

FAIR enzymology: how STRENDA can help you to make your data Findable, Accessible, Interoperable and Reusable

Johann M. Rohwer

Lab. for Molecular Systems Biology, Dept. of Biochemistry, Stellenbosch University, Stellenbosch



STRENDA DB is an online curated database of enzyme function data, which ensures that submitted data conform to the STRENDA guidelines.

Introduction: When reporting results from enzyme function experiments (e.g. enzyme-kinetic parameters), it is crucial that considerable metadata are also reported in order to allow others to replicate the experiments or fully interpret the results. These metadata are essential for making experimental data findable, accessible, interoperable and reusable (<https://fairsharing.org/>).

Internationally agreed upon standards facilitate this process by defining the metadata that must be reported. The sad reality, however, is that crucial metadata (e.g. temperature, pH, ionic strength or buffer composition) are often missing from papers reporting enzyme function data, even in the recent literature [1].

Results: The STRENDA Commission [2] was created in 2004 under the auspices of the Beilstein Institut to formulate Standards for Reporting ENzymology DAta. In a series of annual meetings a set of guidelines has been developed, which have been registered in FAIRsharing.org and are recommended by the majority of journals publishing enzyme function data. To facilitate compliance with the standards, to ensure persistence of the data and metadata, and to allow easy searching according to a variety of input fields, the online database STRENDA DB [3] was developed.

In this presentation I provide a brief history and an overview of the STRENDA Guidelines. By way of an example, I will explain how they can be applied in a workflow when preparing a manuscript reporting on enzyme function data. More recent developments, such as the creation of an XML-based markup language, EnzymeML (based on the Systems Biology Markup Language), will be discussed. This has the potential to significantly ease the burden of compliance to the STRENDA Guidelines by providing an automated seamless workflow for many aspects of the process from experimental data acquisition in the lab, to fitting of kinetic parameters, to creating an EnzymeML document for storage in a standard format, to finally entering the data into STRENDA DB.

Conclusion: In many ways, modern science is in a reproducibility crisis, and adherence to a set of well thought-out, community-accepted standards provides a way out of this crisis. The STRENDA Guidelines, EnzymeML and STRENDA DB fulfil this role for enzyme function data, and researchers working in this field are strongly encouraged to make use of these tools to ensure that their data are FAIR.

References:

1. Halling, P., Fitzpatrick, P. F., Raushel, F. M., Rohwer, J., Schnell, S., Wittig, U., Wohlgemuth, R., Kettner, C. (2018) *Biophys. Chem.* **242**, 22-27.
2. Tipton, K. F., Armstrong, R. N., Bakker, B. M., Bairoch, A., Cornish-Bowden, A., Halling, P. J., Hofmeyr, J.-H., Leyh, T. S., Kettner, C., Raushel, F. M., Rohwer, J., Schomburg, D., Steinbeck, C. (2014) *Perspectives in Science* **1**, 131-137.
3. Swainston, N., Baici, A., Bakker, B. M., Cornish-Bowden, A., Fitzpatrick, P. F., Halling, P., Leyh, T. S., O'Donovan, C., Raushel, F. M., Reschel, U., Rohwer, J. M., Schnell, S., Schomburg, D., Tipton, K. F., Tsai, M.-D., Westerhoff, H. V., Wittig, U., Wohlgemuth, R., Kettner, C. (2018) *FEBS J.* **285**, 2193-2204.

Keywords: database, enzyme function data, enzyme kinetics, FAIRsharing, reproducibility, standards