Åbo Akademi

OPEN ACCESS

This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Distinctive structural motifs co-ordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes

Dimitriou, Polytimi; Denesyuk, Alexandre; Nakayama, T; Johnson, Mark S; Denessiouk, Konstantin

Published in:
Protein Science

DOI:
10.1002/pro.3527

Published: 01/01/2019

Document Version
Accepted author manuscript

Document License
Publisher rights policy

Link to publication

Please cite the original version:
Dimitriou, P., Denesyuk, A., Nakayama, T., Johnson, M. S., & Denessiouk, K. (2019). Distinctive structural motifs co-ordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes. *Protein Science*, *28*(2), 344–364. https://doi.org/10.1002/pro.3527

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

This document is downloaded from the Research Information Portal of ÅAU: 25. Apr. 2025

Protein Science

Conserved structural motifs coordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes

Journal:	*Protein Science*
Manuscript ID	Draft
Wiley - Manuscript type:	Full-Length Papers
Date Submitted by the Author:	n/a
Complete List of Authors:	Dimitriou, Polytimi; Åbo Akademi University, Faculty of Science and Engineering, Biochemistry Denesyuk, Alexander; Åbo Akademi University, Faculty of Science and Engineering, Biochemistry; Institute for Biological Instrumentation Nakayama, Toru; Tohoku University, Biomolecular Engineering Johnson, Mark; Abo Akademi University, Biochemistry and Pharmacy Denessiouk, Konstantin; Abo Akademi Institutionen for naturvetenskaper, Faculty of Science and Engineering, Pharmaceutical Sciences/Biochemistry
Keywords:	alpha/beta-hydrolases, catalytic triad, carboxylesterase, structural motif, structural framework

SCHOLARONE™
Manuscripts

Conserved structural motifs coordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes

Polytimi S. Dimitriou,[1] Alexander I. Denesyuk,[1,2] Toru Nakayama,[3] Mark S. Johnson,[1] Konstantin Denessiouk* [1,4]

[1]Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Turku 20520, Finland
[2]Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino 142290, Russia
[3]Tohoku University, Biomolecular Engineering, Aoba 6-6-11, Aramaki, Aoba-ku, Sendai, Miyagi 980-8579, Japan
[4]Pharmaceutical Sciences Laboratory, Faculty of Science and Engineering, Pharmacy, Åbo Akademi University, Turku 20520, Finland
*Corresponding author: Faculty of Science and Engineering, Biochemistry, Åbo Akademi University, Tykistökatu 6, BioCity 3A, 20520, Turku, Finland
Email address: kdenessi@abo.fi (Konstantin Denessiouk)

Running title: Structural motifs in ABH fold enzymes (50 characters max)

(List of total number of manuscript pages, supplementary material, tables and figures:)

- Manuscript pages: 25
- Supplementary Material pages: 10
- Number of Tables: 3
- Number of Figures: 2

(Description of supplementary material including filenames:)

The Supplementary Material includes 2 tables, Table S1 and Table S2, which contain information about contacts in the periphery of the nucleophile zone and water molecule-based interactions in the oxyanion zone, respectively.

ABSTRACT

The alpha/beta-hydrolases (ABH) are among the largest structural families of proteins that are found in nature. Although they vary in their sequence and function, the ABH enzymes use a similar acid-base-nucleophile catalytic mechanism to catalyze reactions on different substrates. Because ABH enzymes are biocatalysts with a wide range of potential applications, protein engineering has taken advantage of their catalytic versatility to develop enzymes with industrial applications. This study is a comprehensive analysis of 40 ABH enzyme families focusing on two identified conserved substructures: the nucleophile zone and the oxyanion zone, which coordinate the catalytic nucleophile and the residues of the oxyanion hole, and independently reported as critical for the enzymatic activity. We also observed a conserved aromatic cluster near the nucleophile and oxyanion zones, and opposite the ligand-binding site. The nucleophile zone, the oxyanion zone and the residue cluster enriched in aromatic side chains comprise a three-dimensional structural organization that shapes the active site of ABH enzymes and plays an important role in the enzymatic function by structurally stabilizing the catalytic nucleophile and the residues of the oxyanion hole. The structural data support the notion that the aromatic cluster can participate in coordination of the catalytic histidine loop, and properly place the catalytic histidine next to the catalytic nucleophile.

KEY WORDS (4-10 keywords or short phrases): alpha/beta-hydrolases, catalytic triad, structural motif, carboxylesterase, structural framework

(50-75 words statement, written for a broader audience, outlining the importance and/or the impact of the work presented in the manuscript)

For the acid-base-nucleophile catalytic machinery of alpha/beta-hydrolases, we show two conserved structural motifs – termed catalytic zones – that serve to properly position the catalytic machinery and coordinate function. Understanding functional role of these structural motifs can contribute to engineering this large family of enzymes for novel applications and

industrial processes.

ABBREVIATIONS

ABH – structural family of alpha/beta-hydrolases

Sm_{nuc-2} – small residue located two positions before the catalytic nucleophile residue

X_{nuc-1} – residue located adjacent (N-terminal) to the catalytic nucleophile residue

Nuc – catalytic nucleophile residue

X_{oxyI} – residues of the oxyanion hole, first position

X_{oxyII} – residues of the oxyanion hole, second position

PDB – Protein Data Bank

SCOP database – Structural Classification of Proteins database

EstFa_R – recombinant carboxylesterase from *Ferroplasma acidiphilum*

X_{ozI} – residue at position "oxyanion zone I"

X_{ozII} – residue at position "oxyanion zone II"

INTRODUCTION

The alpha/beta-hydrolases (ABH) are a structural family of enzymes that share a common fold structure and a conserved catalytic mechanism, while differing in their degree of sequence similarity and function.[1-4] The ABH enzymes catalyze reactions on different substrates, but some ABH enzymes can also act as transporters, receptors, or have other ancillary functions.[5] The ABH fold enzymes are structurally stable, and thus have been extensively used in protein engineering.[6,7]

The ABH family members maintain a half-barrel structure shaped by eight, mostly parallel β-strands (the strand β2 is antiparallel) and six α-helices (αA-αF) flanking the β-sheet. The β-strands and the α-helices are organized in an α-turn-β supersecondary structure geometry that starts with the αA-helix followed by strand β4. The catalytic machinery of all ABH enzymes is similar, with the catalytic triad consisting of an acid: aspartic or glutamic acid; a base: histidine; and a nucleophile: serine, cysteine or aspartate. This catalytic triad is found at conserved locations within the ABH fold (Fig. 1). In particular, the catalytic acid residue is located on a reverse turn after strand β7 or on a tight turn after strand β6; and the catalytic base is a conserved histidine located on the loop following strand β8.[1]

The catalytic nucleophile is the third key player of the chargerelay system that enables the covalent catalysis by nucleophilic attack. The nucleophile is located at the apex of a sharp turn after strand β5, known as the "nucleophile elbow", a conserved geometry identified by the sequence motif Sm_{nuc-2}–X_{nuc-1}–Nuc–X_{nuc+1}–Sm_{nuc+2} (Sm, small residue; X, any residue; Nuc, nucleophile). The dipeptide Sm_{nuc-2}–X_{nuc-1} is located at the end of strand β5 prior to the nucleophile and followed by the dipeptide X_{nuc+1}–Sm_{nuc+2} located at the N-terminal end of the αChelix.

The helical axis of the αC-helix forms an acute angle with respect to strand β5 and the helical N-terminus is placed close to the C-terminus of strand β3; this architecture leads to the arrangement of the "oxyanion hole" that stabilizes the tetrahedral intermediate during catalysis.[1] The oxyanion hole is situated adjacent to the nucleophile and is mainly shaped by the main-chain nitrogen atoms of two residues that we term here oxyanion I (X_{oxyI}) and oxyanion II (X_{oxyII}). Residue X_{oxyI} is always located after the nucleophile, at position X_{nuc+1} of the nucleophile elbow.[1] Residue X_{oxyII} is located after the start of $loop_{\beta3\text{-}\alpha A}$, from which the main-chain nitrogen atom helps form the oxyanion hole in the majority of ABH enzymes. Sometimes, residue X_{oxyII} is located at an alternate position and the oxyanion hole is formed by a side chain. For example, residue X_{oxyII} can be a tyrosine at the start of $loop_{\beta3\text{-}\alpha A}$,[8,9] an aspartic acid after the start of $loop_{\beta3\text{-}\alpha A}$,[10] or an arginine emerging from an α-helix that is located opposite the active site[11]. For some ABH enzymes the oxyanion hole is formed by the main-chain amide NH group from three residues.[12-14]

The nucleophile and the residues forming the oxyanion hole are essential for the enzymatic activity of ABH enzymes. Site-directed mutagenesis of the nucleophile and residues forming the oxyanion hole results in the loss of enzymatic function, usually with little change in the local arrangement of the catalytic site and without interfering with the binding of the ligand;[8,14-17] however, it is possible that the mutation of residues of the oxyanion hole radically distorts the shape of $loop_{\beta3\text{-}\alpha A}$ and hence inactivate the enzyme.[18]

In addition to the residues of the charge-relay network, other residues around the active site have been shown to be essential for enzymatic activity. For example, Kiss *et al.*[19] have shown experimentally that a residue located in the vicinity of the oxyanion hole affects the activity of the enzyme acylaminoacyl-peptidase from *Aeropyrum pernix*. Specifically, the site-directed mutation of His367 to alanine, located two residue-positions before Gly369 of the

oxyanion II, distorted the shape of the oxyanion hole and consequently, peptidase activity was lost.

In our previous work, we studied the structure surrounding the catalytic acid and identified conserved elements that comprise the catalytic acid zone in ABH enzymes.[20] We have shown that the catalytic acid zone coordinates the position of the catalytic histidine loop relative to the catalytic acid and histidine residues. We have also shown that the catalytic acid zone is sometimes found at domain-domain interfaces and residues of the acid zone are directly involved in the formation of multi-subunit complexes and in protein-protein interactions.

Both the experimental study from Kiss *et al.*[19] and our work on the catalytic acid zone[20] suggested that there were additional elements important for the enzymatic activity in the structure around the active site. This motivated us to focus our attention on the architecture around the nucleophile and the oxyanion hole and describe their structural frameworks. In particular, we examined structures for the presence of conserved elements around each key part of the catalytic machinery of the charge-relay system. Herein, we report the occurrence of conserved substructures, which we refer to as the nucleophile zone and oxyanion zone, and discuss their role as supporting structural scaffolds for the enzymatic activity of ABH enzymes.

RESULTS

From the structural point of view, ABH-fold enzymes maintain a conserved catalytic mechanism, which consists of an acid-base-nucleophile triad and the oxyanion hole. Here, we have extensively studied all residues that directly surround key elements of the catalytic machinery; specifically, those residues adjacent to the nucleophile and the oxyanion hole,

aiming to understand their influence on the activity of the ABH-fold enzymes. We compared 41 structures, obtained from the Protein Data Bank[21] (PDB), representing each of the 40 ABH enzyme families described in the SCOP database[22].

This comprehensive analysis revealed structural conservation in three regions that surround the catalytic site: the nucleophile zone near the catalytic nucleophile and oxyanion I residue X_{oxyI}; the oxyanion zone near oxyanion II residue X_{oxyII}; and a cluster of aromatic residues that often surrounds both zones.

The nucleophile zone

The nucleophile zone is a conserved planar structural organization that is found repeatedly in the ABH enzymes (Table 1). It coordinates two key units of the catalytic mechanism: the catalytic nucleophile and the following residue, which is X_{oxyI} of the oxyanion hole. Structurally, the nucleophile zone is formed by residues that belong to the nucleophile elbow and the $loop_{\beta3\text{-}\alpha A}$, and its hydrogen-bond network consists of two conserved, weak hydrogen bonds that connect residues between the nucleophile elbow and the $loop_{\beta3\text{-}\alpha A}$ [for the definition of weak hydrogen bonds, see Material and Methods].

For example, in the carboxylesterase EstFa_R from *Ferroplasma acidiphilum*[23] (PDB ID: 3WJ2), the shape of the nucleophile zone is defined by residues of the nucleophile elbow Asp155–Ser156–Ala157–Gly158 and by the dipeptide His83–Gly84 at the start of $loop_{\beta3\text{-}\alpha A}$ (Fig. 2A). One of the two connecting weak hydrogen bonds is formed between O/Asp155 at position $X_{nuc\text{-}1}$ and CA/Gly84 that precedes Gly85 at X_{oxyII}. The second weak hydrogen bond is formed between CA/$Gly158_{nuc+2}$ and O/$His83_{oxyII\text{-}2}$; this contact anchors the αC-helix to $loop_{\beta3\text{-}\alpha A}$.

Besides the residues of the nucleophile zone, there are also flanking residues, which in EstFa_R are Tyr82$_{oxyII-3}$, Gly85$_{oxyII}$ and Gly154$_{nuc-2}$. X$_{oxyII-3}$, X$_{oxyII}$ and Sm$_{nuc-2}$ interact with the nucleophile zone and with each other (Fig. 2A), and contribute to the local stability. The distances of these peripheral contacts are detailed in the Supplementary Material (Table S1).

Unlike most ABH families that have the main and supplementary contacts of the nucleophile zone, four structural representatives are exceptions (PDB IDs: 3PUI,[28] 1H2W,[29] 1ORV[30] and 2GZS[11] in Table 1), where the entire contact network of the nucleophile zone is disrupted when residue X$_{oxyII}$ is located at an alternate position in the fold. In these cases, the nucleophile zone is not formed (see Discussion).

The nucleophile zone and its adjacent structural elements that have been described, altogether serve to optimally arrange the oxyanion hole for the 36 of 40 ABH enzyme families. The hydrogen-bonding network of the nucleophile zone, together with the peripheral interactions, fix the conformation of the catalytic nucleophile relative to the residues forming the oxyanion hole; the coordination of the second residue of the oxyanion hole also depends on the oxyanion zone that surrounds loop$_{\beta3-\alpha A}$, as is described in the following section.

The oxyanion zone

Adjacent to the nucleophile zone, there is a conserved, planar structural organization that we call the oxyanion zone. The oxyanion zone stabilizes the shape of loop$_{\beta3-\alpha A}$ and coordinates residue X$_{oxyII}$. With regard to the structure of the active site, the role of the oxyanion zone is complementary to that of the nucleophile zone with which it overlaps through the dipeptide X$_{oxyII-2}$–X$_{oxyII-1}$. Structurally, the oxyanion zone is formed by residues

that belong to strand β3, $loop_{β3\text{-}αA}$ and strand β4, and it is mainly organized around a conserved histidine residue at the start of $loop_{β3\text{-}αA}$; histidine is located two residues before X_{oxyII} and the imidazole side chain faces the interior of the protein in 26 of 40 ABH families (columns 4 and 5, Table 2). The hydrogen-bond network of the oxyanion zone consists of four contacts: two contacts link the imidazole ring with neighboring residues and two contacts are formed among other residues of that zone. In 3 of the 40 ABH families, alternative residues are present having the same conformation as the histidine and the zone is formed; whereas, in 11 of 40 ABH families the oxyanion zone is not present because the hydrogen-bonding network is not properly formed due to short side chains or a side chains having a different conformation from histidine.

In EstFa_R[23], for example, the oxyanion zone is shaped by three elements: i) the tetrapeptide Tyr81–Tyr82–His83–Gly84 at the C-terminus of strand β3 and the start of $loop_{β3\text{-}αA}$; ii) the tripeptide Gly90–Asn91–Ile92 at the C-terminus of $loop_{β3\text{-}αA}$; and iii) the dipeptide Ile113–Glu114 at strand β4 (Fig. 2B). The conserved histidine at the start of $loop_{β3\text{-}αA}$ (His83 in EstFa_R) forms two conserved contacts: a hydrogen bond with the adjacent Gly84 (ND1/His83–O/Gly84), and a weak hydrogen bond with Ile113 at strand β4 (CD2/His83–O/Ile113). The residues Gly84 and Gly90, located at each side of $loop_{β3\text{-}αA}$, are linked by a weak hydrogen bond (O/Gly84–CA/Gly90). Lastly, a hydrogen bond N/Ile92–OE2/Glu114 connects the end of $loop_{β3\text{-}αA}$ with strand β4 at positions that we refer to as "oxyanion zone I" (X_{ozI}; Ile92 in EstFa_R) and "oxyanion zone II" (X_{ozII}; Glu114 in EstFa_R).

In addition to the conserved geometry described above, we frequently observed a conserved structural water molecule positioned near the side chain of the residue at the start of $loop_{β3\text{-}αA}$, and bound by multiple interactions (Fig. 2B). For example, in EstFa_R, the water Wat409

interacts with His83 through a weak hydrogen bond (Wat409–CD2/His83), and with the other residues of the oxyanion zone through standard (Tyr81, His83, Ile113) and weak (Tyr82, Ile92) hydrogen bonds. The interaction distances of the water molecule-based contact network can be found in the Supplementary Material (Table S2).

Besides the conserved histidine at the start of $loop_{\beta3\text{-}\alpha A}$, which is critical for maintaining the shape of the loop, two key residues are highly conserved within the oxyanion zone. Gly84 ($Gly_{oxyII\text{-}1}$) in EstFa_R is located between the conserved histidine and X_{oxyII}. Like histidine, $Gly_{oxyII\text{-}1}$ is a structural element shared between the nucleophile and oxyanion zones present in 30 of the 40 ABH families (Table 1). The second conserved residue of the oxyanion zone is an acidic amino acid – aspartate in 13 families and glutamate in 3 families – located at strand β4 at position X_{ozII} in 16 of the 29 ABH families having the oxyanion zone (Table 2). Alternatively, position X_{ozII} can be occupied by an amide – asparagine in 6 families and glutamine in 1 family and they play a similar role in 7 of the 29 ABH families. In contrast to X_{ozII}, the type of amino acid that occupies position X_{ozI} is variable since it is the main-chain nitrogen atom that participates in the conserved interaction between residues X_{ozI} and X_{ozII} (column N/X_{ozI}–OD/X_{ozII} in Table 2); this interaction likely offers additional stabilization of the geometry of $loop_{\beta3\text{-}\alpha A}$.

Thus, the oxyanion zone is formed by seven basic elements: i) a tetrapeptide at the end of strand β3 and the start of $loop_{\beta3\text{-}\alpha A}$; ii) a tripeptide at the end of $loop_{\beta3\text{-}\alpha A}$; iii) a dipeptide at strand β4; iv) a conserved histidine at position $X_{oxyII\text{-}2}$; v) a conserved glycine at position $X_{oxyII\text{-}1}$; vi) a conserved acidic residue at position X_{ozII}; and vii) a conserved structural water molecule that is positioned at the center of the zone, near the imidazole ring of the conserved histidine at $X_{oxyII\text{-}2}$ (Fig. 2B).

All in all, the oxyanion zone, the nucleophile zone and the previously described catalytic acid zone[20] are three conserved structural organizations of the active site of ABH enzymes that lie on the same plane within the β-sheet and coordinate the residues of the oxyanion hole, the catalytic nucleophile and the catalytic acid residue. The catalytic acid zone interacts with the catalytic histidine loop and coordinates the catalytic histidine next to the catalytic acid residue.[20] The results of that study showed that the conservation of the structure of the active site is extended beyond the plane of the β-sheet. This finding appears to be valid also for the space that surrounds the nucleophile and the oxyanion zones, as described below.

Aromatic residues surround the catalytic nucleophile and residues of the oxyanion hole

The planar structural organization of both the nucleophile zone and the oxyanion zone primarily involve main-chain interactions, and thus, the overall shapes of the zones are based on the protein fold. Indeed, apart from the residues at positions $X_{oxyII-2}$ and X_{ozII}, the side chains of other residues of the nucleophile and oxyanion zones are not involved in the formation of the geometry of the two zones. Within the nucleophile and oxyanion zones, or located nearby, there are, however, conserved residues whose side chains interact outside of the plane of the β-sheet.

For example, in the carboxylesterase EstFa_R[23] the residues X_{nuc-1} (Asp155) from the nucleophile zone and $X_{oxyII-4}$ (Tyr81) from the oxyanion zone interact through a weak hydrogen bond (OD1/Asp155–CE2/Tyr81 in Fig. 2C), and their side-chains lie parallel and above these two zones. A nearby residue, Tyr95, which is located after position X_{ozI} (Ile92$_{ozI}$), has a similar side-chain conformation and interacts with both X_{nuc-1} (OD1/Asp155–CE2/Tyr95) and $X_{oxyII-4}$ (OD2/Asp155–OH/Tyr95) through weak and conventional hydrogen

bonds; the aromatic ring of Tyr95 also forms a CH-π interaction with the side-chain CH_2 group of His83, and seems to play a role in stabilizing and directing the side chain of His83 towards the protein interior. Thus, the side chains of all three residues are linked outside the plane of the nucleophile and the oxyanion zones.

As a unit, the hydrogen-bond network among these three residues (Asp155, Tyr81 and Tyr95 in EstFa_R) creates a "roof"-like structural arrangement, positioned over the plane of the nucleophile and oxyanion zones. In order to establish whether this is a common aspect among the ABH families, we have compared their equivalent positions with the structure of carboxylesterase EstFa_R: namely, i) the residue at position X_{nuc-1} (Asp155), ii) the residue at position $X_{oxyII-4}$ (Tyr81) and iii) the residue (Tyr95) that protrudes from the start of the αA-helix or the end of $loop_{\beta3-\alpha A}$ and lies above the β-sheet and within an interaction distance of X_{nuc-1} and $X_{oxyII-4}$ – two conserved positions of the nucleophile and oxyanion zones.

By comparing the active sites of the 40 ABH family representatives, we observed aromatic residues (Trp, Tyr, Phe, His; Table 3): in 22 of 40 ABH families at position X_{nuc-1}, in 8 of 40 ABH families at $X_{oxyII-4}$, and in 27 of 40 ABH families at the position equivalent to Tyr95 in EstFa_R. This latter position is occupied by tryptophan in 14 of the 40 ABH families and its side chain has a characteristic perpendicular conformation over the β-sheet (column W/Y/F in Table 3). The amino acid type at position X_{nuc-1} is more variable than for the other two locations: it is the only position that is occupied by a histidine residue in addition to tryptophan, tyrosine and phenylalanine (column X_{nuc-1} in Table 3). Residue $X_{oxyII-4}$ has a lower occupancy of aromatic residues but larger hydrophobic amino acids (leucine, isoleucine, methionine) are present in the "aromatic cluster" in 24 of the 40 ABH families (column $X_{oxyII-4}$ of Table 3). Together, we observed that the residues from these three positions often cluster – forming a "roof"-like structural arrangement – frequently enriched in

aromatic side chains positioned over the plane of nucleophile and oxyanion zones: 16 ABH families have aromatic pairs and 3 ABH families have aromatic triads, but only 5 ABH families do not have any aromatic residue at one of these three positions.

On the opposite side of the β-sheet we also observed an aromatic residue packing against the imidazole ring of the conserved histidine (His83 in EstFa_R) in 13 of 26 ABH families. The aromatic residue ($Tyr115_{ozII\pm1}$ in EstFa_R, in Fig. 2C) is typically located either before or after residue X_{ozII} (interaction distances: see column $SC/X_{ozII\pm1}$–$O/H_{oxyII\text{-}2}$ in Table 3).

In summary, the arrangement of the nucleophile and oxyanion zones and the conserved residues that interact and are located outside of the β-sheet is such that an aromatic cluster can occur above the two planar zones and directly over the side chain of the conserved histidine (or the equivalent residue) of the oxyanion zone; an aromatic residue also tends to occur “below” the imidazole ring, opposite to and symmetrically to the aromatic cluster. The conservation of these residues clearly shows that the active site structure is extended outside the plane of the β-sheet, where main elements of the catalytic mechanism are situated.

DISCUSSION

The nucleophile zone and the oxyanion zone are planar, structural organizations that coordinate the catalytic nucleophile and the two residues that form the oxyanion hole. Both zones are located at the catalytic site in the majority of ABH enzymes, while aromatic residues frequently surround the β-sheet region where the key units of the catalytic machinery are situated. Here, we discuss residues of the zones that have structural and functional meaning for the ABH enzymes.

The nucleophile zone coordinates the catalytic nucleophile and the oxyanion I residue, and aids the optimal arrangement of the oxyanion hole in most ABH enzymes

The nucleophile zone is situated in the core of the active site of ABH enzymes of the overwhelming majority of ABH enzymes, where it directly coordinates the catalytic nucleophile residue and the oxyanion I residue. Besides the coordination of these two key units of the catalytic machinery, the weak hydrogen bonds of the zone (between Asp155 and Gly84; Gly158 and His83) bind the nucleophile elbow and the loop$_{\beta 3\text{-}\alpha A}$. Consequently, it seems likely that the nucleophile zone is required for the optimal arrangement of the oxyanion hole adjacent to the catalytic nucleophile. This optimal arrangement is achieved by placement of the catalytic nucleophile and the oxyanion I residues of the nucleophile elbow close to the oxyanion II residue that follows the highly conserved His-Gly dipeptide on loop$_{\beta 3\text{-}\alpha A}$.

The nucleophile zone is only formed when the oxyanion II residue is located two positions after the start of loop$_{\beta 3\text{-}\alpha A}$. The importance of the nucleophile zone for the optimal geometry of the oxyanion hole is further confirmed in that only four ABH families lack important

features of the nucleophile zone. In these four ABH families the oxyanion II residue is located at an alternate position and employs its side chain to assemble the oxyanion hole. The weak hydrogen bonds of the nucleophile zone are apparently not required for the formation of the oxyanion hole in these cases.

While analyzing structural consequences in the four ABH families that lack the nucleophile zone, we deduced that the residue at the start of $loop_{\beta3\text{-}\alpha A}$ – equivalent to His83 in EstFa_R – is a determinant of the local geometry, and the type of amino acid and the side-chain conformation of the residue at the start of $loop_{\beta3\text{-}\alpha A}$ correlates with the existence of the nucleophile zone in the active site of ABH enzymes. The four exceptions fall into two separate cases. In the first case (PDB IDs: 3PUI,[28] 1H2W,[29] 1ORV[30]), the oxyanion II residue is a tyrosine, which emerges from the start of $loop_{\beta3\text{-}\alpha A}$ and provides its hydroxyl group to form the oxyanion hole. This tyrosine occupies the sequence and mainchain structural position of the conserved histidine (His83 in EstFa_R). In the second case (PDB ID: 2GZS[11]), the residue that occurs at the start of $loop_{\beta3\text{-}\alpha A}$ is an aspartic acid. The aspartic acid at this position helps form the oxyanion hole through a salt bridge with arginine – the oxyanion II residue – which is located on an α-helix opposite to the active site and distant from the β-sheet. Thus, in both cases where the nucleophile zone is not formed, the residue at the start of $loop_{\beta3\text{-}\alpha A}$ is either the oxyanion II residue or a residue that forms a salt bridge with it. However, in the remaining 36 ABH enzymes that have the nucleophile zone, the residue at the start of $loop_{\beta3\text{-}\alpha A}$ is mostly a histidine, which is critical for the formation of the oxyanion hole, as we describe below.

In addition to the two weak hydrogen bonds important for the nucleophile zone, there are semi-conserved, flanking residues (Gly154, Tyr82, Gly85 in EstFa_R) that structurally support the rigidity of the nucleophile zone by interacting with residues of the zone and with

each other. Both the principal and the peripheral contacts of the nucleophile zone are located in the area of the active site of the ABH enzymes and are shown in Fig. 2A. However, their logical division into main and flanking types is based on the role of the nucleophile zone. We consider that the two conserved, weak hydrogen bonds of the nucleophile zone actively coordinate the catalytic nucleophile and oxyanion I residues. In contrast, the other contacts which are shown in Fig. 2A, are situated further away from the catalytic site, they are semi-conserved and seem to play supportive roles for the local supersecondary structure rather than a straightforward role in the coordination of the key units of the catalytic machinery and the formation of the oxyanion hole. Our assumption for the supplementary structural role of the flanking elements is also supported by the high conservation of glycines at positions Sm_{nuc-2} and Sm_{nuc+2}. Because of the lack of steric clashes and the existence of the two extra weak hydrogen bonds that can be formed by the CA atoms of both glycine residues (Fig. 2A), these residues may help enhance the local stability. With the β-sheet contact N/Asp155–O/Tyr82 creating an imaginary line of separation, this sets a boundary between the elements of function (nucleophile zone; Fig. 2A, right) and the flanking elements (Fig. 2A, left).

The oxyanion zone is organized around a conserved histidine residue, which is relatively remote from the active site, but structurally important for the enzymatic function

The oxyanion zone is a conserved geometry of the active site in most ABH enzymes, which is situated right next to the nucleophile zone. It includes seven conserved structural elements, which are connected through an extended hydrogen-bonding network that functions to stabilize elements of the catalytic machinery (Fig. 2B). Among the conserved elements of the oxyanion zone, we consider the conserved residues $His_{oxyII-2}$, $Gly_{oxyII-1}$ and the acidic residue

X_{ozII} to be critical for the stability of the overall shape of $loop_{\beta3\text{-}\alpha A}$, that eventually leads to the optimal arrangement of the oxyanion hole. By being critical for the stability of $loop_{\beta3\text{-}\alpha A}$, we mean that the hydrogen-bonding network of the oxyanion zone is committed to coordinating the oxyanion II residue next to oxyanion I residue and the catalytic nucleophile, where the latter two residues are coordinated by the nucleophile zone, as we discussed earlier. The structures of the nucleophile and oxyanion zones are linked through the shared dipeptide $His_{oxyII\text{-}2}$–$Gly_{oxyII\text{-}1}$ that participates in the contact networks of both zones. The association of the two zones through this connecting segment shows that the nucleophile and oxyanion zones have complementary roles in the overall arrangement of the structure of the catalytic site. However, their clear structural separation results in different levels of structural conservation, where the nucleophile zone appears to be much more conserved than the oxyanion zone, that may lead to diverse catalytic activities among the ABH families.

$His_{oxyII\text{-}2}$ appears to be the central organizing unit of the oxyanion zone. Half of the contacts of the zone are formed around its side chain. The imidazole ring is placed centrally in the area defined by $loop_{\beta3\text{-}\alpha A}$ and interacts with the adjacent residue $X_{oxyII\text{-}1}$ and a neighboring residue from strand β4. Other residues than histidine at the position $X_{oxyII\text{-}2}$ can play a similar structural role when their side-chain points towards the center of $loop_{\beta3\text{-}\alpha A}$. For example, in our dataset, we have found that the hydrogen-bonding network of the oxyanion zone is fully maintained when glutamine (PDB ID: 1ZI9[36]), methionine (PDB ID: 1Q0R[39]) or arginine (PDB ID: 1BS9[53]) appear at the start of $loop_{\beta3\text{-}\alpha A}$ (columns 4 and 5 in Table 2). A proline (PDB IDs: 1JFR,[47] 1TCA,[48] 1JMK[52]) at the same position partially maintains the bonds of the oxyanion zone. Outside the dataset, we observed another variation at this position, in the structure of cholesterol esterase from *Candida cylindracea* (PDB ID: 1LLF[63], fungal lipases family): a phenylalanine occurs in the position of the conserved histidine and maintains all of

the contacts of the oxyanion zone.

The conserved histidine at position $X_{oxyII-2}$ was previously suggested to be part of the oxyanion hole, with the well-conserved dipeptide His-Gly termed the "oxyanion pocket".[64] This suggestion was later withdrawn because the histidine does not directly participate in catalysis.[60,65] Here, we show that the conserved histidine (or the equivalent residue) in the middle of oxyanion zone is indirectly involved in the ideal positioning of oxyanion II to complete the shape of the oxyanion hole, due to the hydrogen bond with the adjacent residue that appears to stabilize the position of the oxyanion II backbone amide.

The first four residues of $loop_{\beta3-\alpha A}$ – the conserved histidine being the first residue – has been previously mentioned as a possible sequence identifier of different ABH families.[66-68] The four-residue loop segment is suggested to be involved in substrate binding[69] and in the stabilization of the position of the putative hydrolytic water molecule[66]. This indicates that the area around the oxyanion zone is actively involved in several aspects of catalytic activity, as mentioned above.

Multiple experimental studies confirm the importance of the conserved histidine near the oxyanion hole, and thus, the oxyanion zone in general. Site-directed mutagenesis of histidine to alanine resulted in the complete loss of enzymatic function.[19,64,70,71] It is suggested that this mutation is responsible for the distortion of the shape of $loop_{\beta3-\alpha A}$ and thus, the oxyanion hole fails to be formed.[19,64,71] The mutation of histidine to leucine and asparagine had a similar adverse effect on enzymatic activity.[72] Experimental studies have also been applied to those residues that have a similar side-chain conformation as the conserved histidine. The mutation of glutamine to histidine improved the catalytic activity[73], while the mutation of arginine to alanine and lysine again resulted in a loss of function[74]. Notably, site-directed mutagenesis of

the residue $X_{oxyII-1}$ that follows the conserved histidine also leads to a loss of activity, e.g. mutation of the well-conserved glycine[12,32] or aspartate at this position[36]. Taken together, these mutagenesis studies show that the position $X_{oxyII-2}$ should preferably be occupied by a histidine, since in many structures, mutating histidine leads to loss of function as does the adjacent residue; in contrast, mutating other residues to histidine can improve the enzymatic activity.

Thus, for the oxyanion zone, the residue at position $X_{oxyII-2}$ (His83 in EstFa_R) is critical for the geometry of $loop_{\beta3-\alpha A}$. In those ABH families that do not have the oxyanion zone, the type of amino acid at that position can be part of large-scale variations of the structure of the active site that serve the enzymatic function,[26,31,62] or its specific occurrence can contribute to the substrate binding[30], the regulation of pH optimum[75] or the formation of the oxyanion hole[8,9,11,76]. Nonetheless, regardless the amino acid variation, the residue at position $X_{oxyII-2}$ also affects the enzymatic activity in the ABH enzymes without the oxyanion zone. Indeed, experimental studies have shown that site-directed mutation of the residue at the start of $loop_{\beta3-\alpha A}$ can decrease[9] or eliminate the enzymatic activity entirely.[8,76]

Residue X_{ozII} enhances the stability of the oxyanion zone

The acidic side chain of X_{ozII} faces the exterior of the protein, opposite to the conserved histidine that points towards the core of the protein. In many cases, residue X_{ozII} is an aspartic acid, and its side-chain oxygen atom forms a well-conserved hydrogen bond with the amide of residue X_{ozI} located near the end of $loop_{\beta3-\alpha A}$. This contact seems to stabilize the shape of $loop_{\beta3-\alpha A}$ in the ABH enzymes with the oxyanion zone. Although position X_{ozII} is located distantly from the catalytic site, site-directed mutagenesis of the aspartic acid results in

detrimental loss of activity, which suggests that this amino acid plays an important role in maintaining local folding and stability of the enzyme structure.[70] A second mutational study showed that this conserved aspartic acid is actively involved in the formation of the protein-protein complex in the case of acyl esterase Aes from *Escherichia coli*.[77]

The aromatic residues around the nucleophile and oxyanion zones can affect the stability of the catalytic site

Besides the conservation of the nucleophile and oxyanion zones, we showed that aromatic residues regularly occur around both structural organizations. Particularly, we showed that aromatic residues often cluster over the nucleophile and oxyanion zones (Fig. 2C), while aromatic pairs are formed by the conserved histidine and a residue whose side chain lies opposite to the aromatic cluster with respect to the β-sheet.

The conservation of aromatic residues in the periphery of the catalytic site can prove beneficial for the local structural stability. Several studies have previously analyzed the importance of aromatic-aromatic interactions around the active site and shown their influence on the enzymatic[71,78-81,82] and the thermal[83] stability of the protein, on the lining of the binding pocket[29] and on the non-local stabilization of the protein tertiary structure[84].

We observed aromatic residues at three specific positions forming the aromatic cluster: the residue X_{nuc-1} at the C-terminus of strand β5, the residue $X_{oxyII-4}$ at the C-terminus of strand β3 and a residue at the end of $loop_{\beta3-\alpha A}$ or the start of αA-helix. Our results indicate that the residue X_{nuc-1}, located before the nucleophile, is not a randomly occurring amino acid, but it is very often an aromatic residue. Experimental studies have shown that site-directed mutation of residue X_{nuc-1} results in the reduction[85] or the elimination of the catalytic

efficiency of the enzyme[86]. It is suggested that the residue X_{nuc-1} plays a structural role in the correct geometry of the catalytic site[85-87]. Residue $X_{oxyII-4}$ appears to be the most conserved of the aromatic cluster residues, as it is mostly a hydrophobic or else an aromatic residue.

The third residue of the aromatic cluster does not have a defined sequence position, because its position heavily depends on the length of $loop_{\beta3-\alpha A}$. Nonetheless, it is always linked with the residues X_{nuc-1} and $X_{oxyII-4}$. In the majority of ABH families, this position is occupied by an aromatic amino acid, quite often by a tryptophan residue, whose side chain has a characteristic, nearly perpendicular conformation over the β-sheet, with the indole ring landing opposite to residue Sm_{nuc-2} of the nucleophile elbow. Interestingly enough, this structural phenomenon has been previously observed[88] in a structurally similar part of the *Clostridium MP* flavodoxin[89] structure, which also belongs to α/β fold class. Experimental studies, that refer to the third residue of the aromatic cluster, have showed that its site-directed mutation can lead to decreased catalytic activity[85,90,91] and affect the thermostability of the enzyme[91].

Finally, the location of the aromatic cluster suggests that it would be involved in the coordination of the catalytic histidine loop. Two separate studies have mentioned that residue X_{nuc-1}[87] and the third residue of the cluster[71] form contacts with aromatic residues that occur at the loop fragment after the catalytic histidine, and thus participate in extended aromatic interactions near the catalytic histidine. This fact is rather interesting, because in our previous study[20] we have shown that the catalytic acid zone residues coordinate the loop fragment before the catalytic histidine.

CONCLUSIONS

In this study, we analyzed the structure around the nucleophile elbow and the oxyanion hole in representative structures from 40 ABH families. We have shown that the nucleophile zone, the oxyanion zone and the aromatic cluster comprise a three-dimensional structural organization that shapes the active site of ABH enzymes and plays an important role in the enzymatic function by structurally stabilizing the catalytic nucleophile and the residues of the oxyanion hole. This finding is complementary to the result from our previous study, where we showed that there is a conserved structural organization that coordinates the catalytic acid residue and links it with the catalytic histidine. Overall, we were able to confirm that there is a conserved supporting architecture around each key unit of the catalytic machinery.

Undeniably, position $X_{oxyII-2}$ at the start of $loop_{\beta3-\alpha A}$ is paramount for the local structure of the catalytic site and the activity of ABH enzymes, and is shared by both the nucleophile and the oxyanion zones. Regardless the type of amino acid, position $X_{oxyII-2}$ is typically occupied by a residue that is suitable for the optimal geometry of the oxyanion hole, either indirectly by stabilizing the arrangement of the oxyanion hole, or directly by providing a side-chain group to form the oxyanion hole. In addition to that, a conserved acidic residue at position X_{ozII} of the oxyanion zone proves to be essential for the formation of the oxyanion hole, although it is located remote from the active site.

Finally, we showed that aromatic residues tend to appear around the nucleophile and the oxyanion zones in most ABH families, often forming an aromatic cluster over and an aromatic pair below the plane of the nucleophile and oxyanion zones. Aromatic residues are shown to enhance the stability of the active site, and thus, we consider that they also affect the stability of the ABH enzymes. The aromatic cluster, which is located close to the active site, connects the nucleophile and the oxyanion zones through interactions between residues that belong to both zones. In addition, there is evidence that the aromatic cluster can be

involved in the coordination of the catalytic histidine loop. Further experimental studies on the conserved aromatic residues may provide valuable insight about their role in the enzymatic activity.

MATERIAL AND METHODS

We have used the Structural Classification Of Proteins (SCOP) database[22] to choose representative structures for our dataset. There are 41 ABH families according to SCOP classification, which contain 128 different entries of proteins with the ABH fold. The single entry from the TTHA1544-like family (SCOP Family #39) is not included in our dataset because it is not a hydrolase.[92] From the 127 entries, we have selected 40 representative structures ensuring that they represent each listed ABH family, on the condition that they are mainly unligated and have the highest resolution. In total, we have formed a dataset, which contains 40 structures plus the carboxylestarase EstFa_R from the study of Ohara *et al.*[23] that we use as the reference structure throughout this study. We have obtained all protein structures from Protein Data Bank (PDB).[21]

The results of this study are presented in three tables. Each table has a different number of entries, depending on the subset of representative structures that have each conserved structural motif. Specifically, Table 1 describes 40 representative structures from 40 ABH enzyme families (rows 1-40) plus the carboxylesterase EstFa_R (row A), with the exception of the single-entry protein from the TTHA1544-like family. In Table 1, we show not only the 36 ABH structures and EstFa_R that maintain the contact network of the nucleophile zone, but also the four ABH structures that lack all interactions of the nucleophile zone. Table 2 only includes those structures that maintain the contact network of the oxyanion zone and

specifically, 29 representative structures from the 40 ABH enzyme families (rows 1-29) plus the carboxylesterase EstFa_R (row A). The protein from SCOP Family TTHA1544-like is not taken into account here as well. Table 3 includes all 40 representative structures (rows 1-40) from the 40 ABH families (except the TTHA1544-like family) and the carboxylesterase EstFa_R (row A).

For all structural analysis, such as hydrogen bonds, hydrophobic or other types of weak interactions, we have used the BIOVIA (Accelrys) Discovery Studio[93] (http:// accelrys.com / products / collaborativescience / bioviadiscoverystudio/). We have identified the weak hydrogen bonds from CHO contacts in structures with at least 2.0 Å resolution based on geometrical criteria given in (94) and with distances C...O ≤ 4.1 Å and H...O ≤ 3.0 Å, and the weak hydrogen bonds from CHπ contacts with distance between 3.4 Å and 6Å[82].

Lastly, for the visualization and analysis of the structural data, we have used Discovery Studio[93] and Bodil[95]. Figures are drawn with MolScript.[96]

SUPPLEMENTARY MATERIAL

The Supplementary Material includes two tables: Table S1 and Table S2. Table S1 reports the peripheral contacts of the nucleophile zone for the 40 ABH structures of the dataset. These contacts are shown in Figure 2A for the carboxylesterase EstFa_R. Table S2 reports the contacts that are formed around the water molecule in the 28 ABH families that maintain a water molecule in the oxyanion zone, as shown in Figure 2B for EstFa_R.

ACKNOWLEDGEMENTS

This work is supported by grants from the Otto A. Malm Foundation, Sigrid Juselius Foundation and Joe, Pentti and Tor Borg Memorial Fund. We thank the Biocenter Finland Bioinformatics Network (Dr. Jukka Lehtonen) and CSC IT Center for Science for computational support for the project. P.S.D. is funded by the Åbo Akademi Doctoral Network of Informational and Structural Biology.

REFERENCES

1. Nardini M, Dijkstra BW (1999) α/β-Hydrolase fold enzymes: the family keeps growing. Curr Opin Struct Biol 9:732-737.

2. Mindrebo JT, Nartey CM, Seto Y, Burkart MK, Noel JP (2016) Unveiling the functional diversity of the alpha/beta-hydrolase superfamily in the plant kingdom. Curr Opin Struct Biol 41:233-246.

3. Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:D423-D429.

4. Rauwerdink A, Kazlauskas RJ (2015) How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of α/β-hydrolase fold enzymes. ACS Catalysis 5:6153-6176.

5. Marchot P, Chatonnet A (2012) Enzymatic activity and protein interactions in alpha/beta-hydrolase fold proteins: moonlighting versus promiscuity. Protein Pept Lett 19:132-143.

6. Jochens H, Hesseler M, Stiba K, Padhi SK, Kazlauskas RJ, Bornscheuer UT (2011) Protein engineering of α/β-hydrolase fold enzymes. ChemBioChem 12:1508-1517.

7. Suplatov DA, Besenmatter W, Svedas VK, Svendsen A (2012) Bioinformatic analysis of alpha/beta-hydrolase fold enzymes reveals subfamily-specific positions responsible for discrimination of amidase and lipase activities. Protein Eng Des Sel 25:689-697.

8. Turner JM, Larsen NA, Basran A, Barbas CF, Bruce NC, Wilson IA, Lerner RA (2002) Biochemical characterization and structural analysis of a highly proficient cocaine esterase. Biochemistry 41:12297-12307.

9. Szeltner Z, Renner V, Polgár L (2000) Substrate and pH-dependent contribution of oxyanion binding site to the catalysis of prolyl oligopeptidase, a paradigm of the serine oligopeptidase family. Protein Sci 9:353-360.

10. Ericsson DJ, Kasrayan A, Johansson P, Bergfors T, Sandström AG, Bäckvall JE, Mowbray SL (2008) X-ray structure of *Candida antarctica* lipase A shows a novel lid structure and a likely mode of inter-facial activation. J Mol Biol 376:109-119.

11. Larsen NA, Lin H, Wei R, Fischbach MA, Walsh CT (2006) Structural characterization of enterobactin hydrolase IroE. Biochemistry 45:10184-10190.

12. Ordentlich A, Barak D, Kronman C, Ariel N, Segall Y, Velan B, Shafferman A (1998) Functional characteristics of the oxyanion hole in human acetylcholinesterase. J Biol Chem 273:19509-19517.

13. Zhang Y, Kua J, McCammon JA (2002) Role of the catalytic triad and oxyanion hole in acetylcholinesterase catalysis: an ab initio QM/MM study. J Am Chem Soc 124:10572-10577.

14. Mandrich L, Menchise V, Alterio V, De Simone G, Pedone C, Rossi M, Manco G (2008) Functional and structural features of the oxyanion hole in a thermophilic esterase from *Alicyclobacillus acidocaldarius*. Proteins 71:1721-1731.

15. Lai L, Xu Z, Zhou J, Lee KD, Amidon GL (2008) Molecular basis of prodrug activation by human valacyclovirase, an alpha-amino acid ester hydrolase. J Biol Chem 283:9318-9327.

16. Charavgi MD, Dimarogona M, Topakas E, Christakopoulos P, Chrysina ED (2013) The structure of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into its role as a potential biocatalyst. Acta Crystallogr D Biol Crystallogr 69:63-73.

17. Huang J, Huo YY, Ji R, Kuang S, Ji C, Xu XW, Li J (2016) Structural insights of a

hormone sensitive lipase homologue Est22. Sci Rep 6:28550.

18. Goettig P, Brandstetter H, Groll M, Göhring W, Konarev PV, Svergun DI, Huber R, Kim JS (2005) X-ray snapshots of peptide processing in mutants of tricorninteracting factor F1 from *Thermoplasma acidophilum*. J Biol Chem 280:33387-33396.

19. Kiss AL, Palló A, NáraySzabó G, Harmat V, Polgár L (2008) Structural and kinetic contributions of the oxyanion binding site to the catalytic activity of acylaminoacyl peptidase. J Struct Biol 162:312-323.

20. Dimitriou PS, Denesyuk A, Takahashi S, Yamashita S, Johnson MS, Nakayama T, Denessiouk K (2017) Alpha/beta-hydrolases: a unique structural motif coordinates catalytic acid residue in 40 protein fold families. Proteins 2017;85:1845-1855.

21. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235-242.

22. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536-540.

23. Ohara K, Unno H, Hosoya M, Fujino N, Hirooka K, Takahashi S, Yamasita S, Kusunoki M, Nakayama T (2014) Structural insights into the low pH adaptation of a unique carboxylesterase from *Ferroplasma*: altering the pH optima of two carboxylesterases. J Biol Chem 289:24499-24510.

24. Spiller B, Gershenson A, Arnold FH, Stevens RC (1999) A structural view of evolutionary divergence. Proc Natl Acad Sci U S A 96:12305-12310.

25. Zhu X, Larsen NA, Basran A, Bruce NC, Wilson IA (2003) Observation of an arsenic adduct in an acetyl esterase crystal structure. J Biol Chem 278:2008-2014.

26. Ronning DR, Klabunde T, Besra GS, Vissa VD, Belisle JT, Sacchettini JC (2000) Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. Nat Struct Biol 7:141-146.

27. Murayama K, Shirouzu M, Terada T, Kuramitsu S, Yokoyama S (2005) Crystal structure of TT1662 from *Thermus thermophilus* HB8: a member of the alpha/beta hydrolase fold enzymes. Proteins 58:982-984.

28. Narasimhan D, Collins GT, Nance MR, Nichols J, Edwald E, Chan J, Ko MC, Woods JH, Tesmer JJ, Sunahara RK (2011) Subunit stabilization and polyethylene glycolation of cocaine esterase improves in vivo residence time. Mol Pharmacol 80:1056-1065.

29. Szeltner Z,Rea D, Renner V, Fulop V, Polgar L (2002) Electrostatic effects and binding determinants in the catalysis of prolyl oligopeptidase. Site specific mutagenesis at the oxyanion binding site. J Biol Chem 277:42613-42622.

30. Engel M, Hoffmann T, Wagner L, Wermann M, Heiser U, Kiefersauer R, Huber R, Bode W, Demuth HU, Brandstetter H (2003) The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. Proc Natl Acad Sci U S A 100:5063-5068.

31. Liao DI, Breddam K, Sweet RM, Bullock T, Remington SJ (1992) Refined atomic model of wheat serine carboxypeptidase II at 2.2Å resolution. Biochemistry 31:9796-9812.

32. Roussel A, Canaan S, Egloff MP, Rivière M, Dupuis L, Verger R, Cambillau C (1999) Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. J Biol Chem 274:16995-17002.

33. Goettig P, Groll M, Kim JS, Huber R, Brandstetter H (2002) Structures of the tricorninteracting aminopeptidase F1 with different ligands explain its catalytic mechanism. EMBO J 21:5343-5352.

34. Zhang R, Koroleva O, Collert F, Joachimiak A, Midwest Center for Structural Genomics. 1.5 Å crystal structure of the cephalosporin C deacetylase.

35. Oakley AJ, Klvana M, Otyepka M, Nagata Y, Wilce MC, Damborsky J (2004) Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 A resolution: dynamics of catalytic residues. Biochemistry 43:870-878.

36. Kim HK, Liu JW, Carr PD, Ollis DL (2005) Following directed evolution with crystallography: structural changes observed in changing the substrate specificity of dienelactone hydrolase. Acta Crystallogr D Biol Crystallogr 61:920-931.

37. Horsman GP, Ke J, Dai S, Seah SY, Bolin JT, Eltis LD (2006) Kinetic and structural insight into the mechanism of BphD, a C-C bond hydrolase from the biphenyl degradation pathway. Biochemistry 45:11071-11086.

38. Agarwal V, Lin S, Lukk T, Nair SK, Cronan JE (2012) Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proc Natl Acad Sci U S A 109:17406-17411.

39. Jansson A, Niemi J, Mantsala P, Schneider G (2003) Crystal structure of aclacinomycin methylesterase with bound product analogues: implications for anthracycline recognition and mechanism. J Biol Chem 278:39006-39013.

40. Seetharaman J, Lew S, Wang D, Kohan E, Patel D, Whitehead T, Fleishman S, Ciccosanti C, Xiao R, Everett JK, Acton TB, Montelione GT, Tong L, Hunt JF, Northeast Structural Genomics Consortium. Crystal Structure of Engineered Protein. Northeast Structural Genomics Consortium Target OR94.

41. Zou J, Hallberg BM, Bergfors T, Oesch F, Arand M, Mowbray SL, Jones TA (2000) Structure of *Aspergillus niger* epoxide hydrolase at 1.8 Å resolution: implications for the structure and function of the mammalian microsomal class of epoxide hydrolases. Structure

8:111-122.

42. Hofmann B, Tolzer S, Pelletier I, Altenbuchner J, van Pee KH, Hecht HJ (1998) Structural investigation of the cofactorfree chloroperoxidases. J Mol Biol 279:889-900.

43. Bellizzi III JJ, Widom J, Kemp C, Lu JY, Das AK, Hofmann SL, Clardy J (2000) The crystal structure of palmitoyl protein thioesterase 1 and the molecular basis of infantile neuronal ceroid lipofuscinosis. Proc Natl Acad Sci U S A 97:4573-4578.

44. Devedjiev Y, Dauter Z, Kuznetsov SR, Jones TL, Derewenda ZS (2000) Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 Å. Structure 8:1137-1146.

45. Padmanabhan B, Kuzuhara T, Adachi N, Horikoshi M (2004) The crystal structure of CCG1/TAF(II)250interacting factor B (CIB). J Biol Chem 279:9615-9624.

46. Bourne PC, Isupov MN, Littlechild JA (2000) The atomic-resolution structure of a novel bacterial esterase. Structure 8:143-151.

47. Wei Y, Swenson L, Castro C, Derewenda U, Minor W, Arai H, Aoki J, Inoue K, ServinGonzalez L, Derewenda ZS (1998) Structure of a microbial homologue of mammalian platelet-activating factor acetylhydrolases: *Streptomyces exfoliatus* lipase at 1.9 Å resolution. Structure 6:511-519.

48. Uppenberg J, Hansen MT, Patkar S, Jones TA (1994) The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica*. Structure 2:293-308.

49. Kawasaki K, Kondo H, Suzuki M, Ohgiya S, Tsuda S (2002) Alternate conformations observed in catalytic serine of *Bacillus subtilis* lipase determined at 1.3 Å resolution. Acta Crystallogr D Biol Crystallogr 58:1168-1174.

50. Roussel A, Yang Y, Ferrato F, Verger R, Cambillau C, Lowe M (1998) Structure and activity of rat pancreatic lipase-related protein 2. J Biol Chem 273:32121-32128.

51. Schmidt A, Gruber K, Kratky C, Lamzin VS (2008) Atomic resolution crystal structures and quantum chemistry meet to reveal subtleties of hydroxynitrile lyase catalysis. J Biol Chem 283:21827-21836.

52. Bruner SD, Weber T, Kohli RM, Schwarzer D, Marahiel MA, Walsh CT, Stubbs MT (2002) Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. Structure 10:301-310.

53. Ghosh D, Erman M, Sawicki M, Lala P, Weeks DR, Li N, Pangborn W, Thiel DJ, Jornvall H, Gutierrez R, Eyzaguirre J (1999) Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase. Acta Crystallogr D Biol Crystallogr 55:779-784.

54. Janda IK, Devedjiev Y, Cooper DR, Chruszcz M, Derewenda U, Gabrys A, Minor W, Joachimiak A, Derewenda ZS (2004) Harvesting the high-hanging fruit: the structure of the YdeN gene product from *Bacillus subtilis* at 1.8 angstroms resolution. Acta Crystallogr D Biol Crystallogr 60:1101-1107.

55. Arndt JW, Schwarzenbacher R, Page R, Abdubek P, Ambing E, Biorac T, Canaves JM, Chiu HJ, Dai X, Deacon AM, DiDonato M, Elsliger MA, Godzik A, Grittini C, Grzechnik SK, Hale J, Hampton E, Han GW, Haugen J, Hornsby M, Klock HE, Koesema E, Kreusch A, Kuhn P, Jaroszewski L, Lesley SA, Levin I, McMullan D, McPhillips TM, Miller MD, Morse A, Moy K, Nigoghossian E, Ouyang J, Peti WS, Quijano K, Reyes R, Sims E, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, von Delft F, Wang X, West B, White A, Wolf G, Xu Q, Zagnitko O, Hodgson KO, Wooley J, Wilson IA (2005) Crystal structure of an alpha/beta serine hydrolase (YDR428C) from *Saccharomyces cerevisiae* at 1.85 Å resolution. Proteins 58:755-758.

56. Bartlam M, Wang G, Yang H, Gao R, Zhao X, Xie G, Cao S, Feng Y, Rao Z (2004)

Crystal structure of an acylpeptide hydrolase/esterase from *Aeropyrum pernix* K1. Structure 12:1481-1488.

57. Legler PM, Kumaran D, Swaminathan S, Studier FW, Millard CB (2008) Structural characterization and reversal of the natural organophosphate resistance of a D-type esterase, *Saccharomyces cerevisiae* S-formyl-glutathione hydrolase. Biochemistry 47:9592-9601.

58. Gorman J, Shapiro L, Burley SK, Midwest Center for Structural Genomics. Structural Genomics target NYSGRCT920 related to A/B-hydrolase fold.

59. Osipiuk J, Xu X, Zheng H, Savchenko A, Edwards A, Joachimiak A, Midwest Center for Structural Genomics. Crystal structure of hypothetical protein Atu1826, a putative alpha/beta hydrolase from *Agrobacterium tumefaciens*.

60. Hisano T, Kasuya K, Tezuka Y, Ishii N, Kobayashi T, Shiraki M, Oroudjev E, Hansma H, Iwata T, Doi Y, Saito T, Miki K (2006) The crystal structure of polyhydroxybutyrate depolymerase from *Penicillium funiculosum* provides insights into the recognition and degradation of biopolyesters. J Mol Biol 356:993-1004.

61. Mirza IA, Nazi I, Korczynska M, Wright GD, Berghuis AM (2005) Crystal structure of homoserine transacetylase from *Haemophilus influenzae* reveals a new family of alpha/beta-hydrolases. Biochemistry 44:15768-15773.

62. Schleberger C, Sachelaru P, Brandsch R, Schulz GE (2007) Structure and action of a CC bond cleaving alpha/beta-hydrolase involved in nicotine degradation. J Mol Biol 367:409-418.

63. Pletnev V, Addlagatta A, Wawrzak Z, Duax W (2003) Threedimensional structure of homodimeric cholesterol esterase-ligand complex at 1.4 Å resolution. Acta Crystallogr D Biol Crystallogr 59:50-56.

64. Braaz R, Handrick R, Jendrossek D (2003) Identification and characterisation of the

catalytic triad of the alkaliphilic thermotolerant PHA depolymerase PhaZ7 of *Paucimonas lemoignei*. FEMS Microbiol Lett 224:107-112.

65. Papageorgiou AC, Hermawan S, Singh CB, Jendrossek D (2008) Structural basis of poly(3-hydroxybutyrate) hydrolysis by PhaZ7 depolymerase from *Paucimonas lemoignei*. J Mol Biol 382:1184-1194.

66. Nardini M, Ridder IS, Rozeboom HJ, Kalk KH, Rink R, Janssen DB, Dijkstra BW (1999) The x-ray structure of epoxide hydrolase from *Agrobacterium* radiobacter AD1. An enzyme to detoxify harmful epoxides. J Biol Chem 274:14579-14586.

67. De Simone G, Galdiero S, Manco G, Lang D, Rossi M, Pedone C (2000) A snapshot of a transition state analogue of a novel thermophilic esterase belonging to the subfamily of mammalian hormone-sensitive lipase. J Mol Biol 303:761-771.

68. Bahl CD, Hvorecny KL, Bridges AA, Ballok AE, Bomberger JM, Cady KC, O'Toole GA, Madden DR (2014) Signature motifs identify an *Acinetobacter* Cif virulence factor with epoxide hydrolase activity. J Biol Chem 289:7460-7469.

69. Medrano FJ, Alonso J, García JL, Romero A, Bode W, GomisRüth FX (1998) Structure of proline iminopeptidase from *Xanthomonas campestris* pv. citri: a prototype for the prolyl oligopeptidase family. EMBO J 17:1-9.

70. Lai KK, Stogios PJ, Vu C, Xu X, Cui H, Molloy S, Savchenko A, Yakunin A, Gonzalez CF (2011) An inserted α/β sub-domain shapes the catalytic pocket of *Lactobacillus johnsonii* cinnamoyl esterase. PLoS One 6:e23269.

71. Tyukhtenko S, Karageorgos I, Rajarshi G, Zvonok N, Pavlopoulos S, Janero DR, Makriyannis A (2016) Specific inter-residue interactions as determinants of human monoacylglycerol lipase catalytic competency: a role for global conformational changes. J Biol Chem 291:2556-2565.

72. Bell PA, Kasper CB (1993) Expression of rat microsomal epoxide hydrolase in *Escherichia coli.* Identification of a histidyl residue essential for catalysis. J Biol Chem 268:14011-14017.

73. Porter JL, Boon PL, Murray TP, Huber T, Collyer CA, Ollis DL (2015) Directed evolution of new and improved enzyme functions using an evolutionary intermediate and multi-directional search. ACS Chem Biol 10:611-621.

74. Bordes F, Barbe S, Escalier P, Mourey L, André I, Marty A, Tranier S (2010) Exploring the conformational states and rearrangements of *Yarrowia lipolytica* lipase. Biophys J 99:2225-2234.

75. Rudenko G, Bonten E, D'Azzo A, Hol WGJ (1995) Three-dimensional structure of the human protective protein: structure of the precursor form suggests a complex activation mechanism. Structure 3:1249-1259.

76. Brim RL, Nance MR, Youngstrom DW, Narasimhan D, Zhan CG, Tesmer JJ, Sunahara RK, Woods JH (2010) A thermally stable form of bacterial cocaine esterase: a potential therapeutic agent for treatment of cocaine abuse. Mol Pharmacol 77:593-600.

77. Schiefner A, Gerber K, Brosig A, Boos W (2014) Structural and mutational analyses of Aes, an inhibitor of MalT in *Escherichia coli*. Proteins 82:268-277.

78. Faulds CB, Molina R, Gonzalez R, Husband F, Juge N, SanzAparicio J, Hermoso JA (2005) Probing the determinants of substrate specificity of a feruloyl esterase, AnFaeA, from *Aspergillus niger*. FEBS J 272:4362-4371.

79. Chourasia M, Sastry GM, Sastry GN (2011) Aromatic-Aromatic Interactions Database, A(2)ID: an analysis of aromatic π-networks in proteins. Int J Biol Macromol 48:540-552.

80. Madhusudan Makwana K, Mahalakshmi R (2015) Implications of aromaticaromatic

interactions: From protein structures to peptide models. Protein Sci 24:1920-1933.

81. Truongvan N, Chung HS, Jang SH, Lee C (2016) Conserved tyrosine 182 residue in hyper-thermophilic esterase EstE1 plays a critical role in stabilizing the active site. Extremophiles 20:187-193.

82. Burley SK, Petsko GA (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science 229:23-28.

83. Kannan N, Vishveshwara S (2000) Aromatic clusters: a determinant of thermal stability of thermophilic proteins. Protein Eng 13:753-761.

84. Lanzarotti E, Biekofsky RR, Estrin DA, Marti MA, Turjanski AG (2011) Aromatic-aromatic interactions in proteins: beyond the dimer. J Chem Inf Model 51:1623-1633.

85. Haruki M, Oohashi Y, Mizuguchi S, Matsuo Y, Morikawa M, Kanaya S (1999) Identification of catalytically essential residues in *Escherichia coli* esterase by site-directed mutagenesis. FEBS Lett 454:262-266.

86. Li C, Li JJ, Montgomery MJ, Wood SP, Bugg TDH (2006) Catalytic role for arginine 188 in the C-C hydrolase catalytic mechanism for *Escherichia coli* MhpC and *Burkholderia xenovorans* LB400 BphD. Biochemistry 45:12470-12479.

87. Horsman GP, Bhowmik S, Seah SY, Kumar P, Bolin JT, Eltis LD (2007) The tautomeric half-reaction of BphD, a C-C bond hydrolase. Kinetic and structural evidence supporting a key role for histidine 265 of the catalytic triad. J Biol Chem 282:19894-19904.

88. Janin J, Chothia C (1980) Packing of alpha-helices onto beta-pleated sheets and the anatomy of alpha/beta proteins. J Mol Biol 143:95-128.

89. Burnett RM, Darling GD, Kendall DS, LeQuesne ME, Mayhew SG, Smith WW, Ludwig

ML (1974) The structure of the oxidized form of clostridial flavodoxin at 1.9A resolution. J Biol Chem 249:4383-4392.

90. McKary MG, Abendroth J, Edwards TE, Johnson RJ (2016) Structural basis for the strict substrate selectivity of the mycobacterial hydrolase LipW. Biochemistry 55: 7099-7111.

91. Zhang XF, Yang GY, Zhang Y, Xie Y, Withers SG, Feng Y (2016) A general and efficient strategy for generating the stable enzymes. Sci Rep 6:33797.

92. Xie Y, Takemoto C, Kishishita S, UchikuboKamo T, Murayama K, Chen L, Liu ZJ, Wang BC, Manzoku M, Ebihara A, Kuramitsu S, Shirouzu M, Yokoyama S (2007) Structure of the minimized alpha/beta-hydrolase fold protein from *Thermus thermophilus* HB8. Acta Crystallogr F 63:993-997.

93. Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes; 2016.

94. Derewenda ZS, Lee L, Derewenda U (1995) The occurrence of CH...O hydrogen bonds in proteins. J Mol Biol 252:248-262.

95. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Björklund D, Iftikhar Z, Huhtala M, Repo S, Jussila A, Jaakkola J, Pentikäinen O, Nyrönen T, Salminen T, Gyllenberg M, Johnson M (2004) BODIL: a molecular modeling environment for structure-function analysis and drug design. J Comput Aided Mol Des 18:401-419.

96. Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Cryst 24:946-950.

FIGURE LEGENDS

Figure 1: The active site of the ABH fold enzymes. Elements of the catalytic machinery are arranged as follows: the catalytic acid (“Catalytic Acid”) is located either after strand β7 (Group A) or strand β6 (Group B); the catalytic base (“Catalytic Base”); and the catalytic nucleophile (“Catalytic Nucleophile”). The oxyanion hole is usually formed by two residues: the oxyanion I residue (“XoxyI”), adjacent to the catalytic nucleophile, and the oxyanion II residue (“XoxyII”) in the loop following strand β3. The “Acid zone”, the “Nucleophile zone” and the “Oxyanion zone” designate the positions of three clusters of conserved residues that are linked by hydrogen bonds around the catalytic acid, the catalytic nucleophile and the residues of the oxyanion hole, respectively; the boundaries of each zone are marked by dashed lines. The positions of conserved elements of the zones are shown as: “Water” for structural water molecules; “His” for the conserved histidine of the Oxyanion zone; and “W/Y/F” for the aromatic residues tryptophan, tyrosine or phenylalanine from the aromatic cluster. The positions “oxyanion zone I” (“XozI”) and “oxyanion zone II” (“XozII”) are part of the Oxyanion zone.

Figure 2: The nucleophile zone (A), the oxyanion zone (B) and the aromatic cluster (C) in the active site of ABH enzymes.

(A) The nucleophile zone in the structure of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The zone is shaped by the residues Asp155 Ser156 Ala157 Gly158 of the nucleophile elbow and His83 Gly84 at the start of loopβ3 αA, and involves

two, weak hydrogen bonds, O/Asp155 CA/Gly84 and CA/Gly158 O/His83. The residues Tyr82, Gly85 and Gly154 occupy semi-conserved, peripheral positions and form four, supplementary hydrogen bonds to the nucleophile zone. “Snuc” designates the catalytic nucleophile residue, which is a serine in the majority of ABH enzymes; Main-chain nitrogen atoms at the oxyanion I (“N/XoxyI” from Ala157 in EstFa_R) and oxyanion II (“N/XoxyII” from Gly85 in EstFa_R) positions form the oxyanion hole.

(B) The oxyanion zone in the structure of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The zone is formed by: the residues Tyr81 Tyr82 His83 Gly84 at the end of strand β3 and the start of loopβ3 αA; the Gly90 Asn91 Ile92 at the end of loopβ3 αA; and Ile113 Glu114 from strand β4. All amino acids except XoxyII, and all contacts except the structural β sheet hydrogen bond between Tyr81 and Ile113, belong to the oxyanion zone.“N/XoxyII” designates the main chain nitrogen atom of oxyanion II residue Gly85 in EstFa_R; and “Wat409” designates the conserved structural water molecule forming multiple hydrogen bonds in the oxyanion zone of the ABH enzymes. A hydrogen bond is formed between the main chain nitrogen atom of the residue at “oxyanion zone I” (“N/XozI”, Ile92 in EstFa_R) and the residue at “oxyanion zone II” (“OD,OE/XozII”, side chain oxygen atom of Glu114 in EstFa_R).

(C) Conserved residues at the aromatic cluster near the active site of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The residues Tyr81, Tyr95 and

Asp155 are located at positions where aromatic amino acids (as well as hydrophobic residues) are frequently found in all ABH enzymes; the side chain interactions of the residues at these positions form a "roof"-like structural arrangement over the β sheet. Conservation of an aromatic residue is also observed before or after the residue at the position "oxyanion zone II" ("XozII", Glu114 in EstFa_R); in EstFa_R, the side chain hydroxyl of Tyr115 interacts with the main chain oxygen of His83, and the two residues form an aromatic pair below the β sheet."Snuc" and "XoxyII" designate the catalytic nucleophile and oxyanion II residue of the catalytic machinery, and "Wat409" and "XozI" designate the conserved structural water molecule and the residue at position "oxyanion zone I" of the Oxyanion zone.

Table 1. Geometrical parameters of the conserved bonds that shape the nucleophile zone in 40 ABH fold families.

N	Family	PDB ID Resolution	S_{nuc}	O/X_{nuc-1} – CA/$G_{oxyII-1}$	CA/G_{nuc+2} – O/$H_{oxyII-2}$	N/X_{oxyI} (N/X_{nuc+1})	N/X_{oxyII}
A	Carboxylesterase	3WJ2_A[24] 1.61 Å	S156	O/D155-CA/G84 3.1 (2.7)	CA/G158-O/H83 3.3 (2.7)	N/A157	N/G85
B	Carboxylesterase	3WJ1_A[24] 1.50 Å	S151	O/D150-CA/G79 3.1 (2.6)	CA/G153-O/H78 3.4 (2.8)	N/A152	N/G80
1	Acetylcholinesterase-like	1QE3_A[25] 1.50 Å	S189	O/E188-CA/G105 3.2 (2.5)	CA/G191-O/H104 3.6 (3.2)	N/A190	N/G106
2	Carboxylesterase	1LZL_A[26] 1.30 Å	S160	O/Q159-CA/G87 3.1 (2.7)	CA/G162-O/H86 3.2 (2.6)	N/A161	N/G88
3	Mycobacterial antigens	1DQZ_A[27] 1.50 Å	S124	O/L123-CA/G39 4.1 (3.1)	OG/S126-O/D38 2.7	N/M125	N/L40
4	Hypothetical protein TT1662	1UFO_A[28] 1.60 Å	S113	O/G112-CA/G32 3.2 (2.9)	CA/G115-O/H31 3.4 (2.9)	N/L114	N/L33
5	PepX catalytic domain-like	3PUI_A[29] 1.53 Å	S117	O/V116-CE2/Y44 5.7 (4.7)	CD1/L119-OD1/N42 3.6 (2.5)	N/Y118	OH/Y44
6	Prolyl oligopeptidase, C-terminal domain	1H2W_A[30] 1.39 Å	S554	O/G553-CE1/Y473 3.5 (2.8)	CA/G556-O/Y473 4.2 (3.2)	N/N555	OH/Y473
7	DPP6 catalytic domain-like	1ORV_A[31] 1.80 Å	S630	O/W629-CE1/Y547 3.6 (2.9)	CA/G632-O/Y547 3.4 (2.6)	N/Y631	OH/Y547
8	Serine carboxypeptidase-like	3SC2_A[32] 2.20 Å	S146	O/E145-CA/G52 3.8 (3.2)	CB/A148-O/N51 4.0 (2.9)	N/Y147	N/G53
9	Gastric lipase	1HLG_A[33] 3.00 Å	S153	O/H152-CA/G66 3.1 (2.7)	CA/G155-O/H65 3.3 (2.8)	N/Q154	N/L67
10	Proline aminopeptidase-like	1MTZ_A[34] 1.80 Å	S105	O/S104-CA/G36 3.2 (2.8)	CA/G107-O/H35 3.3 (2.7)	N/Y106	N/G37
11	Acetyl xylan esterase-like	1L7A_A[35] 1.50Å	S181	O/G180-CA/G90 4.0 (3.3)	CA/G183-O/H89 3.6 (3.4)	N/Q182	N/Y91
12	Haloalkane dehalogenase	1MJ5_A[36] 0.95 Å	D108	O/H107-CA/G37 4.2 (3.6)	CA/G110-O/H36 3.4 (3.0)	N/W109	N/N38
13	Dienelactone hydrolase	1ZI9_A[37] 1.50 Å	S123	O/Y122-CA/D36 3.1 (2.3)	CA/G125-O/Q35 3.3 (2.8)	N/L124	N/I37
14	Carbon-carbon bond hydrolase	2OG1_A[38] 1.60 Å	S112	O/N111-CA/G41 3.3 (2.4)	CA/G114-O/H40 3.2 (2.8)	N/M113	N/G42
15	Biotin biosynthesis protein BioH	4ETW_A[39] 2.05 Å	A82	O/W81-CA/G21 3.3 (2.7)	CA/G84-O/H20 3.3 (2.7)	N/L83	N/W22
16	Aclacinomycin methylesterase RdmC	1Q0R_A[40] 1.45Å	S102	O/L101-CA/G31 3.3 (2.7)	CA/G104-O/M30 3.2 (2.7)	N/M103	N/G32
17	Carboxylesterase/ lipase	4DIU_A[41] 2.00 Å	S93	O/L92-CA/G23 3.3 (2.7)	CA/G95-O/H22 3.5 (3.0)	N/L94	N/F24
18	Epoxide hydrolase	1QO7_A[42] 1.80 Å	D192	O/G191-CA/G116 4.4 (3.7)	CA/G194-O/H115 3.4 (2.9)	N/I193	N/W117
19	Haloperoxidase	1BRT_A[43] 1.50 Å	S98	O/F97-CA/G31 3.7 (3.1)	CA/G100-O/H30 3.3 (2.4)	N/T99	N/F32
20	Thioesterases	1EI9_A[44] 2.25 Å	S115	O/F114-CA/G40 3.2 (2.8)	CA/G117-O/H39 3.8 (3.2)	N/Q116	N/M41
21	Carboxylesterase/ thioesterase 1	1FJ2_A[45] 1.50 Å	S114	O/F113-CA/G24 3.3 (2.9)	CA/G116-O/H23 3.4 (2.6)	N/Q115	N/L25
22	Ccg1/TafII250-interacting factor B (Cib)	1IMJ_A[46] 2.20 Å	S111	O/P110-CA/G40 4.2 (3.4)	CB/S113-O/H39 3.6 (2.6)	N/L112	N/I41

23	A novel bacterial esterase	1QLW_A[47] 1.09 Å	S206	O/H205-CA/G70 3.3 (2.7)	OG/S208-O/H69 2.9	N/Q207	N/C71
24	Lipase	1JFR_A[48] 1.90 Å	S131	O/H130-CA/G62 3.3 (3.0)	CA/G133-O/P61 3.4 (2.7)	N/M132	N/F63
25	Fungal lipases	1TCA_A[49] 1.55 Å	S105	O/W104-CA/G39 3.5 (2.8)	CA/G107-O/P38 3.5 (3.1)	N/Q106	N/T40
26	Bacterial lipase	1ISP_A[50] 1.30 Å	S77	O/H76-CA/G11 3.6 (3.3)	CA/G79-O/H10 3.4 (2.9)	N/M78	N/I12
27	Pancreatic lipase, N-terminal domain	1BU8_A[51] 1.80 Å	S152	O/H151-CA/G76 3.6 (2.8)	CA/G154-O/H75 3.3 (2.7)	N/L153	N/F77
28	Hydroxynitrile lyase-like	3C6X_A[52] 1.05 Å	S80	O/E79-CB/T11 3.5 (2.7)	CA/G82-O/H10 4.1 (3.7)	N/C81	N/I12
29	Thioesterase domain of polypeptide, polyketide and fatty acid synthases	1JMK_C[53] 1.71 Å	S80	O/Y79-CA/P26 3.3 (2.3)	CA/G82-O/P25 3.0 (2.5)	N/A81	N/V27
30	Cutinase-like	1BS9_A[54] 1.10 Å	S90	O/Y89-CA/E12 3.4 (2.8)	CA/G92-O/R11 3.6 (3.1)	N/Q91	N/T13
31	YdeN-like	1UXO_A[55] 1.80 Å	S71	O/H70-CA/G10 3.4 (3.1)	CA/G73-O/H9 3.5 (2.9)	N/L72	N/Y11
32	Putative serine hydrolase Ydr428c	1VKH_A[56] 1.85 Å	S110	O/H109-CA/G37 3.2 (2.5)	CA/G112-O/H36 3.5 (3.0)	N/V111	N/G38
33	Acylamino-acid- releasing enzyme, C-terminal domain	1VE6_A[57] 2.10 Å	S445	O/Y444-CA/G368 3.7 (3.3)	CA/G447-O/H367 3.5 (3.2)	N/Y446	N/G369
34	Hypothetical esterase YJL068C	1PV1_A[58] 2.30 Å	S161	O/H160-CA/G57 3.6 (3.2)	CA/G163-O/S56 3.4 (3.0)	N/M162	N/L58
35	Hypothetical protein VC1974	1R3D_A[59] 1.90 Å	S91	O/Y90-CA/G23 3.4 (3.0)	CA/G93-O/H22 3.3 (2.6)	N/L92	N/L24
36	Atu1826-like	2I3D_A[60] 1.50 Å	S108	O/Y107-CA/P33 3.5 (2.8)	CA/G110-O/H32 3.3 (2.9)	N/F109	N/H34
37	PHB depolymerase-like	2D80_A[61] 1.70 Å	S39	O/L38-CA/G249 3.1 (2.5)	CA/G41-O/H248 3.4 (2.6)	N/S40	N/C250
38	IroE-like	2GZS_A[11] 1.40 Å	S189	O/H188-NH2/R130 3.3	CA/G191-OD2/D90 5.1 (4.3)	N/Y190	NH1/R130
40	O-acetyltransferase	2B61_A[62] 1.65 Å	S143	O/G142-CB/A48 3.7 (2.8)	CA/G145-O/H47 3.3 (2.9)	N/F144	N/L49
41	2,6-dihydropseudo- oxynicotine hydrolase-like	2JBW_A[63] 2.10 Å	S217	O/R216-CA/G146 3.6 (3.1)	CA/G219-O/G145 3.0 (2.7)	N/L218	N/L147

The nucleophile zone is found around key elements of the catalytic mechanism, the nucleophile and the oxyanion hole. Column S_{nuc} shows the position of the nucleophile (that is commonly a serine) in each of the 40 ABH families and the two carboxylesterases. Similarly, columns N/X_{oxyI} (N/X_{nuc+1}) and N/X_{oxyII} show the position of the first and the second oxyanion hole residues respectively, along with the atoms or reactive groups that shape the oxyanion hole. In the majority of cases, the oxyanion hole is formed by main chain nitrogen atoms, except for the structures with PBD IDs: 3PUI, 1H2W, 1ORV and 2GZS because they use side chain reactive groups. Columns $O/X_{nuc-1} – CA/G_{oxyII-1}$ and $CA/G_{nuc+2} – O/H_{oxyII-2}$ show the two weak hydrogen bonds of the nucleophile zone in the 40 ABH families and the two carboxylesterases, reporting the residues that interact and the contact distances (the values in parentheses show the distance to the hydrogen bonds).

Table 2. Geometrical parameters of the conserved bonds that shape the oxyanion zone in 29 ABH fold families.

N	Family	PDB ID Resolution	(ND1, CD2)/$H_{oxyII-2}$ – O/$G_{oxyII-1}$	(ND1, CD2)/$H_{oxyII-2}$ – O/X_{ozII-1}	N/X_{ozI-2} – O/$G_{oxyII-1}$	N/X_{ozI} – OD/X_{ozII}
A	Carboxylesterase	3WJ2_A[24] 1.61 Å	ND1/H83-O/G84 2.6	CD2/H83-O/I113 3.9 (3.0)	CA/G90-O/G84 3.9 (2.9)	N/I92-OE2/E114 2.8
B	Carboxylesterase	3WJ1_A[24] 1.50 Å	ND1/H78-O/G79 2.6	CD2/H78-O/V108 3.7 (2.8)	CA/G85-O/G79 3.9 (2.8)	N/V87-OD2/D109 3.0
1	Acetylcholinesterase-like	1QE3_A[25] 1.50 Å	ND1/H104-O/G105 2.7	CD2/H104-O/L134 3.8 (2.7)	CA/G111-O/G105 4.1 (3.1)	N/G113-OD1/N135 2.9
2	Carboxylesterase	1LZL_A[26] 1.30 Å	ND1/H86-O/G87 2.7	CD2/H86-O/V116 4.0 (3.1)	CA/G93-O/G87 4.2 (3.1)	N/A95-OE1/E117 2.8
4	Hypothetical protein TT1662	1UFO_A[28] 1.60 Å	ND1/H31-O/G32 2.8	CD2/H31-O/F57 3.5 (2.7)	N/G35-O/G32 2.9	N/K37-OD2/D58 2.9
9	Gastric lipase	1HLG_A[33] 3.00 Å	CD2/H65-O/G66 3.0 (1.9)	ND1/H65-O/G97 4.2	N/A69-O/G66 3.1	CB/A71-O/G97 3.3 (2.5)
10	Proline aminopeptidase-like	1MTZ_A[34] 1.80 Å	CD2/H35-O/G36 2.9 (1.9)	ND1/H35-O/Y61 3.3	N/M40-O/G36 3.3	N/H42-OD2/D62 2.9
11	Acetyl xylan esterase-like	1L7A_A[35] 1.50Å	ND1/H89-O/G90 2.7	CD2/H89-O/M115 3.4 (2.5)	N/A93-O/G90 3.1	OG/S94-CD2/L116 3.7 (2.8)
12	Haloalkane dehalogenase	1MJ5_A[36] 0.95 Å	ND1/H36-O/G37 2.8	CD2/H36-O/C61 3.2 (2.4)	N/T40-O/G37 3.0	N/S42-OD1/D62 2.9
13	Dienelactone hydrolase	1ZI9_A[37] 1.50 Å	NE2/Q35-O/D36 3.3	OE1/Q35-N/L63 3.1	N/G39-O/D36 3.6	CG2/V40-OD1/D62 4.3 (3.5)
14	Carbon-carbon bond hydrolase	2OG1_A[38] 1.60 Å	ND1/H40-O/G41 2.8	CD2/H40-O/K69 3.1 (2.3)	CB/A46-O/G41 3.4 (3.0)	N/G48-OD1/D70 2.8
15	Biotin biosynthesis protein BioH	4ETW_A[39] 2.05 Å	ND1/H20-O/G21 2.8	CD2/H20-O/V45 3.1 (2.2)	N/L24-O/G21 3.4	N/A26-OD1/D46 2.9
16	Aclacinomycin methylesterase RdmC	1Q0R_A[40] 1.45Å	CB/M30-O/G31 3.9 (2.9)	No contact	N/L34-O/G31 3.4	N/A36-OD1/D58 2.8
17	Carboxylesterase/ lipase	4DIU_A[41] 2.00 Å	ND1/H22-O/G23 2.8	CD2/H22-O/P48 3.4 (2.4)	N/G26-O/G23 2.9	OG/S28-CD1/I49 3.1 (2.0)
18	Epoxide hydrolase	1QO7_A[42] 1.80 Å	CD2/H115-O/G116 3.0 (1.9)	ND1/H115-O/P147 3.4	N/G119-O/G116 3.0	CD1/F121-OG/S148 3.3 (2.7)
19	Haloperoxidase	1BRT_A[43] 1.50 Å	ND1/H30-O/G31 2.8	CD2/H30-O/Y56 3.3 (2.3)	N/L34-O/G31 3.0	N/G36-OD1/D57 2.8
20	Thioesterases	1EI9_A[44] 2.25 Å	CD2/H39-O/G40 3.0 (2.0)	ND1/H39-O/L70 3.7	N/D43-O/G40 3.4	N/C45-OD2/D79 3.0
21	Carboxylesterase/ thioesterase 1	1FJ2_A[45] 1.50 Å	ND1/H23-O/G24 2.7	CD2/H23-O/P49 3.6 (2.8)	N/D27-O/G24 3.5	N/H30-Wat859- -CE1/H50 2.8; 3.4 (2.6)
22	Ccg1/TafII250-interacting factor B (Cib)	1IMJ_A[46] 2.20 Å	ND1/H39-O/G40 2.7	CD2/H39-O/I67 3.3 (2.3)	N/F43-O/G40 3.1	N/S45-OD1/D68 2.9
23	A novel bacterial esterase	1QLW_A[47] 1.09 Å	ND1/H69-O/G70 2.8	CD2/H69-O/I102 3.4 (2.4)	N/L73-O/G70 4.0	N/G75-OD1/D103 2.8
26	Bacterial lipase	1ISP_A[50] 1.30 Å	ND1/H10-O/G11 2.8	CD2/H10-O/V39 3.3 (2.3)	N/G14-O/G11 2.9	N/S16-OD1/D40 2.9
27	Pancreatic lipase, N-terminal domain CE1/H75-OD1/D105 3.9 (2.9)	1BU8_A[51] 1.80 Å	CD2/H75-O/G76 3.2 (2.2)	ND1/H75-O/V104 3.3	CG1/I78-O/G76 3.3 (2.3)	CD1/I78-(CG,CD)/R107 4.0; 4.0
28	Hydroxynitrile lyase-like	3C6X_A[52] 1.05 Å	ND1/H10-O/T11 2.7	CD2/H10-O/L36 3.2 (2.4)	N/H14-O/T11 3.1	N/A16-OD1/D37 2.8

30	Cutinase-like	1BS9_A[54] 1.10 Å	CG/R11-O/E12 3.6 (2.6)	NH1/R11-O/I41 3.7	N/A15-O/E12 3.2	N/G18-Wat320- ND2/N42 3.0; 3.5
31	YdeN-like	1UXO_A[55] 1.80 Å	ND1/H9-O/G10 2.7	CD2/H9-O/L37 3.5 (2.8)	N/A13-O/G10 3.5	N/S15-OD1/N38 2.8
32	Putative serine hydrolase Ydr428c	1VKH_A[56] 1.85 Å	ND1/H36-O/G37 2.8	CD2/H36-O/I71 3.5 (2.7)	CB/N45-O/G37 3.0 (2.5)	CD/P47-OE2/E72 3.2 (2.2)
33	Acylamino-acid-releasing enzyme, C-terminal domain	1VE6_A[57] 2.10 Å	CD2/H367-O/G368 2.6 (1.6)	ND1/H367-O/P395 3.9	CB/A372-O/G368 3.5 (2.6)	N/D374-OD1/N396 2.7
35	Hypothetical protein VC1974	1R3D_A[59] 1.90 Å	CD2/H22-O/G23 2.8 (1.7)	ND1/H22-O/L48 3.3	N/G26-O/G23 3.0	N/G28-OD1/D49 2.8
36	Atu1826-like	2I3D_A[60] 1.50 Å	ND1/H32-O/P33 2.7	CD2/H32-O/F63 3.3 (2.3)	CA/G39-O/P33 3.3 (2.3)	N/M41-OD1/N64 2.7
37	PHB depolymerase-like	2D80_A[61] 1.70 Å	ND1/H248-O/G249 2.8	CD2/H248-O/P281 3.6 (2.7)	N/Q252-O/G249 3.1	N/Y254-NE2/Q282 3.0
40	O-acetyltransferase	2B61_A[62] 1.65 Å	ND1/H47-O/A48 2.7	CD2/H47-O/S85 3.2 (2.4)	N/G51-O/A48 3.0	CA/D52-OD1/N86 3.9 (2.9)

The oxyanion zone is found in 29 of 40 ABH fold families. In the 26 of 29 families of the table, the oxyanion zone is arranged around a conserved histidine ($H_{oxyII-2}$) which is located two residues before the oxyanion II (X_{oxyII}), and forms two conserved contacts with neighboring residues, one electrostatic interaction and one weak hydrogen bond, shown in columns (ND1, CD2)/$H_{oxyII-2}$ – O/$G_{oxyII-1}$ and (ND1, CD2)/$H_{oxyII-2}$ – O/X_{ozII-1}. Three ABH structures (PBD IDs: 1ZI9, 1Q0R and 1BS9) have a residue other than histidine in the equivalent position, maintaining in principle the contact scheme. The residue between the oxyanion II (X_{oxyII}) and the conserved histidine ($H_{oxyII-2}$) is a conserved glycine ($G_{oxyII-1}$) that interacts regularly with a residue of the oxyanion loop (X_{ozI-2}), as shown in column N/X_{ozI-2} – O/$G_{oxyII-1}$. Last, the oxyanion zone has a conserved electrostatic interaction between the positions oxyanion zone I (X_{ozI}) and II (D_{ozII}), shown in column N/X_{ozI} – OD/D_{ozII}. The residue at the end of oxyanion loop (X_{ozI}) is variable, while the residue at the position oxyanion zone II (D_{ozII}) is commonly an aspartic acid. All four columns report the residues that interact and the contact distances between them (the values in parentheses show the distance to the hydrogen bonds).

Table 3. Aromatic cluster residues over the plane of nucleophile and oxyanion zones.

N	Family	PDB ID Resolution	X_{oxyII}	$X_{oxyII-4}$	W/Y/F	X_{nuc-1}	S_{nuc}	$SC/X_{ozII+1} – O/H_{oxyII-2}$
A	Carboxylesterase	3WJ2_A[24] 1.61 Å	G85	Y81	Y95	D155	S156	OH/Y115-O/H83; 2.6
B	Carboxylesterase	3WJ1_A[24] 1.50 Å	G80	Y76	Y90	D150	S151	OH/Y110-O/H78; 2.6
1	Acetylcholinesterase-like	1QE3_A[25] 1.50 Å	G106	W102	Y118	E188	S189	OH/Y136-O/H104; 2.7
2	Carboxylesterase	1LZL_A[26] 1.30 Å	G88	W84	S98	Q159	S160	OH/Y118-O/H86; 2.7
3	Mycobacterial antigens	1DQZ_A[27] 1.50 Å	L40	L36	W49	L123	S124	O/G71-OH/Y77; 4.0 CE1/Y77-O/D38; 3.6
4	Hypothetical protein TT1662	1UFO_A[28] 1.60 Å	L33	A29	I40	G112	S113	CB/A59-π/F88; 3.4 CE2/F88-O/H31; 3.5
5	PepX catalytic domain-like	3PUI_A[29] 1.53 Å	OH/Y44	V40	W52	V116	S117	-
6	Prolyl oligopeptidase, C-terminal domain	1H2W_A[30] 1.39 Å	OH/Y473	Y471	Y473	G553	S554	-
7	DPP6 catalytic domain-like	1ORV_A[31] 1.80 Å	OH/Y547	E545	Y547	W629	S630	-
8	Serine carboxypeptidase-like	3SC2_A[32] 2.20 Å	G53	W49	Y60	E145	S146	O/S95-OH/Y151; 4.4 CE2/Y151-O/N51; 3.1
9	Gastric lipase	1HLG_A[33] 3.00 Å	L67	L63	W74	H152	S153	OG/S99-OE1/Q64; 3.7 NE2/Q64-O/H65; 3.3
10	Proline aminopeptidase-like	1MTZ_A[34] 1.80 Å	G37	T33	L45	S104	S105	OE1/Q63-O/H35; 3.0
11	Acetyl xylan esterase-like	1L7A_A[35] 1.50Å	Y91	K87	-	G180	S181	CG2/V117-NE2/Q182; 3.9 NE2/Q182-O/H89; 2.8
12	Haloalkane dehalogenase	1MJ5_A[36] 0.95 Å	N38	F34	W45	H107	D108	CD1/L63-O/H36; 3.8
13	Dienelactone hydrolase	1ZI9_A[37] 1.50 Å	I37	I33	M44	Y122	S123	CD1/L63-CD1/L128; 3.7 CD1/L128-O/Q35; 5.4
14	Carbon-carbon bond hydrolase	2OG1_A[38] 1.60 Å	G42	M38	Y52	N111	S112	CB/S71-CE/M113; 4.0 CE/M113-O/H40; 4.0
15	Biotin biosynthesis protein BioH	4ETW_A[39] 2.05 Å	W22	L18	W29	W81	A82	CD1/L47-O/H20; 3.5
16	Aclacinomycin methylesterase RdmC	1Q0R_A[40] 1.45Å	G32	L28	W39	L101	S102	CE1/H59-O/M30; 2.9
17	Carboxylesterase/lipase	4DIU_A[41] 2.00 Å	F24	L20	V31	L92	S93	OH/Y50-O/H22; 3.0
18	Epoxide hydrolase	1QO7_A[42] 1.80 Å	W117	L113	F124	G191	D192	CD1/L149-O/H115; 3.6
19	Haloperoxidase	1BRT_A[43] 1.50 Å	F32	L28	W39	F97	S98	CD/R58-O/H30; 3.2
20	Thioesterases	1EI9_A[44] 2.25 Å	M41	I37	M51	F114	S115	CD1/I72-CZ2/W38; 4.3 CD1/W38-O/H39; 3.2
21	Carboxylesterase/ thioesterase 1	1FJ2_A[45] 1.50 Å	L25	F21	W32	F113	S114	CB/A51-π/W67; 3.5 NE1/W67-O/H23; 2.7
22	Ccg1/TafII250-interacting factor B (Cib)	1IMJ_A[46] 2.20 Å	I41	L37	W48	P110	S111	CD2/L69-O/H39; 4.4
23	A novel bacterial esterase	1QLW_A[47] 1.09 Å	C71	L67	W78	H205	S206	NE2/Q104-O/H69; 3.0
24	Lipase	1JFR_A[48]	F63	I59	I70	H130	S131	OG1/T89-NH1/R99; 3.9

		1.90 Å						NH1/R99-O/P61; 2.8
25	Fungal lipases	1TCA_A[49] 1.55 Å	T40	L36	F48	W104	S105	CG/P68-SD/M72; 3.7 CE/M72-O/P38; 3.2
26	Bacterial lipase	1ISP_A[50] 1.30 Å	I12	M8	F19	H76	S77	CE2/F41-O/H10; 3.7
27	Pancreatic lipase, N-terminal domain	1BU8_A[51] 1.80 Å	F77	I73	W85	H151	S152	NE1/W106-O/H75; 3.0
28	Hydroxynitrile lyase-like	3C6X_A[52] 1.05 Å	I12	L8	W19	E79	S80	CD1/L38-O/H10; 3.9
29	Thioesterase domain of polypeptide, polyketide and fatty acid synthases	1JMK_C[53] 1.71 Å	V27	A23	Y34	Y79	S80	CB/F51-π/F155; 4.3 CE1/F155-O/P25; 3.3
30	Cutinase-like	1BS9_A[54] 1.10 Å	T13	G9	S22	Y89	S90	OH/Y43-O/R11; 2.7
31	YdeN-like	1UXO_A[55] 1.80 Å	Y11	I7	F20	H70	S71	CG/M39-π/W50; 3.5 CZ3/W50-O/H9; 3.5
32	Putative serine hydrolase Ydr428c	1VKH_A[56] 1.85 Å	G38	Y34	F50	H109	S110	OH/Y73-O/H36; 2.7
33	Acylamino-acid-releasing enzyme, C-terminal domain	1VE6_A[57] 2.10 Å	G369	L365	F381	Y444	S445	OH/Y397-O/H367; 2.8
34	Hypothetical esterase YJL068C	1PV1_A[58] 2.30 Å	L58	Y54	A65	H160	S161	CG2/T87-O/S56; 3.7
35	Hypothetical protein VC1974	1R3D_A[59] 1.90 Å	L24	L20	W31	Y90	S91	CD1/L50-O/H22; 3.5
36	Atu1826-like	2I3D_A[60] 1.50 Å	H34	I30	V46	Y107	S108	CE2/F65-O/H32; 3.4
37	PHB depolymerase-like	2D80_A[61] 1.70 Å	C250	A246	F261	L38	S39	CB/A283-π/W305; 4.0 NE1/W305-O/H248; 2.9
38	IroE-like	2GZS_A[11] 1.40 Å	NH1/R130	M88	V94	H188	S189	-
40	O-acetyltransferase	2B61_A[62] 1.65 Å	L49	I45	W65	G142	S143	CG1/V87-OE1/Q148; 3.4 NE2/Q148-O/H47; 2.9
41	2,6-dihydropseudo-oxynicotine hydrolase-like	2JBW_A[63] 2.10 Å	L147	M143	S154	R216	S217	N/G173-OE1/Q176; 2.8 NE2/Q176-O/G145; 3.0

Aromatic residues tend to appear in specific positions, close to residues that belong to the nucleophile and oxyanion zones. A three-residue cluster is placed above the plane of the zones, which is shaped often by aromatic side chains and consists of: one residue (column $X_{oxyII-4}$) that is found before the conserved histidine ($H_{oxyII-2}$) of Table 2 and the oxyanion II (column X_{oxyII}), another residue (column X_{nuc-1}) that is found before the nucleophile (column S_{nuc}) and the last residue (column W/Y/F) that is found at the end of oxyanion loop and its side chain lies close to the other two. Phenylalanine, tyrosine and tryptophan occur in all these positions, while histidine is found only in column X_{nuc-1}. Aromatic residues are also found below the plane of the nucleophile and oxyanion zones, where the residue next to oxyanion zone II (D_{ozII}) of Table 2 forms often a contact with the conserved histidine ($H_{oxyII-2}$) of Table 2, shown in column SC/X_{ozII+1} – O/$H_{oxyII-2}$. The symbol “-” indicates that there is no contact.

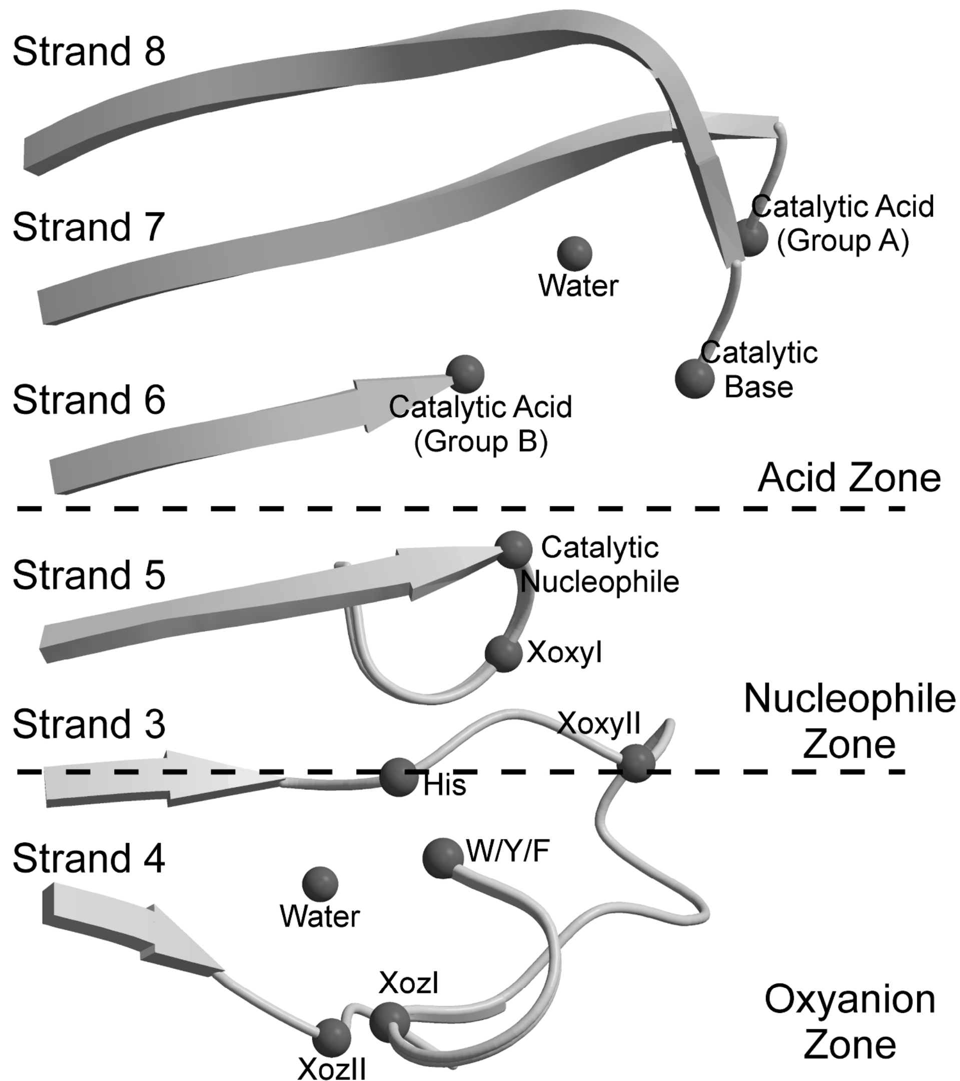

Figure 1: The active site of the ABH fold enzymes. Elements of the catalytic machinery are arranged as follows: the catalytic acid ("Catalytic Acid") is located either after strand β7 (Group A) or strand β6 (Group B); the catalytic base ("Catalytic Base"); and the catalytic nucleophile ("Catalytic Nucleophile"). The oxyanion hole is usually formed by two residues: the oxyanion I residue ("XoxyI"), adjacent to the catalytic nucleophile, and the oxyanion II residue ("XoxyII") in the loop following strand β3. The "Acid zone", the "Nucleophile zone" and the "Oxyanion zone" designate the positions of three clusters of conserved residues that are linked by hydrogen bonds around the catalytic acid, the catalytic nucleophile and the residues of the oxyanion hole, respectively; the boundaries of each zone are marked by dashed lines. The positions of conserved elements of the zones are shown as: "Water" for structural water molecules; "His" for the conserved histidine of the Oxyanion zone; and "W/Y/F" for the aromatic residues tryptophan, tyrosine or phenylalanine from the aromatic cluster. The positions "oxyanion zone I" ("XozI") and "oxyanion zone II" ("XozII") are part of the Oxyanion zone.

247x307mm (300 x 300 DPI)

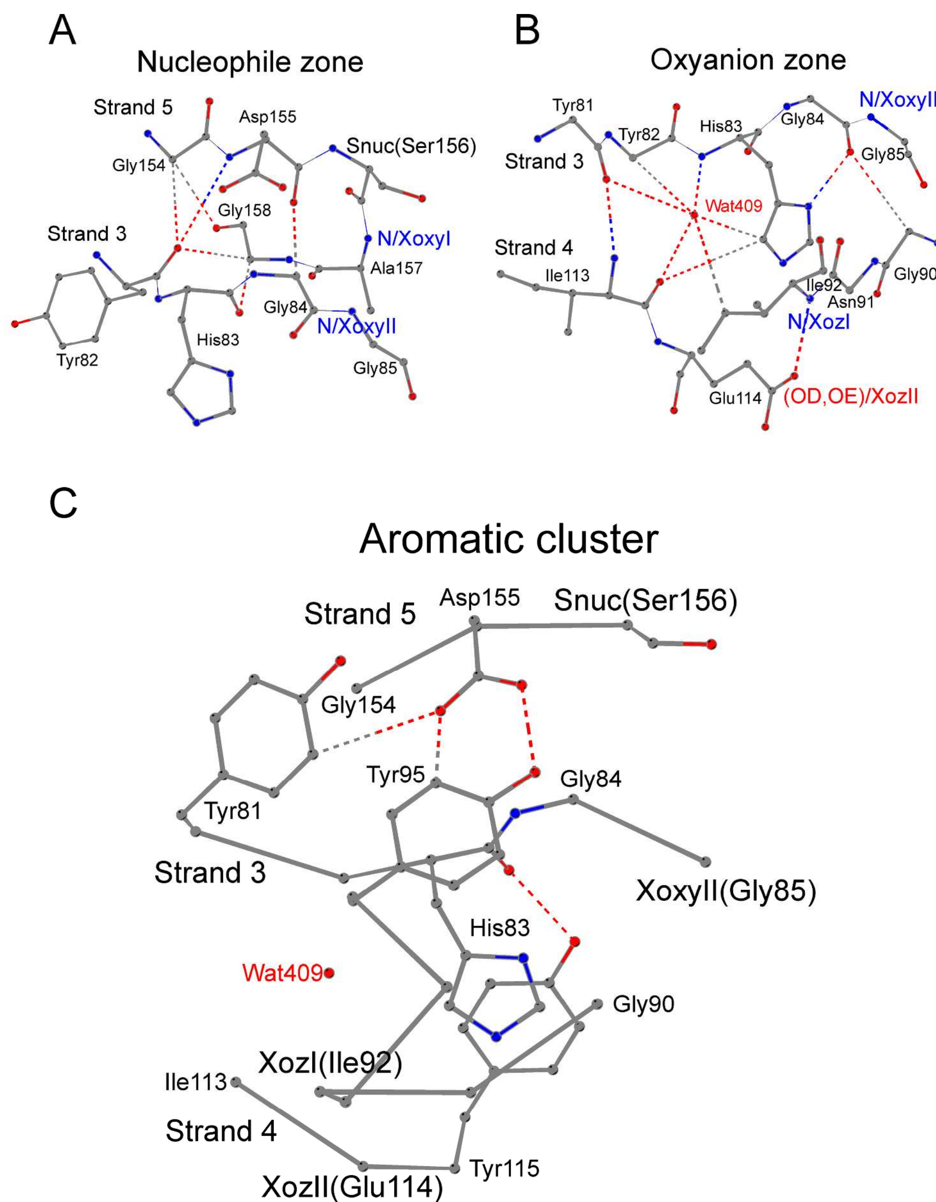

Figure 2: The nucleophile zone (A), the oxyanion zone (B) and the aromatic cluster (C) in the active site of ABH enzymes.

(A) The nucleophile zone in the structure of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The zone is shaped by the residues Asp155 Ser156 Ala157 Gly158 of the nucleophile elbow and His83 Gly84 at the start of loopβ3 αA, and involves two, weak hydrogen bonds, O/Asp155 CA/Gly84 and CA/Gly158 O/His83. The residues Tyr82, Gly85 and Gly154 occupy semi-conserved, peripheral positions and form four, supplementary hydrogen bonds to the nucleophile zone. “Snuc” designates the catalytic nucleophile residue, which is a serine in the majority of ABH enzymes; Main-chain nitrogen atoms at the oxyanion I (“N/XoxyI” from Ala157 in EstFa_R) and oxyanion II (“N/XoxyII” from Gly85 in EstFa_R) positions form the oxyanion hole.

(B) The oxyanion zone in the structure of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The zone is formed by: the residues Tyr81 Tyr82 His83 Gly84 at the end of strand β3 and the

start of loopβ3 αA; the Gly90 Asn91 Ile92 at the end of loopβ3 αA; and Ile113 Glu114 from strand β4. All amino acids except XoxyII, and all contacts except the structural β sheet hydrogen bond between Tyr81 and Ile113, belong to the oxyanion zone."N/XoxyII" designates the main chain nitrogen atom of oxyanion II residue Gly85 in EstFa_R; and "Wat409" designates the conserved structural water molecule forming multiple hydrogen bonds in the oxyanion zone of the ABH enzymes. A hydrogen bond is formed between the main chain nitrogen atom of the residue at "oxyanion zone I" ("N/XozI", Ile92 in EstFa_R) and the residue at "oxyanion zone II" ("OD,OE/XozII", side chain oxygen atom of Glu114 in EstFa_R).

(C) Conserved residues at the aromatic cluster near the active site of the carboxylesterase EstFa_R from Ferroplasma acidiphilum (PDB ID: 3WJ2). The residues Tyr81, Tyr95 and Asp155 are located at positions where aromatic amino acids (as well as hydrophobic residues) are frequently found in all ABH enzymes; the side chain interactions of the residues at these positions form a "roof"-like structural arrangement over the β sheet. Conservation of an aromatic residue is also observed before or after the residue at the position "oxyanion zone II" ("XozII", Glu114 in EstFa_R); in EstFa_R, the side chain hydroxyl of Tyr115 interacts with the main chain oxygen of His83, and the two residues form an aromatic pair below the β sheet."Snuc" and "XoxyII" designate the catalytic nucleophile and oxyanion II residue of the catalytic machinery, and "Wat409" and "XozI" designate the conserved structural water molecule and the residue at position "oxyanion zone I" of the Oxyanion zone.

243x307mm (300 x 300 DPI)

Supplementary Material for "Conserved structural motifs coordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes"

Polytimi S. Dimitriou,[1] Alexander Denesyuk,[1,2] Toru Nakayama,[3] Mark S. Johnson,[1] Konstantin Denessiouk*[1,4]

[1]Structural Bioinformatics Laboratory, Biochemistry,Faculty of Science and Engineering, Åbo Akademi University, Turku 20520, Finland
[2]Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino 142290, Russia
[3]Tohoku University, Biomolecular Engineering, Aoba 6-6-11, Aramaki, Aoba-ku, Sendai, Miyagi 980-8579, Japan
[4]Pharmaceutical Sciences Laboratory, Faculty of Science and Engineering, Pharmacy, Åbo Akademi University, Turku 20520, Finland

*Corresponding author: Faculty of Science and Engineering, Biochemistry, Åbo Akademi University, Tykistökatu 6, BioCity 3A, 20520, Turku, Finland
E-mail address: kdenessi@abo.fi (Konstantin Denessiouk)

Table S1. Inventory of the peripheral interactions around the nucleophile zone in 36 ABH fold families of the SCOP database[1] and the equivalent interactions in the four structures that lack the nucleophile zone (rows 5,6,7 and 38). Column "SCOP Family #/SCOP Family Name" includes the order and the name of each ABH fold family according to the SCOP database[1]; the ABH fold family "39. TTHA-1544-like" is excluded from this table, because its single-entry protein is not a hydrolase. Column "PDB ID/Resolution" shows the Protein Data Bank[2] code of the representative ABH structure and its respective resolution. The remaining four columns show the contact distances of the peripheral interactions around the nucleophile zone, which are shown in Figure 2A, and are reported in the Nucleophile Zone section of the Results. The "-" designates that a contact is not formed. Values in parentheses are distances to the hydrogen atom of the corresponding hydrogen bond. All distances are obtained using Discovery Studio[3].

N	SCOP Family #/ SCOP Family Name	PDB ID Resolution	CA/G_{nuc-2} – O/$X_{oxyII-3}$	CA/G_{nuc-2} – O/G_{nuc+2}	N/X_{nuc-1} – O/$X_{oxyII-3}$	CA/G_{nuc+2} – O/$X_{oxyII-3}$
A	2. Carboxylesterase	3WJ2_A[4] 1.61 Å	CA/G154-O/Y82 3.2 (2.7)	CA/G154-O/G158 3.9 (2.8)	N/D155-O/Y82 3.6	CA/G158-O/Y82 3.2 (2.1)
1	1. Acetylcholinesterase-like	1QE3_A[5] 1.50 Å	CA/G187-O/I103 3.2 (2.7)	CA/G187-O/G191 4.2 (3.3)	N/E188-O/I103 3.6	CA/G191-O/I103 3.2 (2.3)
2	2. Carboxylesterase	1LZL_A[6] 1.30 Å	CA/G158-O/I85 3.1 (2.7)	CA/G158-O/G162 3.6 (2.6)	N/Q159-O/I85 3.3	CA/G162-O/I85 3.3 (2.3)
3	3. Mycobacterial antigens	1DQZ_A[7] 1.50 Å	CA/G122-O/L37 4.3 (3.2)	CA/G122-O/S126 3.4 (2.3)	N/L123-O/L37 4.4	OG/S126-O/L37 3.5
4	4. Hypothetical protein TT1662	1UFO_A[8] 1.60 Å	CA/G111-O/L30 3.3 (2.4)	CA/G111-O/G115 3.6 (2.5)	N/G112-O/L30 2.9	CA/G115-O/L30 3.5 (2.4)
5	5. PepX catalytic domain-like	3PUI_A[9] 1.53 Å	-	CA/G115-O/L119 3.9 (3.2)	-	CA/L119-O/R41 3.6 (2.7)
6	6. Prolyl oligopeptidase, C-terminal domain	1H2W_A[10] 1.39 Å	CA/G552-O/G472 3.7 (3.1)	CA/G552-O/G556 3.3 (2.4)	N/G553-O/G472 4.7	CA/G556-O/G472 3.8 (3.1)
7	7. DPP6 catalytic domain-like	1ORV_A[11] 1.80 Å	CA/G628-O/V546 3.5 (2.9)	CA/G628-CA/G632 3.5 (2.6)	N/W629-O/V546 4.3	CA/G632-O/V546 3.3 (2.5)
8	8. Serine carboxypeptidase-like	3SC2_A[12] 2.20 Å	CA/G144-O/L50 3.0 (2.5)	CA/G144-O/A148 5.4 (4.6)	N/E145-O/L50 3.4	CB/A148-O/L50 3.4 (2.9)
9	9. Gastric lipase	1HLG_A[13] 3.00 Å	CA/G151-O/Q64 3.7 (2.7)	CA/G151-O/G155 4.3 (3.4)	N/H152-O/Q64 3.0	CA/G155-O/Q64 4.0 (2.9)
10	10. Proline aminopeptidase-like	1MTZ_A[14] 1.80 Å	CA/G103-O/M34 3.6 (2.6)	CA/G103-O/G107 3.5 (2.6)	N/S104-O/M34 3.1	CA/G107-O/M34 3.5 (2.4)
11	11. Acetyl xylan esterase-like	1L7A_A[15] 1.50Å	CA/G179-O/Y88 3.3 (2.7)	CA/G179-O/G183 3.5 (2.7)	N/G180-O/Y88 4.0	CA/G183-O/Y88 3.4 (2.6)
12	12. Haloalkane	1MJ5_A[16]	CA/V106-O/Q35	CG1/V106-O/G110	N/H107-O/Q35	CA/G110-O/Q35

	dehalogenase	0.95 Å	3.8 (2.9)	3.5 (3.2)	3.3	3.6 (2.6)
13	13. Dienelactone hydrolase	1ZI9_A[17] 1.50 Å	CA/G121-O/A34 3.2 (2.3)	CA/G121-O/G125 3.4 (2.4)	N/Y122-O/A34 3.1	CA/G125-O/A34 3.6 (2.5)
14	14. Carbon-carbon bond hydrolase	2OG1_A[18] 1.60 Å	CA/G110-O/L39 3.6 (2.6)	CA/G110-O/G114 3.5 (2.6)	N/N111-O/L39 3.0	CA/G114-O/L39 3.7 (2.6)
15	15. Biotin biosynthesis protein BioH	4ETW_A[19] 2.05 Å	CA/G80-O/L19 3.7 (2.8)	CA/G80-O/G84 3.4 (2.5)	N/W31-O/L19 2.8	CA/G84-O/L19 4.0 (2.9)
16	16. Aclacinomycin methylesterase RdmC	1Q0R_A[20] 1.45Å	CA/G100-O/V29 3.6 (2.7)	CA/G100-O/G104 3.6 (2.8)	N/L101-O/V29 3.1	CA/G104-O/V29 4.1 (3.0)
17	17. Carboxylesterase/ lipase	4DIU_A[21] 2.00 Å	CA/G91-O/L21 3.2 (2.4)	CA/G91-O/G95 3.7 (2.6)	N/L92-O/L21 2.9	CA/G95-O/L21 3.8 (2.7)
18	18. Epoxide hydrolase	1QO7_A[22] 1.80 Å	CA/G190-O/L114 3.8 (2.8)	CA/G190-O/G194 3.9 (2.9)	N/G191-O/L114 3.6	CA/G194-O/L114 3.9 (2.8)
19	19. Haloperoxidase	1BRT_A[23] 1.50 Å	CA/G96-O/I29 3.6 (2.7)	CA/G96-O/G100 4.3 (3.5)	N/F97-O/I29 3.0	CA/G100-O/I29 4.4 (3.4)
20	20. Thioesterases	1EI9_A[24] 2.25 Å	CA/G113-O/W38 3.5 (2.6)	CA/G113-O/G117 3.7 (2.7)	N/F114-O/W38 3.0	CA/G117-O/W38 3.7 (2.6)
21	21. Carboxylesterase/ thioesterase 1	1FJ2_A[25] 1.50 Å	CA/G112-O/L22 3.0 (2.6)	CA/G112-O/G116 3.7 (2.7)	N/F112-O/L22 2.9	CA/G116-O/L22 3.3 (2.3)
22	22. Ccg1/TafII250-interacting factor B (Cib)	1IMJ_A[26] 2.20 Å	CA/S109-O/L38 5.3 (4.4)	OG/S109-OG/S113 2.7	CD/P110-O/L38 3.5 (2.8)	OG/S113-O/L38 5.1
23	23. A novel bacterial esterase	1QLW_A[27] 1.09 Å	CA/S204-O/I68 3.4 (2.4)	CB/S204-O/S208 3.1 (2.4)	N/H205-O/I68 3.4	OG/S208-O/I68 3.0
24	24. Lipase	1JFR_A[28] 1.90 Å	CA/G129-O/S60 3.1 (2.3)	CA/G129-O/G133 3.6 (2.6)	N/H130-O/S60 3.0	CA/G133-O/S60 4.0 (2.9)
25	25. Fungal lipases	1TCA_A[29] 1.55 Å	CA/T103-O/V37 3.4 (2.5)	CG2/T103-O/G107 3.6 (2.8)	N/W104-O/V37 3.2	CA/G107-O/V37 3.8 (2.8)
26	26. Bacterial lipase	1ISP_A[30] 1.30 Å	CA/A75-O/V9 3.3 (2.5)	CB/A75-O/G79 3.5 (2.6)	N/H76-O/V9 2.9	CA/G79-O/V9 3.9 (2.8)
27	27. Pancreatic lipase, N-terminal domain	1BU8_A[31] 1.80 Å	CA/G150-O/V74 3.6 (2.6)	CA/G150-O/G154 3.3 (2.4)	N/H151-O/V74 3.0	CA/G154-O/V74 4.2 (3.1)
28	28. Hydroxynitrile lyase-like	3C6X_A[32] 1.05 Å	CA/G78-O/I9 3.4 (2.6)	CA/G78-O/G82 4.5 (3.7)	N/E79-O/I9 3.1	CA/G82-O/I9 4.5 (3.6)
29	29. Thioesterase domain of	1JMK_C[33]	CA/G78-O/F24	CA/G78-O/G82	N/Y79-O/F24	CA/G82-O/F24

	polypeptide, polyketide and fatty acid synthases	1.71 Å	3.8 (2.7)	3.8 (2.9)	3.9	3.9 (2.9)
30	30. Cutinase-like	1BS9_A[34] 1.10 Å	CA/G88-O/A10 3.5 (2.6)	CA/G88-O/G92 3.7 (2.8)	N/Y89-O/A10 3.1	CA/G92-O/A10 3.8 (2.7)
31	31. YdeN-like	1UXO_A[35] 1.80 Å	CA/A69-O/I8 3.2 (2.3)	CB/A69-O/G73 3.8 (3.1)	N/H70-O/I8 2.8	CA/G73-O/I8 4.4 (3.3)
32	32. Putative serine hydrolase Ydr428c	1VKH_A[36] 1.85 Å	CA/G108-O/I35 3.3 (2.3)	CA/G108-O/G112 4.0 (3.0)	N/H109-O/I35 3.2	CA/G112-O/I35 3.5 (2.4)
33	33. Acylamino-acid-releasing enzyme, C-terminal domain	1VE6_A[37] 2.10 Å	CA/G443-O/V366 3.6 (3.1)	CA/G443-O/G447 3.7 (2.7)	N/Y444-O/V366 4.2	CA/G447-O/V366 3.5 (2.6)
34	34. Hypothetical esterase YJL068C	1PV1_A[38] 2.30 Å	CA/G159-O/L55 3.0 (2.3)	CA/G159-O/G163 3.4 (2.4)	N/H160-O/L55 3.1	CA/G163-O/L55 3.3 (2.4)
35	35. Hypothetical protein VC1974	1R3D_A[39] 1.90 Å	CA/G89-O/V21 3.4 (2.4)	CA/G89-O/G93 4.0 (3.0)	N/Y90-O/V21 3.0	CA/G93-O/V21 3.9 (2.9)
36	36. Atu1826-like	2I3D_A[40] 1.50 Å	CA/G106-O/L31 3.4 (2.5)	CA/G106-O/G110 3.6 (2.6)	N/Y107-O/L31 3.0	CA/G110-O/L31 3.8 (2.7)
37	37. PHB depolymerase-like	2D80_A[41] 1.70 Å	CA/G37-O/L247 3.1 (2.6)	CA/G37-O/G41 3.6 (2.7)	N/L38-O/L247 3.2	CA/G41-O/L247 3.8 (2.8)
38	38. IroE-like	2GZS_A[42] 1.40 Å	CA/G187-O/L89 5.5 (4.5)	CA/G187-O/G191 3.6 (2.8)	N/H188-OD2/D90 4.8	-
39	40. O-acetyltransferase	2B61_A[43] 1.65 Å	CA/G141-O/C46 3.5 (2.9)	CA/G141-O/G145 3.7 (2.7)	N/G142-O/C46 4.2	CA/G145-O/C46 3.4 (2.5)
40	41. 2,6-dihydropseudo-oxynicotine hydrolase-like	2JBW_A[44] 2.10 Å	CA/G215-O/L144 3.9 (3.0)	CA/G215-O/G219 3.1 (2.2)	N/R216-O/L144 3.1	CA/G219-O/L144 3.9 (2.8)

Table S2. Inventory of contacts around the conserved, structural water molecule of the oxyanion zone in 28 ABH fold families of the SCOP database[1]. This table does not include the ABH fold families that lack the contact network of the oxyanion zone (SCOP ABH families 3, 5, 6, 7, 8, 24, 25, 29, 34, 38 and 41), the ABH family "27. Pancreatic lipase, N-terminal domain" that maintains the oxyanion zone but lacks the conserved water molecule, and the ABH fold family "39. TTHA-1544-like" single-entry is not a hydrolase. Column "PDB ID/Resolution" shows the Protein Data Bank[2] code of the representative ABH structure and its respective resolution. The remaining six columns show the distances of the contacts that involve the conserved, structural water molecule of the oxyanion zone, which are shown in Figure 2B and discussed in the Oxyanion Zone section of the Results. Values in parentheses are distances to the hydrogen atom of the corresponding hydrogen bond. All distances are obtained using Discovery Studio[3].

N	SCOP Family #/ SCOP Family Name	PDB ID Resolution	O/$X_{oxyII-4}$ - HOH	CA/$X_{oxyII-3}$ - HOH	N/$H_{oxyII-2}$ - HOH	(ND1, CD2)/$H_{oxyII-2}$ - HOH	O/X_{ozII-1} - HOH	CA/X_{ozI} - HOH
A	2. Carboxylesterase	3WJ2_A[4] 1.61 Å	O/Y81-Wat409 3.3	CA/Y82-Wat409 3.9 (3.1)	N/H83-Wat409 2.9	CD2/H83-Wat409 3.1 (2.2)	O/I113-Wat409 2.8	CG1/I92-Wat409 5.1 (4.3)
1	1. Acetylcholinesterase-like	1QE3_A[5] 1.50 Å	O/W102-Wat510 2.7	CA/I103-Wat510 3.5 (2.7)	N/H104-Wat510 3.0	CD2/H104-Wat510 3.5 (2.6)	O/L134-Wat510 2.7	CA/G113-Wat510 3.6 (3.0)
2	2. Carboxylesterase	1LZL_A[6] 1.30 Å	O/W84-Wat324 3.5	CA/I85-Wat324 3.5 (2.7)	N/H86-Wat324 2.8	CD2/H86-Wat324 3.1 (2.3)	O/V116-Wat324 2.8	CA/A95-Wat324 3.8 (2.9)
3	4. Hypothetical protein TT1662	1UFO_A[8] 1.60 Å	O/A29-Wat386 2.8	CA/L30-Wat386 3.4 (2.5)	N/H31-Wat386 3.0	CD2/H31-Wat386 3.2 (2.4)	O/F57-Wat386 2.9	CA/K37-Wat386 4.1 (3.2)
4	9. Gastric lipase	1HLG_A[13] 3.00 Å	O/L63-Wat408 2.6	CA/Q64-Wat408 3.5 (2.8)	N/H65-Wat408 2.9	ND1/H65-Wat408 3.7	O/G97-Wat408 3.0	CA/A71-Wat408 3.7 (2.7)
5	10. Proline aminopeptidase-like	1MTZ_A[14] 1.80 Å	O/T33-Wat295 3.0	CA/M34-Wat295 3.3 (2.3)	N/H35-Wat295 3.1	ND1/H35-Wat295 3.3	O/Y61-Wat295 2.7	CB/H42-Wat301- -Wat295 3.4 (2.6); 3.0
6	11. Acetyl xylan esterase-like	1L7A_A[15] 1.50Å	O/K87-Wat423 2.7	CA/Y88-Wat423 3.8 (3.1)	N/H89-Wat423 3.5	CD2/H89-Wat423 3.8 (2.9)	O/M115-Wat423 2.8	OE1/E98-Wat476 -Wat423 2.7; 2.8
7	12. Haloalkane dehalogenase	1MJ5_A[16] 0.95 Å	O/F34-Wat3004 2.8	CA/Q35-Wat3004 3.5 (2.7)	N/H36-Wat3004 3.0	CD2/H36-Wat3004 3.3 (2.5)	O/C61-Wat3004 2.9	CA/S42-Wat3004 3.7 (2.9)
8	13. Dienelactone hydrolase	1ZI9_A[17] 1.50 Å	O/I33-Wat2726 2.9	CA/A34-Wat2726 3.4 (2.5)	N/Q35-Wat2726 2.9	CG/Q35-Wat2726 3.6 (2.8)	O/P61-Wat2726 2.8	CG2/V40-Wat2726 5.1 (4.2)
9	14. Carbon-carbon bond hydrolase	2OG1_A[18] 1.60 Å	O/M38-Wat829 2.7	CA/L39-Wat829 3.6 (2.9)	N/H40-Wat829 3.1	CD2/H40-Wat829 3.6 (2.9)	O/K69-Wat829 2.8	CA/G48-Wat829 3.0 (2.2)
10	15. Biotin biosynthesis protein BioH	4ETW_A[19] 2.05 Å	O/L18-Wat301 2.7	CA/L19-Wat301 3.5 (2.8)	N/H20-Wat301 3.1	CD2/H20-Wat301 3.5 (2.6)	O/V45-Wat301 2.8	CA/A26-Wat301 3.8 (2.9)
11	16. Aclacinomycin methylesterase RdmC	1Q0R_A[20] 1.45Å	O/L28-Wat3483 2.7	CA/V29-Wat3483 4.2 (3.5)	N/M30-Wat3483 3.9	CB/M30-Wat3483 4.3 (3.4)	O/Y57-Wat3483 3.0	CA/A36-Wat3483 4.3 (3.2)
12	17. Carboxylesterase/	4DIU_A[21]	O/L20-Wat332	CA/L21-Wat332	N/H22-Wat332	CD2/H22-Wat332	O/P48-Wat332	CA/S28-Wat332

	lipase	2.00 Å	2.6	3.3 (2.4)	2.9	3.4 (2.6)	2.7	4.4 (3.5)
13	18. Epoxide hydrolase	1QO7_A[22] 1.80 Å	O/L113-Wat2065 2.7	CA/L114-Wat2065 3.5 (2.6)	N/H115-Wat2065 2.9	ND1/H115-Wat2065 3.0	O/P147-Wat2065 2.7	CA/F121-Wat2065 4.4 (3.4)
14	19. Haloperoxidase	1BRT_A[23] 1.50 Å	O/L28-Wat309 2.8	CA/I29-Wat309 3.6 (2.8)	N/H30-Wat309 3.1	CD2/H30-Wat309 3.4 (2.5)	O/Y56-Wat309 2.8	CA/G36-Wat309 3.3 (2.5)
15	20. Thioesterases	1EI9_A[24] 2.25 Å	O/I37-Wat517 3.3	CA/W38-Wat517 4.0 (3.2)	N/H39-Wat517 3.4	ND1/H39-Wat517 3.1	O/L70-Wat517 2.7	CA/C45-Wat517 3.5 (2.5)
16	21. Carboxylesterase/ thioesterase 1	1FJ2_A[25] 1.50 Å	O/F21-Wat844 2.9	CA/L22-Wat844 3.6 (2.7)	N/H23-Wat844 2.9	CD2/H23-Wat844 3.3 (2.4)	O/P49-Wat844 2.8	CA/G29-Wat844 3.8 (2.9)
17	22. Ccg1/TafII250-interacting factor B (Cib)	1IMJ_A[26] 2.20 Å	O/L37-Wat215 3.0	CA/L38-Wat215 3.7 (3.2)	N/H39-Wat215 3.8	CD2/H39-Wat215 3.5 (2.6)	O/I67-Wat215 2.9	CA/S45-Wat215 3.6 (2.8)
18	23. A novel bacterial esterase	1QLW_A[27] 1.09 Å	O/L67-Wat2163 2.8	CA/I68-Wat2163 3.5 (2.9)	N/H69-Wat2163 2.9	CD2/H69-Wat2163 3.4 (2.6)	O/I102-Wat2163 2.7	CA/G75-Wat2163 3.6 (2.7)
19	26. Bacterial lipase	1ISP_A[30] 1.30 Å	O/M8-Wat306 2.8	CA/V9-Wat306 3.6 (3.0)	N/H10-Wat306 2.9	CD2/H10-Wat306 3.4 (2.6)	O/V39-Wat306 2.8	CA/S16-Wat306 3.8 (2.8)
20	28. Hydroxynitrile lyase-like	3C6X_A[32] 1.05 Å	O/L8-Wat2002 2.7	CA/I9-Wat2002 3.5 (2.9)	N/H10-Wat2002 3.0	CD2/H10-Wat2002 3.4 (2.6)	O/L36-Wat2002 2.7	CA/A16-Wat2002 3.7 (2.8)
21	30. Cutinase-like	1BS9_A[34] 1.10 Å	O/G9-Wat301 3.9	CA/A10-Wat301 3.7 (2.7)	N/R11-Wat301 3.1	NH1/R11-Wat301 3.1	O/I41-Wat301 2.7	O/G18-Wat301 2.8
22	31. YdeN-like	1UXO_A[35] 1.80 Å	O/I7-Wat2076 2.9	CA/I8-Wat2076 3.4 (2.5)	N/H9-Wat2076 3.0	CD2/H9-Wat2076 3.2 (2.4)	O/L37-Wat2076 2.3	CA/S15-Wat2076 4.6 (3.7)
23	32. Putative serine hydrolase Ydr428c	1VKH_A[36] 1.85 Å	O/Y34-Wat648 3.4	CA/I35-Wat648 3.6 (2.8)	N/H36-Wat648 3.0	CD2/H36-Wat648 3.3 (2.5)	O/I71-Wat648 2.6	CA/P47-Wat648 3.4 (2.4)
24	33. Acylamino-acid-releasing enzyme, C-terminal domain	1VE6_A[37] 2.10 Å	O/L365-Wat942 3.5	CA/V366-Wat942 3.3 (2.4)	N/H367-Wat942 2.9	ND1/H367-Wat942 3.4	O/P395-Wat942 2.6	OD2/D374-Wat942 2.6
25	35. Hypothetical protein VC1974	1R3D_A[39] 1.90 Å	O/L20-Wat264 2.6	CA/V21-Wat264 3.4 (2.7)	N/H22-Wat264 3.0	ND1/H22-Wat264 3.4	O/L48-Wat264 2.7	CA/G28-Wat264 3.6 (2.6)
26	36. Atu1826-like	2I3D_A[40] 1.50 Å	O/I30-Wat621 2.8	CA/L31-Wat621 3.9 (3.3)	N/H32-Wat621 3.7	CD2/H32-Wat621 3.8 (3.0)	O/F63-Wat621 2.9	CA/M41-Wat621 3.7 (2.8)
27	37. PHB depolymerase-like	2D80_A[41] 1.70 Å	O/A246-Wat507 2.6	CA/L247-Wat507 3.4 (2.6)	N/H248-Wat507 3.0	CD2/H248-Wat507 3.5 (2.8)	O/P281-Wat507 2.7	OE1/Q282-Wat507 4.0
28	40. O-acetyltransferase	2B61_A[43] 1.65 Å	O/I45-Wat443 2.9	CA/C46-Wat443 4.0 (3.3)	N/H47-Wat443 3.3	CD2/H47-Wat443 3.8 (3.0)	O/S85-Wat443 3.2	O/D52-Wat443 3.8

References

1. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536-540.

2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Research 28:235-242.

3. Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes; 2016.

4. Ohara K, Unno H, Hosoya M, Fujino N, Hirooka K, Takahashi S, Yamasita S, Kusunoki M, Nakayama T (2014) Structural insights into the low pH adaptation of a unique carboxylesterase from *Ferroplasma*: altering the pH optima of two carboxylesterases. J Biol Chem 289:24499-24510.

5. Spiller B, Gershenson A, Arnold FH, Stevens RC (1999) A structural view of evolutionary divergence. Proc Natl Acad Sci U S A 96:12305-12310.

6. Zhu X, Larsen NA, Basran A, Bruce NC, Wilson IA (2003) Observation of an arsenic adduct in an acetyl esterase crystal structure. J Biol Chem 278:2008-2014.

7. Ronning DR, Klabunde T, Besra GS, Vissa VD, Belisle JT, Sacchettini JC (2000) Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. Nat Struct Biol 7:141-146.

8. Murayama K, Shirouzu M, Terada T, Kuramitsu S, Yokoyama S (2005) Crystal structure of TT1662 from *Thermus thermophilus* HB8: a member of the alpha/beta hydrolase fold enzymes. Proteins 58:982-984.

9. Narasimhan D, Collins GT, Nance MR, Nichols J, Edwald E, Chan J, Ko MC, Woods JH, Tesmer JJ, Sunahara RK (2011) Subunit stabilization and polyethylene glycolation of cocaine esterase improves in vivo residence time. Mol Pharmacol 80:1056-1065.

10. Szeltner Z,Rea D, Renner V, Fulop V, Polgar L (2002) Electrostatic effects and binding determinants in the catalysis of prolyl oligopeptidase. Site specific mutagenesis at the oxyanion binding site. J Biol Chem 277:42613-42622.

11. Engel M, Hoffmann T, Wagner L, Wermann M, Heiser U, Kiefersauer R, Huber R, Bode W, Demuth HU, Brandstetter H (2003) The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. Proc Natl Acad Sci U S A 100:5063-5068.

12. Liao DI, Breddam K, Sweet RM, Bullock T, Remington SJ (1992) Refined atomic model of wheat serine carboxypeptidase II at 2.2-Å resolution. Biochemistry 31:9796-9812.

13. Roussel A, Canaan S, Egloff MP, Rivière M, Dupuis L, Verger R, Cambillau C (1999) Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. J Biol Chem 274:16995-17002.

14. Goettig P, Groll M, Kim JS, Huber R, Brandstetter H (2002) Structures of the tricorn-interacting aminopeptidase F1 with different ligands explain its catalytic mechanism. EMBO J 21:5343-5352.

15. Zhang R, Koroleva O, Collert F, Joachimiak A, Midwest Center for Structural Genomics. 1.5 Å crystal structure of the cephalosporin C deacetylase.

16. Oakley AJ, Klvana M, Otyepka M, Nagata Y, Wilce MC, Damborsky J (2004) Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 A resolution: dynamics of catalytic residues. Biochemistry 43:870-878.

17. Kim HK, Liu JW, Carr PD, Ollis DL (2005) Following directed evolution with crystallography: structural changes observed in changing the substrate specificity of dienelactone hydrolase. Acta Crystallogr D Biol Crystallogr 61:920-931.

18. Horsman GP, Ke J, Dai S, Seah SY, Bolin JT, Eltis LD (2006) Kinetic and structural insight into the mechanism of BphD, a C-C bond hydrolase from the biphenyl degradation pathway. Biochemistry 45:11071-11086.

19. Agarwal V, Lin S, Lukk T, Nair SK, Cronan JE (2012) Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proc Natl Acad Sci U S A 109:17406-17411.

20. Jansson A, Niemi J, Mantsala P, Schneider G (2003) Crystal structure of aclacinomycin methylesterase with bound product analogues: implications for anthracycline recognition and mechanism. J Biol Chem 278:39006-39013.

21. Seetharaman J, Lew S, Wang D, Kohan E, Patel D, Whitehead T, Fleishman S, Ciccosanti C, Xiao R, Everett JK, Acton TB, Montelione GT, Tong L, Hunt JF, Northeast Structural Genomics Consortium. Crystal Structure of Engineered Protein. Northeast Structural Genomics Consortium Target OR94.

22. Zou J, Hallberg BM, Bergfors T, Oesch F, Arand M, Mowbray SL, Jones TA (2000) Structure of *Aspergillus niger* epoxide hydrolase at 1.8 Å resolution: implications for the structure and function of the mammalian microsomal class of epoxide hydrolases. Structure 8:111-122.

23. Hofmann B, Tolzer S, Pelletier I, Altenbuchner J, van Pee KH, Hecht HJ (1998) Structural investigation of the cofactor-free chloroperoxidases. J Mol Biol 279:889-900.

24. Bellizzi III JJ, Widom J, Kemp C, Lu JY, Das AK, Hofmann SL, Clardy J (2000) The crystal structure of palmitoyl protein thioesterase 1 and the molecular basis of infantile neuronal ceroid lipofuscinosis. Proc Natl Acad Sci U S A 97:4573-4578.

25. Devedjiev Y, Dauter Z, Kuznetsov SR, Jones TL, Derewenda ZS (2000) Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 Å. Structure 8:1137-1146.

26. Padmanabhan B, Kuzuhara T, Adachi N, Horikoshi M (2004) The crystal structure of CCG1/TAF(II)250-interacting factor B (CIB). J Biol Chem 279:9615-9624.

27. Bourne PC, Isupov MN, Littlechild JA (2000) The atomic-resolution structure of a novel bacterial esterase. Structure 8:143-151.

28. Wei Y, Swenson L, Castro C, Derewenda U, Minor W, Arai H, Aoki J, Inoue K, Servin-Gonzalez L, Derewenda ZS (1998) Structure of a microbial homologue of mammalian platelet-activating factor acetylhydrolases: *Streptomyces exfoliatus* lipase at 1.9 Å resolution. Structure 6:511-519.

29. Uppenberg J, Hansen MT, Patkar S, Jones TA (1994) The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica*. Structure 2:293-308.

30. Kawasaki K, Kondo H, Suzuki M, Ohgiya S, Tsuda S (2002) Alternate conformations observed in catalytic serine of *Bacillus subtilis* lipase determined at 1.3 Å resolution. Acta Crystallogr D Biol Crystallogr 58:1168-1174.

31. Roussel A, Yang Y, Ferrato F, Verger R, Cambillau C, Lowe M (1998) Structure and activity of rat pancreatic lipase-related protein 2. J Biol Chem 273:32121-32128.

32. Schmidt A, Gruber K, Kratky C, Lamzin VS (2008) Atomic resolution crystal structures and quantum chemistry meet to reveal subtleties of hydroxynitrile lyase catalysis. J Biol Chem 283:21827-21836.

33. Bruner SD, Weber T, Kohli RM, Schwarzer D, Marahiel MA, Walsh CT, Stubbs MT (2002) Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. Structure 10:301-310.

34. Ghosh D, Erman M, Sawicki M, Lala P, Weeks DR, Li N, Pangborn W, Thiel DJ, Jornvall H, Gutierrez R, Eyzaguirre J (1999) Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase. Acta Crystallogr D Biol Crystallogr 55:779-784.

35. Janda IK, Devedjiev Y, Cooper DR, Chruszcz M, Derewenda U, Gabrys A, Minor W, Joachimiak A, Derewenda ZS (2004) Harvesting the high-hanging fruit: the structure of the YdeN gene product from *Bacillus subtilis* at 1.8 angstroms resolution. Acta Crystallogr D Biol Crystallogr 60:1101-1107.

36. Arndt JW, Schwarzenbacher R, Page R, Abdubek P, Ambing E, Biorac T, Canaves JM, Chiu HJ, Dai X, Deacon AM, DiDonato M, Elsliger MA, Godzik A, Grittini C, Grzechnik SK, Hale J, Hampton E, Han GW, Haugen J, Hornsby M, Klock HE, Koesema E, Kreusch A, Kuhn P, Jaroszewski L, Lesley SA, Levin I, McMullan D, McPhillips TM, Miller MD, Morse A, Moy K,

Nigoghossian E, Ouyang J, Peti WS, Quijano K, Reyes R, Sims E, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, von Delft F, Wang X, West B, White A, Wolf G, Xu Q, Zagnitko O, Hodgson KO, Wooley J, Wilson IA (2005) Crystal structure of an alpha/beta serine hydrolase (YDR428C) from *Saccharomyces cerevisiae* at 1.85 Å resolution. Proteins 58:755-758.

37. Bartlam M, Wang G, Yang H, Gao R, Zhao X, Xie G, Cao S, Feng Y, Rao Z (2004) Crystal structure of an acylpeptide hydrolase/esterase from *Aeropyrum pernix* K1. Structure 12:1481-1488.

38. Legler PM, Kumaran D, Swaminathan S, Studier FW, Millard CB (2008) Structural characterization and reversal of the natural organophosphate resistance of a D-type esterase, *Saccharomyces cerevisiae* S-formylglutathione hydrolase. Biochemistry 47:9592-9601.

39. Gorman J, Shapiro L, Burley SK, Midwest Center for Structural Genomics. Structural Genomics target NYSGRC-T920 related to A/B hydrolase fold.

40. Osipiuk J, Xu X, Zheng H, Savchenko A, Edwards A, Joachimiak A, Midwest Center for Structural Genomics. Crystal structure of hypothetical protein Atu1826, a putative alpha/beta hydrolase from *Agrobacterium tumefaciens*.

41. Larsen NA, Lin H, Wei R, Fischbach MA, Walsh CT (2006) Structural characterization of enterobactin hydrolase IroE. Biochemistry 45:10184-10190.

42. Hisano T, Kasuya K, Tezuka Y, Ishii N, Kobayashi T, Shiraki M, Oroudjev E, Hansma H, Iwata T, Doi Y, Saito T, Miki K (2006) The crystal structure of polyhydroxybutyrate depolymerase from *Penicillium funiculosum* provides insights into the recognition and degradation of biopolyesters. J Mol Biol 356:993-1004.

43. Mirza IA, Nazi I, Korczynska M, Wright GD, Berghuis AM (2005) Crystal structure of homoserine transacetylase from *Haemophilus influenzae* reveals a new family of alpha/beta-hydrolases. Biochemistry 44:15768-15773.

44. Schleberger C, Sachelaru P, Brandsch R, Schulz GE (2007) Structure and action of a C-C bond cleaving alpha/beta-hydrolase involved in nicotine degradation. J Mol Biol 367:409-418.