Amidohydrolase Superfamily

Aimin Liu, *Georgia State University, Atlanta, Georgia, USA*

Lu Huo, *Georgia State University, Atlanta, Georgia, USA*

Advanced article

Article Contents

Online posting date: 15th August 2014

The amidohydrolase superfamily is a structure-based cluster of enzymes that contain a sturdy and versatile triosephosphate isomerase (TIM)-like $(\beta/\alpha)_8$-barrel fold embracing the catalytic active site. To date, the amidohydrolase superfamily has grown into one of the largest families of enzymes, with tens of thousand of members catalysing a wide range of hydrolytic and nonhydrolytic metabolic reactions which are important in amino acid and nucleotide metabolism as well as biodegradation of agricultural and industrial compounds. Previously, the presence of a mono- or dinuclear *d*-block metal cofactor in the active site was thought to be one of the main characteristics of the members in this superfamily. However, recently new members containing a trinuclear metal cofactors or no cofactor at all were discovered. It has become apparent that activating a well-ordered water molecule by an active site residue for nucleophilic attack on the organic substrate is a common mechanistic feature for all members of the superfamily.

Introduction

The concept of the 'amidohydrolase superfamily' was introduced by Holm and Sander. The striking similarities of the three-dimensional structures of adenosine deaminase (ADA), phosphotriesterase (PTE) and urease (URE) inspired the unification of a broad set of *d*-block metal-dependent hydrolase enzymes into a unique enzyme superfamily (Holm and Sander, 1997). The amidohydrolase superfamily is more than just a class of hydrolase enzymes; it is a functionally diverse group able to catalyse the cleavage of not only C–N, C–C, C–O, C–Cl, C–S and O–P bonds of organic compounds. Most of the characterised members of the amidohydrolase superfamily are important enzymes for metabolism of histidine, tryptophan and lignin-derived compounds metabolism; the *de novo* biosynthesis of purine and pyrimidine nucleotides; and biodegradation of agricultural and industrial materials including rubber chemicals, herbicides, leather, paper and others. Some members of this enzyme superfamily are medically relevant. For instance, ADA catalyses ammonia elimination from the heterocyclic nitrogenous base of the substrate. The deficiency of this enzyme is linked to a common form of severe combined immunodeficiency (SCID). Since the superfamily is not function-designated, an enzyme with amidohydrolase in its name is not necessarily a member of the amidohydrolase superfamily. For instance, penicillin amidohydrolase (also known as penicillin acylase, EC 3.5.1.11) contains neither a catalytic metal ion nor a triosephosphate isomerase (TIM)-barrel domain; thus, it is not a member of the amidohydrolase superfamily. **See also**: Biodegradation of Organic Pollutants

The amidohydrolase superfamily contains more than five related subfamilies in the Pfam database annotation (Finn *et al*., 2014). The amidohydrolase_1 subset (accession number: PF01979) contains a large group of 18 418 protein sequences (http://pfam.janelia.org). Its members catalyse the hydrolysis of a wide range of substrates bearing amide, ester, halogen or other functional groups at carbon and phosphorus centres. This family includes well-characterised enzymes such as ADA, PTE, URE, cytosine deaminase (CDA), D-amino acid deacetylase (AAD), dihydroorotase (DHO), *N*-acetylglucosamine-6-phosphate deacetylase (AGD) and renal dipeptidase (RDP). Among them, DHO and URE also belong to MEROPS peptidase family M38 (*β*-aspartyl dipeptidase, clan MJ), where they are classified as nonpeptidase homologues.

The amidohydrolase_2 subset (accession number: PF04909) contains 6287 aligned protein sequences. These proteins are related to the metal-dependent hydrolases but they appear to be either structurally or functionally divergent from the amidohydrolase_1 group. Members of this branch are known to catalyse nonhydrolytic reactions including decarboxylation and hydration (Liu and Zhang, 2006). The prototypic member of the amidohydrolase_2 is

eLS subject area: Biochemistry

How to cite:
Liu, Aimin; and Huo, Lu (August 2014) Amidohydrolase Superfamily. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020546.pub2

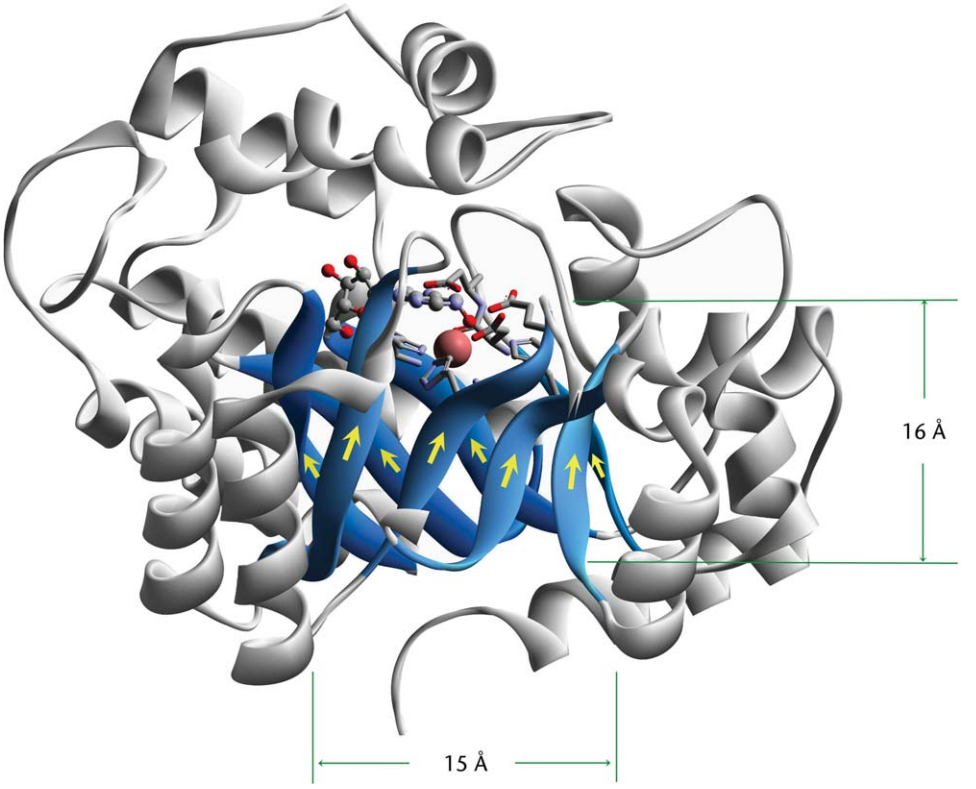


Figure 1 Metal cofactor and the TIM-like parallel barrel core in ADA (PDB entry 1A4M). The metal ion is depicted as CPK sphere, metal ligands are shown in sticks and the propeller structural fold is highlighted in blue colour in the representation. Substrate analogue 6-hydroxy-1,6-dihydro purine nucleoside is represented in scaled ball and stick.

α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD, EC 4.1.1.45), which converts α-amino-β-carboxymuconate-ε-semialdehyde to α-aminomuconate semialdehyde in a nonhydrolytic C–C bond cleavage process (Li *et al*., 2005, 2006, 2007). Despite the similar tertiary structures, it is not clear if members of the amidohydrolase_2 subset arose from a different evolutionary origin than the amidohydrolase_1 due to their low overall sequence similarities.

The amidohydrolase_3 subset (accession number: PF07969) consists of 5270 protein sequences, most of which are hypothetical proteins. Some, including D-aminoacylase, formyltransferase/hydrolase complex Fhc subunit A, are amidohydrolase-like enzymes. This branch of the amidohydrolase superfamily contains the most diverse set of sequences, including a high proportion of outlier sequences that have only low levels of sequence identity to their closest superfamily relatives.

The amidohydolase_4 subset (accession number: PF13147) consists of 10 639 protein sequences. The representative member of this subset includes allantoinase, D-hydantoinase (DYH), dihydropyrimidinase (DHPase), imidazolonepropionase (IPase) and a latent form of dihydroorotase.

The amidohydrolase_5 subset (accession number: PF13594) is a relative small group of enzymes with 3712 sequences to date. Isoaspartyl dipeptidase is its representative member.

More than 30 additional new subsets of amidohydrolases are added, most of which are based on the functional distinctions. These include: creatininase (accession number: PF02633, 1362 sequences), which is a group of urease-related amidohydrolases hydrolysing creatinine to creatine; FGase (accession number: PF05013, 1531 sequences), which consists of a group of *N*-formylglutamate amidohydrolases; and TatD_DNase (accession number: PF01026, 8430 sequences), which is a rapidly growing subgroup of DNAse.

Structure

Most members of the amidohydrolase superfamily consist of a central $(\beta/\alpha)_8$ barrel in which eight parallel β strands are flanked on their outer face by eight α helices (**Figure 1**). However, members that contain a trinuclear metal centre possess a distorted $(\beta/\alpha)_7$ barrel fold. This gives rise to a prominent architectural feature: a sturdy pocket with an internal cavity adjacent to the active site. The pocket has a propeller-like shape made by the seven or eight β strands with a depth of approximately 15–18 Å and a diameter of approximately 14–16 Å that forms a compact thermo-

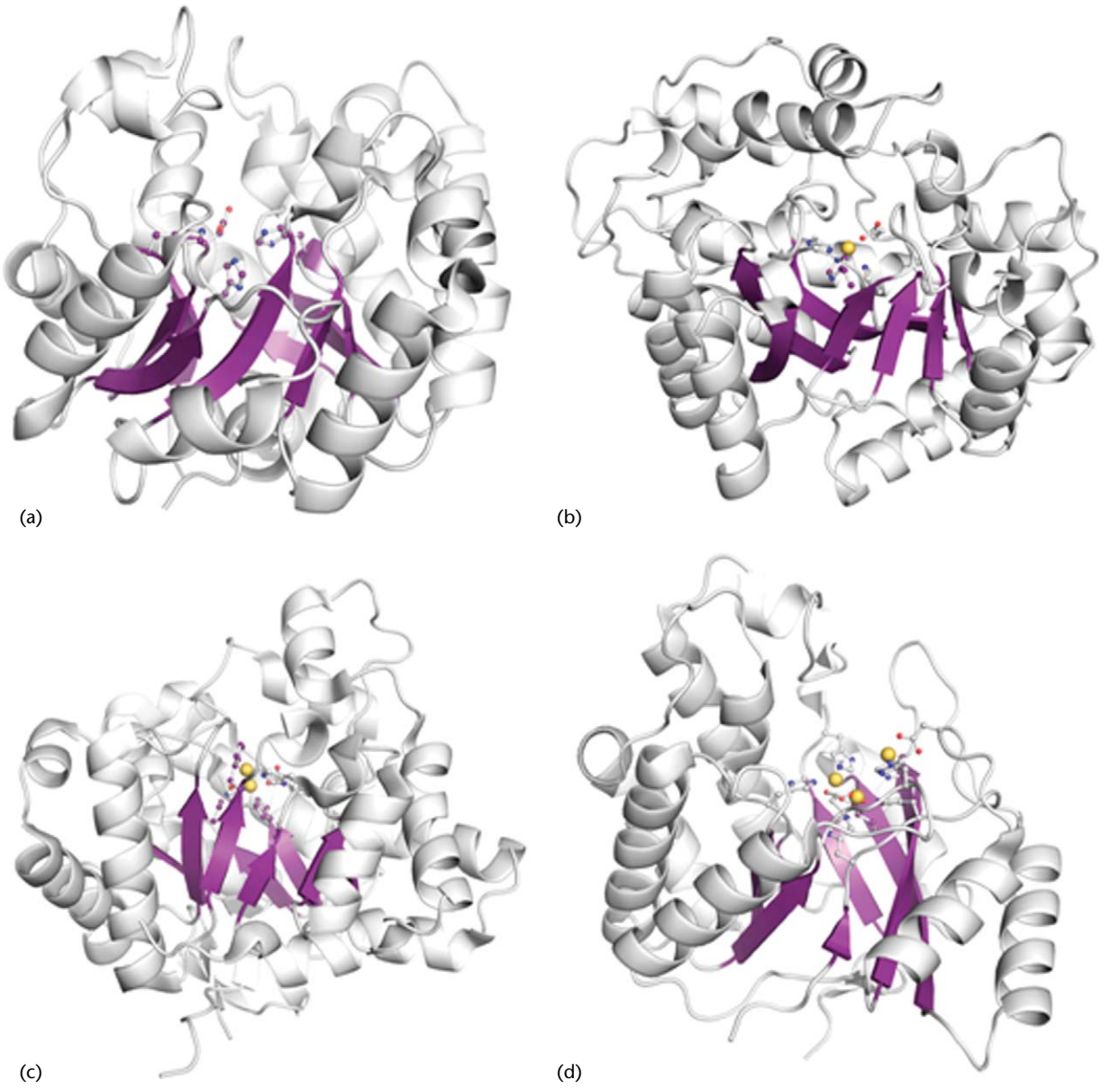


Figure 2 The crystal structures of (a) BmulJ_04915 (PDB entry 4DNM, cofactor free), (b) ADA (PDB entry 1A4M, mononuclear metal cofactor), (c) PTE (PDB entry 1HZY, dinuclear metal cofactor) and (d) HPP (PDB entry 4GC3, trinuclear metal cofactor). Metal ions are shown as yellow spheres. Metal ligand residues are shown in sticks and the propeller structural fold is highlighted in purple.

stable core. The $(\beta/\alpha)_8$-barrel fold is a common structural platform found in approximately 10% of all proteins with known three-dimensional structures and in approximately 33 superfamilies in the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop). It is also known as the TIM-barrel fold because it was first defined from the three-dimensional structure of TIM. The coordination of one or two d-block metal cofactors was previously used to characterize members of the amidohydrolase superfamily. Recently, several new members have been identified either with a trinuclear metal centre or cofactor free, but the four active site residues, three histidines and one aspartate/glutamate, which serve as the metal ligands in the metal-containing amidohydrolase superfamily proteins, are still conserved in the newly classified metal independent members. **Figure 2** shows examples of the TIM-barrel along with an active site that contains no-, mono-, di- or tri- nuclear metal cofactors. There are currently 27 functionally annotated unique members of the amidohydrolase superfamily for which high-resolution X-ray crystal structures are available (**Table 1**).

For the majority of the amidohydrolase superfamily members, the biochemical studies have revealed that the metal cofactor is catalytically essential (Seibert and

Table 1 Functionally annotated and structurally characterised members of the amidohydrolase superfamily

Enzyme	Abbreviation	E.C. number	PDB code
α-Amino-β-carboxymuconate-ε-semialdehyde decarboxylase	ACMSD	4.1.1.45	2HBV
Adenosine deaminase	ADA	3.5.4.4	1A4M
Adenine deaminase	ADase	3.5.4.2	2ICS
N-acetylglucosamine-6-phosphate deacetylase	AGD	3.5.1.25	1O12
Allantoate amidohydrolase	ATase	3.5.2.5	1Z2L
Adenosine 5′-monophosphate deaminase	AMPD	3.5.4.6	2A3L
Cytosine deaminase	CDA	3.5.4.1	1K6W
D-Aminoacylase	DAA	3.5.1.81	1M7J
Dihydroorotase	DHO	3.5.2.3	1J79
Dihydropyrimidinase (L-hydantoinase)[a]	DHPase	3.5.2.2	1GKR
Guanine deaminase	GAH	3.5.4.3	1WKQ
D-Hydantoinase	HYD	3.5.2.2	1NFG
Isoaspartyl dipeptidase	IAD	3.4.19.5	1ONW
Imidazolonepropionase	IPase	3.5.2.7	2BB0
Phosphotriesterase	PTE	3.1.8.1	1HZY
Renal dipeptidase	RDP	3.4.13.11	1ITQ
γ-Resorcylate decarboxylase	RSD	4.1.1.44	2DVT
Urease	URE	3.5.1.5	2UBP
5-Methylthioadenosine/S-adenosylhomocysteine deaminase	MTAD	3.5.4.28	1J6P
4-Oxalomesaconate hydratase	OMAH	4.2.1.83	2GWG
2-Pyrone-4,6-dicarboxylic acid hydrolase	PHD	3.1.1.57	4D8L
N-Acyl-D-glutamate deacylase	AGD	3.5.1.82	3GIP
N-Isopropylammelide isopropyl amidohydrolase	IIA	3.5.99.4	2QT3
Enamidase	ENA	3.5.2.18	2VUN
Isoorotate decarboxylase	IDCase	4.1.1.66	4HK7
L-Histidinol phosphate phosphatase	HPP	3.1.3.15	4GC3
Atrazine chlorohydrolase	ATZA	3.8.1.8	3LS9

[a]Although DHPase and HYD have the same E.C. number, they are different enzymes and have different substrate specificities. DHPase enzymes (including L-hydantoinase) catalyse the reversible hydrolytic ring opening of six- or five-membered cyclic diamides such as dihydropyrimidines and 5′-monosubstituted hydantoins to the corresponding 3-ureido acids and carbamoyl amino acids, respectively. HYD is an industrial enzyme that is widely used in the production of D-amino acids, which are precursors for synthesis of antibiotics, peptides and pesticides.

Raushel, 2005). The metal cofactor is located at the opening of the barrel near the *C*-terminal ends of several β strands, coordinated by several histidine, aspartic acid and in some occasions cysteine residues of the β strands and βα loops. The metal-containing site is referred to as the catalytic face of the barrel, in contrast to the stabilising face of the opposite end of the barrel (Serner and Höcker, 2005). To accommodate the metal centre and substrates, the catalytic face is often wider than the stabilising face, consequently making the other side of the barrel appear more compressed. Thus, the barrel sometimes looks like a conoid.

The most common metal in this enzyme superfamily is a zinc ion or a dinuclear Zn_2 pair. However, transition metals, such as divalent iron, nickel and manganese ions, are also observed in amidohydrolase enzymes. Cobalt ions can often be substituted for zinc *in vitro*, producing a comparable or even higher level of enzyme activity (Li *et al*., 2005). Non *d*-block metal ions, such as Mg^{2+} and Ca^{2+}, are usually unable to perform the biochemical functions at the enzyme active site, suggesting that the role of the *d*-block metal cofactor is beyond simply providing a cationic charge for substrate binding.

The metal cofactor is tethered to the protein through a few protein residues. There are several types of active site architecture known in this enzyme superfamily, and the origin of the metal ligands is a valuable tool to divide the enzyme superfamily into seven subgroups (Seibert and Raushel, 2005). A mononuclear cofactor requires four or five ligands, whereas a dinuclear cofactor demands five or six protein ligands. One notable structural feature is that at least one water molecule or water-derived hydroxide is coordinated to the metal ion(s) in the enzyme structures. The solvent-derived ligand generally remains bound to the metal in the enzyme–substrate complex but it is replaced by substrate on some occasions such as 2,6-dihydroxybenzoate (γ resorcylate) decarboxylase (RSD) (Goto *et al*., 2006).

For a mononuclear metal cofactor of this enzyme superfamily, the fifth β strand invariably provides a histidine ligand for the metal coordination. The sixth β strand contributes a histidine residue either as a metal ligand or a nonligand but an important catalytic component at the enzyme active site. These two histidine residues are the signature amino acids conserved across the mononuclear and dinuclear metal cofactor containing amidohydrolase superfamily members. It is common for the first and/or

eighth β strands to add more protein ligands for metal coordination. An aspartic acid residue from the eighth β strand in many amidohydrolase enzymes plays a dual role, that is, acid/base catalyst and metal ligand.

The two divalent metals are separated by approximately 3.6 Å in the dinuclear metal cofactor. The more buried metal cation is coordinated to two histidine residues from the end of β strand 1 and an aspartate from β strand 8. The more solvent-exposed metal ion is ligated to the protein through two imidazole side chains of histidine from β strands 5 and 6 (Seibert and Raushel, 2005). The two divalent metal ions are bridged by a hydroxide in addition to a carbamate from the posttranslational modification of a lysine residue from β strand 4 in the presence of bicarbonate. In some cases, an unmodified glutamic acid from β strand 4 or a cysteine from β strand 2 is employed as a bridge instead of the modified lysine. Until now, the characterised dinuclear amidohydrolase enzymes have a homogeneous metal preference. However, some mixed metal cofactors may exist. Enamidase, for instance, is a bifunctional enzyme belonging to the amidohydrolase family that mediates hydrolysis of 1,4,5,6-tetrahydro-6-oxonicotinate to ammonia and (*S*)-2-formylglutarate; it reportedly contains an Fe–Zn cofactor (Alhapel *et al*., 2006).

A subfamily under the amidohydrolase superfamily, polymerase and histidinol phosphatase, catalyses the reaction of phosphoester hydrolysis. Members in this subfamily have a trinuclear metal centre buried in a distorted $(\beta/\alpha)_7$ barrel. The β-strand 3 is much longer than other strands composing the barrel and overlays with the β-strand 3 and 4 of the $(\beta/\alpha)_8$ structures. All the metal ligands for the mono- and di- nuclear metal centre are conserved and coordinate the α and β metals. In addition, the third metal is coordinated by an aspartate or histidine from the end of β-strand 1 and two histidine residues from the end of β-strand 2 and 8.

The identification of several metal independent members further expanded the structural diversity of the amidohydroloase superfamily. The overall structure of metal independent members still adopts a $(\beta/\alpha)_8$ TIM-barrel fold and the four conserved residues, three histidine (two from the end of β-strand 1 and one from β-strand 6) and one aspartate (from the end of β-strand 8) ligand residues in the mononuclear metal centre, are still present in the active site. These conserved residues are now proposed to conduct new catalytic roles since they no longer serve as metal ligands.

Other than the histidine residues from the sixth β strands, which are conserved throughout the entire family, the remaining metal ligands exhibit some variations. For example, an H×H metal-binding motif is commonly seen in strand 1 of the β-barrel, but H×D and E×H have been observed from the ACMSD protein subfamily of amidohydrolase_2 (Liu and Zhang, 2006). The histidine from the fifth β-strand is absent in the cofactor free members. Moreover, the metal-binding motif from strand 1 does not always serve as metal ligands. For instance, in the structures of AGD, the H×H motif is present in the active site but it is not ligated directly to the divalent metal.

The overall sequence conservation among the amidohydrolase superfamily is rather low, indicating that the TIM-like barrel fold is not dictated by details of sequences but rather by overall distribution of polar or charged and nonpolar or noncharged residues. Also, it is not unusual that enzymes in this superfamily contain noncatalytic domains in addition to the catalytic domain. Thus, conventional sequence alignments occasionally fail to reveal the general characteristics of a potential new member of the amidohydrolase superfamily. When either the metal-binding motif or the TIM-like barrel is not obvious, an advanced sequence alignment may be required. If this is the case, a secondary structure pattern can be calculated by the PSIPRED protein structure prediction server at http://bioinf.cs.ucl.ac.uk/psipred (Buchan *et al*., 2010). Then a 'secondary structure-based' sequence alignment with those members with known three-dimensional structures may reveal clusters of similar residues at topologically equivalent positions. An example using this strategy was described in a study of ACMSD, which successfully predicted that this is a new member of the amidohydrolase superfamily. A site-directed mutagenesis analysis was followed and the results revealed important information concerning the enzyme such as metal ligand identities and active site residues as well as their possible roles in the catalytic process (Li *et al*., 2006).

However, one should be cautious in drawing conclusions solely based on a sequence study. Sequence, and even structural, similarities do not always translate into functional similarities. One extreme example in the functionally diverse amidohydrolase superfamily is the difference between melamine deaminase and atrazine chlorohydrolase. These two enzymes share 98% sequence identity but catalyse completely different reactions (Seffernick *et al*., 2001).

Structural Diversity: Imperfect Barrel

Approximately 200 amino acids are required to assemble a TIM-like parallel barrel. To form the barrel, a β strand/α helix pair will have to repeat eight times and linked by $\beta\alpha$ loops. Occasionally, there is a deletion or an insertion in a strand that causes a single residue interruption of the secondary structure in the amidohydrolase enzymes. For example, the members of the large radical S-adenosylmethionine (SAM) superfamily typically employ an imperfect TIM barrel fold even though they do not belong to amidohydrolase superfamily. It is not unusual for a member of the superfamily to exhibit a distorted propeller barrel or additional insertion domains and loops for specific structural or functional requirements, such as substrate specificity. **See also**: Radical S-adenosylmethionine (SAM) Superfamily

Owing to high variability in the lengths of barrel strands and helices, evolution has produced highly distorted $(\beta/\alpha)_8$-

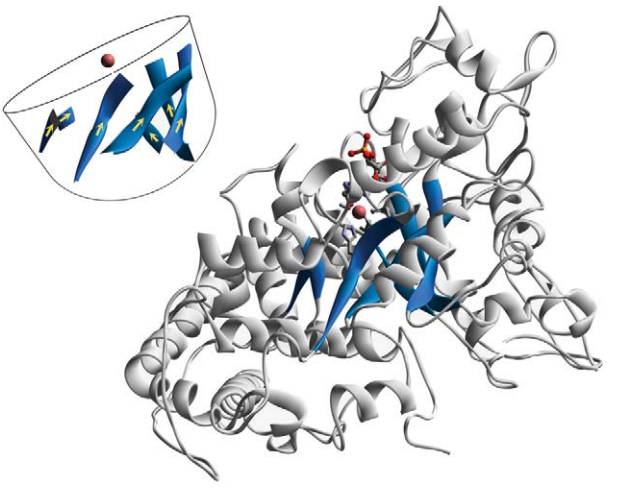
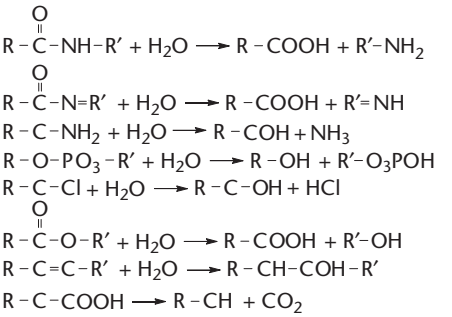
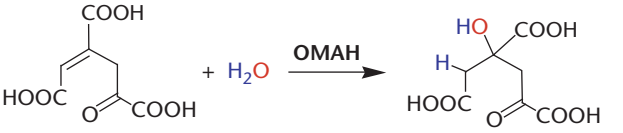


Figure 3 An imperfect TIM-barrel exhibited in the structure of AMPD in complex with coformycin 5′-phosphate (ball and stick) (PDB entry 2A3L). The inset highlights the distorted barrel and the zinc ion (CPK sphere).



Scheme 1 Hydrolytic and nonhydrolytic reactions catalysed by the structurally characterised members of the amidohydrolase superfamily.



Scheme 2 Chemical reaction catalysed by 4-oxalomesaconate hydratase (OMAH). A water molecule is added across the 4,5-carbon double bond.

propeller folds for some special biochemical needs. A small subset of the amidohydrolase superfamily with imperfect folds has been reported (Han *et al*., 2006). In the structure of adenosine 5′-monophosphate deaminase (AMPD), six or seven of the eight β strands can be recognized, but the fifth and sixth are replaced by somewhat irregular loop structures, resulting in an apparently imperfect β-barrel (**Figure 3**). Although the barrel is not composed of eight strands, the catalytic zinc ion is coordinated to an aspartic acid and three histidine residues, as seen in the active site of other amidohydrolase enzymes (Han *et al*., 2006).

One-fold, Numerous Reactions

Hydrolytic reactions

Thus far, members of the amidohydrolase_1 subfamily catalyse only hydrolytic reactions (**Scheme 1**). There are two classes of hydrolytic reactions that differ in their substrate structural properties. Accordingly, members of the amidohydrolase_1 subfamily may be divided into cyclic amidohydrolase enzymes and noncyclic amidohydrolase enzymes. Cyclic amidohydrolase enzymes catalyse the hydrolysis of cyclic C–N bonds. They are commonly found in nucleotide metabolites of purines and pyrimidines. These enzymes share similar catalytic mechanisms and show considerable sequence and structural homogeneity, suggesting that they might have evolved from a common ancestral protein. Enzymes in the noncyclic group catalyse hydrolysis of C–O, P–O, C–Cl and C–S bonds in addition to the common C–N bond cleavage.

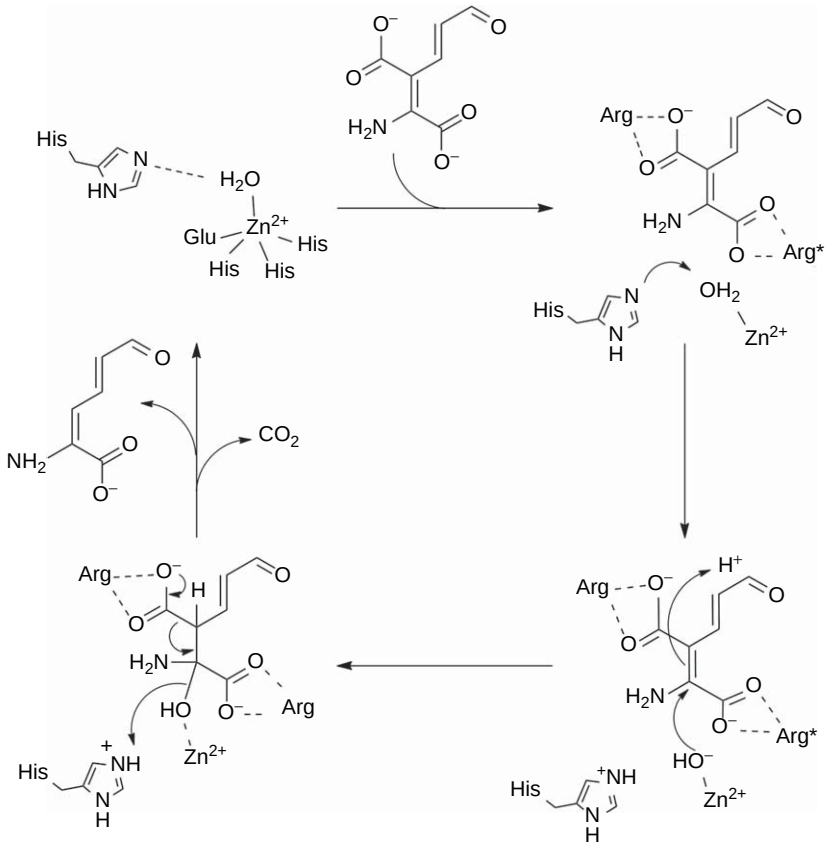
Nonhydrolytic reactions

The more diverse catalytic reactions of the amidohydrolase_2 subfamily are not restricted to hydrolytic reactions (**Scheme 1**) (Li *et al*., 2006). ACMSD is the first structurally characterised member that catalyses a nonhydrolytic reaction in this subfamily (PDB entry 2HBV). This enzyme catalyses a C–C bond cleavage of a nonoxidative decarboxylation reaction. The second described member of this subfamily is RSD, which also catalyses a C–C bond cleavage reaction (Goto *et al*., 2006). A recent bioinformatics study has led to the suggestion that at least 60 nonredundant protein sequences constitute a new ACMSD protein family as a subgroup of the amidohydrolase superfamily (Liu and Zhang, 2006). The members of the ACMSD protein family include several functionally annotated members such as isoorotate decarboxylase (IDCase), 5-carboxyvanillic acid decarboxylase (5-CVD), 4-oxalomesaconate hydratase (OMAH) and RSD in addition to ACMSD (Liu and Zhang, 2006). Decarboxylation is one of the most common reactions in biological systems and decarboxylases adopt very diverse catalytic mechanisms (Li *et al*., 2012). Structural and mechanistic studies on ACMSD reveal a transition metal dependent, oxygen independent decarboxylase that have never been noted. The crystal structures of RSD, 5-CVD, IDCase and OMAH are all available in the protein data bank with high similarity when compared to ACMSD: a mononuclear metal centre buried in a $(\beta/\alpha)_8$ TIM-barrel.

Unlike ACMSD and RSD, the enzyme OMAH catalyses a hydration reaction, adding a water molecule across a carbon-carbon double bond (**Scheme 2**).

Enzyme Mechanisms

The best-characterised members of the amidohydrolase superfamily share a common catalytic mechanism. **Scheme 3** illustrates the proposed mechanism for the mononuclear



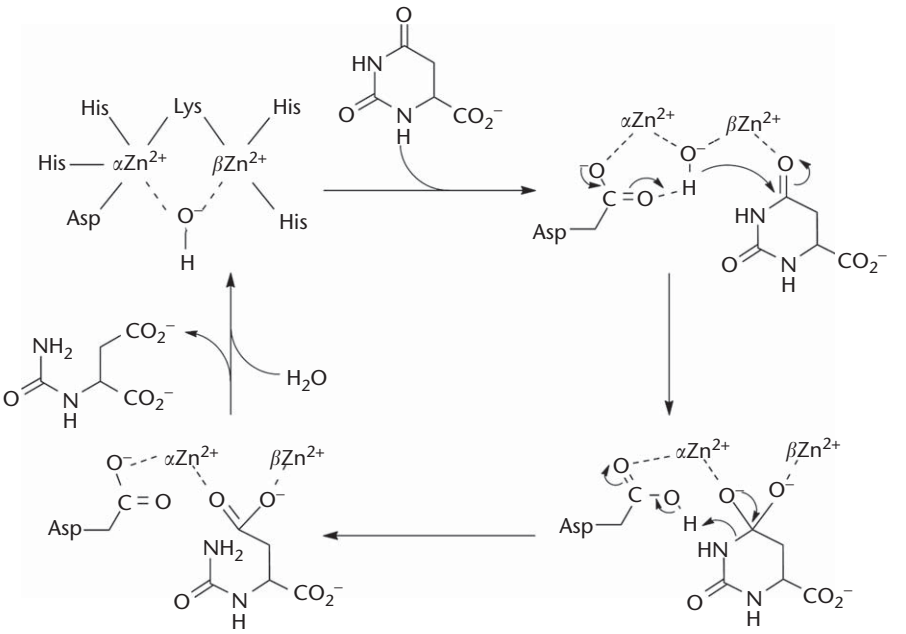
Scheme 3 A working model of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) catalytic cycle for a nonhydrolytic C–C bond cleavage. Arg* represents a conserved arginine residue intruded to catalytic centre from a neighbouring subunit of the homodimeric protein.

Zn-dependent enzyme ACMSD. The common feature of the mechanism is that metal-bound water becomes a hydroxide with the assistance of an active site catalyst, that is, a histidine residue. The hydroxide attacks the substrate-bearing amide or other functional groups at carbon and phosphorus centres and forms a substrate-based tetrahedral carbon (or phosphorus) intermediate. Subsequent collapse of the tetrahedral intermediate leads to the formation of the products (Huo *et al*., 2012, 2013).

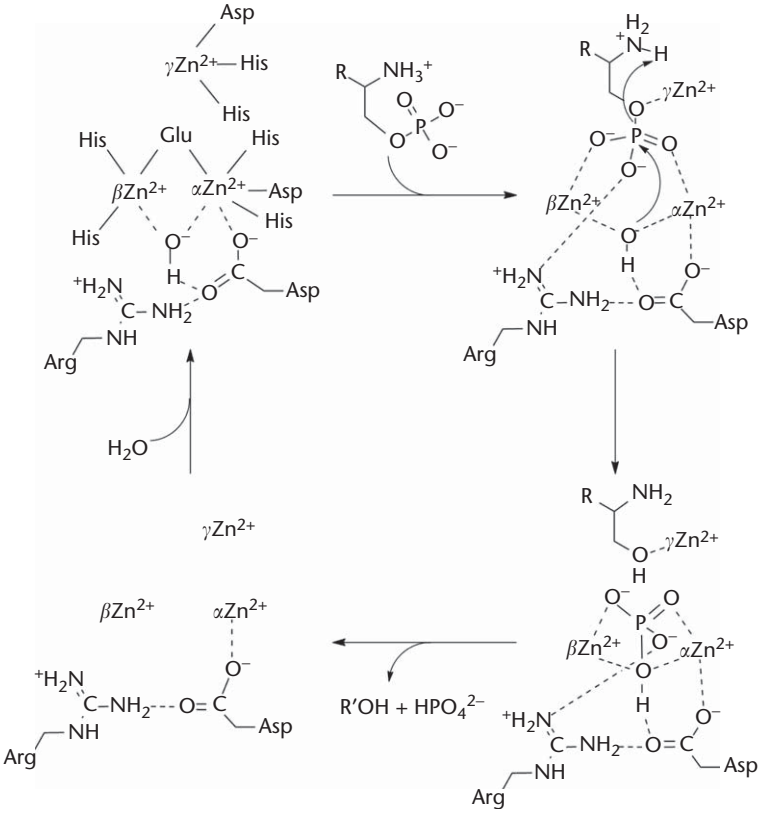
Mechanistic understanding of the nonhydrolytic C–C bond and C=C bond cleavage reactions is still in its infancy. However, recent biochemical and structural studies have yielded evidence supporting the idea that these reactions have mechanisms similar to those proposed for the hydrolytic reactions. The core elements for the required chemistry, a proper Lewis acid metal centre, a water ligand and a nearby histidine (the conserved histidine from the end of β-strand 6), are all in place at the active site of the characterised members of the amidohydrolase_2 branch, suggesting a common mechanistic paradigm for substrate activation at the early steps of the reaction catalysed by the enzyme of this superfamily. From a chemical perspective, the enzyme OMAH of the ACMSD protein family catalyses a mechanistically insightful reaction; it adds a water molecule to the substrate, as if it were the half-reaction in a decarboxylation mechanism proposed for ACMSD (**Scheme 3**). Although there is no characterised member of the amidohydrolase_3 subfamily at the present stage, it is almost certain that some structural variations and new functionalities will soon be described for members of this branch.

An arginine residue, Arg239 in ACMSD, is conserved across the characterised members which catalyse decarboxylation reactions and has been shown to be catalytically essential. Arg239 plays an important role in substrate binding in solution for ACMSD. The crystal structure of IDCase in complex with a substrate analogue further demonstrated that this arginine specifically binds to the leaving carboxyl group (Xu *et al*., 2013). Unlike other active site residues, Arg239 is intruded from a neighbouring subunit, indicating this subfamily of enzymes needs to function as dimers. This is also verified by the gel filtration studies done for OMAH and RDC.

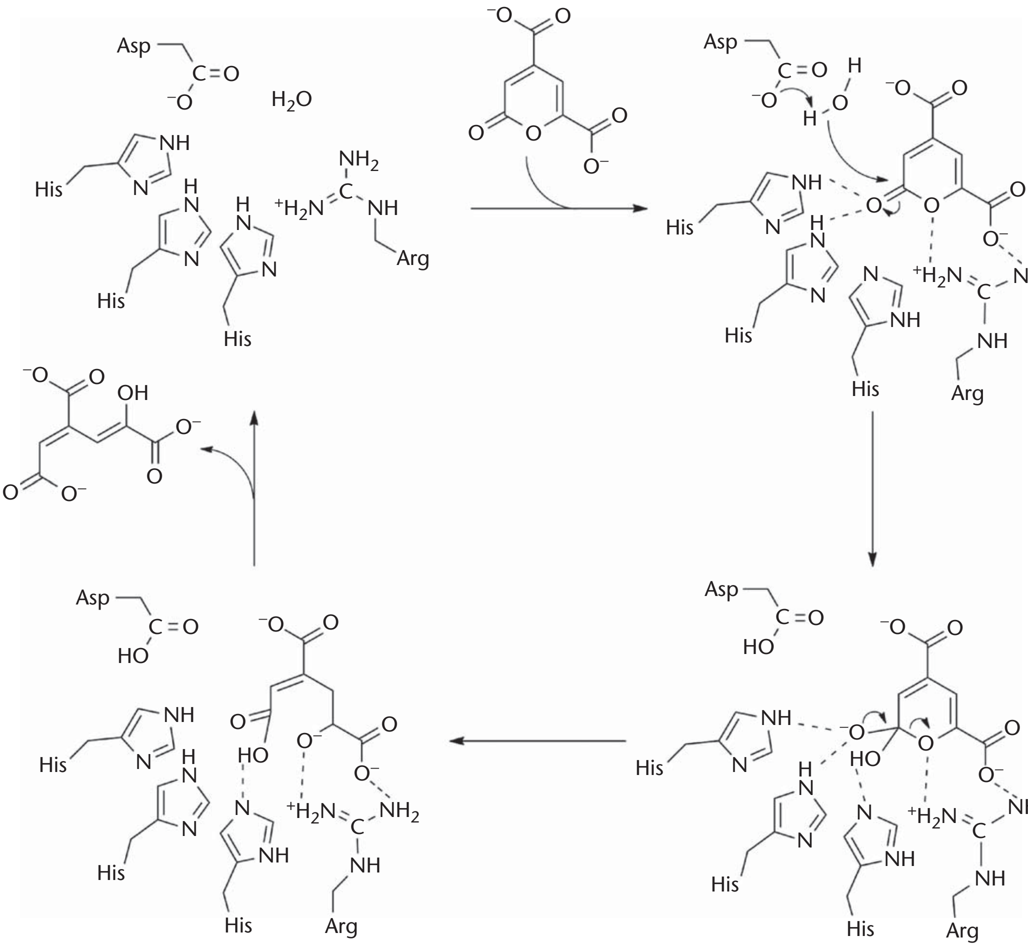
The dinuclear enzymes proceed similarly, with a second divalent metal ion providing additional activation power towards the substrate. DHO is one of the best-characterised dinuclear amidohydrolase enzymes. It is believed that the original *keto* oxygen of the substrate interacts with the more solvent exposed metal ion in DHO, whereas the more buried metal ion activates the nucleophilic



Scheme 4 The proposed catalytic mechanism of dihydroorotase (DHO). DHO contains a dinuclear metal centre and catalyses a ring opening reaction using dihydroorotate as substrate.



Scheme 5 The mechanistic model for a trinucler zinc centre-mediated L-histidinol phosphate phosphatase (HPP) reaction. L-Histidinol and inorganic phosphate are the two reaction products.



Scheme 6 The proposed catalytic cycle of LigI, an enzyme catalyses the reversible hydrolysis of 2-pyrone-4,6-dicarboxylate to 4-oxalomesaconate and 4-carboxy-2-hydroxymuconate in the degradation of lignin. The enzymatic reaction is independent to the presence of a metal ion.

attack during catalysis. A proposed mechanism of DHO catalysed reaction is shown in **Scheme 4**.

The subgroup of polymerase and histidinol phosphatases that contain a trinuclear metal centre is relatively new. Tyrosine phosphatase and L-histidinol phosphate phosphatase (HPP) are two representative members (Hagelueken *et al.*, 2009; Kim *et al.*, 2011; Ghodge *et al.*, 2013). In HPP-catalysed reactions, three metal ions and one arginine residue are involved in binding with the phosphate oxygens of the substrate. The γ-metal binds to the leaving oxygen and is proposed to serve as a Lewis acid to facilitate the O–P bond cleavage. A hydroxide group, which initially serves as a bridge between the α- and β-metal, nucleophilic attacks the phosphorus of the substrate. The amino group of the substrate may function as a general acid to protonate the leaving oxygen. A proposed mechanism is shown in **Scheme 5**.

The enzyme uronate isomerase (URI) catalyses a non-hydrolytic aldose/ketose isomerisation reaction between D-glucuronate and D-fructuronate (Nguyen *et al.*, 2008). This enzyme was initially thought to be the first member of the amidohydrolase superfamily whose activity is independent of the presence of a divalent metal (Williams *et al.*, 2006). However, a careful re-examination of this enzyme revealed that URI is a mononuclear zinc-containing enzyme (Nguyen *et al.*, 2009).

LigI, an enzyme which catalyses the reversible hydrolysis of 2-pyrone-4,6-dicarboxylate from *Sphingomonas paucimobilis*, is the true first member of amidohydrolase superfamily that does not require a divalent metal (Hobbs *et al.*, 2012). Later on, BmulJ_04915 and GLI are found as the second and third examples (Hobbs *et al.*, 2013). Although this group of enzymes does not contain a metal cofactor, the original metal-binding residues, including three histidine and one aspartate, are still conserved in the active site with new catalytic functions. For LigI and BmuIJ_04915, the three histidine residues are proposed to bind and polarise the substrate. The aspartate residue is proposed to activate a water molecule for nucleophilic attack (**Scheme 6**).

Concluding Remarks

The amidohydrolase superfamily has attracted substantial attention as a large structure-based cluster of enzymes with

thousands of members and divergent catalytic functions. In the past few years there have been huge advances in the studies of the amidohydrolase superfamily. An increasing number of sequences are now becoming functionally annotated, and many members that exhibit distinct biological significances are now structurally defined. It is becoming clear that the metal cofactor elaborated by the sturdy and versatile TIM-like β barrel is enormously powerful. The substrate specificity appears to be mostly dictated by the loops, insertions and conformational restrictions of the catalytic face of the TIM-like β barrel. This insight has come from site-directed mutagenesis, kinetics, structural and spectroscopic studies of many members of the superfamily.

Major advances have been made, but a structure-based functional annotation is still challenging. Many principals and strategies related to the design of loops, insertions, subunit interactions and the catalytic impact of protein dynamics remain to be elucidated. Nonetheless, the rapid expansion of knowledge concerning the metal-dependent TIM-barrel enzymes is making the amidohydrolase superfamily a well-characterised large group of enzymes that may enable a reliable evolutionary analysis for the origin of the divergent members. The amidohydrolase superfamily as a whole is emerging as a popular and valuable asset in the enzymology field for studying structure–function relationships and evolution.

Acknowledgement

This work was supported in part or whole by National Science Foundation Grant CHE-0843537, National Institutes of Health grant GM108988, and Georgia Research Alliance Distinguished Scholar Programme (A.L.), and Molecular Basis of Disease Area of Focus graduate fellowship (L.H.).

References

Alhapel A, Darley DJ, Wagener N *et al*. (2006) Molecular and functional analysis of nicotinate catabolism in *Eubacterium barkeri*. *Proceedings of the National Academy of Sciences USA* **103**: 12341–12346.

Buchan DW, Ward SM, Lobley AE *et al*. (2010) Protein structure prediction servers at University College London. *Nucleic Acids Research* **38**: W563–W568.

Finn RD, Bateman A, Clements J *et al*. (2014) The Pfam protein families database. *Nucleic Acids Research* **42**: D222–D230.

Ghodge SV, Fedorov AA, Fedorov EV *et al*. (2013) Structural and mechanistic characterization of L-histidinol phosphate phosphatase from the polymerase and histidinol phosphatase family of proteins. *Biochemistry* **52**: 1101–1112.

Goto M, Hayashi H, Miyahara I *et al*. (2006) Crystal structures of nonoxidative zinc-dependent 2,6-dihydroxybenzoate (γ-resorcylate) decarboxylase from *Rhizobium* sp. strain MTP-10005. *Journal of Biological Chemistry* **281**: 34365–34373.

Hagelueken G, Huang H, Mainprize IL, Whitfield C and Naismith JH (2009) Crystal structures of Wzb of *Escherichia coli* and CpsB of *Streptococcus pneumoniae*, representatives of two families of tyrosine phosphatases that regulate capsule assembly. *Journal of Molecular Biology* **392**: 678–688.

Han BW, Bingman CA, Mahnke DK *et al*. (2006) Membrane association, mechanism of action, and structure of *Arabidopsis* embryonic factor 1 (FAC1). *Journal of Biological Chemistry* **281**: 14939–14947.

Hobbs ME, Malashkevich V, Williams HJ *et al*. (2012) Structure and catalytic mechanism of LigI: insight into the amidohydrolase enzymes of cog3618 and lignin degradation. *Biochemistry* **51**: 3497–3507.

Hobbs ME, Vetting M, Williams HJ *et al*. (2013) Discovery of an L-fucono-1,5-lactonase from cog3618 of the amidohydrolase superfamily. *Biochemistry* **52**: 239–253.

Holm L and Sander C (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins* **28**: 72–82.

Huo L, Davis I, Chen L and Liu A (2013) The power of two: arginine 51 and arginine 239* from a neighboring subunit are essential for catalysis in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase. *Journal of Biological Chemistry* **288**: 30862–30871.

Huo L, Fielding AJ, Chen Y *et al*. (2012) Evidence for a dual role of an active site histidine in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase. *Biochemistry* **51**: 5811–5821.

Kim HS, Lee SJ, Yoon HJ *et al*. (2011) Crystal structures of YwqE from *Bacillus subtilis* and CpsB from *Streptococcus pneumoniae*, unique metal-dependent tyrosine phosphatases. *Journal of Structural Biology* **175**: 442–450.

Li T, Iwaki H, Fu R *et al*. (2006) α-Amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD) is a new member of the amidohydrolase superfamily. *Biochemistry* **45**: 6628–6634.

Li T, Lu H, Pulley C and Liu A (2012) Decarboxylation mechanisms in biological system. *Bioorganic Chemistry* **43**: 2–14.

Li T, Ma J, Hosler JP, Davidson VL and Liu A (2007) Detection of transient intermediates in the metal-dependent non-oxidative decarboxylation catalyzed by α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase. *Journal of the American Chemical Society* **129**: 9278–9279.

Li T, Walker AL, Iwaki H, Hasegawa Y and Liu A (2005) Kinetic and spectroscopic characterization of ACMSD from *Pseudomonas fluorescens* reveals a pentacoordinate mononuclear metallocofactor. *Journal of the American Chemical Society* **127**: 12282–12290.

Liu A and Zhang H (2006) Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry* **45**: 10407–10411.

Nguyen TT, Brown S, Fedorov AA *et al*. (2008) At the periphery of the amidohydrolase superfamily: Bh0493 from *Bacillus halodurans* catalyzes the isomerization of D-galacturonate to D-tagaturonate. *Biochemistry* **47**: 1194–1206.

Nguyen TT, Fedorov AA, Williams L *et al*. (2009) The mechanism of the reaction catalyzed by uronate isomerase illustrates how an isomerase may have evolved from a hydrolase within the amidohydrolase superfamily. *Biochemistry* **48**: 8879–8890.

Seffernick JL, de Souza ML, Sadowsky MJ and Wackett LP (2001) Melamine deaminase and atrazine chlorohydrolase: 98

percent identical but functionally different. *Journal of Bacteriology* **183**: 2405–2410.

Seibert CM and Raushel FM (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* **44**: 6383–6391.

Serner R and Höcker B (2005) Catalytic versatility, stability, and evolution of the ($\beta\alpha$)8-barrel enzyme fold. *Chemical Reviews* **105**: 4038–4055.

Williams L, Nguyen T, Li Y, Porter TN and Raushel FM (2006) Uronate isomerase: a nonhydrolytic member of the amidohydrolase superfamily with an ambivalent requirement for a divalent metal ion. *Biochemistry* **45**: 7453–7462.

Xu S, Li W, Zhu J *et al*. (2013) Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase. *Cell Research* **23**: 1296–1309.

Further Reading

Frey PA and Hegeman AD (2007) *Enzymatic Reaction Mechanisms*, 1st edn. Oxford University Press. ISBN: 0195122585, 848 pages.

Petsko GA and Ringe D (2004) *Protein Structure and Function*. 1st edn. Sinauer Associates. Incorporated, ISBN: 0878936637, 195 pages.

Whitford D (2005) *Proteins: Structure and Function*, 1st edn. John Wiley & Sons, Ltd. ISBN: 978-0-471–49894-0, 542 pages.