Proteomics & Bioinformatics Part II

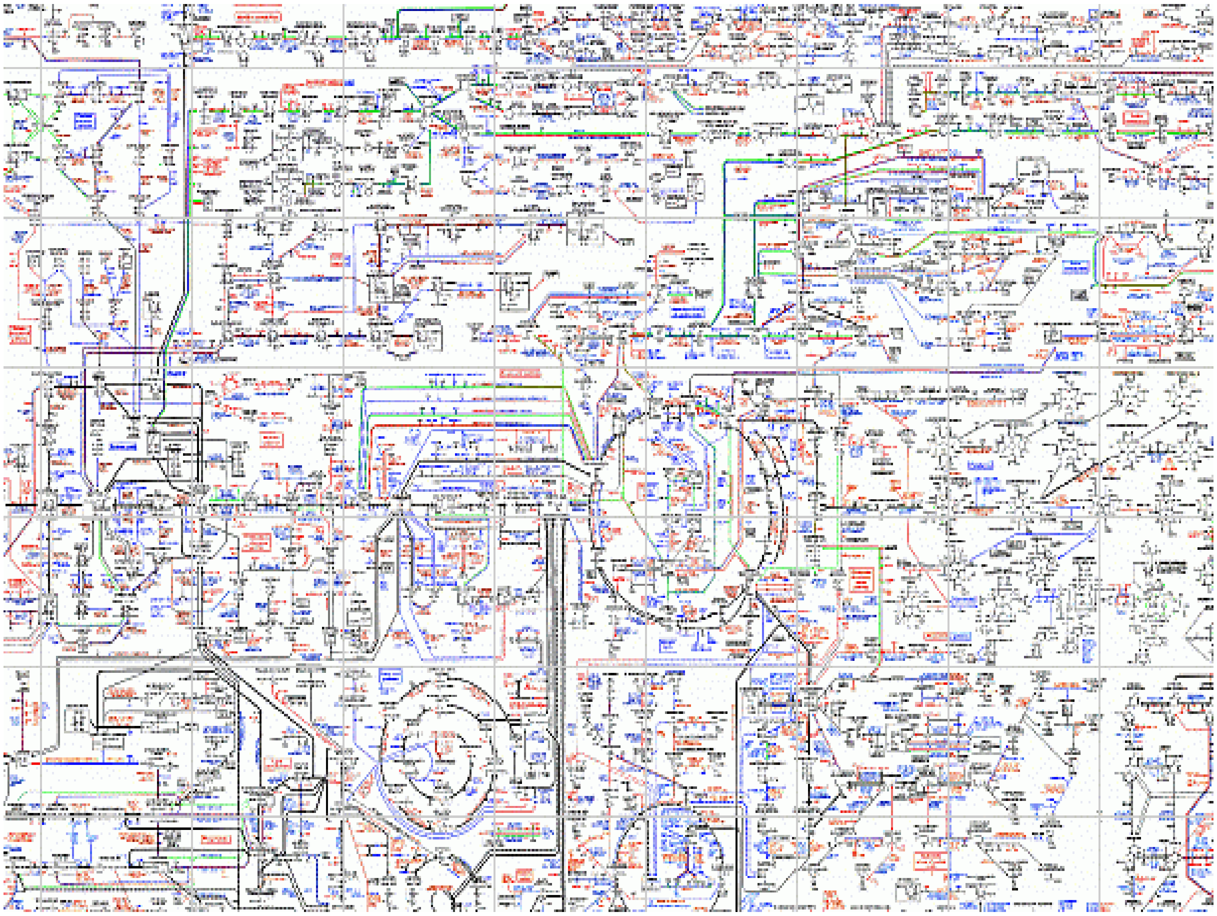
David Wishart
University of Alberta

3 Kinds of Proteomics

- Structural Proteomics
  - High throughput X-ray Crystallography/Modelling
  - High throughput NMR Spectroscopy/Modelling
- Expressional or Analytical Proteomics
  - Electrophoresis, Protein Chips, DNA Chips, 2D-HPLC
  - Mass Spectrometry, Microsequencing
- Functional or Interaction Proteomics
  - HT Functional Assays, Ligand Chips
  - Yeast 2-hybrid, Deletion Analysis, Motif Analysis

Historically...

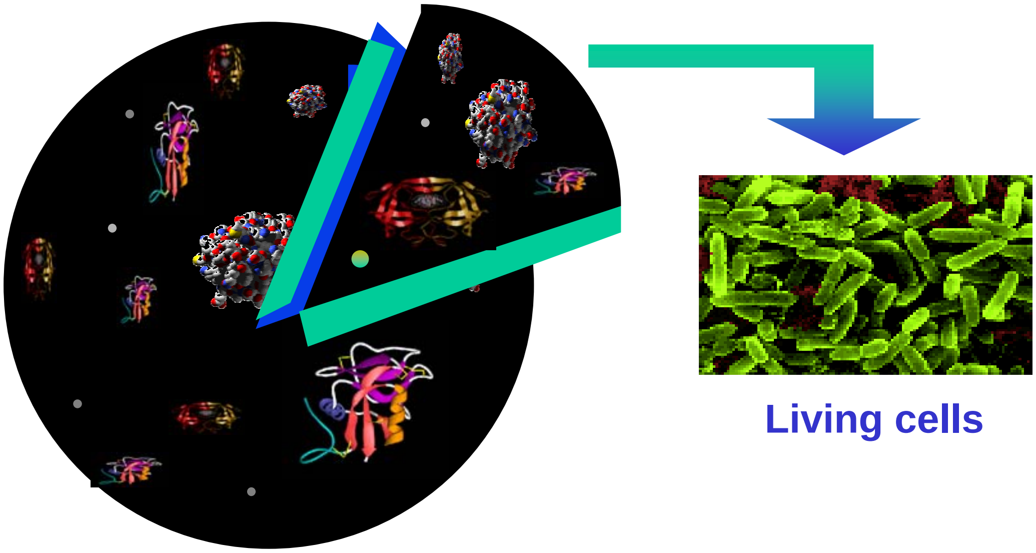
- Most of the past 100 years of biochemistry has focused on the analysis of small molecules (i.e. metabolism and metabolic pathways)
- These studies have revealed much about the processes and pathways for about 400 metabolites which can be summarized with this...



More Recently...

- Molecular biologists and biochemists have focused on the analysis of larger molecules (proteins and genes) which are much more complex and much more numerous
- These studies have primarily focused on identifying and cataloging these molecules (Human Genome Project)

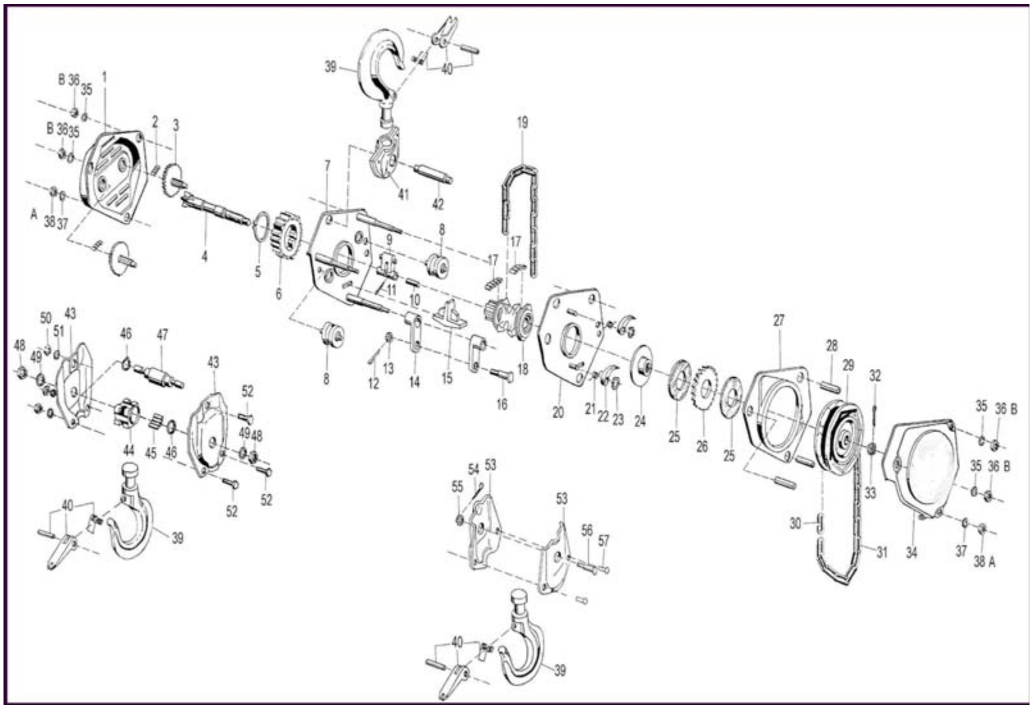
Nature’s Parts Warehouse



Living cells

The protein universe

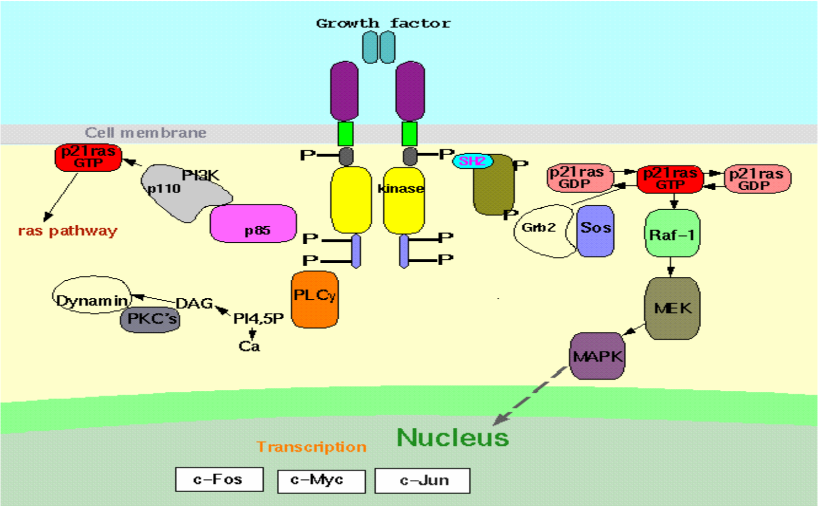
The Protein Parts List



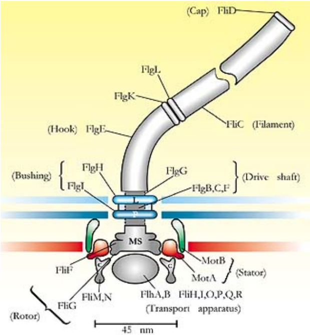
However...

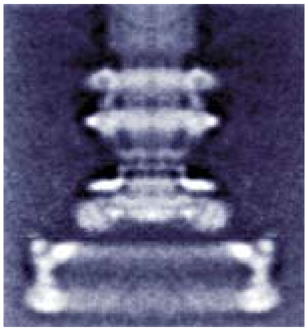
- This cataloging (which consumes most of bioinformatics) has been derogatively referred to as “stamp collecting”
- Having a collection of parts and names doesn’t tell you how to put something together or how things connect -- *this is biology*

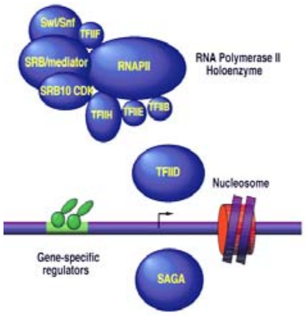
Remember: *Proteins Interact*

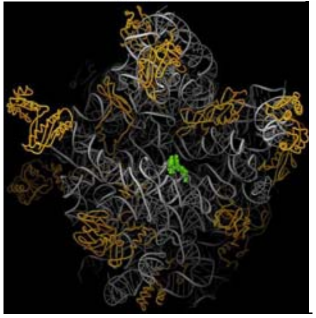


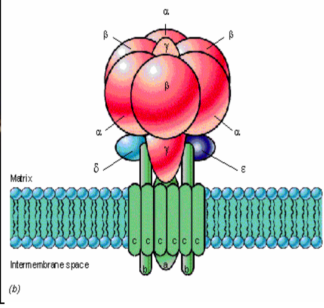
Proteins *Assemble*

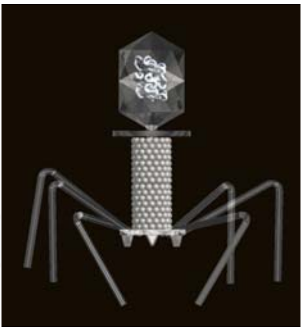






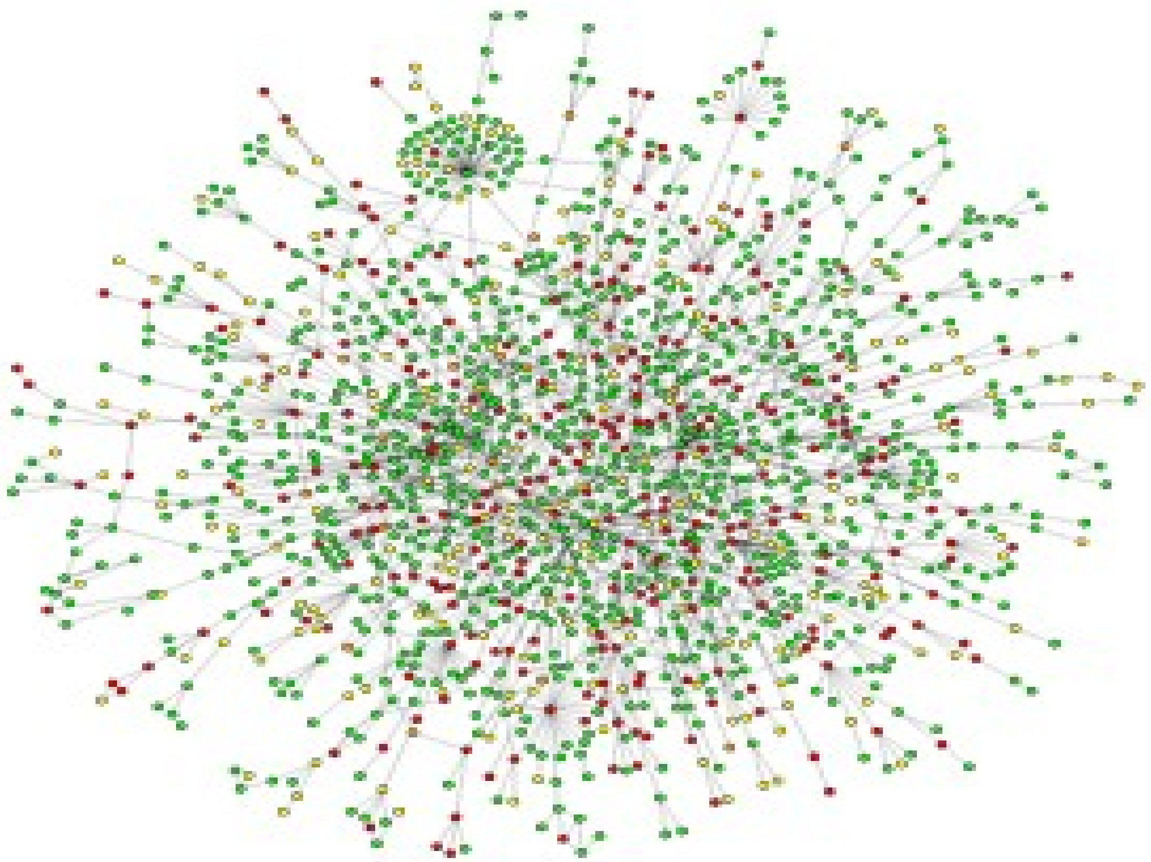






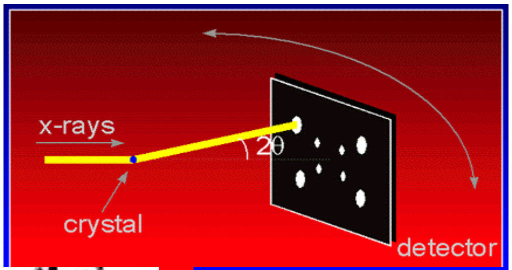
For the Past 10 Years...

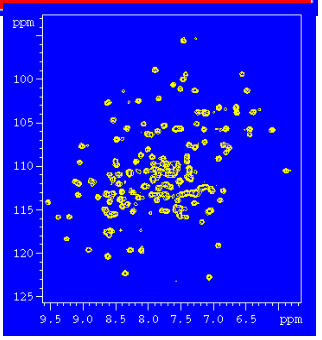
- Scientists have increasingly focused on “signal transduction” and transient protein interactions
- New techniques have been developed which reveal which proteins and which parts of proteins are important for interaction
- The hope is to get something like this..



Protein Interaction Tools and Techniques - Experimental Methods

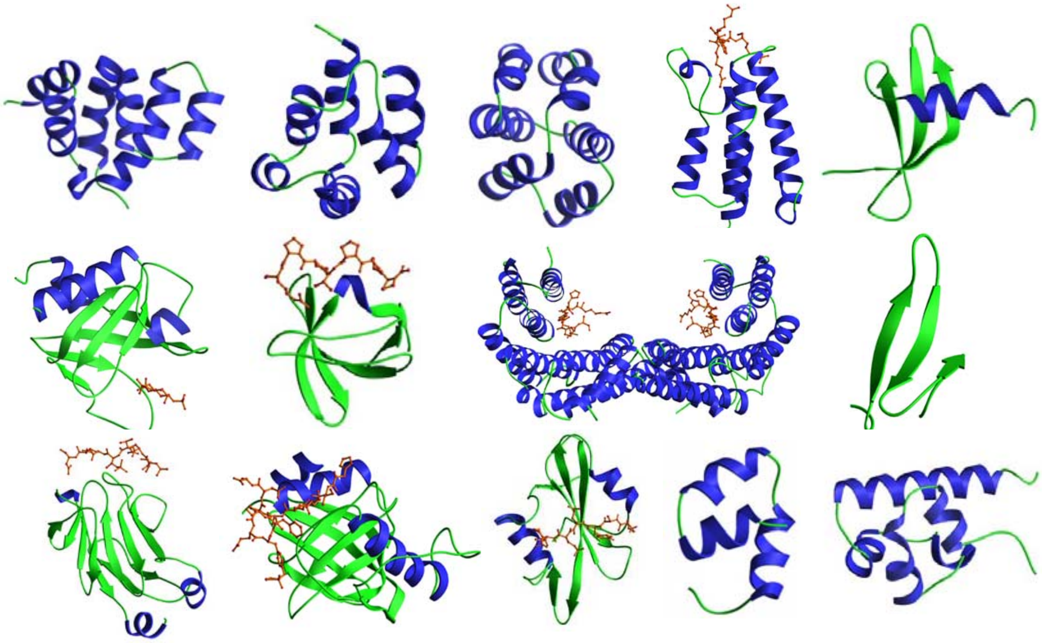
3D Structure Determination





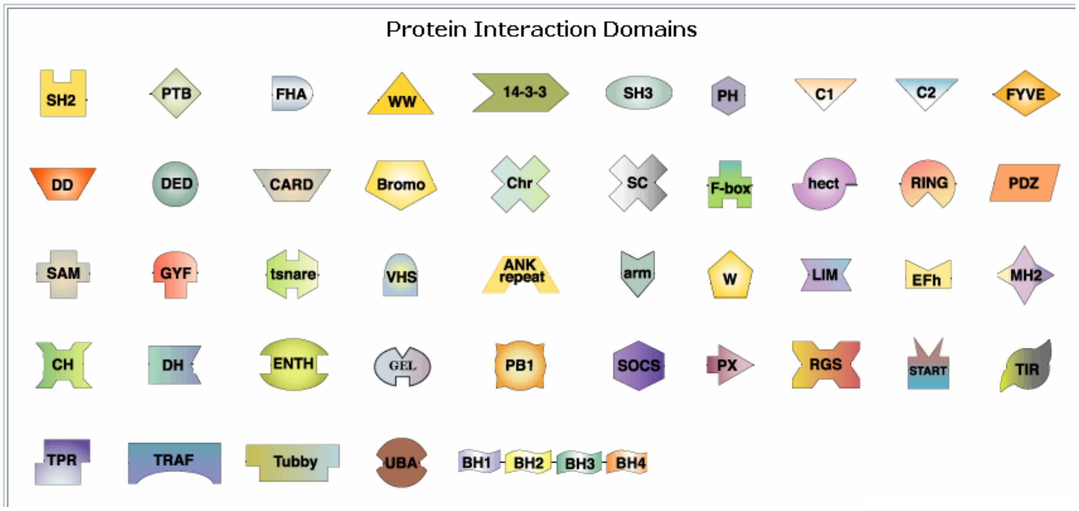
- X-ray crystallography
  - grow crystal
  - collect diffract. data
  - calculate e- density
  - trace chain
- NMR spectroscopy
  - label protein
  - collect NMR spectra
  - assign spectra & NOEs
  - calculate structure using distance geom.

Protein Interaction Domains



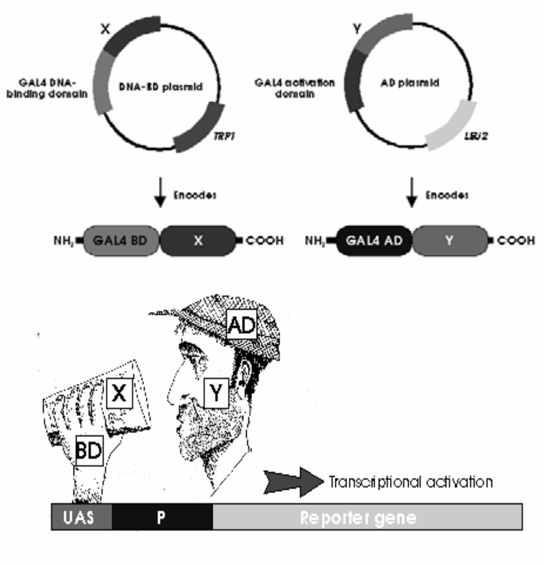
http://www.mshri.on.ca/pawson/domains.html

Protein Interaction Domains



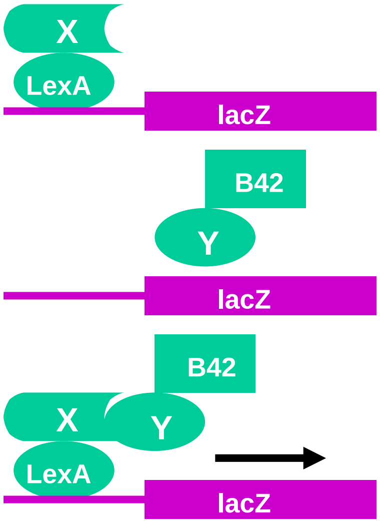
http://www.mshri.on.ca/pawson/domains.html

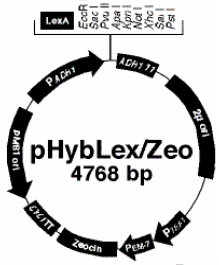
Yeast Two-Hybrid Analysis

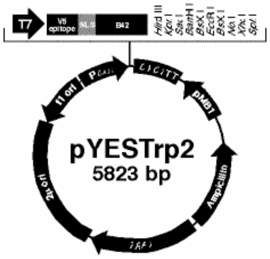


- Yeast two-hybrid experiments yield information on protein protein interactions
- GAL4 Binding Domain
- GAL4 Activation Domain
- X and Y are two proteins of interest
- If X & Y interact then reporter gene is expressed

Invitrogen Yeast 2-Hybrid







Example of 2-Hybrid Analysis

- Uetz P. et al., “*A Comprehensive Analysis of Protein-Protein Interactions in Saccharomyces cerevisiae*” Nature 403:623-627 (2000)
- High Throughput Yeast 2 Hybrid Analysis
- 957 putative interactions
- 1004 of 6000 predicted proteins involved

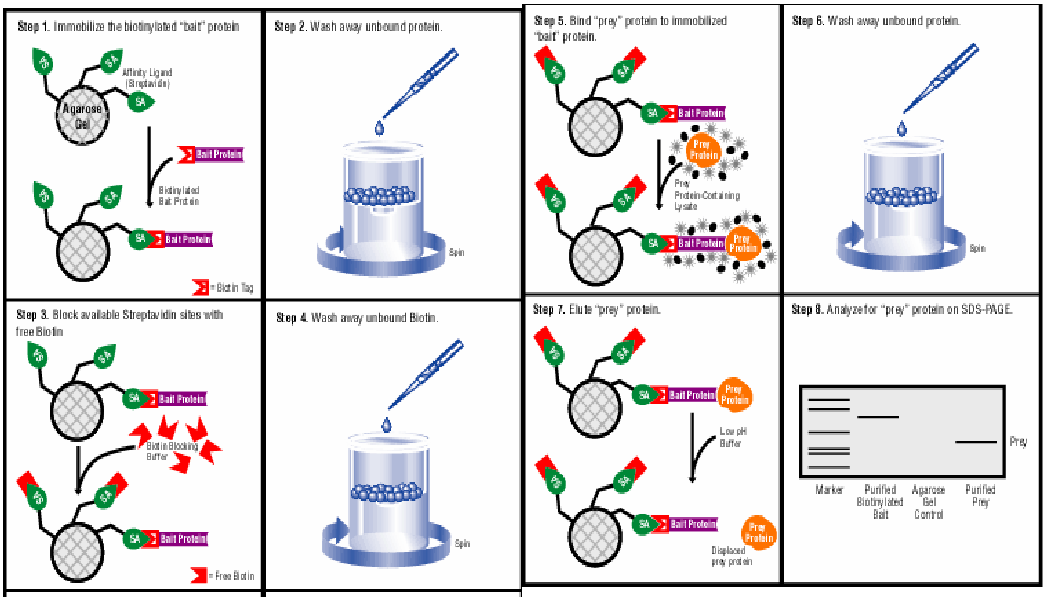
Example of 2-Hybrid Analysis

- Rain JC. et al., “*The protein-protein interaction map of Helicobacter pylori*” Nature 409:211-215 (2001)
- High Throughput Yeast 2 Hybrid Analysis
- 261 H. pylori proteins scanned against genome
- >1200 putative interactions identified
- Connects >45% of the H. pylori proteome

Another Way?

- Ho Y, Gruhler A, et al. *Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry*. Nature 415:180-183 (2002)
- High Throughput Mass Spectral Protein Complex Identification (HMS-PCI)
- 10% of yeast proteins used as “bait”
- 3617 associated proteins identified
- 3 fold higher sensitivity than yeast 2-hybrid

Affinity Pull-down



Step 1. Immobilize the biotinylated "bait" protein

Step 2. Wash away unbound protein.

Step 3. Block available Streptavidin sites with free Biotin

Step 4. Wash away unbound Biotin.

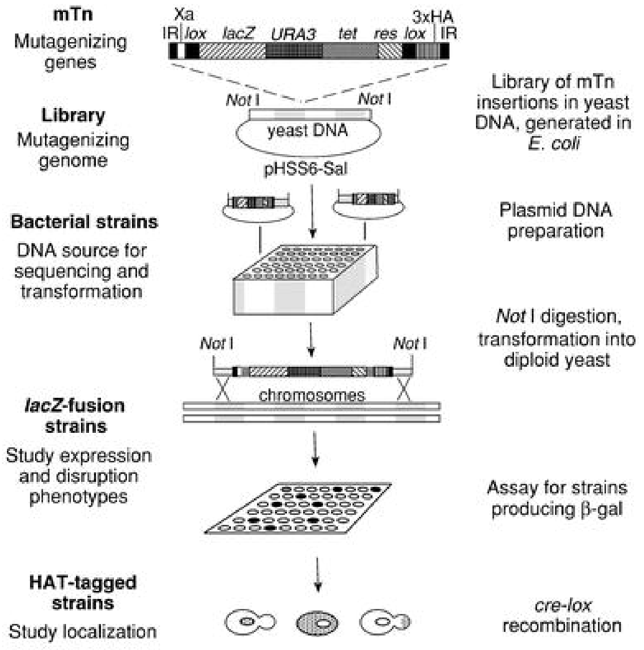
Step 5. Bind “prey” protein to immobilized “bait” protein.

Step 6. Wash away unbound protein.

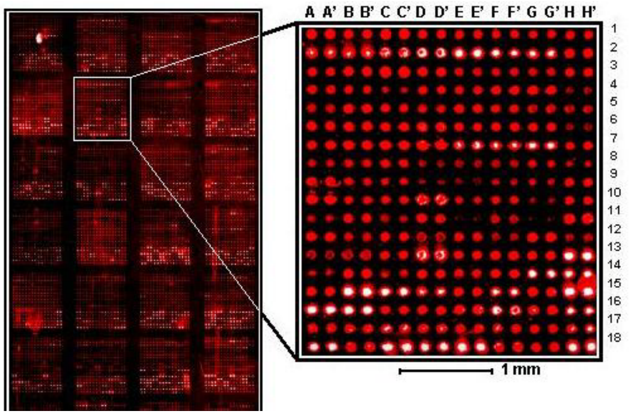
Step 7. Elute “prey” protein.

Step 8. Analyze for “prey” protein on SDS-PAGE.

Transposon Tagging

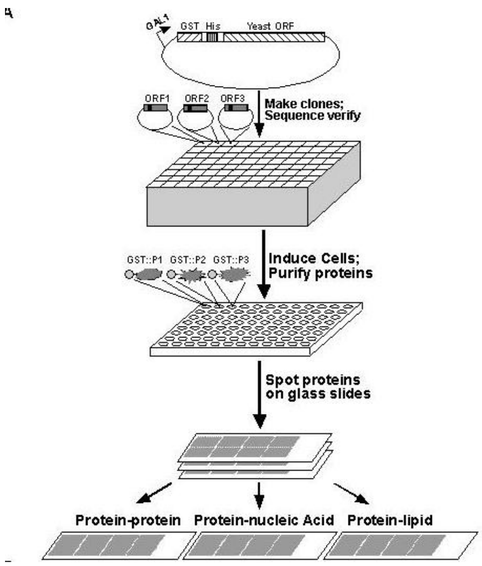


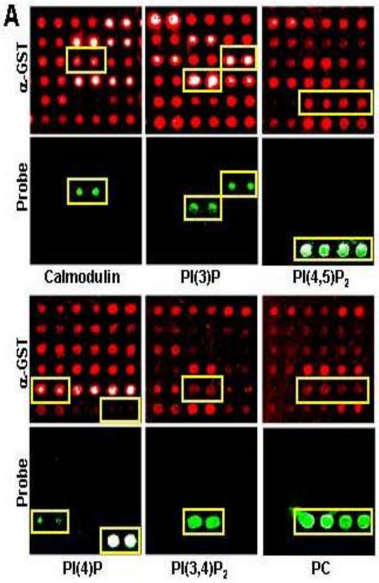
Protein Arrays



H Zhu, J Klemic, S Chang, P Bertone, A Casamayor, K Klemic, D Smith, M Gerstein, M Reed, & M Snyder (2000).**Analysis of yeast protein kinases using protein chips**. Nature Genetics 26: 283-289

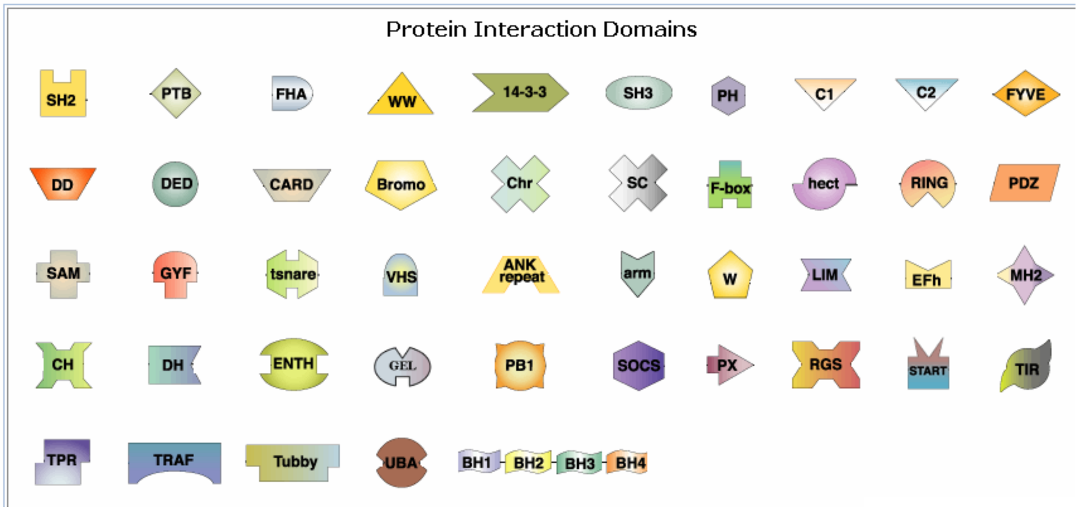
Protein Arrays





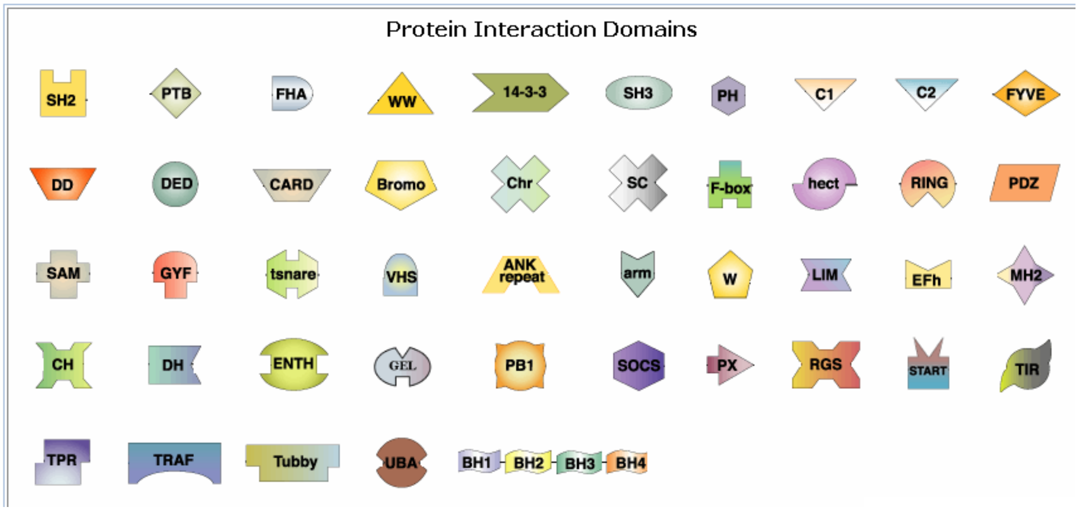
Protein Interaction Tools and Techniques - Computational Methods

Sequence Searching Against Known Domains



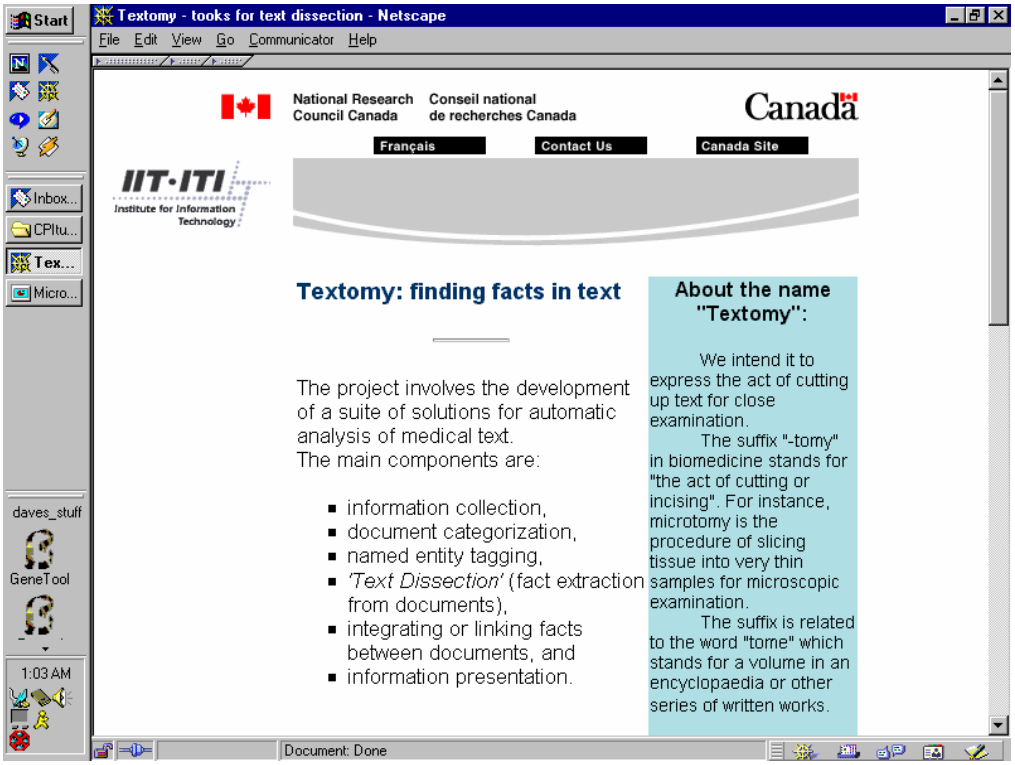
http://www.mshri.on.ca/pawson/domains.html

Motif Searching Using Known Motifs



Text Mining

- Searching Medline or Pubmed for words or word combinations
- “X binds to Y”; “X interacts with Y”; “X associates with Y” etc. etc.
- Requires a list of known gene names or protein names for a given organism
- Sometimes called “Textomy”


Textomy - tooks for text dissection - Netscape
File Edit View Go Communicator Help
Start
Inbox...
CPltu...
Tex...
Micro...
daves_stuff
GeneTool
1:03 AM
Document: Done

National Research Council Canada

Conseil national de recherches Canada

Canada

Français | Contact Us | Canada Site

IIT·ITI Institute for Information Technology

Textomy: finding facts in text

The project involves the development of a suite of solutions for automatic analysis of medical text.
The main components are:

- information collection,
- document categorization,
- named entity tagging,
- *'Text Dissection'* (fact extraction from documents),
- integrating or linking facts between documents, and
- information presentation.

About the name "Textomy":

We intend it to express the act of cutting up text for close examination.

The suffix "-tomy" in biomedicine stands for "the act of cutting or incising". For instance, microtomy is the procedure of slicing tissue into very thin samples for microscopic examination.

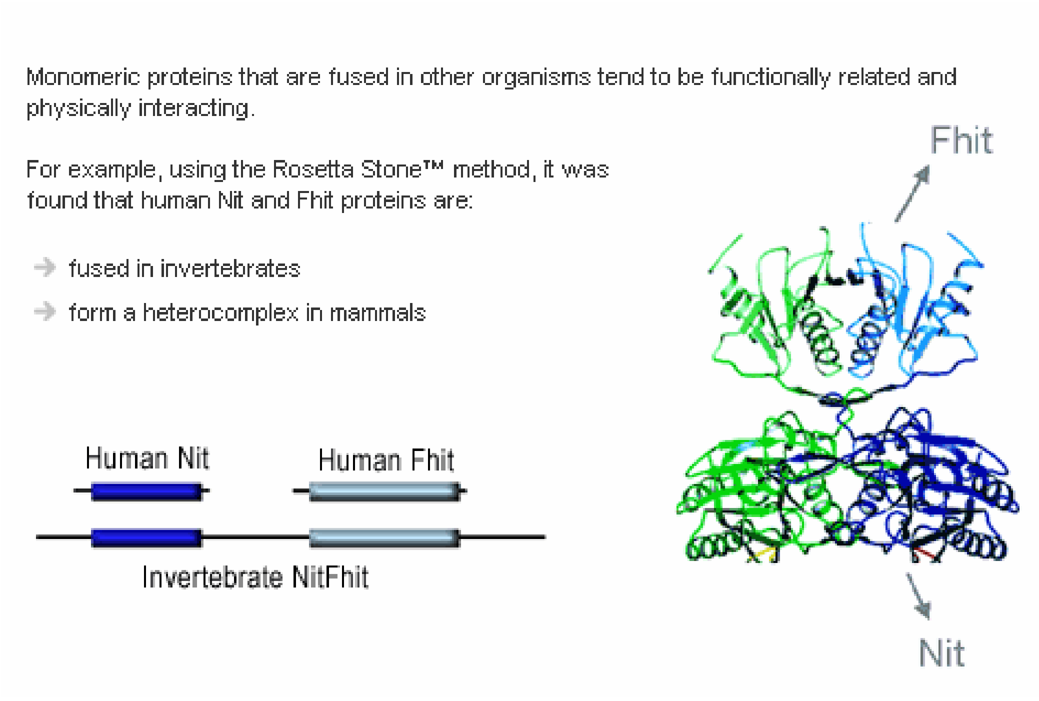
The suffix is related to the word "tome" which stands for a volume in an encyclopaedia or other series of written works.

http://textomy.iit.nrc.ca/

Pre-BIND

- *Donaldson et al. BMC Bioinformatics 2003 4:11*
- **Used Support Vector Machine (SVM) to scan literature for protein interactions**
- **Precision, accuracy and recall of 92% for correctly classifying PI abstracts**
- **Estimated to capture 60% of all abstracted protein interactions for a given organism**

Rosetta Stone Method



Monomeric proteins that are fused in other organisms tend to be functionally related and physically interacting.

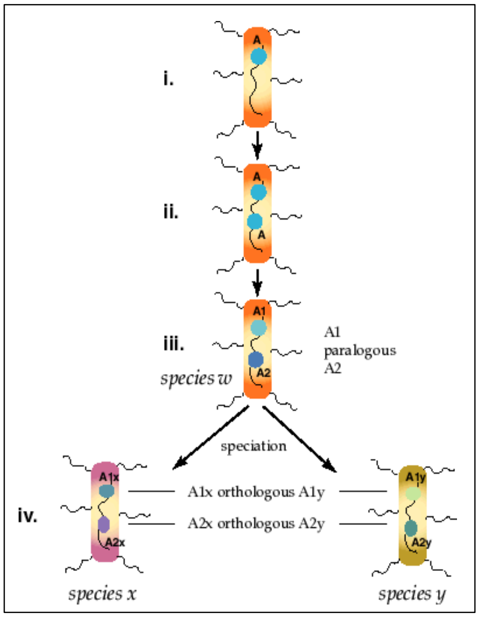
For example, using the Rosetta Stone™ method, it was found that human Nit and Fhit proteins are:

- fused in invertebrates
- form a heterocomplex in mammals

Interologs, Homologs, Paralogs...

- Homolog
  - Common Ancestors
  - Common 3D Structure
  - Common Active Sites
- Ortholog
  - Derived from Speciation
- Paralog
  - Derived from Duplication
- Interolog
  - Protein-Protein Interaction

YM2



Finding Interologs

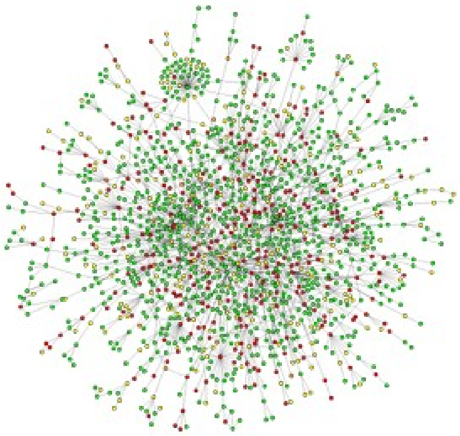
- If A and B interact in organism X, then if organism Y has a homolog of A (A’) and a homolog of B (B’) then A’ and B’ should interact too!
- Makes use of BLAST searches against entire proteome of well-studied organisms (yeast, E. coli)
- Requires list of known interacting partners

A Flood of Data

- High throughput techniques are leading to more and more data on protein interactions
- This is where bioinformatics can play a key role
- Some suggest that this is the “future” for bioinformatics

Interaction Databases

- BIND
  - http://www.blueprint.org/bind/bind.php
- DIP
  - http://dip.doe-mbi.ucla.edu/
- MINT
  - http://mint.bio.uniroma2.it/mint/
- PathCalling
  - http://portal.curagen.com/extpc/com.curagen.portal.servlet.Yeast

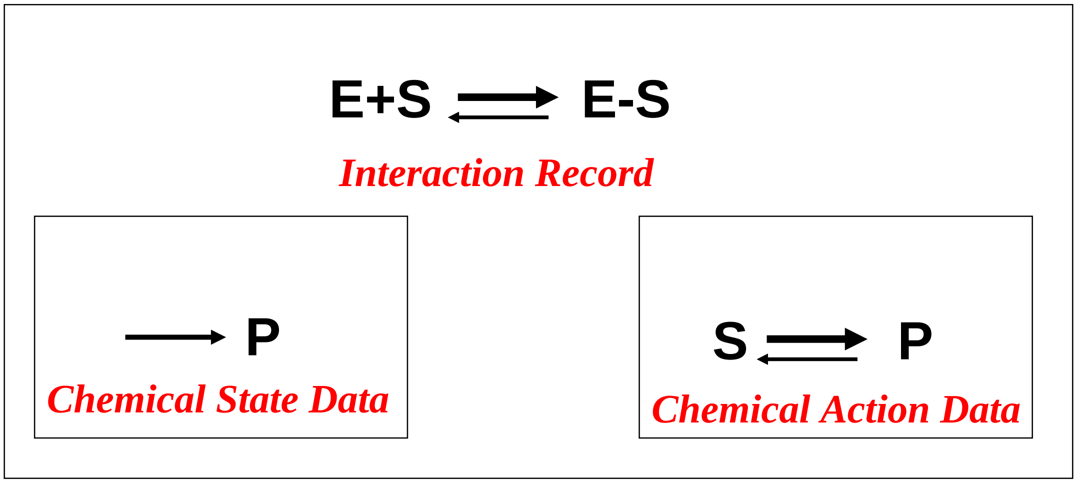


The BIND Database

- BIND - Biomolecular Interaction Network Database
- Conceived and Developed by Chris Hogue, Tony Pawson, Francis Ouellette
- Designed to capture almost all interactions between biomolecules (large and small)
- Largest database of its kind

BIND Data Model

$$S \xrightarrow{E} P$$



$E+S \rightleftharpoons E\text{-}S$

Interaction Record

$\longrightarrow P$

Chemical State Data

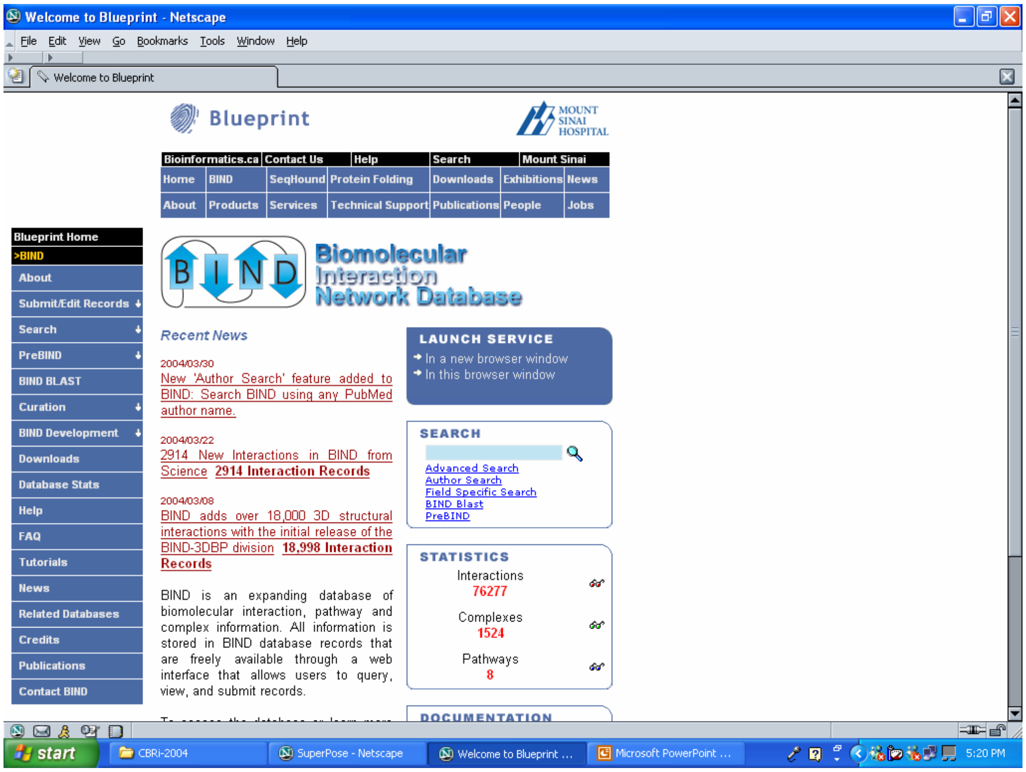
$S \rightleftharpoons P$

Chemical Action Data

BIND Can Encode...

- Simple binary interactions
- Enzymes, substrates and conformational changes
- Restriction enzymes
- Limited proteolysis
- Phosphorylation (reversible)
- Glycosylation
- Intron splicing
- Transcriptional factors

BIND



Recent News

2004/03/30
New 'Author Search' feature added to BIND: Search BIND using any PubMed author name.

2004/03/22
2914 New Interactions in BIND from Science **2914 Interaction Records**

2004/03/08
BIND adds over 18,000 3D structural interactions with the initial release of the BIND-3DBP division **18,998 Interaction Records**

BIND is an expanding database of biomolecular interaction, pathway and complex information. All information is stored in BIND database records that are freely available through a web interface that allows users to query, view, and submit records.

STATISTICS

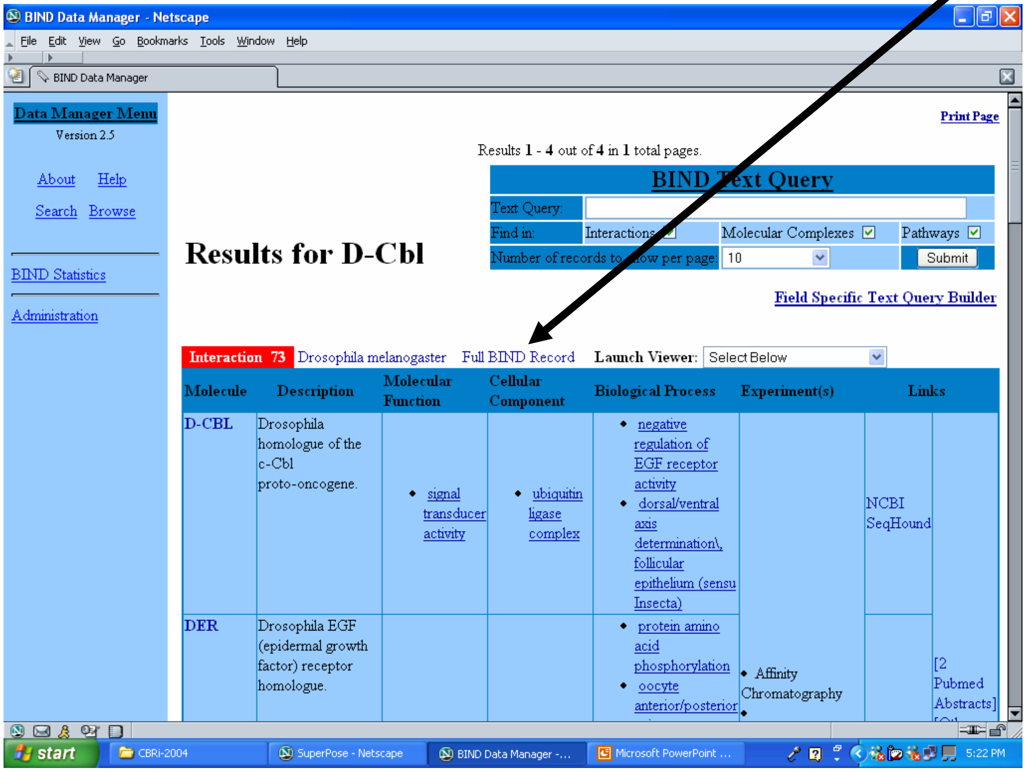
Interactions
76277

Complexes
1524

Pathways
8

BIND Query Result

click



Data Manager Menu
Version 2.5

About Help
Search Browse

BIND Statistics

Administration

Print Page

Results **1** - **4** out of **4** in **1** total pages.

BIND Text Query

Text Query:			
Find in:	Interactions ☑	Molecular Complexes ☑	Pathways ☑
Number of records to show per page:	10	Submit	

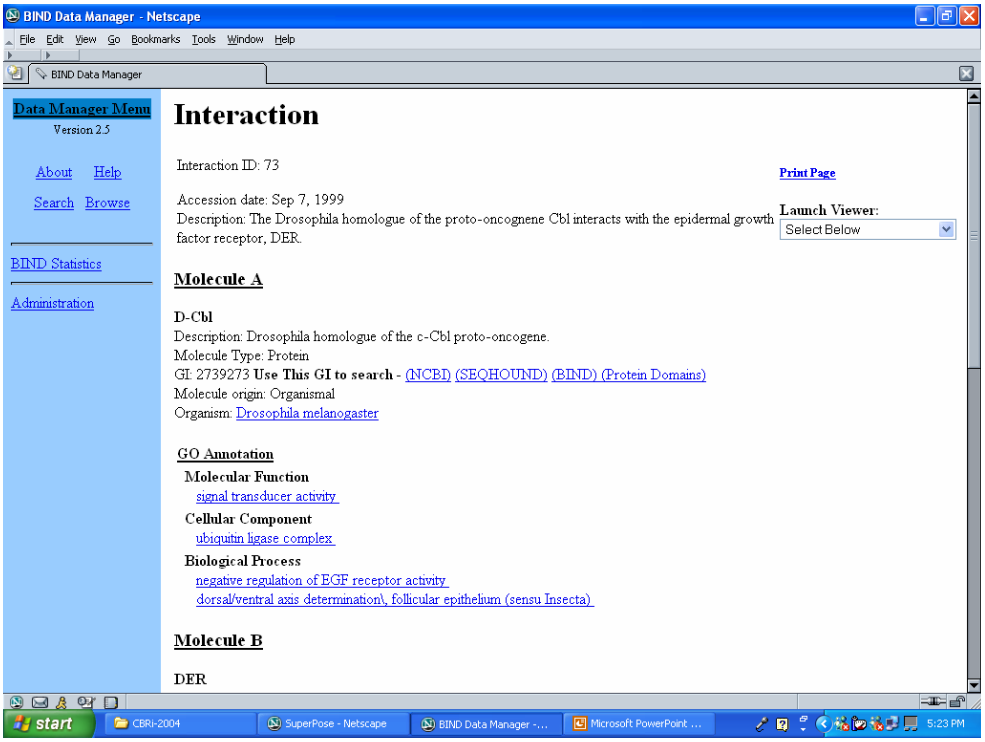
Field Specific Text Query Builder

Results for D-Cbl

Interaction 73 Drosophila melanogaster Full BIND Record **Launch Viewer:** Select Below

Molecule	Description	Molecular Function	Cellular Component	Biological Process	Experiment(s)	Links	
D-CBL	Drosophila homologue of the c-Cbl proto-oncogene.	• signal transducer activity	• ubiquitin ligase complex	• negative regulation of EGF receptor activity • dorsal/ventral axis determination\, follicular epithelium (sensu Insecta)	• Affinity Chromatography •	NCBI SeqHound	[2 Pubmed Abstracts]
DER	Drosophila EGF (epidermal growth factor) receptor homologue.			• protein amino acid phosphorylation • oocyte anterior/posterior			

BIND Details



Data Manager Menu
Version 2.5

About Help
Search Browse

BIND Statistics

Administration

Interaction

Interaction ID: 73

Print Page

Accession date: Sep 7, 1999
Description: The Drosophila homologue of the proto-oncognene Cbl interacts with the epidermal growth factor receptor, DER.

Launch Viewer: Select Below

Molecule A

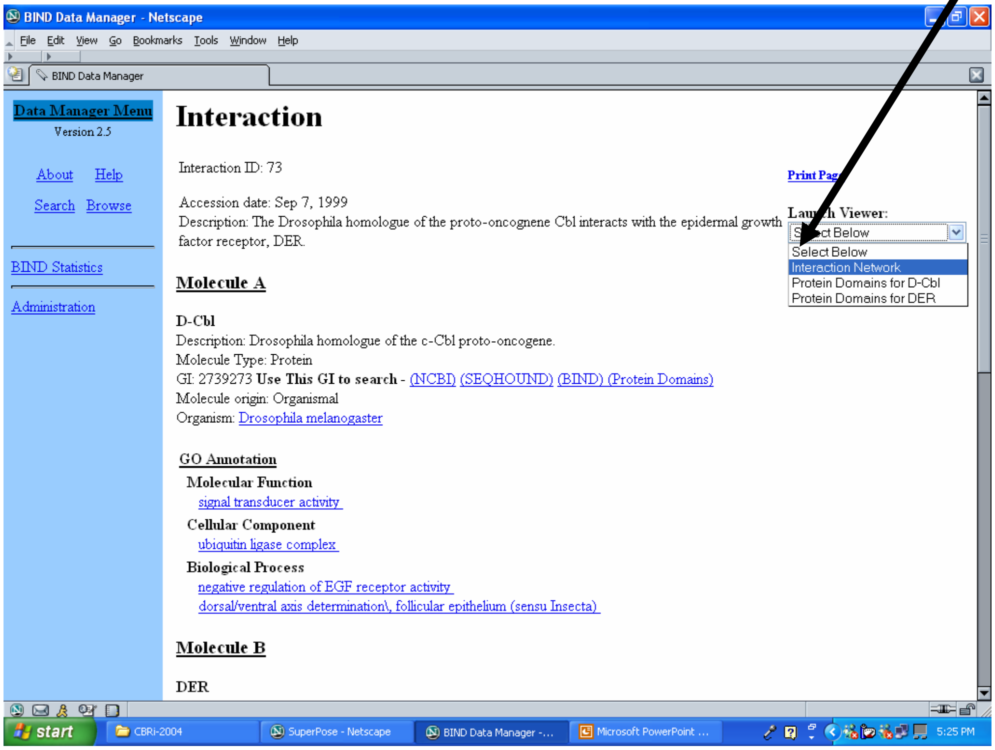
D-Cbl
Description: Drosophila homologue of the c-Cbl proto-oncogene.
Molecule Type: Protein
GI: 2739273 **Use This GI to search -** (NCBI) (SEQHOUND) (BIND) (Protein Domains)
Molecule origin: Organismal
Organism: Drosophila melanogaster

GO Annotation
Molecular Function
signal transducer activity
Cellular Component
ubiquitin ligase complex
Biological Process
negative regulation of EGF receptor activity
dorsal/ventral axis determination\, follicular epithelium (sensu Insecta)

Molecule B

DER

BIND Details



Data Manager Menu
Version 2.5

About Help

Search Browse

BIND Statistics

Administration

Interaction

Interaction ID: 73

Print Page

Accession date: Sep 7, 1999
Description: The Drosophila homologue of the proto-oncogene Cbl interacts with the epidermal growth factor receptor, DER.

Launch Viewer:
Select Below
Select Below
Interaction Network
Protein Domains for D-Cbl
Protein Domains for DER

Molecule A

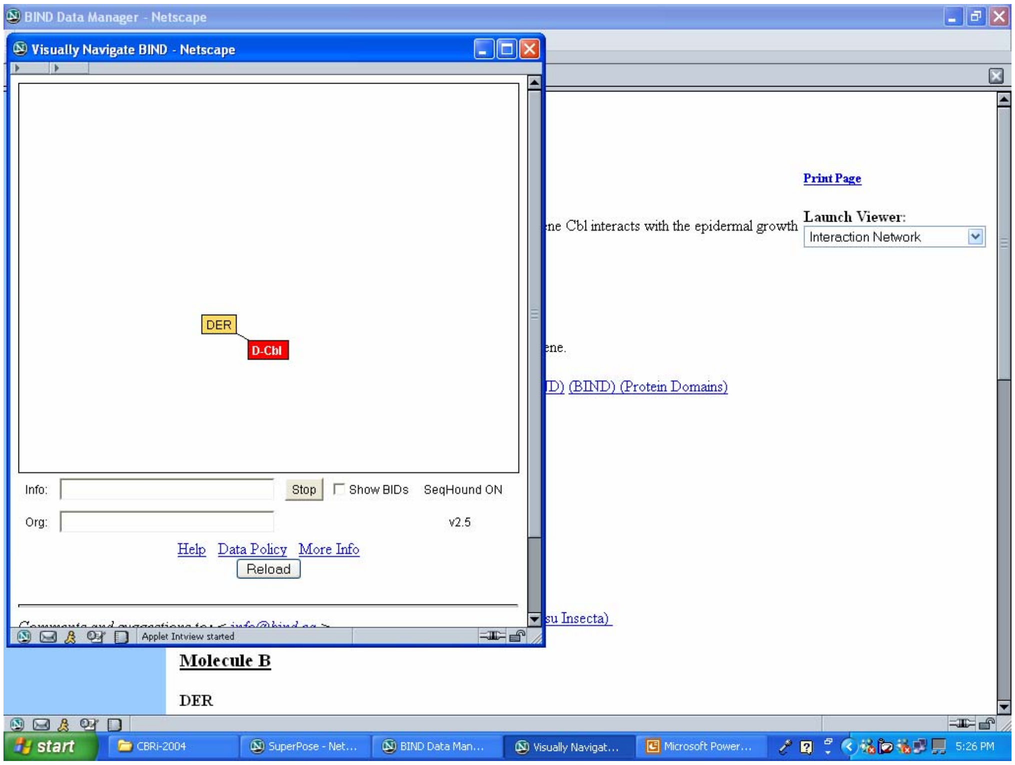
D-Cbl
Description: Drosophila homologue of the c-Cbl proto-oncogene.
Molecule Type: Protein
GI: 2739273 **Use This GI to search -** (NCBI) (SEQHOUND) (BIND) (Protein Domains)
Molecule origin: Organismal
Organism: Drosophila melanogaster

GO Annotation
Molecular Function
signal transducer activity
Cellular Component
ubiquitin ligase complex
Biological Process
negative regulation of EGF receptor activity
dorsal/ventral axis determination\, follicular epithelium (sensu Insecta)

Molecule B

DER

BIND Details



Info: Stop ☐ Show BIDs SeqHound ON

Org: v2.5

Help Data Policy More Info
Reload

Comments and suggestions to: < info@bind.ca >

Print Page

ene Cbl interacts with the epidermal growth

Launch Viewer:
Interaction Network

ene.

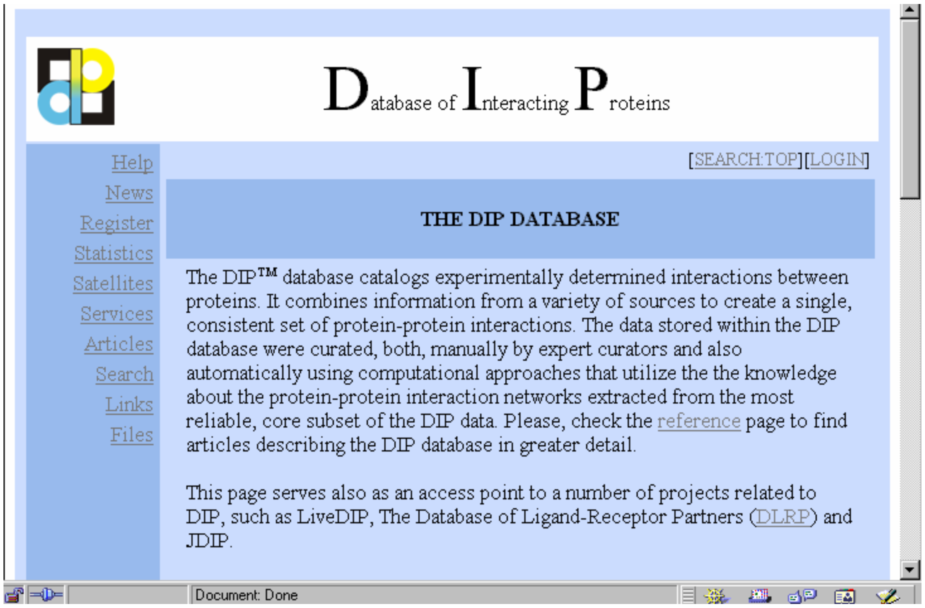
D) (BIND) (Protein Domains)

su Insecta)

Molecule B

DER

DIP Database of Interacting Proteins



Database of Interacting Proteins

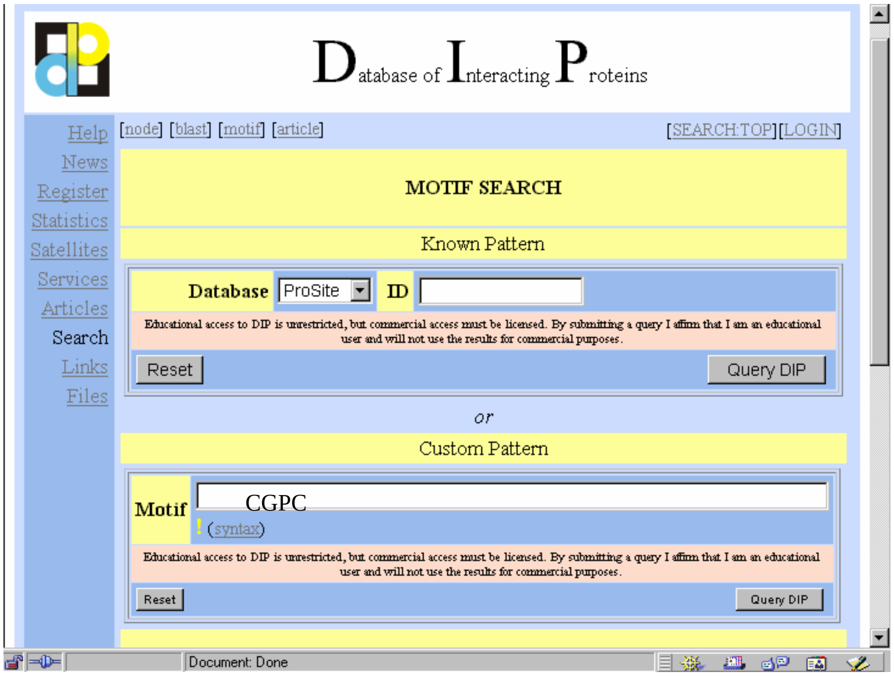
THE DIP DATABASE

The DIP™ database catalogs experimentally determined interactions between proteins. It combines information from a variety of sources to create a single, consistent set of protein-protein interactions. The data stored within the DIP database were curated, both, manually by expert curators and also automatically using computational approaches that utilize the the knowledge about the protein-protein interaction networks extracted from the most reliable, core subset of the DIP data. Please, check the reference page to find articles describing the DIP database in greater detail.

This page serves also as an access point to a number of projects related to DIP, such as LiveDIP, The Database of Ligand-Receptor Partners (DLRP) and JDIP.

http://dip.doe-mbi.ucla.edu/

DIP Query Page



Database of Interacting Proteins

MOTIF SEARCH

Known Pattern

Database: ProSite ID:

Educational access to DIP is unrestricted, but commercial access must be licensed. By submitting a query I affirm that I am an educational user and will not use the results for commercial purposes.

or

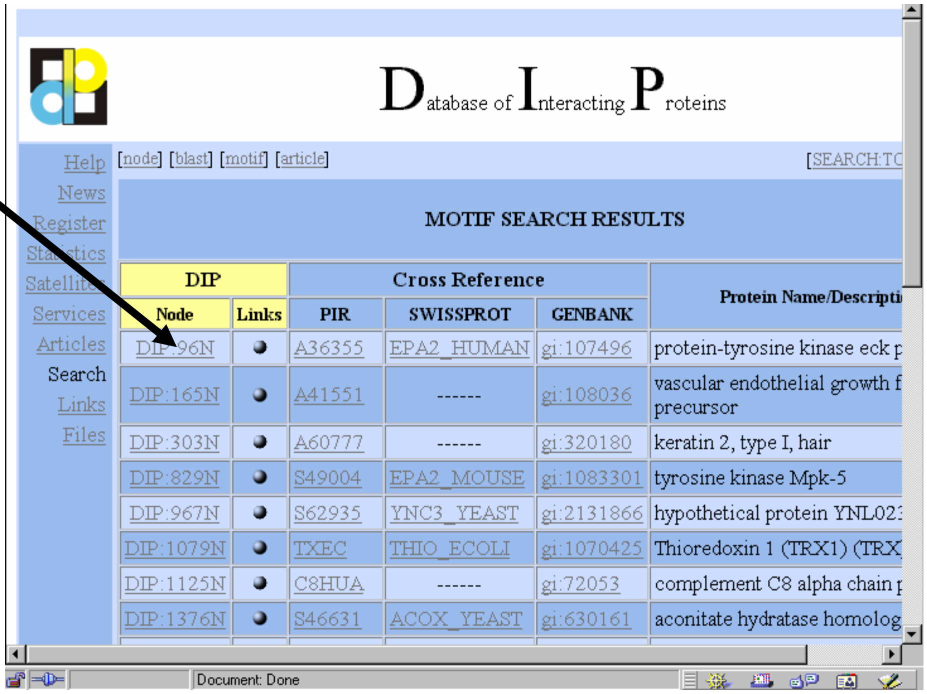
Custom Pattern

Motif: CGPC

! (syntax)

Educational access to DIP is unrestricted, but commercial access must be licensed. By submitting a query I affirm that I am an educational user and will not use the results for commercial purposes.

DIP Results Page



click

Database of Interacting Proteins

Help
News
Register
Statistics
Satellite
Services
Articles
Search
Links
Files

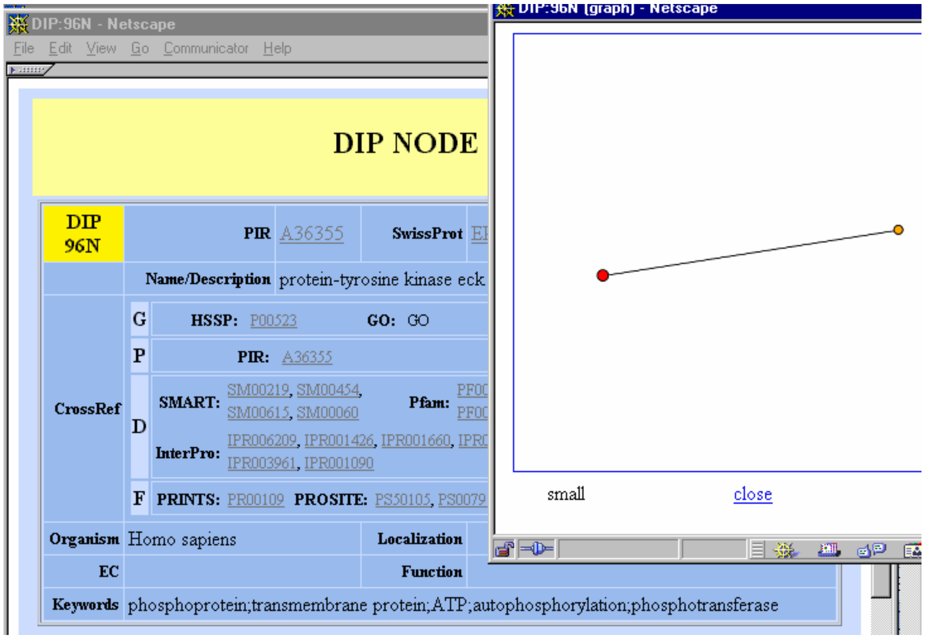
[node] [blast] [motif] [article]　　[SEARCH:TO

MOTIF SEARCH RESULTS

DIP		Cross Reference			Protein Name/Descripti
Node	Links	PIR	SWISSPROT	GENBANK	
DIP:96N	●	A36355	EPA2_HUMAN	gi:107496	protein-tyrosine kinase eck p
DIP:165N	●	A41551	------	gi:108036	vascular endothelial growth f precursor
DIP:303N	●	A60777	------	gi:320180	keratin 2, type I, hair
DIP:829N	●	S49004	EPA2_MOUSE	gi:1083301	tyrosine kinase Mpk-5
DIP:967N	●	S62935	YNC3_YEAST	gi:2131866	hypothetical protein YNL023
DIP:1079N	●	TXEC	THIO_ECOLI	gi:1070425	Thioredoxin 1 (TRX1) (TRX
DIP:1125N	●	C8HUA	------	gi:72053	complement C8 alpha chain p
DIP:1376N	●	S46631	ACOX_YEAST	gi:630161	aconitate hydratase homolog

Document: Done

DIP Results Page



DIP:96N - Netscape
File　Edit　View　Go　Communicator　Help
DIP:96N (graph) - Netscape
small
close

DIP NODE

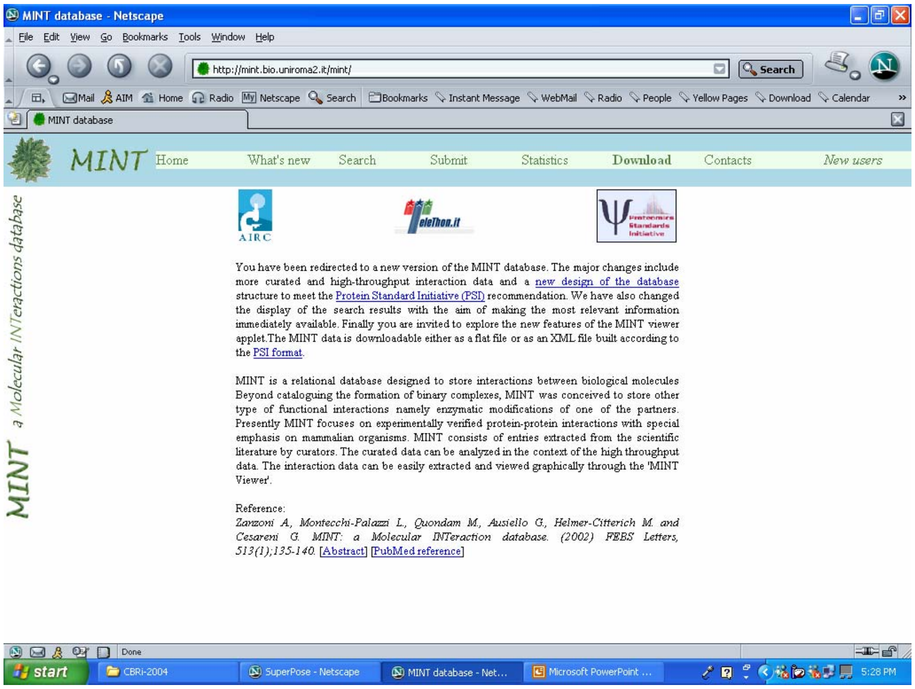
DIP 96N	PIR	A36355	SwissProt	EI
	Name/Description	protein-tyrosine kinase eck		

CrossRef

- G　HSSP: P00523　GO: GO
- P　PIR: A36355
- D　SMART: SM00219, SM00454, SM00615, SM00060　Pfam: PF0C PF0C
  InterPro: IPR006209, IPR001426, IPR001660, IPRC IPR003961, IPR001090
- F　PRINTS: PR00109　PROSITE: PS50105, PS0079

Organism	Homo sapiens	Localization	
EC		Function	
Keywords	phosphoprotein;transmembrane protein;ATP;autophosphorylation;phosphotransferase		

MINT Molecular Interaction Database


MINT database - Netscape
File Edit View Go Bookmarks Tools Window Help
http://mint.bio.uniroma2.it/mint/
Search
Mail AIM Home Radio Netscape Search Bookmarks Instant Message WebMail Radio People Yellow Pages Download Calendar
MINT database
MINT
Home What's new Search Submit Statistics Download Contacts New users
MINT a Molecular INTeractions database
AIRC
TeleThon.it
Proteomics Standards Initiative
Done
start CBRI-2004 SuperPose - Netscape MINT database - Net... Microsoft PowerPoint ... 5:28 PM

You have been redirected to a new version of the MINT database. The major changes include more curated and high-throughput interaction data and a new design of the database structure to meet the Protein Standard Initiative (PSI) recommendation. We have also changed the display of the search results with the aim of making the most relevant information immediately available. Finally you are invited to explore the new features of the MINT viewer applet.The MINT data is downloadable either as a flat file or as an XML file built according to the PSI format.

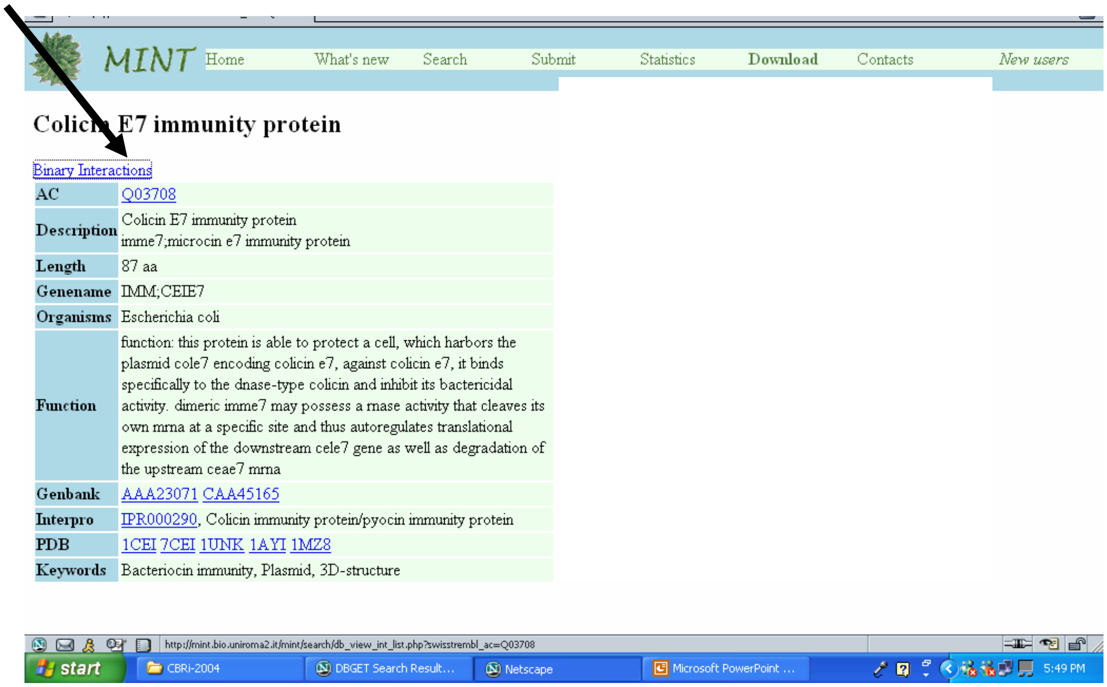
MINT is a relational database designed to store interactions between biological molecules Beyond cataloguing the formation of binary complexes, MINT was conceived to store other type of functional interactions namely enzymatic modifications of one of the partners. Presently MINT focuses on experimentally verified protein-protein interactions with special emphasis on mammalian organisms. MINT consists of entries extracted from the scientific literature by curators. The curated data can be analyzed in the context of the high throughput data. The interaction data can be easily extracted and viewed graphically through the 'MINT Viewer'.

Reference:
Zanzoni A, Montecchi-Palazzi L., Quondam M., Ausiello G., Helmer-Citterich M. and Cesareni G. MINT: a Molecular INTeraction database. (2002) FEBS Letters, 513(1);135-140. [Abstract] [PubMed reference]

http://mint.bio.uniroma2.it/mint/

MINT Results

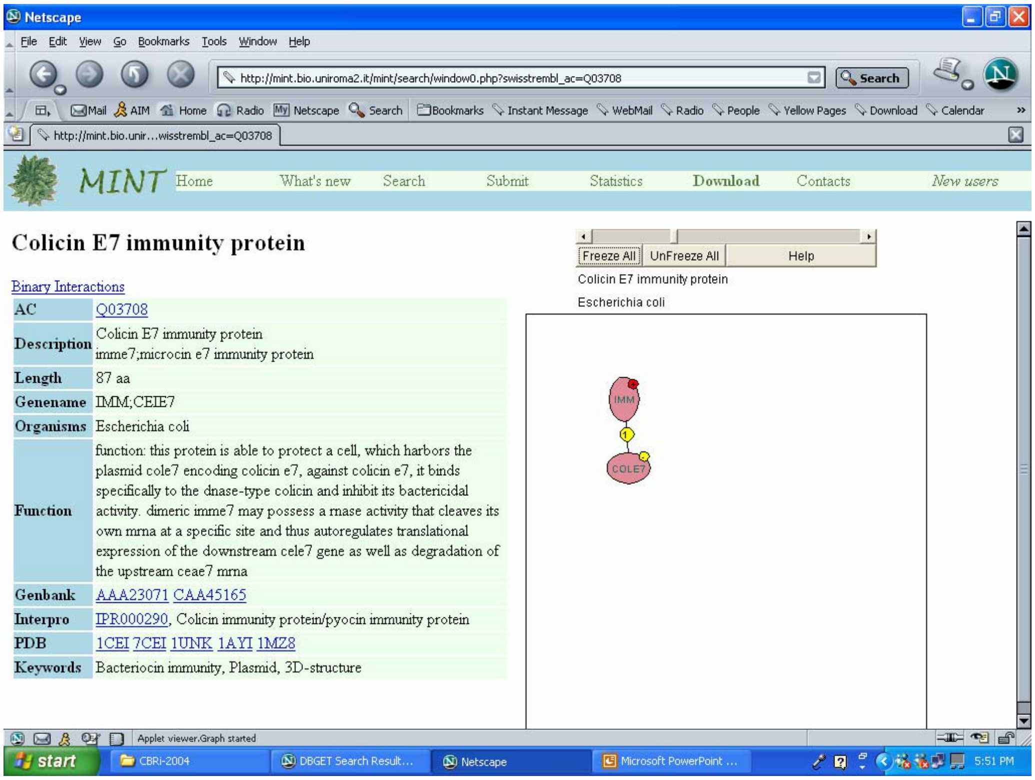
click


MINT
Home What's new Search Submit Statistics Download Contacts New users
http://mint.bio.uniroma2.it/mint/search/db_view_int_list.php?swisstrembl_ac=Q03708
start CBRI-2004 DBGET Search Result... Netscape Microsoft PowerPoint ... 5:49 PM

Colicin E7 immunity protein

Binary Interactions

AC	Q03708
Description	Colicin E7 immunity protein imme7;microcin e7 immunity protein
Length	87 aa
Genename	IMM;CEIE7
Organisms	Escherichia coli
Function	function: this protein is able to protect a cell, which harbors the plasmid cole7 encoding colicin e7, against colicin e7, it binds specifically to the dnase-type colicin and inhibit its bactericidal activity. dimeric imme7 may possess a rnase activity that cleaves its own mrna at a specific site and thus autoregulates translational expression of the downstream cele7 gene as well as degradation of the upstream ceae7 mrna
Genbank	AAA23071 CAA45165
Interpro	IPR000290, Colicin immunity protein/pyocin immunity protein
PDB	1CEI 7CEI 1UNK 1AYI 1MZ8
Keywords	Bacteriocin immunity, Plasmid, 3D-structure

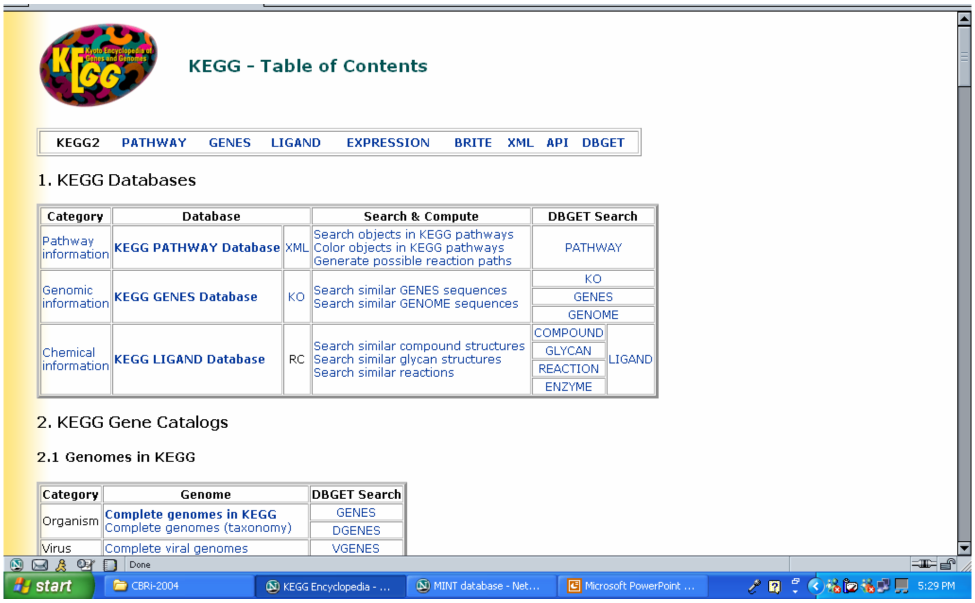

Netscape
File Edit View Go Bookmarks Tools Window Help
http://mint.bio.uniroma2.it/mint/search/window0.php?swisstrembl_ac=Q03708
Search
Mail AIM Home Radio Netscape Search Bookmarks Instant Message WebMail Radio People Yellow Pages Download Calendar
http://mint.bio.unir...wisstrembl_ac=Q03708
MINT
Home
What's new
Search
Submit
Statistics
Download
Contacts
New users
Freeze All
UnFreeze All
Help
Colicin E7 immunity protein
Escherichia coli
IMM
1
COLE7
Applet viewer.Graph started
start
CBRI-2004
DBGET Search Result...
Netscape
Microsoft PowerPoint ...
5:51 PM

Colicin E7 immunity protein

Binary Interactions

AC	Q03708
Description	Colicin E7 immunity protein imme7;microcin e7 immunity protein
Length	87 aa
Genename	IMM;CEIE7
Organisms	Escherichia coli
Function	function: this protein is able to protect a cell, which harbors the plasmid cole7 encoding colicin e7, against colicin e7, it binds specifically to the dnase-type colicin and inhibit its bactericidal activity. dimeric imme7 may possess a rnase activity that cleaves its own mrna at a specific site and thus autoregulates translational expression of the downstream cele7 gene as well as degradation of the upstream ceae7 mrna
Genbank	AAA23071 CAA45165
Interpro	IPR000290, Colicin immunity protein/pyocin immunity protein
PDB	1CEI 7CEI 1UNK 1AYI 1MZ8
Keywords	Bacteriocin immunity, Plasmid, 3D-structure

KEGG Kyoto Encyclopedia of Genes and Genomes



KEGG - Table of Contents

KEGG2 PATHWAY GENES LIGAND EXPRESSION BRITE XML API DBGET

1. KEGG Databases

Category	Database		Search & Compute	DBGET Search	
Pathway information	KEGG PATHWAY Database	XML	Search objects in KEGG pathways Color objects in KEGG pathways Generate possible reaction paths	PATHWAY	
Genomic information	KEGG GENES Database	KO	Search similar GENES sequences Search similar GENOME sequences	KO GENES GENOME	
Chemical information	KEGG LIGAND Database	RC	Search similar compound structures Search similar glycan structures Search similar reactions	COMPOUND GLYCAN REACTION ENZYME	LIGAND

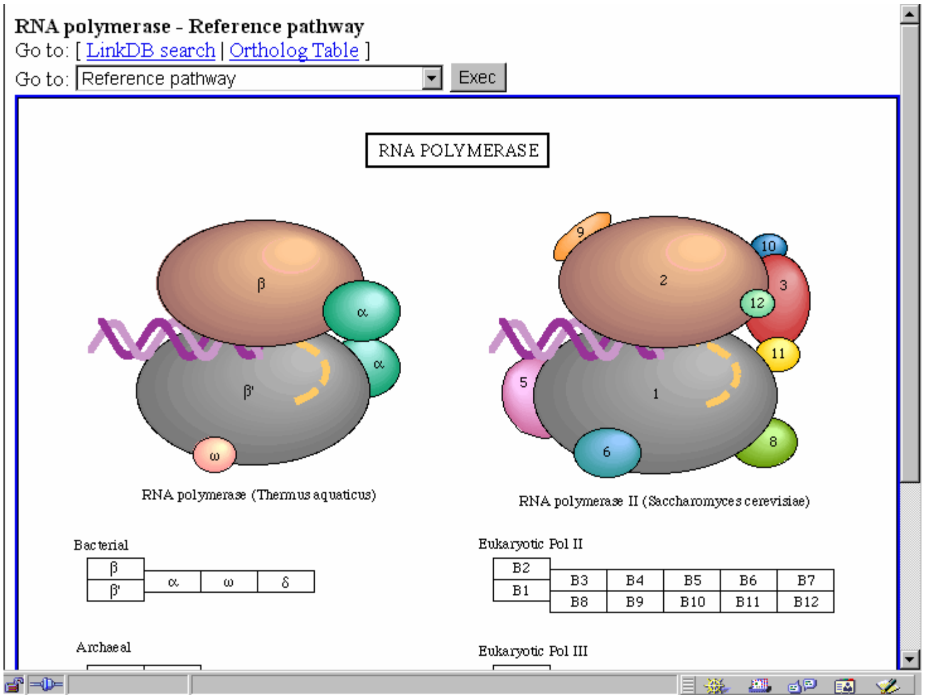
2. KEGG Gene Catalogs

2.1 Genomes in KEGG

Category	Genome	DBGET Search
Organism	Complete genomes in KEGG Complete genomes (taxonomy)	GENES DGENES
Virus	Complete viral genomes	VGENES

http://www.genome.ad.jp/kegg/kegg2.html

KEGG



RNA polymerase - Reference pathway

Go to: [LinkDB search | Ortholog Table]

Go to: Reference pathway Exec

RNA polymerase (Thermus aquaticus)

RNA polymerase II (Saccharomyces cerevisiae)

Bacterial

β			
β'	α	ω	δ

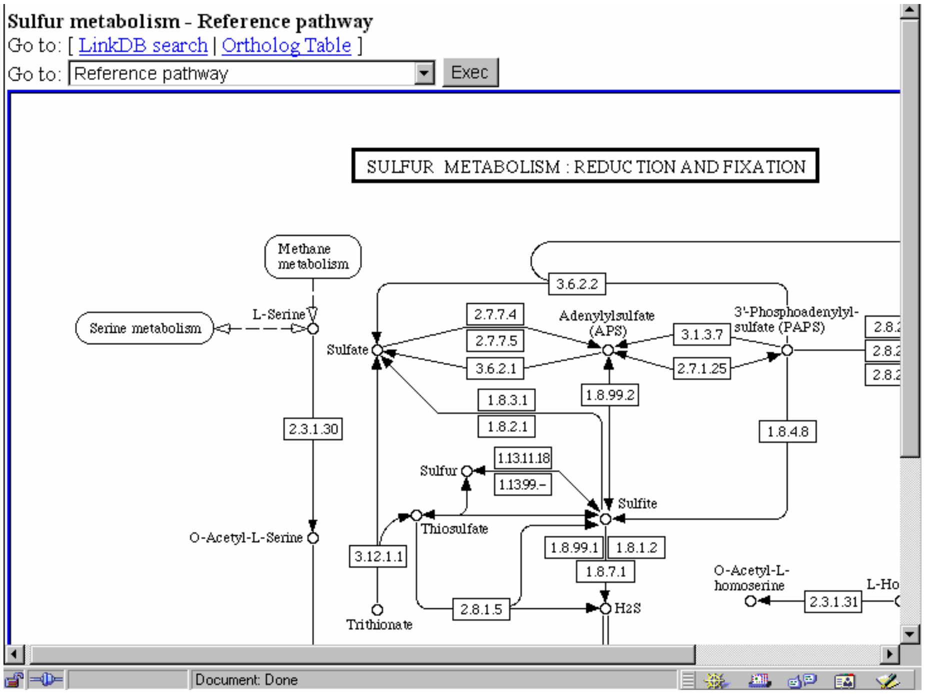
Eukaryotic Pol II

B2	B3	B4	B5	B6	B7
B1	B8	B9	B10	B11	B12

Archaeal

Eukaryotic Pol III

KEGG



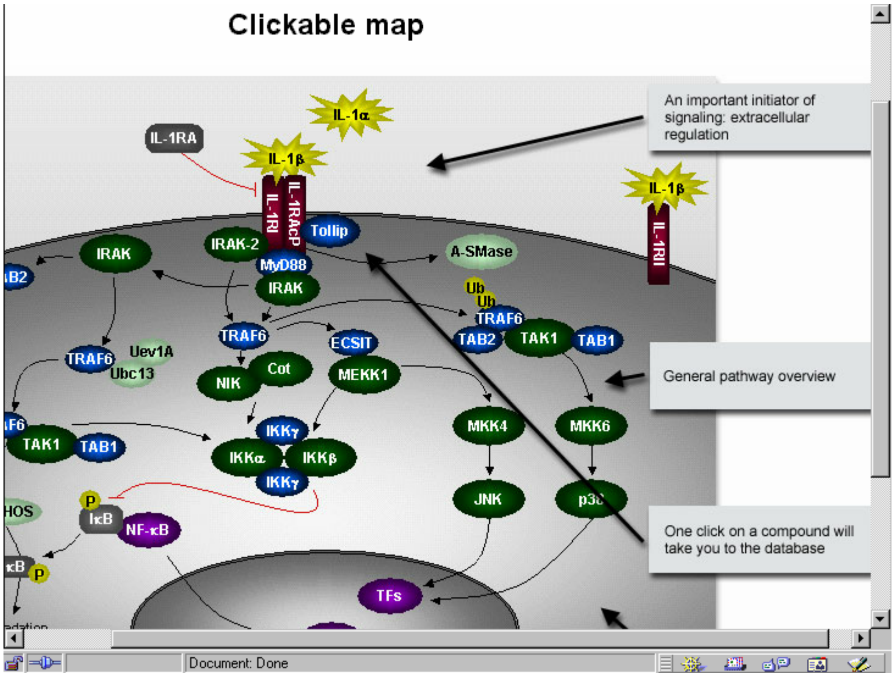
Sulfur metabolism - Reference pathway

Go to: [LinkDB search | Ortholog Table]

Go to: Reference pathway Exec

Document: Done

TRANSPATH



Clickable map

An important initiator of signaling: extracellular regulation

General pathway overview

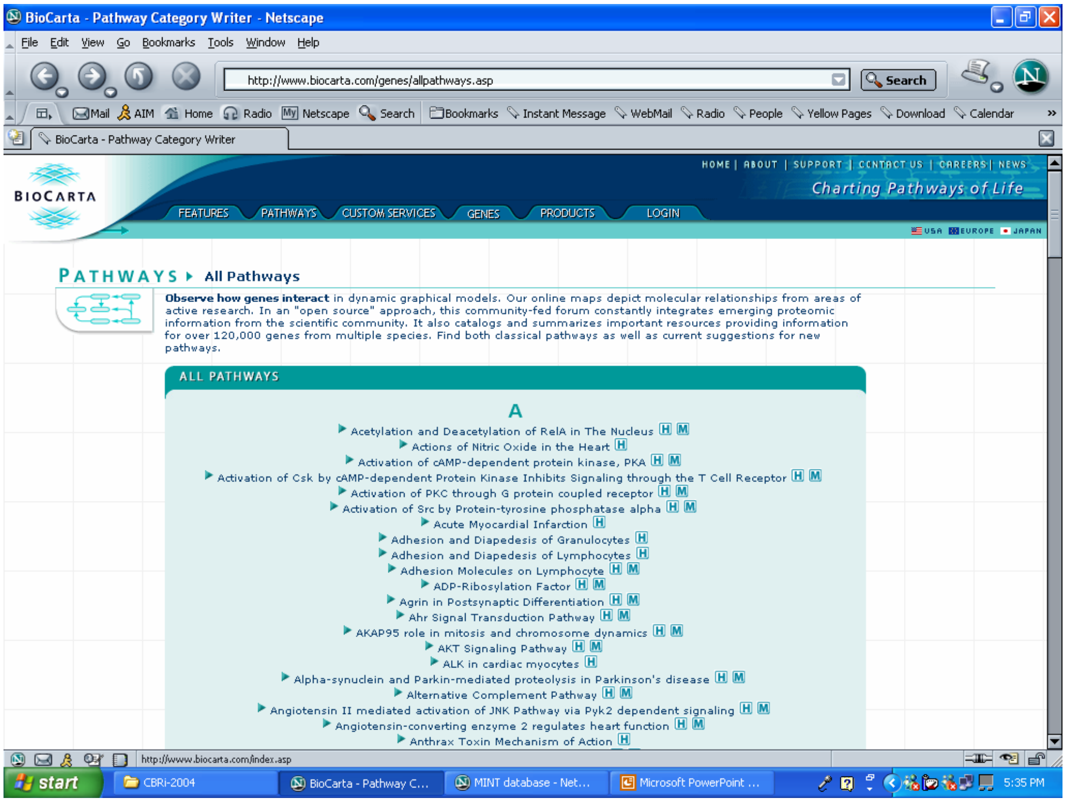
One click on a compound will take you to the database

Document: Done

http://www.biobase.de/pages/products/transpath.html

BIOCARTA

- www.biocarta.com
- Go to “Pathways”
- Web interactive links to many signalling pathways and other eukaryotic protein-protein interactions


BioCarta - Pathway Category Writer - Netscape
File Edit View Go Bookmarks Tools Window Help
http://www.biocarta.com/genes/allpathways.asp
Search
Mail AIM Home Radio Netscape Search Bookmarks Instant Message WebMail Radio People Yellow Pages Download Calendar
BioCarta - Pathway Category Writer
BIOCARTA
HOME | ABOUT | SUPPORT | CONTACT US | CAREERS | NEWS
Charting Pathways of Life
FEATURES PATHWAYS CUSTOM SERVICES GENES PRODUCTS LOGIN
USA EUROPE JAPAN
http://www.biocarta.com/index.asp
start CBRi-2004 BioCarta - Pathway C... MINT database - Net... Microsoft PowerPoint ...
5:35 PM

PATHWAYS ▸ All Pathways

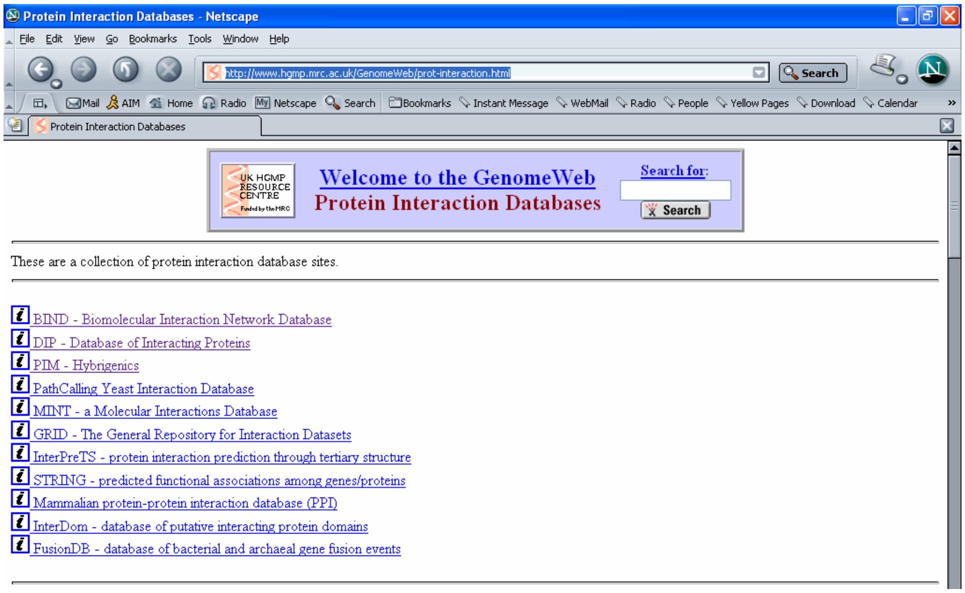
Observe how genes interact in dynamic graphical models. Our online maps depict molecular relationships from areas of active research. In an "open source" approach, this community-fed forum constantly integrates emerging proteomic information from the scientific community. It also catalogs and summarizes important resources providing information for over 120,000 genes from multiple species. Find both classical pathways as well as current suggestions for new pathways.

ALL PATHWAYS

A

- Acetylation and Deacetylation of RelA in The Nucleus H M
- Actions of Nitric Oxide in the Heart H
- Activation of cAMP-dependent protein kinase, PKA H M
- Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor H M
- Activation of PKC through G protein coupled receptor H M
- Activation of Src by Protein-tyrosine phosphatase alpha H M
- Acute Myocardial Infarction H
- Adhesion and Diapedesis of Granulocytes H
- Adhesion and Diapedesis of Lymphocytes H
- Adhesion Molecules on Lymphocyte H M
- ADP-Ribosylation Factor H M
- Agrin in Postsynaptic Differentiation H M
- Ahr Signal Transduction Pathway H M
- AKAP95 role in mitosis and chromosome dynamics H M
- AKT Signaling Pathway H M
- ALK in cardiac myocytes H
- Alpha-synuclein and Parkin-mediated proteolysis in Parkinson's disease H M
- Alternative Complement Pathway H M
- Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling H M
- Angiotensin-converting enzyme 2 regulates heart function H M
- Anthrax Toxin Mechanism of Action H

Other Databases


Protein Interaction Databases - Netscape
File Edit View Go Bookmarks Tools Window Help
http://www.hgmp.mrc.ac.uk/GenomeWeb/prot-interaction.html
Search
Mail AIM Home Radio Netscape Search Bookmarks Instant Message WebMail Radio People Yellow Pages Download Calendar
Protein Interaction Databases
UK HGMP RESOURCE CENTRE Funded by the MRC
Search for:
Search

Welcome to the GenomeWeb
Protein Interaction Databases

These are a collection of protein interaction database sites.

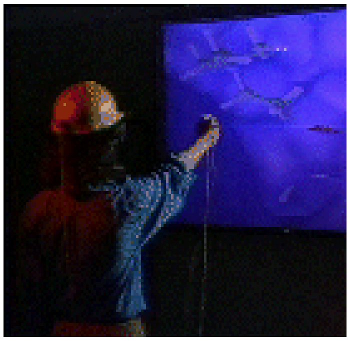
- BIND - Biomolecular Interaction Network Database
- DIP - Database of Interacting Proteins
- PIM - Hybrigenics
- PathCalling Yeast Interaction Database
- MINT - a Molecular Interactions Database
- GRID - The General Repository for Interaction Datasets
- InterPreTS - protein interaction prediction through tertiary structure
- STRING - predicted functional associations among genes/proteins
- Mammalian protein-protein interaction database (PPI)
- InterDom - database of putative interacting protein domains
- FusionDB - database of bacterial and archaeal gene fusion events

http://www.hgmp.mrc.ac.uk/GenomeWeb/prot-interaction.html

Functional Proteomics: A Three-Pronged Process

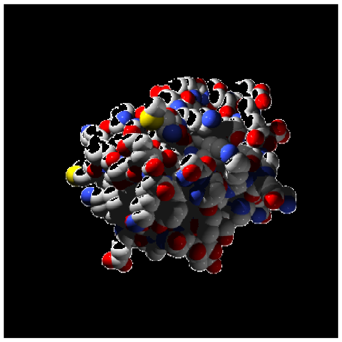
Data Mining Backfilling

Exp. Data Collection



Computer Simulation

Simulation: Three Types of Data (Models)

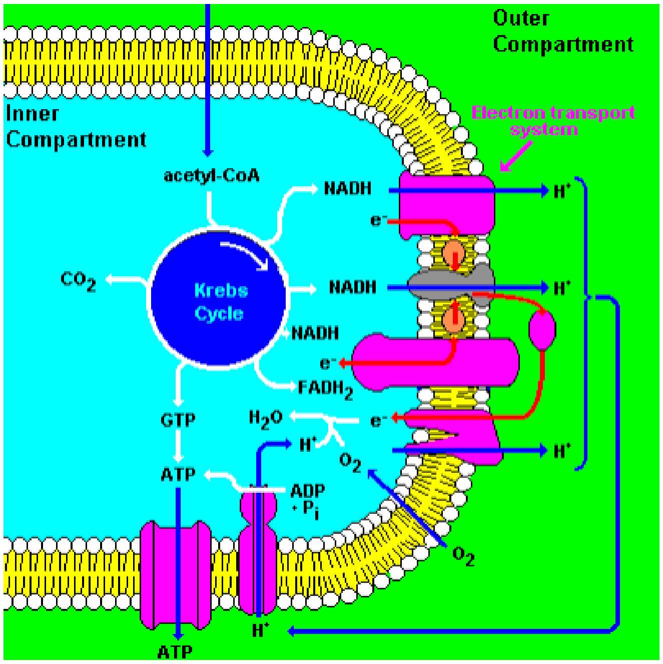


Atomic Scale
0.1 - 1.0 nm
Coordinate data
Dynamic data
0.1 - 10 ns
Molecular dynamics

Meso Scale
1.0 - 10 nm
Interaction data
Kon, Koff, Kd
10 ns - 10 ms
Mesodynamics

Continuum Model
10 - 100 nm
Concentrations
Diffusion rates
10 ms - 1000 s
Fluid dynamics

Cell Simulation with DEs



$$\frac{dx_1}{dt} = k_{11}x_1 + k_{21}x_2 + k_{31}x_3 + \ldots$$

$$\frac{dx_2}{dt} = k_{12}x_1 + k_{22}x_2 + k_{32}x_3 + \ldots$$

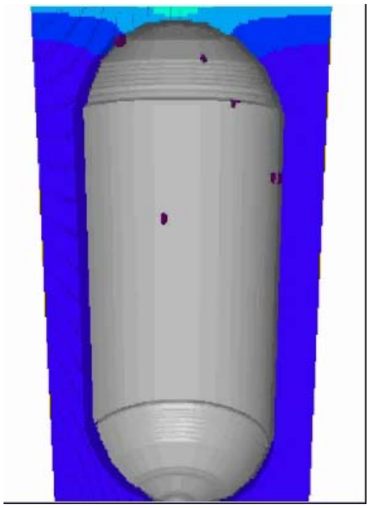
$$\frac{dx_3}{dt} = k_{13}x_1 + k_{23}x_2 + k_{33}x_3 + \ldots$$

$$\frac{dx_4}{dt} = k_{13}x_1 + k_{24}x_2 + k_{34}x_3 + \ldots$$

Continuum Modelling

- Desire to simulate spatially and temporally (to make movies)
- Use techniques developed for oil and gas resevoir simulation (pumping, diffusion, reaction, pressure -- CMG Inc.)
- Uses theory of non-turbulent fluid dynamics, discretized over small volumes
- Based on measured parameters of real cells, real metabolites, proteins

Continuum Simulation

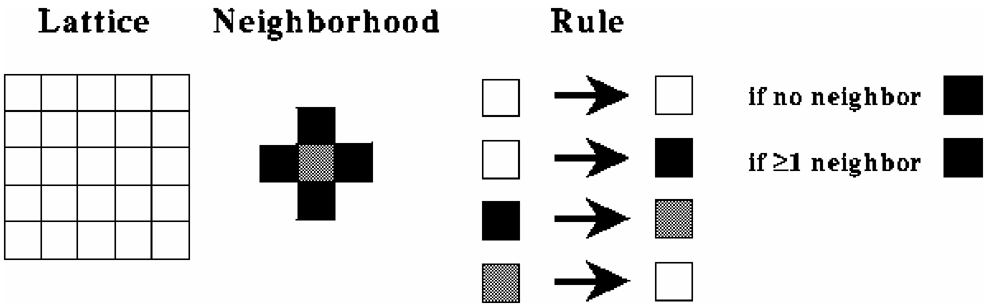


movie

Cellular Automata (CA)

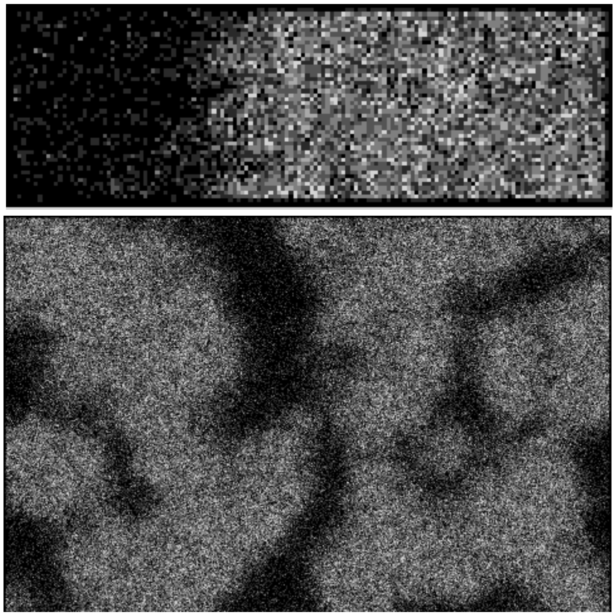
- Computer modelling method that uses lattices and discrete state “rules” to model time dependent processes
- No differential equations to solve, easy to calculate, more phenomenological
- Simple unit behavior -> complex group behavior
- Can be used to create Mandelbrot figures
- Used to model fluid flow, percolation, reaction + diffusion, traffic flow, ecology

Cellular Automata



Can be extended to 3D lattice

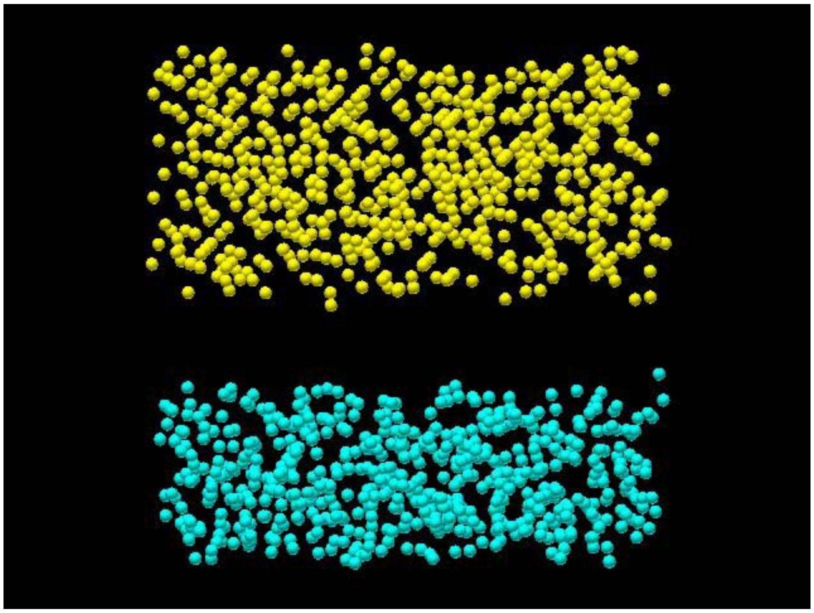
Reaction/Diffusion with Cellular Automata



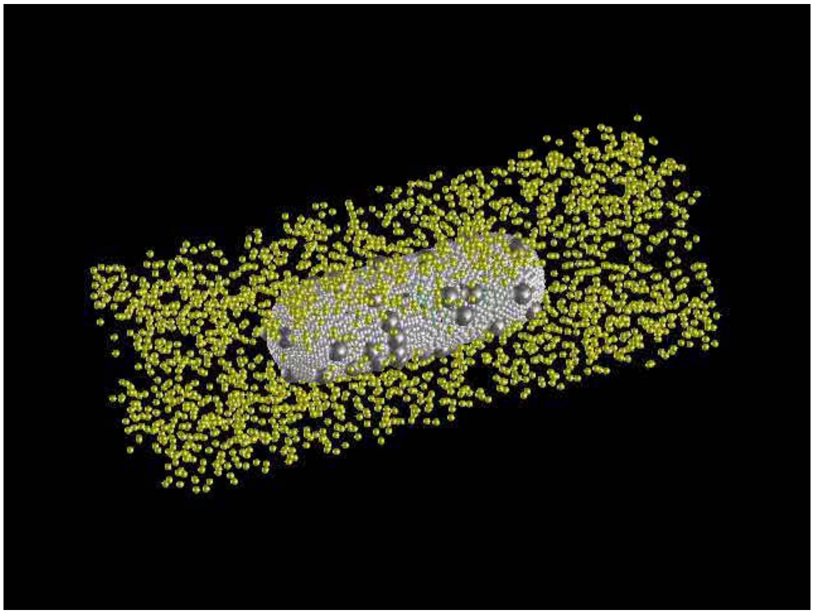
Another Example of CA

SimCity 2000

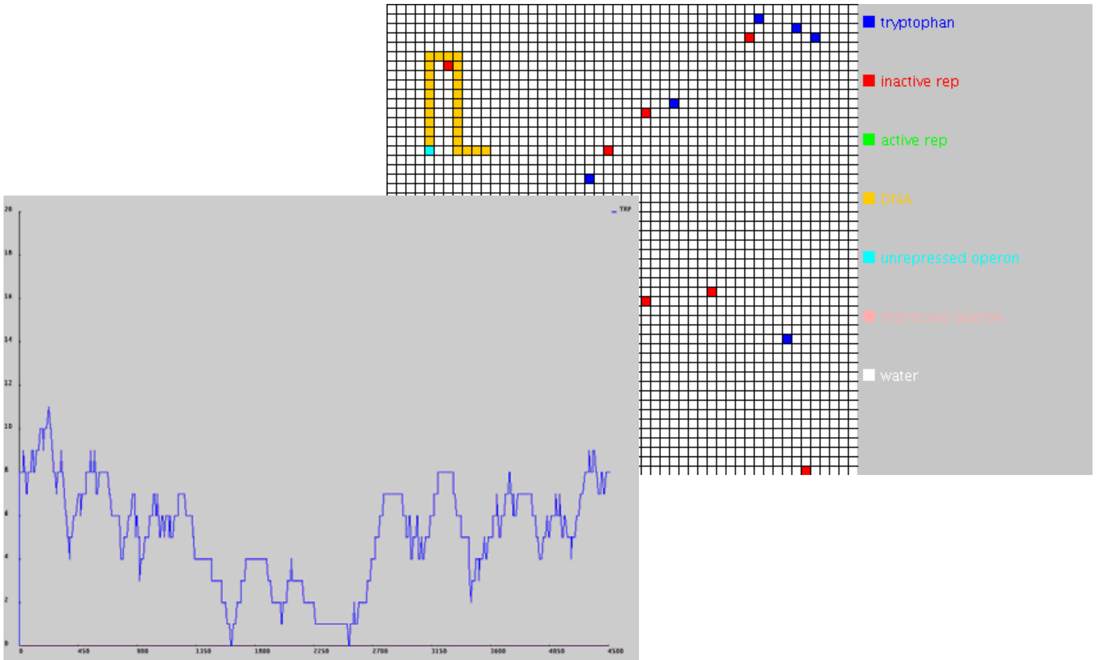
CA Simulations of Diffusion + Reaction



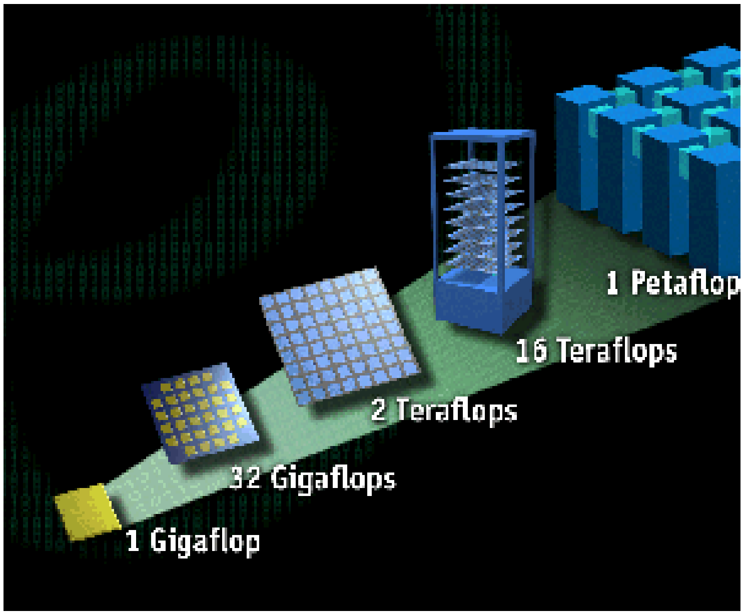
CA Simulations of Transport



CA for Trp Repressor



How Big A Computer?



Functional Proteomics

- Mixture of experimental and computational techniques
- Trying to reach a point where functions and interactions can be predicted and modelled
- The future of proteomics (and bioinformatics)