Using Receptor-Site Similarity to LigandCross into New Diversity

Steven Muskal, Ph.D.

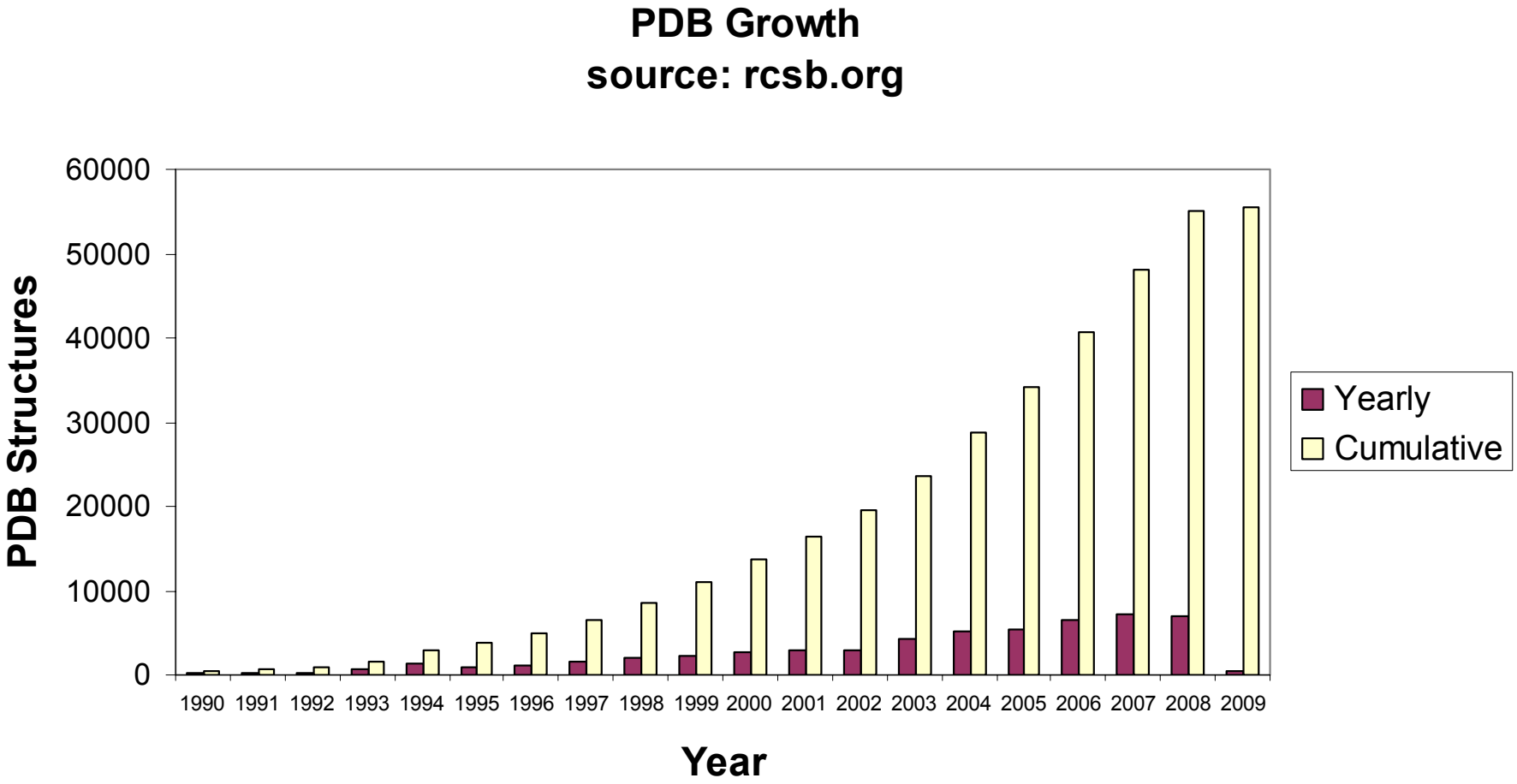
Chief Executive Officer
Eidogen-Sertanty, Inc.

smuskal@eidogen-sertanty.com

Eidogen·Sertanty

Protein Structure Growth Continues

> 50K Structures/co-complexes (Apr-2008)
> 600 deposits per month → **>150/week!**

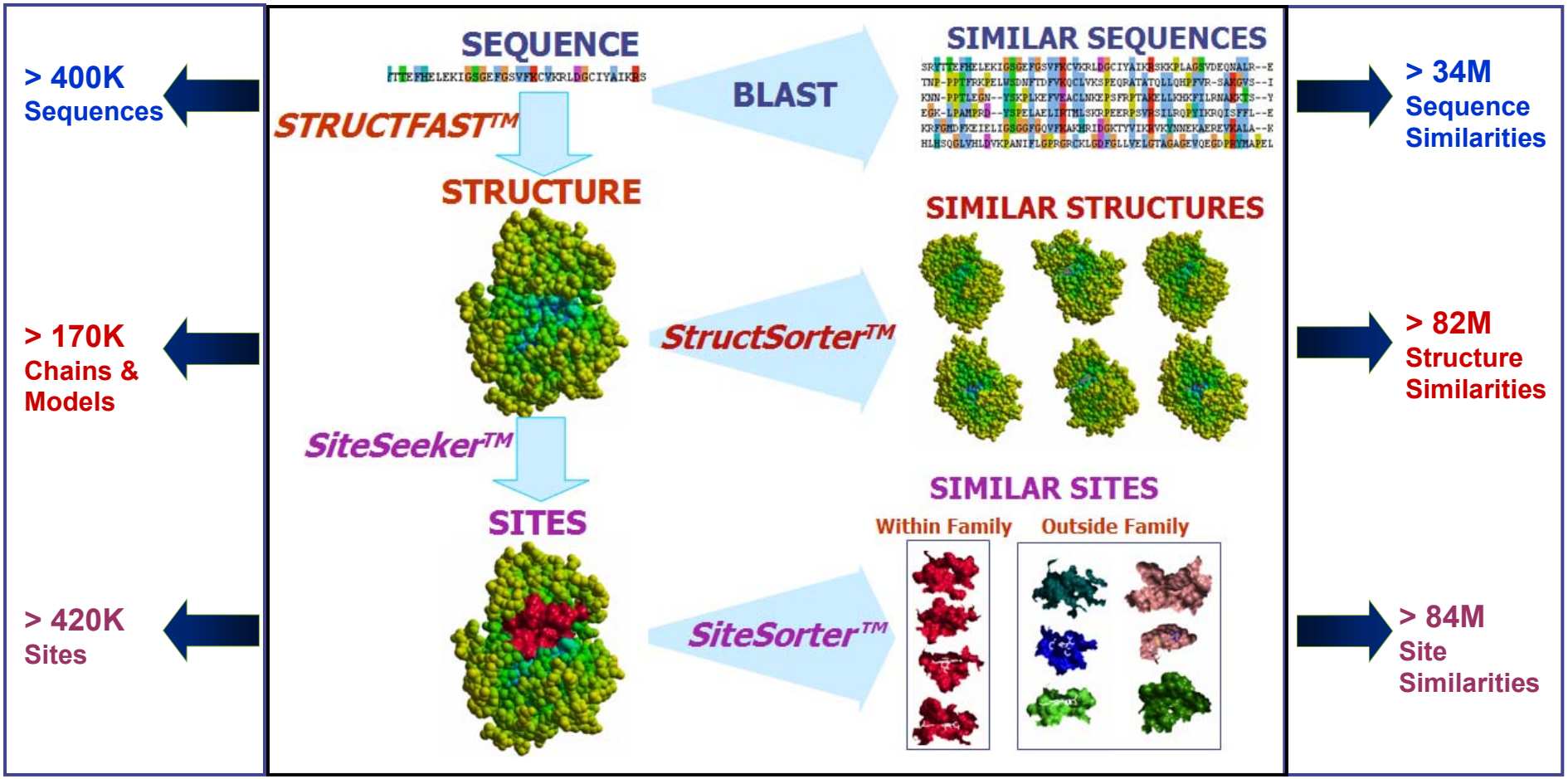


Drugs Developed using Structural Knowledge

Inhibitor/Drug	Disease	Company(s)	Protein targeted	Enzyme Family
STI-571/Gleevec	Chronic Myeloid Leukemia	Novartis	c-Abl kinase	Tyrosine kinase
Fluoroquinolone/Ciprofloxacin	Bacterial infection	Bayer	Gyrase	ATP Hydrolase
Saquinavir/Invirase, Ritonavir/Norvir, Indinavir/ Crixivan, Nelfinavir/Viracept, Amprenavir/Agenerase, Fosamprenavir/Lexiva,	AIDS	Roche, Abbott, Agouron, Merck, Vertex	HIV-1 Protease	Aspartylprotease
Trusopt	Glaucoma	Merck	Carbonic Anhydrase	Lyase
Thymitaq	Cancer	Agouron	Thymidylate synthase	Methyl transferase
Celecoxib/Celebrex, Rofecoxib/Vioxx	Inflammation, rheumatoid arthritis	Searle, Merck	Cox-2	Oxidoreductase
AG3340/Prinomastat	Cancer	Agouron	Matrix metalloprotease	Metalloprotease
Oseltamivir phosphate/Tamiflu, Zanamivir/Relenza	Influenza	Roche	Neuraminidase	Glycosidase

Source: http://www.active-sight.com/science/sbdd.html

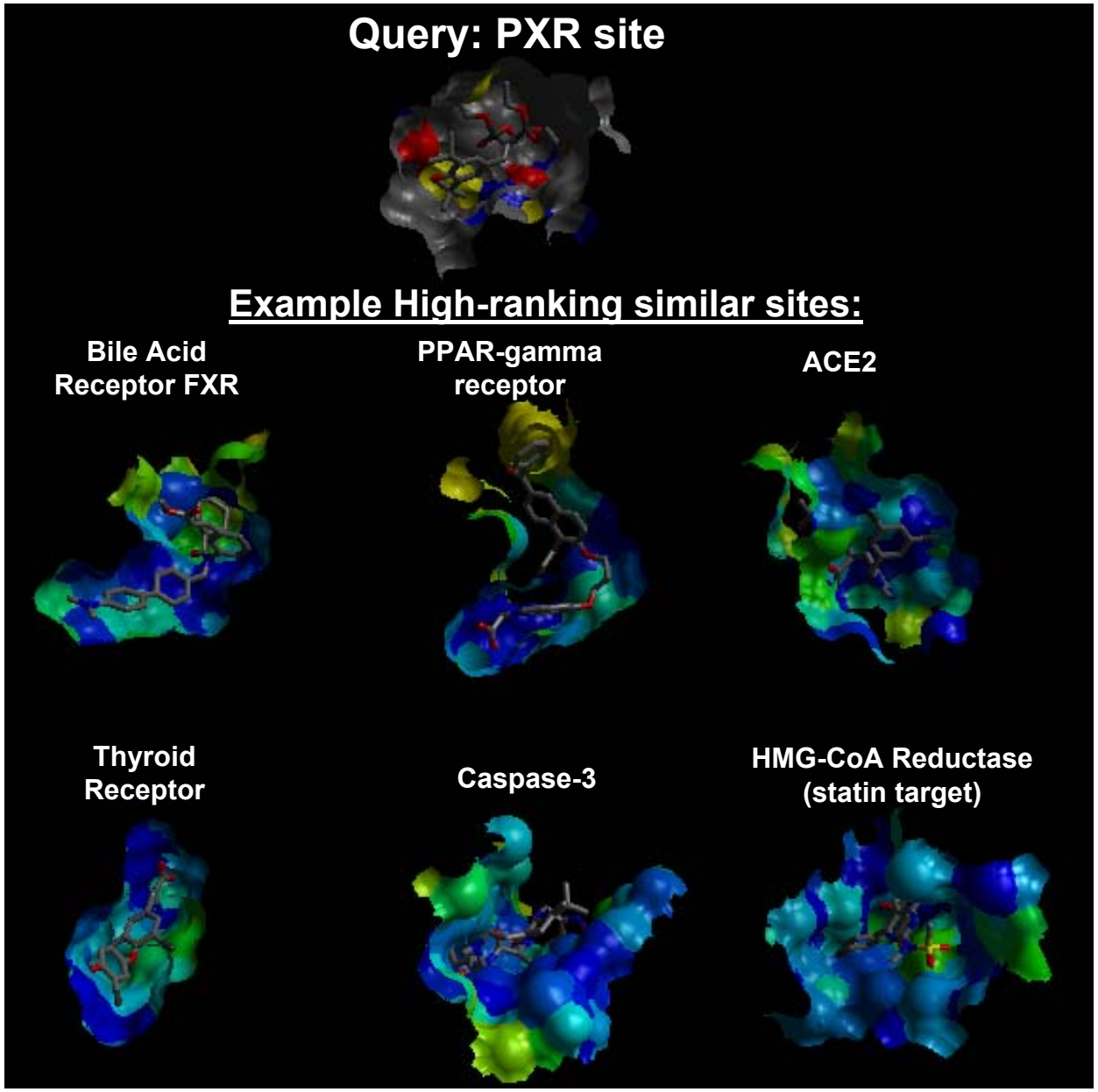
TIP Content and Algorithm Engine



- Interrogating the druggable genome with structural informatics **MolecularDiversity (2006)**
- STRUCTFAST: Protein Sequence Remote Homology Detection and Alignment Using Novel Dynamic Programming and Profile-Profile Scoring **Proteins. 2006 64:960-967**
- StructSorter: A Method for Continuously Updating a Comprehensive Protein Structure Alignment Database **J. Chem. Inf. Model. 2006, 46, 1871-1876**
- Convergent Island Statistics: A fast method for determining local alignment score significance. **Bioinformatics, 2005, 21, 2827-2831.**

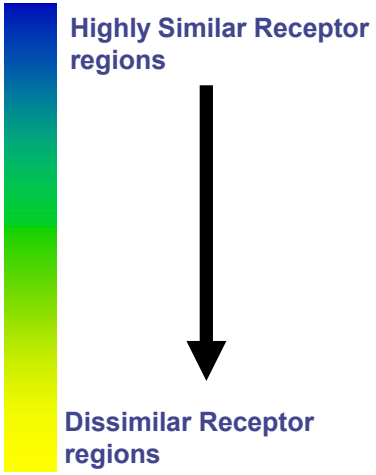
Eidogen-Sertanty

Nature Exploits Site Similarity...



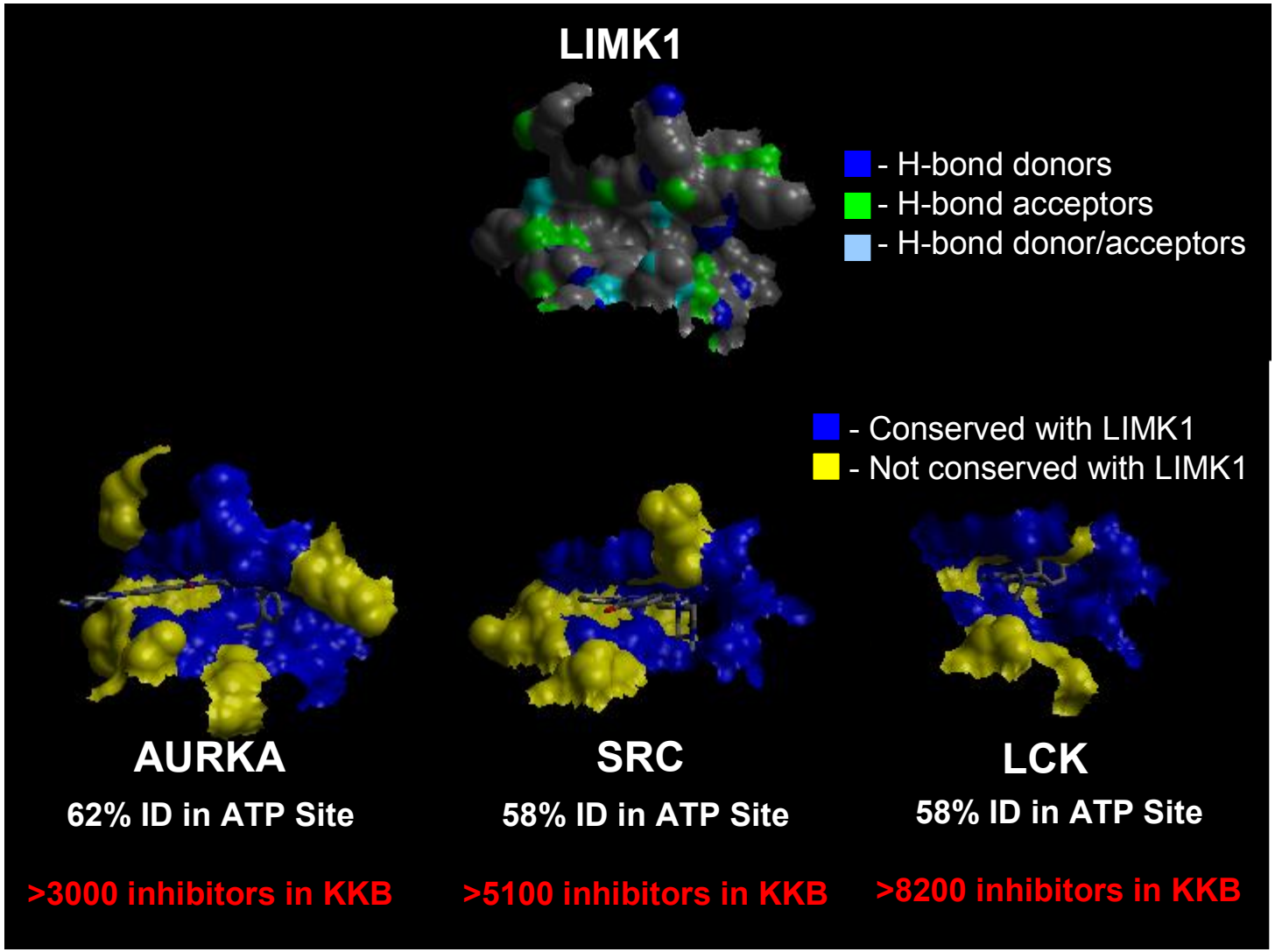
Pregnane X-receptor – PXR ("sensor)" →CYP3A4 ("executioner")
PXR Binds > 50% drugs
Including some bile acids, statins, herbal components, a selection of HIV protease inhibitors, calcium channel modulators, numerous steroids, plasticizers and monomers, organochlorine pesticides, a peroxisome proliferator-activated receptor-ãantagonist, xenobiotics and endobiotics…

Site Similarity Coloring



Eidogen-Sertanty

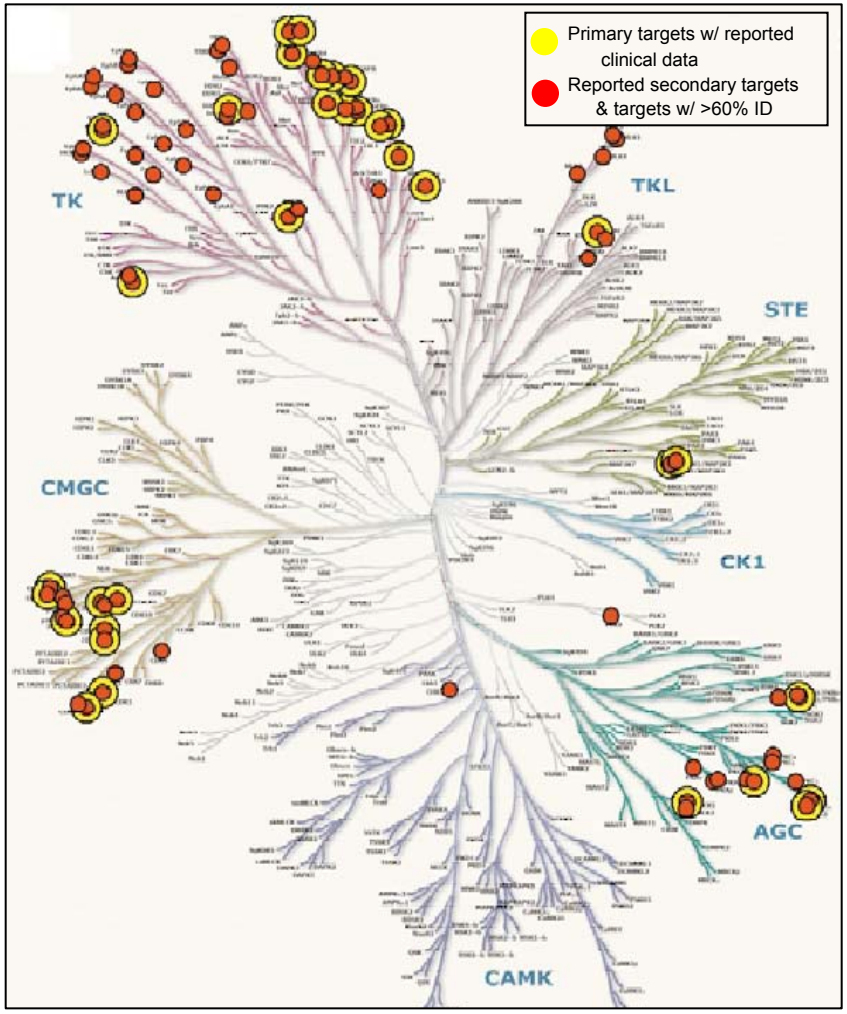
Borrowing Matter Ideas using Site Similarity



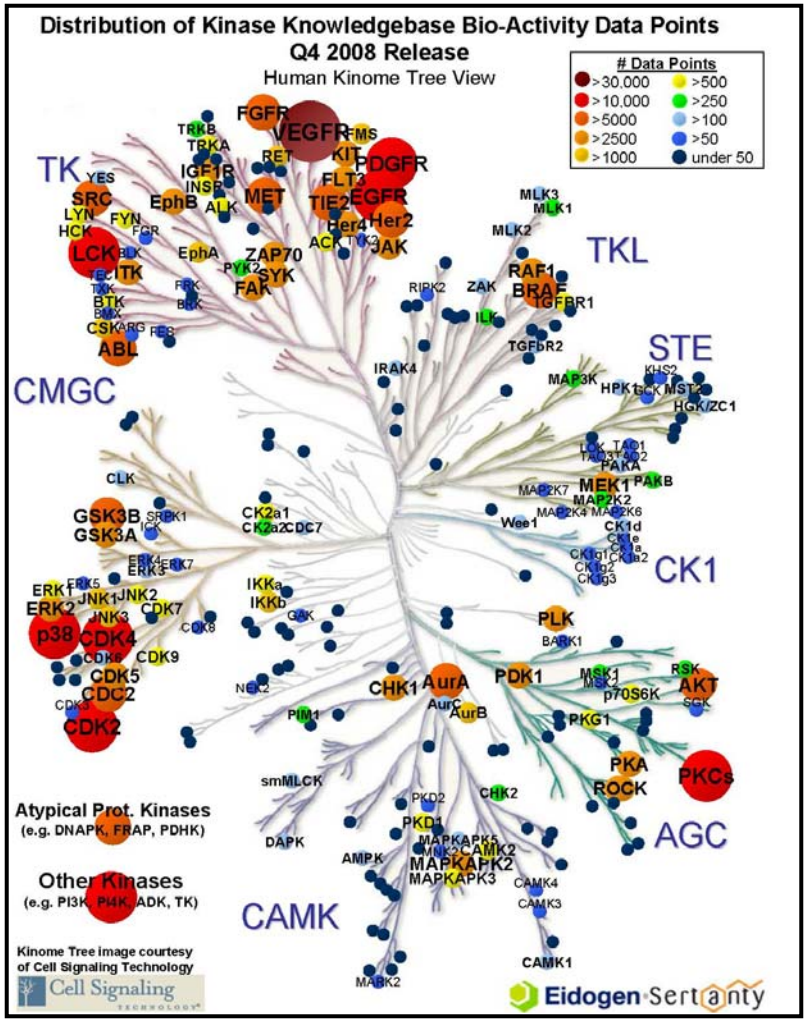
Kinase SAR Knowledgebase (KKB) – Hot Targets

Kinase Targets of Clinical Interest

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



Eidogen-Sertanty KKB
SAR Data Point Distribution



> 384,000 SAR data points curated from
> 5100 journal articles and patents

Eidogen Sertanty

Kinase Knowledgebase (KKB)

Kinase inhibitor structures and SAR data mined from

Kinase Validation Set

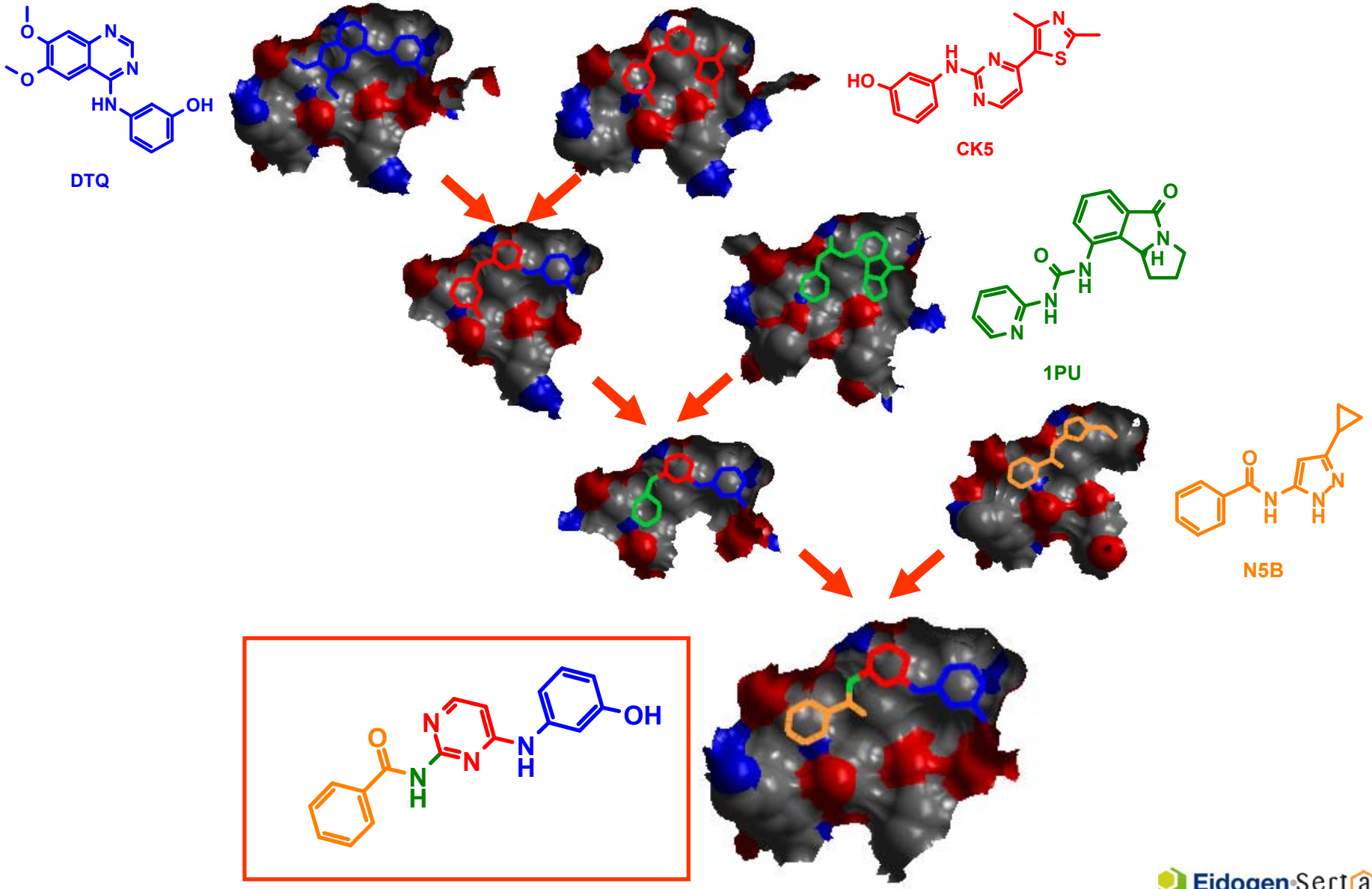
Three sizable datasets freely available to the research community

http://www.eidogen-sertanty.com/kinasednld.php

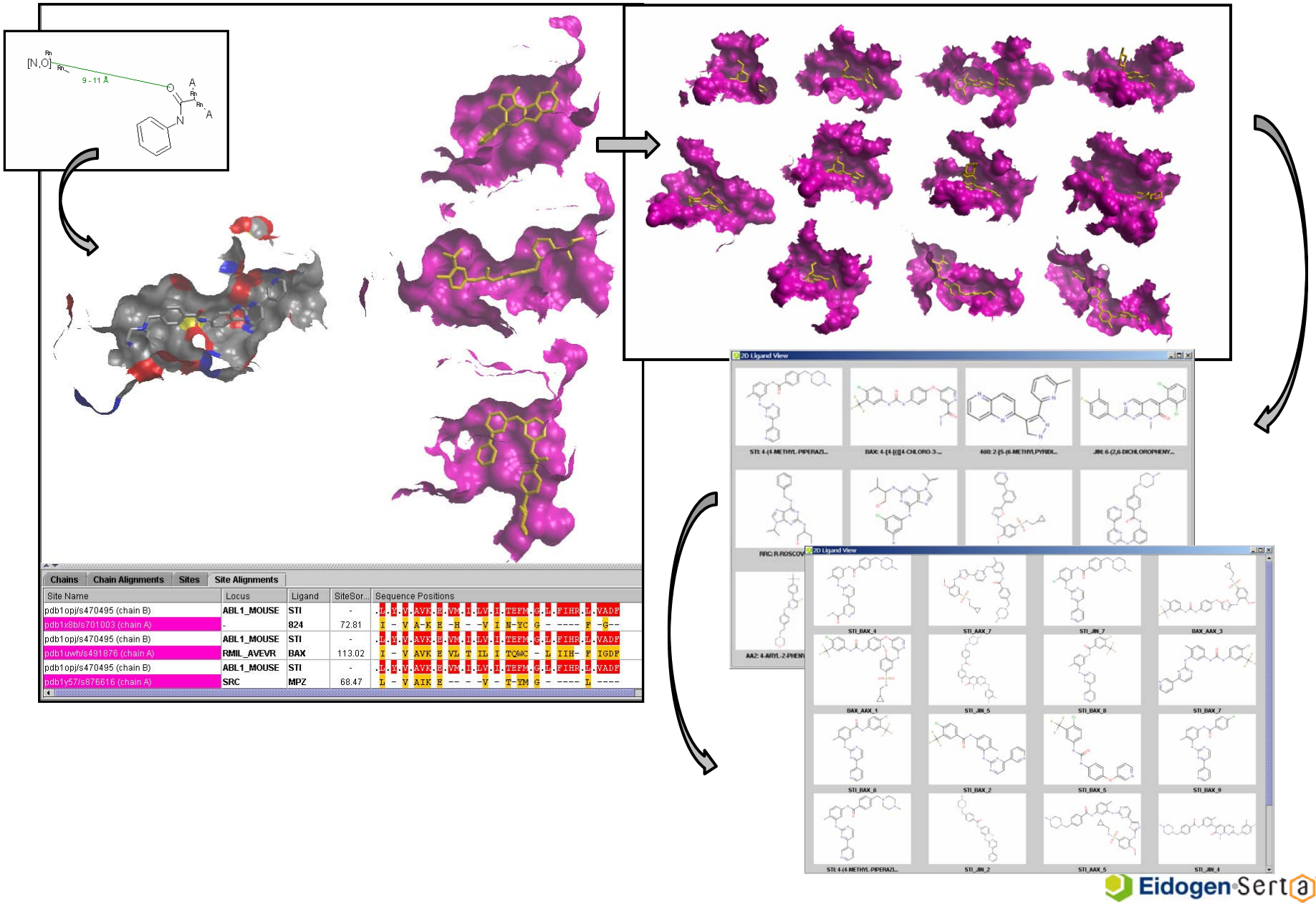
Eidogen-Sertanty

LigandCross: Shuffling Ligand Functionality

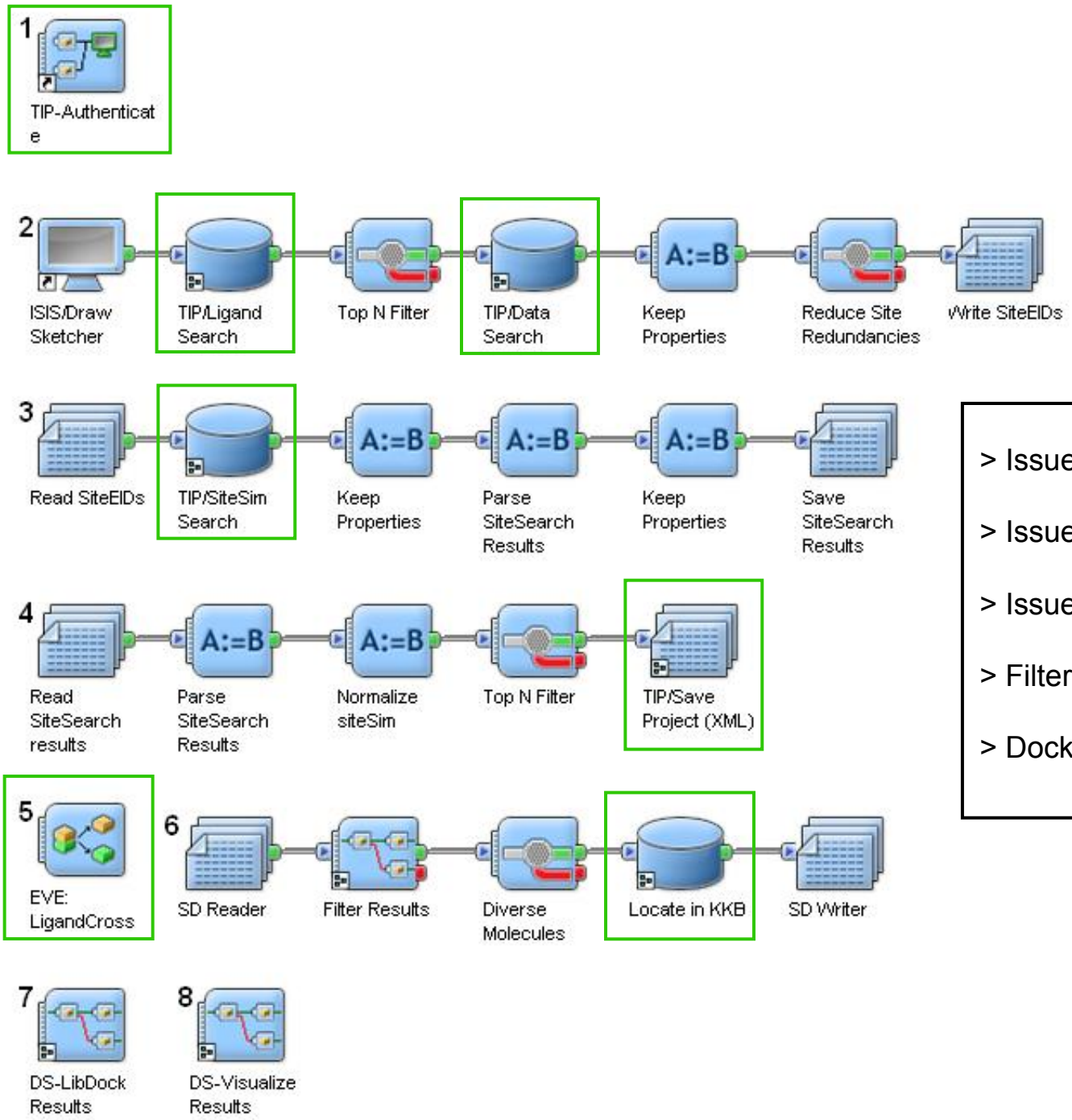
Similar to Vertex's BREED: J. Med. Chem. **47**, 2768 (2004)



From Ligand Query to Sites to New Ligand Ideas



Eidogen-Sertanty



> Issue TIP/LigandSearch

> Issue TIP/SiteSimSearch

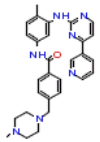
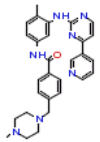
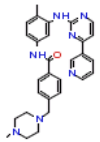
> Issue LigandCross

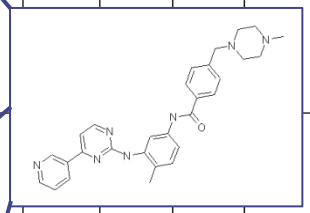
> Filter and locate results in KKB

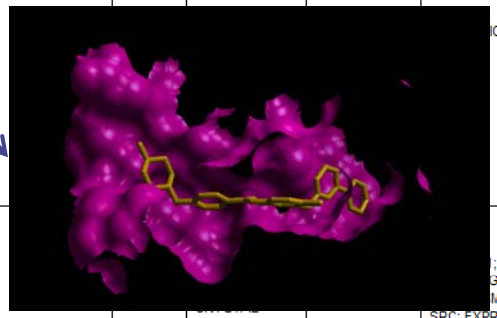
> Dock and visualize results

Eidogen Sertanty

Step 1: Find Co-complexes and Sites from Ligand-Structure-Search

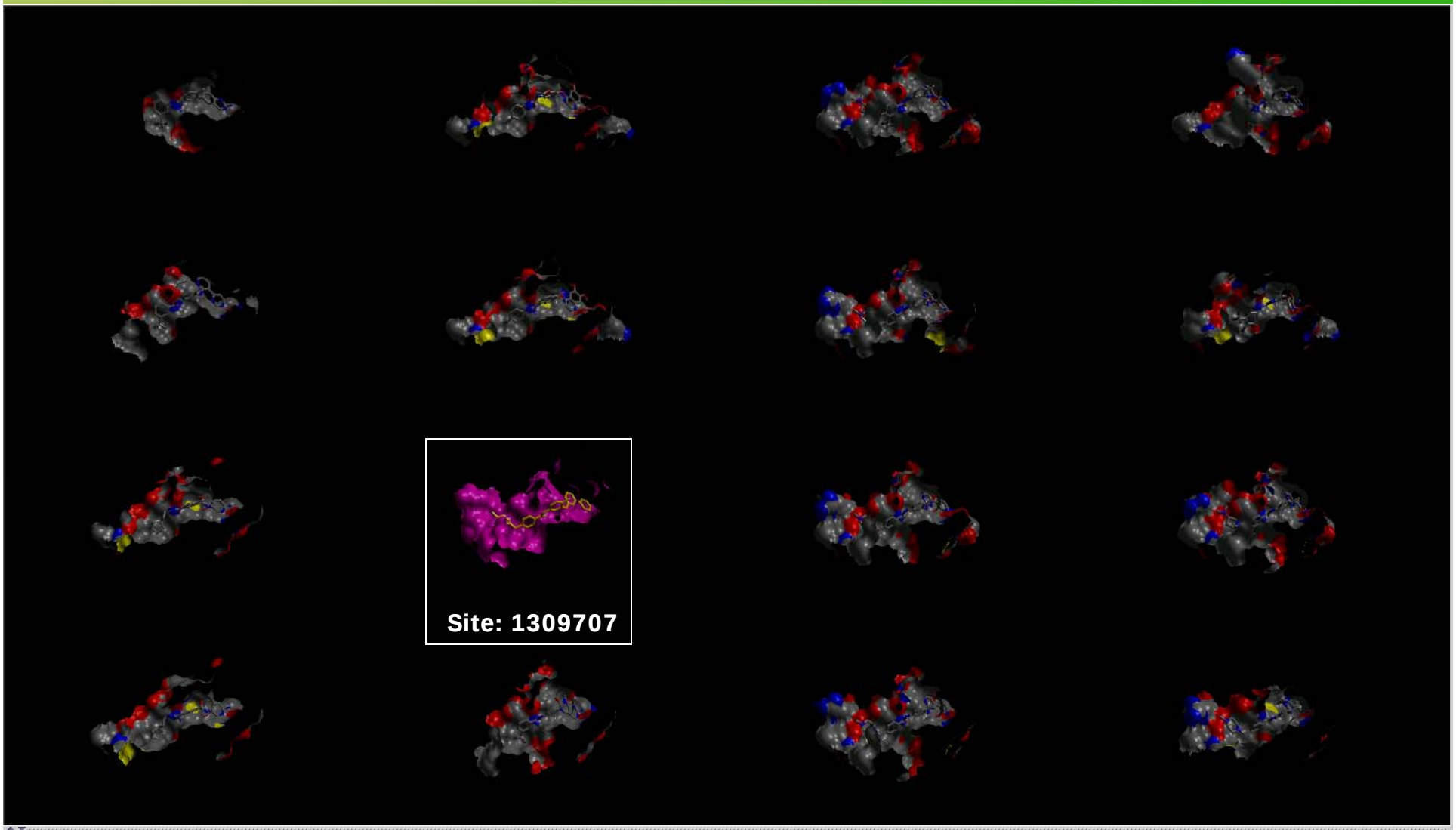
Molecule	ligname	similarity	pdbcode	siteeid	FourCode	pdbID	pdbBnxNumber	proteinId	title	classification	source	compound	releaseDate	journalTitle	journalReference	exptype
	STI	1	2pl0A	1309707	2pl0	2pl0	1305799	42526	LCK BOUND TO IMATINIB	TRANSFERASE	MOL_ID: 1; ORGANISM_SCIENTIFIC: HOMO SAPIENS; ORGANISM_COMMON: HUMAN; GENE: LCK; EXPRESSION_SYSTEM: SPODOPTERA FRUGIPERDA; EXPRESSION_SYSTEM_COMMON: FALL ARMYWORM; EXPRESSION_SYSTEM_VECTOR_TYPE: ION_SYSTEM_PLASMID:	MOL_ID: 1; MOLECULE: PROTO-ONCOGENE TYROSINE-PROTEIN KINASE LCK; CHAIN: A; FRAGMENT: PROTEIN KINASE; SYNONYM: P56-LCK, LYMPHOCYTE CELL-SPECIFIC PROTEIN-TYROSINE KINASE, LSK, T CELL-SPECIFIC PROTEIN-TYROSINE KINASE; EC: 2.7.10.2; ENGINEERED: YES	09-OCT-07	CLASSIFYING PROTEIN KINASE STRUCTURES GUIDES USE OF LIGAND-SELECTIVITY PROFILES TO PREDICT INACTIVE CONFORMATIONS: STRUCTURE OF LCK/IMATINIB COMPLEX	PROTEINS 2007	XRAY DIFFRACTION
	STI	1	2oiqA	1146914	2oiq	2oiq	1125109	26318	STRUCTURE OF CHICKEN C-SRC KINASE DOMAIN IN COMPLEX WITH THE CANCER DRUG IMATINIB.	TRANSFERASE	; ORGANISM_SCIENTIFIC: GALLUS; M_COMMON: CHICKEN; GENE: SRC; EXPRESSION_SYSTEM: ESCHERICHIA COLI; EXPRESSION_SYSTEM_COMMON: BACTERIA; EXPRESSION_SYSTEM_STRAIN: BL21DE3; EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID; EXPRESSION_SYSTEM_PLASMID: PET28	MOL_ID: 1; MOLECULE: PROTO-ONCOGENE TYROSINE-PROTEIN KINASE SRC; CHAIN: A, B; FRAGMENT: KINASE DOMAIN; SYNONYM: P60-SRC, C-SRC, PP60C-SRC; EC: 2.7.10.2; ENGINEERED: YES	20-MAR-07	C-SRC BINDS TO THE CANCER DRUG IMATINIB WITH AN INACTIVE ABL/C-KIT CONFORMATION AND A DISTRIBUTED THERMODYNAMIC PENALTY.	STRUCTURE V. 15 299 2007	XRAY DIFFRACTION
	STI	1	2hyyA	918207	2hyy	2hyy	904013	16961	HUMAN ABL KINASE DOMAIN IN COMPLEX WITH IMATINIB (STI571, GLIVEC)	TRANSFERASE	MOL_ID: 1; ORGANISM_SCIENTIFIC: HOMO SAPIENS; ORGANISM_COMMON: HUMAN; GENE: ABL1; EXPRESSION_SYSTEM: SPODOPTERA FRUGIPERDA; EXPRESSION_SYSTEM_COMMON: FALL ARMYWORM	MOL_ID: 1; MOLECULE: PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL1; CHAIN: A, B, C, D; SYNONYM: P150, C-ABL, ABELSON MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1; EC: 2.7.10.2;	16-JAN-07	STRUCTURAL BIOLOGY CONTRIBUTIONS TO THE DISCOVERY OF DRUGS TO TREAT CHRONIC MYELOGENOUS LEUKAEMIA.	ACTA CRYSTALLOGR.,SECT.D V. 63 80 2007	XRAY DIFFRACTION





Eidogen-Sertanty

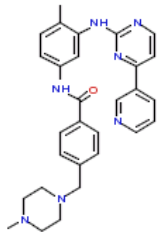
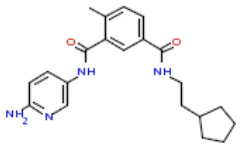
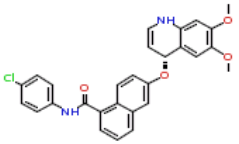
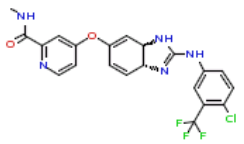
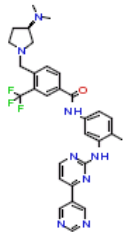
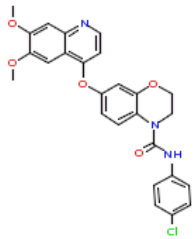
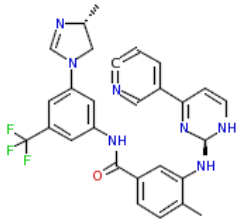
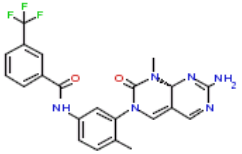
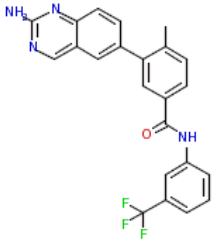
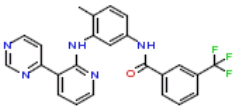
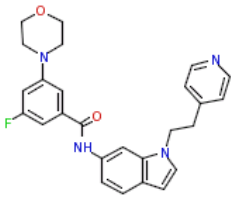
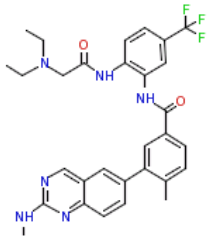
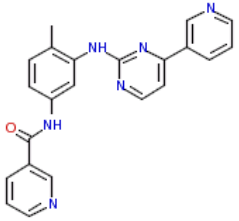
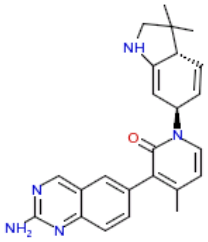
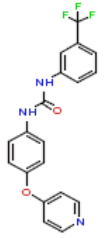
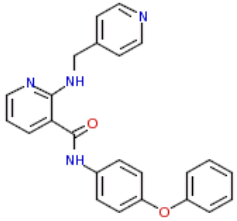
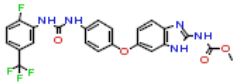
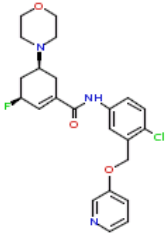
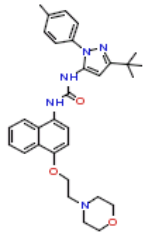
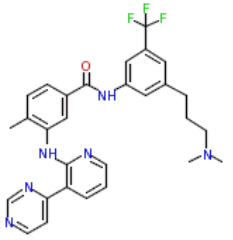
Step 2: Find Other Receptor Sites from Site-Similarity Search



Chains | Chain Alignments | Sites | Site Alignments

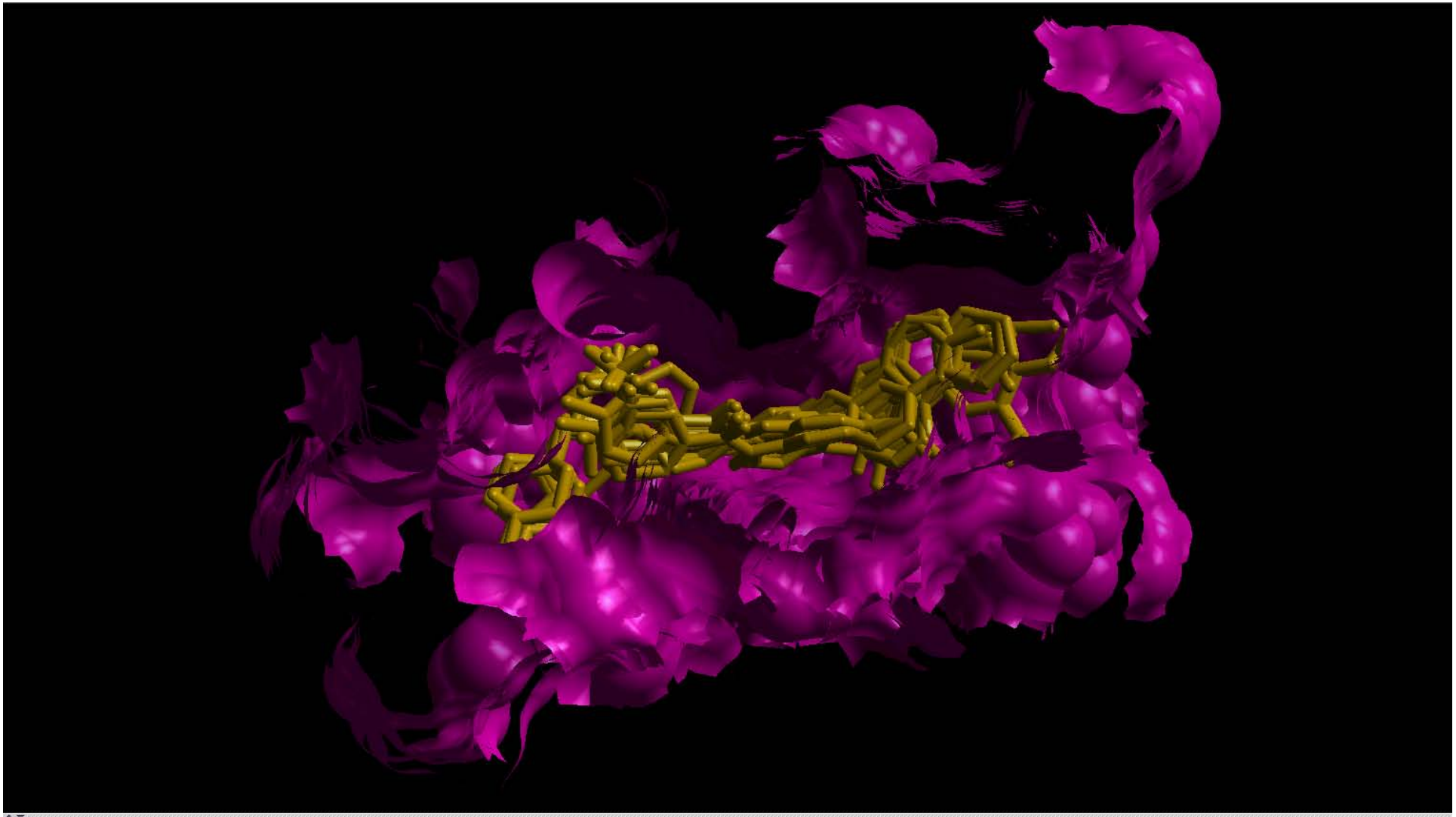
Site Name	Locus	Ligand	%Conf	Sequence Positions
pdb2pl0/s1309707 (chain A)	LCK	STI	100	.L.V.AVK.E.LM.L.LV.I.TEYM.GS.I.YIHR.L.IADF
pdb2ofw/s916548 (chain B)	LCK	242	100	.L.V.AVK.E.LM.L.LV.I.TEYM.G.I.Y.H.L.IADF.I
pdb2rl5/s1396160 (chain A)	-	2RL	100	.LG.V.AVK.L.E.LL.L.VV.V.TEFCKFGN.L.CIH.L.ICDF
pdb2e2b1/s1284639 (chain B)	ABL	406	100	.L.Y.V.A.K.E.VM.L.LV.I.TEFMT.G.L.FIHRD.L.VADF

Example Ligands Extracted from Similar Sites

 STI	 C92	 900	 276	 406
 2RL	 NIL	 7MP	 242	 MUH
 WBT	 1N8	 PRC	 C19	 KIN
 608	 GIG	 L11	 B96	 RAJ

Eidogen-Sertanty

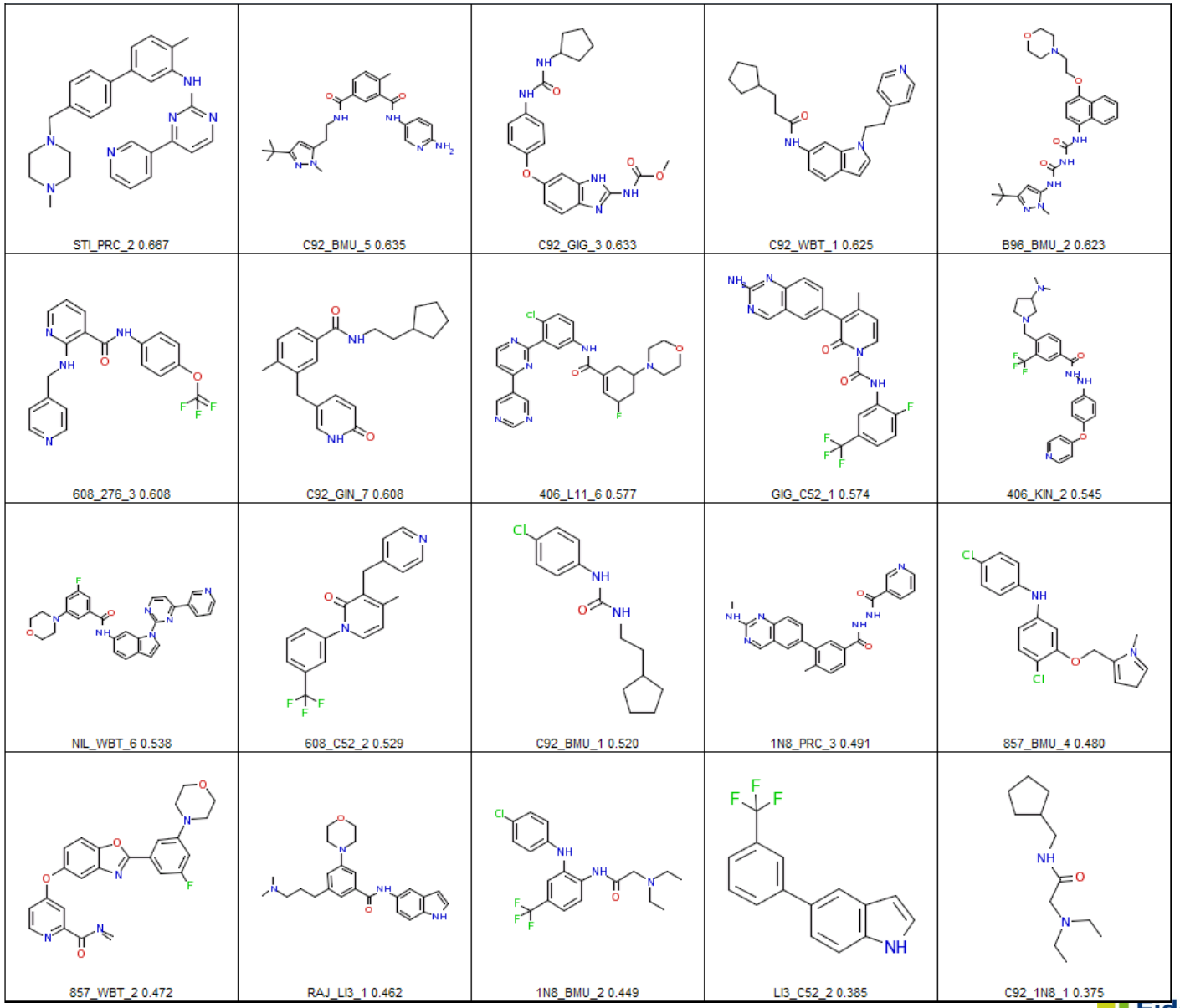
Step 3: LigandCross – Mixing Ligand Features from Aligned Sites



Chains | Chain Alignments | Sites | Site Alignments

Site Name	Locus	Ligand	%Conf	Sequence Positions
pdb2pl0/s1309707 (chain A)	LCK	STI	100	.L.V.AVK.E.LM.L.LV.I.TEYM.GS.I.YIHR.L.IADF
pdb2of/s916648 (chain B)	LCK	242	100	.L.V.AVK.E.LM.L.LV.I.TEYM.G.I.Y.H.L.IADF.I
pdb2rl5/s1396160 (chain A)	-	2RL	100	.LG.V.AVK.L.E.IL.I.VV.V.TEFCKFGN.L.CIH.L.ICDF
pdb2e2b1/s1284639 (chain B)	ABL	406	100	.L.Y.V.A.K.E.VM.I.LV.I.TEFMT.G.L.FIHRD.L.VADF

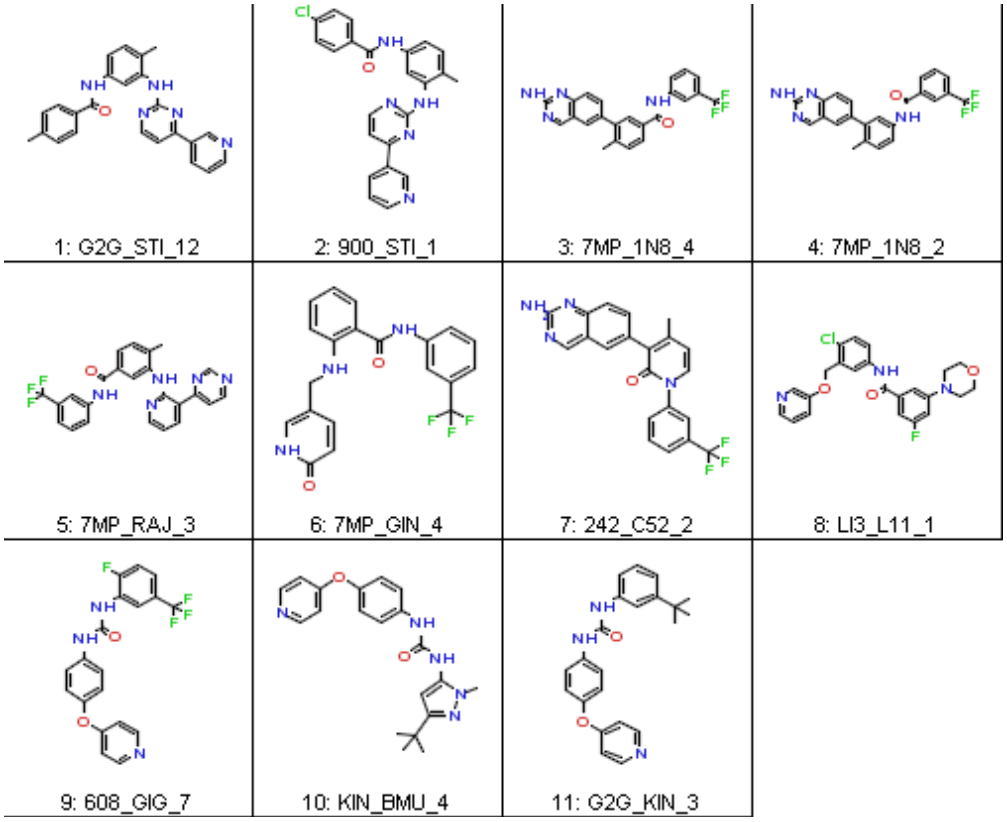
Example LigandCross Results



Eidogen-Sertanty

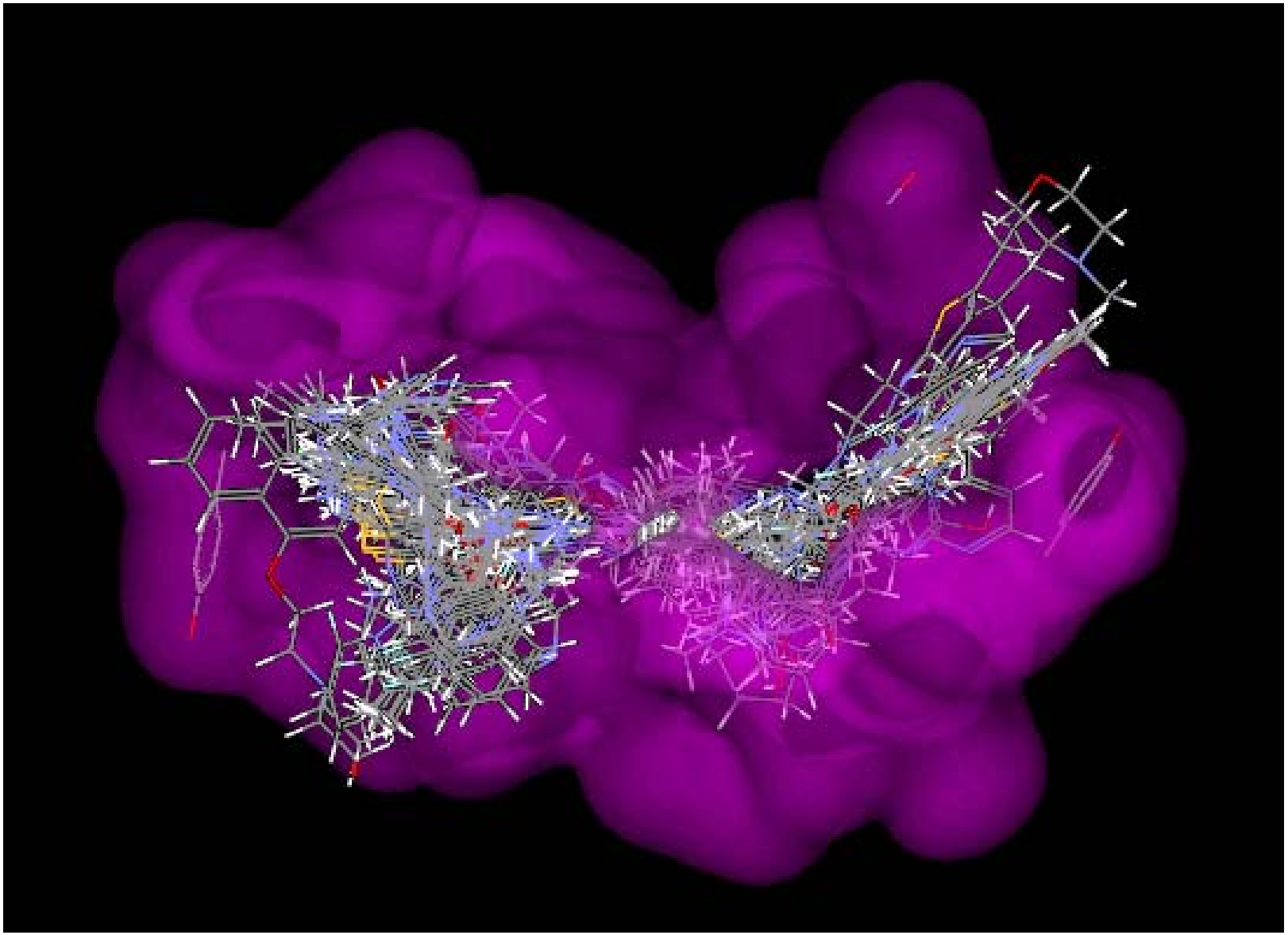
Step 4: LigandCross Ligands with Reported Biological Activity

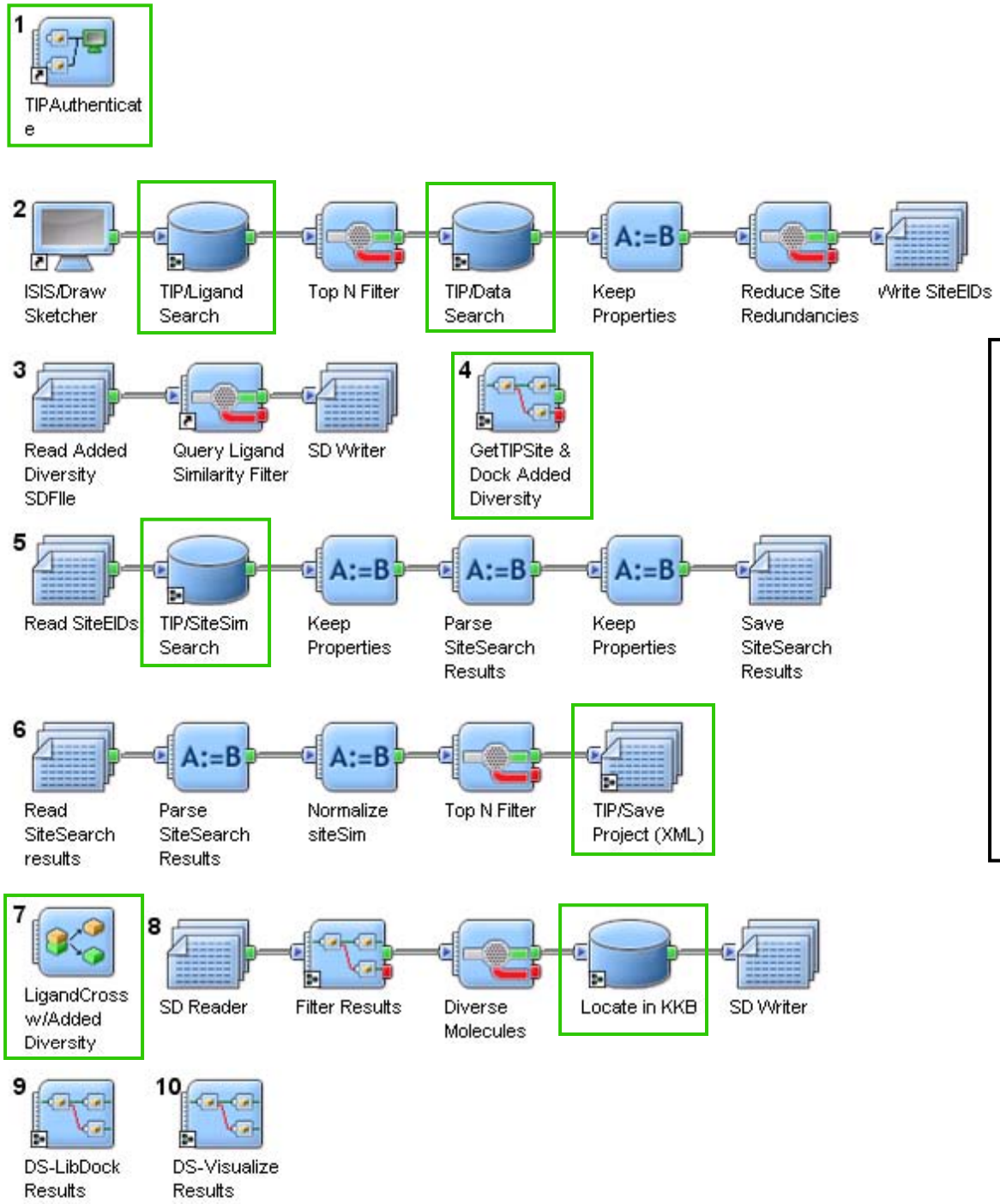
	Kinase Knowledgebase (pIC50)										Bayesian Model Predictions (PP)									
LC-ID	ABL	PDGFR	PDGFRB	JAK3	KDR	LCK	MAPK14	TEK	KIT	RAF1	ABL	PDGFR	PDGFRB	JAK3	KDR	LCK	MAPK14	TEK	KIT	RAF1
G2G_STI_12	6.7	8	8								0.40	0.90	0.76	0.81	0.59	0.15	0.89	0.45	0.70	0.37
900_STI_1	6.1	8	8								0.38	0.91	0.76	0.72	0.55	0.16	0.88	0.42	0.71	0.55
7MP_1N8_4				7.8	9	9.5	8.7				0.36	0.49	0.34	0.32	0.94	1.00	0.95	0.67	0.86	0.39
7MP_1N8_2				6.8	8.3	9.5	9				0.37	0.46	0.31	0.44	0.92	1.00	0.92	0.69	0.84	0.45
7MP_RAJ_3					8.4			8.4			0.35	0.73	0.50	0.49	0.92	0.81	0.86	0.94	0.74	0.37
7MP_GIN_4					7.6						0.16	0.50	0.40	0.82	0.95	0.67	0.70	0.41	0.76	0.51
242_C52_2									7.9		0.30	0.28	0.29	0.74	0.80	0.66	0.74	0.31	1.00	0.43
LI3_L11_1							7.2				0.31	0.73	0.55	0.84	0.74	0.69	0.62	0.36	0.76	0.85
608_GIG_7										6.1	0.28	0.61	0.57	0.69	0.93	0.50	0.60	0.68	0.85	0.50
KIN_BMU_4										6.1	0.31	0.43	0.45	0.78	0.75	0.57	0.77	0.33	0.81	0.25
G2G_KIN_3										6.1	0.25	0.51	0.52	0.75	0.89	0.59	0.64	0.43	0.84	0.43



Eidogen-Sertanty

Step 5: LigandCross Ligands reDocked into s1309707





> Issue TIP/LigandSearch

> Identify/Dock “AddedDiversity”

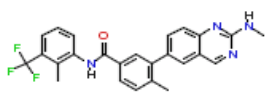
> Issue TIP/SiteSimSearch

> LigandCross w/AddedDiversity

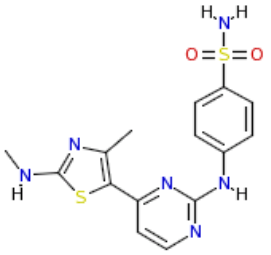
> Filter and locate results in KKB

> Dock and visualize results

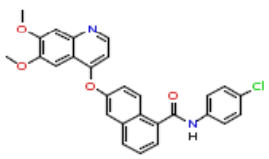
Example Potent Kinase Inhibitors ("Added Diversity")



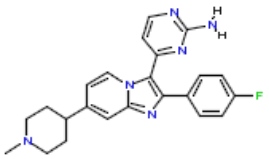
4336533 LCK pval: 11.00



4302493 CDK9 pval: 10.54

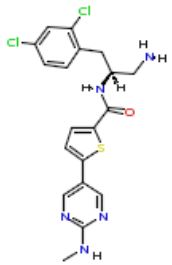


4332561 KDR pval: 10.52

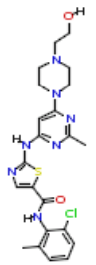


4318145 PKG pval: 10.40

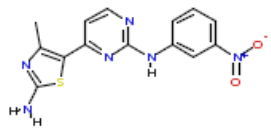
Chiral



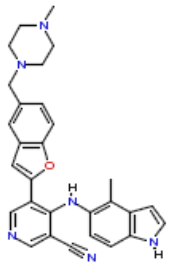
4336686 PKA pval: 10.00



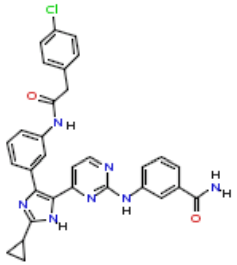
4272835 ABL1 pval: 10.00



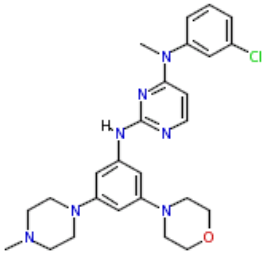
894611 CDK2 pval: 9.70



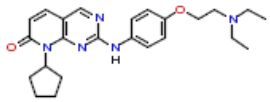
4358565 PRKCQ pval: 9.70



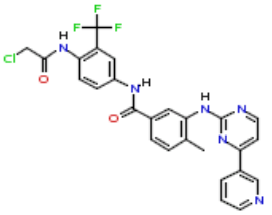
4363734 RAF1 pval: 9.30



4369892 EPHB4 pval: 9.24

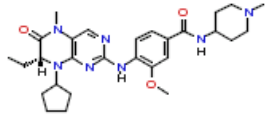


809 CDK4 pval: 9.15

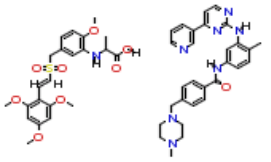


4374385 PDGFRA pval: 9.14

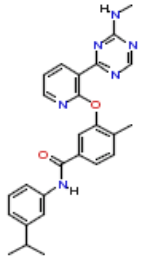
Chiral



4366691 PLK1 pval: 9.10

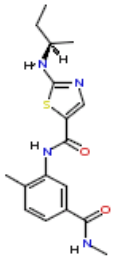


4301886 BCR_ABL pval: 9.08

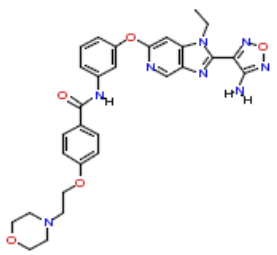


4307551 TEK pval: 9.00

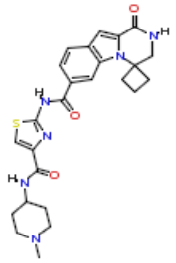
Chiral



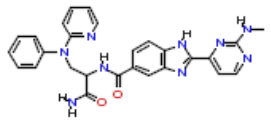
4363016 MAPK11 pval: 8.82



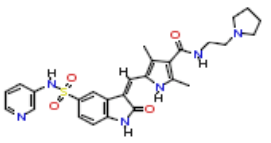
4343448 ROCK1 pval: 8.74



4363247 MAPKAPK2 pval: 8.70

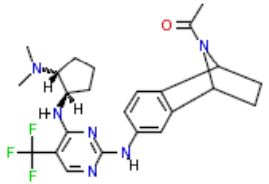


4291996 IKB pval: 8.70



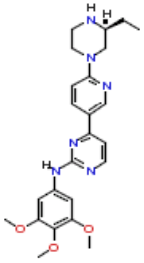
4208857 FAK2 pval: 8.22

Chiral

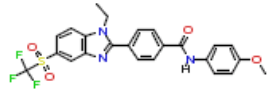


4373725 PTK2B pval: 8.22

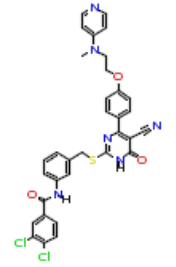
Chiral



1788 ZAP70 pval: 8.10

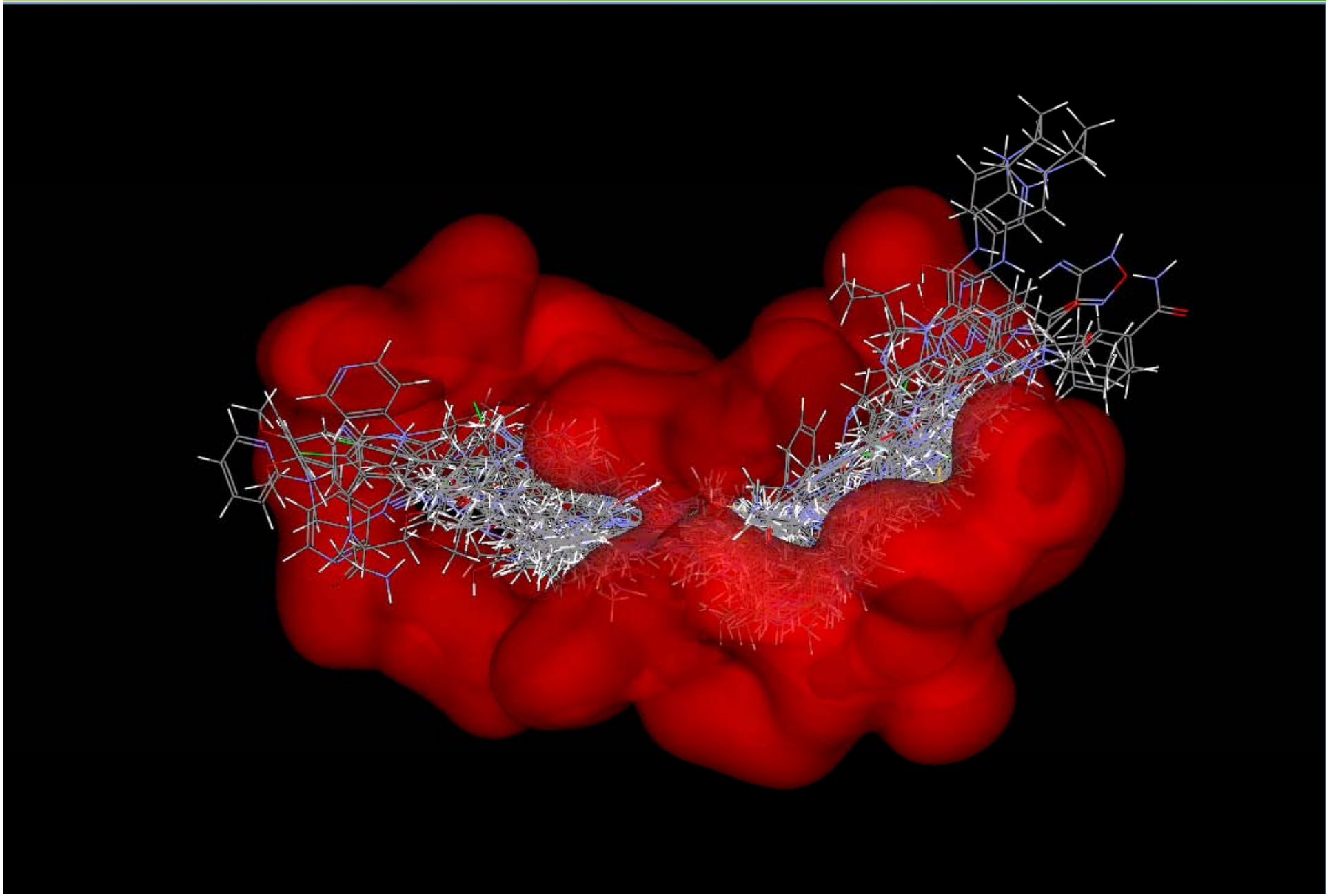


2425813 PTPN9 pval: 5.96

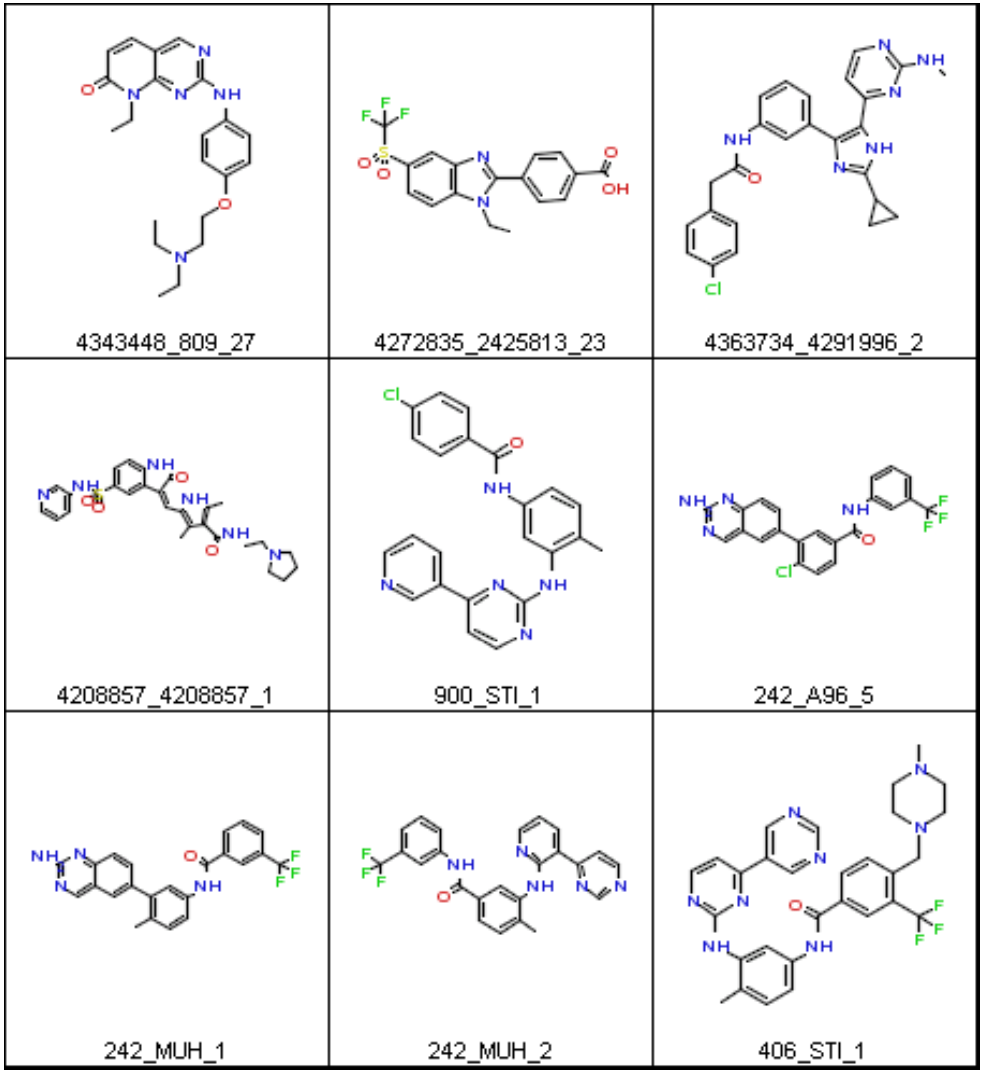


4303129 MAP3K2 pval: 4.70

Potent Kinase Inhibitors Docked (s1309707)



LigandCross Examples using “Added Diversity”



4343448_809_27:
CDK4: 6.80 CDK2: 5.63 CDK2: 6.12 CDC2: 5.58 CSK: 5.99 CDK5: 6.81
CDK4: 6.80 CDK2: 5.63 CDK2: 6.12 CDC2: 5.58 CDK4: 6.80

4272835_2425813_23:
PTPN1: 4.24 PTPRA: 4.21

4363734_4291996_2:
RAF1: 9.00 MAPK1: 5.29 BRAF: 8.05 BRAF: 8.52

4208857_4208857_1:
FAK2: 8.22 KDR: 5.86 PDGFRB: 4.90 EGFR: 4.17 ERBB2: 5.23

900_STI_1:
PDGFR: 8.00 PDGFR: 8.00 ABL: 6.10 PDGFRB: 8.00 PDGFR: 8.00
ABL: 6.10

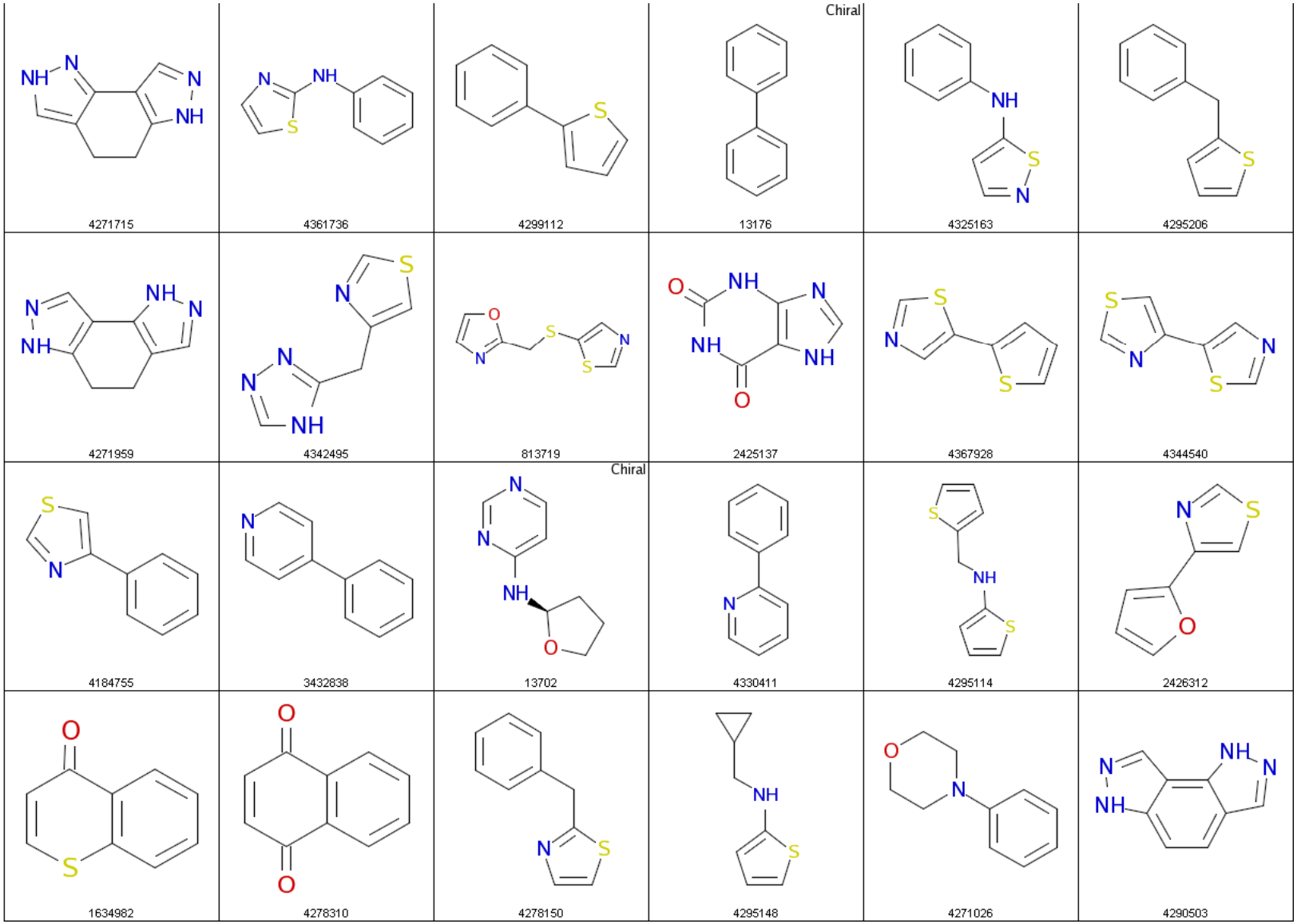
242_A96_5:
LCK: 9.40

242_MUH_1:
LCK: 9.40 TEK: 7.68 KDR: 8.22 MAPK14: 9.00 JAK3: 6.81

242_MUH_2:
KDR: 8.40 TEK: 8.40 TEK: 8.40 KDR: 8.40 TEK: 8.40 KDR: 8.40

406_STI_1:
BCR_ABL: 8.40 BCR_ABL: 5.30 LYN: 8.06 ABL1: 8.07 ABL1: 8.40

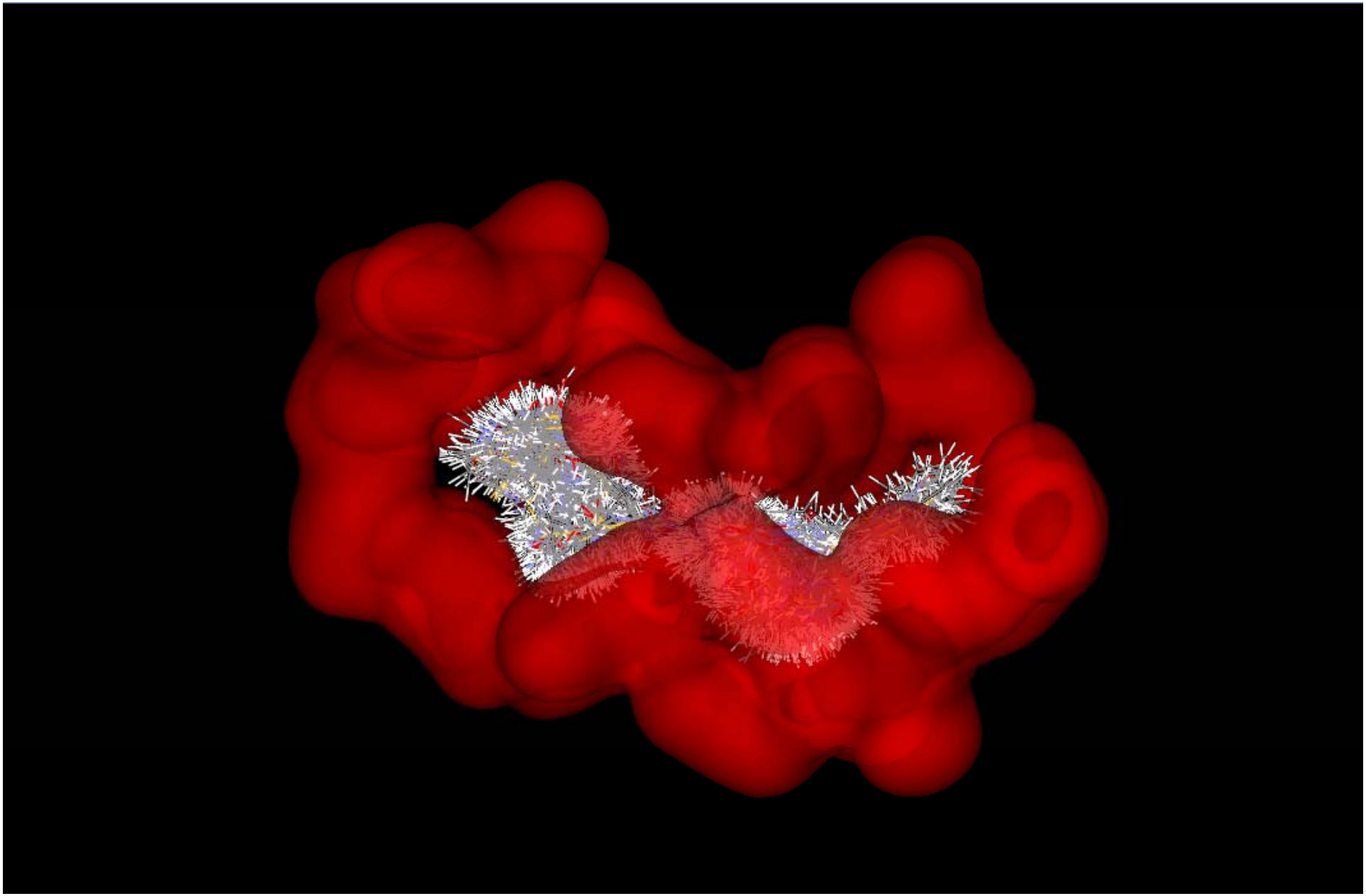
Murcko Assemblies Found in Kinase Inhibitors



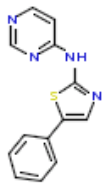
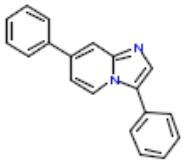
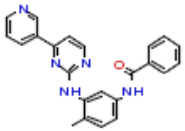
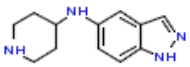
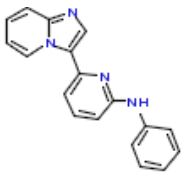
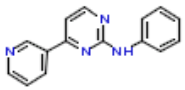
Murcko Assemblies: Contiguous ring systems plus chains that link two or more rings

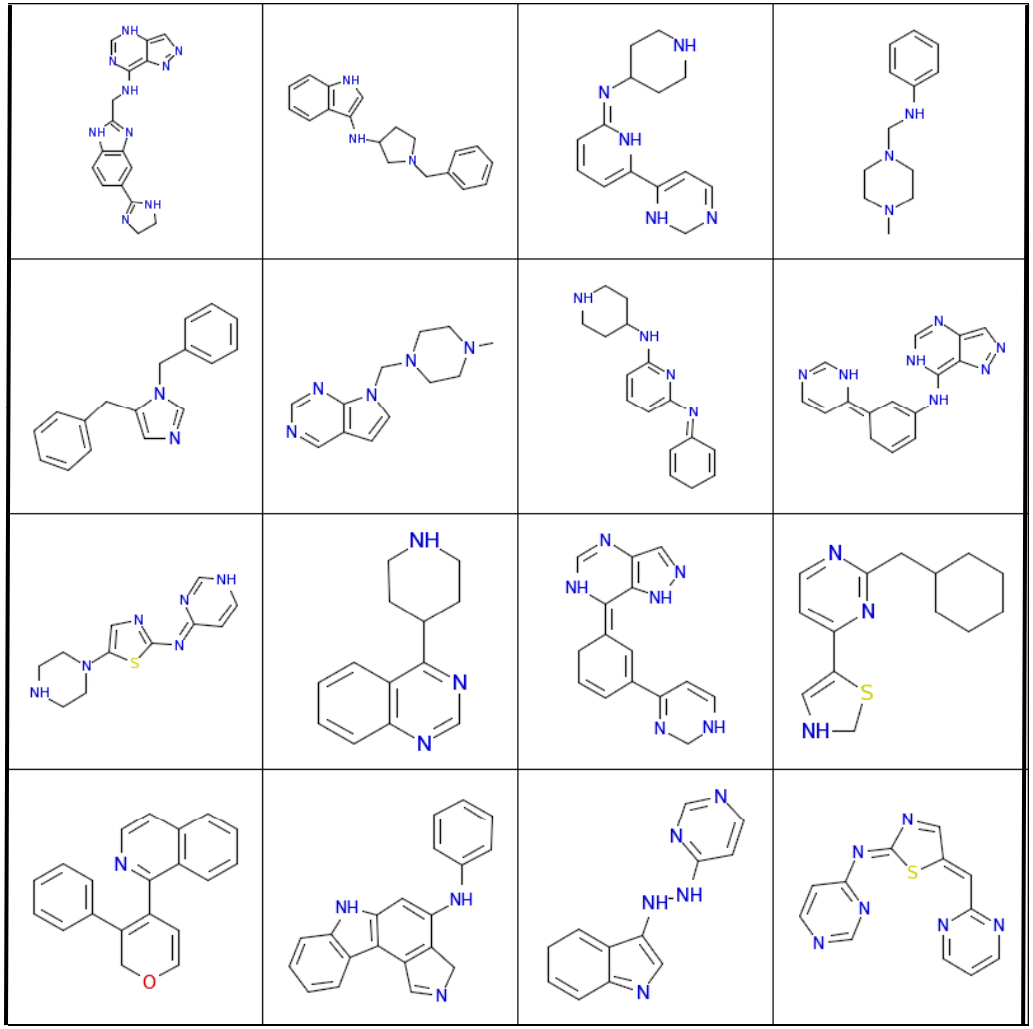
"The Properties of Known Drugs. 1. Molecular Frameworks", Guy W. Bemis and Mark A. Murcko, *J. Med. Chem.* 1996, 39, 2887-2893.

Positional Murcko Assemblies (parent inhibitors docked into s1309707)



LigandCross Results: Positional Murcko Assemblies
from docked Kinase inhibitors (s1309707)

	KDR KDR	Enzyme Assay Enzyme Assay	7.4437 7.4437
	KDR	Enzyme Assay	7.0088
	PDGFR PRKCA PRKCA ABL EGFR PDGFR PDGFR PDGFRB ABL PDGFRB PDGFRB PDGFRA ABL PDGFRB	Enzyme Assay Enzyme Assay Enzyme Assay Enzyme Assay Enzyme Assay Enzyme Assay Enzyme Assay Cell-Based Assay Enzyme Assay Cell-Based Assay Enzyme Assay Enzyme Assay Enzyme Assay Cell-Based Assay	7 4.1427 4.1427 6.3979 4.1871 7 7 7.1871 6.3979 6.2218 7 5.2218 6.3979 6.7696
	ROCK ROCK1	Enzyme Assay Enzyme Assay	6.5421 6.5229
	IRAK4	Enzyme Assay	5.9370
	PRKCA PRKCD ABL EGFR	Enzyme Assay Enzyme Assay Enzyme Assay Enzyme Assay	4.9788 4.4089 5.7447 4



Kinase Activity ????

Eidogen-Sertanty

Conclusions

- Significant receptor-site similarities exist within and across target families

- The structurally resolved and modelable proteome is a very rich source for new matter ideas

- LigandCross can be an effective approach to generating novel, bioactive matter using co-complexes, known inhibitors, and/or fragment-based information.

Acknowledgements

- Stephan Schürer
- Kevin Hambly
- Joe Danzer
- Brian Palmer
- Derek Debe
- Aleksandar Poleksic

- Accelrys/Scitegic - Shikha Varma-O'Brien/Ton van Daelen

About Eidogen-Sertanty

- Knowledge-Driven Solutions Provider
  - Sertanty established in 2003, acquired Libraria assets
  - Sertanty acquired Eidogen/Bionomix in 2005→ Eidogen-Sertanty
  - $20M invested: Libraria ($6M), Eidogen/Bionomix ($12M), Sertanty/ES ($2M)
  - 14 distributed FTE's (4 US and 10 India)
  - Worldwide (bio)pharmaceutical customer base
  - Cash-positive since 2006

- Databases & Software – Annual Subscriptions
  - *TIP*™ - Protein Structural Informatics Platform
  - *KKB*™ - Kinase SAR and Chemistry Knowledgebase
  - CHIP™ - Chemical Intelligence Platform

- DirectDesign™ Fee-For-Service
  - In Silico Target Screening ("Target Fishing" and Repurposing)
  - Target and compound prioritization services
  - Fast Follower Design: Novel, Patentable Leads

Eidogen·Sertanty