

INTERNET RESOURCES

PREDICTION OF PHYSICAL PROPERTIES

Compute pI/MW	http://www.expasy.ch/tools/pi_tool.html
MOWSE	http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse
PeptideMass	http://www.expasy.ch/tools/peptide-mass.html
TGREASE	ftp://ftp.virginia.edu/pub/fasta/
SAPS	http://www.isrec.isb-sib.ch/software/SAPS_form.html

PREDICTION OF PROTEIN IDENTITY BASED ON COMPOSITION

AACompIdent	http://www.expasy.ch/tools/aacomp/
AACompSim	http://www.expasy.ch/tools/aacsim/
PROPSEARCH	http://www.embl-heidelberg.de/prs.html

MOTIFS AND PATTERNS

BLOCKS	http://blocks.fhcrc.org
Pfam	http://www.sanger.ac.uk/Software/Pfam/
PRINTS	http://www.bioinf.man.ac.uk/dbbrowser/PRINTS/PRINTS.html
ProfileScan	http://www.isrec.isb-sib.ch/software/PFSCAN_form.html

PREDICTION OF SECONDARY STRUCTURE AND FOLDING CLASSES

nnpredict	http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html
PredictProtein	http://www.embl-heidelberg.de/predictprotein/
SOPMA	http://pbil.ibcp.fr/
Jpred	http://jura.ebi.ac.uk:8888/
PSIPRED	http://insulin.brunel.ac.uk/psipred
PREDATOR	http://www.embl-heidelberg.de/predator/predator_info.html

PREDICTION OF SPECIALIZED STRUCTURES OR FEATURES

COILS	http://www.ch.embnet.org/software/COILS_form.html
MacStripe	http://www.york.ac.uk/depts/biol/units/coils/mstr2.html
PHDTopology	http://www.embl-heidelberg.de/predictprotein
SignalP	http://www.cbs.dtu.dk/services/SignalP/
TMpred	http://www.isrec.isb-sib.ch/ftp-server/tmpred/www/TMPRED_form.html

STRUCTURE PREDICTION

DALI	http://www2.ebi.ac.uk/dali/
Bryant-Lawrence	ftp://ncbi.nlm.nih.gov/pub/pkb/
FSSP	http://www2.ebi.ac.uk/dali/fssp/
UCLA-DOE	http://fold.doe-mpi.ucla.edu/Home
SWISS-MODEL	http://www.expasy.ch/swissmod/SWISS-MODEL.html
TOPITS	http://www.embl-heidelberg.de/predictprotein/



APPENDIX 2

Useful web sites

Good starting points for 'surfing' in molecular biology:

National Center for Biotechnology Information

<http://www.ncbi.nlm.nih.gov/>

European Bioinformatics Institute

<http://www.ebi.ac.uk>

Expasy (Swiss Institute of Bioinformatics)

<http://www.expasy.ch/>

GenomeNew (Kyoto University and University of Tokyo)

<http://www.genome.ad.jp/>

Collection of on-line analysis tools, including database searches

<http://www-biol.univ-mrs.fr/english/logligne.html>

<http://www.fccc.edu/research/labs/dunbrack/molecularmodeling.html>

Index of web sites in molecular biology, including specialized databases

<http://www.cbs.dtu.dk/biolink.html>

BCM search launcher—various database searches and associated tools

<http://kiwi.imgen.bcm.tmc.edu:8088/search-launcher/launcher.html>

Collections of protein analysis tools

<http://www.graylab.ac.uk/cancerweb/research/protanal.html>

<http://www-biol.univ-mrs.fr/english/logligne.html>

General information:

Primer on molecular genetics

http://www.ornl.gov/TechResources/Human_Genome/publicat/primer/prim1.html

Human genome project information

<http://www.ornl.gov/hgmis/>

Genome sequencing project information:

<http://www.mcs.anl.gov/home/gaasterl/genomes.html>

<http://www-biol.univ-mrs.fr/english/genome.html>

(Organelles): <http://megasun.bch.umontreal.ca/ogmpproj.html>

Growth of databanks http://www.genome.ad.jp/dbget/db_growth.gif
http://www.genome.ad.jp/dbget/db_growth.html

Database of metabolic pathways

http://www-c.mcs.anl.gov/home/compbio/PUMA/Production/puma_graphics.html

Electronic scholarly publishing <http://www.esp.org/>

Bibliographic database: PubMed (U.S. National Library of Medicine)

<http://www.ncbi.nlm.nih.gov/PubMed/>

Primarily sequence information:

Sequence retrieval:

SRS <http://srs.ebi.ac.uk/>

Entrez <http://www.ncbi.nlm.nih.gov/Entrez/>

Oligopeptide dictionary at PIR

<http://www-nbrf.georgetown.edu/pirwww/search/patmatch.html>

Calculation of multiple sequence alignments

<http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html>

Prosite database of sequence motifs:

<http://www.expasy.ch/prosite/>

Scan Prosite: <http://www.expasy.ch/tools/scnpsit1.html>

Collections of multiple sequence alignments:

<http://www2.ebi.ac.uk/dali/fssp/fssp.html>

<http://www.sanger.ac.uk/Pfam> or <http://pfam.wustl.edu/>

Analysis of genome sequences:

PEDANT <http://pedant.mips.biochem.mpg.de>

Primarily structure information:

Polypeptide conformation <http://www.chem.qmw.ac.uk/iupac/misc/biop.html>

Protein Data Bank <http://www.rcsb.org>

ReLiBase (receptor-ligand complexes)

<http://www.pdb.bnl.gov:8081/home.html>

Promise: prosthetic groups and metal ions in protein active sites

<http://bmbgsi11.leeds.ac.uk/promise/MAIN.html>

Secondary structure assignments:

DSSP <http://ara.ebi.ac.uk/dssp/>

STRIDE http://www.emblheidelberg.de/cgi/stride_serv

Protein loop classification <http://bonsai.lif.icnet.uk/bmm/loop/>

Protein modules <http://www.bork.embl-heidelberg.de/Modules/>

Classifications of protein structures:

SCOP <http://scop.mrc-lmb.cam.ac.uk/scop/>

CATH <http://www.biochem.ucl.ac.uk/bsm/cath/>

DALI <http://www2.embl-ebi.ac.uk/dali/>

FSSP (Fold classification based on Structure-Structure alignment of

Proteins) <http://www2.embl-ebi.ac.uk/dali/fssp/>

Indices of other protein structure classifications

<http://www.bioscience.org/urlists/protdb.htm>

<http://msd.ebi.ac.uk/add/Links/fold.shtml>

Databases of protein sequences homologous to those of known structures:
FSSP (Fold classification based on Structure-Structure alignment of Proteins)

<http://www2.ebi.ac.uk/dali/fssp/>

HSSP (Homology-derived secondary structure)

<http://www.sander.embl-heidelberg.de/hssp/>

Access to structural databases at University College London

<http://www.biochem.ucl.ac.uk/bsm/biocomp/index.html#data bases>

Calculation of accessible surface area

<http://www.bork.embl-heidelberg.de/ASC/scr1-form.html>

Calculation of hydrophobicity profile

<http://bmbsgi11.leeds.ac.uk/bmb5dp/profiles.html>

Sites specialized to specific protein families:

Index: <http://msd.ebi.ac.uk/add/Links/family.html>

Globins: <http://bmbsgi11.leeds.ac.uk/promise/GLOBINS.html>

Protein kinases <http://www.sdsc.edu/kinases/>

MEROPS database of peptidases:

<http://www.bi.bbsrc.ac.uk/Merops/Merops.htm>

IUBMB EC list for peptidases:

<http://www.chem.qmw.ac.uk/iubmb/enzyme/EC34>

AIDS-related information

<http://www-fbsc.ncifcrf.gov/HIVdb/>

Proteins of the immune system:

General antibody-related material: <http://www.antibodyresource.com>

Searching the Kabat databank of antibody sequences and specificities

<http://immuno.bme.nwu.edu/>

Compilation of links to on-line databases and resources of immunological interest, (as well as much other material related to computational molecular biology)

<http://www.infobiogen.fr/services/deambulum/english/db5.html>

Immunogenetics database, plus links to sequence analysis tools

<http://www.genetik.uni-koeln.de/dnaplot/>

Conformational changes:

Database of conformational changes in proteins:

<http://bioinfo.mbb.yale.edu/MolMovDB/db/ProtMotDB.main.html>

Protein structure prediction:

Protein structure prediction centre

<http://predictioncenter.llnl.gov/>

Homology modelling server

<http://www.expasy.ch/swissmod/SWISS-MODEL.html>

Threaders:

UCLA server

<http://www.doe-mpi.ucla.edu/people/frsvr/frsvr.html>EMBL server <http://www.embl-heidelberg.de/predictprotein/predictprotein.html>Rotamer libraries: <http://www.fccc.edu/research/labs/dunbrack/index.html><http://duc.urbb.jussieu.fr/rotamer.html>

Major data archive projects in molecular biology

Name of data bank (home URL)	Type of data	Location
GenBank www.ncbi.nlm.nih.gov/	Nucleic acid sequences	National Library of Medicine, Washington, D.C., U.S.A.
EMBL Data Library www.ebi.ac.uk/ebi_docs/embl_db/ebi/topembl.html	Nucleic acid sequences	European Bioinformatics Institute, Hinxton, U.K.
DNA Data Bank of Japan www.ddbj.nig.ac.jp/	Nucleic acid sequences	National Institute of Genetics, Mishima, Japan
Protein Identification Resource www-nbrf.georgetown.edu/pir/	Amino acid sequences	Georgetown University, Washington, D.C., U.S.A.
Munich Information Center for Protein Sequences (MIPS) speedy.mips.biochem.mpg.de/	Amino acid sequences	Max-Planck-Institute für Biochemie, Martinsried, Germany
International Protein Information Database in Japan (JIPID)	Amino acid sequences	Science University of Tokyo Noda, Japan
SWISS-PROT www.expasy.ch/sprot/	Amino acid sequences	Geneva, Switzerland and Hinxton, U.K.
Protein Data Bank www.rcsb.org	Protein structures	Rutgers University, New Jersey, U.S.A.
Nucleic Acid Data Bank ndbserver.rutgers.edu/	Nucleic acid structures	Rutgers University, New Jersey, U.S.A.
BioMagResBank www.bmrb.wisc.edu/	NMR Structure determinations	Madison, Wisconsin, U.S.A.
Cambridge Structural Database www.ccdc.cam.ac.uk/	Small-molecule crystal structures	Cambridge, U.K.

Web Sites for database searches

Task	Program Name
Web site	
Retrieve one sequence	
http://srs.ebi.ac.uk/	Sequence Retrieval System (SRS)
http://www.ncbi.nlm.nih.gov/Entrez/	Entrez

Retrieve one structure www.rcsb.org	Protein Data Bank (PDB)
Match one sequence to one sequence http://www-hto.usc.edu/software/seqaln/seqaln-query.html http://vega.igh.cnrs.fr/bin/align-guess.cgi	ALIGN
Multiple sequence alignment http://www.ebi.ac.uk/clustalw	ClustalW
Probe sequence databank with sequence http://www.ncbi.nlm.nih.gov/blast/psiblast.cgi or http://www.sanger.ac.uk/Software/Pfam/search.shtml http://www.cse.ucsc.edu/research/compbio/HMM-apps/HMMapplications.html	PSI-BLAST Hidden Markov models
Probe structure databank with structure http://www2.embl-ebi.ac.uk/dali/	DALI
Probe structure databank with sequence (Fold recognition, threaders) http://www.embl-heidelberg.de/predictprotein/predictprotein.html http://www.doe-mpi.ucla.edu/people/frsvr/frsvr.html	
Bibliographical search http://www.ncbi.nlm.nih.gov/	U.S. National Library of Medicine (PubMed)