



GILEAD

Advancing Therapeutics.
Improving Lives.

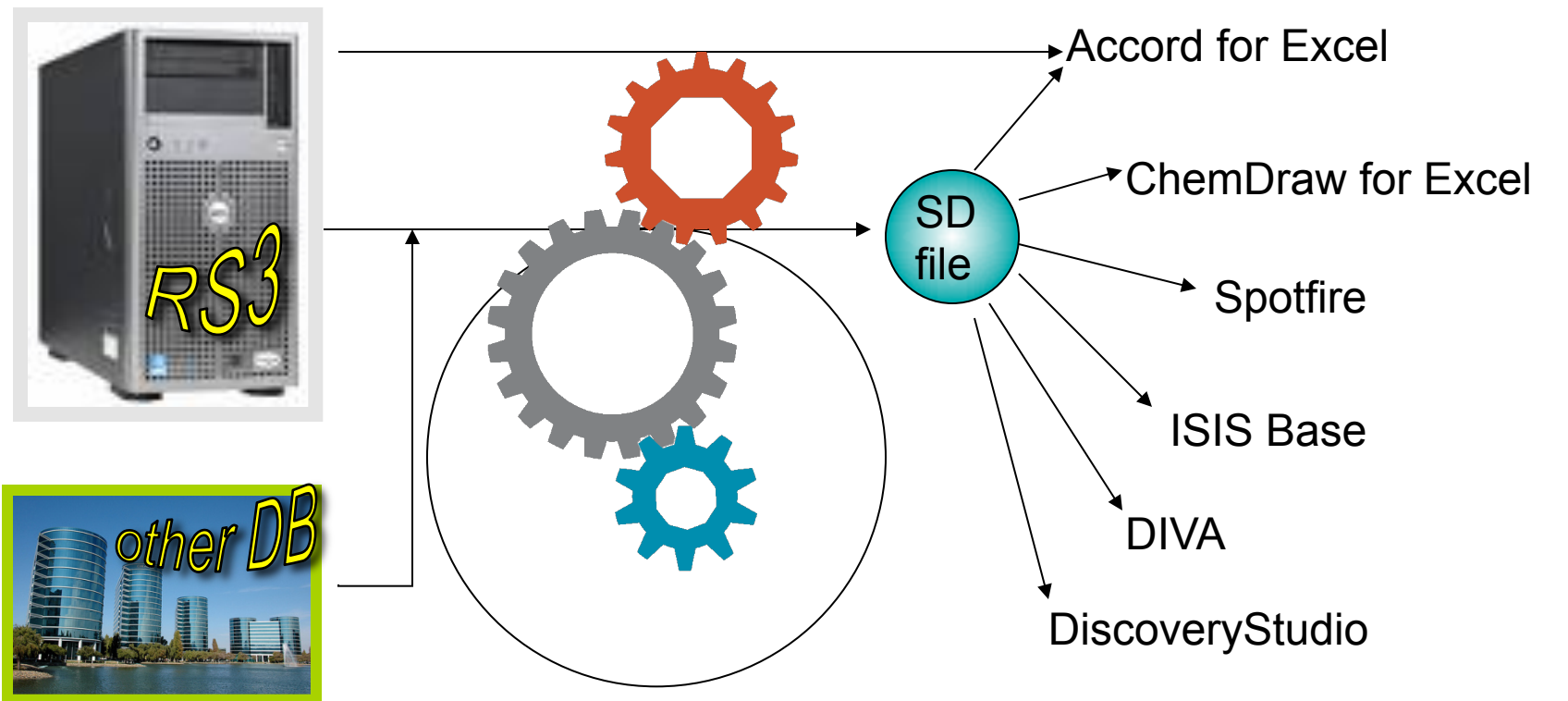


Presenting Eidogen/Sertanty Kinase Knowledge Base (KKB) via Dotmatics browser

Kerim Babaoglu

September 25th, 2012

How we used to look at data....

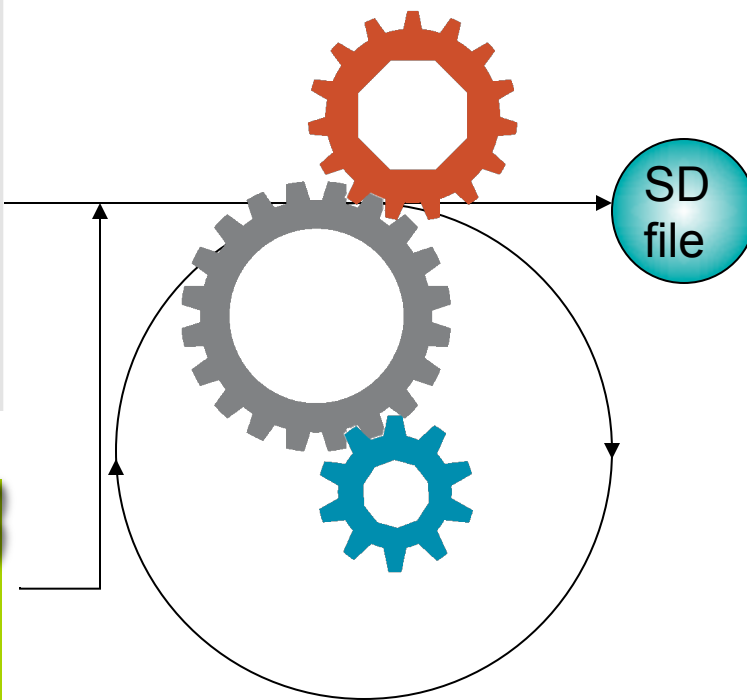


Processing
Annotation
With PipelinePilot

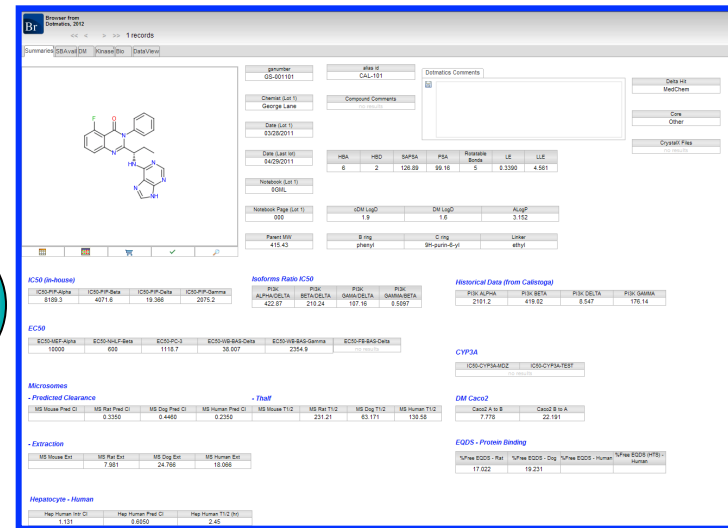
Examples:

Arithmetics (e.g. ratios, LE, LLE)
R-group analysis
Grouping by cores or substructures
Calculate fitness functions
(i.e. TPP match)

Dotmatics serves many needs



Processing
Annotation
With PipelinePilot



The screenshot displays a complex web application interface for chemical data analysis. It includes a chemical structure of a complex molecule, several data tables, and various search and filter options. The tables contain numerical data and chemical identifiers.

IC50 (h-Axoal)	IC50 (h-PP-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)
8159.3	4071.6	19.368	2075.2	

IC50	IC50 (h-PP-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)	IC50 (h-P-2a)
10000	605	1118.7	38.007	2354.9		

Microsomes	Predicted Clearance	Extraction	Hepaticity - Human
10000	0.3350	0.4485	0.2350

All-in-one data visualization

FDA Approved Protein Kinase Inhibitors (as of March 2012)

Table 1. FDA Approved Protein Kinase Inhibitors (as of March 2012)

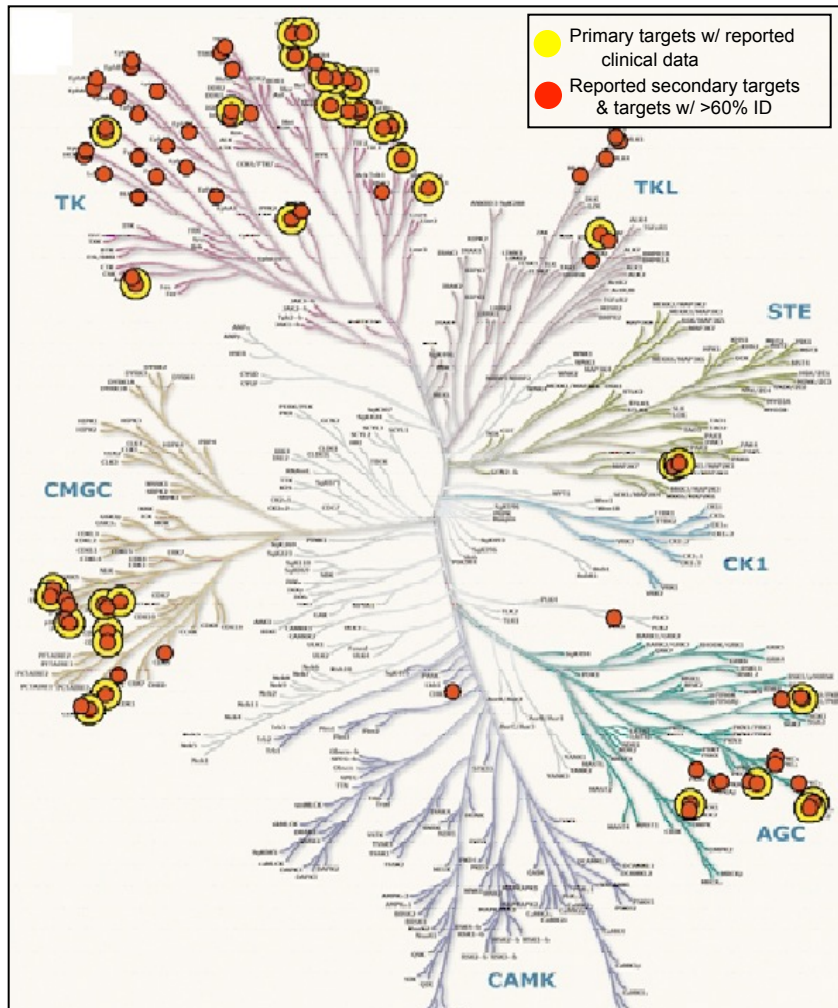
generic (brand) name	year of approval	company	indication	target kinase
imatinib (Gleevec)	2001	Novartis	chronic myeloid leukemia (CML)	Abl, c-Kit, PDGFR α/β
gefitinib (Iressa)	2003	AstraZeneca	non-small-cell lung carcinoma (NSCLC)	EGFR
erlotinib (Tarceva)	2004	Genetech, OSI	NSCLC, pancreatic cancer	EGFR
sorafenib (Nexavar)	2005	Bayer, Onyx	hepatocellular carcinoma, renal cell carcinoma (RCC)	Raf, VEGFR2/3, c-Kit, PDGFR β
sunitinib (Sutent)	2006	Pfizer	gastrointestinal stromal tumor (GIST), RCC	c-Kit, VEGFR, PDGFR, FLT3
dasatinib (Sprycel)	2006	Bristol-Myers Squibb	CML	Abl, c-Kit, PDGFR, Src
nilotinib (Tasigna)	2007	Novartis	CML	Abl, c-Kit, PDGFR, Src, ephrin
lapatinib (Tykerb)	2007	GlaxoSmithKline	breast cancer	EGFR, ErbB2
pazopanib (Votrient)	2009	GlaxoSmithKline	RCC	VEGFR, PDGFR α/β , c-Kit
vandetanib (Caprelsa)	2011	AstraZeneca	thyroid cancer	VEGFR, EGFR, RET
vemurafinib (Zelboraf)	2011	Roche, Plexxicon	CML	Abl, c-Kit, PDGFR, Src, ephrin
crizotinib (Xalkori)	2011	Pfizer	NSCLC (ALK +ve)	ALK, MET
ruxolitinib (Jakafi)	2011	Incyte	myelofibrosis	JAK1/2
axitinib (Inlyta)	2012	Pfizer	RCC	VEGFR, PDGFR β , c-Kit

Reference: [dx.doi.org/10.1021/jm3003203](https://doi.org/10.1021/jm3003203) | J. Med. Chem. 2012, 55, 6243–6262

Kinase SAR Knowledgebase (KKB) – Hot Targets

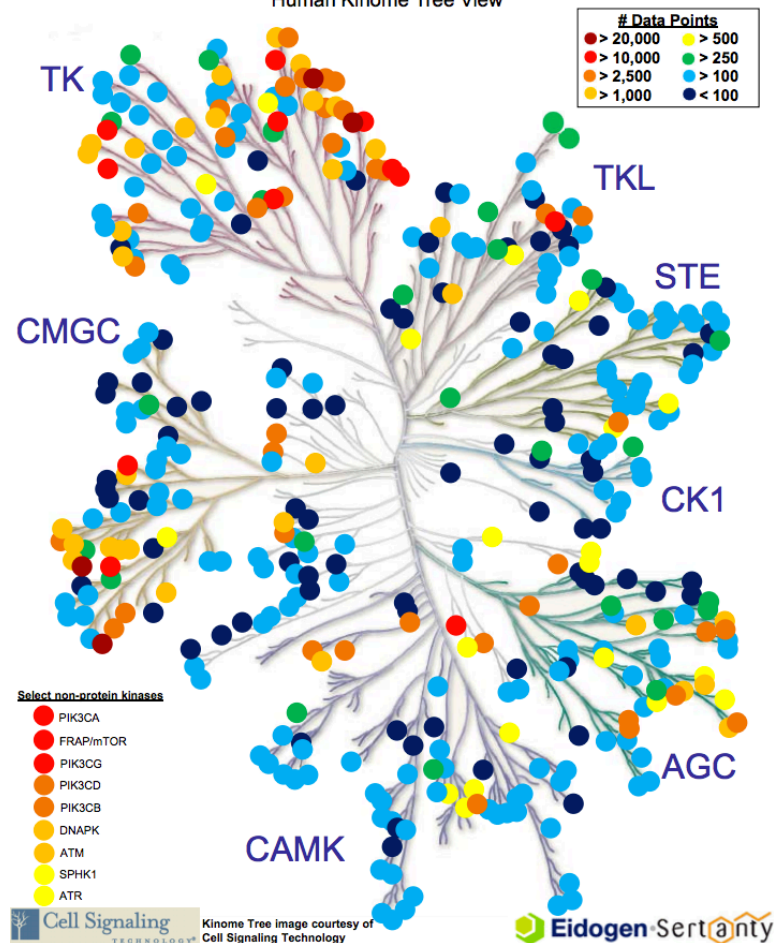
Kinase Targets of Clinical Interest

from Vieth *et al. Drug Disc. Today* **10**, 839 (2005).



Eidogen-Sertanty KKB SAR Data Point Distribution

Kinases with Biological Activity Data in Kinase Knowledgebase
 Q2 2012 Release
 Human Kinome Tree View

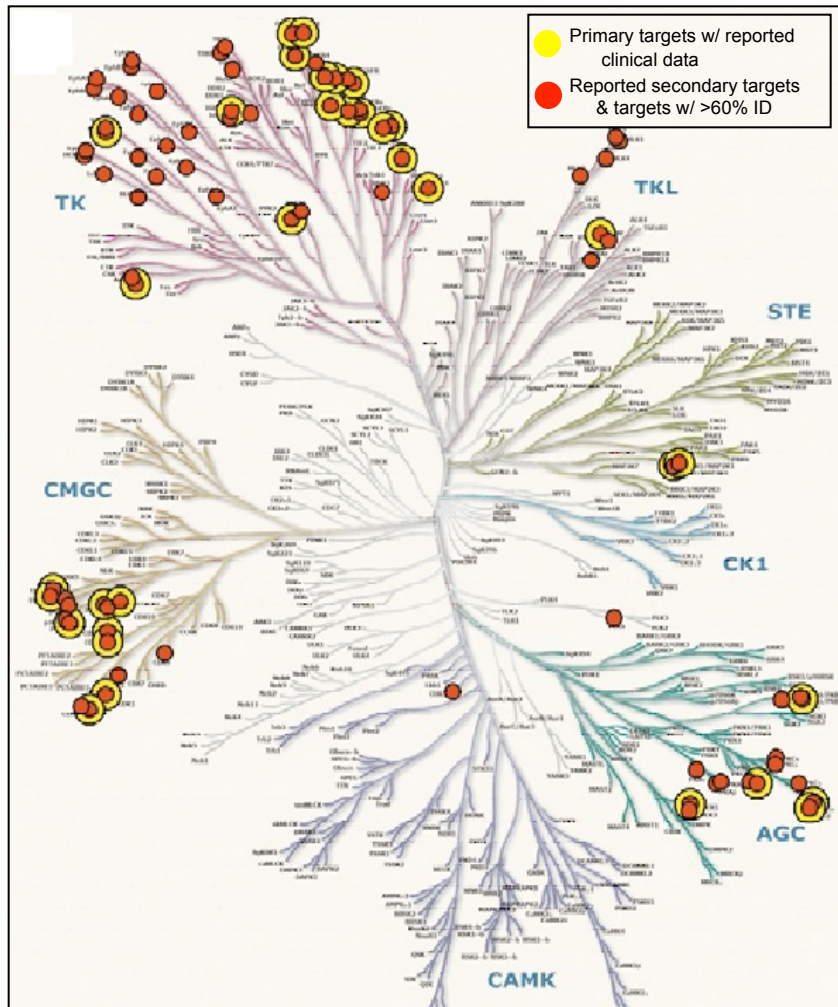


> 649,000 SAR data points curated from
 > 7915 journal articles and patents

Kinase SAR Knowledgebase (KKB) – Hot Targets

Kinase Targets of Clinical Interest

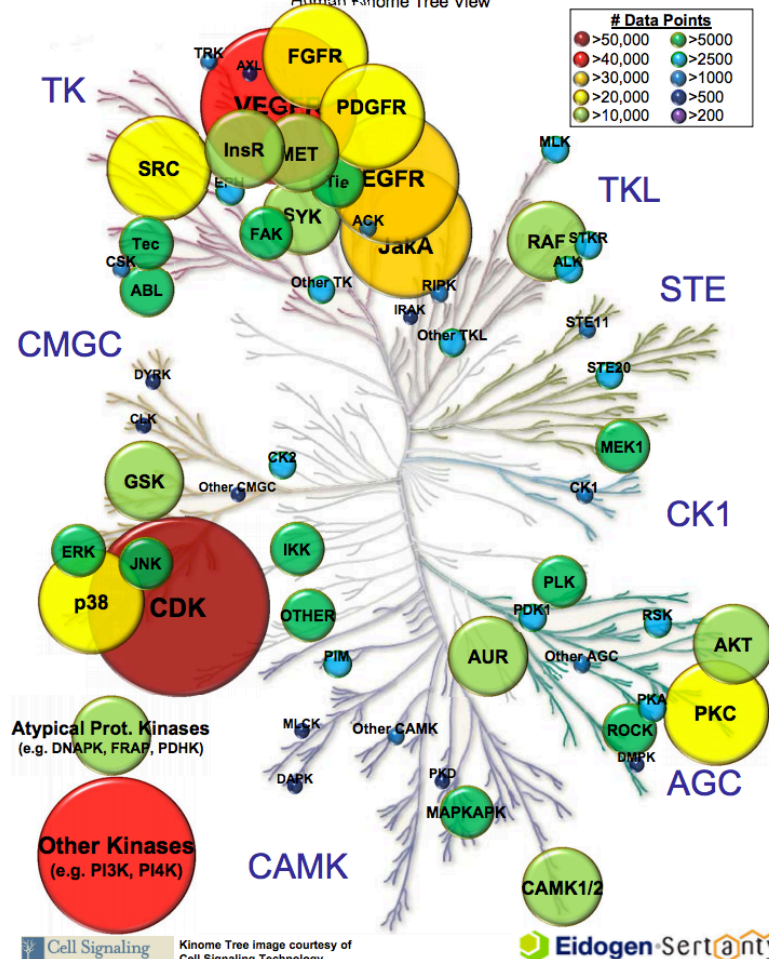
from Vieth *et al. Drug Disc. Today* **10**, 839 (2005).



Eidogen-Sertanty KKB SAR Data Point Distribution

Distribution of Kinase Knowledgebase Bio-Activity Data Points By Family
Q2 2012 Release

Human Kinome Tree View



> 649,000 SAR data points curated from
> 7915 journal articles and patents

Steven Muskal

Kinase inhibitor structures and SAR data mined from
> 7915 journal articles/patents

▪ **KKB Content Summary (Q2 2012):**

of kinase targets: **> 480**

of SAR Data points: **> 649,000**

of **unique** kinase molecules with SAR data: **>241,000**

of annotated assay protocols: **>25,472**

of all kinase inhibitors (with or without bio-activity data): **> 586,000**

▪ **KKB Growth Rate:**

- Average **15-20K** SAR data points added per quarter
- Average **20-30K** unique structures added per quarter

KKB Summary Statistics – Q2 2012

Articles covered:	2,307	(+ 30)
Patents and patent applications covered:	5,608	(+ 93)
Total Number of Bio-activity data points:	649,384	(+ 31,602)
Total Number of unique molecules:	586,610	(+ 8601)
Total Number of unique molecules w/ assay data:	241,680	(+ 8601)
Total Number of assay protocols:	25,472	(+ 322)

Targets with largest increase in Data Points in Q2-12

Target	# Data Points added
FGFR3	4626
KDR	4482
FGFR1	4446
FLT3	3047
TTK	1634
FGFR2	1428
FGFR4	1364
PIK3CA	1254
PIK3CD	955
JAK3	920
MTOR	827
JAK2	533
PTK2	463
RPS6KB1	425
JAK1	387
ALK	361
AKT1	357
ROCK2	335
SYK	305
BRAF	268
GSK3B	251
LRRK2	219
EGFR	211
BTK	197
TYK2	188
IRAK4	178
PIK3CB	158
PIK3CG	147
PIM1	147
IKBKB	130
CDK2	117
MAPK1	108
ERBB2	94
CSF1R	73
MET	72
TGFBR1	68
PLK1	55
PIM3	54
CDK9	52

Why Dotmatics?

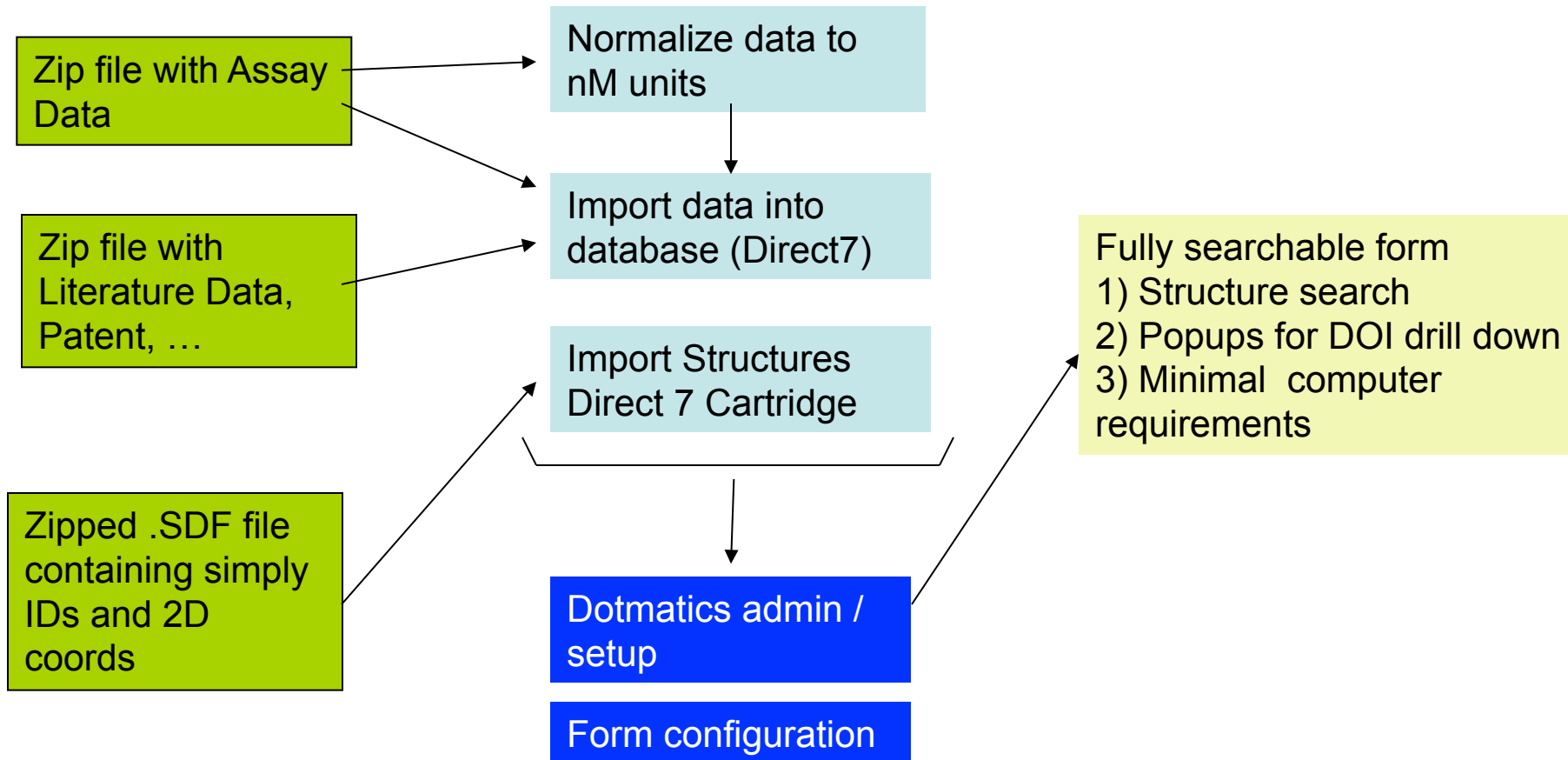
- **System already familiar to internal researchers**
 - Main format for project team data
- **Allows for viewing of all data at once**
 - Structures, data, references all on one page
 - Lookup functionalities make searches easier
 - Popups/Hyperlinks for increased access to references
 - Allows set-up of pre-canned views for simplicity and consistency of data viewing
- **Speed of searches**
 - Previous set up involved time consuming join protocols in Pipeline Pilot (~1-2 minutes)
 - Use of data cartridges vastly improves query times (~1-3 seconds)
- **Practical reasons**
 - Batteries Included. A lot of free features (lookups, tables, etc.)
 - Plays well with others. Various software of various origin can all work together under the hood.

Eidogen into Dotmatics browser

Eidogen provides

What we do

End Result



Main Form

Br

Browser from
Dotmatics, 2012

ID	=	<input type="text"/>
----	---	----------------------

Literature Source

eidotype	=	journal	=	volume	=	title	=	eidoyear	=	pages	=	DOI
=		=		=		=		=		=		=

Patent info

assignee	=	<input type="text"/>
inventors	=	<input type="text"/>
patentnumber	=	<input type="text"/>
pt keyword	=	<input type="text"/>
title	=	<input type="text"/>

SSS	EXACT	,60)=1	,70)=1	,80)=1
-----	-------	--------	--------	--------

Normalized to nM ***Run Potency Queries Here***

assaytarget	=	assaymeasure	=	assaytype	=	assayvalue nm	=	assayvaluemax nm	=	assayvaluemin nm	=	assayrelation
=		=		=		=		=		=		=

Please Note: The upper table has only normalized data with units in nM. The table below contains the raw data as in the database but the units vary (e.g. a query for 10nM will not find 0.01uM). % inhibition data can be seen and queried in the lower table.

Raw data from Eidogen ***NOTE units vary, this will affect queries***

assaytarget	=	assaymeasure	=	assaytype	=	assayvalue	=	assayvaluemax	=	assayvaluemin	=	assayunit	=	assayrelation
=		=		=		=		=		=		=		=

Also note: A query in one table will return just the queried data in that table. However, the other table returns all data for the compound. This data may be useful to observe, but be sure to note which datasource you are using when creating views.

Lookups with dropdown lists

Browser from Dotmatics, 2012

Literature Source

eidotype	journal	volume	title
----------	---------	--------	-------

-- Webpage Dialog

paste values or choose from list: cancel apply

- AAK1
- ABL
- ABL1
- ABL2
- ACK1
- ACP5
- ACVR1
- ACVR1B
- ACVR2A
- ACVR2B
- ACVRL1
- ADA
- ADCK3
- ADCK4
- ADCY
- ADK
- ADORA3
- ADRBK1
- ADRBK2
- AHCY
- AHR
- AK
- AK1
- AK2

Double click

SSS	EXACT	,60)=1	,70)=1	,80)=1
-----	-------	--------	--------	--------

Normalized to nM ***Run Potency Queries Here***

assaytarget	assaymeasure	assaytype	assay
-------------	--------------	-----------	-------

Raw data from Eidogen ***NOTE units vary, this will affect queries***

assaytarget	assaymeasure	assaytype	assayvalue
-------------	--------------	-----------	------------

http://fcdotprd01:8080/browser/query/lookupContro Local intranet

- Easy access to alias lists to find kinase of interest

Results of ABL target search: Data

Browser from Dotmatics, 2012
Kerim Babaoglu (admin) EDOGENKKB
forms queries list views tools clear edit query browse

543 records

*Normalized to nM ***Run Potency Queries Here****

assaytarget	assaymeasure	assaytype	assayvalue nm	assayvaluemax nm	assayvaluemin nm	assayrelation
ABL	IC50	Enzyme Assay	410	0	0	=

*Raw data from Eidogen ***NOTE units vary, this will affect queries****

assaytarget	assaymeasure	assaytype	assayvalue	assayvaluemax	assayvaluemin	assayunit	assayrelation
ABL	IC50	Enzyme Assay	0.4100			uM	=
EGFR	IC50	Cell-Based Assay	14			uM	=
CSK	IC50	Enzyme Assay	6			uM	=
EGFR	IC50	Cell-Based Assay	10			uM	>
EGFR	IC50	Enzyme Assay	0.1600			uM	=
CSK	IC50	Enzyme Assay	15			uM	=
EGFR	IC50	Cell-Based Assay	35			uM	=
PRKCA	IC50	Enzyme Assay	24			uM	=

PRKCA IC50 Enzyme Assay 24 uM =

- Dual tables allow for return of both queried data and all data present for said entry

Results of ABL target search: Literature

Browser from Dotmetrics, 2012

1 >> 543 records

Kerim Babaoglu (admin) EDOGENKKB

forms queries list views tools clear edit query browse

Literature Source

eidotype	journal	volume	title	eidoyear	pages	DOI
A	J. Med. Chem.	40	Use of a Pharmacophore Model for the Design of EGF-R Tyrosine Kinase Inhibitors: 4-(Phenylamino)pyrazolo [3,4-d]pyrimidines	1997	3601-3616	10.1021/jm970124v
A	Chem. Rev.	101	Comparative QSAR Study of Tyrosine Kinase	2001	2573-2600	10.1021/cr010154c

Patent info

assignee	
inventors	
patentnumber	
pt keyword	
title	

Raw data from Eidogen ***NOTE: units vary, this will affect queries***

assaytarget	assaymeasure	assaytype	assayvalue	assayvaluemax	assayvaluemin	assayunit	assayrelation
ABL	IC50	Enzyme Assay	0.4100			uM	=
EGFR	IC50	Cell-Based Assay	14			uM	=
CSK	IC50	Enzyme Assay	6			uM	=
EGFR	IC50	Cell-Based Assay	10			uM	>
EGFR	IC50	Enzyme Assay	0.1600			uM	=
CSK	IC50	Enzyme Assay	15			uM	=
EGFR	IC50	Cell-Based Assay	35			uM	=
PRKCA	IC50	Enzyme Assay	24			uM	=

the other table returns all data for the compound. This data may be useful to observe, but be sure to note which datasource you are using when creating views.

Addition of doi allows easy hyperlinking

Browser from Dortmund, 2012

http://pubs.acs.org/ - Use of a Pharmacophore Model for the Design of EGF-R Tyrosine Kinase Inh...

Log In Register Cart Website Demos ACS Journals

ACS Publications MOST TRUSTED. MOST CITED. MOST READ.

Publications A-Z Home | Authors & Reviewers | Librarians | ACS Member

Search Citation DOI Subject Search

Anywhere Search

J. Med. Chem. All Publications/Website

Subscriber access provided by Literature Resources

Browse the Journal Articles ASAP Current Issue Submission & Review Subscribe About

Article

Use of a Pharmacophore Model for the Design of EGF-R Tyrosine Kinase Inhibitors: 4-(Phenylamino)pyrazolo[3,4-d]pyrimidines

Peter Traxler,* Guido Bold, Joerg Frei, Marc Lang, Nicholas Lydon, Helmut Mett, Elisabeth Buchdunger, Thomas Meyer, Marcel Mueller, and Pascal Furet

Novartis Pharmaceuticals, Therapeutic Area Oncology, Novartis Limited, CH-4002 Basel, Switzerland

J. Med. Chem., 1997, 40 (22), pp 3601-3616
DOI: 10.1021/jm970124v
Publication Date (Web): October 24, 1997
Copyright © 1997 American Chemical Society

Section: Pharmacology

Abstract

In the course of the random screening of a pool of CIBA chemicals, the two pyrazolopyrimidines **1** and **2** have been identified as fairly potent inhibitors of the EGF-R tyrosine kinase. Using a pharmacophore model for ATP-competitive inhibitors interacting with the active site of the EGF-R protein tyrosine kinase (PTK), the class of the pyrazolo[3,4-d]pyrimidines was then optimized in an interactive process leading to a series of 4-(phenylamino)-1H-pyrazolo[3,4-d]pyrimidines as highly potent inhibitors of the EGF-R tyrosine kinase. The most potent compounds **13**, **14**, **15**, **17**, **19**, **22**, **26**, **28**, and **30** of this

Tools

- Add to Favorites
- Download Citation
- Email a Colleague
- Permalink
- Order Reprints
- Rights & Permissions
- Citation Alerts

SciFinder Links

SciFinder®

- Get Reference Detail
- Get Substances
- Get Cited
- Get Citing

Explore by:

- Author of this Article
- Any Author
- Research Topic (Now with patent search)

Related Content

4-(Phenylamino)pyrazolo[3,4-d]pyrimidines and Selective Inhibitors of the EGF-Receptor Tyrosine Kinase

Soda Can Cap: Unexpectedly High Molecular Weight

Two-Stage Synthesis of a Cyclohexane Macromolecule

Other ACS authors:

Peter Traxler
Guido Bold
Joerg Frei
Marc Lang
Nicholas Lydon
Helmut Mett

pages	DOI
3601-3616	10.1021/jm970124v
2573-2600	10.1021/cr010154c

Normalized to nM ***Rur

assay/target	
ABL	

Raw data from Eidogen

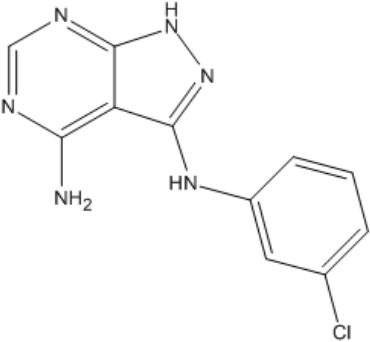
assay/target	assay
ABL	
EGFR	
CSK	
EGFR	
EGFR	
CSK	
EGFR	
PRKCA	

normalized data with units in nM. The table is in the database but the units vary (e.g. a query for EGFR data can be seen and queried in the lower table).

not just the queried data in that table. However, this data may be useful to observe, especially when creating views.

Substructure searches

Browser from Dotmatics, 2012



Chemical structure of 4-(4-chlorophenyl)-1H-imidazo[4,5-b]pyridin-2-amine is displayed. The structure consists of a fused imidazo[4,5-b]pyridine ring system with an amino group (-NH₂) at the 2-position and a 4-chlorophenyl group (-NH-C₆H₄-Cl) at the 4-position.

Search criteria: SSS EXACT ('60')=1 ('70')=1 ('80')=1

Normalized to nM ***Run Potency Queries Here***

assaytarget	assaymeasure	assaytype	assayvalue nm	assayvaluemax nm	assayvaluemin nm
= ABL	=	=	=	=	=

Tools: ID, Literature Source (eidotype, journal), Patent info (assignee, inventors, patentnumber, pt keyword, title)

Vertical toolbar: Drawing tools (lines, circles, polygons, text, arrows, eraser, lasso, zoom, pan, etc.)

- Allows user familiar drawing tools
- Can be combined with other queries (e.g. target)
- Speed!

Substructure results stock views

Browser from Dotmatics, 2012

1-7 of 7

Records large structures

Show Excluded:

forms queries list views tools clear edit query browse

17

18

20

21

22

23

43

ASSAYMEASURE: IC50
ASSAYVALUE_NM: 410

ASSAYMEASURE: IC50
ASSAYVALUE_NM: 1900

ASSAYMEASURE: IC50
ASSAYVALUE_NM: 1360

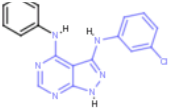
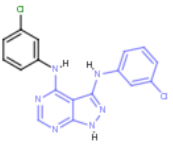
ASSAYMEASURE: IC50
ASSAYVALUE_NM: 1980

ASSAYMEASURE: IC50
ASSAYVALUE_NM: 1360

ASSAYMEASURE: IC50
ASSAYVALUE_NM: 1660

ASSAYMEASURE: IC50IC50
ASSAYVALUE_NM: 1500010000

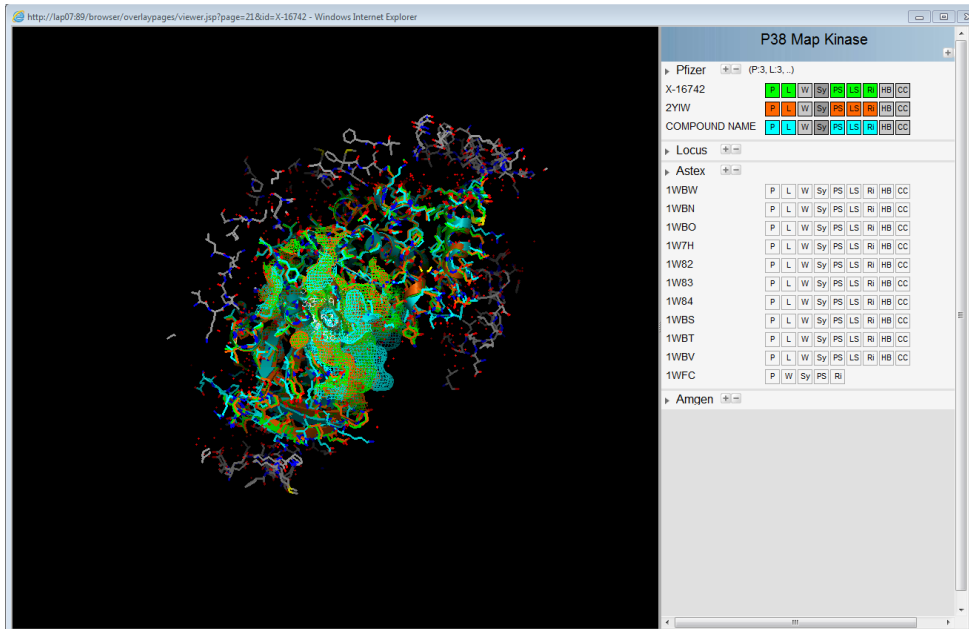
Exclude: Include:

				
4	21		IC50	1980

Decision Thumbnails

- Admin created project views make unified data views one click away

Future directions: PDBs and Kinomes



Link out to public pdb's or embed viewer on form

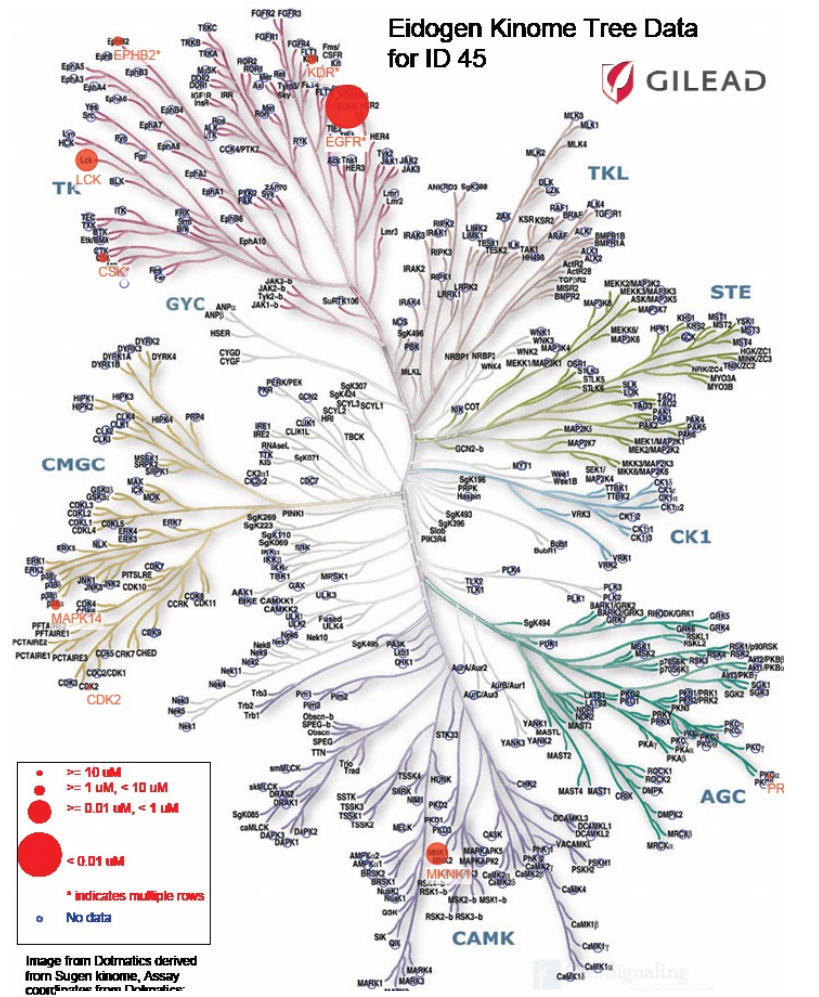


Image from Dolmatics derived from Sugen kinome, Assay coverage from Dolmatics

ASSAYTARGET	ACTIVITY	ASSAYRELATION
EGFR	5.0000e-08	=
LCK	0.080000	=
MKNNK1	0.63000	=
CSK	1.4000	=
KDR	1.4890	=
MAPK14	6.3000	=
EPHB2	8.1000	=
PRKCA	100	=
CDK2	250	=

Summary

- **Eidogen KKB currently being served up to internal researchers in a format they are used to using**
- **Allows for easy querying and viewing of all data at once**
 - Structures, data, references all one one page
 - Searches return results in seconds vs. minutes
 - Pre-canned views simplify user experience
 - Popups and hyperlinks allow easy tie in with other internal resources
- **Future Directions/Wish list**
 - Create Popup for Patent information similar to current doi Popup
 - Kinome plots using dotmatics tools
 - Links to PDB for easy access of public structures

Acknowledgements

Gilead

- **George Lane**
- **Uli Schmitz**

Eidogen/Sertanty

- **Steven Muskal**

Dotmatics

- **Shikha O' Brien**
- **Philip Mounteney**