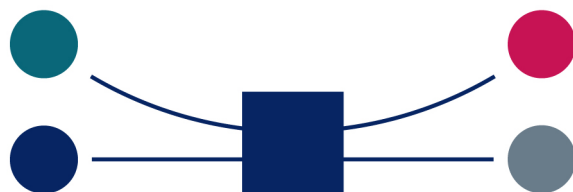# User Guide for Data Input into

STRENDA DB

Beilstein-Institut
Dr. Carsten Kettner

Version: 0.91

Date: 15/09/14

Table of Content

Welcome!

Thank you for using STRENDA DB, the electronic data submission system for functional enzyme data.

After the development of the STRENDA guidelines through extensive interactions with the biochemical community and the subsequent recommendation by to date more than 30 biochemistry journals, STRENDA DB was developed to provide the community with a web-based tool for the validation of functional enzyme data in compliance with these guidelines to pursue the following major objectives:

Support of quality assessment of manuscripts. Authors enter the relevant functional enzyme data from the manuscript into the data submission system. The data entry requires the description of the minimum information on both materials, methods, assay conditions and the experimental results based on the corresponding experimental conditions. As this minimum information is defined by the STRENDA guidelines it is also determined by the compulsory fields in the entry section of STRENDA DB. The system validates automatically the data entered in the compulsory fields. After the successful data input the author receives a STRENDA Registry Number (SRN) as a proof of compliance with the STRENDA guidelines which also indicates the data set unambiguously – comparable to the UniProtKB AC for protein data sets. Please note that the data is validated only with regards to formal completeness. The submission system does not check the quality of the data which is up to the reviewers of the respective journal.

Database for functional protein data. The data entered are stored in a database resulting in a repository that provides the kinetic data of enzyme reactions along with the assay conditions under which the experiments were performed. The data-acquisition model used by the existing protein-function databases, such as KEGG, UniProtKB and BRENDA is retrospective – annotators retrieve data, often by hand, from the published literature and add it piece-wise into the database. The direct electronic submission of data by the authors during or prior to the publication process has proven to be the gold standard for comprehensive data acquisition in e.g. structural biology (see also PDB: www.pdb.org). The data will be made accessible to the public as soon as the manuscript data are reviewed and published in a journal.

The STRENDA Commission seeks for support from both the scientific community and the journals to integrate STRENDA DB into the publication practices of the community.

This document provides a step-by-step manual for the input of data into STRENDA DB. If there are further questions, please feel free to contact the Commission by email strenda@beilstein-institut.de.

Features implemented in Version 1.0

With the release of STRENDA DB in version 1.0 this web-based tool marks an unique data submission system based on the structure of scientific manuscripts. A large number of useful functionalities is implemented which are characterized by focusing on the most used kinds of use cases and data types such as

- for the STRENDA DB website:
    - start page providing links e.g. for data input, database access, registration and login, imprint, guidelines;
    - profile management: definition of user roles (author, admin,...), user profile management (changing login data, password, password forgotten etc.), administrator (access to all user data and experimental data,
    - overview pages for data input, author-specific, all by admin, imprint, terms and licences for the use of data,
    - STRENDA guidelines
- for the data submission section
    - condensation of the diverse compounds entry tables into one providing pull-down menu for the selection of the appropriate compound used;
    - reorganisation of the pages on which the assay conditions are defined
    - deletion of the confusion caused by 'constant/varies' terms for changing compounds as introduced in the prototype,
    - all entries remain in access, new conditions are defined after copying previous conditions,
    - Proofread editing from the overview page
    - reuse of the assay conditions concept; reload of the selected assay condition together with the result set and change of one or both, ensuring consistency of assay condition and result;
    - Export data as PDF by printing the overview to virtual PDF printer
    - creation of STRENDA ID
    - save data in database
- for the database retrieval
    - easy to use database query field ('Google-like') which allows quick access to data

Features (not yet) implemented in Version 1.0

Clearly, this first version of STRENDA DB cannot fulfil all demands and requirements that might be desirable – there are still some limitations that should be known by the users before entering data and giving up frustrated. However, bear in mind that there is many commercial software on the market which keep old and annoying software errors with increasing version numbers.

So, please make allowances for still missing features – the Commission is working on the open issues which are:

- the registration process requires no confirmation using a token sent to the user.

- simple database interface to query data; data retrieval will be made possible by providing a "Google"-like search mask which allows the user to query for SRN, proteins, author names, organisms but not for numerical data such as K_M values between x and y mM.

- simple user or manuscript administration; the focus is set on the enzyme data input;

- special cases of data types can not be treated such as unknown protein or protein name, protein without UniProtKB entry, or only the reaction is known (e.g. crude extracts);

- multi-component enzymes cannot be entered; data for the catalytic active subunits can be entered;

- missing cross check of post-translational modifications against original protein sequence (e.g. validation whether phosphorylation at the correct amino acid residue);

- restricted user support for macromolecules; automatic search in PubChem only. Further databases can be selected in subsequent versions.

- Macromolecules can be entered but in the Results Set they are not considered neither as substrates or products nor as inhibitors or activators, this feature will be implemented in subsequent versions.

- entry of cell lines (different from cell types) along with the definition of the organ;

- validation of names of organisms (at NCBI Taxonomy) with TaxonID in particular if one of each entry is changed.

- Validation and use of standardized termes of e.g. tissues (using BTO).

- bi-substrate reactions; V_{max} and K_M values of a specific depends on the concentration;

- grid datasets with the concentration of both/all substrates changed;

- progress curves or time-course data

Brief introduction in the overall concept

STRENDA DB consists of two main parts; the data submission section for the input of data and the actual database for data retrieval. The access to the data submission section requires the author to be logged in (after registration) but no authentication is required when just querying the database. The data submission system includes an administration section for the administration of user data as well as of the manuscript data. From this administration section the author has access to any data sets which are still in work but not yet completed (finalized). In addition, this allows the author to enter data from different manuscripts at the same time. In particular, for larger working groups this concept becomes essential when many different manuscript data are to be entered subsequently or at the same time.
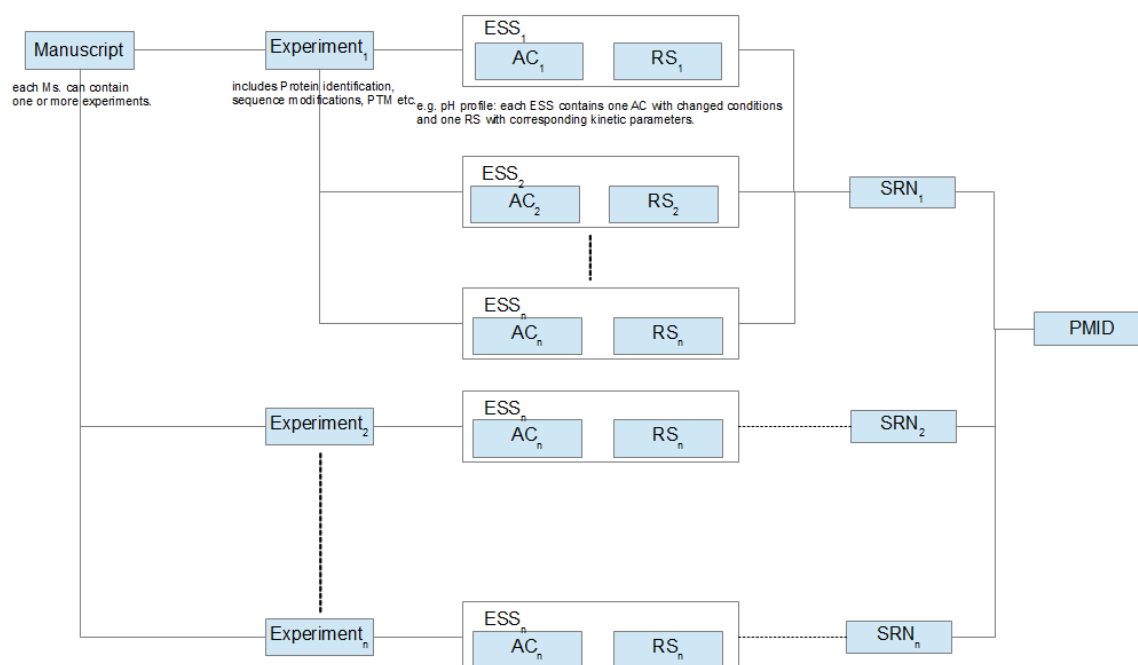
Reflecting the structure of manuscripts

The overall concept of STRENDA DB follows the idea to reflect the structure of a manuscript since this is the view of the author on his data which are supposed to be published. The author knows best which data are relevant for entering in STRENDA DB in terms of functional enzyme data. In addition, the author also know where to find the data that link the experimental conditions data with the corresponding experimental results data in his manuscript.

The structure of a manuscript follows an usual schema: introduction, materials and methods, results, discussion, references. The input concept of STRENDA DB reflects this schema with the focus on both sections the materials and methods and results (Fig. 1).

A manuscript can contain one or more experiments, whereas each experiment includes the study on one specific protein. This structure allows the user to enter data from the comparison of, e.g., the activity of two isozymes. By each isozyme a new experiment is defined which in turn can consist of one or more so called Experimental Subsets (ESS). The "core" of the experiment is characterized by the basic data on the protein, such as protein identification, protein sequence modifications (PSMs), post-translational modifications (PTMs), source and reaction carried out by the protein. The ESS consists of one defined assay condition linked to the experimental result(s) as for example kinetic parameters at a defined pH. In the case of a pH profile the experiment consists of several ESS each with different pH values but with the same assay components connected to the pH dependent kinetic parameters.

In consequence, when entering such tabular data (since pH profiles can be represented in tables), the author needs to enter the whole enzyme assay once and varies the specific parameters for the subsequent assay conditions.

Figure 1: Basic concept of STRENDA DB reflecting the usual structure of manuscripts.

Example 1: A very simple experiment is described in the manuscript: The kinetics of human hexokinase is explored using varied initial concentrations of substrate a.

Experiment 1: human HXK

> → ESS_1: components used in this assay with varied initial substrate a concentrations ranging from a minimum to a maximum value and corresponding kinetics values.

Example 2: The kinetics of the yeast pyruvate kinase a and b (PK a and PK b) were investigated at both various pH (3 to 9) and temperatures (10 – 40 °C, rising by 5 °C steps). Following the concept implemented in STRENDA DB this means:

Experiment 1: yeast PK a

> → ESS_1: components used in this assay (Assay Conditions 1, AC_1), @pH3 and corresponding kinetics values(Results Set 1, RS_1)

$\rightarrow$ ESS$_2$: AC$_2$ from AC$_1$, pH4 and RS$_2$

$\rightarrow$ ESS$_n$: ACn from AC$_2$, pHx and RSn

Experiment 2: yeast PK a

$\rightarrow$ ESS$_1$: components used in this assay (Assay Conditions 1, AC$_1$), @10 °C and corresponding kinetics values(Results Set 1, RS$_1$)

$\rightarrow$ ESS$_2$: assay conditions from ESS1, 15 °C and corresponding kinetics

$\rightarrow$ ESS$_n$: assay conditions from ESS2, xx °C and corresponding kinetics

Experiment 3:

yeast PK b $\rightarrow$ ESS$_1$: as for yeast PK a in Experiment 1 et seq.

Experiment 4:

yeast PK b $\rightarrow$ ESS$_1$: as for yeast PK a in Experiment 2 et seq.

In addition, if inhibitors or activators are used in the experiment, the first ESS, for example, includes the kinetic parameters without the inhibitor or activator. The subsequent ESS provides the kinetics which is dependent of the inhibitor or activator added. Furthermore, if several inhibitors were tested the number of ESS corresponds to the number of inhibitors. Finally, a STRENDA Registry Number (SRN) is assigned to each complete experiment (which includes one or more ESS) which marks the functional properties of a given enzyme under defined conditions. Thus, multiple SRNs can be linked to one manuscript if this contains a corresponding number of experiments on enzyme functions whereas the manuscript is identified by one PubMed ID (PMID). The user can query the database for PMIDs and obtain the number of SRNs and experiments along with the assay conditions and experimental results respectively.

Nomenclature

In order to avoid phrases such as "click on button xyz" or "click on link nmol" or "the page called Manuscript Overview opens", a number of labels is used throughout as following:

Button: small capitals of the action. Example: CANCEL;

Action link: underlined. Example: Edit;

Page name: Arial font used. Example: Manuscript Overview

Section name: italics and bold. Example: **Search for Protein**;

Field name: Tahoma font, 11pt used. Example: Tissue;

Data Input Step by Step

Registration

A typical user who is just interested in querying data in STRENDA DB or in browsing the STRENDA DB website for information such as the STRENDA guidelines or other useful information needs not acquire certain rights for his usage. Thus, there is not need to enroll in STRENDA DB. However, if the user is an author who is going to deposit functional enzyme data from his manuscript registration and login is required. This kind of personalization is required to avoid spamming STRENDA DB on the one hand and to determine responsibility for data entered in terms of naming a corresponding author to allow for data integrity in the data base. A registered and logged-in user is allowed to do both deposition of data and querying the database.

In addition, the log in procedure starts a session on the STRENDA server that allows the author to leave the entry process at any defined site without loosing data. The data once entered are stored on the server and can be restored at resumption of the data input (e.g. leaving the entry process for lunch and continuing the data input after lunch). Furthermore, when entering data in the data submission form a session is generated on the server using cookies. This session collects and stores the data temporarily for about 20 minutes. This 20 min period will restart after each activity of the user. Activity means click on buttons, initializing requests to the server etc. However, after a longer period of inactivity (e.g. lunch) due to security reasons the system closes the session automatically and all data not stored at defined points will be lost.

The first step towards logging in STRENDA DB is the registration in the system. In accordance to the Privacy Policy only the most required user data are requested to be entered and are stored in a separate part of the database.

Click on the button REGISTER in the upper menu bar, the registration form opens. Select a useful user name (please note, that this user name cannot be changed by the user) and enter the data as required for the registration process:

(1) First name: the (real) first name of the user, up to 100 characters;

(2) Last name: the (real) last name of the user, up to 100 characters;

(3) Email address: the (real) email address of the user.

(4) Affiliation: the affiliation at which the user usually works or lives, typically, the name of the organization and institute. Details such as addresses, zip codes and names of the city are not required.

(5) Password: the user can choose any password with a length of up to 100 characters. For safe passwords see e.g. here: http://www.wikihow.com/Create-a-Password-You-Can-Remember

(6) Re-type Password: in order to avoid typos in the password the user is requested to repeat the entry of his password. The validation procedure checks whether the passwords entered matches with each other

(7) REGISTER: click on the button starts the validation procedure and saves the data in the database if no validation error occurs. The user returns to the homepage displaying a confirmation message and which allows the user now to log in either via the provided link or via the button in the navigation bar.

(8) CANCEL: the user returns to the start page without saving any data.

After Registration

Administration of user profile

The registered user is allowed to administer his own user profile which includes data on his real name, user name, affiliation, valid email address as well as his password. With the exception of the user name all entries of the profile can be changed and even deleted by the author. However, deletion will be possible only if no published manuscript has been assigned to the author yet.

After login the user is presented the page Manage Manuscripts which is the main page for the administration of already entered data on relevant manuscripts. Now, the navigation bar displays EDIT PROFILE. Click on this button leads the user to his profile. Instructions on how to modify this data please can be found in Modification of User Data.

Lost password or user name

If the author is unable to recall the password or the user name he can send a request to the STRENDA DB team (strenda-db@beilstein-institut.de) for being resent his user name or password provided that one of the identification attributes (password or user name) can be stated by the user. In this first version an automatic system-generated user name or password will not be implemented. The request will be sent to the administrator who provides the missing data 'manually' by email within reasonable time.

Users' Rights

When editing a manuscript the author has a number of rights to create a new manuscript, to add and edit manuscript data and to delete these data. He will be able to finalize the data input which also "closes" the further access to these manuscript data since the manuscript has been submitted to a journal and thus the data cannot be changed unless the review causes further modifications.

Please note that this does not affect the experimental data export for each STRENDA Registry Number (SRN).

The author has no access to any other authors' data.

Login

The access to the data submission system requires the user to log in the system with user name and password. Prerequisite for the access is the registration in the system with user name, password and some additional personal details beforehand.

Click on the button LOGIN in the upper navigation bar, enter user name and password in the form and click on LOGIN. The registered user is guided into his pages to manage his manuscript data: Manage Manuscripts.

Logout

The user can stop his input work at defined points within the input procedure at which the data entered are stored. Then, the user logs out and leaves the system by clicking on the button LOGOUT in the upper navigation bar.

Modification of User Data

Sometimes it is necessary for the user to change some data e.g. in case of moving or changing email addresses. The click on EDIT PROFILE in the menu bar opens the page User Profile of <user name> opens providing a list with the following entries:

User Name, First Name, Last Name, Email, Affiliation.

The majority of these entries do not require any further explanation since these data were entered by the user during the registration process and thus are self declarative.

In the lower navigation bar the following buttons are provided:

EDIT: opening the page Edit User <user name> which provides the form known from the registration process. Any data can be modified with the exception of the user name. Clicking on SAVE returns the user to User Profile of <user name> still logged in with the new data stored.

Clicking on CANCEL returns the user to User Profile of <user name> without any changed stored.

CHANGE PASSWORD: opening the page Change Password for <user name> providing a form with just two fields: Password and Re-enter password. Please note that at this stage this feature is kept

pretty simple; neither the old password needs to be confirmed nor the new password requires confirmation by the user by sending an email which contains a token that in turn is confirmed. There is only a validation process implemented that checks the new password against the re-entered one.

CANCEL returns the user to User Profile of <user name> without saving any data.

SAVE returns the user to User Profile of <user name> with saving the new password.

From here, the user returns to Manage Manuscripts by clicking on the corresponding button in the navigation bar.

Manage Manuscripts

This section provides the user with an overview of all manuscripts in input process and finalized (and presumably published).

A manuscript refers to a complete set of all data from a study that is or might be the subject of a manuscript submitted for publication and that is relevant for STRENDA DB. Please note, that only functional enzyme data should be entered but not structural information about the protein.

First-time users may start with Add New Manuscript.

Already existing manuscripts

Manuscripts for which the data input process is still ongoing, can be handled as following:

Edit: Start with or resume the input of experimental data. After the user has clicked on "Edit", a new page opens and the user is guided to the overview of this manuscript together with the list of experiments that are included in this manuscript.

Delete: the user can delete the entire manuscript along with all experimental data already entered.

Finalize: the user can finalize the manuscript after all experiments created in the List of Experiments have been finalized and obtained a SRN. Please note, that finalization means that the user won't have further access to the manuscript's data. However, if necessary the manuscript can be reopened on request.

Finalized manuscripts

After the manuscripts has been finalized, the user has only restricted access to the manuscript's data and Show replaces all other links:

Show: After finalization the author only has read-only access to the manuscript's data. This means that all the data can be either displayed online or exported as PDF and XML. If necessary, e.g. in

the case of major revisions recommended by the reviewers, the author can request for reopening the manuscript to be enabled to modify any data. After this process, both the experiment and the manuscript are finalized and read-only.

Reopen manuscripts

After the finalization of the manuscript, the user has no rights to edit the data. If editing becomes required, the user needs to contact the administrator to request for reopening the manuscript. The contact is carried out by email: strenda-db@beilstein-institut.de.

After reopening the author has full access to his manuscript's data and can change, add and modify his data. Usually, the author needs to carry out this action after these changes also have been done with the manuscript that is subjected to be re-submitted to the journal. Please note that the SRN does not change when experimental data are modified. If experiments are added additional SRNs are generated.

Set PMID

After the publication of the manuscript the paper is indexed by indexing services such as PubMed Central and thus gains a PubMed Identifier (PMID). This PMID is added by the admin to the data set that corresponds to the manuscript data entered previously. Since there is no automatic alerting service the curators of STRENDA DB will scan PubMed for papers that fit with the entries in STRENDA DB to gather the corresponding PMIDs. This will be done regularly.

However, it would be appreciated if the authors would notify the curators of STRENDA DB when their manuscript is published: strenda-db@beisltein-institut.de.

Publish data

The data entered in STRENDA DB will be made public as soon as the manuscript was published and a PMID is obtained. The publication of the manuscript's data is carried out by the administrator of STRENDA DB; no action is expected from the author.

Add New Manuscript

To enter the data of a new manuscript, click on the button ADD NEW MANUSCRIPT on the navigation bar at the bottom of the page. The user is guided to a new page, Add Manuscript, and requested to enter the title of the manuscript and the authors' names.

Please note: the format for the authors' names is in accordance to the format of PubMed, i.e. <surname1> <initial of the first name>, <surname2> <initial of the first name>.

Example: Doe J, Smith KM

A click on the SAVE button leads back to the Manage Manuscripts page. To start with the input of data click on Add/Edit Data within the column 'Actions' on the right-hand side of the page. A new page opens called Manuscript Overview.

Manuscript Overview

This page consists of two sections: The ***Manuscript Data*** box and a table called ***List of Experiments***.

The box ***Manuscript Data*** provides any data available on the progress of the manuscript.

If you become aware that any details on the manuscript need to be changed, click on Edit on the bottom of the Manuscript Data box. A "new" page (this is the same page as you have entered when adding a new manuscript) opens and the user can change/modify the title and the authors' names of the manuscript.

Entering new experimental data

An experiment refers to a set of all data for the same enzyme, perhaps studied under a variety of conditions.

At the beginning the ***List of Experiments*** is empty and no links to actions are provided. Start with the input of data with a click on ADD EXPERIMENT.

If there are already experiments stored in the system, the ***List of Experiments*** provides the names of these experiments along with links in the 'Actions' column on the right-hand side of the table to edit or delete individual experimental data sets.

Add/Edit Data: Start with or resume the input of experimental data.

Delete:the user can delete the entire experiment along with all experimental data already entered.

Add (new) experiment

After clicking on ADD NEW EXPERIMENT a new page opens, Add Experiment, providing the head data on the manuscript and, in the section *Experiment Details* showing two compulsory fields:

1. Please name the following Experiment:

   Enter any name for this experiment. Please note that that the name of the experiment is just a mnemonic aid for the author to distinguish between the experiments presented and thus must not necessarily appear in the manuscript.

2. Please describe the methodology used in your experiment:

   In this field any information about the methodology used and techniques applied should be given. A literature reference may suffice for an established procedure but any modification should be detailed.

Following information should be given in addition (see also the STRENDA guidelines):

- *Type of assay;* e.g., continuous or discontinuous, direct or coupled,

- *Reaction of stopping procedure*; in the case of discontinuous assays,

- *Direction of the assay*; with respect to the reaction equation provided, e.g., NAD reduction by alcohol dehydrogenase; alcohol + NAD+ --> aldehyde or ketone + NADH + H^{+},

- *Reactant determined*; e.g., NADH formation, O$_2$ utilization.

Click on either CANCEL to abort the entry and return to Manuscript Overview without saving any data or on SAVE EXPERIMENT NAME. The data is stored and the author is led to the new page Add Protein.

Add protein data

The page Add Protein displays on the top the head data on the manuscript (left-hand side) and the head data on the experiment (right-hand side). The latter is completed in parallel to the completion of the corresponding data fields.

The fastest and easiest way to complete all compulsory fields in this section is to query UniProtKB for the protein's name. All relevant data will be downloaded from UniProtKB and automatically entered in the corresponding fields. Check the data entered, determine the protein sequence modifications if changed and the post-translational modifications if necessary and proceed by clicking either on CANCEL to abort the entry process and return to Manuscript Overview with the *List of Experiments* not updated or on SAVE & SHOW EXPERIMENT with the data entered stored. The user then returns to Manuscript Overview with the experiment listed in the table *List of Experiments*.

In the following the fields are described step by step.

Search for Proteins

Enter the name of the protein investigated, the UniProtKB AC if known or the name of the author who has published on this protein. It is even possible to perform combined search queries such as protein name and author name or protein name and organism. After entering the search term and clicking SEARCH a hit list opens and provides all hits found in UniProtKB. The user then can select the appropriate hit and returns to Add Protein.

For example, the query "hexokinase yeast" results in a smaller hit list than searching just for "hexokinase".

Protein Description

UniProtKB AC

> Automatically inserted, alternatively can be entered manually but please note that in this case all further compulsory fields need to be completed manually. It is recommended first to search for the protein in UniProtKB using Search for Proteins and then check and/or add the data manually.

Protein Name

> The common name of the protein along with the synonyms (in brackets) are shown as stored in UniProtKB.

Sequence

> The protein's sequence is displayed in one-letter-code as comes from UniProtKB. If the sequence exceeds a certain number of amino acid residues it is possible that not all residues will be shown at a glance. In this case the user can increase the field by drawing the resize handle located in the lower right corner of the text field. Please note that sequence data in this field are read-only.

Protein Sequence Modification

First the author is asked whether the proteins contains sequence modifications.

Click YES, if modifications are known and/or experimentally carried out. The section expands by additional fields:

> Please modify sequence: The sequence from *Protein Description*/Sequence is displayed, again, and allows for editing. Please note, in this version 1.0 modifications made are not obvious. For example, if a number of amino acids were deleted from the sequence, the user can delete them in this field but this deletion is not indicated. This is the same for insertions; a user-friendly feature could be to indicated inserted amino acids by various colours. However, this feature is planned for later versions of

STRENDA DB.

Modifications of the sequence are defined in the subsequent fields following the request "Please specify the type(s) of modification(s)".

Use UniProtKB numbering if numbering system is clear, give number, alternatively give part of surrounding sequence as follows:

• **N-terminal tag/fusion**, click in the box. The box is ticked and a free-text field appears. This field requires the entry of the tagging site in the following format: N-AGHL...MLGYH. Amino acids to the left of the dots represent the sequence of the tag, amino acids to the right of the dots represent the first five amino acids of the protein to which the tag is attached.

• **C-terminal Tag/fusion**: Format: MLGYH...AGHL-C

Amino acids to the left of the dots represent the five amino acids of the protein to which the tag is attached, amino acids to the right of the dots represent the sequence of the tag.

• **Insertion**: Format: MLGYH|AGHL|PLVHG

Amino acids to the left of the dots represent the five amino acids of the protein that are immediately N-terminal to the insertion, amino acids between the dots represent the insertion, amino acids to the right of the dots represent the five amino acids of the protein that are immediately C-terminal to the insertion. It would be helpful to indicate the site number of the last amino acid in the block of five before the insertion , e.g. 25, MLGYH|AGHL|PLVHG

• **Deletion**: Format: MLGYH|...|PLVHG

Amino acids to the left of the dots represent the five amino acids of the protein that are immediately N-terminal to the deletion, the dots represent the deletion, amino acids to the right of the dots represent the five amino acids of the protein that are immediately C-terminal to the deletion. It would be helpful to indicate the site number of the last amino acid in the block of five before the deletion, e.g. 25, MLGYH|...|PLVHG

• **Mutation**: Format: MLGYH|X->Y|PLVHG

Amino acids to the left of the dots represent the five amino acids of the protein that are immediately N-terminal to the mutation, the X->Y notation indicates that amino acid X was mutated to amino acid Y, amino acids to the right of the dots represent the five amino acids of the protein that are immediately C-terminal to the mutation. It would be helpful to indicate the site number of the last amino acid in the block of five before the mutation, e.g 25, MLGYH|X->Y|PLVHG

• **Other modification**: Please specify the modification by a short description or provide a reference.

Click No if the proteins used was without any sequence modifications.

Click Unknown, if it is not known whether the sequence has been modified. However, for this case a new field opens asking the author for providing information why this modification is not known.

Post-translational Modifications (PTM)

As for the ***Protein Sequence Modification***, first the author is asked whether the proteins contains sequence modifications.

Click Yes, if PTMs are known and/or experimentally carried out. The section expands by additional fields accompanied by a check box for each field:

Phosphorylation, Glycosylation, Acetylation, Hydroxylation, Methylation, Other Modifications.

With these five types of PTS more than 90% of all modifications are covered and can be entered. For the remaining 10% the field Other Modifications provides space for specification by a few words, 200 characters maximum.

The way of data entry is for all types of modifications the same: Select the modification, e.g. phosphorylation, the field expands and the author is asked at which amino acid residue the modification occurs. If one or more serine residues are phosphorylated, the corresponding check box is ticked and another field expands. In this field just the site number(s) of the phosphorylated serine residue is given.

> Example: 25, 100, 333.

Please note that in the current version of STRENDA DB there is no cross-validation between the protein sequence and site numbers of amino acids for PTMs. This means that the user can inadvertently indicate a PTM at a asparagine residue at position 150 even though there is no asparagine residue in the sequence. That's why the author is requested for entering this data with high accuracy.

Click No if the protein was not chemically modified. Nothing further happens but this indication will be annotated to the protein description in the summary report.

Click Unknown, if it is not known whether the protein was modified. In this case a new field opens asking the author for providing information why this information is not known.

Protein Source

The first question the author needs to answer is whether the protein was cloned and heterologously expressed from a plasmid.

Click Yes, if the protein was cloned and expressed in your lab. Detailed information need to be entered in Comments. This field is then compulsory.

Click No if it is a native protein, isolated and purified in your lab.

Click Unknown: The protein may be purchased from a commercial vendor. The user is requested to specify where the protein was purchased, and to provide the lot number etc.. The field Comments becomes compulsory.

Tissue: Specify the tissue from which the protein was isolated, e.g. skeletal muscle, liver, epithelium etc.

TaxonID: usually auto-completed. However, if this field appears blank it will be necessary to query this ID from the taxonomy browser provided by NCBI Taxonomy (http://www.ncbi.nlm.nih.gov/taxonomy) in connection with the organism from which the protein was isolated.

Organism: usually auto-completed. If not, the appropriate name of the organism from which the protein was isolated can be retrieved from NCBI Taxonomy (http://www.ncbi.nlm.nih.gov/taxonomy).

Strain: if the protein was isolated from microorganisms or from subspecies of plants, please specify the strain using the common nomenclature

Cell Type: please specify the cell type from which the protein was isolated, e.g. hepatocyte. If in doubt it is recommended to retrieve appropriate ontologies which can be accessed at http://cellontology.org/.

Localization: Please specify from which subcellular compartment the protein was isolated. Examples for subcellular compartments are cytosol, mitochondria, vacuole, lysosome etc.

When using the Search Protein functionality of STRENDA DB any compulsory fields in this section are automatically completed, i.e. Taxon ID and Organism.

Please note if either Taxon ID or Organism is changed there is no cross-check between these fields and thus must be manually checked. In addition, the user can feel free to enter further information, e.g. on Strain, Cell Type and Localization of the protein investigated. These fields are free text fields but the author is requested to use the commonly accepted terminologies as can be obtained from ontologies.

Protein Reaction

This section consists usually of three fields of which EC number and Reaction as in ExploreEnz are automatically completed after a successful UniProtKB search i.e. a protein has been found in UniProtKB and selected for STRENDA DB. The data neither in EC number nor in Reaction as in ExploreEnz are editable. If it is not possible to obtain any or complete EC number because the

protein investigated has not been listed in UniProtKB yet or several EC numbers are assigned to the protein or there is only a partial EC number available or due to other reasons, the user needs to comment this fact in the field Comments on Protein Reaction. This is a free text field restricted to 2000 characters.

In addition, if no unambiguous EC number is available or only a partial EC number was delivered from UniProtKB the field Reaction as in ExploreEnz is turned off and the author is requested to provide detailed information in Comments on Protein Reaction.

At this point the data input for the protein has been finish and can be aborted by clicking on CANCEL. The user returns to Manuscript Overview without storing the data entered previously.

Alternatively, the data input can be finalized by clicking on SAVE & SHOW EXPERIMENT. The author returns to Manuscript Overview with the data entered stored and a new row generated in the "List of Experiments" including the data on the protein added previously. For changes or additions, the user may click on Edit in the column 'Actions'.

Assay Conditions

The author returns to Manuscript Overview after saving the data entered in the page Add Protein. In the *List of Experiments* a new experiment is added along with a number of facts is displayed:

Description: the name of the experiment as entered previously;

Protein: the name (and common names) of the protein;

Organism: the name of the organism from which the protein was isolated;

UniProtKB AC: the accession number of the protein characterized in the experiment and as assigned to the protein;

SRN: STRENDA Registry Number but this field is still empty since the experiment has not been finalized yet and thus no SRN could be assigned to the experiment;

Finalized: since the input for this experiment has not been finalized, there is no finalization date;

Status: in this case still open, since the input of the experimental data is not finalized yet;

Actions: at this state, there is only one action provided, i.e. Add/Edit Data.

The author may create a new experiment by clicking on ADD EXPERIMENT. This operation leads the author to another Add Protein page asking for the input of the basic information about the protein

used in the experiment. The user also may click on MANAGE MANUSCRIPTS which leads him back to the Manuscripts Overview page.

In order to proceed with the data input for the protein investigated and for which the data where entered in the section ADD PROTEIN click on the corresponding Add/Edit Data in the column 'Actions'.

Experiment Overview

As Manuscript Overview is the central page for displaying all experiments that are going to be entered in STRENDA DB, Experiment Overview is the central page that presents all experimental subsets that each consist of a pair of assay conditions and results set.

After clicking on Add/Edit Data in the column 'Actions' within the *List of Experiments* the page Experiment Overview opens displaying the *Manuscript Data* box on the left-hand side and the *Experiment* box" on the right-hand side. Please note, that the *Experiment* box now shows the previously chosen name of the experiment along with the essential data on the protein. This data can be modified by clicking on Edit on the bottom of this box. The author returns to Add Experiment and by clicking on SAVE EXPERIMENT NAME to Add Protein where the protein's data can be changed.

After clicking on SAVE & SHOW EXPERIMENT and then after clicking on Add/Edit Data in the corresponding line of the experiment the author comes back to Experiment Overview.

The author may return to Manuscript Overview by clicking on MANUSCRIPT OVERVIEW on the bottom of the page. One also could click on FINALIZE EXPERIMENT. However, in this case, an error message is displayed explaining that no further information about the experiment was given.

Click on CANCEL to return to Experiment Overview.

In order to enter the assay conditions along with the result data click on ADD NEW EXPERIMENTAL SUBSET.

Add New Experimental Subset: Add Assay Conditions

First, the author is requested to name the subsequent experimental subsets. An experimental subset (ESS) refers to a set of all data for the same protein and consists of the assay conditions used and the results set obtained. For further details please refer to the Brief introduction in the overall concept.

The name of the experimental subset (ESS) just supports the author in the organization of data within STRENDA DB and looking about data already entered and still missing. This name does not

necessarily match with corresponding information given in the manuscript. The author may feel free also just to number the ESS.

Either click on CANCEL to return to Experiment Overview or

click on EDIT ASSAY CONDITIONS which leads the user to Assay Conditions.

Assay Conditions for Experimental Subset

This page defines the details of the assay.

After entering a name for the ESS the page Assay Conditions changes slightly by providing the user two tables (which are empty at the beginning): *Small Assay Component*s and *Macromolecular Components*. In addition the user is requested to enter the *Physical Properties* of the assay.

Small Assay Components are small molecules with a low molecular weight of up to 200 Daltons, organic compounds such as carbohydrates, amino acids, short-chained lipids as well as compounds from the secondary metabolism but also inorganic molecules such as salts.

Macromolecular components are defined here as large molecules with molecular weights above 200 Daltons such as polypeptides, glycans, lipids, auxiliary enzymes necessary for the assay etc.

Concentration of the assayed enzyme followed by the a field whose name is determined by the protein used in the assay. This field defines the basis on which kinetic data may be calculated. Not only real concentrations can be given but also masses and units.

Physical Properties mean pH, pD and Temperature either in °C or K.

After the input of the assay conditions the author clicks on EDIT RESULTS to enter the results data and to complete the ESS or on CANCEL to abort the data input. He then returns to Experiment Overview.

Add Small Assay Components

At the beginning the table on *Small Assay Components* is empty but after adding components it lists the following data:

Role: definition of the role of the component in the assay in dependence whether it is a substrate, product, inhibitor, activator, buffer, salt or other component. The role is defined for each component individually when adding compounds (see Edit Compound).

Compound Name: commonly used name of the compound;

ChEBI ID: Identifier of this component if found in The database and ontology of Chemical Entities of Biological Interest (ChEBI, http://www.ebi.ac.uk/chebi/);

PubChem CID: Identifier for this component found in PubChem (http://pubchem.ncbi.nlm.nih.gov/). Please note that the PubChem Project will be queried automatically only by STRENDA DB thus ChEBI IDs are provided only if the chosen compound is also indexed by PubChem.

Stoich. (=Stoichiometry): Stoichiometry of the component in the enzyme reaction. This number is 1 by default but can be modified by the user (see Edit Compound).

Concentration(s): initial concentration(s) used in the assay. These can be fixed concentrations as used for e.g. buffers and salts. But they also can be varied and thus are displayed as ranges with a minimum and maximum concentration in this field.

Actions: unless compounds are listed no actions are provided. With compound listed in this table there are following actions: Edit opens Edit Compound and enables the author to modify entries or even to change the compound previously selected completely. Delete deletes the compound from the table.

Click on ADD COMPONENT to add small assay compounds.

Edit Compound

This page is a subordinated section which however is a central page either to define the compounds used in the assay manually or to query PubChem for these compounds to become supported with the the major information about the corresponding molecules. In addition, on this page both the role for each compound used and the initial concentrations applied in the assay are defined. Since this procedure is carried out for any component in the assay, the author returns to this page many times.

It is recommended to query PubChem for the compounds (Search PubChem) used unless they are expected not to be indexed by this database since it is a newly synthesized compound. The advantage of this procedure is that the compound's definitions such as correct compound's name, the IUPAC name and the unambiguous identifier (either ChEBI ID or PubChem CID) is given. However, it is also possible to enter the compound's definitions manually but this may be very time-consuming.

Compound Search

Search PubChem: enter the compound's name used in the assay and click on SEARCH. A dialogue window opens providing a list of hits that match with the search string. The compounds found are accompanied by a ChEBI ID (if available) and a PubChem CID. Click on Select to choose the appropriate compound. Please note in case of no hits it is possible that there is no match between the search string and the PubChem entry. However, this does not necessarily mean that this compound is not indexed by PubChem but can be caused by the use of

different names or nomenclature. In this case it is recommended to query PubChem directly. After the selection of an appropriate compound the author returns to Edit Compound with the compulsory fields completed, i.e. Compound Name, and additional fields also completed, i.e. InChi, IUPAC, ChEBI ID (if available) and PubChem CID.

Compound Definitions

Role: Definition of the role the compound plays in the assay. Following roles are available: Substrate, Product, Buffer, Salt, Other Compound, Inhibitor and Activator. Any compounds that do not match with the defined roles given are supposed to be Other Compound. Substrate, product, inhibitor and activator refer to the enzyme reaction.

Roles once set for a compound cannot be changed later. If the author discovers that there is a mismatch between the compound used and the role it plays in the assay he then needs to delete this compound from the list of small assay components and to restart the entry.

Please also note that some compounds can play multiple roles, i.e. some are substrates and allosteric inhibitors. Some compounds activate at low concentration but inhibit at high concentrations. In this case the same compound needs to be added twice as assay component, once as substrate and once as inhibitor or once as activator and once as inhibitor, depending on the experiment carried out.

There is additional information needed for buffers. To enter a buffer, select from PubChem an appropriate salt form of the buffer compound with the correct counter-ion (e.g. Tris hydrochloride, potassium acetate). Then give the total concentration of the buffer compound in both protonation states, as in the usual convention (i.e. not the actual concentration of the individual PubChem compound). Together with the pH value entered elsewhere, this will fully characterise the buffer system used. In the case of buffers like phosphate, either the mono- or di-salt can be entered. If a buffer is made by, e.g. adjusting Tris hydrocholoride with NaOH, the additional NaCl generated may be declared as an added salt. Note that Tris base is named 'Tromethamine' in PubChem.

Stoichiometry: set to 1 by default but can be changed manually if necessary.

Initial Concentration is FIXED or VARIED: Fixed concentrations are usual e.g. for salts and buffers whereas substrate concentrations vary to obtain kinetic parameter. If clicked on FIXED a concentration value is entered in Value. If clicked on VARIED the Minimum Value and the Maximum Value for the concentration requires input. If 'Product' is assigned to a compound often any concentrations are not determined, thus the button UNKNOWN is activated preventing the author from entering any concentration values.

Do not use commas as thousands separators in large numbers. For decimal values decimal points must be used; e.g. 5.0.

Unit: Select the corresponding concentration unit from the pull down menu, i.e. M, mM, nM, µM.

Compound Name: automatically completed when the PubChem search was used otherwise input of the correct commonly used name.

InChi: is the IUPAC International Chemical Identifier which is a textual identifier for chemical structures that enables to encode molecular information and to facilitate the search in databases. If the InChi is not available the user needs not to generate it manually or by computing means. This field is automatically completed after a successful PubChem search.

IUPAC: name of the compound in accord to the IUPAC rules for naming chemical structures. If the entry of the compound was carried manually, the user should try to create an appropriate IUPAC name. This field is automatically completed after a successful PubChem search.

ChEBI ID: data identifier for molecules indexed by ChEBI. This field is automatically completed after a successful PubChem search.

PubChem CID: data identifier for molecules indexed by PubChem. This field is automatically completed after a successful PubChem search.

CANCEL: the user returns to Assay Conditions without any data stored in the table "Small Assay Components".

SAVE COMPOUND: Return to Assay Conditions with the data on the previously selected compound entered in the table "Small Assay Components".

Add Macromolecular Components

At the beginning the table on *Macromolecular Components* is empty but after adding components the following data are listed:

Role: definition of the role of the component in the assay in dependence whether it is a substrate, product, inhibitor, activator, buffer, salt or other component. The role is defined for each component individually when adding compounds (see Edit Macromolecule).

Class: definition of the chemical class of the component, i.e. peptide, lipid, carbohydrate, nucleotide, and other.

Compound Name: commonly used name of the compound;

Database used: Acronym or full name of the database used for the entry of macromolecular components' data, e.g. ChEBI, UniProt etc.;

Identifier: database identifier of the component used;

Stoich. (=Stoichiometry): Stoichiometry of the component in the enzyme reaction. This number is 1 by default but can be modified by the user (see Edit Macromolecule).

Concentration(s): initial concentration(s) used in the assay. These can be fixed concentrations as used for e.g. buffers and salts. But they also can be varied and thus are displayed as ranges with a minimum and maximum concentration in this field.

Actions: unless compounds are listed no actions are provided. With compound listed in this table there are following actions: Edit opens Edit Macromolecule and enables the author to modify data entries or even to change the compound previously selected completely. Delete deletes the compound from the table.

In order to add information about macromolecular components used in the assay click on ADD COMPONENT on the bottom of the table "Macromolecular Components".

Edit Macromolecule

This page is a subordinated section which however is a central page either to define the macromolecular compounds used in the assay manually or to query (in this version) PubChem for these compounds to become supported with the the major information about the corresponding molecules.

Please note this section is still provisional. For user support, currently for data just PubChem is queried. For later versions it is planned to link to additional databases.

In analogy to the input of small assay component data in this page both the role for each compound used and the initial concentrations applied in the assay is defined. Since this procedure is carried out for any macromolecular component in the assay, the author can return to this page many times.

It is recommended first to query PubChem for the compounds (Search PubChem) used unless they are expected not to be indexed by this database since it is obvious that this component is listed by other databases as for example for nucleotides. The advantage of going through PubChem is that the compound's definitions such as correct compound's name, the IUPAC name and the unambiguous identifiers (either ChEBI ID or PubChem CID) are given.

For macromolecules, however, presumably the compound's definitions need to be entered manually. This is especially true for auxiliary enzymes or coupling enzymes need to be given which are subsumed under Other Compounds.

Compound Search

Search PubChem: if it is obvious that the macromolecule should be indexed by PubChem enter the compound's name used in the assay and click on SEARCH. A dialogue window opens providing a list of hits that match with the search string. The compounds found are accompanied by a ChEBI ID (if available) and a PubChem CID. Click on Select to choose the appropriate compound. Please note in case of no hits it is possible that there is no match

between the search string and the PubChem entry. This does not necessarily mean that this compound is not indexed by PubChem but can be caused by the use of different names or nomenclature. In this case it is recommended to query PubChem directly. After the selection of an appropriate compound the author returns to Edit Macromolecule with the compulsory fields completed, i.e. Compound Name, and additional fields also completed, i.e. InChi, IUPAC, ChEBI ID (if available) and PubChem CID.

If the component used is not indexed by PubChem and the author is aware of the chemical class such as carbohydrate, peptide, lipid or nucleotide he is recommended to query those databases that appear more suitable for such searches.

Macromolecules Definitions

Role: Definition of the role the compound plays in the assay. Following roles are available: Substrate, Product, Buffer, Salt, Other Compound, Inhibitor and Activator. Any compounds that do not match with the defined roles given are supposed to be Other Compound. Substrate, product, inhibitor and activator refer to the enzyme reaction.

Roles once set for a compound cannot be changed later. If the author discovers that there is a mismatch between the compound used and the role it plays in the assay he then needs to delete this compound from the list of small assay components and to restart the entry.

Please also note that some compounds can play multiple roles, i.e. some are substrates and allosteric inhibitors. Some compounds activate at low concentration but inhibit at high concentrations. In this case the same compound needs to be added twice as assay component, once as substrate and once as inhibitor or once as activator and once as inhibitor, depending on the experiment carried out.

There is additional information needed for buffers. To enter a buffer, select from PubChem an appropriate salt form of the buffer compound with the correct counter-ion (e.g. Tris hydrochloride, potassium acetate). Then give the total concentration of the buffer compound in both protonation states, as in the usual convention (i.e. not the actual concentration of the individual PubChem compound). Together with the pH value entered elsewhere, this will fully characterise the buffer system used. In the case of buffers like phosphate, either the mono- or di-salt can be entered. If a buffer is made by, e.g. adjusting Tris hydrocholoride with NaOH, the additional NaCl generated may be declared as an added salt. Note that Tris base is named 'Tromethamine' in PubChem.

Stoichiometry: set to 1 by default but can be changed manually if necessary.

Initial Concentration is FIXED or VARIED: Fixed concentrations are usual e.g. for salts and buffers whereas substrate concentrations vary to obtain kinetic parameter. If clicked on FIXED a concentration value is entered in Value. If clicked on VARIED the Minimum Value and the

Maximum Value for the concentration requires input. If 'Product' is assigned to a compound often any concentrations are not determined, thus the button UNKNOWN is activated preventing the author from entering any concentration values.

Do not use commas as thousands separators in large numbers. For decimal values decimal points must be used; e.g. 5.0.

Unit: Select the corresponding concentration unit from the pull down menu, i.e. M, mM, nM, µM, mg ml^{-1}, katal, IU.

Class: definition of the chemical class of this component. There are following classes pre-defined: peptide, carbohydrate, nucleotide, lipid, other.

Compound Name: automatically completed when the PubChem search was used otherwise input of the correct commonly used name.

InChi: is the IUPAC International Chemical Identifier which is a textual identifier for chemical structures that enables to encode molecular information and to facilitate the search in databases (IUPAC International Chemical Identifier Project Page). If the InChi is not available the user needs not to generate it manually or by computing means. This field is automatically completed after a successful PubChem search.

IUPAC: name of the compound in accord to the IUPAC rules for naming chemical structures. If the entry of the compound was carried manually, the user should try to create an appropriate IUPAC name if possible. This field is automatically completed after a successful PubChem search.

Database used: determination of the database that was used for querying and input of data in STRENDA DB. Acronyms such as ChEBI, PubChem or UniProt can be used but also full names if appropriate along with home URL of the database. This field is automatically completed after a successful PubChem search.

Identifier: data identifier for the molecule indexed in the database determined above. This field is automatically completed after a successful PubChem search.

CANCEL: the user returns to Assay Conditions without any data stored in the table "Macromolecular Components".

SAVE COMPOUND: Return to Assay Conditions with the data on the previously selected compound entered in the table "Macromolecular Components".

Both tables, **Small Assay Components** and **Macromolecular Components** may fill up to several entries in dependence of the complexity of the assay. The last definitions which need to be given are the physical properties.

Concentration of the assayed protein

This very brief section refers to the amount of protein used in the assay which will be the basis for the calculations of the kinetic parameters. The field name is defined by the name of the protein as approved in and obtained from Add Protein. The user just needs to enter the number together with the appropriate unit. Due to the variety of amounts of enzyme used in assays the unit drop down menu provides the corresponding variety of units such as mass per volume, molar concentrations, units per volume as well as percentage amounts of enzyme.

The enzyme concentration is a mandatory field.

Physical Properties

In this section pH, pD and the temperature at which the enzyme was assayed are defined. It is possible either to enter pH or pD or both. Temperature can be given in degrees centigrade (°C) or Kelvin (K). Please note that there is no inter-conversion between these temperature scales which means that a value of 25 °C won't be calculated to 293 K automatically and vice versa.

On the bottom of Assay Conditions there are two buttons: CANCEL and ADD/EDIT RESULTS.

CANCEL returns the user to Experiment Overview without saving any data.

He can resume the data entry by clicking on...

EDIT RESULTS: the author is led to Result Set and requested for entering the kinetic parameters that are obtained under the experimental conditions previously defined.

Result Set

The Result Set is the second part of an Experimental Subset which provides the experimental results that were obtained under the experimental conditions used as defined in Assay Condition.

On the top of this page the usual information is provided about the manuscript data and the experiment data (defined by the protein data) in the corresponding boxes. Following in the section **Results for <title of ESS>** the kinetic parameters for both the substrates (left-hand side table) and products (right-hand side table) are shown in tabular format. The procedure for both is the same that's why the entry process is not described separately but for both together.

When arriving here at first time the tables only show the substrates and products previously defined that are displayed in the column "Name". The neighboured column "Values" is empty and the column 'Actions' provide the links to Edit Data and Clear (after data were entered). The column 'Values' will display all kinetic data entered.

Click on <u>Edit Data</u> to proceed with the input of the kinetic parameters for either the substrate for which the kinetic data were obtained or for the corresponding product.

Edit Kinetic Parameters

The page Edit Kinetic Parameter first displays the both information boxes on the manuscript data and the experiment following by the fields for the entry of the kinetic parameters. On top of the input section the name of the substrate/product for which the parameters are going to be entered is recalled to make sure that the author enters the right data to the corresponding compound.

Please note the following requirements for a successful data input and restrictions that derive from this early implementation stage:

- at least one kinetic parameter together with the standard error needs to be entered to pass the subsequent validation process;
- there is no validation between interconnected kinetic parameters such as K_M and k_{cat}/K_M. In the case of unintentional, erroneous data entry these data can be inconsistent;
- in this version a normalization for values has not been implemented yet. This means that there is neither a feature implemented that could calculate values such as V/K_M from the single values V and K_M nor that there is an interconversion of units, e.g. 0.05 M to 50 mM.

After the entry of values the author either can click on CANCEL to abort the data input and to return to Result Set without any data stored or click on SAVE to store the data entered and to return to Result Set displaying these kinetic data for the selected compound.

This data can be removed completely by clicking on <u>Delete</u> or modified/corrected/changed by clicking on <u>Edit Data</u> which returns the author to Edit Kinetic Parameter to review and change the data.

On the page Edit Kinetic Parameter the author can save the experimental results after all kinetic data had been entered by clicking on SAVE. This returns the author to Experiment Overview which now displays the complete ESS.

Edit Inhibition Parameters

If the assay conditions include a compound with the role inhibitor, the user is requested to enter inhibition parameters. This becomes obvious when an additional table below the table Kinetic Parameters appears: Inhibition Parameters.

The author may click on <u>Edit Data</u> in the column 'Actions' within this table in order to enter inhibition data.

The page Edit Inhibition Parameter opens and first asks the author whether the inhibition was reversible.

If YES, the table displays the compounds defined in the assay as substrates and products and data can be entered for the corresponding compound by clicking on Edit Data in the column 'Actions'.

A new page opens asking in a pull down menu whether the inhibition was competitive, uncompetitive or mixed. After the selection of the appropriate inhibition type the user is requested to enter the kinetics data for the inhibition along with the statement whether the inhibition was linear or hyperbolic (Math. function). After completion of the mask, the user returns to Edit Inhibition Parameter by clicking on SAVE. Now, he finds the table filled with the data entered previously.

If NO, the author is requested to select the appropriate inhibition type from the pull down menu which provides the following types: tight, covalent, other. After selection, the author is requested to enter the kinetics of the inhibition in the fields. CANCEL aborts the action and returns him to Results Set without storing any data. SAVE returns the author to Results Set with storing and displaying the data entered.

If UNKNOWN, the author is requested to provide the values for IC_{50} inhibition. CANCEL aborts the action and returns him to Results Set without storing any data. SAVE returns the author to Results Set with storing and displaying the data entered.

On Result Set any of the data previously entered can be modified by clicking on Edit Data or deleted by clicking on Clear in the 'Actions' column. The data entry can be finalized by clicking on SAVE which returns the user to Experiment Overview which displays the newly added ESS.

Edit Activation Parameters

If the assay conditions include a compound with the role inhibitor, the user is requested to enter inhibition parameters. This becomes obvious when an additional table below the table Kinetic Parameters appears: Activation Parameters.

The author may click on Edit Data in the column 'Actions' within this table in order to enter activation data. The page Edit Activation Parameter opens and displays data fields that are arranged in three sections.

The first section asks for values for *Affinity Constant* and whether this data is true or apparent. Click on one of either buttons to indicated whether the data was obtained under TRUE or APPARENT conditions.

The second sections requests the author to enter data on the *Reaction Velocity* at zero concentration of the activator (control experiment) and maximum amount of activator used in the experiment, both connected with the indication whether this data is TRUE or APPARENT. In addition, the user is

asked whether the maximum concentration used in the experiment was saturating.

If yes, click on YES.

If no, click on NO.

If unknown, click on UNKNOWN. No further details are required.

Please note that this is just a qualitative statement.

The third section refers to *influence of the activator on the Michaelis-Menten constant* of the substrate accompanied by the buttons YES, NO, UNKNOWN.

If YES, the user is requested to select the appropriate substrate and click on Edit Data to enter further details on the activation. The page Edit Michaelis Constant opens and requires the data on the effect on the Michaelis-Menten constant with minimum and maximum concentration of the activator for the substrate previously selected. For both data it needs to be indicated whether they were TRUE or APPARENT.

If NO or UNKNOWN, no further action is required from the author.

Last but not least, the author is requested to make a statement on whether the activator was used with saturating concentration. The statement can be given by just clicking on either of the buttons YES, NO, UNKNOWN.

The data on the effect on the Michaelis-Menten constant can be either dropped by clicking on CANCEL (return to Edit Activation Parameter without storing any data) or stored by clicking on SAVE (return to Edit Activation Parameter with displaying this data in the table that follows the question whether the activator influences Michaelis-Menten constant.

CANCEL returns the user to Result Set without storing any data.

SAVE returns the user to Result Set storing and displaying all data entered previously.

Any data can be modified after review by clicking on Edit Data in the column 'Actions'. In addition any data can be deleted by clicking on Clear in the same column.

Click on SAVE to return to Experiment Overview which displays the additional ESS (on the bottom) along with the ESS created previously.

Review, add new ESS or finalization of the experiment?

On Experiment Overview the author has several options (Fig. 2):

1. review and modify data,

2. add new ESS,

3. finalize the experiment,

4. or to return to the manuscript overview.
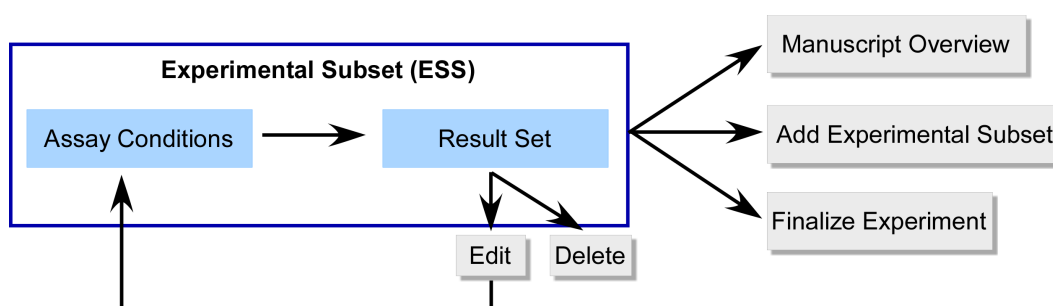
Figure 2: Editing loop within an Experimental Subset. For the modification of any data either within the assay conditions or the result set the author always starts with the form for the assay conditions.

These options are discussed below.

1. Review and modify data

In the case of an error detected in either the assay conditions or the experimental results the author can change the data by clicking on Edit. This guides the user through a loop consisting of return to Assay Conditions with the request for entering a name for the ESS (Fig. 2). This can be skipped by clicking on SAVE SUBSET NAME (since usually the name needs not to be changed) to reach the following form to enter the assay conditions for the corresponding ESS. This can be skipped by clicking on EDIT RESULTS if modifications need to be carried out with the kinetic parameters only.

Please note this loop is the sole way for any modifications within the ESS to ensure consistency between the assay conditions and the kinetic data and to keep integrity of the entire ESS.

2. Add new ESS

Further experimental subsets that are based on the pre-defined experiment that includes the same protein can be created by clicking on ADD EXPERIMENTAL SUBSET. The author returns to Assay Conditions where he is requested to enter a new name of the new ESS followed by the same process as described above. First, the assay conditions need to be entered. Here, the author is provided with those data that have been entered previously and thus only those parameters need to be changed (e.g. pH, temperature) that were already defined.

Alternatively, parameters can be added (inhibitors, activators, etc.) or replaced (substrates, products, etc.). Thus, the definition of a new assay condition is limited to alterations which speeds up the entire entry process.

Click on EDIT RESULTS to proceed with the input of the kinetic parameter. Here, the author won't find any data from previous result sets and thus has to enter those data that correspond with the assay conditions. If the review process requires that errors need to be corrected, the author runs through the ESS loop (see Fig. 2).

The author now has the choice between clicking on FINALIZE EXPERIMENT and MANUSCRIPT OVERVIEW. Since the finalization of an experiment after the entry of one or more ESS is supposedly the typical way this will be described first in the following. In the case of clicking on MANUSCRIPT OVERVIEW before the finalization of the experiment, please see Manuscript Overview -1-.

3. Finalize Experiment

After all ESS had been entered (and reviewed for accuracy and comprehensiveness by the author!) the author completes the current experiment by clicking on FINALIZE EXPERIMENT. The page Experiment Overview changes by showing a new section: *Finalize Experiment*.

Beneath the section title a box appears, ideally with green background stating that the data input complies with the STRENDA guidelines and thus the author can complete the experiment finally by clicking on FINALIZE. Alternatively, the author can abort the finalization by clicking on CANCEL which leads him back to Experiment Overview.

After clicking on FINALIZATION, the completion of the data input is confirmed and a STRENDA registry number (SRN) is displayed. The user does not need write this SRN down since the entire data set together with the SRN can be exported (see Export data)

On the bottom of the page there are several options with regard to the export function but also to return to Manuscript Overview without exporting data. At any time it is possible to return to this page in order to have access on the export formats provided.

Manuscript Overview -1-

After clicking on MANUSCRIPT OVERVIEW the author returns to Manuscript Overview showing the *List of Experiments* without any changes for this experiment in all columns. Return to either edit or finalize the experiment by clicking on Edit. The author arrives on Experiment Overview displaying all experimental subsets previously entered. Here, the buttons MANUSCRIPT OVERVIEW, ADD EXPERIMENTAL SUBSET and FINALIZE EXPERIMENT are provided, again.

4. Manuscript Overview -2-

After the finalization of the experiment the user may return to Manuscript Overview. Now, some items were added to the corresponding experiment in the *List of Experiments*:

- a SRN is assigned to the experiment,
- finalization date is added,
- 'Status' was changed from open to finalized,
- 'Actions' do not give writing access to the data by Edit but reading access by the following links:
  - Show: opening of Experiment Overview displaying all data entered;
  - Export PDF: opening a dialogue window asking for either saving the file or opening the file with the default PDF viewer;
  - Export XML: opening a dialogue window asking for either saving the file or opening the file with the default XML viewer.

On the bottom the additional buttons allow for returning to Manuscript Overview and to add a new experiment (ADD EXPERIMENT).

Export data

Data from finalized experiments can be exported via two ways. The typical way would be after finalization and assignment of the SRN to the experiment when the buttons EXPORT PDF and EXPORT XML are displayed.

After clicking on EXPORT PDF a dialogue window opens asking for either saving the file or opening the file with the default PDF viewer. This PDF can be stored on the local computer and submitted to the journal together with the original manuscript.

After clicking on EXPORT XML a dialogue window opens asking for either saving the file or opening it using the default XML viewer. This file also can be stored on the local computer and in

principal used for the upload to third party's databases. However, since this XML file would need to be parsed for other's purposes at this stage this is just an indication for STRENDA DB being open for data exchange. Any further requirements need to be discussed with the interested community.

Finalizing Manuscript

After the finalization of all experiments the author returns to Manage Manuscripts and can either add further experiments by Edit or delete all data that derive from this manuscript by Delete (and the manuscript is deleted, too). The last step is to finalize the manuscript itself which means that all experimental data from the manuscript were entered and the manuscript is ready for submission to the journal along with the exported PDF file(s) as supplementary information.

Click on FINALIZE a dialogue window opens asking whether the author is really willing to finalize this manuscript. CANCEL returns the author to Manage Manuscripts without any action taken. FINALIZE returns the author to Manage Manuscripts. The author now has read-only access to this manuscript via Show. Clicking on Show guides the author to Manuscript Overview providing him read-only access to all experiments. In addition, all data can be exported as PDF and XML file.

In the case of revisions of these entered data necessary, the manuscript can be reopened on request at strenda-db@beilstein-institut.de.

Subject Index