SBML: Where It’s Been and Where It’s Going

Michael Hucka

Senior Research Fellow—Control and Dynamical Systems
Co-director—Biological Network Modeling Center (BNMC), Beckman Institute
California Institute of Technology
Pasadena, California, USA

Background

Background

- Conviction that computational modeling is crucial
  - Enables **quantitative** hypothesis testing

Background

- Conviction that computational modeling is crucial
  - Enables **quantitative** hypothesis testing
- Not a new idea—dates to 1940’s if not earlier
  - Theoretical & technological advances made since then

Background

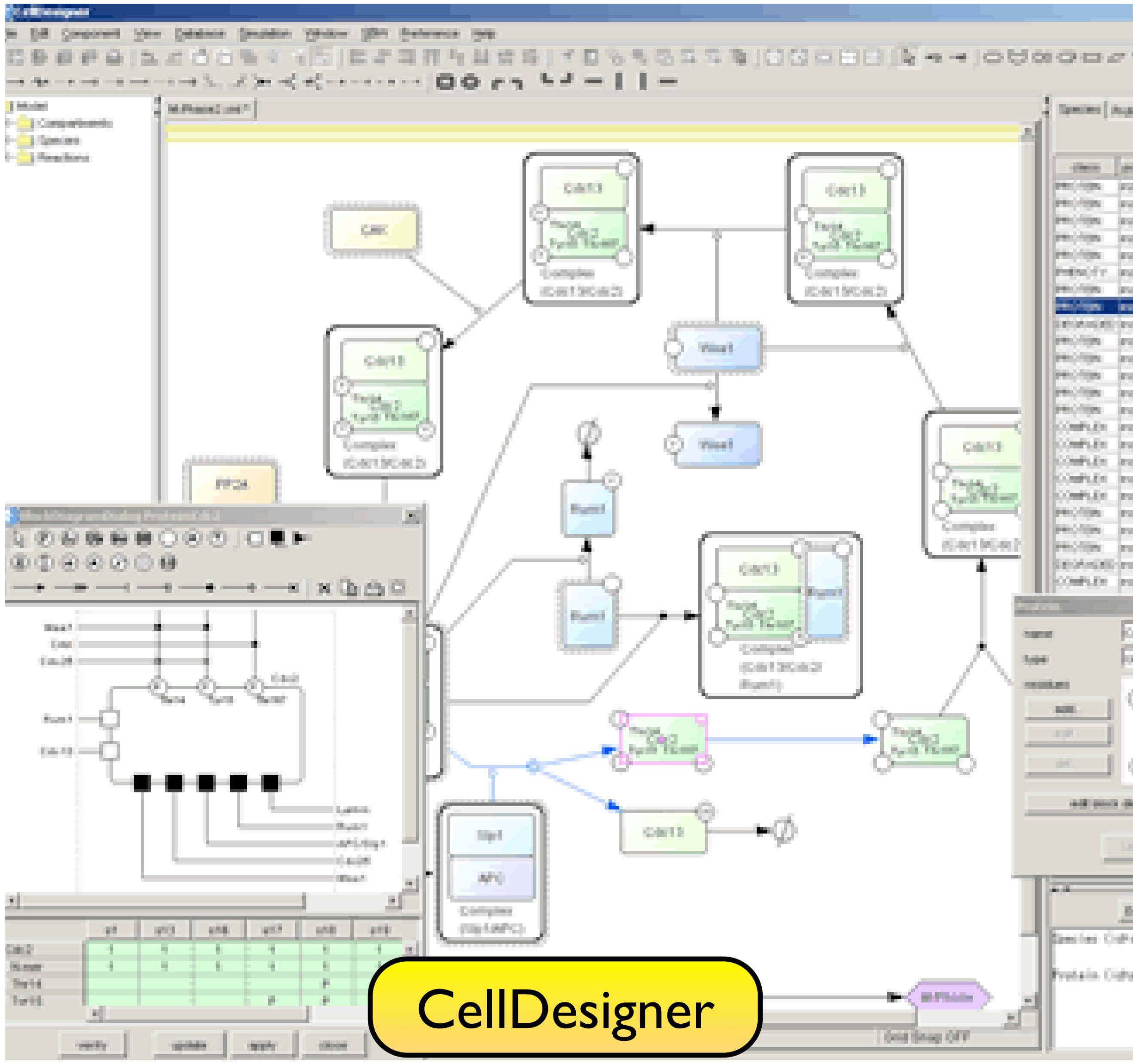
- Conviction that computational modeling is crucial
  - Enables **quantitative** hypothesis testing
- Not a new idea—dates to 1940’s if not earlier
  - Theoretical & technological advances made since then
- Support is better than ever
  - General mathematical environments
  - Special-purpose software tools

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing

Specialized software tools for computational modeling in biology

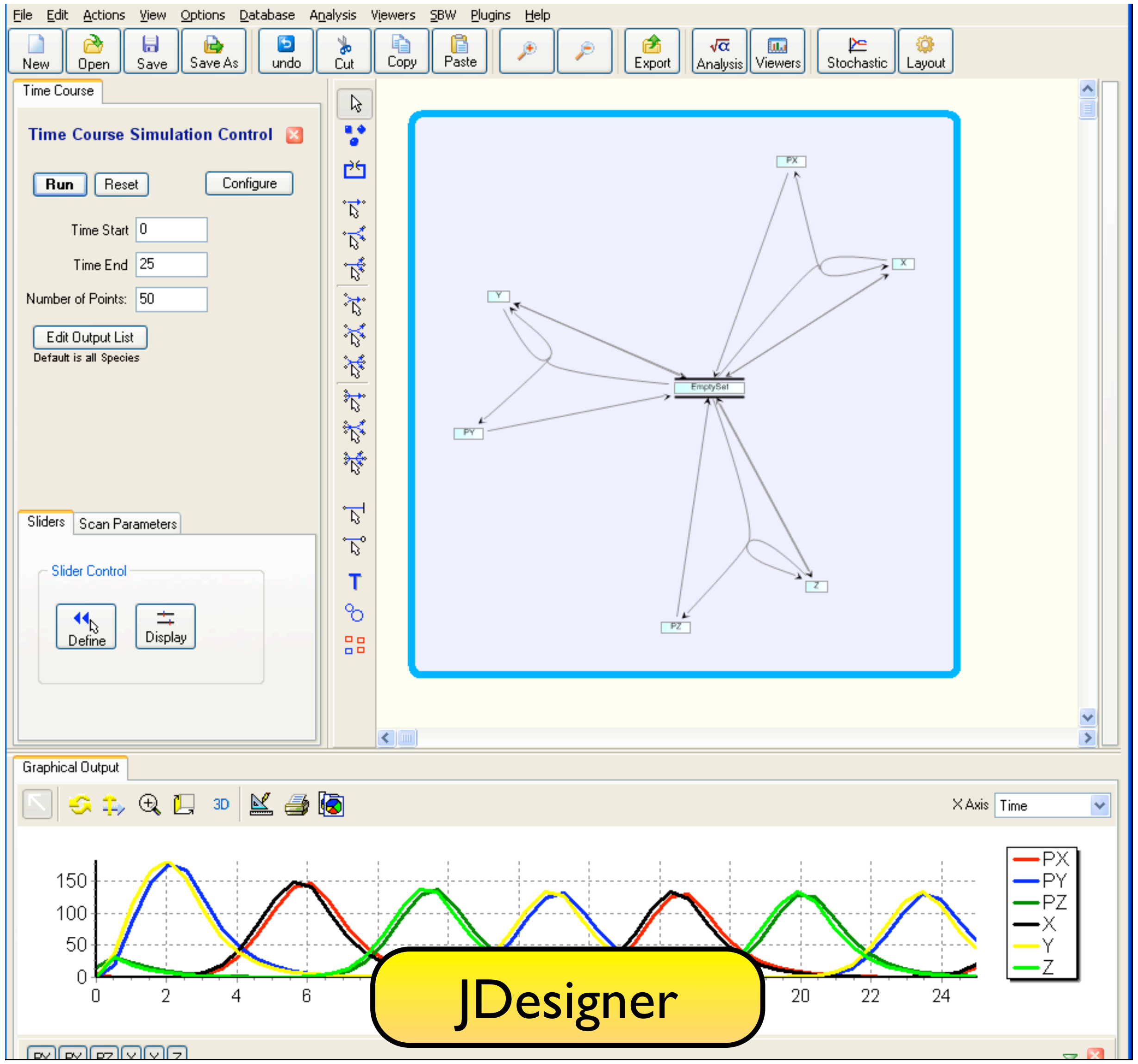
- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing



CellDesigner

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing



Time Course Simulation Control

Time Start 0

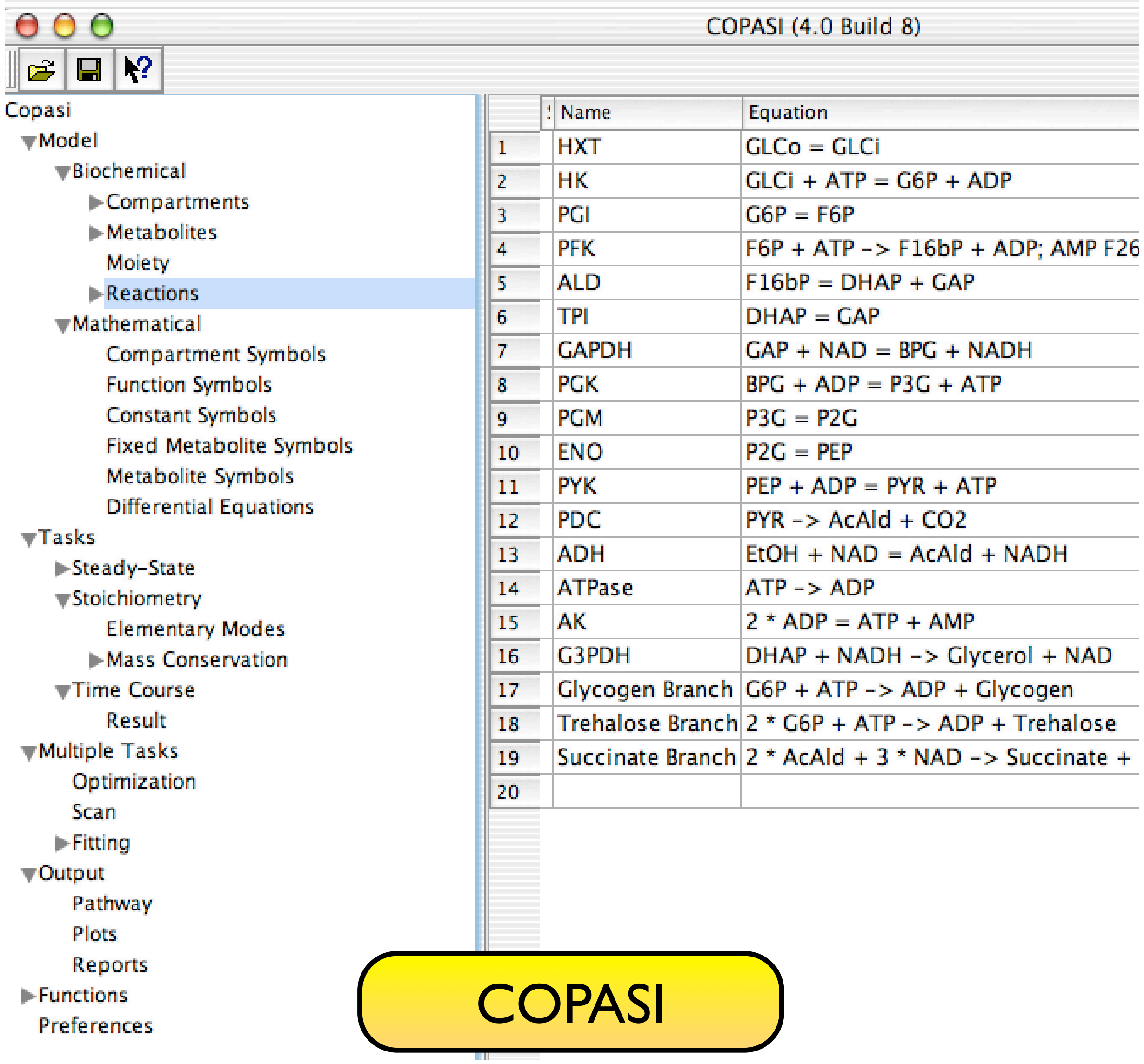
Time End 25

Number of Points: 50

Default is all Species

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing

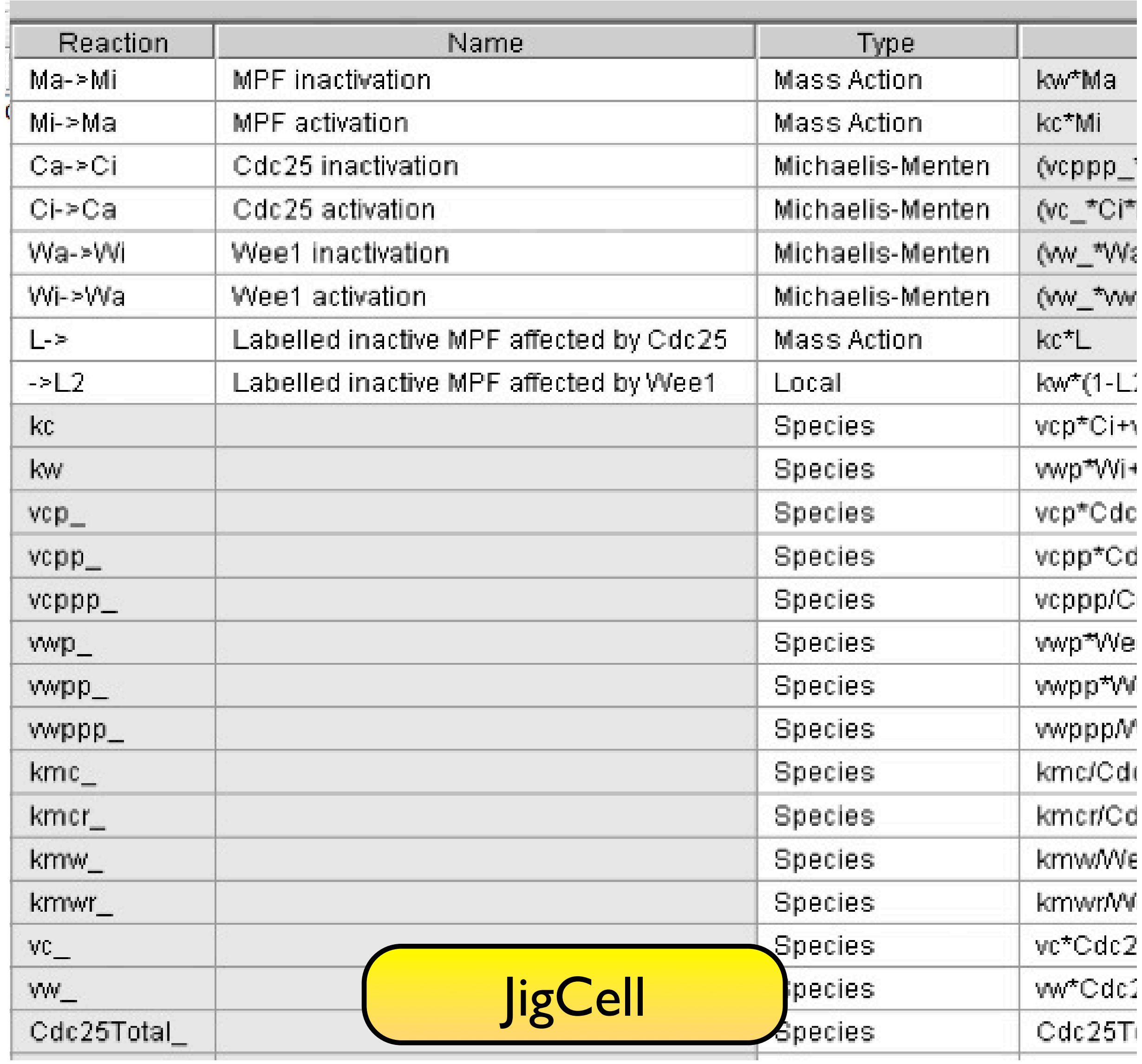


	Name	Equation
1	HXT	GLCo = GLCi
2	HK	GLCi + ATP = G6P + ADP
3	PGI	G6P = F6P
4	PFK	F6P + ATP -> F16bP + ADP; AMP F26
5	ALD	F16bP = DHAP + GAP
6	TPI	DHAP = GAP
7	GAPDH	GAP + NAD = BPG + NADH
8	PGK	BPG + ADP = P3G + ATP
9	PGM	P3G = P2G
10	ENO	P2G = PEP
11	PYK	PEP + ADP = PYR + ATP
12	PDC	PYR -> AcAld + CO2
13	ADH	EtOH + NAD = AcAld + NADH
14	ATPase	ATP -> ADP
15	AK	2 * ADP = ATP + AMP
16	G3PDH	DHAP + NADH -> Glycerol + NAD
17	Glycogen Branch	G6P + ATP -> ADP + Glycogen
18	Trehalose Branch	2 * G6P + ATP -> ADP + Trehalose
19	Succinate Branch	2 * AcAld + 3 * NAD -> Succinate +
20		

COPASI

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing

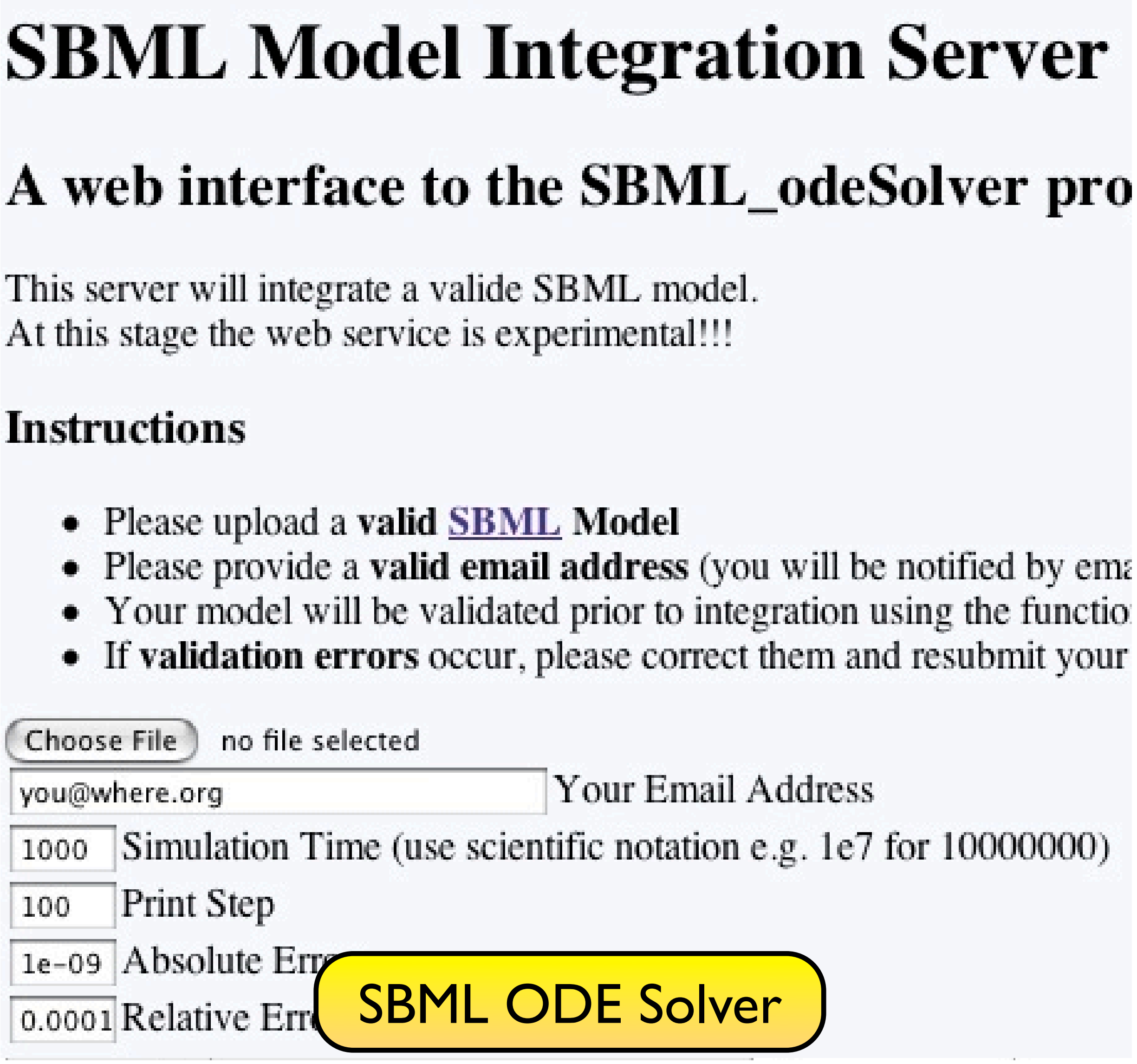


Reaction	Name	Type	
Ma->Mi	MPF inactivation	Mass Action	kw*Ma
Mi->Ma	MPF activation	Mass Action	kc*Mi
Ca->Ci	Cdc25 inactivation	Michaelis-Menten	(vcppp_
Ci->Ca	Cdc25 activation	Michaelis-Menten	(vc_*Ci*
Wa->Wi	Wee1 inactivation	Michaelis-Menten	(vw_*W
Wi->Wa	Wee1 activation	Michaelis-Menten	(vw_*vw
L->	Labelled inactive MPF affected by Cdc25	Mass Action	kc*L
->L2	Labelled inactive MPF affected by Wee1	Local	kw*(1-L
kc		Species	vcp*Ci+
kw		Species	vwp*Wi
vcp_		Species	vcp*Cdc
vcpp_		Species	vcpp*Cd
vcppp_		Species	vcppp/C
vwp_		Species	vwp*We
vwpp_		Species	vwpp*W
vwppp_		Species	vwppp/W
kmc_		Species	kmc/Cd
kmcr_		Species	kmcr/Cd
kmw_		Species	kmw/We
kmwr_		Species	kmwr/W
vc_		Species	vc*Cdc2
vw_		pecies	vw*Cdc
Cdc25Total_		pecies	Cdc25T

JigCell

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing



SBML Model Integration Server

A web interface to the SBML_odeSolver pro

This server will integrate a valide SBML model.
At this stage the web service is experimental!!!

Instructions

- Please upload a **valid** SBML **Model**
- Please provide a **valid email address** (you will be notified by ema
- Your model will be validated prior to integration using the functio
- If **validation errors** occur, please correct them and resubmit your

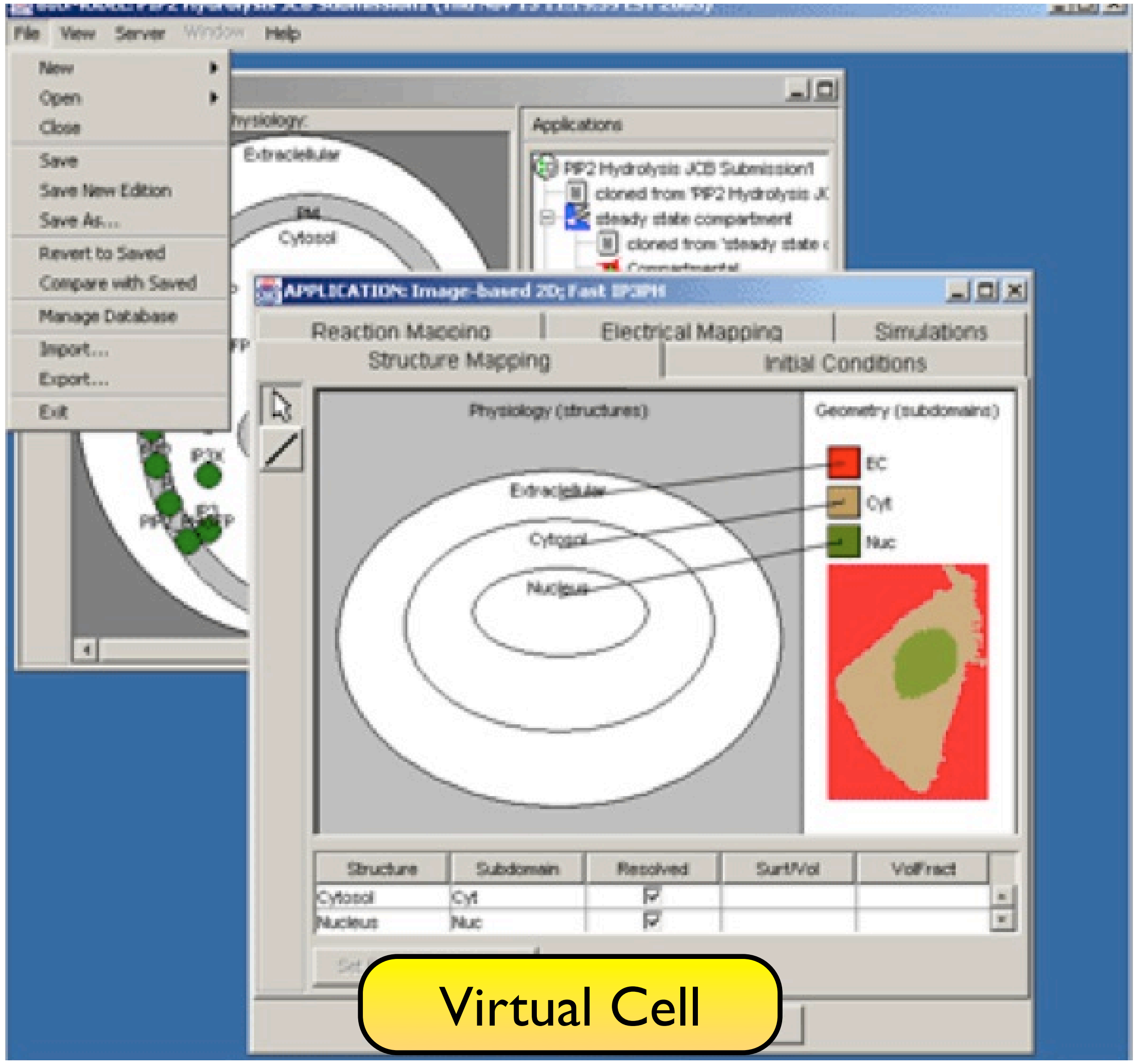
no file selected

you@where.org Your Email Address
1000 Simulation Time (use scientific notation e.g. 1e7 for 10000000)
100 Print Step
1e-09 Absolute Err
0.0001 Relative Err

SBML ODE Solver

Specialized software tools for computational modeling in biology

- > 100 available
- Range of capabilities
  - Editing/creating models
  - Simulating/analyzing
  - Visualizing
  - Databasing



Structure	Subdomain	Resolved	Surf/Vol	VolFract
Cytosol	Cyt	☑		
Nucleus	Nuc	☑		

Virtual Cell

Ability to exchange models is critical

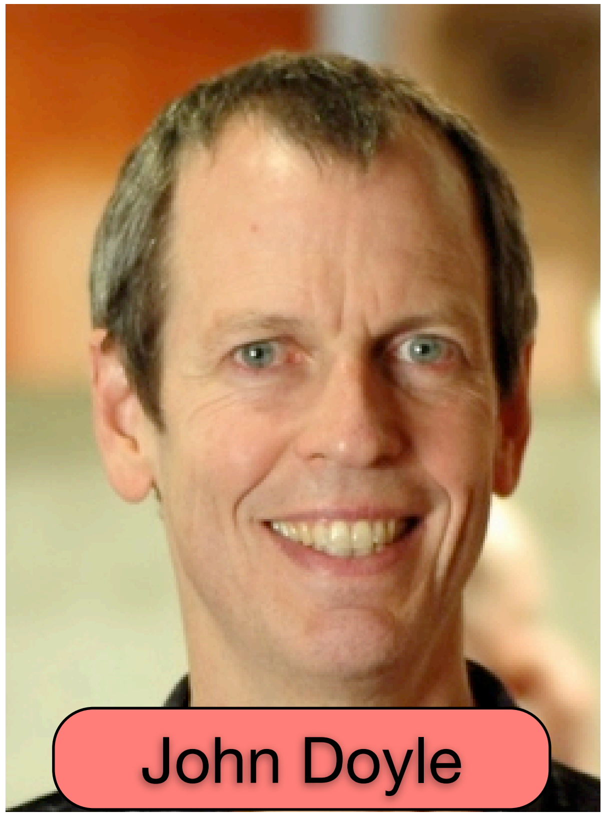
Ability to exchange models is critical

- Simply publishing equations is not enough
  - Don’t want to transcribe equations from papers
  - You want a common file format

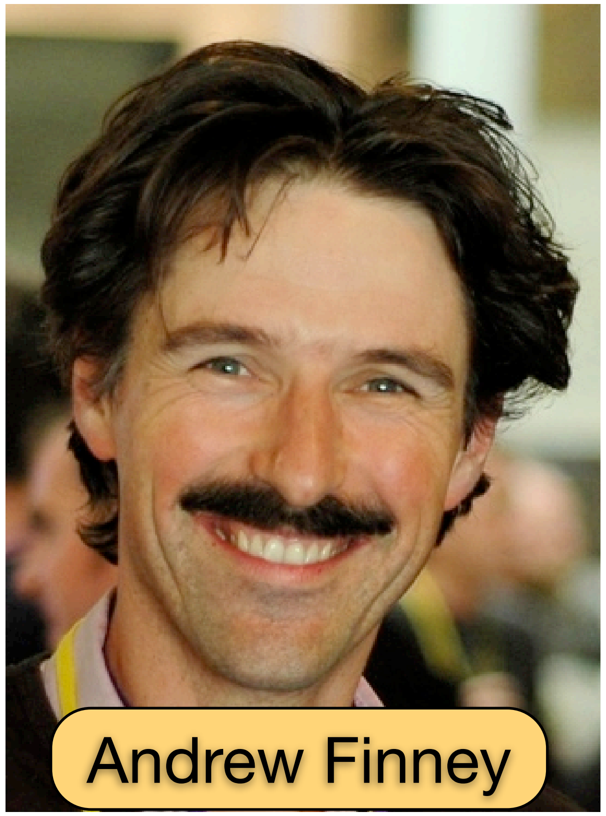
Ability to exchange models is critical

- Simply publishing equations is not enough
  - Don’t want to transcribe equations from papers
  - You want a common file format
- **Not** a new idea—seems obvious
  - Still, a format hadn’t existed before year 2000
    - Each tool had its own unique proprietary format
    - (Fewer tools too)

Hiroaki Kitano



John Doyle



Andrew Finney

Herbert Sauro

Hamid Bolouri

(the speaker)

JST ERATO Kitano Project

One initial component: get 8–10 software systems interacting

SBML =
Systems Biology Markup Language

SBML =
Systems Biology Markup Language

- Machine-readable format for computational models

SBML = Systems Biology Markup Language

- Machine-readable format for computational models
- Suitable for reaction networks
    - Arbitrary rate functions

$$2\,A + B \rightharpoonup C$$

$$C \rightleftharpoons D + F$$

…

SBML =
Systems Biology Markup Language

- Machine-readable format for computational models
- Suitable for reaction networks
  - Arbitrary rate functions
- Models can also include
  - Compartments
  - Mathematical “extras”

$$2\ A + B \rightharpoonup C$$

$$C \rightleftharpoons D + F$$

...

SBML =
Systems Biology Markup Language

- Machine-readable format for computational models
- Suitable for reaction networks
  - Arbitrary rate functions
- Models can also include
  - Compartments
  - Mathematical “extras”
- Declarative, not procedural

$$2A + B \rightharpoonup C$$

$$C \rightleftharpoons D + F$$

…

SBML is an XML format

- SBML defined using UML and XML Schema
  - Targeted at XML, but mostly independent of it
- A **lingua franca** for software, not humans
  - Think HTML

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2" level="2" version="1
      xmlns:math="http://www.w3.org/1998/Math/MathML">
    <model>
        <listOfCompartments>
            <compartment id="cell"/>
        </listOfCompartments>
        <listOfSpecies>
            <species id="X0" compartment="cell" initialConcentrati
            <species id="X1" compartment="cell" initialConcentrati
            <species id="T" compartment="cell" initialConcentratio
            <species id="S1" compartment="cell" initialConcentrati
            <species id="S2" compartment="cell" initialConcentrati
        </listOfSpecies>
        <listOfParameters>
            <parameter id="Keq" value="2.5"/>
        </listOfParameters>
        <listOfRules>
            <assignmentRule variable="S1">
                <math xmlns="http://www.w3.org/1998/Math/MathML">
                    <apply>
                        <divide/>
                        <ci> T </ci>
                        <apply>
                            <plus/>
                            <cn> 1 </cn>
                            <ci> Keq </ci>
                        </apply>
                    </apply>
                </math>
            </assignmentRule>
```

Where is SBML today?

Now the *de facto* standard

The Systems Biology Markup Language (SBML) is a computer-readable format for represe **models of biochemical reaction networks**. SBML is applicable to metabolic networks, c signaling pathways, regulatory networks, and many others.

Internationally Supported and Widely Used

SBML has been evolving since mid-2000 through the efforts of an international group of so developers and users. Today, SBML is **supported by over 100 software systems**, includi following (where "*" indicates SBML support in development):

BALSA
BASIS
BIOCHAM
BioCharon
ByoDyn
BioCyc
BioGrid
BioModels
BioNetGen
BioPathway Explorer
Bio Sketch Pad
DBsolve
Dizzy
E-CELL
ecellJ
ESS
FluxAnalyzer
Fluxor
Gepasi
Gillespie2
HSMB
HybridSBML
MMT2
Modesto
Moleculizer
Monod
Narrator
NetBuilder
Oscill8
PANTHER Pathway
PathArt
PathScout
PathwayLab
SBMLmerge
SBMLR
SBMLSim
SBMLToolbox
SBliD
SBToolbox
SBW
SCIpath
Sigmoid*
SigPath
SigTran

- Supported by >100 systems
- Accepted by journals
  - Nature
  - PLoS
  - BMC
- Used in texbooks & courses

A community of modelers and software developers

- **sbml-discuss** (275+ people), **sbml-announce**
- Annual **SBML Forum** meeting (at ICSB)
- Annual **SBML Hackathon**

A community of modelers and software developers

- **sbml-discuss** (275+ people), **sbml-announce**
- Annual **SBML Forum** meeting (at ICSB)
- Annual **SBML Hackathon**

SBML development process so far

- **Informal discussions** lead to proposals for change
  - Fix errors, lacunae, and niggling issues
  - Self-organized community efforts for significant extensions
    - Whitepapers, discussions, software implementations
- **SBML editors**: Hucka, Andrew Finney, Nicolas Le Novère
  - **Reconcile** proposals for changes
  - **Write** final specifications

Support by SBML Team

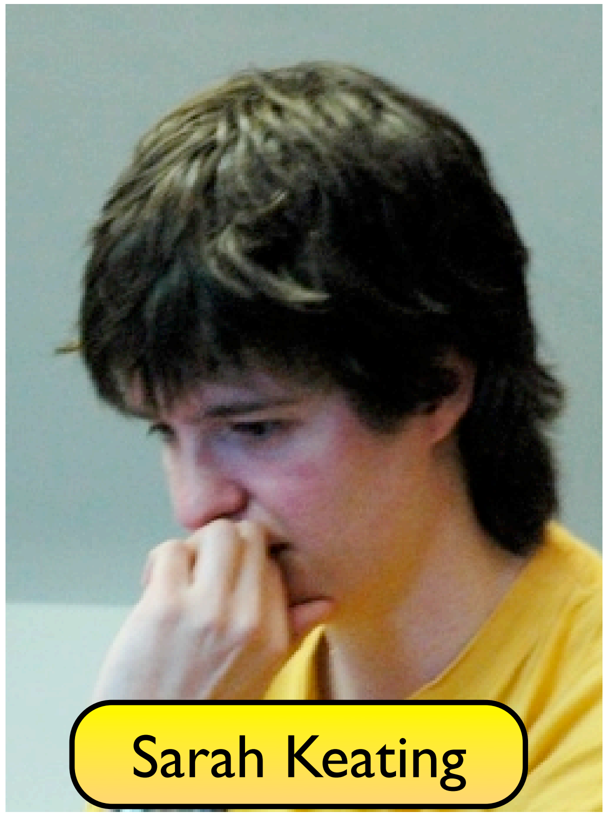
- Writing grants for core development
- Writing infrastructure software
  - libSBML
  - MathSBML, SBMLToolbox
- Maintaining web & mailing list resources
- Organizing workshops & other events

Software for working with SBML

Software for working with SBML

- libSBML: API library


Ben Bornstein

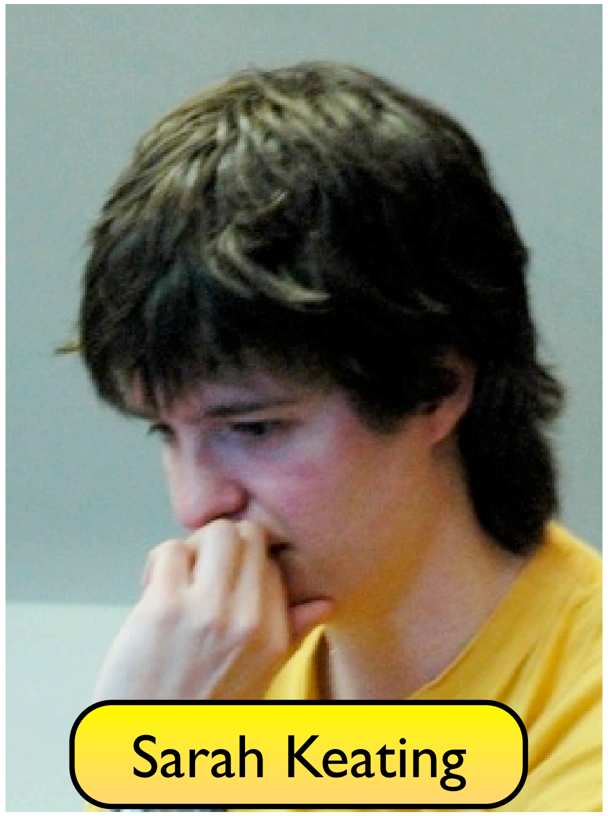

Sarah Keating

Software for working with SBML

- libSBML: API library
- MathSBML: Mathematica user package

Software for working with SBML

- libSBML: API library
- MathSBML: Mathematica user package
- SBMLToolbox: basic MATLAB interface


Sarah Keating

Software for working with SBML

- libSBML: API library
- MathSBML: Mathematica user package
- SBMLToolbox: basic MATLAB interface
- Online SBML validator at http://sbml.org

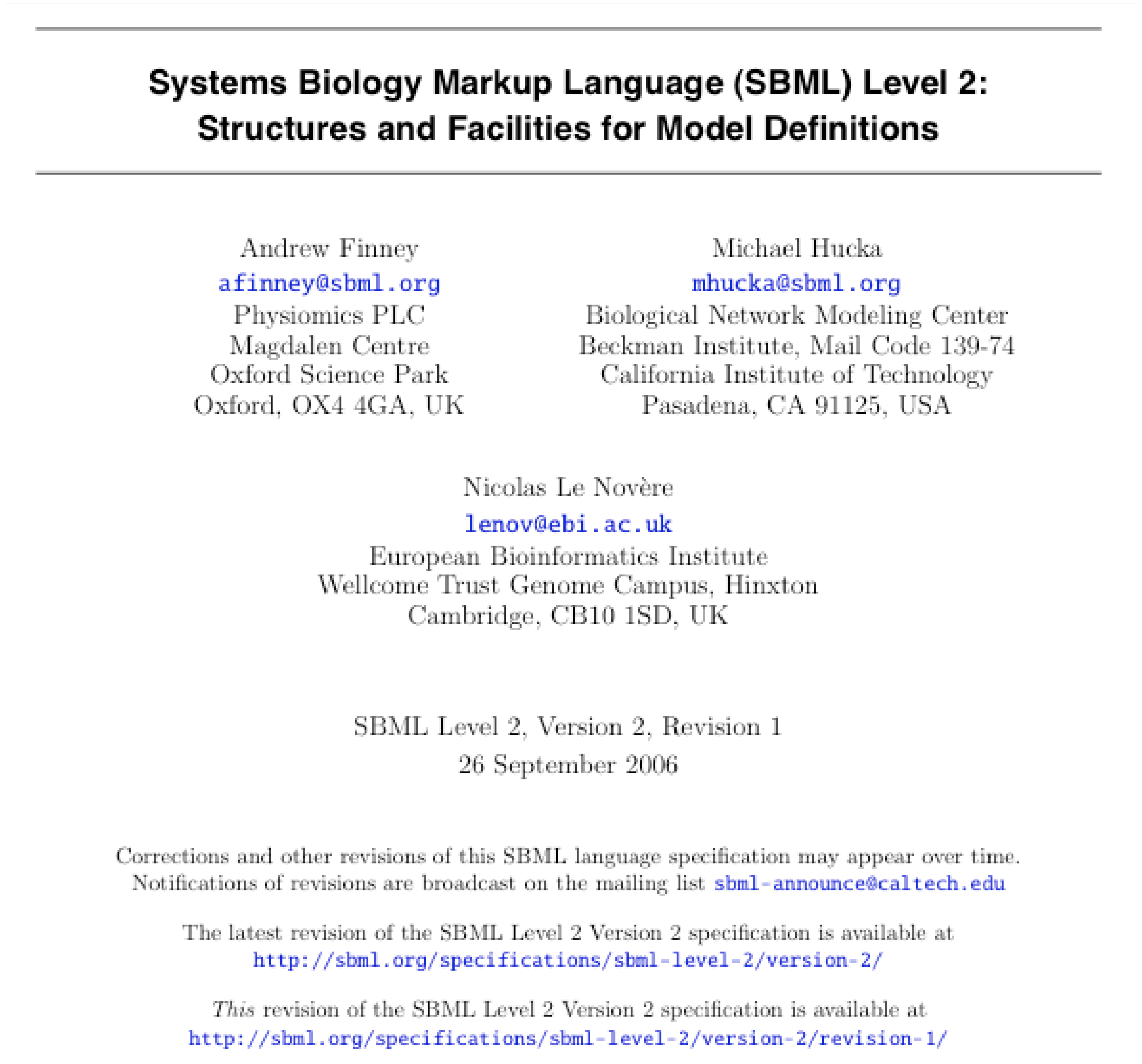

Ben Bornstein

SBML “Levels”

- Levels are meant to coexist
- ***Level 1***: mostly basic compartmental modeling
- ***Level 2***: significantly more features—e.g.:
  - User-defined functions
  - Events
  - “Types” for chemical species and compartments
  - Initial conditions, constraints, other “fiddly bits”
- ***Level 3***: now (back) in development

Latest: SBML Level 2 Version 2

- Final version released September 26



Systems Biology Markup Language (SBML) Level 2:
Structures and Facilities for Model Definitions

Andrew Finney
afinney@sbml.org
Physiomics PLC
Magdalen Centre
Oxford Science Park
Oxford, OX4 4GA, UK

Michael Hucka
mhucka@sbml.org
Biological Network Modeling Center
Beckman Institute, Mail Code 139-74
California Institute of Technology
Pasadena, CA 91125, USA

Nicolas Le Novère
lenov@ebi.ac.uk
European Bioinformatics Institute
Wellcome Trust Genome Campus, Hinxton
Cambridge, CB10 1SD, UK

SBML Level 2, Version 2, Revision 1
26 September 2006

Corrections and other revisions of this SBML language specification may appear over time.
Notifications of revisions are broadcast on the mailing list sbml-announce@caltech.edu

The latest revision of the SBML Level 2 Version 2 specification is available at
http://sbml.org/specifications/sbml-level-2/version-2/

This revision of the SBML Level 2 Version 2 specification is available at
http://sbml.org/specifications/sbml-level-2/version-2/revision-1/

Examples of significant changes

- Much clarified explanation of interpreting reactions
- Simplification to units system
- Species types, compartment types
- “Constraints”
- Support for the Systems Biology Ontology (SBO)
- Recommended standard format for annotations
- “Revisions” process for handling errata

What lies ahead?

Full SBML Test Suite

- Allows developers to test implementation of SBML support
- Critical for improving software interoperability
- Currently have a partial “SBML semantic test suite”
- Needs further work to—
  - Complete coverage of SBML features
  - Improve ease of use
  - Update for Level 2 Version 2 and Level 3
  - Add web system for reporting results, comparisons, etc.

SBML Level 3

SBML Level 3

- Modular language extensions
  - Core expected to be based “mostly” on Level 2 Version 2

SBML Level 3

- Modular language extensions
  - Core expected to be based “mostly” on Level 2 Version 2
- Layered on top of core: feature sets for—
  - Diagram storage
  - Multicomponent species
  - Models composed of submodels
  - Arrays and/or sets of components
  - Spatial geometry
  - *Other capabilities*

Modular extensions support in libSBML & SBML Test Suite

- Goal: enable libSBML to be extended through plug-ins
  - Proposals for SBML extensions can come with libSBML add-on
  - Developers can pick & choose which ones are compiled in
- Goal: enable Test Suite to be similarly extended
  - Proposals for SBML extensions can come with Test Suite add-ons

Collateral standardization efforts

- **Systems Biology Ontology** (SBO)
  - For computational models
  - Add annotations about roles & meanings of the math
- “**Parameter sets**”
  - Single model, multiple sets of numerical values

Revised SBML governance & development process

- Borrow ideas from W3C & other organizations
- Implement a better-defined, **regimented process**
  - Calls for proposals, etc.
  - Voting, etc.
  - Issue tracking system
- Have an **architectural board** to steer development
- Have **more SBML Editors**
  - Elect SBML Editors for **limited terms**

Standards body recognition

- Currently not recognized by a standards-making body
  - At some point in the future, it probably should be
    - Of special interest to commercial efforts
- Questions remain
  - **When** to seek standardization
  - **Which agency**? ISO? OMG? W3C?

Closing

The funding

- JST ERATO Kitano Symbiotic Systems Project (Japan) (to 2003)
- National Institute of General Medical Sciences (USA)
- National Science Foundation (USA)
- International Joint Research Program of NEDO (Japan)
- JST ERATO-SORST Program (Japan)
- Japanese Ministry of Agriculture
- Japanese Ministry of Educ., Culture, Sports, Science and Tech.
- BBSRC e-Science Initiative (UK)
- DARPA IPTO Bio-SPICE Bio-Computation Program (USA)
- Air Force Office of Scientific Research (USA)
- STRI, University of Hertfordshire (UK)
- Beckman Institute, Caltech (USA)
- Molecular Sciences Institute (USA)

The SBML Team people

- Hamid Bolouri
- Herbert Sauro
- Andrew Finney
- Maria Schilstra
- Jo Matthews
- Akira Funahashi
- Ben Bornstein
- Ben Kovitz
- Bruce Shapiro
- Sarah Keating

A million thanks to the SBML Community too

Where to learn more

- **http://sbml.org**
- *Upcoming*:
  - SBML Forum 2006 in Tokyo, Oct. 12-13 (after ICSB 2006)
  - SBML Hackathon 2007 in June at U. Newcastle, UK
  - SBML Forum 2007 in Long Beach, CA, USA, Oct. 5-6 (ICSB 2007)
- **Thank you**!