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….

Accord for Excel

ChemDraw for Excel

SD file

Spotfire

ISIS Base

DIVA

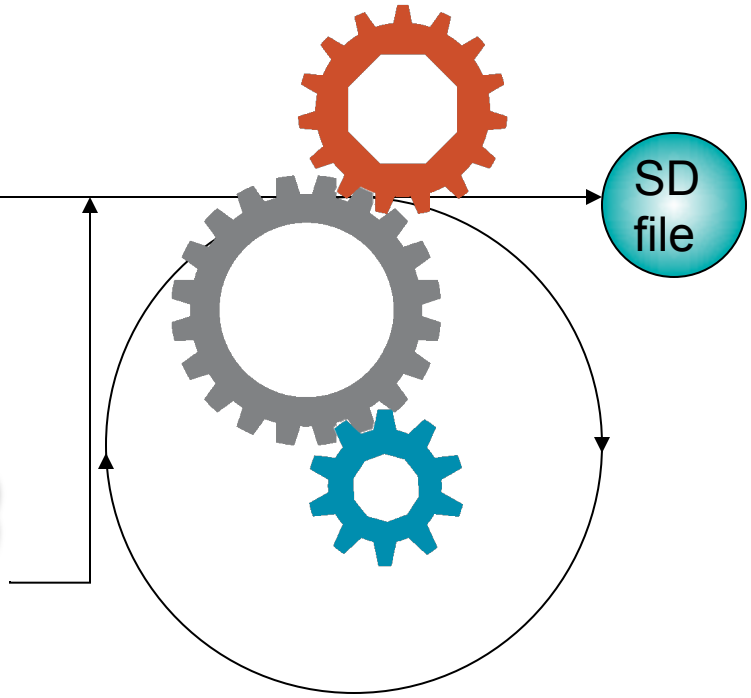
DiscoveryStudio

Processing
Annotation
With PipelinePilot

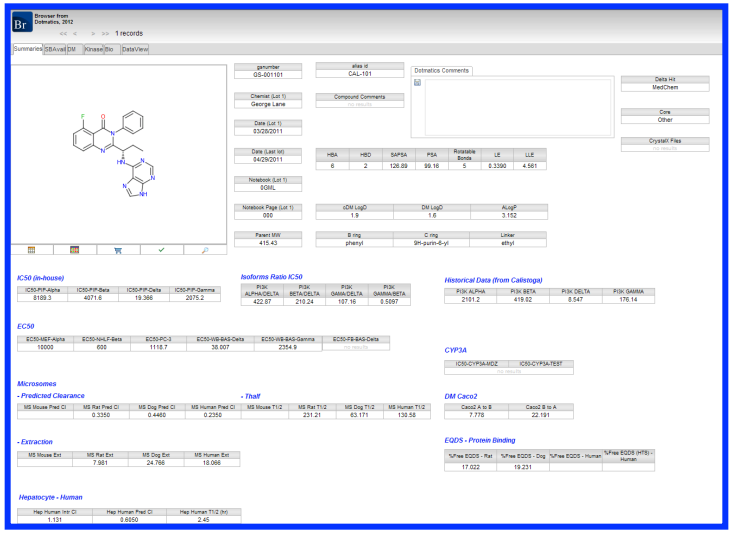
Examples:

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



All-in-one data visualization

FDA Approved Protein Kinase Inhibitors (as of March 2012)

Journal of Medicinal Chemistry Perspective

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

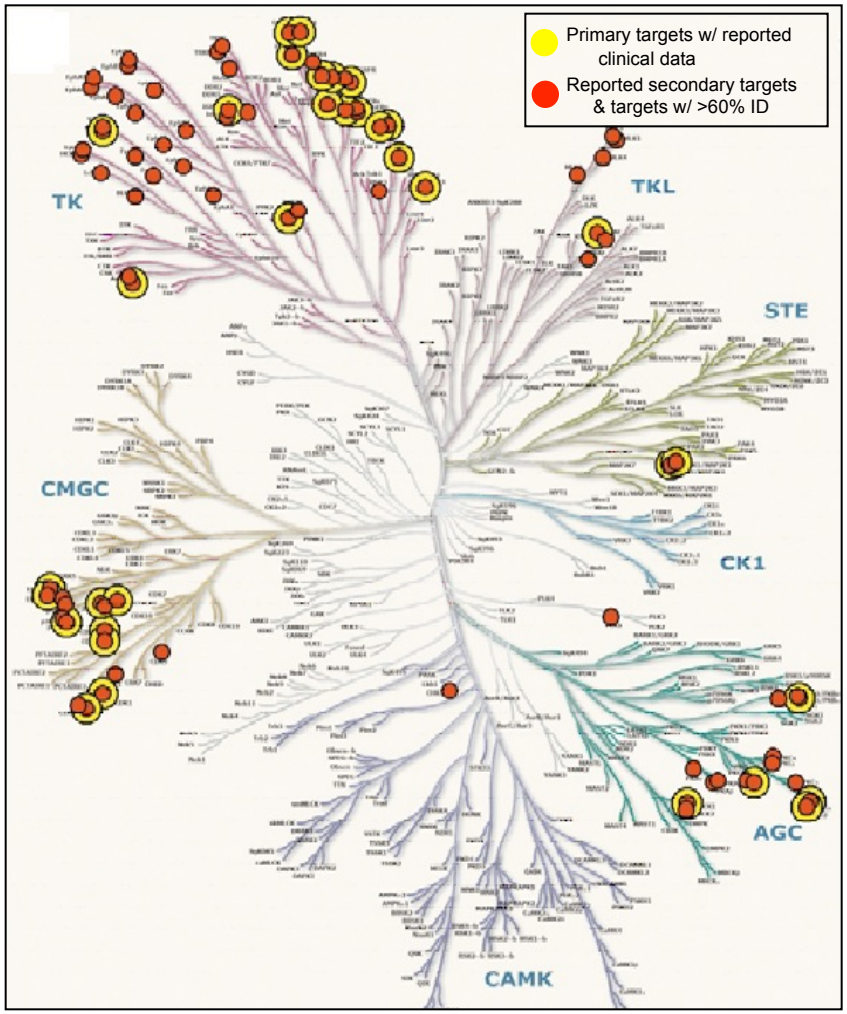
generic (brand) name	year of approval	company	indication	target kinase
imatinib (Gleevec)	2001	Novartis	chronic myeloid leukemia (CML)	Abl, c-Kit, PDGFRα/β
gefitinib (Iressa)	2003	AstraZeneca	non-small-cell lung carcinoma (NSCLC)	EGFR
erlotinib (Tarceva)	2004	Genetech, OSI	NSCLC, pancreatic cancer	EGFR
sorafenib (Nexavar)	2005	Bayer, Onyx	hepatocellular carcinoma, renal cell carcinoma (RCC)	Raf, VEGFR2/3, c-Kit, PDGFRβ
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, PDGFRα/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, PDGFRβ, c-Kit

Reference: dx.doi.org/10.1021/jm3003203 | J. Med. Chem. 2012, 55, 6243–6262

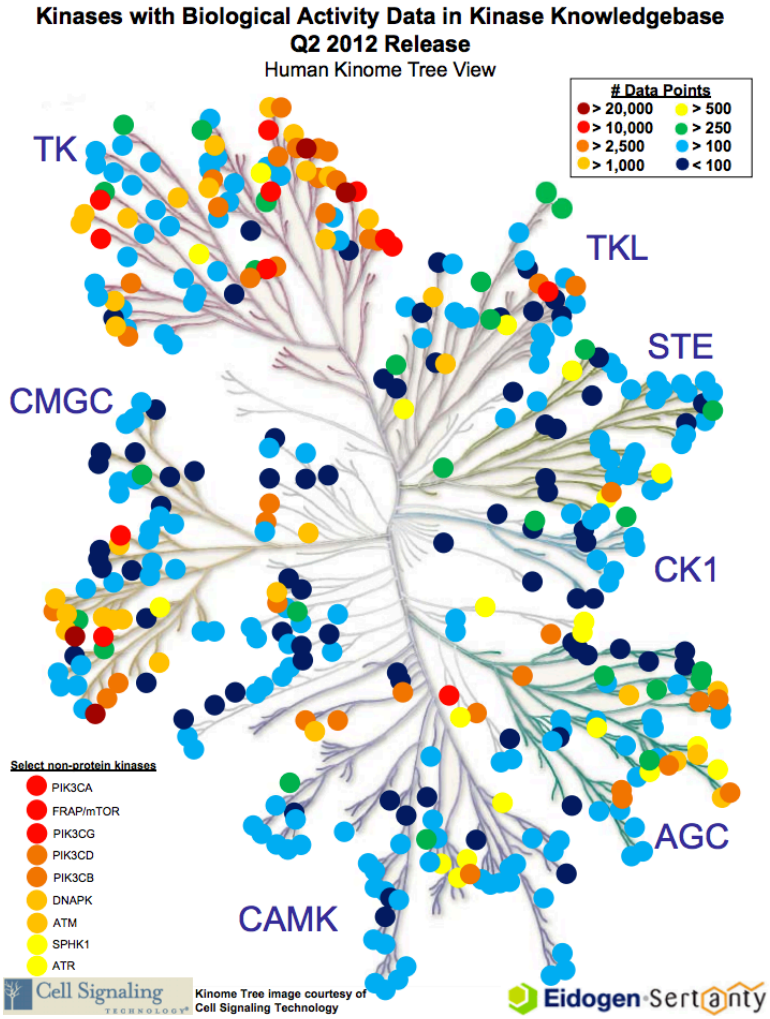
Kinase SAR Knowledgebase (KKB) – Hot Targets

Kinase Targets of Clinical Interest

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



Eidogen-Sertanty KKB SAR Data Point Distribution



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

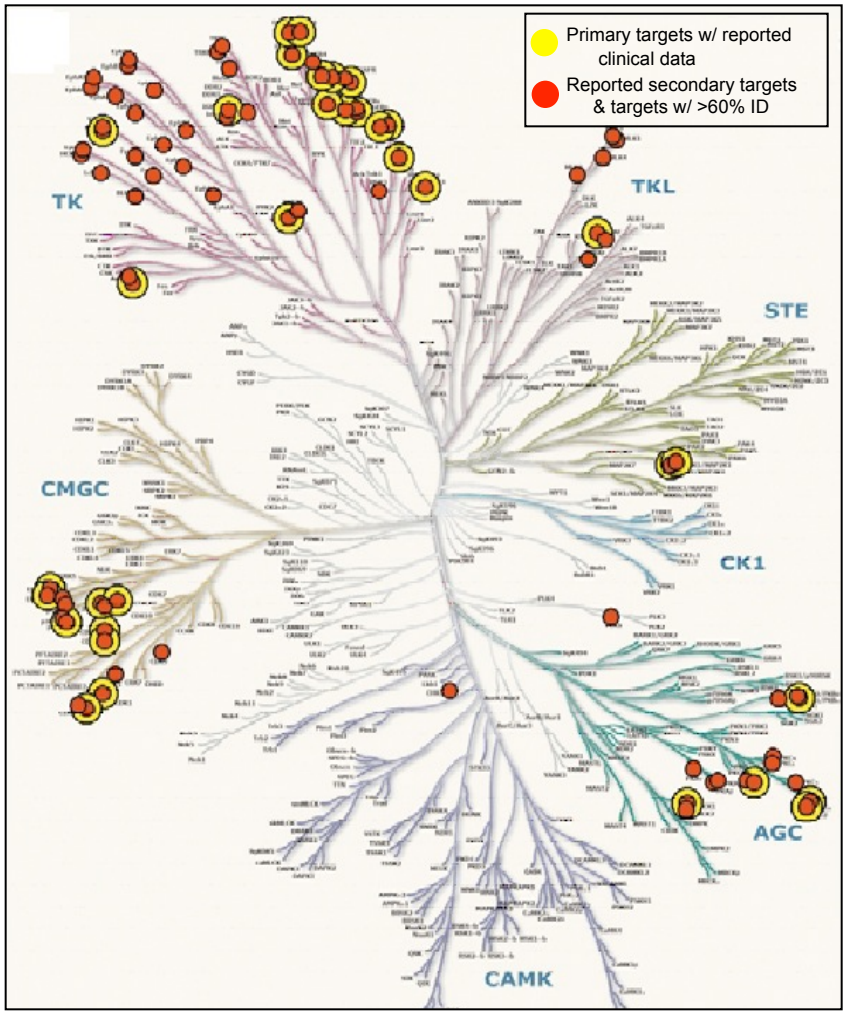
Eidogen Sertanty

Steven Muskal

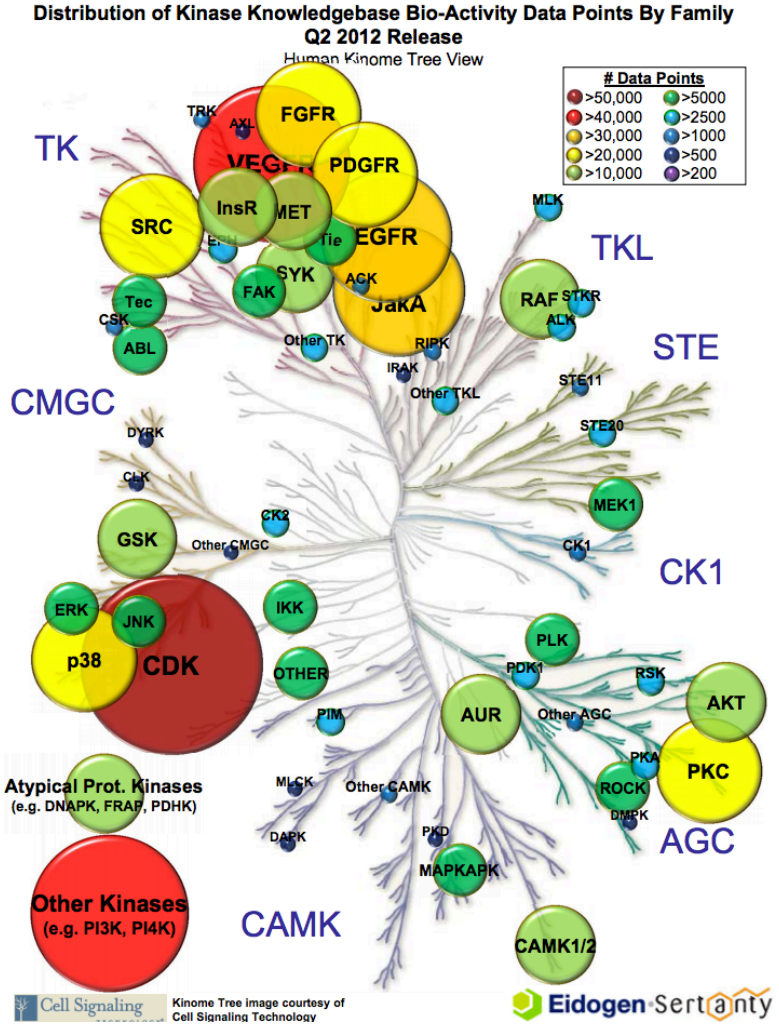
Kinase SAR Knowledgebase (KKB) – Hot Targets

Kinase Targets of Clinical Interest

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



Eidogen-Sertanty KKB SAR Data Point Distribution



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

Eidogen Sertanty

Steven Muskal

Kinase Knowledgebase (KKB)

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

Steven Muskal

Why Dotmatics?

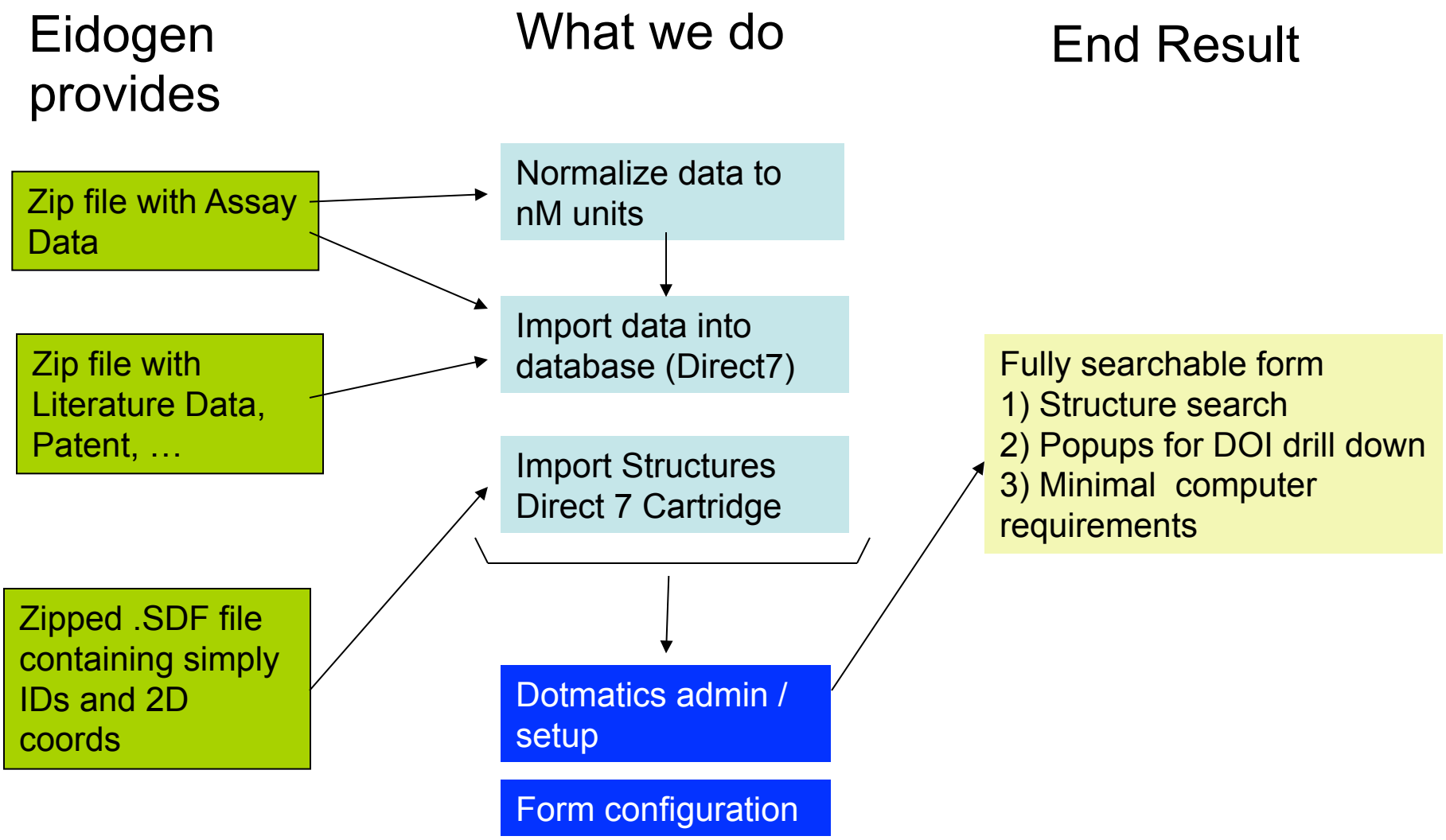
- **System already familiar to internal researchers**
  - Main format for project team data

- **Allows for viewing of all data at once**
  - Structures, data, references all one 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



George Lane

Main Form

Browser from Dotmatics, 2012

ID
=

Literature Source

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

Patent info

assignee	=
inventors	=
patentnumber	=
pt keyword	=
title	=

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
=	=	=	=	=	=	=

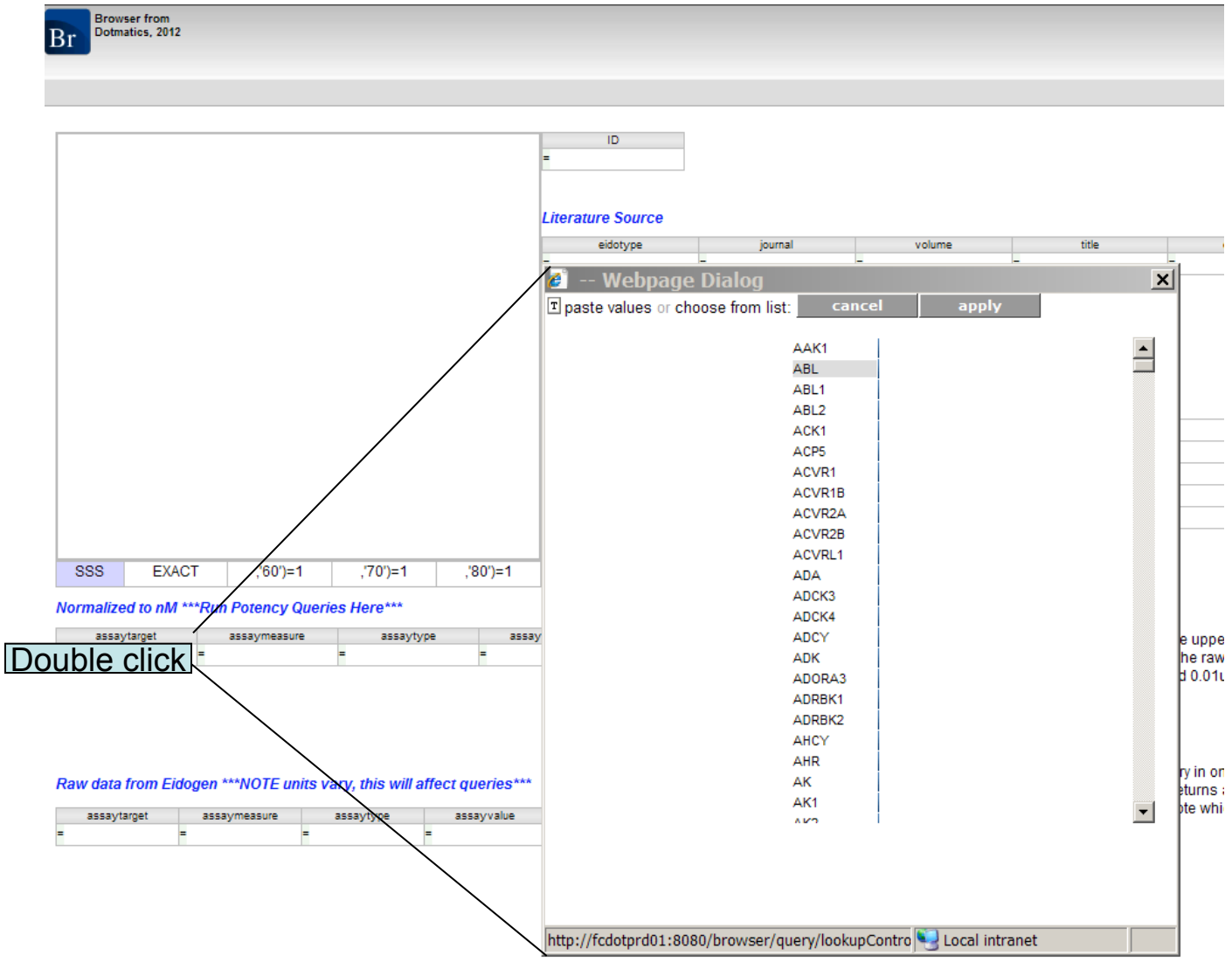
Please Note: The upper table has only normalized data with units in nM. The table below contains the raw data as in the database but the units vary (e.g. a query for 10nM will not find 0.01uM). % inhibition data can be seen and queried in the lower table.

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

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

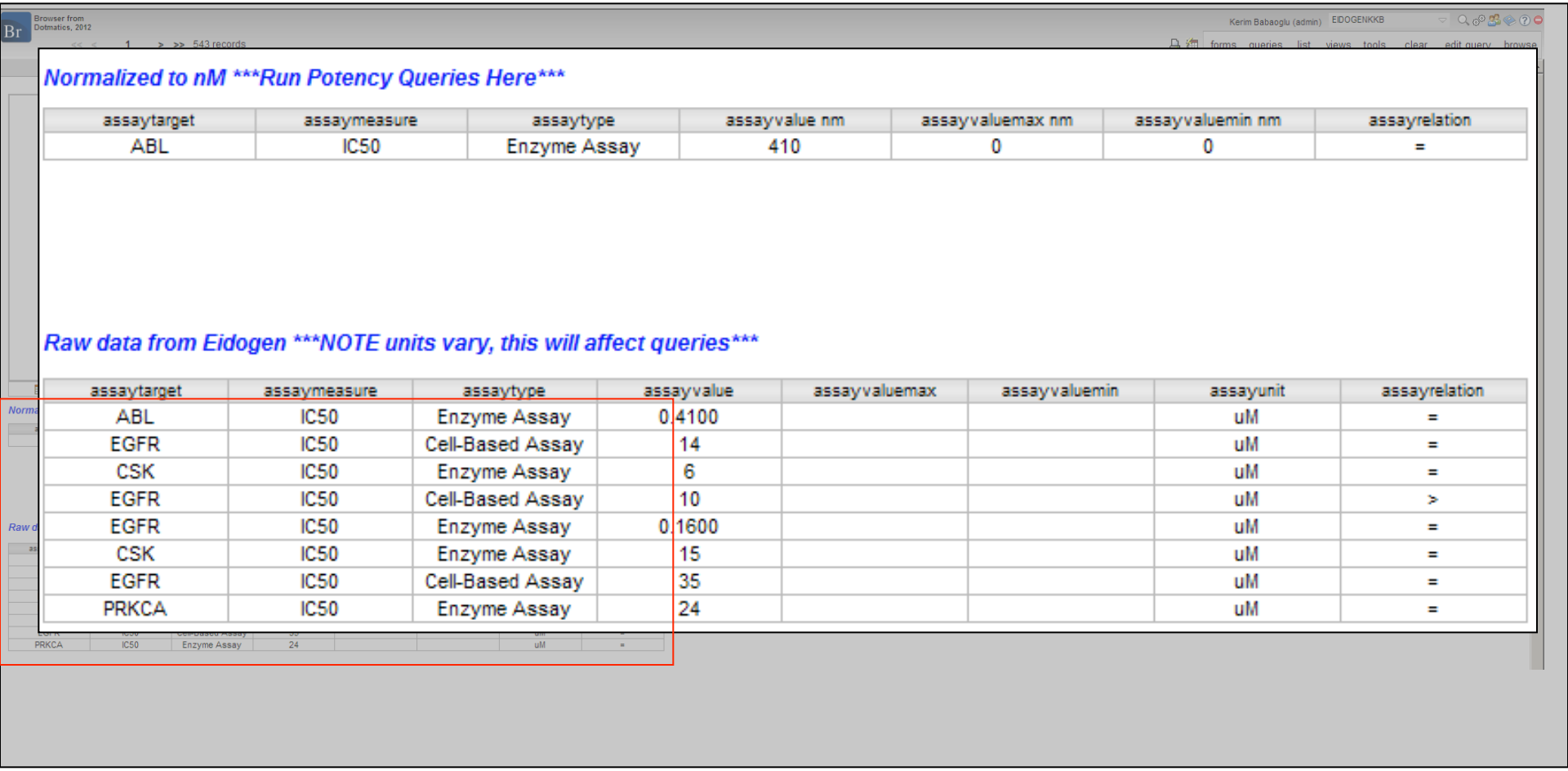
Also note: A query in one table will return just the queried data in that table. However, the other table returns all data for the compound. This data may be useful to observe, but be sure to note which datasource you are using when creating views.

Lookups with dropdown lists



- **Easy access to alias lists to find kinase of interest**

Results of ABL target search: Data



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

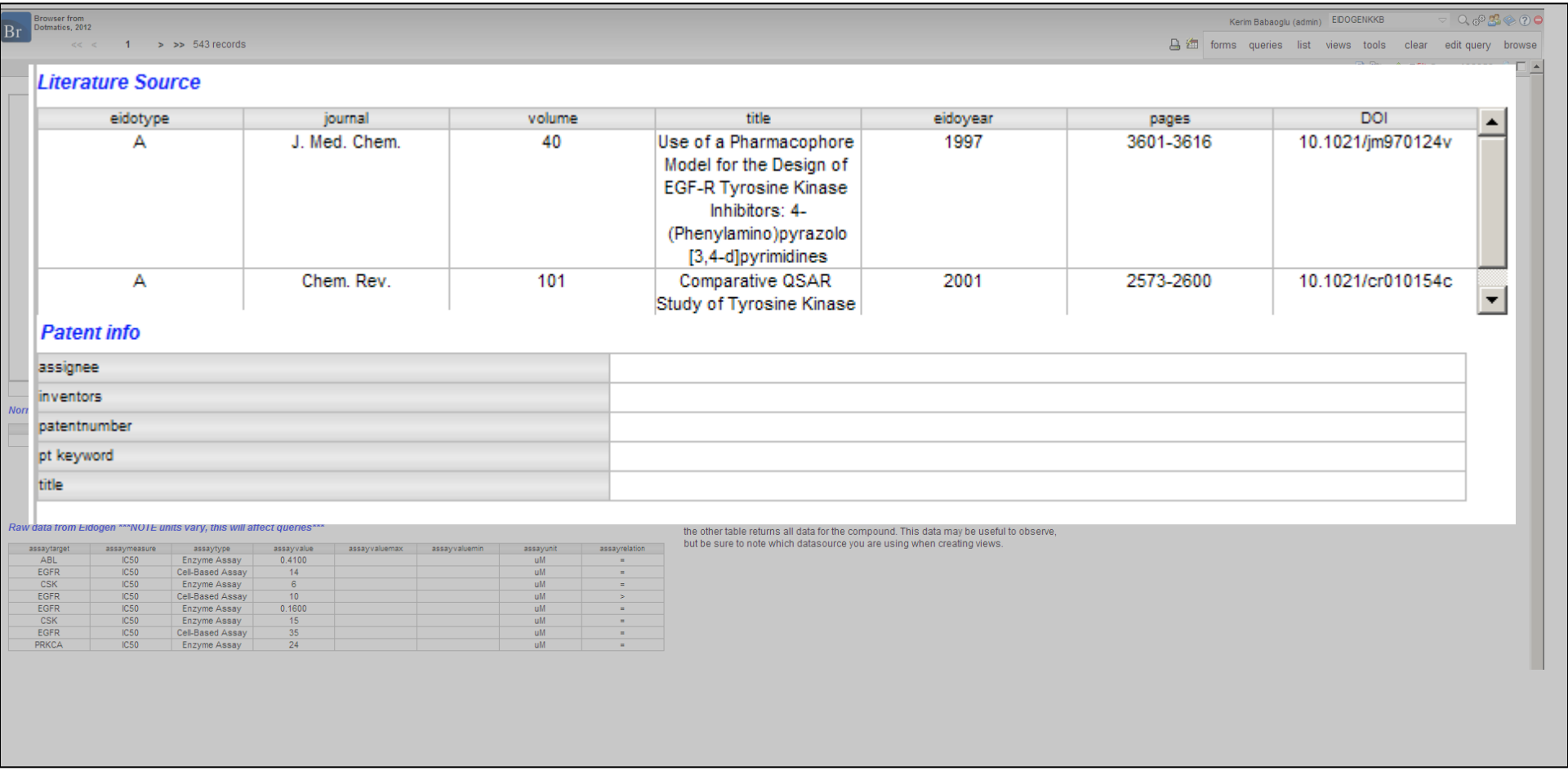
assaytarget	assaymeasure	assaytype	assayvalue nm	assayvaluemax nm	assayvaluemin nm	assayrelation
ABL	IC50	Enzyme Assay	410	0	0	=

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

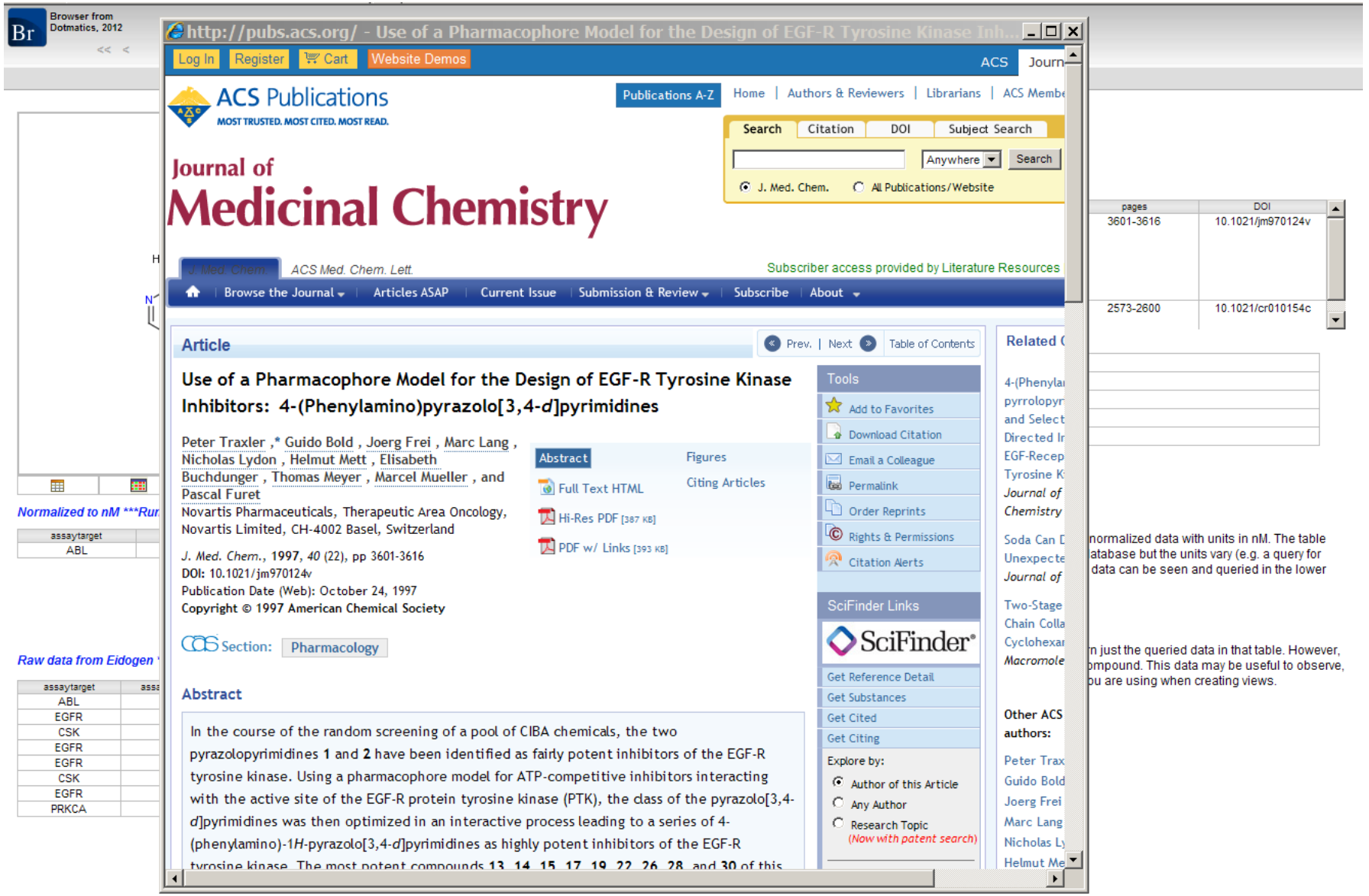
assaytarget	assaymeasure	assaytype	assayvalue	assayvaluemax	assayvaluemin	assayunit	assayrelation
ABL	IC50	Enzyme Assay	0.4100			uM	=
EGFR	IC50	Cell-Based Assay	14			uM	=
CSK	IC50	Enzyme Assay	6			uM	=
EGFR	IC50	Cell-Based Assay	10			uM	>
EGFR	IC50	Enzyme Assay	0.1600			uM	=
CSK	IC50	Enzyme Assay	15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

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

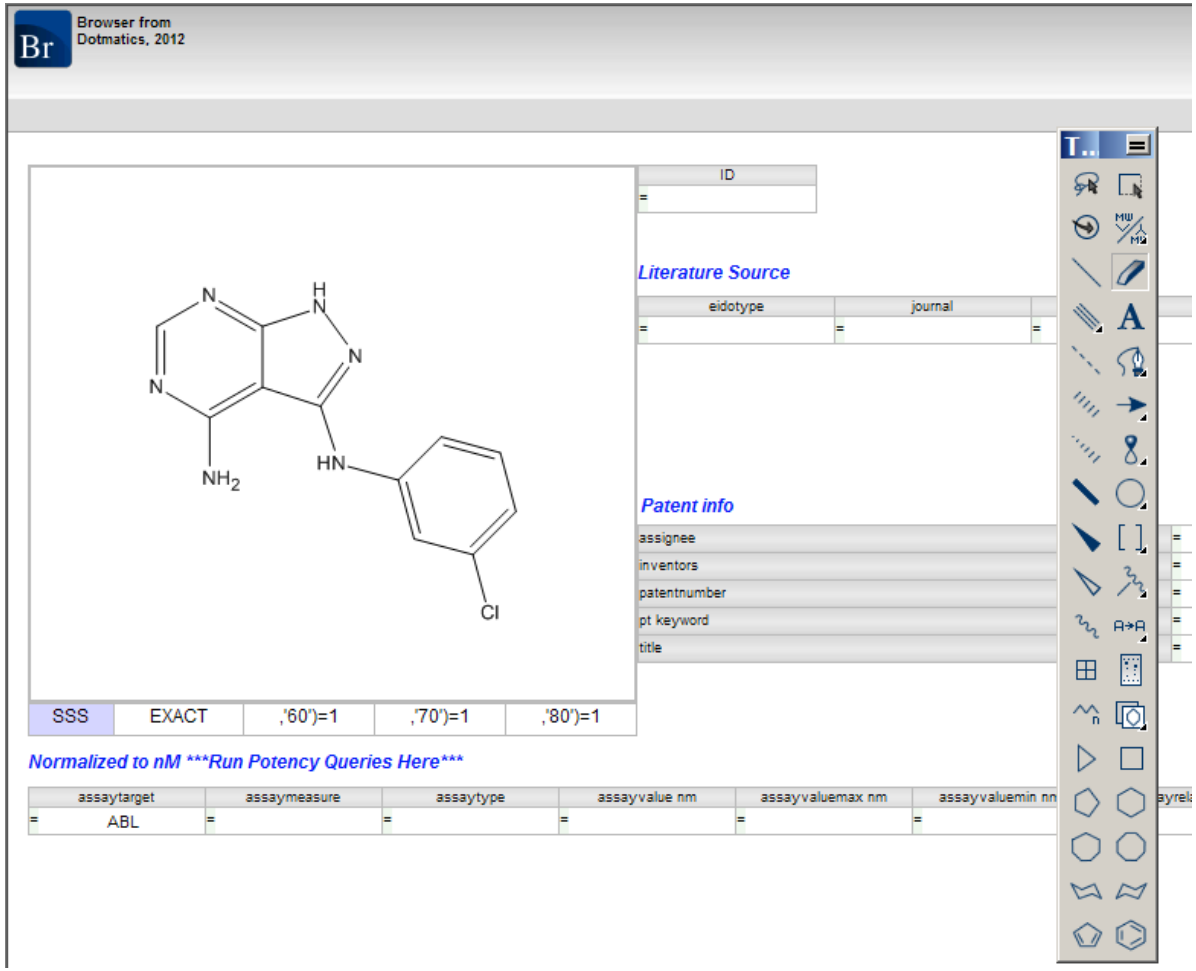
Results of ABL target search: Literature



Addition of doi allows easy hyperlinking

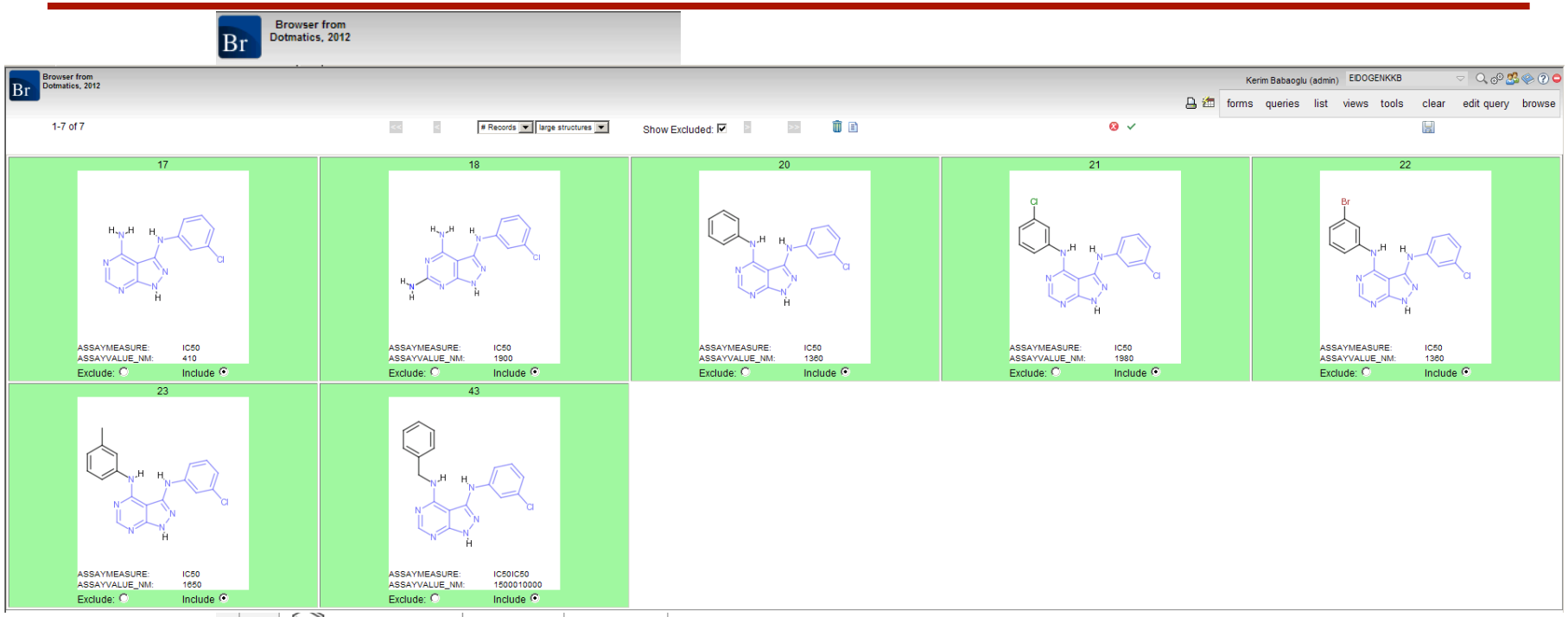


Substructure searches



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

Substructure results stock views

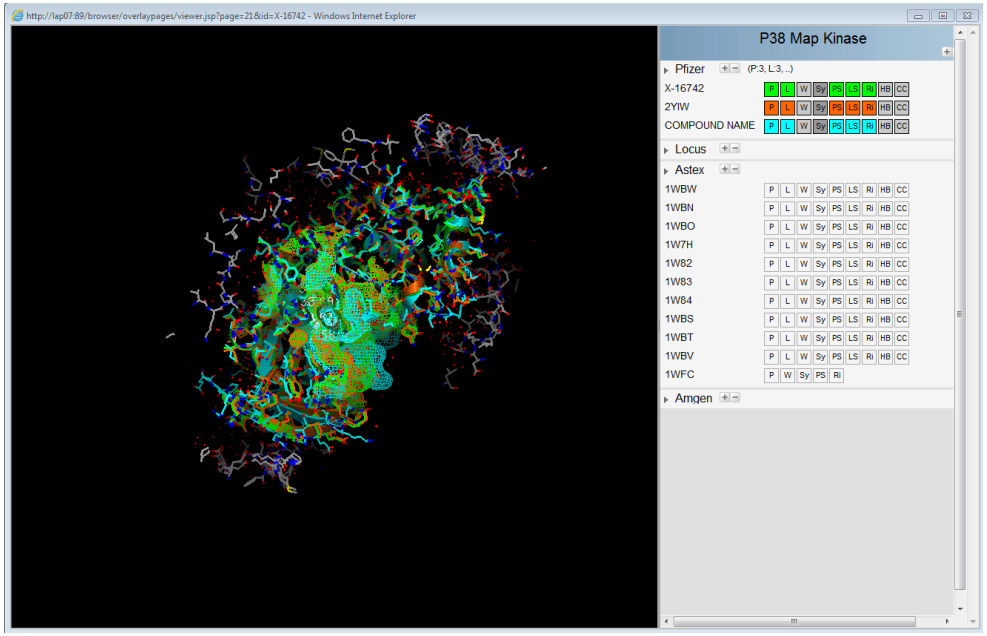


Decision Thumbnails

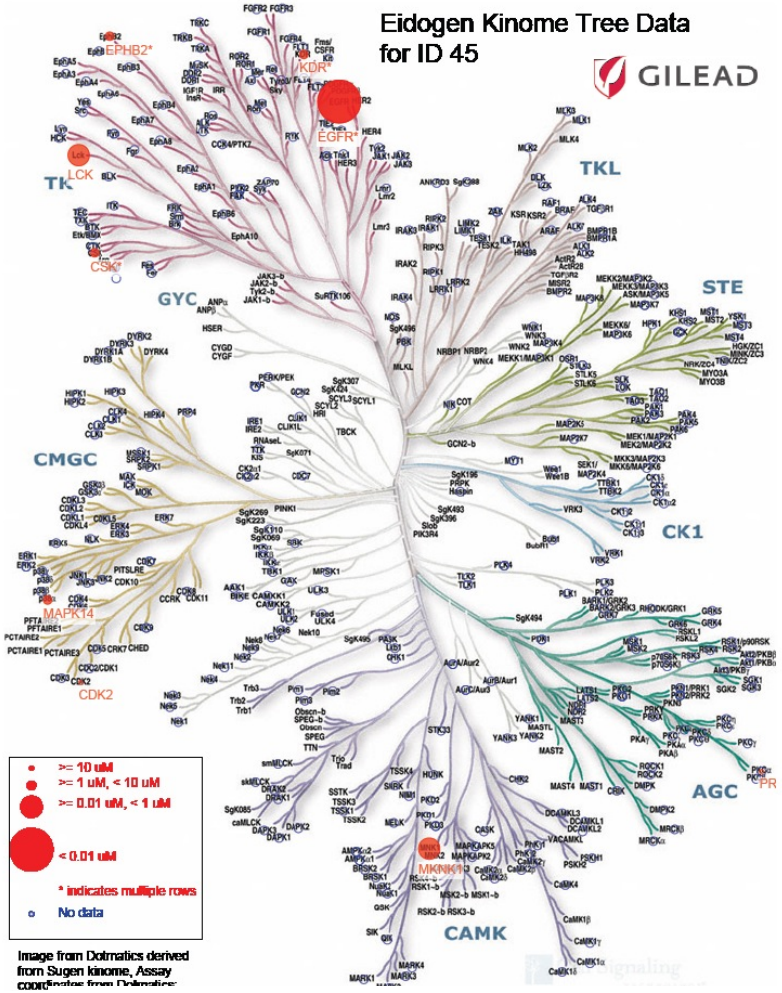
4	21		IC50	1980

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

Future directions: PDBs and Kinomes



Link out to public pdbs or embed viewer on form



ASSAYTARGET	ACTIVITY	ASSAYRELATION
EGFR	5.0000e-06	=
LCK	0.060000	=
MKNK1	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**