# On Enzyme Nomenclature and Classification

4

## 4.1 What Is in the Name?

The word *enzyme* (*ενζυμη* meaning *in yeast* in Greek), first used by Kuhne in 1877, is now well accepted to describe a biological catalyst. Majority of enzyme names today carry the suffix "-ase" as recommended for all enzyme names by Duclaux in 1898. Proteolytic enzymes are a significant exception to this generally accepted norm. Some of them have retained the older tradition of usually ending with "-in," for example, trypsin, chymotrypsin, papain, and subtilisin.

All enzymes are proteins but not all proteins are enzymes. Catalysis by RNA molecules (the so-called ribozyme) has expanded the realm of biological catalysts to beyond proteins. While few RNA catalysts have been recognized, the vast majority of enzymes that we come across in biology are proteins. Evolution has selected L-isomers of 20 amino acids to build proteins. This has put limits on the number and nature of available reactive chemical groups/functions that could be recruited for catalysis. Proteins are rich in nucleophilic groups, but electrophiles are poorly represented. In some instances, the protein component alone is inadequate to catalyze a given reaction. Nature therefore recruited many nonprotein components called *cofactors* to generate a functional catalyst. In such enzymes, the inactive protein component without the cofactor is termed the *apoenzyme* and the active enzyme, including the cofactor, the *holoenzyme*. The cofactors may be either metal ions (e.g., Mn[II] in arginase, Ni[II] in urease, and Ca[II] in DNase I) or coenzymes (organic molecules like pyridine nucleotides $NAD^+$ and $NADP^+$, flavin adenine dinucleotide (FAD), pyridoxal phosphate (PLP), thiamine pyrophosphate (TPP), biotin, cobamide, and heme). Binding of a cofactor to its cognate apoenzyme could exhibit a range of binding strength. A very tightly bound cofactor – which is difficult to remove without damaging the enzyme – is also known as a *prosthetic group*. Quite often prosthetic groups are covalently bound to the apoenzyme – lipoamide of transacylase is an example.



N. S. Punekar, *ENZYMES: Catalysis, Kinetics and Mechanisms*,
https://doi.org/10.1007/978-981-13-0785-0_4

$$\text{Holoenzyme} = \text{Apoenzyme} + \text{Cofactor}$$

Non-covalent interactions between a protein (such as an apoenzyme) and a cofactor may be weak or strong. This is amply obvious from the list of enzymes that require divalent metal ions for activity. For instance, enzymes with a tightly bound metal ion are termed *metalloenzymes* (such as urease), while those with a weakly bound metal ion are grouped as *metal-activated enzymes* (such as Fe [II]-catechol dioxygenase). Enzymes with metal ion dissociation constants ($K_D$) of the order of $10^{-8}$ M or higher are generally grouped as metal-activated enzymes, while those with $K_D$ values lower than $10^{-8}$ M are considered metalloenzymes. This boundary is artificial, and obviously there is a continuum of binding strengths observed in nature. It may be noted in passing that even non-covalent interactions can be very strong – essentially irreversible in few cases like avidin–biotin complex ($K_D = 10^{-15}$ M; $t_{1/2}$ of 2.5 years) or for that matter the two strands of a double-stranded DNA!

## 4.2 Enzyme Diversity and Need for Systematics

The middle of the twentieth century saw an exponential increase in research on enzymes. Sooner than later the number of new enzymes reported crossed manageable limits for an individual. As a consequence, in some cases, the same (or similar) enzyme activities were given different names. Some names like catalase give very little indication of the nature of the reaction they catalyze. Systematic classification, cataloging, and nomenclature of enzymes therefore became a necessity. This was easier recognized than done. Enzymes could be grouped according to any of the following considerations:

(a) Occurrence and/or source of the enzyme: Laccase obtained from Japanese lacquer tree, papain from papaya, and horseradish peroxidase are examples of three such plant enzymes. Similarly a number of digestive enzymes are isolated from pancreatic juice such as trypsin, chymotrypsin, carboxypeptidase, lipase, etc. the common source of lysozyme is from hen egg white (HEW).
(b) Nature of the substrate on which the enzyme acts: They could be classified into enzymes hydrolyzing (or acting on) proteins, carbohydrates, lipids, etc.
(c) Based on cofactor requirement: Typically many enzymes are simply proteinaceous in nature. However those depending on a cofactor could be listed into separate groups like thiamine pyrophosphate (TPP) enzymes, pyridoxal phosphate (PLP) enzymes, metalloenzymes, etc.
(d) Common functional context: One could, in principle, group enzymes belonging to discrete pathways like glycolytic enzymes, enzymes of histidine biosynthesis, etc. they may also be grouped as soluble, membrane-bound, or belonging to organelles such as mitochondria, etc.

(e) Nature of the overall reaction catalyzed: An enzyme can be assigned to a group by considering the type of the reaction it catalyzes. For instance, they may catalyze oxidation, hydrolysis, etc.
(f) The mechanism of reaction: The intimate mechanism of the reaction at the enzyme active site and the nature of intermediate complexes with the enzyme may be considered. For example, proteases may be classified depending upon whether an enzyme-bound acyl-enzyme intermediate is formed or not.

It should be obvious from the above list that a systematic, meaningful classification and cataloging of all the enzymes has not been easy. The problem is compounded by the enormous diversity of enzyme structures and activities. A typical RNA hydrolysis is achieved through a protein (RNase A), a protein-RNA complex (RNase P) or RNA alone (ribozyme). The peptide bond hydrolysis is possible with enzymes that are efficient in acidic pH or alkaline pH, require a divalent metal ion, contain a serine -OH or cysteine -SH, etc. Enzymatic decarboxylation of histidine may recruit pyridoxal phosphate or in a more primitive form may simply use a bound pyruvate. Pyridoxal phosphate bound to glycogen phosphorylase serves more of a structural role rather than function as a cofactor. Alkyldihydroxyacetonephosphate synthase (an enzyme involved in the biosynthesis of ether phospholipids) uses FAD for a non-redox reaction. Clearly, no single criterion listed above would be satisfactory. To address these issues, an international commission was set up in 1955 which presented its first report in 1961 (Table 2.1).

## 4.3 Enzyme Commission: Recommendations

Considering the diversity of enzyme sources, reactions, and mechanisms, it became apparent that a formal system of nomenclature and classification was required. "Enzyme Commission" was appointed by the International Union of Biochemistry to address this issue. Its first report, published in 1964, forms the basis of present system of classification. This system of classification is being updated periodically with updates made in 1972, 1978, 1984, and 1992. There are also many electronic supplements such as Supplement 14 of 2008 (Enzyme Nomenclature 1992). The most recent information and guidelines on enzyme nomenclature may be found at the official web site of the International Union of Biochemistry and Molecular Biology (IUBMB) – http://www.chem.qmul.ac.uk/iubmb/enzyme/. The EC classification is universally accepted with a unique name and EC number for each enzyme. By this means names for every enzyme (listed with trivial name) could be rationalized and also given an EC catalog number (McDonald et al. 2009).

The nature of the overall reaction catalyzed by the enzyme – expressed by the formal equation – forms the basis of EC classification. Clearly the intimate mechanism of the reaction and the formation of intermediate complexes with the enzyme, if any, are not considered. The Enzyme Commission defined six general categories of reactions, thereby assigning an enzyme to one of the six classes. A number of unique

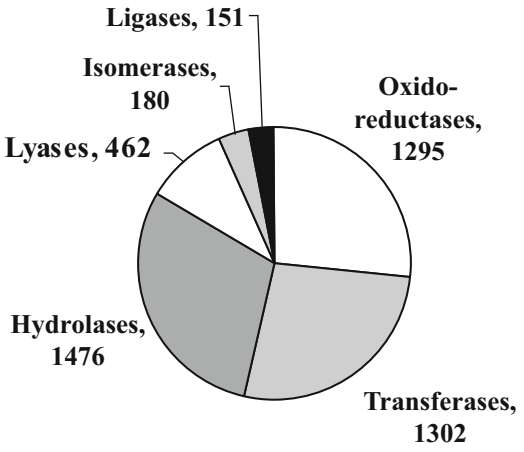


**Fig. 4.1 Distribution of enzymes into six different classes according to EC classification**. Data from BRENDA database (January 2009 release) is shown schematically

enzymes, represented in each of these six classes (BRENDA database release January 2009), are shown in Fig. 4.1. Expectedly, enzymes for the redox and hydrolytic reactions are the most represented group. These six classes are oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases. Each of these six classes is further divided into a number of subclasses and sub-subclasses, according to the nature of the reaction catalyzed. In this classification each enzyme is given a code number consisting of a four-number system. On this system the first number indicates the main class and the second and third show the subclass and sub-subclass, respectively, thus defining the type of reaction. The fourth number is the actual number of that enzyme within its sub-subclass. For example, alcohol dehydrogenase is given the code "EC 1.1.1.1." The first number indicates that it belongs to oxidoreductase class (EC 1.x.x.x). Within this class, enzymes acting on CH-OH group of donors bear the same subclass number (EC 1.1.x.x). Within this subclass, enzymes that use NAD or NADP as electron acceptor are given the number EC 1.1.1.x. Since alcohol dehydrogenase is the first enzyme in this category, it gets its fourth number (EC 1.1.1.1). Similarly, all carboxylesterases have the same first three digits in their EC code (EC 3.1.1.x). The fourth digit however distinguishes them by the actual carboxylic ester they hydrolyze.

Systematic name is assigned to each enzyme by the Commission, in addition to an accepted trivial name. This name includes the name or names of substrates followed by a reaction name that ends in "-ase." Because such systematic names can at times be too long and unwieldy, the Commission has also made recommendations for the use of trivial names. However, for the group (EC 3.3.3.x) of common proteases like pepsin, trypsin, papain, etc., it has not yet been possible to find acceptable systematic names. Enzyme Commission nomenclatures for enzymes representative of each class and those enzymes commonly referred to in this book are given in Table 4.1.

The universally accepted EC classification and enzyme codes are finding place (and direct utility) in a number of databases describing enzymes, genes, genomes, and metabolic pathways. Some of these databases are listed in Table 4.2.

**Table 4.1** Enzyme commission nomenclatures for representative/common enzymes

| EC No. | Systematic name | Trivial name | Reaction |
|---|---|---|---|
| 1. *Oxidoreductases: Loss or gain of electrons by substrates* | | | |
| 1.1.1.1 | Alcohol:$NAD^+$ oxidoreductase | Alcohol dehydrogenase | Alcohol+$NAD^+$⇄ Aldehyde+NADH+$H^+$ |
| 1.1.1.27 | L-lactate:$NAD^+$ oxidoreductase | Lactate dehydrogenase | L-lactate+$NAD^+$⇄ Pyruvate+NADH+$H^+$ |
| 1.2.1.2 | Formate:$NAD^+$ oxidoreductase | Formate dehydrogenase | Formate+$NAD^+$⇄ $CO_2$ + NADH+$H^+$ |
| 1.4.1.4 | L-Glutamate:$NADP^+$ oxidoreductase (deaminating) | NADP-glutamate dehydrogenase | L-Glutamate+$H_2O$ + $NADP^+$⇄ 2-Oxoglutarate+$NH_3$ + NADPH +$H^+$ |
| 1.11.1.6 | Hydrogen peroxide:Hydrogen peroxide oxidoreductase | Catalase | $H_2O_2$+ $H_2O_2$⇄ $O_2$ + $2H_2O$ |
| 2. *Transferases: Transfer of a reactive group from a donor substrate to an acceptor substrate* | | | |
| 2.1.2.1 | L-Serine:Tetrahydrofolate 5,10-hydroxymethyl-transferase | Serine hydroxymethyltransferase | L-Serine+Tetrahydrofolate⇄ Glycine+5,10-Methylenetetrahydrofolate |
| 2.1.3.1 | Methylmalonyl-CoA: Pyruvate carboxytransferase | Methylmalonyl-CoA carboxytransferase | Methylmalonyl-CoA + Pyruvate⇄ Propionyl CoA + Oxaloacetate |
| 2.4.1.8 | Maltose:Orthophosphate glucosyltransferase | Maltose phosphorylase | Maltose+Orthophosphate⇄ β-D-Glucose 1-phosphate+D-Glucose |
| 2.6.1.1 | L-Aspartate:2-Oxoglutarate aminotransferase | Aspartate aminotransferase | L-Aspartate+2-Oxoglutarate⇄ Oxloacetate+L-Glutamate |
| 2.7. 1.1 | ATP:D-Hexose 6-phosphotransferase | Hexokinase | ATP + D-Hexose⇄ ADP + D-Hexose 6-phosphate |
| 2.7.7.16 | Ribonucleate pyrimidine-nucleotido-2′-transferase (cyclizing) | Ribonuclease | Transfers 3′-phosphate of a pyrimidine nucleotide residue of polynucleotide from 5′ position of the adjoining nucleotide to the 2′ position of the pyrimidine nucleotide itself, forms a cyclic nucleotide |
| 3. *Hydrolases: Introduction of elements of water into a substrate* | | | |
| 3.1.1.1 | Carboxylic ester hydrolase | Carboxylesterase | Carboxylic ester+$H_2O$⇄ Alcohol+Carboxylate |
| 3.1.3.1 | Orthophosphoric monoester phosphohydrolase | Alkaline phosphatase | Orthophosphoric monoester $H_2O$⇄ Alcohol+Orthophosphate |
| 3.2.1.1 | α-1,4-Glucan 4-glucanohydrolase | α-Amylase | Hydrolyzes α-1,4-glucan links in starch |
| 3.2.1.17 | Mucopeptide N-acetylmuramylhydrolase | Lysozyme | Hydrolyzes β-1,4-glucan links in peptidoglycan |
| 3.4.4.4 | Not possible yet | Trypsin | Hydrolyzes peptides, amides and esters of aromatic L-amino acids |

(continued)

**Table 4.1** (continued)

| EC No. | Systematic name | Trivial name | Reaction |
|---|---|---|---|
| 3.5.3.1 | L-Arginine amidinohydrolase | Arginase | L-Arginine+$H_2O$⇄ L-Ornithine+Urea |
| 3.7.1.1 | Oxaloacetate acetylhydrolase | Oxaloacetase | Oxaloacetate+$H_2O$⇄ Oxalate+Acetate |
| 4. *Lyases: Elimination of a group with double bond formation or addition of group to double bond* | | | |
| 4.1.1.1 | Pyruvate carboxy-lyase | Pyruvate decarboxylase | Pyruvate⇄ Acetaldehyde+$CO_2$ |
| 4.2.1.1 | Carbonate hydro-lyase | Carbonic anhydrase | $H_2CO_3$⇄ $CO_2$ + $H_2O$ |
| 4.3.1.1 | L-Aspartate ammonia-lyase | Aspartate ammonia-lyase | L-Aspartate$^+$⇄ Fumarate+$NH_3$ |
| 4.3.2.1 | L-Argininosuccinate arginine-lyase | Argininosuccinate lyase | L-Argininosuccinate⇄ Fumarate+L-arginine |
| 4.99.1.1 | Protoheme ferro-lyase | Ferrochelatase | Protoporphyrin+$Fe^{2+}$⇄ Protoheme+$2H^+$ |
| 5. *Isomerases: Intramolecular rearrangements* | | | |
| 5.1.1.4 | Proline racemase | Proline racemase | L-Proline⇄ D-Proline |
| 5.1.3.2 | UDP-glucose 4-epimerase | UDP-glucose epimerase | UDP-glucose⇄ UDP-galactose |
| 5.3.1.1 | D-Glyceraldehyde-3-phosphate ketol-isomerase | Triosephosphate isomerase | D-Glyceraldehyde 3-phosphate⇄ Dihydroxyacetone phosphate |
| 5.3.1.5 | D-xylose ketol-isomerase | Xylose isomerase | D-Xylose⇄ D-Xylulose |
| 5.4.99.2 | Methylmalonyl-CoA CoA-carbonylmutase | Methylmalonyl-CoA mutase | Methylmalonyl-CoA⇄ Succinyl-CoA |
| 6. *Ligases (synthetases): Joining of molecules by covalent bond formation* | | | |
| 6.1.1.1 | L-Tyrosine:tRNA ligase (AMP) | Tyrosyl-tRNA synthetase | ATP + L-Tyrosine+tRNA⇄ AMP + Pyrophosphate+L-Tyrosyl-tRNA |
| 6.3.1.1 | L-Aspartate:Ammonia ligase (AMP) | Asparagine synthetase | ATP + L-Aspartate+$NH_3$⇄ AMP + Pyrophosphate+L-Asparagine |
| 6.3.1.2 | L-Glutamate:Ammonia ligase (ADP) | Glutamine synthetase | ATP + L-Glutamate+$NH_3$⇄ ADP + Orthophosphate+L-Glutamine |
| 6.3.4.5 | L-Citrulline:L-Aspartate ligase (AMP) | Argininosuccinate synthetase | ATP + L-Citrulline+L-Aspartate⇄ AMP+ Pyrophosphate+L-Argininosuccinate |
| 6.4.1.2 | Acetyl-CoA: Carbon dioxide ligase (ADP) | Acetyl-CoA carboxylase | ATP + Pyruvate+$CO_2$ + $H_2O$⇄ ADP+ Orthophosphate+Malonyl-CoA |

Enzyme examples selected in this table are those often referred to in this book

**Table 4.2 Databases with enzyme EC numbers**

| Database | Website |
|---|---|
| BRENDA (The Comprehensive Enzyme Information System) | http://www.brenda-enzymes.info/ |
| ExPASy (Enzyme nomenclature Database) | http://www.expasy.ch/enzyme/ |
| KEGG (Kyoto Encyclopedia of Genes and Genomes) | http://www.genome.ad.jp/kegg/kegg2.html |
| MetaCyc (Pathway/Genome Databases) | http://metacyc.org/ and http://biocyc.org/ |
| SYSTERS (Protein Family Database) | http://systers.molgen.mpg.de/ |
| InterPro (database of protein families, domains, and functional sites) | http://www.ebi.ac.uk/interpro/ |
| Protein Mutant Database | http://pmd.ddbj.nig.ac.jp/ |
| BioCarta (Pathways of Life) | http://www.biocarta.com/ |
| ExplorEnz (The Enzyme Database) | https://www.enzyme-database.org/ |
| *IUBMB (Enzyme Nomenclature)* | http://www.chem.qmul.ac.uk/iubmb/enzyme/ |

## 4.4 Some Concerns

The EC system of classification and nomenclature was arrived from a broad consensus with clear emphasis on the total reaction in question. The systematic names in a given class may be based on a written reaction, even if only the reverse reaction is experimentally demonstrated. This has created some situations that are less than perfect. While all enzyme-catalyzed reactions are reversible in principle (at least micro-reversibility at the active site!), the reaction-based classification for the forward direction would not be the same as that for the reverse direction. This was recognized very early by JBS Haldane, and according to him, calculation from thermodynamic data shows that catalase may act in the direction of $H_2O_2$ synthesis only under an $O_2$ pressure of many billions of atmosphere! It would, therefore, be perverse, if logical, to describe it as water oxidase or for that matter the peroxidase as water dehydrogenase. To address such issues, the Commission has recommended that the more important direction of the overall reaction, from a biochemical view point, be used. It may be noted from Table 4.1 that reaction involving interconversion of NADH and $NAD^+$ is all written in the direction where $NAD^+$ is reduced by the other substrate. Also, when an overall reaction involves two types of reactions, then the second function is indicated in brackets. For instance, an oxidoreductase (decarboxylating) means the enzyme catalyzes an oxidation-reduction reaction in which one of the substrates is being decarboxylated.

Apart from issues related to the direction of overall reaction considered, a major concern is that the reaction mechanism is given less importance or completely ignored! Functionally distinct reactions are catalyzed by the following four enzymes:

| | |
|---|---|
| Fumarate + $H_2O$ ⇄ Malate | (Fumarate hydratase; EC 4.2.1.2) |
| Fumarate + $NH_3$ ⇄ Aspartate | (Aspartate ammonia-lyase; EC 4.3.1.1) |
| Fumarate + Arginine ⇄ Argininosuccinate | (Argininosuccinate lyase; EC 4.3.2.1) |
| Fumarate + AMP ⇄ Adenylosuccinate | (Adenylosuccinate lyase; EC 4.3.2.2) |

All four enzymes are part of an evolutionary super-family of proteins; mechanistically they add elements of H-X to fumarate, using Michael reaction. This is a typical case where EC classification matches both in terms of overall reaction and the reaction mechanism. While UDP-glucose epimerase – based on the overall reaction – rightfully belongs to "isomerases" (Class 5), its intimate reaction mechanism tells a different story. The enzyme has tightly bound $NAD^+$, and the reaction at C4 of the hexose involves redox chemistry. The reaction mechanism of glucosamine 6-phosphate isomerase is similar to that of ketose-aldose isomerase:

$$\text{D-Glucosamine 6-phosphate} + H_2O \rightleftarrows \text{D-Fructose 6-phosphate} + NH_3$$

It was therefore formerly grouped under Class 5 (with the code EC 5.3.1.10) as an isomerase. Subsequently this reaction was recognized as a C-N bond hydrolysis and reassigned as glucosamine 6-phosphate deaminase (EC 3.5.99.6).

**Multiple Enzyme Forms and Isozymes** At times we find that the same enzyme activity in an organism (or different organisms) is displayed by different protein forms. Enzymes that catalyze the same overall reaction but follow different mechanistic paths are not uncommon. Some examples of different protein forms are listed below (Table 4.3).

When these multiple molecular forms are coded by different but related genes (having different primary structure – amino acid sequence!), they are termed isozymes. Different isozyme forms of an enzyme are easily distinguished through their characteristic electrophoretic mobilities. The muscle and the heart forms of lactate dehydrogenase are the best studied examples of isozymes. All isozymes are examples of multiple forms of an enzyme, but all multiple forms need not be isozymes.

Multiple molecular forms of the same enzyme may also arise due to other reasons. They may occur as (a) interconvertible forms through covalent modifications (e.g., phosphorylation of glycogen phosphorylase and adenylylation of *E. coli* glutamine synthetase), (b) proteolytic variants (chymotrypsinogen and chymotrypsin), (c) different oligomeric states of the same monomer (e.g., bovine liver glutamate dehydrogenase and avian liver acetyl CoA carboxylase), and (d) distinct conformational states of the same enzyme protein (e.g., R and T states of aspartate carbamoyltransferase). Occurrence of multiple enzyme forms is often associated with their role in metabolic regulation.

**Table 4.3 Enzyme examples with different mechanistic forms**

| Enzyme | EC number | Mechanistic difference |
|---|---|---|
| Ribonucleotide reductase | EC 1.17.4.1 | An iron protein |
| | EC 1.17.4.2 | Requires a cobamide coenzyme |
| Methionine synthase | EC 2.1.1.13 | Contains cobalamin |
| | EC 2.1.1.14 | Does not contain cobalamin |
| Proteases | EC 3.4.4.x | Serine, cysteine, carboxylate, or metal ion at active site |
| Histidine decarboxylase | EC 4.1.1.22 | Pyridoxal-phosphate (mammalian) |
| | | Pyruvoyl prosthetic group (bacterial) |
| Fructose bisphosphate aldolase | EC 4.1.2.13 | Class I – Forms a protonated imine |
| | | Class II – Zinc polarized (microbial) |
| Dehydroquinase | EC 4.2.1.10 | Type I – Forms a protonated imine |
| | | Type II – Forms an enolate intermediate |

All multiple enzyme forms share the same EC number and a formal name. However these names need to be suitably prefixed or suffixed to indicate the modification, organ source, or organelle source. Much effort goes into characterizing any new enzyme. Therefore great care must be taken to define what it does and not repeat the name or indicate another reaction. In the final outcome, one needs not get bogged down by issues of semantics, for the excitement of enzymology beckons the uninitiated and the specialist alike.

## References

McDonald AG, Boyce S, Tipton KF (2009) ExplorEnz: the primary source of the IUBMB enzyme list. Nucleic Acids Res 37:D593–D597

Enzyme Nomenclature 1992 [Academic Press, San Diego, California, ISBN 0-12-227164-5 (hardback), 0-12-227165-3 (paperback)] with Supplement 1 (1993), Supplement 2 (1994), Supplement 3 (1995), Supplement 4 (1997) and Supplement 5 (in Eur J Biochem, 223:1–5 (1994); Eur J Biochem, 232:1–6 (1995); Eur J Biochem, 237:1–5 (1996); Eur J Biochem, 250:1–6 (1997) and Eur J Biochem, 264:610–650 (1999); respectively), IUBMB, Academic Press, Orlando