Minutes of

the 15th STRENDA Meeting

9 September, 2019

Hotel Jagdschloss Niederwald, Rüdesheim, Germany

By Dr. Carsten Kettner

Beilstein-Institut, Trakehner Str. 7 – 9, 60487 Frankfurt/Main, Germany

ckettner@beilstein-institut.de

Table of Contents

List of Participants

1. Barbara M. Bakker (BB), University Medical Center Groningen, The Netherlands
2. Paul F. Fitzpatrick (PF), The University of Texas Health Science Center at San Antonio, TX, USA
3. Carsten Kettner (CK), Beilstein-Institut, Frankfurt am Main, Germany
4. Thomas S. Leyh (TL), The Albert Einstein College, Bronx, NY, USA
5. Jürgen Pleiss (JP), University of Stuttgart, Germany
6. Frank M. Raushel (FM), Texas A&M University, College Station, TX, USA
7. Johann Rowher (JR), University of Stellenbosch, South Africa
8. Santiago Schnell (SS), University of Michigan, Ann Arbor, MI, USA
9. Neil Swainston (NS), University of Liverpool, UK
10. Ming-Daw Tsai (MDT), Academia Sinica Taipei, Taiwan
11. Hans V. Westerhoff (HW), Vrije Universiteit Amsterdam, The Netherlands
12. Ulrike Wittig (UW), Heidelberg Institute for Theoretical Studies, Germany
13. Roland Wohlgemuth (RW), Lodz University of Technology, Poland

Unfortunately absent

- Toni Baici, University of Zurich, Switzerland
- Athel Cornish-Bowden, CNRS-BIP, Marseille, France
- Peter Halling (PH), University of Glasgow, UK
- Claire O'Donovan, EBI, Cambridge, UK
- Dietmar Schomburg, Technical University Braunschweig

The Agenda

(as approved by the participants)

Monday, 9 September

9.00   Welcome and Opening - CK

Proposal for appointments to STRENDA Commission - CK

Report about activities since the previous meeting, including:

- Conferences and contacts – CK, RW, et al-.
- Progress with STRENDA DB - CK
- NFDI4Chem - CK
- EnzymeML - JP

Reconciliation of actions planned – CK and task leaders
(task lists are included in the Appendix below)

12.30   Lunch

13.30   Proposal: Measurement Protocols - SS

Development of actions for both STRENDA and STRENDA DB, plans

Re-arrangement of previous action plan, new plan for 2019/2020

17.00   Close

19.30   Welcome reception for the Beilstein Enzymology Symposium

Results

Appointments / Memberships

Prof. Hans Westerhoff, Universities of Amsterdam, was appointed to the STRENDA Commission. He accepted the invitation.

The appointment of Andrew McDonald, Dublin, was discussed but the decision has been postponed to the next meeting. Andrew's role in maintaining ExploreEnz was highly appreciated and as such it is desirable to have a link to the EC Commission. In addition, there are plans to involve him in the further development of STRENDA DB.

Dietmar Schomburg stepped back from his STRENDA engagements as he will be retired by the end of this year. In addition, he doesn't believe to be able to contribute meaningful to the STRENDA project in the future. He will be in contact with CK to find another link/representative for BRENDA.

Overview

CK gave a brief overview of past activities and developments since the last STRENDA meeting, held in Manchester, September 2018. During his 'road-show' in November 2018, he attended the #scidata2018 conference on Better Science through Better Data with a presentation, and spent BMC (London) and PLoS (Cambridge) a visit.

In the following, a short summary:

SciData2018 – Better science through better data, London, Nov. 2018
https://www.symplur.com/healthcare-hashtags/scidata18/

- organized by SpringerNature,
- cover the benefits, challenges and practicalities of managing and publishing research data
- long chat with Magdalena Skipper, Editor in Chief *Nature*
- long chat with Andrew Hufton, Editor in Chief *Scientific Data*
- STRENDA is on the list of recommended repositories of *Scientific Data* which is also the list for the other *Nature* journals. However, STRENDA DB is arranged into the chemistry section rather than into the biochemistry part. This is due to the fact that the metadata delivered by STRENDA DB is automatically used for this arrangement in the list –> needs to be fixed!

Visit with BioMedCentral, London, Nov. 2018

- Christian Matheou, *BMC Biology*, Senior Editor
- Alison Cuff, Senior Editor for six BMC series, including *BMC Biochemistry*, *BMC Systems Biology* (closed down!), *BMC Ecology*
- Both are aware of STRENDA but not sure what STRENDA really does
- BMC is interested in structured and standardized data
- but not sure how to implement the incorporation in workflow
- first step: recommendation to authors to use STRENDA DB
- journals approach authors

Chat with *eLife*'s Senior Editor

- Susanna Richmond, Senior Editor (Mark Patterson on previous Beilstein Symposium)
- aware of STRENDA but not sure (again), what STRENDA really does
- more discussion in the team planned
- meanwhile Editorial in *eLife* reporting about the implementation of a minimal standards group
- contact Andy Collings, happy about contribution, first draft of a description in Q3/2019.

Visit with PLOS One, Cambridge, Nov. 2018

- Long chat with senior editors from *PLOS*, including Iratxe Puebla, Leonie Mück, Alejandra Clark, Miquel Vall-llosera Camps, Carla Pegoraro
- Aware of STRENDA and aware of the issues with the implementation of community-based reporting guidelines, considered to be very important to the edtiors but there is not enough time to refer to them efficiently,
- "lunch-and-learn" presentation with the entire PLOS team in Cambridge, great discussion after presentation, the 'carrot-and-stick' issue was intensively discussed,
- Opportunity to advertise for STRENDA via the EveryOne PLOS Blog (https://blogs.plos.org/everyone/2019/05/07/strenda/)

Micropublications

CalTech is publisher of *Micropublications*, an online journal for small research datasets. Publications are peer-reviewed and they are using electronic forms to capture the content in much the same way as STRENDA DB does. However, *Micropublications* is not structured and yet not searchable. *Micropublications* is willing to have enzymology micro-publications but it would require an editor to handle the entries. STRENDA DB output is "already a micro-publication" and it just needs an introductory and an conclusive/discussion section.

Task: get STRENDA DB on the list of recommended repositories

Molecular Catalysis

In contact with Frank Hollmann, TU Delft, member of the editorial board. He's discussed STRENDA and STRENDA DB at the previous editorial board meeting and received the agreement to recommend their authors to use STRENDA DB. CK contacted Rafael Luque, Editor in Chief, as well and provided both with material describing STRENDA DB in more detail.

Both the STRENDA Guidelines and STRENDA DB are now recommended by *Molecular Catalysis* (https://www.elsevier.com/journals/molecular-catalysis/2468-8231/guide-for-authors#87500).

It would be great to get feedback on how the authors respond.

More Journals

RW is in contact with a number of biotechnology journals, a.o. *Biotechnology and Bioengineering* (Danny Clark), *Biotechnology Journal*, *New Biotechnology* and *Biocatalysis and Biotransformation*.

The *Biotechnology Journal* already describes STRENDA and is willing to include the Guidelines into the instructions for authors.

The Editor in Chief, Antonio Ballesteros, from *Biocatalysis and Biotransformation* has been contacted recently and we are keeping in touch.

Presentations

- eScienceDays, Heidelberg, March: with a short presentation on STRENDA DB
- BioCuration, Cambridge, April: Poster presentation by UW. Actually not the perfect target group as this conference is mainly for curators and database editors but it is important to raise awareness in this community as well. In addition, one can find many interesting new technical ideas and solutions...
- Trends in Enzymology and Biocatalysis, Rome, May: Poster presentation and opportunity to run two Workshops on Enzyme function data monitored from origin to reuse. A short presentation on STRENDA DB followed by a discussion about the pitfalls, challenges and requirements for complete enzyme reporting. Interesting outcome: young scientists embraced STRENDA DB primarily (altruistically) because of the completeness of data aspect of the deposition.
- Nature as a Teacher, Stuttgart, July: Presentation of STRENDA DB with embedded moderated discussion
- BioRoboost, Madrid, July, presentation of STRENDA DB by RW. The participants appreciated clearly the importance of STRENDA for synthetic biology. However, this community is not yet fully utilizing initial-rate parameters in their designs.

General

Athel Cornish-Bowden has published a nice overview of STRENDA on the FEBS Network: Making Enzyme Data Fully Accessible, May 2018 (https://tinyurl.com/yxk219tk).

NFDI4Chem: CK reported that Beilstein is involved in the National Research Data Infrastructure consortium for chemistry with their expertise in open access and standardization projects. The goal of this large project is to establish a research data infrastructure by developing exchange formats, metadata, standards and creating a culture of research data management and sharing. Five Taskgroups have been established, i.e. Management, smartlab, repositories, standards and formats, community and training.

EnzymeML: JP presents the recent development of EnzymeML, a data exchange format that mediates the flow of enzyme function data from recording and modelling and publishing and databases. The goal is to acquire reaction conditions, time-course data, etc. and to transfer these to a modelling package such as BioCatNet or Copasi. From there, the data can be transferred as published data into STRENDA DB or SABIO-RK. EnzymeML is based on SBML enriched by extensions from the STRENDA Guidelines that

are required for enzymology data capturing. Currently, modelling is the weakest branch of the system. The idea is also to accelerate the data flow into STRENDA DB with an exchange format in hand.

Proposal: setting up a simple proof-of-principle experimental design that generates "real" data that can be published and transferring this data into STRENDA DB via EnzymeML. This EnzymeML file appears simple and is restricted to these simple experiments but can be used to learn from the data transfer procedure and to extend by further data types and complexity of the experiments.

STRENDA DB

Status

- 42 datasets published
- ~ 30 datasets started (but not yet completed) by a student from SS's group
- SS and MDT still continuing to try to hire someone with the intention of entering data, and helping others to do so
- Retrospective data entry is currently allowed
- Some bug fixing has been carried out in the background
- Matomo (former Piwik) analytics being implemented by 7 December, 2019.

Next Steps

XML metadata

- Metadata schema used for DOI registration with DataCite needs to be reviewed and updated.

EuropePMC

- Links from abstract to dataset
- Prerequisite is the use of a specific metadata template
- Goal: embedding a link to the STRENDA DB record at the corresponding abstracts indexed by EuropePMC

Developments and Extensions Planned, e.g.

- User profile (ORCID)
- DOI of article and direct link to the publication
- Query/overview/PDF: native protein vs. modified protein
- Resolving the decimals "issue"
- Protein purification (as stated in the guidelines)
- Macromolecules (access to specific databases, assignment of role which allows users to enter the kinetics of enzymes with macromolecules as substrates or modifiers)
- Kinetic parameters (fitted, ones, progress curves, initial rate values)
- Kinetics of modifiers
- Multi-subunit enzymes
- Validation of sequence modificaitons
- Duplication of experiments (comparable to Experimental Datasets)

Proposal – Measurement Protocols

SS presents his work on the lack of reliability of initial-rate parameters. He presents the remarkable spread in K_M data of a certain enzyme (forgotten which one). Idea of SS was to develop measurement protocols as part of the guidelines.

There were lots of discussions regarding how to explain this spread, e.g. error in measurements, i.e. typical assay errors. Following suggestions were made:

- Cluster the papers with similar K_M
- Identify outliers
- Check the conditions of outliers against those in other clusters
- Identify which information is missing to understand the results, e.g. [S] <<]E]

However, as there were many discussions on his study, there was the decision to postpone his proposal to the next meeting.

Reconciliation of Actions Planned and Agreed at the 14th Meeting

No 1: Increasing the number of journals actively supporting STRENDA DB

Status Date: 9 September, 2019

Task	Definition and Status	Who? By when / when to start?	Status
"Package Deal"	Prerequisite: increased number of datasets in STRENDA DB		Prerequisite not met
Focus initially on 3-4 key journals, perhaps JBC, FEBS J, Biochemistry, ABB, Biochem J.	Select journals, based on which criteria Proposal for criteria: • enzyme kinetic data; • protein functional characterization; • reuse of published data (?)	HW	started
Count number of enzyme function papers not using STRENDA and/or DB	• Access journal(s) • Retrieve/read papers • count	PH (2), JR	DONE see Appendix below
	Invite key editors to our meeting	Alanna Shepartz for Enzymology Symp.	Invited but not accepted due to time constraints

Task	Definition and Status	Who? By when / when to start?	Status

	A STRENDee presents to Editorial Board meeting – we have members on many
	• Contact with journal
	• invitation,
	• presentation and discussion
	• follow-up discussion until decision
		After Paul's trial phase, data number has increased	ONGOING
			
			Dec., 2018:
			    Direct contact with journals, i.e. *PLOS One* and *Biology*, *BMC Systems Biology*, *BMC Biology*, *eLife*. All journals just "a bit" aware of STRENDA, personal discussions helped to sort out how to proceed. Follow-ups still required.
			    BMC Systems Biology may become obsolete as this journal does not accept any submissions!
			
			Jan., 2019
			    Molecular Catalysis: Instructions in preparation that advise authors to refer to STRENDA DB. Authors will have to explain in the cover letters if they did not use STRENDA DB.

Task	Definition and Status	Who? By when / when to start?	Status
For trial period, offer help to enter data to STRENDA DB from submitted papers	- select journal(s) to support, - create a workflow, - define what to do - contact the journal	PF	OPEN
Keep publishing "white papers", letters etc. on STRENDA and DB	- What to write? - Where to submit to? - Does it include blogging and social media?	NS, SS, UW, JP: e.g. Why we need a data exchange format?	<u>ONGOING</u> <u>May, 2019</u> Blog post on STRENDA DB on Everyone Blog of PLOS <u>Jan., 2019</u> - first draft of a white-paper on EnzymeML Draft is on Google docs, pls. request for link. <u>August, 2019</u> - Manuscript on a proof-of-concept circulating <u>September, 2019</u> 2ⁿᵈ draft circulating with some extensions regarding the schema

Task	Definition and Status	Who? By when / when to start?	Status
Tick box for authors saying where (whether) you have deposited data (and if yes, where).	Comment: not clear where this tick box will appear.	PH: preparation paragraph to be sent to journals	OPEN
Details of presentation in instructions to authors that make it easy to check, e.g. submit STRENDA DB pdf, accession code		PH: preparation of paragraph	OPEN
Reward junior authors (cash? recognition?)	Define the threshold, Specify award (T-shirt, talk at conference, travel grant,…) Define "junior"	CK	OPEN

No.2 Increasing the data rate in STRENDA DB

Status data: 9 September, 2019

Task	Definition and Status	Who? By when / when to start?	Status
Upload of published data by community	- Check papers with enzymology recently and convince people to submit! - Potential source: SABIO-RK could offer support (can provide contact details, Ulrike has overview of scientists with e-mail addresses since 2-3 years) - E-mail people who recently published enzymology data. - Carsten plus commission can draft a Newsletter. It should be personalised, with the name of the recent publication. They may upload already published data, but can also be made aware for uploading future datasets.	CK, UW, TL, BB, NS	<u>DONE</u> ➢ >70 messages sent out ➢ Rare responses (?)
Advocacy for STRENDA DB	We need enzymologists with a strong voice (Ken Johnson and ..). They can advocate STRENDA by: - Writing a letter to many journals - Writing editorials - Uploading a dataset and very explicitly state this in their paper	Requirement: more data in DB CK, RW (identification	OPEN

Task	Definition and Status	Who? By when / when to start?	Status
	Other stakeholders: biochemical engineers or systems biologists with a strong voice • Identification required • Encourage to engage • Eventually coordination of wording in letters and editorials required. •	persons, contact) Next step: encourage persons to vote for STRENDA DB	
Upload (new) data	Hire a student to support people in the submission process. (Ming Daw offers that he can also hire a person). Aligned with "Package Deal" (s. task list 'Increasing number of journals')	MDT: entering data from ABB CK	Student from SS has entered ~30 datasets into STRENDA DB, only a few finalized; she's left SS's group, most datasets still pending, for some papers problems as STRENDA DB could not cover all information, sometimes not clear where to enter, some authors not responding on request. CK not able to finalize since no access to literature used. SS will send CK selected papers after the most promising datasets have been identified.

Task	Definition and Status	Who? By when / when to start?	Status
Education	Teaching about STRENDA DB in enzymology/DMP courses.	All (teaching), CK: STRENDA "play ground" online?	<u>ONGOING</u> CK: Introduction into STRENDA DB on several workshops, including online polls and intensive discussion – aim: awareness rising
Education (ctd)	Development of tutorial, →compilation material (but any of us could present this at a conference/course) Deposition and using enzyme data	BB, UW, JR	Started and ongoing
	Beilstein could even offer a summer school for enzyme kinetics. ➔ Organization (place, date, invitations, etc) by Beilstein • Teaching by STRENDees • Participation based on application, • Travel grants • Low fee + boarding and lodging	CK: Check for potential clashes (FEBS Course)	Tbd by Beilstein

Tasks for 2019/2010

Task	Definition	Lead
OPEN tasks from previous meeting	Tick box for authors saying where (whether) you have deposited data (and if yes, where):	PH
	• Preparation paragraph to be sent to journals	PH
	Details of presentation in instructions to authors that make it easy to check, e.g. submit STRENDA DB pdf, accession code.	
	• Preparation of paragraph	PF
	For trial period, offer help to enter data to STRENDA DB from submitted papers, e.g.	
	• select journal(s) to support, • create a workflow, • define what to do • contact the journal	CK
	Reward junior authors • Define threshold • Define "junior" • Specify award	RW, CK
	Advocacy for STRENDA DB • Strong voices write letters to journals • Upload a dataset and very explicitly states this in their paper • Identify and encourage these voices	
ONGOING tasks from previous meeting	White papers EnzymeML	NS, SS, JP, UW, CK
	Sugg't: analysis on FAIRness	CK, (PH?)

Task	Definition	Lead
Teaching/Education	Support online courses with material, e.g. videos, • Compile material • Include in regular lectures, • Collect materials on STRENDA website • STRENDA Presentation	BB, UW, JR CK
Awareness rising	Acceptance letter from Journal: • Contact STRENDA DB to enter data in DB, • Create a sentence for this	CK
Upload published data	Hire a student to capture data from publications retrospectively, ~30 dataset entered by TK:	SS, MDT SS sends selected articles that allow CK to finalize some oft he datasets
	Frontiers journals	RW
EnzymeML	• Prototype pipeline, • Design of EnzymeML for defined real experiments, • Upload into STRENDA DB (needs to be developed)	JP, CK, NS, UW
Additional data	From College Station, spring 2020?	FR
Advertisement	Video on STRENDA DB, why, what, how?	CK
Google Search	Check SEO Link from your website to STRENDA DB	CK All with websites
Wikipedia	STRENDA article	JP, UW

Task	Definition	Lead
Awareness rising	• Letter to key journals (FEBS, JBC, Biochemistry, Biochem J., Int'l Journal of biological macromolecules) • Rise awareness for STRENDA DB • After recommendation of guidelines, now DB • Offer authors help for data entry	HW
First contact	Identify journals not on the list of recommending journals First contact	JR, SS CK
	In contact with: • Biotechnology & Bioengineering • Biotechnology Journal • New Biotechnology • Biocatalysis and Biotransformation	RW
Funding	Proposal reused for new applications?	

Appendix

List of references from 01 January 2018 to September 2019 with more than 10 hits (search criteria = "enzyme kinetics" or KM or kcat or Vmax in coomplete record. Total hits = 1983 journals.

The entire list can be found here: https://tinyurl.com/y2v43vs2

PubMed

# references	Journal Title
170	International journal of biological macromolecules
103	The Journal of biological chemistry
47	Biochemistry
40	Applied microbiology and biotechnology
39	Scientific reports
31	Protein expression and purification
30	Molecules (Basel, Switzerland)
28	Enzyme and microbial technology
24	PloS one
23	Bioorganic chemistry
21	Archives of biochemistry and biophysics
21	Applied biochemistry and biotechnology
20	Xenobiotica; the fate of foreign compounds in biological systems
19	Journal of bioscience and bioengineering
19	Biochemical and biophysical research communications

# references	Journal Title
18	Journal of agricultural and food chemistry
17	International journal of molecular sciences
16	The FEBS journal
15	The Biochemical journal
15	Methods in molecular biology (Clifton, N.J.)
14	Journal of microbiology and biotechnology
13	Journal of biotechnology
13	Drug metabolism and disposition: the biological fate of chemicals
13	Chembiochem : a European journal of chemical biology
13	Biochimica et biophysica acta. Proteins and proteomics
12	Marine drugs
12	Dalton transactions (Cambridge, England : 2003)
12	Chemico-biological interactions
12	Biochemical pharmacology
11	Preparative biochemistry & biotechnology
11	Frontiers in microbiology
11	Food chemistry
11	Bioresource technology
11	Biochimie
10	The Journal of steroid biochemistry and molecular biology

Scopus

# references	Journal titles
78	International Journal of Biological Macromolecules
19	Biocatalysis and Agricultural Biotechnology
15	Applied Biochemistry and Biotechnology
15	Protein Expression and Purification
12	Process Biochemistry
11	3 Biotech