Heidelberg Institute for Theoretical Studies

HITS

Reaction kinetics database SABIO-RK

Andreas Weidemann
HITS Heidelberg, Germany
SDBV group

COMBINE 15. October 2015

SDBV group @ HITS
Data management for scientists in life sciences

Current major projects:

- SABIO-RK: reaction kinetics database
- SEEK: for sharing heterogeneous scientific research datasets, including models, data sheets, images...
- OperationExplorer: visualization of health data from German hospitals; in collaboration with investigative journalists
- Standards: NormSys (Martin)
- ...

Motivation for a reaction kinetics database

- quantitative data on reaction kinetics are required e.g. for modellers
- problem: hard to find in literature

=> 2005 / 2006 start of collecting reaction kinetics data, storage in a public accessible database

http://sabiork.h-its.org/

HITS

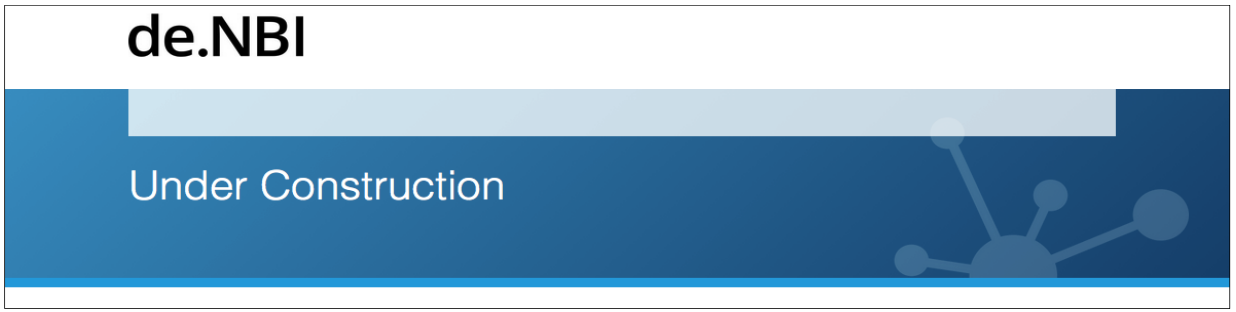
Motivation for a reaction kinetics database

- quantitative data on reaction kinetics are required e.g. for modellers
- problem: hard to find in literature

=> 2005 / 2006 start of collecting reaction kinetics data, storage in a public accessible database

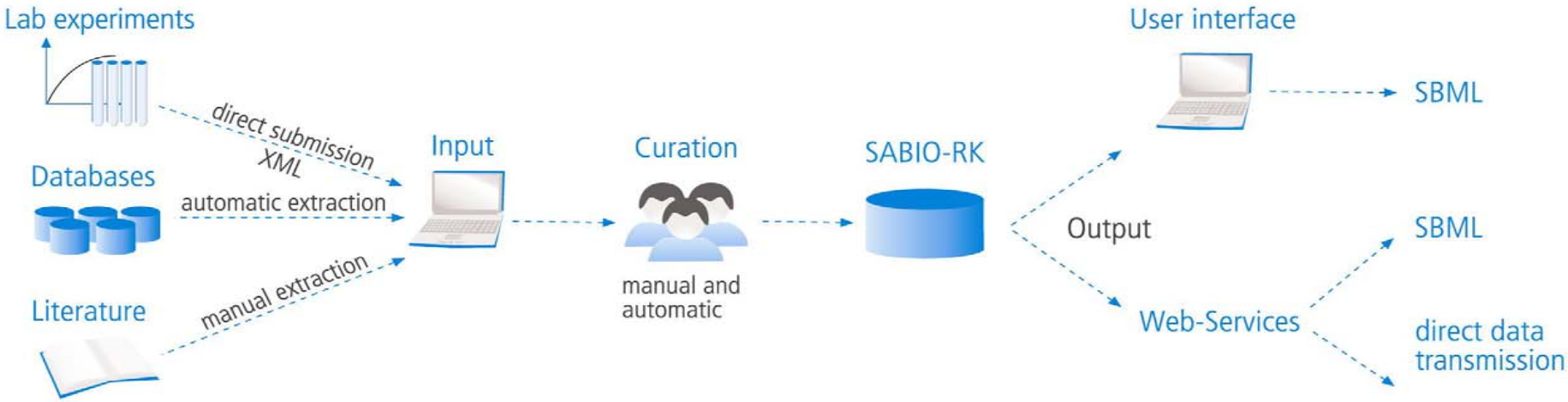
http://sabiork.h-its.org/

- SABIO-RK is now part of 'The German Network for Bioinformatics Infrastructure' (de.NBI) which is still in consolidation phase:



HITS

SABIO-RK: data workflow



HITS

Data extraction from literature

- kinetics data from literature are manually entered via a web form => standardization

- the data are manually annotated, e.g. EC number and UniProtKB ID are added

- further automatic annotations performed: KEGG, PubChem,...

- in agreement with @ Pedro: many data are found in figures / figure legends and tables => not searchable in PubMed abstracts

HITS

Stored data I

General information	
Organism	Homo sapiens
Tissue	erythrocyte
EC Class	2.7.1.1
SABIO reaction id	793
Variant	wildtype
Experiment Type	in vitro
Pathways	Glycolysis classical Glycolysis/Gluconeogenesis Starch and Sucrose metabolism
Event Description	-

Substrates		
name	**location**	**comment**
ATP	-	-
D-Glucose	-	-

Products		
name	**location**	**comment**
ADP	-	-
D-Glucose 6-phosphate	-	-

Modifiers				
name	**location**	**effect**	**comment**	**protein complex**
N-Acetylglucosamine	-	Modifier-Inhibitor	-	-
hexokinase(Enzyme)	-	Modifier-Catalyst	-	HXK;

Enzyme (protein data)				
	UniProtKB_AC	**name**	**mol. weight (kDa)**	**deviation (kDa)**
subunit	-	-	-	-
complex	-	-	132.0	8.0

HITS

Stored data II

Kinetic Law		
type	**formula**	**annotation**
Competitive inhibition	-	SBO:0000260

Parameter							
name	**type**	**species**	**start val.**	**end val.**	**deviat.**	**unit**	**comment**
A	concentration	ATP	5.0	-	-	mM	-
Ki	Ki	N-Acetylglucosamine	0.3	0.5	-	mM	versus Glucose
Km2	Km	D-Glucose	0.048	0.08	-	mM	-
Km1	Km	ATP	0.57	1.0	-	mM	MgATP

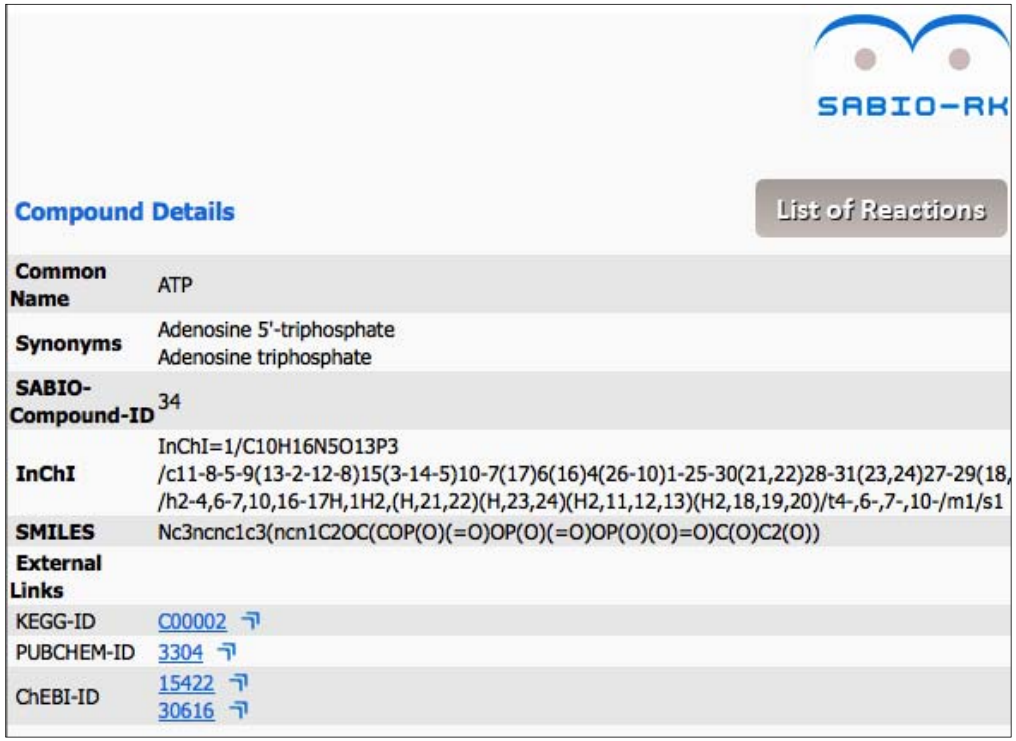
Experimental conditions			
	start value	end value	unit
temperature	37.0	-	°C
pH	7.25	-	-
buffer	0.033 M Tris/HCl, 10 mM MgCl2, 0.33 mM NADP+, 0.15 I.U. Glucose-6-phosphate dehydrogenase, 0.15 I.U. 6-Phosphogluconate dehydrogenase		
comment	-		

Reference						
title	**author**	**year**	**journal**	**volume**	**pages**	**PubMed**
Purification and some properties of human erythrocyte hexokinase	Rijksen G, Staal GE	1976	Biochim Biophys Acta	445	330-41	953036

HITS

Data annotation

example 'ATP' as compound



SABIO-RK

Compound Details List of Reactions

Common Name	ATP
Synonyms	Adenosine 5'-triphosphate Adenosine triphosphate
SABIO-Compound-ID	34
InChI	InChI=1/C10H16N5O13P3 /c11-8-5-9(13-2-12-8)15(3-14-5)10-7(17)6(16)4(26-10)1-25-30(21,22)28-31(23,24)27-29(18, /h2-4,6-7,10,16-17H,1H2,(H,21,22)(H,23,24)(H2,11,12,13)(H2,18,19,20)/t4-,6-,7-,10-/m1/s1
SMILES	Nc3ncnc1c3(ncn1C2OC(COP(O)(=O)OP(O)(=O)OP(O)(O)=O)C(O)C2(O))
External Links	
KEGG-ID	C00002
PUBCHEM-ID	3304
ChEBI-ID	15422 30616

HITS

Public search interface

provides many options for detailed search

HITS

Search results

search for 'ATP' and 'homo sapiens'

Total number of kinetic law entries found: **1627**

1 2 3 4 5 6 7 8 9 10 .. 109 Next

display 15 entries per page

Kinetic data	Reaction	Enzyme			Tissue	Organism	Parameter (besides concentration)	Environment		Add to export cart?
		ECNumber	Protein	Variant				°C	pH	
▼	D-Glucose + ATP = ADP + D-Glucose 6-phosphate	2.7.1.1		wildtype	erythrocyte	Homo sapiens	Ki Km	37.0	7.25	☐

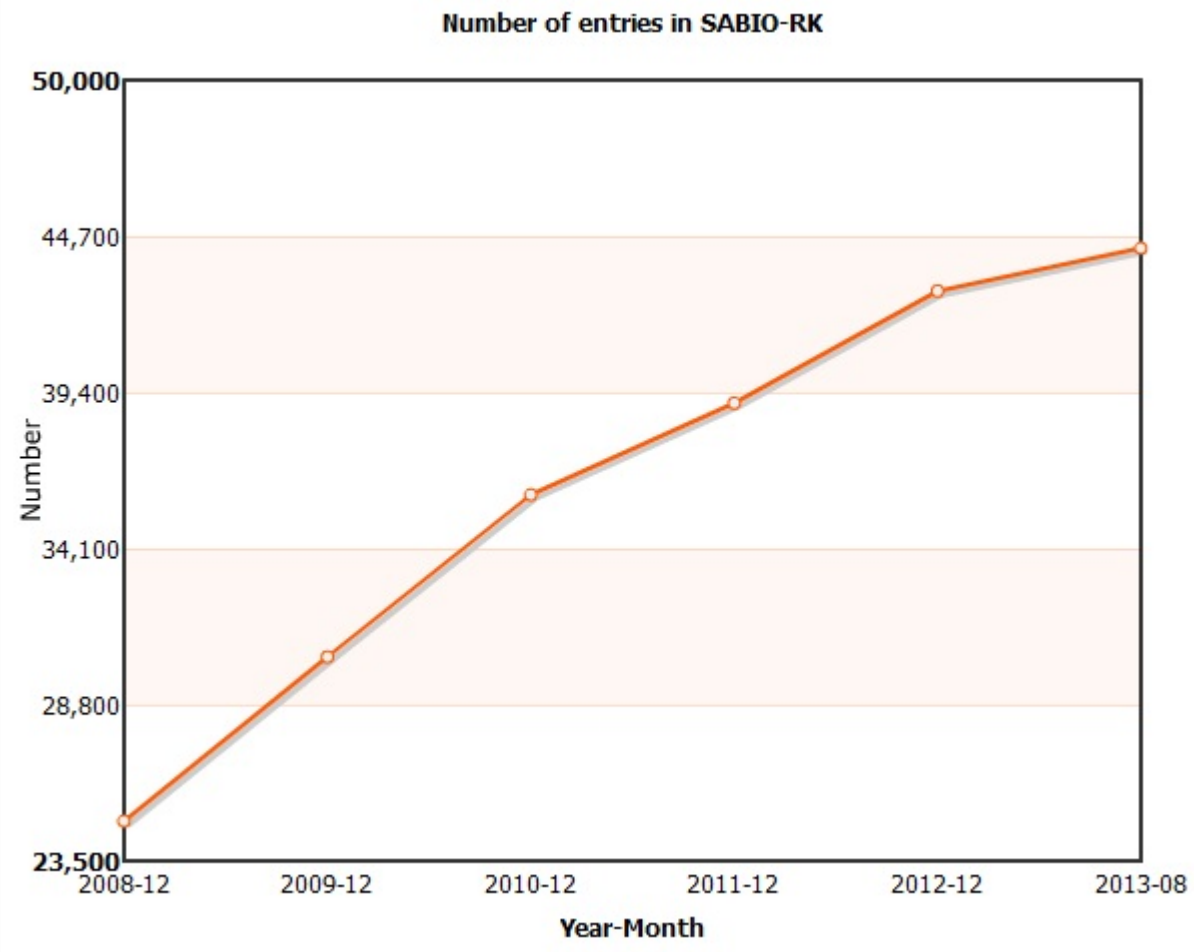
Entry ID: 10716

General information	
Organism	**Homo sapiens**
Tissue	erythrocyte
EC Class	2.7.1.1
SABIO reaction id	793
Variant	wildtype
Experiment Type	in vitro
Pathways	Glycolysis classical Glycolysis/Gluconeogenesis Starch and Sucrose metabolism
Event Description	-

Substrates		
name	**location**	**comment**
ATP	-	-
D-Glucose	-	-

HITS

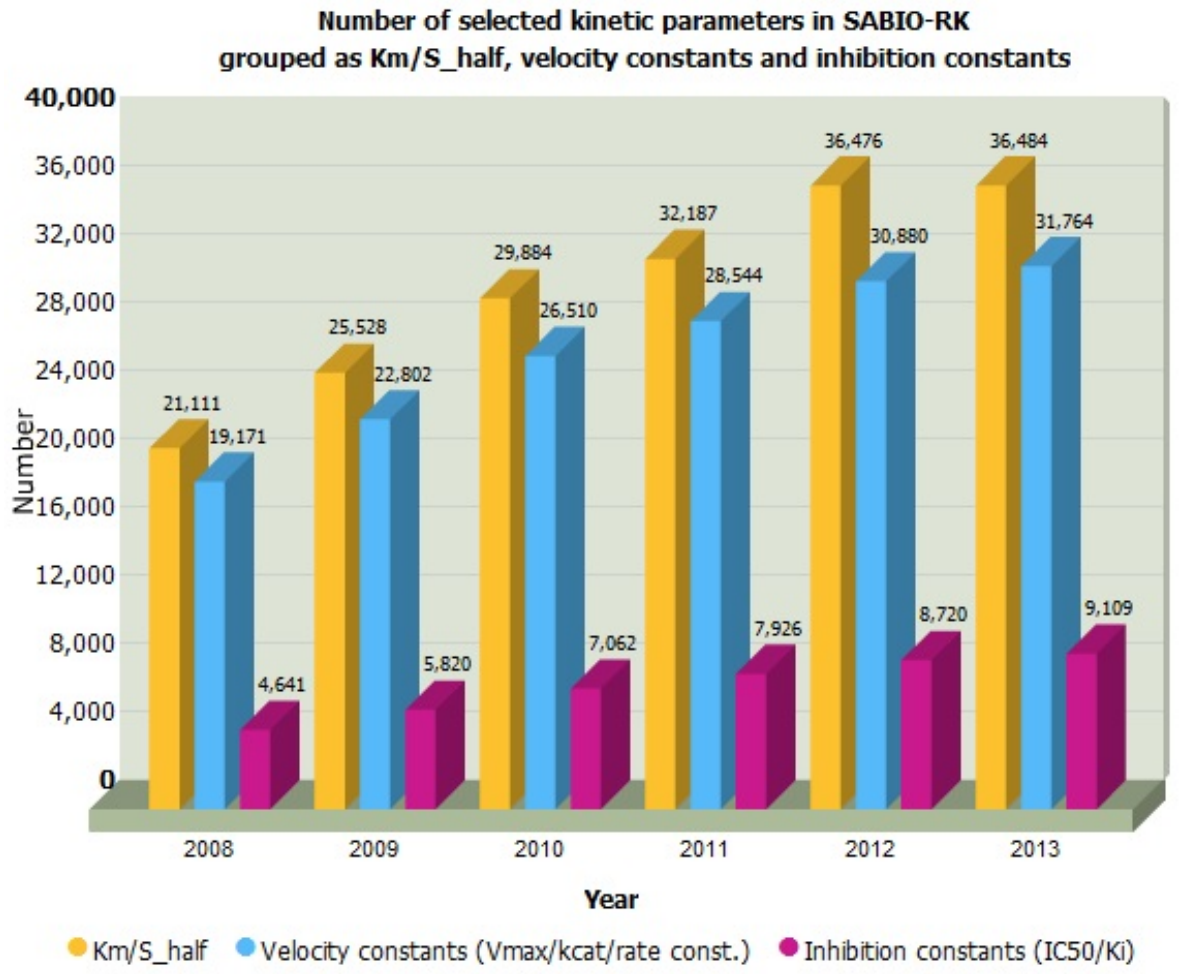
Statistics I



as of October 2015: 52.672 single entries

HITS

Statistics II



kinetic parameter as of October 2015: 116.287

HITS

Data upload and data import into SABIO-RK

- direct submission of experimental data

- kinetics data in SBML format:

upload of SBML file

↓

parsing & mapping of entities

↓

annotation

↓

storage in database

HITS

RESTfull webservices I

- allows programmatic access of data

- example: get the list of SABIO kinetic law entries:

http://sabiork.hits.org/sabioRestWebServices/searchKineticLaws/entryIDs?q=Tissue:”spleen” AND Organism:"Homo sapiens"

- returns SBML document

HITS

RESTfull webservices II

application and tools using SABIO-RK webservices:

- CellDesigner
- JWS Online (Jacky Snoep)
- Virtual Cell (VCell, Ion Moraru)
- CySabioRK-Plugin for Cytoscape (Matthias König)
- BioUML (Mandrik, Kiselev,...)
- Path2Models
- Bio2RDF
- Virtual Liver Network (Matthias König)
- University of Manchester (integrated into Taverna workflows)
- Systems Biology Metabolic Modeling Assistant SBMM (Univ. Malaga)
- SBMLsqueezer (Andreas Dräger)
- ….

HITS

Data export

- spreadsheets (tables, PDF)

- BioPax (main entities, mapping coded by Oliver Ruebenacker)

- SBML:

HITS

Data export in SBML format – problems

- issues encountered eg:

  - concentration ranges as boundary conditions
    - enzyme
    - compounds

  - experimental conditions (pH, temp., ...) only in SABIO_RK namespace exported

  - description of enzyme complexes

HITS

Data import from SBML files – challenges

- most of the information in SBML is optional (e.g. units, names)

- global and local attributes (e.g. global parameters for whole model and local parameters for single reactions)

- sometimes missing organism (e.g. „cellular organisms“ in 74 curated models in BioModels DB)

- frequently missing annotations to UniprotKB for enzymes/catalysts (only EC number)

- missing parameter definitions (Km, Vmax etc.)

- abbreviations for compound names (used only within the model – no standard controlled vocabulary)

HITS

Your data ...

if you are interested …

- we could focus on your specific research topic and insert relevant publications

- please feel free to send us your data as SMBL

- talk to us about uploading and storage of your data in other formats

HITS

HITS SDBV Group

SABIO-RK

KTS
Klaus Tschira Stiftung gemeinnützige GmbH

DFG

SysMO
Systems Biology of Microorganisms

de.NBI

HITS