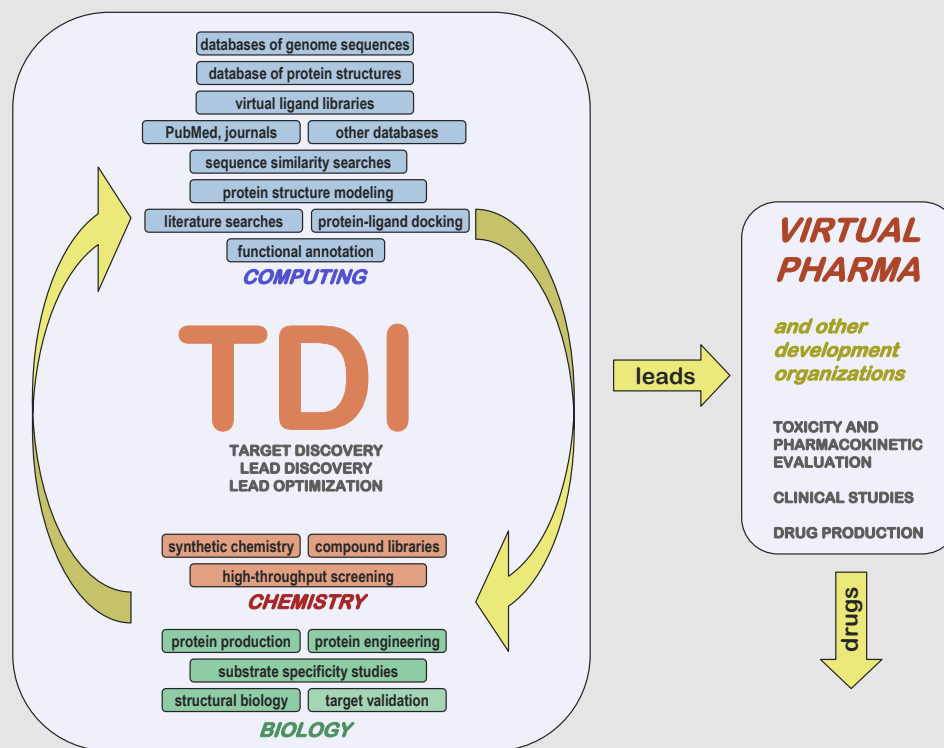


A “kernel” for the Tropical Disease Initiative

An open source approach to drug discovery



Marc A. Marti-Renom

<http://sgu.bioinfo.cipf.es>

Structural Genomics Unit
Bioinformatics Department
Prince Felipe Research Center (CIPF), Valencia, Spain



Bioinformatics and Genomics Department (CIPF)

<http://bioinfo.cipf.es>



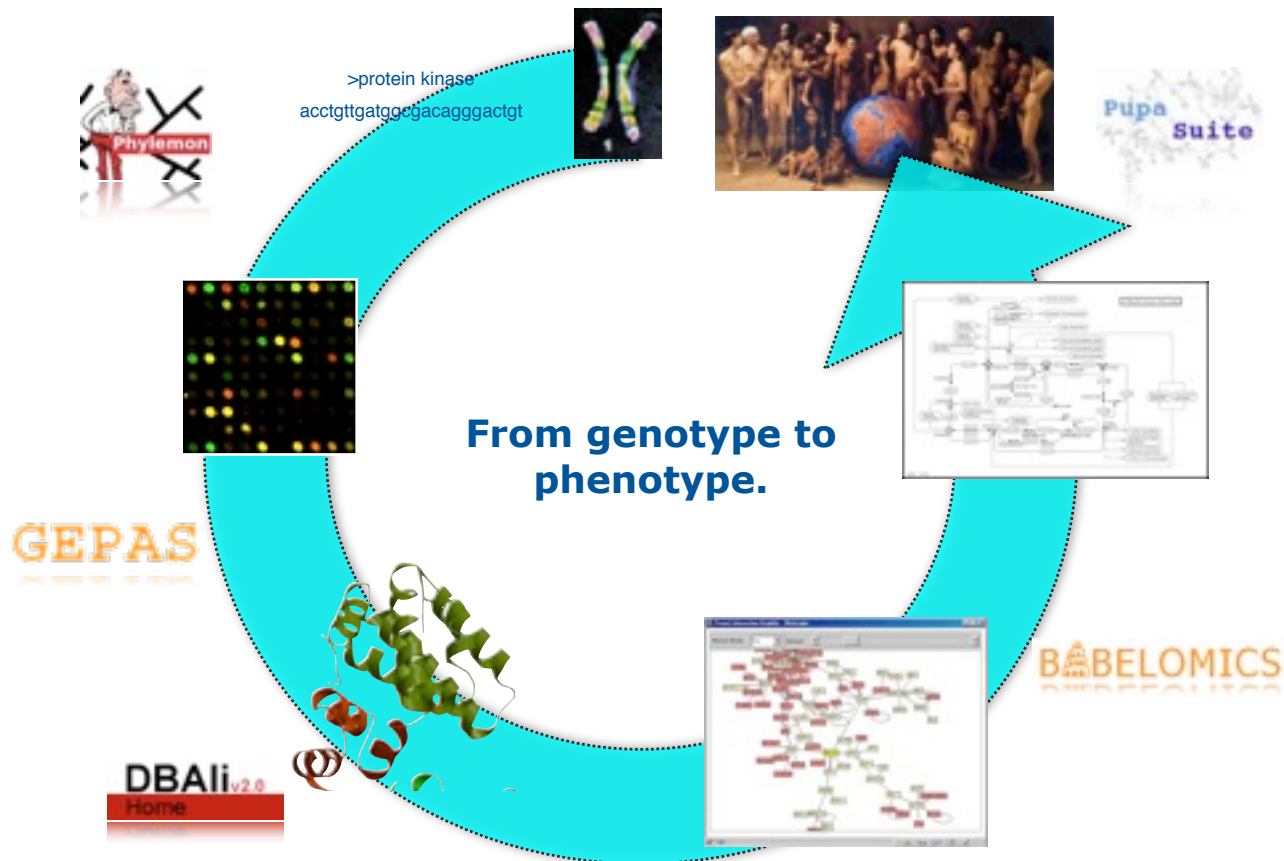
Functional Genomics
Dr. Joaquín Dopazo



Comparative Genomics
Dr. Hernán Dopazo

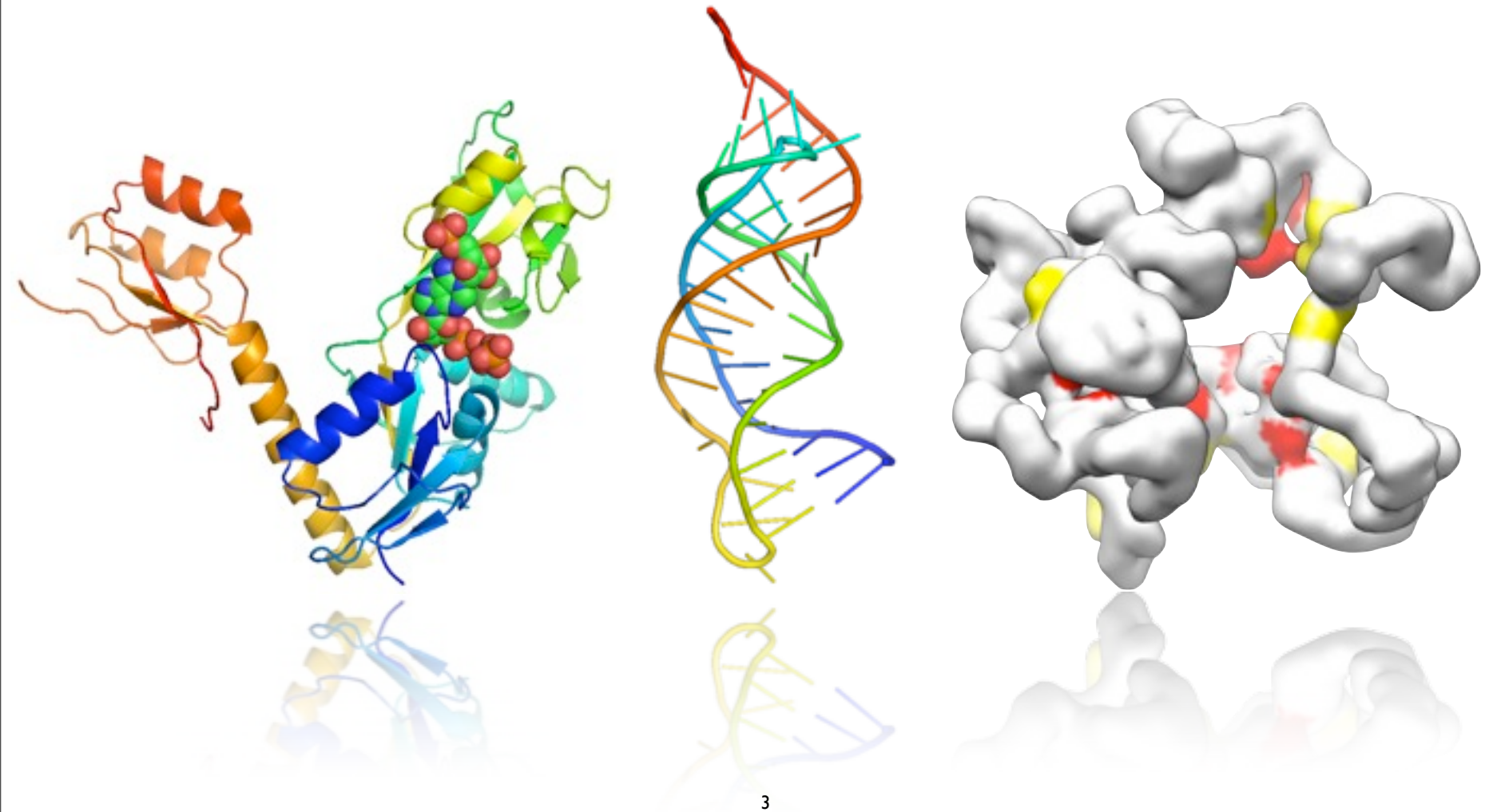


Structural Genomics
Dr. Marc A. Martí-Renom



Structural Genomics Unit

Bioinformatics Department, CIPF



TDI *a story*



2004

.Steve Maurer (Berkeley) and Arti Rai (Duke)
.PLoS Medicine, Dec. 2004. Vol 1(3):e56

2005

.TDI web site <http://TropicalDisease.org>
.Ginger Taylor and The Synaptic Leap



2006

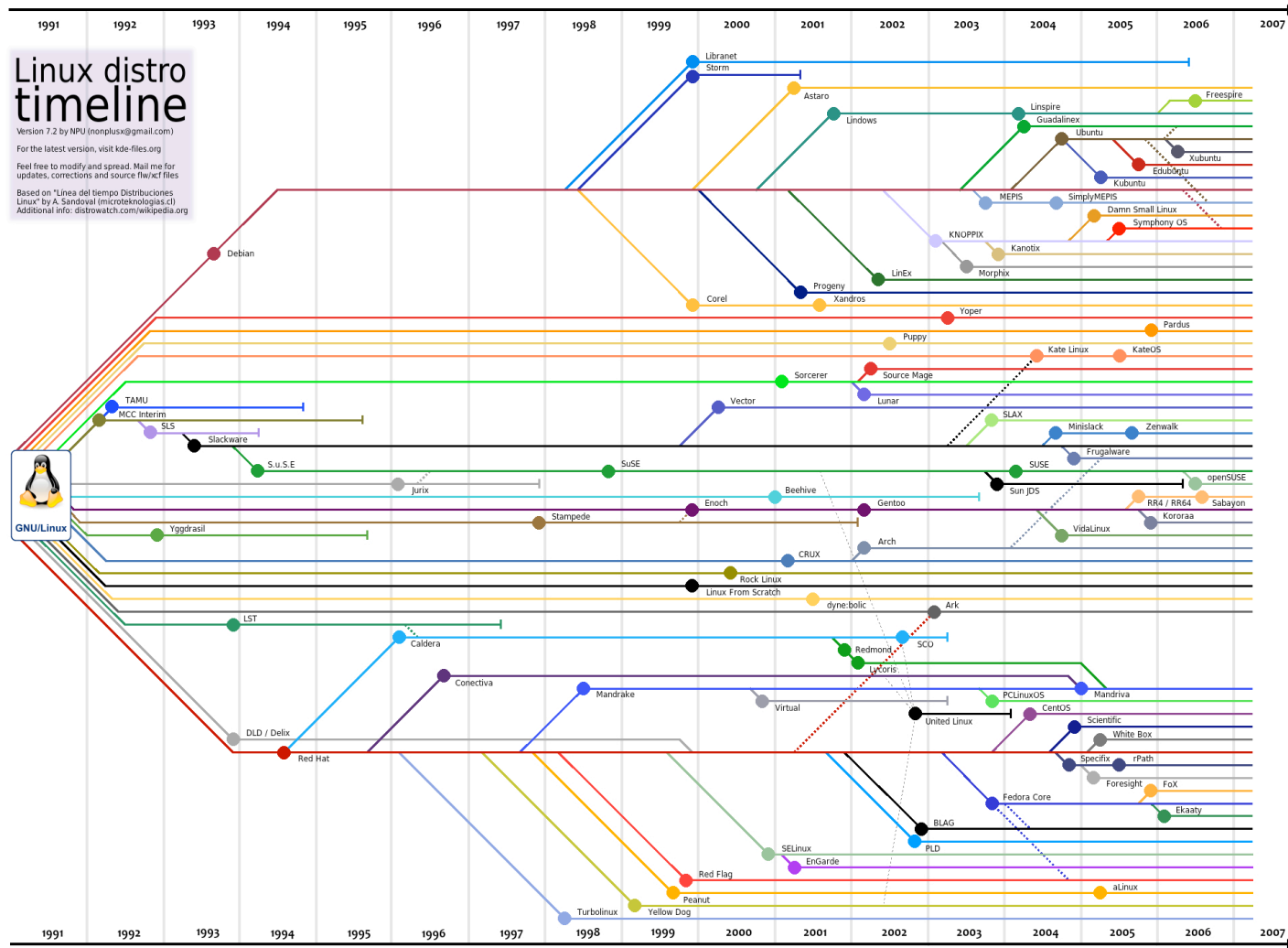
.Maurer and Sali 41th in “50 Who Matter”
.TSL web site <http://TheSynapticLeap.org>



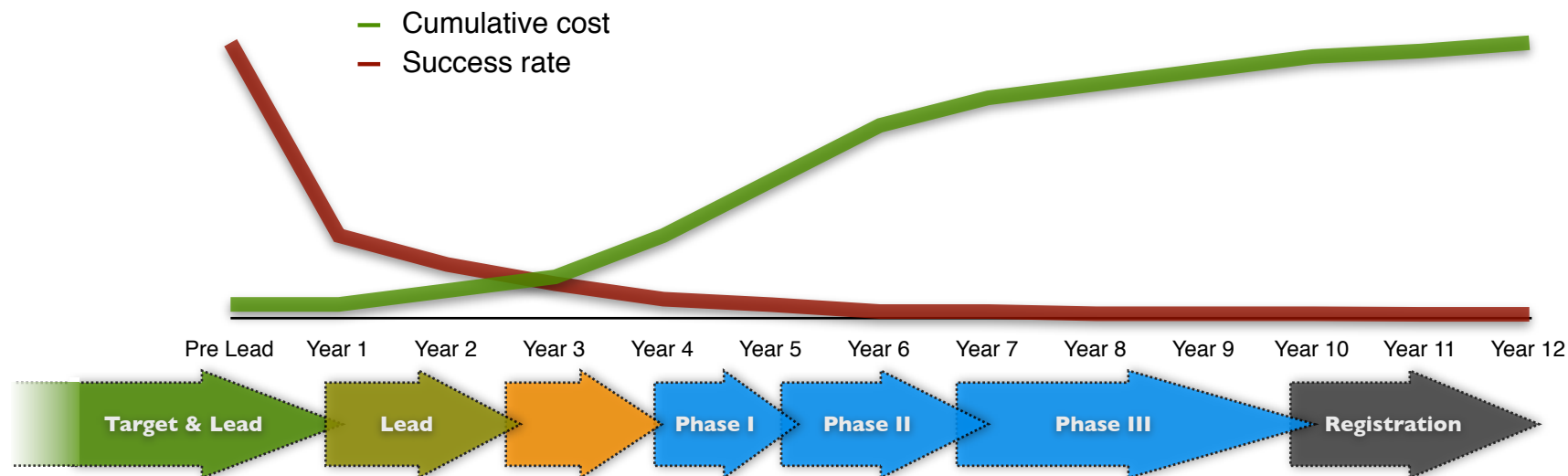
2008

.TDI kernel <http://TropicalDisease.org/kernel>

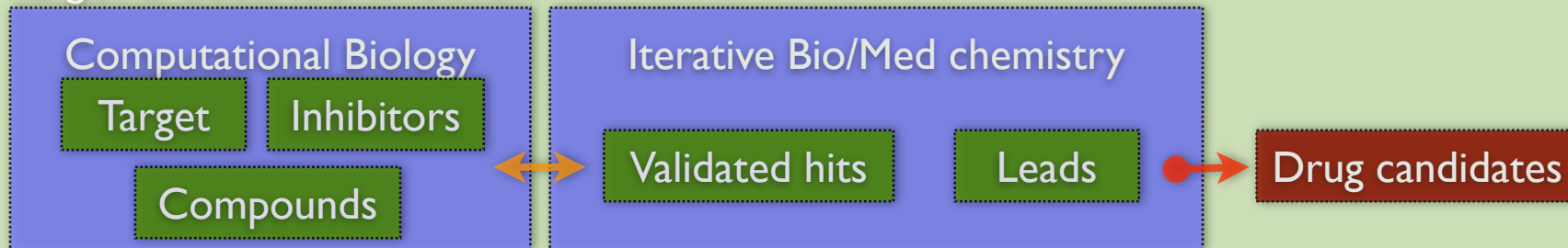
Open Source without a Kernel?



Drug Discovery pipeline

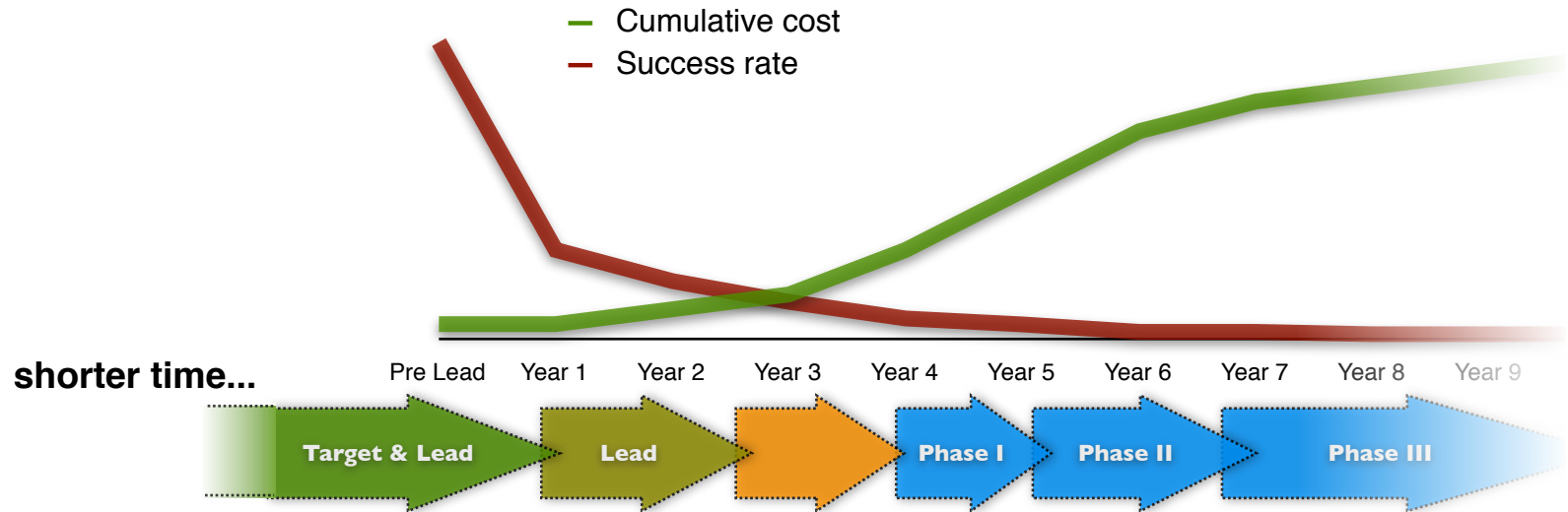


Target & Lead identification



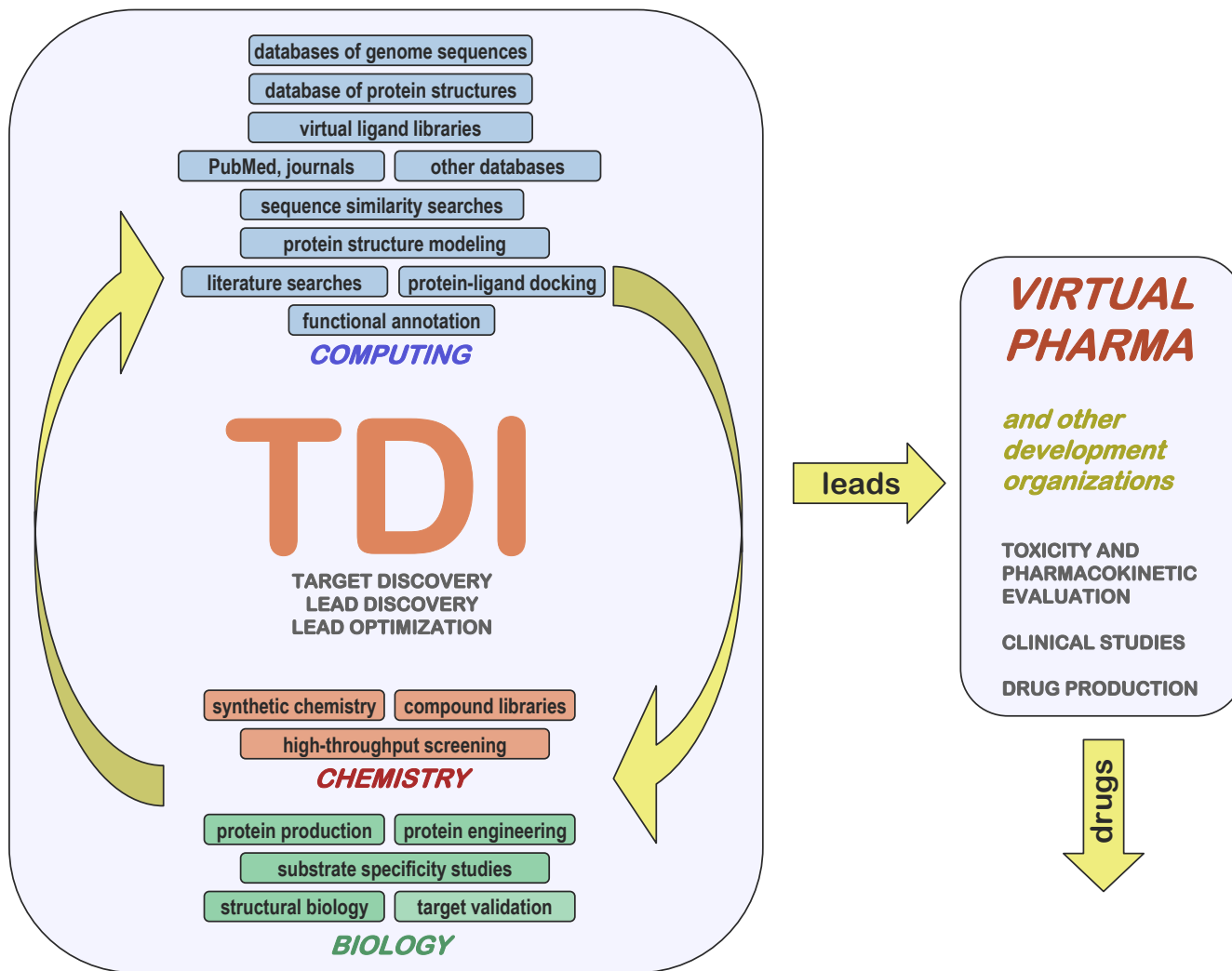
Adapted from: - Nwaka & Ridley. (2003) *Nature Reviews. Drug Discovery*. 2:919
 - Austin, Brady, Insel & Collins. (2004) *Science*. 306:1138

Drug Discovery pipeline



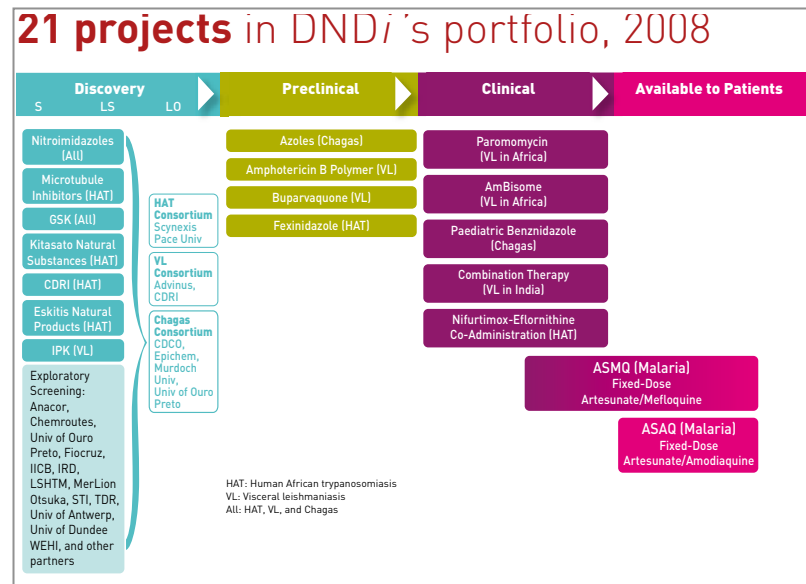
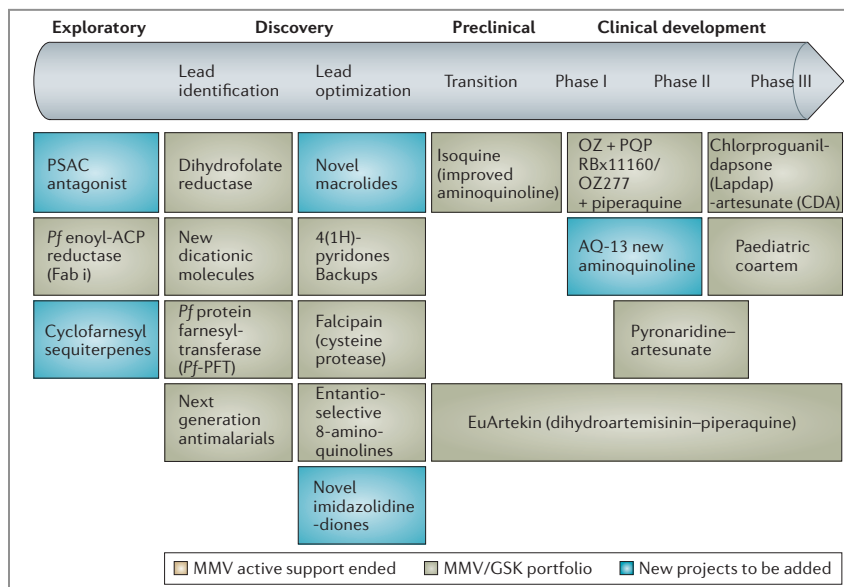
- + Completeness of genome projects (eg, Malaria)
- + New and more complete biological databases
- + New software and computers (cheaper and faster)
- + Internet == more people == less cost

TDI flowchart



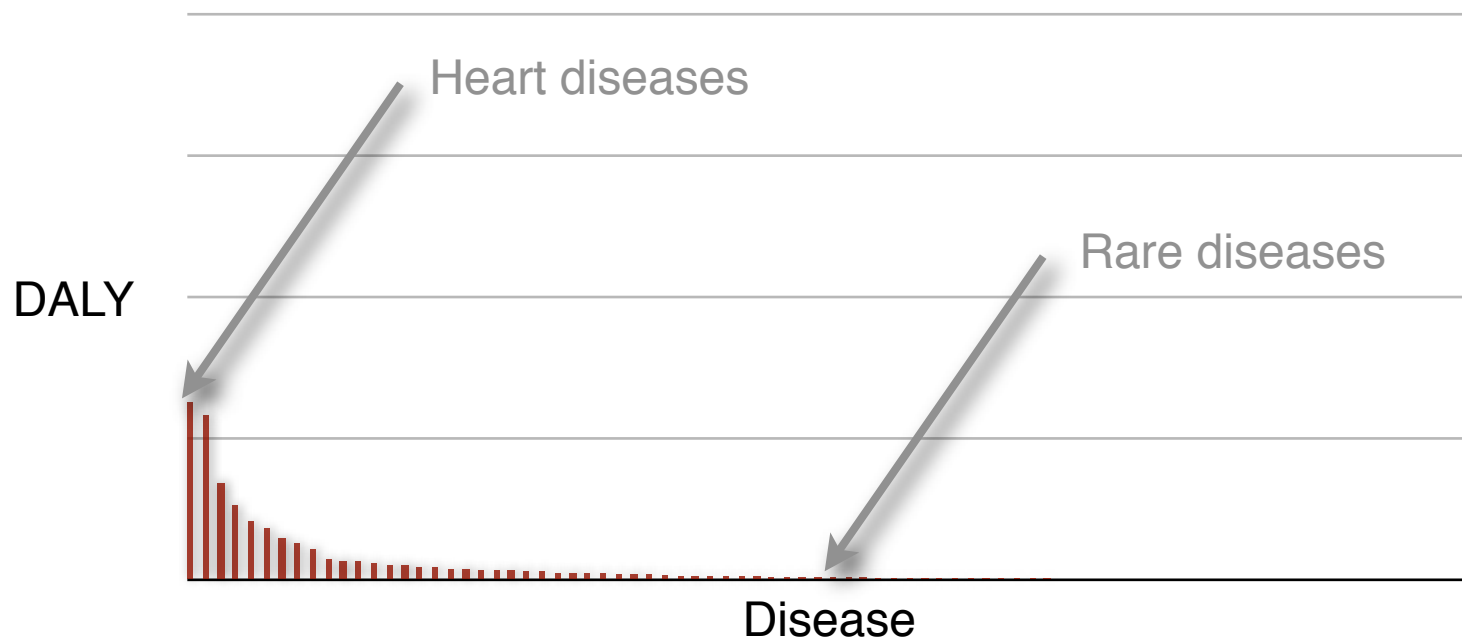
Non-Profit organizations

Open-Source + Out-Source = low cost business model



Need is High in the Tail

- DALY Burden Per Disease in Developed Countries
- DALY Burden Per Disease in Developing Countries



Disease data taken from WHO, *World Health Report 2004*

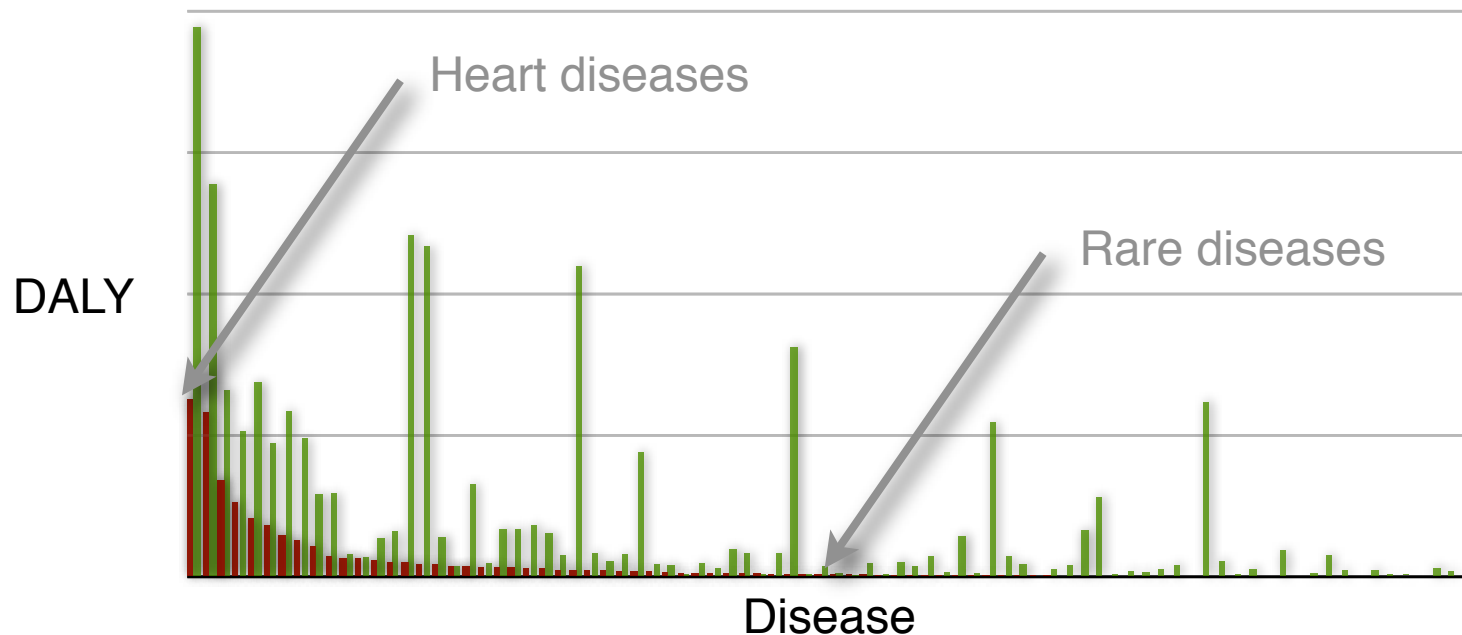
DALY - Disability adjusted life years

DALY is not a perfect measure of market size, but is certainly a good measure for importance.

DALYs for a disease are the sum of the years of life lost due to premature mortality (YLL) in the population and the years lost due to disability (YLD) for incident cases of the health condition. The DALY is a health gap measure that extends the concept of potential years of life lost due to premature death (PYLL) to include equivalent years of 'healthy' life lost in states of less than full health, broadly termed disability. One DALY represents the loss of one year of equivalent full health.

Need is High in the Tail

- DALY Burden Per Disease in Developed Countries
- DALY Burden Per Disease in Developing Countries



Disease data taken from WHO, *World Health Report 2004*

DALY - Disability adjusted life years

DALY is not a perfect measure of market size, but is certainly a good measure for importance.

DALYs for a disease are the sum of the years of life lost due to premature mortality (YLL) in the population and the years lost due to disability (YLD) for incident cases of the health condition. The DALY is a health gap measure that extends the concept of potential years of life lost due to premature death (PYLL) to include equivalent years of 'healthy' life lost in states of less than full health, broadly termed disability. One DALY represents the loss of one year of equivalent full health.

“Unprofitable” Diseases and Global DALY (in 1000’s)

Malaria*	46,486
Tetanus	7,074
Lymphatic filariasis*	5,777
Syphilis	4,200
Trachoma	2,329
Leishmaniasis*	2,090
Ascariasis	1,817
Schistosomiasis*	1,702
Trypanosomiasis*	1,525

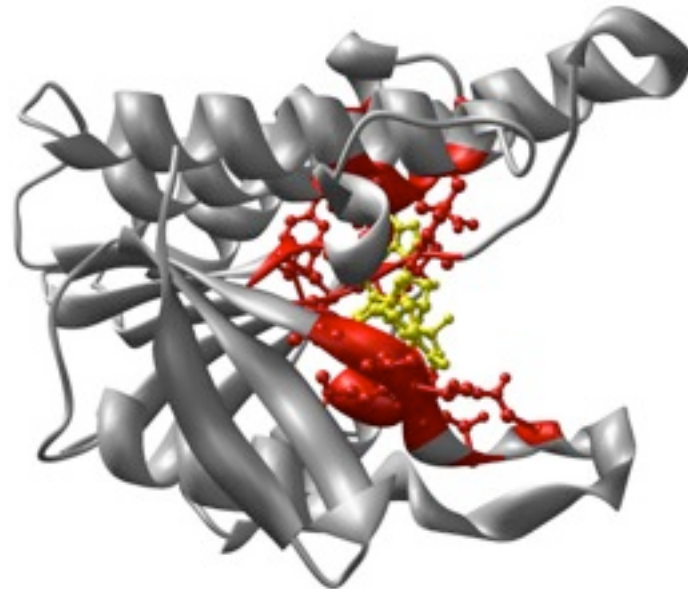
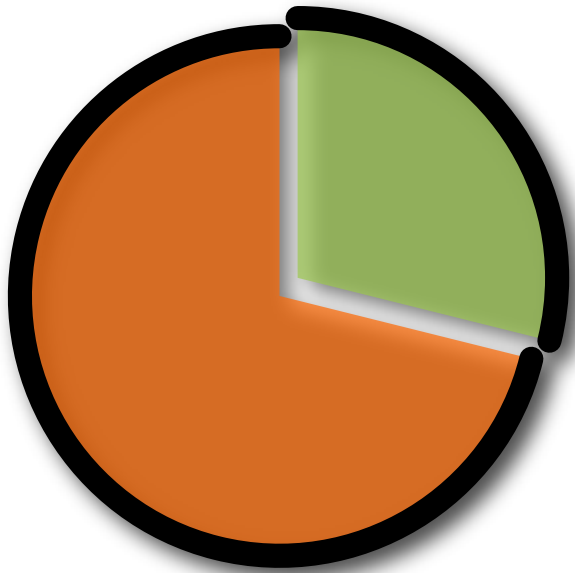
Trichuriasis	1,006
Japanese encephalitis	709
Chagas Disease*	667
Dengue*	616
Onchocerciasis*	484
Leprosy*	199
Diphtheria	185
Poliomyelitis	151
Hookworm disease	59

Disease data taken from WHO, *World Health Report 2004*

DALY - Disability adjusted life year in 1000’s.

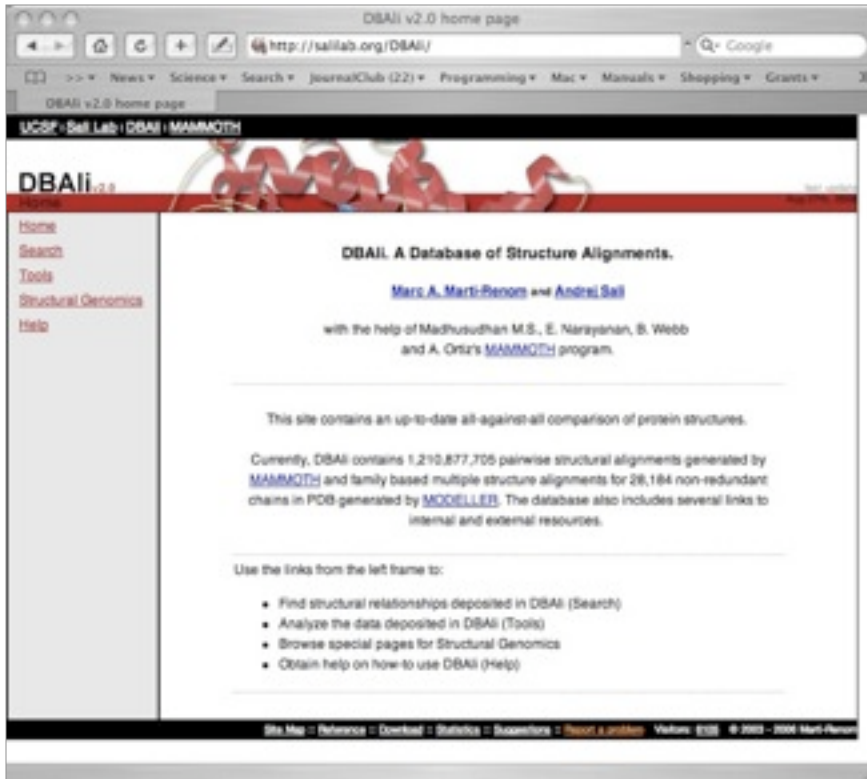
* Officially listed in the WHO Tropical Disease Research [disease portfolio](#).

Predicting binding sites in protein structure models.



DBAli_{v2.0} database

<http://www.dbali.org>



- ✓ Fully-automatic
- ✓ Data is kept up-to-date with PDB releases
- ✓ Tools for “on the fly” classification of families.
- ✓ Easy to navigate
- ✓ Provides tools for structure analysis

Does not provide a stable classification similar to that of CATH or SCOP

Pairwise structure alignments	
Last update:	October 6th, 2007
Number of chains:	96,804
Number of structure-structure comparisons:	1,748,371,897
Multiple structure alignments	
Last update:	August 1st, 2007
Number of representative chains:	34,637
Number of families:	12,732

Uses MAMMOTH for similarity detection

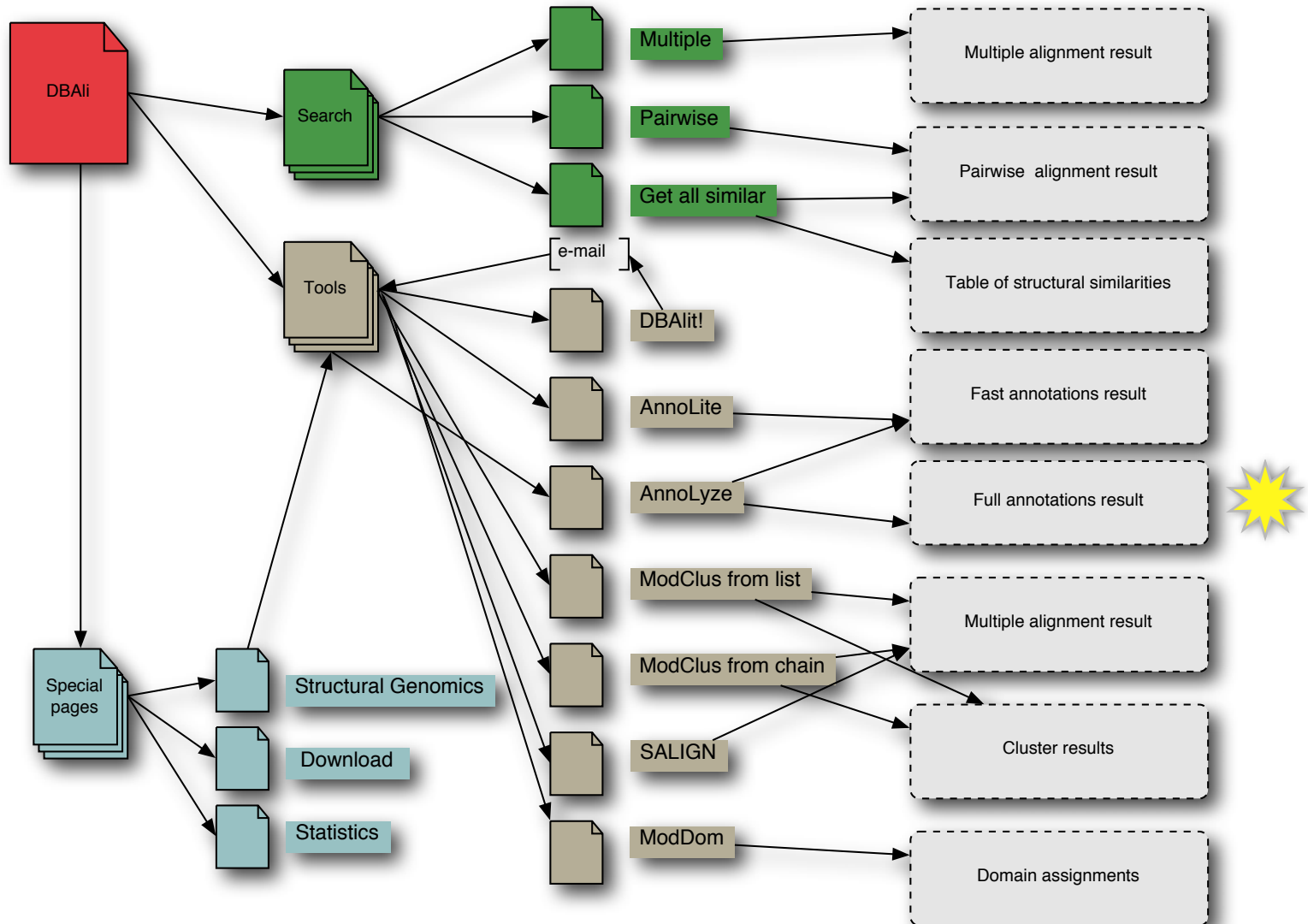
- ✓ VERY FAST!!!
- ✓ Good scoring system with significance

Ortiz AR, (2002) *Protein Sci.* 11 pp2606

Marti-Renom et al. 2001. *Bioinformatics.* 17, 746

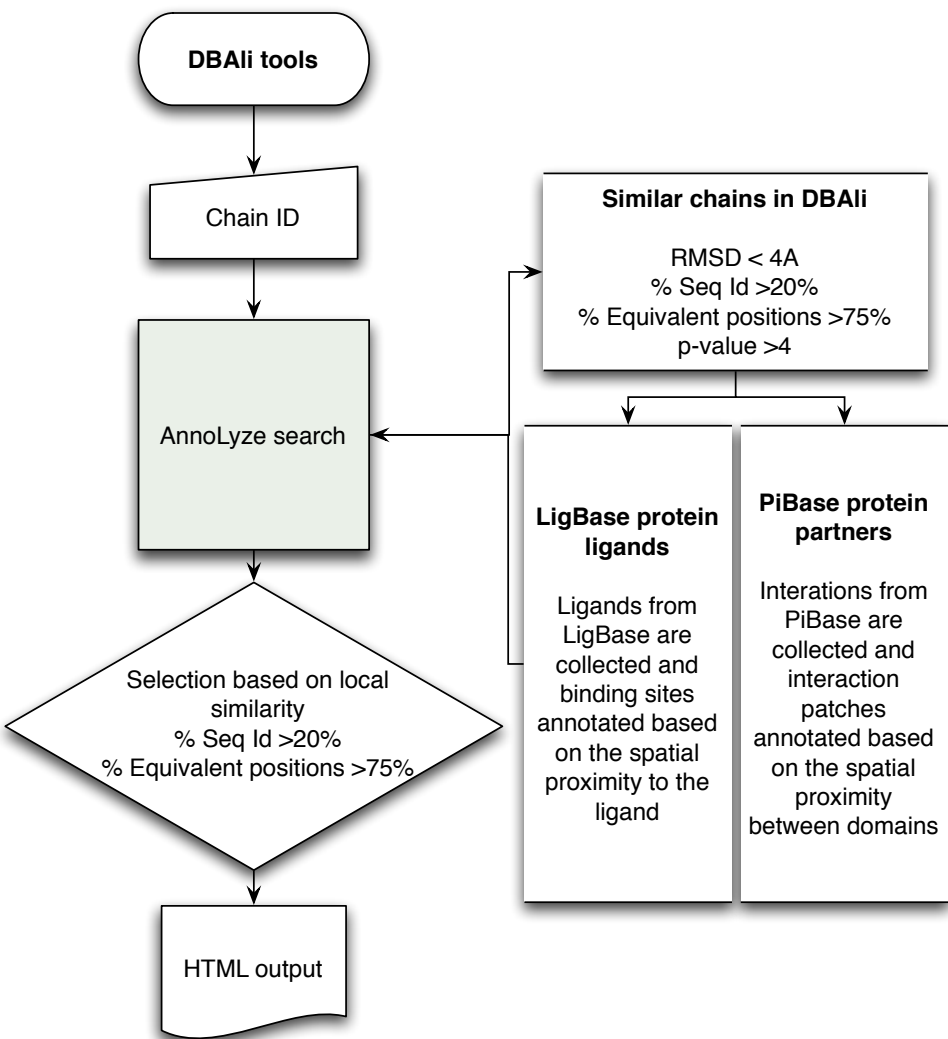
DBAli_{v2.0} database

<http://www.dbali.org>



Marti-Renom et al. BMC Bioinformatics (2007) Volume 8. Suppl S4

Method



Inherited ligands: 4

Ligand	Av. binding site seq. id.	Av. residue conservation	Residues in predicted binding site (size proportional to the local conservation)
MO2	59.03	0.183	48 49 52 62 63 66 67 113 116
CRY	20.00	0.111	23 29 31 37 44 48 49 63 85 94 96 103 121
BOG	20.00	0.111	19 20 21 48 49 51 96 98 136
ACY	15.87	0.163	23 29 31 37 44 45 81 83 85 94 96 98 103 121 135



Inherited partners: 1

Partner	Av. binding site seq. id.	Av. residue conservation	Residues in predicted binding site (size proportional to the local conservation)
d.113.1.1	23.68	0.946	19 20 50 51 52 53 54 55 56 57 58 77 78 79 80 81 82 83 84 85 93 95 97 99 134 135 138 142 145



Sensitivity .vs. Precision

	Optimal cut-off	Sensitivity (%) Recall or TPR	Precision (%)
Ligands	30%	71.9	13.7

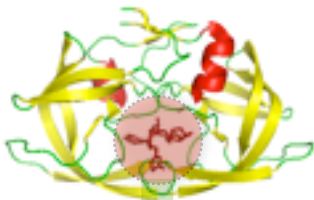
$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad \text{Precision} = \frac{TP}{TP + FP}$$

~90-95% of residues correctly predicted

Comparative docking

Expansion

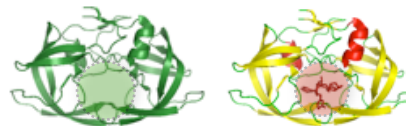
co-crystallized protein/ligand



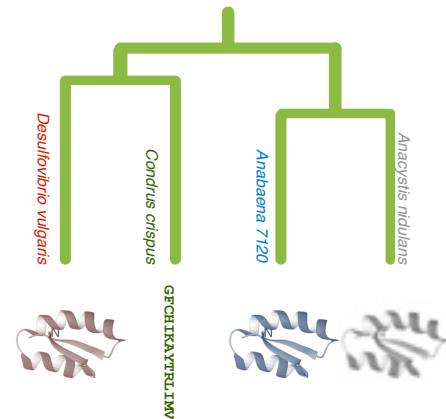
crystallized
protein

2. Inheritance

model



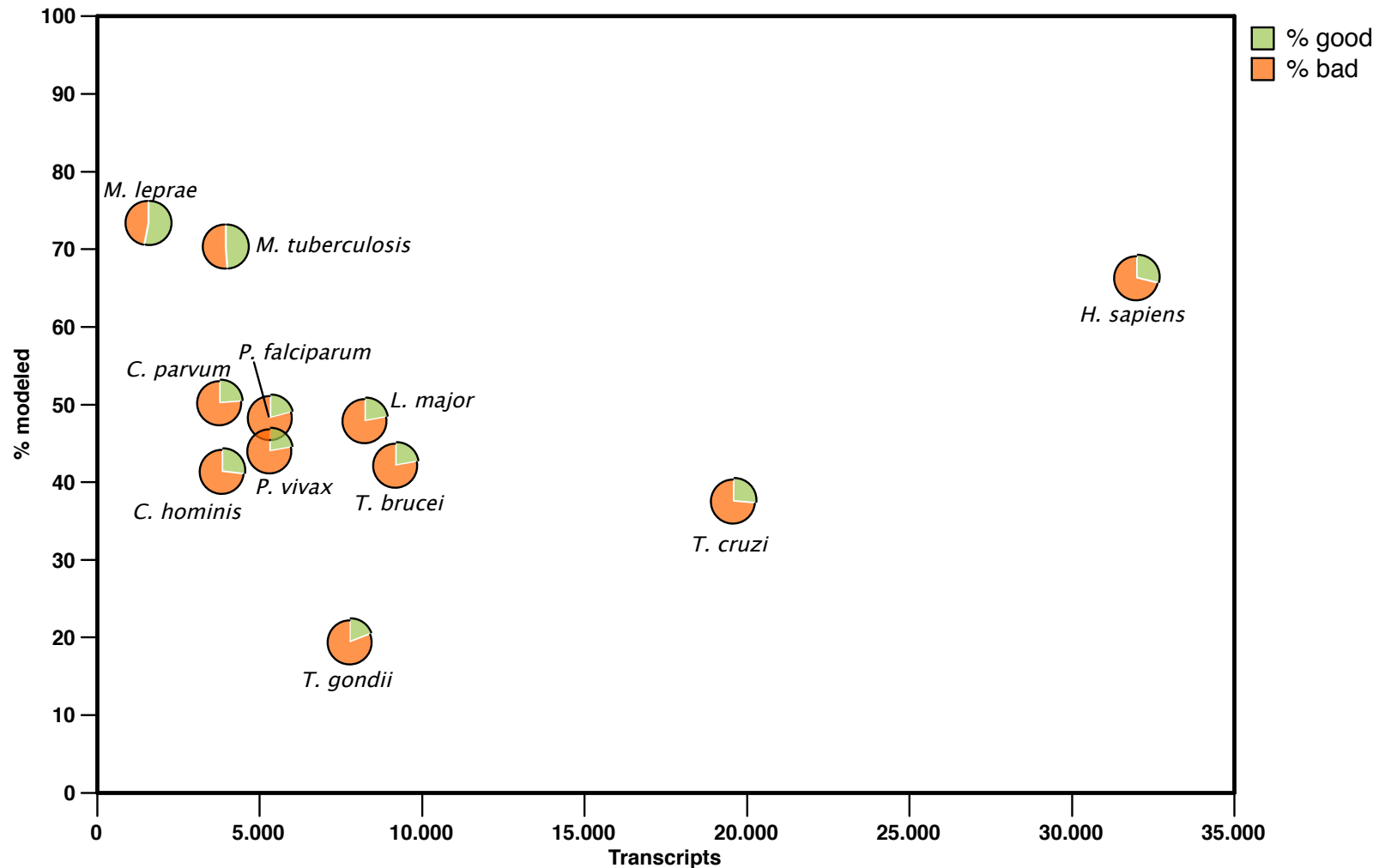
template



1. Modeling

Modeling Genomes

data from models generated by ModPipe (Eswar, Pieper & Sali)



A good model has MPQS of 1.0 or higher

Summary table

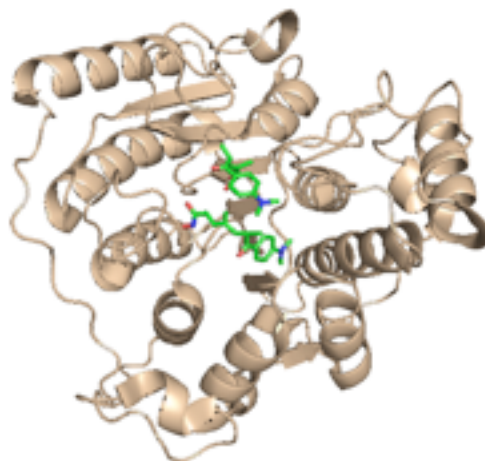
models with inherited ligands

29,271 targets with good models, 297 inherited a ligand/substance similar to a known drug in DrugBank

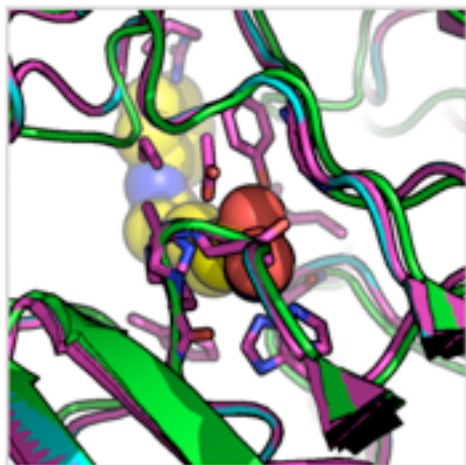
	Transcripts	Modeled targets	Selected models	Inherited ligands	Similar to a drug	Drugs
<i>C. hominis</i>	3,886	1,614	666	197	20	13
<i>C. parvum</i>	3,806	1,918	742	232	24	13
<i>L. major</i>	8,274	3,975	1,409	478	43	20
<i>M. leprae</i>	1,605	1,178	893	310	25	6
<i>M. tuberculosis</i>	3,991	2,808	1,608	365	30	10
<i>P. falciparum</i>	5,363	2,599	818	284	28	13
<i>P. vivax</i>	5,342	2,359	822	268	24	13
<i>T. brucei</i>	7,793	1,530	300	138	13	6
<i>T. cruzi</i>	19,607	7,390	3,070	769	51	28
<i>T. gondii</i>	9,210	3,900	1,386	458	39	21
TOTAL	68,877	29,271	11,714	3,499	297	143

L. major Histone deacetylase 2 + Vorinostat

Template 1t64A a human HDAC8 protein.



PDB	EO	Template	Seq	Model		Ligand	Exact	SupStr	SubStr	Similar
1c3sA	83.33/80.00	1t64A	36.00/1.47	LmjF21.0680.1.pdb	90.91/100.00	SHH	DB02546	DB02546	DB02546	DB02546



[DB02546](#) Vorinostat

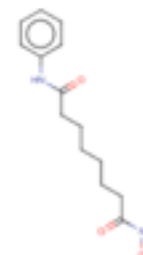
Small Molecule; Approved; Investigational

Drug categories:

Anti-Inflammatory Agents, Non-Steroidal
Anticarcinogenic Agents
Antineoplastic Agents
Enzyme Inhibitors

Drug indication:

For the treatment of cutaneous manifestations in patients with cutaneous T-cell lymphoma who have progressive, persistent or recurrent disease on or following two systemic therapies.



L. major Histone deacetylase 2 + Vorinostat

Literature

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 13143–13147, November 1996
Medical Sciences

Apicidin: A novel antiprotozoal agent that inhibits parasite histone deacetylase

(cyclic tetrapeptide/*Apicomplexa*/antiparasitic/malaria/coccidiosis)

SANDRA J. DARKIN-RATTRAY*[†], ANNE M. GURNETT*, ROBERT W. MYERS*, PAULA M. DULSKI*, TAMI M. CRUMLEY*, JOHN J. ALLOCCO*, CHRISTINE CANNOVA*, PETER T. MEINKE[‡], STEVEN L. COLLETTI[‡], MARIA A. BEDNAREK[‡], SHEO B. SINGH[§], MICHAEL A. GOETZ[§], ANNE W. DOMBROWSKI[§], JON D. POLISHOOK[§], AND DENNIS M. SCHMATZ*

Departments of *Parasite Biochemistry and Cell Biology, [‡]Medicinal Chemistry, and [§]Natural Products Drug Discovery, Merck Research Laboratories, P.O. Box 2000, Rahway, NJ 07065

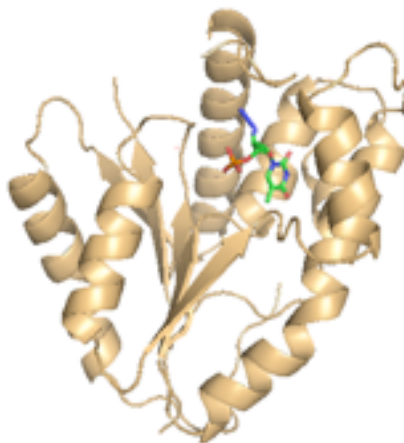
ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Apr. 2004, p. 1435–1436
0066-4804/04/\$08.00+0 DOI: 10.1128/AAC.48.4.1435–1436.2004
Copyright © 2004, American Society for Microbiology. All Rights Reserved.

Vol. 48, No. 4

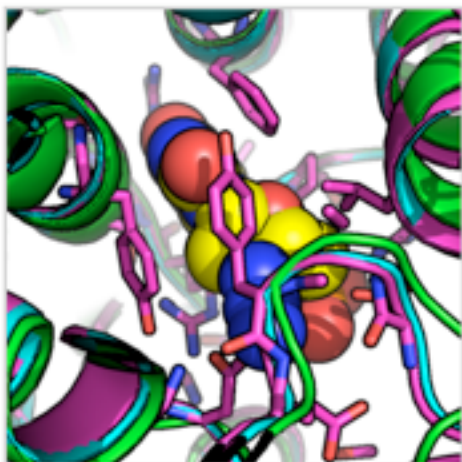
Antimalarial and Antileishmanial Activities of Aroyl-Pyrrolyl-Hydroxyamides, a New Class of Histone Deacetylase Inhibitors

P. falciparum thymidylate kinase + zidovudine

Template 3tmkA a yeast thymidylate kinase.



PDB	iQ	Template	iQ	Model	iQ	Ligand	Exact	SupStr	SubStr	Similar
2tmkB	100.00/100.00	3tmkA	41.00/1.49	PFL2465c.2.pdb	82.61/100.00	ATM		DB00495		DB00495



[DB00495](#) Zidovudine

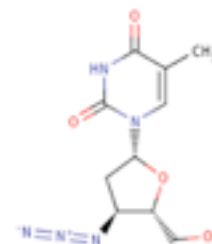
Small Molecule; Approved

Drug categories:

Anti-HIV Agents
Antimetabolites
Nucleoside and Nucleotide Reverse Transcriptase Inhibitors

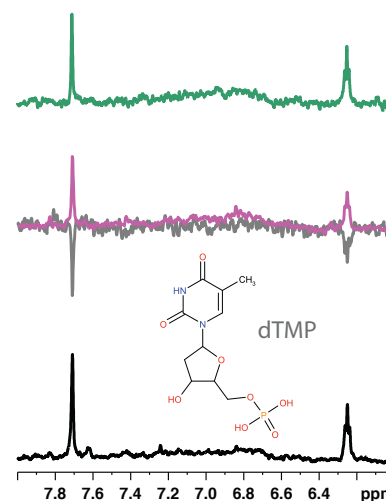
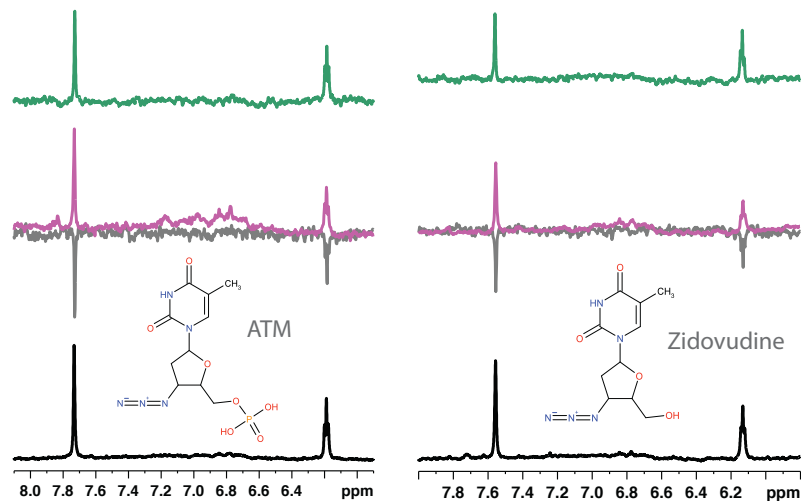
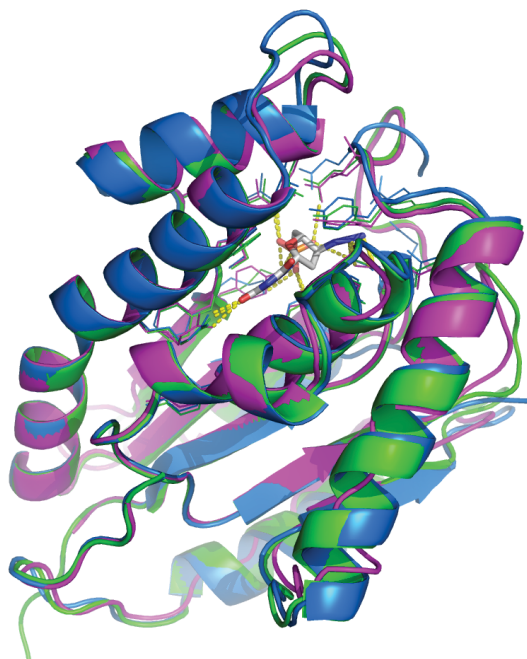
Drug indication:

For the treatment of human immunovirus (HIV) infections.



P. falciparum thymidylate kinase + zidovudine

NMR Water-LOGSY and STD experiments



Leticia Ortí, Rodrigo J. Carbajo, and Antonio Pineda-Lucena

TDI's kernel

<http://tropicaldisease.org/kernel>

Ortí et al . “A kernel for open source drug discovery in tropical diseases”. Submitted.
Ortí et al . “A Kernel for the Tropical Disease Initiative”. Submitted.

the Tropical Disease Initiative
an open source drug discovery project

You are browsing version 1.0 (2008/05/01) of the TDI Kernel.

Posted on 05/07/08 to Target. Only the feed. No comments yet. Add your thoughts or trackback from your own site. Edit this entry.

Putative histone deacetylase, predicted to bind 1 ligands [SHH]

UniPort id: Q9GU59 [C. parvum]

Target keywords: Anticarcinogenic Agents, Antineoplastic Agents, Transcription, Chromatin regulator, Anti-inflammatory Agents, Non-Steroidal, Enzyme Inhibitors, Q9GU59, Transcription regulation, Nucleus

Do you consider this target suitable for drug discovery? 0 0 0 0 0 (No Ratings Yet)

Binding site prediction to approved drugs (need help reading this page?):

PDB	ID	Template	dt	Model	Ligand	Exact	Super	Subst	Similar
1c3aA	41.1240.00	554A	37.001.47	cpd_1385.1.pdb	SHH	0802346	0802346	0802346	0802346

0802346 Vorinostat
Small Molecule, Approved, Investigational

Drug categories:
Anti-inflammatory Agents, Non-Steroidal
Anticarcinogenic Agents
Antineoplastic Agents
Enzyme Inhibitors

Drug Indication:
For the treatment of cutaneous manifestations in patients with cutaneous T-cell lymphoma who have progressive, persistent or recurrent disease on or following two systemic therapies.

Shown ligand SHH
OCTANE(D)CACC(Y)ORCKYMDOPHENYLAMIDE
expanded from [SHH] to template [SHH] used for building a 3D model of cpd_1385.1.pdb. Download the coordinates: data/0802346/0802346.001.001.pdb

2008 - Open Access.
Powered by WordPress.
Theme by Upstart Blogger.

Acknowledgments

<http://sgu.bioinfo.cipf.es>

<http://tropicaldisease.org>



COMPARATIVE MODELING

Andrej Sali

Ursula Pieper
M. S. Madhusudhan
Narayanan Eswar
Min-Yi Shen
Ben Webb
Maya Topf



b-GLOBIN DOMAIN

Job Dekker

Bryan Lajoie
Ye Zhan
Mark Umbarger



FUNDING

Centro de Investigación Príncipe Felipe
Ministerio de Educación y Ciencia
STREP UE Grant
Marie Curie Reintegration Grant

MODEL ASSESSMENT

Francisco Melo (CU)
Alejandro Panjkovich (CU)

NMR

Antonio Pineda-Lucena

Leticia Ortí
Rodrigo J. Carbajo

MAMMOTH

Angel R. Ortiz

FUNCTIONAL ANNOTATION

Fatima Al-Shahrour
Joaquin Dopazo

BIOLOGY

Jeff Friedman (RU)
James Hudsped (RU)
Partho Ghosh (UCSD)
Alvaro Monteiro (Cornell U)
Stephen Krilis (St. George H)

Tropical Disease Initiative

Stephen Maurer (UC Berkeley)
Arti Rai (Duke U)
Andrej Sali (UCSF)
Ginger Taylor (TSL)
Matthew Todd (U Sydney)

CCPR Functional Proteomics

Patsy Babbitt (UCSF)
Fred Cohen (UCSF)
Ken Dill (UCSF)
Tom Ferrin (UCSF)
John Irwin (UCSF)
Matt Jacobson (UCSF)
Tack Kuntz (UCSF)
Andrej Sali (UCSF)
Brian Shoichet (UCSF)
Chris Voigt (UCSF)

EVA

Burkhard Rost (Columbia U)
Alfonso Valencia (CNB/UAM)

CAMP

Xavier Aviles (UAB)
Hans-Peter Nester (SANOFI)
Ernst Meinjohanns (ARPIDA)
Boris Turk (IJS)
Markus Gruetter (UE)
Matthias Wilmanns (EMBL)
Wolfram Bode (MPG)