multisite model, in-line with our previous simulations, and using the parameters presented in Ben-David et al. (2015) and Blaha-Nelson et al. (2017). Na^+ and Cl^- ions were added to the system until a neutral system with a salt concentration of 0.15 mol/l was obtained. The resulting systems were then subjected to 5,000 steps of each of steepest decent and conjugate gradient minimizations, after which they were heated up from 0 to 300 K over a 500 ps of simulation time in an NVT ensemble, and then equilibrated for a further 500 ps in an NPT ensemble. Positional restraints (1,000 kJ $mol^{-1} nm^{-2}$) were applied to all heavy atoms in each of the *xyz* directions during this equilibration step, and the positional restraints where then gradually reduced to 100 kJ $mol^{-1} nm^{-2}$ over 5 steps of 500 ps NPT equilibration. Afterward, a 20 ns NPT equilibration was performed with no restraints on the DDM micelle, but keeping weak (100 kJ $mol^{-1} nm^{-2}$) restraints on the protein, to allow the micelle to equilibrate to the protein. During this 20 ns equilibration, the DDM detergents were free to assemble around the interface with the protein (interacting with the H1 and H2 helices). Finally, a last 5 ns NPT MD simulation was carried out without any restraints on either the protein or the micelle, so that the complex of the protein and the micelle would be fully equilibrated and stable (see supplementary figs. 7 and 8, Supplementary Material online). The final configuration of the system was then used as the starting structure for the HREX-MD simulations, which were performed without any restraints on either the DDM micelle or the protein during the simulations.

All HREX-MD simulations were performed using GROMACS v. 5.1.4, interfaced with the PLUMED 2.3.0 plugin (Tribello et al. 2014). A 1 fs integration step was used in the simulations, and all bonds were constrained with the P-LINCS algorithm (Hess 2008). The cutoff for the short-range nonbonded interactions was set to 10 Å, whereas the long-range electrostatic interactions were calculated using the particle mesh Ewald method (Darden et al. 1993), in combination with periodic boundary conditions. All HREX-MD simulations were performed in an NPT ensemble, which was maintained

using the velocity rescale thermostat and the Parrinello–Rahman barostat (Parrinello and Rahman 1981). Residues 68–73, 115, and 134 were chosen as the "hot" region for the HREX-MD simulations, and the simulations were performed using eight replicas, with the scaling factor, λ, exponentially ranging from 1.0 to 0.6, which corresponds to an effective temperature range from 300 to 500 K. An exchange between neighboring replicas was attempted every 2 ps, and each replica was sampled for 50 ns with configurations being saved every 2 ps. This resulted in an average exchange rate of 40–50%. Only the replica without scaling ($\lambda = 1.0$) was used for further analysis. For all the eight systems, the accumulated sampling time is 3.2 μs.

EVB Simulations

EVB simulations of TBBL hydrolysis by mutant and revertant PON1 variants were performed using the crystal structures of the wild-type (PDB ID: 3SRG [Ben-David et al. 2012]) and H115W mutant PON1 (PDB ID: 4HHO [Ben-David et al. 2013]) as starting points for each trajectory. That is, for each of the mutant and revertant trajectories, the relevant variants for each clone considered were manually inserted into either the wild-type or the H115W structure, using the Richardson rotamer library (Lovell et al. 2000) as implemented in Chimera (Pettersen et al. 2004). The wild-type structure was used for all revertants; in the case of the mutants, the H115W structure was used for all clones with an H115W or H115C mutation (1H1, 2H2, 3H4, and 4H1) and the wild-type structure was used for all clones with an H115G mutation (6H15, 7H17, 9H10, and 9H8; for sequences see supplementary table 3, Supplementary Material online). This is because, as discussed in Ben-David et al. (2013), substitution of H115 to a bulky tryptophan causes a major rearrangement in the active site and in particular the position of catalytic metal ion, but this shift is unlikely to occur in the case of the H115G mutation, making the wild-type crystal structure a much better starting point. We have modeled TBBL hydrolysis through the "classical" H115-mechanism by wild-type and mutant PON1 variants in great detail elsewhere, and the protocol and parameters used for the current study mirror our recent work, and therefore we will only describe the protocol in brief and refer the readers to Blaha-Nelson et al. (2017) for further details. We note that the only major difference between the current simulation setup and that in Blaha-Nelson et al. (2017) is that we now use a larger 30 Å simulation sphere (see below), in order to capture the effect of mutations further away from the active site, which requires some small modifications in our EVB calibration and system setup.

Our EVB simulations were performed using a multilayer model, in which all atoms within the inner 85% of a simulation sphere of 30 Å, centered on the catalytic calcium ion (again described using a multisite model as in the HREX-MD), were fully mobile in our simulations, atoms within the outer 15% of the sphere were restrained to their crystallographic positions using 10 kcal mol^{-1}Å^{-2} positional restraints, and all atoms outside this region were restrained to their crystallographic positions using 100 kcal mol^{-1}Å^{-2} positional restraints. Only residues within the mobile region were ionized, as described in Blaha-Nelson et al. (2017), to prevent system instabilities by having charged residues outside the mobile region. This model, which is standard for multiscale simulations of chemical reactivity in spherical boundary conditions, both allows for faster convergence of the calculations through not needing the entire system to be flexible, and also mimics to some extent the stabilizing effect of the micelle/membrane, as described in Ben-David et al. (2015) and Blaha-Nelson et al. (2017). The protonation patterns of all histidines in the system as well as all ionizable residues within the mobile region were determined by PROPKA 3.1 (Olsson et al. 2011) coupled with visual inspection, and the key ionized residues are shown in supplementary table 7, Supplementary Material online. The TBBL substrate and nucleophilic water molecule were manually placed in the active site, guided by the conformation of the 2-hydroxyquinoline inhibitor in the wild-type crystal structure, in order to obtain an optimal reactive geometry between the substrate, the nucleophilic water molecule, and the relevant general base (either H115 or D269), depending on whether the H115- or the D269-mechanism was being modeled, note that as H115 and D269 are in different parts of the active site, this results in different substrate positions in the Michaelis complex. The resulting structure was then immersed in a 30 Å droplet of TIP3P water molecules centered on the catalytic calcium ion, described using the surface constrained all atom solvent model (Warshel and King 1985), before performing equilibration and subsequent EVB simulations.

For each system, three individual equilibration runs were performed, during which time the system was gradually heated from 0 to 300 K over 240 ps simulation time, and during which time strong (200 kcal mol^{-1}Å^{-2}) positional restraints on all crystallographic coordinates were gradually dropped to 0 until only 0.5 kcal mol^{-1}Å^{-2} restraints remained on the TBBL substrate and reacting water molecule. These weak restraints on only the reacting fragments were retained for the remainder of the equilibration and EVB simulations. Once the system had been heated to 300 K, a further 22.5 ns of equilibration was performed at 300 K for each system (supplementary fig. 9, Supplementary Material online), and the end-point of each equilibration was used as the starting point for five independent EVB trajectories, generated by equilibrating the system for a further 400 ps using five independent random seeds. Thus, 15 independent EVB trajectories in total were generated based on the initial 3 equilibration runs. Each EVB simulation was performed in 51 EVB free energy perturbation/umbrella sampling windows of 200 ps in length each, leading to a total of 10.2 ns of simulation time per EVB trajectory, 153 ns of simulation time per system, and 4.131 μs of EVB simulations over all systems and mechanisms. All simulations were performed using the OPLS-AA force field with a 1 fs step size, and with a 10-Å cutoff for nonbonded interactions, with the exception of reacting atoms, which were described using a 99 Å cutoff (i.e., essentially no cutoff). Long-range interactions were described using the local reaction field approach (Lee and Warshel 1992). The EVB states and parameters used to describe TBBL hydrolysis

are presented in detail in Blaha-Nelson et al. (2017, supporting information), as are the experimental and computational data used for the calibration of the EVB parabola to the uncatalyzed reaction in aqueous solution. Note that in the case of the D269-based mechanism shown in figure 4, the EVB parameters for protonated and unprotonated D269 are those presented for the analogous D269-based hydrolysis of paraoxon in (Blaha-Nelson et al. 2017) (the other parameters are identical for both mechanisms). The only parameters that have changed are the EVB mapping parameters, due to the larger 30 Å sphere size, and these are presented in supplementary table 8, Supplementary Material online. For comprehensive simulation details, see Blaha-Nelson et al. (2017).

Finally, note that we also attempted to simulate TBBL and paraoxon hydrolysis by 5m-OPH and 5m-OPH-115H using the EVB approach. Here, however, all our simulations, starting from multiple starting configurations for the different substrates and for the nucleophilic water molecule gave very high activation free energies. The reasons for failure are unclear, especially in view of the success in simulating the neolactonase trajectory and its revertants (described in the Results and the Discussion), as well as simulations of wild-type PON1 and several PON1 mutants in previous work (Ben-David et al. 2015; Blaha-Nelson et al. 2017). However, our EVB simulations suggest that the calcium ion in the upward displaced location observed in the crystal structures of H115W, 3m-OPH, and 5m-OPH (fig. 1A) is extremely floppy, leading to rearrangement of the metal position and metal coordinating residues in the EVB simulations compared with the crystal structures. It is unclear to us therefore whether these mutations cause either a change in substrate binding position or a change in mechanism, that we do not capture in simulations of these particular enzyme variants.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank the Israel Structural Proteomics Centre, for access to their protein purification and crystallization facilities, and the Swedish National Infrastructure for Computing (SNIC) for computational time (SNIC 2017/12-11 and 2018/2-3). Funding by the Sasson & Marjorie Peress Philanthropic Fund, the Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellowships, KAW 2013.0124 and KAW 2018.0140), and in part, by the U.S. Defence Threat Reduction Agency (DTRA) (contract number HDTRA11710057) is gratefully acknowledged. D.S.T. is the incumbent of the Nella and Leon Benoziyo Professorial Chair. Finally, we thank Dennis M. Krüger for assistance with initial simulation setup, and Peter M. Kasson for assistance with data analysis.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. 2015. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25.

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 66(2):213–221.

Afriat L, Roodveldt C, Manco G, Tawfik DS. 2006. The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. *Biochemistry* 45(46):13677–13686.

Afriat-Jurnou L, Jackson CJ, Tawfik DS. 2012. Reconstructing a missing link in the evolution of a recently diverged phosphotriesterase by active-site loop remodeling. *Biochemistry* 51(31):6047–6055.

Aharoni A, Gaidukov L, Yagur S, Toker L, Silman I, Tawfik DS. 2004. Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization. *Proc Natl Acad Sci U S A*. 101(2):482–487.

Amitai G, Gupta RD, Tawfik DS. 2007. Latent evolutionary potentials under the neutral mutational drift of an enzyme. *HFSP J*. 1(1):67–78.

Bar-Rogovsky H, Hugenmatter A, Tawfik DS. 2013. The evolutionary origins of detoxifying enzymes: the mammalian serum paraoxonases (PONs) relate to bacterial homoserine lactonases. *J Biol Chem*. 288(33):23914–23927.

Ben-David M, Elias M, Filippi JJ, Dunach E, Silman I, Sussman JL, Tawfik DS. 2012. Catalytic versatility and backups in enzyme active sites: the case of serum paraoxonase 1. *J Mol Biol*. 418(3–4):181–196.

Ben-David M, Sussman JL, Maxwell CI, Szeler K, Kamerlin SCL, Tawfik DS. 2015. Catalytic stimulation by restrained active-site floppiness—the case of high density lipoprotein-bound serum paraoxonase-1. *J Mol Biol*. 427(6):1359–1374.

Ben-David M, Wieczorek G, Elias M, Silman I, Sussman JL, Tawfik DS. 2013. Catalytic metal ion rearrangements underline promiscuity and evolvability of a metalloenzyme. *J Mol Biol*. 425(6):1028–1038.

Blaha-Nelson D, Kruger DM, Szeler K, Ben-David M, Kamerlin SC. 2017. Active site hydrophobicity and the convergent evolution of paraoxonase activity in structurally divergent enzymes: the case of serum paraoxonase 1. *J Am Chem Soc*. 139(3):1155–1167.

Bridgham JT, Ortlund EA, Thornton JW. 2009. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* 461(7263):515–519.

Bussi G. 2014. Hamiltonian replica exchange in GROMACS: a flexible implementation. *Mol Phys*. 112(3–4):379–384.

Cheema J, Faraldos JA, O'Maille PE. 2017. Epistasis and dominance in the emergence of catalytic function as exemplified by the evolution of plant terpene synthases. *Plant Sci*. 255:29–38.

Damm W, Frontera A, Tirado-Rives J, Jorgensen WL. 1997. OPLS all-atom force field for carbohydrates. *J Comput Chem*. 18(16):1955–1970.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: an *N*·log(*N*) method for Ewald sums in large systems. *J Chem Phys*. 98(12):10089–10092.

Dasmeh P, Girard E, Serohijos A. 2017. Highly expressed genes evolve under strong epistasis from a proteome-wide scan in *E. coli*. *Sci Rep*. 7(1):15844.

Draganov DI, Teiber JF, Speelman A, Osawa Y, Sunahara R, La Du BN. 2005. Human paraoxonases (PON1, PON2, and PON3) are lactonases with overlapping and distinct substrate specificities. *J Lipid Res*. 46(6):1239–1247.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 60(12):2126–2132.

Erijman A, Shifman JM, Peleg Y. 2014. A single-tube assembly of DNA using the transfer-PCR (TPCR) platform. *Methods Mol Biol*. 1116:89–101.

Gong LI, Bloom JD. 2014. Epistatically interacting substitutions are enriched during adaptive protein evolution. *PLoS Genet*. 10(5):e1004328.

Gupta A, Adami C. 2016. Strong selection significantly increases epistatic interactions in the long-term evolution of a protein. *PLoS Genet*. 12(3):e1005960.

Gupta RD, Goldsmith M, Ashani Y, Simo Y, Mullokandov G, Bar H, Ben-David M, Leader H, Margalit R, Silman I, et al. 2011. Directed

evolution of hydrolases for prevention of G-type nerve agent intoxication. *Nat Chem Biol*. 7(2):120–125.
Harel M, Aharoni A, Gaidukov L, Brumshtein B, Khersonsky O, Meged R, Dvir H, Ravelli RB, McCarthy A, Toker L, et al. 2004. Structure and evolution of the serum paraoxonase family of detoxifying and antiatherosclerotic enzymes. *Nat Struct Mol Biol*. 11(5):412–419.
Harms MJ, Thornton JW. 2013. Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet*. 14(8):559–571.
Hess B. 2008. P-LINCS: a parallel linear constraint solver for molecular simulation. *J Chem Theory Comput*. 4(1):116–122.
Jorgensen WL, Chandresekhar J, Madura JD. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 76:926.
Jorgensen WL, Maxwell DS, Tirado-Rives J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc*. 118:11225–11236.
Kaltenbach M, Burke JR, Dindo M, Pabis A, Munsberg FS, Rabin A, Kamerlin SCL, Noel JP, Tawfik DS. 2018. Evolution of chalcone isomerase from a noncatalytic ancestor. *Nat Chem Biol*. 14(6):548–555.
Kaltenbach M, Jackson CJ, Campbell EC, Hollfelder F, Tokuriki N. 2015. Reverse evolution leads to genotypic incompatibility despite functional and active site convergence. *eLife* 4:e06492.
Khersonsky O, Tawfik DS. 2005. Structure–reactivity studies of serum paraoxonase PON1 suggest that its native activity is lactonase. *Biochemistry* 44(16):6371–6382.
Khersonsky O, Tawfik DS. 2006. Chromogenic and fluorogenic assays for the lactonase activity of serum paraoxonases. *ChemBioChem* 7(1):49–53.
Krüger DM, Kamerlin S. 2017. Micelle Maker: an online tool for generating equilibrated micelles as direct input for molecular dynamics simulations. *ACS Omega* 2(8):4524–4530.
La Du BN, Aviram M, Billecke S, Navab M, Primo-Parmo S, Sorenson RC, Standiford TJ. 1999. On the physiological role(s) of the paraoxonases. *Chem Biol Interact*. 119–120:379–388.
Lee FS, Warshel A. 1992. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. *J Chem Phys*. 97(5):3100–3107.
Lovell SC, Word JM, Richardson JS, Richardson DC. 2000. The penultimate rotamer library. *Proteins* 40(3):389–408.
Marelius J, Kolmodin K, Feierberg I, Åqvist J. 1998. Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. *J Mol Graph Model*. 16(4–6):213–225.
McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr*. 40(4):658–674.
Natarajan C, Inoguchi N, Weber RE, Fago A, Moriyama H, Storz JF. 2013. Epistasis among adaptive mutations in deer mouse hemoglobin. *Science* 340(6138):1324–1327.
Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. 2011. PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J Chem Theory Comput*. 7(2):525–537.
Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 276:307–326.
Pareek V, Samanta M, Joshi NV, Balaram H, Murthy MR, Balaram P. 2016. Connecting active-site loop conformations and catalysis in triosephosphate isomerase: insights from a rare variation at residue 96 in the plasmodial enzyme. *ChemBioChem* 17(7):620–629.
Parrinello M, Rahman A. 1981. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 52(12):7182–7190.
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 25(13):1605–1612.
Povolotskaya IS, Kondrashov FA. 2010. Sequence space and the ongoing expansion of the protein universe. *Nature* 465(7300):922–926.
Shah P, McCandlish DM, Plotkin JB. 2015. Contingency and entrenchment in protein evolution under purifying selection. *Proc Natl Acad Sci U S A*. 112(25):E3226–E3235.
Siu SWI, Pluhackova K, Böckmann RA. 2012. Optimization of the OPLS-AA force field for long hydrocarbons. *J Chem Theory Comput*. 8(4):1459–1470.
Soylemez O, Kondrashov FA. 2012. Estimating the rate of irreversibility in protein evolution. *Genome Biol Evol*. 4(12):1213–1222.
Starr TN, Thornton JW. 2016. Epistasis in protein evolution. *Protein Sci*. 25(7):1204–1218.
Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. 2014. PLUMED 2: new feathers for an old bird. *Comput Phys Commun*. 185(2):604–613.
van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen H. 2005. GROMACS: fast, flexible, and free. *J Comput Chem*. 26(16):1701–1718.
Warshel A, King G. 1985. Polarization constraints in molecular dynamics simulations of aqueous solutions: the surface constraint all atom solvent (SCAAS) model. *Chem Phys Lett*. 121(1–2):124–129.
Warshel A, Weiss RM. 1980. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J Am Chem Soc*. 102(20):6218–6226.
Wellner A, Raitses Gurevich M, Tawfik DS. 2013. Mechanisms of protein sequence divergence and incompatibility. *PLoS Genet*. 9(7):e1003665.
Wu NC, Thompson AJ, Xie J, Lin CW, Nycholat CM, Zhu X, Lerner RA, Paulson JC, Wilson IA. 2018. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. *Nat Commun*. 9(1):1264.