EC Numbers Are Everywhere

Assignment of EC Numbers to Enzymatic Reactions with MOLMAP Reaction Descriptors and Random Forests

Diego A. R. S. Ladeão and João Alves-da-Souza

Computational Bioinformatics, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

Received March 17, 2008

---

EC Bioinformatics

Open Access

---

SRI International Bioinformatics

---

help!!!

they're everywhere!!!

---

Three-field EC numbers

Four-field EC numbers

---

SRI International Bioinformatics
Historical Background

Back in the 1950s

- The number of known enzymes was increasing rapidly
- No guiding authority
- The same enzymes became known by several different names, and
- The same name was sometimes given to different enzymes
- Names often conveyed little or no idea of the nature of the reactions catalyzed
The Situation Was Chaotic…

- Catalase (also known as equilase, caperase, optidase…)
- Diaphorase (dehydrogenase)
- Zwischenferment (glucose-6-phosphate dehydrogenase)
The First Enzyme Commission

In August 1955 M. Dixon and O. Hoffmann-Ostenhof convinced the president of the International Union of Biochemistry (IUB) to set up an International Enzyme Commission to tackle the problems

Members included:

- M. Dixon, U.K. (president)
- A.E. Braunstein, U.S.S.R.
- S.P. Colowick, U.S.A
- P.A.E. Desnuelle, France
- V.A. Engelhardt, U.S.S.R
- E.F. Gale, U.K
- O. Hoffmann-Ostenhof, Austria
- A.L. Lehninger, U.S.A.
- (K. Linderstrom-Lang, Denmark) E.C. Webb, UK
- F. Lynen, Germany
The Basic Concept

Enzymes are classified and named by the reactions they catalyze.
The Reports of the First and Second Commissions

- The first EC list was presented in 1961 at the General Assembly of the IUB in Moscow
- Introduction of the Fundamental Concepts for classifications (to be discussed soon)
  712 entries
- Following this publication, the commission was dissolved, and the Standing Committee on Enzymes (only 4 of the original members) formed
- Published the second version in 1964 - 875 entries
The Expert Committee on Enzymes

- Formed in 1969 to revise the list
- Published the third document in 1972 - 1770 entries

Members:
- A.E. Braunstein, U.S.S.R.
- J.S. Fruton, USA
- O. Hoffmann-Ostenhof, Austria
- B.L. Horecker, USA
- W.B. Jakoby, USA
- P. Karlson, Germany
- B. Keil, France
- E.C. Slater, Holland
- E.C. Webb, United Kingdom
- W.J. Whelan, Australia
1977: Move to NC-IUB

- A more permanent solution was needed

- In 1977 two new nomenclature committees were formed:
  - The Nomenclature Committee of IUB (NC-IUB)
  - The IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN)

- NC-IUB (now NC-IUBMB) assumed responsibility for the EC list

- The 1978 the 4th EC list was published with 2122 entries
Current Status

- Ongoing curation by the NC-IUBMB since 1977

- Transition from print to online content
  - A few supplements were published in Eur. J. Biochem (up to 1999)
  - All newer data is only available electronically. Currently there are 4314 entries

Current active full members:
- K.F. Tipton, Ireland (Trinity College Dublin)
- R. Cammack, UK (King's College London)
- G.P. Moss, UK (Queen Mary University of London)
- D. Schomburg, Germany (chairman) (BRENDA)

Active associate members:
- A. McDonald, Ireland (Trinity College Dublin) – computer support
- K. Axelsen, Denmark (UniProt)
- R. Caspi, USA (MetaCyc)
- I. Schomburg, Germany (BRENDA)

Curator (at BRENDA):
- C. Munaretto
DraftEnz is a MySQL database developed by Andrew McDonald from Trinity College that permits EC curators to enter, edit, and review enzyme entries.

Following initial curation in DraftEnz, each entry goes through a few weeks of private review and a month of public review in CurrEnz.
The EC Number

Each enzyme is given a unique four-digit code, known as the Enzyme Commission, or EC, number.
# The Six Main Classes of Enzymes

<table>
<thead>
<tr>
<th>Class</th>
<th>Name</th>
<th>Reaction catalyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Oxidoreductases</td>
<td>$AH_2 + B = A + BH_2$ or $AH_2 + B+ = A + BH + H^+$</td>
</tr>
<tr>
<td>2.</td>
<td>Transferases</td>
<td>$AX + B = A + BX$</td>
</tr>
<tr>
<td>3.</td>
<td>Hydrolases</td>
<td>$A–B + H_2O = AH + BOH$</td>
</tr>
<tr>
<td>4.</td>
<td>Lyases</td>
<td>$A–B + X–Y = A–B$</td>
</tr>
<tr>
<td>5.</td>
<td>Isomerasees</td>
<td>$A = B$</td>
</tr>
<tr>
<td>6.</td>
<td>Ligases</td>
<td>$A + B + \text{NTP} = A–B + \text{NDP} + P$ or $A + B + \text{NTP} = A–B + \text{NMP} + \text{PP}$</td>
</tr>
</tbody>
</table>
Sub Classes and Sub-Subclasses

- Each of the six main classes is further subdivided
- The subclass generally contains information about the type of compound or group involved

EC 1.1.1.1

(e.g. 1.1. acts on the CH–OH group of donors whereas 1.3. acts on the CH–CH group of donors)

- The sub-subclass further specifies the type of reaction involved. (e.g. for the oxidoreductases, 1.-.1. indicates that NAD or NADP is the acceptor, 1.-.2. has cytochrome as the acceptor, etc)

EC 1.1.1.1

- The fourth digit is a serial number that is used to identify the individual enzymes within a sub-subclass
Sub Classes of Class 1

EC 1
- Oxidoreductases
  - EC 1.1
    - Acting on the CH-OH group of donors
  - EC 1.2
    - Acting on the aldehyde or oxo group of donors
  - EC 1.3
    - Acting on the CH-CH group of donors
  - EC 1.4
    - Acting on the CH-NH₂ group of donors
  - EC 1.5
    - Acting on the CH-NH group of donors
  - EC 1.6
    - Acting on NADH or NADPH
  - EC 1.7
    - Acting on other nitrogenous compounds as donors
  - EC 1.8
    - Acting on a sulfur group of donors
  - EC 1.9
    - Acting on a heter group of donors
  - EC 1.10
    - Acting on dihydroxyfurans and related substances as donors
  - EC 1.11
    - Acting on a peroxide as acceptor
  - EC 1.12
    - Acting on hydrogen as donor
  - EC 1.13
    - Acting on single donors with O₂ as oxidant and incorporation of oxygen into the substrate (oxygenases). The oxygen incorporated need not be derived from O₂
  - EC 1.14
    - Acting on paired donors, with O₂ as oxidant and incorporation or reduction of oxygen. The oxygen incorporated need not be derived from O₂
  - EC 1.15
    - Acting on superoxide as acceptor
  - EC 1.16
    - Oxidizing metal ions
  - EC 1.17
    - Acting on CH or CH₂ groups
  - EC 1.18
    - Acting on iron-sulfur proteins as donors
  - EC 1.19
    - Acting on reduced flavodoxin as donor
  - EC 1.20
    - Acting on phosphorus or arsenic in donors
  - EC 1.21
    - Acting on X-H and Y-H to form an X-Y bond
  - EC 1.22
    - Acting on halogen in donors
  - EC 1.97
    - Other oxidoreductases

EC 2
- Transferases

EC 3
- Hydrolases

EC 4
- Lyases

EC 5
- Isomerases

EC 6
- Ligases
**Reaction Direction**

- For consistency, the reaction direction is the same for all enzymes in a given class.

- The *systematic* names, on which the classification and code numbers are based, may be derived from the written direction, even though only the reverse of this has been actually demonstrated experimentally.

- Ideally, a comment would indicate that…
The Format

EC 1.13.12.17

**Accepted name:** dichloroarylaflavin A synthase

**Reaction:** dichlorochromopyrrolate + 4 O₂ + 4 NADH + 4 H⁺ = dichloroarylaflavin A + 2 CO₂ + 6 H₂O + 4 NAD⁺

For diagram of rebeccamycin biosynthesis, click here

**Glossary:** dichloro-arylaflavin A = rebeccamycin aglycone

**Systematic name:** dichlorochromopyrrolate:NADH:oxygen 2,5-oxidoreductase (dichloroarylaflavin A-forming)

**Comments:** The conversion of dichlorochromopyrrolate to dichloroarylaflavin A is a complex process that involves two enzyme components. RebP is an NAD-dependent cytochrome P₄₅₀ oxygenase that performs an aryl-aryl bond formation yielding the six-ring indolocarbazole scaffold [1]. Along with RebC, a flavin-dependent hydroxylase, it also catalyses the oxidative decarboxylation of both carboxyl groups. The presence of RebC ensures that the only product is the rebeccamycin aglycone dichloroarylaflavin A [2]. The enzymes are similar, but not identical, to StaP and StaC, which are involved in the synthesis of staurosporine [3].

**Links to other databases:** BRENDA, EXPASY, IUBMB, KEGG

**References:**

[EC 1.13.12.17 created 2010]
EC Numbers Define Enzymes, Not Reactions!

More accurately, an EC number stands for an active site. Enzymes with multiple active sites (e.g., if several genes fuse to encode a single polypeptide) should receive multiple EC numbers.
Limitations

- No enzyme can be tested with all potential substrates…

- Enzymes that perform very complex reactions
  - pyridoxal 5’-phosphate synthase (glutamine hydrolyzing)

- Enzymes with a very broad substrate range (liver alcohol dehydrogenase)

- Old enzymes with a single reference - are they real?
Where Is the EC List?

- The primary source is a MySQL database available at enzyme-database.org

- Another database, prepared by Gerry Moss, is available at http://www.chem.qmul.ac.uk/iubmb/enzyme/

- A copy of the EC list is available via the ENZYME DB (SIB) at http://www.expasy.ch/enzyme/

- Yet another one is IntEnz at (EBI-SIB) http://www.ebi.ac.uk/intenz/index.jsp

- The EC list is also included in databases such as MetaCyc, BRENDA, KEGG etc.
EC Numbers and Pathway Tools

Each EC class, sub class and sub-sub class is implemented as a class in MetaCyc

MetaCyc Reaction: 1.97.1.3

Species Comparison

Superclasses: Reactions-Classified-By-Conversion-Type -> Simple-Reactions -> Chemical-Reactions -> EC-Reactions -> 1 -> Oxidoreductases -> 1.97 -> Other oxidoreductases -> 1.97.1 -> Solo sub-subclass for oxidoreductases that do not belong in the other subclasses

Reactions-Classified-By-Substrate -> Small-Molecule-Reactions

Enzymes and Genes:
- sulfur reductase: gec, gzeII, gzeA (Acidianus ambivalens)
- H2 sulfur oxidoreductase (Pyrococcus abyssii)

In Pathway: sulfur reduction

Note that this reaction equation differs from the official Enzyme Commission reaction equation for this EC number, which can be found here.
Reactions with full EC numbers can be marked “official” or “not official”.

A non-official reaction is one that matches the definition in the EC entry, yet differs from the exact reaction equation specified in the list.

MetaCyc contains over 9000 reactions, out of which 5580 have a full EC number.
Partial EC Numbers

- Partial EC numbers look like EC numbers except the last number is replaced by a dash, e.g. 2.1.1.-
- Partial EC numbers should not be used for functional assignment!
- Partial EC numbers are used for two primary reasons:
  - Partial knowledge (2.1.1.- is the general class of methyltransferases)
  - A well characterized enzyme that has not received an EC number yet
- The use of EC 2.3.4.? Vs. EC 2.3.4.n (Green and Karp 2005)
When To Assign A Full EC Number?

One simple rule: Assign a full EC number to a reaction only if you want the name matcher to attribute this reaction to every enzyme, in every genome, that is annotated with this number.
EC Numbers and Pathway Tools - Problems

- Currently, EC numbers are associated with Pathway Tools reactions rather than enzymes.

- This leads to reaction duplication.

When several EC enzymes are characterized with overlapping reactions, we need to have duplicate identical reactions, each with a different EC number.
**Another Problem: Incorrect Interpretation**

- The E. coli YdiB protein is EC 1.1.1.282, quinate dehydrogenase

\[
\text{L-quinate} + \text{NAD(P)}^+ \rightleftharpoons \text{3-dehydroquinate} + \text{NAD(P)H} + \text{H}^+
\]

- Pathway Tools automatically expands that reaction to the two following reactions and links them to the enzyme.

\[
\begin{align*}
\text{L-quinate} + \text{NADP}^+ &= \text{3-dehydroquinate} + \text{NADPH} + \text{H}^+ \\
\text{L-quinate} + \text{NAD}^+ &= \text{3-dehydroquinate} + \text{NADH} + \text{H}^+
\end{align*}
\]

- Problem is, these two reactions are associated with the EC numbers EC 1.1.1.25 and EC 1.1.1.24, which describe other enzymes.
What We Can Do About It

- Separate the reactions from the EC numbers, permitting multiple EC numbers per reaction and multiple reactions per EC number
Conclusion Remarks

- EC Numbers are very useful
- There are thousands of characterized enzymes w/o EC numbers
- Expansion of the EC list is slow
- Urgent Need to accelerate
- Why so little funding?
- Should we ask NIH to step up?