

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Enzyme active-site prediction

ArticleInfo		
ArticleID	:	4289
ArticleDOI	:	10.1186/gb-2001-3-1-reports0003
ArticleCitationID	:	reports0003
ArticleSequenceNumber	:	21
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2001-11-12 Received : 2001-11-12 OnlineDate : 2001-12-20
ArticleCopyright	:	BioMed Central Ltd2001
ArticleGrants	:	
ArticleContext	:	130593311

Abstract

A new method for identifying the active site of an enzyme from structural information alone

Significance and context

The growing amount of protein sequence and structural information makes being able to predict protein function and active sites even more important. With an overwhelming number of proteins to analyze, traditional methods need to be complemented by computational analysis to cope with the large volumes of data and to generate working hypotheses. Ondrechen *et al.* used the programs UHBD and HYBRID to calculate the electrostatic potentials of the ionizable groups on proteins (the side chains of the amino acids lysine, arginine, asparagine, glutamine, histidine, tyrosine and cysteine, and the amino and carboxyl termini) and the net charges for each group. Using this information they were able to generate theoretical microscopic titration curves (THEMATICS), showing net charge as a function of pH. A visual comparison of the predicted titration curves (separated for each amino acid) enabled the authors to identify abnormally shaped (perturbed) curves, which indicate the location of the enzymes active site. This method can be used not only for analysis of new enzymes but also for the analysis of protein-protein interactions and other ligand-protein interactions. THEMATICS goes one step further than previous attempts at predicting protein function on the basis of similarity to other proteins by allowing identification of the active residues not just the region conserved between various proteins.

Key results

Ondrechen *et al.* show detailed analysis of the theoretical titration curves of ionizable amino acids for three enzymes: triosephosphate isomerase (TIM), aldose reductase (AR) and phosphomannose isomerase (PMI). The first two of these share similar folds but catalyze different reactions, whereas TIM and PMI catalyze similar reactions but have different structures. Identification of the active site of each enzyme was mainly by visual assessment of the theoretical titration curve, spread over a pH range wider than is actually achievable. The authors looked for a shift in the shapes of the titration curves and pK_a values for the amino acids, which would suggest partial protonation of a particular group over a wide pH range. These changes were interpreted as a hallmark of the active-site residue. Analysis of the predicted sites using a three-dimensional model confirmed that amino acids in most of the predicted active sites are clustered together - for example, His95, Glu165 and Tyr164 in TIM. There were, however, some

amino acids that had perturbed curves but were away from the active site - for example, Lys112 of TIM. Such 'false positives' were occasionally identified in each of the enzymes. Another anomaly was the apparent lack of abnormality in the shape of the curves for His110 in AR, which, from site-directed mutagenesis, is important for catalysis. Occasionally the titration curves were of very abnormal shape, for example, Glu294 in PMI, with a highly perturbed curve that has slope over a wide pH range and partial charge between -2.0 and 7.0, creating problems in its interpretation. Such curves were excluded from further analysis. Active sites for ten enzymes were identified and a non-enzymatic protein (winged bean albumin), which gave negative results, was also used to validate the method.

Links

Details of the [University of Houston Brownian Dynamics \(UHBD\) program](#) can be found online.

Conclusions

The use of THEMATICS enabled identification of the active sites of selected enzymes in the absence of biochemical data. Some of the results were supported by crystallography or NMR results and mutagenesis studies.

Reporter's comments

THEMATICS is database-independent and is a fast computational tool, but it requires good biochemical knowledge and the ability rapidly to identify observed anomalies. There are possible problems with the method related to the false positives and the possibility of missing some amino acids at the active site. These anomalies were attributed by the authors to the anomalies in charge distribution within the protein. Finally, it has to be remembered that the active sites predicted are only theoretical until they are confirmed by studies *in vitro*. THEMATICS is nevertheless a very useful method to identify potential active sites and extract valuable information from increasingly large amounts of genomic data. As suggested by Ondrechen *et al.* the method can be automated for high-throughput screening.

Table of links

[Proceedings of the National Academy of Sciences of the United States of America](#)

References

1. Ondrechen MJ, Clifton JG, Ringe D: THEMATICS: a simple computational predictor of enzyme function from structure. Proc Natl Acad Sci USA. 2001, 98: 12473-12478. 0027-8424