





Classification of Enzyme Reaction Mechanisms

Noel M. O'Boyle,¹ Gemma L. Holliday,² Daniel E. Almonacid,¹ Peter Murray-Rust,¹ John B.O. Mitchell¹ and Janet M. Thornton²

¹Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Rd, Cambridge, CB2 1EW, U.K.

²EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, U.K.

Background

The current system of enzyme nomenclature and classification was developed by the Enzyme Commission (EC) [1]. The EC system is a hierarchical classification based upon the overall reactions catalysed by enzymes. This system does not take into account differences in mechanism: in some cases, enzymes with different mechanisms have the same EC number, and vice versa.

Results

The pairwise similarity of all of the 99 entries in the dataset was calculated. Table 1 shows the ten most similar pairs, as well as the correspondance between the EC codes for each pair

Aim

Here, we describe the first quantitative method to measure the similarity of reactions based on their explicit mechanisms. This is the first step in the development of an enzyme classification system complementary to the EC, which will take into account differences and similarities in mechanism.

Method

- 1. Our dataset is MACiE [2,3], a database of enzyme reaction **Discussion** mechanisms.
- 2. Similarity between steps of a mechanism is measured in one of two ways:
 - a) The Fingerprint method (FP) uses the Euclidean distance between fingerprints containing key information about bond,

the conespondance between the EC coues for each pa
--

		FP			BC	
Rank	Pair ^a	Score	same EC ^b	Pair	Score	same EC
1	32, 33	0.94	0	39, 71	1	3
2	39, 71	0.84	3	32, 33	1	0
3	62, 63	0.78	3	10, 55	0.75	3
4	69, 83	0.75	0	11, 58	0.71	0
5	55, 61	0.74	0	2, 29	0.67	2
6	11, 58	0.74	0	74, 75	0.56	2
7	36, 45	0.72	0	62, 63	0.54	3
8	16, 33	0.71	0	7, 21	0.51	3
9	32, 99	0.71	0	69, 83	0.5	0
10	16, 32	0.71	0	55, 61	0.5	0

Table 1 - The 10 most similar pairs of reactions in the dataset. ^aReaction nn refers to MACiE entry M00nn. ^bThe level up to which the two reactions share an EC code (ranges from 0 to 4).

Both the FP and BC methods are able to detect similarities between reaction mechanisms. For example, the pair of enzymes M0039 and M0071 differ only in their substrate specificity, and have very similar reaction mechanisms, except that the mechanism of M0039 is Ca^{2+} -dependent.

charge, and molecular changes

b) The Bond Change method (BC) uses the Tanimoto coefficient to measure similarity between bond changes (including cleavages and formations) (Equation 1)

 $S_{xv} = (no. of bond changes in common)$ Equation 1 (total no. of bond changes)

- 3. A global alignment is performed for each pair of reaction mechanisms using the Needleman-Wunsch algorithm. This finds the best match between mechanistic steps in two different reactions, and gives a score for the overall similarity.
- 4. The scores from (3) are adjusted based on the maximum possible score. This penalises unmatched steps.

Example

Similarity of MACiE entries, M0002 and M0029 (BC method).

M0002, beta-lactamase

M0029, glutaminase **Bond changes**

P D	P D	
R. R.	R R	

The two methods share 6 of their respective top 10 (out of 4851) reaction pairs (shown in bold in Table 1). Because the FP and BC methods emphasise different aspects of the steps, they give complementary measures of similarity. For example, according to the BC method, the two reaction mechanisms M0002 and M0029 (see above) are very similar, but the FP method gives them a lower score as it is sensitive to the difference between cleavage of intramolecular versus intermolecular bonds.

It is also possible to use these similarity methods to search the dataset for mechanisms similar to a particular query. This could be used by an experimentalist to find out whether there are known enzymes with similar mechanisms, as a step in the validation of a proposed mechanism.

Conclusions

Although the EC system plays a very important conceptual role in the classification of enzyme reactions, researchers should look beyond the hierarchical picture that the EC imposes. Based on the similarity methods described here, we hope to develop a complementary classification system for enzymes that takes into account the similarity of their reaction mechanisms.



The best alignment between the 5-step reaction Figure mechanism M0002 and the 4-step reaction mechanism M0029.

Acknowledgements

This work is funded by BBSRC grant BB/C51320X/1. We also thank Unilever for funding the Centre for Molecular Science Informatics.

References

[1] EC System, http://www.chem.qmul.ac.uk/iubmb/enzyme/ [2] MACiE, http://www-mitchell.ch.cam.ac.uk/macie [3] G.L. Holliday, G.J. Bartlett, D.E. Almonacid, N.M. O'Boyle, P. Murray-Rust, J.M. Thornton, J.B.O. Mitchell, *Bioinformatics*, 2005, 21, 4315.