



## Integrating Structural and Kinetic Enzymatic Information in Systems Biology

Matthias Stein, Razif R. Gabdoulline, Rebecca C. Wade

published in

*NIC Workshop 2006,*  
*From Computational Biophysics to Systems Biology,*  
Jan Meinke, Olav Zimmermann,  
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)  
John von Neumann Institute for Computing, Jülich,  
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,  
ISBN-13: 978-3-9810843-0-6, pp. 129-132 , 2006.

© 2006 by John von Neumann Institute for Computing  
Permission to make digital or hard copies of portions of this work for  
personal or classroom use is granted provided that the copies are not  
made or distributed for profit or commercial advantage and that copies  
bear this notice and the full citation on the first page. To copy otherwise  
requires prior specific permission by the publisher mentioned above.

<http://www.fz-juelich.de/nic-series/volume34>

# Integrating Structural and Kinetic Enzymatic Information in Systems Biology

Matthias Stein, Razif R. Gabdoulline, and Rebecca C. Wade

EML Research gGmbH, Molecular and Cellular Modeling Group,  
Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Germany  
*E-mail: matthias.stein@eml-r.villa-bosch.de*

The successful modelling of metabolic and signalling pathways in systems biology requires a detailed and consistent set of enzymatic kinetic parameters. Sometimes, these parameters are not available from the literature or were obtained under different experimental conditions. In many cases, the enzymes' activities are related to the molecular interaction field of the enzyme. PIPSA (Protein Interaction Property Similarity Analysis) is a method that enables the comparison of molecular interaction fields (for example the electrostatic potential) to classify proteins. Results obtained using the quantitative PIPSA (qPIPSA) method to derive estimates of missing enzymatic data are presented for a number of proteins involved in glycolysis.

## 1 Introduction

In systems biology, one aims at an understanding of biochemical processes in the context of a cell or organ. One approach to this goal is the mathematical modelling of protein networks by a set of coupled differential equations describing the variations of metabolite concentrations over time. Enzymatic kinetic parameters such as substrate binding affinities, catalytic turnover and initial metabolite concentrations are required when setting up a kinetic model for a signalling or metabolomic network. Despite recent achievements in large scale proteomics, many of the enzyme kinetic data have not been determined or are available only for a different organism or under different environmental conditions, see for example reference<sup>1</sup>.

The derivation of quantitative structure-function relationships for the enzyme is one of the challenges of modern enzymology. Structure-based systems biology provides insight into complex enzymatic reactions at a molecular level. We here present quantitative Protein Interaction Property Analysis (qPIPSA) which relates electrostatic potentials of enzymes to their kinetic constants.

## 2 The Quantitative PIPSA Method

The PIPSA method has been described in detail in references<sup>2,3</sup>. The molecular interaction fields of two proteins are compared in a skin region around the protein surface (see Figure 1).

PIPSA relies on protein structures and therefore a semi-automated procedure was set up to derive protein structural models by homology modelling when the structures have not been determined experimentally. For PIPSA, an experimentally well-characterized training set of enzymes is required. The retrieval of amino acid sequences, experimental enzymatic kinetic data from related organisms and the choice of an appropriate template structure is followed by multiple sequence alignment and comparative protein structure modelling.

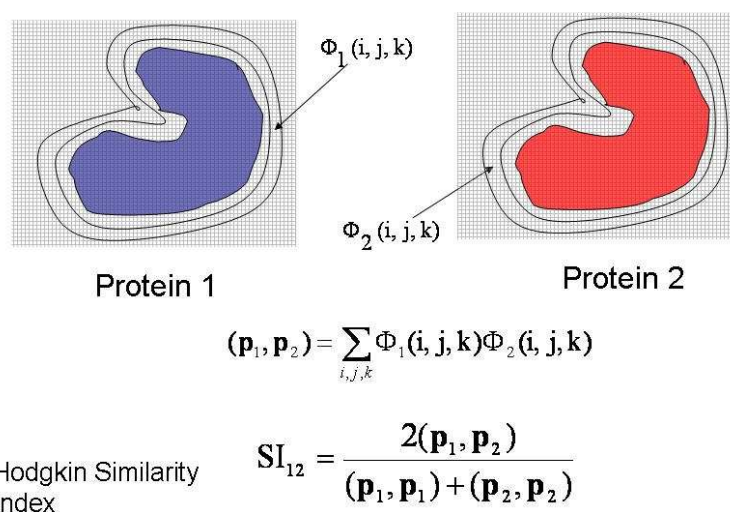


Figure 1. Schematic description of the PIPSA method. The molecular interaction fields of two proteins on a three-dimensional cubic grid are compared in a skin region. The similarity of the molecular interaction fields of the two proteins is calculated using the Hodgkin similarity index.

PIPSA can be applied to individual families of enzymes to generate kinetic parameters that are required for systems biology simulations or to a larger number of enzymes to derive a functional classification of enzymes.

### 3 Application to Glycolytic Enzymes

#### 3.1 Glucokinases

Glucokinases catalyze the first step in glycolysis, namely the phosphorylation of glucose at the 6 position to yield glucose-6-phosphate. The substrate binding affinity (as given by  $K_m$  values) between enzymes from mammals, like *Homo sapiens* and *Rattus norvegicus*, and non-mammalian enzymes, from e.g. *Saccharomyces cerevisiae*, differ by several orders of magnitude although this is not mirrored in the variation of amino acid sequences. When the electrostatic potentials of a set of 8 experimentally characterized glucokinases from various organisms are compared, large variations can be detected. The analysis of the open (apo) form of the enzyme gives the best correlation with experimental  $K_m$  values, indicating that it is the open form of the enzyme not the closed form that is determining the substrate binding kinetics.

#### 3.2 Triose Phosphate Isomerases

The electrostatic potentials of a series of twelve experimentally characterized organisms were analyzed<sup>4</sup>. This set was used to derive a correlation between the logarithm of  $k_{cat}/K_m$  values and the differences in electrostatic potentials. This correlation can be used to predict

the kinetic parameters of other enzymes. The correlation between predicted and experimental  $k_{cat}/K_m$  values indicated experimental outliers from rabbit and *Giardia lamblia*. In addition, patches of amino acid residues on the protein surface could be identified that are functionally important for substrate binding and catalysis, respectively.

### 3.3 The Full Glycolytic Pathway

The classical glycolysis consists of a series of ten chemical reactions to physiologically make use of the chemical energy stored in glucose molecules. The PIPSA method has been applied to the series of ten enzymes of the Emden-Meyerhof pathway and compared across a set of eleven organisms for which sequence information was available and for which sufficient sequence and structural homology was present in all ten enzymes of the pathway. This comparative metabolic pathway analysis shows that the interaction fields around the active sites are rather conserved across these organisms despite some amino acid sequence variation. The largest deviation from human enzymes is observed for glycolytic enzymes from the plants rice and *Arabidopsis*.

## 4 Concluding Remarks

Protein structure-based molecular systems biology provides insight into physiological processes at a very detailed level. This insight is complementary to the more abstract mathematical modelling and allows the functional classification of individual enzymes or the comparison of full pathways across various organisms. The generation of enzymatic parameters from protein structural modelling aids the construction and interpretation of kinetic models of networks of proteins for which not all of the critical parameters have been determined experimentally.

## Acknowledgments

This work was supported by the Klaus Tschira Foundation and the BMBF within the German systems biology network ‘Hepatosys’ (Grant 0313076).

## References

1. R. R. Gabdoulline, U. Kummer, L. F. Olsen, R. C. Wade, *Concerted Simulations Reveal How Peroxidase Compound II Formation Results in Cellular Oscillations*, *Biophys. J.* **85**, 1421–1428 (2003).
2. N. Blomberg, R. R. Gabdoulline, M. Nilges, R. C. Wade, *Classification of Protein Sequences by Homology Modeling and Quantitative Analysis of Electrostatic Similarity*, *Protein* **37**, 379–387 (1999).
3. R. C. Wade, R. R. Gabdoulline, F. De Rienzo, *Protein Interaction Property Similarity Analysis*, *Int. J. Quant. Chem.* **83**, 122–127 (2001).
4. R. R. Gabdoulline, M. Stein, R. C. Wade, *Relating Kinetic Parameters of Enzymes and Their Interaction Fields*, manuscript in preparation, (2006).