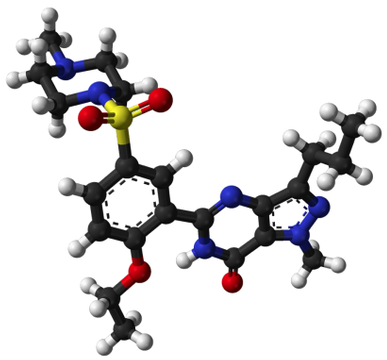
nAnnolyze : ligand-target prediction by structural network biology

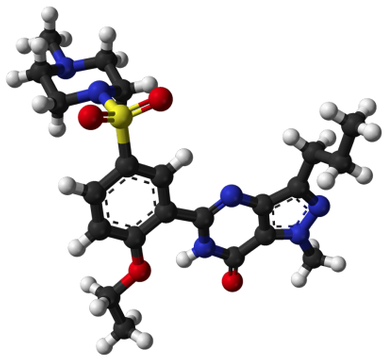
Francisco Martínez-Jiménez

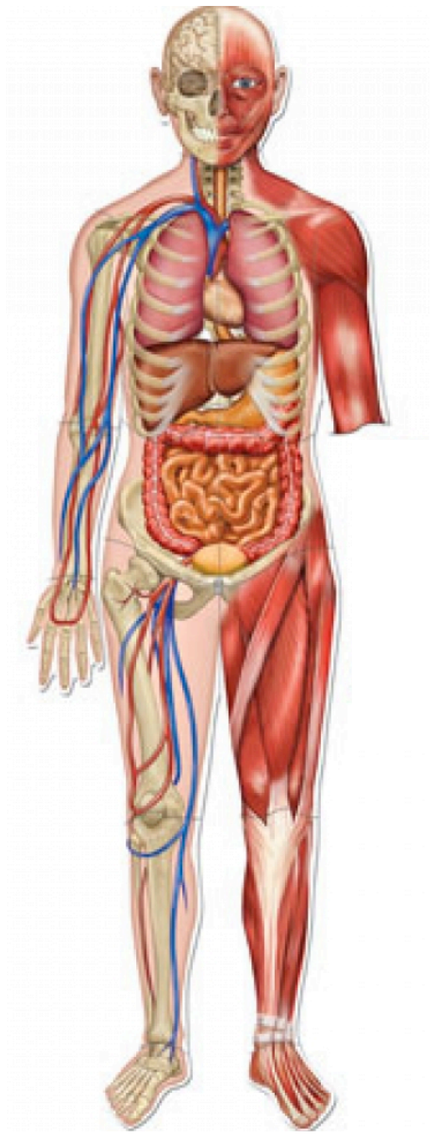
Student Symposium, JdBI2014, Sevilla

Finding out drugs mode of action...

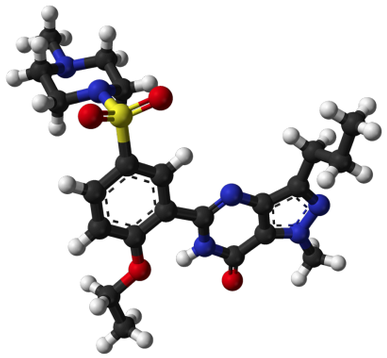


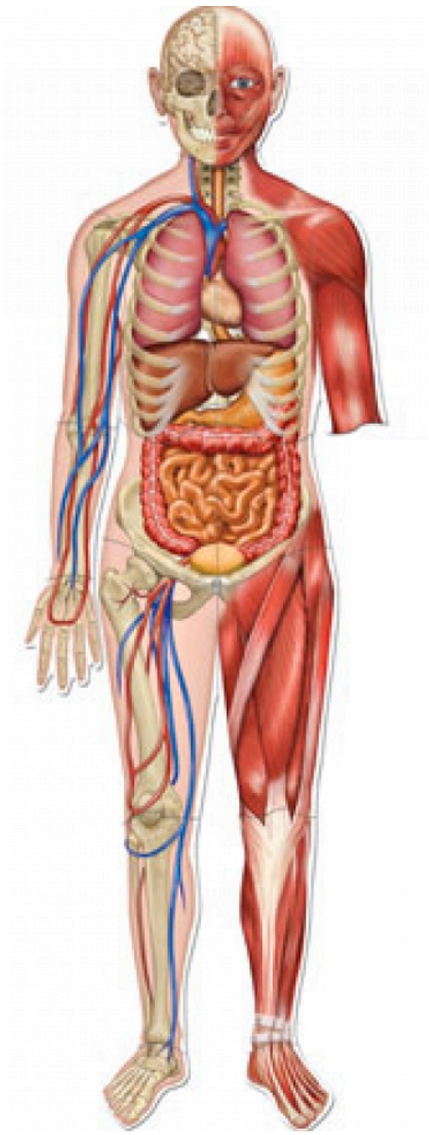
Finding out drugs mode of action...



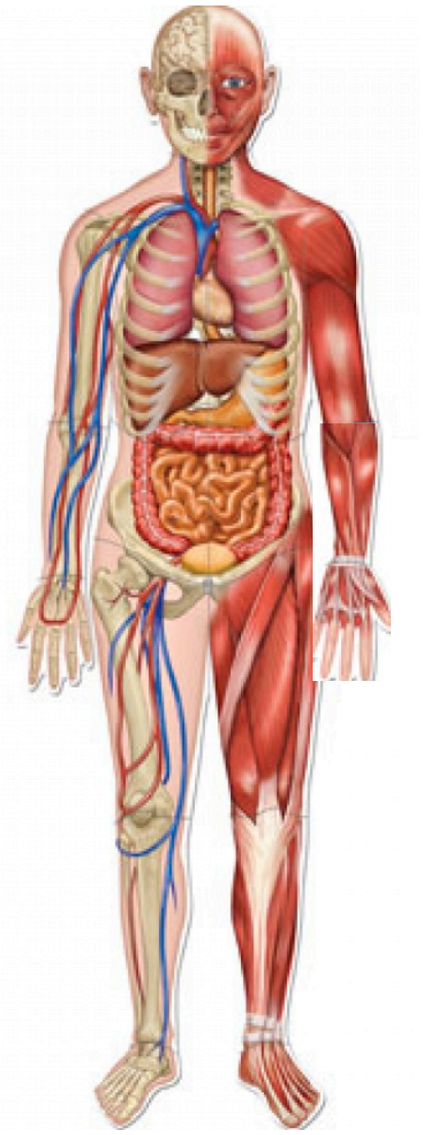


Finding out drugs mode of action...

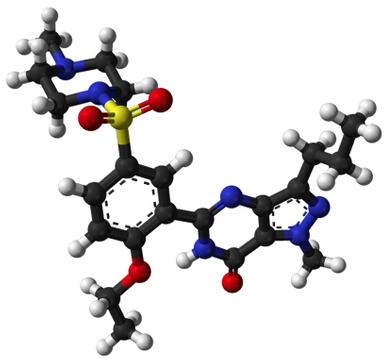




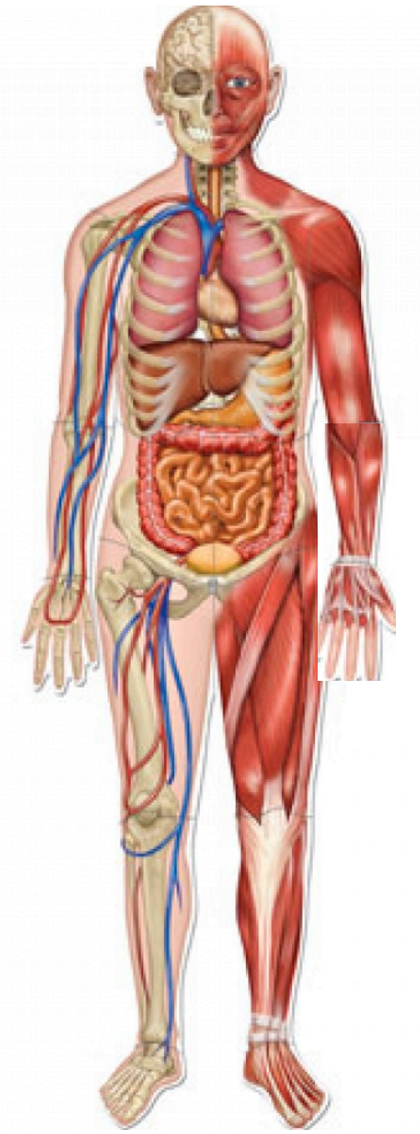
Phenotype



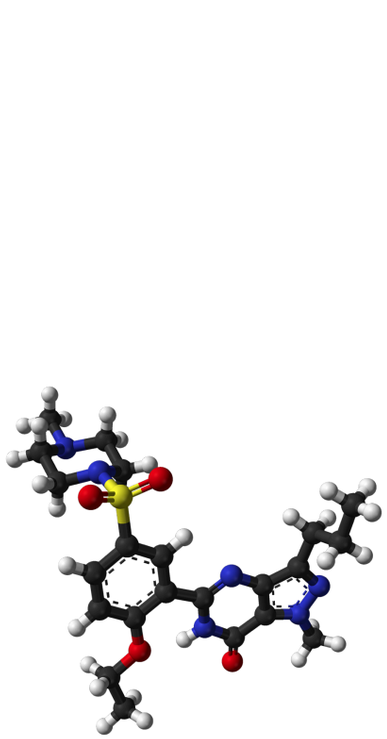
Finding out drugs mode of action...



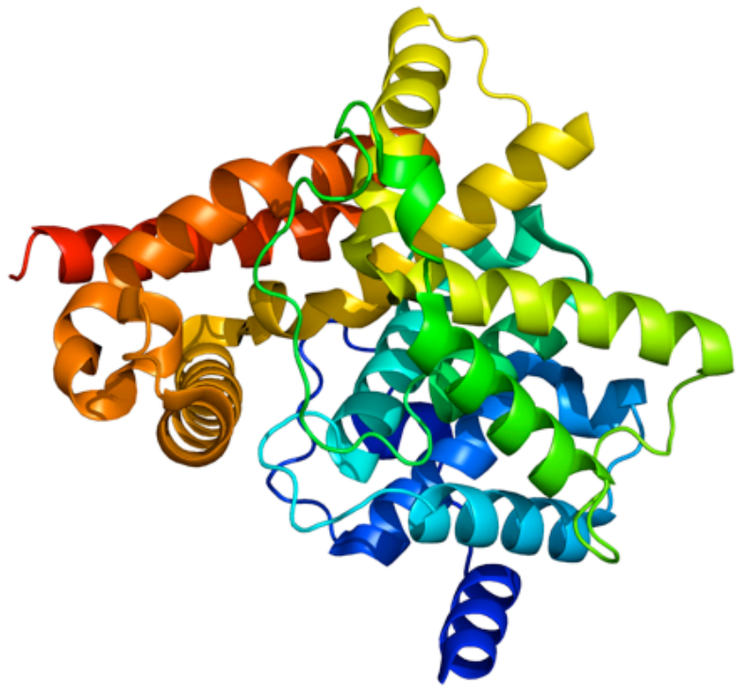
Phenotype



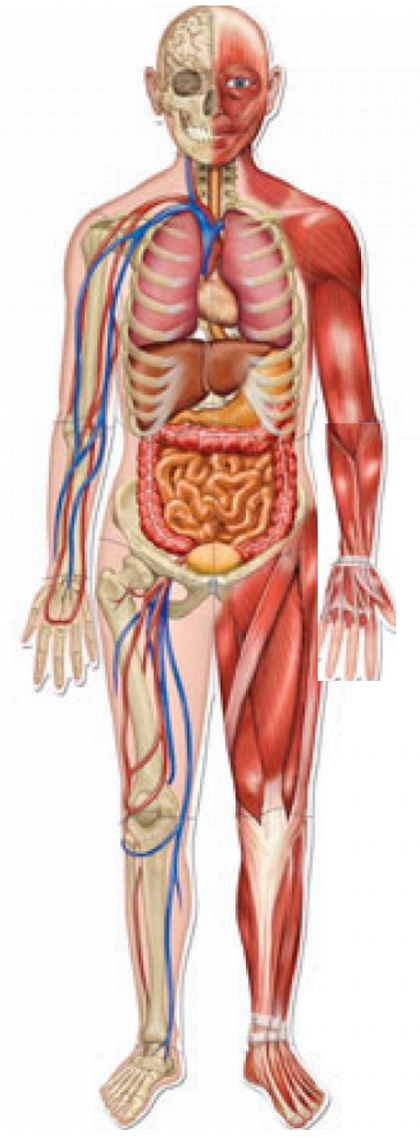
Finding out drugs mode of action...



?



Phenotype



Existing computational methods

Prediction details & accuracy

Computational time

Existing computational methods

Prediction details & accuracy

free structure methods

★Based on previous knowledge.
★Many different methods.
★Good performance.
★Poor information about the interaction.

Computational time

Existing computational methods

Prediction details & accuracy

free structure methods

★Based on previous knowledge.
★Many different methods.
★Good performance.
★Poor information about the interaction.

structure based methods

Virtual Docking

★Very precise. Ligand and receptor orientation.
★Needs the binding-site.
★Needs the structure or a reliable 3D-model.
★Not applicable at wide scale.

Computational time

Existing computational methods

Prediction details & accuracy

free structure methods

★Based on previous knowledge.
★Many different methods.
★Good performance.
★Poor information about the interaction.

structure based methods

Comparative Docking

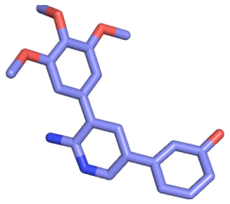
★Outputs binding-site localization.
★Based on structural comparisons.
★Applicable at wide scale.
★Needs the structure or a reliable 3D-model.

Virtual Docking

★Very precise. Ligand and receptor orientation.
★Needs the binding-site.
★Needs the structure or a reliable 3D-model.
★Not applicable at wide scale.

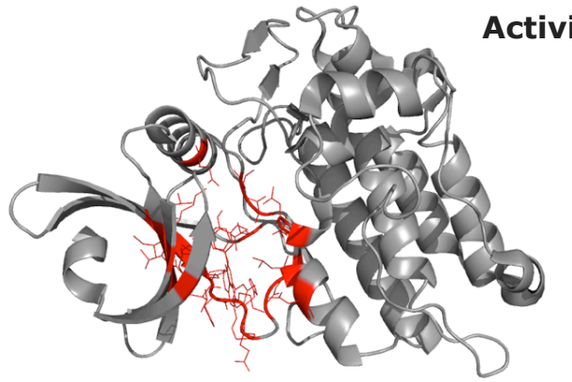
Computational time

Comparative Docking



A3F

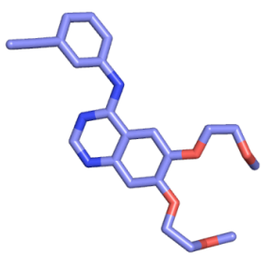
co-crystallized



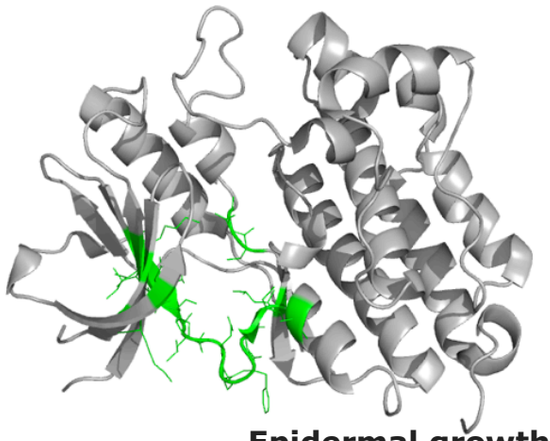
Activin receptor type-1

Similar binding-sites

AQ4

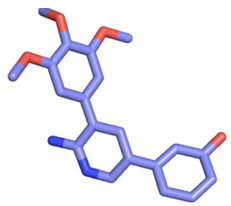


co-crystallized



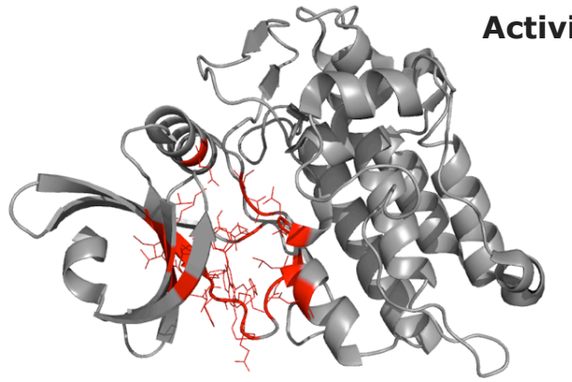
Epidermal growth factor receptor

Comparative Docking

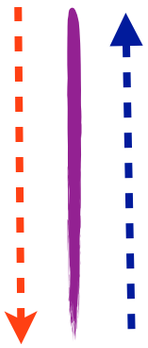


A3F

co-crystallized

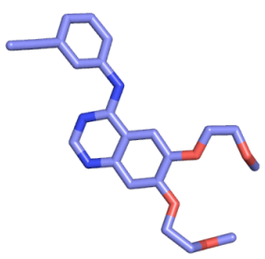


Activin receptor type-1

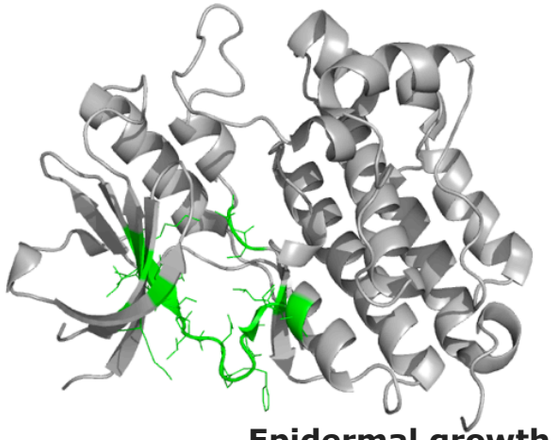


Similar binding-sites

AQ4

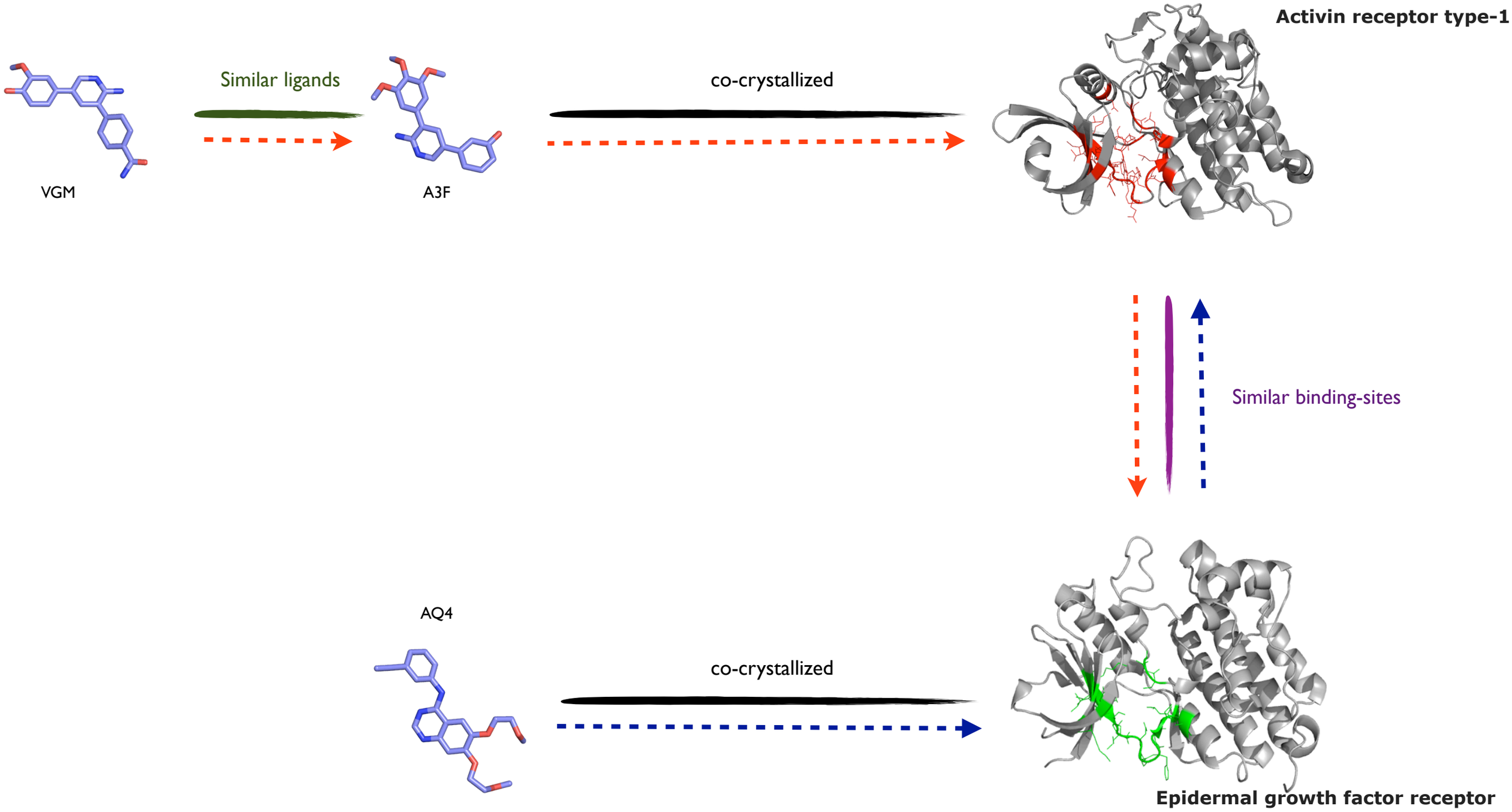


co-crystallized

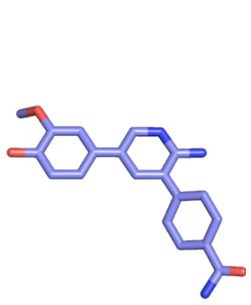


Epidermal growth factor receptor

Comparative Docking

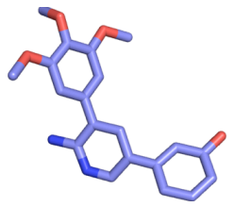


Comparative Docking



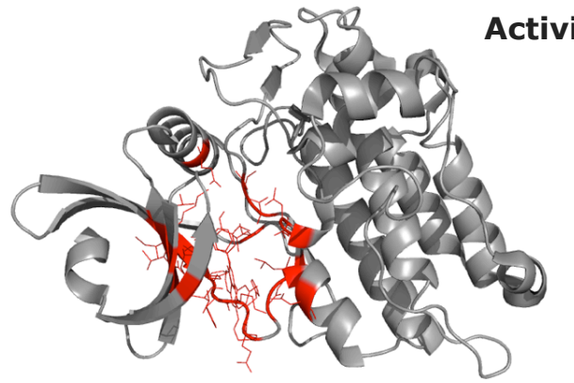
VGM

Similar ligands



A3F

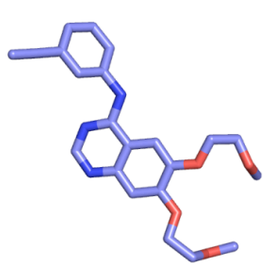
co-crystallized



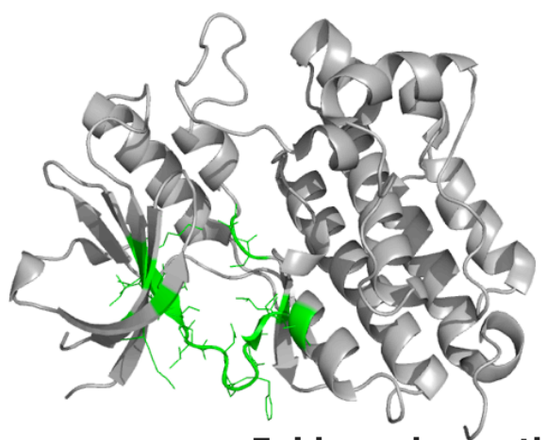
Activin receptor type-1

Similar binding-sites

AQ4



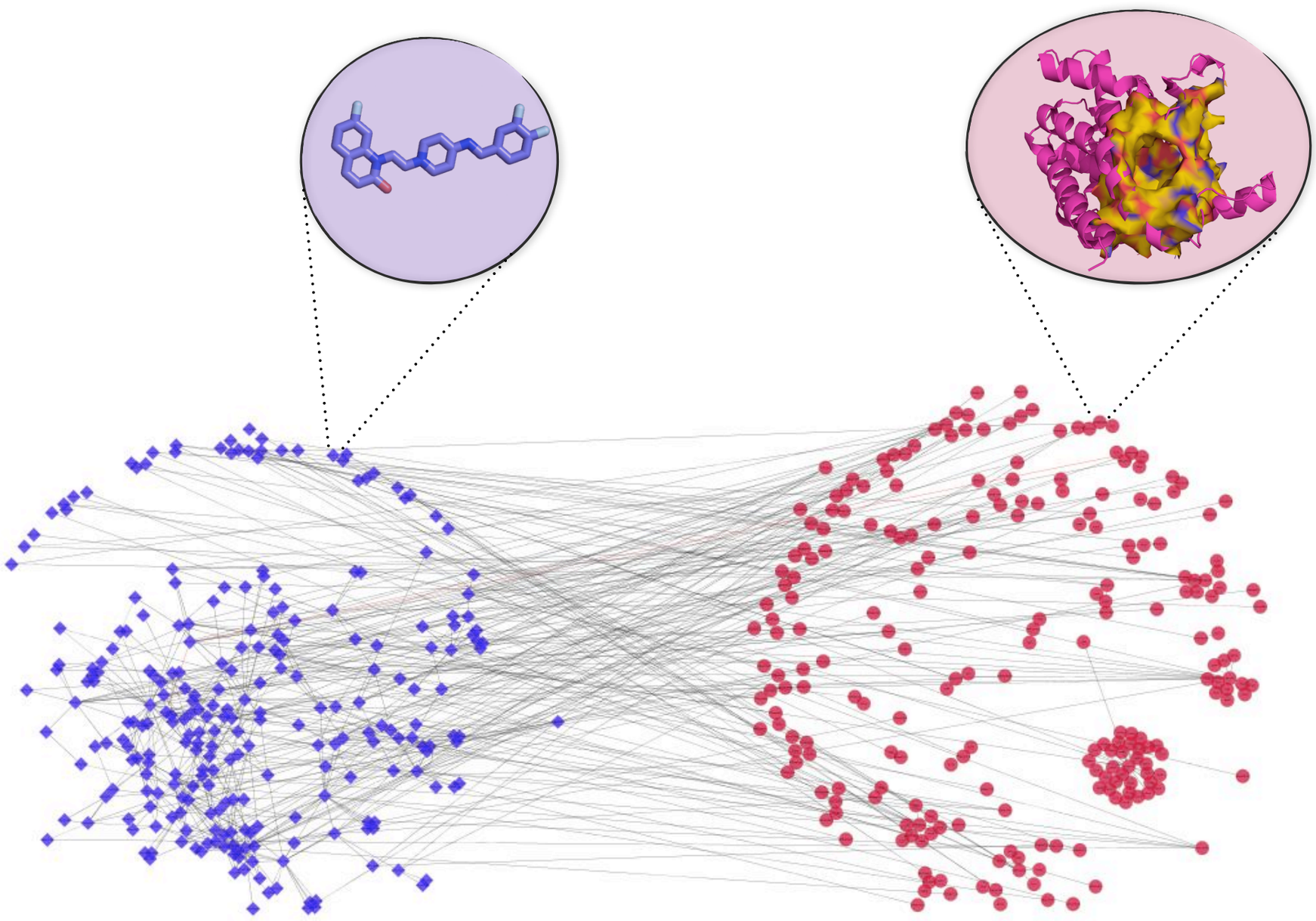
co-crystallized



Epidermal growth factor receptor

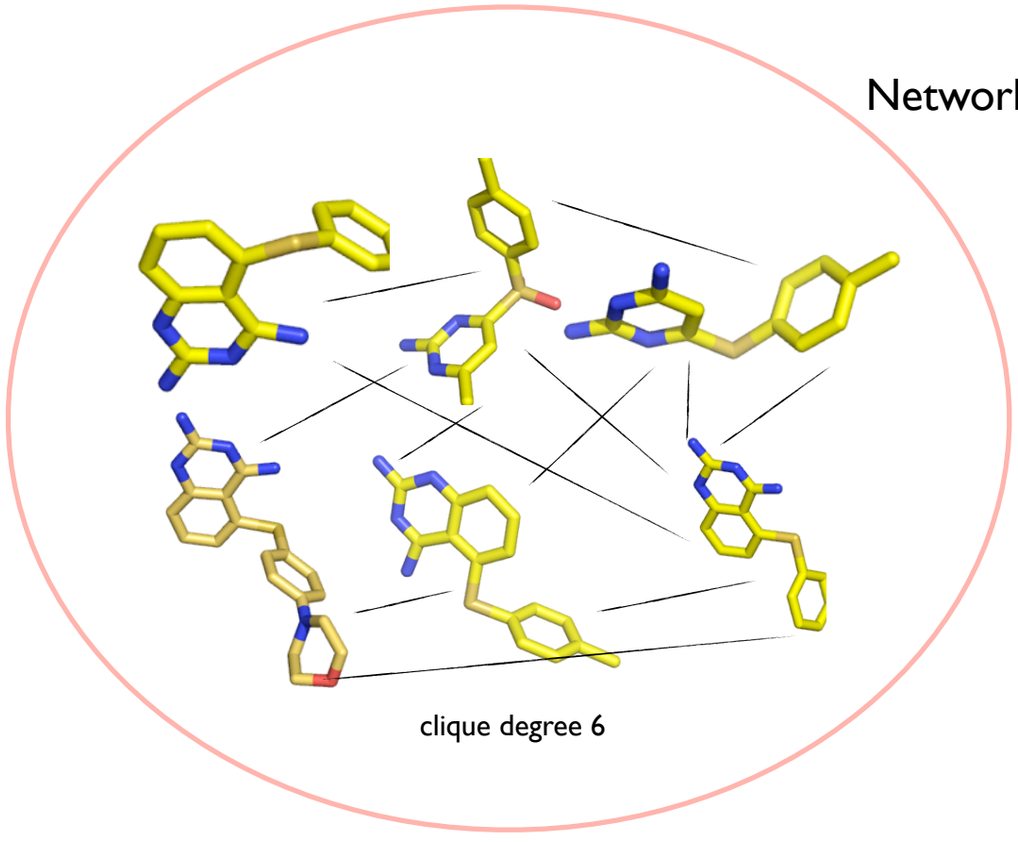
Network-based Method

nAnnolyze



Ligand subnetwork

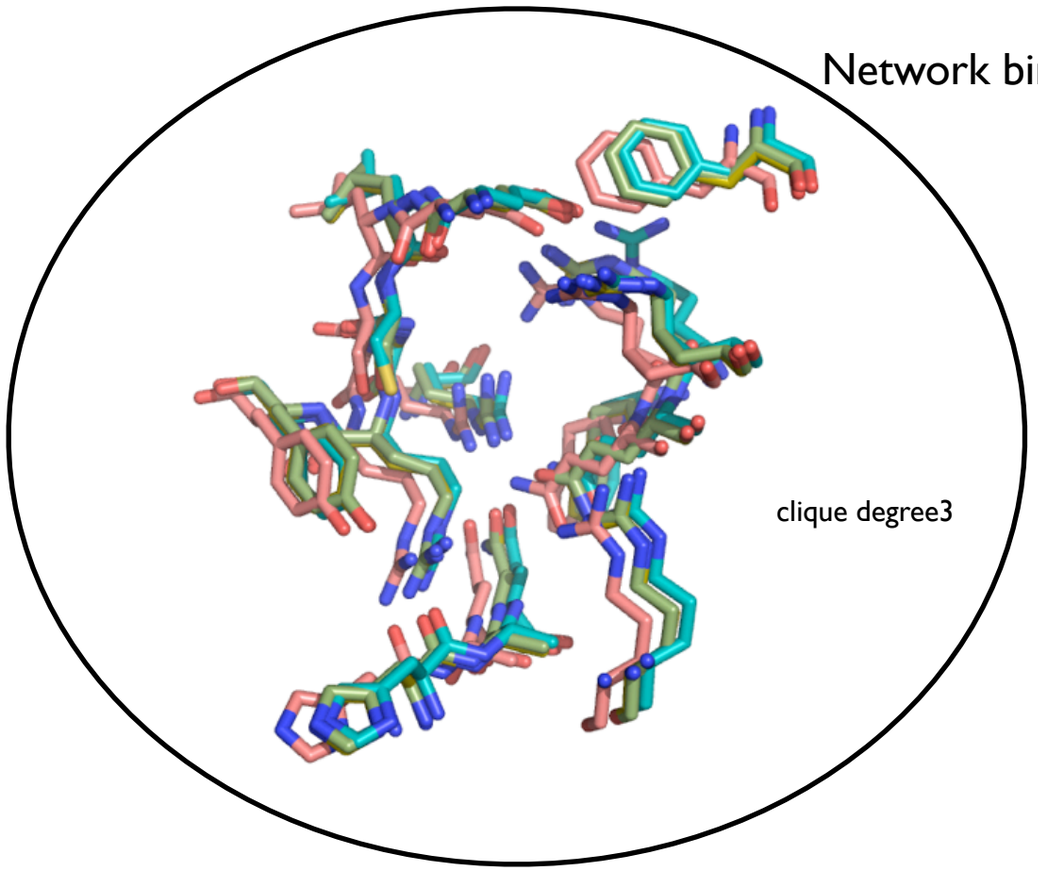
- Retrieved 7,609 high drug-likeness* compounds from PDB.
- Nodes of highly similar compounds: cliques of similarities.
- 4,101 **nodes** of ligand clusters and 24,856 **edges**.
- Edges weight = normalized similarity score.



* Bickerton, G. R., Paolini, G. V, Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. *Nature chemistry*, *4*(2), 90–8.

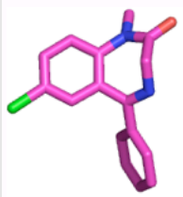
Protein binding-site network

- Binding-sites for the 7,609 compounds: 28,299 binding-sites.
- Similarities between proteins by structural comparisons of the binding-site.
- 19,483 **nodes** of binding-sites and 29,811 **edges**.
- Edges weight = normalized binding-site similarity score.

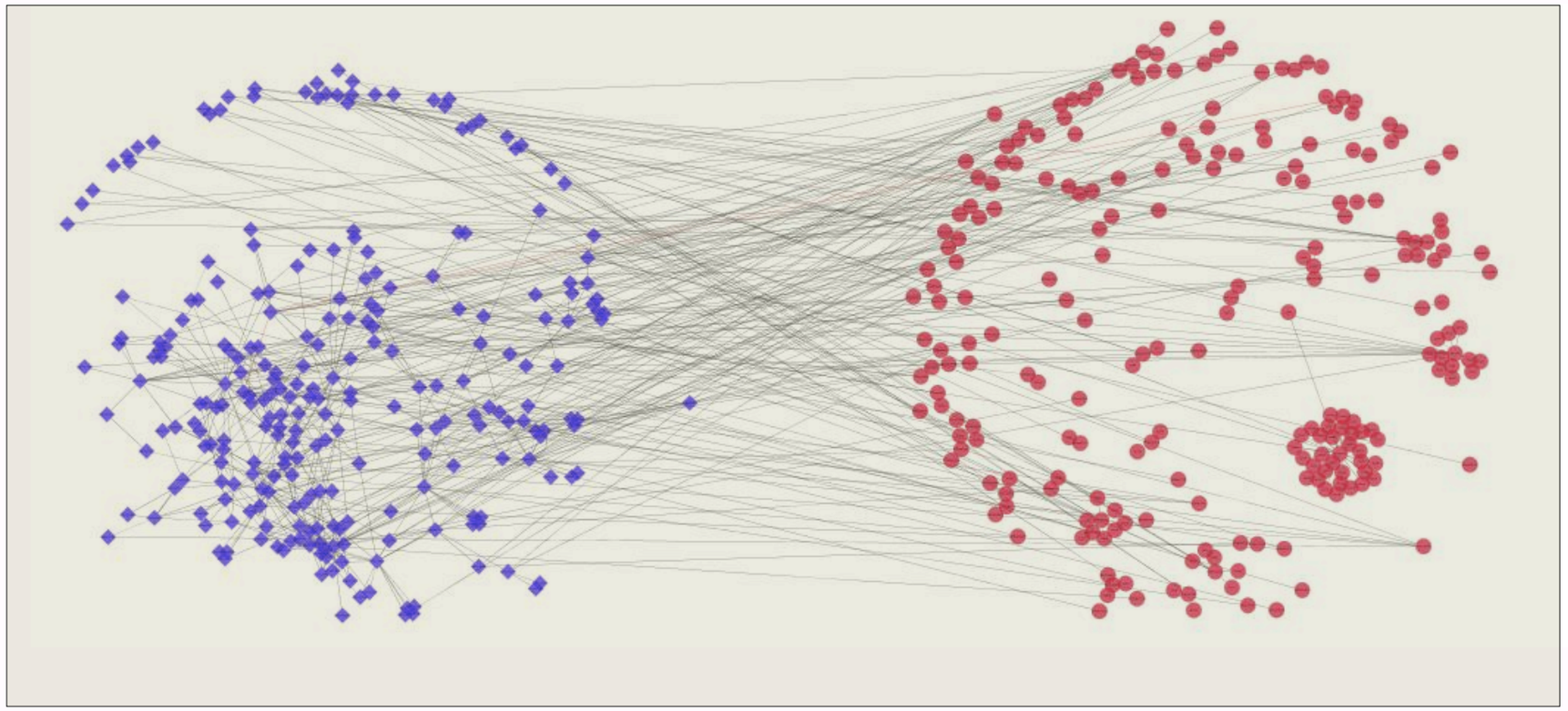


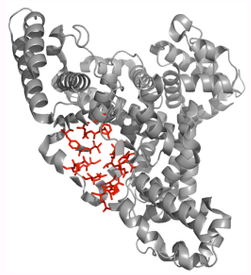
Link the two subnetworks by edges between protein structures and their co-crystallized ligands.

Looking for targets...

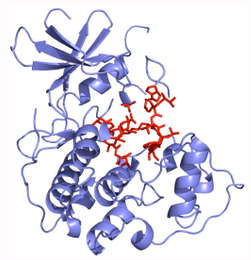


Query
DZP





t1

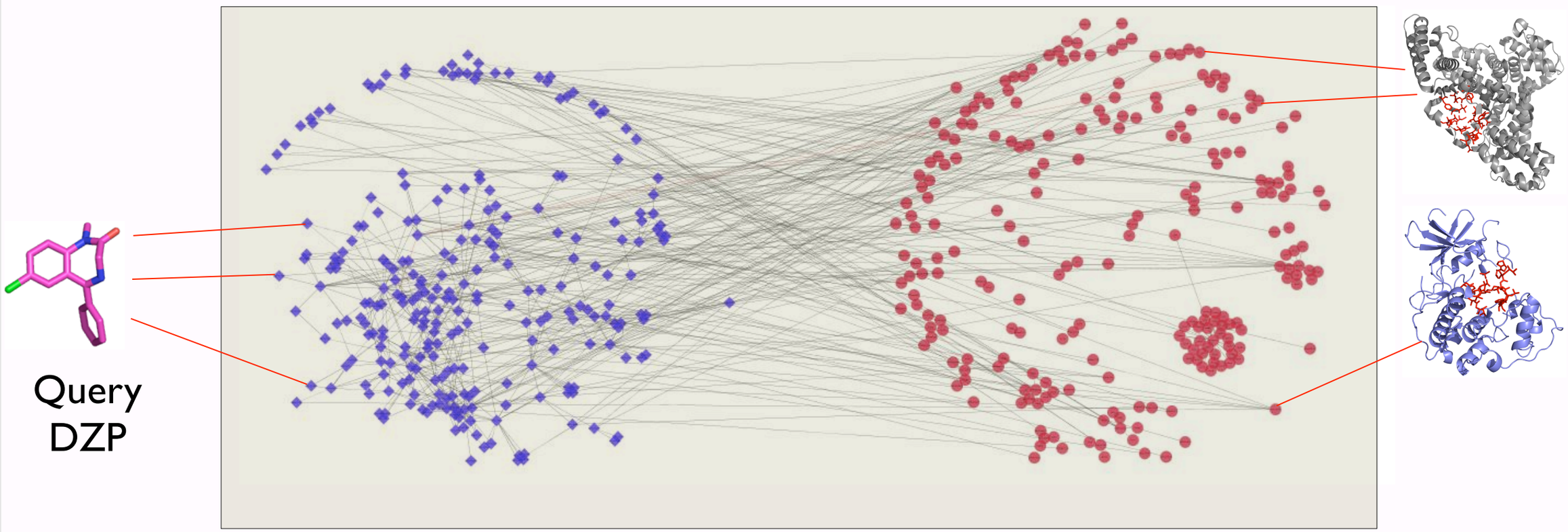


t2

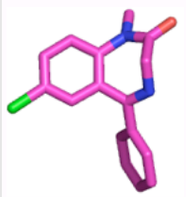
.
.
.

tN

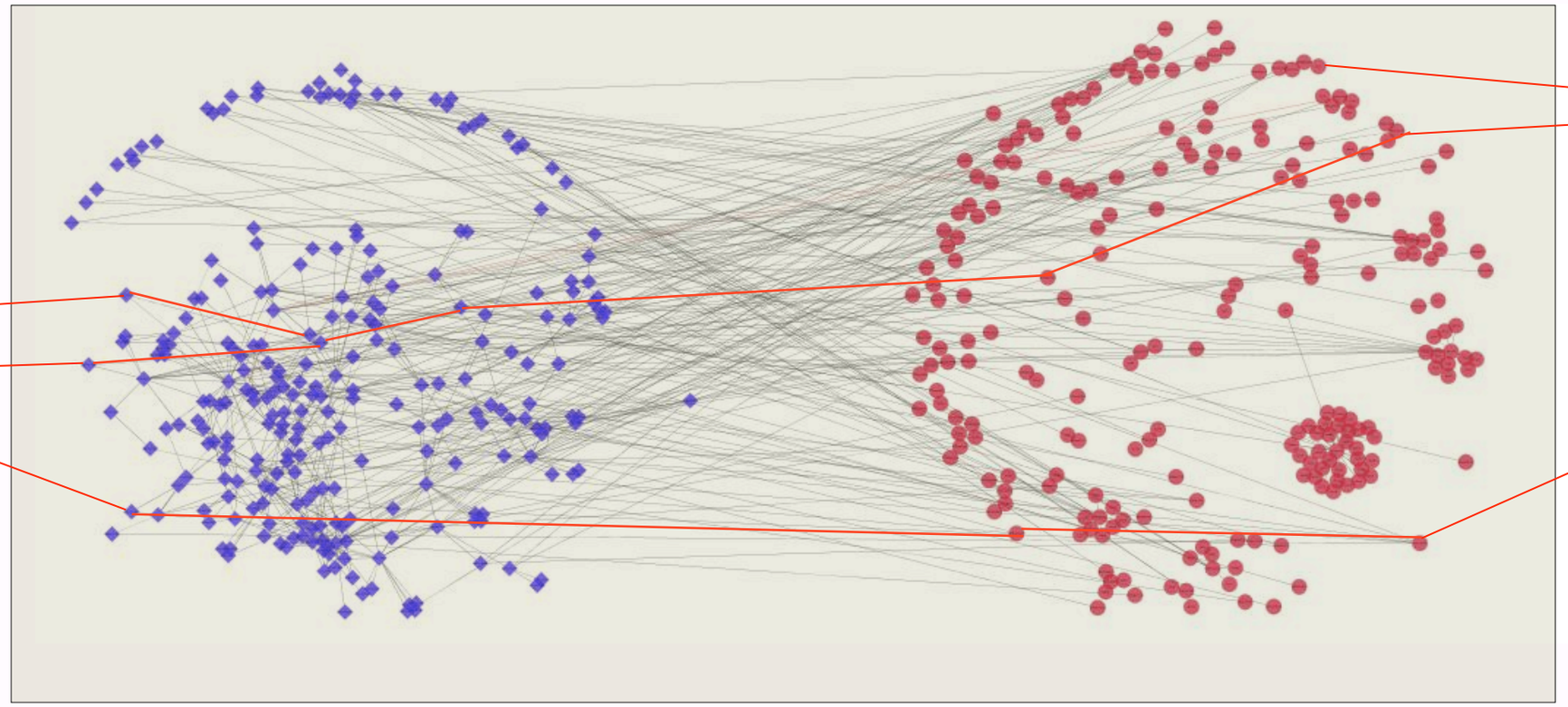
Looking for targets...

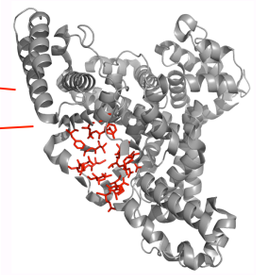


Looking for targets...

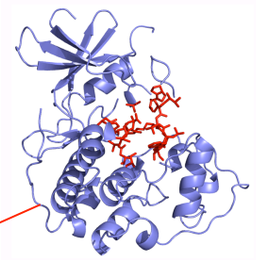


Query
DZP





t1

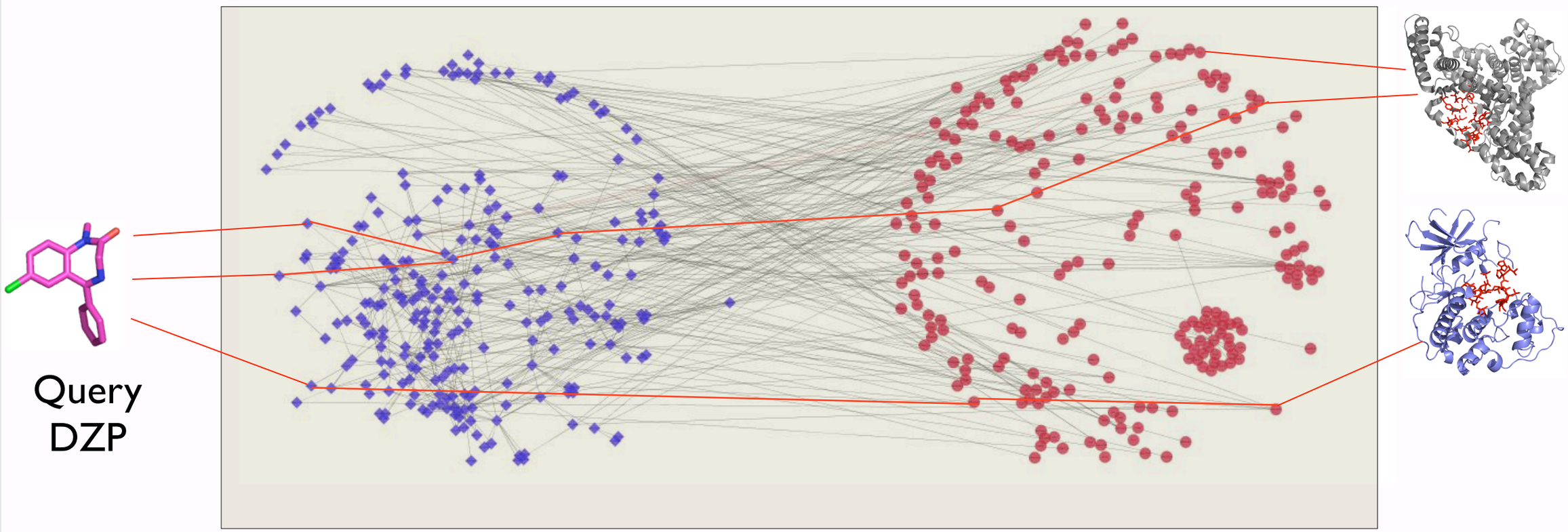


t2

.
.
.

tN

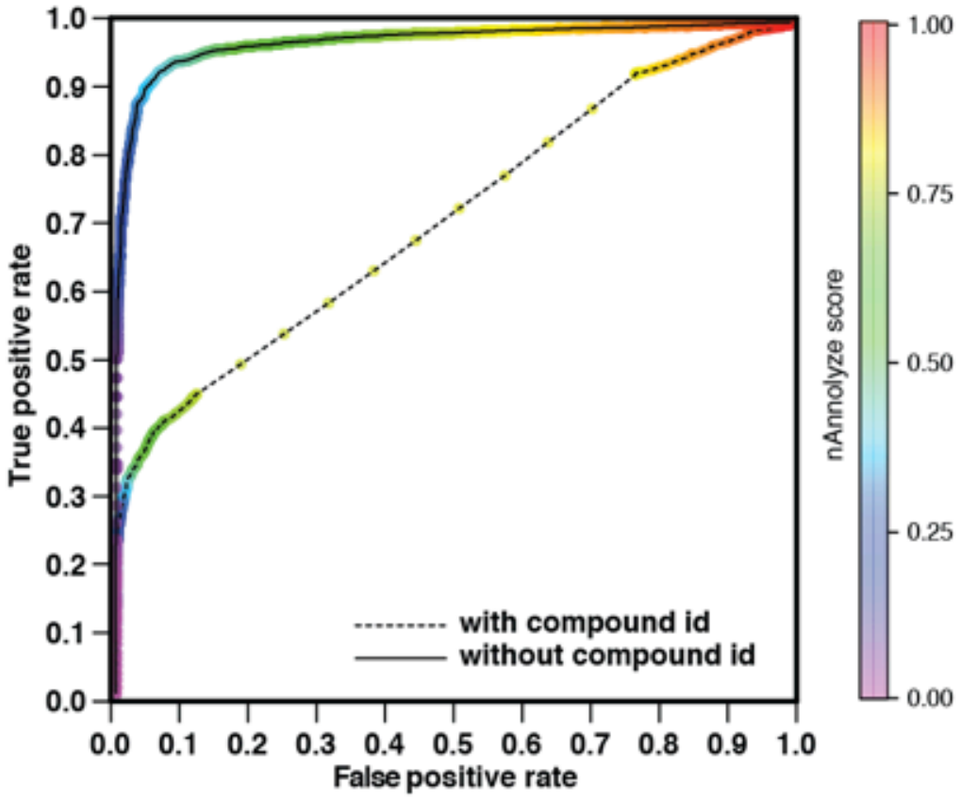
Looking for targets...

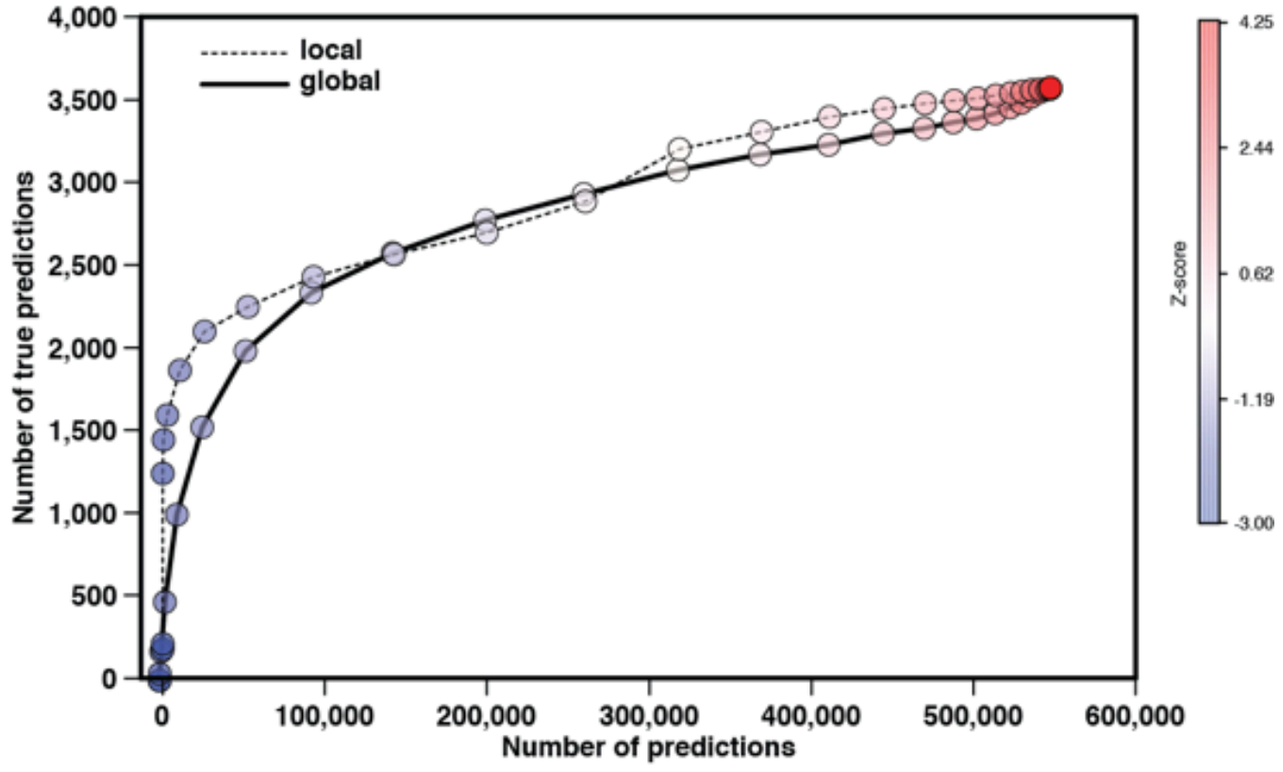


Ligand	Target	Distance	Global Z-score	Local Z-score
DZP	t1	1.3	-1.6	-2.5
DZP	t2	2.5	2.3	1.02
DZP	tM	1.9	-1.6	-3.16
DZP	tN	2.6	2.42	2.97

Benchmarking

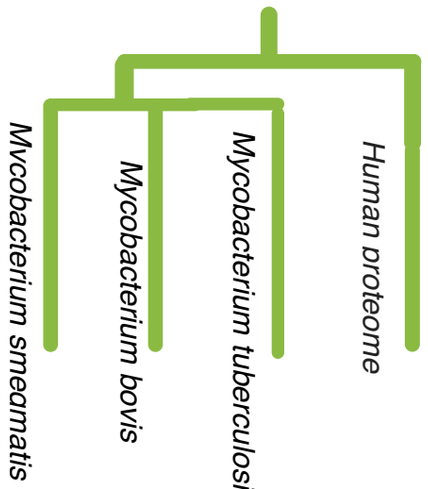
- 232 approved FDA drugs co-crystallized with a protein.
- Test-set = 6,282 true drug-protein pairs and 5,981 negative pairs.
- Drug ID = 0.97 AUC
- Anonymous compounds = 0.73 AUC





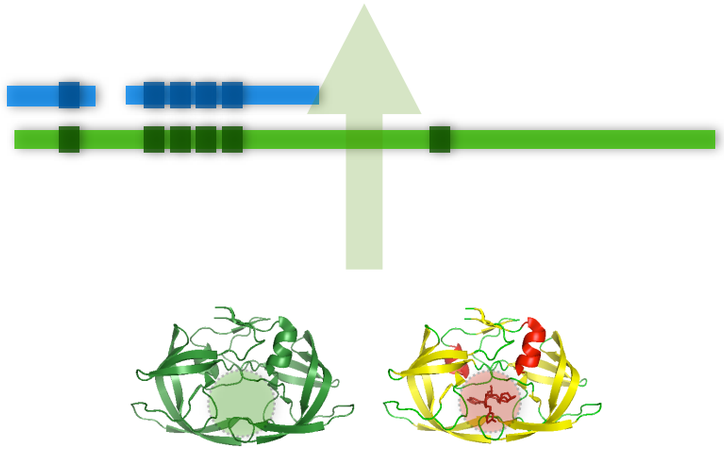
Applying the method, modeling genomes...

1. Modeling



2. Binding-site inheritance

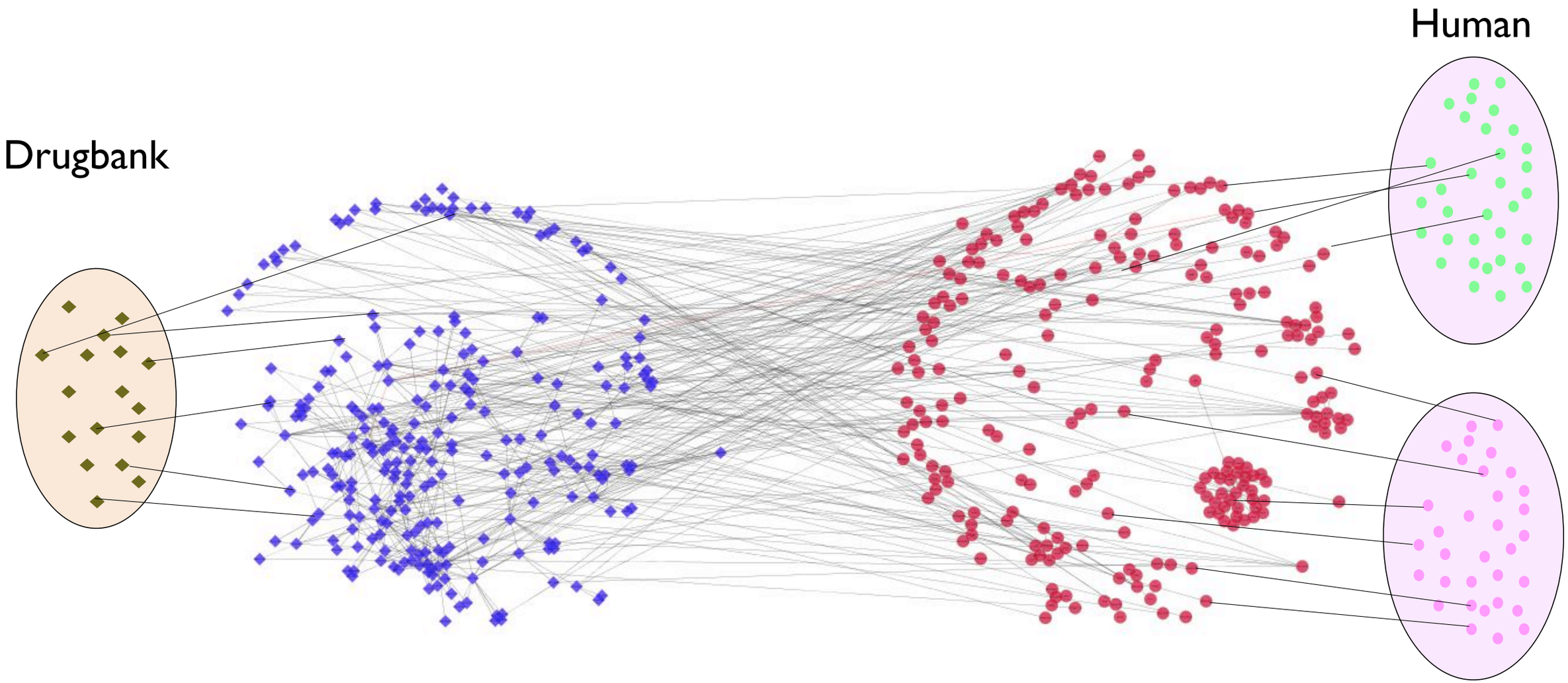
3D model



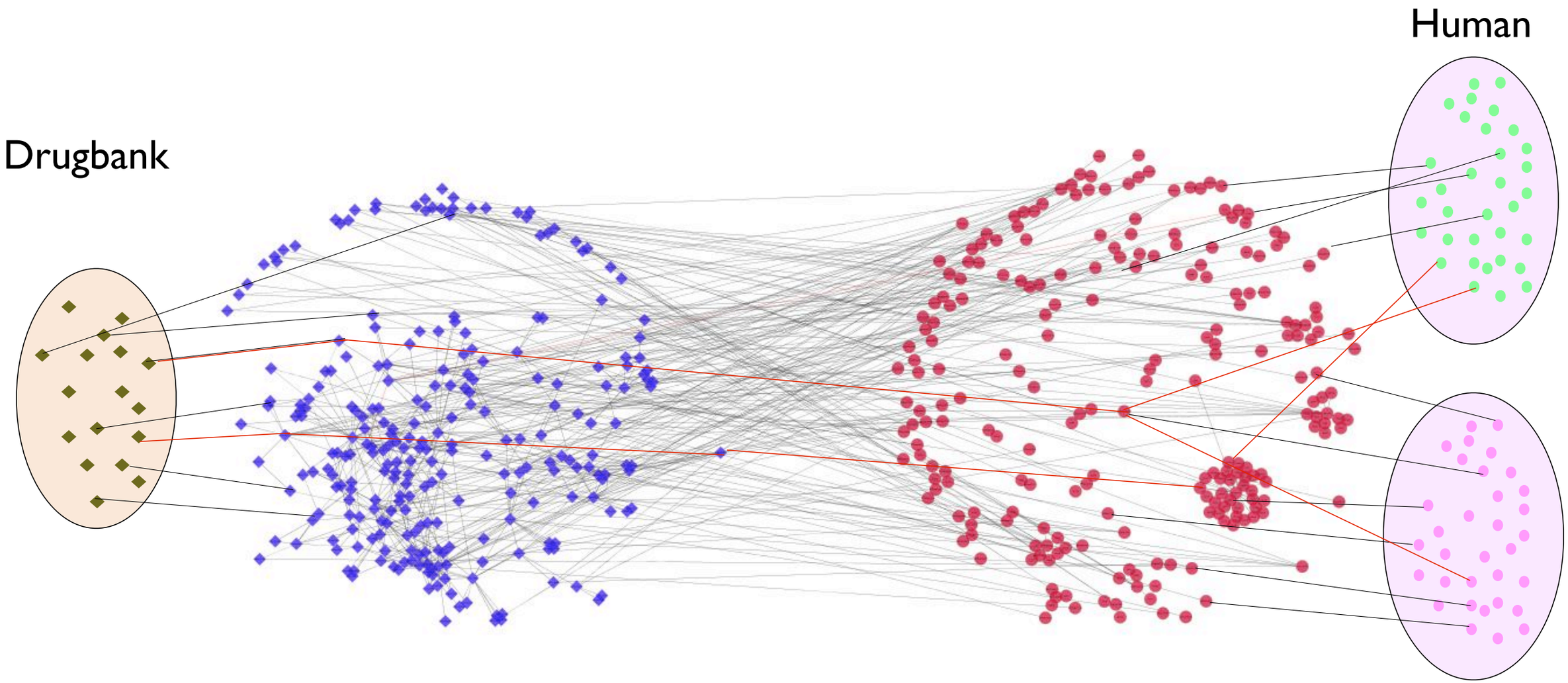
PDB templates

	Human	Bacterial proteomes
3D reliable models	31,734 with overlapping	5,008 no overlapping
Different Proteins	14,000	5,008 different proteins
Inherited binding-sites	64,000	30,000

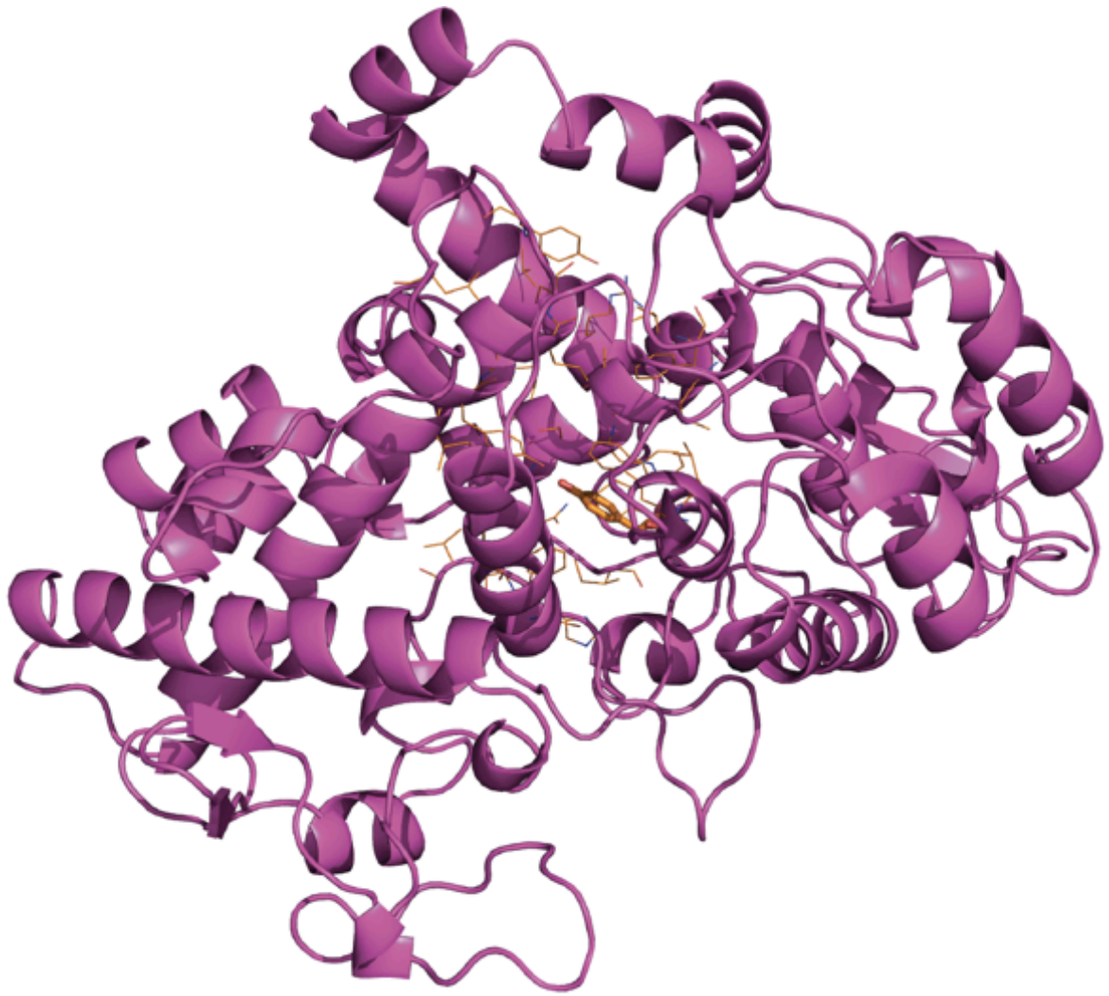
Searching for Drugbank drugs interactions...



Searching for Drugbank drugs interactions...



Human Cyclooxygenase-1 targeted by NSAID drugs



- 21 out of the 44 approved FDA drugs against COX-1 (score > 0.85).
- Human structure model from the sheep COX-1.
- Predicted binding site includes Tyrosine 385.

Drug ID	Drug name	nAnnoLyze score
DB00712	Flurbiprofen	0.97
DB00328	Indomethacin	0.97
DB01600	Tiaprofenicacid	0.96
DB00870	Suprofen	0.96
DB00821	Carprofen	0.96
DB00788	Naproxen	0.96
DB00500	Tolmetin	0.94
DB00465	Ketorolac	0.94
DB00963	Bromfenac	0.92
DB00586	Diclofenac	0.91
DB06802	Nepafenac	0.90
DB01283	Lumiracoxib	0.90
DB00784	Mefenamicacid	0.89
DB00861	Diflunisal	0.88
DB04552	NiflumicAcid	0.88
DB00991	Oxaprozin	0.88
DB01050	Ibuprofen	0.87
DB00939	Meclofenamicacid	0.86
DB01399	Salsalate	0.86
DB01009	Ketoprofen	0.86
DB00605	Sulindac	0.85

Sorafenib pathway targeting through binding of several protein

Sorafenib pathway targeting through binding of several protein

Target	Score	Structure	KEGG Pathway
MAPK 14	**0.99**	**Yes**	MAPK signaling Fox0 signaling VEGF signaling Rap1 signaling RIG-I-like receptor signaling Acute myeloid leukemia
CDK19	**0.97**	**No**	-
FLT1	**0.90**	**Yes**	Ras signaling pathway
RAF 1	**0.89**	**Yes**	MAPK signaling Ras signaling Rap1 signaling VEGF signaling Fox0 signaling pathway Acute myeloid leukemia
ARAF	**0.88**	**Yes**	Fox0 signaling Acute myeloid leukemia
CDK10	**0.88**	**No**	-
BRAF	**0.88**	**Yes**	MAPK signaling Rap1 signaling Fox0 signaling Acute myeloid leukemia
CDK8	**0.87**	**Yes**	-
FLT3	**0.86**	**Yes**	Acute myeloid leukemia
MAPK 15	**0.86**	**No**	-

Annotated (Chembl, PubChem, Drugbank, PDB)

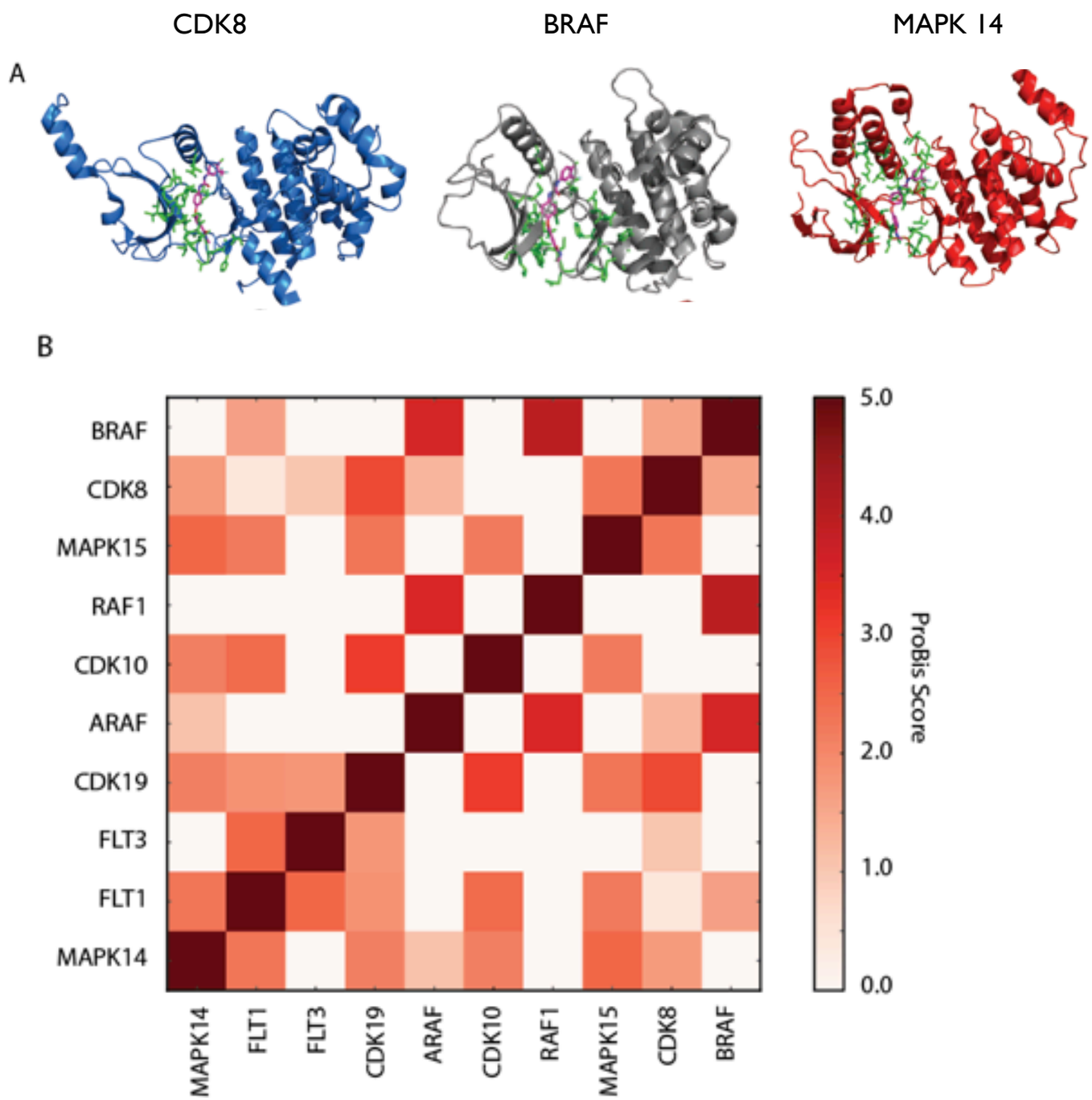
Not Annotated

Sorafenib pathway targeting through binding of several protein

Target	Score	Structure	KEGG Pathway
MAPK 14	**0.99**	**Yes**	MAPK signaling Fox0 signaling VEGF signaling Rap1 signaling RIG-I-like receptor signaling Acute myeloid leukemia
CDK19	**0.97**	**No**	-
FLT1	**0.90**	**Yes**	Ras signaling pathway
RAF 1	**0.89**	**Yes**	MAPK signaling Ras signaling Rap1 signaling VEGF signaling Fox0 signaling pathway Acute myeloid leukemia
ARAF	**0.88**	**Yes**	Fox0 signaling Acute myeloid leukemia
CDK10	**0.88**	**No**	-
BRAF	**0.88**	**Yes**	MAPK signaling Rap1 signaling Fox0 signaling Acute myeloid leukemia
CDK8	**0.87**	**Yes**	-
FLT3	**0.86**	**Yes**	Acute myeloid leukemia
MAPK 15	**0.86**	**No**	-

Annotated (Chembl, PubChem, Drugbank, PDB)

Not Annotated



Antimicrobial drugs against Mycobacterium tuberculosis

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Target Prediction for an Open Access Set of Compounds Active against *Mycobacterium tuberculosis*

Francisco Martínez-Jiménez[1,2], George Papadatos[3], Lun Yang[4], Iain M. Wallace[3], Vinod Kumar[4], Ursula Pieper[5], Andrej Sali[5], James R. Brown[4]*, John P. Overington[3]*, Marc A. Marti-Renom[1,2]*

1 Genome Biology Group, Centre Nacional d'Anàlisi Genòmica (CNAG), Barcelona, Spain, **2** Gene Regulation Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain, **3** European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom, **4** Computational Biology, Quantitative Sciences, GlaxoSmithKline, Collegeville, Pennsylvania, United States of America, **5** Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, United States of America

Highlights oral presentation

Tuesday,23

nAnnolyze

Home Help About Download

Introduce your query molecule

ID of the molecule DB00398

e.g: DB00712

SMILE CNC(=O)C1=NC=CC(OC2=CC

e.g: CC(C(O)=O)C1=CC(F)=C(C=C1)C1=CC=CC=C1

Molecular Weight 464.82

e.g: 244

Input an example!

Upload your File with the SMILES Upload a File

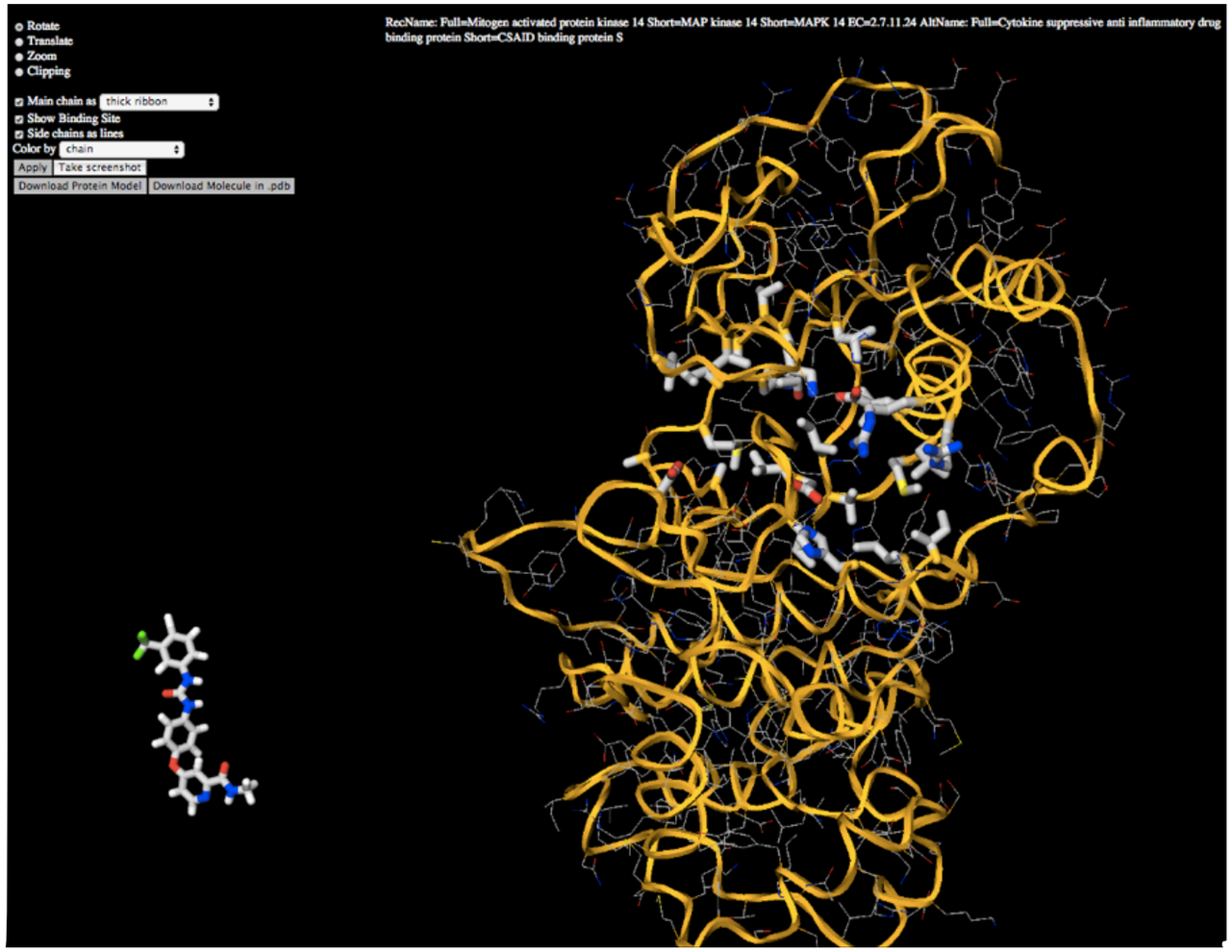
Choose Files No file chosen ?

Maximum 10 compounds per search

Select the organism

- ○ *Mycobacterium tuberculosis, Mycobacterium bovis, Mycobacterium smegmatis*
- ◉ *Homo Sapiens*

Submit



nAnnolyze

Home Help About Download

+ Toggle Statistics

Graphics Table

Visualize:

- ◉ *Overrepresented GO terms and Pathways*
- ○ *All Go terms and Pathways*

Filter by global z-score: < -2

Statistics for Compound DB00398

Download .csv

GO terms ?

Pathways ?

Best Worst

Molecule ID	External ID	Target UP	Score	Z-Score local	Z-Score global	Description	Binding-Site residues	Visualize...
DB00398	NP_6205	Q13083	0.99	-3.23	-3.30	RecName: Full=Mitogen-activated protein kinase 14;Short=MAP kinase 14;Short=MAPK 14;EC=2.7.11.24;AltName: Full=Cytokine suppressive anti-inflammatory drug-binding protein;Short=CSAID-binding protein;S	[:A and (30;38;51;52;53;67;70;71;74;	
DB00398	NP_0558	Q5JQZ7	0.98	-4.14	-3.24	RecName: Full=Cyclin-dependent kinase 19;EC=2.7.11.22;AltName: Full=CDC2-related protein kinase 6;AltName: Full=Cell division cycle 2-like protein kinase 6;AltName: Full=Cell division protein kinase 1	[:A and (35;50;51;52;66;69;70;73;78;	
DB00398	NP_0012	G1T9S5	0.97	-4.39	-3.14	SubName: Full=Uncharacterized protein;	[:A and (27;28;32;35;50;51;52;66;69;	
DB00398	GENSCAN		0.96	-3.21	-3.06		[:A and (924;925;932;950;952;969;97	
DB00398	NP_0013	L7RSM2	0.94	-2.91	-2.95	SubName: Full=Mitogen-activated protein kinase 14;SubName: Full=Mitogen-activated protein kinase 14 isoform CRA_e;	[:A and (30;35;38;51;70;74;83;84;85;	

http://nannolyze.cnag.cat

Acknowledgments

Davide Baù
Gireesh K. Bogu
François le Dily
Marc A. Marti-Renom
David Dufour
François Serra
Michael Goodstadt
Yasmina Cuartero

COLLABORATORS

Jim Brown (GSK)
LLuís Ballell (GSK)
John Overington (EBI-EMBL)
Andrej Sali (UCSF)
Anna Tramontano (Sapienza University)

http://marciuslab.org
http://integrativemodeling.org
http://cnag.cat · http://crg.cat

ERA-NET PathoGenoMics
HFSP
GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
cnag
CRG Centre for Genomic Regulation