STRENDA DB - The 'PDB' for Enzyme Function Data?

STRENDA Commission* and Carsten Kettner
Trakehner Str. 7-9, 60487 Frankfurt/Main, Germany; ckettner@beilstein-institut.de

Difficulties with Enzyme Data

Enzyme activity data can be found in large quantities in the scientific literature and databases. However, samples show that the data were measured under different experimental conditions (e.g., temperatures, pH, ionic strength, enzyme and substrate concentrations, activators and inhibitors) which makes any comparisons difficult.

Without a complete description of the experiment including materials and methods, the comparison, interpretation and reproduction of enzyme activity data is not possible [1,2]. This seriously hampers the progress of the research enterprise due to the lack of rigors and reproducibility of enzyme data.

The difficulties become even more acute for those wishing to use published data to model the behavior of metabolic systems, cellular behavior and the interaction of cells within tissues and organs. In particular, this is the case for systems and synthetic biology, which requires reliable data to create high-quality simulation data.

How to Make this Data Useful?

The integration of enzyme functional data with data from genomic, transcriptomic, metabolomic and proteomic analyses requires that

- the data must be of high quality and should be accompanied with information on statistical variability.
- the data must have been obtained under comparable experimental conditions, which requires a definition of minimum experimental information.
- the data reported in the literature must be unambiguous, which requires the proper description of materials and methods.
- the data should be comprehensive and readily accessible by other scientists.

STRENDA Provides Assistance

STRENDA (**St**andards for **R**eporting **En**zymology **Da**ta) is a data standardization project supported by the Beilstein-Institut.

The STRENDA Commission focuses on three main aims:

(A) definition of standardized assay conditions,
(B) establishment of publication standards for enzyme activity data,
(C) proposal of STRENDA DB, an electronic validation and storage system for functional enzyme data.

(A) Standardization of Assay Conditions

The derivation of uniform assay reporting protocols for the standardization of data for single enzymes and groups of enzymes presents a great challenge since the conditions under which an enzyme operates depend on the organism and organelle in which it occurs.

The basis of initial assay standards can be the physiological conditions, which are those conditions in which cells, tissues, organs, or even the whole organism are present. However, these conditions need to be determined.

First important steps toward the definition of these conditions have been carried out in collaboration with a number of various Dutch working groups: A standard assay for the enzymes from the glycolysis of baker's yeast has been defined and tested resulting in obtaining the essential kinetics of all enzymes involved in this pathway.

This approach can be regarded as proof-of-principle and can be applied with modifications for the characterization of additional metabolic pathways [3].

(B) Publication Standards for Functional Enzyme Data

The STRENDA Guidelines were developed through extensive interactions with the biochemistry community to define the minimum information that is needed to with rigorously describe assay conditions (List Level 1A) and enzyme activity data (List Level 1B).

However, the STRENDA Guidelines neither dictate or limit the experimental techniques used in enzymology experiments nor establish a metric for judging the quality of experimental data, but rather ensure that data sets are complete and validated, allowing scientists to review, reuse and verify them. The emphasis is on providing useful and reliable information [4].

With the aim to support authors to comprehensively report kinetic and equilibrium data from their investigations of enzyme activities, currently **more than 50 international biochemistry journals** include the STRENDA Guidelines in their Instructions for Authors.

List Level 1A:
Data required for a complete description of an experiment. This information should render the results reproducible.
Version 1.7, September 22, 2016; doi:10.3762/strenda.17

Data	Comments
Identity of the enzyme	
Name of the reaction catalyst	name, preferably the accepted name from the IUBMB Enzyme list
EC number	
Sequence accession number	
Organism/species & strain	NCBI Taxonomy ID
Additional information on the enzyme	
Isoenzyme	naturally occuring variant
Tissue	
Organelle	
Localization	within cell. Specify what localization is based on
Post-translational modification	add only when determined
Preparation	
Description	*e.g.*, commercial source, procedure used or reference along with modifications
Artificial modification	*e.g.*, truncated, His-tagged, fusion protein, lacking native glycosylation
Enzyme or protein purity	purity defined by which criteria. Specify whether protein or enzyme was purified
Metalloenzyme	mutant, content, cofactors
Storage Conditions	
Storage temperature	If frozen, freezing method, *e.g.*, -20 °C flash
Atmosphere if not air	
pH	*e.g.*, 7.0
At which temperature was the pH measured?	*e.g.*, 25 °C
Buffer & concentrations (including counter-ion)	*e.g.*, 200 mM potassium phosphate, 100 mM HEPES-KOH
Metal salt(s) & concentrations	*e.g.*, 10 mM KCl, 1.0 mM $MgSO_4$
Other components	*e.g.*, 1.0 mM EDTA, 1.0 mM dithiothreitol, 10% glycerol, 20% DMSO, 1 mg/ml PEG2000, 2 mg/ml BSA, peptidase inhibitors
Enzyme/protein concentration	molar concentration if known, otherwise mass concentration; *e.g.*, mg ml^{-1} or better: µM
Optional: Statement about observed loss of activity under the above conditions	*e.g.*, less than 10% loss after 1 month
Statement about the thawing procedure	*e.g.*, on ice

List Level 1B:
Description of Enzyme Activity Data. This is the minimum information required to describe enzyme activity data.
Version 1.7, September 22, 2016, doi:10.3762/strenda.27

Information required	Comments
Required data for all enzyme functional data	
Number of independent experiments	any problems of reproducibility should be stated
Precision of measurement	*e.g.*, standard error of the mean, standard deviation, confidence limits, quartiles
Specification whether relative to subunit or oligomeric form	
Data necessary for reporting kinetic parameters	
k_{cat}	V_{max} may be divided by the specific activity units, measured in s^{-1} or min^{-1}
V_{max}	V_{max} given as units, as defined in List Level 1A
k_{cat}/K_m	k_{cat}/K_m given as concentration per time, *e.g.*, $mM^{-1}\ s^{-1}$
K_m	units or concentration necessary, *e.g.*, mM
$S_{0.5}$	concentration, *e.g.*, mM
Hill coefficient, saturation ratio (RS) or other coefficients of cooperativity	
How was the given parameter obtained?	*e.g.*, non-linear curve fitting using least squares, non-parametric method such as direct linear plot, linear regression to tranformed form of rate equation. Note: if commercial computer programs are used, determine which were used.
Model used to determine the parameters	with explanation of why is the chosen model considered to be the "right" model
High-substrate inhibition, if observed, with Ki value	
Data required for reporting inhibition data	
Time-dependence and reversibility	with method described
Inhibition	K_i units necessary
types:	
reversible	*e.g.*, competitive, uncompetitive, etc. with units and how values were determined
tight-binding	association/disscociation rates

(C) STRENDA DB - an Electronic Validation and Storage System for Functional Enzyme Data

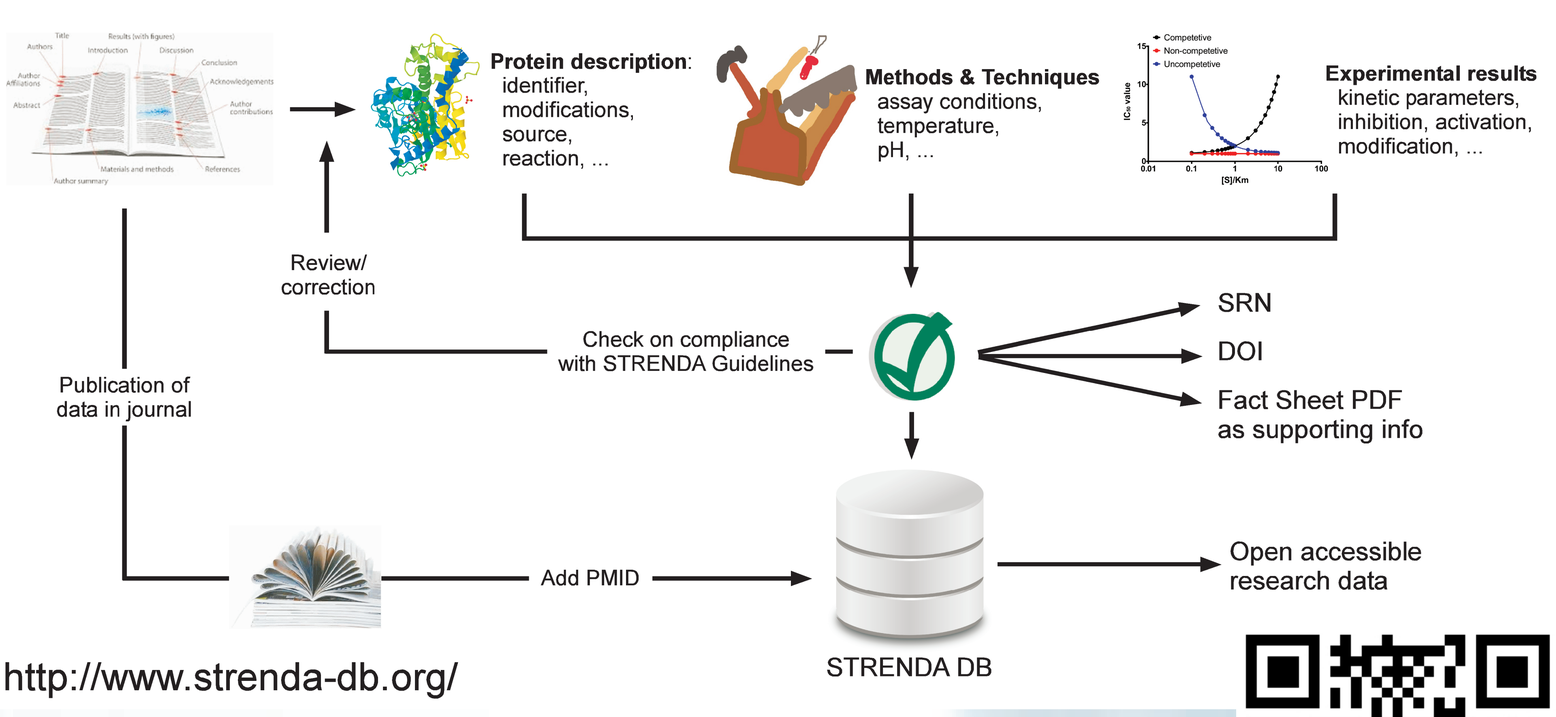
Authors (and journals) benefit from the use of STRENDA DB since it

- is an online storage and search platform,
- incorporates the STRENDA Guidelines,
- checks submitted manuscript data on compliance with Guidelines,
- ensures that data are complete and valid,
- points to missing protocol information.

A successful formal compliance is followed by

- issuance of a STRENDA Registry Number (SRN),
- generation of a fact sheet (PDF) containing all input data that can be submitted with the manuscript to the journal,
- assignment of a DOI for each dataset that allows reference and tracking of the data,
- public availability of data in the database only after the corresponding article has been peer-reviewed and published in a journal.

Many journals are already recommending their authors to store their enzyme assay and activity data in STRENDA DB, among them are *eLife*, *Journal of Biological Chemistry*, *Beilstein Journal of Organic Chemistry*, *Research Data*, *Nature*, and *Archives of Biochemistry and Biophysics*.



References

[1] Kettner, C. & Hicks, M.G. (2005) The Dilemma of Modern Functional Enzymology. *CEI* **1**:171-181.
[2] Apweiler, R. *et al.* (2005) The importance of uniformity in reporting protein-function data. *TiBS* **30**(1):11-12.
[3] van Eunen, K. *et al.* (2010) Measuring enzyme activities under standardized *in vivo*-like conditions for Systems Biology. *FEBS J.* **277**(3):749-760.
[4] Tipton et al. (2014) Standards for Reporting Enzyme Data: The STRENDA Consortium: What it aims to do and why it should be helpful. *Persp. in Science* **1**:131-137

* Members of the STRENDA Commission: Richard N. Armstrong †, Vanderbilt University, USA; Amos Bairoch, Swiss Institute of Bioinformatics; Barbara Bakker, University of Groningen, The Netherlands; Athel Cornish-Bowden, CNRS Marseille, France; Paul F. Fitzpatrick, The University of Texas Health and Science Center, USA; Peter Halling, Strathclyde University, UK; Thomas S. Leyh, Albert Einstein College of Medicine, USA; Claire O'Donovan, EMBL - European Bioinformatics Institute, UK; Frank M. Raushel, Texas A&M University, USA; Johann M. Rohwer, University of Stellenbosch, South Africa; Santiago Schnell, University of Michigan, USA; Dietmar Schomburg, Technical University Braunschweig, Germany; Neil Swainston, University of Manchester, UK; Ming-Daw Tsai, Academia Sinica, Taiwan; Ulrike Wittig, Heidelberg Institute for Theoretical Studies, Germany; Roland Wohlgemuth, Sigma-Aldrich, SwitzerlandKeith F. Tipton, Dublin, Ireland; Carsten Kettner (co-ordination), Frankfurt, Germany;

For more information, visit: www.beilstein-strenda.org. The STRENDA project is primarily supported by the Beilstein-Institut.