Minutes of the

11th STRENDA Meeting

14 September 2015

Jagdschloss Niederwald Hotel
Rüdesheim, Germany

by Carsten Kettner

Agenda

(as approved by the participants)

- Report about past and future STRENDA activities
- STRENDA Commission: Organization
- STRENDA DB
- STRENDA Guidelines Level 2 – first steps

Participants

(in alphabetical order)

Barbara Bakker	University Medical Center Groningen, University of Groningen, The Netherlands	B.M.Bakker@med.umcg.nl
Paul Fitzpatrick	The University of Texas Health Science Center at San Antonio, TX, USA	fitzpatrick@biochem.uthscsa.edu
Peter Halling	University of Strathclyde, Glasgow, Schottland, UK	p.j.halling@strath.ac.uk
Johann Rohwer	University of Stellenbosch, South Africa	jr@sun.ac.za
Frank M. Raushel	Texas A&M University, College Station, TX, USA	raushel@chem.tamu.edu
Neil Swainston	The University of Manchester, UK	Neil.Swainston@manchester.ac.uk
Ming-Daw Tsai	Academia Sinica, Taipeh, Taiwan	mdtsai@gate.sinica.edu.tw
Carsten Kettner	Beilstein-Institut, Frankfurt am Main, Germany	ckettner@beilstein-institut.de

Unfortunately absent

Thomas S. Leyh	The Albert-Einstein-College, Bronx, NY, USA	tom.leyh@einstein.yu.edu
Claire O'Donovan	EMBL Outstation, European Bioinformatics Institute, Hinxton, UK	odonovan@ebi.ac.uk
Jan-Hendrik S. Hofmeyer	University of Stellenbosch, South Africa	jhsh@sun.ac.za
Amos Bairoch	Swiss Institute of Bioinformatics, Geneva, Switzerland	ab@isb-sib.ch
Dietmar Schomburg	TU Braunschweig, Germany	d.schomburg@tu-bs.de

Athel Cornish-Bowden	CNRS-BIP, Marseille, France	acornish@imm.cnrs.fr
Richard Armstrong†	Vanderbilt University, Nashville, TN, USA	r.armstrong@vanderbilt.edu

Report

Since the last STRENDA meeting in September 2014 some promoting and advertising activities have been continued. CK reported from the ACS National Meeting in Denver, March 2015, at which the two Beilstein standardization initiatives, STRENDA and MIRAGE, were presented and discussed (Session: Research Results: Reproducibility, Reporting, Sharing & Plagiarism). In addition, as co-organizer together with Martin Hicks, CK also published a meeting report that includes a short summary of all contributions to this session in the CINF Bulletin, 2015, 67(2).

Data quality and reproducibility was also the topic of another special session, organized by CK together with Laszlo Fesus (University of Debrecen, Hungary) at the FEBS Congress in Berlin (July 2015). Here, a small panel of speakers presented their views on the publication reality but also discussed some possible ways towards the improvement of both the assessment of data quality and the data themselves.

The STRENDA Book (http://www.sciencedirect.com/science/journal/22130209/1/1-6), published in 2014, is obviously still very interesting for the community; from May 2014 to April 2015 the articles have been downloaded more than 42,000 times and there are good reasons to expect even more full-text requests in the future. This shows that the efforts spent by the authors were worth it.

In the late summer, CK became a member of the DIN Standards Committee 'Food and Agricultural Products' in the section biotechnology, subgroup 'Data processing and integration'. The aim of this task group is to develop standards for traceable, searchable and interoperable data together with integrated data processing for biotechnology and life sciences. CK will provide his expertise and experiences from more than ten years community-driven standardization work.

STRENDA Commission - Organization

The Commission commemorated Richard N. Armstrong, an enthusiastic supporter of STRENDA, who passed away in June 2015.

After more than twelve years of developing and maintaining guidelines, discussing and developing a software solution for the assessment of manuscripts' data on compliance with the STRENDA guidelines and many more activities raised the question if the Commission would need to undergo some re-organization.
CK presented his view on the major challenges for the Commission, i.e.

STRENDA DB:

- development (new features, bug fixing, UI, ontologies...)
- addressing issues such as data citation, micropublishing, commenting, RDF (resource description framework),...
- promotion towards acceptance and making it mandatory (as for the guidelines)
- publication on STRENDA DB
- ...

Guideline Level 2

- collection of best practice references vs. „real“ guidelines
- online presentation
- maintenance and curation
- ...

Networking regarding

- Quality, reproducibility, transparency

Currently, the Commission consists of one working group with members from a variety of biochemical fields providing various expertises. Due to individual work-loads the members are available for STRENDA to a limited degree. However, in order to promote the STRENDA concerns more efficiently (journals, scientific networks, funding agencies) and to increase productivity on request, CK proposed two possible solutions: extension of the Commission by additional hands or re-organization of the Commission into a working group that is active hands on and consist of experienced Postdocs and an advisory board that reviews the outcome of the working group and discusses the results with the wider community, journals and funding agencies.

After close examination of options and possibilities **the Commission decided that the organization and structure will be kept as it is** but the members agreed to involve interested members of their working groups in hands-on activities. The contributors will be then acknowledged on the STRENDA web site.

STRENDA DB

CK gave a short overview of what has been done since the discussions at the previous STRENDA meeting in September 2014.

The following tasks have been finished:

- Name of ESS in the Experiment box
- Renaming: was the protein heterologously expressed?
- Protein concentration in the assay: units/ml
- Rearrangement of „questions“ in the inhibition section
- Buffers: additional statement: refer to help text
- Activation section fine
- Macromolecules: replace peptides by proteins and nucleotides by DNA/RNA
- Set default for units
- Scientific notations required
- Product not mandatory
- Kinetic parameters: delete Ki
- Finalization experiment: dialogue, the same as for finalization manuscript
- Methodology box at the end of assay conditions

Following change requests were re-classified to feature requests:

- Inclusion of equations?
- Databases in mind: ChEBI for lipids, UniProtKB for proteins
- Copy experiment #1 to experiment #2
- Buffers: additional statement: refer to help text

Since March 2015 two test phases have been carried out. Following results have been reported (and commented and mostly fixed):

- Wording: titles, buttons, links, fields
- Descriptions: direct at the field
- Fields: protein source, enzyme concentration, enzyme reaction
- User guide: help texts close to the fields and in general for each section
- Units: corrections and additions
- Bugs: e.g. rounding errors, values ≥ 0
- Open Issues:
    - cooperativity, normalization...
    - macromolecules: list of concentrations/amounts
    - macromolecules: which database to choose for automatic query? (peptides, nucleic acids, oligosaccharides, lipids...)
    - special experimental cases,
    - cross-checks, validation procedures
    - ontologies,
    - bi-substrate reactions, grid datasets, progress curves etc.,
    - security issues, query interface

The results of the second phase have been reported closely before this STRENDA meeting and were discussed at this meeting.

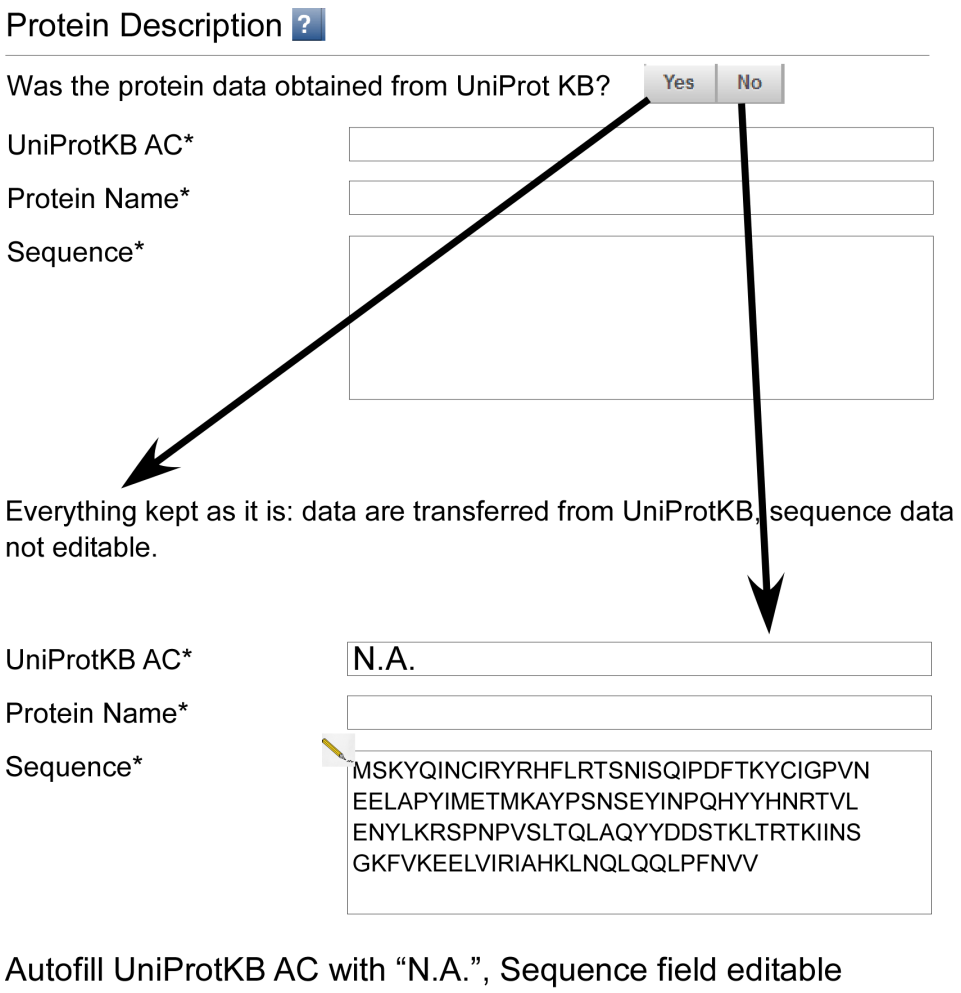
The current version of STRENDA DB is 0.9.1 RC6 (release candidate 6).

Field Description	Change requested	Specification	Degree of complexity [high/average /low]
e.g. Manage Manuscripts	A useful help box appeared on one of the initial pages, but then disappeared again before I had time to read it!	Extend display time and allow user to close (closing cross should be better visible)	low
Add Experiment, Experiment Details: Methodology	Add recommendation such as fill in later	Below "Please describe the methodology…": "For first time users it is recommended to leave this field blank and fill in the details when the experiment is going to be finalized."	low
Add Experiment, Experiment Details: Methodology	The Experiment Details free text box that comes up early in describing an experiment. I am asked to include details that I haven't specified below before I can see exactly what I can specify below. For release soon I would add some help text like	Add to help text "The intention here is to provide the user with a free text box to give details that are not currently captured in the more structured fields elsewhere. At present we are not sure whether the positioning and availability of this box is best, and would appreciate comments on this."	low
Add Protein		Add: "Search for Proteins in UniProtKB"	low
Protein Description	Synthetic proteins cannot be described using aa sequence since the sequence field is non-editable	In the section protein description, first ask whether it is a UniProt-indexed protein. If yes, everything is kept as it is. If no, make the sequence field editable, see graphical specification #1	high
Protein Reaction – Reaction as in ExplorEnz	Bug: Change of UniProt Search using UniProtKB AC does not result in an update of "Reaction as in ExplorEnz"	Description: First search and select hit from UniProtKB hit list ⇢ autofill all mandatory fields. Second search with UniProtKB AC in search field, select hit from hit list ⇢ autofill but reaction field is not updated	
Protein Sequence Modification	Bug: Letters only, no numbers or other characters		low

Field Description	Change requested	Specification	Degree of complexity [high/average /low]
Protein Sequence Modification	Less strict validation on 5 aa before / 5 aa after modification. Change to 1 aa before / 1 aa after		low
Protein Source	Re-order fields	(Host) organism, Taxon ID, Strain, Cell type, Tissue, Localization	low
Protein Source	Re-name field "Organism"	(Host) organism	low
Protein Source, Was the protein heterologously expressed	Improved specification of yes/no/unknown	No: keep everything as is Unknown: Comment field only Yes: (Host) organism, Taxon ID, Strain, Cell type, ~~Tissue~~, Localization. (Host) organism without autofill	high
Experiment Overview, Help text	Define ESS in help text		low
Assay Conditions, Protein concentration	Deletion of some units Add units	Delete units/mol, %, vol% mg of pure enzyme • ml^{-1} mg of total enzyme • ml^{-1}	low
Units	Units with negative power should be expressed uniformly	Protein concentration: units ml^{-1}	low
Units	Units with negative powers – these all appear with the minus sign as a superscript, but not the digit after it. I'm guessing these are done by Unicode codes in the font, rather than a superscript marking, but there are Unicode digit superscripts e.g. 00B9 for superscript 1	Uniform representation: e.g., mg ml^{-1}	low
Assay Conditions – Macromolecular Compounds	Change unit delete units	unit ml^{-1} and uniform representation of units with negative powers units/mol, %m vol%	low
Assay Conditions, Protein concentration	Add comment field below of auto-filled name of protein	"How was the protein concentration measured?"	average

Field Description	Change requested	Specification	Degree of complexity [high/average /low]
Results, units	Delete some units: change some units:	k_{cat}: µM s^{-1} V: M s^{-1} to mM min^{-1} and add µM min^{-1} V: nkat to nkat mg^{-1} and add kat kg^{-1}	low
Result set	Add another field	Reaction monitoring, e.g. ADP formation	average
LogIn	Password forgotten	Send email request to strenda@beilstein-institut.de with the following text: "Dear STRENDA Commission, my username is [...]. I have forgotten my passwort. Please provide me with a new one. Thank you."	average
Finalize Experiment	Provide a pre-PDF for checks before closing the experiment by Finalization	Add Button : Export Pre-PDF. PDF can be displayed in PDF-Reader for proof reading.	high

Graphical specification #1



Back to table entry

Time line for the next steps

The Commission agreed on releasing STRENDA DB by the end of 2015. Before the release can be announced some tasks need to be finished beforehand:

- Bug fixing and implementation of requests as listed above
- Test (STRENDA internal) and another bug fixing if necessary
- Input of real data by the members of the STRENDA Commission
- Collection of comments and experiences
- Bug fixing if necessary
- Release of STRENDA DB

It is also worth to think about whether to invite external scientists who support STRENDA to enter their real data in order to extend the experience base and to widen the circle of scientiest who could promote STRENDA DB in the editorial boards.

STRENDA Guidelines Level 2

(i.e. guidelines for uniform enzyme assays)

The Commission discussed this issue in depth regarding its impact, capability and feasibility. Several ways for developing these guidelines have been considered but at the end the Commission decided first to organize a workshop with experts from diverse fields to sound out the opportunities to real implement such guidelines which could include a general part and an organism-related part. This workshop could be held in conjunction with the next STRENDA meeting in fall 2016. The result of this workshop will either lead to the recognition that such guidelines are either already existing or essential (but not yet implemented) or needless. The Commission will make a decision about its next steps in this matter in dependence of the outcome of the workshop.