

Collaborative Drug Discovery – UCSF 3rd Annual Community Meeting

Archive, Mine, and (selectively) Collaborate



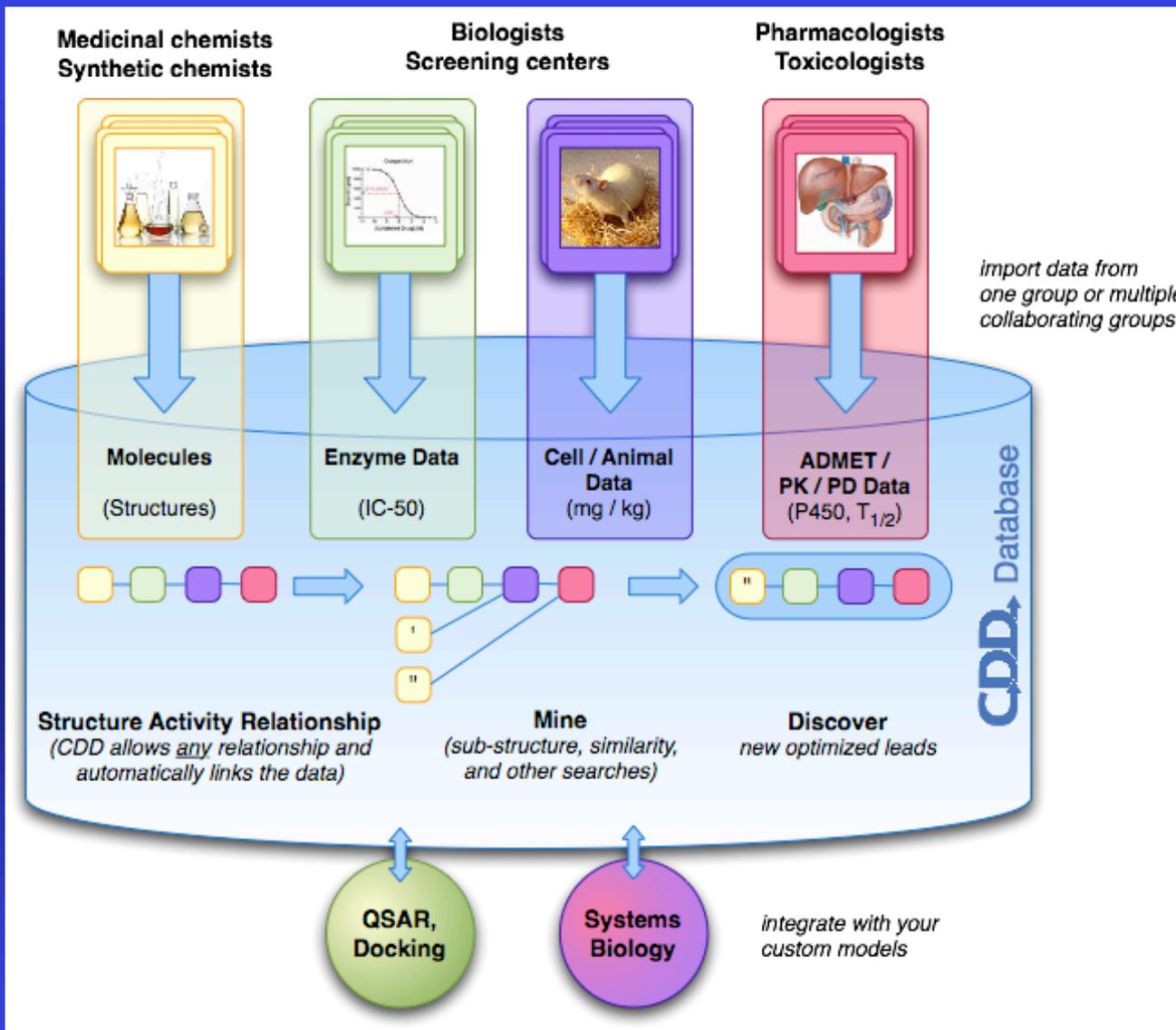
COLLABORATIVE
DRUG
DISCOVERY

What if?

- Technology
- Collaborations



- **Modules:**
 - **CDD Vault** – Secure place for private data – private by default
 - **CDD Collaborate** – *Selectively* share subsets of data
- **Datasets:**
 - **CDD Public** – Expanding public data sets
 - Published compounds
 - Data from community members
 - Vendor libraries
- **Unique to CDD** – *Simultaneously* query your private data, collaborators' data, & public data





COLLABORATIVE
DRUG
DISCOVERY

Collaborate Mode 1 – Big Pharma with single login to multiple CROs/Biotech groups

**Consolidated
CDD DB**

Company A

Project A

Project B

Project C

Project D



Collaborator 1

Collaborator 2

Collaborator 3

Collaborator 4



COLLABORATIVE
DRUG
DISCOVERY

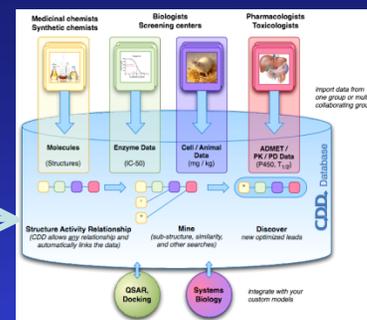
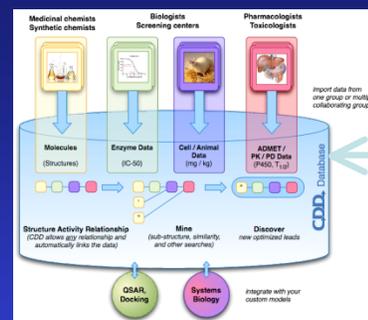
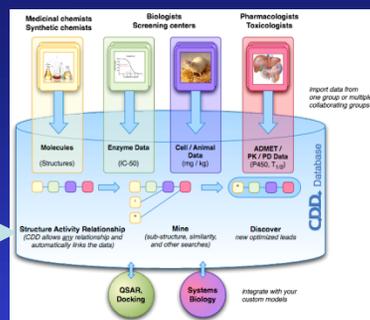
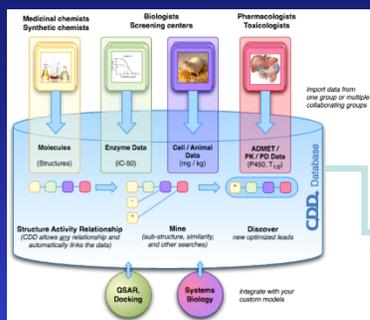
Collaborate Mode 2 – Academic + Foundations securely sharing data subsets

CDD DB
group 1

CDD DB
group 2

CDD DB
group 3

CDD DB
group 4





COLLABORATIVE
DRUG
DISCOVERY

Use CDD Vault™, Collaborate™, Public™ for
Global R&D Community

New Types of Offerings for a Rapidly Changing Industry

- CDD Community: tap into global R&D via “crowd-sourcing”
- CDD Alliance: better management of your collaborations
- CDD Customize: fast, good, and cost-effective = agile



COLLABORATIVE DRUG DISCOVERY
Barry Bunin's Site

Welcome, Barry: [Account Preferences](#) | [Logout](#)

Archive

Mine

Collaborate

Currently viewing my group's data and 41 shared data sets

[Protocols](#) | [Molecules](#) | [Plates](#) | [References](#) | [Files](#)

In vivo productivity assay ...

Approved Drugs - JHU- 1581

FDA/Orphan Drugs - caffeine

Malaria/Trypanosome: St. Ju...

Maximum recommended daily d...

MLSMR - SMR000326667

PDSP Ki Database - 19520

TimTec Natural Product Libr...

Molecules /

Caffeine

Privately Shared Molecule

Owner: [Moses Hohman](#)

Created: 10/3/2008

Updated: 10/3/2008

Moses Hohman

Email: moses@collaborativedrug.com

Group: Hohman Lab

PI: Moses Hohman (moses@moseshohman.com)

Institution: Collaborative Drug Discovery, Inc.

Definition

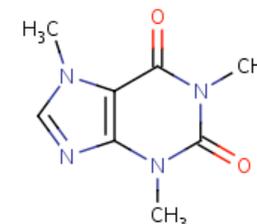
Name Caffeine

Synonyms

Description

Structure [SMILES](#) | [CXSMILES](#) | [InChi](#) | [InChiKey](#)

CDD-90



[Find other molecules containing this structure](#)

[ChemSpider page](#)

Lipinski properties

Molecular 194.191 g/mol



COLLABORATIVE
DRUG
DISCOVERY

CDD Public

CDD • Barry Bunin's Site

Barry Bunin: [Your Account](#) · [Log out](#)

[Dashboard](#)

[Group Data](#)

[Import Data](#)

[Share Data](#) **4**

Currently using **Barry Bunin's Site** data and **2 shared data sets** ⓘ

[Choose data sets ...](#)

[Search](#)

[Saved Searches](#)

[Molecules](#)

[Protocols](#)

[Plates](#)

[References](#)

[Files](#)

Search references ...

any field

Go

Help ⓘ

[+ Add a new reference](#)

[📄 Search and import from PubMed](#)

1 Reference · Showing per page

[Previous](#) · Page of 1 · [Next](#)

Reference

Keywords

PUBLIC REFERENCE TB Efficacy Data from Published Literature

Synthesis and antitubercular activity of 7-(R)- and 7-(S)-methyl-2-nitro-6-(S)-(4-(trifluoromethoxy)benzyloxy)-6,7-dihydro-5H-imidazo[2,1-b][1,3]oxazines, analogues of PA-824.

Li, X.; Manjunatha, U. H.; Goodwin, M. B.; Knox, J. E.; Lipinski, C. A.; Keller, T. H.; Barry, C. E. 3rd; Dowd, C. S.

[Bioorg Med Chem Lett](#) (2008) Vol 18, No 7, pp 2256-2262

[Abstract](#) · [Full text](#) · [View in PubMed](#)

[Previous](#) · Page of 1 · [Next](#)

Mapping template: - Select a mapping template -

Fields	Molecule name	SMILES	SYNONYM	SOLVENT	T. Cruzei IC50 (nM)	T. Cruzei ...Su
	CDD121	<chem>Oc1cc2ccc...cc5)cc34</chem>	ABC121	DMSO	100	1
	CDD122	<chem>COc1ccc2c...cc5)cc34</chem>	ABC122	DMSO	101	2
Data preview	CDD123	<chem>CC(=O)n1c...4cccc14</chem>	ABC123	DMSO	102	3
	CDD124	<chem>Brc1ccc(c...[nH]3)o2</chem>	ABC124	DMSO	103	4
	CDD125	<chem>O=C(NN=C...cc4[nH]3</chem>	ABC125	DMSO	104	5
Map to	Molecule Name or Synonym	Molecule Structure	Unmapped	Unmapped	Unmapped	Unmapped

▼ Molecule Fields

 Molecule Name or Synonym

 Molecule Structure

 User-defined Field

► Batch Fields

► Plate and Well

 Readouts

 Leave Unmapped

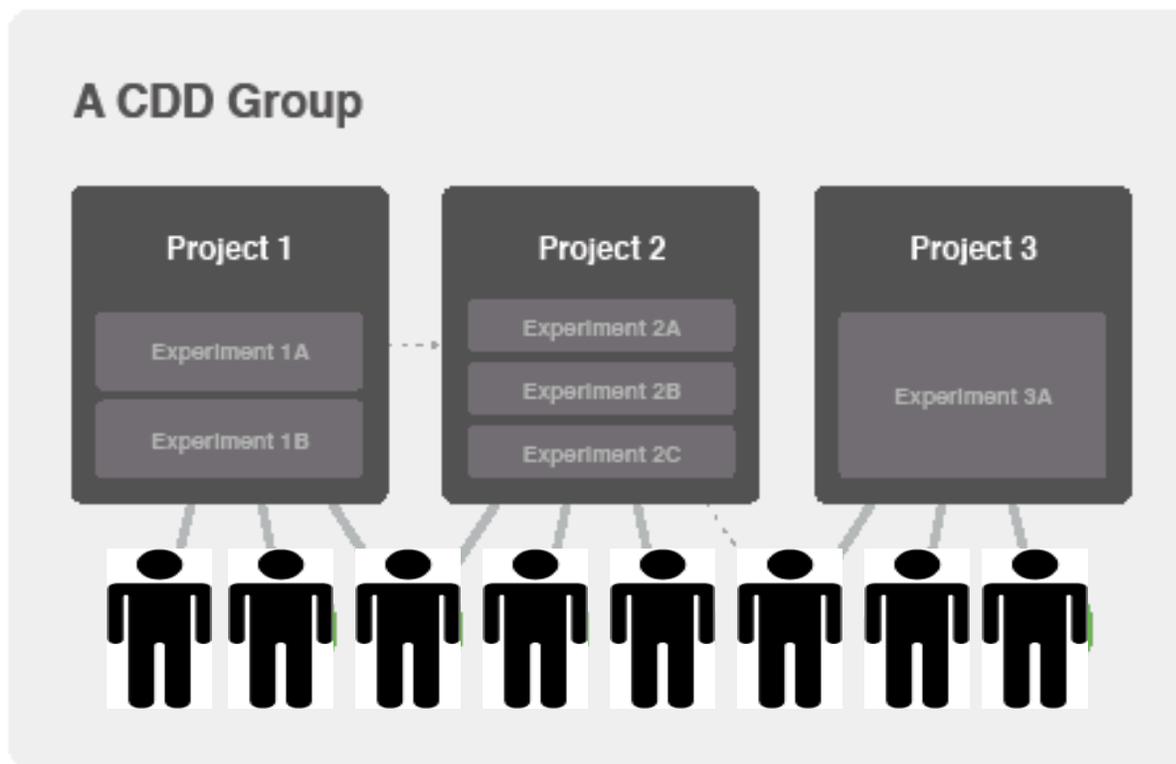
Molecule name is mapped to Molecule Name or Synonym

You can use this field to map any molecule name or synonym.

 Use the values of this field as the names of any newly created molecules



- **Mapper Overhaul** – more intuitive to batch upload complex data, save mappings
 - **Projects** – more advanced collaborative features beyond private-to-private data
 - Editable data
 - Refined data viewing selections
 - Multi-group functionality requested by screening centers, foundations, & biopharma
 - **Configurable Terminology** –
 - Configure CDD w/ your own labels (alts to molecules, protocols...like objects, antibodies, sequences, screens, etc)
 - **Advanced Filters & Alerts for Compound Liabilities**
 - Reactive Functionality likely to cause toxicity, alkylate DNA, BBB models, etc
 - What Lipinski calls Alerts (for example with Abbotts’ ALERT SMARTS filters).
- NOTE: For each feature we do a “bang per buck” analysis, and truly agonize over which item is best to do first, given lots of competing requests



Projects provide subdivisions of data and data privacy



COLLABORATIVE
DRUG
DISCOVERY

Configurable Terminology: “User-Defined” Objects, Subjects, Molecules

CDD • Bunin Group ▾

Barry Bunin: [Admin](#) · [Your Account](#) · [Help](#) · [Log out](#)

[Dashboard](#)

[Group Data](#)

[Import Data](#)

[Share Data](#)

[Recent Activity](#)

[Messages 2](#)

[Group Settings](#)

[People](#)

Your Recent Screens

[See all screens](#)

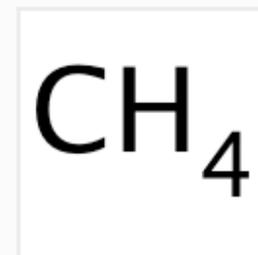
[Create a Screen](#)

Screen	Type	Substances
Falcipain IC50	enzyme	80

Your Recent Substances

[See all substances](#)

[Create a Substance](#)



[ABC-001](#)

[▶ All Recent Activity](#)



COLLABORATIVE
DRUG
DISCOVERY

Configurable Terminology: “User-Defined” Objects, Subjects, Biologicals

CDD • Bunin Group

Barry Bunin: [Admin](#) · [Your Account](#) · [Help](#) · [Log out](#)

Dashboard

Group Data

Import Data

Share Data

Recent Activity

Messages 2

Group Settings

People

Your Recent Screens

[See all screens](#)

Create a Screen

Screen	Type	Substances
Falcipain IC50	enzyme	80

Your Recent Substances

[See all substances](#)

Create a Substance

```
GAATTC TCTTTGGTATCCAATGAAGAAATCGAATCCA  
TTCAGGAGAAAATAAGACCGAAGCTGCTCAATTAGGC  
GTGAAACTTGCCAGCTTACTTCGGCATGTCCTGGTCA  
CAACCATTATTTAAAGTCGCATTTAAAAAACTTGTTG
```

[ABC-001](#)

► [All Recent Activity](#)



COLLABORATIVE
DRUG
DISCOVERY

Drug Toxicity Alerts

Caffeine - Collaborative Drug Discovery

file:///Users/moses/code/design-svn/templates/molecule.html

Google

CDD Lipinski Group

Chris Lipinski: [Your Account](#) · [Help](#) · [Log out](#)

[Dashboard](#)

[Group Data](#)

[Import Data](#)

[Share Data](#)

Currently using **Lipinski Group** data and **6 shared data sets** ⓘ

[Choose data sets ...](#)

[Back to Dashboard](#)

CAS 957217-65-1 PUBLIC MOLECULE

Available in 3 data sets. Now viewing:

TB MIC Prathipati NIAID - 37532

[Overview](#)

[Batches 2](#)

[Plates 1](#)

[Protocols 3](#)

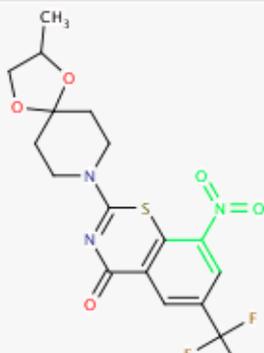
[Associated Files 2](#)

[Alerts 2](#)

2 Alerts

Structure Match Preview

Alert



⚠ Abbott Alarm NMR failed

- CSC
- Nc1cccc2aaaaa12

⚠ Pfizer Lint failed

- O~N(=O)-c(:*):* (aromatic NO2)



COLLABORATIVE
DRUG
DISCOVERY

Implementing Toxic Functionality Alerts (see Lipinski's presentation)

- **Basic Capability = 2 FTE weeks**
 - Store list of SMARTS alerts with intuitive description
 - Associate list of alerts with structures
 - Set up alerts to be automatically run against new and existing structures
 - Display alerts
 - **End User Flexibility = 1 FTE week**
 - **Advanced Implementation = 3 FTE weeks**
- NOTE: Generally any feature can have about 25% of unforeseen features



- **Advanced visualization tools** – for anything mined.
 - **Mine page extensions** – e.g., more list logic on query results
 - **Performance-Speed Optimization (mapper/slurper/mining)**
 - **Processing data (averaging, SD, etc) and advanced calculations** – “Excel” manipulations
 - **API** – program your own capabilities and integrate nicely with other technologies
 - **Multi-group, Multi-dimensional MedChem Collaborations** –
 - Social Networking
 - Community of models to support complex, collaborative projects
 - **Engage Big pharma**
 - Data, models, compounds, & capabilities
 - “Pre-competitive”
- **Reminder:** For each feature we do a “bang per buck” analysis



COLLABORATIVE
DRUG
DISCOVERY

Collaborative Dashboard for Secure To-the-Minute Global R&D Tracking

Dashboard - Collaborative Drug Disco...



COLLABORATIVE DRUG DISCOVERY

Barry Bunin's Site

Welcome, Barry: [Account](#) [Preferences](#)

Archive

Mine

Col

Currently viewing my group's data and [51 shared data sets](#)

[Dashboard](#) | [Message Board](#)

Dashboard

Latest activity in the [Barry Bunin's Site](#) ([switch group](#))

There are pending publication requests for you to [review...](#)

Last week

Molecule	Barry Bunin created molecule GA Analog 7	Thursday, September 3
Molecule	Barry Bunin updated molecule 10 from Table 1	Thursday, September 3
Molecule	Barry Bunin updated molecule 10 from Table 1	Wednesday, September 2
Protocol	Barry Bunin updated protocol Aug 2009 meeting	Wednesday, September 2
Protocol	Barry Bunin created protocol Aug 2009 meeting	Wednesday, September 2
Molecule	Barry Bunin updated molecule Dundee 6890	Wednesday, September 2
Protocol	Barry Bunin created protocol Dundee consortium Assay 5	Wednesday, September 2
Molecule	Barry Bunin updated molecule Dundee 6890	Wednesday, September 2
Molecule	Barry Bunin created molecule Dundee 6890	Wednesday, September 2
Molecule	Barry Bunin updated molecule 10 from Table 1	Wednesday, September 2
User	Barry Prom joined Barry Bunin's Site	Tuesday, September 1

Your Groups

✓ Barry Bunin's Site

- Barry Bunin (group admin)
- Krishna Dole (full access)
- Graham Douglas (read-add)
- Sean Ekins (read-export)
- Sylvia Ernst (read-add)
- Sylvia Ernst (readonly)
- Sylvia Ernst (full access)
- Lauren Ferris (read-add)
- Andrew Hutchison (read-export)
- Navin Kadaba (full access)
- Lixin Liu (read-export)
- William Nelson (read-export)
- Barry Prom (read-export)
- Dan Robertson (read-add)
- Anna Spektor (read-add)
- Demo User (read-export)
- Anatoly Volovik (read-export)
- Tony Williams (full access)
- Todd Wipke (read-add)

ASINEX

ChemBridge

Done

Save this search...

Show **first 1**

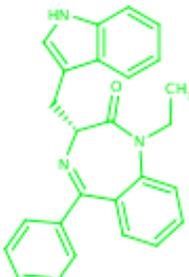
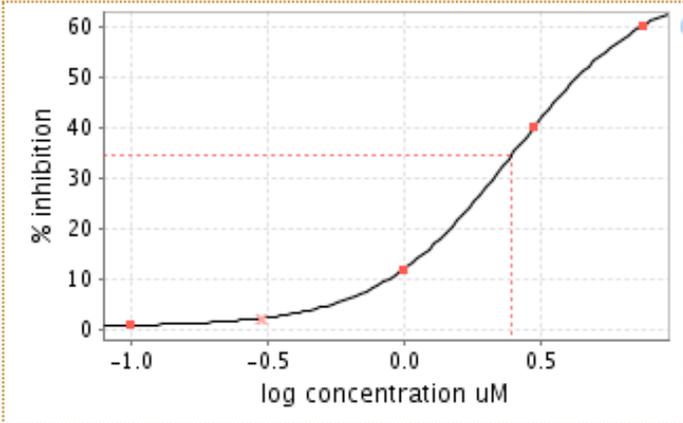
Search Results

9 matching results

Run: 2008-03-24 — Batch: unspecified

IC ₅₀	2.51 uM
Hill slope	1.76
R ²	1.0
Baseline (percent)	0.767
Maximum (percent)	68.6
# of data points used	N = 4 (and 1 outlier)

[Display options](#) | [Export 9 res](#)

Molecule	CCK-A Dose	Dose-response Plot	CCK A receptor binding a	
	IC ₅₀ (uM)		IC ₅₀ (uM)	IC ₅₀ +/-
<p>summary details</p>  <p>10 from Table 1 Barry Bunin's Site SMILES InChI</p>	2.51	 <p>Mark outliers</p>	0.12	+/- 0.02
			0.12	

Heat map for CCK A receptor binding...

Heat Map

Readout Definition:

To view the entire plate, please scroll the table from left to right (scrollbar at bottom).

	1	2	3	4	5	6	7	8	9	10	11	12
A	30.0 z: -0.078	8.9 z: -0.744	23.0 z: -0.299	4.0 z: -0.899	0.2 z: -1.019	78.0 z: 1.438	50.0 z: 0.554	37.0 z: 0.143	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
B	81.0 z: 1.533	65.0 z: 1.027	58.0 z: 0.806	89.0 z: 1.785	79.0 z: 1.469	81.0 z: 1.533	83.0 z: 1.596	67.0 z: 1.091	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
C	53.0 z: 0.649	4.0 z: -0.899	5.4 z: -0.854	8.0 z: -0.772	1.0 z: -0.993	6.0 z: -0.835	7.0 z: -0.804	1.2 z: -0.987	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
D	10.0 z: -0.709	4.0 z: -0.899	2.7 z: -0.940	0.61 z: -1.006	4.3 z: -0.889	6.7 z: -0.813	2.1 z: -0.959	1.9 z: -0.965	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
E	34.0 z: 0.049	20.0 z: -0.393	6.0 z: -0.835	11.0 z: -0.678	0.12 z: -1.021	8.2 z: -0.766	60.0 z: 0.870	2.2 z: -0.955	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
F	94.0 z: 1.943	84.0 z: 1.627	77.0 z: 1.406	61.0 z: 0.901	59.0 z: 0.838	71.0 z: 1.217	91.0 z: 1.848	68.0 z: 1.122	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
G	24.0 z: -0.267	58.0 z: 0.806	11.0 z: -0.678	67.0 z: 1.091	29.0 z: -0.109	44.0 z: 0.364	66.0 z: 1.059	37.0 z: 0.143	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025
H	2.2 z: -0.955	5.8 z: -0.842	1.4 z: -0.981	8.9 z: -0.744	3.6 z: -0.911	4.2 z: -0.892	9.3 z: -0.731	7.5 z: -0.788	100.0 z: 2.133	98.0 z: 2.069	0.01 z: -1.025	0.01 z: -1.025

Z'-factor: 0.969 (controls)

Z-factor: -1.928

Legend

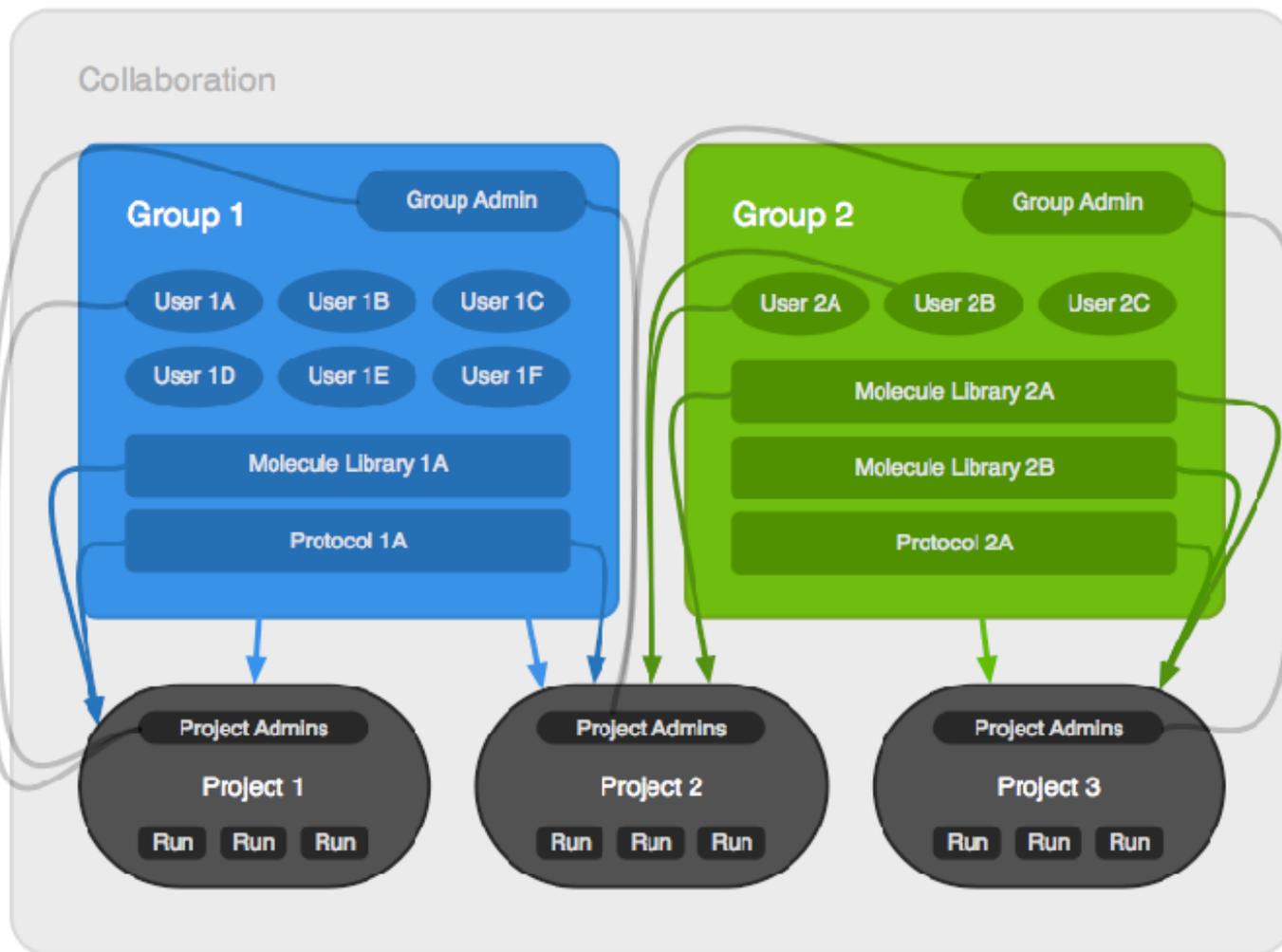
- Positive control
- Negative control



COLLABORATIVE
DRUG
DISCOVERY

What if?

- Technology
- Collaborations





COLLABORATIVE
DRUG
DISCOVERY

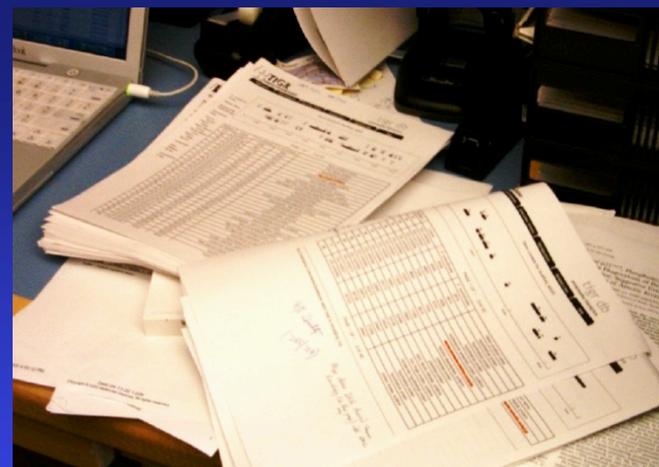
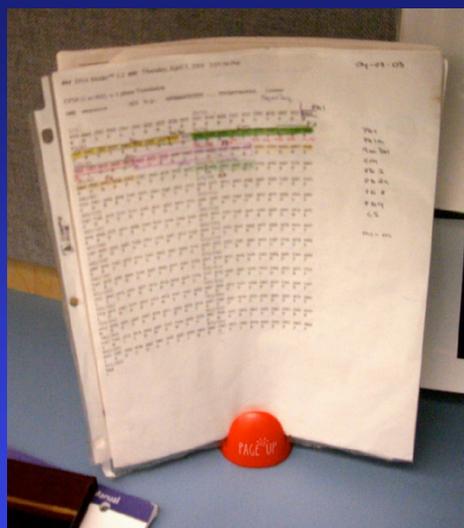
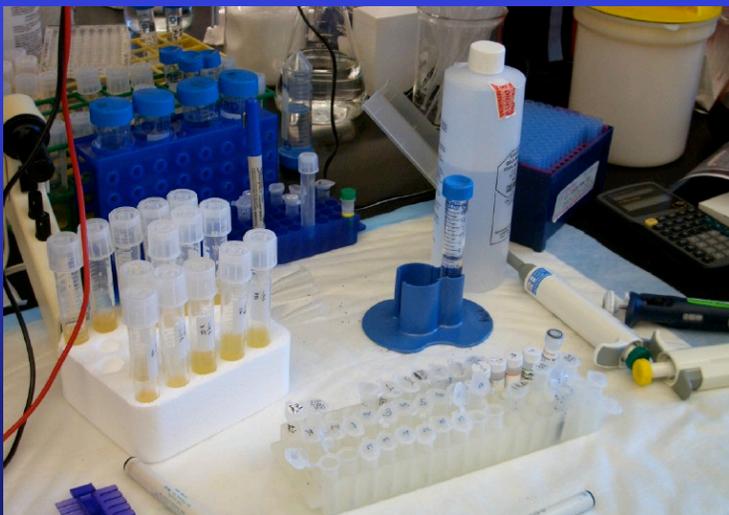
Who We Are

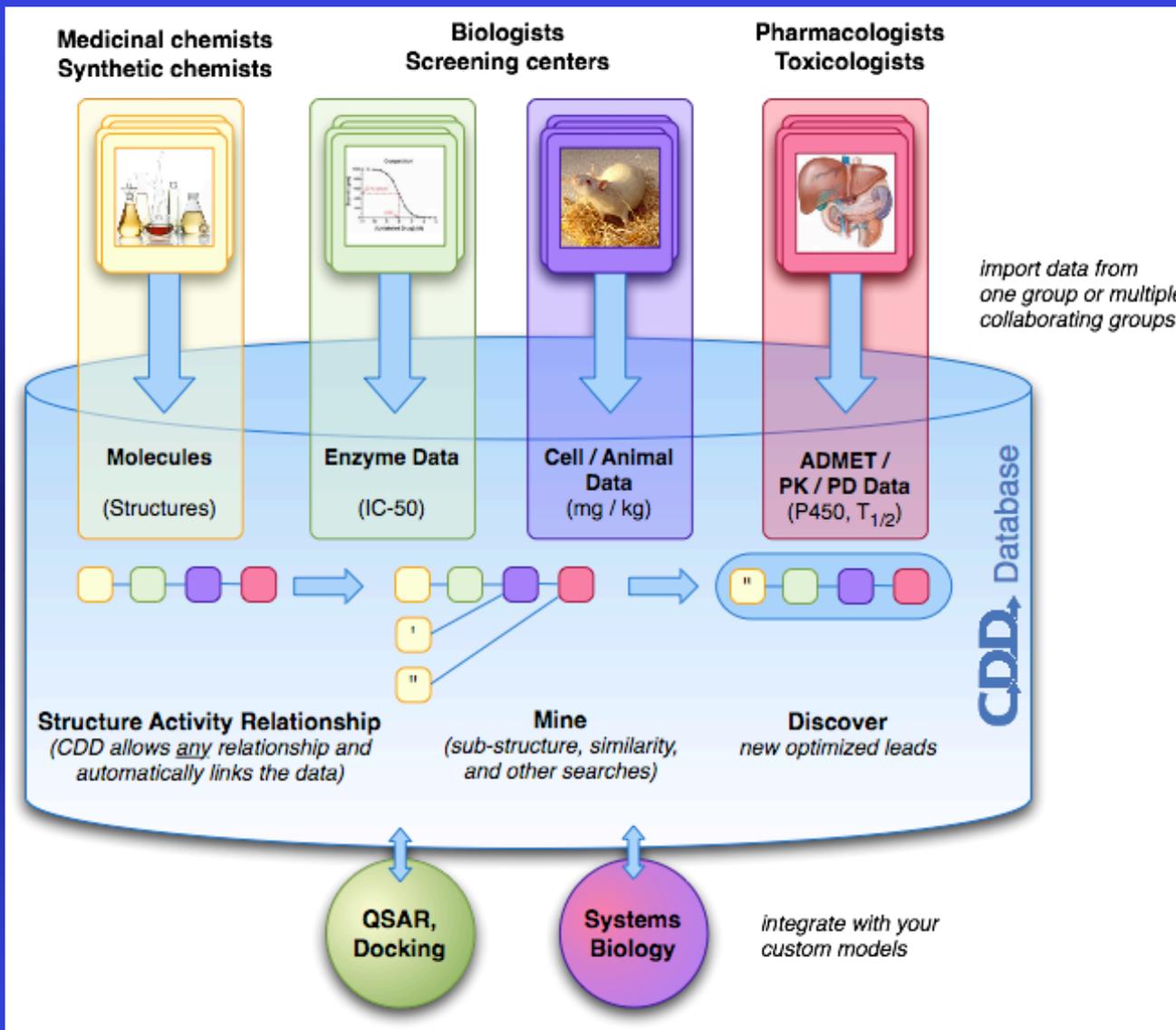
- **A private company, founded in 2004.**
- **Product: CDD database and custom services**
- **Working with thousands of leading researchers using CDD on six continents**
- **Provides secure “cloud-based” collaborative drug discovery embracing academic, non-profit, and for-profit company researchers (cultures) in invitation-only groups**
 - Affordable, easy-to-use, yet extremely powerful
- **SAB:**
 - Christopher Lipinski, PhD (ex Pfizer)
 - James McKerrow, MD PhD
 - David Roos, PhD
 - Adam Renslo, PhD
 - Wes Van Voorhis, MD PhD
 - Jim Wikel (ex Eli Lilly)



COLLABORATIVE
DRUG
DISCOVERY

Typical Lab: The Data Explosion Problem



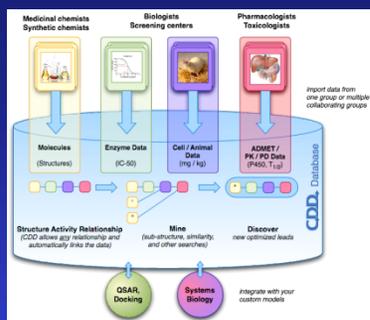




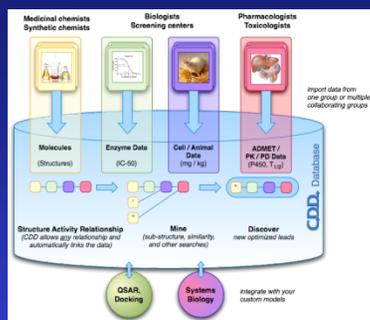
COLLABORATIVE
DRUG
DISCOVERY

Collaborate Mode 1 – Big Pharma with single login to multiple CROs/Biotech groups

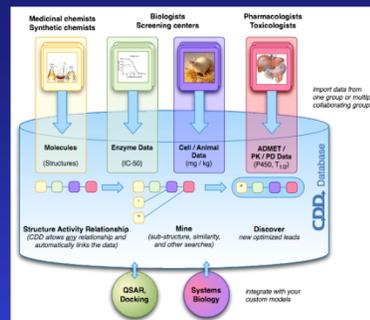
Consolidated
CDD DB



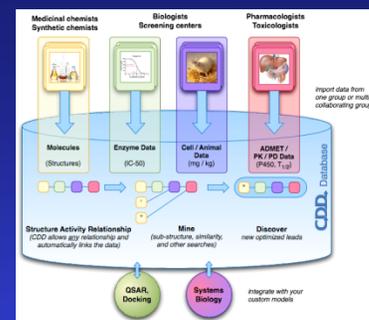
Collaborator 1



Collaborator 2



Collaborator 3



Collaborator 4



COLLABORATIVE
DRUG
DISCOVERY

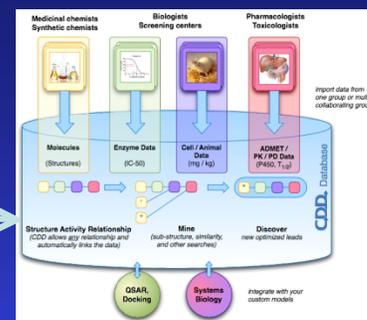
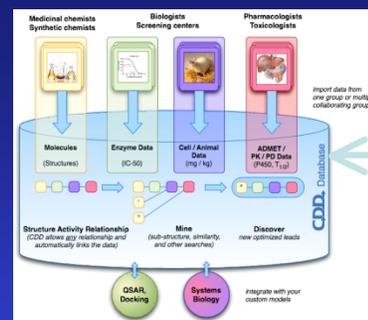
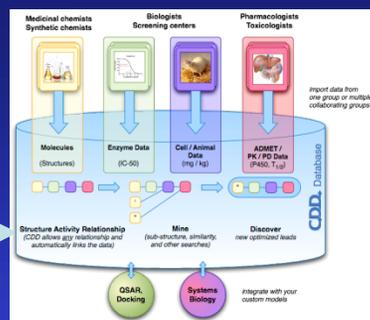
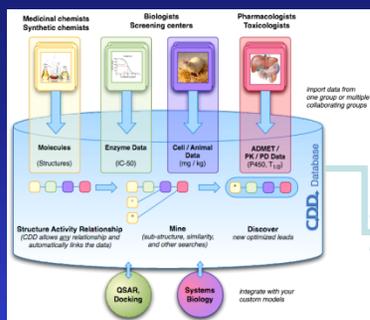
Collaborate Mode 2 – Academic + Foundations securely sharing data subsets

CDD DB
group 1

CDD DB
group 2

CDD DB
group 3

CDD DB
group 4





CDD database system

- Web based (log in securely into to your database from any computer).
- We host the server for you
- Highly secure, all traffic encrypted, server in a secure professionally hosted environment.



COLLABORATIVE
DRUG
DISCOVERY

Back pocket



COLLABORATIVE
DRUG
DISCOVERY

Implementing Toxic Functionality Alerts (see Lipinski's presentation)

- **Basic Capability = 2 FTE weeks**
- **End User Flexibility = 1 FTE week**
- **Advanced Implementation = 3 FTE weeks**
- **NOTE: Generally any feature can have about 25% of unforeseen features**



Implementing Toxic Functionality Alerts (see Lipinski's presentation)

- **Basic Capability = 2 FTE weeks**
 - Store list of SMARTS alerts with the SMARTS, a name, and a text description (back end) - 1 point
 - Associate list of alerts that apply with structures - 1 point
 - Set up alerts to be automatically run when adding a new structure - 1 point
 - Run alerts against existing structures - 1 point
 - Ability to display alerts on the show molecule page - 2 points
 - **End User Flexibility = 1 FTE week**
 - Allow super-users to manage this list - 2 points
 - Rerun alerts after they're changed against all structures - 1 point
 - **Advanced Implementation = 3 FTE weeks**
 - Ability to display alerts in mine results - 3 points
 - Ability to mine by presence/absence of alerts in general - 2 points
 - Ability to mine by presence/absence of specific alerts - 3 points
- NOTE: Generally any feature can have about 25% of unforeseen features

Edit Group Settings

Group name: Bunin Group

Primary contact: Barry Bunin

Principal investigator:

Principal investigator email:

Institution:

Website URL:

External link setting: Ban external links Allow external links Allow external links and do not show a warning page

Experimental subject name: Substance

Experiment name: Screen

Run name: Test

Save changes or cancel

- Published Review in Drug Discovery Today 2009 (14: 261-270).

Drug Discovery Today • Volume 14, Number 02 • February 2009

REVIEWS



Novel web-based tools combining chemistry informatics, biology and social networks for drug discovery

Moses Hohman¹, Kellan Gregory¹, Kelly Chibale², Peter J. Smith³, Sean Ekins^{4,5,6} and Barry Bunin¹

¹ Collaborative Drug Discovery, Inc., 1818 Gilbreth Road, Suite 220, Burlingame, CA 94010, USA
² Institute of Infectious Disease and Molecular Medicine and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa
³ Division of Pharmacology, Department of Medicine, University of Cape Town, Medical School, K48, OMS, Grote Schuur Hospital, Observatory, 7925, South Africa
⁴ Collaborations in Chemistry, Jerkintown, PA 19046, USA
⁵ University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway NJ 08854, USA
⁶ Department of Pharmaceutical Sciences, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA

A convergence of different commercial and publicly accessible chemical informatics, databases and social networking tools is positioned to change the way that research collaborations are initiated, maintained and expanded, particularly in the realm of neglected diseases. A community-based platform that combines traditional drug discovery informatics with Web2.0 features in secure groups is believed to be the key to facilitating richer, instantaneous collaborations involving sensitive drug discovery data and intellectual property. Heterogeneous chemical and biological data from low-throughput or high-throughput experiments are archived, mined and then selectively shared either just securely between specifically designated colleagues or openly on the Internet in standardized formats. We will illustrate several case studies for anti-malarial research enabled by this platform, which we suggest could be easily expanded more broadly for pharmaceutical research in general.

The networked revolution

Recent research suggests that open collaborative drug discovery will be the future paradigm of biomedical research [1–3]. Reviews in this journal have provided a perspective on the many publicly accessible, open access chemistry databases and Internet-based collaborative tools [4,5] that are likely to enhance scientific research in future. Some of these public databases are already being used for structure activity relationship (SAR) development [6] and rapid lead identification [7]. It takes a combination of biology and chemistry insight, however, to translate molecules into potential drugs and there has been little, if any, discussion of how collaborations between chemists and biologists are to be facilitated [8]. The challenges associated with bringing chemists and biologists together for virtual drug discovery projects for neglected diseases [8] provide an arena for testing new approaches that can perhaps be expanded more broadly to commercial drug discovery projects. The biological data available for sharing are frequently stored in single document or Excel™ files. Compilation of data is sporadic with no depth and little, if any, standardization of the data formats or crucial information such as experimental procedures and statistical analysis to quantify data quality to allow reproducibility and comparisons between groups. Before collaborations begin, data security and integrity should always be considered while intellectual property arrangements (Material Transfer and Intellectual Property (IP) Rights Agreements) are often (at least in academia) seen as necessary, but generally as a hindrance to progress. As a collaboration progresses the needs of data users may change, so it is important to have flexibility in the use of systems for tracking or storage of data and between systems [8].

Any tool that can tap into a growing community of researchers becomes more valuable as a function of Metcalfe's law, which simply states the value of a network is equal to the square of the

Corresponding author: Bunin, B. (bbunin@collaborativedrug.com)

1874-0066/09/\$ - see front matter © 2009 Elsevier Ltd. All rights reserved. doi:10.1016/j.drudis.2009.01.006

www.drugdiscoverytoday.com 1



COLLABORATIVE
DRUG
DISCOVERY

CDD Successfully Builds Collaborative Drug Discovery Communities

- Published Review in Drug Discovery Today 2009 (14: 261-270).
- \$1.89M Tuberculosis B&MGF Project
- Thousands of scientists engaged:

Drug Discovery Today • Volume 14, Number 02 • February 2009

REVIEWS



Novel web-based tools combining chemistry informatics, biology and social networks for drug discovery

Moses Hohman¹, Kellan Gregory¹, Kelly Chibale², Peter J. Smith³, Sean Ekins^{4,5,6} and Barry Bunin¹

[Register for Public Access](#) | [Login](#)



COLLABORATIVE
DRUG
DISCOVERY

[Home](#)

[Products](#)

[Services](#)

[Benefits](#)

[Community](#)

[Blog](#)

[Events](#)

[Company](#)

[Home](#) /

Organizations that benefit from CDD

- ASINEX
- BioSeek, Inc.
- Broad Institute
- Cedars-Sinai Medical Center
- ChemBridge Corporation
- Columbia University
- Cornell University
- Drugs for Neglected Diseases initiative
- Fred Hutchinson Cancer Research Center
- Harvard University
- Indiana University
- Indiana University Purdue University
- Purdue University
- The Rockefeller University
- San Francisco VA Medical Center
- Scripps Research Institute
- Seattle Biomedical Research Institute
- Semafore Pharmaceuticals, Inc.
- St. Jude Children's Research Hospital
- Stanford University
- STATegics, Inc.
- TimTec, Inc.
- Torrey Pines Institute for Molecular Studies
- UCSF General Hospital

Who Uses CDD?

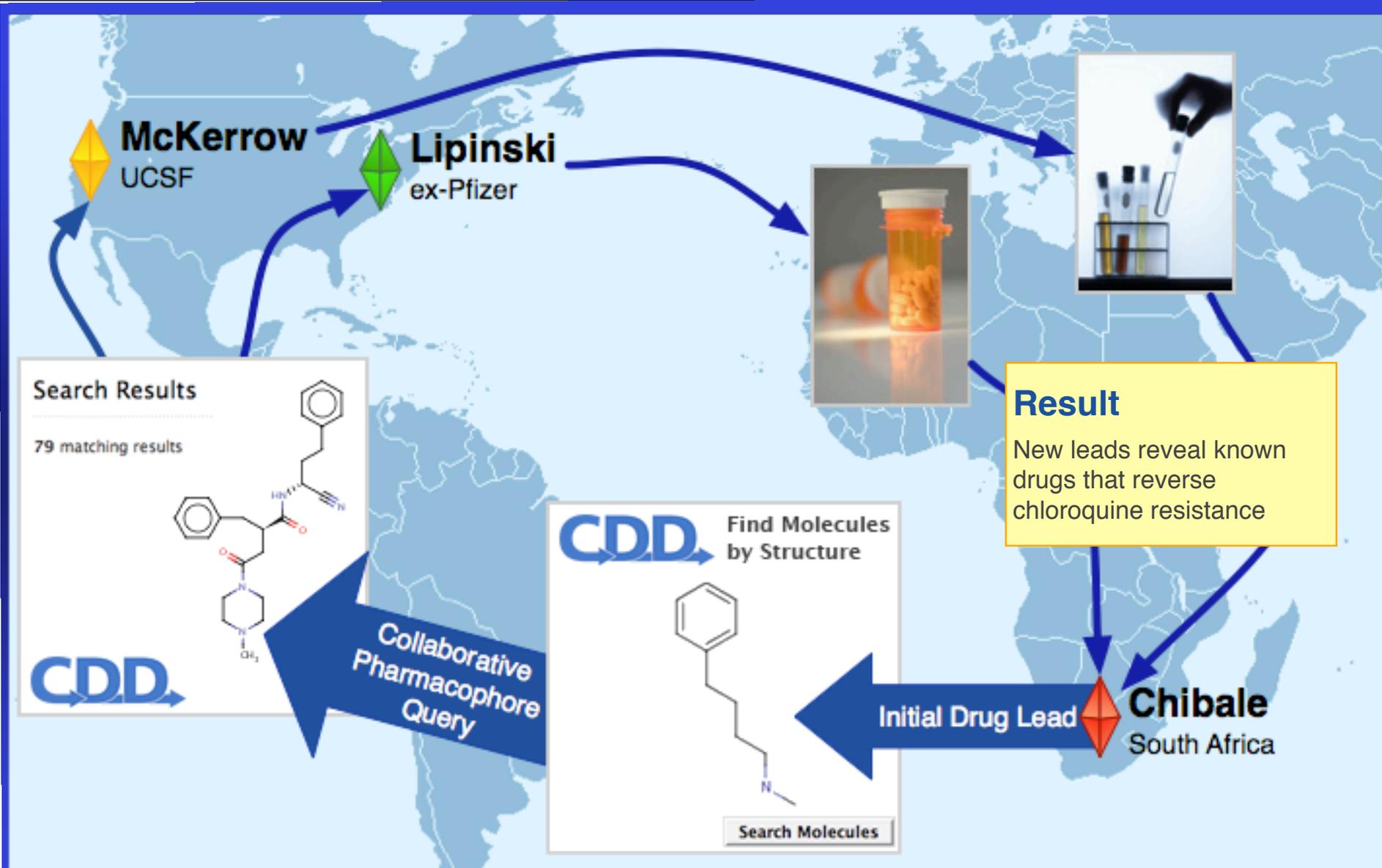
[Testimonials](#)

"The CDD Database is an extremely elegant platform. I highly recommend it for anyone generating drug discovery data."

— Bryan Roth, M.D., Ph.D., University of



- **Archive**
 - Organize and upload your experimental data
- **Mine**
 - Analyze your data to suggest new drug candidates
- **Collaborate**
 - Keep your data 100% private,
 - Exchange confidentially with collaborators, *or...*
 - Share openly within the CDD community





COLLABORATIVE
DRUG
DISCOVERY

CDD Database system

- CDD works right “out of the box”, no implementation hassle. Login & go.
- User can access (with Firefox browser) from Windows, Mac, even Linux/Unix
- A wide variety of data formats are supported including Chemical structure formats like SD format or SMILES. Files (images, pdf etc. can be associated to multiple entities)
- Easy to use - no special knowledge required; navigation with familiar web-browser tools like hyperlinks, tabs, dropdown menus etc.
- Different user privileges can be set according to your user structure.
- Supports Collaborations



CDD Database system – Distributed Research

- Ability to tap into hundreds of external researchers data with no lengthy business/legal cycles
 - Save 6-12 months on a new discovery
- No need to “poke a hole” into your Firewall
- No need for users to have any IT expertise
- 3rd party for hosting – liability relief



- You decide which data to share with whom
(we currently manage this on our server via group management)
- Store and share protocols
- Via a unique identifier (e.g. a structure, a name, a unique label) teams from diverse disciplines as Chemistry, Toxicology, Assay design, animal studies etc. can share their data in real time.
- Collaborate across different organizations. Our technology and pricing enables and supports a variety of collaboration models.
- Publication mechanism: Specialized community data collections are available for no charge. Current collections – see next slide



- Easy batch import of your data is supported by a mapper and slurper.
- Add molecules, analytical, or biological data manually, all required tools are integrated.
- Export your data with Chemical structures into Excel without the need for extra plugins or extra software.
- Chemical drawing tool integrated which is powerful and very easy to use.
- Plates, heatmaps, property calculations, 3D views and much more (similarity search by end of Jan/08)



COLLABORATIVE
DRUG
DISCOVERY

CDD “Agile” Development Advantage

- Traditional software development life cycle (6-12 months)
- CDD Agile development includes automated testing:
 - Biweekly minor updates
 - Monthly major updates
- Recent examples:
 - Logins to multiple groups for screening centers
 - Dashboard for intra-group coordination
 - Customize mine results



CDD COLLABORATIVE DRUG DISCOVERY
McKerrow Group

Files Tools

Assays

Assays 21-40 of 40 ◀ 1 2 ▶

Name	Type	# Compounds
S. mansoni CB2 Ki (nM) (GSK 2000)	enzyme	23
S. mansoni (elastase) IC50 (uM)	enzyme	0
S. mansoni (in vivo) (mouse)	Animal	1
S. mansoni legumain (Sm AE) IC50 (uM)	enzyme	0

Molecules /

9 from Table 1 (Merck Compound)

Owner: Barry Bunin
Created: 7/30/2007
Updated: 1/9/2008

Find other molecules containing this structure

Delete this molecule

Definition | [Edit definition](#) | [Add synonym](#) | [Add user defined field](#)

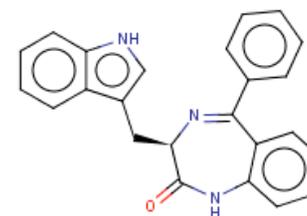
Name 9 from Table 1 (Merck Compound)

Synonyms AA = R-Trp, RX = H

Description Note - literature compound used as Internal Standard.

SMILES O=C1Nc2ccccc2C(=N[C@@H]1Cc3c[nH]c4ccccc34)c5ccccc5

Solubility 10 uM



Batches: 1 | [Add a batch](#)

Name	Date	Location	Who	Notes	#Protocols	Actions
1	2004-02-14	-20 Freezer	Barry	>99% Pure by HPLC 214 nM	0	

Plates: 1

Name	Size	Protocol Run
BAB-003	96	<ul style="list-style-type: none"> CCK A receptor binding affinities

Molecular weight 365.427 g/mol

log P 5.0459

H-bond donors 2

H-bond acceptors 2

Lipinski Rule of 5 One violation
3 of 4 within desirable range

Formula C₂₄H₁₉N₃O

pK_a 3.2221

Exact mass 365.153 g/mol

Atom count 47

Mine Protocols and Molecules

Find Molecules by Protocol

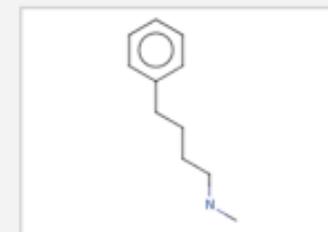
(any protocol type) (any protocol)

Add a term

Delete term

Keywords

Structure

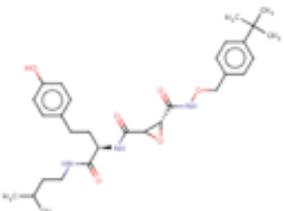


Save this search as: Save

Show first 100 results Search Molecules

Search Results

First 100 of 230 matching results

Molecule	Cruzain (rec) IC50 (uM)	T. brucei rhodesiense Rhodesain (rec) IC50 (uM)
 WRR-476	0.028	< 0.01
	< 0.01	< 0.01

“Five years ago our lab tested 20 compounds a year with no way to handle 200. Today, with CDD’s software, we can fully process 3000 new compounds per year and advance the best compounds to late-stage development.”

— Dr. James McKerrow, UCSF



- Low total cost of ownership: You do not need to buy and maintain your own server, run updates, have an IT specialist; minimal time for training; reasonable license fees. Lots of tools integrated which otherwise need to be purchased extra (structure editor, export with structures into regular Excel etc.).
- Superior customer support (support schedule including regular trainings and pings; collaboration support)
- Chemaxon tools integrated.
- New features added regularly and do not need to wait for a maintenance schedule.



- **"Five years ago our lab tested 20 compounds a year with no way to handle 200. Today, with CDD's software, we can fully process 3,000 new compounds per year and advance the best compounds to late-stage development."**
 - James McKerrow, MD, PhD, UCSF
- **"One of the biggest barriers for academic drug discovery is the poor access to chemical data represented in an intelligent format. CDD presents data and associated tools that capture the relationship between chemical structure and biological activity. Structure-Activity Relationship (SAR) data substantially improve the distributed drug discovery process."**
 - Christopher Lipinski, PhD, Pfizer, Retired
- **"It's a great tool for mining and sharing data with collaborators. I've used a couple of other programs, but they don't put everything together so intuitively."**
 - Mary Lynn Baniecki, PhD, Clardy Laboratory, Harvard Medical School
- **"I was very impressed by the system's simplicity, responsiveness and functionality. It certainly looks better than anything I've seen so far."**
 - Wayne Best, PhD
- **"The CDD Database is an extremely elegant platform. I highly recommend it for anyone generating drug discovery data."**
 - Bryan Roth, MD, PhD, University of North Carolina

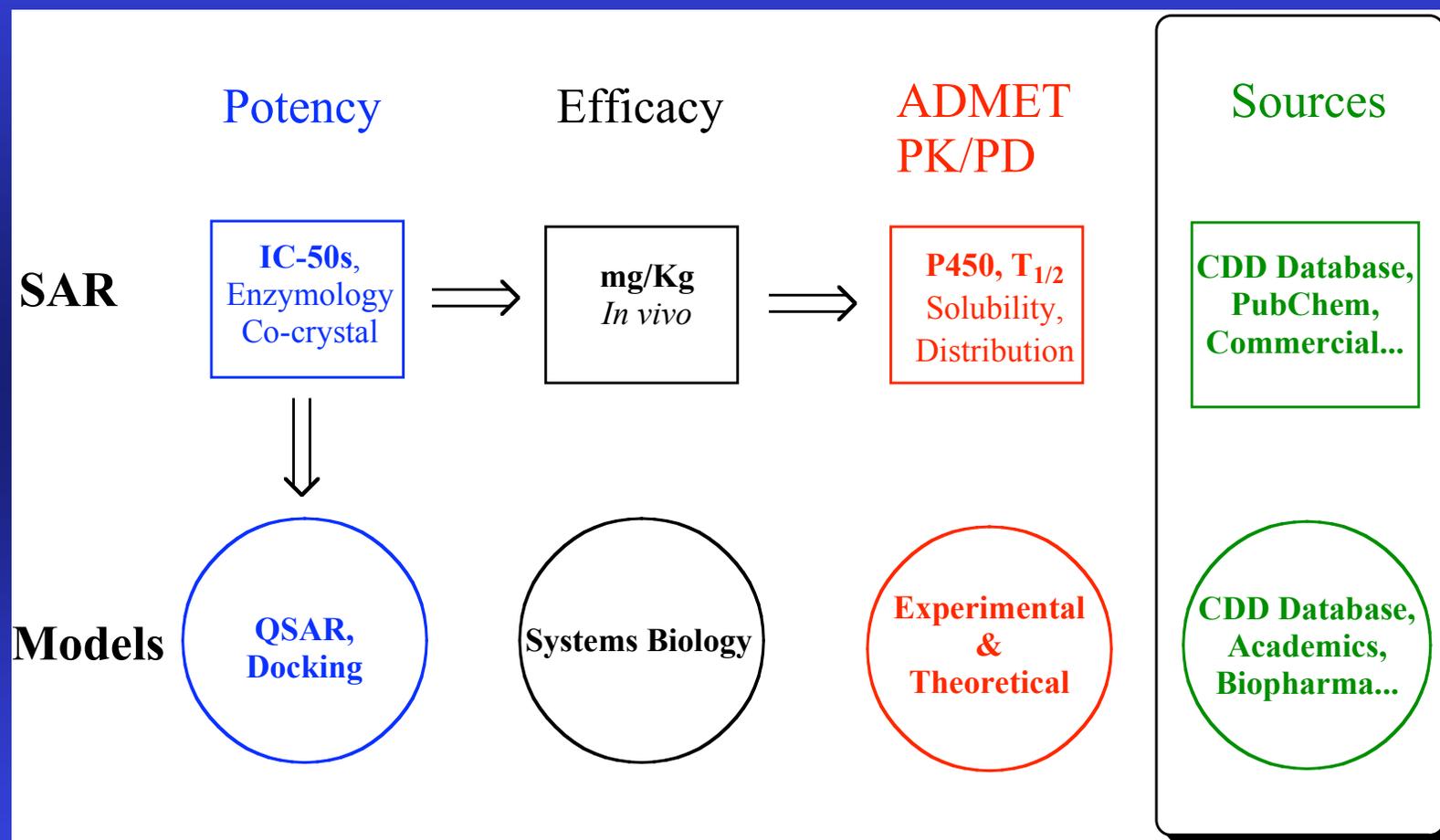


COLLABORATIVE
DRUG
DISCOVERY

CDD Current Community & Strategic Relationships

- ASINEX, Inc.
 - Broad Institute
 - Cedars-Sinai Medical Center
 - Columbia University
 - Cornell University
 - Drugs for Neglected Diseases initiative
 - Fred Hutchinson CRC
 - Gates Foundation (BMGF)
 - Harvard University
 - Indiana University–Purdue University Indianapolis
 - Johns Hopkins
 - Louisiana State University
 - Marine Biological Laboratory
 - MIT
 - Myelin Repair Foundation
 - Purdue University
 - San Francisco VA Medical Center
 - Seattle Biomedical Research Institute
 - Semafore Pharmaceuticals, Inc.
 - Stanford University
 - St. Jude Children’s Research Hospital
 - SureChem
 - TimTec, Inc.
 - UCSF General Hospital
 - University of California, Berkeley
 - University of California, Los Angeles
 - University of California, San Francisco
 - University of California, Santa Cruz
 - University of Cape Town
 - University of Mississippi
 - University of North Carolina
 - University of Pennsylvania
 - University of Sydney
 - University of Texas
 - University of Washington
- All at: www.collaboratedrug.com/who

Archive Preclinical SAR (with any Relationship)





Edit Molecule - K11777 - Collaborative Drug Discovery - Windows Internet Explorer

http://qa.collaborativedrug.com/molecules/914892;edit

File Edit View Favorites Tools Help

Google G Go 8 blocked Check AutoLink AutoFill Send to Settings

Edit Molecule - K11777 - Collaborative Drug Discovery Edit Home Feeds (1) Print Research Page Tools

Welcome, **Joey**: Logout | Change password

Archive Mine Collaborate

Assays | **Molecules** | Plates

Edit Molecule

OK Cancel delete molecule and its properties and readouts

Name: K11777

Synonym: McKCel-4;CRA-003316;K777;McKmisl-016;KH
Please separate synonyms with a semicolon.

Description: HCl salt, Originally from Khrepi (1994), large batches from Seres (2004), CRA # is a part of 2200 compound library from Celera;
Insert: α β γ μ Δ \circ show/hide more

User Defined Fields: Add New

SMILES: CN1CCN(CC1)C(=O)N[C@@H](Cc2ccccc2)C(=O)N[C@@H](CCc3ccccc3)C=C\S(=O)(=O)C

File Edit View Insert Tools Help

H C N O F React Select Erase Paste Undo Redo Zoom

- + P S Cl \rightarrow \leftrightarrow \otimes \otimes \leftarrow \rightarrow \otimes \otimes

More Br I \leftarrow \rightarrow \otimes \otimes \otimes \otimes \otimes \otimes

Done

Internet 100%

MarvinView

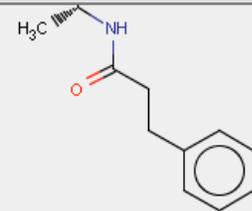
File Edit View Tools Help

search -

File Edit View Favorites Tools Help

Address <http://www.collaborativedrug.com/cdd/page/search.jsp#resultAnchor>

Google wsgr Search 14 blocked Check AutoLink AutoFill Options wsgr



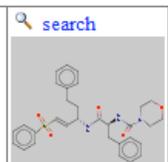
Search Molecules

Display Option:
Number of search results per page 20

Clear Result Export to Excel Export to SDFFile Save Query as: >>

Results: 1 molecules containing C[C@H](CCc1ccccc1)NC(=O)CCc2ccccc2
having T. cruzi: In vitro assay run with readout Days Survival > 5 (days)
and having Cruzain (rec) IC50 (uM) run with any readouts

table rosetta show structures

Molecule	Structure	Cruzain (rec) IC50 (uM) DU4-71 (%)	Cruzain (rec) IC50 (uM) IC50 (nM)	Cruzain (rec) IC50 (uM) IC50 (uM)	Cruzain (rec) IC50 (uM) Ki (uM)	Cruzain (rec) IC50 (uM) kobs/I (see excel) (M)	T. cruzi: In vitro assay Days Survival (days)
K11002				0.01			8.0--25.0

Internet



COLLABORATIVE
DRUG
DISCOVERY

CDD – Community Open Access Data

1. **Dr. Christopher Lipinski**

Chemoinformatics data on FDA and Orphan approved drugs

2. **Prof. Roos (U. Penn)**

Modern Malaria literature data linked to genes

3. **Prof. Gelb (UW)**

Modern Malaria literature data linked to assays

4. **Prof. Guy (St. Jude CRH)**

>15,000 Army Malaria screening data dating back to WW-II after escrow period

5. **Prof. McKerrow (UCSF)**

“Open-Content” Drug discovery – T. Brucei, T. Cruzi, S. Mansoni, Leishmania, P. falciparum



Assays

Assays 21-40 of 40

◀ 1 2 ▶

Name	Type	# Compounds
S. mansoni CB2 Ki (nM) (GSK 2000)	enzyme	23
S. mansoni (elastase) IC50 (uM)	enzyme	0
S. mansoni (in vivo) (mouse)	Animal	1
S. mansoni legumain (Sm AE) IC50 (uM)	enzyme	0
T. brucei brucei extracts IC50 (uM)	enzyme	94
T. brucei Cathepsin B (rec) IC50 (uM)	enzyme	137
T. brucei Cathepsin B (rec) % Inhibition (1 uM) HTS	enzyme	1342
T. brucei (in vitro)	cell	74
T. brucei (in vitro) % Growth-Inhibition HTS (1uM)	cell	1159
T. brucei (in vitro) parasitemia	cell	34

[Add a new assay](#)

Search assays

Search by keyword within assay names, types and descriptions.

K11777 - Collaborative Drug Discovery - Windows Internet Explorer

http://qa.collaborativedrug.com/molecules/914892

File Edit View Favorites Tools Help

Google Go Bookmarks 8 blocked Check AutoLink AutoFill Send to Settings

K11777 - Collaborative Drug Discovery Edit Home Feeds (1) Print Research Page Tools >>

Name K11777

Synonyms McKCel-4;CRA-003316;K777;McKmisI-016;KH11777

Description HCl salt, Originally from Khrepi (1994), large batches from Seres (2004), CRA # is a part of 2200 compound library from Celera; Delivered=Celera 2006

SMILES CN1CCN(CC1)C(=O)N[C@@H](Cc2ccccc2)C(=O)N[C@@H](CCc3ccccc3)C=C\S(=O)(=O)c4ccccc4

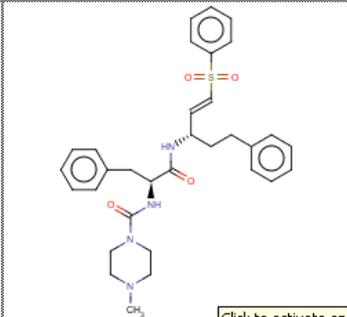
Batches: 3

Name	Date	Location	Who	Notes	#Assays
	2006-05-24				8
Batch-01					2
BATCH-1	2006-05-25				1

Plates: 0

Assays: 12

Assay Name	Run Date	Readout	Value
Cruzain (rec) IC50 (uM)	2005-02-17	IC50 (uM)	0.1
T. cruzi: In vitro assay	2005-04-04	Days survival minus control (days)	



Click to activate and use this control

Double-click image above to see alternate views



COLLABORATIVE
DRUG
DISCOVERY

CDD database system: Security overview

- *CDD encrypts all traffic between a user's web browser and our servers once the user has logged in.*
- *CDD does thorough automated testing of our application code, including security tests to make sure one group cannot access another group's private data.*
- *CDD uses two levels of firewalling ("defense in depth"), hardware and software, for each sensitive server in our server environment.*
- *CDD encrypts all database and customer file backups using public key cryptography, and the private keys are backed up securely.*
- *CDD stores user passwords in our database using strong, one-way encryption hashes.*
- *CDD stays current with all security updates to our system software, firmware and operating systems.*
- *CDD uses Linux systems in production.*
- *A number of application security enhancements around password security are being made mandatory.*
- *Access to production servers and superuser accounts is limited and audited within CDD.*