

SYLLABUS

for the

MSc (Bioinformatics & Biotechnology)

**offered by the
INSTITUTE OF BIOINFORMATICS
AND APPLIED BIOTECHNOLOGY
(IBAB)**

**BIOTECH PARK
ELECTRONICS CITY, PHASE I
BANGALORE 560 100**

Tel: 080-285 289 00/285 289 01/285 289 02

Fax: 080-285 289 04

<http://www.ibab.ac.in>

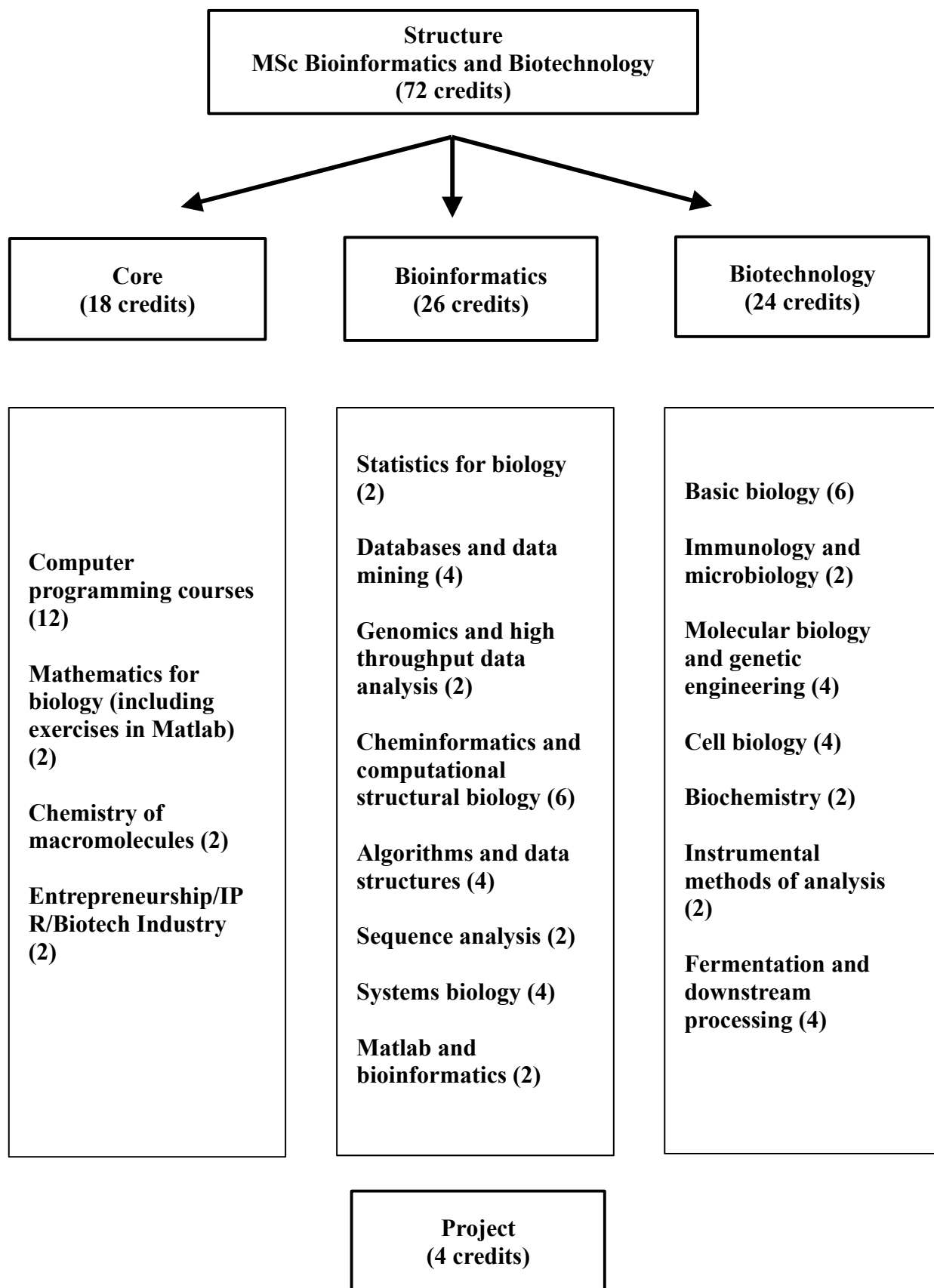
recognized by the

**INDIRA GANDHI NATIONAL OPEN
UNIVERSITY (IGNOU)
NEW DELHI**

Preamble

Biological information has been accumulating over the last 50 years from a variety of investigations combined with an unprecedented deluge of data from the sequencing and subsequent analysis of a large number of genomes. This together with the projected cascade of information anticipated from personal genomics is creating an enormous information overload challenging life scientists in terms of data storage, dissemination, analysis, comparison and interpretation at various levels. This has created unique opportunities for exploring areas of fundamental importance to the biological sciences and also the basis for understanding a variety of disease states at the molecular level that will be critical for preventive or therapeutic measures. Although the necessity of computing in biological science has long been understood, application of the concepts and tools of information science/computer science in the last couple of decades has brought forth a paradigm shift in the way life science research and teaching programs are implemented. These have enabled new scientific advances through new kinds of experiments. These in turn generate new kinds of data, complex both in volume and in analysis, so much so that today databases containing life science informatics have become an essential part of the infrastructure of life science research. At the same time it continues to

present a huge challenge to accomplish the goal of being able to use, exploit and share these data most effectively. Amidst these, it has also thrown open enormous opportunities for innovations overlapping Bio and IT. Thus, there is vast scope and need to forge an alliance between the two disciplines – life science and IT/computer science – to create a new kind of professional who can address challenges in the overlapping Bio-IT area. Today practically every branch of life-science makes use of some computational approach or the other and these are well integrated into the programmes of advanced study and research. The proposed two year Masters' program in Bioinformatics and Biotechnology is a step to further integrate these two branches for preparing a new generation of young students and researchers who are well trained to address challenges in the evolving post genomic era of biology.



Distribution of credits

I Core = 18 credits

Theory: 12 credits; Lab courses: 6 credits

Theory: 12 credits

- Computer programming 6 credits
- Mathematics for biology (including exercises in Matlab) 2 credits
- Chemistry of macromolecules 2 credits
- Entrepreneurship/IPR/Biotech industry 2 credits

Lab courses: 6 credits

- Computer programming (with application to the Life Sciences) 6 credits

II Bioinformatics = 26 credits:

Theory: 18 credits; Lab courses: 8 credits

Theory: 18 credits

- Statistics for biology 2 credits
- Databases and data mining 2 credits
- Genomics and high throughput data analysis 2 credits
- Computational structural biology 2 credits
- Cheminformatics 2 credit
- Algorithms and data structures 2 credits
- Sequence analysis 2 credit
- Systems biology 2 credits
- Matlab and bioinformatics 2 credits

Lab courses: 8 credits

- Databases and data mining 2 credits
- Cheminformatics and computational structural biology 2 credits
- Algorithms and data structures 2 credits
- Systems biology 2 credits

III Biotechnology = 24 credits

Theory: 16 credits; Lab courses: 8 credits

Theory: 16 credits

- Basic biology 4 credits
- Molecular biology and genetic engineering 2 credits
- Cell biology 2 credits
- Biochemistry 2 credits
- Microbiology and Immunology 2 credits
- Instrumental methods of analysis 2 credits
- Fermentation and downstream processing 2 credits

Lab courses: 8 credits

- Foundations of experimental biology 2 credits
- Mammalian cell culture and protein experiments 2 credits
- Experiments with RNA, DNA and genetic engineering 2 credits
- Fermentation and downstream processing 2 credits

IV Project: 4 credits

TOTAL CREDITS: (I +II +III +IV) = 72

Breakup:

I	Core	18
II	Bioinformatics	26
III	Biotechnology	24
IV	Project	4
TOTAL		72

SUMMARY

There are 4 semesters. Each semester has 18 credits' worth of teaching, including both theory and practicals.

SEMESTER I

Theory

Basic biology	4 credits
Mathematics for biology (including exercises in Matlab)	2 credits
Computer programming (Linux, Shell scripting & C)	2 credits
Computer programming (Perl, SQL/MySQL)	2 credits
Chemistry of macromolecules	2 credits
Biochemistry	2 credits

Lab courses

Foundations of experimental biology	2 credits
Computer programming (with applications to Life science)	2 credits

TOTAL 18 credits

SEMESTER II

Theory

Cell biology	2 credits
Immunology and microbiology	2 credits
Statistics for biology	2 credits
Computer programming (C++, Java)	2 credits
Databases and data mining	2 credits
Genomics and high throughput data analysis	2 credits

Lab courses

Mammalian cell culture and protein experiments	2 credits
Computer Programming (with applications to Life science)	2 credits
Databases and data mining	2 credits

TOTAL 18 credits

SEMESTER III

Theory

Computational structural biology	2 credits
Molecular biology and genetic engineering	2 credits
Algorithms and data structures	2 credits
Cheminformatics	2 credit
Sequence analysis	2 credit

Lab courses

Computer programming (HTML, XML, CGI/PERL)	2 credits
Experiments with RNA, DNA and genetic engineering	2 credits
Algorithms and data structures	2 credits
Cheminformatics and computational structural biology	2 credits

TOTAL **18 credits**

SEMESTER IV

Theory

Systems biology	2 credits
Matlab and bioinformatics	2 credits
Instrumental methods of analysis	2 credits
Fermentation and downstream processing	2 credits
Entrepreneurship, intellectual property and the biotech industry	2 credits

Lab courses

Systems biology	2 credits
Fermentation and downstream processing	2 credits
Project work	4 credits

TOTAL **18 credits**

Credit hours:

Theory:	46 (1 credit = 1hr lecture/week)
Lab courses:	26 (1 credit = 3hrs practicals/week)

SEMESTER SCHEDULE

	Start date	Exam preparation and Exams	Vacation
Sem I, III	1 August	10 Dec to 22 Dec	23 Dec to 21 Jan
Sem II, IV	22 January	15 June to 1 July	2 July to 31 July

FACULTY

Courses	Faculty
Semester I	
Basic biology	Dr. K. Acharya, Dr. R. Grover & Visiting faculty
Molecular biology and genetic engineering - I	Dr. K. Acharya & Dr. R. Grover
Mathematics for biology (including exercises in Matlab)	Visiting Faculty: Dr. B. S. Sudhindra from IITB
Computer programming	Mr. Nagesh Rao (Visiting faculty)
Chemistry of macromolecules	Dr. V. Ramakrishnan & Dr. N. Yathindra
Semester II	
Cell biology	Dr. K. Acharya & Dr. R. Grover
Biochemistry	Dr. N. Yathindra, Dr. V. Ramakrishnan, Dr. M. Banerjee & Visiting faculty from IISc
Statistics for biology	Dr. S. Bhogle & Dr. A. Mathur (Visiting Faculty)
Computer programming	Mr. Nagesh Rao (Visiting faculty)
Databases and data mining	Dr. S. Agrawal
High throughput data analysis	Dr. S. Agrawal & Dr. K. Acharya
Semester III	
Cheminformatics	Dr. V. Ramakrishnan & Dr. B. S. Sudhindra
Computational structural biology	Dr. V. Ramakrishnan & Dr. N. Yathindra
Microbiology and Immunology	Dr. G. Ramananda Rao, Dr. A. A. Karande, Dr. M. S. Shaila & other Visiting Faculty
Algorithms and data structures	Visiting Faculty: IITB
Sequence analysis	Visiting faculty: IISc & NCBS
Semester IV	
Systems biology	Dr. S. Agrawal and Visiting faculty
Molecular biology and genetic engineering - II	Dr. K. Acharya & Dr. R. Grover
Matlab and bioinformatics	Dr. B. S. Sudhindra & Visiting Faculty
Instrumental methods of analysis	Dr. V. Ramakrishnan & Visiting Faculty
Fermentation and downstream processing	Dr. K. Acharya & Visiting faculty
Entrepreneurship/IPR/Biotech Industry	Dr. G. Saberwal

TABLE OF CONTENTS

	Page No.
SEMESTER I	12
SEMESTER II	29
SEMESTER III	41
SEMESTER IV	56

SEMESTER I

Theory

Basic biology	4 credits
Mathematics for biology (including exercises in Matlab)	2 credits
Computer programming (Linux, Shell scripting & C)	2 credits
Computer programming (Perl, SQL/MySQL)	2 credits
Chemistry of macromolecules	2 credits
Biochemistry	2 credits

Lab courses

Foundations of experimental biology	2 credits
Computer programming (with applications to Life science)	2 credits

THEORY

BASIC BIOLOGY	CREDITS: 4
----------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- Identify major kingdoms, a few important phyla and their main characteristics
- List the taxonomical features of organisms that are commonly used as research models and discuss the probable relevance of data from such models to various other species.
- List the major pathogens affecting lives of humans, and important plants and animals of importance to human needs, directly or indirectly; also explain the important taxonomical characters, as well as brief their life cycles
- Relate the diversity to the evolution concepts
- Explain the interactions within populations and the surrounding abiotic/biotic factors
- Understand the relationship between population dynamics and evolution (genetic drift and related concepts)
- Describe the significance of balance of various interactions and recycling of chemical constituents within a system
- Identify major animal and plant physiological systems, their constituents (organs and tissues) and

- the mechanisms of operation
- Discuss a few important examples of common disorders affecting various systems in humans and plants
 - Explain the peculiarities of germ cell production in the context of genetic diversity, and general physiology
 - Explain concepts in genetics and developmental biology
 - Establish the relationship between reproduction, cell divisions, genetics, development and differentiation.
 - Describe the basic principles of inheritance and their discovery; explain genetic linkage and its applications
 - List the major human genetic disorders and explain their characters

Unit I: Taxonomy and biodiversity

Broad basis of classification of major kingdoms: protista, plants, animals; reference to NCBI sources; research models from different major taxa; taxonomy and brief life cycles of examples (major ones only) pests in plants, animals and humans.

Unit II: Ecology

Introduction to ecosystem concepts; brief description of abiotic and biotic factors influencing dynamics of ecosystem; population ecology; urban ecosystems and impact of modern human lifestyles on various ecosystems.

Unit III: Plant and animal physiology

Basics of concepts in cellular vs. general physiology; an overview of plant physiology: plant tissues and functions, water balance, solute transport, translocation (phloem), plant growth, respiration and photosynthesis; human physiology: digestive system, circulation, excretion, muscle structure and functions, and neural tissues and functions; examples of common disorders affecting different systems in humans and plants.

Unit IV: Genetics and evolution

Mendelian laws of inheritance, examples of multiple alleles governing one phenotype; overview of cytogenetics and genetic linkage; brief overview of molecular genetics; types of mutations and their role in

evolution; major human genetic disorders; Darwins theory of evolution, origin of life (current theory), natural selection and concept of speciation, brief history of species, phylogeny, cladistic systematics, molecular basis of evolution

Unit V: Reproduction and developmental biology

Significance of meiosis in sexual reproduction, differences between male and female gamete generation, fertilization, embryogenesis, early development (– plants and humans), role of hormones in germ cell production and early development and puberty; brief discussion on diversity in these processes across other important species, including plants.

Textbooks and Reference Books:

1. **Typologies and taxonomies: An introduction to classification techniques (quantitative applications in the social sciences);** *K.D. Bailey*, Sage Publications, 1994.
2. **Evolution.** *M.W. Strickberger*. Jones and Bartlett publishers Inc., London, 1996
3. **Organizing knowledge: taxonomies, knowledge and organizational effectiveness;** *P. Lambe*; Chandos Publishers, 2007.
4. **Ecology, concepts and applications;** *M.C. Molles*; McGraw-Hill Higher Education, 2nd edition, 2002.
5. **Ecology: from individuals to ecosystems;** *M. Begon, C.R. Townsend and J.L. Harper*; Blackwell Publishing, 4th edition, 2005.
6. **Plant physiology;** *L. Taiz and E. Zeiger*; Palgrave Academic Publisher, 4th edition, 2006.
7. **Principles of anatomy and physiology;** *G.J. Tortora and B.H. Derrickson*; Wiley Higher Education, 11th edition, 2006.
8. **Endocrinology: An Integrated Approach** *Nussey, S.S. and Whitehead, S.A.* London: Taylor & Francis; 2001
9. **Genes IX;** *B. Lewin*; Jones & Bartlett Publishers, Inc., 9th edition, 2008.
10. **An introduction to genetic analysis;** *J.F. Griffiths et.al*; W. H. Freeman & Co, 8th edition, 2000.
11. **Developmental Biology** *Gilbert, Scott F. Sunderland (MA):* Sinauer Associates, Inc.; 2000

MATHEMATICS FOR BIOLOGY (INCLUDING EXERCISES IN MATLAB)	CREDITS: 2
--------------------------------------------------------------------	-------------------

OBJECTIVES:

This course is intended to refresh the basic concepts in mathematics that are frequently applied in different areas of bio-computing and chemical-computing. Students who successfully complete this course will have:

- Studied the concept of functions and derivatives.
- Done basic operations in differential and integral calculus.
- Done calculations with vectors, linear spaces and linear transformations.
- Done binary operations on two vectors in 3D Euclidean space.

Unit I: Basic concepts:

Logarithms, Functions, domain and range. Limits of a function. Continuity and differentiability.

Unit II: Derivatives:

Derivative of a function, Differentiation rules, Derivatives of trigonometric functions. Applications of derivatives

Unit III: Integration:

Indefinite Integrals, Differential equations. Properties, Area and Mean value theorem. Applications.

Unit IV: Linear algebra:

Basic concepts, Matrix addition and multiplication. Determinants, Cramers Rule. Eigen values and Eigen vectors; Applications. Vector spaces, Linear transformations. Dot product and cross product. Vector and scalar functions and fields, derivatives. Curves, tangents, arc lengths.

Textbooks and Reference Books:

1. **Mathematics: from the birth of numbers;** *J. Gullberg*; W.W. Norton & Company, Inc, 1997.
2. **Painless algebra** (Barron's Painless) *L. Long*. Barron's Educational Series. 1998.

3. **Practical algebra: a self-teaching guide**, 2nd edition. *P. H. Selby and S. Slavin*. John Wiley & Sons. 1991.
4. **Calculus made easy**. *S. P. Thompson and M. Gardner*. St. Martin's Press. 1998.
5. **Calculus for dummies**. *M. Ryan*. Wiley Publishing. 2003.
6. **Calculus & analytic geometry**; *G.B. Thomas, R.L. Finney*; Pearson Education Publications, 2001.

COMPUTER PROGRAMMING (Linux Shell scripting and C) (THEORY AND LAB COURSES)	CREDITS: 2
------------------------------------------------------------------------------------	-------------------

OBJECTIVES:

- Appreciate the history, growth, architecture and features of UNIXes, learn basic commands to get simple tasks done on Linux.
- Deal with files in detail, use of text editors and filters to process and edit text.
- Use Master regular expressions and tools that use them and to understand the Linux process model and process related commands.
- Understand the support provided by bash to build scripts, Implement decisions and loop controls within shell scripts.
- Customize the environment and build custom tools, automate tasks and work flows using scripts.
- Realize the scripting model and prepare to learn other scripting languages like Perl.
- Get a start into the programming world and equip oneself with logic building skills that are used in programming.
- Deal with the modular programming paradigm and organize code into functions, write programs that manage memory and data using arrays, strings, pointers and structures.
- Operate on files using file handling routines and gain confidence in programming

Unit I: Introduction & Basic Commands

Introduction to Computers, software, operating systems; History and features of UNIX and GNU/Linux. Basic commands and concepts of GNU/Linux: who, id, tty, logname, date, passwd, cal, bc, uname, man, info.

Unit II: Files & Advanced Files

Files: Concept of files; Types of files; File commands – cat, less, more, I/O redirection and piping, touch, ls, pathnames, pwd, mkdir, rmdir, chdir, wildcards, cp, mv, rm. Advanced Files: Concept of permissions; ls -l, chmod; the ext2 filesystem; concept of hard links and symbolic links, ln; file utilities – file, type, wc, split, cmp.

Unit III: Text Editing with vim & Advanced vim

Text Editing with vim: Introduction to vi; modes of operation; navigation commands; inserting/ appending/ deleting; basic ex-commands. Advanced vim: Buffers, Windows and Files; Clipboards – yanking, deleting, pasting; Keyboard customisation – key maps, abbreviations; Configuration and customisation – resource files.

Unit IV: Filters & Regular Expressions

Role of filters; types of flat file databases; head, tail, cut, paste, sort, uniq. Regular Expressions: Operators; framing regular expressions; grep, egrep and fgrep.

Unit V: Processes

Processes: Concept; the ps command; jobs and background jobs; nohup processes; signals and kill; priority and nice; scheduling processes – at, batch, cron.

Unit VI: I/O Redirection & Environment

Input redirection; output redirection; error redirection; combination of redirections; null redirection; combined redirection; restricted input; merging file descriptors; tee. Environment: Standard variables; accessing, changing and listing variable values; creating, deleting and listing aliases; history; prompt customisation; profile files.

Unit VII: Introduction to Shell Scripting

Significance of quotes; escaping; semicolons and multiple commands; sub-shell execution; simple shell script creation and execution.

Unit VIII: First Steps in Shell Programming

Creating, modifying and deleting variables; output using echo; accepting input from user using read; read only variables; terminal formatting using echo and tput, evaluating arithmetic expressions using expr and \$[].

Unit IX: Decisions & Loops in Shell Programming

The if-fi construct; if-else-fi; if-elif-else-fi; multiple commands as conditions; the \$? variable and exit status; the test command; test operators; equivalence of test and [].Loops: while-done; until-done; for-done; break and continue; infinite loops.

Unit X: Functions & Trapping Signals

Defining functions; calling functions; passing arguments; receiving parameters; local variables; returning values from functions; unsetting functions.
Handling signals; ignoring signals.

Unit XI: Introduction to C programming, Data Types & Operators

Program and programming; types of computers; bases– decimal, binary, octal, hexadecimal; types of languages; importance and features of C. Numeric data types – char, int, short, long, float, double, long double; memory organisation and range; variables and constants. Arithmetic operators; relational operators; Boolean operators; bitwise operators; assignment operators; increment/decrement operators; conditional operator; miscellaneous operators.

Unit XII: Input/Output, Decisions & Loops in C Programming

Formatted I/O using printf() and scanf(); character I/O using getchar() and putchar(). Comments; simple decision making using if and if-else; nested if and else-if ladder; switch-case. Need for loops; simple loop using goto; structured loops – while, do-while, for; break and continue; series generation and pattern generation.

Unit XIII: Functions, Recursion, Macros and pre-processor directives

Role of functions; function definition, call and declaration; arguments and parameters; stack frames. Meaning and need of recursion; solution space; classical examples. The pre-processor; simple macros; macros with arguments; macros v/s functions; #include; conditional compilation.

Unit XIV: Storage classes, Arrays, Strings and Pointers

extern, static, auto and register; linking modules. Need for arrays; defining and using arrays; defining and using

2D arrays; multi-dimensional arrays; array operations. Definition of strings; storage of strings in arrays; string operations using string.h. Concept and use of pointers; declaring and using pointers; operations on pointers; void pointers; NULL pointers; function pointers; pointers and functions – passing by value, passing by pointers; pointers and arrays – array storage and properties, array and pointer conversions; pointers and strings – string operations using pointers; passing arrays to functions.

Unit XV: Command Line Arguments, Structures and Unions

Need; receiving arguments in main; validating arguments; argument type conversions. Need for structures; defining structures; instantiating structures; operating on structures and pointers to structures; structure memory organisation; nested structures; arrays of structures; bit-fields; need for unions; memory organisation of unions.

Unit XVI: Type definition, Dynamic Memory Allocation and File handling

Type definition; enumerated data type. Static memory allocation; need for dynamic memory allocation; malloc() and calloc(); free(); memory problems – memory leaks, dangling pointers, invalid deallocation. Opening files in various modes; closing files; reading and writing characters; reading and writing strings; formatted I/O on files; raw I/O on files; seeking in files.

Textbooks and Reference Books:

1. **Red Hat linux 7.2 bible**; *C. Negus*; Wiley Dreamtech India Pvt Ltd., 2004.
2. **Learning Red Hat linux**; *B. McCarty*; O'Reilly & Associates Inc., 1999.
3. **Computer programming in C**; *V. Rajaraman*; Prentice-Hall of India, 1994.
4. **The C programming language**; *D. Ritchie*; Pearson Education, 2007.
5. **Let us C**; *Y. Kanetkar*; BPB Publications, 1999.

COMPUTER PROGRAMMING (PERL, SQL/MY SQL) (THEORY AND LAB COURSES)	CREDITS: 2
-------------------------------------------------------------------------	-------------------

OBJECTIVES:

- Use Perl for text and file processing, use arrays and array functions to solve a wide variety of problems with ease
- Use the powerful regular expression capabilities of Perl and generate formatted textual reports
- Use hashes to manage data in memory and take advantage of Perl's powerful system interface
- Write programs that solve many common system administration problems
- Download and use modules from the standard Perl distribution and use Perl references to build extra-powerful programs
- Use Perl to execute other programs from within the Perl environment and process their result.
- Comprehend terms and concepts surrounding a relational database, Create databases and tables within them
- Design tables and organize data within them, define and manage constraints on data stored in tables
- Write basic SQL queries and retrieve data from tables selectively based on given requirements
- Insert, update and delete rows of data from/to a table
- Create joins and views to simplify data retrieval

Unit I: Introduction to Perl and working with Scalars

Features of perl; simple perl scripts; I/O in perl using print, printf and <STDIN>; chopping using chop and chomp. Operators and expressions; contexts; the default scalar.

Unit II: Decisions, Loops, Lists and Arrays in Perl Programming

if; unless; && and ||; 'and' and 'or'; modifiers. while; until; do-while; for; foreach; last, next and redo; labelling and goto; labelling and loops. Creating lists; list assignments; creating arrays; array assignments; accessing array elements; the default array; array operations – push, pop, shift, unshift, splice, sort, reverse, chop, delete; split and join; map.

Unit III: Hashes, Strings and Regular Expressions in Perl Programming

Need and concept; lists and hashes; keys and values; iterating through hashes using foreach and each; deleting and element using delete. String functions – length, index

and rindex, substr; case conversion using uc, lc, ucfirst, lcfirst and escape characters. Match; substitute; translate; special operators and escape sequences; extracting matches; back-references; split revisited.

Unit IV: Subroutines, File handling, Files and Directories

Defining subroutines; calling subroutines; argument passing; receiving parameters, local and my variables; returning values; wantarray; declaring subroutines. Obtaining file handles; closing file handles; reading from file handles; writing to file handles; the default file handle; implementing filters; providing input to external commands; collecting output of external commands. Finding the current directory; chdir, mkdir and rmdir; file globbing; file tests; file operations - unlink, chmod.

Unit V: UNIX Features and Report Generation

system(); command substitution using back-quotes; fork. The format definition; picture elements; special variables - \$~, \$%, \$=, \$-; write; select.

Unit VI: References, Packages and Modules

Concept; references to scalars; references to arrays; references to hashes; references to subroutines; anonymous arrays; anonymous hashes; anonymous subroutines. Meaning of packages, namespaces and symbol tables; the package statement; modules; creating modules; using modules.

Unit VII: Introduction to SQL, Tables, Databases and Database design

Introduction to DBMS, RDBMS, SQL. Creating a database; altering a database; creating tables; altering tables; dropping tables. Entity-Relationship diagrams; normalization; primary keys and foreign keys.

Unit VIII: SQL Queries, Datatypes and Expressions

INSERT, UPDATE, DELETE, SELECT; conditions using WHERE, AND & OR, sorting using ORDER BY, grouping using GROUP BY, selecting unique records using DISTINCT. Integers, Character strings, Points, Date; boolean expressions, date expressions, numerical expressions.

Unit IX: SQL Operators and Functions

Mathematical operators, logical operators, max(), min(), sum(), mid(), len(), round(), first(), count(), average().

More about SQL: joins, views, miscellaneous SQL concepts.

Textbooks and Reference Books:

1. **Beginning Perl for bioinformatics**; *J. Tisdall*; Wrox Press Ltd., 2000.
2. **Genomic Perl**; *R.A. Dwyer*; Cambridge University Press, 2003.
3. **MySQL and mSQL**; *Yarger*; O'Reilly Media Publishers, 1999.
4. **Mastering SQL**; *M. Gruber*; BPB Publications, 2000.
5. **CGI/Perl**; *D. Zak*; Thomson Learning, 1st edition, 2002.

CHEMISTRY OF MACROMOLECULES	CREDITS: 3
-----------------------------	------------

OBJECTIVES:

This course is intended to brush up chemistry fundamentals which are frequently applied in key areas of computational and molecular biology. Students who successfully complete this course will have:

- A good understanding of atomic structure and chemical bonding, in particular the importance of different types of bonds in macromolecules and macromolecular assemblies.
- A good understanding of fundamental concepts of stereochemistry and their particular relevance to biological phenomena.
- A good understanding of the fundamental principles governing the structure of the 4 macromolecules, viz., proteins, nucleic acids, lipids and carbohydrates.
- A good understanding of the basics of thermodynamics and how it influences macromolecular structure and function.

Unit I: Atomic structure and chemical bonding

Chemical bonding, atomic and molecular structure, atomic and van der waal's radii of atoms, Valence bond and Molecular orbital theory; hybridization; Symmetry elements and operations. Molecular interactions: van der Waals, electrostatic, different types of hydrogen bonding and hydrogen bond like interactions, hydrophobic interaction, Structure of water; solvent solute interaction

Unit II: Elements of Stereochemistry

Stereoisomerism, asymmetric centers, different kind of stereoisomers, geometrical and optical isomers, enantiomers, meso forms and racemic mixture and optical rotation, Klyne and Prelog priority rules, Disymmetry, Bond rotational potentials and energy barrier; configuration vs. conformation. Rotational isomers and concept of hindered rotation; anti,syn,gauche conformations; nonbonded atom pairs.

Unit III: Basics of thermodynamics

Concepts of information and entropy, information as negative entropy, entropy change, enthalpy, free energy and free energy change.

Unit IV: Chemistry of proteins

Structure properties and classification of amino acids; single and three letter symbols; pH titration of amino acids; L and D configuration of amino acids; modified amino acids; asymmetric centers beyond alpha carbon; Resonance structure of amide bond and planar peptide unit; peptide geometry, dimensions; polar character of peptide and dipole moment of peptide; peptide virtual bond-amino acid residues; proteins as polypeptides-disulphide bonds; energetics of *cis* and *trans* peptide, special properties of prolyl residues, , conformation of dipeptides and Ramachandran map.

Unit V: Chemistry of nucleic acids

Chemistry of Nucleic acids structure, nucleic acids as polyelectrolytes, structure of nucleic acids bases, tautomerism, neutral and charged forms of bases; Chemical structure nucleosides, nucleosides, nucleoside diphosphates and nucleoside triphosphates, dinucleoside monophosphates and nucleoside triphosphates; modified nucleosides; variety of base pairing interactions; base triples and base quartets; Unique features of Watson & Crick pairs; their implication to Watson & Crick duplex and foundation to molecular biology; sugar ring conformations; *anti* and *syn* conformations in nucleosides and nucleotides; backbone conformations; Preferred nucleotide and internucleotide conformations.

Unit VI: Chemistry of carbohydrates

Open and closed form of Furanoses, pentapyranoses and hexapyranoses; stereoisomerism in penta and hexapyranose sugars; anomeric carbon and

characterisation of alpha and beta sugars; different types of linkages; Hexapyranose derivatives; different types of disaccharides; lactose, maltose, cellobiose, laminaribiose etc.; chair and boat conformations for sugars; axial and equatorial glycosidic linkages; hydrogen bonding properties of sugars; description of disaccharide conformation; polysaccharides; bacterial cell wall polysaccharides; proteoglycans and glycoproteins.

Unit VI: Chemistry of lipids

Chemical structure of different lipids, fatty acids and steroids, structure of lipids sphingolipids, glycolipids, organization of lipids in membrane structure.

Textbooks and Reference Books:

1. **Lehninger - Principles of biochemistry**, Fifth Edition, D.L. Nelson and M.M. Cox, Freeman & Co. NY, 2008
2. **Physical chemistry**; *A. G. Whittaker*; Viva Books Pvt. Ltd., 2002.
3. **Chemistry for biologists**; *J. Fisher*; Viva Books Pvt. Ltd., 2002.
4. **Organic chemistry**; *R.T. Morrison and R.N. Boyd*; Pearson Education, 6th edition, 2005.
5. **Chemistry: Molecules, matter and change**; *L. Jones and P. Atkins*; W.H. Freeman and Company, 4th edition, 2000.
6. **Biophysical chemistry, I,II & III**, *C. Cantor and P.Schimmel*, W. H. Freeman and Co., San Francisco, 1980
7. **Principles of Biochemistry**, *D.Voet, J.G.Voet and C.W.Pratt*, John Wiley, 2007
8. **Introduction to Protein Structure**, *C. Branden and J. Tooze*, Garland Publishing, NY, 1991.
9. **Principles of Nucleic Acid Structure**, *W. Saenger*, Springer-Verlag, NY, 1984.
- 1.10. **Proteins: Structure and Molecular Properties**, *T.E.Creighton*, W.H. Freeman and Co. San Francisco, 1992.

BIOCHEMISTRY	CREDITS: 2
---------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- describe the chemical nature and structure of amino acids, nucleotides, proteins, nucleic acids, lipids and carbohydrates
- explain the basic principles of biochemical reactions and related energy dynamics
- describe major metabolic pathways, and their role in cell homeostasis and survival
- describe the nature of enzymes and principles of enzyme kinetics
- provide examples of structure-function relations between major enzymes

Unit I: Basics of molecular structure of proteins

Principles of protein structure; anatomy of proteins – Hierarchical organization of protein structure; primary, secondary, supersecondary, tertiary and quaternary structure; internal coordinates of proteins; the Ramachandran Map, protein energetics; protein folding.

Unit II: Basics of molecular structure of DNA and RNA

Helical structures of DNA and RNA; types of base pairing – Watson-Crick and Hoogsteen; A, B, Z types of double helices; RNA secondary structure motifs, tRNA structure; various types of interactions of DNA with proteins and small molecules.

Unit III: Basics of molecular structure of carbohydrates

Structural diversity of polysaccharides; glycoconjugates: various types (glycolipids, glycoproteins etc); their structures and interactions with various biomolecules, structure of bacterial polysaccharides.

Unit IV: Major metabolic pathways

Basic principles of biochemical reactions: relationship between cell metabolism, biological order, heat energy and oxidation of organic molecules (glycolysis as a model); fermentation reactions; citric acid cycle and electron transport system.

Unit V: Enzymology

Principles of catalysis, enzymes and enzyme kinetics, enzyme regulation, mechanism of enzyme catalysis, isozymes; enzyme immobilization.

Textbooks and Reference Books:

1. **Enzymes: Biochemistry, biotechnology, clinical chemistry**; *T. Palmer*; Affiliated East-West Press Private Limited, 2004.
2. **Biochemistry**; *L. Stryer*; W H Freeman & Co, 5th edition, 2002.
3. **Lehninger principles of biochemistry**; *D.L. Nelson, M. M. Cox*; W.H Freeman & Company, 5th edition, 2009.
4. **Biochemistry**; *D. Voet, J.G. Voet*; Wiley, 3rd edition, 2008.

LAB COURSES

FOUNDATIONS OF EXPERIMENTAL BIOLOGY	CREDITS: 2
--------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- list the factors that can influence the accuracy and precision of experimental results in general
- state the specific range of values for accuracy and precision of results of chosen experiments
- list the potential laboratory hazards and demonstrate 'safe' practices for carrying out any identified general laboratory procedures
- perform basic calculations with ease and accuracy, and prepare any type of stock or working solutions with ease
- systematically design an experiment for at least a few given simple experimental problems
- create a detailed flow chart by referring to a brief protocol, identify the equipment required (along with duration of use) and list the necessary reagents and consumables (along with the quantities) required.
- list potential suppliers of major suppliers of major laboratory consumables, reagents and equipment;
- state the procedure of obtaining quotations, comparing technical and commercial aspects, and make a decision to purchase a required item
- explain the general procedure and principles of management of major lab-equipment and reagent stocks
- follow good record keeping habits, including preparedness for making detailed observations during experiments

- perform quantitative estimations and electrophoretic analysis of DNA, RNA and proteins

Unit I: Basics of laboratory work and safety

Safe laboratory practices (handling reagents and instrumentation; preparing to handle biohazards, and hazards by radioactivity, electricity and fire), calculations, designing experiments, purchasing procedures, potential suppliers of lab-requirements, stock maintenance (sensitive reagents, and routine chemicals and consumables) and record keeping.

Unit II: Precision and accuracy

Study of factors influencing precision and accuracy of results in general.

Unit III: Experiment designing

Designing experiments to find answers to specific questions; exercises in relating scientific questions to appropriate techniques, sample-details, selection of controls, statistical aspects, etc.

Unit IV: Microscopy

Basics of microscopy and types of microscopes, histology: major plant animal tissues, cell counting, chromosomal preparations.

Unit V: Concepts in cleanliness, sterility vs. molecular stability

Basic procedures for clean vs. sterile vs. RNase/DNase free consumables and bench space; solution preparations for routine purposes vs. cell culture and molecular biology work; reagent types for different purposes (study of catalogs); significance of temperature control and swiftness in experiments: experiments with molecular stability at different temperature

Textbooks and Reference Books:

1. **Laboratory reference**; *J. Roskams and L. Rodgers*; I K International Pvt Ltd, 2004.
2. **Lab math**; *D. Adams and L. Spencer*; I K International Pvt Ltd, 2004.

COMPUTER PROGRAMMING	CREDITS: 2
-----------------------------	-------------------

EXERCISES WITH APLPLICATIONS TO LIFE SCIENCE	
---------------------------------------------------------	--

The contents will be similar to the corresponding theory section described above.

SEMESTER II

Theory

Cell biology	2 credits
Immunology and microbiology	2 credits
Statistics for biology	2 credits
Programming (C++, Java)	2 credits
Databases and data mining	2 credits
Genomics and high throughput data analysis	2 credits

Lab courses

Mammalian cell culture and protein experiments	2 credits
Programming (with applications to life science)	2 credits
Databases and data mining	2 credits

THEORY

CELL BIOLOGY	CREDITS: 2
--------------	------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- schematically represent the eukaryotic cell organelles, list their major chemical constituents and state their functions
- provide examples to illustrate the diversity in number and distribution of cell organelle across cell-types
- give examples to describe the role of differentiation in generating various cell-types in higher organisms
- identify and explain all major differences between prokaryotic and eukaryotic cells, as well as animal and plant cells
- explain the mechanisms of cell division during growth and reproduction
- elucidate the link between meiotic processes and genetic diversity
- identify and describe various types of interactions with other cells and the extra-cellular matrix, in the context of at least a few specific functions
- explain how extra-cellular signals are transduced into intracellular signals and functions
- summarize the characters of a cancerous cell and current trends in cancer research

Unit I: Basic structure and function of cells and their organelles

Differences in the basic structure and composition of prokaryotic cells and eukaryotic animal and plant cells; structure and function of eukaryotic (plant and animal) cell organelles.

Unit II: Diversity in structure and function across tissues/species

Diversity in the size and shape of cells depending on functions within different tissues; variations in the number and structure of organelles depending on the type of cells (e.g., rich smooth endoplasmic reticulum in lipid secreting cells).

Unit III: Cell signaling

Principles of signal transduction; Classification of signaling mechanisms; hormones and their receptors classes; signaling through G-protein coupled receptors and their effectors; Tyrosine Kinases and Ras, MAP Kinase Pathways; second messengers; steroid hormone receptors, interaction and regulation of signaling pathways; bacterial and plant two-component signaling systems; bacterial chemotaxis and quorum sensing.

Unit IV: Cell-cell interactions

Cellular communication: General principles of cell communication; cell adhesion and roles of different adhesion molecules; gap junctions, extracellular matrix, integrins; neurotransmission and its regulation.

Host-parasite interaction: Recognition and entry processes of different pathogens like bacteria and viruses into animal and plant host cells; alteration of host cell behavior by pathogens, virus-induced cell transformation, pathogen-induced diseases in animals and plants; cell-cell fusion in both normal and abnormal cells.

Unit V: Cell cycle and divisions

Introduction to cell cycle, cell division and types: mitosis and meiosis; typical phases of mitosis and meiosis; brief discussions on the molecular mechanisms, relationship with growth, differentiation, reproduction and cancer; cancer cell biology

Textbooks and Reference Books:

1. **Molecular cell biology**; *H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell*; W.H Freeman & Comp., 6th edition, 2007.
2. **Molecular biology of the cell**; *B. Alberts et. al.*; Taylor & Francis Publishers, 5th edition, 2008.
3. **The cell: A molecular approach**; *G. M. Cooper and R. E. Hausman*; ASM Press, 5th edition, 2009.

MICROBIOLOGY AND IMMUNOLOGY	CREDITS: 2
------------------------------------	-------------------

Objectives

Students who successfully complete this course will be able to

- Explain the types and constituents of the human immune system, their specific functions and mechanism of actions
- List the major pathogens affecting lives of important species, directly or indirectly, and explain the important features, which make them successful parasites
- Describe the cell and molecular details of living and propagation of microbial organisms

Unit I: Immunology - part I

Innate and adaptive immune system: Cells and molecules involved in innate and adaptive immunity, antigens, antigenicity and immunogenicity. humoral and cell-mediated immune responses,

Unit II –Immunology – part II

B and T cell epitopes, structure and function of antibody molecules, generation of antibody diversity, monoclonal antibodies, antibody engineering, antigen-antibody interactions, MHC molecules, antigen processing and presentation, activation and differentiation of B and T cells, B and T cell receptors, primary and secondary immune modulation, the complement system, Toll-like receptors, cell-mediated effector functions, inflammation,

Unit III – Immunology – part III

Hypersensitivity and autoimmunity, immune response during bacterial (tuberculosis), parasitic (malaria) and viral (HIV) infections, congenital and acquired

immunodeficiencies, vaccines.

Unit IV: General microbiology

Broad taxonomy of microbes; chemotaxonomy (cell wall components, isoprenoid - quinones, amino acid sequence of proteins, protein profiles, cytochrome composition, ribosomal RNA, etc); genetic methods in taxonomy (PCR and DNA fingerprinting as identification tools for bacteria, DNA base composition and hybridisation, rDNA sequencing, etc); Archaeobacteria: taxonomic position (relatedness to eucaryotes and procaryotes, unique molecular and biochemical features);

Unit V: Pathogenic microbes

Overview of virology; prokaryotic genomes; characteristics of major pathogens and related epidemiology in humans, economically important animals and plants.

1. **Kuby immunology**; *R.A. Goldsby, T.J. Kindt and B.A. Osborne*; W. H. Freeman & Co, 6th edition, 2000.
2. **Immunology: understanding the immune system**; *K.D. Elgert*; Wiley-Blackwell Publication, 2nd edition, 2009.
3. **Janeway's immunobiology**; *K. Murphy, P.Travers and M. Walport*; Taylor & Francis Publishers, 7th edition, 2008.
4. **Brock biology of microorganisms**; *M.T. Madigan, J.M. Martinko, P.V.Dunlap and D.P. Clark*; Pearson Education, 12th edition, 2009.
5. **Prescott/Harley/Klein's microbiology**; *J. Willey, L. Sherwood and C.Woolverton*; McGraw-Hill Higher Education, 7th edition, 2008.

STATISTICS FOR BIOLOGY	CREDITS: 2
------------------------	------------

OBJECTIVES:

This course is meant to educate students on the basic principles of statistics. Students who successfully complete this course will be able to:

- apply basic principles of statistics to molecular biology and structural biology problems.
- understand the theoretical background in the usage of statistical packages like SAS and SPSS.

- have a strong foundation to pursue a career in the area of biostatistics if they wish.

Unit I: Data Analysis

Data Representation, Average, Spread, Probability, Permutations and Combinations. Frequency distributions; Normal, Binomial, Poisson, Central Limit Theorem, Bayesian Statistics.

Unit II: Basic Mathematical Statistics

Random sampling, Estimation of parameters, Confidence intervals, Testing of Hypothesis, decisions, Quality control, Goodness of fit, Significance tests, Correlation and Multiple regression analysis, Hypothesis testing.

Unit III: Advanced Mathematical Statistics

ANOVA and Markov Chains, Hidden Markov models.

Unit IV: Data reduction methods

Multivariate and Principal Component Analysis.

Textbooks and Reference Books:

1. **Principles of biostatistics;** *M. Pagano*; Cengage Learning Publishers, 2nd edition, 2008.
2. **Schaum's outline of probability and statistics;** Tata McGraw Hill, 2004.
3. **Basic biostatistics for geneticists & epidemiologists;** *R.C. Elston, W.D. Johnson*; Wiley, 2008.
4. **Data reduction and error analysis for the physical sciences.** P. R. Bevington McGraw Hill. 1969.

<p>COMPUTER PROGRAMMING EXERCISES (C++ & JAVA) WITH APPLICATIONS TO LIFE SCIENCES (THEORY AND PRACTICALS)</p>	<p>CREDITS: THEORY 2, PRACTICALS 2</p>
--------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------

OBJECTIVES:

- Understand Object-Oriented Programming concepts and techniques
- Convert an Object-Oriented design into a C++ program and separate the Interface from Implementation and improve usage of code
- Create custom Abstract Data Types

- Upgrade, enhance and reuse already existing applications without modifying code using Inheritance
- Develop platform-independent code using Java
- Understand how Object Oriented Programming can be implemented in Java
- Develop multithreaded applications to improve efficiency
- Use ready-made packages of Java to write programs faster
- Build graphical user interfaces using the Swing package
- Deploy applets and add more power to websites
- Use streams to communicate and connect to databases and write database-driven Java applications

Unit I: Introduction to OOP and C++

Introduction to Object Oriented Programming: The OOP programming model; enhancements over structured programming. C++ as a better C: Streams; manipulators; comments; variable definitions; type casts; scope resolution; improved syntax for structures, unions, enums; anonymous unions and enums; void pointers; constants and pointers; references; constant references.

Unit II: Procedural and Object-Based Programming in C++

Prototype; inline functions; default arguments; function overloading; function overload resolution.

Classes; objects; class definition; memory organisation of objects; member functions; constructors; destructors; the this pointer; static data members; static member functions; constant member functions; friend functions; friend classes.

Unit III: C++ Free Store Management and Inheritance

new and delete; benefits over malloc() and free(); free store exhaustion.

Class relationships; structural types of inheritance; private, public and protected inheritance; properties of inheritance; function overriding; memory organisation of objects; constructors and destructors; the is-a relationship; class containership and the has-a relationship; late binding and virtual functions; pure virtual functions and abstract classes; virtual inheritance and virtual base classes.

Unit IV: Introduction to Java, Java Quick Start and OOP in Java

History of Java, Features of Java, JVM, JRE and JDK.
Java Quick Start: Learning the basics of Java in comparison to C++. Object Oriented Programming in Java: Classes, interfaces and packages; access modifiers; constructors; the this and super references; inner classes and nested classes; anonymous classes.

Unit V: Exception Handling, Garbage Collection and Multithreading

Concept of exception handling; throwing and catching exceptions; handling and declaring exceptions; re-throwing of exceptions. Concept of garbage collection; overriding the finalize() method. Concept of multithreading; services provided by Thread class; implementing multithreading by extending Thread class; implementing multithreading by implementing Runnable interface; sleeping and interrupting threads; synchronized methods and statements; inter-thread communication.

Unit VI: The java.lang and java.util Packages

Object class; Class class and the Reflection API; System and Runtime classes; Properties class; Wrapper classes; overview of other classes. Vector, ArrayList, LinkedList, HashSet, LinkedHashSet, TreeSet, HashMap, LinkedHashMap, TreeMap; Iterator and Enumeration; Collections and algorithms.

Unit VII: The javax.swing Package

GUI development using the javax.swing Package: Class hierarchy; working with containers and layout managers; working with components; event handling; menu system; drawing and painting.

Unit VIII: Applets

Applications of applet; The applet lifecycle; loading applets.

Unit IX: The java.io Package and Database access using the java.sql Package

InputStream, OutputStream, Reader, Writer, and their sub-classes; File class; RandomAccessFile class, StreamTokenizer. Connecting to databases using JDBC; creating and executing statements; working with result sets.

Textbooks and Reference Books:

- 1. The complete reference C++**, *H. Schildt*, Tata Mcgraw Hill, 1998.
- 2. The complete reference Java 2**, *H. Schildt*, 5th edition, Tata Mcgraw Hill, 2002.

DATABASES AND DATA MINING	CREDITS: 2
----------------------------------	-------------------

OBJECTIVES:

This course will train the students to develop biological databases related to direct academic/industrial applications. It will facilitate knowledge on the following:

- Building relational biological databases
- Develop and write the code for mining large datasets and identify the important features
- The students would be able to design small biological databases and able to connect the small databases to develop a databank.

Unit I:

Design biological databases: relational data models, semi structured data models.

Unit II:

Text databases sequence databases, pattern databases, Conceptual database modeling, Database management system architectures, Database Normalization, The Structured Query Language.

Unit III:

Database design principles: brief mention of normalization theory query paradigms, query languages (SQL, pSQL and mySQL) similarity searches, approx.retrieval integrated data analysis & applications VIEWS across databases, pathway modeling, network queries

Unit IV:

Data mining: general paradigms: clustering, classification, biclustering, structure discovery in biosequences, rule finding, probabilistic graphical models, signaling pathways

Textbooks and Reference Books:

1. **Data structures and algorithms**; *A.V. Aho*; Pearson Education, 2002.
2. **Database system concepts**; *A. Silberschatz*; McGraw-Hill, 4th edition, 2002.
3. **Fundamentals of database systems**; *R. Elmasari, S.B. Navathe*; 5th edition, 2006.

GENOMICS AND HIGH THROUGHPUT DATA ANALYSIS	CREDITS: 2
---------------------------------------------------	-------------------

OBJECTIVES:

This course is intended to introduce the high throughput molecular data, its applications, how it is generated and its analysis. The focus will be different types of high density array data like expression, exons, SNP arrays, DNA microarray, ChIP on ChIP data, protein mass spectrometry data etc. Students who successfully complete this course will be able to:

- comprehensively understand the generation and usage of high throughput data in drug designing and biomarker discovery.
- choose and design the high throughput data study and analysis.
- apply different statistical methods and algorithms to analyze different types of data.
- interpret the analysis output.

Unit 1: Gene expression profiles

Transcriptomics data processing and analysis: microarray and other expression profiling data and scoring differential expression.

Unit II: Genome wide DNA-protein interactions

High-throughput TF-DNA interaction measurements and modeling the sequence-specificity of TFs and analysis

Unit III: Genomics

Genome projects, genome structure, next generation sequencing; Metagenomics, population genomics, comparative genomics of non-coding sequence using high throughput methods; Analysis of the data to discover cis-regulatory elements in 3' UTRs.

Unit IV: Transcriptomics

Post transcriptional world, the transcriptome, the RNA

world, mapping and analysis of non-coding RNA, micro RNAs, piRNAs etc.

Unit V: Proteomics

Proteome profiling methods, 2-D electrophoresis image comparisons, identifying signal transduction pathways using protein-protein interaction; yeast two-hybrid, mass spectrometry data processing and analysis

Textbooks and Reference Books:

1. **An introduction to data structures with applications**; *J.P. Tremblay and J.Paul*; T. McGraw Hill, 2nd edition, 1984.
2. **Experimental design and data analysis for biologists**; *G.P. Quinn and M.J. Keough*; Cambridge University Press, 2002.
3. **Guide to analysis of DNA microarray data**; *S. Knudsen*; Wiley, 2nd edition, 2004.
4. **Mass spectrometry: Principles and applications**; *E.D. Hoffmann & V. Stroobant*; Wiley, 3rd edition, 2007.
5. **Advanced analysis of gene expression microarray data**; *A. Zhang*; World Scientific Publishing, 2006.

LAB COURSES

MAMMALIAN CELL CULTURE AND PROTEIN EXPERIMENTS	CREDITS: 2
-------------------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- perform protein extractions from bacteria and tissues (solid and liquid, plant and animal), immuno-precipitation and dialysis
- analyze at least a few samples using 2-dimensional electrophoresis and multiple staining procedures
- purify at least a few proteins by chromatography and salt precipitation methods
- identify given protein(s) by western blotting method and ELISA; use ELISA to quantify proteins as well
- perform enzyme kinetics studies

- seed, culture and passage mammalian cells in flasks, without contamination
- maintain cell lines using correct culturing and storage methods
- prepare culture media and sterilize, and count the cells using microscopes
- assess the quality of cells in culture at different stages
- perform staining of the cells and assess their quality
- transfect mammalian cells using more than one procedure

Unit I: Basic cell culture procedures

Concepts of cell lines, transformation, sterility and culturing mammalian cells; practicing basic procedures: seeding, passaging, sterile media preparations etc.

Unit II: Cell culture

Passaging through multiple lines of generations using at least 2 different types of mammalian cells

Unit III: Transfection

Transfecting with lipofectamine and other methods using a standard gene and mammalian cells.

Unit IV: Basic protein handling

Basic protein extraction methods, electrophoretic and spectrophotometric quantitative and qualitative analysis methods, basic methods of purification of proteins by salt precipitation.

Unit V: Immunological techniques

Western blotting analysis, quantitative analysis of specific antigens using ELISA, immunoprecipitations, flow cytometry.

Units VI: Enzyme studies

Basic experiments in enzyme assays and kinetics; study of factors influencing enzyme kinetics.

Textbooks and Reference Books:

1. **Purifying proteins for proteomics: A laboratory manual**; *R. J. Simpson*; Cold Spring Harbor Laboratory Press, 2004.
2. **Human molecular biology laboratory manual**; *S. Surzycki*; Blackwell Publications, 2003.

3. **Enzymes: Biochemistry, biotechnology, clinical chemistry**; *T. Palmer*; Affiliated East-West Press Private Limited, 2004.
4. **Culture of animal cells: A manual of basic technique**, *R. I. Freshney*, Wiley-Liss, 2000
5. **The condensed protocols from the molecular cloning: A lab manual** – *J. Sambrook*; Cold Spring Harbour Lab. Press, 2006.

DATABASES AND DATA MINING	CREDITS: 2
----------------------------------	-------------------

OBJECTIVES:

This course will train the students to develop biological databases related to direct academic/industrial applications. It will facilitate knowledge on the following:

- Building biological relational databases
- Develop and write code for mining large datasets and identify important features
- Design small biological databases and connect the small databases to develop a databank.

Developing data architecture using MySQL and oracle. Principles and developing modules on biological databases using relational system, developing genome annotation datasets, parsing, developing tables and relational systems.

Textbooks and Reference Books:

1. **A first course in database systems**; *J. D. Ullman and J. Widom*; Pearson Education, 3rd edition, 2008.
2. **Database system concepts**; *A. Silberschatz, H.F. Korth and S. Sudarshan*; McGraw-Hill, 4th edition, 2002.

PROGRAMMING EXERCISES WITH APPLICATIONS TO LIFE SCIENCE	CREDITS: 2
--------------------------------------------------------------------	-------------------

The contents will be similar to the corresponding theory sessions mentioned above.

SEMESTER III

Theory

Computational structural biology	2 credits
Molecular biology and genetic engineering	2 credits
Algorithms and data structures	2 credits
Cheminformatics	2 credits
Sequence analysis	2 credits

Lab courses

Computer Programming (HTML, XML, CGI/PERL)	2 credits
Experiments with RNA, DNA and genetic engineering	2 credits
Algorithms and data structures	2 credits
Cheminformatics and computational structural biology	2 credits

THEORY

COMPUTATIONAL STRUCTURAL BIOLOGY	CREDITS: 2
----------------------------------	------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- understand the basics of molecular modelling and simulations.
- study the concept of force fields and its individual components.
- study the concept of energy minimisation and dynamics.
- modify molecular parameter files to meet prescribed objectives.
- design an MD protocol.
- simulate small bio-molecules using standard MC and MD protocols.

Unit I: Empirical force field models

Molecular mechanics force fields, bond stretching, bond angle bending, torsional terms, improper torsions and out-of-plane bending motions. Nonbonded interactions: Electrostatic interactions, pointcharge electrostatic models, partial atomic charges, solvent dielectric models,

van der Waals interactions: Dispersive interactions and repulsive contribution. Many body effects in empirical potentials: Effective pair potentials, hydrogen bonding in molecular mechanics, force field models for the simulation of liquid water, simple water models and polarisable water models. United atom force fields and reduced representations. Energy minimization: Derivative and nonderivative minimization: simplex, steepest descent and conjugate gradient methods; second derivative methods: the Newton-Raphson method.

Unit II: Computer simulation methods

Historical background, time averages and ensemble averages. Calculation of simple thermodynamic properties: energy, heat capacity, pressure, temperature, radial distribution functions, phase space.

Unit III: Molecular dynamics

MD algorithm, fundamental issues – time steps, MD protocols, analysis of MD results, using MD for Conformational sampling, limitations and perspectives, applications.

Unit IV: Monte Carlo simulation methods

Theoretical background, implementation of Metropolis Monte Carlo Method, configurationally biased Monte Carlo, Monte Carlo Vs MD, applications.

Textbooks and Reference Books:

1. **Structural bioinformatics.** *P. E. Bourne and J. Gu.* 2nd Edition. John Wiley and Sons. 2009.
2. **Structural bioinformatics: An algorithmic approach.** *F. J. Burkowski.* CRC Press. 2008
3. **Computational molecular biology: an introduction** *P. Clote and R. Backofen,* Wiley & Sons, 2000.
4. **Bioinformatics vol:2: Structure, function & applications** *J. M. Keith* Humana Press. 2008
5. **Compact handbook of computational biology.** *A. K. Konopka and M. J. C. Crabb,* Marcel Dekker, New York, 2004

MOLECULAR BIOLOGY AND GENETIC ENGINEERING	CREDITS: 2
------------------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- explain the central dogma of molecular biology and its significance to life in general
- explain in detail, the following processes and their mechanisms: replication, transcription and translation
- list the differences in the above-mentioned processes, between prokaryotes and eukaryotes
- list and explain the mechanism of the post-transcriptional and post-translational modifications, in eukaryotes
- explain DNA repair, splicing mechanisms, and molecular transport within cells
- describe the mechanism of regulation of gene-expression and the role of small RNA molecules in cellular functions
- list the steps involved in recombinant DNA technology and genetic engineering; explain the principles of the same
- list the major enzymes required for in vitro handling of DNA and RNA; explain the differences of such enzymes as compared to the enzymes in vivo
- explain the principles of types of vectors used for general cloning and expressing proteins
- state and discuss the ethical dilemma associated with genetic engineering and associated applications
- design primers with 5' addition of restriction sites or mutations
- prepare competent bacterial cells
- perform restriction digestion and ligation of given DNA and vectors, and transformation of bacteria
- select the transformed colonies and further clone the target gene to express the proteins
- create cDNA libraries and screen expression libraries

Unit I: DNA replication, repair and recombination

General features of chromosomal replication, DNA replication machinery, fidelity of replication, extra chromosomal replicons. DNA damage and repair

mechanisms and their role in carcinogenesis. Recombination between homologous DNA sites.

Unit II: RNA synthesis and processing

Transcription factors and machinery, RNA polymerases, formation of initiation complex, transcription activators and repressors, molecular mechanisms of transcription control. Transcription termination, processing of eukaryotic mRNA, regulation of mRNA processing, signal-mediated transport through nuclear pore complexes. Processing of rRNA and tRNA and other mechanisms of post-transcriptional control.

Unit III: Protein synthesis and processing

Ribosome, formation of initiation complex, initiation factors and their regulation, elongation and elongation factors, termination, genetic code, amino-acylation of tRNA, tRNA identity, aminoacyl tRNA synthetase, translational proofreading and translation inhibitors. Post-translational modifications of proteins, protein sorting to organelles and secretion.

Unit IV: Regulation of gene expression

Regulation of gene expression in prokaryotes and eukaryotes, role of chromatin structure and DNA methylations in regulating gene expression and gene silencing.

Unit V: Concepts in recombinant DNA technology

Basic principles; introduction to types of vectors used in recombinant DNA technology, their specific uses and comparison of their features; an overview of various enzymes used in recombinant DNA technology.

Unit VI: Cloning and subcloning strategies

Preparation of competent cell-transformation, transfection; recombinant selection and screening; genomic DNA library; cDNA synthesis strategies – linkers, adapters; making genomic and cDNA libraries in plasmids and phages. PCR product cloning (TA cloning), cloning strategies in yeast, *E. coli* and *B. subtilis*.

Unit VII: Expression vectors

Examples of expression vectors and case studies in recombinant proteins of commercial success.

Unit VIII: Genetic engineering

Mutagenesis: deletion mutagenesis, oligonucleotide derived mutagenesis, site directed mutagenesis; Applications of rDNA technology in diagnostics; pathogenesis; genetic diversity; therapeutic proteins-vaccines, molecular probes (production, labeling and uses), PCR.

Text and Reference Books:

1. **Molecular cell biology**; H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell; W.H Freeman & Comp., 6th ed., 2007.
2. **Molecular biology of the cell**; B. Alberts et. al.; Taylor & Francis Publishers, 5th ed., 2008.
3. **The cell: a molecular approach**; G.M. Cooper and R.E. Hausman; ASM Press, 5th ed., 2009.
4. **Genes IX**; B. Lewin; Oxford University Press, 2008.
5. **An introduction to genetic analysis**; A.J.F. Griffiths, W.H.Freeman & Co., 2008.
6. **Recombinant DNA**; J.D. Watson; Scientific American Books, 1992.
7. **Molecular biology of the gene**; J.D. Watson et.al.; Pearson Education, 6th edition, 2008.
8. **Principles of gene manipulation and genomics**; S.B. Primrose and R.M. Twyman; Blackwell Publishing, 7th edition, 2006.
9. **Improving nature: The science and ethics of genetic engineering**; M. J. Reiss, R. Straughan, Cambridge University Press, 2002.

ALGORITHMS AND DATA STRUCTURES (THEORY AND PRACTICALS)	CREDITS: 2
-------------------------------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will develop the skills and understanding of the following:

- Basic terminology and mathematical terms of computer algorithms with some of the standard algorithms as examples
- Developments and understanding of basic representation of algorithms as flow charts and understanding of notion of complexity analysis of algorithms

- Various data structures in general which are used in the development and implementation of algorithms in various computer languages such as 'C', 'C++' and 'JAVA' etc. but not specific to any such programming language
- Familiarization and implementation (practical sessions) of various standard algorithms for biological applications with learning on complexity analysis of recursive and non-recursive algorithms
- Introduction to the more advance concepts such as parallel algorithms and dynamic programming

Unit I: Basic concepts of algorithms

Introduction; Notion of an algorithm (what is an algorithm), Notation for Programs (Flow charts, terminology like repeat, begin end etc); Basic Mathematical Notations (like integers, real numbers, basics of set theory, functions and relations); Fundamentals of Algorithmic Solving; Important problem types in general.

Unit II: Elementary algorithmics

Introduction, Fundamentals of the Analysis Framework; Asymptotic Notations, Efficiency of Algorithms, Basic idea of computational complexity, Average and Worst case analysis, Elementary operations, Basic Examples; Determinant calculations, Sorting, Multiplication of large integer numbers, Calculating the Fabonacci Sequence etc.

Unit III: Basic data structure and control structures

Arrays, stack, queues, pointers, Lists, Graphs, Trees, heaps and other elementary data structures, Analyzing control structures like For, while, repeat etc.

Unit IV: Mathematical aspects and analysis of algorithms

Mathematical Analysis of Non-recursive Algorithm; Mathematical Analysis of Recursive Algorithm; Example: Fibonacci Numbers; Empirical Analysis of Algorithms; Algorithm Visualization.

Unit V: Analysis of sorting and searching algorithms

Brute Force; Selection Sort and Bubble Sort; Sequential Search and Brute-force string matching; Divide and conquer; Merge sort; Quick Sort; Binary Search; Binary tree; Traversal and Related Properties; Decrease and

Conquer; Insertion Sort; Depth first Search and Breadth First Search, Basics of Greedy algorithms.

Unit VI: Other related topics

Basics of Parallel Algorithms, Introductory Graph Algorithms, Dynamic programming

Textbooks and Reference Books:

1. **Introduction to algorithms**; *T.H. Cormen*; Prentice Hall of India Pvt Ltd, 2nd edition, 2001.
2. **Genetic algorithms**; *D.E. Goldberg*; Pearson Education, 2006.
3. **Bioinformatics: Data, sequence analysis & evolution**; *J.M. Keith*; Humana Press, 2008.
4. **Algorithms in bioinformatics**; *G. Benson, and R. Page*; Springer, 2003.
5. **Bioinformatics: Databases, tools and algorithms**; *O. Basu and S.K. Thukral*; Oxford Press, 2007.

CHEMINFORMATICS	CREDITS: 2
-----------------	------------

OBJECTIVES:

This course will introduce the discipline of chemoinformatics and its applications in computational drug discovery. Students who successfully complete this course will be able to grasp the following areas of the drug discovery cycle

- drug life cycle, stages of drug discovery, and strategic issues in drug discovery and development.
- the principles of computational methods involved in Lead generation
- principles of virtual screening and quantitative structure activity relationship. Principles and practice of docking.
- concepts and algorithms involved in rational and automated peptide and protein design.

Unit I: Drug discovery pipeline

Drug life cycle, stages of drug discovery, and strategic issues in drug discovery.

Unit II: Lead generation

2D and 3D molecular structures; molecular descriptors and fingerprints; molecular similarity and diversity

Unit III: Physical and chemical properties

(solubility and partition coefficients), pharmacological properties (absorption, distribution and toxicity), and global properties (oral bioavailability and drug-likeness)

Unit IV: Overview of drug development

HTS; Clinical trials; Data analysis, classification methods, quantitative structure-property relationships. Applications of chemoinformatics in drug research (chemical libraries, virtual screening, protein-ligand interactions, prediction of pharmacological properties).

Unit V: Computer aided drug design

Methods of computer aided drug design, ligand design methods, docking algorithms, drug design approaches, computer based tools for drug design. pharmacophore generation.

Textbooks and Reference Books:

1. **Handbook of chemoinformatics**; *J. Gasteiger*; Wiley-VCH, 2003.
2. **Cheminformatics in drug discovery**; *Ed. T.I. Opera*; Wiley Publishers, 2004.

SEQUENCE ANALYSIS	CREDITS: 2
-------------------	------------

OBJECTIVES:

This course has been designed to introduce and explain all aspects of nucleotide and protein sequence analysis. The course will focus on pairwise and multiple sequence alignments and their different applications in understanding biology. Students will acquire knowledge on the following:

- A comprehensive understanding of sequence alignment and the related statistics
- All possible applications of sequence alignment that includes the prediction of overall similarity and dissimilarity across gene and protein families, domain identification, basic scoring functions and the relevance of confidence
- Understanding of popular sequence analysis algorithms – different versions of nucleotide and pro-

tein BLAST and the softwares for multiple sequence alignment

- Identification of remote homologs etc.
- Ability to write programs for the pairwise and multiple sequence alignments
- Ability to install and configure different versions of BLAST
- Ability to write programs for basic scoring and statistical functions to be used for homology searches
- Ability to write programs for sequence homology based domain searches

Unit I:

Various file formats for bio-molecular sequences: genbank, fasta, gcg, msf, nbrf-pir etc. Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues. Sequence patterns and profiles: Basic concept and definition of sequence patterns, motifs and profiles, various types of pattern representations viz. consensus, regular expression (prosite-type) and profiles; profile-based database searches using PSI-BLAST, analysis and interpretation of profile-based searches.

Unit II:

Dot matrix analysis Substitution matrices. Pairwise sequence alignments (Algorithmic concepts); Local Vs Global Multiple sequence alignment (Algorithmic concepts); Dynamic programming and progressive sequence alignment. Sequence analysis and dynamic programming.

Unit III:

Scoring matrices: basic concept of a scoring matrix, Matrices for nucleic acid and proteins sequences, PAM and BLOSUM series, principles based on which these matrices are derived. Differences between distance & similarity matrix, Sequence-based database searches: what are sequence-based database searches, BLAST and FASTA algorithms, Various versions of basic BLAST and FASTA, Use of these methods for sequence analysis including the on-line use of the tools and interpretation of results.

Unit IV:

Multiple sequence alignments (MSA): the need for MSA,

basic concepts of various approaches for MSA (e.g., progressive, hierarchical etc.). Algorithm of CLUSTALW and Pileup and their application for sequence analysis (including interpretation of results), concept of dendograms and their interpretation.

Unit V:

Relationship between basic concepts in systematics, phylogeny and concepts in molecular evolution; nature of data used in taxonomy and phylogeny; definition and description of phylogenetic trees and various types of trees; case studies in phylogenetic sequence analysis

Textbooks and Reference Books:

1. **Bioinformatics vol.1: Data, sequence analysis & evolution;** *J. M. Keith*; Humana Press, 2008.
2. **Biological sequence analysis;** *R. Durbin*; Cambridge University Press, 1998.
3. **A cell biologists' guide to modeling and bioinformatics;** *R. M. Holmes*; Wiley Interscience, 2007.

LAB COURSES

<p>COMPUTER PROGRAMMING (HTML, XML, CGI/PERL) EXERCISES WITH APLPLICATIONS TO LIFE SCIENCE</p>	<p>CREDITS: 2</p>
-------------------------------------------------------------------------------------------------------	--------------------------

OBJECTIVES:

- Appreciate the role played by HTML in today's Internet-driven world and understand the tags that are used in HTML documents
- Design and build static websites based on the given requirements and Link documents together using hyperlinks
- Appreciate the need and use of XML as a standard for data exchange and communication.
- Understand the syntactic requirements of XML documents and Create custom XML documents based on given requirements.
- Document the structure of XML documents using DTDs.
- Document the structure of XML documents using XML Schema and understand the concept of XML Parsing.
- Understand the CGI architecture and CGI scripting requirements.

- Create HTML pages with forms and submit forms to web servers passing form data.
- Write CGI scripts using Perl to read form data.
- Design CGI scripts that respond appropriately to user requests.
- Develop CGI scripts using Perl that interact with the MySQL database and responds to user requests.

Unit I: Overview of Internet, HTML & Basic HTML Tags

Overview of the Internet: Internet; web server; web browser; HTTP and HTML; domains and IP addresses. HTML Tags; pairing tags; attributes. Basic HTML Tags: html; head; body; title; headings; paragraphs; bold, italics and underline, horizontal rule.

Unit II: Working with colours & frames

Colour models and representations; coloured text and background. Working with frames, images and tables: using frames; embedding images; creating tables; formatting tables.

Unit III: Hyperlinks, Miscellaneous tags and Forms

Creating hyperlinks to other pages; creating local hyperlinks; images as hyperlinks. Abbreviations, pre-formatted text; lists – ordered and unordered. Forms: need, form elements.

Unit IV: Introduction to XML and XML basics

Need for XML; differences between XML and HTML; application of XML; XML examples. XML basics: elements; attributes; validation.

Unit V: DTD

DTD declarations-internal and external; use of DTD; DTD-XML building blocks; DTD basics: elements, attributes, entity- internal and external; validation; examples.

Unit VI: Schema

Introduction; use of schema; examples; Simple types: elements, attributes, restrictions, examples; Complex types: elements- empty, elements only, text only, mixed, examples; indicators: order, occurrence, group, examples; data types: string, date, misc.

Unit VII: Introduction to CGI and HTML forms

History of CGI; architecture of CGI; application of CGI.

HTML forms revisited: HTML form elements and attributes; text fields; text areas; radio buttons; check boxes; submit button; reset button; drop down boxes; how to create an HTML form.

Unit VIII: Simple input/output to CGI

CGI environment variables; simple responses.

Unit IX: Interface between HTML forms and CGI

Sending data to the server; GET and POST method; use of CGI.pm module.

Textbooks and Reference Books:

1. **Complete reference HTML**, *T. A. Powell*, Tata Mcgraw Hill, 1999.
2. **Beginning XML**, *D. Hunter*, Shroff publishers & Distributors, 2000

BIOLOGY LAB: EXPERIMENTS WITH RNA, DNA AND GENETIC ENGINEERING	CREDITS: 2
-------------------------------------------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- perform total genomic and plasmid DNA extractions from bacteria
- perform total genomic DNA extractions from a few tissues (solid and liquid, plant and animal)
- isolate total RNA and mRNA from a few important tissues and cells
- assess the quantify and quality of RNA/DNA using gel-documentation and spectrophotometric method; state the relative advantages and disadvantages of these methods
- amplify a given region of DNA and RNA by PCR and RT-PCR, respectively
- analyze DNA and RNA by using Southern and Northern blotting procedures, respectively
- list the steps involved in recombinant DNA technology and genetic engineering; explain the principles of the same
- list the major enzymes required for in vitro handling of DNA and RNA; explain the differences of such enzymes as compared to enzymes in vivo

- explain the principles of types of vectors used for general cloning and expressing proteins
- state and discuss the ethical dilemma associated with genetic engineering and associated applications
- design primers with 5' addition of restriction sites or mutations
- prepare competent bacterial cells
- perform restriction digestion and ligation of given DNA and vectors, and transformation of bacteria
- select the transformed colonies and further clone the target genes to express proteins
- create cDNA libraries and screen expression libraries

Unit I: DNA handling

Isolation and quantification of genomic (plant and animal) and plasmid DNAs; analysis of DNA by electrophoresis and gel documentation methods, primer designing, PCR amplification of a target DNA.

Unit II. RNA handling

Isolation of mRNA from at least 2 types of sources, analysis by electrophoresis and gel documentation methods, RT-PCR, qRT-PCR, microarray.

Unit III: Transforming bacteria

Bacterial culture techniques, preparing competent cells, transformation of bacteria and selection of transformed cells.

Unit IV: Cloning

Designing primers with mutations and restriction sites; cloning a RT-PCR product in simple vectors using restriction digestion and ligation; purification of cloned target and sequencing the same; gateway cloning.

Unit V: Expression and down-stream processing of proteins

Expression of cloned sequences in *E. coli* and *B. subtilis*; purification of proteins by standard column and batch chromatography, and filtering methods; comparison of protocols for cell disruption and protein purification methods.

Textbooks and Reference Books:

1. **The condensed protocols from the molecular cloning: A lab manual** – *J. Sambrook*; Cold Spring

Harbour Lab. Press, 2006.

2. **Methods in biotechnology**, Taylor & Francis, 1997

3. **Molecular biotechnology: Principles and application of recombinant DNA**, *B. R. Glick and J. J. Pasternak*, ASM PRESS, 2004.

4. **Methods in biotechnology**, *H.P. Schmauder*, Taylor & Francis, 1997.

5. **Protein expression: A practical approach**; *S. J. Higgins and B. D. Hames*, Oxford University Press, 2004

ALGORITHMS AND DATA STRUCTURES	CREDITS: 2
--------------------------------	------------

As mentioned above, together with theory

CHEMINFORMATICS AND COMPUTATIONAL STRUCTURAL BIOLOGY	CREDITS: 2
------------------------------------------------------	------------

OBJECTIVES:

This lab course will make use of Discovery studio, Pipeline pilot and MOE program suites to fulfil the following objectives.

- Characterizing protein binding site.
- Dock and score ligand molecules with the help of Autodock and DS.
- Use and customise SAR tools
- Design targeted chemical libraries.
- Learn to use data analysis and reporting platforms like Pipeline pilot or knime.
- Perform basic operations like building & editing models of peptides, proteins, carbohydrates, and nucleic acids and visualizing complex macromolecular structures. Convert molecules from one structural format to other and perform basic molecular engineering in a computing machine.
- Perform energy minimization and molecular dynamics simulations using GROMOS and CHARMM force-fields.
- Perform homology and comparative modelling of protein 3D structures using 'HOMOLOGY' and 'MODELLER' programs.

Unit I: Molecular docking

Characterization of protein binding site, docking, QSAR

Unit II: Molecular modeling

Building and editing peptide, nucleic acid, representations, Geometrical measurement Construction of models of peptides, proteins, carbohydrates, and nucleic acids for visualizing complex macromolecular structures. Molecular format conversion, molecular visualization program for displaying, animating, and analyzing large bio-molecular systems using 3-D graphics. Comparing structures, mutations, studying interactions creating electrostatic potential diagrams. Exercise from sequence to structure.

Unit III: Molecular dynamics

Molecular Mechanics-Energy minimization; Molecular dynamics. Simulation of proteins using GROMOS and CHARMM forcefields.

Unit IV: Homology modeling

Homology and comparative modeling of protein 3D structures using 'HOMOLOGY' and 'MODELLER'.

Textbooks and Reference Books:

1. **Handbook of chemoinformatics**; *J. Gasteiger*; Wiley-VCH, 2003.
2. **Cheminformatics in drug discovery**; *Opera*; Wiley Publishers, 2004.
3. **Molecular modeling - principles and applications**, *A. R. Leach*, 2nd Edition, Pearson Education Limited, U.K, 2001.
4. **Molecular dynamics simulation: elementary methods**, *J. M. Haile*, Wiley-Interscience, New York, 1992.

SEMESTER IV

Theory

Systems biology	2 credits
Matlab and bioinformatics	2 credits
Instrumental methods of analysis	2 credit
Fermentation and downstream processing	2 credit
Entrepreneurship, intellectual property and the biotech industry	2 credits

Lab

Systems biology	2 credit
Fermentation and downstream processing	2 credit
Project work	4 credits

THEORY

SYSTEMS BIOLOGY	CREDITS: 2
------------------------	-------------------

OBJECTIVES:

This course will introduce the discipline of molecular and biochemical flux based systems biology and its applications in drug target identification and biomarker discovery. Students who successfully complete this course will be able to grasp the following:

- Interaction of signaling and metabolic pathways
- Static and dynamic systems biology models (normal vs. disease)
- Systems biology models constructed from protein-protein interaction data
- Systems biology models constructed from microarray expression data
- Systems biology models constructed from enzyme kinetics parameters.
- Biological interpretation of all types of models towards understanding an example disease process

Unit I:

Analysis of DNA microarray data: computing condition-specific patterns of co-expression; identifying sets of co-regulated genes. Constructing co-expression networks.

Unit II:

Protein protein interactions: databases such as DIP, PPI server and tools for analysis of protein protein

interactions, Bioinformatics Resources at the species level. Constructing and analyzing protein interactions networks using system wide approaches.

Unit III:

High-throughput biological data: experimental techniques for detecting protein-protein interactions, protein-DNA binding, protein phosphorylation, metabolite concentrations; assessing data quality.

Unit IV:

Flux based analysis of pathways in prokaryotes and eukaryotes. Cross talk analysis of pathways at the systems level.

Unit V:

Regulatory networks: determining their basic building blocks and structural properties. Interaction networks underlying infectious diseases: prediction and analysis. Obtaining biological insights by combining diverse sources of data: functional annotation of proteins; improving the reliability of the data; refining our current knowledge about biochemical pathways.

Textbooks and Reference Books:

1. **Advanced analysis of gene expression microarray data**; *A. Zhang*; World Scientific Publishing, 2006.
2. **Systems biology: A textbook**; *E. Klipp et.al.*; Wiley, 2009
3. **DNA microarrays**; *M. Schena*; Scion Publishing, 2006.

MATLAB AND BIOINFORMATICS	CREDITS: 2
----------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- Understand and handle the open MATLAB language
- Read, analyze and visualize data from genomics and proteomics experiments
- Prototype new algorithms in the MATLAB language or customize existing ones.
- Build applications related to genomics and proteomics.

Unit I: Understanding MATLAB

Understand the basic concepts of MATLAB.

Unit II: Handling MATLAB

Reading, visualizing and analyzing large data sets generated by genomics and proteomics experiments.

Unit III: Algorithms and applications

Modifying existing algorithms in MATLAB; writing new algorithms

Textbooks and Reference Books:

1. **Matlab: A practical introduction to programming and problem solving**, *S. Attaway*, Butterworth-Hienemann, 2009
2. **Matlab for neuroscientists: An introduction to scientific computing in matlab**, *P Wallisch, M Lusignan, M Benayoun, and T. I. Baker*, Elsevier, 2009
3. **Essentials of MATLAB programming**, *S. J. Chapman*, 2nd Edition, BAE SYSTEMS, Australia.

INSTRUMENTAL METHODS OF ANALYSIS	CREDITS: 2
-----------------------------------------	-------------------

OBJECTIVES:

This course will support the wet-lab experiments by having a comprehensive account of the principles and instrumentation of chromatographs and spectrophotometers that are used in the laboratory. Students who successfully complete this course will

- be able to explain the principles of operation of major equipment in the biotech lab
- be able to prepare the standard operating procedures for each equipment, which would ensure safety and correct results
- be able to operate all of the equipment and identify the sensitivity ranges, when applicable
- have an understanding of the theory of chromatography, and hands-on experience in thin layer chromatography (TLC), and HPLC.
- have an awareness of the theoretical aspects of X-ray crystallography.
- be able to analyze UV-VIS, NMR, Fluorescence, CD, and Mass spectrum of organic and bio-molecules;

identify organic molecules from spectral data and predict protein conformational aspects from NMR and CD.

Unit I: Chromatographic methods

Theory of chromatography. Retention, plate theory, separation with a stationary bed column. Thin layer chromatography, Liquid chromatography, ion exchange, HPLC, GPC and application to proteomics.

Unit II: X-ray crystallography

X-ray diffraction and Bragg's law, crystal systems, point groups and space group, application for molecular structure determination.

Unit III: Biomolecular spectroscopy

Principles, time and frequency domain, UV-VIS, fluorescence, infrared and Raman spectroscopy, CD.

Unit IV: NMR

Chemical shifts, coupling constants, Karplus equation and torsion angles, application to structure determination of biomolecules.

Unit V: Mass spectrometry

Ionization techniques, ESI-MS, MALDI –TOF, Tandem MS, applications in proteomics.

Unit VI: Microscopy

Principles of operations of different types of microscopes: regular and confocal microscopes, electron microscopy.

Textbooks and Reference Books:

1. **Biomolecular crystallography**; *B. Rupp*; Garland Science, 2009.
2. **Spectroscopy for biological sciences**; *G.G. Hammes*; Wiley, 2005.
3. **Principles of fluorescence spectroscopy**; *J. R. Lakowicz*; Springer, 3rd edition, 2008.
4. **The condensed protocols from the molecular cloning: A lab manual**; *J. Sambrook*; Cold Spring Harbour Lab. Press, 2006.
5. **Outline of crystallography for biologists**; *D. Blow*; Oxford University Press, 2002.

6. **Chromatography: Concepts and contrasts**; *J. M. Miller*; Wiley, 2nd edition, 2009.
7. **Protein purification** – principles and practice; R.K. Scopes, Springer Verlag.
8. **BioCromatography**-Theory and practice; MA Vijayalakshmi, Taylor and Francis.

FERMENTATION AND DOWNSTREAM PROCESSING	CREDITS: 2
-------------------------------------------	------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- explain the principles, advantages and limitations of different types of fermentation
- perform basic bacterial culture and maintain bacterial stocks
- assess the growth characteristics of bacteria in laboratory scale as well as large-scale, using fermentors
- perform sonication as well as other procedures to extract expressed proteins of interest, from bacteria
- perform processes in preparation, fermentation and downstream processes; identify factors that will affect fermentation technically and economically
- isolate the expressed proteins, and assess the purity and functionality of the same
- explain and perform laboratory as well as industrial scale protein purification
- perform FPLC experiments on at least a few samples

Unit I: Fermentation and fermentors

Introduction (theory): Basic types and principles of fermentation and their applications. Anatomy, operation and maintenance of fermentors for bacteria and mammalian cells.

Unit II: Introduction to down stream processing

Significance of downstream processing in biotechnology; principles, procedures and economics of downstream processing, physico-chemical principles involved in bio-separation processes.

Unit III: Product purification

Intracellular and extracellular fermentation products; options for cell disruption for intracellular products,

removal of insoluble, debris/biomass separation techniques, flocculation and sedimentation, centrifugation and filtration methods. membrane-based separations (micro and ultrafiltration).

Unit IV: Precipitation and extraction processes

Precipitation using salts, organic solvents, and polymers; aqueous two-phase extraction, supercritical extraction; dialysis, crystallization, etc.

Textbooks and Reference Books:

1. **Rate controlled separations;** *P C Wankat*; Elsevier, 1990.
2. **Bioseparations;** *P A Belter and E Cussler*; Wiley 1985.
3. **Product recovery in bioprocess technology,** BIOTOL Series, VCH, 1990.
4. **Separation processes in biotechnology;** *J. M. Asenjo*; 1993.
5. **Batch fermentation - modeling, monitoring and control;** *A. Cinar, S. J Parulekar, C. Undey and G. Birol*; 2003.
6. **Culture of animal cells;** *R. Ian Freshney*; John Wiley & Sons, Inc, 1998.
7. **Downstream processing of proteins: methods and protocols (methods in biotechnology);** *M.A. Desai*; Humana Press, 2000.
8. **Principles of fermentation technology;** *P.F. Stanbury et.al*; Elsevier Science, 2nd edition, 2003.
9. **Fermentation and biochemical engineering handbook: Principles, process design and equipment;** *H.C. Vogel and C.C. Haber*; Noyes Publications, 2nd edition, 2007.

ENTREPRENEURSHIP, INTELLECTUAL PROPERTY AND THE BIOTECH INDUSTRY	CREDITS: 2
-----------------------------------------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will:

Unit I: Intellectual property

- Understand the different types of intellectual property
- Understand the particular relevance of IP in the area of biotechnology
- Understand the basic structure of a patent and be able

to do an elementary level drafting of patent specification.

- Know about different patent offices in major countries and the information available from these offices online.
- Understand the particular importance of traditional knowledge and bioethics in India
- Understand the major treaties that govern IP internationally
- Understand the concepts of prior art and freedom to operate.

Unit II: Entrepreneurship

- Distinguish an opportunity as appropriate for a social venture or a business venture
- Identify an opportunity for leveraging resources
- Identify an opportunity for creating value
- Identify different types of customer needs, and slot these into different market segments.
- Realize the nature of the proposed business

Unit III: The biotech industry

- Understand the different types of companies in the biotech industry, both in terms of area of work and in terms of stage of work
- Understand the major ethical issues in this industry
- Understand the major regulatory issues in this industry
- Understand the kinds of analysis that are required to assist in rational policy making to support the industry.

Unit I: Intellectual property

Introduction to intellectual property; Biotechnology and IP; anatomy of a patent; analyzing patentability of biotech inventions; research and IP; major treaties that govern IP; patent search; drafting patent specifications; introduction to Indian and US patent offices; traditional knowledge and bioethics; freedom to operate/licensing.

Unit II: Entrepreneurship

Introduction to social and business entrepreneurship.; the importance of spotting the opportunity, leveraging resources and creating value; customer needs and market segmentation; categories of value: enhancements, extensions and specializations.

Unit III: The biotech industry

The structure of the industry: companies working in different areas of specialization (agri-biotech, pharma-biotech etc.) and different stages (R&D, manufacture, sales and marketing); ethical issues, regulatory issues, the changing nature of the pharma-biotech industry; policy making relevant to the biotech industry

Textbooks and Reference Books (The books change every year. The following is an illustrative list):

1. **Entrepreneur journeys Vol. 1**; *S. Mitra*; Hachette Publishing Group, 2009.
2. **Entrepreneurship**; *R.D. Hisrich*; Tata McGraw-Hill Publishing Co Ltd, 6th edition, 2006.
3. **Strategic entrepreneurial growth**; *D.F. Kuratko, H.P. Welsch*; Harcourt College Publishers, 2001.
4. **Intellectual property rights**; *P. Ganguli*; Tata McGraw Hill Publishing Co Ltd, 2009.
5. **Biotechnology, IPRs and biodiversity**; *M.B. Rao*; Pearson Publications, 2008.

LAB COURSES

SYSTEMS BIOLOGY	CREDITS: 2
------------------------	-------------------

OBJECTIVES:

This lab course will make use of Cytoscape and other free softwares to develop models:

- Developing different types of systems biology models
- Stastical and topological analysis of static and dynamic systems biology models
- Selection of statistically significant models
- Biological analysis and interpretation of these models to describe some diseases examples
- Ability to select drug targets from each type of model
- Ability to identify biomarkers

Unit I:

Analysis of protein-protein interaction data

Unit II:

Analysis of microarray expression data

Unit III:

Learning on statistical models for pathway modelling using molecular data.

Unit IV:

Developing network models using expression data, interaction data and SNP data.

Textbooks and Reference Books:

1. **Advanced analysis of gene expression microarray data**; *A. Zhang*; World Scientific Publishing, 2006.
2. **Systems biology: A textbook**; *E. Klipp et.al.*; Wiley, 2009
3. **DNA microarrays**; *M. Schena*; Scion Publishing, 2006.

FERMENTATION AND DOWNSTREAM PROCESSING	CREDITS: 2
-----------------------------------------------	-------------------

OBJECTIVES:

Students who successfully complete this course will be able to:

- explain the principles, advantages and limitations of different types of fermentation
- perform basic bacterial culture and maintain bacterial stocks
- assess the growth characteristics of bacteria during laboratory scale experiments as well as during large-scale production, using fermentors
- perform sonication as well as other procedures to extract the expressed proteins of interest from bacteria
- perform steps in preparation, fermentation and downstream processing; identify factors that will affect fermentation technically and economically
- isolate the expressed proteins, and assess the purity and functionality of the same
- explain and perform laboratory as well as industrial-scale protein purification
- perform FPLC experiments on at least a few samples

Unit I: Fermentor operation

Anatomy, operation and maintenance of fermentors for bacteria and mammalian cells.

Unit II: Bacterial growth dynamics

Thorough bacterial growth studies and preliminary fermentation exercises

Unit III: Fermentation and product purification

Case studies in fermentation and product purification

Unit IV: Basic laboratory procedures

Use of laboratory scale fermentors to overexpress a cloned protein.

Unit V: Down stream processing

FPLC use to purify the protein of interest.

Textbooks and Reference Books:

1. **Batch fermentation - modeling, monitoring and control**; *A. Cinar, S. J. Parulekar, C. Undey and G. Birol*; Marcel Dekker. 2003.
2. **Culture of animal cells**; *R. I. Freshney*; John Wiley & Sons, Inc, 1998.
3. **Downstream processing of proteins: Methods and protocols** (methods in biotechnology); *M. A. Desai*; Humana Press, 2000.

PROJECT WORK	CREDITS: 4
---------------------	-------------------

Students will work with faculty on research projects during the 4th semester.

OBJECTIVES:

Students who successfully complete the project work would have practical, long term experience in (almost independently) handling a real-time research problem/industrial task. The students would have gained expertise in a very specific domain and thereby be able to differentiate between acquiring the preliminary abilities during regular course sessions vs. learning through actual task handling in the long run. The work would provide a clear understanding of their capacities and limitations in different aspects – knowledge, skills and attitude/personality.

The students will be confident about their abilities. They would have also realized the importance of planning, patience, team work and communication in

accomplishing a given task. In general, students will also be able to:

- explain in detail about several aspects related to the project work
- demonstrate the technical skills used extensively in the project
- work independently to formulate a strategy to address the given problem/task
- complete specific tasks within the estimated time
- systematically report any completed task verbally and formally, in front of an audience, as well as in formal writing