Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

CURRICULUM STRUCTURE AND DETAILED SYLLABI

FOR

MASTER OF SCIENCE

IN

BIOINFORMATICS

(Applicable from the academic year 2021-2022)

Department of Bioinformatics
Maulana Abul Kalam Azad University of Technology, West Bengal
(Formerly West Bengal University of Technology)
Haringhata-741249, Nadia, West Bengal, India

Web: https://makautwb.ac.in

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Vision of the Department

To emerge as a world class center of education for building the best human resource in the Bioinformatics and allied areas through teaching, training, innovation, and research.

Mission of the Department

M1: To create opportunities for multi-disciplinary education, training, and research in the area of Bioinformatics.

M2: To educate the students to choose professions in industry, academia, and entrepreneurship.

M3: To impart environmental, legal, and ethical awareness to the students for the inclusive development of the society.

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Program Educational Objectives (PEOs)

The PEOs for the graduating students of M.Sc. in Bioinformatics program will be able to

PEO1: Pursue advanced education and research in Bioinformatics to lifelong learning.

PEO2: Flourish in their trained professional career in industry and academics globally.

PEO3: Inculcate spirit of work, constructive thinking, professionalism, enhancing entrepreneurship skills and ethical behavior.

Program Outcomes (POs)

The graduates of M.Sc. in Bioinformatics program will have

PO1: An ability to acquire knowledge to solve biological problems using Information technology.

PO2: Apply appropriate techniques, resources, and software to predict and model complex bioinformatics systems.

PO3: An ability to expertise in interdisciplinary research and collaborate towards achieving common goals.

PO4: Deploy a wide range portfolio of advanced analytical and Bioinformatics techniques using contemporary software tools for the purpose of solving real-life problems.

Program Specific Outcomes (PSOs):

After successful completion of the program, the graduates of M.Sc. in Bioinformatics program shall be able to

PSO1: Utilize modern software tools for mathematical modeling and solution of problems with high throughput in the domain of Bioinformatics.

PSO2: Use theoretical and practical knowledge of the program appropriately in academics and industrial fields.

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

M.Sc in Bioinformatics: 2021-2022
Curriculum Structure

For 1st Semester: Total 21 Credits

Sl. No	Category	Subject Code	Subject Name	Contact Hrs./wk	Credits	Total
		Theory		L-T-P		
1	**Theory**	MSBIN 101	Molecular Biology	3-0-0	3	21
2	**Theory**	MSBIN 102	Computational Biochemistry	3-0-0	3	
3	**Theory**	MSBIN 103	Mathematics and Statistics	3-0-0	3	
4	**Theory**	MSBIN 104	Application of Bio tools and Bio database	3-0-0	3	
5	**Theory**	MSBIN 105	Data Structure and Application	3-0-0	3	
		Practical				
1	**Practical**	MSBIN 192	Computational Biochemistry Lab	0-0-4	2	
2	**Practical**	MSBIN 194	Bioinformatics Lab	0-0-4	2	
3	**Practical**	MSBIN 195	Data Structure and Application Lab	0-0-4	2	

For 2nd Semester : Total 21 Credits

Sl. No	Category	Subject Code	Subject Name	Contact Hrs./wk	Credits	Total
		Theory		L-T-P		
1	**Theory**	MSBIN 201	Structural Bioinformatics	3-0-0	3	**21**
2	**Theory**	MSBIN 202	Genomics and Proteomics	3-0-0	3	
3	**Theory**	MSBIN 203	Molecular Modeling and Molecular Dynamics	3-0-0	3	
4	**Theory**	MSBIN 204	Computer language(Python)	3-0-0	3	
		Practical				

1	**Practical**	MSBIN 291	Structural Bioinformatics Lab	0-0-6	3	
2	**Practical**	MSBIN 292	Genomics and Proteomics Lab	0-0-6	3	
3	**Practical**	MSBIN 294	Computer language (Python) Lab	0-0-6	3	

For 3rd Semester: Total 21 Credits

Sl. No	Category	Subject Code	Subject Name	Contact Hrs./wk	Credits	Total
		Theory		**L-T-P**		
1	**Theory**	MSBIN 301	Computational Drug Design	3-0-0	3	**21**
2	**Theory**	MSBIN 302	System Biology	3-0-0	3	
3	**Mandatory Learning Course**	MSBIN 303	Research Methodology and IPR	3-0-0	3	
4	**Theory (Elective)**	MSBIN 304(A)	Omics Technology	3-0-0	3	
		MSBIN 304(B)	Next Generation Sequence Technology			
		MSBIN 304(C)	Embryology and Human Genetics			
		Practical				
1	**Practical**	MSBIN 391	Computational Drug Design Lab	0-0-6	3	
2	**Practical**	MSBIN 394	Elective Lab	0-0-6	3	
3	**Practical**	MSBIN 395	Computer language (R) Lab	0-0-6	3	

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

For 4th Semester: Total 21 Credits

Sl. No	Category	Subject Code	Subject Name	Credits	Total
Sessional					**21**
1	**Sessional**	MSBIN 481	Project Work	15	
	Sessional	MSBIN 482	Grand Viva	3	
	Sessional	MSBIN 483	Communication Skill	3	
Total Credits for the Program					**84**

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

M.Sc in Bioinformatics: 2021-2022

Curriculum Structure

Subject Name: Molecular Biology	Category: Theory
Subject Code: MSBIN 101	Semester: Frist
L-T-P: 3-0-0	Credit: 3

Course Outcome: On completion of this course, the students are expected to:

Sl No	Course Outcome	Mapped Unit
01	MSBIN101.1: Describe the dynamic organization of cells.	I
02	MSBIN101.2: Outline the basic structure of chromatin.	II
03	MSBIN101.3: Summarize the details of various signal transduction processes in correlation with the biological impact like onset of cancer.	III
04	MSBIN101.4: Apply the knowledge of Cellular processes techniques.	IV
5	MSBIN101.5: Define Manipulating and studying cells method	V
6	MSBIN101.6: Demonstrate the Genome instability and cell transformation process.	VI

Course Contents:

Unit I: Dynamic organization of cell (6 lectures)
Universal features of cells; cell chemistry and biosynthesis: chemical organization of cells; internal organization of the cell - cell membranes: structure of cell membranes and concepts related to compartmentalization in eukaryotic cells; intracellular organelles: endoplasmic reticulum and Golgi apparatus, lysosomes and peroxisomes, ribosomes, cellular cytoskeleton, mitochondria, chloroplasts and cell energetics; nuclear compartment: nucleus, nucleolus and chromosomes.

Unit II: Chromatin structure and dynamics (12 lectures)
Chromatin organization - histone and DNA interactome: structure and assembly of eukaryotic and prokaryotic DNA polymerases, DNA-replication, repair and recombination; chromatin

control: gene transcription and silencing by chromatin-Writers,-Readers and –Erasers; Transcriptional control: Structure and assembly of eukaryotic and prokaryotic RNA Polymerases, promoters and enhancers, transcription factors as activators and repressors, trancriptional initiation, elongation and termination; post-transcriptional control: splicing and addition of cap and tail, mRNA flow through nuclear envelope into cytoplasm, breakdown of selective and specific mRNAs through interference by small non-coding RNAs (miRNAs and siRNAs), protein translation machinery, ribosomes-composition and assembly; universal genetic codes, degeneracy of codons, Wobble hypothesis; Iso-accepting tRNA; mechanism of initiation, elongation and termination; co- and post-translational modifications, mitochondrial genetic code.

Unit III: Cellular signalling, transport and trafficking (3 lectures)
Molecular mechanisms of membrane transport, nuclear transport, transport across mitochondria and chloroplasts; intracellular vesicular trafficking from endoplasmic reticulum through Golgi apparatus to lysosomes/cell exterior.

Unit IV: Cellular processes (8 lectures)
Cell cycle and its regulation; cell division: mitosis, meiosis and cytokinesis; cell differentiation: stem cells, their differentiation into different cell types and organization into specialized tissues; cell-ECM and cell-cell interactions; cell receptors and trans-membrane signalling; cell motility and migration; cell death: different modes of cell death and their regulation.

Unit V: Manipulating and studying cells (3 lectures)
Isolation of cells and basics of cell culture; observing cells under a microscope, different types of microscopy; analyzing and manipulating DNA, RNA and proteins.

Unit VI: Genome instability and cell transformation (8 lectures)
Mutations, proto-oncogenes, oncogenes and tumour suppressor genes, physical, chemical and biological mutagens; types of mutations; intra-genic and inter-genic suppression; transpositions-transposable genetic elements in prokaryotes and eukaryotes, role of transposons in genome; viral and cellular oncogenes; tumor suppressor genes; structure, function and mechanism of action; activation and suppression of tumor suppressor genes; oncogenes as transcriptional activators.

Learning Resources:

1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Molecular Biology of the Cell. New York: Garland Science.

2. Lodish, H. F. (2000). Molecular Cell Biology. New York: W.H. Freeman.

3. Krebs, J. E., Lewin, B., Kilpatrick, S. T., & Goldstein, E. S. (2014). Lewin's Genes XI. Burlington, MA: Jones & Bartlett Learning.

4. Cooper, G. M., &Hausman, R. E. (2009). The Cell: a Molecular Approach. Washington: ASM; Sunderland.

5. Hardin, J., Bertoni, G., Kleinsmith, L. J., & Becker, W. M. (2012). Becker's World of the Cell. Boston: Benjamin Cummings.

6. Watson, J. D. (1987). Molecular Biology of the Gene (7th ed.). Menlo Park, CA: enjamin/Cummings.

Subject Name: Computational Biochemistry	Category: Theory
Subject Code: MSBIN 102	Semester: Frist
L-T-P: 3-0-0	Credit:3

Course Outcome: On completion of this course, the students are expected to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN102.1: *Memorize* the basic knowledge on biomolecules	Unit-I and Unit-II
2	MSBIN102.2: *Describe* the knowledge about chemical bonds	Unit-III
3	MSBIN102.3: *Explain* advance knowledge about the enzyme kinetics and its reactions mechanism.	Unit-IV
4	MSBIN102.4: *Describe* the new hypothesis on designing and engineering of enzyme for industrial application.	Unit-V
5	MSBIN102.5:*Outline* new insights on enzyme at the molecular level, their interactions, and design of biocatalyst.	Unit-V1

Course Contents:

Unit: I (8 Lectures)
Basic Concept of Protein and Nucleic acid: General structure and properties of proteins; amino acid composition of proteins; structures of common amino acids, amino acid sequence in protein molecules, physicochemical properties of proteins, analysis of dihedral angle of residues and Ramachandran plot, denaturation of protein. Molecular structure of DNA, denaturation and renaturation of DNA, molecular organization of messenger ribonucleic acid, transfer ribonucleic acid, ribosomal ribonucleic acid.

Unit: II (8 Lectures)
Biochemistry and Functions of Carbohydrates and Lipids - Introduction of carbohydrates, chemical structures of Carbohydrates-Glucose and Fructose, Classification of carbohydrates, biomedical role of carbohydrates.

Biochemistry and Functions of Lipids- Introduction of Lipids, lipid nomenclature, chemical structures of different lipids, Classification of lipids, Bioinformatics tools for lipidomic analysis, role of lipids in biological membranes.

Unit: III (4 Lectures)

Nature of Chemical Bonds in Biomolecules: Types of chemical bonds-covalent bonds, inter-atomic distances for fractional bond, characters of partial double bond, metallic bond-inter-atomic distance and bond number in metals, hydrogen bond-nature, physical properties of H-bonds, H-bonds in protein and nucleic acids, electrostatic and VDW interaction in biomolecules.

Unit: IV (8 Lectures)

Enzyme Kinetics and inhibition: Concept of enzymes, Enzyme names and classification, Enzyme structure and substrate binding, mechanism of enzyme actions, factors affect rate of enzyme action- pH and temperature, enzyme specificity, enzyme kinetics, regulation of enzyme activity, Enzymes inhibition-reversible and Irreversible inhibition, competitive, uncompetitive and noncompetitive, Industrial enzymology.

Unit: V (6 Lectures)

Enzyme Design and Its Computational Applications: Engineering protein solubility and aggregation, engineering enzyme activity and specificity, engineering protein stability, engineering multiple properties of enzyme. Structural organization, function, and identification method of the enzyme tunnels and channels, function and identification method of the enzyme gates and their molecular engineering. In silico construction of protein mutants and prediction of their activities.

Unit: VI (6 Lectures)

Design of Biocatalyst using Computational tools: Applications of computational design of biocatalysts: design of efficient biocatalysts, design of biocatalysts with novel activity, design of specific and enantio selective biocatalysts, design of biocatalysts with tailored flexibility. Computational Protein Engineering: Computational tools for de novo design of active sites of enzyme, Computational tools for design of ligand exchange pathways, Computational tools for design of protein–protein interfaces, computational tools for analysis of mutation.

Learning Resources:

1. Computational Biochemistry and Biophysics by Oren M. Becker (Editor), Alexander D. MacKerell Jr. (Editor), Publisher:CRC Press, ISBN-13: 978-0824704551.

2. Computational Methods in Physics, Chemistry and Biology: An Introduction by Paul Harrison, Publisher: Wiley, John & Sons, Incorporated, ISBN: 0471495638.

3. Fundamentals of Enzymology by Nicholas C. Price, Publisher: Oxford/Panima Pub Corp, ISBN: 9780198064398.

4. Protein Design Methods And Applications by Guerois, Springer, ISBN:9781588295859.

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Mathematics and Statistics	Category: Theory
Subject Code: MSBIN 103	Semester: Frist
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
01	MSBIN103.1: Apply basic statistical concepts commonly used in Biotechnology.	I
02	MSBIN103.2. Use basic analytical techniques to generate results.	II
03	MSBIN103.3. Demonstrate the design and analysis of statistical methods.Demonstrate and understand the central concepts of modern statistical theory and their probabilistic	III
04	MSBIN103.4. Demonstrate and understand the central concepts of modern statistical theory and their probabilistic foundation.	IV
05	MSBIN103.5. Apply biostatistician knowledge to real-life problems in research	V

Course Contents:

Unit I: Algebra (10 lectures)
Linear equations, functions: slopes-intercepts, forms of two-variable linear equations; constructing linear models in biological systems; quadratic equations (solving, graphing, features of, interpreting quadratic models etc.), introduction to polynomials, graphs of binomials and polynomials; Symmetry of polynomial functions, basics of trigonometric functions, Pythagorean theory, graphing and constructing sinusoidal functions, imaginary numbers, complex numbers, adding-subtracting-multiplying complex numbers, basics of vectors, introduction to matrices.

Unit II: Calculus (4 lectures)
Differential calculus (limits, derivatives), integral calculus (integrals, sequences and series etc.)

Unit III: Mathematical models in biology(6 lectures)
Population dynamics; oscillations, circadian rhythms, developmental patterns, symmetry in biological systems, fractal geometries, size-limits & scaling in biology, modelling chemical reaction networks and metabolic networks.

Unit IV: Statistics (5 lectures)

Probability: counting, conditional probability, discrete and continuous random variables; Error propagation; Populations and samples, expectation, parametric tests of statistical significance, nonparametric hypothesis tests, linear regression, correlation & causality, analysis of variance, factorial experiment design.

Unit IV: Numerical methods (15 lectures)

Solution of non-linear equations: Bisection method, Regula-Falsi method, Newton Raphson.

Solution of linear equations: Gauss Jordan, Gauss-Seidal iterative method, LU Factorization method.

Interpolation: Lagrange's method, Newton's Forward and Backward, Symbolic operatos

Numerical integration: Trapezoidal method, Simpson 1/3 method.

Solution of differential equation: Taylor's series, Euler's method, Runge-Kutta method.

Learning Resources:

1. Stroud, K. A., & Booth, D. J. (2009). Foundation Mathematics. New York, NY: Palgrave Macmillan.

2. Aitken, M., Broadhursts, B., &Haldky, S. (2009) Mathematics for Biological Scientists. Garland Science.

3. Billingsley, P. (1986). Probability and Measure. New York: Wiley.

4. Rosner, B. (2000). Fundamentals of Biostatistics. Boston, MA: Duxbury Press.

5. Daniel, W. W. (1987). Biostatistics, a Foundation for Analysis in the Health Sciences. New York: Wiley.

Subject Name: Application of Bio tools and Bio database	Category: Theory
Subject Code: MSBIN 104	Semester: Frist
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN104.1: Distinguish different databases according to their function.	I
2	MSBIN104.2: Select data retrieval system of chromosome structure, transcript components in 3 kingdom of life.	II
3	MSBIN104.3: Explain the algorithms behind different tools.	III
4	MSBIN104.4: Describe the methods of protein modelling	IV

Course Contents:

Unit I-Introduction to Bioinformatics
A word on Bioinformatics, Introduction, Branches of Bioinformatics, Aims of Bioinformatics, Scope/research areas of bioinformatics.

Unit II-Biological tools and databases
Sequence and molecular file form and ats, Introduction, Sequence file formats, Sequence conversion tools, Molecular file formats, Molecular file format conversion,

Databases in bioinformatics & introduction: - Introduction, Biological databases, Classification schema of biological databases, Biological database retrieval system Biological sequence databases, National Centre for biotechnology information (NCBI), Introduction, Tools and databases of NCBI, Database retrieval tool, Sequence submission to NCBI, BLAST, PSI-BLAST, RPS-BLAST, Specialized tools, Nucleotide database, Literature database, Protein database, Gene expression database, GEO, Structural database, Chemical database, Other databases, EMBL Nucleotide Sequence Database, Introduction, Sequence retrieval, Sequence submission at EMBL, Resources of EMBL, Biological annotation and data curation, Sequence analysis tools, Features of database, DNA databank of Japan, Introduction, Resources of DDBJ, Data submission to DDBJ, Protein information resource, Introduction, Resources of PIR, Data retrieval of PIR, Databases of PIR, Swiss-Prot: Introduction, Features of Swiss-prot Protein 3D structure and classification databases, Protein databank, Introduction, Harnessing data from PDB , Data deposition tools, PDB beta, RCSB PDB structural genomics information portal, Molecular modelling databases, Introduction, Retrieval of structural data from MMDB, Conserved domain database, E-MSD, Introduction, Resources of E-MSD, Data submission at E-MSD, Search system of E-MSD, 3D genomics, Introduction, Assessing 3D genomics, Gene 3D, Introduction, Retrieving data from gene 3D, Protein structural classification Databases, Introduction, CATH-Class, Architecture, Topology, Homologous, SCOP(Structural Classification of Proteins)

Unit III: Bio-algorithms and tools Sequence alignments
Introduction, Concept of alignment, Scoring Matrices, PAM, BLOSUM, Alignment of pairs of sequences, Alignment algorithms, Heuristic methods, Multiple sequence Alignment (MSA) Gene prediction methods: principles and challenges, Introduction, Biological overview, What is gene prediction? Computational methods of gene prediction, Combination of two methods, Why is gene prediction difficult? Molecular Phylogeny Introduction, Phenotypic phylogeny and molecular phylogeny, Molecular clocks, Methods of phylogeny, Statistical Evaluation of the obtained phylogenetic trees or validation methods, Software for phylogenetic analysis, Reliability of molecular phylogenetic prediction Molecular Viewers, Introduction, A few molecular viewers,RasMol, Deep view- The Swiss-PDB viewer (SPDBV), Cn3D.

Unit IV: Protein Modelling Protein structure and modelling Protein and secondary structure prediction: Introduction, Levels of protein structure, Conformational parameters Secondary structure of a protein, Secondary structure types, Secondary structure prediction, Software for secondary structure prediction, Limitations of secondary structure prediction.

Protein modelling: Introduction, Methods of protein modelling, Homology or comparative modelling, Model refinement, Evaluation of the model, Hands on in comparative modelling using Swiss-model, Threading or fold recognition, Ab initio/De novo method.

Learning Recourses:

1. Bioinformatics: a Textbook, Wiley Online Library
2. Mount, D. W. (2001). Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
3. Bourne, P. E., &Gu, J. (2009). Structural Bioinformatics. Hoboken, NJ: Wiley-Liss.
4. Lesk, A. M. (2004). Introduction to Protein Science: Architecture, Function, and Genomics. Oxford: Oxford University Press.
5. Campbell, M &Heyer, L. J. (2006), Discovering Genomics, Proteomics and Bioinformatics, Pearson Education.
6. Oprea, T. (2005). Chemoinformatics in Drug Discovery, Volume 23. Wiley Online Library.
7. Gasteiger, J. &Engel,T. (2003), Chemoinformatics: a Textbook, Wiley Online Library

Subject Name: Data Structure and Application	Category: Theory
Subject Code: MSBIN 105	Semester: Frist

L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
01	MSBIN105.1: Describe the basics of data structure and recall the prerequisite of the data structure and algorithm.	UNIT I
02	MSBIN105.2: Apply the knowledge of linear data structure, like array, linked list, stack and queue.	UNIT II
03	MSBIN105.3: Apply the knowledge of non-linear data structure, like graph and tree.	UNIT III
04	MSBIN105.4: Apply different types of the searching and sorting algorithms.	UNIT IV

Course Contents:

UNIT I (10 lectures):

Algorithms: Introduction, basics of time and space analysis of algorithms – order notations.

Introduction: Why we need data structure? Linear and non-linear data structure.

Array: Concepts,1D & 2D- dimensional array representation in memory, String.

Function and Pointer: Function, Recursion, Tower-of-Hanoi problem, Pointer concept, Array and Pointer, Function and Pointer, Call by value, Call by address.

UNIT II (10 lectures):

Linear list – Singly linked list implementation, insertion, deletion and searching operations on linear list, circular linked list implementation, Applications of linked lists.

Stacks- Operations, representations of stacks, stack applications-infix to postfix conversion, postfix expression evaluation, recursion implementation.

Queues- Operations, representations of queue, Circular Queue operations, Dequeues, applications of queues.

UNIT III (12 lectures):

Graph theory: Concepts – Connected graph, regular graph, undirected graph, directed graph, complete graph, null graph, isomorphic graph, multi-graph and weighted graph. Hamiltonian cycle. Degree of vertex, in-degree & out-degree. Representation of graphs (adjacency matrix & list representation), Application of graph theory in Biological Science.

Trees: Definitions, tree representation, properties of trees, Binary tree, Binary tree representation, binary tree properties, binary tree traversals, binary tree implementation, applications of trees in Biological Science.

UNIT IV (8 lectures):

Searching and Sorting –Sorting- selection sort, bubble sort, insertion sort, merge sort, Searching-linear and binary search methods, comparison of sorting and searching methods.

Learning_Resources:

1. Fundamentals of Data Structures of C –by Ellis Horowitz, SartajSahni, Susan Andersonfreed.
2. Data Structures in C –by Aaron M. Tenenbaum.
3. Data Structures Using C –by Reema Thareja.
4. Data structures and Algorithm Analysis in C, 2nd edition, M.A. Weiss, Pearson.
5. Data structures and Program Design in C, 2nd edition, R. Kruse, C.L. Tondo and B. Leung, Pearson.

Subject Name: Computational Biochemistry Lab	Category: Practical
Subject Code: MSBIN 192	Semester: Frist
L-T-P: 0-0-4	Credit:2

Course Outcomes: On completion of this course, students should be able to:
MSBIN192.1: *Analyze* biomolecular crystallographic data.

MSBIN192.2: *Test* the conformation of biomolecules.

MSBIN192.3: *Analyze* of Ramachandran plot and protein secondary structure.

MSBIN192.4: *Measuring* different chemical bonds in biomolecules.

MSBIN192.5: *Explain* structural properties of enzyme and lipids.

MSBIN192.6: *Design* of smart libraries for engineering protein stability.

Course Contents:

1. Visualization and analysis of protein databank. The basic concept of pr.tein crystallographic data, Detail structural analysis of protein, DNA, RNA, tRNA, protein-DNA, and protein-RNA complexes. Measuring inter-atomic distances of protein-ligand structures.
2. Calculation of dihedral angles of important residues in protein and identification of their conformation. Analysis of Ramachandran plot and protein secondary structure.
3. Computation of H-bonds and VDW interactions in protein and protein-ligand complexes, active site analysis of enzyme, protein-water-ligand, and protein-water-metal interactions study.
4. Identification of enzyme tunnels, channels, gates, and ligands' transport by computational tools. *In silico* analysis of protein mutants.
5. Prediction of protein solubility from protein primary sequence. Transportation of substrate, a product, an inhibitor, a co-factor, or a co-solvent from the outside environment into the binding site and vice versa using computational tools.
6. Identification of 'hot spots' and design of smart libraries for engineering protein stability, catalytic activity, substrate specificity.
7. Etc.

Subject Name: Bioinformatics Lab	Category: Practical
Subject Code: MSBIN 194	Semester: Frist
L-T-P: 0-0-4	Credit:2

Course Outcomes:

On completion of this course, students should be able to:

MSBIN194.1: Apply different commands to the U/nix operating system.

MSBIN194.2: Apply UNIX operating system in Linux platform.

MSBIN194.3: Apply different commands in Shell Script.

MSBIN194.4: Describe the concept of Mac operating system and server handling.

Course Contents:

1. Conceptual background, Understanding the Unix command.

2. General purpose: Utilities, the file system, Handling ordinary files, Basic file attributes.

3. The vi editor, the shell, the process, Customize the environment, Simple filters, Essential shell programming.

4. Essential system administrator, The X window system.

5. All bio tool applications.

Subject Name: Data Structure and Application Lab	Category: Practical
Subject Code: MSBIN 195	Semester: Frist
L-T-P: 0-0-4	Credit:2

Course Outcomes: On completion of this course, students should be able to:
MSBIN195.1: Apply appropriate logic and data structure to solve given problem.
MSBIN195.2: Code on linear data structure respective problems.
MSBIN195.3: Code on non-linear data structure related problems.
MSBIN195.4: Apply appropriate logic and data structure to solve searching and sorting problems.

Course Contents:

1. C language syntax revision with the solution of different problems.

2. Implementation of different 1-D and 2-D related problems using C language.

3. Implementation of different Singly and Circular linked list related problems (creation, insertion, deletion, traversing etc.) using C language.

4. Implementation of different Recursive problems using C language.

5. Implement the representations of Stacks and its application using C language.

6. Implement the representations of different types of Queues using C language.

7. Implementation of different types of Tree problems using C language.

8. Implement selection sort, bubble sort, insertion sort, quick sort, merge sort methods using C language.

9. Implement linear and binary search methods using C language.

Subject Name: Structural Bioinformatics	Category: Theory

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Code: MSBIN 201	Semester: Second
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN201.**1**: *Outline* basic knowledge of protein structure and function.	Unit-I
2	MSBIN201.**2**: *Summerize* new methods of analysing and characterization of macromolecules for solving problems in protein chemistry.	Unit-II
3	MSBIN201.**3**: *Describe* accurate and comprehensive information about the classification of protein structures.	Unit-III
4	MSBIN201.**4**: *Apply* the computational chemistry tools, and their the underlying theory.	Unit-IV
5	MSBIN201.**5**: *Discuss* the dynamic behaviour of proteins by computational molecular dynamics simulations method.	Unit-V

Course Contents:

Unit I: (10 Lectures)

Advances in Protein Structure Prediction: Structural organization and function of Greek key, Jelly roll motif, and Rossmann fold, Protein structure prediction- template-based modelling; template-free modelling; homology modelling, Chou and Fasman and GOR method, Loop modelling, ab-initio modelling, fragment assembly. Methods for predicting irregular secondary structures –coil, different types of turns, omega loop, and beta hairpin.

Unit II: (5 Lectures)

Protein Design by Computational Study: De novo protein design, optimizing the protein sequence, validates computational predictions, applications of protein design-de novo interface design, scaffold design, ligand binding site design, and design of protein switches.

Unit III: (5 Lectures)

Classification of Protein 3D structure: Concept of 3D structure of protein; Algorithm-DALI, VAST, and CE. RMSD and Z-score for structural comparison, Database for structure based classification, CATH, SCOP.

Unit IV: (10 Lectures)
Computational Chemistry and Molecular Mechanics: Molecular mechanics, its advantage, disadvantage, and limitation. Force field- bond stretching, angle bending, cross-term, and non-bonded interaction, parameterization of force field. Different types of water model and its application. Wave–particle duality,Schrödinger wave equation,particle in a box model,simple harmonic oscillation, potential barrier,Expectation value, Aspect of computational chemistry, potential energy surface, single point energy calculation, geometry optimization, local minima, global minima, and saddle point, basic concept of energy landscape.

Unit V: (10 Lectures)
Molecular Dynamic Simulation of Biomolecules: Basic concept of MD-simulation, molecular interactions, configuration file of MD-simulation, equation of motion in MD simulation, finite difference method, integrating equation of motion, relationship between the solution of equation of motions and ensembles, solvation of protein, simulation with explicit solvent, constraints of MD-simulation, multiple time step dynamics, transport properties of MD simulation, multiple trajectory analysis.

Learning Resources:

1. Structural Bioinformatics 2008 Edition by Forbes J. Burkowski, publisher Taylor & Francis Ltd, ISBN: 9781584886839.

2. Structural Bioinformatics Methods And Protocols 2021 Edition by ZoltanGaspari, Publisher-Springer, ISBN: 9781071602720.

3. Structural Bioinformatics, 2nd Edition by GuJenne et.al, Publisher: Wiley India, ISBN: 9788126574711, 8126574712.

4. Understanding Molecular Simulation: From Algorithms to Applications (Computational Science Series, Vol 1) by DaanFrenkel. ISBN: 978-0122673511

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Genomics and Proteomics	Category: Theory
Subject Code: MSBIN 202	Semester: Second
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN202.1: Select different tools to submit and retrieve genomics and proteomics data from widely used public repositories.	I, II
2	MSBIN202.2: Use tools to perform functional annotation of lists of genes and proteins.	III, IV
3	MSBIN202.3: Use bioinformatics tools to analyse genomics and proteomics data, involving identification and quantification approaches.	V, VI
4	MSBIN202.4: Choose statistical tools for Genome and Proteome data analysis.	VII, VIII

Course Content:

Genomics:

Unit:I (5 lectures)
Genomics: Structural and Functional Genomics, Sequence based approach, EST and dbEST, SAGE analysis Microarray based approach, Software: Arrayplot, SNOMAD, Mutational genomics, Comparative genomics, Organization of Genome, Strategies of Genome Sequencing, Model Plant Genome Project, Functional Analysis of Genes.

Unit:II (5 lectures)
Human Genome Project: GENOME. FLOW CHART OF HUMAN GENOME. NUCLEAR GENOME. MITOCHONDRIAL GENOME. REPETITIVE DNA

Unit:III (5 lectures)
Single Nucleotide Polymorphism, Concept, Classification, Search Method, Assay Method, Clinical application, SNP mapping, SNP and cancer

Unit:IV (5 lectures)

The Human Genome Project, International HapMap Project, www.hapmap.org, The 1000 Genomes Project –www.1000.genomes.org, Encyclopedia of DNA Elements (ENCODE) Project, The Cancer Genome Atlas (TCGA), Human Microbiome Project (HMP) – www.hmpdacc.org, The eMERGE (Electronic Medical Records and Genomics) Network Common Features of Projects, High throughput, Use of technology, in particular – Automation (Robotics, AI) – Databases – Visualization, simulation/computational models – Groupware: Coordination and communication, Public domain tools, Open sharing of data.

Proteomics:

Unit V: (4 lectures)
Proteomics Data and Databases; UniProtKB, IntAct, Reactome, PRIDE, peptides Atlas,Software and open source tools; Reusing shared data, Application, challenges and opportunities of proteomics, Human Proteome Project; The human protein atlas

Unit VI: (7 lectures)
Protein identification and quantification; PeptideShaker; Post processing of proteomics data; Post-translational modifications identification; Biomarker candidate identification; Utilities of Galaxy server

Unit VII: (7 lectures)
Protein-Protein interaction data analysis; PPI databases (primary, secondary and predictive databases); Network analysis of PPI data (Cytoscape); proteins to pathways (Reactome); Proteogenomics: Database creation, Database search, Novel peptide analysis

Unit VIII: (2 lectures)
Differential expression; ProStaR; Statistical methods for expression analysis of proteomics data; FDR; PCA; Enrichment analyses

Learning Resources:

1. Principles of Genome Analysis and Genomics. Primrose SB &Twyman RM. 2007. Blackwell. ISBN-10: 1405101202

2. Introduction to Genomics. A.M Lesk, Oxford University press, 2007. ISBN-10: 0199557489

3. A Primer of Genome Science. Greg Gibson and Spencer V. Muse. 2nd ed. 2004. SINAUER Associates Inc. ISBN-10: 0878932364

4. Introduction to Proteomics: Tools for the New Biology. Daniel C. Liebler, Humana Press Inc., 2002. ISBN-10: 0896039919

5. Other research papers, review articles and web resources

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Molecular Modeling and Molecular Dynamics	Category: Theory
Subject Code: MSBIN 203	Semester: Second
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1.	MSBIN203.**1**: *Summarize* basic knowledge about the replica exchange technique in computational chemistry.	**Unit-I**
2.	MSBIN203.**2**: *Apply* the advance MD Simulation techniques to solve complex biochemical problems.	**Unit-II**
3.	MSBIN203.**3**:. *Produce* dynamic models by computational method and justify with experimental data.	**Unit-III**
4.	MSBIN203.**4**: *Describe* insights into biological systems using multi-scale engineering methods to analyze data in non-intuitive ways.	**Unit-IV**
5.	MSBIN203.**5**: *Explain* how computer simulations of biomolecules can be used to investigate biological problems at the electronic level.	**Unit-V**

Course Contents:

Unit I: **(8 Lectures)**
Basic Concept of Replica Exchange Molecular Dynamics Simulation: Replica Exchange Molecular Dynamics Simulation, Modified Replica Exchange MD Simulation, Hamiltonian Replica Exchange Molecular Dynamics, Multiplexed-replica exchange molecular dynamics, application of replica exchange MD Simulation.

Unit II: **(8 Lectures)**
Advance MD Simulation techniques and its application: Meta Dynamics, Constant-pH Molecular Dynamics Simulations, Temperature-accelerated MD Simulation and ab-initio MD Simulation. Basic idea of Inhomogeneous Solvation Theory (IST), Application of IST in WaterMap, structure-based ligand discovery and Grid Inhomogeneous Solvation Theory (GIST), GIST algorithm and its application in Molecular Dynamics Simulation, convergence of entropy and energy, end state analysis.

Unit III: **(8 Lectures)**
Validation of Simulation Results: Umbrella Sampling method, Steered Molecular Dynamics (SMD) Simulations, adaptive steered molecular dynamics (ASMD) and Nudged Elastic Band (NEB) simulations. Validation and verification of MD Simulation results and their correlation with experimental data.

Unit IV:**(8 Lectures)**
Multi scale Modelling of complex Biological System: Coarse-Grained Protein Models and their Applications: Brief History, Sampling Schemes, Examples of Protein Coarse-Grained Models, Applications of Coarse-Grained Models in Multiscale Modeling Pipelines, Multiscale Modeling Simulation: Example and Strategies, Reconstruction of Atomic Representation from Coarse-Grained Models.

Unit V: **(8 Lectures)**
Energy Landscape study and electron transfer mechanism of Bio-molecules: Free energy and potential energy landscape study of biomolecules and its application, Protein Dynamics in the Living Cytoplasm, Electron transfer in peptides and proteins: Electron transfer by super exchange, Electron hopping, Amino acid side chains as intermediates in electron hopping, Electron tunneling.

Learning Resources:

1. Multiscale Approaches to Protein Modeling by AndrzejKolinsski, ISBN: 9781441968883, 1441968881, Publisher: Springer-Verlag New York Inc.

2. Multiscale Modelling and Simulation by Sabine Attinger, Publisher: Springer-Verlag Berlin and Heidelberg GmbH & Co. KG., ISBN: 9783540211808, 9783540211808.

3. Innovations in BiomolecularModeling and Simulations (Volume 1) Editor: Tamar Schlick.Print ISBN:978-1-84973-461-5.

4. Theory and Practice in Replica-Exchange Molecular Dynamics Simulation Paperback – Import, 1 May 2012 by Cooke Isard, Benjamin Perry Publisher: Proquest, Umi Dissertation Publishing (1 May 2012), ISBN-13: 978-1248997017.

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Computer Language (Python)	Category: Theory
Subject Code: MSBIN 204	Semester: Second
L-T-P: 3-0-0	Credit:3

Course Outcomes:

By the end of this course students will be able to:

Sl No	Course Outcome	Mapped Unit
01	MSBIN204.1: Describe the basic concepts of python installation, operators and control structures programming syntax, and debugging.	UNIT I
02	MSBIN204.2: Apply several types of String operation, List, and Function used in programming.	UNIT II
03	MSBIN204.3: Describe and apply the concepts of File, Dictionaries, Tuples, and different useful Packages	UNIT III
04	MSBIN204.4: Apply python in the bioinformatics field.	UNIT IV

Course Contents:

UNIT I: 20L

Introduction to Python, installation of Python, character set, Constants, variables and keywords. Type of variables & constants. Rules of constructing variable identifier.

Types of Instructions (Type declaration, Arithmetic & Control Instructions), Data Types, Operators, Keywords, Hierarchy of operators, Associativity of operators, Type conversion (explicit and implicit), Different types of operators, Control Instructions: if-else, switch case, conditional operator. Loops (for, while), break & continue statement.

UNIT II:15L

String operations- Asking the user for input, Comments, String slices, String length, Strings are immutable, in operator, String comparison, String methods, Parsing strings, Format operator,

Function calls- Built-in functions, Type conversion functions, Math functions, adding new functions, Function Definitions and uses, Flow of execution, Parameters and arguments, Fruitful functions and void functions, Lists- list is a sequence, Lists are mutable, traversing a list, List operations, List slices, List methods, deleting elements, Lists and functions, Lists and strings, Objects and values, List arguments.

UNIT III: 10L

Dictionaries- Dictionary as a set of counters, Dictionaries and files, Looping and dictionaries.

Tuples- Tuples are immutable, comparing tuples, Tuple assignment, Dictionaries and tuples, Multiple assignment with dictionaries, Using tuples as keys in dictionaries.

Files- Persistence, opening files, Text files and lines, reading files, searching through a file, User choose the file name, using try-except-open, Writing files.

Unit IV: 5L

Application of python packages in Bioinformatics (Bio-data handling and application).

Learning Resources:

1. Introduction to Python Programming language- Chaitanya Singh
2. Python programming language- by G van Rossum
3. Python Programming: An Introduction to Computer Science, by John Zelle.

Subject Name: Structural Bioinformatics Lab	Category: Practical
Subject Code: MSBIN 291	Semester: Second
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:

MSBIN291.1: Analyze secondary structure and motif of protein.

MSBIN291.2: Apply different computational modelling tools and techniques for design protein.

MSBIN291.3: Describe and demonstrate the force field of biomolecules.

MSBIN291.4: Experiment with energy minimization and molecular dynamics methods in biomoleulces.

MSBIN291.5: Explain and apply Free energy calculation method for ligand.

Course Contents:

1. Visualization and analysis of Greek key, Jelly Roll motif, and Rossmann fold in protein. Homology and template-free modeling of protein structure. Modeling the loop region of the protein, analysis of different types of turns. Computational tools of scaffold design, ligand binding site design, and design of protein switches.
2. Analysis of force field of protein, nucleic acids, and ligands, computational visualizations of different water molecules.

3. Energy minimization of protein, protein-water complexes, and identification of local minima, global minima, and saddle point.
4. Molecular dynamics simulation of protein, protein-water, protein-nucleic acid, and protein-ligand complexes. Parameterization technique of new ligand. Water dynamics study.
5. Free energy calculation of liagnd using LIE and MM/PBSA method.

Subject Name: Genomics and Proteomics Lab	Category: Practical
Subject Code: MSBIN 293	Semester: Second
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:
MSBIN293.1: Apply knowledge to retrieve functional omics data, check quality of the data and processing the data.
MSBIN293.2: Identify the abundance and localization of RNA and proteins.
MSBIN293.3: Apply different pipelines and analyze network modeling for scientific use.
MSBIN293.4: Apply their bioinformatics knowledge and skills to critically assess the data characteristics, experimental design for efficient progress in high throughput biology.

Course Contents:

Genomics Lab:

1. Data format (FASTQ, BED, GTF, GFF3, SAM, BAM and VCF);
2. Quality control, FastQC, Trimmomatic
3. Genome assembly,
4. Read mapping TopHat, Bowtie, Picard, Cufflinks, CuffMerge
5. Read Count, HTSeq-count, RSEM
6. Differential expression analysis, edgeR, DESeq
7. Biological interpretation: DAVID, Reactome, Gene Ontology
8. Biological Network analysis using Cytoscape
9. Utilities of Galaxy Server and GEO2R

Proteomics Lab:

1. Protein identification and quantification; MaxQuant; SearchGUI and PeptideShaker;
2. Post-translational modifications identification
3. Utilities of Galaxy Server: ProteoRE
4. Protein-Protein interaction network analysis using Cytoscape
5. Etc.

Subject Name: Computer Language (Python) Lab	Category: Practical
Subject Code: MSBIN 294	Semester: Second
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:
MSBIN294.1: Describe the basic concepts of python installation, operators and control structures programming syntax, and debugging.
MSBIN294.2: Understand and apply several types of String operation, List, and Function used in programming.
MSBIN294.3: Describe and apply the concepts of File, Dictionaries, Tuples, and different useful Packages.
MSBIN294.4: Apply python in bioinformatics.

Course Contents:

1. Implement different types of operators, control instructions: if-else, switch case, conditional operator. loops (for, while), break & continue statement based problems using python language.
2. Implement different string operations and string methods using python language.
3. Implement different user defined function calls and its application using python language.
4. Implement different list operations, lists and functions, lists and strings etc. using python language.
5. Implement different Dictionaries and Tuples based problems using python language.
6. Implement different problems based on files and its application using python language.
7. Application of python packages in bioinformatics domain.

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Computational Drug Design	Category: Theory
Subject Code: MSBIN 301	Semester: Third
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN301.**1**: *Explain* the different techniques of structure based computational drug discovery methods.	**Unit-I**
2	MSBIN301.**2**: *Apply* advance concepts in Bioinformatics and Pharmacophores Modelling methods.	**Unit-II**
3	MSBIN301.**3**: *Summarize the* knowledge about the physicochemical Properties and the techniques involved in the computational QSAR method. Thereby quantification and statistical representation of it.	**Unit-III**
4	MSBIN301.**4**: *Explain* the knowledge about a drug candidate's safety, efficacy, and toxicity for regulatory approval.	**Unit-IV**
5	MSBIN301.**5**: *Outline* the concept of pharmacokinetic and pharmacodynamic profile of dug discovery process.	**Unit-V**

Course Contents:

Unit I: (8 Lectures)

Molecular Docking and Virtual Screening: Drug-likeness and Lipinnki's 'rule of five'. -Docking: Introduction, basic steps of docking, Types of docking-Rigid and flexible docking Search algorithm: Random searching, systematic searching, simulation based searching, Scoring/evaluatio functions- Force-Field Based scoring, Empirical Scoring Function, Knowledge-Based scoring function, Virtual screening: Ligand based virtual screening (LBSV)-Pharmacophore based screening.

Unit II: (8 Lectures)

Pharmacophores Modelling: Modeling and mapping of pharmacophore- constrained systemic search approach, Clique detection , pharmacophore generation: Ligand based pharmacophore generation, receptor based pharmacophore generation. Pharmacophore applications: Pharmacophore searching, Pharmacophore based de-novo design of ligands, pharmacophore based Lead optimization. Salt-bridge-complete and incomplete slat bridge in protein and its role in drug discovery.

Unit III: **(8 Lectures)**

QSAR: molecular descriptors: Based on 2D structure Physiochemical properties, Molar refractivity, topological indices, 2D fingerprints. Based on 3D structure3D fragment screens, Pharmacophore keys. Data verification and manipulation. QSAR models: Free Wilson analysis, Hansch constant, Hammett Substituent Constant 3D QSAR: common 3D molecular field, 3D-QSAR model- Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Index Analysis (CoMSIA)

Unit V: **(8 Lectures)**

In silico **ADMET**: Prediction of ADMET properties, Toxicity prediction, Introduction to combinatorial chemistry and Chemoinformatics, Combinatorial libraries: Diverse and focused libraries, Library enumeration, combinatorial library design strategies, Approaches to product based library design, multi-objective library design .

Unit VI: **(8 Lectures)**

Basic concept of drug discovery: Drug discovery methods and basic steps of drug designing process, drug target identification, fundamentals of pharmacokinetics, absorption- bio-availability, distribution-volume of distribution, Metabolism, Elimination-drug elimination kinetics, Drugs affinity and efficacy, Response of Drug on receptor: Agonist, antagonist.

Learning Resources

1. An introduction to chemoinformatics- Andrew R. Leach and Valerie J. Gillet (springer)
2. Molecular Modeling-Principles and Applications- Andrew R. Leach.
3. Bioinformatics-from Genomes to drugs- Thomas lengauer.
4. Fundamentals of Medicinal Chemistry by Gareth Thomas.

Subject Name: Systems Biology	Category: Theory
Subject Code: MSBIN 302	Semester: Third
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN302.1: Describe systems based approaches in biological sciences.	Unit I, V
2	MSBIN302.2: Select Systems Biology web-resources and tools that will help them in reconstructing and re-defining complex biological processes.	Unit II, III

3	MSBIN302.3: Choose an appropriate modeling technique to be used in a complex biological system.	IV, VI

Course Contents:

Unit I: Introduction to Systems Biology (6 lectures)

Basic concepts of Systems Biology, Top-down and bottom-up approaches for Systems Biology, Databases for Systems Biology, Multi-omics data integration, Robustness in biological systems, Potential limitations of systems biology.

Unit II: Fundamentals of biological networks (10 lectures)

Basics of network and its properties (node, edge, density, degree, shortest path, diameter, characteristic path length, degree distribution, directed and undirected network, hub), Network motifs, Feed Forward Loops, Auto-regulation, Network topologies (random, small-world, scale-free/power-law), Network perturbation, Cluster detection

Unit III: Reconstruction of biological networks (8 lectures)

Gene regulatory networks (GRN), Noise in GRN, Protein networks, Signalling networks, Metabolic network from genome information and its structural and functional analysis, Metabolic Flux Analysis, Biochemical reaction networks and stoichiometry

Unit IV: Dynamic nature of biological networks (3 lectures)

Mathematical models in biology, Dynamical modeling, Parameter Estimation, Simulation. Deterministic versus Stochastic approaches.

Unit V: Multiscale representation of cells and emerging phenotypes (5 lectures)

Spatio-temporal Systems Biology, The IUPS Physiome project, Cytomics - From cell states to predictive medicine, Synthetic Biology, Recent developments and trends of Systems Biology.

Unit VI: Tools for Systems Biology (8 lectures)

SBML, Cytoscape, Reactome, Ontologies in systems biology, E-cell, V-cell, ChEBI, Metaboanalyst, COBRA toolbox, CellDesigner, ReconMap, 3Omics

LearningResources:

1. Systems Biology PROPERTIES OF RECONSTRUCTED NETWORKS. By Bernhard Ø. Palsson. Published by Cambridge University Press, ISBN: 9780521859035
2. Uri Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits, 2/e, CRC Press, (2006)
3. Systems Biology: Applications and Perspectives. By Peter Bringmann. Published by Springer, 2007, ISBN 3540313389
4. Systems Biology: Principles, Methods and Concepts. By Andrzej K, Konopka. Published by CRC Press, 2007, ISBN 0824725204

5. Systems Biology: Definitions and Perspectives. By Lilia Alberghina, Hans V. Westerhoff. Published by Birkhauser, 2005, ISBN 354022968X
6. R.Heinrich and S. Schuster, The Regulation of Cellular Systems, Chapman & Hall, 1996.ISBN 978-1-4613-1161-4
7. E.O. Voit, Computational Analysis of Biochemical Systems. Cambridge University Press, 2000.
8. Other research papers, review articles and web resources

Subject Name: Research Methodology and IPR	Category: Theory
Subject Code: MSBIN 303	Semester: Third
L-T-P: 3-0-0	Credit:3

Course Outcomes: On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN303.1: Select research problem formulation.	**I**
2	MSBIN303.2: Explain research related information.	**II**
3	MSBIN303.3: Describe research ethics.	**III**
4	MSBIN303.4: Demonstrate the importance of computer in Bioinformatics	**IV**
5	MSBIN303.5: Review the patent rights and licensing process.	**V**
6	MSBIN303.6: Infer advanced IPR concept	**VI**

Course Contents:

Unit I: Meaning of research problem, Sources of research problem, Criteria Characteristics of a good research problem, Errors in selecting a research problem, Scope and objectives of research problem. Approaches of investigation of solutions for research problem, data collection, analysis, interpretation, Necessary instrumentations

Unit II: Effective literature studies approaches, analysis Plagiarism, Research ethics,

Unit III: Effective technical writing, how to write report, Paper Developing a Research Proposal, Format of research proposal, a presentation and assessment by a review committee

Unit IV: Nature of Intellectual Property: Patents, Designs, Trade and Copyright. Process of Patenting and Development: technological research, innovation, patenting, development. International Scenario: International cooperation on Intellectual Property. Procedure for grants of patents, Patenting under PCT.

Unit V: Patent Rights: Scope of Patent Rights. Licensing and transfer of technology. Patent information and databases. Geographical Indications.

Unit VI: New Developments in IPR: Administration of Patent System. New developments in IPR; IPR of Biological Systems, Computer Software etc. Traditional knowledge Case Studies, IPR and IITs.

Learning Resources:

Stuart Melville and Wayne Goddard, “Research methodology: an introduction for science & engineering students’”

Wayne Goddard and Stuart Melville, “Research Methodology: An Introduction”

Ranjit Kumar, 2 nd Edition, “Research Methodology: A Step by Step Guide for beginners”

Halbert, “Resisting Intellectual Property”, Taylor & Francis Ltd, 2007.

Mayall, “Industrial Design”, McGraw Hill, 1992.

Niebel, “Product Design”, McGraw Hill, 1974.

Asimov, “Introduction to Design”, Prentice Hall, 1962.

Introduction to Research, NPTEL online certification course.

Subject Name: Omics Technology	Category: Theory (Elective)
Subject Code: MSBIN 304(A)	Semester: Third
L-T-P: 3-0-0	Credit:3

Course Outcomes:
On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN304(A).1: Describe how to browse and retrieve functional omics data.	I
2	MSBIN304(A).2: Locate the information on the abundance and localisation of RNA (and proteins) in different species and biological conditions.	II, III
3	MSBIN304(A).3: Recognize and validate the plans towards annotation, analysis pipelines and network	IV

	modeling for scientific use.	
4	MSBIN304(A). 4: Extend the concepts of sequencing, transcriptomics and other ‘omics' methods applied and deduce logical interpretations from the data analysis.	V
5	MSBIN304(A).5: Apply their bioinformatics knowledge and skills to critically assess the data characteristics, experimental design for efficient progress in high throughput biology.	VI, VII

Course Contents:

Unit I: Genome, Genomics and Metagenomics (10 lectures)

Organization of genome structure, Variations in genome, Genome Molecular markers, First, second and third generation sequencing techniques; De novo and reference based Whole genome sequencing using Next Gen sequencing; Data format (FASTA, BED, GTF, GFF3, and VCF); Quality control, Genome assembly, Read mapping, BWA, SAM, BAM, Taxonomic investigation, Variant calling, Genome annotation; Variants of interest; Personal genomics; Genome-Wide Association Studies (GWAS); Metagenomics; Analysis of Metagenomic Data; QIIME, PICRUSt; Epigenomics; Genomics resources: TCGA, Ensembl, Genome Browser, The ENCODE project;

Unit II: Transcriptomics (10 lectures)

Quantitative transcriptomics: qRT-PCR; Delta-Delta Ct method; Advanced transcriptomics: gene expression microarrays, Normalization and experimental design models for microarray data; Next-generation sequencing in transcriptomics: RNA-seq experiments; Quality control of raw reads: FASTQC and Trimmomatic; Read alignment: TopHat, Bowtie, HISAT2; Quantification of expression (HTSeq-count) and differential expression analysis (edgeR, limma-voom and DESeq2); Isoform Discovery and Alternate Expression; Analysis of transcription factor binding sites; miRNomics; Functional genomics resources: ArrayExpress, Expression Atlas, Single Cell Expression Atlas

Unit III: Proteomics (8 lectures)

Protein identification and quantification; PeptideShaker; Post processing of proteomics data; Post-translational modifications identification; Biomarker candidate identification; Shotgun proteomics; Proteogenomics: Database creation, Database search, Novel peptide analysis; Proteomics resources: UniProtKB, IntAct, PRIDE, peptides Atlas, GPMDB and neXtProt; Software and open source tools

Unit IV: Interactomics and metabolomics (3 lectures)

Interactomics; Chromatin immunoprecipitation (ChIP), Chip-Seq data analysis; Protein-Protein interaction and network; Metabolomics

Unit V: Multiscale representation of cells and emerging phenotypes (2 lectures)

The IUPS Physiome project, Cytomics - From cell states to predictive medicine, Recent developments and trends of omics technologies.

Unit VI: EDA and Network Analysis of -omics Data (3 lectures)

Exploratory data analysis: tools and techniques; Basics of network and its properties (node, edge, density, degree, shortest path, diameter, degree distribution, directed and undirected network, hub etc.), Network motifs, Cluster detection

Unit VII: Bioinformatics Tools for Omics Technology (4 lectures)

Cytoscape, Reactome, Gene Ontologies, ChEBI, Metaboanalyst, CellDesigner, 3Omics

<u>Learning Resources:</u>

1. High Throughput Screening: Methods and Protocols (Methods in Molecular Biology) 29 Nov 2011 - William P. Janzen (Editor), Paul Bernasconi (Editor) ISBN-13: 978-1617794919.
2. RNA Bioinformatics (Methods in Molecular Biology) Hardcover – 22 Jan 2015 - Ernesto Picardi (Editor) ISBN-13: 978-1493922901.
3. Discovering Genomics, Proteomics and Bioinformatics, 2nd Edition. Campbell AM &Heyer LJ, Benjamin Cummings 2007; CSH Press, NY. ISBN-10: 8131715590
4. Principles of Proteomics. R.M Twyman (2004). (BIOS Scientific publishers). ISBN-10: 1859962734
5. Principles of Gene Manipulation and Genomics- Primrose S &Twyman R, 7th Edition, Blackwell, 2006. ISBN-10: 1405135441
6. Principles of Genome Analysis and Genomics. Primrose SB &Twyman RM. 2007. Blackwell. ISBN-10: 1405101202
7. Introduction to Genomics. A.M Lesk, Oxford University press, 2007. ISBN-10: 0199557489
8. A Primer of Genome Science. Greg Gibson and Spencer V. Muse. 2nd ed. 2004. SINAUER Associates Inc. ISBN-10: 0878932364
9. Introduction to Proteomics: Tools for the New Biology. Daniel C. Liebler, Humana Press Inc., 2002. ISBN-10: 0896039919
10. Other research papers, review articles and web resources

Maulana Abul Kalam Azad University of Technology, West Bengal
Master of Science in Bioinformatics Program: 2021-2022

Subject Name: Next Generation Sequence Technology	Category: Theory (Elective)
Subject Code: MSBIN 304(B)	Semester: Third
L-T-P: 3-0-0	Credit:3

Course Outcomes:
On completion of this course, students should be able to:

Sl No	Course Outcome	Mapped Unit
1	MSBIN304(B).1: Identify the different NGS technologies available in the market.	I
2	MSBIN304(B).2: Estimate good quality data.	II
3	MSBIN304(B).3: Explain the chemistry behind NGS applications.	III
4	MSBIN304(B).4: Evaluate the main parameters influencing the design of a sequencing project.	IV
5	MSBIN304(B).5: Apply the different pipelines in the analysis of NGS data and Infer the results obtained from the different NGS applications.	V

Course Contents:

Unit I (5L)
History of sequencing, Generations of sequencing, Principle and theories behind NGS technologies, Accessing public sequencing repositories, Galaxy server

Unit II (8L)
NGS data formats and tools, Unix command line, FASTQ, Readalignment+QC, FASTQC, Picard, Bowtie, Samtools, BEDTools, SNP/indel, Structural variation, Integrated Genomics Viewer (IGV)

Unit III (12L)
Next generation sequencing: Methods and applications, 454 sequencing, Illumina sequencing, Ion Torrent: Proton / PGM sequencing, ONT nanopore sequencing, Whole genome sequencing

Unit IV (10L)

RNA-seq experimental design, Comparison of different RNA sequencing platforms, Tophat, Cufflinks, Differential gene expression, Downstream analysis: pathway analysis, clustering and Gene ontology

Unit V (5L)

ChiP-seq analysis, MACS, functional annotation and detect potential binding sites (motif)

Subject Name: Computational Drug Design Lab	Category: Practical
Subject Code: MSBIN 391	Semester: Third
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:

MSBIN391.1: Experiment with molecular docking techniques between different biomolecules.
MSBIN391.2: Analyze the ligand conformation, binding energy, and scoring function using different molecular docking programs.
MSBIN391.3: Apply computational tools and Identify salt bridge interactions in protein.
MSBIN391.4: Experiment with pharmacophore modelling and Virtual screening of ligand.
MSBIN391.5: Apply computational tools for analyzing of QSAR properties, ADMET and toxicity test of ligands.
MSBIN391.6: Design ligand for structure based drug discovery.

Course Contents:

1. Preparation of receptor and ligand, molecular docking study-protein with ligand, protein with protein. protein with DNA, protein with RNA.
2. Comparative analysis of ligand conformation using different molecular docking programs.
3. Analysis of scoring function and validation of docking results with crystal data.
4. Binding energy calculation of ligand and correlation coefficient analysis between calculated and experimental data.
5. Identifications of salt bridge in protein using computational program and calculation their geometry.
6. Pharmacophore modelling, measuring the volume of ligand binding cavity in protein and Virtual screening from different chemical libraries and selection of specific ligand.
7. Computational tools for analysis of QSAR properties of ligands. Calculations of molecular descriptor using different computational tools.
8. Modification of ligand and structure based, water based, and salt-bridge based ligand design.
9. *In silico* analysis of ADMET and toxicity test of new identified ligands.

Subject Name: Omics Technology Lab	Category: Practical
Subject Code: MSBIN 394(A)	Semester: Third
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:
MSBIN394(A).1: Apply knowledge to browse and retrieve functional omics data.
MSBIN394(A).2: Identify the information on the abundance and localisation of RNA (and proteins) in different species and biological conditions.
MSBIN394(A).3: Recognize and validate the plans towards annotation, analysis pipelines and network modeling for scientific use.
MSBIN394(A).4: Extend the concepts of sequencing, transcriptomics and other ‘omics' methods applied and deduce logical interpretations from the data analysis.
MSBIN394(A).5: Apply their bioinformatics knowledge and skills to critically assess the data characteristics, experimental design for efficient progress in high throughput biology.

Course Contents:

1. Data format (FASTA, FASTQ, BED, GTF, GFF3, SAM, BAM and VCF);
2. Quality control, FastQC, Trimmomatic
3. Genome assembly
4. Read mapping,TopHat, Bowtie, Picard, Cufflinks, CuffMerge
5. Read Count, HTSeq-count, RSEM
6. Differential expression analysis, edgeR, DESeq
7. Biological interpretation: DAVID, Reactome, Gene Ontology
8. Biological Network analysis using Cytoscape
9. Utilities of Galaxy Server and GEO2R
10. Protein identification and quantification; MaxQuant; SearchGUI and PeptideShaker;
11. Post-translational modifications identification
12. Utilities of Galaxy Server: ProteoRE
13. Protein-Protein interaction network analysis using Cytoscape
14. Exploratory data analysis: Data visualization, Plots

Learning Resource:

Lesk, A. M. (2004). Introduction to Protein Science: Architecture, Function, and Genomics. Oxford: Oxford University Press.
2. Campbell, M & Heyer, L. J. (2006), Discovering Genomics, Proteomics and Bioinformatics, Pearson Education.
3. OMICS: Applications in Biomedical, Agricultural, and Environmental Sciences, by Debmalya Barh, Vasudeo Zambare, et al.
4. Principles of Genome Analysis and Genomics. Primrose SB & Twyman RM. 2007. Blackwell. ISBN-10: 1405101202

5. Introduction to Genomics. A.M Lesk, Oxford University press, 2007. ISBN-10: 0199557489
6. Discovering Genomics, Proteomics and Bioinformatics, 2nd Edition. Campbell AM & Heyer LJ, Benjamin Cummings 2007; CSH Press, NY. ISBN-10: 8131715590
7. Principles of Proteomics. R.M Twyman (2004). (BIOS Scientific publishers). ISBN-10: 1859962734

Subject Name: Computer Language (R) Lab	Category: Practical
Subject Code: MSBIN 395	Semester: Third
L-T-P: 0-0-6	Credit:3

Course Outcomes: On completion of this course, students should be able to:
MSBIN395.1: Apply basic syntax of R language to solve simple problems.
MSBIN395.2: Apply R language syntax to solve operators, data types and control structures related problems.
MSBIN395.3: Apply R language syntax to solve list, matrices, function related problems.
MSBIN395.4: Apply R packages to solve statistical problems and File concept related problems.
MSBIN395.5: Apply R language to visualize graphs.

Course Contents:

1. Construct and execute basic programs in R using elementary programming techniques, e.g. import/export of data from file or Internet, assign and manipulate data structures, create user-defined functions, loops, condition statements and debugging.

2. Use R for statistical calculations.

3. Implement and describe Monte Carlo techniques as well as perform simulation studies with analysis and evaluation of result

4. Graphically visualize data and results of statistical calculations

5. Use external R-packages in statistics and data mining

6. Application of R packages in bioinformatics domain.

Learning Resources:

1. Data Analysis with R: A comprehensive guide to manipulating, analyzing, and visualizing data in R, 2nd Edition, by Anthony Fischetti

2. Data Analysis Using R: A Primer for Data Scientist, by Dr. Dhaval Maheta

3. Hands-On Exploratory Data Analysis with R: Become an expert in exploratory data analysis using R packages, by Radhika Datar and Harish Garg.

Subject Name: Computational Drug Design Lab	**Category:** Practical
Subject Code: MBSIN 391	**Semester:** Third
L-T-P: 0-0-6	**Credit:** 3

Course Outcomes:

MBSIN391.1: *Compute the* molecular docking techniques by different biomolecules.

MBSIN391.2: *Analyze* the ligand conformation, binding energy, and scoring function using different molecular docking programs.

MBSIN391.3: *Apply* computational tools and Identify salt bridge interactions in protein.

MBSIN391.4: *Employ* pharmacophore modelling and Virtual screening of ligand.

MBSIN391.5: *Apply* computational tools for analyzing of QSAR properties, ADMET and toxicity test of ligands.

MBSIN391.6: *Design* ligand for structure based drug discovery.

Course Contents:

Preparation of receptor and ligand, molecular docking study-protein with ligand, protein with protein. protein with DNA, protein with RNA.

Comparative analysis of ligand conformation using different molecular docking programs.

Analysis of scoring function and validation of docking results with crystal data.

Binding energy calculation of ligand and correlation coefficient analysis between calculated and experimental data.

Identifications of salt bridge in protein using computational program and calculation their geometry.

Pharmacophore modelling, measuring the volume of ligand binding cavity in protein and Virtual screening from different chemical libraries and selection of specific ligand.

Computational tools for analysis of QSAR properties of ligands. Calculations of molecular descriptor using different computational tools.

Modification of ligand and structure based, water based, and salt-bridge based ligand design.

In silico analysis of ADMET and toxicity test of new identified ligands.

Subject Name: Project Work	Category: Sessional
Subject Code: MSBIN 481	Semester: Fourth
	Credit: 15

Course Outcomes

On successful completion of the course, students will able to

MSBIN481.1: *Identify* issues that must be addressed within framework of specific thesi*s*.

MSBIN481.2: *Plan* research design

MSBIN481.3: *Critically* evaluate different technical solutions

MSBIN481.4:Analyze techniques/experimental methods.

MSBIN481.5: *Develop* project management skills.

MSBIN481.*6:Build* report writing skills

Course Contents: A project dissertation would be of one semester duration and one project would be allotted to one student. The topic of the research shall be in the domain of **advance area of bioinformatics** to understand real life problem. Final presentation and viva-voce of the project will be based on the project thesis submission. The evaluation of the thesis will be done by the panel of examiners.

Syllabus: Project Proposal: Preparation Selection of research lab and research topic: Students should first select a lab wherein they would like to pursue their dissertation. The supervisor or senior researchers should be able to help the students to read papers in the areas of interest of the lab and help them select a topic for their project. The topic of the research should be hypothesis driven.

Review of literature: Students should engage in systematic and critical review of appropriate and relevant information sources and appropriately apply qualitative and/or quantitative evaluation processes to original data; keeping in mind ethical standards of conduct in the collection and evaluation of data and other resources. Writing Research Proposal: With the help of the senior researchers, students should be able to discuss the research questions, goals, approach, methodology, data collection, etc. Students should be able to construct a logical outline for the project including analysis steps and expected outcomes and prepare a complete proposal in scientific proposal format for dissertation.

Writing Research Proposal: With the help of the senior researchers, students should be able to discuss the research questions, goals, approach, methodology, data collection, etc. Students should be able to construct a logical outline for the project including analysis steps and expected outcomes and prepare a complete proposal in scientific proposal format for dissertation.

Syllabus: Thesis writing At the end of their project, thesis has to be written giving all the details such as aim, methodology, results, discussion and future work related to their project. Students may aim to get their research findings published in a peer-reviewed journal. If the research findings have application-oriented outcomes, the students may file patent application.

Subject Name: Grant Viva	Category: Sessional
Subject Code: MSBIN 482	Semester: Fourth
	Credit: 03

Course Outcomes

At the end of this course, students will be able to:

MSBIN482.**1**. *Develop* their understanding of different subjects learnt in previous semesters.

MSBIN482.**2**. *Apply* fundamental concepts which they learn in different subjects.

MSBIN482. **3**.*Develop* their interview facing skills.

Course Contents:

This course is designed to test student's learning and understanding of different subjects which they learned across their Bioinformatics course. The key objective of this course is to prepare students to face interviews both technical/non-technical and academic/industrial. This course will also help students in comprehending their Bioinformatics knowledge and refreshing fundamental concepts of different subjects.

Subject Name: Communication Skill	Category: Sessional
Subject Code: MSBIN 483	Semester: Fourth
	Credit: 03

Subject Name: Communication Skill
Subject Code: MSBIN 483
Course Outcomes

At the end of this course, students will be able to:

MSBIN483.1: Demonstrate critical and innovative scientific thinking.

MSBIN483.2: Display competence in oral, written, and visual scientific communication.

MSBIN483.3: Apply Bioinformatics theories.

MSBIN483.4: Show an understanding of opportunities in the field of Bioinformatics.

MSBIN483.5: Use current technology related to the Bioinformatics field.

MSBIN483.6: Communicate ethically.

Course Contents:

Presentations of Research work in the area of Bioinforrmatics and communicate to all students in the form of seminars.