DEPARTMENT OF BIOTECHNOLOGY & DEPARTMENT OF CHEMISTRY

M.Sc. Program in Computational Molecular Sciences

Eligibility: Bachelors' degree in any Science, Pharmacy and Engineering (with physics, chemistry, biology)

Program Structure (Applicable to 2025 admission onwards)

YEAR	FIRST SEMESTER						SECOND SEMESTER					
	SUB CODE	**SUBJECT NAME**	**L**	**T**	**P**	**C**	**SUB CODE**	**SUBJECT NAME**	**L**	**T**	**P**	**C**
	BIO 5115	Biochemistry and Medicinal Chemistry	4	0	0	4	BIO 5213	Computer-aided Molecular Design	4	0	0	4
	CHM 5115	Computational Chemistry	4	0	0	4	BIO 5214	Multiscale Modelling and Simulations	4	0	0	4
I	BIO 5116	Cheminformatics	4	0	0	4	CHM 5216	Computational Design of Materials	4	0	0	4
	BIO 5117	Computational Biology [FYZ]	4	0	0	4	BIO 5215	AI/ML in Chemical Biology	4	0	0	4
	BIO 5143	Lab-I: Biomolecular Modelling [NSG]	0	0	4	2	BIO 5248	Lab-III: Molecular Docking & Dynamics	0	0	4	2
	BIO 5144	Lab-II: Programming for Beginners	0	0	4	2	BIO 5249	Lab-IV: AI/ML in molecular sciences	0	0	4	2
						20						**20**
	Total											**40**
	THIRD AND FOURTH SEMESTER											
III		Third Sem (MOOC's): RMTC (**3** credits), Intellectual property rights (**2** credits) Electives-I, II, III (**3** credit each) -----> related to the project (**9** credits) Seminar -**2** Credits										**16**
IV	BIO 6092	Project Work (including industry and collaborative projects; Scopus indexed publication)										**24**
		Total Credits							**0**	**0**	**0**	**80**

SEMESTER – I

BIOCHEMISTRY AND MEDICINAL CHEMISTRY [4 0 0 4]

Cell structure and function, Biomolecules-water, carbohydrates, lipids, proteins, nucleic acids, chemistry of cell membranes and permeability. Classification of carbohydrates, stereoisomers, epimers, etc., biological importance of monosaccharides, disaccharides and polysaccharides, Inter conversion of sugars, Carbohydrates as drug targets. Food and drug metabolism, Glycolytic pathway, Citric acid cycle, Gluconeogenesis, Xenobiotic metabolism: phase I and phase II drug metabolism, Classification of lipids, simple lipids, compound lipids, derived lipids, Lipids as drug targets. Classification of aminoacids, Structure and properties of aminoacids, Classification of proteins; proteins as drug targets; Monoclonal antibodies, peptide drugs; enzymes as drug targets; structural proteins as drug targets; Receptors as drug targets; nucleic acid as drug targets. Drug classification, Agonists and Antagonists, prodrugs, drug transport, receptors, ion-channels, and carrier proteins; pharmacodynamics and pharmacokinetics; Approaches in ligand design and discovery.

References:

1. Nelson, D. L., Cox, M. M., Hoskins, A. A. (2021). Lehninger Principles of Biochemistry. United Kingdom: Macmillan Learning.
2. Voet, D., Pratt, C. W., Voet, J. G. (2018). Voet's Principles of Biochemistry Singapore: Wiley.
3. Crowe, J., Bradshaw, T. (2021). Chemistry for the Biosciences: The Essential Concepts. United Kingdom: Oxford University Press.
4. Patrick, G. L. (2023). An Introduction to Medicinal Chemistry. United Kingdom: Oxford University Press.

COMPUTATIONAL CHEMISTRY [4 0 0 4]

Quantum Mechanics Foundations, Schrödinger Equation, Quantum Mechanical Operators, Postulates of Quantum Mechanics, Wave Functions, Ab Initio Methods (Hartree-Fock, DFT), Semi-empirical Methods,, Basis Sets, Molecular Orbital Theory, Potential Energy Surface (PES), Geometry Optimization, Optimization Algorithms, Time-Dependent Density Functional Theory (TD-DFT), Computational Chemistry Software (ORCA), Calculation Setup, Data Analysis.

References:

1. Leach, AR., Molecular Modelling: Principles And Applications, 2/E. (2009). India: Pearson Education.
2. Frenkel, D., Smit, B. (2023). Understanding Molecular Simulation: From Algorithms to Applications. Netherlands: Elsevier Science.
3. Jensen, J. H. (2017). Molecular Modeling Basics. United Kingdom: CRC Press LLC.
4. Szabo, A., & Ostlund, N. S. (2012). Modern quantum chemistry: introduction to advanced electronic structure theory. Courier Corporation.
5. Pauling, L., Wilson, E. B. (2012). Introduction to Quantum Mechanics with Applications to Chemistry. United States: Dover Publications.

CHEMINFORMATICS [4 0 0 4]

Chemical Databases: Data curation, relational chemical databases, handling Markush structures, processing SMILES, InChI, and fingerprints. Library Design: Strategies for generating chemical libraries, partitioning, diverse subset selection, and focused compound library construction from chemical fragments. Libraries: ChEMBL, PubChem, DrugBank, BindingDB. QSAR/QSPR: Molecular descriptors, data analysis, classification methods, data modeling, model validation, regulatory applications of QSARs. Computational toxicology analysis (OPERA, TEST), Nanoinformatics and statistical modelling; nanodescriptors; unsupervised techniques, computational models/tools/software packages to compute/model activity and toxicity, Pharmacophore Modeling: Generation and application of pharmacophore models for virtual screening, case studies. Linear Regression Analysis, Logistic Regression, Bayes Classified Clustering Algorithm, Query and Visualization Techniques, Machine Learning and Statistical Techniques, Chemical Search and Filtering: RDKit-based chemical search, multilogic filtering, common substructure and scaffold analysis, chemical similarity metrics (Tanimoto coefficient, bioisoster similarity). Molecular Graphs and Visualization.

References:

1. Leach, A. R., Gillet, V. (2007). An Introduction to Chemoinformatics. Netherlands: Springer Netherlands.
2. Roy, K. (Ed.). (2023). Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development. Academic Press.
3. Oprea, T. I. (2004). Chemoinformatics in drug discovery. Wiley-VCH.
4. Varnek, A. (Ed.). (2017). Tutorials in chemoinformatics. John Wiley & Sons, Inc.

COMPUTATIONAL BIOLOGY [4 0 0 4]

Algorithms and complexity, Algorithm Design Techniques, Data Storage, Relational Databases, Biological Data Types, Biological Data Mining, Machine learning Methods, Identification of Functional Sites in Sequences, Pattern Matching, Gene and Domain Prediction, Next Generation Sequencing: Whole Genome, Transcriptome, and Metagenome Analysis, Genome Rearrangement, Genome Assembly, GWAS, Biomarkers Identification from Genome Proteome & Metabolome, Enrichment Analysis, Graph Theory, Biological Networks, Protein Secondary and Tertiary Structure Prediction, RNA Structure Prediction, Structure Similarity Search.

References:

1. Neil Jones & Pavel Pevzner (2004), “Introduction to Bioinformatics Algorithms”, MIT Press.
2. Heitor Silvério Lopes & Leonardo Magalhães Cruz (2011), “Computational Biology and Applied Bioinformatics”, InTech.
3. Dongqing Wei, Qin Xu, Tangzhen Zhao, Hao Dai, Bryan Bergeron (2014), “Advance in Structural Bioinformatics”, Springer.
4. David W Mount (2001), “BIOINFORMATICS: Sequence and Genome Analysis”, Cold Spring Harbor.
5. Genome-Wide Association Studies: From Polymorphism to Personalized Medicine. (2016). United Kingdom: Cambridge University Press
6. Stram, D. O. (2013). Design, Analysis, and Interpretation of Genome-Wide Association Scans. Netherlands: Springer New York.

Lab I **BIOMOLECULAR MODELLING [0 0 4 2]**

Sequence Analysis using BioEdit; Sequence-Structure Correlation; Correlating sequence conservation with structural features (e.g., conserved motifs and binding sites). Sequence analysis and alignment.

Molecular interaction visualization; predicting interaction sites (protein-ligand, protein-protein); Techniques for analyzing binding pockets and interaction surfaces; advanced Chimera Command-Line Operations; distance measurements, angle measurements, and hydrogen bond detection. Rendering and saving high-quality images for publications.

Protein structure visualization; UCSF Chimera; Homology modeling (template-based prediction); AlphaFold Models; Analyze the quality of the model using structural assessment tools (e.g., Ramachandran plots, RMSD comparisons). Applications in drug discovery, functional genomics, and molecular recognition, Real-world examples of structure prediction use cases.

Software/Resources:
UCSF Chimera/ChimeraX (latest version).
BioEdit (latest version).
Online Databases.
PDB files, pdb-tools and open-babel
Computers with both software installed.

References:

1. Bladon, P., Gorton, J. E., Hammond, R. B. (2012). Molecular Modelling: Computational Chemistry Demystified. United Kingdom: RSC Publishing.
2. Burkowski, F. J. (2014). Computational and Visualization Techniques for Structural Bioinformatics Using Chimera. United Kingdom: CRC Press, Taylor & Francis Group.
3. Jensen, J. H. (2017). Molecular Modeling Basics. United Kingdom: CRC Press LLC.

Lab II **PROGRAMMING TOOLS FOR BEGINNERS [0 0 4 2]**

Foundational concepts in Unix/Linux, High-Performance Computing (HPC), Python, and R scripting. The Unix/Linux module covers basics like file management, directory handling, file permissions, and environment settings. Students will learn essential commands (cp, mv, cd, ls, rm, cat, pwd) along with advanced utilities such as grep, sed, and awk. Key topics include shell programming, conditional statements, loops, and process handling. The HPC section introduces concepts of parallel computing, job scheduling (PBS, SLURM), and managing tasks on computing clusters. Python programming covers keywords and identifiers, data types (lists, tuples, dictionaries), I/O operations, operators, control structures (if-else, for, while loops), file handling, and exception handling. Students will also explore functions, recursion, modules, and packages, along with an introduction to object-oriented programming (OOP) concepts like classes, inheritance, and method overriding. The R scripting module emphasizes data types, data manipulation, statistical functions, and creating visualizations.

References:

1. Tansley, D. S. W. (2000). Linux and UNIX Shell Programming. Switzerland: Addison-Wesley.
2. David M, Beazley DM (2009) Python Essential Reference (4th Edition). Pearson Addison-Wesley Professional
3. Lutz M (2009) Learning Python (5th Edition). O'Reilly
4. Hager, G. (2010). Introduction to High Performance Computing for Scientists and Engineers. CRC Press.
5. Keller, S. (2016). R Programming for Beginners: Fast and Easy Learning R. United States: CreateSpace Independent Publishing Platform.

SEMESTER – II

COMPUTER-AIDED MOLECULAR DESIGN [4 0 0 4]

Drug Discovery Process, Target Identification and Validation, Lead Identification and Optimization, Pre-clinical and Clinical Testing, biophysical characterization techniques; Structure and Ligand-Based Drug Design, Protein Structure Prediction, Docking, Scoring Algorithms, De Novo Drug Design, Virtual Screening, Pharmacophore Modeling, Computational Tools in Structure-Based Drug Design, Targeting Protein-Protein Interactions, Free Energy Calculations (MM-GBSA, FEP, LIE), Solvation Energies, Computational Enzyme Design, Enzyme Catalysis Mechanisms, Enzyme Kinetics, Rational Design, Directed Evolution, Emerging Trends in Computer-Aided Drug Design (CADD), Recent trends in immunoinformatics and glycoinformatics; Ethical and Regulatory Issues in Drug Design.

References:

1. Canales, A. (Ed.). (2017). Biophysical techniques in drug discovery. Royal Society of chemistry.
2. Young, D. C. (2009). Computational Drug Design: A Guide for Computational and Medicinal Chemists. Germany: Wiley.
3. Sehgal, S. A. (2021). Quick Guideline for Computational Drug Design (Revised Edition). Singapore: Bentham Science Publishers.
4. Ghosh, A. K., Gemma, S. (2014). Structure-based Design of Drugs and Other Bioactive Molecules: Tools and Strategies. Germany: Wiley.
5. Reetz, M. T., Sun, Z., Qu, G. (2023). Enzyme Engineering: Selective Catalysts for Applications in Biotechnology, Organic Chemistry, and Life Science. Germany: Wiley.
6. Lee, S. Y., Nielsen, J., & Stephanopoulos, G. (2021). Protein engineering: tools and applications. John Wiley & Sons.

MULTISCALE MODELLING AND SIMULATIONS [4 0 0 4]

Molecular Mechanics: Historical development, forces between molecules, molecular mechanics force fields (CHARMM, AMBER, GROMOS), potential energy surfaces, parameterization, and conformational analysis. Statistical mechanics concepts: microstates, partition functions, thermodynamic properties.

Molecular Dynamics (MD) and Monte Carlo Simulations: Basics of MD, force fields, integration algorithms (Verlet, leapfrog), non-bonded interactions, temperature and pressure control, Monte Carlo methods, periodic and non-periodic simulations, and MD trajectory analysis. Advanced MD Techniques: Comparison of MD and Monte Carlo, periodic boundary conditions, replica exchange MD, accelerated MD, and targeted MD. Coarse-Grained MD Simulations: Coarse-graining strategies, CG force fields (Martini, UNRES), and applications in membrane and polymer modelling.

References:

1. Leach, A., Molecular Modelling: Principles And Applications, 2/E. (2009). India: Pearson Education.

2. Frenkel, D., Smit, B. (2023). Understanding Molecular Simulation: From Algorithms to Applications. Netherlands: Elsevier Science.
3. Hinchliffe, A. (2008). Molecular Modelling for Beginners. United Kingdom: Wiley.
4. Li, G. (2024). Chemical Theory and Multiscale Simulation in Biomolecules: From Principles to Case Studies. Netherlands: Elsevier.
5. Alavi, S. (2020). Molecular Simulations: Fundamentals and Practice. Germany: Wiley.
6. Zhou, K., Liu, B. (2022). Molecular Dynamics Simulation: Fundamentals and Applications. Netherlands: Elsevier Science.
7. Kholmurodov, K. (2007). Molecular Simulation Studies in Material and Biological Sciences. United States: Nova Science Publishers.
8. Editors: Chipot C. , Christophe A. et al., (2007) Free Energy Calculations, Springer Series in Chemical Physics, Vol. 86

COMPUTATIONAL DESIGN OF MATERIALS [4 0 0 4]

Foundations of Materials Modelling: Importance of materials modelling, atomic and continuum modelling techniques, crystallographic structures, unit cells, Bravais lattices, reciprocal space, diffraction, and interfaces (grain boundaries, surfaces). Advanced Crystallography and Microstructure: Stereographic projections, plastic deformation mechanisms, texture evolution, and martensitic transformations. Thermodynamics and Phase Transitions: Thermodynamic principles, free energy, phase diagrams, and kinetic Monte Carlo simulations for diffusion and microstructure evolution. Non-equilibrium Molecular Dynamics Methods: Classical mechanics, force fields, non-equilibrium simulations, constant temperature/pressure ensembles, and microstructure analysis. Software Tools for Materials Modelling: Open-source tools like LAMMPS, VASP, GROMACS, Quantum Espresso for electronic structure calculations, and deMon2K for quantum chemistry applications.

References:

1. Zhou, K., Liu, B. (2022). Molecular Dynamics Simulation: Fundamentals and Applications. Netherlands: Elsevier Science.
2. Nanomaterials: Design and Simulation. (2007). Netherlands: Elsevier Science.
3. Westmoreland, P. R. (2013). Applying Molecular and Materials Modeling. Netherlands: Springer Netherlands.
4. Pal, S., Ray, B. C. (2020). Molecular Dynamics Simulation of Nanostructured Materials: An Understanding of Mechanical Behavior. United States: CRC Press.
5. Simulations in Nanobiotechnology. (2012). United Kingdom: Taylor & Francis.

AI/ML IN CHEMICAL BIOLOGY [4 0 0 4]

Introduction to AI and ML: Basic concepts of Artificial Intelligence (AI) and Machine Learning (ML), their roles in advancing research in chemistry and biology, fundamental algorithms (supervised, unsupervised, semi-supervised, reinforcement learning), and generative AI. Databases like Kaggle. Data Acquisition and Preprocessing: Methods for acquiring and preprocessing chemical and biological data, including data cleaning, normalization, feature engineering, and working with molecular fingerprints and descriptors. Feature Selection and Model Evaluation: Techniques for feature selection, performance evaluation metrics for models. Theory and Applications to Molecular Systems: Inverse molecular design, molecular property prediction, molecular dynamics (MD) simulations, materials discovery and design, and applications of deep learning in chemistry (Deep Chem).

References:

1. Dutt, S., Chandramouli, S., & Das, A. K. (2019). Machine learning, Pearson India.
2. Karthikeyan, A., Priyakumar, U.D. Artificial intelligence: machine learning for chemical sciences. J Chem Sci 134, 2 (2022). https://doi.org/10.1007/s12039-021-01995-2
3. Ramsundar, B., Eastman, P., Walters, P., Pande, V. (2019). Deep Learning for the Life Sciences: Applying Deep Learning to Genomics, Microscopy, Drug Discovery, and More. United States: O'Reilly Media.
4. Artificial Intelligence and Machine Learning in Drug Design and Development. (2024). United Kingdom: Wiley.
5. Maasch, J. R. M. A., Fuente Nunez, C. d. l. (2022). Machine Learning for Drug Discovery. United States: American Chemical Society.
6. Cartwright, H., Machine Learning in Chemistry: The Impact of Artificial Intelligence. (2020). United Kingdom: Royal Society of Chemistry.

Lab-III: <u>MOLECULAR DOCKING AND DYNAMICS [0 0 4 2]</u>

Molecular Docking using AutoDock Vina and Colab: Basics of molecular docking, calculating receptor-ligand interactions and binding affinity, Applications of docking in drug discovery. Overview of AutoDock Vina and hands-on for docking simulations. Using Google Colab for cloud-based docking. File formats: PDB, PDBQT, ligand and receptor preparation. Visualizing docking poses using software like PyMOL or Chimera. Comparing docking scores and identifying the best binding mode. Discussion on the biological significance of docking results.

Molecular Dynamics Simulations using GROMACS: Overview of GROMACS for MD simulations. File formats: PDB, GRO, and topology files. Setting up and running MD simulations. Using GROMACS tutorials: System setup and solvation of a protein in a water box, Energy minimization and equilibration, Performing a short production run. Visualizing MD trajectories using VMD/UCSF Chimera or PyMOL. Calculating key parameters: root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration. Interpretation of MD simulation results to understand system stability and dynamics.

Hands on training with Charmm-GUI server: Building different systems including nanomaterials

Software/Resources:
VMD/Pymol/UCSF Chimera.
GROMACS.
PDB files and pdb-tools
Online servers for input processing

References:

1. Molecular Docking for Computer-Aided Drug Design: Fundamentals, Techniques, Resources and Applications. (2021). Netherlands: Academic Press.
2. Li, G. (2024). Chemical Theory and Multiscale Simulation in Biomolecules: From Principles to Case Studies. Netherlands: Elsevier.
3. Huey, R., Morris, G. M., & Forli, S. (2012). Using AutoDock 4 and AutoDock vina with AutoDockTools: a tutorial. The Scripps Research Institute Molecular Graphics Laboratory, 10550(92037), 1000.

Lab-IV: **AI/ML IN MOLECULAR SCIENCES [0 0 4 2]**

Hands on Google Colab and Jupyter Notebooks for AI/ML: Setting up Google Colab for cloud-based AI/ML computations, Using Jupyter Notebooks for data analysis and scripting, Installation of essential libraries: scikit-learn, TensorFlow, rdkit.

Hands-on Tutorial: Drug Design with RDKit: Using RDKit for cheminformatics and molecular manipulation, Generating molecular descriptors and fingerprints using RDKit, Building a simple QSAR model for drug activity prediction with scikit-learn.Train a machine learning model (e.g., Random Forest, Support Vector Machine) to predict drug-like properties.

Building and Validating AI/ML Models for Drug Design: Model training and validation: train-test split, cross-validation, Performance evaluation: metrics like accuracy, precision, recall, ROC curves, Application of models to screen a virtual drug library for potential hits.

Advanced Tools: ChemProp, DeepChem, Aspirin Resources, Integrating ML potentials with molecular docking and dynamics simulations. Using pre-trained models for force field predictions in molecular simulations, retrosynthesis and reaction prediction.

Software/Resources:
VMD/Pymol/UCSF Chimera.
Google Colab
Jupyter notebooks

References:

1. Alkhalifa, S. (n.d.). Machine Learning in Biotechnology and Life Sciences: Build Machine Learning Models Using Python and Deploy Them on the Cloud. Germany: Packt Publishing.
2. Ramsundar, B., Eastman, P., Walters, P., Pande, V. (2019). Deep Learning for the Life Sciences: Applying Deep Learning to Genomics, Microscopy, Drug Discovery, and More. United States: O'Reilly Media.