the FEBS Journal

REVIEW ARTICLE

Fifty-five years of enzyme classification: advances and difficulties

Andrew G. McDonald and Keith F. Tipton

School of Biochemistry and Immunology, Trinity College, Dublin, Ireland

Keywords
EC number, enzyme database, ExplorEnz, IUBMB, IUPAC names

Correspondence
A. G. McDonald, School of Biochemistry and Immunology, Trinity College, Dublin 2, Ireland
Fax: +353 1 6772 400
Tel: +353 1 896 3959
E-mail: amcdonld@tcd.ie

(Received 22 July 2013, revised 10 September 2013, accepted 12 September 2013)

doi:10.1111/febs.12530

Since the publication of a list of enzymes classified according to the reactions that they catalysed, by Dixon and Webb in 1958, its content and presentation have undergone a number of significant changes. These have been necessitated by new information, as well as the need to improve clarity. The move from printed versions to the online environment, through the ExplorEnz website, has allowed the process of adding newly reported enzymes to be automated and the information content to be enriched. Search and output facilities have also been enhanced. These and the problems attendant on the use of the Enzyme Commission classification system for some groups of enzymes are the subject of this review.

Introduction

The year 2013 is the 100th anniversary of the seminal paper describing the Michaelis–Menten equation [1], which has recently appeared in an English translation [2], and also the 110th anniversary of Victor Henri's proposal of the reaction scheme on which that equation was based [3], but there are also several other anniversaries that have relevance to enzyme studies. These include the 90th anniversary of the Hartridge and Roughton [4] rapid flow apparatus, and the 70th anniversary of the stopped-flow studies on peroxidase by Chance [5]. It is 55 years since Koshland introduced the theory of induced fit [6], and 50 years since the use of the term allosteric in the title of a paper by Monod *et al.* [7]. Cleland's [8] papers on kinetic nomenclature were published 50 years ago, and the first publication of the Enzyme List by Dixon and Webb [9] appeared 55 years ago. The earlier history of the Enzyme List has been reviewed elsewhere [10], and this review will concentrate on more recent developments and applications.

Naming and classifying enzymes

The naming of enzymes has often been a haphazard and confusing affair. The names used often meant little or were ambiguous. Several names were given to the same enzyme, and several different enzymes were given the same name, so that it became difficult to be sure what others were talking about. Often, the suffix '-ase' was added to the name of the substrate, as in the case of urease (the first enzyme to be crystallized), or else the name gave some indication of the reaction catalysed, e.g. glucose oxidase, but even that enzyme was also known by other names. Several enzymes were

Abbreviations
EC, Enzyme Commission; IUBMB, International Union of Biochemistry and Molecular Biology; IUPAC, International Union of Pure and Applied Chemistry; NC, Nomenclature Committee.

given names that were decidedly unhelpful, as shown in the examples in Table 1. The glycolytic enzyme that split fructose bisphosphate into glycerone phosphate plus D-glyceraldehyde 3-phosphate was often simply called aldolase, because it catalysed a reverse aldol condensation reaction. However, there are now 35 other aldolases that have been recognized. The clinical chemists developed their own shorthand names for the enzymes important to them, as shown in Table 1, sometimes making it difficult for others to understand their publications. The tendency to coin apparently meaningless names still persists. Some names are indicative of the species of origin; for example, barnase is an extracellular ribonuclease from *Bacillus amyloliquefaciens*, and sirtuin 1 from humans is sirtuin 6 from the mouse. More recently, often ambiguous gene names have been used to describe some enzymes.

Table 1. Some alternative enzyme names. Chemical pathology names are abbreviated names that are, or have been, used in clinical biochemistry (CP).

Other name(s)	CP name(s)	NC-IUBMB name
Aldolase		Fructose bisphosphate aldolase
Alkaline phenyl phosphatase	ALP	Alkaline phosphatase
AspT; AAT	ALT; AST; GOT	Aspartate transaminase
Citrulline phosphorylase	OCT	Ornithine carbamoyltransferase
DT-diaphorase		NAD(P)H dehydrogenase (quinone)
Galactowaldenase		UDP-glucose 4-epimerase
γ-Glutamyltranspeptidase	GGT	γ-Glutamyltransferase
Glycolaldehydetransferase	TKT	Transketolase
Malic enzyme		Malate dehydrogenase (decarboxylating)
Glycerophosphatase	ACP	Acid phosphatase
Guanase	GDS	Guanine deaminase
New yellow enzyme (NYE)		D-Amino acid oxidase
Notatin; glucose aerodehydrogenase	GOD	Glucose oxidase
Old yellow enzyme (OYE)		NADPH dehydrogenase
Phosphosaccharomutase	GPI	Glucose-6-phosphate isomerase
Phosphocreatine kinase	CK	Creatine kinase
Pyrazinamidase		Nicotinamidase
Pyruvate transaminase	ALT	Alanine transaminase
Rhodanese		Thiosulfate sulfurtransferase
Steapsin; triacetinase; tweenase	LPS	Triacylglycerol lipase
Sirtuin 1 (SIRT 1)		Histone deacetylase
Taka-amylase A	AMS	α-Amylase

Although such arcane names may help to keep one's subject area esoteric, the rapid increase in the number of enzymes being discovered necessitated the development of some method for naming them systematically. In the 1950s, two groups of enzymologists set about addressing this problem by classifying enzymes in terms of their function, rather than by their structures; an approach that was radically different from that used in other branches of nomenclature. Otto Hoffmann-Ostenhof classified enzymes using a system based on the number of molecules involved in the reaction [11]. He proposed the following three general classes of enzymes: (a) hydrolases, transferases, and oxidoreductases (reaction type A + B = C + D); (b) lyases and synthases (A = B + C); and (c) racemases (A = B).

Malcolm Dixon and Edwin Webb, who were compiling a list of all known enzymes for their influential book *Enzymes*, noted that, despite the relatively large number of enzymes, the number of types of reaction involved was quite small. Enzymes were classified into three broad groups according to the type of reaction catalysed: the hydrolysing enzymes, the transferring enzymes, and other enzymes. The last of these groups group included the enzymes catalysing synthetic reactions linked with the breakdown of ATP or GTP, the stereoisomerases, and the enzymes that added groups to double bonds. The oxidizing enzymes were thought of as transferring hydrogen to an acceptor, and were therefore categorized with the other transferring enzymes. The published list [9] contained 659 enzymes, which were numbered sequentially: hydrolysing enzymes (1–221); transferring enzymes (222–579); and other enzymes (580–659). Each of these main groups was further subdivided according to the nature of the substrates or type of reaction involved. This was the beginning of the current enzyme classification and nomenclature system, with enzymes divided into groups and subgroups according to the nature of the reaction catalysed. However, the combination of such a division with a consecutive numbering system would not allow for any enzyme subsequently identified to be assigned a meaningful number in its class unless all those in subsequent classes were renumbered.

An International Commission on Enzymes was established by the International Union of Biochemistry [now termed the International Union of Biochemistry and Molecular Biology (IUBMB)] in 1956 to address the problems of enzyme classification and nomenclature. This resulted in the formulation of the enzyme classification system, which was developed by the IUBMB through its Nomenclature Committee (NC-IUBMB) into the NC-IUBMB Enzyme List, which was published in a report in 1961 [12], and in

expanded form in the second edition of *Enzymes* in 1964 [13]. In this, the earlier system was extended by classifying enzymes into six groups according to the type of reaction catalysed, as shown in Fig. 1.

Each of these classes was further subdivided, with each enzyme being given a unique four-digit code, the Enzyme Commission (EC) number. The second digit (the subclass), generally contains information about the type of compound or group involved. For the oxidoreductases, the subclass indicates the type of group in the donor that undergoes oxidation (for example, 1.1 acts on the CH–OH group of donors, whereas 1.4 acts on the $CH–NH_2$ group of donors). The third digit, the sub-subclass, further specifies the type of reaction involved. 1×1 indicates that NAD^+ or $NADP^+$ is the acceptor, and 1×2 that cytochrome is the acceptor, etc. The fourth digit is a serial number that is used to identify the individual enzyme within a sub-subclass. Figure 2 illustrates the use of this system for the transferases, with EC 2.1 and EC 2.1.3 expanded to show the complete sub-subclasses. A list of the numbers for different enzyme classes can be found at http://www.enzyme-database.org/class.php.

The NC-IUBMB Enzyme List, or, to give it its full title, *Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes by the Reactions they Catalyse*, was updated in printed editions in 1965, 1972, 1978, 1984 and 1992, and also in the 3rd edition of Dixon and Webb [14], with Edwin Webb having overall responsibility. Meanwhile, new material was published in a series of supplements (see http://www.chem.qmul.ac.uk/iubmb/enzyme/). An insightful 'Personal Retrospective' on the development of enzyme nomenclature was published by Webb 20 years ago [15]. Although the 1992 printed version of the Enzyme List ran to over 850 pages, space limitations necessitated the removal of some data. For example, some of the alternative names for enzymes were removed from the Enzyme List prior to publication of that edition. This could cause problems for people looking for information about an enzyme if they knew it by some name other than those used in the Enzyme List. For example, the archaic name 'old yellow enzyme' was dropped, which proved to be unfortunate, as that name has recently returned to use in the literature. A restricted list of alternatives for reactant names was provided as a separate glossary.

CLASS	NAME	REACTION CATALYSED
1.	Oxidoreductases	$AH_2 + B = A + BH_2$
		or $AH_2 + B^+ = A + BH + H^+$
2.	Transferases	$AX + B = A + BX$
3.	Hydrolases	$A–B + H_2O = AH + BOH$
4.	Lyases	$A = B + X–Y = \begin{matrix} A– & B \\ \vert & \vert \\ X & Y \end{matrix}$
5.	Isomerases	$A = B$
6.	Ligases	$A + B + NTP = A–B + NDP + P$
		or $A + B + NTP = A–B + NMP + PP$

Fig. 1. Enzyme classes of the NC-IUBMB enzyme list.

Figure 3 shows how the number of recognized and classified enzyme has increased since the first published Enzyme List. This would clearly have caused difficulties for further printed versions. For example, a printout of the current Enzyme List data would take over 2000 A4 pages, and would soon become out of date. Space limitation is no longer a factor for online versions, and comprehensive search facilities can be incorporated to make finding information much easier.

Online versions

For over 10 years, G. P. Moss of Queen Mary University of London has made the Enzyme List available as a series of flat files at http://www.chem.qmul.ac.uk/iubmb/enzyme/. This has restricted search facilities and requires manual data input. ExplorEnz [16,17] was developed in order to facilitate curation of the enzyme nomenclature data, to maintain correct data formatting, and to facilitate its use by other databases. Options for both simple substring searching and more complex searches were also provided. The enzyme data and their associated literature references are stored in MySQL databases on a dedicated server, and are accessed through a web interface located at http://www.enzyme-database.org. This database is now the primary repository for all enzymes classified by the IUBMB, and is also reproduced by many other databases, including BRENDA [18], ExPASy [19], KEGG [20] and MetaCyc [21], as well as at the Queen Mary University of London Enzyme Nomenclature website. Associated sites [17] allow draft entries to be prepared and automated formatting of new and revised entries (including references) with date-stamping, global changes, etc.

Enzyme fields

Figure 4 shows an example of an entry from ExplorEnz to indicate the type of data contained and so that the individual fields can be considered.

Accepted name

In earlier versions of the list, the name used for an enzyme was termed the 'trivial name'. This was subsequently changed to 'recommended name', but, in the 1992 book, in order to save space, the name was simply

EC 1 [+] **Oxidoreductases**
EC 2 [–] **Transferases**
EC 2.1 [–] Transferring one-carbon groups
EC 2.1.1 [+] Methyltransferases
EC 2.1.2 [+] Hydroxymethyl-, formyl- and related transferases
EC 2.1.3 [–] Carboxy- and carbamoyltransferases
EC 2.1.3.1 methylmalonyl-CoA carboxytransferase
EC 2.1.3.2 aspartate carbamoyltransferase
EC 2.1.3.3 ornithine carbamoyltransferase
EC 2.1.3.4 malonyl-CoA carboxyltransferase
EC 2.1.3.5 oxamate carbamoyltransferase
EC 2.1.3.6 putrescine carbamoyltransferase
EC 2.1.3.7 3-hydroxymethylcephem carbamoyltransferase
EC 2.1.3.8 lysine carbamoyltransferase
EC 2.1.3.9 *N*-acetylornithine carbamoyltransferase
EC 2.1.3.10 malonyl-*S*-ACP:biotin-protein carboxyltransferase
EC 2.1.3.11 *N*-succinylornithine carbamoyltransferase
EC 2.1.4 [+] Amidinotransferases
EC 2.2 [+] Transferring aldehyde or ketonic groups
EC 2.3 [+] Acyltransferases
EC 2.4 [+] Glycosyltransferases
EC 2.5 [+] Transferring alkyl or aryl groups, other than methyl groups
EC 2.6 [+] Transferring nitrogenous groups
EC 2.7 [+] Transferring phosphorus-containing groups
EC 2.8 [+] Transferring sulfur-containing groups
EC 2.9 [+] Transferring selenium-containing groups
EC 2.10 [+] Transferring molybdenum- or tungsten-containing groups
EC 3 [+] **Hydrolases**
EC 4 [+] **Lyases**
EC 5 [+] **Isomerases**
EC 6 [+] **Ligases**

Fig. 2. The enzymes of EC 2, the transferases, with subclass EC 2.1 and sub-subclass EC 2.1.3 expanded to show their content.

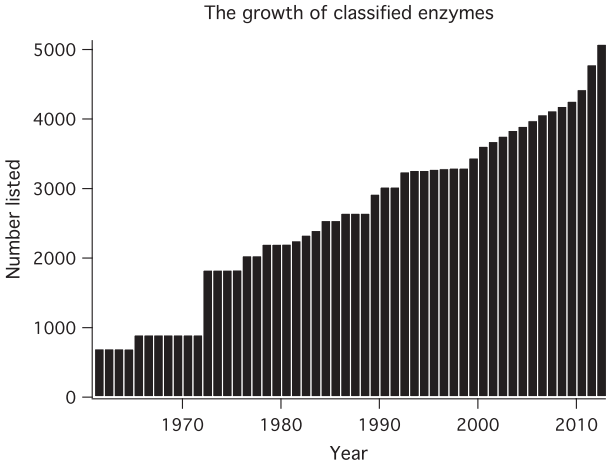
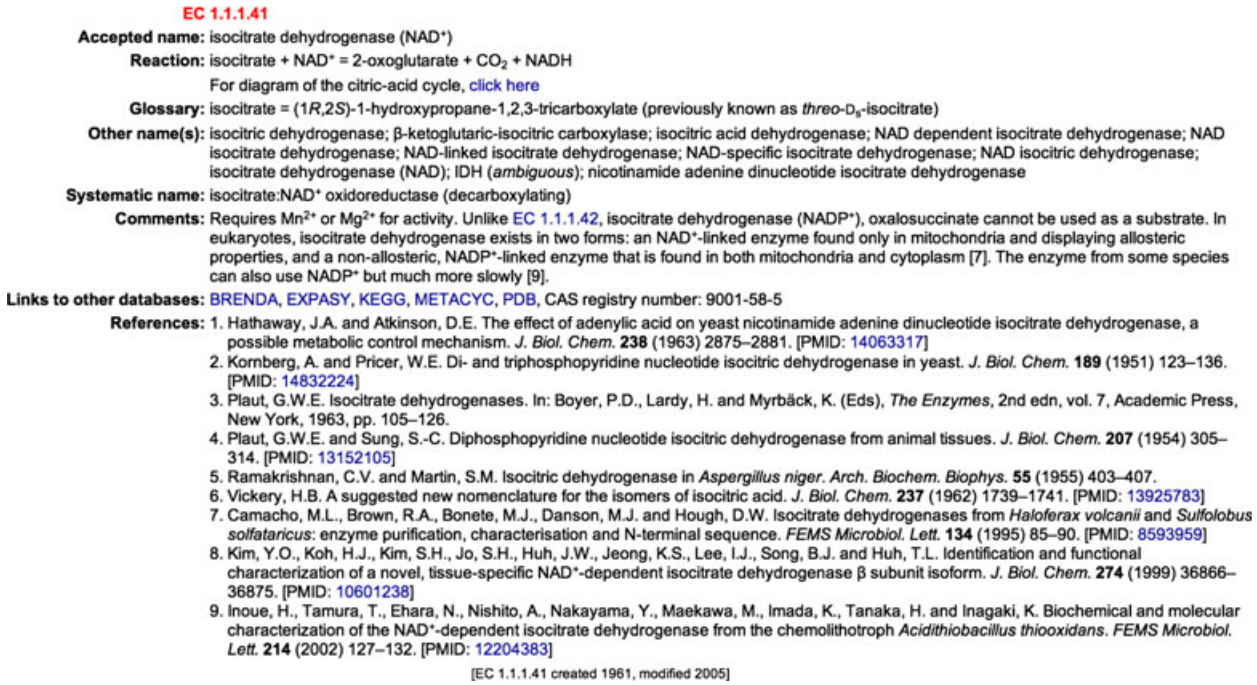


Fig. 3. The numbers of classified enzymes since the published list of 1961.

given after the EC number. The designation has now been changed to 'accepted name'. This is generally the most commonly used name for the enzyme, provided that it is neither misleading nor ambiguous. Such names may not always give any indication of what the enzyme catalyses. For example, the name catalase (EC 1.11.1.6) gives no indication that the enzyme breaks down hydrogen peroxide to oxygen plus water, but it was regarded to be too well established to suggest changing it. Sometimes, the accepted name is qualified by a parenthetic term to distinguish it from other enzymes catalysing similar reactions, such as isocitrate dehydrogenase (NAD^+) and isocitrate dehydrogenase ($NADP^+$).

Reaction

The reactions presented in the Enzyme List are, as far as possible, mass-balancing equations (that is, an equals sign is used rather than an arrow). They are not necessarily charge-balanced, as it would not be meaningful to do this for several reactants, such as phosphate and its derivatives, unless the pH and metal ion composition were also specified. The reactions are not meant to indicate the equilibrium position of the reac-



EC 1.1.1.41

Accepted name: isocitrate dehydrogenase (NAD$^+$)

Reaction: isocitrate + NAD$^+$ = 2-oxoglutarate + CO_2 + NADH

For diagram of the citric-acid cycle, click here

Glossary: isocitrate = (1*R*,2*S*)-1-hydroxypropane-1,2,3-tricarboxylate (previously known as *threo*-D_s-isocitrate)

Other name(s): isocitric dehydrogenase; β-ketoglutaric-isocitric carboxylase; isocitric acid dehydrogenase; NAD dependent isocitrate dehydrogenase; NAD isocitrate dehydrogenase; NAD-linked isocitrate dehydrogenase; NAD-specific isocitrate dehydrogenase; NAD isocitric dehydrogenase; isocitrate dehydrogenase (NAD); IDH (*ambiguous*); nicotinamide adenine dinucleotide isocitrate dehydrogenase

Systematic name: isocitrate:NAD$^+$ oxidoreductase (decarboxylating)

Comments: Requires Mn^{2+} or Mg^{2+} for activity. Unlike EC 1.1.1.42, isocitrate dehydrogenase (NADP$^+$), oxalosuccinate cannot be used as a substrate. In eukaryotes, isocitrate dehydrogenase exists in two forms: an NAD$^+$-linked enzyme found only in mitochondria and displaying allosteric properties, and a non-allosteric, NADP$^+$-linked enzyme that is found in both mitochondria and cytoplasm [7]. The enzyme from some species can also use NADP$^+$ but much more slowly [9].

Links to other databases: BRENDA, EXPASY, KEGG, METACYC, PDB, CAS registry number: 9001-58-5

References:

1. Hathaway, J.A. and Atkinson, D.E. The effect of adenylic acid on yeast nicotinamide adenine dinucleotide isocitrate dehydrogenase, a possible metabolic control mechanism. *J. Biol. Chem.* **238** (1963) 2875–2881. [PMID: 14063317]
2. Kornberg, A. and Pricer, W.E. Di- and triphosphopyridine nucleotide isocitric dehydrogenase in yeast. *J. Biol. Chem.* **189** (1951) 123–136. [PMID: 14832224]
3. Plaut, G.W.E. Isocitrate dehydrogenases. In: Boyer, P.D., Lardy, H. and Myrbäck, K. (Eds), *The Enzymes*, 2nd edn, vol. 7, Academic Press, New York, 1963, pp. 105–126.
4. Plaut, G.W.E. and Sung, S.-C. Diphosphopyridine nucleotide isocitric dehydrogenase from animal tissues. *J. Biol. Chem.* **207** (1954) 305–314. [PMID: 13152105]
5. Ramakrishnan, C.V. and Martin, S.M. Isocitric dehydrogenase in *Aspergillus niger*. *Arch. Biochem. Biophys.* **55** (1955) 403–407.
6. Vickery, H.B. A suggested new nomenclature for the isomers of isocitric acid. *J. Biol. Chem.* **237** (1962) 1739–1741. [PMID: 13925783]
7. Camacho, M.L., Brown, R.A., Bonete, M.J., Danson, M.J. and Hough, D.W. Isocitrate dehydrogenases from *Haloferax volcanii* and *Sulfolobus solfataricus*: enzyme purification, characterisation and N-terminal sequence. *FEMS Microbiol. Lett.* **134** (1995) 85–90. [PMID: 8593959]
8. Kim, Y.O., Koh, H.J., Kim, S.H., Jo, S.H., Huh, J.W., Jeong, K.S., Lee, I.J., Song, B.J. and Huh, T.L. Identification and functional characterization of a novel, tissue-specific NAD$^+$-dependent isocitrate dehydrogenase β subunit isoform. *J. Biol. Chem.* **274** (1999) 36866–36875. [PMID: 10601238]
9. Inoue, H., Tamura, T., Ehara, N., Nishito, A., Nakayama, Y., Maekawa, M., Imada, K., Tanaka, H. and Inagaki, K. Biochemical and molecular characterization of the NAD$^+$-dependent isocitrate dehydrogenase from the chemolithotroph *Acidithiobacillus thiooxidans*. *FEMS Microbiol. Lett.* **214** (2002) 127–132. [PMID: 12204383]

[EC 1.1.1.41 created 1961, modified 2005]

Fig. 4. Example of an entry from ExplorEnz to show the type of data contained.

tion or the direction in which the reaction is believed to operate *in vivo*. By convention, the direction chosen for the reaction in any given sub-subclass is the same for all enzymes, even if a reaction has only been observed in the reverse direction. Although this might seem to be less helpful than it could be, it should be remembered that, for some enzymes, such as glutamate dehydrogenase (EC 1.4.1.2) and fructose bisphosphate aldolase (EC 4.1.2.13), the preferred reaction direction varies with cellular conditions. Thermodynamic data for many enzymes can be found in the GTD Thermodynamics of Enzyme-catalysed Reactions database [22]. However, the equilibrium constant of the reaction may be misleading in terms of the direction in which it actually operates *in vivo*, as it is the thermodynamic properties of the overall metabolic system, not of the individual reaction, that are important in determining the flux direction. For example, the equilibrium for the oxidation of ethanol by alcohol dehydrogenase (EC 1.1.1.1).

$$CH_3CH_2OH + NAD^+ = CH_3CHO + NADH + H^+$$

greatly favours ethanol formation under physiological conditions, but ethanol oxidation is the dominant direction *in vivo*, because acetaldehyde (ethanal) produced is rapidly removed to form acetate in the essentially irreversible reaction catalysed by aldehyde dehydrogenase (NAD$^+$; EC 1.2.1.3). Accepted names do not necessarily indicate the direction in which the enzyme may operate; for example, the name dehydrogenase does not mean that the reaction always proceeds in the oxidative direction, but reductase may be used in cases where the reaction is known to overwhelmingly favour the reductive direction. The direction in which an enzyme is known to work in a specific metabolic pathway may be mentioned in the Comments.

General reactions

In the case of enzymes with broad substrate specificities where the number of substrates, or potential substrates, is large, the Enzyme List often gives a single generic reaction. For example, the reaction for D-amino acid oxidase (EC 1.4.3.3) is given as

$$\text{a D-amino acid} + H_2O + O_2 = \text{a 2-oxo carboxylate} + NH_3 + H_2O_2$$

The problem with this approach is that it does not indicate which D-amino acids are substrates and which are not. Other databases, such as BRENDA [18] and KEGG [20], may contain such information, although the absence of a compound from the sub-

strate/product list does not necessarily mean that it is not a substrate, but may simply mean that nobody has yet tried it.

Markush terms are also used for some general reactions. For example, the reaction catalysed by primary amine oxidase (EC 1.4.3.21) is given as

$$RCH_2NH_2 + H_2O + O_2 = RCHO + NH_3 + H_2O_2$$

and that of carbonyl reductase (NADPH; EC 1.1.1.184) is given as

$$R\text{-}CHOH\text{-}R' + NADP^+ = R\text{-}CO\text{-}R' + NADPH + H^+$$

Such formulations are somewhat more informative than the general reactions described above, and are searchable in ExplorEnz.

In some cases, it is not practicable to describe the reaction in a simple reaction equation. Examples of this include many reactions involving oligomeric substrates, such as the peptides, polynucleotides, and polysaccharides. For example, the reaction catalysed by α-amylase (EC 3.2.1.1) is given as

> Endohydrolysis of (1→4)-α-D-glucosidic linkages in polysaccharides containing three or more (1→4)-α-linked D-glucose units

and the reaction of exodeoxyribonuclease I (EC 3.1.11.1) is given as

> Exonucleolytic cleavage in the 3′- to 5′-direction to yield nucleoside 5′-phosphates.

Alternative reactions

Many enzymes are active towards more than one substrate. In such cases, the alternative reactions may be listed separately. For example, alcohol dehydrogenase (EC 1.1.1.1) oxidizes both primary and secondary alcohols, and the reaction is given as

(1) a primary alcohol + NAD^+ = an aldehyde + NADH + H^+
(2) a secondary alcohol + NAD^+ = a ketone + NADH + H^+

The enzyme taurochenodeoxycholate 6α-hydroxylase EC 1.14.13.97 is known to catalyse the hydroxylation of two biologically important bile acids, and its reaction is given as

(1) taurochenodeoxycholate + NADPH + H^+ + O_2 = taurohyocholate + $NADP^+$ + H_2O
(2) lithocholate + NADPH + H^+ + O_2 = hyodeoxycholate + $NADP^+$ + H_2O

In some cases, it is thought to be helpful to give the general reaction catalysed followed by specific reactions, as, for example, for monoterpene ε-lactone hydrolase (EC 2.1.1.70):

(1) isoprop(en)ylmethyloxepan-2-one + H_2O = 6-hydroxyisoprop(en)yl-methylhexanoate (general reaction)
(2) 4-isopropenyl-7-methyloxepan-2-one + H_2O = 6-hydroxy-3-isopropenylheptanoate
(3) 7-isopropyl-4-methyloxepan-2-one + H_2O = 6-hydroxy-3,7-dimethyloctanoate

Consecutive reactions

Enzyme classification is based on the overall reaction catalysed, and thus any intermediate products that are not released from the enzyme are not considered. However, there are cases where intermediate products are released. An example of this is tricin synthase (EC 2.1.1.175), which catalyses two successive methylation reactions. These are designated as (1a) and (1b) to distinguish them from alternative reactions:

2 *S*-adenosyl-L-methionine + tricetin = **2** *S*-adenosyl-L-homocysteine + 3′,5′-*O*-dimethyltricetin (overall reaction)
(1a) *S*-adenosyl-L-methionine + tricetin = *S*-adenosyl-L-homocysteine + 3′-*O*-methyltricetin
(1b) *S*-adenosyl-L-methionine + 3′-*O*-methyltricetin = *S*-adenosyl-L-homocysteine + 3′,5′-*O*-dimethyltricetin

Some enzymes can catalyse reactions in which a substrate for the subsequent reaction is generated by a preceding reaction catalysed by the same enzyme, as in the case of CTP synthase (glutamine hydrolysing; EC 6.3.4.2):

ATP + UTP + L-glutamine = ADP + phosphate + CTP + L-glutamate (overall reaction)

(1a) L-glutamine + H_2O = L-glutamate + NH_3
(1b) ATP + UTP + NH_3 = ADP + phosphate + CTP

In this case, the ammonia required for step (1b) is formed by glutamine hydrolysis in step (1a), but step (1b) can also occur in the presence of added ammonia.

In some cases, the immediate product of the reaction is unstable, and is rapidly converted to another compound by an uncatalysed reaction, such as hydrolysis or decarboxylation, as, for example, for putrescine aminotransferase (EC 2.6.1.82):

putrescine + 2-oxoglutarate = 1-pyrroline + L-glutamate + H_2O (overall reaction)
(1a) putrescine + 2-oxoglutarate = 4-aminobutanal + L-glutamate
(1b) 4-aminobutanal = 1-pyrroline + H_2O (spontaneous)

Multienzyme systems

Multienzyme complexes and single polypeptides that catalyse disparate activities are listed under their separate activities, with Comments that they exist as multienzyme complexes or multifunctional proteins in some species. For example, the multifunctional protein catalysing the synthesis of tryptophan is listed as its component activities (EC 2.4.2.18, EC 4.1.1.48, EC 4.2.1.20, EC 4.1.3.27, and EC 5.3.1.24), with a note in the Comments for each indicating that, in some organisms, the enzyme is part of a multifunctional protein, together with one or more of the other enzymes. This treatment has the advantages that it avoids uncertainties about where it should be classified and the fact that what is a multifunctional protein in some species may exist as separate components in another. Furthermore, it avoids the need for complex systematic names. Another advantage of listing the components separately is that the overall reaction catalysed by a multifunctional protein may obscure its actual functions. For example, the combined reaction of the multifunctional 6-phosphofructo-2-kinase (EC 2.7.1.105) and fructose-2,6-bisphosphate-2-phosphatase (EC 3.1.3.46) would simply be ATP hydrolysis.

Incomplete reactions

Although the reaction catalysed should be fully characterized before an enzyme is included in the Enzyme List, there are a few exceptions. For example, in the case of *trans*-feruloyl-CoA synthase (EC 6.2.1.34), the reaction is given as

ferulic acid + CoA + ATP = feruloyl-CoA + products of ATP breakdown

and the Comments state 'It has not yet been established whether AMP + diphosphate or ADP + phosphate are formed in this reaction'. It is intended to rectify this when further information becomes available. There is also a group of oxidoreductases that have only been studied with artificial acceptors, such as 2,6-dichloroindophenol or methylene blue. These are classified as EC 1.x.99.y (with other acceptors). An example is 3-hydroxycyclohexanone dehydrogenase (EC 1.1.99.26), where the reaction is given in the general form:

3-hydroxycyclohexanone + acceptor = cyclohexane-1,3-dione + reduced acceptor

Such entries are transferred to other subclasses when the nature of the physiological acceptor is known.

Diagrams

As shown in the example of Fig. 4, many of the enzyme entries are linked to diagrams that show the involvement of the enzyme in a metabolic pathway and/or its reaction mechanism. These diagrams were developed by H. B. F. Dixon and G. P. Moss, and are now being maintained and expanded by G. P. Moss.

Glossary

The Glossary is used to relate the common names of the compounds shown in the reaction field or accepted names with their International Union of Pure and Applied Chemistry (IUPAC) names, and any alternative names that may be used. This can be accessed separately at http://www.enzyme-database.org/glossary.php, where the entries are each linked to the ChemSpider database [23] to allow access to their structures and other chemical information. The NC-IUBMB works closely with the IUPAC on the Joint Committee on Biochemical Nomenclature to ensure that these names accord with the IUPAC system [24].

Other name(s)

As mentioned above, this section is intended to list the alternative names that are, or have been, used for each enzyme, The aim is to be inclusive, and, where other names are ambiguous, misleading, or incorrect, they are marked as being such. Gene names that have been used to describe an enzyme are also indicated as such.

Systematic name

This attempts to describe in unambiguous terms what reaction the enzyme actually catalyses. Systematic names consist of two parts. The first contains the name of the substrate or, in the case of a bimolecular reaction, of the two substrates separated by a colon. The second part, ending in '-ase', indicates the nature of the reaction. A number of generic words indicating a type of reaction may be used in either accepted or systematic names: oxidoreductase, oxygenase, transferase (with a prefix indicating the nature of the group transferred), hydrolase, lyase, racemase, epimerase, isomerase, mutase, or ligase. Where additional information is needed to make the

reaction clear, a phrase indicating the reaction or a product should be added in parentheses after the second part of the name, e.g. (ADP-forming), (dimerizing), and (CoA-acylating). As systematic names use, as far as practicable, the IUPAC names for compounds, they can be somewhat unwieldy; for example, the systematic name for dihydroneopterin aldolase (EC 4.1.2.25) is 2-amino-4-hydroxy-6-(D-*erythro*-1,2,3-trihydroxypropyl)-7,8-dihydropteridine glycolaldehyde-lyase (2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine-forming).

Comments

This field may contain information on the nature of the reaction catalysed, possible relationships with other enzymes, species differences, metal ion and cofactor requirements, metabolic role(s), etc. Where appropriate, the direction in which the reaction is known to operate in a specific metabolic pathway may be indicated.

Links to other databases

These allow direct access to sites that provide further information on the enzyme in question.

References

Key references on the identification, nature, properties and function of the enzyme are listed. Whereas the references in the published versions of the Enzyme List were given without titles, these are provided in the web version of Enzyme Nomenclature, where references are also linked to PubMed, so that the user can be easily directed to abstracts and related material.

History

The last field in an enzyme entry provides a history of the enzyme entry since its creation, e.g. 'EC 1.1.1.205 created 1961 as EC 1.2.1.14, transferred 1984 to EC 1.1.1.205'.

Conventions and procedures

Inclusion criteria

It is important to note that the characterization of an enzyme-catalysed reaction is a prerequisite for its inclusion in the Enzyme List. Direct experimental evidence is required that the proposed enzyme actually catalyses the reaction claimed; speculation is not acceptable. Close sequence similarity is not sufficient without evidence for the reaction catalysed, because only a small change in sequence is sufficient to change the activity or specificity of an enzyme. Furthermore, because classification is based solely on the reaction catalysed, there are cases where proteins of very different sequences catalyse the same reaction. The postulation that an enzyme must exist to fill an apparent gap in a biochemical pathway is not adequate without direct evidence for its existence. Artificially constructed catalysts or enzymes engineered to have altered specificities are, at present, not considered.

Adding new enzymes

Proposals for addition to the Enzyme List, or for the modification of existing entries, can be found, along with guidelines and more detailed classification rules, which will shortly be updated, at http://www.enzyme-database.org/faq.php. On receipt of a proposed new entry, it is considered by a subgroup of the NC-IUBMB to ensure that it meets the criteria for a new enzyme. The entry is then correctly formatted and, after consideration by the Joint Committee on Biochemical Nomenclature, it is made available for public review at http://www.enzyme-database.org/newenz.php. Any suggested changes are then incorporated before the entry is added to the existing Enzyme List.

Alterations and revisions

As more information becomes available, it may be necessary to alter the classification of a particular enzyme or even remove it from the Enzyme List. Several new subclasses and sub-subclasses have been added over the past few years, necessitating enzymes being moved. An important principle is that any given EC number is never reused. Thus, if an enzyme is found to have been incorrectly classified, or not to exist, the EC number is left unoccupied and recorded as a 'deleted or transferred entry'. One example of this concerns pancreatic ribonuclease. In some earlier editions of the Enzyme List, this was listed as a transferase (EC 2.7.7.16), as it 'transfers the 3′-phosphate of a pyrimidine nucleotide residue of a polynucleotide from the 5′-position of the adjoining nucleotide to the 2′-position of the pyrimidine nucleotide itself, forming a cyclic nucleotide'. It then spent some time as a phosphoric diester hydrolase (EC 3.1.4.22), before reaching its current resting place among the hydrolases (EC 3.1.27.5) as one of the 'endoribonucleases producing other than 5′-phosphomonoesters', as it catalyses 'Endonucleolytic cleavage to nucleoside 3′-phosphates and 3′-phosphooligonucleotides ending in Cp or Up with 2′,3′-cyclic phosphate intermediates'. The EC numbers 2.7.7.16 and 3.1.4.22

now merely contain a note that the entry for pancreatic ribonuclease has been transferred to EC 3.1.27.5. Whereas this current classification may be consistent with the system being based on the overall reaction catalysed, and perhaps with the way in which most biochemists think about the enzyme, it upsets several of those working on ribonucleases who believe that it should, by right, have remained as a transferase. In fact, inspection of the process indicates that the reaction is, in fact, a lyase reaction that produces the cyclic products that may be subsequently hydrolysed. This has not found favour with some of those working in the field, and to date we have taken the majority view and left it as a hydrolase.

The criteria for removing an enzyme from the list are stringent. There must be positive evidence that the entry is incorrect, such as the reported product having been incorrectly identified or the reported activity being attributable to the combined activities of two separate enzymes. It is not sufficient that the enzyme was reported many years ago and that there have been no more recent publications on it. This is reflected in the relatively small number of deleted entries shown in Table 2.

Problems and limitations

The aims of the NC-IUMB enzyme nomenclature system are quite modest: to provide, as far as possible, correct and unambiguous information on enzymes classified according to the reaction(s) catalysed. As such, it is intended to be of use to other researchers and databases. Protein structure information is available through the linked databases, and even information on cofactor requirements is often given in somewhat legalistic form. For example, a phrase such as 'the enzyme from *E. coli* requires Mg^{2+}' leaves open the possibility that an enzyme catalysing the same reaction from another source may not require that metal ion. This is, for example, illustrated by the entry for fructose bisphosphate aldolase (EC 4.1.2.13), which represents enzymes that operate by very different chemical mechanisms. This is indicated by the Comments, which state, 'The yeast and bacterial enzymes are zinc proteins. The enzymes increase electron attraction by the carbonyl group, some (Class I) forming a protonated imine with it, others (Class II), mainly of microbial origin, polarizing it with a metal ion, e.g. zinc'. Other sources, such as the MACIE database [25], should be consulted for more detailed information on reaction mechanisms. Data on the kinetic parameters of individual enzymes, which depend on conditions, are also outside the scope of the Enzyme List, but may be found from other sources, such as BRENDA [18] and SABIO-RK [26].

Table 2. Enzyme count per class in the NC-IUBMB list.

Class	1	2	3	4	5	6	All
Current	1440	1505	1249	547	231	162	5134
Transferred	198	68	279	66	7	4	622
Deleted	75	66	103	29	8	6	287
Total	1713	1639	1631	642	246	172	6043

The functional classification may not readily accommodate isoenzymes. For example, there are several isoenzymes of alcohol dehydrogenase in human liver. However, as most oxidize primary alcohols and have a strong preference for NAD^+ as the coenzyme, they are all grouped together under the general heading of EC 1.1.1.1. Similarly, species differences in protein structure are only taken into account when they result in significantly different substrate specificities. Although the number of alternative reactions catalysed by some enzymes is being expanded, the Enzyme List does not aim to be exhaustive in this respect. It is possible to trace some metabolic pathways by use of the enzyme list [27], but the fact that not all substrates for an enzyme are listed is a limiting factor. However, there are many other reaction or reactant databases that may be used to fill any gaps, such as BRENDA [18], the R-Pair system in KEGG REACTION [20], MetaCyc [21], Rhea [28], and UM-BBD [29].

Enzymes hydrolysing peptide bonds in proteins and oligopeptides (the peptidases; EC 3.4.-.-) represent a group of enzymes that do not fit in with the enzyme classification system, as all of these can be regarded as catalysing the same reaction: hydrolytic cleavage of a peptide bond. Furthermore, it is often difficult to define their specificity, as it depends on the nature of several amino acids around the peptide bond to be hydrolysed, and the conformation of the substrate polypeptide chain. There are, for example, several enzymes with different names that hydrolyse peptide bonds on the *C*-terminal side of arginine or lysine residues, which, under the rules, should all be classified as the same enzyme. Early on, it was decided to break the rules and to divide the endopeptidases, for example, into sub-subclasses according to catalytic mechanism: serine endopeptidases (EC 3.4.21), cysteine endopeptidases (EC 3.4.22), aspartic endopeptidases (EC 3.4.23), metalloendopeptidases (EC 3.4.24), etc. However, even this approach has proven inadequate to cope with the large number of known peptidases, some of which differ only in their sensitivities towards activators, such as calcium ions. Because of that, this group remains inadequately covered by the Enzyme

List, but a fuller coverage can be found at the MEROPS database [30], which does not operate under the same constraints.

Conclusions

The Enzyme List and its operation have undergone a number of changes in division and operation to adapt to new information and procedures, but its basic design and the use of the information-rich EC number as a simple identifier persists. Despite the limitations discussed above, it is a tribute to the insights and perseverance of Edwin Webb and Malcolm Dixon that this system, which they played key roles in introducing, remains in widespread use over 50 years later.

Acknowledgements

This work was partially supported by Science Foundation Ireland (Grant No. SFI 02/IN.1/B043-Tipton).

References

1 Michaelis L & Menten ML (1913) Die kinetik der invertinwirkung. *Biochem Z* **49**, 333–360.
2 Goody RS & Johnson KA (2011) The original Michaelis constant: translation of the 1913 Michaelis–Menten Paper. *Biochemistry* **50**, 8264–8269.
3 Henri V (1903) Lois générales de l'action des diastases. Hermann, Paris.
4 Hartridge H & Roughton FJW (1923) A method of measuring the velocity of very rapid chemical reactions. *Proc R Soc Lond A* **104**, 376–394.
5 Chance B (1943) The kinetics of the enzyme–substrate compound of peroxidase. *J Biol Chem* **151**, 533–577.
6 Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* **44**, 98–104.
7 Monod J, Changeux J-P & Jacob F (1963) Allosteric proteins and cellular control systems. *J Mol Biol* **6**, 306–329.
8 Cleland WW (1963) The kinetics of enzyme-catalyzed reactions with two or more substrates or products. I. Nomenclature and rate equations. *Biochim Biophys Acta* **867**, 104–137.
9 Dixon M & Webb EC (1958) Enzymes. Longmans Green, London, and Academic Press, New York.
10 Tipton KF & Boyce S (2000) History of the enzyme nomenclature system. *Bioinformatics* **16**, 34–40.
11 Hoffman-Ostenhof O (1953) Suggestions for a more rational classification and nomenclature of enzymes. *Adv Enzymol* **14**, 219–220.
12 Report of the Comission on Enzymes of the International Union of Biochemistry (1961) IUB Symposium Series, Vol. 20. Pergamon Press, New York and Oxford.
13 Dixon M & Webb EC (1964) Enzymes, 2nd edn. Longmans Green, London, and Academic Press, New York.
14 Dixon M, Webb EC, Thorne CJR & Tipton KF (1979) Enzymes, 3rd edn. Longmans Green, London, and Academic Press, New York.
15 Webb EC (1993) Enzyme nomenclature: a personal retrospective. *FASEB J* **7**, 1192–1194.
16 McDonald AG, Boyce S, Moss GP, Dixon HBF & Tipton KF (2007) ExplorEnz: a MySQL database of the IUBMB enzyme nomenclature. *BMC Biochem* **8**, 14.
17 McDonald AG, Boyce S & Tipton KF (2009) ExplorEnz: the primary source of the IUBMB enzyme list. *Nucleic Acids Res* **37**, D593–D597.
18 BRENDA – The Comprehensive Enzyme Information System. http://www.brenda-enzymes.org/.
19 ExPASy – Enzyme Nomenclature Database. http://www.expasy.ch/enzyme/.
20 KEGG – Kyoto Encyclopedia of Genes and Genomes. http://www.genome.jp/kegg/.
21 MetaCyc Database of Metabolic Pathways and Enzymes. http://metacyc.org/.
22 GTD – Database of Thermodynamic Parameters of Enzyme Catalyzed Reactions. http://xpdb.nist.gov/enzyme_thermodynamics/.
23 ChemSpider – the Free Chemical Structure Database. http://www.chemspider.com/.
24 Panico R, Powell WH & Richer J-C (1994) A Guide to IUPAC Nomenclature of Organic Compounds. Blackwell Science, Oxford.
25 MACIE – Mechanism, Annotation and Classification in Enzymes Database. http://www.ebi.ac.uk/thornton-srv/databases/MACiE/.
26 SBAIO-RK Biochemical Reaction Kinetics Database. http://sabio.villa-bosch.de/.
27 McDonald AG, Tipton KF & Boyce S (2009) Tracing metabolic pathways from enzyme data. *Biochim Biophys Acta* **1794**, 1364–1371.
28 Rhea – Database of Chemical Reactions. http://www.ebi.ac.uk/rhea/.
29 University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD). http://umbbd.ethz.ch/.
30 MEROPS – the Peptidase Database. http://merops.sanger.ac.uk/.