

Enzyme kinetics informatics: from instrument to browser

Neil Swainston^{1,*}, Martin Golebiewski^{2,*}, Hanan L. Messiha¹, Naglis Malys¹, Renate Kania², Sylvestre Kengne², Olga Krebs², Saqib Mir², Heidrun Sauer-Danzwith², Kieran Smallbone¹, Andreas Weidemann², Ulrike Wittig², Douglas B. Kell¹, Pedro Mendes^{1,3}, Wolfgang Müller², Norman W. Paton¹ and Isabel Rojas²

1 Manchester Centre for Integrative Systems Biology, University of Manchester, UK

2 Heidelberg Institute for Theoretical Studies, Germany

3 Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA, USA

Keywords

data analysis; database; enzyme; kinetics; metadata

Correspondence

N. Swainston, Manchester Centre for Integrative Systems Biology, University of Manchester, Manchester M1 7DN, UK
Fax: +44 161 306 8918
Tel: +44 161 306 5146
E-mail: neil.swainston@manchester.ac.uk
Website: <http://www.mcisb.org>

*These authors contributed equally to this work

(Received 31 May 2010, revised 20 June 2010, accepted 13 July 2010)

doi:10.1111/j.1742-4658.2010.07778.x

A limited number of publicly available resources provide access to enzyme kinetic parameters. These have been compiled through manual data mining of published papers, not from the original, raw experimental data from which the parameters were calculated. This is largely due to the lack of software or standards to support the capture, analysis, storage and dissemination of such experimental data. Introduced here is an integrative system to manage experimental enzyme kinetics data from instrument to browser. The approach is based on two interrelated databases: the existing SABIO-RK database, containing kinetic data and corresponding metadata, and the newly introduced experimental raw data repository, MeMo-RK. Both systems are publicly available by web browser and web service interfaces and are configurable to ensure privacy of unpublished data. Users of this system are provided with the ability to view both kinetic parameters and the experimental raw data from which they are calculated, providing increased confidence in the data. A data analysis and submission tool, the KINETICS-WIZARD, has been developed to allow the experimentalist to perform data collection, analysis and submission to both data resources. The system is designed to be extensible, allowing integration with other manufacturer instruments covering a range of analytical techniques.

Introduction

The field of systems biology is heavily reliant on reliable experimental data in order to create predictive models. With the establishment of high-throughput technologies in genomics, proteomics and metabolomics over the past decade, the amount of data available to the biochemistry community is increasing exponentially. However, the collection and dissemination of experimental data can be a labour-intensive process, such that much acquired data never becomes available to the community in an accessible and utilizable form. Thus, the data flow from the experiment to the consumer performing the analysis, the comparison

or the set-up of computer models can still constitute a bottleneck. This problem calls for systems that capture the data directly from the experimental instrument, process and normalize it to agreed standards and finally transfer these data to publicly available databases to make them accessible.

To facilitate the dissemination of data, a number of initiatives have been developed to advise on the minimum requirements to follow in the storage and dissemination of experimental data in fields such as transcriptomics and proteomics, which will ultimately allow data to be easily and freely shared between

Abbreviations

SBML, Systems Biology Markup Language; SBRML, Systems Biology Results Markup Language; STRENDA, Standards for Reporting Enzymology Data; XML, Extensible Markup Language.

laboratories worldwide [1–3]. The enzyme kinetics community is active in this area with the development of both standardized experimental operating procedures [4] and recommendations on data storage in the form of the Standards for Reporting Enzymology Data (STRENDa, <http://www.strenda.org>) [5] guidelines.

There exist several publicly available databases containing enzyme kinetics data that adhere to these recommendations, with BRENDA [6], a database for enzyme functional data, and the biochemical reaction kinetics database, SABIO-RK [7], as the two most comprehensive and most used examples. The data gathered in these resources were typically manually extracted from the biochemical literature and entered into the databases by hand, a labour-intensive and time-consuming process. To support this manual work SABIO-RK offers a tailored input interface [8], which allows users to manually enter kinetic data and corresponding metadata, utilizing standardized terms in the form of controlled vocabularies and references to external resources. BRENDA recently also introduced the support of kinetic parameter submission. These input interfaces could in principle assist experimenters in submitting their kinetic data to the databases. However, entering each dataset manually can be a tedious and error-prone process and is unlikely to be accepted as standard practice by the scientific community. To date there has been no support for automated submission of kinetic data or for storage of the original raw experimental data from which these constants were calculated.

We here introduce an automated system to support the whole workflow of deriving kinetic data from the laboratory instrument and make it accessible in the web. The task of managing enzyme kinetics data involves four steps: data capture, analysis, submission and querying/visualization. The first three tasks have been integrated in a unified tool, the KINETICSWIZARD. Data querying and visualization are provided by web browser interfaces for manual access and web services for automated access to both the newly developed MeMo-RK and the existing SABIO-RK databases.

The KINETICSWIZARD, introduced here, provides a unified interface for capturing and fitting raw kinetics time series data along with sufficient metadata to allow these data to be queried, such as detailed and unambiguous descriptions of the reactions studied, their reactants and modifiers, and experimental conditions. Data can then be automatically submitted to the relevant data repositories. By collecting this in a principled manner, the intention is that any data collected and submitted to the repositories will be complete, consistent and adhere to defined standards, such as the STRENDa recommendations.

Although much of this work has been developed in the context of systems biology, the tools described are sufficiently generic to be used in other fields, such as molecular enzymology and drug discovery.

Results

Data capture, analysis and submission

KINETICSWIZARD data capture

The key to ensuring that resources storing enzyme kinetics data can be usefully employed in a systems biology environment is in the richness and accuracy of the metadata associated with the kinetic constants. Specifically, for instance, we need to know the experimental conditions under which *in vitro* assays were performed, such as pH, temperature and buffer. Additionally, the components of the assay, such as enzyme variants, substrates, products and modifying molecules, must be unambiguously defined.

Designed to be used by experimentalists rather than bioinformaticians, the KINETICSWIZARD is intended to hide much of the more technical aspects of data management from the user and present an intuitive, user-friendly interface from which this necessary metadata can be obtained. The KINETICSWIZARD can be launched automatically from the instrument software, allowing data to be captured, analysed and submitted to databases immediately upon acquisition. The KINETICSWIZARD has been developed initially to integrate with a BMG Labtech NOVostar instrument (Offenburg, Germany). A generic version, which reads experimental data from a spreadsheet, along with an example of experimental data in this spreadsheet format, is also available (<http://www.mcisb.org/resources/kinetics/>). The system has been designed in a modular manner to allow the support of different instruments and experimental techniques (see Fig. 1).

In a typical experimental set-up, the user runs several time series assays, in which a reactant concentration is varied. The WIZARD allows the user to specify these varying reactant concentrations, which are then associated with the experimental data, and used in the subsequent fitting step to calculate kinetic parameters. A number of assays can be 'grouped' together, supporting experimental set-ups in which numerous reactions are assayed on a single plate.

To provide this functionality, the tool draws heavily on the use of existing data resources that are relevant to the task, and queries these resources via web service interfaces where possible. Exploiting existing data resources has the advantages of greatly reducing the volume of metadata that the experimentalist must

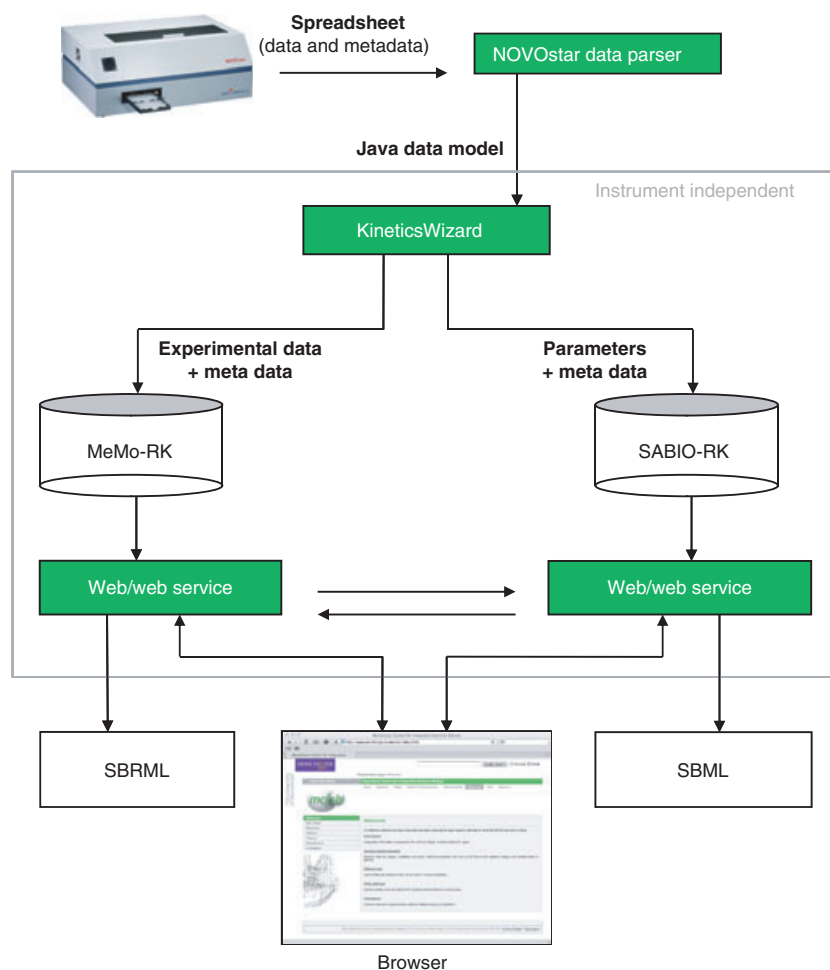


Fig. 1. Enzyme kinetics from instrument to browser. Data are extracted from the NOVOstar instrument as a Microsoft Excel spreadsheet. They are parsed into a data model and imported into the KINETICSWIZARD. The KINETICSWIZARD provides a graphical user interface that allows the experimenter to associate metadata to the experimental data. Kinetic constants are then calculated and the data submitted to appropriate repositories: MeMo-RK (<http://www.mcisb.org/MeMo-RK/>) for the experimental raw data, and SABIO-RK (<http://sabio.h-its.org/>) for the derived kinetic parameters, equations and appropriate metadata. Links are maintained between the repositories allowing both raw data and parameter sets to be accessed through web browser interfaces and web services. Kinetic data can be exported from SABIO-RK in SBML format and experimental data exported from MeMo-RK in SBRML format.

submit, while also annotating the submitted data with standard ontological terms to facilitate subsequent querying.

An example is in the specification of the reaction itself. It has been reported that relying on textual descriptions of small molecules and enzymes can result in inconsistencies, as the naming of such species is largely subjective and can differ greatly from individual to individual [9]. The KINETICSWIZARD ensures the consistent specification of reaction components by utilizing libAnnotationSBML [10], a library that provides an interface to the KEGG database [11]. The user specifies an organism and a gene name, from which the KEGG web service is queried and all reactions catalysed by the enzyme encoded by the supplied gene are returned (see Fig. 2). An individual reaction can then be selected, and the corresponding database entry queried to harvest a number of terms that would otherwise have to be specified manually by the user, such as EC term and the identity of reactants, products and enzyme. By utilizing KEGG reactions in this way,

reaction participants are specified internally as entries in either the KEGG or the ChEBI [12] databases, and enzymes as UniProt [13] terms. Accurate stoichiometry of each of the reaction participants is also gathered. This provides an unambiguous, computer-readable 'signature' for the specified reaction, which facilitates subsequent querying of the data themselves.

Situations may arise in which reactions are being studied that are not in the KEGG database. Future iterations could query other sources containing such data, such as Reactome [14], BRENDA or SABIO-RK itself. Alternatively, the user interface could be extended to allow the user to specify the reaction manually. However, this approach would put a greater burden on the user, and would increase the likelihood that inconsistent reactants, enzymes, EC terms, etc., would be input.

After defining the reaction, the user is provided with the facility to specify buffer reagents and coupling enzymes, along with other metadata values, including the environmental conditions, such as pH and temperature, under which the assays were performed.

Fig. 2. Specifying the reaction components. Upon specification of an organism and a gene, a search is performed against the KEGG web service, allowing the user to select from a list of reactions. The user can then specify the direction of the reaction, and which substrate concentration was varied during the assays.

In order to ensure that a given parameter is used in the intended manner, it is also necessary to specify the kinetic mechanism and equation that was used to determine the parameter. The initial version of the KINETICSWIZARD assumes that all reaction mechanisms follow irreversible, steady-state Henri–Michaelis–Menten kinetics [15]. Future releases of the KINETICSWIZARD will support more complex mechanisms, for example in cases where inhibition or allosterism is observed. The kinetic mechanism and all kinetic parameters are specified internally, and later archived with, unambiguous terms from the Systems Biology Ontology [16].

Utilizing existing bioinformatics resources provides the twin advantages of reducing the burden on the experimentalist in redefining metadata that are already present digitally elsewhere, while also ensuring the consistency of the metadata, aiding subsequent comparisons, analyses and reuse of data from different experiments or different laboratories.

v_{\max} parameters are often specified without any indication of the enzyme concentration contained within the term. To prevent this, the KINETICSWIZARD captures the enzyme concentration used in the assay, allowing the kinetic parameter to be submitted as a k_{cat} value. This decouples the parameter from the enzyme concentration and increases the usability of the value. To facilitate this further, standard units are specified for all parameters, with substrate and enzyme concentrations input in mM and nM, respectively.

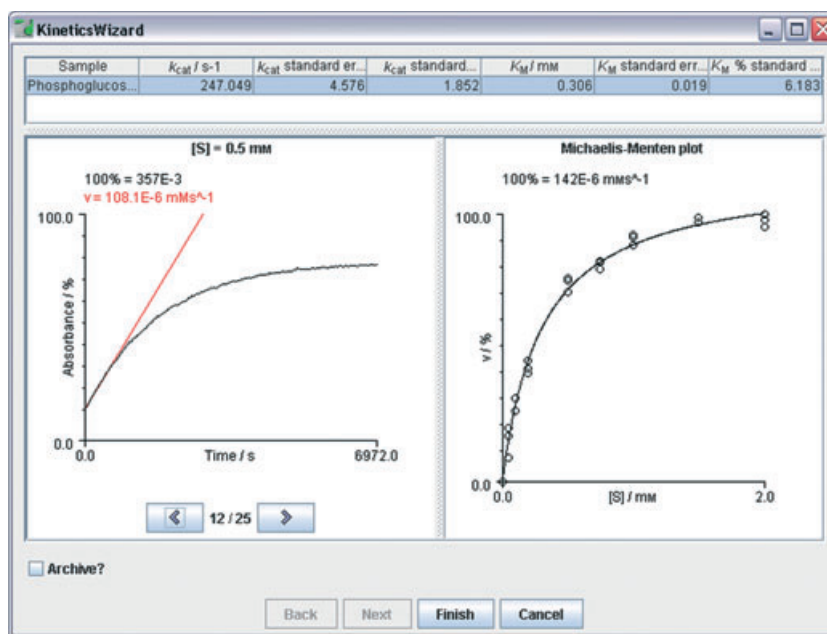
Finally, a free text field is available, allowing the user to assign notes and comments to the dataset.

KINETICSWIZARD data analysis

Following the data capture phase, the next step before data submission is data analysis [17], whereby kinetic parameters are determined by applying an appropriate fitting algorithm to the experimental time series data. By default, the initial version of the KINETICSWIZARD provides a fitting algorithm that assumes irreversible Henri–Michaelis–Menten kinetics. As the tool develops further, fitting to other more complex kinetic mechanisms will be supported.

During the experimental set-up, individual assays may be specified as being either samples or blanks. Blanks are assays that contain all components apart from the enzyme under investigation, and if present their data are subtracted from those of the sample assays. A straight-line fit is then used to estimate initial reaction rates. These values are then fed into the Eadie–Hofstee linearized version of the Michaelis–Menten equation [18,19] to provide estimates of k_{cat} and K_{M} . More accurate parameter values are subsequently obtained through nonlinear regression via the Levenberg–Marquardt algorithm [20,21]. Although the curve-fitting algorithm is automated, the user is provided with a visual representation of the fit from which the initial rate is calculated. The user may then perform a manual refit by dragging the initial rate line; a feature that can be utilized to correct for lag times of coupling enzymes, for example. Overriding the automated initial rate calculation will update the calculated k_{cat} and K_{M} parameters in real time (see Fig. 3).

Fig. 3. Displaying and manipulating the results of the curve-fitting algorithm. The left-hand panel allows the user to view each assay in the data set and its automatically fitted initial rate. The red initial-rate line may be manually corrected by dragging, allowing the default fit to be overridden for noisy or anomalous data. These initial rates are plotted against substrate concentration in the right-hand panel, which shows the Michaelis–Menten curve. The top panel shows the calculated kinetic parameters k_{cat} and K_{M} , together with their standard errors. Manually correcting an initial rate updates both the Michaelis–Menten curve and the calculated kinetic parameters in real time.



In order to test and validate the KINETICSWIZARD fitting algorithm, home-produced enzymes have been assayed (see Materials and methods). A number of time series assays were acquired for each enzyme, and the data captured and analysed using the KINETICSWIZARD. The calculated kinetic parameters were comparable with those calculated by the GRAFIT software package (Erithacus Software Ltd, Horley, UK), version 5.0 (see Table 1).

KINETICSWIZARD submission tool

The data submission task is two-fold: submission of the raw experimental data to MeMo-RK and submission of derived kinetic equations with their kinetic parameters and corresponding metadata to SABIO-RK.

MeMo-RK is a derivation of the MeMo database, originally constructed for storage of metabolomics

data [22]. It has been amended to store raw, experimental kinetics data and associated metadata, including submitter, laboratory, instrument settings and experiment type, such as absorbance or fluorescence.

Derived, secondary data in the form of kinetic parameters and equations, definitions of the reactions being studied and relevant metadata describing the experimental and environmental conditions such as temperature, pH, buffer solution, coupling enzymes are represented in an Extensible Markup Language (XML) document and submitted directly to the SABIO-RK submission web service. SabioML, a novel XML schema, has been developed for this purpose and could also serve as a kinetic data transfer format between sources other than SABIO-RK. Derived from the SABIO-RK database schema [23], it comprises kinetic laws, parameters and relevant metadata in a structured and standardized format, exploiting a controlled vocabulary and appropriate

Table 1. Comparison of kinetic parameters calculated by the KINETICSWIZARD and GRAFIT. Detailed views of the reaction, parameters and metadata can be found at the appropriate SABIO-RK records, <http://sabio.h-its.org/kineticLawEntry.jsp?kinlawid=29371>, <http://sabio.h-its.org/kineticLawEntry.jsp?kinlawid=29401> and <http://sabio.h-its.org/kineticLawEntry.jsp?kinlawid=29390>, respectively).

Enzyme	KINETICSWIZARD	GRAFIT
Fructose-bisphosphate aldolase (ALF1_YEAST, EC: 4.1.2.13)	k_{cat} : $4.14 \pm 0.061 \text{ s}^{-1}$ K_{M} : $0.451 \pm 0.024 \text{ mM}$	k_{cat} : $4.27 \pm 0.097 \text{ s}^{-1}$ K_{M} : $0.442 \pm 0.037 \text{ mM}$
Pyruvate decarboxylase isozyme 2 (PDC5_YEAST, EC: 4.1.1.1)	k_{cat} : $1.78 \pm 0.037 \text{ s}^{-1}$ K_{M} : $11.4 \pm 0.65 \text{ mM}$	k_{cat} : $1.79 \pm 0.029 \text{ s}^{-1}$ K_{M} : $11.3 \pm 0.51 \text{ mM}$
Glucose-6-phosphate isomerase (G6PI_YEAST, EC: 5.3.1.9)	k_{cat} : $247 \pm 5.1 \text{ s}^{-1}$ K_{M} : $0.307 \pm 0.021 \text{ mM}$	k_{cat} : $253 \pm 5.1 \text{ s}^{-1}$ K_{M} : $0.304 \pm 0.020 \text{ mM}$

ontologies. Upon submission, the data are held in a gatekeeper database that can only be accessed by the submitter and curators of SABIO-RK. Upon formal curation and release by the submitter, the data are then made public in the database. This process ensures consistency and completeness of the data and provides data confidentiality, such that data can remain privately accessible before publication.

The KINETICSWIZARD can be configured to perform these submission steps automatically, ensuring that both experimental data and derived kinetic parameters are captured and stored immediately upon acquisition and analysis.

Data access

Access to the submitted data utilizes the two data repositories, MeMo-RK for experimental raw data and SABIO-RK for derived kinetic equations with their parameters and corresponding metadata. This approach is consistent with a distributed, loosely coupled system [24], in which a number of independent data resources are populated, and then later queried via web browser or web service interfaces. The key to the development of such a distributed system is to ensure a consistent means of identifying species, reactions and parameters across each of these data resources. Data submitted from the KINETICSWIZARD populates both databases, and from this, each resource can subsequently cross-reference the other, providing a link from kinetic parameters to raw data and vice versa.

An advantage of this approach is that it uncouples the storage of raw data from the storage of derived kinetic parameters, such that users have a single interface to query and retrieve kinetic parameters, irrespective of whether they have been extracted from literature or submitted by the KINETICSWIZARD. Also, this separation facilitates submission of kinetic parameters to other repositories, such as BRENDA, without affecting the raw data storage in MeMo-RK.

Web browser interface

Both MeMo-RK and SABIO-RK have web browser interfaces. SABIO-RK provides an interface for performing sophisticated searches for kinetic parameters, based on a combination of reactants, enzymes, organisms, tissues, pathways, experimental conditions, etc. Pages displaying a set of kinetic parameters link to the original data source, e.g. to the PubMed reference of the paper from which the data have been extracted, or to the corresponding page in MeMo-RK displaying the raw experimental data where the data have been

submitted from the KINETICSWIZARD (see Table 1 and Fig. 4). Similarly, MeMo-RK provides a link to the associated kinetic parameters in SABIO-RK, and contains a searchable interface to the raw experimental data (see Fig. 5).

Web service interface

The SABIO-RK web services (<http://sabio.h-its.org/web/service.jsp>) provide flexible programmatic access to the data, allowing users to write clients to customize and automate access directly from their simulation software, systems biology platforms, tools or databases [25]. The web services provide customizable points of entry and thereby an extensive search capability for kinetic data and corresponding metadata stored in SABIO-RK. The task of automatically finding parameters and associated data is aided by specifying and storing metadata using controlled vocabularies and ontological terms. As in the web browser interface, reactions with their kinetic data can be exported in Systems Biology Markup Language (SBML) [26]. An example of direct access to kinetic data through these web services has been implemented in CELLDESIGNER, a modelling tool for biochemical networks [27].

Once a given set of kinetic parameters has been discovered from the SABIO-RK web services, the user may then retrieve associated raw data in Systems Biology Results Markup Language (SBRML) [28] format via the MeMo-RK web services, allowing the data to be viewed or refitted. Such a query across distributed web services can be performed with specialized workflow software, such as TAVERNA [29].

Discussion

The development of this system was driven by the need to exchange kinetic data between experimentalists and consumers, particularly in the context of high-throughput assays and the integration of their results into biochemical computer models for simulation. Such a system had the following requirements: to provide a means of calculating kinetic parameters from raw experimental data; to store these parameters in a standardized and consistent way, such that they can readily be queried and used in systems biology studies [30,31]; and to archive the raw experimental data such that it could be reused if required, e.g. for quality control or for refitting. Furthermore, the system was to be applicable to data from a number of instruments using different experimental techniques, and the intended users of the system were experimental biologists, not bioinformaticians.

General information							
Organism	Saccharomyces cerevisiae						
Strain	BY4700 transformed in Y258						
Tissue	-						
EC Class	5.3.1.9						
SABIO reaction id	1123						
Variant	wildtype Glucose-6-phosphate isomerase						
Recombinant	expressed in Saccharomyces cerevisiae						
Substrates							
name	location	comment					
D-Fructose 6-phosphate	-	-					
Products							
name	location	comment					
D-Glucose 6-phosphate	-	-					
Modifiers							
name	location	effect	comment	protein complex			
Glucose-6-phosphate isomerase(Enzyme)	-	Modifier-Catalyst	-	(P12709)*2;			
Enzyme (protein data)							
	UniProt-ID	name	mol. weight (kDa)	deviation (kDa)			
subunit	-	-	-	-			
complex	-	-	-	-			
Kinetic Law							
	type	formula					
Michaelis-Menten		$k_{cat} * Et * S / (K_s + S)$					
Parameter							
name	type	species	start val.	end val.	deviat.	unit	comment
k_{cat}	k_{cat}	-	247.20936123	-	5.0661300497	$s^{(-1)}$	-
Et	concentration	Enzyme	6.666666E-7	-	-	mm	-
K_s	K_m	D-Fructose 6-phosphate	0.30698334365	-	0.020968017498	mm	-
S	concentration	D-Fructose 6-phosphate	-	2	-	mm	-
Experimental conditions							
	start value	end value	unit				
temperature		30	°C				
pH		6.5	-				
buffer	0.0 mM Glucose-6-phosphate 1-dehydrogenase, 100.0 mM 2-(N-morpholino)ethanesulfonic acid, 5.0 mM magnesium dichloride, 100.0 mM potassium chloride, 0.4 mM NADP						
Reference							
SABIORK id	title	author	year	direct submission link			
2459	Glycolysis	Hanan Messiha and Naglis Malys, Manchester Centre for Integrative Systems Biology (MCISB), UK	2009	http://beaconw.cs.manchester.ac.uk:8092/mcisb-web/index.jsp?application=MeMo-RK&file=index.jsp&directory=Browse&homepage=index.jsp&experimentId=14795743_9a81_45eb_ab79_d779a61fa41b			

Fig. 4. Screen capture of the web browser interface to SABIO-RK (<http://sabio.h-its.org/>), showing a coherent set of kinetic parameters submitted from the KINETICSWIZARD. A cross-link to the corresponding experimental raw data in MeMo-RK is shown at the bottom.

The KINETICSWIZARD addresses many of these issues by providing an interactive tool that integrates with instrumentation software and allows kinetic parameters to be calculated from experimental data, also providing the facility to manually correct the automated fit for noisy or anomalous data. The data model representing raw experimental data is a simple one that can be applied to many experimental techniques.

The tool manages the collection of metadata and the submission of these data to appropriate resources. In order to facilitate both the querying of these resources

and subsequent data integration, standardized terms or references to external resources are associated with the data, and these can be assigned in an intuitive, user-friendly manner. Considering systems biology studies, the task of parameterizing models with kinetic parameters is greatly simplified with data in this form, as both the SBML file containing the model and the underlying data stored in the resources can be annotated with the same terms for metabolites, enzymes, EC codes, parameter types, etc. This task is facilitated by the storage of kinetic data in SABIO-RK, from which data

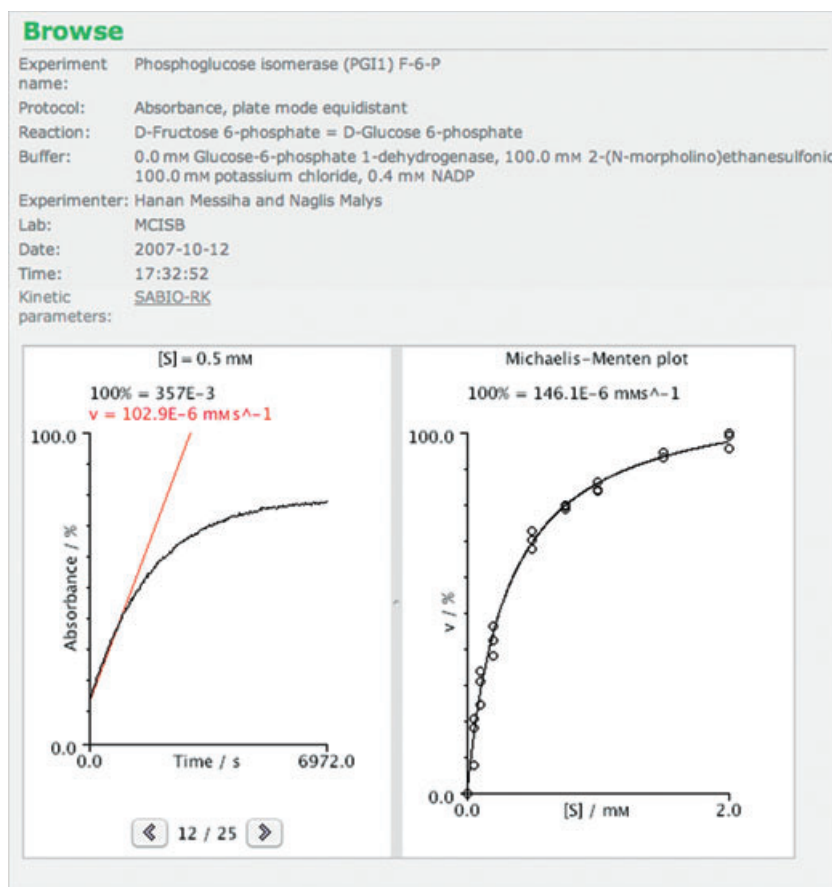


Fig. 5. Screen capture of the web browser interface to MeMo-RK (<http://www.mcisb.org/MeMo-RK/>), showing instrument raw data, the Michaelis–Menten curve and a link to parameter data in SABIO-RK.

can be exported in SBML format either through a web browser or web service interface.

Beyond the calculation, storage and dissemination of kinetic parameters, another primary focus of the work is on the management and distribution of raw experimental data. It is hoped that the introduction of a system for the storage and retrieval of raw enzyme kinetics assay data will encourage the community to share such data and to make it available in tandem with any kinetic parameters that are published. The proteomics community have made progress in this area in recent years, both with the development of standards for representing data [32] and encouraging major journals to advise that instrument data be shared in addition to derived results [33,34]. Crucially, such efforts have been supported by the development of software tools to aid experimentalists in making their data available [35–37]. It is hoped that the introduction of such a system here, along with the standardization efforts of the STREND A commission, will encourage comparable behaviour in the enzyme kinetics community, such that the publication of enzyme kinetic parameters without the sharing of associated experimental data becomes the exception rather than the norm.

Materials and methods

Enzyme expression, purification and quantification

Enzymes were expressed in *Saccharomyces cerevisiae* strains containing either overexpression plasmid [38] or chromosome-integrated gene fusion [39] and purified essentially as described previously [40]. Enzyme purity was analysed by SDS/PAGE according to Laemmli [41]. The amount and concentration of purified enzyme was determined using a standard method [42] and preparation quality confirmed with the 2100 Bioanalyzer (Agilent Technologies, Foster City, CA, USA).

Kinetic assays

Kinetic time course data of purified enzymes were determined in high-throughput measurements using a NOVostar plate reader in 384-well format plates. All measurements were carried out at 30 °C in 60 μ L reaction volumes in a reaction buffer that consisted initially of 100 mM Mes, pH 6.5, 100 mM KCl and 5 mM free magnesium chloride plus other reagents and substrates that were specific for each individual enzyme.

Assays were automated so that all reagents in the reaction buffer were in 45 μL , enzyme in 5 μL and substrate in 10 μL volumes. In almost all cases, the enzyme was incubated in the reaction mixture and the reactions were started by the addition of the substrate.

Assays for each individual enzyme were either developed or modified from previously published methods to be compatible with the conditions of the reactions (e.g. pH compatibility or unavailability of commercial substrates). For each individual enzyme, the forward and the backward reaction were measured whenever applicable, depending on the possibility of the production of active enzyme, the availability of substrates as well as the suitability of the assays at the specified pH. Some assays were modified, altering the concentration of coupling enzymes or other reagents to ensure that the rate measured was the rate of the reaction of interest.

All assays were coupled to enzymes where NAD(P) or NAD(P)H was a product or substrate whose formation or consumption could be followed spectrophotometrically at 340 nm using an extinction coefficient ($\Sigma_{340\text{ nm}}$) of $6.620\text{ mm}^{-1}\cdot\text{cm}^{-1}$.

All measurements were based on at least duplicate determination of the reaction rates at each substrate concentration. For all assays, control experiments were run in parallel to correct for any unwanted background activity.

Implementation and distribution

The KINETICSWIZARD, MeMo-RK web browser interface and web service interface are written in JAVA 1.6. MeMo-RK has been tested on POSTGRESQL 8.3. All are supported in Windows and MacOS X and are distributed as source code and associated build files. They are distributed under the open source Academic Free Licence v3.0 from <http://mcisb.sourceforge.net>. An example version of the KINETICSWIZARD, and usage instructions, can be found at <http://www.mcisb.org/resources/kinetics/>, together with links to the MeMo-RK web browser and web service interfaces. The SABIO-RK web browser and web service interfaces, submission tool and the transfer procedures are written in JAVA 1.6 and owned by HITS gGmbH (Heidelberg Institute of Theoretical Studies, Heidelberg, Germany). The SABIO-RK database system is currently implemented in Oracle 10 g and is owned by HITS gGmbH. Free access to data in SABIO-RK is granted for academic use via web browser interface or web services. Terms and conditions can be found at the SABIO-RK homepage (<http://sabio.h-its.org/>).

Acknowledgements

The authors thank the EPSRC and BBSRC for their funding of the Manchester Centre for Integrative Systems Biology (<http://www.mcisb.org>), BBSRC/EPSRC

grant BB/C008219/1, and the Klaus Tschira Foundation (KTF) and the German Federal Ministry of Education and Research (BMBF) for funding the Scientific Databases and Visualization group at the Heidelberg Institute for Theoretical Studies (<http://www.h-its.org/>). NS also thanks Joseph Dada for assistance with the SBRML export.

References

- 1 Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC *et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* **29**, 365–371.
- 2 Taylor CF, Paton NW, Lilley KS, Binz PA, Julian RK Jr, Jones AR, Zhu W, Apweiler R, Aebersold R, Deutsch EW *et al.* (2007) The minimum information about a proteomics experiment (MIAPE). *Nat Biotechnol* **25**, 887–893.
- 3 Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, Ashburner M, Ball CA, Binz PA, Bogue M, Booth T *et al.* (2008) Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat Biotechnol* **26**, 889–896.
- 4 van Eunen K, Bouwman J, Daran-Lapujade P, Postmus J, Canelas AB, Mensonides FI, Orij R, Tuzun I, van den Brink J, Smits GJ *et al.* (2010) Measuring enzyme activities under standardized *in vivo*-like conditions for systems biology. *FEBS J* **277**, 749–760.
- 5 Apweiler R, Cornish-Bowden A, Hofmeyr JH, Kettner C, Leyh TS, Schomburg D & Tipton K (2005) The importance of uniformity in reporting protein-function data. *Trends Biochem Sci* **30**, 11–12.
- 6 Schomburg I, Chang A & Schomburg D (2002) BRENDA, enzyme data and metabolic information. *Nucleic Acids Res* **30**, 47–49.
- 7 Wittig U, Golebiewski M, Kania R, Krebs O, Mir S, Weidemann A, Anstein S, Saric J & Rojas I (2006) SABIO-RK: integration and curation of reaction kinetics data. Proceedings of the 3rd International workshop on Data Integration in the Life Sciences 2006 (DILS'06), Hinxton, UK. *Lect Notes Bioinformatics* **4075**, 94–103.
- 8 Rojas I, Golebiewski M, Kania R, Krebs O, Mir S, Weidemann A & Wittig U (2007) SABIO-RK (System for the Analysis of Biochemical Pathways Reaction Kinetics). *Proceedings of the 2nd International Symposium on "Experimental Standard Conditions of Enzyme Characterizations"*, 2006, Ruedesheim am Rhein, Germany, 189–202, Logos-Verlag, Berlin.
- 9 Herrgård MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Blüthgen N, Borger S, Costenoble R, Heinemann M *et al.* (2008) A consensus yeast metabolic network reconstruction obtained from a

- community approach to systems biology. *Nat Biotechnol* **26**, 1155–1160.
- 10 Swainston N & Mendes P (2009) libAnnotationSBML: a library for exploiting SBML annotations. *Bioinformatics* **25**, 2292–2293.
 - 11 Kanehisa M & Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30.
 - 12 Degtyarenko K, de Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, Alcántara R, Darsow M, Guedj M & Ashburner M (2008) ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res* **36**, D344–D350.
 - 13 The UniProt Consortium (2010) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* **38**, D142–D148.
 - 14 Vastrik I, D'Eustachio P, Schmidt E, Joshi-Tope G, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, Lewis S *et al.* (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biol* **8**, R39.
 - 15 Michaelis L & Menten ML (1913) Die Kinetik der Invertinwirkung. *Biochem Z* **49**, 333–369.
 - 16 Le Novère N (2006) Model storage, exchange and integration. *BMC Neurosci* **7**, S11.
 - 17 Mendes P & Kell DB (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* **14**, 869–883.
 - 18 Eadie GS (1942) The inhibition of cholinesterase by physostigmine and prostigmine. *J Biol Chem* **146**, 85–93.
 - 19 Hofstee BHJ (1959) Non-inverted versus inverted plots in enzyme kinetics. *Nature* **184**, 1296–1298.
 - 20 Levenberg K (1944) Method for the solution of certain non-linear problems in least squares. *Q Appl Math* **2**, 164–168.
 - 21 Marquardt D (1963) An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math* **11**, 431–441.
 - 22 Spasić I, Dunn WB, Velarde G, Tseng A, Jenkins H, Hardy N, Oliver SG & Kell DB (2006) MeMo: a hybrid SQL/XML approach to metabolomic data management for functional genomics. *BMC Bioinform* **7**, 281.
 - 23 Rojas I, Golebiewski M, Kania R, Krebs O, Mir S, Weidemann A & Wittig U (2007) Storing and annotating of kinetic data. *In Silico Biol* **7**(Suppl 2), S37–44.
 - 24 Kell DB. (2006) Metabolomics, modelling and machine learning in systems biology: towards an understanding of the languages of cells. The 2005 Theodor Bücher lecture. *FEBS J* **273**, 873–894.
 - 25 Golebiewski M, Mir S, Kania R, Krebs O, Weidemann A, Wittig U & Rojas I (2007) Integration of SABIO-RK in workbenches for kinetic model design. *BMC Syst Biol*, **1**(Suppl 1), P4.
 - 26 Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531.
 - 27 Funahashi A, Jouraku A, Matsuoka Y & Kitano H (2007) Integration of CellDesigner and SABIO-RK. *In Silico Biol* **7**(Suppl 2), S81–90.
 - 28 Dada J, Spasić I, Paton N & Mendes P (2010) SBRML: a markup language to associate systems biology data with models. *Bioinformatics* **26**, 932–938.
 - 29 Hull D, Wolstencroft K, Stevens R, Goble C, Pocock MR, Li P & Oinn T (2006) Taverna: a tool for building and running workflows of services. *Nucleic Acids Res* **34**, W729–W732.
 - 30 Li P, Dada JO, Jameson D, Spasić I, Swainston N, Carroll K, Dunn WB, Khan F, Messiha HL, Simeonidis E *et al.* (2010) Systematic integration of experimental data and models in systems biology. *BMC Bioinform* (Under consideration).
 - 31 Swainston N, Jameson D, Li P, Spasić I, Mendes P & Paton NW (2010) Integrative information management for systems biology. *Data Integration in the Life Sciences, Proceedings, 7th International Workshop, DILS 2010* (In press).
 - 32 Vizcaino JA, Côté R, Reisinger F, Foster JM, Mueller M, Rameseder J, Hermjakob H & Martens L (2009) A guide to the Proteomics Identifications Database proteomics data repository. *Proteomics* **9**, 4276–4283.
 - 33 Anon. (2007) Democratizing proteomics data. *Nat Biotechnol* **25**, 262.
 - 34 Anon. (2007) Time for leadership. *Nat Biotechnol* **25**, 821.
 - 35 Jones P & Côté R (2008) The PRIDE proteomics identifications database: data submission, query, and dataset comparison. *Methods Mol Biol* **484**, 287–303.
 - 36 Siepen JA, Swainston N, Jones AR, Hart SR, Hermjakob H, Jones P & Hubbard SJ (2007) An informatic pipeline for the data capture and submission of quantitative proteomic data using iTRAQ. *Proteome Sci* **5**, 4.
 - 37 Barsnes H, Vizcaino JA, Eidhammer I & Martens L (2009) PRIDE Converter: making proteomics data-sharing easy. *Nat Biotechnol* **27**, 598–599.
 - 38 Gelperin DM, White MA, Wilkinson ML, Kon Y, Kung LA, Wise KJ, Lopez-Hoyo N, Jiang L, Piccirillo S, Yu H *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev* **19**, 2816–2826.
 - 39 Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK & Weissman JS (2005) Global analysis of protein expression in yeast. *Nature* **425**, 737–741.
 - 40 Malys N & McCarthy JEG (2006) Dcs2, a novel stress-induced modulator of m7GpppX pyrophosphatase

- activity that locates to P bodies. *J Mol Biol* **363**, 370–382.
- 41 Laemmli UK. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T1. *Nature* **227**, 680–685.
- 42 Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ & Klenk DC (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**, 76–85.