



# Bioinformatics up to Date

(Bioinformatics Center, Biotechnology Division)  
North-East Institute of Science & Technology  
Jorhat - 785 006, Assam

## CONTENTS

Cover Story	1
Bioinfy Quiz	1
Computational Chemistry	1
Proteomics	2
Bioserver	3
Upcoming Events	3
Computers for Biologists	4
Genomics	4
Molecule of the month	5
Bioinfy Animator	5
Contact Us	5
Answers to quiz	5

**Adviser:**  
Dr.P.G.Rao

**Editors:**  
Dr. R.L.Bezbaruah  
Pompi Sharma  
Dhrubjyoti Gogoi

## COVER STORY

In what is being billed as the largest study of its kind, a consortium of 44 research institutions identified four new genes that indicate Alzheimer's disease risk. Another group of investigators from the United States and Europe reported a fifth gene. This news doubles the numbers of genes now thought to contribute to Alzheimer's disease.

Rudolph Tanzi, a neurology professor at Harvard Medical School, told *Bloomberg* that he is most excited by one particular newfound Alzheimer's gene: CD33. It might be responsible for a failure of the body to eliminate as much beta amyloid as it should. Beta amyloid buildup has long been thought to play a key role in degradation of nerve cells in the brain.

Still, it will be a decade or more before any of these discoveries will translate into drugs, but it has the Alzheimer's research community buzzing over new avenues of possible inquiry. The research can be found in the journal *Nature Genetics*.

[www.fiercebiotechresearch.com](http://www.fiercebiotechresearch.com)

## COMPUTATIONAL CHEMISTRY

### Open BaBel : The Open S0urce chemiSTry tool box

Open Babel is a chemical toolbox designed to speak the many languages of chemical data. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas.

#### Goals of the Open Babel project

Open Babel is a project to facilitate the interconversion of chemical data from one format to another – including file formats of various types.

#### This is important for the following reasons:

- Multiple programs are often required in realistic workflows. These may include databases, modeling or computational programs, visualization programs, etc.
- Many programs have individual data formats, and/or

## BIOINFY QUIZ

1.SWISSPROT protein sequence database began in -

- A) 1987
- B)1985
- C)1986

2.An example of Homology & similarity tool?

- A)EMBOSS
- B)BLAST
- C)PROSPECT

3. First molecular biology server Expsy in the year?

- A)1991
- B)1993
- C)1992

4.Deposition of cDNA into inert structure is

- A)DNA microarray
- B)DNA fingerprinting
- C)DNA probe

5.Characterizing molecular component is -

- A)Genomics
- B)Chemiinformatics
- C)Bioinformatics

Answers on page 5

- Chemical representations often vary considerably:
  - Some programs are 2D. Some are 3D. Some use fractional k-space coordinates.
  - Some programs use bonds and atoms of discrete types. Others use only atoms and electrons.
  - Some programs use symmetric representations. Others do not.
- Some programs specify all atoms. Others use “residues” or omit hydrogen atoms.
- Individual implementations of even standardized file formats are often buggy, incomplete or do not completely match published standards.

## PROTEOMICS

### Protein information resource

The Protein Information Resource (PIR), located at Georgetown University Medical Center, Washington, D.C. is an integrated public bioinformatics resource to support genomic and proteomic research, and scientific studies. It was established in 1984 by the National Biomedical Research Foundation (NBRF) as a resource to assist researchers in the identification and interpretation of protein sequence information. PIR has provided many protein databases and analysis tools freely accessible to the scientific community, including the Protein Sequence Database (PSD), the first international database, which grew out of Atlas of Protein Sequence and Structure. In 2002, PIR along with its international partners, EBI (European Bioinformatics Institute) and SIB (Swiss Institute of Bioinformatics), were awarded a grant from NIH (National Institute of Health) to create UniProt, a single worldwide database of protein sequence and function, by unifying the PIR-PSD, Swiss-Prot, and TrEMBL databases. PIR offers a wide variety of resources mainly oriented to assist the propagation and standardization of protein annotation:

**PIRSF** - Protein family classification system

**iProClass** - Integrated protein knowledgebase

**iProLINK** - Literature, information & knowledge



## BIOSERVER

### APSSP2: AdvAnced Protein SecondAry Structure Prediction Server

This server allow to predict the secondary structure of protein's from their amino acid sequence. This is an advanced version of PSSP server, which participated in CASP3 and in CASP4. PSSP was also part of CAFASP2. Raghava, G. P. S. (2000) Protein secondary structure prediction using nearest neighbor and neural network approach. CASP4: 75-76. This server is also participating in world-wide Live-Bench competition EVA, so one can get the performance of methods including APSSP2 from EVA Server. This server is also part of Meta II Prediction server.

[Request Form](#)

Target/name of protein (optional):

Paste your sequence data here:

Please Select the Format of your Sequence (File or Paste)

Single sequence - amino acids only

Please Enter your Email:

(Please enter your email address if you want to receive your result via email)

[www.imtech.res.in](http://www.imtech.res.in)

## UPCOMING EVENTS

6th International Symposium on Health Informatics and Bioinformatics - HIBIT 2011 from May 2-5, 2011 @ Izmir, Turkey

## COMPUTERS FOR BIOLOGISTS

### MobyLe@RbPS

MobyLe is a framework and web portal specifically aimed at the integration of bioinformatics software and databanks.

#### Functionalities:

- **data reusability:** the tagging of the user data facilitates the reuse of input values or results between different programs.
- **automatic data validation and format conversion:** the description of the expected data and their format allows to verify and convert input values if necessary.
- **service discovery and workflow authoring assistance:** services are provided through a searchable menu; furthermore, type compatibility mechanisms between results and potential program inputs let users either interactively pipe tasks or build complete workflows before to run them.

www.mobyle.rpbs.univ-paris-diderot.fr

## GENOMICS

### CMR

The Comprehensive Microbial Resource (CMR) is a free website used to display information on all of the publicly available, complete prokaryotic genomes. In addition to the convenience of having all of the organisms on a single website, common data types across all genomes in the CMR make searches more meaningful, and cross genome analysis highlight differences and similarities between the genomes. CMR offers a wide variety of tools and resources, all of which are available off of our menu bar at the top of each page.

Genomes: Search for *Bacillus anthracis* Download

723 Total Genomes  
 6 Complete, 5 Incomplete, 0 Eukarya, 0 Archaea, 11 Bacteria, 0 Viruses  
 [G] = GenBank FTP [T] = JCVI FTP [S] = Sequencing Center Genome Page [N] = NCBI Genome Page [P] = Publication

Organism Name (sort)	Kingdom (sort)	Taxon ID (sort)	Size (sort)	Complete Genome (sort)	Sequencing Center (sort)	Links (sort)
<a href="#">Bacillus anthracis A0029</a>	Bacteria	261620	5.22 Mb	No	J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>
<a href="#">Bacillus anthracis A0248</a>	Bacteria	592021	5.50 Mb	Yes	Los Alamos National Laboratory, J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>
<a href="#">Bacillus anthracis Ames</a>	Bacteria	198094	5.22 Mb	Yes	J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>
<a href="#">Bacillus anthracis Ames Ancestor</a>	Bacteria	261624	5.50 Mb	Yes	J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>
<a href="#">Bacillus anthracis CNH14-9095</a>	Bacteria	280354	5.24 Mb	No	J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>
<a href="#">Bacillus anthracis Sterne</a>	Bacteria	260739	5.22 Mb	Yes	Los Alamos National Laboratory DOE, J. Craig Venter Institute	<a href="#">G</a> <a href="#">T</a> <a href="#">S</a> <a href="#">N</a> <a href="#">P</a>

www.cmr.jcvi.org

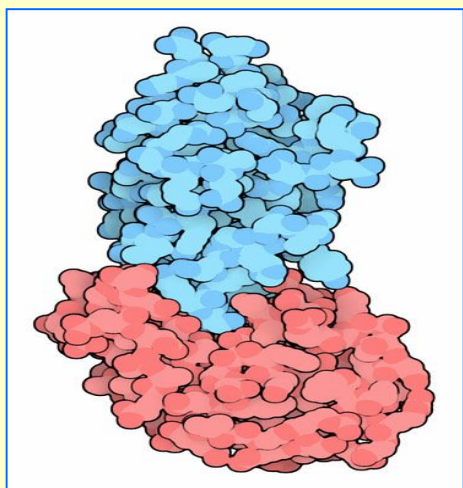
## MOLECULE OF THE MONTH NANOBODIES

Nanobodies are similar fragments taken from the unusual camel antibodies. The arms are cut off, creating a smaller, more stable molecule that, like an antibody Fab, is highly specific for a given target molecule. The one shown here binds specifically to lysozyme. The nanobody is shown in blue and lysozyme is shown in red. Researchers have already used them in research, and are currently exploring them for use as drugs.

PDB ID: 1mel

Method: X-RAY DIFFRACTION

R-Value: 0.221 (obs.)

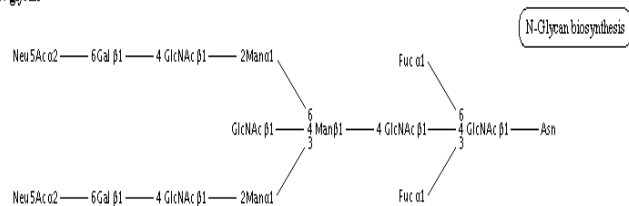


www.pdb.org

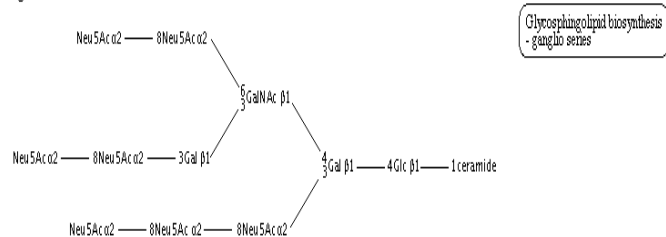
## OTHER GLYCAN DEGRADATION- Reference pathway

### OTHER GLYCAN DEGRADATION

#### N-glycan



#### Ganglioside



www.genome.jp

## Answers of Bioinfy Quiz

1) 1987 2)BLAST 3)1992 4)DNA microarray 5)Cheminformatics

### CONTACT US:

Pompi Sharma, Project Assistant (Level – II)

E-mail: pompi.sharma86@gmail.com

Dhrubajyoti Gogoi, Project Assistant (Level – II)

E-mail: dhruba.bio.du@gmail.com