

Bridging data and models using VCell (the “Virtual Cell”)

<http://vcell.org>

Michael Blinov

COMBINE 2011, September 7th, HITS



University of Connecticut
Health Center



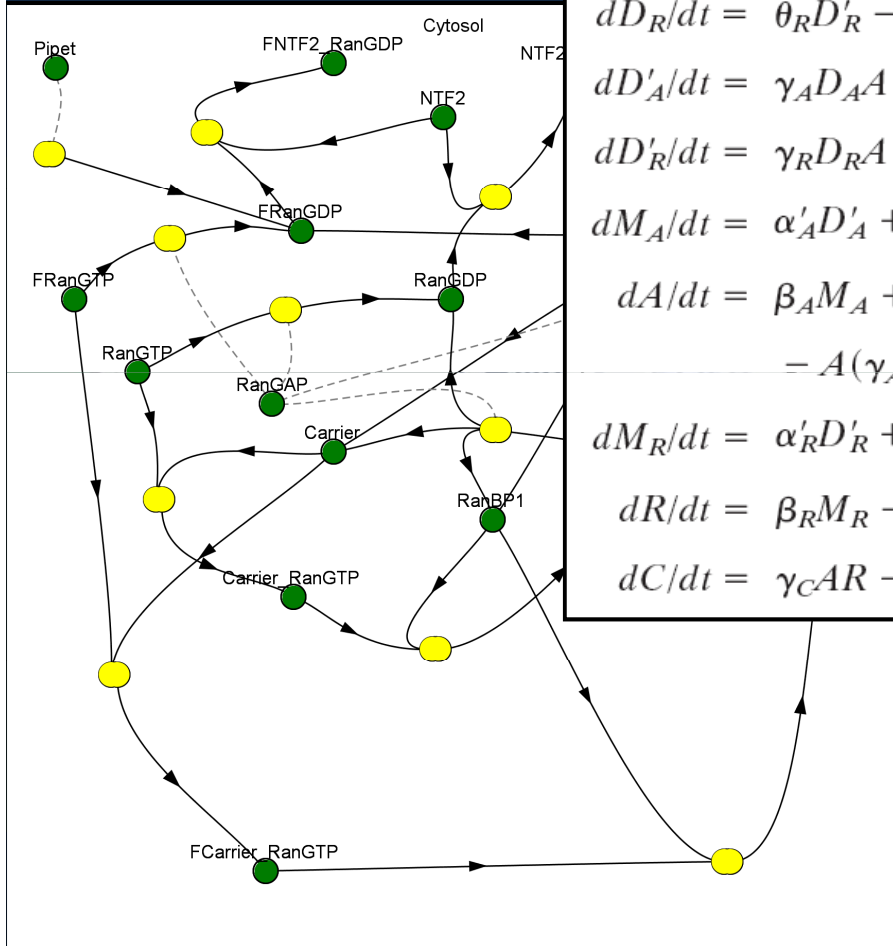
Richard D. Berlin
Center for Cell Analysis and Modeling

National Technology Center for Networks and Pathways

Outline

- The Virtual Cell (VCell) capabilities
- The VCell model architecture and standards
- Data-driven modeling

What is a *Model* ?



$$dD_A/dt = \theta_A D'_A - \gamma_A D_A A$$

$$dD_R/dt = \theta_R D'_R - \gamma_R D_R A$$

$$dD'_A/dt = \gamma_A D_A A - \theta_A D'_A$$

$$dD'_R/dt = \gamma_R D_R A - \theta_R D'_R$$

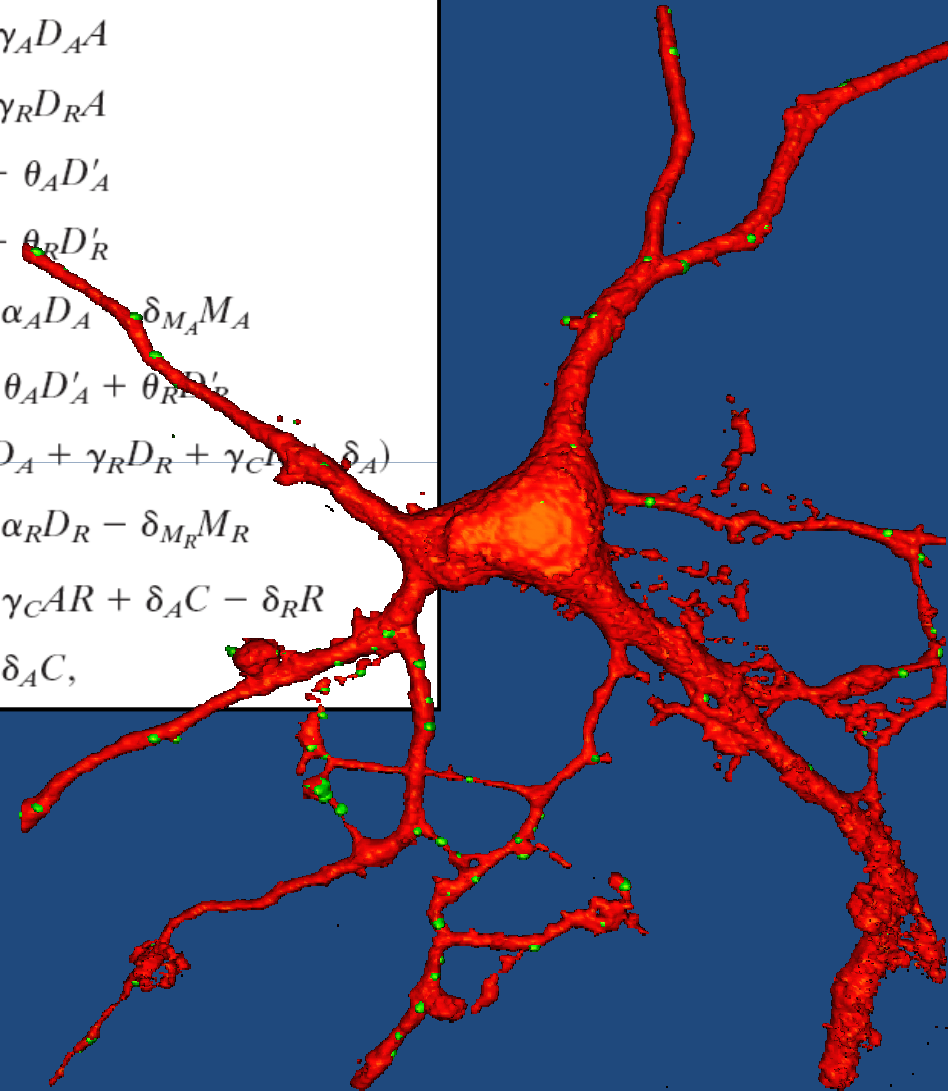
$$dM_A/dt = \alpha'_A D'_A + \alpha_A D_A - \delta_{M_A} M_A$$

$$dA/dt = \beta_A M_A + \theta_A D'_A + \theta_R D'_R - A(\gamma_A D_A + \gamma_R D_R + \gamma_C AR + \delta_A)$$

$$dM_R/dt = \alpha'_R D'_R + \alpha_R D_R - \delta_{M_R} M_R$$

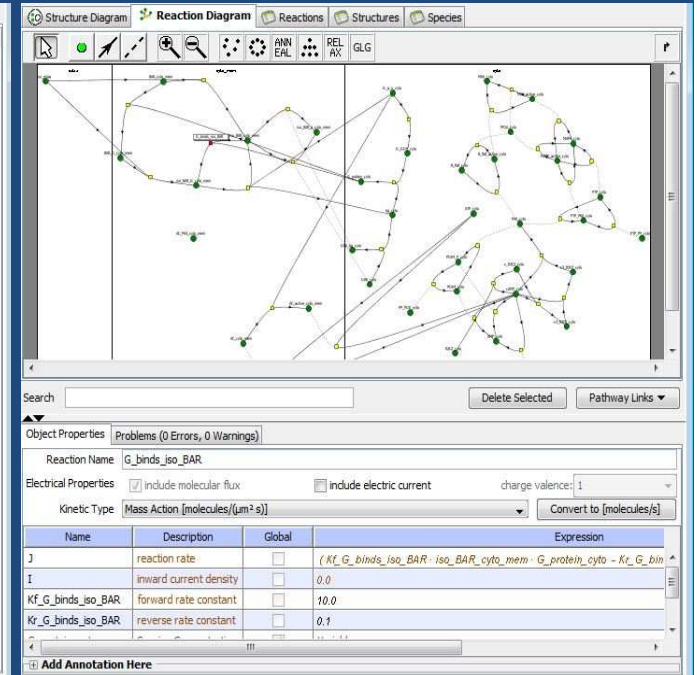
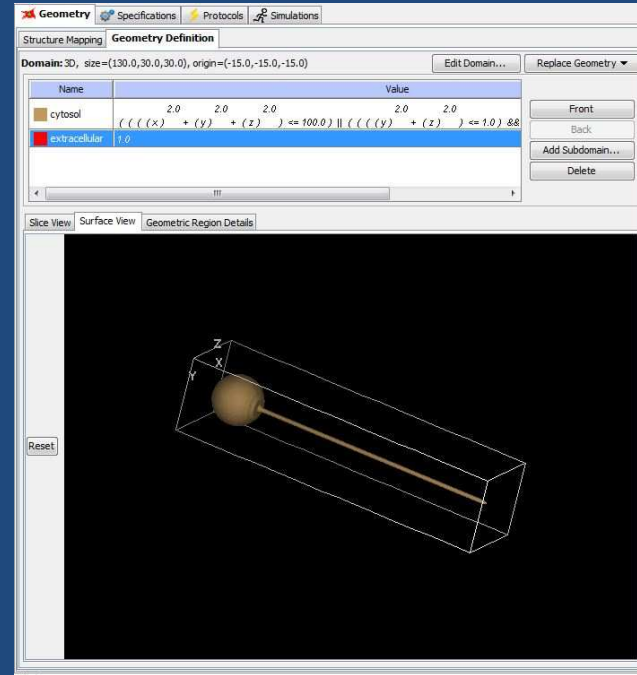
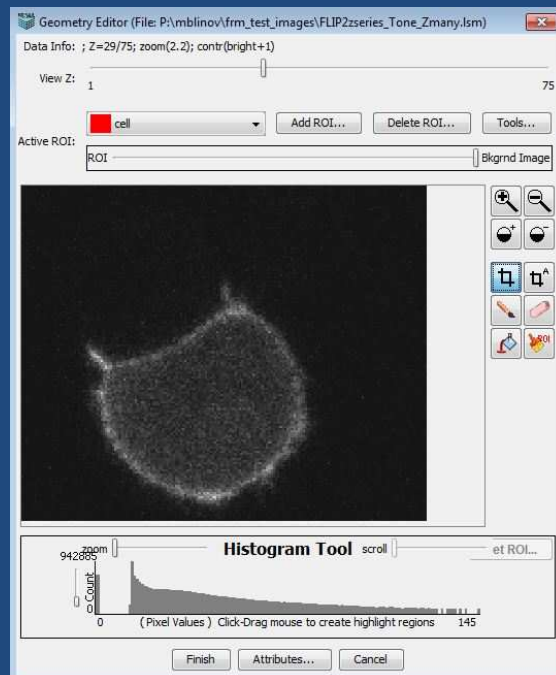
$$dR/dt = \beta_R M_R - \gamma_C AR + \delta_A C - \delta_R R$$

$$dC/dt = \gamma_C AR - \delta_A C,$$



The Virtual Cell (VCell)

- The VCell is a simulation environment where you can:
 - Import and segment 2D or 3D images (including microscopic z-stacks .lsm)
 - Create an analytic 1D, 2D or 3D geometry
 - Build a reaction network and place it into this geometry.

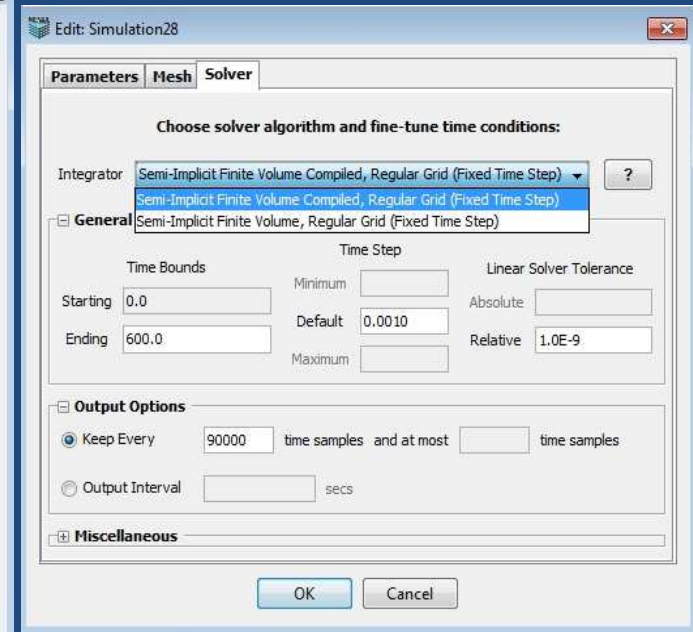
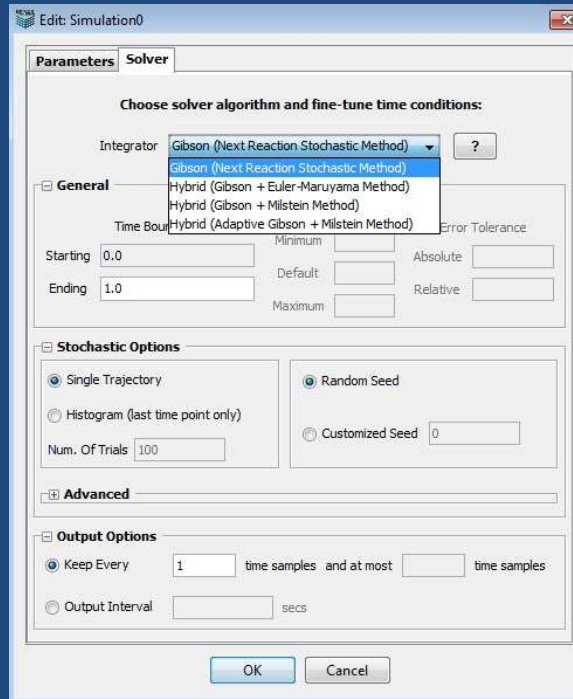
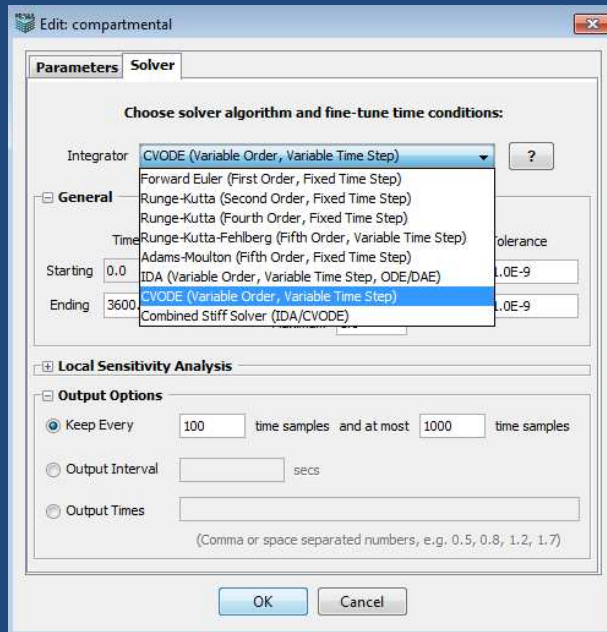
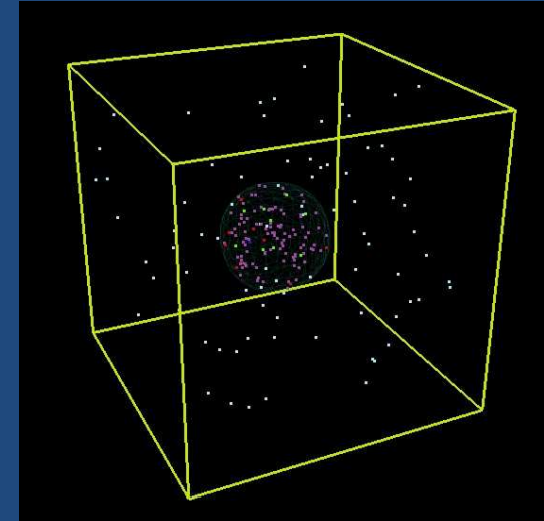


Virtual Cell Modeling Capabilities

- Many biological problems, e.g.
 - Signaling and metabolic pathways
 - Intracellular trafficking
 - Ion channels
 - Fluorescent indicators and probe redistribution
- Reaction - Diffusion - Advection - Electrophysiology
- 0, 1, 2 or 3 D including geometries from microscope images
- Rule-based modeling (BioNetGen)
- Links to external resources (BioModels, Pathways Common, KEGG, PSLID, etc)

Virtual Cell Simulation Capabilities

- Simulators:
 - Many ODE solvers
 - Many stochastic and hybrid solvers
 - PDE solvers
 - 3D spatial agent-based simulations (Smoldyn)
- Parameter scans and parameter estimation (Copasi)



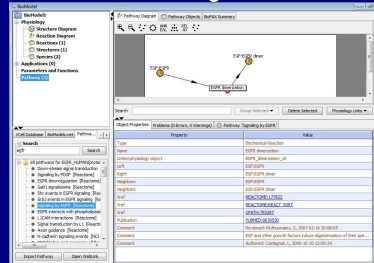
VCell interoperability

- Import/export:
 - SBML import/export,
 - CellML import,
 - MatLab export
 - BNGL import
 - Results export (xls, images, movies)
 - Import/export BioPAX/SBPAX
- Databases:
 - Import from BioModels.net database
 - Import from Pathway Commons
 - Linking reactions to KEGG
- Software integration (+50):
 - libSBML for import/export SBML
 - BioNetGen for rule-based modeling
 - Smoldyn for spatial agent-based stochastics
 - Copasi for parameter estimation
 - VisIt for visualization
- Open Science Grid

Outline

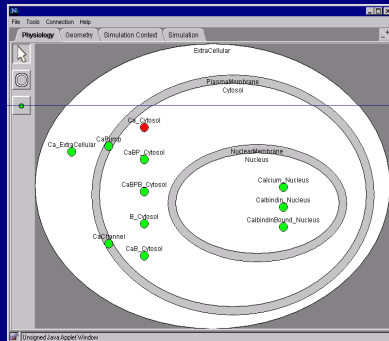
- The Virtual Cell (VCell) capabilities
- The VCell model architecture and standards
- Data-driven modeling

Pathway Model

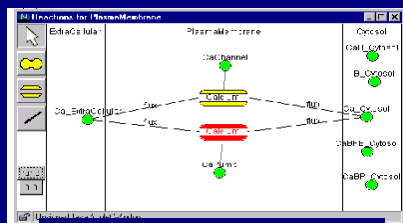


Imported BioPAX objects

Physiology

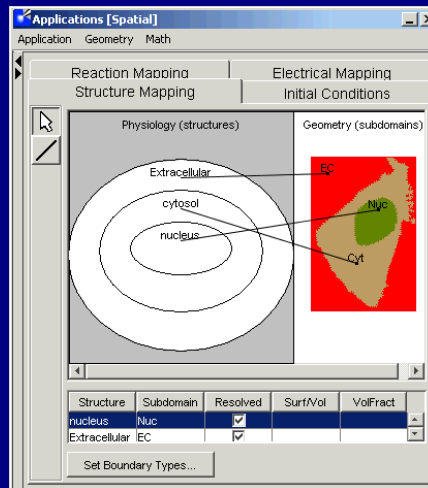


Molecular Species
Compartment Topology

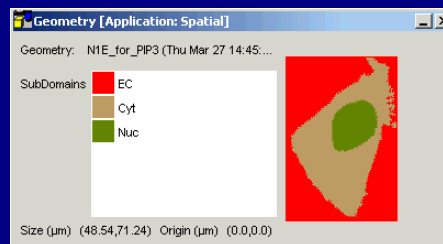


Reactions and Fluxes

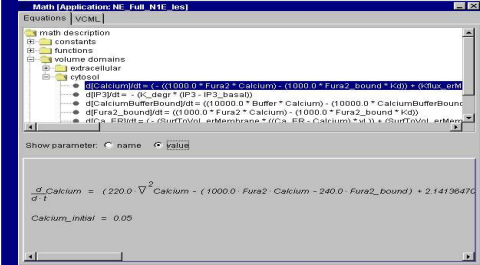
Applications



Topology → Geometry,
Initial Conditions, Boundary
Conditions, Diffusion
Coefficients, Pseudo-steady,
Enable/Disable Reactions
Electrophysiology Protocols
ODE/PDE/Stochastic



Math Description



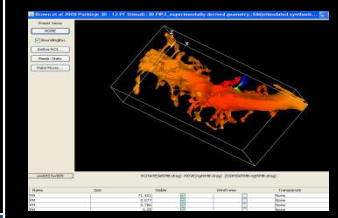
Simulations

Timestep,
Mesh Size,
Parameter
Searches,
Sensitivity

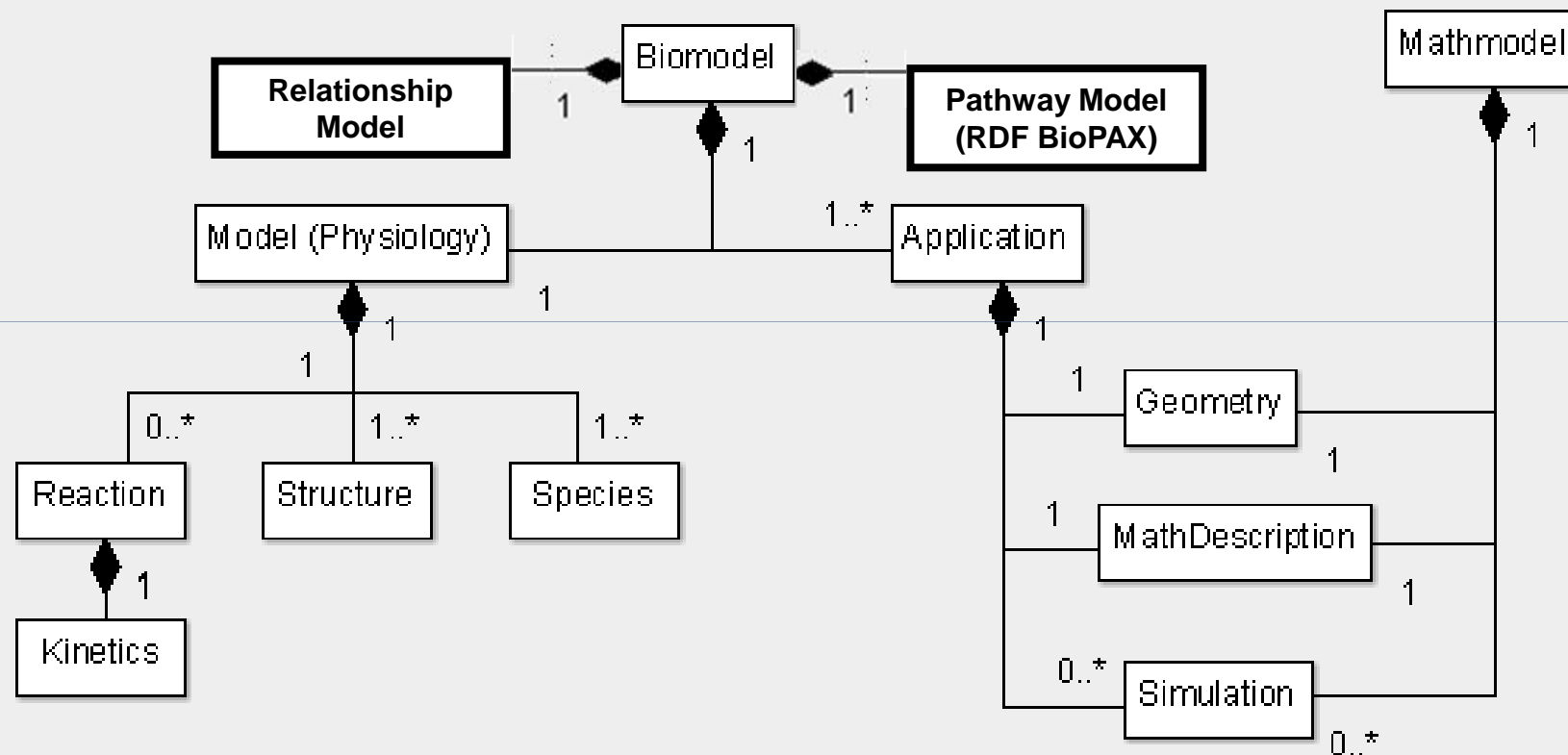


Results

Results



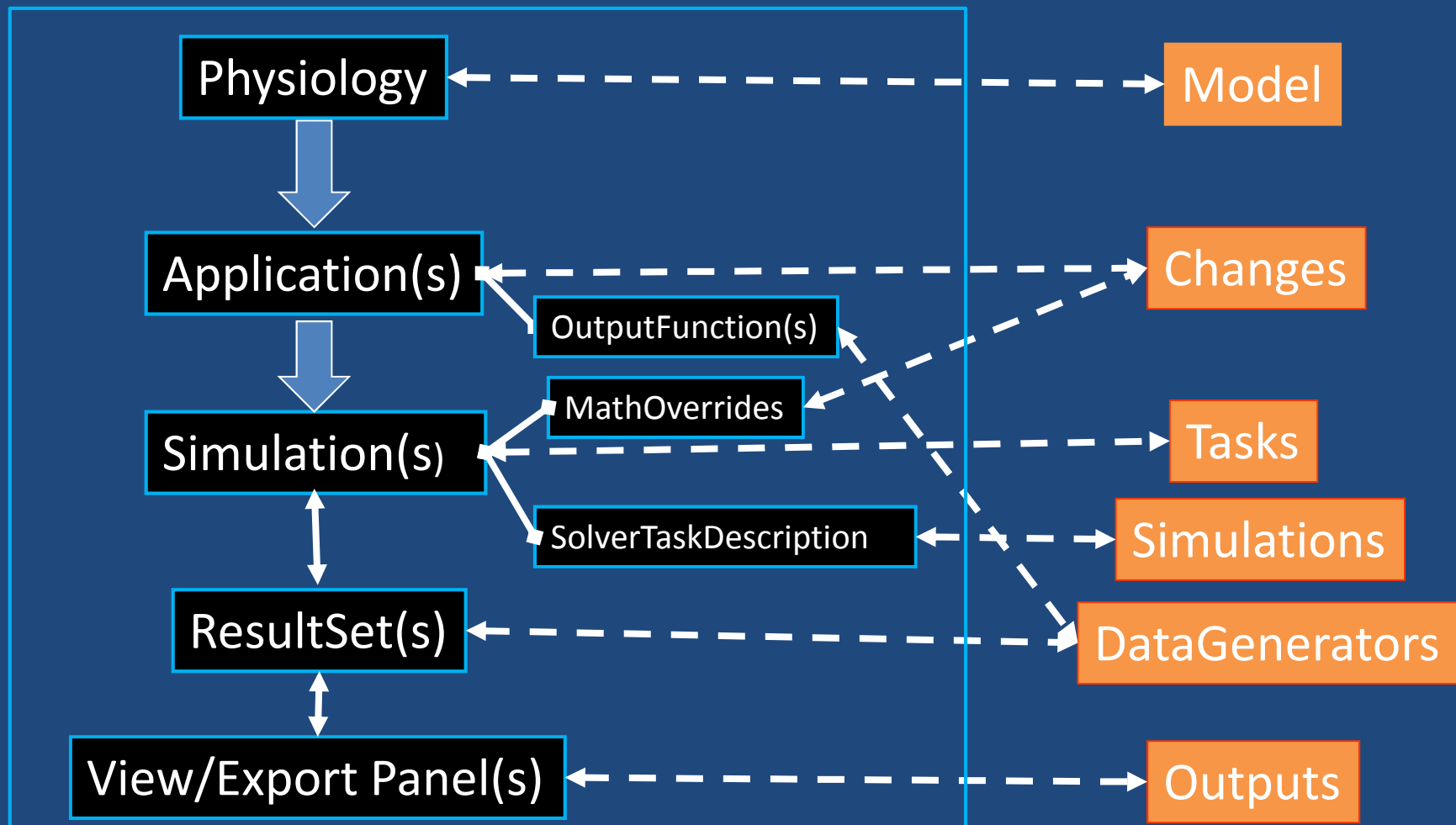
VCML = SBML + SED-ML + BioPAX



Currently stored models – 45,859
Currently stored applications – 62,091
Currently stored simulations – 263,533

VCML – SED-ML Mappings

VCell BioModel \longleftrightarrow SED-ML elements



VCell Database

- Model types: BioModels, MathModels, Geometries
- Models privacy:
 - Private
 - Shared with specific users
 - Public
- Access/import Biomodels.net database
- Search/import Pathways Common databases, including Reactome, NCI/PID, BioCyc collection, etc

Total Registered VCell Users - 14959

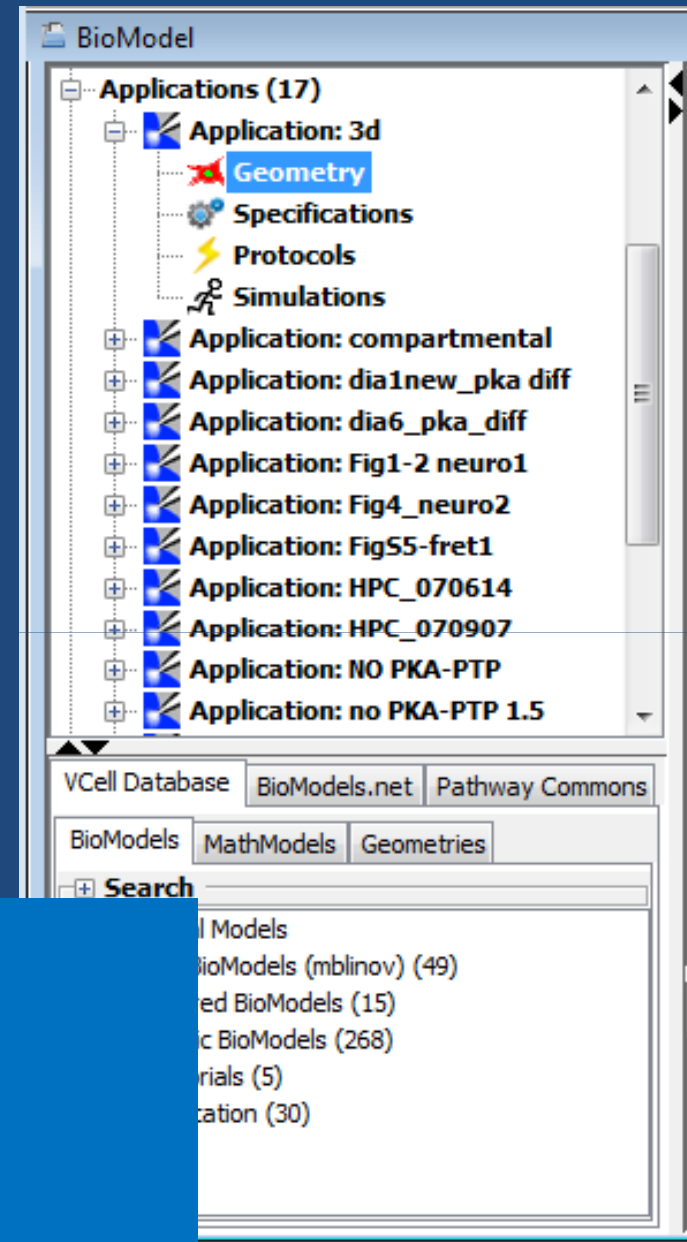
Users Who Ran Simulations - 3289

Total Models - 45859

Total Applications - 62091

Public Models - 507

Public Simulations - 3314



VCell: web-based vs client-based

- Internet connection is required for
 - Initial installation & updates
 - Database access
 - Running spatial or long simulations
 - Viewing saved results
 - Advantage – VCell is independent of user installation history
 - Always running whatever is currently deployed at <http://vcell.org/>
 - No backwards compatibility issues
- Registration is required for database access and server simulations
 - Free and instantaneous
- Model creation and ODE/stochastic simulations – can be run locally (VCell 5.1 local PDE/Smoldyn)

Outline

- The Virtual Cell (VCell) capabilities
- The VCell model architecture and standards
- Data-driven modeling

Bridging models and data: PROBLEMS

- Models are always created “ad-hoc”, multiple trials and errors. Users rarely use detailed species names, annotations, SBO terms, etc. Even curated and annotated models rarely have entity names that are crystal clear.

http://www.ebi.ac.uk/biomodels-main/BIOMD0000000250

<input type="checkbox"/> Compartments	<input type="checkbox"/> Species		
<input type="checkbox"/> default	<input type="checkbox"/> EGF	<input type="checkbox"/> HRG	
<input type="checkbox"/> cytoplasm	<input type="checkbox"/> A1	<input type="checkbox"/> A1_2	<input type="checkbox"/> A2
	<input type="checkbox"/> A2_2	<input type="checkbox"/> A3	<input type="checkbox"/> A3_2
	<input type="checkbox"/> DUSPmRNA	<input type="checkbox"/> ERK_c	<input type="checkbox"/> pERK_c
	<input type="checkbox"/> ppERK_c	<input type="checkbox"/> F	<input type="checkbox"/> c_FOS_c
	<input type="checkbox"/> pc_FOS_c	<input type="checkbox"/> c_FOSmRNA	<input type="checkbox"/> FmRNA
	<input type="checkbox"/> Kin	<input type="checkbox"/> Kin_2	<input type="checkbox"/> pMEK
	<input type="checkbox"/> MEK	<input type="checkbox"/> DUSP_c	<input type="checkbox"/> pDUSP_c
	<input type="checkbox"/> RSK_c	<input type="checkbox"/> pRSK_c	<input type="checkbox"/> RsD
	<input type="checkbox"/> RsT		
<input type="checkbox"/> nucleus	<input type="checkbox"/> CREB_n	<input type="checkbox"/> pCREB_n	<input type="checkbox"/> ERK_n
	<input type="checkbox"/> pERK_n	<input type="checkbox"/> ppERK_n	<input type="checkbox"/> Elk1_n

Bridging models and data: PROBLEMS

- Models are always created “ad-hoc”, multiple trials and errors. Users rarely use detailed species names, annotations, SBO terms, etc. Even curated and annotated models rarely have entity names that are crystal clear.
- SBML annotation schemas are often inadequate to identify model elements

<input type="checkbox"/> EGF_EGFR	Initial concentration: 0.0
Compartment: cytoplasm	
Annotations:	set#1 bqbiol:hasPart UniProt Q9QX70_RAT UniProt EGF_RAT
<input type="checkbox"/> (EGF_EGFR)2	Initial concentration: 0.0
Compartment: cytoplasm	
Annotations:	set#1 bqbiol:hasPart UniProt EGF_RAT UniProt Q9QX70_RAT

Bridging models and data: infrastructure

- Pathway databases, like Reactome, have a huge collection of entities and interactions.
- Some databases (SABIO-RK, BioCyc) have quantitative information.
- Some databases (UNIPROT, Bind, MmM, etc) have information about binding and modification sites, interactions and binding partners.
- Many databases provide SBML export. However, a modeler may want a simultaneous access to many resources.

Bridging models and data: VCell approach

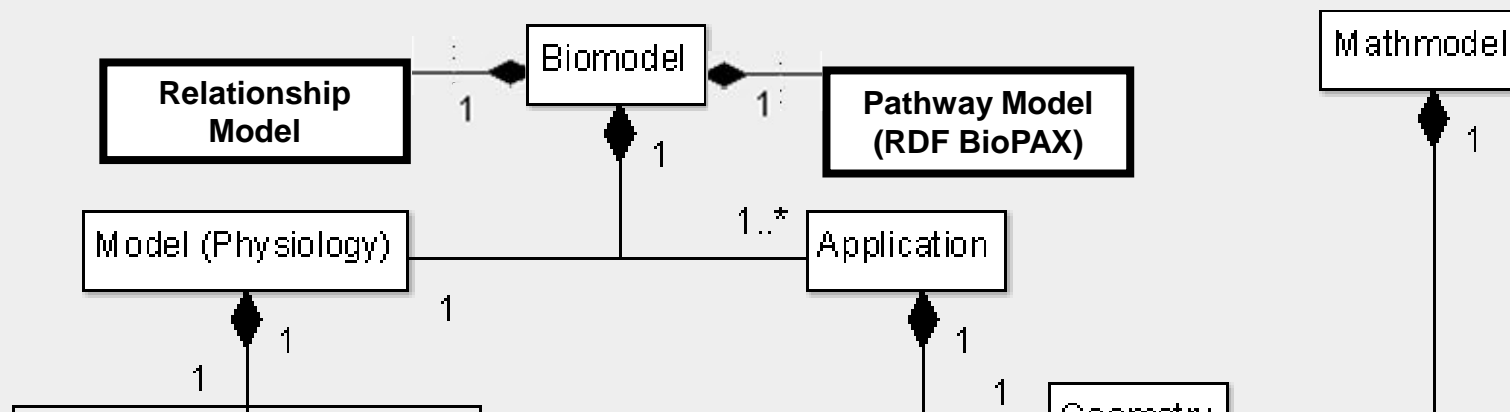
Search the WWW of pathways – Pathway Commons



Bring selected entities into VCell and store them as a Pathway Model



Convert elements of PM into new VCell model elements, or link existing VCell elements to PM elements.



BioModel

BioModel1

- Physiology
 - Structure Diagram
 - Reaction Diagram
 - Reactions (0)
 - Structures (1)
 - Species (0)
- Applications (0)
- Parameters and Functions
- Pathway (3)**

Pathway Diagram | Pathway Objects | BioPAX Summary

Search: Group Selected ▼ Delete Selected Physiology Links ▼

Object Properties | Problems (0 Errors, 0 Warnings) | Pathway "Signaling by EGFR"

Select	Entity Name	Type
<input type="checkbox"/>	Dimerisation of phospho-ERK-1	Biochemical Rea...
<input checked="" type="checkbox"/>	EGFR dimerization	Biochemical Rea...
<input type="checkbox"/>	Dimerisation of phospho-ERK-2	Biochemical Rea...
<input type="checkbox"/>	Nuclear translocation of phospho-ERK-2 dimer	Biochemical Rea...
<input type="checkbox"/>	Nuclear translocation of phospho-ERK-1 dimer	Biochemical Rea...
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer:CBL:Phospho-Sprouty ubiquitinated	Complex
<input type="checkbox"/>	SHP2-GRB2:Phospho GAB1-EGF-Phospho-EGFR dimer	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer:Phospho-CBL:CIN85 ubiquitinated:Endophilin:Epsin:Eps15R:Eps...	Complex
<input type="checkbox"/>	EGF:EGFR dimer	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer:CBL:CIN85	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer:Phospho-CBL	Complex
<input type="checkbox"/>	GRB2:Phospho GAB1-EGF-Phospho-EGFR dimer	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer Ubiquitinated:Phospho-CBL:GRB2	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR dimer:CBL:Phospho-Sprouty	Complex
<input type="checkbox"/>	EGF:Phospho-EGFR (Y1045) dimer:Phospho-CBL:CIN85 ubiquitinated:Endophilin:Epsin:Eps15R:Eps...	Complex

Search: dimer Add Selected

Search: Search

69 pathways for EGFR_HUMAN(protein)

- Down-stream signal transduction
- Signaling by PDGF [Reactome]
- EGFR downregulation [Reactome]
- Gab1 signalosome [Reactome]
- Shc events in EGFR signaling [Reactome]
- Grb2 events in EGFR signaling [Reactome]
- Signaling by EGFR [Reactome]**
- EGFR interacts with phospholipase
- L1CAM interactions [Reactome]
- Signal transduction by L1 [Reactome]
- Axon guidance [Reactome]
- N-cadherin signaling events [NCI]

Import Pathway Open Weblink

BioModel1

Physiology

Structure Diagram

Reaction Diagram

Reactions (0)

Structures (1)

Species (0)

Applications (0)

Parameters and Functions

Pathway (3)

Search

egfr

Search

69 pathways for EGFR_HUMAN(protein)

- Down-stream signal transduction
- Signaling by PDGF [Reactome]
- EGFR downregulation [Reactome]
- Gab1 signalosome [Reactome]
- Shc events in EGFR signaling [Reactome]
- Grb2 events in EGFR signaling [Reactome]
- Signaling by EGFR [Reactome]
- EGFR interacts with phospholipase
- L1CAM interactions [Reactome]
- Signal transduction by L1 [Reactome]
- Axon guidance [Reactome]
- N-cadherin signaling events [NCI]

Import Pathway

Open Weblink

Pathway Diagram

Pathway Objects

BioPAX Summary

+

-

•••

•••

ANN

EAL

•••

REL

AX

•••

Search

Group Selected

Delete Selected

Physiology Links

Object Properties

Problems (0 Errors, 0 Warnings)

Pathway "Signaling by EGFR"

Property	Value
Type	Biochemical Reaction
Name	EGFR dimerization
Left	EGF:EGFR
Right	EGF:EGFR dimer
Neighbors	EGF:EGFR
Neighbors	EGF:EGFR dimer
Xref	REACTOME:177922
Xref	REACTOME:REACT_9397
Xref	CPATH:793297
Publication	PUBMED:8639530
Comment	Reviewed: Muthuswamy, S, 2007-02-16 20:08:05
Comment	EGF and other growth factors induce oligomerization of their specific receptors. Inactive EGFR monomers ar...
Comment	Authored: Castagnoli, L, 2006-10-10 13:09:34

BioModel

BioModel1

Physiology

Structure Diagram

Reaction Diagram

Reactions (0)

Structures (1)

Species (0)

Applications (0)

Parameters and Functions

Pathway (3)

Pathway Diagram

Pathway Objects

BioPAX Summary

+

-

•••

•••

ANN

EAL

•••

REL

AX

•••

Search

Group Selected

Delete Selected

Physiology Links

VCell Database

BioModels.net

Pathwa...

Search

egfr

Search

69 pathways for EGFR_HUMAN(protei

Down-stream signal transduction

Signaling by PDGF [Reactome]

EGFR downregulation [Reactome]

Gab1 signalosome [Reactome]

Shc events in EGFR signaling [Rea

Grb2 events in EGFR signaling [Re

Signaling by EGFR [Reactome]

EGFR interacts with phospholipase

L1CAM interactions [Reactome]

Signal transduction by L1 [Reacto

Axon guidance [Reactome]

N-cadherin signaling events [NCI

Import Pathway

Open WebLink

Import into Physiology

Interaction	Type	Entity Name	Entity Type	Stoich. Coef.	Location/Compartment	ID
		EGF:EGFR	Complex	1	c0	EGF_...
EGFR dimeriza...	Conversion	EGFR dimeriz...	Biochemical ...	1	c0	EGFR...
EGFR dimeriza...	Product	EGF:EGFR di...	Complex	1	c0	EGF_...
EGFR dimeriza...	Reactant	EGF:EGFR	Complex	2	c0	EGF_...

Search

OK

Cancel

BioModel

BioModel1

- Physiology
 - Structure Diagram
 - Reaction Diagram
 - Reactions (1)
 - Structures (1)
 - Species (2)
- Applications (0)
- Parameters and Functions
- Pathway (3)**

VCell Database | BioModels.net | Pathwa...

Search
 egfr

69 pathways for EGFR_HUMAN(protei...
 • Down-stream signal transduction
 • Signaling by PDGF [Reactome]
 • EGFR downregulation [Reactome]
 • Gab1 signalosome [Reactome]
 • Shc events in EGFR signaling [Rea...
 • Grb2 events in EGFR signaling [Re...
 • **Signaling by EGFR [Reactome]**
 • EGFR interacts with phospholipase...
 • L1CAM interactions [Reactome]
 • Signal transduction by L1 [Reacto...
 • Axon guidance [Reactome]
 • N-cadherin signaling events [NCI...
 • Stabilization and expression of the...

Pathway Diagram | Pathway Objects | BioPAX Summary

Search

Object Properties | Problems (0 Errors, 0 Warnings) | **Pathway "Signaling by EGFR"**

Property	Value
Type	Biochemical Reaction
Name	EGFR dimerization
Linked physiology object	EGFR_dimerization_c0
Left	EGF:EGFR
Right	EGF:EGFR dimer
Neighbors	EGF:EGFR
Neighbors	EGF:EGFR dimer
Xref	REACTOME:177922
Xref	REACTOME:REACT_9397
Xref	CPATH:793297
Publication	PUBMED:8639530
Comment	Reviewed: Muthuswamy, S, 2007-02-16 20:08:05
Comment	EGF and other growth factors induce oligomerization of their spe...
Comment	Authored: Castagnoli, L, 2006-10-10 13:09:34

Bridging models and data: quantitative

- SBPAX: providing pathway databases with BioPAX extension capable of encoding quantitative information
- Bringing some quantitative data from BioCyc
- Retrieving kinetic data from SABIO-RK
- Problems:
 - Match retrieved data with existing BioModel entities and BioPAX data in Pathway Model
 - Change/update reactions in BioModel (add /remove ATP/ADP transfer)
 - Deal with parameter ranges
 - Each time new data is brought from Pathway Commons, it should be compared with data retrieved from SABIO-RK

Bridging models and data: SABIO-RK

Search SABIO-RK



Retrieve a list of reactions/pathways



Retrieve a list of kinetic laws AND experimental conditions



Compare with existing entities in Pathway Model by KEGG, Chebi or UniPROT id



Create new entries in Pathway Model or link with existing entries
AND
Create new linked species/reactions in BioModel or assign kinetic law
to existing reactions

Each time new data is brought from Pathway Commons,
it should be compared with data retrieved from SABIO-RK

BioModel

BioModel1

- Physiology
 - Reactions (0)
 - Reaction Diagram
 - Structures (1)
 - Structure Diagram
 - Species (0)
- Applications (0)
- Parameters and Functions
- Pathway (0)

VCell Database | BioModels.net | Pathway Commons | SabioRK

Search

JAK

Search

☒ Pathway ☐ Compound ☐ Enzyme

Reaction	Pathway
1 ((IFNRJ2_activated)(SHP2)) <-> 1 IFNRJ2 ...	JAK/STAT signal tran...
1 ((IFNRJ2_activated)(STAT1)) <-> 1 IFNRJ...	JAK/STAT signal tran...
1 ((STAT1_activated)(P)) <-> 1 Phosphat...	JAK/STAT signal tran...
1 ((STAT1_activated)(STAT1_activated)) <-> 1...	JAK/STAT signal tran...
1 ((STAT1_activated)(STAT1_activated)) + 1...	JAK/STAT signal tran...
1 ((STAT1_activated)(STAT1_activated)) <->...	JAK/STAT signal tran...
1 Ciliary neurotrophic factor receptor alpha ...	JAK/STAT signal tran...
1 IFNRJ2 <-> 1 IFNRJ2_activated	JAK/STAT signal tran...
1 IFNRJ2_activated + 1 SHP2 <-> 1 ((IFNRJ...	JAK/STAT signal tran...

Import SabioRK Data

Open Link

Reactions

Reaction Diagram

Structures

Structure Diagram

Species

Equation	Name	Structure	Kinetics
(add new here, e.g. $a+b \rightarrow c$)			

Search

Add New

Delete Selected

Pathway Links

Object Properties

Problems (0 Errors, 0 Warnings)

Organism: [Mus musculus](#)

Tissue: liver

EC Class: -

wildtype

Substrates

name	location	comment
Ciliary neurotrophic factor receptor alpha	cytosol	-
JAK-1	cytosol	-

Products

name	location	comment
Receptor JAK complex	cytosol	-

Modifiers

name	location	effect	comment	protein complex
Enzyme	cyt	modif	-	-

Enzyme (protein d

name	UniProt	Enzyme	Mol. weight (kDa)	deviation (kDa)
subunit	-	-	-	-
complex	-	-	-	-

Kinetic Law

type	formula
Mass action rate law for first order irreversible reactions	$cytoplasm * (kf1 * JAK - kb1 * RJ)$

Parameter

name	type	spec	sta	nd	al.	id	un	comment
cytoplasm	volume	-	-	-	-	-	-	-
R	concentration	Ciliary neurotrophic factor receptor alpha	12.0	-	-	-	-	nM
RJ	concentration	Receptor JAK complex	-	-	-	-	-	-
kb1	unimolecular rate constant	-	0.05	-	-	-	-	s^{-1}
kf1	bimolecular rate constant	-	0.1	-	-	-	-	$nM^{-1} * s^{-1}$
JAK	concentration	JAK-1	12.0	-	-	-	-	nM

Signalling reaction event

event description	type	event type name	event type description	annotation
Receptor-JAK Binding	-	-	protein binding	GO:0005515
Receptor-JAK Binding	-	-	interferon-gamma receptor binding	GO:0005133

Reference

title	author	year	journal	volume	pages	PubMed
Control mechanism of JAK/STAT signal transduction pathway	Yamada S, Shiono S, Joo A, Yoshimura A	2003	FEBS Lett	534	190-6	12527385

SABIO-RK screenshot

Virtual Cell data perspectives

- MIRIAM-compliant annotations
- More interacting with databases:
 - Better unified search among Pathway Commons and SABIO-RK
 - Getting parameters from ByoCyc and SABIO-RK databases,
 - Getting compartments from GO and FMA
 - Complex queries among multiple RDF resources (Jena, Oliver Ruebenacker)
 - NEW IDEA: Queries among BioPAX, BioModels and simulation data – Michel Dumontier
 - Working with experimental data-sets.
- Integration with rule-based modeling: common interface for specifying reaction networks and rules
- Visualization of models as SBGN-PD and SBGN-ER
- Possibility to identify and reuse modeling components

The Virtual Cell Project



Richard D. Berlin

Center for Cell Analysis and Modeling

National Technology Center for Networks and Pathways

Les Loew
Jim Schaff
Ion Moraru
Boris Slepchenko

Michael Blinov

Ed Boyce
Fei Gao
Anu Lakshminarayana
Frank Morgan
Igor Novak
Diana Resasco

Oliver Ruebenacker

Dan Vasilesku

Xintao Wei

Li Ye



NIH Roadmap



**National Center for
Research Resources**



**National Institute of
General Medical Sciences**



**Cell
Migration
Consortium**