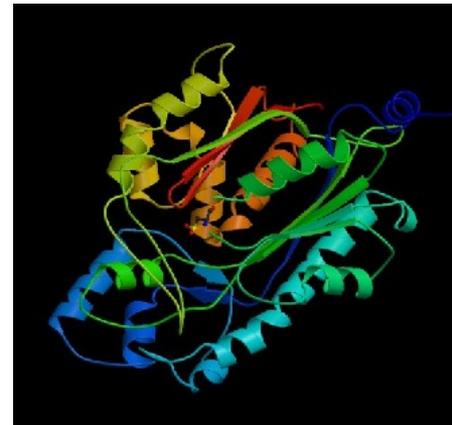


GENOTYP

```
>chs1  
atgacagaat  
acaggatgac  
tatgacgtga  
cggcttatat  
gatgacc...
```

```
>chs1  
MFVDDHLA  
VNQNFYLR  
SHRQL...
```

GEN.KÓD



STRUKTURA

FENOTYP



FUNKCE

Levels of protein structure: 0,1

- Zeroth: amino acid composition – no structural information
- Primary
 - This is simply the order of covalent linkages along the polypeptide chain, i.e. the sequence itself

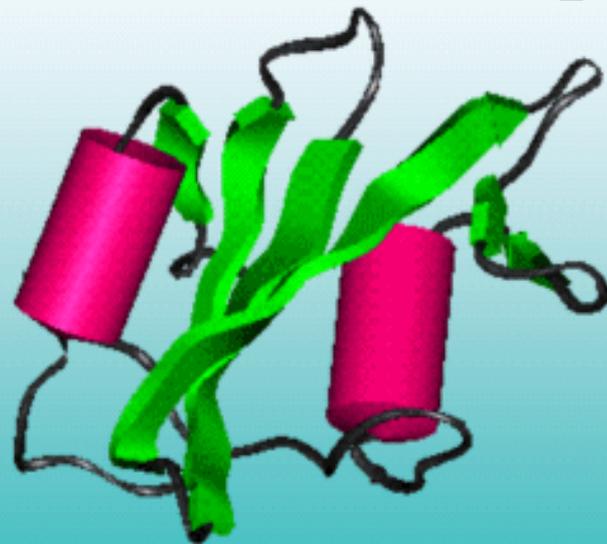
MHGAYRTPRSKTDAYGCQILETRAS

Levels of protein structure: 2

- Secondary
 - Local organization of the protein backbone: α -helix, β -strand (which assemble into β -sheets), turn and interconnecting loop



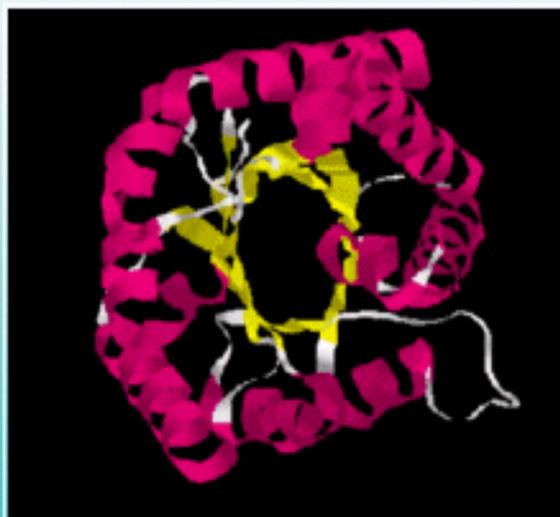
Levels of protein structure: 3



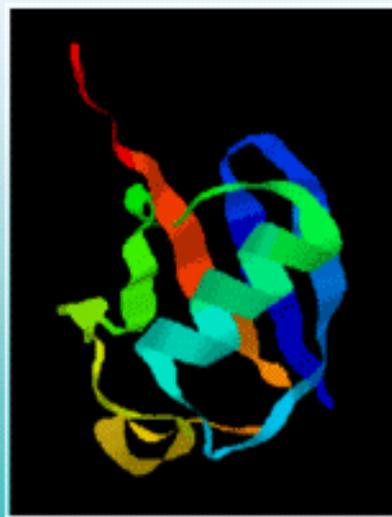
- Tertiary
 - packing of secondary structure elements into a compact spatial unit
 - “Fold” or domain – this is the level to which structure prediction is currently possible

Structural classes: 2

α/β (parallel β -sheet)



$\alpha+\beta$ (antiparallel β -sheet)



Most popular class!

Analýza sekvence proteinů

- statistická analýza
- identifikace motivů (vzorů) typických pro vybrané funkce
- strukturní modelování (ab initio, fragmentové metody, threading, homologní modelování)
- další nástroje pro předpovídání funkce a struktury

Statistická analýza

Zastoupení dipeptidů
v proteinových
sekvencích

24174 LL	1384 MM
18928 SG	1348 HW
18914 SL	1345 PC
18843 AL	1339 QW
18785 LA	1337 NC
18675 AA	1324 WF
18629 LS	1267 CN
17616 SS	1241 CF
17397 GL	1175 WQ
16309 GG	1156 QC
16120 LG	1138 CY
15790 AS	1087 WM
15769 LE	1026 HM
15510 LV	985 WP
15416 AG	929 HC
15390 AV	883 CH
15328 GS	718 WH
15319 VL	718 CC
	640 CM
	553 MC
	526 WW

Statistická analýza

Zložení dipeptidů lze využít k hodnocení podobnosti, které je určitým způsobem dokonalejší než Needleman-Wunsch

MASAQSFYLLFNMVLADHS HQ

MA, AS, SA, AQ, QS, FY, YL, LL, LF, FN, NM, MV, VL, LA, AD, DH, HS, SH, HQ

FNMVLADHS HQMASAQSFYLL

MA, AS, SA, AQ, QS, FY, YL, LL, QM, FN, NM, MV, VL, LA, AD, DH, HS, SH, HQ

Algoritmus pro globální srovnání pomocí DP

	A	B	C	N	J	R	Q	C	L	C	R	P	M
A	1												
J					1								
C			1					1	1				
J				1									
N					1								
R						1						1	
C			1					1	1				
K													
C			1					1	1				
R						1					1		
B	1	2	1	1	1	1	1	1	1	1	1	0	0
P	0	0	0	0	0	0	0	0	0	0	0	1	0

	A	B	C	N	J	R	Q	C	L	C	R	P	M
A	1												
J					1								
C			1					1	1				
J				1									
N					1								
R						1						1	
C			1					1	1				
K													
C			1					1	1				
R						1					2	0	0
B	1	2	1	1	1	1	1	1	1	1	1	0	0
P	0	0	0	0	0	0	0	0	0	0	0	1	0

	A	B	C	N	J	R	Q	C	L	C	R	P	M
A	1												
J					1								
C			1					1	1				
J				1									
N					1								
R						1						1	
C			1					1	1				
K													
C			1					1	1				
R						1					1	2	0
B	1	2	1	1	1	1	1	1	1	1	1	0	0
P	0	0	0	0	0	0	0	0	0	0	0	1	0

	A	B	C	N	J	R	Q	C	L	C	R	P	M
A	1												
J					1								
C			1					1	1				
J				1									
N					1								
R						1						1	
C			1					1	1				
K													
C			1					1	1				
R						1						1	
B	1	2	1	1	1	1	1	1	1	1	1	0	0
P	0	0	0	0	0	0	0	0	0	0	0	1	0

sequence 1

	A	B	C	N	J	R	Q	C	L	C	R	P	M
A	8	7	6	6	5	4	4	3	3	2	1	0	0
J	7	7	6	6	6	4	4	3	3	2	1	0	0
C	6	6	7	6	5	4	4	4	3	3	1	0	0
J	6	6	6	5	6	4	4	3	3	2	1	0	0
N	5	5	5	6	5	4	4	3	3	2	1	0	0
R	4	4	4	4	4	5	4	3	3	2	2	0	0
C	3	3	4	3	3	3	3	4	3	3	1	0	0
K	3	3	3	3	3	3	3	3	3	2	1	0	0
C	2	2	3	2	2	2	2	3	2	3	1	0	0
R	2	1	1	1	1	2	1	1	1	1	2	0	0
B	1	2	1	1	1	1	1	1	1	1	1	0	0
P	0	0	0	0	0	0	0	0	0	0	0	1	0

sequence 2

sequence 1 ABCNJ-RQCLCR-PM
 sequence 2 AJC-JNR-CKCRBP-

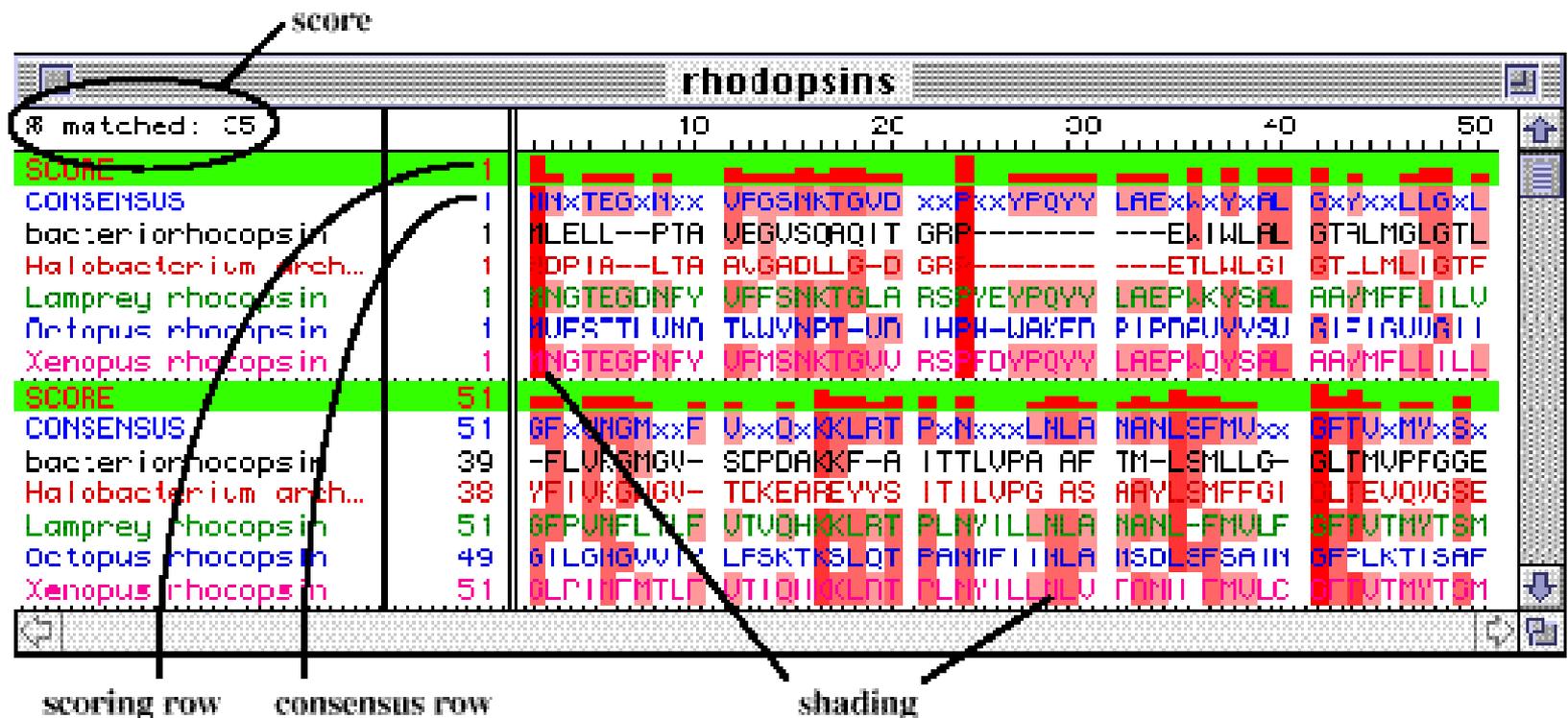
sequence 1 ABC-NJRQCLCR-PM
 sequence 2 AJCJN-R-CKCRBP-

Vícenásobné zarovnání proteinů

Snaha minimalizovat součet skóre všech dvojic

$$S(v) = \text{SUM}(s_m, s_n), \text{ kde } m < n$$

Vícenásobné zarovnání proteinů



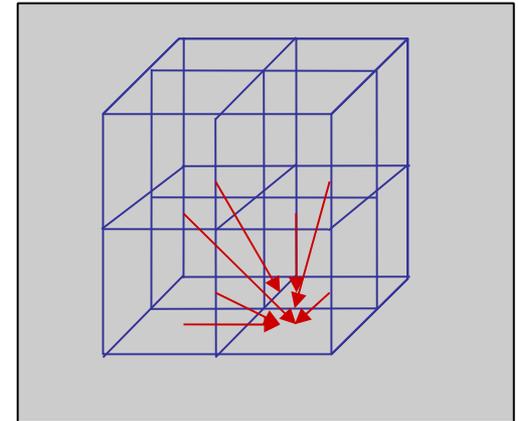
Vícenásobné zarovnání proteinů

- metoda vícerozměrného DP
- metoda progresivního zarovnávání
- celá sada různých heuristik pro zrychlení

Vícenásobné zarovnání proteinů

Metoda vícerozměrného DP

$$F(i,j,k) = \max\{ \begin{aligned} &F(i-1,j-1,k-1)+S(x_i, x_j, x_k), \\ &F(i-1,j-1,k) +S(x_i, x_j, -), \\ &F(i-1,j, k-1)+S(x_i, -, x_k), \\ &F(i-1,j, k) +S(x_i, -, -), \\ &F(i, j-1,k-1)+S(-, x_j, x_k), \\ &F(i, j-1,k) +S(-, x_j, x_k), \\ &F(i, j, k-1)+S(-, -, x_k) \} \end{aligned}$$



$$O(2^N * L^N)$$

Vícenásobné zarovnání proteinů

Metoda progresivní (funguje dobře pokud známe fylogenetické vztahy a začínáme sekvencemi nejvíce příbuznými). Známý nástroj CLUSTALW.

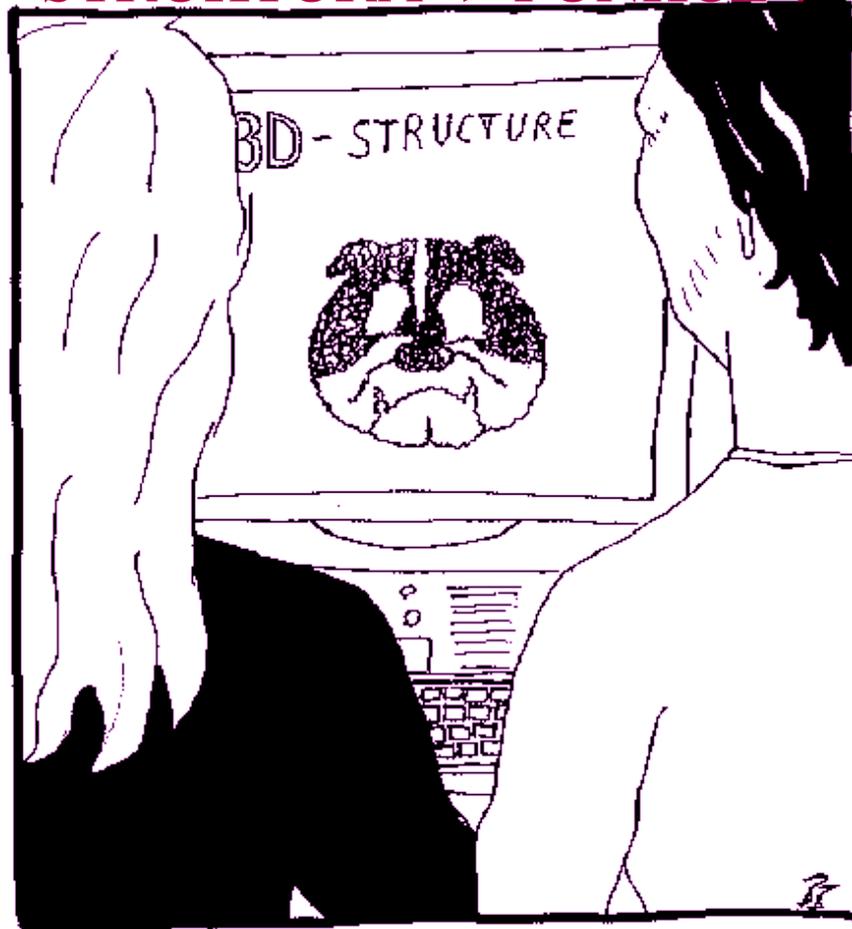
ALGORITMUS:

- přidej sekvenci
- zarovnej 2 sekvence
- opakuj postup s dalšími sekvencemi $O(N*L^2)$

Nástroje pro analýzu

Predikce lokalizace	SIGNALP
Predikce sekundární struktury	PSIPRED
Predikce transmembránové topologie	MEMSAT2
Predikce domén	DOMPRED
Predikce typu skládání	GenTHREADER
Predikce struktury	ROSETTA

STRUKTURA -> FUNKCE ?



Brigitte Boeckmann / 1994

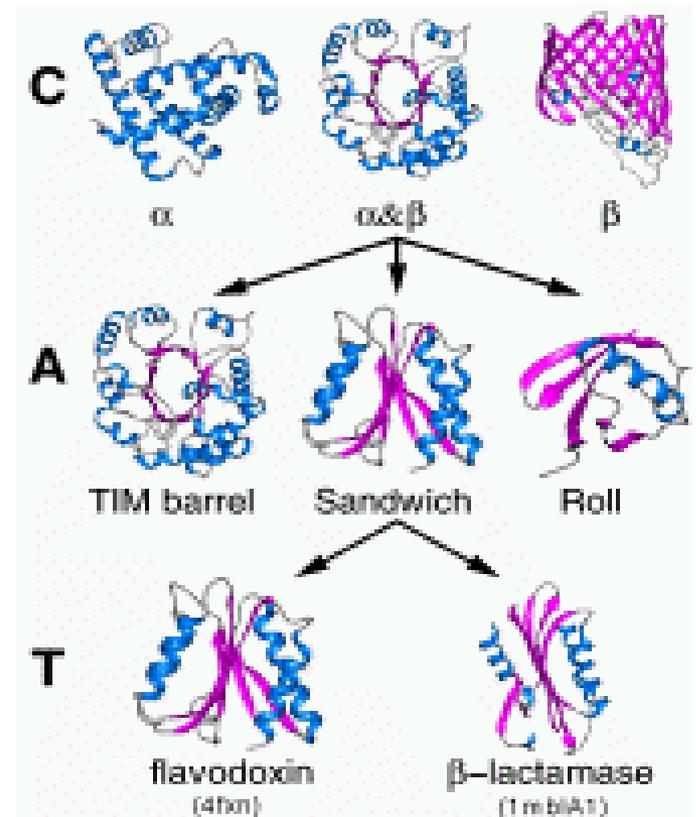
" It's a restriction enzyme."

Strukturní data

- PDB
- PDBsum

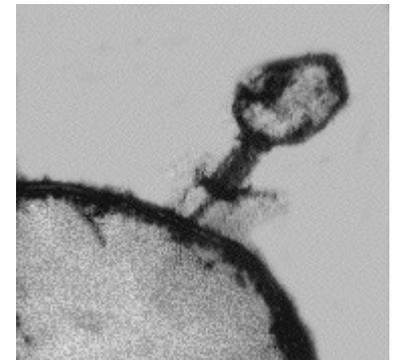
Odvozená data

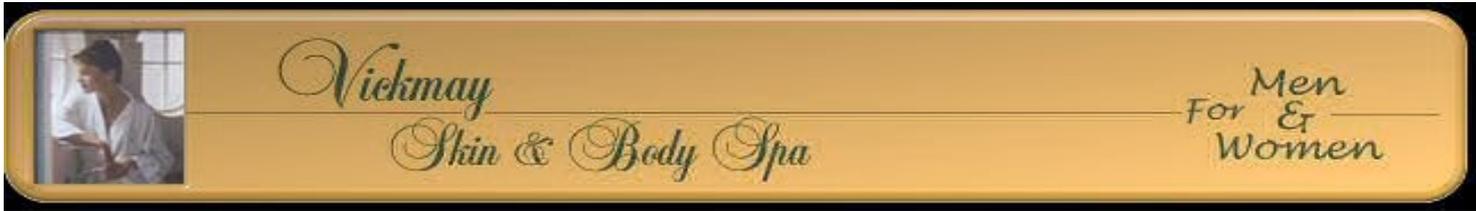
- SCOP (Class, Fold, Superfamily, Family)
- CATH (Class, Architecture, Topology, Homologous superfamily)



Strukturní data

- 1HEW, 1AM7
- Lysozym – enzym hydrolyzující (štěpící) vazbu mezi cukry polysacharidů, které se nacházejí v buněčné stěně některých bakterií – zdroje: vejce, slzy, bakteriofág T4





For oily, acne-prone and problem skin conditions.

- OIL-FREE
- LYSOZYME / LIPOPEPTIDE ANTIMICROBIAL COMPLEX
- MOISTURIZES BY SEALING IN MOISTURE
- FRAGRANCE-FREE
- CONTAINS ANTI-BACTERIAL ENZYME LYSOZYME

PRICE: \$20.00

<http://www.rcsb.org/pdb/>

PDB soubor - hlavička

HEADER	HYDROLASE(O-GLYCOSYL)	20-JAN-92	1HEW	1HEW	2
COMPND	LYSOZYME (E.C.3.2.1.17) COMPLEXED WITH THE INHIBITOR			1HEW	3
COMPND	2 TRI-N-ACETYLCHITOTRIOSE			1HEW	4
SOURCE	HEN (GALLUS GALLUS) EGG WHITE			1HEW	5
AUTHOR	J.C.CHEETHAM,P.J.ARTYMIUK,D.C.PHILLIPS			1HEW	6
REVDAT	1 31-JAN-94 1HEW 0			1HEW	7
JRNL	AUTH J.C.CHEETHAM,P.J.ARTYMIUK,D.C.PHILLIPS			1HEW	8
JRNL	TITL REFINEMENT OF AN ENZYME COMPLEX WITH INHIBITOR			1HEW	9
JRNL	TITL 2 BOUND AT PARTIAL OCCUPANCY. HEN EGG-WHITE			1HEW	10
JRNL	TITL 3 LYSOZYME AND TRI-N-ACETYLCHITOTRIOSE AT 1.75			1HEW	11
JRNL	TITL 4 ANGSTROMS RESOLUTION			1HEW	12
JRNL	REF J.MOL.BIOL.	V. 224	613 1992	1HEW	13
JRNL	REFN ASTM JMOBAK UK ISSN 0022-2836		070	1HEW	14
REMARK	1			1HEW	15
REMARK	1 REFERENCE 1			1HEW	16
REMARK	1 AUTH L.N.JOHNSON,J.C.CHEETHAM,P.J.MC*LAUGHLIN,			1HEW	17
REMARK	1 AUTH 2 K.R.ACHARYA,D.BARFORD,D.C.PHILLIPS			1HEW	18
REMARK	1 TITL PROTEIN-OLIGOSACCHARIDE INTERACTIONS: LYSOZYME,			1HEW	19
REMARK	1 TITL 2 PHOSPHORYLASE, AMYLASES			1HEW	20
REMARK	1 REF CURR.TOP.MICROBIOL.IMMUNOL.	V. 139	81 1988	1HEW	21

PDB soubor – primární struktura

```
REMARK      5 THE THREE SUGAR UNITS OF THE INHIBITOR MOLECULE ARE BOUND      1HEW  56
REMARK      5 IN THE UPPER THREE SITES (A TO C) OF THE LYSOZYME ACTIVE      1HEW  57
REMARK      5 SITE CLEFT.  NAG MOLECULES, NUMBERED 203, 202, AND 201, ARE  1HEW  58
REMARK      5 BOUND IN SITES A, B, AND C, RESPECTIVELY.                      1HEW  59
SEQRES      1      129  LYS VAL PHE GLY ARG CYS GLU LEU ALA ALA ALA MET LYS  1HEW  60
SEQRES      2      129  ARG HIS GLY LEU ASP ASN TYR ARG GLY TYR SER LEU GLY  1HEW  61
SEQRES      3      129  ASN TRP VAL CYS ALA ALA LYS PHE GLU SER ASN PHE ASN  1HEW  62
SEQRES      4      129  THR GLN ALA THR ASN ARG ASN THR ASP GLY SER THR ASP  1HEW  63
SEQRES      5      129  TYR GLY ILE LEU GLN ILE ASN SER ARG TRP TRP CYS ASN  1HEW  64
SEQRES      6      129  ASP GLY ARG THR PRO GLY SER ARG ASN LEU CYS ASN ILE  1HEW  65
SEQRES      7      129  PRO CYS SER ALA LEU LEU SER SER ASP ILE THR ALA SER  1HEW  66
SEQRES      8      129  VAL ASN CYS ALA LYS LYS ILE VAL SER ASP GLY ASN GLY  1HEW  67
SEQRES      9      129  MET ASN ALA TRP VAL ALA TRP ARG ASN ARG CYS LYS GLY  1HEW  68
SEQRES     10      129  THR ASP VAL GLN ALA TRP ILE ARG GLY CYS ARG LEU      1HEW  69
HET        NAG      201      15      N-ACETYL-D-GLUCOSAMINE      1HEW  70
HET        NAG      202      14      N-ACETYL-D-GLUCOSAMINE      1HEW  71
HET        NAG      203      14      N-ACETYL-D-GLUCOSAMINE      1HEW  72
FORMUL     2  NAG      3(C8 H15 N1 O6)      1HEW  73
FORMUL     3  HOH     *103(H2 O1)      1HEW  74
```

PDB soubor – sekundární struktura

```

HELIX 1 A ARG 5 HIS 15 1 1HEW 75
HELIX 2 B LEU 25 GLU 35 1 1HEW 76
HELIX 3 C CYS 80 LEU 84 5 1HEW 77
HELIX 4 D THR 89 ILE 98 1 1HEW 78
HELIX 5 E VAL 109 ASN 113 1 1HEW 79
SHEET 1 S1 2 LYS 1 PHE 3 0 1HEW 80
SHEET 2 S1 2 PHE 38 THR 40 -1 N THR 40 O LYS 1 1HEW 81
SHEET 1 S2 3 ALA 42 ASN 46 0 1HEW 82
SHEET 2 S2 3 SER 50 GLY 54 -1 O SER 50 N ASN 46 1HEW 83
SHEET 3 S2 3 GLN 57 SER 60 -1 O ILE 58 N TYR 53 1HEW 84
TURN 1 T1 MET 12 HIS 15 TYPE III 1HEW 85
TURN 2 T2 LYS 13 GLY 16 TYPE I 1HEW 86
TURN 3 T3 LEU 17 TYR 20 TYPE II 1HEW 87
TURN 4 T4 ASN 19 GLY 22 DISTORTED TYPE II 1HEW 88
TURN 5 T5 TYR 20 TYR 23 TYPE I ' 1HEW 89
TURN 6 T6 SER 24 ASN 27 TYPE III 1HEW 90
TURN 7 T7 LEU 25 TRP 28 TYPE III 1HEW 91
TURN 8 T8 SER 36 ASN 39 TYPE III ' 1HEW 92
TURN 9 T9 ASN 46 GLY 49 TYPE I 1HEW 93

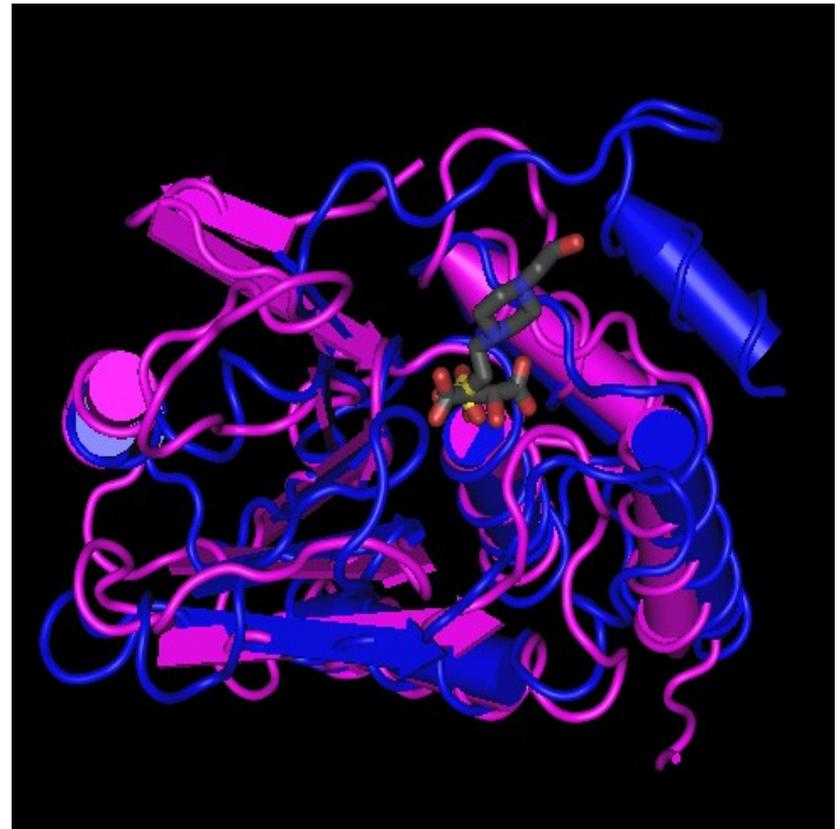
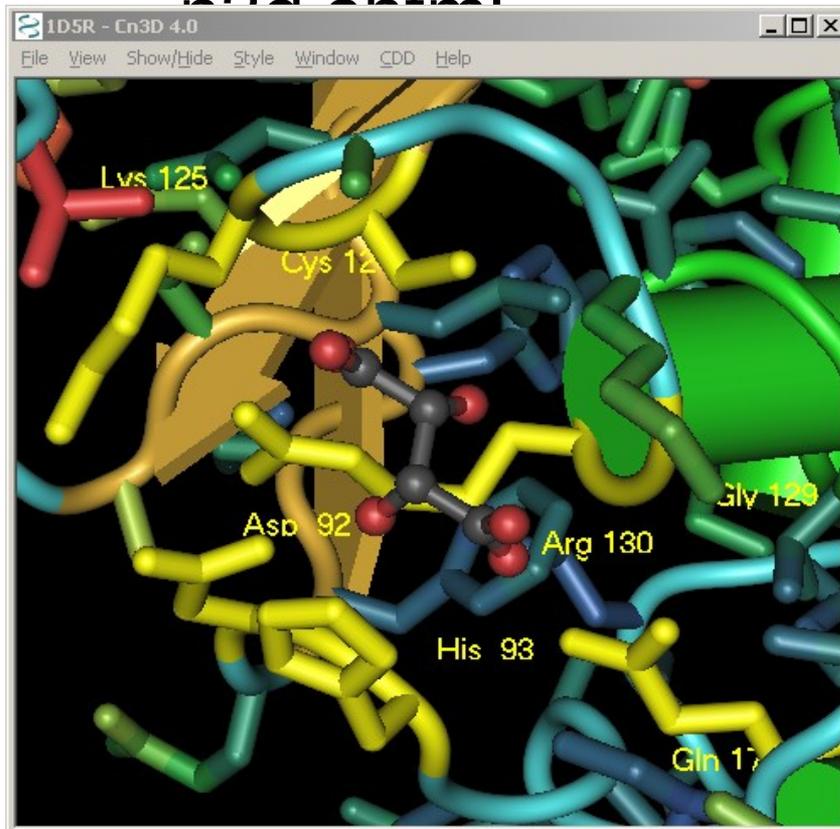
```


PDB soubor – terciární struktura

```
CRYST1    78.860    78.860    38.250    90.00    90.00    90.00 P 43 21 2      8  1HEW  113
ORIGX1          1.000000    0.000000    0.000000          0.000000          1HEW  114
ORIGX2          0.000000    1.000000    0.000000          0.000000          1HEW  115
ORIGX3          0.000000    0.000000    1.000000          0.000000          1HEW  116
SCALE1         0.012681    0.000000    0.000000          0.000000          1HEW  117
SCALE2          0.000000    0.012681    0.000000          0.000000          1HEW  118
SCALE3          0.000000    0.000000    0.026144          0.000000          1HEW  119
ATOM         1  N  LYS      1      3.398    9.981    10.408    1.00  30.48    1HEW  120
ATOM         2  CA LYS      1      2.459   10.365     9.364    1.00  28.03    1HEW  121
ATOM         3  C  LYS      1      2.458   11.880     9.149    1.00  21.93    1HEW  122
ATOM         4  O  LYS      1      2.481   12.672    10.100    1.00  14.10    1HEW  123
ATOM         5  CB LYS      1      1.026    9.935     9.695    1.00  30.54    1HEW  124
ATOM         6  CG LYS      1      0.028   10.169     8.558    1.00  37.93    1HEW  125
ATOM         7  CD LYS      1     -1.415   10.089     9.048    1.00  33.23    1HEW  126
ATOM         8  CE LYS      1     -2.357   10.822     8.082    1.00  32.17    1HEW  127
ATOM         9  NZ LYS      1     -3.661   10.090     8.025    1.00  31.92    1HEW  128
ATOM        10  N  VAL      2      2.429   12.232     7.880    1.00  17.30    1HEW  129
ATOM        11  CA VAL      2      2.395   13.653     7.465    1.00  14.47    1HEW  130
ATOM        12  C  VAL      2      0.977   13.868     6.903    1.00  17.58    1HEW  131
ATOM        13  O  VAL      2      0.642   13.368     5.826    1.00  32.65    1HEW  132
ATOM        14  CB VAL      2      3.533   14.012     6.536    1.00  22.88    1HEW  133
```

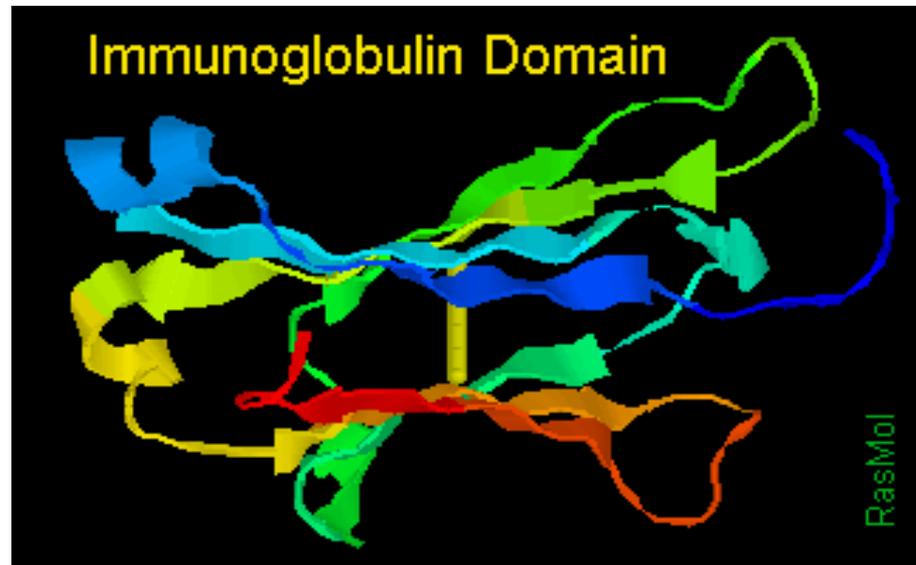
Cn3D

<http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>



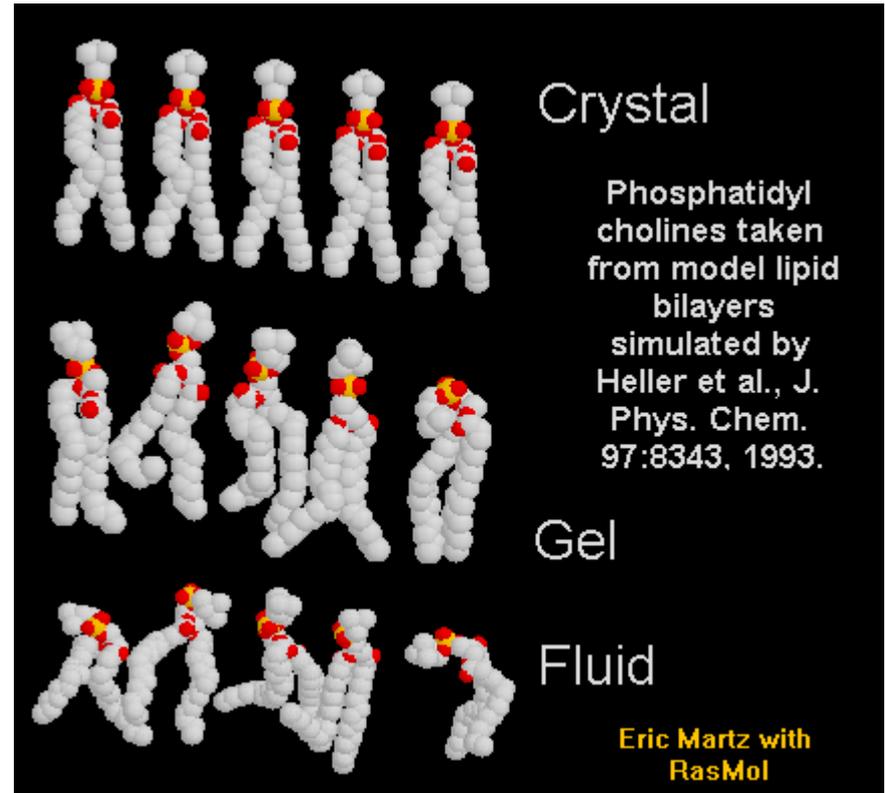
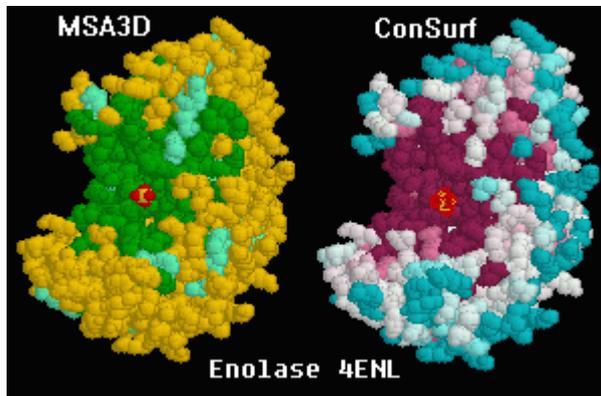
RasMol

<http://www.umass.edu/microbio/rasmol>



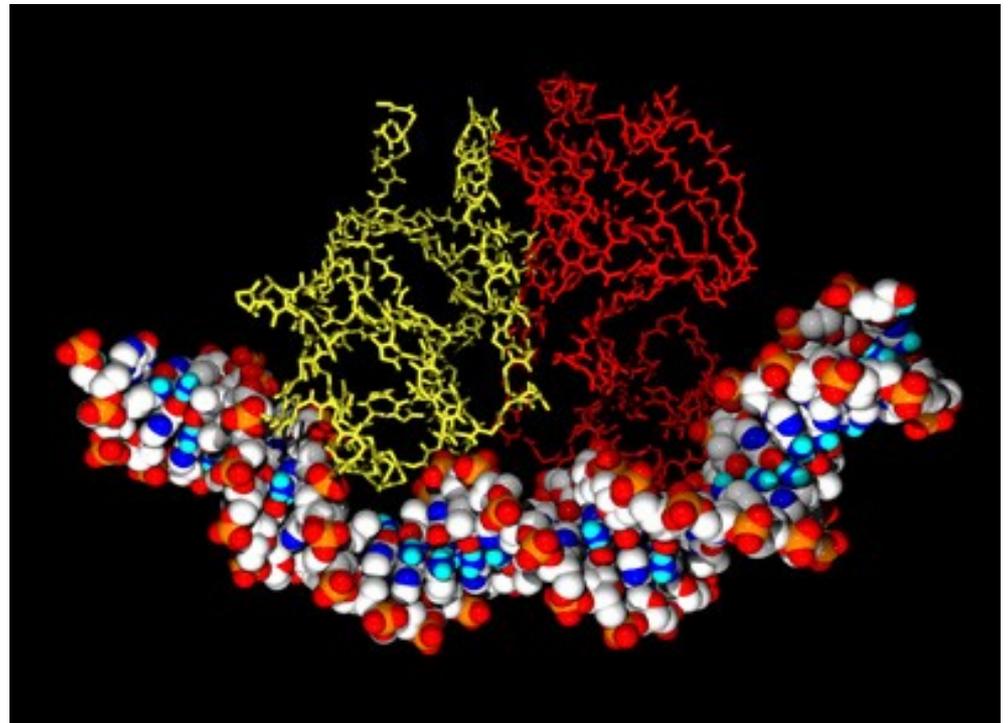
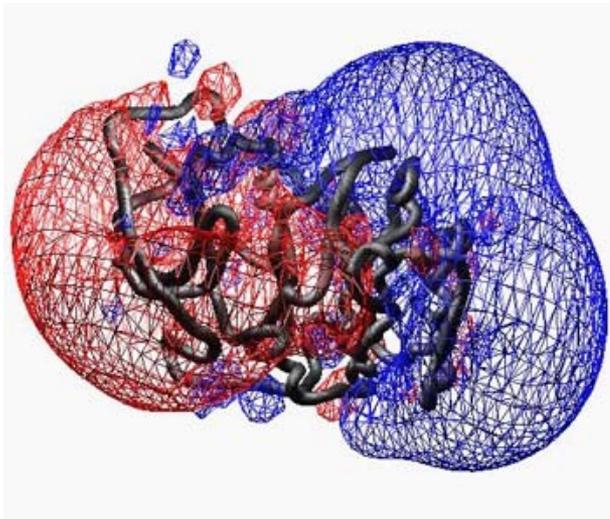
Chime/Protein Explorer

<http://www.umass.edu/microbio/chime/explorer/preview.htm>



Swiss PDB Viewer

<http://www.expasy.org/spdbv>



VMD <http://www.ks.uiuc.edu/Research/vmd/>
PyMol <http://pymol.sourceforge.net/>

Dílčí funkce proteinů

- ✦ Enzymy (katalyzátory, substrát se proměňuje v produkt, aktivní místo)
- ✦ Interakce protein-protein
- ✦ Interakce protein-DNA
- ✦ Interakce protein-ligand
- ✦ Transdukce signálu, regulace
- ✦ Strukturální proteiny (vlákna, glykoproteiny)
- ✦ Motory

Gene Ontology

Funkce genů/proteinů jsou zjišťovány experimentálně a publikovány v časopisech.

Terminologie není zdaleka jednoznačná:
protein synthesis – translation – ribosomal complex –
peptide chain elongation

Ontologie jsou vytvářeny ve snaze zavést do popisu funkcí určitý systém

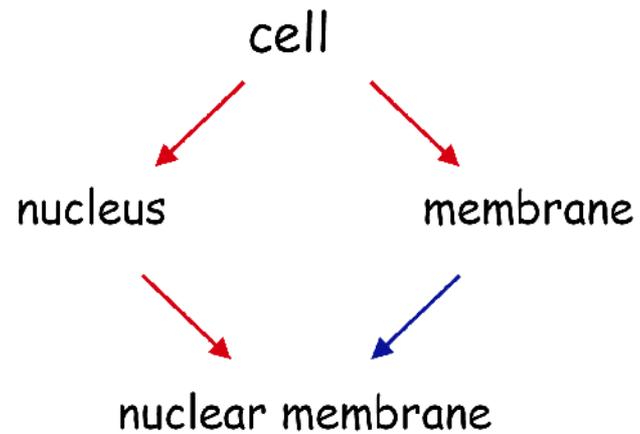
Gene Ontology

- definice ontologií – strukturovaných sad termínů (DAG) pro popis
 - biologické funkce
 - molekulární funkce
 - lokalizace
- přiřadování uzlů v ontologiích genům/proteinům
- vytváření nástrojů pro využití dat
- <http://www.geneontology.org/>



Structure of the GO

is-a
part-of



b

molecular function

nucleic acid binding

enzyme

DNA binding

helicase

adenosine triphosphatase

chromatin binding

DNA helicase

ATP-dependent helicase

DNA-dependent adenosine triphosphatase

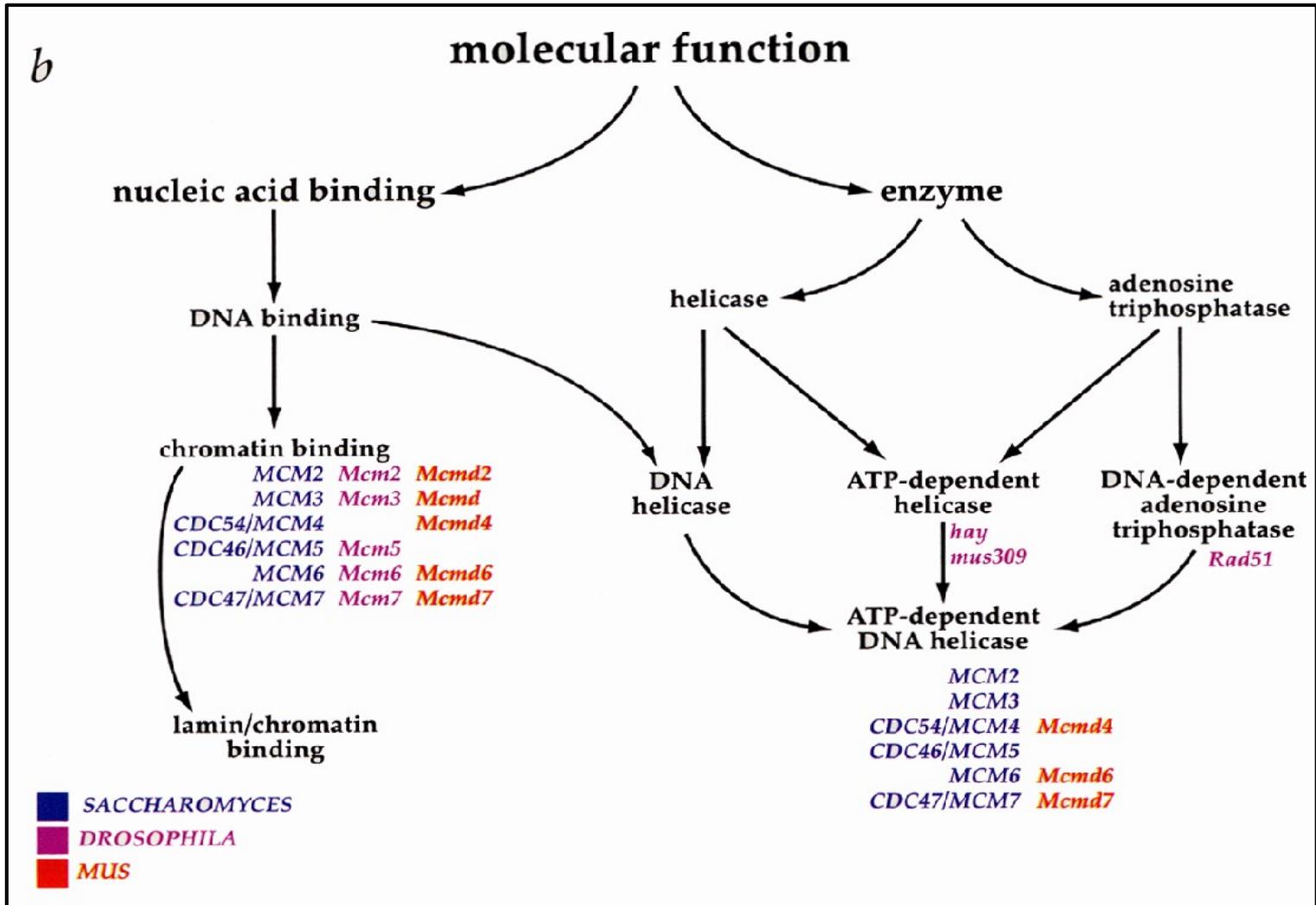
MCM2 Mcm2 Mcmd2
MCM3 Mcm3 Mcmd3
CDC54/MCM4 Mcm4 Mcmd4
CDC46/MCM5 Mcm5 Mcmd5
MCM6 Mcm6 Mcmd6
CDC47/MCM7 Mcm7 Mcmd7

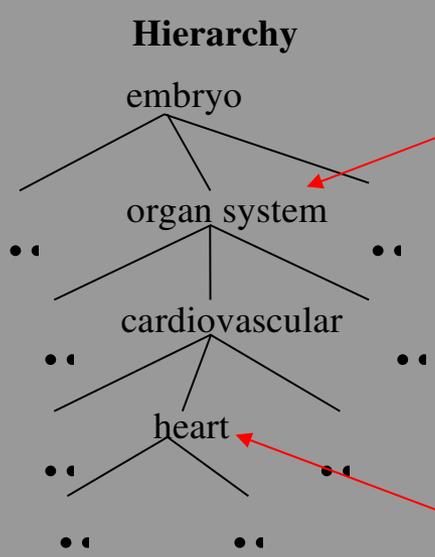
ATP-dependent DNA helicase

MCM2
MCM3
CDC54/MCM4 Mcmd4
CDC46/MCM5
MCM6 Mcmd6
CDC47/MCM7 Mcmd7

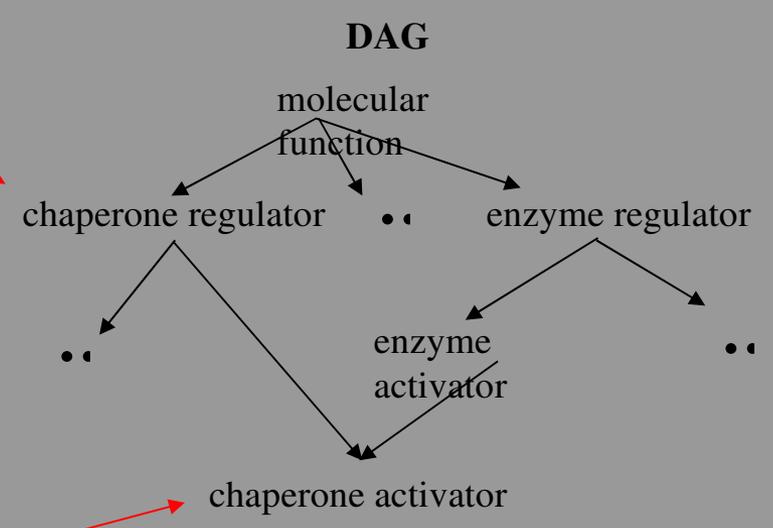
lamin/chromatin binding

SACCHAROMYCES
 DROSOPHILA
 MUS





Query for this term



Returns things annotated to descendents

Gene Ontology - ontologie

[Term]

id: GO:0006903
name: vesicle targeting
namespace: biological_process
def: "Targeting of a vesicle to a specific destination membrane." [GO:jic]
relationship: part_of GO:0016192 ! vesicle-mediated transport

[Term]

id: GO:0006904
name: vesicle docking during exocytosis
namespace: biological_process
def: "The initial attachment of a vesicle membrane to a target membrane\, mediated by proteins protruding from the membrane of the vesicle and the target membrane\, during exocytosis." [GO:jic]
subset: gosubset_prok
is_a: GO:0048278 ! vesicle docking
relationship: part_of GO:0006887 ! exocytosis

[Term]

id: GO:0006905
name: vesicle transport
namespace: biological_process
def: "OBSOLETE (was not defined before being made obsolete)." [GO:curators]
comment: This term was made obsolete because the meaning of the term is ambiguous. To update annotations\, consider the biological process term 'vesicle-mediated transport ; GO\:0016192'.
is_obsolete: true

Gene Ontology

- alternativní formát ontologie

```
%polyol catabolism ; GO:0046174 % polyol metabolism ; GO:0019751
%alditol catabolism ; GO:0019405 % alditol metabolism ; GO:0019400
%hexitol catabolism ; GO:0019407 % hexitol metabolism ; GO:0006059
%galactitol catabolism ; GO:0019404
%mannitol catabolism ; GO:0019592
%sorbitol catabolism ; GO:0006062
%pentitol catabolism ; GO:0019527 % pentitol metabolism ; GO:0019519
%arabitol catabolism ; GO:0051157 % arabitol metabolism ; GO:0051161
%arabitol utilization ; GO:0019591
%D-arabitol catabolism ; GO:0051159 % D-arabitol metabolism ; GO:0051163
%D-arabitol catabolism to xylulose 5-phosphate ; GO:0019528
```

Gene Ontology - anotace

TAIR gene:1944535 ERS2 GO:0004673 TAIR:Communication:1675000
ISS F ETHYLENE RESPONSE SENSOR 2 ERS2 PROTEIN|AT1G04310|
ETHYLENE RESPONSE SENSOR 2 gene taxon:3702 20020827 TAIR

TAIR gene:1944536 ETR1 GO:0005783 TAIR:Publication:1547355|
PMID:11916973 IDA C ETHYLENE RESPONSE 1 ETR|HISTIDINE KINASE
ETR1|AT1G66340|EIN1|ETHYLENE INSENSITIVE 1|ETHYLENE RESPONSE 1 gene taxon:3702
20020904 TAIR

TAIR gene:1944536 ETR1 GO:0009727 TAIR:Publication:1795|
PMID:9974395 IMP P ETHYLENE RESPONSE 1 ETR|HISTIDINE KINASE
ETR1|AT1G66340|EIN1|ETHYLENE INSENSITIVE 1|ETHYLENE RESPONSE 1 gene taxon:3702
20020904 TAIR

TAIR gene:1944536 ETR1 GO:0004673 TAIR:Communication:1675000
ISS F ETHYLENE RESPONSE 1 ETR|HISTIDINE KINASE ETR1|AT1G66340|
EIN1|ETHYLENE INSENSITIVE 1|ETHYLENE RESPONSE 1 gene taxon:3702 20020827
TAIR

TAIR gene:1944538 ETR2 GO:0004673 TAIR:Communication:1675000
ISS F ETHYLENE RESPONSE 2 ETHYLENE RESPONSE 2|AT3G23150|ETR2
gene taxon:3702 20020827 TAIR



DictyBase



The Wellcome Trust
Sanger Institute

The *Schizosaccharomyces pombe*
Genome Sequencing Project

The Pathogen Group



FlyBase



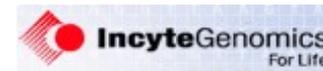
GENE ONTOLOGY™ CONSORTIUM

InterPro

<http://www.geneontology.org>



Genome
Knowledge Base



WormBase



ZFIN

AstraZeneca



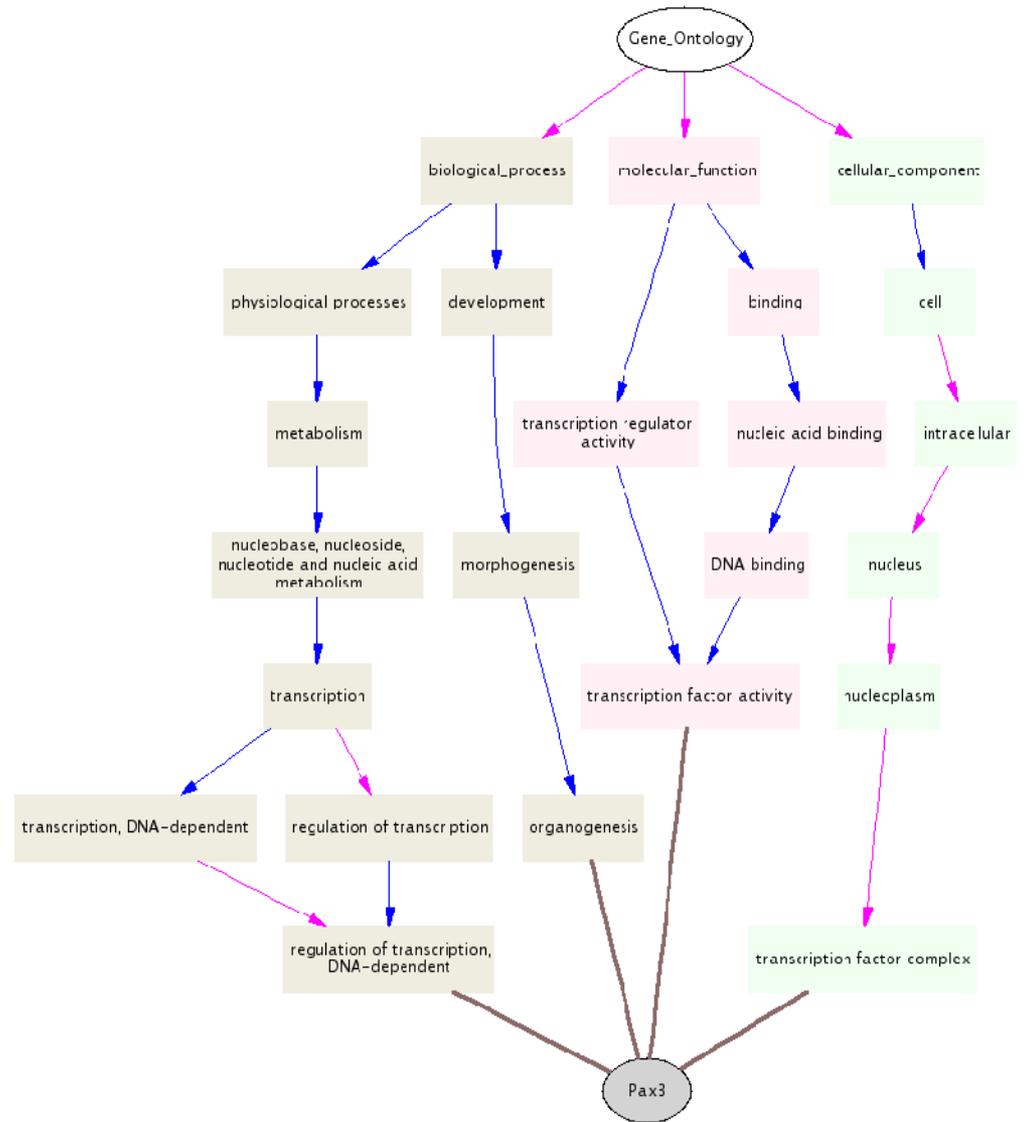
TJ

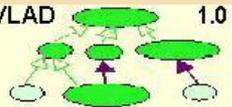


paired box gene 3

paired domain gene 3; PAX3/FKHR fusion gene; paired domain gene HuP2; paired box homeotic gene 3

Plays a critical role during fetal development. Mutations are associated with Waardenburg syndrome, craniofacial-deafness-hand syndrome and alveolar rhabdomyosarcoma.



VLAD 1.0
**VLAD - Visual Annotation Display**

VLAD is a tool for visualizing GO annotations. The annotation data are a subset of those available from the Gene Ontology project. VLAD will produce a graphical summary of the annotations.

Using the GO for data analysis...is there a functional “theme” in your set of genes?

[Help](#) [Feedback](#) Download
(Some...) (problems? suggestions? requests?) (coming soon)

Vocabularies:
 Biological Process
 Molecular Function
 Cellular Component

Annotation Set: MGI

Annotations Filters:
 Exclude IEA annotations.

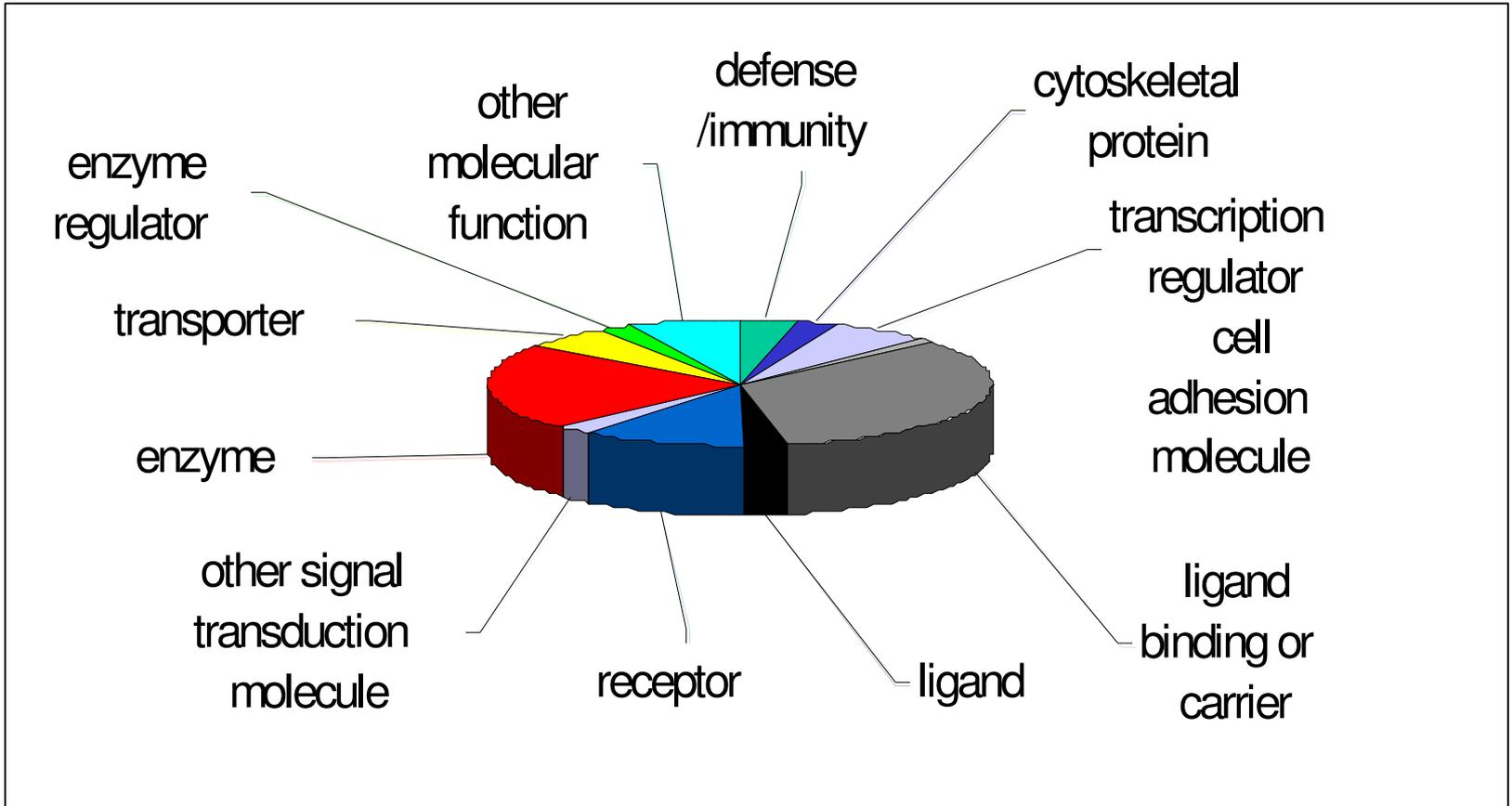
Query Set:
(gene symbols and/or IDs; use textarea or upload file)

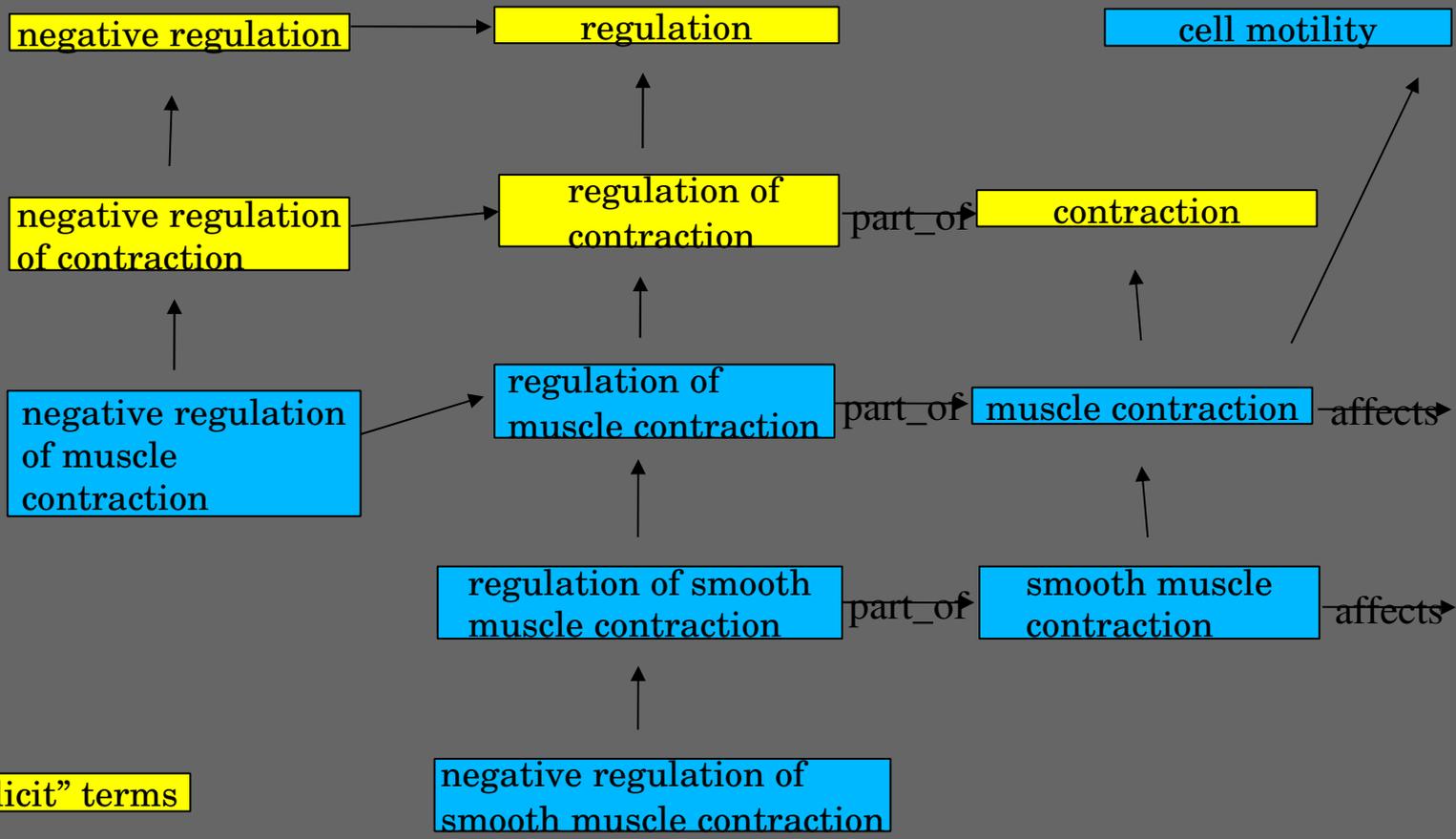
Upload file of symbols/IDs:

Scoring:
 Percentages
 P-values

Display Settings:
20 Collapsing threshold.

<http://proto.informatics.jax.org/prototypes/vlad/>





“implicit” terms

actual GO terms

AmiGO

<http://www.godatabase.org/>

The screenshot shows a Netscape browser window titled "AmiGO : GO:0000724 details - Netscape". The page features the AmiGO logo and navigation links: "GOst Search", "Get this GO term as RDF XML.", and "Get this data as a GO flat file.". The main heading is "double-strand break repair via homologous recombination". Below this, it lists "Accession:GO:0000724" and "Synonyms: None.". The "Definition:" section states: "The error-free repair of a double-strand break in DNA in which the broken DNA molecule is repaired using homologous sequences. The restoration of two intact DNA molecules results in the exchange, reciprocal or non-reciprocal, of genetic material between the intact DN...". The "Term Lineage" section shows a hierarchical tree of terms: "GO:0003673 : Gene Ontology (72737)" is the root, followed by "GO:0008150 : biological process (49448)", "GO:0007582 : physiological processes (35542)", "GO:0008152 : metabolism (24141)", "GO:0006139 : nucleobase, nucleoside, nucleotide and nucleic acid metabolism (9104)", "GO:0006259 : DNA metabolism (2822)", "GO:0006310 : DNA recombination (947)", "GO:0006312 : mitotic recombination (52)", "GO:0000725 : recombinational repair (32)", "GO:0000724 : double-strand break repair via homologous recombination (32)", "GO:0006281 : DNA repair (733)", "GO:0006302 : double-strand break repair (77)", "GO:0000724 : double-strand break repair via homologous recombination (32)", "GO:0000725 : recombinational repair (32)", and "GO:0000724 : double-strand break repair via homologous recombination (32)". The "External References" section is empty. The "Associated Genes" section includes "Gene Filters: Filter by database:" with a dropdown menu showing "All", "FlyBase", and "SGD", and "Filter by Evidence for Association:" with a dropdown menu showing "Curator Approved". The page number "Page 1" is visible at the bottom left. The browser's status bar at the bottom indicates "Document: Done (5,208 secs)".