Curriculum Vita

Name	**: Shrish Tiwari**
Date of Birth	: 09-11-1965
Present Address	: Dr. Shrish Tiwari Bldg. 37, Rm. 4B14, National Institutes of Health Bethesda, MD 20892, USA Phone: (301) 402-0437, FAX: (301) 402-1344 Email: shrish@nih.gov Homepage: http://www.geocities.com/shrish.geo/
Permanent Address	: Shrish Tiwari Sri Aurobindo Ashram Pondicherry – 605 002, India

Academic Qualifications:

Ph.D.	1997	Jawaharlal Nehru University, New Delhi, India. Thesis title: **Studies in complexity: Applications to Dynamical systems and genomic sequences.** Thesis supervisor: **Prof. R. Ramaswamy.**
M.Sc.	1990	Indian Institute of Technology, Kanpur, India.
B.Sc.	1986	Sri Aurobindo International Centre of Education Pondicherry, India.

Research Experience:

Presently	Visiting Fellow at the National Institutes of Health, Bethesda, USA.
1997-1999	Postdoctoral Fellow at Centre for Cellular and Molecular Biology, Hyderabad, India.
1996-1997	Research Associate at National Chemical Laboratory, Pune, India

Additional Skills/Qualifications:

Programming languages	C, Fortran and Pascal
Operating environments	DOS, Windows 95/98/NT, Mac OS, UNIX, VMS
Web related	HTML, Java Script, CGI script
Databases used	GenBank, PDB, SwissProt, SCOP, VAST
Programs used	BLAST/PSIBLAST, Insight II, Rasmol

Algorithms developed:

GeneScan	Algorithm for gene identification. Available at http://202.41.10.146/GS.html
MEICPS	Algorithm to predict mutations to alter *in vivo* protein stability
run_psi	Script to identify similar protein structures with low sequence homology

Brief resume of my research work

My main area of interest is in Bioinformatics, which now refers to the computational efforts to understand the mechanisms in a living cell. This involves identifying genes, predicting the native structure of proteins, defining their function, determining the protein-protein and protein-DNA interactions which make the machinery of the living cells. I will briefly describe my small contributions to this field.

Gene Identification

Aim here was to identify protein coding regions in genomic DNA. An extensive study of genes from a variety of organisms revealed a consistent 3-periodicity. This observation was quantified by a measure defined as the signal-to-noise ratio at frequency $f = 1/3$ in the Fourier transform of the sequence. This measure was used to develop an algorithm, GeneScan, to scan for genes in complete genomes. The program was capable of identifying genes in prokaryotes and exons in eukaryotes. This problem automated gene identification has great significance since it requires a lot of time, effort and ingenuity to identify genes experimentally.

Protein Stability Analysis

An understanding of *in vitvo* or *in vitro* stability of a protein will help us alter the stability of industrially and medically important proteins. A set of experimental mutations was collected from literature where the effect of the mutation on the stability of the protein was qualitatively determined. Only mutations which were performed on proteins of known structure were selected. The structural environment, as defined by secondary structure, solvent accessible area, for each of the wild type residue was determined. We also selected a set of nonhomologous high-resolution structures. The propensity of all 20 amino acids, in the environment defined by each residue in the previous set was computed. It was found that the propensity of the mutant residue was higher if the mutation stabilised the protein and vice-versa, in most cases.

Protein Secondary Structure Prediction

Prediction of the secondary structure of a protein from knowledge of sequence alone is the first step towards predicting the 3-dimensional conformation of the protein. The latter is closely related to the function of the protein. Statistical profiles of different amino acids were computed from an analysis of a set of nonhomologous high-resolution structures. Profile matrices were built using single and dipeptide frequencies and incorporating near-neighbour effects. These matrices were used to predict structure of a test set of structures, which were homologous to structures in the first set.

Fold recognition

Recognition of similar structures with low sequence homology has important implications in predicting protein structure using homology modeling. BLAST and PSIBLAST rely on good statistics for a good alignment. Using BLAST to identify the fold/structure of a new protein sequence by running the alignment against the PDB database has low success rate because of the small size of the database. To overcome this drawback, the alignment was first performed against SwissProt database using PSIBLAST. The homologous regions of all the hits were extracted, and each of them was aligned with sequences in the PDB database. A significant improvement in identifying similar folds with low sequence identity was noticed. Folds/domains have been identified to be more basic than complete structures. Two protein structures may be quite different but they may have some folds in common. Thus identification of folds seem to be more important.

***Ab initio* protein folding**

This work is in progress. Our aim is to design an energy potential for ab-initio folding of proteins. We are currently looking at one of the interactions reported to be important in the folding of proteins, namely the electrostatic interaction between the atoms of the main-chain. We are designing appropriate weight factors for different residues to get it to correlate to the preference exhibited by the different amino acids to specific structural environment. Design of an energy potential which can fold a protein is bound to lead to insight into the physical mechanism involved in the process.

Publications

1. **A statistical analytical approach to predict secondary structure of proteins from amino acid sequence information**. S. Tiwari and B.V.B. Reddy, *Theoretical Chemical Accounts*, **101**, 46-50 (1999).
2. **Use of propensities of amino acids to the local structural environments to understand effect of substitution mutations on protein stability**. B.V.B. Reddy, S. Datta and S. Tiwari, *Protein Engineering*, **11**, 1137-1145 (1998).
3. **MEICPS: A program to suggest substituted point mutations to engineer intr-cellular protein stability**. B.V.B. Reddy, P. Ramesh and S. Tiwari, *Bioinformatics*, **14**, 225-226 (1998).
4. **Prediction of probable genes by Fourier analysis of genomic sequences**. S. Tiwari, S. Ramachandran, A. Bhattacharya, S. Bhattacharya and R. Ramaswamy, *CABIOS*, **13**, 263-270 (1997).
5. **Gene identification *in silico***. S. Tiwari, A. Bhattacharya, S. Bhattacharya and R. Ramaswamy, *Curr. Sci. (India)*, **71**, 12-24 (1996).
6. **Adaptive control in a resource management model**. S. Tiwari, R. Ramaswamy and J. Subba Rao, *Ecological Modelling*, **84**, 53-62 (1996).
7. **Nose-Hoover dynamics of a nonintegrable system**. S. Tiwari and R. Ramaswamy, *J. Mol. Struc. (Theochem)*, **361**, 111-116 (1996).

Referees

1. **Dr. B.K. Lee**
   Bldg. 37, Rm. 4B15
   National Institutes of Health
   Bethesda, MD 20892, USA
   Email: bkl@helix.nih.gov

2. **Dr. B.V.B. Reddy**
   San Diego Supercomputer Center
   P.O. Box 85608, San Diego
   CA 92186-5608, USA
   Email: breddy@sdsc.edu

3. **Prof. Ramaswamy**
   School of Physical Sciences
   Jawaharlal Nehru University
   New Delhi – 110 067, INDIA
   Email: rama@jnuniv.ernet.in

4. **Prof. A. Bhattacharya**
   School of Life Sciences
   Jawaharlal Nehru University
   New Delhi – 110 067, INDIA
   Email: alok@jnuniv.ernet.in

5. **Dr. Somdatta Sinha**
   1st Floor, South Wing
   Centre for Cellular and Molecular Biology
   Hyderbad – 500 007, INDIA
   Email: sinha@ccmb.ap.nic.in