MACiE

A Database for Classification of Enzyme Reaction Mechanisms

Noel M. O'Boyle,[1] Gemma L. Holliday,[1] Gail J. Bartlett,[2] Daniel E. Almonacid,[1] Peter Murray-Rust,[1] John B.O. Mitchell[1] and Janet M. Thornton[2]

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

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

Introduction

MACiE (Mechanism, Annotation and Classification in Enzymes) is a database of enzyme reaction mechanisms [1,2]. Version 1 of MACiE contains 100 non-homologous enzymes of known structure and well-described mechanism. Reaction steps are annotated in terms of chemistry and amino acid function, and graphical animations of the reaction mechanisms have been created.

Problem

The current system of enzyme nomenclature and classification was developed by the Enzyme Commission (EC) [3]. The EC system is a hierarchical classification based upon the overall reaction catalysed by a particular enzyme. The EC system does not take into account differences in mechanism: in some cases, enzymes with different mechanisms have the same EC number, and *vice versa*.

Aim

Here, we describe a method to measure similarity of enzyme reactions based on the bond changes accompanying individual steps in the mechanism: bond formations and cleavages, and changes in bond order.

Method

1. Similarity values were calculated for each pair of reaction steps, using the Tanimoto coefficient:

$$S_{xy} = \frac{\text{(no. of bond changes in common)}}{\text{(total no. of bond changes)}}$$

2. A global alignment was performed for each pair of reactions, using the Needleman-Wunsch algorithm, giving a reaction similarity value.

Needleman-Wunsch algorithm

This algorithm is used in bioinformatics to find the best global alignment between two nucleic acid or protein sequences. Two related sequences may be almost identical except for mutations of bases (or amino acids), as well as insertions or deletions. The global alignment procedure uses dynamic programming to find the best alignment of one whole sequence against another, and gives a score for the quality of the alignment.

Case Study

A comparison of MACiE entries:
M0002 - a beta-lactamase (E.C. 3.5.2.6)
M0029 - glutamin-(asparagin)-ase (E.C. 3.5.1.38)

EC 3 - Hydrolases (perform hydrolysis of bonds)
- .5 - Acting on Carbon-Nitrogen Bonds, other than Peptide bonds
  - .1 - Substrate is a linear amide
  - .2 - Substrate is a cyclic amide

Both enzymes belong to the same EC subclass, and perform similar chemistry.

M0002 is a five-step reaction, and M0029 a four-step reaction. The result of the alignment is represented in Figure 1.

The same bond changes occur in the aligned reaction steps, with the result that the similarity value for this pair of reactions is very high. M0029 is the most similar reaction to M0002, and *vice versa*; overall in MACIE, this pair of reactions is ranked 7th out of 4950 pairs.

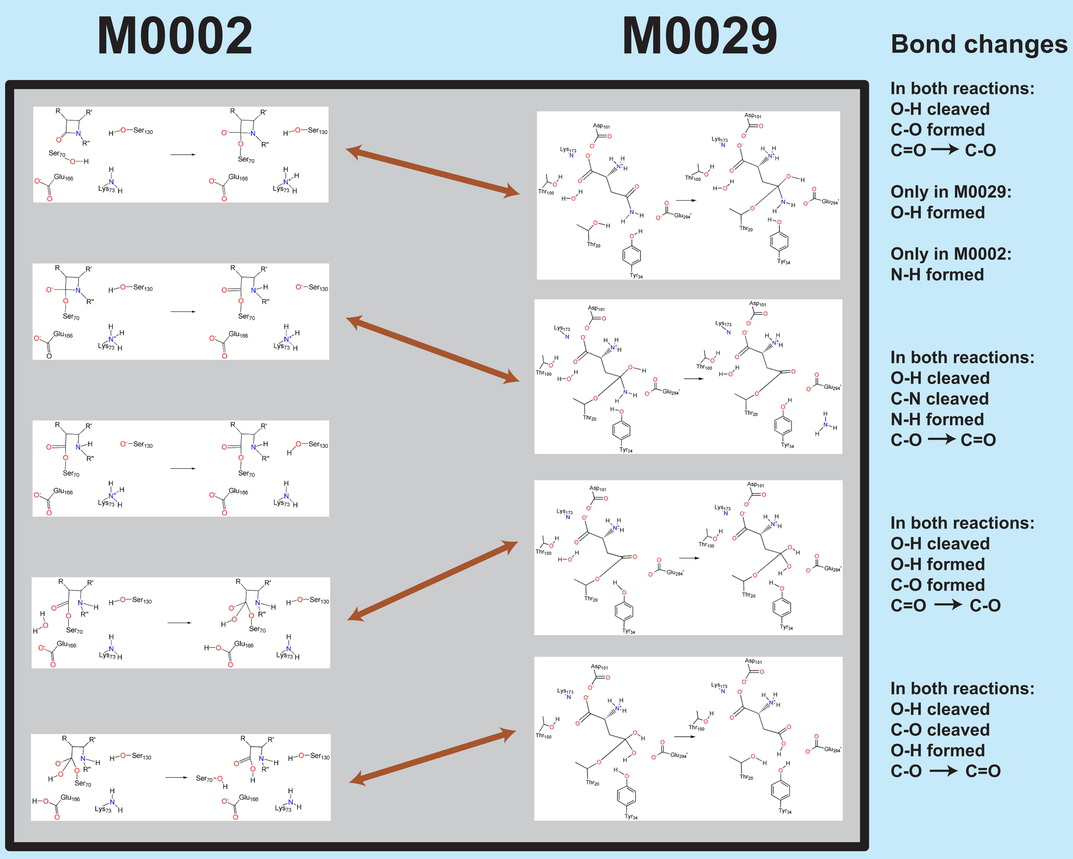

Figure 1 - The best global alignment of M0002 and M0029

Conclusions

Simply by considering bond changes in reaction steps, our method classifies as similar enzymes with similar EC numbers, in many cases. Future work will examine enzymes with similar mechanisms, which are far apart in the EC classification.

References

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

Acknowledgements

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