Interdisciplinary M. Tech. Programme in Bioinformatics

February, 2005

A Committee consisting of (i) Prof. S. N. Maheshwari (Chairman); (ii) Prof. B. Jayaram (Convenor), (iii) Prof. M. N. Gupta, (iv) Prof. Saroj Mishra, (v) Prof. Sunil Nath, (vi) Prof. Jayadeva and (vii) Prof. R. Ramaswami (JNU) was constituted by the Institute to examine the feasibility of starting academic programmes in Bioinformatics at IITD with the objective of contributing to the national requirements for human resource development in Bioinformatics. The committee has met several times during the past one and a half years. Also, the committee benefited from the inputs received from the members of the Bioinformatics Thrust Area Working Group (comprising Prof. M. N. Gupta, Prof. Saroj Mishra, Prof. G. P. Agarwal, Prof. S. K. Koul, Prof. Sneh Anand, Prof. Suresh Chandra, Prof. B. Chandra, Prof. Baisya, Prof. Sunil Nath, Prof. Jayadeva, Dr. Charusita Chakravarty, Dr. Lipika Dey and Prof. B. Jayaram (Coordinator)). Taking into account the existence of research and teaching activity in areas relating to Bioinformatics in many Departments and Centres at the Institute the Committee feels that it is feasible to initiate an Interdisciplinary M.Tech. programme in Bioinformatics at the Institute. The following is a detailed report.

The Committee has deliberated on the academic contents of such a programme, and the deliberations have resulted in a document, appended as annexure, that contains some salient features of the programme semester-wise. The Committee has been influenced in its deliberations by the following considerations.

Bioinformatics has a strong interdisciplinary character. It can be considered to be a confluence of Biology, Computer Science, Information Technology, Mathematics, Chemistry, Physics, and Medicine with the objectives of developing tools to analyze biological, biochemical, biophysical data and to generate new knowledge in these areas. The ideal person for research in this area will be one who is well versed in techniques of modern biology and biochemistry, has strong skills in combinatorial computing and search and retrieval techniques in large biological databases, and understands statistical and probabilistic methods so as to be able to build viable and analyzable models of complex biological phenomena. It is also a fact that persons trained and skilled in these

multifarious ways do not exist, and if this area is to develop in our country these persons will have to be trained and produced.

The Committee recognizes that, in principle, any person trained in physical sciences, or biological sciences, or computing and mathematical sciences is well suited for training in Bioinformatics. However, academic programmes in these areas in our country tend to be narrow and very focused. Students know very little other than the intricacies of their specialization. Keeping in view the limited duration and the credit structure of M.Tech. programmes and the need to keep the preparatory part of the programme small and manageable, Committee is of the view that in the beginning input to the programme be limited to those who have strong biological sciences or strong computing sciences background. This will limit the task of taking care of deficiencies to teaching mathematics and computing to biological sciences students and modern biology and biochemistry to computing sciences students. Beyond taking care of these deficiencies both sets of students will require sound training in statistical and probabilistic methods, information organization and retrieval techniques and computational biology and bioinformatics to enable them to follow and contribute at the frontiers of their chosen area of research during their masters project and beyond.

The programme that is detailed out in the Annexure-1 follows the above philosophy. First semester academic consists of six credits of remedial and preparatory coursework for students of the two streams. It should be noticed that all the students will also undergo training in laboratory techniques in biological data generation via a Wet Lab. in Molecular Biology, and at the same time will also get hands-on experience of the current state of Bioinformatics tools and their usage through a heavily laboratory oriented course on usage of Bioinformatics and data access and mining tools.

The focus of the second semester course work is on building up core skills in computing, computational biology, information storage and retrieval techniques and in providing opportunity to students to start specializing through the process of choosing three elective courses.

The Committee has identified several possible electives from the existing courses being offered by the various Departments in the Institute and has also suggested several new elective courses that could be offered by the participating Departments and Centres.

These new electives reflect both the needs of the programme as well as those that, in the Committee's point of view, are in line with the research being carried out by faculty in the Institute. It should be pointed out that these lists are tentative only and serve at this time the primary purpose of establishing that enough is happening in the Institute for running a viable academic programme at the M.Tech. level. Annexure-2 contains the suggested syllabii for the new courses (core and elective) that will require to be offered for the programme.

The Committee also feels that the academic content can be further strengthened by putting in place protocols that enable students to take for credit courses being offered at nearby sister Institutes like the Jawaharlal Nehru University. The Committee will like to point out that Jawaharlal Nehru University has been running a one year post-graduate programme in Bioinformatics for sometime now. Such a protocol has the added benefit that such interaction can be the foundation of putting in place collaborations in research and project activities at the inter-Institute levels in an area that is highly multidisciplinary and requires large groups with diverse backgrounds for taking up challenging projects at the forefront of a rapidly developing technology. The Committee would like to point out that within the Institute, Bioinformatics is one area where a number of faculty members with diverse backgrounds are collaborating at the research level. Annexure-3 contains a list of faculty members who are known to the Committee members to be collaborating and/or have research interests in various facets of Bioinformatics. It must be mentioned that this lists is not complete and there may be others who are actively researching in this area. It is expected that faculty members listed in Annexure-3 will together take on the task of running this programme and be the initial core faculty of the programme.

The Committee feels that the viable intake for the programme in the beginning would be 20 full time students divided equally between those with biological sciences background and those with computational sciences background. Students who have qualified in the GATE in biological sciences or computer sciences and IT will be eligible to apply. Applicants with NET qualifications in the above areas may also be considered eligible for the programme. Committee feels that biological science background students should have studied mathematics at the +2 level or have had equivalent training in mathematics as a part of their Board/University coursework.

It is expected that the process of stabilizing the programme in time and experience gained thereof will lead to a better understanding of how such a programme should be structured. It is recommended that such a formal review after three batches have passed out be conducted. This review should be the basis for mid-course correction in all aspects of the programme including the credit structure, definition of the core programme, in-take philosophy, size of in-take etc..

Annexure-1

Scheduling of courses

Total : 4 Semester Programme: 60 credits; 9 Core courses (27 credits): 5 Electives (15 credits); Project (18 credits)

I Semester **(15 credits)**

Core courses for students with Biological Sciences background

1. Introduction to Programming and Data Structures (3-0-2, 4 credits) (MAL 701)
2. Mathematics for Biologists (3 credits) [#]

Core courses for students with Non-biology background

1. Foundation Course in Modern Biology (3 credits) [!]
2. Introductory Biochemistry (3 credits) [@]

Common core courses

3. Biological Sciences Wet Lab. [0-0-6; 3 credits] [$]
4. Statistical & Probabilistic Methods in Bioinformatics [3-0-0; 3 credits][#]
5. Open Elective I (3 credits)

II Semester **(18 credits)**

6. Computational Biology [3-0-0; 3 credits] [&]
7. Programming Lab. (C++, Java, Perl etc.) [1-0-4; 3 credits] [!]
8. Database Management [3-0-0; 3 credits][%]
9. Bioinformatics and Biological Databases [2-0-4; 3 credits] [!]
10. Programme Elective I [3 credits]
11. Open Elective II [3 credits]

III Semester **(15 credits)**

12. Biomolecular Modelling & Simulation [2-0-2; 3 credits] [@]
13. Programme Elective II (3 credits)
14. Programme Elective III [3 credits]
15. Major Project (Part I) [6 credits]

IV Semester **(12 credits)**

16. Major Project (Part II) [12 credits]

(Total : 60 credits)

Notes

#	New course of the Mathematics Department
!	New course of the Biochemical Engineering and Biotechnology Department
@	New course of the Chemistry Department
$	New course shared between Chemistry and Biochemical Engineering and Biotechnology Departments
&	New course of the Computer Science and Engineering Department
%	Existing course of the Mathematics/Computer Science and Engineering Departments

---

Some Existing Courses Proposed as Programme Electives

1. BE 722 Protein Engineering [3-0-0, 3 credits]
2. BE 720N Combinatorial Biotechnology [3-0-0, 3 credits]
3. BEL 721 Bionanotechnology [3-0-0, 3 credits]
4. BEL 722 Genomics & Proteomics [, 3-0-0, 3 credits]
5. BEL 723 Data Analysis in DNA Microarrays [3-0-2, 4 credits]
6. CYL 704 Chemical Computations [2-0-2, 3 credits]
7. CYL 542 Biochemistry II-Intermediatory Metabolism [2-1-4, 5 credits]
8. CYL 643 Biochemistry III –Intermediatory metabolism [2-1-4, 5 credits]
9. CYL 626 Medicinal Chemistry [3-1-0, 4 credits]
10. CSL 671 Artificial Intelligence [3-0-2, 4 credits]
11. CSL 860 Special Topics in Parallel Computing [3-0-0, 3 credits]
12. EEL 709 Pattern Recognition [3-0-0, 3 credits]
13. EEL 781 Neural networks [3-0-0, 3 credits]
14. MAL720 Neuro Computing and Applications [3-0-0, 3 credits]

Some Possible New Programme Electives

1. Special Topics in Molecular Biology (Chemistry / DBEB)
2. Special Topics in Computational Biology (CSE / Chemistry)
3. In Silico Drug Design (Chemistry)
4. Biomems (EE / CARE / CBME)
5. Special Topics in Systems Biology (DBEB)

Some Existing Courses Proposed as Open Electives

1. CYL 715 Bioanalytical Chemistry [3-0-0, 3 credits]
2. CYL 645 Immunochemistry [3-1-0, 4 credits]
3. CSL 781 Computer Graphics [3-0-3, 4.5 credits]
4. CSL783 Digital Image Analysis [3-0-3, 4.5 credits]
5. CSL785 Approximation Algorithms [3-0-0, 3 credits]
6. EEL 715 Image Processing [3-0-2, 4 credits]
7. MAL803 Pattern Recognition [3-0-0, 3 credits]
8. Special Topics in Chemical Informatics (Chemistry)
9. Special Topics in Structural Biology (Chemistry / DBEB)
10. Special Topics in Genomic Networks and Metabolomics (DBEB)

Annexure-2

Suggested Syllabus for Proposed New Core Courses

Core Courses for Students with Biological Sciences Background

1. Mathematics for Biologists [3-0-0, 3 credits] (Mathematics)

Elementary Set Theory, Introductory Differential and Integral Calculus, Ordinary Differential Equations of First and Higher Order, Fourier and Laplace Transforms, Linear Algebra and Matrix Theory.

Core Courses for Students with Non-biology Background

1. Foundation Course in Modern Biology [3-0-0, 3 credits] (DBEB)

Biology-Technology Interface, Cell Structure and Function, Non-covalent Interactions in Living Cells, Molecules in Cells, Enzymes: Structure, Catalysis, Industrial Applications, Membrane Transport, Bioenergetics, Introduction to Metabolism, Information Storage and Processing in Cells, Cell Signaling, Nerve Cells and Electrical Properties, Technique in Cell and Molecular biology, Cell Evolution, Biochemical Capacities, Biodiversity.

2. Introductory Biochemistry [3-0-0, 3 credits] (Chemistry)

Introduction- Aims & Scope, Molecules of life: nucleic acids, proteins, carbohydrates and lipids, their structure and function; Genome organization and replication; Gene expression and regulation; Biomolecular interactions; Protein Purification Techniques, Introduction to Enzymes, Vitamins, Coenzymes and Biological Membranes, Metabolic Pathways for Breakdown of Carbohydrates – Glycolysis, Pentose Phosphate Pathway, Citric Acid Cycle, Electron Transport Chain, Photo-Phosphorylation, Oxidation of Fatty Acids, Gluconeogenesis and Control of Glycogen Metabolism, Signal Transduction.

Common Core Courses

3. Biological Sciences Wet Lab. [0-0-6, 3 credits] (DBEB/Chemistry)

Hands on experiments about micro-organisms and their applications; training in isolation, purification, and characterisation of biomolecules.

4. Bioinformatics and Biological Databases [2-0-2, 3 credits] (DBEB)

   Introduction to Bioinformatics and its Application, Molecular Biology for bioinformatics (Central Dogma), Biological data bases (primary, secondary hybrid etc), and their Annotation, Protein and Nucleotide (DNA) sequencing techniques, Pair-wise and multiple sequence alignment algorithm, Phylogenetic Analysis, Hidden Markov Model (HMM) and its Application, Microbial Genomics, Metabolic Flux Analysis.Tools for DNA sequence analysis, protein sequence analysis; Usage of sequence alignment and searching tools for Gene Identification, Genome Annotation, ORFs, ESTs, Codon biases, Redundancy, Search engines; Conserved motifs, patterns, blocks, domains, Secondary and tertiary Structure prediction tools; FASTA, BLAST, PSI-BLAST, CLUSTALW, Multalign, Dialign, GeneBee, MotifScan, TMPred, GOR, Chou-Fasman, NNPredict, Promoterfinder, NEBcutter, Genscan, ORF Finder, IntronExon finder etc. Using Biological databases.

5. Statistical & Probabilistic Methods in Bioinformatics [3-0-0, 3 credits] (Mathematics)

   Probability, Conditional probability, random variables, expected value, Specific discrete and continuous distributions eg. Binomial, Poisson, Geometric, Pascal, Hypergeometric, Uniform, Exponential and Normal, Poisson process, Multidimensional random variables, Multinomial and bivariate normal distributions, Moment generating function, Law of large numbers and central limit theorem, Sampling distributions, Point and interval estimation, Testing of hypothesis, Goodness of fit and contingency tables, Linear regression, Principal Component Analysis; Clustering; Applications of the Above Techniques to Problems in Population Biology and Genomics.

6. Computational Biology [3-0-0; 3 credits] (New Course, CSE)

   Basics of Molecular Biology; Basics of String Matching, Suffix Trees and Suffix Arrays, Introduction to Sequence Similarity; Alignment by Dynamic programming; Local Alignment and Gap Penalties; Multiple Sequence

Alignment; Finding Instances of Known Sites; Relative Entropy and Binding Energy; Finding Instances of Unknown Sites; Correlation of positions in Sequences; Maximum Subsequence problem; Markov Chains; using Interpolated Context Models to Find Genes; Start Codon Prediction; RNA Secondary Structure Prediction.

7. Programming Lab. [1-0-4; 3 credits] (New Course, DBEB)

The laboratory exercises will emphasise biological applications development in C, C++, Perl, Java etc.

8. Biomolecular Modelling & Simulation [2-0-2; 3 credits] (Chemistry)

Introduction to *ab-initio*, semi-empirical & molecular mechanical methods, Theory and Practice of Energy minimization, Monte Carlo and Molecular Dynamics simulations. Structure visualization, Generation of Molecular Electrostatic Potential, Field maps and Surfaces, Theoretical methods to calculate binding free energies and rate constants. Methods to model Proteins, Nucleic Acids (DNA & RNA), Carbohydrates and Membranes.

Suggested Syllabii for Some of the Proposed New Electives

1. Special Topics in Molecular Biology [3-0-0] (DBEB/Chemistry)

Topics will be chosen from current issues in Structural and Functional Genomics and Proteomics.

2. *In Silico* Drug Design [2-0-2] (Chemistry)

Disease / disorder and Drug targets. Concept of receptor / target site. Concepts in molecular recognition. Drug-like properties and associated empirical rules. structure based drug design; Applications of QM methods; Molecular descriptors in QSAR studies, Small molecule force field parameters (charges), potentials, Active site identification algorithms, ligand docking algorithms, thermodynamics & kinetics of protein-drug binding. Drug stability, synthesizability and drug delivery.

3. Special Topics in Systems Biology [3-0-0, 3 credits] (DBEB)

   Emerging new ideas on treating biological systems as systems of molecular networks. Elements of system modeling and mathematical methods to formulate system's response to perturbations. New directions in metabolic pathways and cellomics to better understand organization of tissues, organs and organisms.

4. Special Topics in Computational Biology [3-0-0, 3 credits] (Chemistry, CSE, DBEB)

   New Directions in Nucleic Acid and Protein Sequence Analysis, Protein Folding, Parallelisation and Clustering in Computional Biology.

5. Special Topics in Structural Biology [3-0-0, 3 crdits](Chemistry / DBEB)

   Topics will include recent trends in X-ray, NMR, Electron Microscopy and computer modelling, in elucidating Structural and Functional Aspects of Biomolecular Assemblies and their Interractions.

6. Special Topics in Genomic Networks and Metabolomics [3-0-0, 3 credits](DBEB)

   Topics will include spatial and temporal aspects of genome organization, gene expression and metabolic pathways.

Annexure-3

Some Faculty of the Institute with Research Interests in Bioinformatics

1. Prof. Saroj Mishra, Head, DBEB
2. Prof. G. P. Agarwal, DBEB
3. Dr. Sunil Nath, DBEB
4. Dr. Aradhana Srivastava, DBEB
5. Dr. Tapan Chaudhuri, DBEB
6. Dr. James Gomes, DBEB
7. Dr. J. K. Deb, DBEB
8. Dr. B. Kundu, DBEB
9. Prof. M. N. Gupta, Chemistry
10. Prof. B. Jayaram, Chemistry
11. Dr. C. Chakravarty, Chemistry
12. Dr. Nalin Pant, Chemistry
13. Dr. S. K. Khare, Chemistry
14. Dr. S. Pandey, Chemistry
15. Dr. P. S. Pandey, Chemistry
16. Prof. B. Chandra, Mathematics
17. Prof. Suresh Chandra, Mathematics
18. Dr. Lipika Dey, Mathematics
19. Dr. N. Chatterjee, Mathematics
20. Dr. B. S. Panda, Mathematics
21. Prof. S. K. Koul, Head, CARE
22. Prof. Sneh Anand, CBME
23. Dr. Jayadeva, EE
24. Prof. S. D. Joshi, EE
25. Dr. R. K. P. Bhatt, EE
26. Prof. S. N. Maheshwari, CSE
27. Dr. Naveen Garg, CSE
28. Dr. Amit Kumar, CSE
29. Dr Prasoon Tiwari, CSE

Annexure

Template for New M.Tech. Programme

Name of the M.Tech. Programme	M.Tech. Programme in Bioinformatics

Distribution of Total Credits

Program Core (PC)	Program Elective (PE)	Open Electives (OE)	Total Credits
45	9	6	60

Semester wise Distribution of credits

Scheduling of courses

Semester I					
	Course No.	**Title**	**Type**	**L-T-P**	**Credits**
1A	BE	Foundation Course in Modern Biology	PC	3-0-0	3
2A	CY	Introductory Biochemistry	PC	3-0-0	3
1B	MAL 701	Introduction to Programming and Data Structures	PC	3-0-2	4
2B	MA	Mathematics for Biologists	PC	3-0-0	3
3	BE/CY	Biological Sciences Wet Lab.	PC	0-0-6	3
4	MA	Statistical & Probabilistic Methods in Bioinformatics	PC	3-0-0	3
5		Open Elective - 1	OE		3
		Total Credits			**15**
Semester II					
1	CS	Computational Biology	PC	3-0-0	3
2	BE	Programming Lab. (C++, Java, Perl etc.)	PC	1-0-4	3
3	MA/CS	Database Management	PC	3-0-0	3
4	BE	Bioinformatics & Biological Databases	PC	2-0-4	3
5		Programme Elective - 1	PE		3
6		Open Elective - II	OE		3
7		**Total Credits**			**18**

Semester III					
	CY	Biomolecular Modelling & Simulation	PC	2-0-2	3
		Programme Elective - II	PE		3
		Programme Elective – III	PE		3
		Major Project (Part I)	PC		6
		Total credits			**15**
Semester IV					
		Major Project (Part II)	PC		**12**

A. List of Proposed Program Electives for M.Tech. Programme (from existing courses)

Sr. No.	Course No.	Title	L-T-P	Credits
1	BE 722	Protein Engineering	3-0-0	3
2	BE 720N	Combinatorial Biotechnology	3-0-0	3
3	BEL 721	Bionanotechnology	3-0-0	3
4	BEL 722	Genomics & Proteomics	3-0-0	3
5	BEL 723	Data Analysis in DNA Microarrays	3-0-2	4
6	CYL 704	Chemical Computations	2-0-2	3
7	CYL 542	Biochemistry II-Intermediatory Metabolism	2-1-4,	5
8	CYL 643	Biochemistry III -Intermediatory Metabolism	2-1-4	5
9	CYL 626	Medicinal Chemistry	3-1-0	4
10	CSL 671	Artificial Intelligence	3-0-2	4
11	CSL 860	Special Topics in Parallel Computing	3-0-0	3
12	EEL 709	Pattern Recognition	3-0-0	3
13	EEL 781	Neural networks	3-0-0	3
14	CYL 715	Bioanalytical Chemistry	3-0-0	3
15	CYL 645	Immunochemistry	3-1-0	4
16	CSL 781	Computer Graphics	3-0-3	4.5
17	CSL783	Digital Image Analysis	3-0-3	4.5
18	CSL785	Approximation Algorithms	3-0-0	3
19	EEL 715	Image Processing	3-0-2	4
20	MAL803	Pattern Recognition	3-0-0	3

List of NEW courses if any being proposed
(Details of each new course in annexure-I)

Course No.	Title	L-T-P	Credits
MA	Mathematics for Biologists	3-0-0	3
MA	Statistical & Probabilistic Methods in Bioinformatics	3-0-0	3
BE	Foundation Course in Modern Biology	3-0-0	3
BE	Bioinformatics and Biological Databases	2-0-4	3
BE	Programming Lab	1-0-4	3
CY	Introductory Biochemistry	3-0-0	3
CY	Biomolecular Modelling & Simulation	2-0-2	3
BE/CY	Biological Sciences Wet Lab	0-0-6	3
CS	Computational Biology	3-0-0	3
CS/MA	Database Management	3-0-0	3
CY	Special Topics in Chemical Informatics	3-0-0	3
CY/BE	Special Topics in Structural Biology	3-0-0	3
BE	Special Topics in Genomic Networks and Metabolomics	3-0-0	3