

COURSE-M.Sc. Botany Part-II

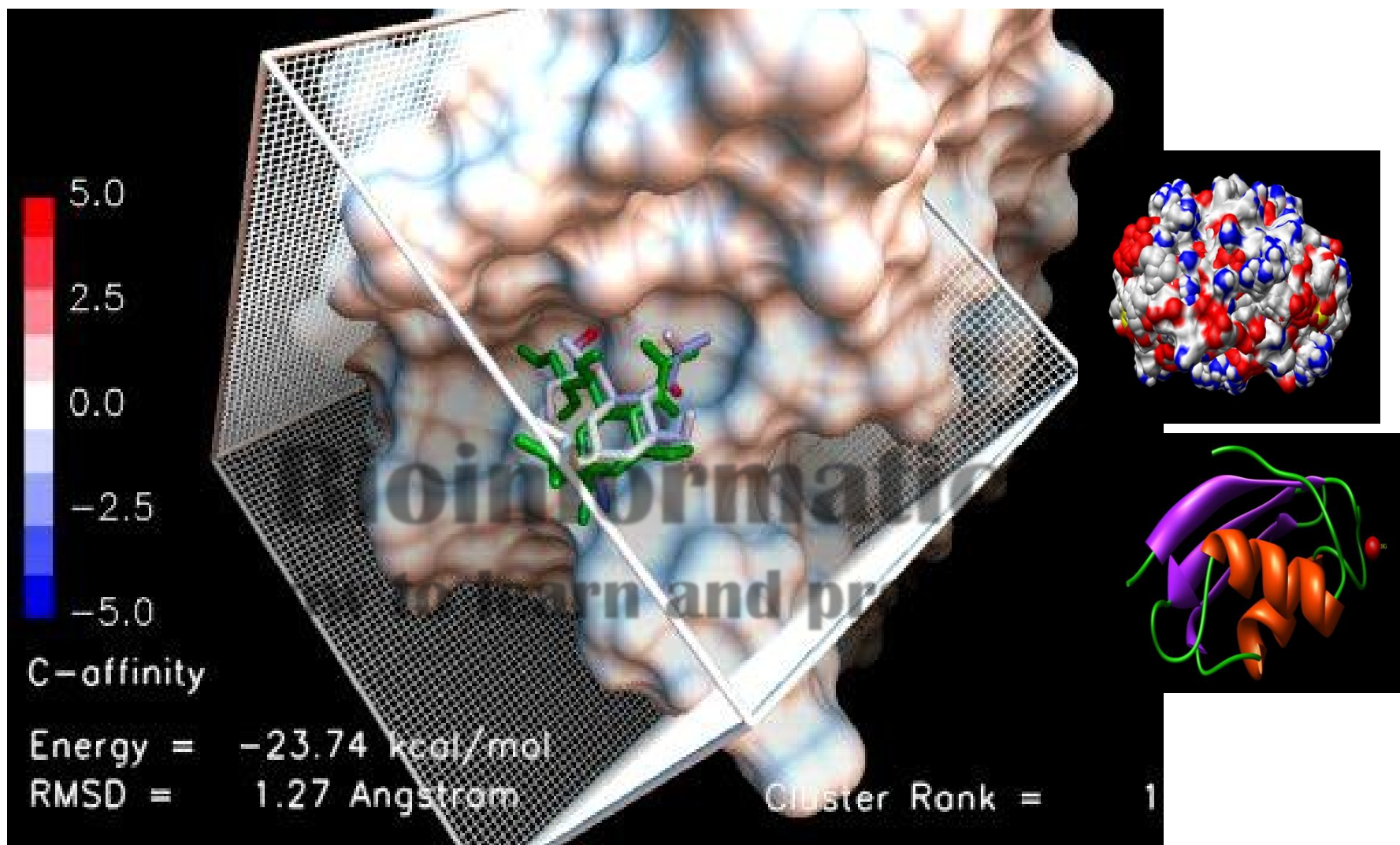
Paper-XVI

Topic – Bioinformatics

How to learn and practice ?

Prepared by -Prof. (Dr) Jainendra Kumar

Co-ordinated by- Prof. (Dr) Shyam Nandan Prasad



Jainendra Kumar

Being an interface between modern biology and informatics (computer science), the area of bioinformatics is meant to discover, develop and implement computational algorithms and software tools that facilitate better understanding of the biological processes with the goal to serve agriculture, environment, food processing, chemical and healthcare sectors

In the area of agriculture, bioinformatics can be used for increasing the nutritional content and yield of the crops, inducing disease resistance into them etc..

In the pharmaceutical sector, it can be used to reduce the time and cost involved in drug discovery process, to custom design drugs and to develop personalized medicine . . .

In other sectors too, the goal of bioinformatics is to enhance the quality and increase the volume of the products.

How to learn the skills of bioinformatics?

Follow these steps –

- 1. Learn computer basics – the operational components**
- 2. Learn applications of MS office especially Ms word and Power point (later MS Excel and MS Access)**
- 3. Learn the efficient use of internet**
- 4. Learn efficient use of Google and other search engines**

You may have your own PC or access to Internet outlets

5. Learn to download files from websites directly and through FTP (File transfer Protocol) from FTP sites

6. Learn C++ and *perl* programming (if not now, then, later)

7. Read the book **Understanding DNA and Gene cloning by **Karl Drlica** published by John Wiley & Sons**

8. Read the book **Introduction to Bioinformatics by **Teresa K. Attwood and David J. Parry-Smith** published by Pearson Education limited (Indian reprint) at least twice (you can read other books later)**

9. If you have done or doing computer application course such as BCA or learnt programming by yourself, read the book **Computational Molecular Biology (An algorithmic approach) by **Pavel A. Pevzner** published by Prentice Hall of India Pvt. Ltd. (Then, other books, of course!)**

10. Know about the Databases and their applicability before you go ahead!

What is database?

A database is a computerized archive where biological (or other) data are stored and organized in a unique way with the help of specially written programs/applications so that the information can be retrieved as and when required by anybody.

Database System

It is a collection of files and data that are inter-related representing a particular field of study. It contains all details about their storage, organization, interpretation, usability and maintenance. It is executed by a software called “Database Management System” (DBMS).

A database system also protects the data from unauthorized access.

Types of databases

- 1. DBs of nucleotide sequences**
- 2. DBs of protein sequences**
- 3. DBs of proteins sequence patterns or motifs**
- 4. DBs of macromolecular 3D structures**
- 5. DBs of ligands and micromolecules**
- 6. DBs of gene expressions**
- 7. DBS of genomic information**
- 8. DBs of proteomic information**
- 9. DBs of metabolic pathways**
- 10. Secondary databases (that contain the results of analysis of the primary databases)**

NCBI (National Centre of Biotechnological Information)

Entrez is the search engine of NCBI

After I finish, I suggest you to try it by going to the NCBI site <http://www.ncbi.nlm.nih.gov/>

Now, we start with structural bioinformatics learning to download 3D atomic coordinate files of proteins from www.rcsb.org/ (Protein data bank site)

11. Download the pdb files of the enzymes of glycolysis having 4-letter codes –

1ig8 (of yeast hexokinase), **1dgk** (human hexokinase), **1hox** (phospho glucoisomerase), **4pfk** (phospho fructokinase), **4ald/1j4e** (human muscle aldolase), **1zen** (bacterial aldolase), **2ypi** (triose phosphate isomerase), **1gpd** (phosphoglyceraldehyde dehydrogenase), **1vpe** (phosphoglycerate kinase), **1eqj/ 1e58** (Phosphoglycero mutase) and **2one** (enolase), **1a3w** (pyruvate kinase)

12. You can use Rasmol, Pymol and other numerous softwares to render 3D structure of these enzymes and construct the glycolytic pathway.

13. You may learn about modern tools and techniques of bioinformatics on-line by visiting the bioinformatics website - <http://www.bioinformatics.org/education/>

14. Then....., go for other tools, applications, techniques and devices of bioinformatics

15. Join a course to learn more advance techniques

Congratulations !

We hope you have learnt something!

... good wishes