OSCAT-2012

International Conference

On

Open Source for Computer-Aided Translational Medicine

February 22-25, 2012

Convenor:

Dr. G.P.S Raghava

Bioinformatics Centre

Institute of Microbial Technology

Chandigarh, India

CONTENTS

I. PROGRAMME

22 February 2012 (Wednesday)		
9:00-11:00	Registration- OSCAT 2012	
11:00-11:10	WELCOME ADDRESS	
11:10-11:50	Prof. Samir K Brahmachari (DG CSIR & Mentor OSDD)	Inaugural lecture
11:50-12:20	HIGH TEA	
SESSION- I: Inaugural Session *Chairperson: Vani Brahmachari*		
12:20-13:00	Bernard Munos	Open-Source Therapeutics: An Opportunity for creative disruption
13:00-14:00	LUNCH	
SESSION-II: Bio-Therapeutics & Bioinformatics *Chairperson: T.S. Balganesh*		
14:00-14:40	Girish Sahni	New drug development from old work horses: Lessons from Protein Thrombolytics
14:40-15:20	Georg Sczakiel	The role of local RNA target structure for small interfering RNA (siRNA) and antisense nucleic acids: A computational predictive approach
15:20-16:20	TEA/Poster Session(P1 - P42)	
16:20-17:00	K. K Bhutani	Open source Drug Discovery: The role of natural product databases
17:00-17:40	Michael Galperin	Molecular biology databases as essential tools in biomedical research
17:40-18:20	Geoff Barton	New tools for the analysis of biological sequences
18:20-19:00	Oral Presentations O1: Dr. Amit Arora O2: Mr Ankit Tripathi O3: Ms. Reema Singh O4: Dr Zahoor Ahmad Parry	
19:00-20:00	Poster Session(P1 - P42)	
20:00	DINNER	

23 February 2012 (Thursday)

SESSION-III: Therapeutic challenges in M. tuberculosis *Chairperson: Robert Glen*		
9:00-9:30	Poster Session(P1 - P42)	
9:30-10:10	T.S. Balganesh	Finding new drugs for the treatment of TB-challenges and opportunities
10:10-10:50	Andris JC Steyn	Environmental free radicals and drug-resistant *Mycobacterium tuberculosis*
10:50-11:10	TEA/Poster Session(P1 - P42)	
11:10-11:50	Abu Salim Mustafa	*In-silico* analysis and wet-lab experiments to identify - proteins and peptides of *Mycobacterium tuberculosis* for immunological diagnosis and vaccine development
11:50-12:30	Akhilesh Pandey	Proteomics: From genome annotation to understanding bacterial pathogenesis
12:30-13:10	Oral Presentations O5: Mr. Saravanan P O6: Dr. Dibyajyoti Banerjee O7: Mr. Shailesh Kumar O8: Mrs Abhilasha Thakur	
13:10-14:10	LUNCH	
SESSION-IV: In-silico drug targets *Chairperson: Geoff Barton*		
14:10-14:50	Joel Sussman	Function and structure of Intrinsically Flexible Proteins (IFPs)
14:50-15:30	Vani Brahmachari	Integration of *in-silico* and experimental analysis: Mining the human genome for the tool kit for developmental gene regulation
15:30-15:50	TEA/Poster Session(P1 - P42)	
SESSION-V: OSDD *Chairperson: Bernard Munos*		
15:50-16:10	Zakir Thomas	Introduction to OSDD
16:10-16:30	S.Ramachandran	1. Computational screening for new inhibitors of *M. tuberculosis* mycolyltransferases antigen 85 group of proteins as potential drug targets 2. ParallelVSR: a pipeline for parallel virtual screening on R – platform.
16:30-16:50	David Wild	Exploring semantic networks of public data for drug-target prediction
16:50-17:10	Anshu Bhardwaj	Open Source Drug Discovery: Community collaborative experiments and resources
17:10-17:30	Vinod Scaria	Genomics of the Unknown: Translational applications of Hologenome sequencing
17:30-17:50	Andrew Lynn	Using grid computing to provision computational resources for community use in OSDD.

17:50-18:10	Haridas Rhode	OSDD Chemistry
18:10-18:30	U C A Jaleel	Successes and challenges of online Cheminformatics research
18:30-18:50	G P S Raghava	Chemoinformatics tools on CRDD portal
18:50-19:10	Oral Presentations O9: Mr. Manjeet Singh Chalga O10: Dr. M Karthikeyan	
19:10-20:00	Poster Session(P1 - P42)	
20:00	DINNER	

24 February 2012 (Friday)

SESSION-VI: Chemoinformatics & Drug Design *Chairperson: S K Brahmachari*		
9:00-9:30	Poster Session(P43 - P83)	
09:30-10:10	Stephen H. Bryant	PubChem: An open repository for chemical structure and biological activity information
10:10-10:50	Christoph Steinbeck	Latest developments in the Chemistry Development Kit (CDK)
10:50-11:10	TEA/Poster Session(P43 - P83)	
11:10-11:50	Manoj Bhasin	Integration of *in-silico* and experimental approaches to drive discovery of biomarkers and therapeutics
11:50-12:30	Ahmed Kamal	Development of new antitubercular agents
12:30-13:00	Oral Presentations O11: Dr. Jaya Shukla O12: Mrs. Indrani Mitra O13: Mr. Harinder Singh	
13:00-14:00	LUNCH	
SESSION-VII: Translational Medicine *Chairperson: J G Frey*		
14:00-14:40	Ena Wang	Paradigm shift in translational medicine powered by computation and bioinformatics
14:40-15:20	Robert Glen	Probing the biology of the Apelin (APJ) receptor through simulation, design, synthesis and pharmacological evaluation
15:20	TEA/Poster Session(P43 – P83)------**CITY TOUR**	
20:00	DINNER	

25 February 2012 (Saturday)

SESSION-VIII: Computer-aided therapeutics *Chairperson: Ena Wang*		
9:00-9:30	Poster Session(P43 - P83)	
9:30-10:00	G P S Raghava	Web services for bioinformatics and chemoinformatics
10:10-10:50	PV Bharatam	Integrated approach to discover y-shaped PPAR γ agonists
10:50 -11:10	TEA/Poster Session(P43 - P83)	
11:10-11:50	Dinkar Salunke	Molecular mimicry and ligand specificity: Experiences from Immunological investigations
11:50-12:30	Ram Vishwakarma	Efforts on antibacterial drug discovery at CSIR-IIIM
12:30-13:10	Oral Presentations O14: Mr. Sandeer Kumar O15: Ms. Rashmi Kumari O16: Mr. Rudra Narayan Das O17: Mr. Ajay Dara	
13:10-14:10	LUNCH	
SESSION-IX: Web services for translational science *Chairperson: PV Bharatam*		
14:10-14:50	J G Frey	The Semantic Web meets Wet and Dry Chemists
14:50-15:30	Sourav Pal	Hydration behavior of different head groups of phospholipids using Density Functional Theory: A validation from Fukui functions for prediction of active sites of hydration and drugs.
15:30-16:00	TEA	
16:00	Concluding session	
20:00	DINNER	

MESSAGE FROM DG

Prof. Brahmachari's message

I am glad that the nucleation of OSCAT-2012 started a decade ago, while I had the privilege to coordinate a team-CSIR project on '*In Silico* Biology on Drug Target Identification' as Director of Institute of Genomics and Integrative biology (IGIB). Since, then we have moved a long way, and on 15th September 2008, Open Source Drug Discovery was launched as a CSIR-led team India initiative with global partnerships. OSCAT is an outcome of such initiative under the leadership of Dr. G.P.S Raghava.

I am sure that these four days of deliberation under OSCAT 2012 will open up new vistas to explore, through computational biology, the biomedical therapies of tomorrow. The remarkable efforts of CSIR- Institute of Microbial Technology (IMTECH), towards promoting open source computational tools and their applications have drawn global attention. I am personally privileged to be an integral part of this exciting journey of new biology in the post-genomic era. I believe this will move into the new era of fourth paradigm of science, so that there is exciting collaboration between participants in the cyber space. I believe OSCAT will help in *in silico* biology efforts for drug target identification in an open source mode.

MESSAGE FROM DIRECTOR

It gives me immense pleasure to send this message for the "2nd International conference on open source for computer-aided translational medicine" which CSIR-IMTECH is organizing on February 22-25, 2012.

I wish that the objective of having intensive interaction between researchers and user agencies towards identification of R&D efforts and sharing of data, experience and infrastructure is fulfilled, culminating in business development and future course of action plans.

OSDD is a CSIR Team India Consortium with Global Partnership with a vision to provide affordable healthcare to the developing world. The OSDD concept aims to synergize the power of genomics, computational technologies and facilitate the participation of best minds. It provides a forum for collaboration & participation so as to solve the complex problems associated with discovering novel therapies for neglected tropical diseases like Malaria, Tuberculosis, Leshmaniasis, etc.

The principles of open innovation through open collaboration and sharing underlies OSDD approaches. In this open innovation platform, all the projects and the research results are reported on the web based platform based on the three cardinal principals of Collaborate, Discover & Share**.** To ensure affordability, the drugs that come out of the OSDD platform will be made available like a generic drug, without Intellectual Property encumbrances. OSDD thus relies on the already established business models of generic industry for delivery of drugs.

It is matter of great pleasure and pride for Institute of Microbial Technology to host this International conference. This conference will be useful for scientific community working in the field of drug discovery particularly on *Mycobacterium tuberculosis*. The objectives of this conference include promotion of open source in therapeutics, integrating *in silico* & wet-lab expertise in addition to discussions on the current status of therapeutics in tuberculosis. This conference is proposed to extend the deliberations on translational research as well.

On this occasion, I on behalf of CSIR-IMTECH send my very best wishes to all the participants and organizers and wish the conference a grand success.

INAUGURAL LECTURE

Inaugural Lecture

Open Source Drug Discovery: A New Paradigm for Distributed Co-Creation

Samir K Brahmachari[1, 2] and OSDD Consortium[2]

[1]CSIR-Institute of Genomics and Integrative Biology, Delhi, India

[2]Council of Scientific and Industrial Research, New Delhi, India

Historically drug discovery has remained a closed door phenomenon for the pharmaceutical companies. Although a large amount of biomedical research and many computational tools are available in the public domain, the cost of drug development is increasing continuously. The human genome sequencing, along with model organisms, held the promise of delivering therapeutic solutions at faster rates. However, while the cost of human genome sequencing has drastically come down from a billion dollars to a few thousand dollars over a decade, the cost of drug discovery has moved upwards and is reaching the billion dollar mark. To convert drug discovery into a similar open source global initiative like the Human Genome Project (HGP) or the World Wide Web (WWW), we have pioneered an innovative new path to drug discovery, namely the Open Source Drug Discovery (OSDD), which is a CSIR-led team India consortium with global partnerships with the vision to provide affordable healthcare to the developing world. This has the potential of providing a new way of discovering drugs by creating a web-based platform for collaborations from around the world.

The present talk will give an overview of how such an approach has worked so far towards the development of a systems biology platform for *Mycobacterium tuberculosis* (Mtb), for the discovery of new therapeutics, targets and potential lead molecules.

INVITED TALKS

Bernard Munos

Bernard Munos is the founder of the InnoThink Center for Research in Biomedical Innovation, which was created to translate innovation research into better business models. Before that, he was Advisor for corporate strategy at Eli Lilly and Company where he focused on disruptive innovation and the radical redesign of the R&D model. His work, which has been published in Nature and Science and was recently profiled by Forbes Magazine, has helped stimulate a broad rethinking of the pharmaceutical business model by industry, investors, policy-makers, regulators, and patient advocates. He has presented his findings at numerous meetings sponsored by the National Academies, the Institute of Medicine, the President's Cancer Panel, the NIH Leadership Forum, the World Health Organization, the OECD, the Kauffman Foundation, the US Patent and Trademark Office, Genome Canada, the American Chemical Society, as well as leading universities and think-tanks in the US and Europe. He received his MBA from Stanford University, and holds other graduate degrees in economics and animal science from the University of California at Davis, and the Paris Institute of Technology for Life, Food and Environmental Sciences in France.

Open-Source Therapeutics: An Opportunity for Creative Disruption

Dr. Bernard Munos

The R&D spending per new drug at major pharmaceutical companies now averages $9 billion. Under these conditions, one would expect the development of new medicines for rare and neglected diseases to dry up since the market for such drugs is small and offers poor returns. Yet, almost 30 NMEs have been approved for rare and neglected diseases by FDA since 2002, the majority in the last 5 years, and most of them have been sponsored by small companies or non-profit organizations with limited resources. An examination of their accounts shows they are bringing new drugs to market for about 1% to 2% of what it would take a big pharma (about $50m to $150m). It is tempting to see there a textbook illustration of what scholars have variously called "low-end disruption" (Clayton Christensen) or "creative destruction" (Schumpeter), which basically means that industry incumbents, slow to adapt and unable to compete, are poised to be displaced by younger, more agile organizations which run on economics that the bigger companies cannot hope to duplicate.

Dr.Girish Sahni

Dr.Girish Sahni, is the Director (2005-present) of Institute of Microbial Technology (IMTECH). He completed his Ph. D. in 1984 from Indian Institute of Science (IISc), Bengaluru. He was a postdoctoral trainee (1984-'86) at University of California, Santa Barbara, USA and continued his career as Senior Research Associate & Adjunct Faculty ('86-'88) at Albert Einstein College of Medicine, New York before joining IMTECH in 1987. His areas of specialization include protein engineering, molecular biology and enzymology. His prominent contributions are in the field of protein cardiovascular drugs especially on 'clot blusters' and his research led to the production of India's first indigenous "clot bluster drug", a natural streptokinase (under brand name 'STPase' marketed by Cadila Pharmaceuticals Ltd., Ahmedabad), and recombinant streptokinase (produced by Shasun Drugs, Chennai) marketed under brand names 'Klotbuster' (Alembic) and 'LupiFlo' (Lupin). He developed a novel life-saver drug (Clot-specific streptokinase) which has been patented worldwide and licensed to a US Pharma company in 2006 and its commercial launch is expected in 2013. He recently developed fourth-generation 'Anti-thrombolytic' clot busters that have been out-licensed for US $ 150 million dollars. He has published many papers in reputed journals. He is a Fellow of Indian Academy of Science, National Academy of Sciences, Association of Microbiologists of India, and Member of Guha Research Conference. He has been honored with many awards: The Vasvik Industrial Award in 2000; CSIR Technology Shield (2001-2002); National Biotechnology Product Development Award in 2002 and Ranbaxy Award in Pharmaceutical Sciences in 2003.

New Drug Development from 'Old Workhorses': Lessons from Protein Thrombolytics

Dr. Girish Sahni

Traditional approaches to new drug discovery depends either on serendipity, or an intuitive and information based design of new chemical entities for major diseases/targets. In case of proteins, our experience, especially in developing countries, has been more or less limited to developing new and improved processes for the production of already well established 'biosimilars'—an experience that has conferred a much-needed confidence to the nascent biotech industry. The presentation will touch upon IMTECH's experiences in developing biosimilars in the Thrombolytic protein area, and the challenges encountered during the "lab-to-market" process. Besides this, however, a new paradigm has also been explored: this involves the development of smarter, more beneficial proteins with clear- cut clinical and commercial advantages. Since commercial exploitation of Therapeutic Proteins is a Big Boys' Game with a global market of hundreds of billions US dollars, the vital experiences gleaned in this foray by IMTECH provides an interesting case study to replicate with other useful proteins.

Georg Sczakiel

Georg Sczakiel is a Full Professor, Director of "Institute of Molecular Medicine", University of Lübeck and University Hospital (UK S-H). He did his Ph.D. (1983-1986) in "Expression and biochemistry of protein domains: ras-p21 and rabbit myosin" at the Max-Planck-Institute of Medical Research, Heidelberg under the suprevision of Prof. Dr. A. Wittinghofer. Then did his PostDoc (1986-1990) at the German Cancer Research Center – Angewandte Tumorvirologie under the mentorship of Prof. Dr. H. zur Hausen – Nobel laureate 2008. From 1991-2000 he was Head of a research group at the German Cancer Research Center – Angewandte Tumorvirologie (Head: Prof. Dr. H. zur Hausen). His research interests include areas of Biochemistry, biology and medical application of nucleic acid-based methodology and of oligonucleotide-based drugs including ´antisense´, ribozymes, siRNA, and microRNA; Delivery of nucleic acid-based tools and drugs to mammalian cells; RNA-based diagniostics. He was also the Chair of the BMBF program "Development of non-viral vector systems for the tumor therapy and gene therapy of cervical cancer" from 1994-2000. From 1998-2000 he was Project leader of the "Steinbeis-Transfer Center of Genome Informatics", Heidelberg. In 2000, he Co-founded "Oligonucleotides Therapeutics Society" (OTS), Boston, USA and was the vice president of the same. Presently he is a Member of the editorial board of "*Oligonucleotides*"; Executive Editor of "*Nucleic Acids Research*"; Member of the Executive Board of "Chemie in unserer Zeit", VCH-Wiley, Weinheim.

The role of local RNA target structure for small interfering RNA (siRNA) and antisense nucleic acids: a computational predictive approach

Dr. Georg Sczakiel

Efficient suppression of gene expression at the post-transcriptional level can be achieved by tools such antisense nucleic acids, catalytic RNA, and small interfering RNA (siRNA) or microRNA. All kinds of such nucleic acid-based tools recognize their target nucleic acid via Watson-Crick base pairing which suggests that structural target accessibility matters.

Here, experimental evidence supporting the above assumption will be provided which adds to the parameters to be considered for the design of biologically effective siRNA-based tools and drugs.

Dr. Kamlesh Kumar Bhutani

Dr. Kamlesh Kumar Bhutani, *Ph.D, FRSC* was born on 25th December 1951 at Ambala in Haryana (India). He completed his MSc (1973) and PhD (1978) from Panjab University, Chandigarh, Punjab (INDIA). He has contributed more than 190 original research papers in International Journals of scientific repute in the fields of Natural Products, Medicinal Chemistry, Phytochemical Analysis, Phytotherapy, Pharmaceutical biology and bioassays, 8 books and 7 book chapters and has written more than 2 dozen review articles and presented 2 documentaries in public domain. His group has presented more than 250 research papers in various conferences held in India and Abroad. Prof Bhutani has delivered more than 100 plenary/ invited / guest lectures to various distinguished scientific and professional bodies globally. He has been granted 19 Indian Patents. He is a member of American Chemical Society, American Society of Pharmacognosy, Phytochemical Society of Europe, Society of Medicinal Plants (GA), Germany, Indian Pharmaceutical Association and Indian Science Congress Association. He has been elected International Fellow Sigma XI of Sigma XI - The Scientific Research Society of North America for noteworthy achievements in research.

Prof. Bhutani is exposed to mixed background qualifications and researches in the areas of medicinal and natural products chemistry, pharmaceutical and biological sciences encompassing diverse fields of SAR studies, isolation of natural products, structure elucidations by spectroscopic as well as X-ray crystallography, determination of pharmacological and biological activities and cellular and molecular mode of action studies. He is consciously researching traditional medicines with an objective to know where the *terra incognita, i.e., ignorance lies*. He has integrated research processes of new drug discovery and complementary medicine in natural products with interdisciplinary approach (Bhutani KK, 2001; Bhutani KK, 2003a; Kalia

et. al., 2004). He has identified and characterised bioactive molecules, responsible for antiinflammatory (C. Selvam and Bhutani KK, 2004,), adaptogenic (Pawar RS and Bhutani KK, 2006), anti-cancer (Chitra G and Bhutani KK, Indian Patents, 2004), anti-diabetic, antioxidant (Hemant J and Bhutani KK, 2001), antileshminial activities (Bharate S and Bhutani KK, 2005) from a number of traditional medicinal plants and developed processes for their purifications and separations. The marker concept has been fully exploited by his group in standardising the herbal medicinal products, using his elucidated new structures of 167 isolated plant compounds (that have found entry into the Dictionary of Natural Products on CD-ROM version 9.2, copyright 2001-2006, Chapman and Hall/CRC), in addition to more than 500 known compound libraries isolated in his lab. He has contributed more than 100 standard and researched Indian Systems of Medicine monographs to National Pool for codification purposes of scientific knowledge in Traditional Medicines and developed marker and biomarker concepts for R & D in traditional medicine. Prof Bhutani has also developed a database of more than 500 medicinal plants and his merit is reflected in successful sponsored vast number of projects from Industry, Government agencies and international organizations to his team. He is the most sought after consultant in his area of expertise.

Dr Bhutani is Officiating Director of NIPER, S.A.S. Nagar since December 31, 2009 (A.N.). He is also heading the Department of Natural Products since its inception. He served as the Dean of the National Institute of Pharmaceutical Education and Research (NIPER) from 2002- 2008 and looked after the research and academic interests of the whole institute. He was involved in developing newer education strategies for NIPER - an Institute of National importance to make it an Institute of global importance. The contributions of Prof. Bhutani in establishing NIPER as Institute of Excellence in India is widely known.

Open source Drug Discovery: the role of Natural product Databases

Dr. Kamlesh Kumar Bhutani

Microbial diseases like malaria, tuberculosis and leishmaniasis are the major cause of mortality and morbidity amongst the developing countries like India. However, when one considers the amount of research that is being done in these disease areas, it comes out to be negligible, mostly because none of these diseases are endemic in the developed countries.

Exploring natural products represents a major strategy for discovering and developing new drugs in the Pharmaceutical world. Plants offer excellent source for the discovery of novel compounds for the treatment of such microbial diseases. Nevertheless, much of the work that is being done in the pharmaceutical industry or academia in the field of natural products remains in the records or is unpublished, there by impairing the transfer of knowledge in the open source. Our group at NIPER has been instrumental in building of databases that are freely available to the public domain. In this attempt, we have published two databases which represent a repository of plants and their traditional medicinal usage, found in the untapped areas of North East India and Greater Bihar. These can serve as a starting point for a number of drug discovery programs. Besides these, we have also compiled a reference library for fingerprinting of Indian medicinal plants as well as their formulations using advanced techniques like HPLC, HPTLC and LC-MS. This not only helps the budding young scientists doing research at the primitive stages, but also for setting benchmarks for the Indian Herbal drugs industry and thus making it at par with the standards that are followed for the conventional medicines.

Michael Y. Galperin

Michael Galperin is graduated from the Biological Faculty of the MV Lomonosov Moscow State University. He worked on the role of proton-motive force in bacterial membrane transport at the Department of Bioenergetics of the A.N. Belozersky Laboratory of Molecular Biology. His subsequent work at the Institute for Genetics and Selection of Industrial Microorganisms centered on improving industrial strains of Bacillus thuringiensis and Gluconobacter oxydans. Later, he spent a year at the Russian National Collection of Industrial Microorganisms (aka VKPM). In 1991, he moved to the United States. After postdoctorates at the University of Louisville, Kentucky (with the late Ron J. Doyle) and at the University of Connecticut at Storrs (with Kenneth M. Noll and Antonio H. Romano), he joined the National Center for Biotechnology Information (NCBI), a division of the National Library of Medicine at the National Institutes of Health in Bethesda, Maryland, and the home of GenBank, PubMed and many other databases, as a GenBank Fellow. In Jan. 1999, he was appointed as a Staff Scientist at the NCBI Computational Biology Branch. His current research interests include microbial genomics particularly in the functions of the remaining 'conserved hypothetical' proteins; protein family annotation for the COG database; analysis and reconstruction of biochemical and signaling pathways in poorly characterized microorganisms; using the genome data for the analysis of enzyme evolution, enzyme diversity and identification of potential drug targets. For the past 5 years, he has been curating the Nucleic Acids Research online Molecular Biology Database collection. In 2008, he became the Executive Editor of the NAR Database Issue. He is also a member of the Editorial Board of Environmental Microbiology for which he used to write a bi-monthly review of genomics news. He keeps on editing papers on microbial genomics and enzyme evolution for FEMS Microbiology Letters and FEMS Microbiology Reviews and serves on the editorial boards of J. Mol. Microbiol. Biotechnol., BMC Genomics, BMC Microbiology, OMICS and Microbial Biotechnology. In 2007, he became a member of the Faculty of 1000, which aims at highlighting the most interest papers published in biomedical literatures.

Molecular biology databases as essential tools in biomedical research

Dr. Michael Y. Galperin

In the past several years, the availability of complete genome sequences gave rise to a variety of high-throughput technologies that generate vast amounts of DNA and RNA sequence, gene expression, proteome, and mass-spectrometry data that are being deposited to the public archival databases at an ever-increasing pace. As a result, there is an increasing gap between data deposition and data analysis, which makes bioinformatics an increasingly attractive approach in all kinds of biomedical research. However, even the results of bioinformatics analyses are often too voluminous for a single journal publication and their proper presentation requires an online supplement. This situation has led to the proliferation of web-based online databases covering various aspects of modern molecular biology. Most of these databases result from long and difficult work of numerous scientists, providing an indispensable research and reference tools for the whole community.

Although there is no easy way to predict whether a particular database is going to be widely used, most successful public databases share several key attributes: clarity of purpose and the target audience, use of the well-established and widely available analysis tools, manual curation of the presented data by highly qualified biologists; thorough documentation of data sources and the analysis results; availability of the key results for downloading. However, due to the budget constraints, even the most successful public databases often face uncertain future. I will briefly discuss the issues of quality and sustainability of the public databases. I will then present several examples from our own research on data hunting that resulted in interesting biological insights: prediction of the universal role of c-di-GMP as bacterial second messenger, role of the sodium-motive force in bacterial pathogens, and elucidation of the key role of house-cleaning in cell metabolism.

Dr. Geoff Barton

Geoff Barton is Professor of Bioinformatics at the College of Life Sciences, University of Dundee. His core research is in the computational analysis and prediction of protein structure and function. He has published over 100 papers with total citations in excess of 8000. Software and web services from his group are widely distributed and used by the scientific community. These include the JPred/JNet secondary structure predictor which performs up to 100,000 predictions/month for scientists in more than 100 countries, and the Jalview multiple alignment analysis workbench which is installed on over 20,000 computers, and appears as an applet on over 150,000 web pages worldwide. Current research includes the analysis and prediction of protein-protein interactions and the analysis of large datasets from high throughput "Next Generation" sequencing technologies and protein mass-spectrometry. Geoff's first degree was in Biochemistry at the University of Manchester followed by Ph.D. research at Birkbeck College London where he developed the first practical technique for protein multiple sequence alignment. After two years as a postdoctoral Fellow at the Imperial Cancer Research Fund Laboratories, London, he was awarded a Royal Society University Research Fellowship to establish a research group in the Laboratory of Molecular Biophysics, University of Oxford. In 1997 he moved to be Research Team Leader and Head of the Protein Data Bank Europe (PDBe) at the European Bioinformatics Institute, Hinxton Cambridge. Geoff took up his current position in 2001.

See: www.compbio.dundee.ac.uk and www.jalview.org for further details.

New tools for the analysis of biological sequences

Dr. Geoff Barton

In this talk he will give a brief summary of the diverse research in bioinformatics that is going on in his group at the University of Dundee. He'll then focus on techniques for the analysis of protein multiple sequence alignments, in particular the program Jalview and the new portable webservices framework, JABAWS. Jalview (www.jalview.org) is a GPL-licenced multiple sequence alignment editor and functional analysis workbench that is widely regarded as the standard desktop tool and web applet for these tasks. The current version of Jalview is installed on over 20,000 computers world-wide, and is mentioned on over 100,000 web pages. The two papers describing Jalview have attracted over 1,100 citations and the most recent paper was flagged by ISI as a "hot paper". Jalview is exploited by major databases such as Pfam/Rfam and the European Bioinformatics Institute (EBI) services. Jalview has used SOAP-webservices technology to provide remote-access to multiple alignment algorithms and protein secondary structure prediction from the University of Dundee since 2004. Recently, we have developed a way to make it easy to deploy these services on your own server or desktop/laptop computer and to provide faster throughput and bigger jobs to be submitted on your own resources. The JABAWS portable webservices framework (www.compbio.dundee.ac.uk/jabaws) simplifies the deployment of web services for bioinformatics. Version 1.0 of JABAWS provided services for five multiple sequence alignment (MSA) methods (Probcons, T-coffee, Muscle, Mafft and ClustalW), while Version 2.0 has added Clustal Omega, three methods for the prediction of protein disorder and a package of methods for calculating protein conservation. JABAWS is the system employed by the Jalview multiple sequence analysis workbench since version 2.6. A fully functional, easy to set up server is provided as a Virtual Appliance (VA) which can be run on most operating systems that support a virtualization environment such as VMware or Oracle VirtualBox. JABAWS is also distributed as a Web Application aRchive (WAR) and can be configured to run on a single computer and/or a cluster managed by Grid Engine, LSF or other queuing systems that support DRMAA. JABAWS provides clients full access to each application's parameters, allows administrators to specify named parameter preset combinations and execution limits for each application through simple configuration files. The JABAWS command-line client allows integration of JABAWS services into conventional scripts.

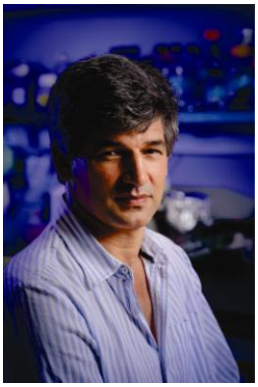
Dr. T.S.Balganesh,

He did his Ph.D from the University of Calcutta followed by Post doctoral Trainings at the Brookhaven National Lab, USA and the Max Planck Institute for Molecular Genetics, West Berlin, Germany respectively. Talking about his achievements: He led AstraZeneca India (AZI) to be recognized as one of the leading pharmaceutical organisations in TB drug discovery. Also led AZI to deliver the first anti-TB compound (AZD5847) for development, which is going through advanced clinical trial phases. He built a drug discovery culture at AstraZeneca India, Bangalore and is responsible for building an impressive pipeline of projects which is expected to deliver new compounds into the clinic every 18 months. He led AstraZeneca in the European Union Frame Work 6 consortium entitled "New Medicines for Tuberculosis". He was responsible for AstraZeneca India being a key player in the conceptualisation for the grant application co-ordinated by Prof Stewart Cole for the Frame Work 6 programme. The proposal was rated 'first' among the various applications and a grant of 16M Euros was awarded for the period 2005-2010. He is also the Steering Committee member for the European Union Frame Work 7 consortium entitled "More medicines for Tuberculosis" & led AZI in drafting the proposal for the European Union Frame Work 7. He has been Instrumental for leading the AZI – GATB collaboration towards finding novel drugs for TB. The collaboration involves identification of novel targets and chemical scaffolds. With his efforts AZI got Wellcome Trust Translation Science award. He was Prinicpal Investigator of a Wellcome grant proposal to investigate NAD mimetic library as a suitable starting point for lead generation. The envisaged project involves the design and construction of a library of NAD mimetics and test for activity on Mycobacterium tuberculosis. He has been invited to deliver lectures by several leading organisations like the Global Alliance for TB, New York, Gordon lecture series and the Keystone Symposia series. He has also been awarded an Honorary Doctorate from the University of Uppsala recently in 2011 for his outstanding work in the field of TB and Malaria drug discovery.

Finding new drugs for the treatment of TB- challenges and opportunities.

Dr. T.S.Balganesh

TB drug discovery and development has been and is a major challenge for both the academia and the pharmaceutical world. The major breakthroughs in 'omics' have still to make an impact on this problem. It is time to revisit our thinking and learn from the microbe, the host as well as the changing scenarios being presented as disease. What assumptions have we made which are not true and what does the decade of science on Mycobacterium tuberculosis tell us in terms of the need towards realigning our efforts. The presentation will discuss the changing aspects of the microbe, the patient and the environment in which we live. The different approaches that have been applied, the limitations and fresh initiatives that are being applied will be discussed. The presentation will also throw light on where we have succeeded and what needs to be built upon these successes.



Dr Adrie JC Steyn

Adrie JC Steyn, Associate Professor, Department of Microbiology, received his B.Sc. and M.Sc. degrees in Microbiology from the University of Stellenbosch, South Africa. He completed his thesis in the lab of Professor Julius Marmur at the Department of Biochemistry at the Albert Einstein College of Medicine, New York and was awarded a Ph.D. degree from the University of Stellenbosch in 1995. Dr. Steyn subsequently conducted postdoctoral research as a Howard Hughes Fellow in the lab of Dr. Barry Bloom at the Albert Einstein College of Medicine. In 1999 he moved with Dr. Bloom to the Harvard School of Public Health where he became a research associate and eventually research scientist. He joined the Department of Microbiology at UAB as an Assistant Professor in 2003 and became an Associate Professor in 2009.

In 2011 he accepted a position as Full Investigator at the KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH) in Durban, South Africa. He maintained his UAB faculty status and continues managing his BSL2 and BSL3 laboratories at UAB. K-RITH is a groundbreaking collaboration between the Howard Hughes Medical Institute (HHMI) and the University of KwaZulu-Natal (UKZN) in South Africa. K-RITH's mission is to conduct outstanding basic science research on TB and HIV, translate the scientific findings into new tools to control TB and HIV, and expand the educational opportunities in the region. The goal is that discoveries made in the heart of the TB and HIV epidemics will drive innovation to control the deadly diseases.

The major research interest in his laboratory is to study the role of *Mycobacterium tuberculosis* (*Mtb*) redox homeostasis in virulence and persistence, in particular the role of host/environmental gases such as nitric oxide (NO), carbon monoxide (CO), oxygen (O_2), and free radicals, and use this knowledge to develop pharmacological approaches that allow the control of TB. He is also studying the genetic basis of how a major risk factor for TB, cigarette smoke (CS), exacerbates disease.

Environmental free radicals and drug-resistant *Mycobacterium tuberculosis*

Dr Adrie JC Steyn

Tuberculosis (TB) is one of the leading causes of death worldwide and claims 1.8 million lives annually. Approximately 2 billion people worldwide are asymptomatically infected with *Mtb* and constitute a major impediment to worldwide public health control measures. However, the mechanism(s) that allows *Mtb* to enter, maintain and emerge from a latent state is unknown. Despite recent successes, no new effective anti-tuberculosis drugs have been developed in the past 40 years. TB can be cured with existing drugs; however, the 6-9 months of treatment lead to patient noncompliance, which enhances drug resistance. Approximately 50 million people are already infected with MDR-TB. While drug-sensitive TB can be cured with isoniazid, rifampicin, ethambutol and pyrazinamide in a 6 month regimen, treatment of MDR-TB can exceed 2 years, thus dramatically increasing treatment costs.

Currently, there are an estimated 1.3 billion smokers in the world and global tobacco related deaths are projected to become the world's single largest health problem and may cause 10 million deaths each year by 2020. Exposure of adults to side-stream or second hand smoke (SHS) has immediate adverse effects on the cardiovascular system that causes lung cancer and coronary heart disease. On the other hand, children exposed to SHS are at increased risk of sudden infant death syndrome, acute respiratory infections, ear problems and severe asthma. In India, a quarter of deaths among middle-aged men are caused by smoking (WHO Report on Global Tobacco Epidemic). In the US, the annual economic losses due to tobacco-related deaths are estimated at $92 billion (WHO). Approximately half of all smokers will develop a serious smoking-related illness, such as chronic obstructive pulmonary disease, which is characterized by irreversible airway obstruction,or cardiovascular disease. Furthermore, about 1–5% of smokers will develop a smoking-related malignancy, mostly lung adenocarcinoma or other epithelial cell tumors.

Smoking may induce progression or reactivation of disease in those infected. Exposure to tobacco smoke, and alcoholism and smoking are loosely associated with MDR TB. However, it is likely that MDR TB patients that remain noncompliant can acquire XDR TB, which can lead to super-XDR TB, which is a troublesome development. Recent studies have suggested that cigarette smoke (CS) may also be associated with faster progression to AIDS. Intriguingly, a link between CS, a known mutagen/carcinogen and drug-resistant *Mtb* has not yet been

considered. This is surprising, considering the devastating impact of both TB and tobacco use on the lives of millions of people. Also, since live *Mtb* in sputum of TB patients is directly exposed to CS, an obvious inference is that the known mutagens in CS are the cause of drug resistance. Alternatively, since many mutagens in CS first have to be metabolized to exert their mutagenic effect (*e.g*., as is evident by the "Ames *Salmonella* reverse mutagenicity assay"), it can also be argued that inhaled and metabolized CS in lungs cause MDR*Mtb*. Tobacco smoke contains at least 5,000 compounds including CO, NOx, nicotine, DNA-damaging agents (*e.g.*, reactive oxygen intermediates [ROI], peroxynitrite, ethylating agents and unidentified compounds), and major classes of carcinogens such as polycyclic aromatic hydrocarbons, aromatic amines and tobacco-specific nitrosamines, and toxic compounds such as acrolein, formaldehyde, acetaldehyde, and short-lived radicals that may contribute to the toxic and mutagenic effects of cigarette smoke. Approximately 10^{15} free radicals are present in a single puff of smoke. Not surprisingly, using the Ames assay, a vast number of studies have shown that CS is highly mutagenic in bacteria[2,3][2,3]. Similarly, CS causes carcinogenic and genotoxic effects in mammalian cells. However, to date, not a single study has reported that CS can generate drug resistant *Mtb*.

Since *Mtb* grows under conditions of stress within human cells, it has been speculated that these conditions may provide a hypermutable background leading to the development of drug resistance. The fact that MDR*Mtb* strains demonstrate a range of virulence, suggested that changes in "fitness" likely occurupon acquisition of drug resistance. However, differences in virulence can also be explained by mutations in genes other than those for antibiotic resistance. The mechanism of how drug resistant mutations in *Mtb* are generated is still unknown and is a fundamentally important question.

In this study, we have established that the free radicals in CS induce drug resistance in *Mtb* towards first line (isoniazid, rifampicin and ethambutol) and second line (streptomycin and kanamycin) TB drugs. CS exposed strains demonstrated an altered colony morphology, growth characteristic and infectivity. Transcriptome profiling showed that *Mtb*H37Rv yields a highly discrete response upon exposure to CS, which we term the 'smoke regulon'. The smoke regulon is particularly enriched in genes encoding PE/PPE proteins, mce operons, lipid metabolism, transporter proteins and components of the electron transport chain, which show a remarkable sequential expression. To characterize the genetic alterations in CS-exposed *Mtb* H37Rv, we

exploited next-generation genome sequencing and comprehensively mined the genome sequences of seven CS-induced single, MDR and XDRH37Rv strains. A novel discovery was the identification of a highly characteristic *Mtb* "smoke signature", which consists of ~100 common single nucleotide mutations found in all CS-exposed strains. Furthermore, each strain contained between 7 and 32 unique mutations responsible for the observed drug-resistance profile, differential morphology, growth and virulence properties. Synonymous and non-synonymous mutations, as well as insertions/deletions were characterized. Noticeably, the largest insertion was 6-bp, present in*rpoB*. Functional analysis and pathway mapping of the genes comprising the smoke signature showed that most of the single nucleotide changes were present in genes involved in intermediary metabolism and the PE-PGRS family.We also developed a novel *in vivo* cellular bio-assay to detect oxidative stress-mediated DNA damage and confirmed that $OH^{\bullet}$ radicals in CS directly mutate *Mtb* DNA. Comparative genome analysis with several clinical *Mtb*strains is currently underway to identify clinically relevant mutations.

This study provides a genetic mechanism for the first time, to understand how smoking may influence TB prognosis, and establishes CS as a major factor for the development of drug resistance in *Mtb*. Our findings may lead to a paradigm shift in understanding how environmental factors and social habits contribute to emergence of bacterial drug-resistance, and have global implications in terms of TB diagnostics and therapy.

Prof. Abu Salim Mustafa

Abu Salim Mustafa, molecular biologist and immunologist; BSC (1972) and MSc (1974) from Aligarh Muslim University, Aligarh, and PhD (1979) from the All India Institute of Medical Sciences, New Delhi, India.

I have been a Scientist (1980-1987) at the Norwegian Cancer Research Institute, Oslo, Norway (1980 to 1987), Research Associate (1987-1989) at the Whitehead Institute for Biomedical Research, MIT, Cambridge, USA, and Assistant Professor (189-192), Associate Professor (1992-1996) and Full Professor (1996-till date) at the Faculty of Medicine, Kuwait University, Kuwait.

At Kuwait University, I became Research Coordinator in 1992, Graduate Program Director in 1998, Member of Medical Research Council in 2001 and Director of Research Core Facility in 2012. Till date, I have published 255 full-length papers, 375 conference abstracts and edited 2 books. I am a recipient of 25 national and international awards in recognition of excellence in research. I became a Fellow of the Royal College of Pathologists, UK in 1998 and was Visiting Professor (Resource Person Program) of American Society for Microbiology/UNESCO in 2001 and 2004.

I am a recipient of more than 50 funded research grants from national and international funding, and have refereed more than 60 research projects submitted for funding to national and international institutions.

I am a member of the editorial board of two international journals, member of 11 Scientific Societies, supervised more than 25 graduate research thesis, invited speaker in 50 scientific conferences and chaired scientific sessions in 15 conferences.

In the recent years, the focus of my research has been molecular biology and immunology of infectious diseases with specific reference to diagnosis and identification of new vaccine candidates against mycobacterial diseases.

In Silico Analysis And Wet-Lab Experiments To Identify Proteins And Peptides Of Mycobacterium Tuberculosis For Immunological Diagnosis And Vaccine Development

Dr. Abu Salim Mustafa

Background: The failures of Mycobacterium bovis BCG as a vaccine and purified protein derivate (PPD) of M. tuberculosis as a diagnostic reagent in controlling the world-wide problem of tuberculosis have accelerated the research to identify M. tuberculosis-specific antigens and peptides, which could be useful as new vaccines and diagnostic reagents against tuberculosis.

Methods and Results: The comparative analyses of M. tuberculosis genome with the genomes of other mycobacteria have led to the identification of several genomic regions of difference (RD) between M. tuberculosis and M. bovis BCG. The identification of immunodominant and HLA-promiscuous antigens and peptides encoded by these RDs could be useful in the diagnosis and developing new vaccines against TB. The analysis of RD proteins and peptides by in silico methods and wet-lab experiments identified several major antigens and peptides of diagnostic and vaccine potentials. To evaluate them for in vivo reactivity, the genes of immunodominant antigens were cloned and expressed in DNA vaccine vectors (pUMVC6 and pUMVC7) and mycobacterial hosts, i.e. BCG, M. vaccae and M. smegmatis. Immunizations of mice and guinea-pigs with the recombinant constructs induced antigen-specific cellular and humoral immune responses in these animal models of TB. Each of these proteins had several T and B cell epitopes scattered throughout the sequences, which confirmed their strong immunogenicity and appropriateness for diagnostic and vaccine applications .

Conclusions: The results of this work suggest that bioinformatics-based in silico identification of promiscuous antigens and peptides of M. tuberculosis is a useful approach to identify new candidates important for diagnosis and vaccine applications. This type of analysis could further be extended to include the complete proteome of M. tuberculosis to identify all possible candidates for diagnostic and vaccine applications against tuberculosis.

The work was supported by Kuwait University Research Administration, grants MI01/10, YM08/07 and YM01/03.

Dr. Akhilesh Pandey

Akhilesh Pandey, M.D., Ph.D. is a Professor at the Institute of Genetic Medicine and the Departments of Biological Chemistry, Oncology and Pathology at the Johns Hopkins School of Medicine. He obtained his M.D. from Armed Forces Medical College, Pune and completed his residency in Pathology at the Brigham and Women's Hospital, Harvard Medical School. He obtained his Ph.D. in the laboratory of Vishva Dixit at the University of Michigan, Ann Arbor in 1995 and carried out his Postdoctoral work in the laboratory of Harvey Lodish at the Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology from 1996-1999. He was a Visiting Scientist with Matthias Mann at the University of Southern Denmark from 1999-2002 before joining Johns Hopkins in 2002. Dr. Pandey developed the SILAC method for quantitative proteomics and continues to apply this technique in many areas including signaling and biomarker discovery. His group has also developed a number of highly accessed and cited web-based databases and resources for human proteins and signal transduction pathways. He currently serves as an Editorial Board member of Molecular and Cellular Proteomics, Proteomics, Clinical Proteomics, Journal of Proteomics and DNA Research. He is also Director of the Institute of Bioinformatics, a non-profit systems biology research institute that he founded in 2002 in Bangalore, India.

Proteomics: From Genome Annotation to Understanding Bacterial Pathogenesis

Dr. Akhilesh Pandey

High resolution mass spectrometry-based proteomics has the potential to revolutionize genome annotation. We have initiated a systematic proteogenomics analysis to carry out genome annotation of various sequenced and unsequenced genomes of biomedical importance using high resolution mass spectrometry. This initiative includes the analysis of *Mycobacterium tuberculosis* and *Candida glabrata* whose genomes have been sequenced and *Anopheles stephensi, Leishmania donovani and Mangifera indica* whose genomes are unsequenced thus far. Gene prediction programs are known to suffer from both false negative and false positive predictions, leading to incorrect exons/exon boundaries, erroneous assignment of translational start/stop sites and missed genes/exons. I will discuss how we have carried out phosphoproteomic analysis to obtain novel insights into the signaling machinery of even well-characterized organisms such as *E. coli*. To gain insights into the extent and biological importance of tyrosine phosphorylation in *Escherichia coli,* we used immunoaffinity-based phosphotyrosine peptide enrichment combined with mass spectrometry to comprehensively identify low abundance tyrosine phosphorylated proteins, and accurately map phosphotyrosine sites. We identified a total of 513 unique phosphotyrosine sites on 343 proteins in *E. coli* K12 and the human pathogen enterohemorrhagic *E. coli* (EHEC) O157:H7, representing the largest phosphotyrosine proteome reported to date in bacteria. The abundance of tyrosine phosphorylation events allowed us to define a number of tyrosine phosphorylation motifs in bacteria. Tyrosine phosphorylated proteins belong to various functional classes such as metabolism, gene expression and virulence. Several proteins of the type three secretion system, required for the attaching and effacing (A/E) lesion phenotype characteristic for intestinal colonization by certain EHEC strains, contained tyrosine phosphorylation sites. Yet, A/E lesion and metabolic phenotypes were unaffected by the absence of the only two known tyrosine kinases in *E. coli* - Etk and Wzc. Substantial residual tyrosine phosphorylation present in an *etk wzc* double mutant strongly indicates the presence of hitherto unknown tyrosine kinases in *E. coli*. Thus, our study reveals tyrosine phosphorylation in bacteria more prevalent than previously recognized, and suggests the involvement of phosphotyrosine-mediated signaling in a broad range of cellular functions. With a growing number of genome sequencing projects, we advocate establishment of Proteogenomic Centers to replace the current model of having Genome Sequencing Centers carrying out initial characterization of genomes.

Dr.Joel L. Sussman,

Joel L. Sussman is Pickman Chair in Structural Biology. He earned his Ph.D. in 1972 from MIT, Cambridge, in Biophysics, under the supervision of Prof. C. Levinthal. From 1976-80, he worked as a Senior Scientist at Deptt. of Structural Chemistry, WIS. He became Professor in 1992 in Deptt.of Structural Biology at Weizmann Institute of Science. He remained Head of Protein Data Bank, Brookhaven National Laboratory, Upton, NY from 1994-99. In 2002, he became the Director of Israel Structural Proteomics Center. In 2005, he was the Member of Editorial Board of journals Proteins, PEDS, PLoS ONE and Evolutionary Bioinformatics. From 2005-08, he was the Coordinator of EC Project – Forum for European Structural Proteomics. From 2006-09, he remained a Member of Executive Board, Int'l Structural Genomics Org. From 2008-11, he was a Member EMBO committee Young Investigators Programme. In 2009, 11, 13 he was the Chairman of ERC Advanced Grant Program - Life Science. He was awarded Bergmann Prize in 1979 for his Outstanding Research in Chemistry in Israel. In 1991, he was given U.S. Army Science Conference Award for outstanding research. In 1994, he was elected Member of European Molecular Biology Organization (EMBO). In 2000, he was awarded Professor-Honoris Causa, by Chinese Academy of Sciences, Shanghai. He got "Elkeles Prize" in 2005 for being Outstanding Scientist in Medicine in Israel, with I. Silman. In the following year, he was awarded "Teva Founders Prize" for Breakthroughs in Molecular Medicine, with I. Silman. He gave Keynote Lecture in 2009 in Asian Crystallography Conf, Beijing. He delivered EMBO Keynote Lecture in 2009 at 10th Int'l Cholinesterase Conf, Sibenik. In 2010, he delivered EMBO Keynote Lecture at 3DSig Int'l Conference, Boston, MA. Bioinformatics in specific: He was a Member of the NIH Task Force & Council on Structural Genomics Initiative (2000). In 2002, he was the Chairman of Israel Council of Higher Education Bioinformatics Committee. From 2000-2011, he remained the Member of SESAME Beam Line Committee.

Function and Structure of intrinsically *Flexible* Proteins (IFPs)

Dr. Joel L. Sussman

Studies during the last decade have identified a family of neural cell adhesion proteins, which are single-pass transmembrane proteins, with substantial sequence similarity to cholinesterases (ChEs). The regions of sequence similarity correspond to only part of their complete sequences, thus establishing the ChE domain as a modular domain incorporated into different proteins, *i.e.* cholinesterase-like adhesion molecules (CLAMs)1,2. CLAMs, however, are devoid of catalytic activity, since they lack one or more residues of the catalytic triad. They appear to play key roles in the earliest stages of the development of the CNS and mutations, in the ChE domain of one of them, *i.e.* neuroligin, has been associated with autism.

The cytoplasmic domains of CLAMs bear no sequence homology to any known protein, and physicochemical studies show that they are, in fact, '*Intrinsically Flexible Proteins* (IFP)3 (often referred to as *Intrinsically Disordered Proteins*, or IDPs) when expressed in *E. coli*1,2. It has been estimated that a large percentage of cellular proteins exist in this *flexible* state; *e.g.*, for mammals, about half of their total proteins are predicted to contain long flexible regions (>30 residues) and ~25% of their proteins are predicted to be fully flexible (disordered)3. We developed a web-tool, *FoldIndex*© (http://bioportal.weizmann.ac.il/fldbin/findex)4, which has proven to be very useful in predicting regions of a protein sequence that are likely to be highly flexible. We have applied *FoldIndex*© to examine the CLAMs family as well as cholinesterase molecules. These 'in silico' studies will be compared with our recent solution studies on CLAMs and their adhesion partners, as well as our studies, in general, on the life-time of IFPs *in vivo*5-7. FoldIndex© is also being used in a routine way in the ISPC (http://www.weizmann.ac.il/ISPC) to aid in crystallization of proteins by first predicting which regions of a protein sequence are likely to be intrinsically flexible (disordered) and, thus, not including these regions in the construct that is cloned. Examples of IFPs will be shown in a new web tool, which we've developed, *Proteopedia*, the collaborative 3D encyclopedia of proteins & other molecules8 (http://www.proteopedia.org).

Dr. Vani Brahmachari

Dr. Vani Brahmachari completed her Ph.D. under the guidance of (late) Prof. T. Ramakrishnan from Indian Institute of Science, Bangalore. Working on post-transcriptional modification of transfer RNA in *M.tuberculosis* and *M.smegmatis*, she sequenced initiator tRNA to decipher the absence of thymidine in TΨC loop and the unique sequence of the anti-codon stem. After brief post-doctoral research at Wistar Institute and MRC, Mammalian Development Unit, London on Royal Society Fellowship, she became a faculty at the Indian Institute of Science, Bangalore, in the department of the then Developmental Biology and Genetics Department (current MRDG). Presently she is a Professor at Dr.B.R. Ambedkar Centre for Biomedical Research, at the University of Delhi.

Her research interests are in the area of mechanisms of genomic imprinting, epigenetic regulation and the impact of epigenetic modulation on human genetic diseases. Her group combines *in silico* analysis and wet experiments in hypothesis driven research. She is an elected Fellow of National Academy of Sciences, India.

Integration of in silico and experimental analysis:
Mining the human genome for the tool kit for developmental gene regulation

Dr. Vani Brahmachari

The cellular/tissue identity is preserved through characteristic signatures of expression profile, maintained through development and differentiation. The conservation of developmental strategies observed across phyla makes it possible to mine the whole genome sequences for components of the developmental memory modules. The analysis of the molecular basis of development in model systems like the Drosophila has demonstrated that the initiation of regulated transcription and its maintenance through development are accomplished by different mechanisms. One such group of trans-acting factors for maintenance, are the protein complexes like polycomb and trithorax group of proteins that are conserved across phyla, from Drosophila to humans. Several homologues of these proteins have not only been identified in human cells, but their role in differentiation on one hand and maintenance of stem cell identity on the other has been demonstrated. However the cis-elements where these factors interact, namely the polycomb/trithorax response elements (PRE/TRE), were not detected in the mammalian genome till recently. We have mined the human genome for such elements by a combination of in silico analysis and experimental validation. We have functionally characterized these cis-elements in vivo. The genetic interactions show their dual ability to promote either repression or activation by interacting with different protein complexes. Our results will be discussed in the light of contextual function of these memory modules that ameliorate the demand and supply imbalance of the genome.

Dr. Stephen Bryant

Stephen Bryant is a Senior Investigator at the Computational Biology Branch of the National Center for Biotechnology Information at the National Library of Medicine. He received his BA in chemistry and English from the University of Virginia and a PhD in biophysics from Johns Hopkins University. Dr. Bryant's previous positions include research scientist at the Wadsworth Center for Laboratories and Research, New York State Department of Health, and Assistant Professor, Department of Biomedical Sciences, School of Public Health, State University of New York at Albany. From 1986-1988 he served as a Senior Research Associate, Protein Data Bank, Chemistry Department, Brookhaven National Laboratory. Postdoctoral Advisor: T. Koetzle.

PubChem: An Open Repository for Chemical Structure and Biological Activity Information

Dr. Stephen Bryant

PubChem is an online public information resource from the National Center for Biotechnology Information (NCBI). The system provides information on the biological activities of chemical substances, linking together results from multiple sources on the basis of chemical structure and/or bioactivity-profile similarity. Following the deposition model introduced by the GenBank genomic sequence database, PubChem's content is derived from user depositions of chemical structure and bioactivity results, including high-throughput biological screening results from the NIH Molecular Libraries program. PubChem’s information retrieval tools support basic searches by chemical names and structures, as well as searches by target names or other terms used in descriptions of bioactivity experiments. Additional tools provide structure-activity analysis within and between multiple sets of bioactivity results. PubChem provides further information on biological activities via links to other NCBI information resources, such as the PubMed biomedical literature database and NCBI's protein 3D structure database, as well as via links to depositor web sites.

Christoph Steinbeck

Christoph Steinbeck studied chemistry at the University of Bonn, where he received his diploma and doctoral degree at the Institute of Organic Chemistry. In 1996, he joined the group of Prof. Clemens Richert at Tufts University in Boston, USA, where he worked in the area of biomolecular NMR on the 3D structure elucidation of peptide-nucleic acid conjugates. In 1997 Christoph Steinbeck became head of the Structural Chemo- and Bioinformatics Workgroup at the Max-Planck-Institute of Chemical Ecology in Jena, Germany. In 2002 he moved to Cologne University Bioinformatics Center as head of the Research Group for Molecular Informatics. He is past chairman of the Computers-Information-Chemistry (CIC) division of the German Chemical Society, past trustee of the Chemical Structure Association (CSA) Trust, a lifetime member of the World Association of Theoretically Oriented Chemists (WATOC), a director of the Metabolomics Society and member of various editorial boards and committees. Today, Christoph is head of chemoinformatics and metabolism at the European Bioinformatics Institute (EBI) in Hinxton, Cambridge, UK. His group develops a number of the leading open source software packages in Chemo- and Bioinformatics, including the Chemistry Development Kit (CDK), a Java library for chemo- and bioinformatics, Bioclipse, an Eclipse-based Rich Client for everything and nothing in particular, and OrChem, a chemical database cartridge for the Oracle database system. They further develop open chemistry databases for the biosciences, such as ChEBI, the dictionary and ontology of Chemical Entities of Biological Interest, IntEnz and Rhea, databases dealing with Enzyme nomenclature and biochemical reactions, and the MetaboLights database, a repository and reference database for Metabolomics. The Steinbeck group's research is dedicated to natural products research, the elucidation of metabolomes by means of computer-assisted structure elucidation and other prediction methods, the reconstruction of metabolic networks and algorithm development in chemoinformatics.

Latest Developments in the Chemistry Development Kit (CDK)

Dr. Christoph Steinbeck

The Chemistry Development Kit (CDK) is a freely available open-source Java library for Structural Chemo- and Bioinformatics. Its architecture and capabilities as well as the development as an open-source project by a team of international collaborators from academic and industrial institutions is described. The CDK provides methods for many common tasks in molecular informatics, including

- 2D and 3D rendering of chemical structures,
- I/O routines,
- SMILES parsing and generation,
- ring searches,
- isomorphism checking,
- structure diagram generation,
- and much more.

The CDK forms the basis of a number of applications, like the open source workflow engine CDK-Taverna, the EBI's open chemistry database ChEBI, the Bioclipse workbench for molecular biology and the search engine OrChem.

Dr Manoj K. Bhasin

Dr. Bhasin is an Assistant Professor of Medicine at Harvard Medical School and Director of the Bioinformatics at Genomics and Proteomics Center of Beth Israel Deaconess Medical Center (Boston, MA, USA). He got PhD from Institute of Microbial Technology (Chandigarh, INDIA) in the area of Computational Biology. Prior to joining the BIDMC Genomics Center, he completed a postdoctoral fellowship at the Dana-Farber Cancer Institute in the area of genome wide prediction of methylation sites and T cell epitope identification. Dr. Bhasin is an experienced Bioinformatician with a strong track record in computational vaccine design, systems biology, functional genomics and proteomics. He is applying and developing systematic analysis strategies for high-throughput transcriptional profiling, epigenomics, genotyping, proteomics, and drug screening data to define disease mechanisms at a molecular level and to identify novel prognostic and predictive biomarkers as well as new drug targets in human disease such as cancer, cardiovascular, and diabetes. In collaboration with leading investigators at Harvard Dr. Bhasin had worked on variety of approaches to identify gene/proteins and microRNAs that are currently being be evaluated as biomarkers and/or targets for therapeutic inventions. In addition to collaborative projects, Dr. Bhasin has strong interests in development of computational methods for systems level integrative analysis of multidimensional genome level data (e.g. mRNA, miRNA, methylation profiling and proteomics) and identifying novel diagnostics, progression and therapeutic biomarkers for renal and pancreatic cancers. To date, Dr. Bhasin has >50 peer-reviewed scientific publications/conference abstracts. He will be giving a talk at OSCAT-2012 about his research in different diseases (e.g. Cancer, Diabetes) showing importance of integration of experimental and informatics approaches to accelerate the discovery of potential biomarkers and therapeutics.

Integration of *Insilico* and Experimental Approaches to Drive Discovery of Biomarkers And Therapeutics.

Dr Manoj K. Bhasin

In the last decade, completion of the human genome project fueled the advent of advanced functional genomics and proteomics technologies (e.g. microarray, next generation sequencing and global proteomics). These have considerably increased our understanding of molecular mechanisms of cancers and sparked a great deal of interest in new targets for evaluation as potential biomarkers. Despite these studies, only a handful of new biomarkers have been validated and translated into clinical application. The validation of clinical usefulness of large number of markers yielded from genome level experiments needs time, money and labor intensive experiments. A major bottleneck is making sense of vast amounts of data generated in a typical "omics" experiment and integrating the complex biological information while utilizing multiple "omics" approaches. Systems biology methodology is the only solution to hasten the discovery of the biomarkers on the basis of clues or evidences from multilevel omics data, data mining, experimental and clinical information. We have performed omics experiments and devised systematic systems-level approaches to identify the potential biomarkers for i) early diagnosis ii) response to therapy iii) onset of resistance to therapy for renal and pancreatic cancers. All these biomarkers are currently undergoing experimental validation in xenograft mouse models to priorities the biomarkers for clinical trials. The first part of my talk will be focused on functional genomics experiments and systematic analysis approaches utilized to identify new biomarkers for early diagnosis, onset of resistance to therapies and identify novel therapeutics for treatment of cancer. Second part of the talk will provide overview of our work to identify biomarkers for the treatment of autoimmune type 1 diabetes (T1D). Major challenges in Type 1 diabetes (T1D) include the lack of reliable biomarkers to predict T1D onset and new therapeutics that efficiently prevent T1D onset or cure frank T1D. There is an unmet need to better understand how therapies reverse T1D and to develop more effective therapies targeting key pathways linked to T1D reversal. We have determined the effects of the different treatments on the transcriptome in several tissues, in order to identify T1D biomarkers that respond to therapy as well as gene expression changes that are affected by T1D and reversed by treatments. The innovative systems level analysis of the genome level signatures identified TNF-alpha as a novel target for the treatment T1D.

Dr Ahmed Kamal

Ahmed Kamal graduated from Osmania University, Hyderabad (India) and obtained his Masters degree in Organic Chemistry. He subsequently did his Ph.D. research in the area of Medicinal Chemistry from the Aligarh Muslim University. He later joined as a Scientist at the Indian Institute of Chemical Technology (IICT), Hyderabad. He carried out his post-doctoral research work at the University of Portsmouth, England and was a visiting scientist at the University of Alberta, Edmonton, Canada. For the last 25 years, he has pursued his research career at IICT, Hyderabad and is presently working in the capacity as Scientist-H in IICT. He also holds an additional charge as Project Director of NIPER, Hyderabad. His research interests mainly focus on the design and synthesis of gene-targeting compounds as new anti-cancer agents, and also design and synthesis of anti-tubercular compounds and biocatalytic transformations. He has won several honours and awards for his research ; CSIR Young Scientist Award in Chemical Sciences (1991), Fellow of the National Academy of Sciences, India (1999), Ranbaxy Research Award in the field of Pharmaceutical Sciences (2005), OPPI Scientist Award from the Organization of Pharmaceutical Producers of India (2009) and Fellow of the A. P. Academy of Sciences (2010). He has over 270 publications, 11 review papers, and 3 book chapters in the areas of biotransformations and medicinal chemistry. He has filed over 70 patents and five of his US patents have been licensed and some compounds from these patents are undergoing preclinical studies. He was instrumental in setting up of Biotechnology Incubation centre at S.P. Biotech Park at Hyderabad. About 50 students have completed their Ph.D. work under his guidance and over 20 of them are currently working for their Ph.D. programme in different universities. He is an editorial advisory board member for the journal “Letters in Drug Design and Discovery”.

Development of New Antitubercular Agents

Dr Ahmed Kamal

Tuberculosis has been considered as an important chronic communicable disease in a number of countries. In the last decade, increase of tuberculosis coinciding with AIDS epidemic has resulted in additional drug resistant isolates of *Mycobacteriumtuberculosis*. This and lack of new drugs available for the treatment have generated a renewed interest in a strategic search for prototype leads. Recent advances have been pointed out some additional drug targets that have potential for improved therapy.

There is a dire need to develop new, faster acting chemotherapeutics with lower toxicity that can be administered with other drugs. In recent years, some new class of compounds based on fluoroquinolones, nitroimidazoles, phenazines, thiazoles, oxazolidinones, benzothiadiazines and thiolactones have emerged as potential compounds for the treatment of tuberculosis. Some compounds are presently in clinical development, while others are being investigated preclinically in an attempt to explore new molecules for the target based treatment of TB. Simultaneously, some new drugs are being identified and validated for their practical usefulness. As a part of investigations in this laboratory on the development of antitubercular agents, new class of different scaffolds have been designed, synthesized and evaluated for their activities. Some of these molecules have shown promising activity, particularly, for multi-drug resistant tuberculosis and are at different stages of preclinical studies.

Dr. Ena Wang

Dr. Ena Wang obtained her Medical Doctor degree at Hebei Medical University, China in 1983 and a Master of Medicine in 1989 at Shanghai Medical University (Fudan University). She came to the United States in 1991 as Visiting Research Scientist at the Department of Microbiology and Immunology, AHSC, University of Arizona and continued her academic career as Assistant research scientist in the same department till 1997. In 1998, she was granted a Cancer Research Fellow position at the Surgery Branch, National Cancer Institute, National Institutes of Health. She was promoted to Staff Scientist in 2001 and appointed as Director of Molecular Sciences in 2007 at the Infectious Disease and Immunogenetics Section (IDIS), Department of Transfusion Medicine, Clinical Center, National Institutes of Health. She is also one of the pioneers of Trans NIH initiative, the Associated Director of Center for Human Immunology and Omic Facility Head at NIH since 2009. She has contributed more than 22 book chapters and published more than 170 peer reviewed articles. The focus of her research is the identification of genetic traits in humans that could explain the relationship between pathogens and the host with particular interest in cancer and chronic infections. This analysis spans the study of germ line characteristics that could explain the influence of genetic background of the host in modulating disease evolution; the analysis also includes of the study of genetic alterations of cancer cells or pathogens at times relevant to disease outcome or response to therapy. Her work has focused extensively on the variation of cancer cells during its natural evolution or in response to therapy and developed technologies for the real-time analysis of evolving cancer phenotypes ex vivo in patients with melanoma and other cancers.

Paradigm shift in Translational Medicine powered by computation and bioinformatics

Dr. Ena Wang & Francesco M Marincola

Fundamental strides in the understanding of the molecular basis of tumor rejection were made in the last decade thanks to observational studies performed at relevant time points in human cancerous tissues and application of cutting edge high throughput technologies. As overwhelming multi-dimensional data accumulated and the complexity of variables in study human make computation biology and bioinformatics the ultimate tool in meaningful data generation and stratification. In the past decade, we have been using system biology approach to study the significance of transcriptional signatures observed in pre-treatment biopsies as predictive of responsiveness to biological therapy and found that the transcriptional signatures observable during and after therapy documenting the switch from chronic to acute inflammation that leads to tumor rejection. Those observations have also been reported in chemotherapy and viral oncolytic therapy, both believed to eliminate tumors exclusively through direct cytotoxicity may play an adjuvant role in stimulating this inflammatory switch. Moreover, mechanisms leading to immune mediated tumor rejection largely overlap those associated with other aspects of immune-mediated tissue-specific destruction such as allograft rejection, graft versus host disease, acute clearance of pathogen and autoimmunity. Immune-mediated rejection of human cancer is a reality that occurs reproducibly in specific model systems such as the regression of melanoma and RCC in response to systemic IL-2 therapy, BCC treated locally with TLR-7 agonists or the adoptive transfer of EBV-specific CTL in lymphoproliferative disorders. It is likely that each model system has its own idiosyncrasies but, at the same time, commonalities dictate the final outcome of rejection. Similarly, other tissue destructive immune pathologies seem to share common effector pathways. Understanding the basic mechanisms that can switch a chronic inflammatory process incapable of eradicating its cause into an acute reaction with the power of destroying completely the triggering cause, may shed insights that may guide the development of novel therapeutic strategies. Even more importantly, identifying the mechanisms that lead to this final common pathway in individual tumors may define a better rational for targeted therapies that may take advantage of each individual cancer’s biology.

Dr. Robert Glen

Robert Glen gained his Ph.D. in X-ray Crystallography and Organic Synthesis from the University of Stirling (Scotland). One of the highlights was the first co-crystallisation of a reactant and product of a chemical reaction in a single crystal. At the Wellcome Foundation he built the Computer-aided Molecular Design group. This included Protein Crystallography, Molecular Transport properties and Electrochemistry. He invented the GASP and GOLD computer programs (a BBSRC funded grant, which has over 2500 citations) which are used extensively in the pharmaceutical industry, was a co-inventor of Zomig (AstraZeneca, a drug for migraine with $6Billion in sales) and invented two other compounds that have entered Phase-2/3 clinical development. He assisted in setting up three biotechnology companies (Arena Pharmaceuticals, Phase-1 Molecular Toxicology and Signase), obtained and directed a BBSRC grant of $3.2M with University College London (discovering activators of soluble guanylate cyclise and working with The Technology Partnership to develop the Baseplate robot) and managed collaborative research and contract research in drug discovery with many large Pharmaceutical companies including the design and synthesis of the first commercial screening library. In 1999 he moved to the University of Cambridge as Director of the Unilever Centre. He has personally published about 150 papers and patents, served as deputy chairman of Lhasa Ltd., on the SAB of a number of biotechnology companies and international academic advisory and grant awarding bodies (Netherlands Genomics Initiative, NIH Roadmap, etc.), on the editorial board of Journal of Chemical Information and Modelling and the Encyclopaedia of Computational Chemistry, a fellow of the Royal Society of Chemistry. He is governor of the CCDC, an honorary fellow of the American Association for Cancer Research and a Fellow of Clare College Cambridge. He has been a consultant to a number of pharmaceutical companies.. He has additionally received over £12M in grants over the last ten years, from EPSRC, BBSRC, DTI and Pharmaceutical and software companies.

Probing the biology of the Apelin (APJ) receptor through simulation, design, synthesis and pharmacological evaluation"

Dr. Robert Glen

Much of early drug discovery depended on leads from natural sources e.g. willow bark yielded aspirin, which led to numerous non-steroidal anti-inflammatory drugs. However, when presented with a newfound biological target, with no small molecule leads, where do we start? How do we probe the target without suitable pharmaceutically active molecules? One approach is to model the receptor and associated endogenous ligands, using a range of methods from molecular dynamics simulations to pharmacophore analysis, to understand the criteria for binding (and mechanism of action) combined with compound design and selection to elucidate the Structure Activity Relationships (SAR). We have taken this approach to design peptide ligands for the Apelin(APJ) receptor, the most potent vasoconstrictor receptor yet discovered. We have combined this methodology with access to ethically sourced human tissue, an approach which eliminates many of the problems associated with animal testing and which also allows investigation of not only healthy, but diseased tissue. Drugs can then be targeted at the diseased state, which is more relevant in a clinical setting. We will describe the SAR of Apelin, the discovery of the first competitive antagonist and introduce the new concept of biased agonism and disclose compounds showing this property at the APJ receptor.

Dr Sourav Pal

He has done his Masters from Indian Institute of Technology (Kanpur) in 1977. He received his doctorate from Indian Association for the cultivation of sciences(Calcutta)(IACS) under the supervision of Debashis Mukherjee. he has done his Post-Doc in 1986 in University of Florida with Prof. R. J. Bartlett. He is currently director of National Chemical Laboratory, Pune. He is also an adjunct professor at the Indian Institute of Science Education and Research (IISER), Pune. Recepient of the Shanti Swarup Bhatnagar Prize in Chemical Sciences, 2000 Recepient of JC Bose National Fellowship of DST, 2008 Recepient of Chemical Research Society of India Silver Medal, 2009 Elected as a Fellow of the Indian National Science Academy, New Delhi, 2003 Elected as a Fellow of the National Academy of Sciences, India, Allahabad, 1998 Elected as a Fellow of the Indian Academy of Sciences, Bangalore, 1996 Elected as a Fellow of the Royal Society of Chemistry, 2011 Received Dr. Jagdish Shankar Memorial Lecture of the Indian National Science Academy, 2006 Recepient of Bimla Churn Law memorial Lecture Award of IACS, Kolkata, 2005 Dai-Ichi Karkaria Endowment Fellow of UICT, 2004-05 Recepient of the Chemical Research Society of India medal, 2000 Elected as a Fellow of the Maharashtra Academy of Sciences, 1994 Recepient of the NCL Research Foundation Scientist of the year (1999) award Recepient of the P.B.Gupta Memorial lecture Award of the Indian Association for the Cultivation of Science, Calcutta for 1993 Received Council of Scientific and Industrial Research (CSIR) Young Scientist award in Chemical Sciences for 1989 Received Indian National Science Academy (INSA) medal for Young Scientist 1987 Received NCL Research Foundation Best Paper Award in Physical Sciences for the year 1995, 1996, 1997, 1999, 2000, 2002 Delivered Prof. R. P. Mitra Memorial Lecture, Delhi University, 2010 Delivered Prof. N.R.Dhar Memorial Lecture, 2011, University of Allahabad, 2011 Recepient of "Science Councillor" Award -2011 of The Indian Society of Health Environment, Education & Research (ISHEER) Jodhpur centre.

Hydration Behaviour of different head groups of phospholipids using Density Functional Theory: A validation from Fukui functions for prediction of active sites of hydration and drugs.

Deepti Mishra, Susanta Das, Sailaja Krishnamuthy and Sourav Pal

The hydration behaviour of different head groups of phospholipids has been studied. Due to the presence of different functional groups the intramolecular interactions of each phospholipid varies as electrostatics involved changes. We attempt here to study change in electrostatics and hydration behaviour of two neutral phospholipid head group models viz; Phosphatidyl ethanolamine (PE) and Phosphatidyl choline (PC) and one negatively charged head group viz; Phosphatidyl Glycerol (PG) in the presence of counter ion. The conformation of the head group noticeably changes with the addition of water molecule majorly near the phosphate group. There is a formation of clatherate like structure of water molecules around the polar groups of head region. It was also noticed that when more number of water molecules are present, both the hydrogen atoms of water molecule gets involved in H-bonding and makes a ice like structure. The structural parameters and charges calculated in this paper can be used for the reparamaterization of the present force field of the phospholipids.

We also validate the ability of local reactivity descriptors viz; Fukui functions for the prediction of active site of hydration on the head group of PE model system. The prediction of the site of attack using condensed Fukui functions for the nucleophile or elecrophile can be used to predict the site of drug interaction.

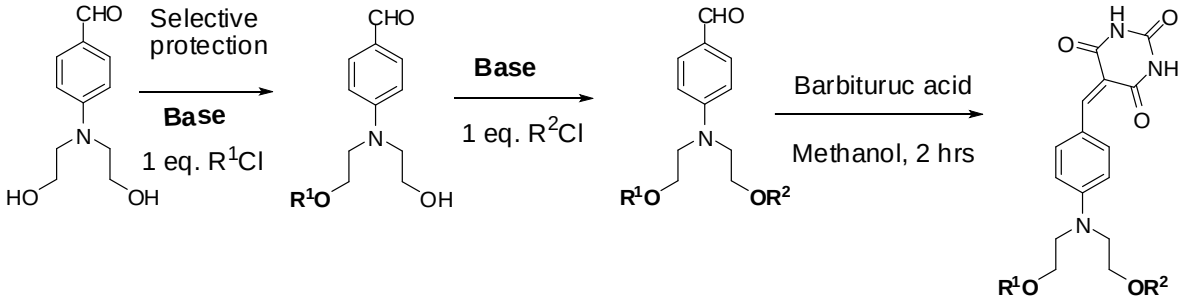
Prof. Prasad V. Bharatam

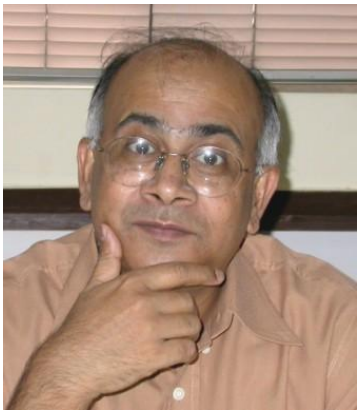
Professor Prasad V. Bharatam, FRSC, graduated from the chemistry department, of Visva Bharati, Santiniketan, India, in 1984. He obtained Ph.D. degree from University of Hyderabad, Hyderabad, India in 1990. After carrying out post-doctoral research at University of Alabama at Birmingham, USA, he took up research-cum-teaching position at Guru Nanak Dev University, Amritsar, India; later moved to NIPER. His research work involves the quantum chemical analysis of drugs, computer aided design and synthesis of anti-diabetic agents and anti-malarial agents. He is a recipient of AvH fellowship (2002), IBM Faculty award (2007), CRSI medal (2008). In his lab, students carry out synthesis of compounds designed by molecular modeling techniques and provide proof of concept for the computational recognized hypotheses.

Integrated approach to discover y-shaped PPAR γ agonists

Dr. Prasad V. Bharatam

PPAR γ is a well known target anti-diabetic activity. The cavity of this target has been found to be Y- shaped, flexible, expandable into sub pockets and various varieties of ligands can bind. Studies on the chemistry, biochemistry and basics of drug action of the glitazone series of compounds helped in the design of many new compounds. We have designed many new Y-shaped leads for the PPARγ agonistic activity. 3D QSAR methods have been employed to design several new leads. The designed molecules have been synthesized and biologically evaluated. The pharmacophoric features of metformin and other related drugs have been explored. The design of GSK3 inhibitors also was taken up. The chemistry and biochemistry of several of the anti-diabetic agents was explored using quantum chemical methods. The results of computer aided design, synthesis and biological evaluation of PPARγ activators will be presented in this lecture.





Dr. Dinakar M. Salunke

Dr Dinakar M Salunke did his Ph.D. from Indian Institute of Science, Bangalore, India in 1983.

Fellowship/ Membership of Professional bodies

- Fellow, Indian National Science Academy (2004)
- Fellow, Indian Academy of Sciences (2001)
- Fellow, National Academy of Sciences (India) (1995)
- Member, Molecular Immunology Forum (1995)
- Member, Guha Research Conference (1993)

Awards and Honours

- National Bioscience Award (1999)
- Shanti Swarup Bhatnagar Prize for Biological Sciences (2000)
- Ranbaxy Research Award for Basic Research in Medical Sciences (2002)
- Outstanding Scientific Achievements, National Institute of Immunology (2000 & 2001)
- Prof RC Shah Memorial Award (2000)
- Dr CR Krishnamurthi Oration Award (2004)
- Dr AT Varute Oration Award (2005)
- JC Bose National Fellowship Award (2007)
- SK Mitra Birth Centenary gold medal (2009)
- Professor GN Ramachandran 60th Birthday Commemoration Medal (2009)

Molecular mimicry and ligand specificity: experiences from immunological investigations

Dr. Dinakar M. Salunke

Antigen recognition and subsequent affinity maturation interface physico-chemical principles of molecular interactions with the physiological processes associated with self-nonself discrimination. We have addressed antigen-antibody interaction in the context of breakdown in the antigenic discrimination using crystallographic approaches. The results show intriguing new aspects of antigen recognition with implications in the context of ligand specificity in receptor binding. Carbohydrate-peptide mimicry in the context of humoral antibody response was extensively investigated using a sugar moiety, methyl α-D-mannopyranoside and a dodecapeptide, DVFYPYPYASGS. Thermodynamic and kinetic binding analyses of mimicry recognizing monoclonal antibodies, generated against the sugar, provided interesting insights and possible mechanistic model for molecular mimicry based on flexibility of the antigen combining site. Comparative crystallographic analyses revealed that the antigen combining site to be capable of specifically recognizing chemically disparate sugar and peptide molecules interacting at the overlapping surface using mutually independent interactions implying a functional equivalence in the absence of structural mimicry. Thus, the existence and prevalence of multi-specific, degenerate mature antibodies in the secondary immune response was evident, as also the receptor plasticity while accommodating a diverse ligand topologies. While this could be one of the mechanisms adopted by the immune system to neutralize the escape mutants generated during pathogenic insult by generating multi – specific receptors, it also highlights complexities involved in rational drug design approaches based on topological principles of receptor specificity.

Dr. Ram Vishwakarma

Dr. Ram Vishwakarma is Director of the CSIR - Indian Institute of Integrative Medicine Jammu (India) since March 2009. Prior to joining IIIM, he was working as Vice-President (Medicinal Chemistry) at Piramal Life Sciences Ltd, Mumbai (2005-2009) on NCE discovery projects on cancer, inflammation and drug-resistant infections. Prior to this, he was a staff-scientist at the National Institute of Immunology, New Delhi (1994-2005) working on the chemical-biology of GPI anchors. Ram did his Ph.D. in medicinal chemistry from Central Drug Research Institute, Lucknow followed by the post-doctoral studies at the University of Cambridge on biosynthesis of cyanocobalamin (vitamin B_{12}). He has over 28 years of experience in new drug discovery, medicinal chemistry, organic synthesis, chemical biology and glycobiology. Ram has worked in England, Canada and USA and has over 100 research publications and 20 patents to his credit.

Efforts on Antibacterial drug discovery at CSIR-IIIM

Dr. Ram Vishwakarma

In our efforts to identify novel anti-TB and antibacterial scaffolds, we have recently completed the whole cell screening of library of drug-like small molecules against *M. tuberculosis* and drug resistant bacteria. This screening led us to the identification of 9 new as well as few new-use bactericidal scaffolds. Medicinal chemistry programs have been directed towards optimization of some of these hits to the lead candidates. The institute is also pursuing the target based programs for a few validated targets of *M. tuberculosis*. These targets include 20S Mtb proteasome, ATP synthase and shikimate kinase. We have also initiated a discovery program focused on Gram negative bacterial target, UDP-3-*O*-(*R*-3-hydroxyacyl)-*N*-acetylglucosamine deacetylase (LpxC). This enzyme catalyzes the committed reaction of lipid A (endotoxin) biosynthesis in Gram-negative bacteria. Natural products chemistry and medicinal chemistry efforts are on to identify novel class of broad-spectrum potent LpxC inhibitors. Overview of some of these drug discovery programs at IIIM will be presented in the talk.

Dr. Jeremy Frey

Jeremy Frey is A Professor of Physical Chemistry at the University of Southampton, UK He is committed to a collaborative and interdisciplinary approach to chemical research. The interactions with Physics, The Opto-Electronics Research Centre (ORC), and Electronics and Computer Science have been particularly fruitful. His research is based on the use of laser spectroscopic techniques to probe molecular structure reactivity and dynamics and organization in a variety of environments from single molecules, molecular beam kinetics and photochemistry, to the study of interfaces and surfaces with interfacial non-linear spectroscopy. As part of his current research he is involved with the UK e-Science programme as PI of the CombeChem project looking at the ways in which e-Science and Grid infrastructure can be developed to provide support for and carry out chemical research, for example in Electronic Laboratory Notebooks (ELNs) with the Smart Tea Project generating and applying a "Semantic Chemical Grid" and applying Web 2.0 & Social Network ideas with Chemical Blogs and relatedtechnologies. Fundamental to the ideas of "Publication @ Source" for scientific data is his work on the interaction of e-print repositories with chemistry in the work on the e-Bank & e-Crystals projects His most recent laser research, involving higher order non-linear effects, is as the PI of a Basic Technology project to generate a nanoscale ultra short pulse of x-ray source using ultrashort-pulsed lasers and fibre technology aimed at probing the shape of single large molecules of biological significance, such as enzymes, using EUV and soft x-ray coherent diffraction imaging and x-ray spectroscopy. He was the chair of the UK e-Science User Group (2005-7) and in 2005/6 held a visiting Fellowship at the Centre for Mathematics and its Applications at ANU, Canberra. He has recently been appointed as the champion for the RCUK Digital Economy Theme on IT as a Utility.

The Semantic Web meets Wet and Dry Chemists
Dr. Jeremy Frey

The talk will cover his work in the e-Science programme on the way in which the semantic web can help chemists work more productively and share their results more widely while maintaining the proper context for their work. In particular he will discuss our new generation of versatile electronic laboratory notebooks and the general way in which semantic technologies help with the linking of chemical information.

OSDD Talks

Dr. G.P.S.Raghava

Dr. G. P. S. Raghava is a scientist working at Bioinformatics Centre, Institute of Microbial Technology (IMTECH), and Chandigarh, India. He did M.Sc. form Meerut University, M.Tech from IIT Delhi and PhD from IMTECH Chandigarh. He worked as Postdoctoral fellow at Oxford University UK (1996-98), Bioinformatics specialist at UAMS, USA (2002-3 & 2006) and visiting professor at POSTECH, South Korea (2004). His group developed more than 100 web servers, 100 research papers, 50 Copyrights, 10 databases and mirror sites. He is responsible for setting Bioinformatics infrastructure at IMTECH and at UAMS Little Rock, USA (http://bic.uams.edu/). He got following major awards/recognition i) Shanti Swarup Bhatnagar Award in Biological Science, 2008 ii) National Bioscience Award for Carrier Development, for year 2005-2006 (by Department of Biotechnology, Govt. India); iii) NASI-Reliance Industries Platinum Jubilee Award, 2009; iv) Thomson Reuters Research Excellence - India Research Front Awards, 2009; v) Fellow of National Academy of Sciences (F.N.A.Sc); and vi) Fellow of Indian Academy of Science (F.A.Sc.) Bangalore.

Chemoinformatics tools on CRDD portal

Dr. G. P. S. Raghava and CRDD Team,

Open Source for Drug Discovery (OSDD) Forum is an initiative with a vision to provide affordable healthcare (http://www.osdd.net/). Computational Resources for Drug Discovery (CRDD) is an important in silico modules of OSDD. It is being initiated with a notion to reduce the cost of drug discovery. CRDD provides a global platform where the best brains can collaborate and collectively endeavor to solve the complex problems associated with identification of potential targets to lead discovery for neglected diseases like Malaria, Tuberculosis, Leshmaniasis, etc. just under one roof. The use of bioinformatics has allowed the identification of some hidden potential targets; hence, CRDD has implemented the computational resources for experimentalists/researchers working in the field of computer-aided drug design to discover drugs against them. Moreover, CRDD offers web backed servers for most of the chemoinformatics and bioinformatics based java libraries which otherwise a complicated task for the biologist to implement.

Development of chemoinformatics tools is not as easy bioinformatics tools, it is because fundament software used for developing chemoinformatics tools are not part of public domain. For example most of the software packages used for computing descriptors of chemicals are commercial cannot be used for developing open source software. In order to overcome this challenge we initiate a mission for developing free ware in chemoinformatics. In most of our inhibitor prediction software we are using software packages Chemistry Development Kit (CDK) and PADEL for computing descriptor of chemicals.

Dr. David Wild

David Wild is an Assistant Professor of Informatics and Computing at Indiana University School of Informatics. He directs the Cheminformatics and Chemogenomics Research Group (CCRG), with approximately 15 students focused on large scale data mining and aggregation of chemical and biological information. He also directs educational programs in cheminformatics at the University, including an innovative distance education program He has been PI or CoPI on over $1.7m of funding, and has over 30 scholarly publications. He is Editor-in-Chief (along with Chris Steinbeck at the EBI) of the Journal of Cheminformatics, and works as editorial advisor or reviewer to many journals. He is involved in several cheminformatics organizations including the Chemical Structure Association Trust and the American Chemical Society. He has helped organize many conferences and symposia in this field, and has recently acted as an expert witness in cheminformatics. He is also the director of Wild Ideas Consulting, a small scientific computing company specializing in informatics and cheminformatics.

Exploring semantic networks of public data for drug-target prediction
Dr. David Wild

We have developed a systems chemical biology data resource called Chem2Bio2RDF (www.chem2bio2rdf.org) that integrates publicly available datasets pertaining to chemical compounds, drugs, drug side effects, targets, genes, pathways, diseases and scholarly publications. The dataset is semantically annotated using ontologies including a new chemogenomic ontology called Chem2Bio2OWL. We have developed a variety of graph-based and other network algorithms to look for chemogenomic and other associations in this data, a method called SLAP for prediction of drug-target association, polypharmacology modeling, and profiling. In this talk, I will describe how the Chem2Bio2RDF resource and its associated algorithms can be applied in drug discovery problems.

Dr. Zakir Thomas

Zakir Thomas is the Project Director of the Open Source Drug Discovery and Head of the Director General's Technical Cell. He has over 22 years of experience in executive positions in government, including as Registrar of Copyrights. He has been involved in implementing major e-governance projects of the government of India.

Introduction to OSDD

Dr. Zakir Thomas

The lack of innovation in diseases that primarily affect the tropical regions has been a subject of discussion for some time. Several solutions were proposed but none has emerged with an alternate model of innovation. Pharmaceutical industry has not yet found a model that is sustainable, though in sheer numbers the disease affects a large number. But in terms of market, the numbers are small. Open Source Drug Discovery (OSDD) proposes an approach that is not patent driven but caters to the market forces to ensure accessibility and affordability of drugs. This talk will focus on the rationale of OSDD and the approach it practices.

Dr. S. Ramachandran

He joined the CSIR-IGIB (then CBT) in 1999 and started his investigations into genomics. His interest had been in investigating the evolution of host-pathogen interactions. He was also interested in investigating the role of repeats in human transcription and also in pathogens. Recently, he is using the powerful methods developed in R platform for analyzing big data flowing through the NGS platforms. He has been fortunate delivering over 80 lectures in National and International conferences and workshops and credited with more than 40 publications including granted patents and copyrights. Many students trained with him are now well placed in various Universities or Institutions all over the World.

Computational screening for new inhibitors of M. *tuberculosis* mycolyltransferases antigen 85 group of proteins as potential drug targets

Shachi Gahoi, Rahul Shubhra Mandal, Nikita Ivanisenko, Priyanka Shrivastava, Sriyans Jain, Ashish Kumar Singh, Muthukurrusi Varieth Raghunandanan, Swarna Kanchan, Bhupesh Taneja, Chhabinath Mandal, Vladimir A. Ivanisenko, Anil Kumar, Rita Kumar [Open Source Drug Discovery Consortium], Srinivasan Ramachandran

The group of antigen 85 proteins of *Mycobacterium tuberculosis* is responsible for converting trehalosemonomycolate to trehalosedimycolate, which contributes to cell wall stability. Here we have used a serial enrichment approach to identify new potential inhibitors by searching the libraries of compounds using both 2D atom pair descriptors and binary fingerprints followed by molecular docking. Three different docking softwares, AutoDock Vina, GOLD and LigandFit were used for docking calculations. In addition, we applied the criteria of selecting compounds with binding efficiency close to the starting known inhibitor, and showing potential to form Hydrogen bonds with the active site amino acid residues. The starting inhibitor was ethyl-3-phenoxybenzyl-butylphosphonate, which had IC_{50} value of 2.0 µM in mycolyltransferase inhibition assay. Our search from more than 34 million compounds from public libraries yielded 49 compounds. During this time we tried to automate this 2D descriptor based similarity searching process and consequent docking exercise and developed ParallelVSR a pipeline for parallel virtual screening on R-platform which is faster as it runs on multiple processors and need minimal manual interventions. Subsequently selection was restricted to compounds conforming to the Lipinski rule of 5 and exhibiting hydrogen bonding to any of the amino acid residues in the active site pocket of all three proteins of antigen 85A, 85B and 85C. Finally we selected those ligands which were ranked top in the table with other known decoys in both the docking results. The compound 415032 from Tuberculosis Antimicrobial Acquisition and Coordinating Facility was further examined using molecular dynamics simulations for 10 nanoseconds. These results showed that the binding is stable, although some of the hydrogen bond atom pairs varied through the course of simulation. The NIH415032 has anti-tubercular properties with IC_{90} at 20 µg/ml. These results will helpful to the medicinal chemists for developing new anti-Tubercular molecules for testing.

ParallelVSR: a pipeline for parallel virtual screening on R – platform

Srikant Prasad Verma, Ashish Kumar Singh, Vikas Pandey, Bhanwar Lal Puniya, Open Source Drug Discovery Consortium, Srinivasan Ramachandran

The rapidly growing size of small molecule databases has significantly broadened the chemical search space and therefore new strategies need to be developed for efficient virtual screening. In recent past, a large number of screening tools have been developed to address the issue. Besides accurate identification of novel therapeutic molecules, speed and technical barrier of the existing virtual screening tools are major bottlenecks. Recent developments in this field have shown that an integrated approach of using different screening techniques together can allow efficient study of large chemical libraries**.** Thus a freely available, fast and easy to use pipeline, which is based on this integrated approach, has potential for use in the field of drug discovery.

We present ParallelVSR, a cheminformatics pipeline which can perform virtual screening on R platform by sequential integration of ligand based virtual screening (LBVS) and structure based virtual screening (SBVS). Given an already known inhibitor of a target protein, it performs fast LBVS using ChemmineR, which uses 2D atom pair descriptors based similarity, to reduce a large chemical library of small molecules to a manageable size, followed by relatively slow SBVS using AutoDock Vina to rank the screened molecules according to the free binding energy of the corresponding protein-ligand complex. To further enhance the overall speed of this screening study, we have used data level parallelization for both LBVS and SBVS using Message Passing Interface (MPI) as communication protocol. We have analysed the performance of our pipeline at LBVS using chemical libraries consisting of 1,601,185 drug-like molecules from ChemDiv and ChemBridge Corporation. We observed significant improvement in the speed of MPI enabled LBVS with increase in number of CPU cores. We offer both multi-threaded and MPI enabled SBVS. Furthermore, ParallelVSR comes with easy to run R scripts which can run on MPI – enabled system.With the added advantage of simplicity in using this pipeline, ParallelVSR is designed to meet the challenge of speed in virtual screening.

Anshu Bhardwaj

Dr. AnshuBhardwaj did her Ph.D. in 2008 from the Center for Cellular and Molecular Biology, Hyderabad, India. During her Ph.D. she developed algorithms for prioritizing single nucleotide markers for disease association specifically in context of mitochondrial genome. Right after her Ph.D., she joined the CSIR's Open Source Drug Discovery (OSDD) Project as a Principal Investigator and worked towards designing and implementing community collaborative platforms and was a key player in community building. She has developed several databases and applications involving large number of students and researchers through online virtual collaboration. One of the major highlights of her contribution to the OSDD project is the conceptual design and implementation of the Connect to Decode Project where a large number of researchers converged their efforts to developing a systems biology map for *Mycobacterium tuberculosis*. As part of the community building, she has presented the concept of the OSDD approach at various national and international platforms.

Open Source Drug Discovery: Community Collaborative Experiments and Resources

Dr. Anshu Bhardwaj, Open Source Drug Discovery Consortium

The prohibitive cost of drug development, high mortality rate, absence of fast acting drugs, current enduring chemotherapeutic treatment, poor diagnosis and ability of tuberculosis (TB) bacilli to adopt latent state and show variable resistance to existing therapy clearly demands new interventions and approaches to find better, fast acting drugs against TB. Given the complexity and the enormity of the challenge, we have devised a community driven approach to address this problem under the aegis of the Open Source Drug Discovery (OSDD) consortium. Keeping in view the fact that a systems level approach may lead to identification of potential drug targets in *Mycobacterium tuberculosis* (Mtb) a comprehensive collation, curation and analysis of protein-protein interaction and metabolic map has been attempted in order to identify potential drug targets and elucidate mechanisms of latency and resistance. Through a large community driven Connect to Decode (C2D) project, the information for generating the interactome and metabolome has been captured manually from literature. The accuracy and scope of assigning function was enhanced by using a combined evidence annotation based approach. This has been accomplished through an online collaborative platform, using Web 2.0 applications, which has successfully generated the most comprehensive Protein-protein functional interaction map and metabolic map of Mtb. The project brought together more than 800 students and researchers online to a common platform. This effort has been able to annotate more than 85% of the Mtb genome. Using this new data set we have built a protein interaction map of Mtb including more than 1400 proteins and identified a set of proteins, which are essential in maintaining the integrity of this protein network for the growth and survival of Mtb. We have also developed a platform for filtering out protein targets based on potential risk for the off target interaction of ligands with human genome and human microbiome. The metabolic map of Mtb has also been analyzed further for prioritizing potential targets. The talk will delve into the details of how the interactome and metabolic map were generated and results of the analysis of the same. It will also touch upon the other important resources developed by the OSDD community.

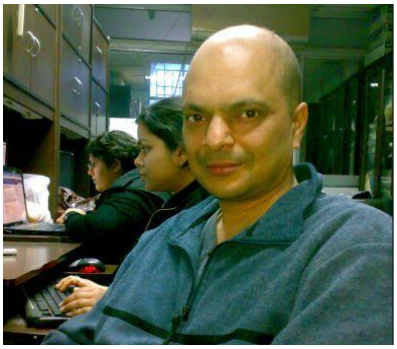
Dr. Vinod Scaria

Vinod Scaria is a scientist at the CSIR Institute of Genomics and Integrative Biology. He is a clinician turned computational biologist with research interests in translational genomics and genome informatics. He is interested in understanding the function and organisation of non-coding RNAs, and how genomic variations could potentially impact them. He is also involved in creating novel methods and resources for analysis and annotation of genomes and understanding the functional impact of genomic variations. He has used adopted novel and creative strategies like the use of social media and participation of large number of undergraduate students in collaborative projects to accelerate genome annotation and co-creation resources for genome annotation.

Genomics of the Unknown: Translational Applications of Hologenome sequencing

Ashok Patowary, Rajendra K Chauhan, Meghna Singh, Shamsudheen K Vellarikkal, Vinita Periwal, Sourav Ghosh, Kushwaha K P,Gajanan N Sapkal, Vijay P Bondre, Milind M Gore, Shahul Hameed, Urvashi Singh, Sridhar Sivasubbu and Vinod Scaria

Identification of pathogens from clinical samples has been the mainstay of clinical microbiology. Conventional processes for isolation and identification of the pathogens when no prior information exists is tedious, time-consuming and often costly, making it of limited applications in special settings like epidemics of new and emerging pathogens. High throughput sequencing technologies provide an alternate solution to the problem by offering the possibility of circumventing the problems of traditional pathogen isolation. In addition the costs of DNA sequencing have been exponentially dropping over the last decade with enormous improvements in the scale and ease of application. The widespread application of DNA sequencing in such settings had been majorly limited by the fact that traditional sequencing methods often required pure isolates for genome analysis, while clinical samples usually exists as mixed genome samples or hologenomes with 2 or more components. We show that effective modifications of the sample processing and analysis protocols would enable identification and assembly of pathogens from mixed genome samples. As a proof of concept we have applied this technique for the identification of Japanese encephalitis virus from a mixed population of genomes. Further we have extended this technique for the identification of diverse human and animal pathogens from their respective hologenomes. We show that this could be extended to special settings including epidemics. Apart from identifying pathogens, the technology would enable to potentially understand the dynamics of host-pathogen interactions. We have also in addition tried incorporating social media to track and predict epidemics, further improving the applicability of the technology to understand dynamics of epidemics and track emerging epidemics.



Dr. Andrew Lynn

Andrew Lynn is a volunteer with the Open Source Drug Discovery (OSDD) program, being involved from its inception. Within OSDD he has shared responsibilities associated with the technical infrastructure and project management through the Technical, Science and Budget committees. His group supports OSDD access to grid infrastructure, and provides methods for target, as well as ligand, based lead identification.

Outside of OSDD, Dr. Lynn presently heads the Information Technology infrastructure and services at Jawaharlal Nehru University. He also runs a research group focused on computational Biology, within the School of Computational and Integrative Sciences. His research interests involve cyber infrastructure for the biological systems, high performance computing and computational methods for biological sequence, structure and systems analysis.

Using grid computing to provision computational resources for community use in OSDD.

Dr. Andrew Lynn and Anmol Hemrom

The Open Source Drug Discovery is a platform to translate research as a result of social production into drugs for neglected tropical diseases. Social Production requires a modular, granular structure with a low cost of integration. The modules for upstream processes in drug discovery are data and computational intensive, and contributors in these areas require extensive resources especially in the areas of target-based virtual screening and chemoinformatics.

GARUDA is a collaboration of science researchers and experimenters on a nationwide grid of computational nodes, mass storage and scientific instruments. The OSDD community has been established as a Virtual Organization within GARUDA, with access to significant computation resources. Access to these resources has been provided through a portal based on the work-flow system GALAXY, after obtaining the requisite grid certification.

GARUDA has provided scalable processor and storage requirements for the OSDD community - however, there are limited applications available on the grid for drug discovery. To establish a process to continuously provide applications on to the grid, a pipeline to port applications required by the OSDD community has been developed. Using a staging system, members of the community can request or install applications of their choice in their home areas. Scripts implementing standard practices have been developed to port applications onto a staging cluster and then made accessible through the grid scheduler, finally making them available for community use through the portal.

This talk with cover the accessing of the grid resources, application porting and community building to translate computational predictions.

Dr. Haridas B. Rode

Haridas B. Rode did his Ph.D. (2002-2006) on the topic of "Medicinal Chemistry of Elastase inhibitors" at Ernst-Moritz-Arndt University, Greifswald, Germany under the supervision of Prof. Dr. Hans-Hartwig Otto. He has been a Post-Doctoral associate (2006-2010) at Chemical Genomics Centre of the Max Planck Society Dortmund, Germany under the Mentorship of Prof. Daniel Rauh. There his work was on "Chemical Biology of kinases". Awards include PhD fellowship from Ernst-Moritz-Arndt-University, Greifswald; Post-Doctoral fellowship from the Max-Planck Society, Dortmund; TVÖD position from the Max-Planck Society, Dortmund. He is a Member of the German Pharmaceutical Society (DphG).Presently he is a senior scientist (from april 2011 till date) at Council of Scientific and Industrial Research, HQ New Delhi. His scientific work includes Medicinal Chemistry related activities of OSDD and OSDDChem related activities of OSDD.

OSDD Chemistry

Dr. Haridas B. Rode

“Open source” has played a commendable role in many software products used worldwide, which includes Linux operating system, Firefox web browser and Internet resources such as Wikipedia. Research carried out in open source is faster as compared to the closed system where scientific community collaborates with a defined goal and with handful of scientists. The hierarchical ladder is absent or at its minimal in the open source, which helps scientist to discuss, collaborate and achieve the results with high efficiency. More importantly the project does not stop upon generation of some results but go on to fulfill overall mandate of the project as other members of open source community drive the activity to an end. This acceleration put-in open system is one of the biggest advantages of practicing science in open source. Pharmaceutical industry is commerce driven and drugs for diseases like tuberculosis, malaria and leishmaniasis may not generate attractive revenues for them. Hence, there is a need for an alternative drug discovery model in these disease areas. Open Source Drug Discovery (OSDD) may be an alternative model in drug discovery. It is utmost important to have the compounds for screening in drug discovery program. In order to have the variety of molecules to be synthesized and screened against TB, malaria, and other neglected diseases; an OSDD outreach program was launched in last year. The aim of outreach program is to impart practical training to a large number of master students in various universities and institutes across the length and breadth of the country. The synthesized compounds are stored in the national repository that has been created at CSIR-CDRI. OSDDChem portal is used for all the activities of outreach program. The different chemistry programs of OSDD will be presented.

Dr U.C.A. Jaleel

Dr U.C.A. Jaleel is an Asst Professor in hemoinformatics at the MCC calicut .Calicut university. He is currently working on the topic related with development of quantum chemical descriptors for QSAR analysis. He is also working in collaboration on a project for parameterization of computational tools of organo metallic compounds. He leads the extremely successful chemoinformatics connect-to-decode programme on OSDD.

Successes and challenges of online Cheminformatics research

Dr UCA Jaleel **.**

We attempts to do a systematic review of the successes and challenges of the online collaborative research progarmme of OSDD-Cheminformatics and Insilco drug discovery process. Problems encountered are classified into technical, scientific and administrative types. This paper also attempts to suggest immediate solutions to some of the problems and bottle-necks faced in the management of such a large open community. The future plan of OSDD based cheminformatics research, planning, management and discussion methodologies include the participation of various communities like OSDD-Garuda, OSDD-e learning, OSDD-WSF and OSDD-JSF. It also attempts to reach out to the open research community to come forward and help the amateur researchers with their expertise and proficiency to realize the lofty goals of the noble OSDD mission.

ORAL PRESENTATION

Microrna Targets: Potential Candidates for Indirect Regulation by Drugs

Amit Arora[1], GPS Raghava[2]

[1]Bioinformatics Centre, Institute of Microbial Technology,Chandigarh, India.

[2]Head, Bioinformatics Centre, Institute of Microbial Technology,Chandigarh, India.

Background:

Regulation of gene expression is central to maintaining homeostasis and to avoid disease. This regulation is achieved initially during the process of gene transcription by transcription factors and thereafter the gene expression is fine tuned by small non-coding RNAs like microRNAs (miRNAs). Whereas transcription factors are DNA binding proteins that inhibit or enhance the transcription of DNA into RNA, miRNAs are small non-coding RNAs that bind to the 3' regions of mRNAs and negatively regulate gene expression in many cellular processes. Not only are transcription factors and miRNAs distinct in their molecular composition and mode of action, differences lie in their gene structure and genomic locations as well. While transcription factors are distinct gene having their own regulatory apparatus, as much as 55% of miRNAs lie in the intronic regions of protein coding genes and are therefore under direct control of the host gene transcription mechanism.

Pharmacogenomics aims to predict molecules that could be safely and effectively used as therapeutic agents. These predictions are based on the genomic and transcriptomic profiles of diseased individuals and primarily focus on the role of single nucleotide polymorphisms (SNPs), copy number variations (CNVs) or relative gene expression levels of drug target or drug metabolizing genes.

It has been previously shown that ~13% of the FDA approved drugs target nuclear receptors-- a group of transcription factors that regulate a multitude of cellular processes from cell growth to differentiation to metabolism. These drugs, like most others, were isolated using screens designed to produce specific biological effects such as immunosuppression, rather that by screens that were meant to study their affect on the target transcription mechanism.

Knowing that complex regulatory networks, involving transcription factors and miRNAs along with RNA splicing machinery , are associated in producing the final gene product-- it is required to investigate how these regulatory networks and therefore the final expression of genes may be indirectly influenced by drug action.

This integrative approach, involving gene regulatory processes and drug actions, could unravel the hidden aspects of pharmacodynamics and could possibly expose distal genes whose expression may be influenced by drugs. Applied to pharmacogenomics, this has the potential to add another layer of investigations-- composed of drug induced gene regulatory processes.

Methodology:

As an initial study we investigate the drug induced perturbation of miRNA target genes. We undertake this study by carefully mining drug and gene regulatory process data-- targeting transcription factors and miRNAs. We start by initially focusing on the drugs that target transcription factors. Knowing that transcription factors target genomic regions thereby resulting in gene transcription-- drug based targeting of these transcription factors has the potential to also alter the targets of these transcription factors. Thereafter we manually create links between the target genes of drug targeted transcription factors (DTTF), and the miRNAs that reside between the genomic loci of these target genes of DTTF. Any alteration of transcription factors by drugs now also has the potential to perturb the miRNAs that reside in the genomic loci of transcription factor targeted genes. This mapping results in the creation of a network where information flows from external drugs to a multitude of genes.

In-house R scripts were used to extract data, find patterns and also for statistical analysis used in this study. miRBase (release 17) was used used to extract human and murine mature miRNAs. Transcription data was mined from literature and from online data sources.

DrugBank contains curated information about FDA approved drugs. Information regarding drug targets, its mode of action, etc was subsequently extracted from DrugBank data.

miRNA targets were predicted using different combinations of miRNA target gene sets from TargetScan miRanda and Diana.

NCBI GEO[http://www.ncbi.nlm.nih.gov/geo/] contains information about gene expression experiments. Gene expression data was downloaded from GEO after being suitably identified as pertaining to Drugged/Normal experiments by using suitable text mining scripts applied to literature.

Results:

We studied the targets of approved drugs (n=1569) only. As many drugs can have same target, there were 1477 drugs that had unique targets. It has been reported before that ~13% of the FDA approved drugs target transcription factors. This figure was verified as 194 drugs targeted

transcription factors (13.1%). The 194 drug targeted transcription factors further targeted 1218 unique genes. Some of the targets were themselves transcription factors and resulted in sub-networks.

Genomic location in the form of miRNA coordinate data was used to compute which miRNAs loci were located within the ORFs of protein coding genes. The 1218 target genes of drug targeted transcription factors (DTTF) revealed a set of 28 miRNAs that resided within their ORFs. When considering transcription factors, these 1218 genes corresponded to 23 drugs as well as 23 unique DTTF and that also have miRNAs in their genomic loci.

In order to detect any affect drugs might have on the miRNA target gene expression, it is required to undertake a relative gene expression analysis. Gene expression datasets, after being downloaded from GEO were normalized, compared with each other and filtered to enable cross platform and cross series analysis. Due to the skewed nature of expression datasets, non-parametric rank tests were employed to detect relative differences between drugged and control sets. Taking into consideration the mode of action of drugs on their targets (agonist, antagonist, substrate function, inducer, inhibitor, cleavage inducer), differences between median gene expression levels of miRNA target genes was observed. It was interesting to note the correlation between the mode of drug action and the difference observed in relative gene expression levels.

Conclusion:

This analysis shows that it could be beneficial while taking into consideration gene regulatory processes during screening for drug targets. Although only 23 drugs were targeting about 28 miRNA, the off-target affect of these drugs could potentially be vast, considering the hundreds of genes that are targeted by each one of these miRNAs. Although relative differences were observed between drugged and control expression of miRNA targeted genes, further analysis is required to verify the significance of these results. With next generation transcriptome sequencing techniques like RNA-Seq, better quality data may reveal more insights into the off-target affects of drugs and applied in conjugation with pharmacogenomic approaches, could reveal more insights to design personalized drug therapies.

Molecular Docking of Antitubercular Drug Tryptanthrin against DNA Recombination Protein Rec-A

Ankit Tripathi, Nupur Wadia, Deepak Bindal, Tarakanta Jana
CSIR-NISCAIR, 14-S.V.Marg, Qutub Institutional Area, New Delhi, India.

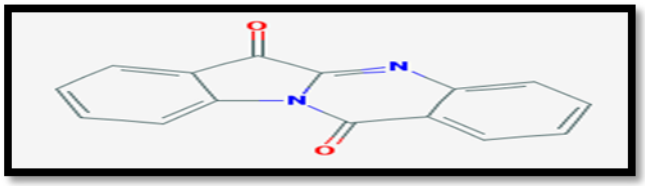
Abstract

DNA recombination protein RecA in *Mycobacterium tuberculosis* H37Rv, a protein plays an important role in the growth, regulation, differentiation of tuberculosis and is considered an important therapeutic target. In this study, we made *In-silico* molecular docking to see the interaction of known natural indoloquinoline alkaloid compound tryptanthrin against RecA protein *M. tuberculosis* H37Rv.

Introduction

Tuberculosis (TB) is an infectious disease that mainly afflicts populations in third-world countries [1]. Tuberculosis is a major public health problem in India. An increasing incidence of deaths due to TB and the known drawbacks of the current existing drugs including the emergence of multi drug-resistant (MDR) and extremely resistant (XDR) strains have led to a renewed interest in the discovery of new anti-tubercular agents with novel modes of actions. Recently, natural products have shown a useful and potentially rich source of anti-tubercular drug candidates, where alkaloids have been found more effective [2, 3].

DNA recombination protein RecA of *M. tuberculosis* H37Rv is a highly conserved, ubiquitous and multi-functional protein [4] involved in the co-ordinated expression of the SOS regulatory system in response to DNA damage and it also medicates DNA repair and



M.W.248.1

Figure 1. Structure of tryptanthrin
(indolo – [2, 1-b] – quinazoline -6, 12 -dione)

homologous recombination. Tryptanthrin (indolo [2, 1-b]-quinazoline-6, 12-dione) is a

compound (Figure 1) with a long history and is well documented to possess antibacterial activity against a variety of pathogenic bacteria [5]. The present study is to explore the affinity and interaction of tryptanthrin against *M. tuberculosis* recombination protein RecA.

Material & Methods:

Natural compound: Tryptanthrin (Indolo[2,1-b]-quinazoline-6,12-dione); Protein Preparation: The crystal structure of *M. tuberculosis* RecA protein taken in this study was retrieved from Protein Data Bank (PDB). Protein charges were calculated with the Gasteiger-Heckel method;

Ligand Preparation: Geometry optimization was analysed using PM6 method. Charges were analysed using PM6 charge calculation method and a pH 7.0. Docking: The natural inhibitor tryptanthrin taken for the study was subjected to dock within active site of RecA using docking server; Dock Analysis: Docking calculations were carried out using Docking Server [6]. The MMFF94 force field (Halgren, 1998) was used for energy minimization of ligand molecule tryptanthrin. Gasteiger partial charges were added to the ligand atoms. Non-polar hydrogen atoms were merged, and rotatable bonds were defined. Docking calculations were carried out on RecA (recombinase) protein model. Essential hydrogen atoms, Kollman united atom type charges, and solvation parameters were added with the aid of AutoDock tools (Morris, Goodsell et al., 1998). Affinity (grid) maps of 20×20×20 Å grid points and 0.375 Å spacing were generated using the Autogrid program (Morris, Goodsell et al., 1998). AutoDock parameter set- and distance-dependent dielectric functions were used in the calculation of the van der Waals and the electrostatic terms, respectively; Docking Validation: Docking simulations were performed using the Lamarckian genetic algorithm (LGA) and the Solis & Wets local search method (Solis and Wets, 1981). Initial position, orientation, and torsions of the ligand molecules were set randomly. Each docking experiment was derived from 10 different runs that were set to terminate after a maximum of 250000 energy evaluations. The population size was set to 150. During the search, a translational step of 0.2 Å, and quaternion and torsion steps of 5 were applied.

RESULT AND DISCUSSIONS

The natural inhibitor tryptanthrin docked and was found to bind with best efficacy. In brief, all the result validation is based on estimated free binding energy, estimated inhibition constant Ki, Van-dar-wall+H-bond+dissolve energy, electrostatic energy, total intermolecular energy, frequency and interaction surface (Table 1). All the comparison was made on basis of

Table 1 Docking studies between *M. tuberculosis* RecA protein and tryptanthrin

Rank	Estimated Free Energy of Binding	Estimated Inhibition Constant (Ki)	vdW + H-bond + dissolve Energy	Electro-static Energy	Total Intermolecular Energy	Frequency	Interact. Surface
1	-6.20 kcal/mol	28.76um	-6.18 kcal/mol	-0.02 kcal/mol	-6.20 kcal/mol	90%	599.808
2	4.02 kcal/mol	1.12 mM	-4.05 kcal/mol	+0.03 kcal/mol	-4.02 kcal/mol	10%	606.835

dock score given by Docking Server. The amino acid residues were found to play important role in binding of inhibitor within active site of the protein (Figure 2a-c).

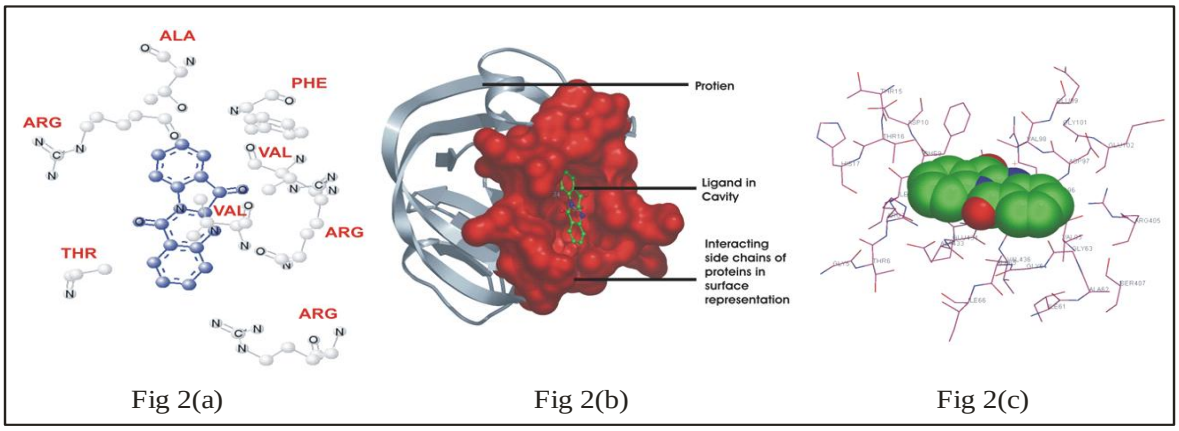


Fig 2(a) Fig 2(b) Fig 2(c)

Fig. 2(a) Amino acid residues interaction with ligand tryptanthrin; Fig. 2(b) Interacting side chains of RecA (in surface view) with ligand tryptanthrin; Fig. 2(c) Protein RecA in stick view interacting with ligands in sphere

The docking result showed there is a greater binding affinity of ligand tryptanthrin to *M. tuberculosis* RecA protein. Active site residues interacting with tryptanthrin are Arg96, Ala65, Phe9, Arg7, Thr16, Val98, Arg405 and Val432. In binding mode, polar bond found between O2 (11) with Arg96 of bond length (3.38 Å), hydrophobic bond between Ala 65 with C11 (15) and important pi-pi bond between C9 (13) with Phe9. These bonding showing stability of tryptanthrin in active residue and suggest it, as potential drug.

Meta-analysis of *Mycobacterium tuberculosis* pathogen and host-infected microarray data

Reema Singh[1,2], Meera Singh[1] and Andrew M. Lynn[2]

[1]Indian Council of Medical Research

[2]School of Computational and Integrative Sciences, Jawaharlal Nehru University, India.

Combining information from multiple studies generates more powerful statistics through meta-analysis, which voids artifacts present in individual studies. Meta-analysis falls into two major categories, one involves combining statistical significance and the other involves combining effects sourced from individual studies. However, due to difficulties in comparing gene-expression data across multiple platforms, several meta-analysis methods has been introduced. Meta-analytic methods can systematically combine data from different platforms to gain a clearer understanding of genes relationship to specify condition of interest. Here, we demonstrate a method developed for meta- analysis which has been validated by implementing it on both artificial data and earlier published results using meta-analysis. The method was then applied on publicly available *Mycobacterium tuberculosis* microarray data-sets.

Publicly available tuberculosis microarray data-sets derived from different platforms were downloaded from NCBI Gene Expression Omnibus (GEO). The aim of this study is to identify consistently up or down regulated genes between different conditions as determined by the sample information provided with the data-sets. For individual data-set annotation we used annotation files provided by vendors along side their microarray platforms. Analysis was performed using in house scripts written in R. GSE soft files and corresponding platform annotation files were downloaded locally. Conversion of GSE soft file into expression set obtained by using the GEOquery from bio-conductor software, each file was parsed to extract normalized gene expression matrix and metadata. All of the expression values were log base-two log transformed. Missing data were allowed. Significance of differential expression was determined using an empirical Bayes approach (Limma) for controlling the standard error of intensity of each probe set based on the standard errors of the intensities of all other probe sets in the comparison . After *p* values were obtained from each gene, up-regulated and down-regulated genes were separated on the basis of their log fold change.

The computational pipeline for individual study analysis has been shown in Figure 1 A and B. For each-possible combination of studies, a meta-analysis was performed to test the null hypothesis that positive results from individual studies do not correspond to the same group.

Meta-analysis pools multiple data sets; increases statistical power and provide an effective way of improving reproductively. For each gene that was present in all the studies within a given meta-analysis, q values (or gene-specific false discovery rates) were calculated using Fisher method for combining p-values, with n being the number of studies.

$$S = \Sigma_{i=1}^{n} -2 * \log (p_i)$$

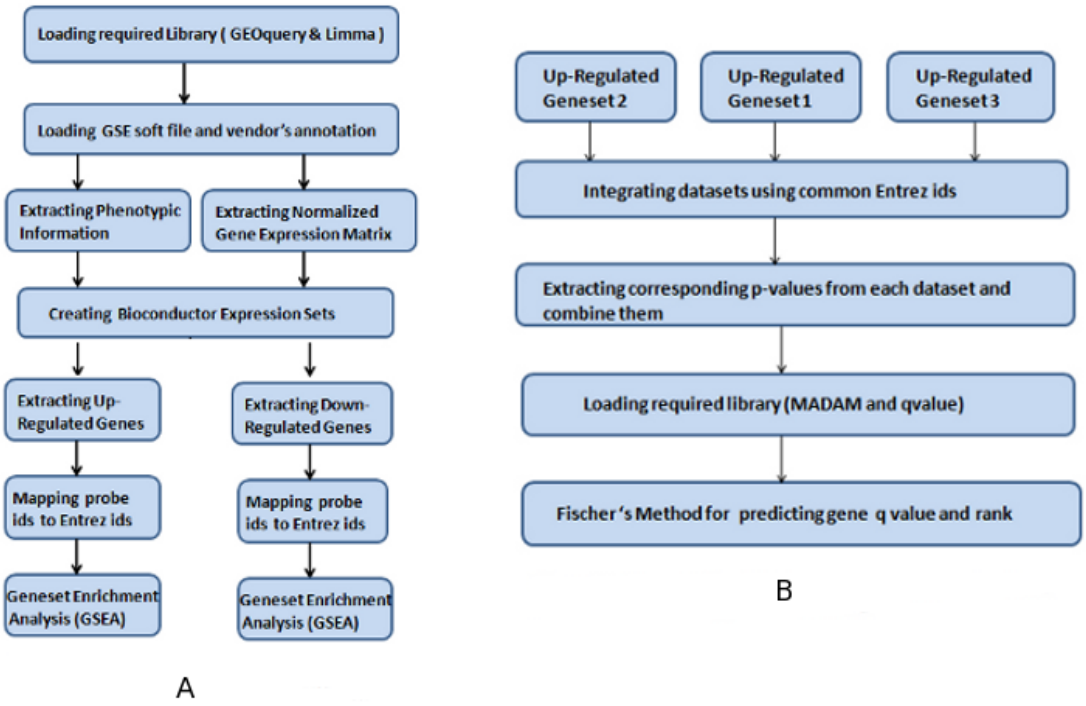


Figure 1:- A) Pipeline for parsing gene expression matrix,phenotypic information and finding differentially expressed genes.
B) Computational pipeline for combining p-values across datasets.

We report differentially expressed genes from publically available mycobacterium tuberculosis dataset, and examine their impact on gene networks to reveal functionally coherent groups of genes. The resulting coexpression and differential expressed gene data represent an important resource for guiding future functional characterization experiments of proteins of unknown functions and proteins of known functions.

Time-dependent antituberculosis activity of PA-824 against non-multiplying *Mycobacterium tuberculosis* persisters in mice.

Ahmad Z.[1,2], Tyagi S.[2], Minkowski A.[2], Peloquin C.[3], Ginsberg A.[4], Grosset J.[2] and Nuermberger E[2].

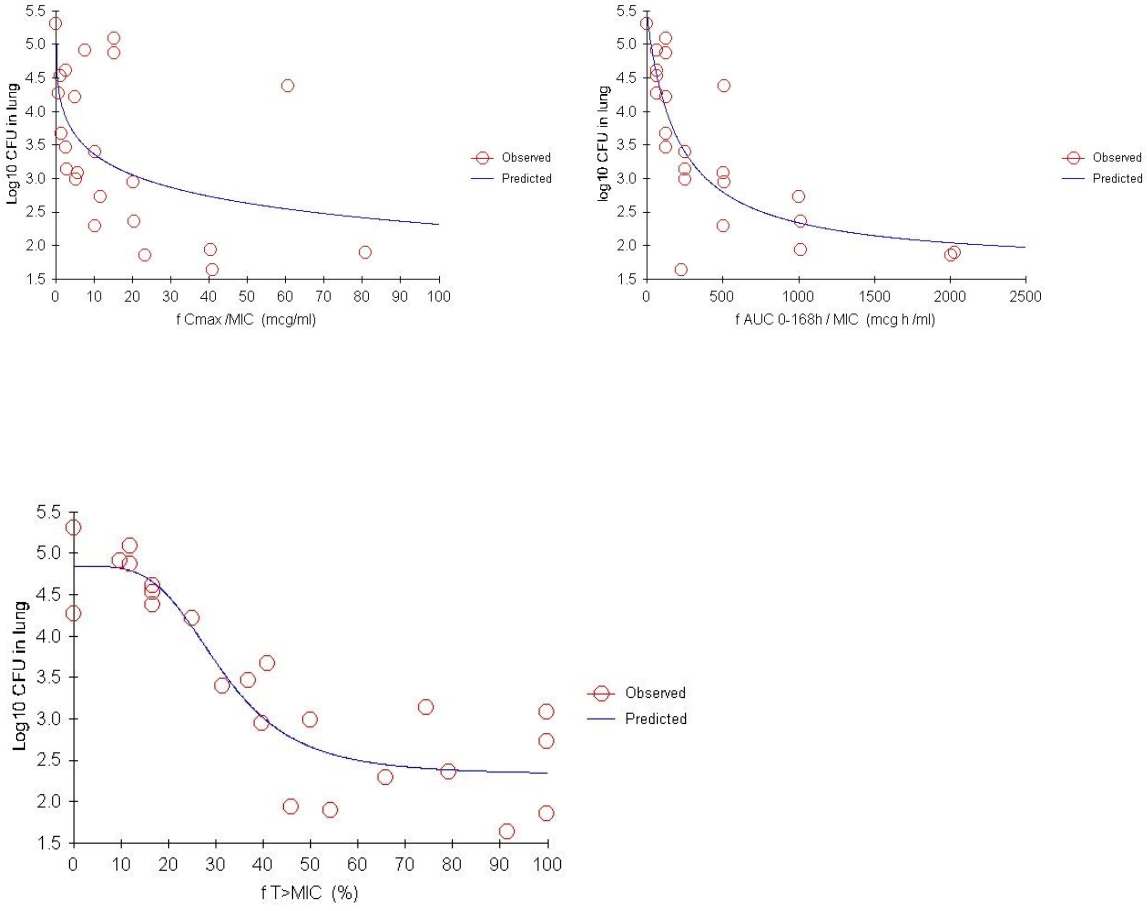
[1]Infectious Disease PK/PD Lab. Indian Institute of Integrative Medicine Branch Srinagar J&K-India, [2]Center for Tuberculosis Research, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, [3]College of Pharmacy, University of Florida, Gainesville, Florida and [4]Global Alliance for TB Drug Development, New York.

Background: PA-824 has shown potent anti-TB activity during Phase II clinical trials for TB treatment. It has potent *in vitro* activity against drug susceptible and resistant isolates of *Mycobacterium tuberculosis*. PA is bactericidal against actively replicating bacilli as well as non-replicating bacilli under hypoxic or prolonged culture conditions. The activity of PA against slowly or non-replicating bacilli was recently attributed to its capacity to donate nitric oxide during enzymatic nitro-reduction within the tubercle bacillus and thereby poison the respiratory apparatus. In murine models of TB, PA has shown substantial dose-dependent bactericidal activity during both phases of treatment with a minimal effective dose (MED) of 12.5 mg/kg and the minimal bactericidal dose (MBD) of 100 mg/kg. In combination with moxifloxacin and pyrazinamide, PA at MBD contributes a sterilizing effect equivalent to that of rifampin. These results in mice suggest PA may have the potential to shorten the duration of treatment for drug-susceptible, as well as MDR-TB in humans. Extended early bactericidal activity (EBA) in humans have shown that patients receiving PA once daily at doses of 200, 600, 1000 and 1200 mg for 14 consecutive days resulted in substantial reduction of sputum CFU counts at a rate of approximately 0.1 $\log_{10}$ CFU/mL/day; in contrast to mice no dose-response was observed in the EBA trial. To better understand the relationship between drug exposure and effect, we recently undertook a dose fractionation study and showed that the anti-TB activity of PA in mice strongly correlated with the free drug T(>MIC) (r^2 = 0.87) and correlated with the free drug AUC/MIC (r^2 = 0.60), but not with the free drug C(max)/MIC (r^2 = 0.17). Free drug T(>MIC) of 22, 48, and 77% were associated with bacteriostasis, a 1-log kill, and a 1.59-log kill (or 80% of the maximum observed effect), respectively. Based on our study and human pharmacodynamic simulations based on phase I data predict 200 mg/day produces free drug T(>MIC) values near

the target for maximal observed bactericidal effect and thus explains lack of a dose-response between 200 and 1,200 mg/day. Our results further support that dose dependent activity in humans is possible below 200 mg/day and indeed second EBA study in humans demonstrated dose dependent activity of PA in humans from 50-200 mg/day. To shorten the duration of TB therapy, a new drug also must kill slowly or non-multiplying bacilli with phenotypic tolerance to existing drugs. Therefore, we sought to identify the pharmacodynamic parameter most closely related to activity against slowly or non-multiplying bacilli with phenotypic tolerance to existing drugs. Dose-ranging single- and multi-dose serum PK in mice was established in first study. A 1-compartment model best fits the data and was used to simulate the serum concentration-time profile for the dosing regimens used in the study. We evaluated the activity of PA during the continuation phase as it measures a drug's activity against a population of "persister" bacilli surviving 1 months of treatment with rifampin (40 mg/kg), isoniazid (10mg/kg) and pyrazinamide (150 mg/kg) (RHZ), the continuation phase assessment may provide greater insight into a drug's treatment-shortening potential. PA-824 was kindly provided by Global Alliance for TB Drug Development. All procedures involving animals were approved by the Institutional Animal Care and Use Committee. Methods: Beginning 2 wks after aerosol infection with 3.5 $\log_{10}$ CFU of *M. tuberculosis* H37Rv, mice received high-dose rifampin (R) 40 mg/kg, isoniazid (H) 10 mg/kg, and pyrazinamide (Z) 150 mg/kg/d for 3 wks to select for persisters. Thereafter, treatment included PA monotherapy in total doses of 144, 288, 576, 1152, 2304 and 4608 mg/kg, divided into 4, 8 12, 24 or 48 doses over 24 d. Lung CFU were counted after treatment. WinNonlin was used for standard pharmacokinetic (PK) analyses and PD analysis using an inhibitory effect sigmoid E_{max} model, estimating 7.5% free drug. Human PK/PD simulations used published Phase I data. Results: RHZ pre-treatment reduced the lung CFU count from 6.62±0.11 to 4.46±0.16 $\log_{10}$. Thereafter, limited re-growth (to 5.30 $\log_{10}$ CFU) occurred without treatment. Dose dependent activity of PA in mice against these persisting bugs was observed from 12-192 mg/kg/day. Daily dose 192 mg/ kg was more effective than 96 mg/kg/day, this was in contrast to activity of PA against actively multiplying bugs where activity saturated at 100 mg/kg/day. Maximal killing (≥2.6 log reduction) was observed with PA at 96 mg/kg q12h and 192 mg/kg qd. Various estimated PD parameters that include E_{max}, E_0 and γ exhibited values 4.84 $\log_{10}$ CFU, 2.32 $\log_{10}$ CFU and 3.96 respectively (Figure 1). CFU counts correlated well with free drug time above MIC (T>MIC) (r^2= 0.84) and fAUC/MIC (r^2= 0.75)

but not fCmax/MIC (r^2=0.26) (Figure 1). EC_{50} was observed with T>MIC value of 31.20% free drug and EC_{90} with 53% free drug T>MIC. T>MIC values of 28% and 45% were associated with 1 and 2 log killing, respectively (Figure 1). At a given T>MIC, higher AUC/MIC was associated with a greater effect. PA has time-dependent bactericidal activity against slowly or non-multiplying persisters in this murine model. As the T>MIC target for a 2 log kill (0.083 log CFU/day) is attainable in humans with doses ≥200 mg, PA-824 may be useful to shorten TB therapy. Conclusions: PA has clear-cut bactericidal activity against persisting *M. tuberculosis* in mice. As previously shown for the initial phase of therapy, T>MIC had the strongest correlation with PA activity. PA may be useful for shortening the duration of treatment for drug-susceptible and M/XDR-TB. Additional studies with respect to PD of PA in other *in vitro* or animal models of persistence (especially hypoxia based) are warranted to confirm these findings.

Figure 1: Correlation of PA-824 activity with pharmacodynamic parameters against persisting *Mycobacterium tuberculosis*.



Targeting essential cell wall lipase rv3802c to combat tuberculosis: homology modeling, virtual screening and comparative docking studies

Parameswaran Saravanan, Avinash Hindupur, Vikash Kumar Dubey, Sanjukta Patra*

Department of Biotechnology, Indian Institute of Technology Guwahati, Guwahati, Assam, India.

Background:

Despite long and widespread usage of vaccine and drugs for tuberculosis, one-third of world's population is still infected. Tuberculosis caused by *Mycobacterium tuberculosis* continues to be the most deadly disease and India is the highest burdened country. The complex, unique cell wall and lipid metabolism are very critical for the survival and infection of mycobacteria. Mycolic lipids are not only essential for survival; they also contribute to the pathogenesis. Rv3802c is a phospholipase/thioesterase with unique PE/PPE domain and participates in cell wall remodeling, suggesting its distinct role in mycobacteriology. Rv3802c gained importance for its location in a mycolic acid gene cluster, experimental studies on essentiality, minimal genome concept and its proven role in cellular integrity from *M. smegmatis*. The recent discovery that tetrahydrolipstatin, binds irreversibly to Rv3802c drives us to search for effective reversible inhibitors. Considering the recent increase of MDR and XDR tuberculosis, finding a new drug target like Rv3802c to cripple pathogen's survival and infection is of high importance.

Results

Homology Modeling

The functional homolog of Rv3802c in *M. smegmatis*, MSMEG_6394 with crystal structure of 2.9 Å was identified as the template for homology modeling. Being cutinase-like, Rv3802c is significantly divergent from human proteome. Rv3802c might be serine hydrolase integrated with the PE/PPE domain which might be essential for the survival of the pathogen. The modelled structure confirmed that Rv3802c is a member of α/β hydrolases. Rv3802c has α/β hydrolase domain with parallel beta sheets covered by helices and "lid" sits atop of active site. The catalytic nucleophilic serine (Ser175) was found at the "nucleophilic elbow" (Gly173-Gly177) and other catalytic residues (Asp268 and His299) are positioned adjacent to each other within the active site. Strict conservation was observed at the active site namely Trp84, Glu85, Ser86,

Thr127, Ala128, Gln129, Met140, Phe174 and Ala300 which might be involved in substrate binding and recognition. The oxyanion hole which is involved in substrate binding was identified as Thr83 and Gln176 by structural comparison studies. The validated modelled structure of Rv3802c has acceptable statistics of backbone dihedral angle distribution of amino acids in Ramachandran plot. The fold quality of Rv3802c from ProsaII was comparable with experimentally determined MSMEG_6394 indicates the acceptability of the Rv3802c model. In order to check the stability of Rv3802c model, RMSD of backbone atoms from MD production run was plotted as time-dependent function and it indicates that the energy minimized Rv3802c model represents a stable conformation and suitable for virtual screening.

Virtual Screening

Structure-based virtual screening (SBVS) is one of the most preferred techniques to identify novel inhibitors against protein of interest. SBVS was used to identify potential inhibitors targeting Rv3802c active site to serve as starting point for translational research. Initial virtual screening was carried out using NCI diversity dataset against Rv3802c model to discover all possible structurally diverse potential hits. Screening was carried out again with closest human structural homolog, monoglyceride lipase (MGL). The molecules with higher binding affinity towards Rv3802c than human MGL was set as the criterion to identify specific inhibitors against Rv3802c. 60% similarity search of 10 top hits from initial screening was performed on ZINC database and the virtual screening with above protocol was performed again with 20 GA runs on the similar molecules to find potential inhibitors which are specific towards Rv3802c. Second approach identified the top hits from similarity screening considering the potent compounds reported by West *et al*. as initial dataset.

Top hits of computational studies were visually inspected and analyzed for all hydrogen bonds and hydrophobic interactions with Rv3802c using PyMOL. The top hits were bound at the active site and share most of the conserved residues at contact region of 4 Å, which are Thr83, Trp84, Glu85, Lys100, Ala101, Leu102, Lys105, Phe174, Ser175, Gly292, Gly293, His299, Ala300, Met301 and Tyr302. Key differences in contact residues were observed in Lys105 and Gly293 with exceptions in MSMEG_6394 which has asparagine and serine respectively. Significant differences in the contact residues were observed, notably Thr83, Trp84 and Lys105 whose counterparts in human MGL are alanine, glycine, and glutamate respectively in the structurally

aligned region. The interactions of the top hits with the key residues of active site were clustered around the substrate binding pocket. The mode of inhibition is expected to be competitive and reversible. The hydrogen bond interaction mode of the best hits from approach 1 and 2 are energetically favorable which also formed backbone hydrogen bonds at the substrate binding pocket of Rv3802c and may provide effective non-covalent inhibition. Best hit ZINC26726377 from first approach forms hydrogen bond interactions with backbone atoms of Thr84, Ser175, and Ala300, and also with side chain atoms of Thr83, Lys100 and His299. Best hit ZINC43866786 from second approach forms hydrogen bond interactions with backbone atoms of Lys100, and Ala300, and also with side chain atom of Glu85. The best hits were also stabilized by the hydrophobic residues of the cavity.

Conclusions

Targeting new drug targets which participates in both cell wall and lipid metabolism of *M. tuberculosis* like Rv3802c might be effective way to combat tuberculosis. Several evidences suggest that essential cell wall lipase, Rv3802c can be attractive drug target and worth pursuing to identify potential inhibitors. Molecular modeling and SBVS identified structurally diverse mycobacteria-specific reversible inhibitors against Rv3802c, ZINC26726377 and ZINC43866786 with difference in predicted binding free energy of -3.99 and -3.28 respectively. Further, the identified inhibitors may have inhibitory effect as anti-tuberculosis drugs against active, MDR and XDR tuberculosis and synergistic effects with DOTS therapy which can be validated by biological tests. Our current study on Rv3802c might be another step towards understanding and spur novel drug to eradicate tuberculosis.

GenomeABC: A platform for benchmarking of genome assemblers

Shailesh Kumar and Gajendra Pal Singh Raghava

Bioinformatics Centre (BIC), Institute of Microbial Technology, Chandigarh, India.

Whole genome sequencing by the second generation sequencing technologies or the Next Generation Sequencing technologies (NGS) provides a new way to study whole genome of an organism. These technologies include Illumina's Solexa technology, ABI's SOLiD and Roche's 454 GS FLX, have a significant impact on the genomics. Though cost of sequencing is coming down due to advancement in sequencing technology like Illumina but assembling of genomes from millions short reads pose a major challenge to bioinformatics community.

Recently, several algorithms have been developed for assembling of whole genome from short reads, freely available for public use in form of software packages such as Velvet, SOAPdenovo, AbySS, Euler-sr, Edena and SSAKE. Presently, it is difficult for a user to choose appropriate assembler for their genomes due to lack of benchmarking of these genome assemblers [3]. These assemblers have been trained and tested on different set of data so performance reported separately cannot be compared. Thus, it is important to benchmark existing assemblers on a standard dataset in order to understand their comparative performance. In past, limited attempts have been made to compare performance assemblers, for example Farrer *et, al.* (2009) evaluate performance of few assemblers on a genome of Pseudomonas syringae pv. syringae B728.

This study is an attempt to provide benchmarking of major assemblers, setting guidelines for benchmarking and for providing facility to evaluate performance of a genome assembler. In addition to assess performance of major assemblers, we have developed a platform "GenomeABC" for the benchmarking of genome assemblers.

We compared six popular genome assemblers (e.g. Velvet, SOAPdenovo, AbySS, Euler-sr, Edena and SSAKE) for the task of genome assembly on Pseudomonas syringae sequencing data (Illumina's Solexa data) i.e. GenomeR and on hypothetical genome data sets. Pseudomonas syringae short read sequencing data have been downloaded from European Nucleotide Archive at European Bioinformatics Institute of Project ID 32555.

At GenomeABC server, we created four random genomes named Genome A, Genome B, Genome C and Genome D of equal size (i.e. 6Mb each). Simulated reads with coverage values of 10X, 20X, 30X and 40X were generated by Genome A, Genome B, Genome C and Genome D

respectively. We wanted to find out the best coverage and best assembler for the purpose of whole genome assembly. Although, simulated reads generated by our method does not have sequencing errors as realistic Illumina's data of Pseudomonas syringae contain some sequencing errors . We also wanted to test the effect of paired end sequencing for the whole genome assembly of an organism. Genome assembly by each assembler produced a different set of contigs at different hash length (K) value. Evaluations of contigs have been done by the alignment of contigs to reference genome with the help of BLAT. For the evaluation of assembly, we have used a standard metrics (discussed in methods) and set criteria of maximum N 50 contig length and minimum error rate (% of total contig size sum) used in previous study.

In case of single end assembly, performance of Euler-sr 1.1.2 is good as compare to other assemblers with simulated reads at coverage of GenomeA (10X) , GenomeB (20X), and GenomeC (30X).But with GenomeD (40X) Velvet 1.0.12 and Euler-sr 1.1.2 perform equally well with N 50 contig length of 5999737. With realistic data of Pseudomonas syringae (GenomeR), Velvet 1.0.12 performs best with N50 contig length of 12421. In this case N 50 contig length for Euler-sr 1.1.2 is 11555 nucleotide. Hence, we can say that low coverage single end sequencing of genome can be assembled efficiently by the software like Euler-sr 1.1.2 or Velvet 1.0.12. SOAPdenovo1.04 does not perform scaffolding in case of single end assembly. So, the performance of SOAPdenovo, in case of single end assembly is poor as compare to assemblers like Velvet 1.0.12 and Euler-sr 1.1.2.

With all synthetic genomes i.e. GenomeA (10X), GenomeB (20X), GenomeC (30X) and GenomeD (40X), error rate (%) is 0 for Velvet 1.0.12. But with GenomeR Velvet 1.0.12 have error rate (%) 1.299, which is more than the Edena 2.1.1 (error rate =0.26%).

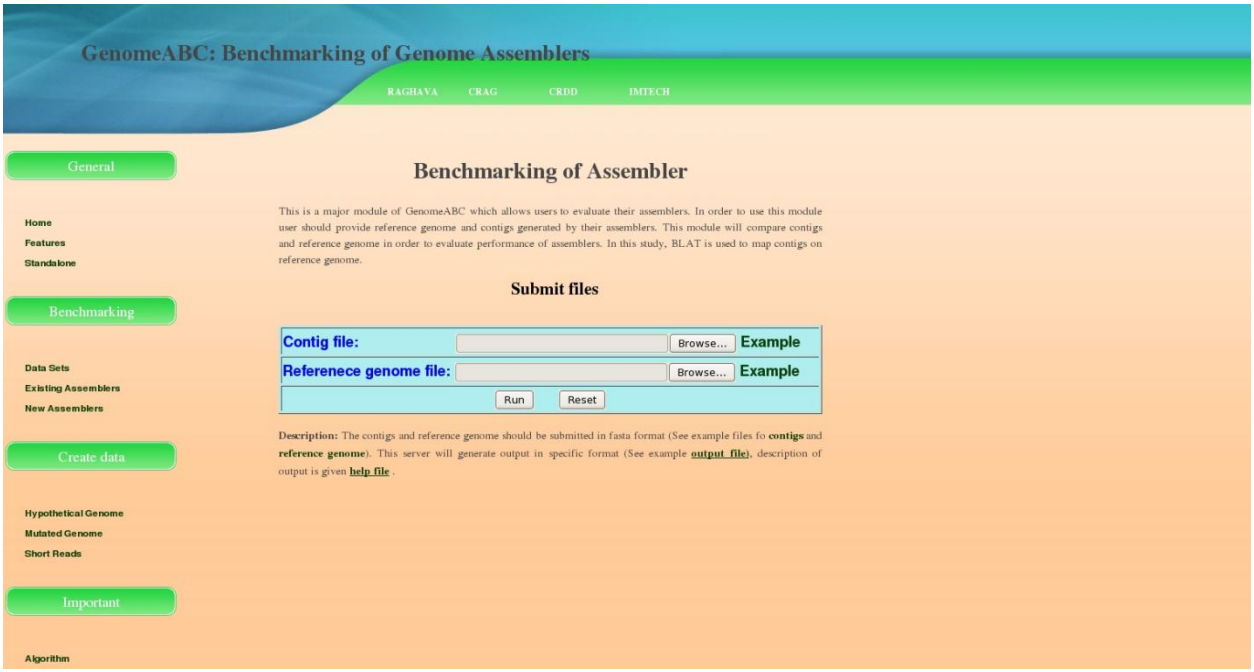
In case of paired end sequencing, SOAPdenovo 1.04 performs good scaffolding for all type of data, i.e. synthetic and realistic data. With the data set of GenomeA (10X) SOAPdenovo 1.04 have the largest N 50 value of 5714124. But with GenomeB (20X), GenomeC (30X) and GenomeD (40X), Velvet 1.0.12 have highest N 50 values. In case of GenomeR (realistic dataset) N 50 is best with SOAPdenovo 1.04 i.e. 164240. On the basis of these results, we can conclude that the assemblers like SOAPdenovo 1.04 can be used for the paired end assembly data. Velvet is good for the high coverage data assembly.

Error rate (%) is another criterion for the assembly. With all synthetic genomes i.e. GenomeA (10X), GenomeB (20X), GenomeC (30X) and GenomeD (40X), error rate (%) is 0 for SSAKE

3.5. At high coverage data of GenomeC and GenomeD other software i.e. AbySS 1.2.4, SOAPdenovo 1.04 and Euler-sr 1.1.2 also produce 0 error rate (%). Velvet 1.0.12 produces some error rate in case of all synthetic data. In case of GenomeR, minimum error rate (%) of 3.93 is produced by SOAPdenovo 1.04.

On the basis of all observations, we can conclude that the Paired end assembly increases the assembly size but the error rate also increases with this. Software like Euler-sr 1.1.2 and Velvet 1.0.12 can be used for the single end assembly. In case of paired end assembly SOAPdenovo 1.04 is best.

Figure: GenomeABC server for benchmarking of genome assembler.



Glycation: the factor for association of increased occurrence of tuberculosis in diabetic state

Dr Dibyajyoti Banerjee, Department of Experimental Medicine and Biotechnology, PGIMER, Chandigarh, India.

Long standing diabetes mellitus is a known risk factor for occurrence of tuberculosis and the cause of such association is largely unknown. Active research in this field is important since both diabetes and tuberculosis are endemic diseases with particular reference to Indian subcontinent. Mycobacterium tuberculosis is an intracellular pathogen which resides inside the stable phagosome in cases of successful tuberculosis infection. It is known that Mycobacterial factors inhibit the effective phagocytic killing of the bacterium which is attributed as a cause of successful chronic tubercular infection. With this background we feel that in chronic diabetes there may be glycation induced inhibition of important proteins responsible for phagocytic killing of microorganisms and that may cause the increased occurrence of tuberculosis at chronic diabetic state. In this connection it is pertinent to mention that the majority of complication of diabetes mellitus is associated with increased extent of protein glycation (commonly estimated by HbA1c estimation) and therefore we feel understanding of glycation of proteins that are important for phagocytosis process is important to understand the patho-biology of increased infections in diabetic state. With this background we have studied glycation pattern of some relevant proteins expected to be glycated more in uncontrolled hyperglycemia using tools of computational biology and observed that there is high probability of glycation of some Lysine residues of NADPH oxidase and iNOS resulting in inhibition of the enzyme activity of these enzymes at chronic hyperglycemic state. At the present moment we have also accumulated experimental evidences to support the above in-silico observation. We believe scrupulous control of blood sugar and development of anti-glycation molecules will help to reduce the incidence of tuberculosis in the diabetic host.

Mannosylated solid lipid nanoparticles encapsulating rifabutin for effective management of tuberculosis

Abhilasha Thakur, Navneet Nimje, Prof. G.P. Agrawal

Department of Pharmaceutical Sciences, Dr. H.S. Gour University Sagar M.P. India.

Background:

Prevention of diseases has been a long-standing challenge for the mankind specially the communicable diseases including tuberculosis (TB), leprosy etc. TB once referred to as the white death is an ubiquitous, highly contagious, chronic, granulomatous, debilitating human disease caused by *Mycobacterium tuberculosis* (*M.tb*) and some other species of same genera. Today, TB tends to be concentrated among inner city dwellers, ethnic minorities and recent immigrants from areas of the world where the disease is still common. However it can occur anywhere, and no one is exempted from the threat of infection.

Although five decades of tuberculosis control programmes using potentially efficacious drugs, together with improvements in economics, medicine and public health, have led to a gradual decrease in the incidence of TB in industrialized nations. However, worldwide TB continues to be a leading cause of death. Even in India nearly 2 million people succumb to the infection loosing their lives. Furthermore, TB has again become a concern as the situation has exacerbated because of the presence of numerous complicating factors like HIV co-infection, lack of patient compliance with chemotherapy, variable efficacy of BCG vaccine and emergence of drug resistant (DR), multidrug resistant (MDR) and more recently extremely resistant (XRD) strains of *M.tb*.

A variety of therapeutic agents are available for the treatment of TB but the access of antimycobacterial agents to *M.tb* inside host macrophages is limited due to the low levels of drug permeation, making it difficult to achieve effective drug concentrations. All these factors warrant the need to develop a drug carrier which would provide an effective, stable and controlled delivery of drugs. Hence it was proposed to prepare targeted solid lipid nanoparticles encapsulating antitubercular drug and thenafter investigate the potential of this novel polysaccharide anchored SLNs as an effective drug delivery system.

Experimental work:

Rifabutin encapsulated SLNs were prepared by employing solvent injection method and then mannosylated by coupling the amine groups present on the surface of the SLNs with mannose. To characterize the synthesis of mannosylated SLNs, FTIR spectroscopic studies were performed. Further the prepared systems (both conjugated and unconjugated) were characterized for shape and surface morphology using scanning electron microscopy (SEM), size and size distribution by photon correlation spectroscopy using a Zetasizer, zeta potential by laser Doppler anemometry using a Malvern Zetasizer and drug entrapment efficiency by using G-50 Sephadex mini column. *In-vitro* drug release profile of both the formulations was carried out using dialysis tube method.

To evaluate the performance of the developed system *ex-vivo* studies and *in-vivo* studies were also performed. Cellular uptake of fluorescein isothiocyanate (FITC)-loaded formulation by macrophage cell lines "J774" was determined using fluorescence activated cell sorter (FACS) instrument. Hematological studies were performed on albino rats in which collected blood samples were analyzed for hemoglobin, RBC count, and WBC count before and after treatment with the formulations. After administration of different formulations, biodistribution of the drug in organs (lungs, liver spleen) and in plasma was also determined in albino rats.

Results:

The objective of the present study was to evaluate the prospective of engineered nanoparticles for selective delivery of an antituberculosis drug, rifabutin, to alveolar tissues. The results show that the mannosylated SLNs entrapping rifabutin could be successfully prepared using the selected method within desired size range (389±2.3 nm), low PDI (0.357) and high drug entrapment (82.6±1.2 %). SEM photomicrographs showed spherical nature of both the formulations. *In vitro* release studies of unconjugated SLNs showed a drug release profile of 89.21±3.8% after 120 hrs in PBS (pH 7.4) whereas rifabutin-loaded mannosylated SLNs showed a drug release profile of 65.23±1.9%. This significant decline in the release rate of rifabutin from mannosylated SLNs showed a more sustained release nature of the formulation. Further the cellular uptake study clearly indicated an enhanced cellular uptake of FITC-loaded mannosylated SLNs by macrophage cell line. Hematological studies results indicated that hemoglobin (Hb) content, RBC count and WBC count were undergoing a deviation from normal value. The Hb content of blood collected from rats treated with free drug, unconjugated SLNs and

mannosylated SLNs was found to decrease from 14.89±0.3 g/100 mL to 12.06±0.2 g/100 mL, 12.78±0.2 g/100 mL, and 13.57±0.4 g/100 mL, respectively. On studing the plasma drug concentration profile the blood level of the drug was sustained to a greater extent in the case of conjugated formulation than that of an unconjugated system. In biodistribution study, initially the free rifabutin content was found to be highest in the kidney and secondarily in liver and lungs. However, the free drug concentration in various organs was observed to decrease rapidly when compared to both the formulations. In the case of unconjugated formulation, the concentration of drug was found to be highest in the liver as compared to other organs. On the contrary, the concentration of drug in the mannosylated formulation was found to be greatest in lungs and lower in liver and kidney.

Conclusion:

Nanoparticles-based vehicles for controlled, sustained, and targeted delivery hold a considerable potential for the delivery of antitubercular drug. SLNs are the class of carrier systems which have been widely explored for the delivery of various pharmacological moieties. The present study reveals the efficacy and suitability of mannosylated solid lipid nanoparticles for the delivery of rifabutin. *In vitro* studies depict the sustained release nature of the formulation. At the same time, mannosylation promotes their selective uptake by lung tissues. This promotes site-specific delivery, thereby enhancing the therapeutic efficacy and reducing the untoward side effects. Thus, the formulation holds a promising prospective for ferrying large doses of drug, providing targeted delivery with the minimal side effects. The prospective of this carrier system in a variety of categories of antitubercular drugs for site-specific delivery is under study.

Development of a System for Monitoring Disease Occurrence by Networking of different ICT Technologies

Chalga M.S. [a], Dixit A. K.[a]

[a] Desert Medicine Research Centre (ICMR), New Pali Road , Jodhpur, Rajasthan, India.

Implementation of Information and Communication Technologies (ICT) in the health sector is facing many challenges due to inadequate infrastructure and limited resources. Health Sector is having Information Technology (IT) facilities up to Block level. However, telecommunication facilities are available up to village level. This Paper describes a system which can be developed by networking various ICT techniques for quick, efficient and feasible health information and disease incidence monitoring system.

Introduction

India is facing various complex health problems such as Malaria, Dengue, H1N1, HIV, Tuberculosis, Nutritional deficiencies etc and faced many severe epidemics in the past[1]. IT based Efficient and fast reporting surveillance system is required to monitor and check the epidemic situation at the right time. IT is widely being used for health care services in the world such as development of epidemiological surveillance databases, hospital management system, use of Geographic Information System, E-health, Tele-Medicine, network of health care institutions, Grid Technology, etc. Although the health sector in India is applying various IT techniques in the fields of health research, health sector is still far behind in applying new information technologies as compared to commercial sectors. Implementation of ICT in the health sector is facing many challenges due to insufficient infrastructure and limited resources[2]. This Paper describes a system which can be developed by networking of various Information and Communication Technologies using optimum resources.

Methodology

Various Health Centres were visited to study their functioning and literature was reviewed through internet. Certain ICT technologies were selected based on their salient features. The networking of the selected ICT technologies was explored. Finally a design was envisaged in which various ICT technologies may be configured to serve as efficient disease surveillance

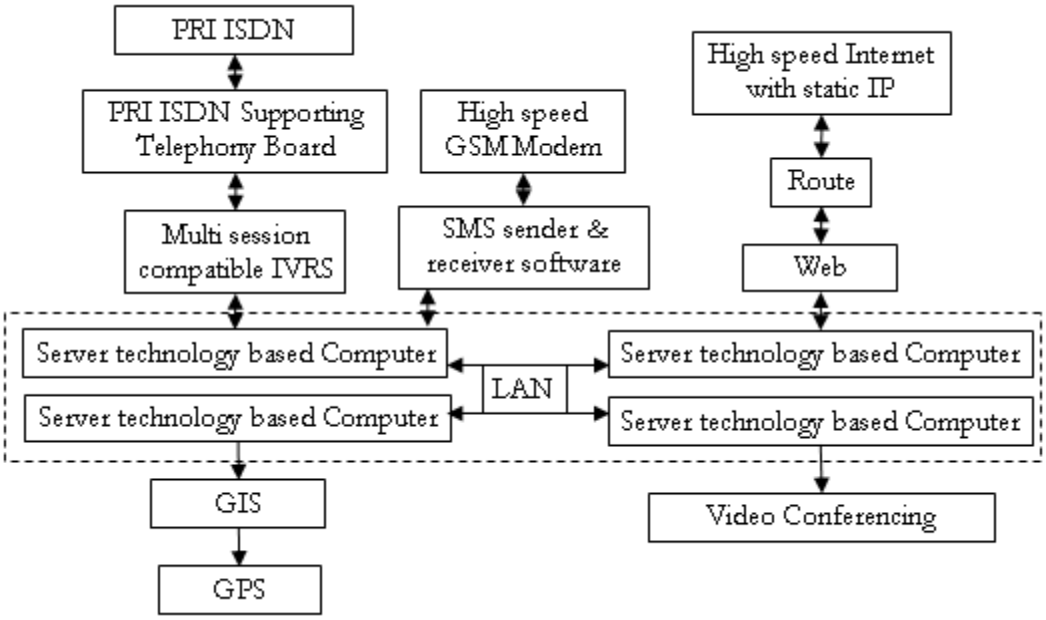
system.

Results and discussion

After visiting various health functionaries and review of literature it was found that the government of India has developed a hierarchy of health functionaries which includes health zone constituting of a cluster of districts, district level hospitals, Block level health monitoring, Community Health Centres, Primary Health Centres and Sub Centres[3]. The Government is attempting to computerize all the health units, however mostly computers are installed up to Block health functionaries only. However, Telephones and Mobile network were found having better coverage as compared to other ICT technologies. It was also explored that the common ICT techniques which are used for data collection and data reporting in health sector abroad were various software, E-mail services, Fax, handheld computers, web based techniques, Mobile services, SMS techniques and Interactive Voice Response System etc. However, techniques like handheld computers, Mobile services, SMS techniques and Interactive Voice Response System are rarely used in India.

An ICT system is envisaged by networking various ICT techniques into a single system such as PRI ISDN, PRI modems, IVRS, telephony board, SMS system, GSM modem, Database server, Web server, GIS server, GPS device etc. (Fig. 2). In this design, PRI ISDN telecommunication line is connected to the telephony board through PRI modem which is further connected to a computer having IVRS software. Other Server computers are configured to serve as Web Server, GIS Server and Video conferencing Server. This system can collect data through IVRS, SMS and Internet and can report information through Fax, SMS, Internet online reports and can perform video conferencing and generate informative GIS maps. The PRI ISDN empowers to collect data through 30 parallel telephonic callers. In one of our study, we have developed a similar system at Desert Medicine Research Centre (ICMR), Jodhpur, India[4]. The system had successfully collected data from root level health functionaries through their mobile phones and reported online disease occurrences reports through system Web Server.

Fig 2. Networking of ICT tools for Health Research



Conclusion

The system can collect information through Internet, SMS and IVRS and can report through Fax, SMS, IVRS and Internet. PRI ISDN can be implemented over a wider area. The system is cost effective, fast, easy to implement and efficient. This system can suffice most of the requirements of the health functionaries of developing countries.

Cloud computing for chemistry and biology and its application in Drug Discovery

M Karthikeyan, Deepak Pandit, Sunil Nalwade, Arvind Bhavsar

National Chemical laboratory, Pune, India.

Today due to advancement in informatics and robotics the amount of experimental data available to public domain increased enormously. The computing paradigm changed over a period of time and it is time for us to cope up with the modernization and use efficient tools and communication systems to stay active in our profession. Storage of data moved from floppy to cloud. Preferred email servers are in open domain. Scientific data are moving from personal computers from corporates to public domain for the benefit of the society due to change in mindset and policies. Chemical or Molecular informatics is no exception in this change in trends and today they play an important role in drug discovery and materials design. The information era brings with it the challenge to handle massive amounts of data being generated in every scientific discipline.

Public databases like Pubmed, PubChem, PDB are flooded with biological and chemical data including scientific text, chemical structures, reaction and activity profile, images, microarray expression profiles to quote a few. Community demands more rapid annotation and analysis tools and approaches to make sense of the ever-growing data without it the data would be a big liability. Cloud Computing comes to the rescue by offering variety of software services, platform and infrastructures with pay as you use or scale as you need model. While being cost effective the scalability it provides has no match. This fits perfectly to the dynamic experimental needs of academic research which totally removes the need for in-house HPC requirements for individual laboratories. Distributed Computing with the performance equivalent to High Performance Clusters can be implemented easily with the use of cluster made out of commodity hardware locally available or on public clouds with infrastructure. In this work, we have highlighted how research in chemoinformatics and bioinformatics got an acceleration to solve the “BIG DATA” problem with opensource based cloud computing and other Distributed Computing Environments as a framework of choice especially to handle data intensive tasks which lies at the core of our discussion.

In-Silico Studies to Understand the Role of ^{99m}Tc-Hydroxypropyl β-Cyclodextrin as an Infection Imaging Agent

Jaya Shukla[1], BR Mittal[2], GP Bandopadhyaya[1]

1. Department of Nuclear Medicine, PGIMER, Chandigarh, India.
2. Department of Nuclear Medicine, AIIMS, New Delhi, India.

Background

Imaging infection is a big a challenge. Presently ^{111}In / ^{99m}Tc labeled WBC, ^{18}F-FDG and ^{99m}Tc-MDP are used with varying success. Labelled WBC is gold standard. But, ^{111}In/ ^{99m}Tc WBC have drawbacks, such as the risk of infection from exogenous microorganisms during in vitro labeling, cross-contamination with other patient/s blood, and the considerable cost of the technique. This necessitates the need of a infection imaging agent. The adherence of bacteria to human cells is mediated by bacterial membrane, produces inflammatory responses to bacterial infections and can be utilized in detection or elimination of invading microorganism. Breast milk protect newborn child from infection is well known. But the reason behind this is the presence of more than 130 oligosaccharides, which act as soluble receptors for different pathogens. The presence of oligosaccharides binding proteins (e.g. Maltose binding protein; MBP) on bacterial membranes is also documented. Maltose binding protein (MBP) is a monomeric periplasmic binding protein on bacterial cell membrane and plays an important role in active transport and bacterial chemotaxis. Low molecular weight oligosaccharides can block of binding sites on bacterial surfaces and prevent bacterial adherence. Cyclodextrins are water-soluble cyclic carbohydrate compounds with a hydrophobic cavity due to the specific orientation of the glucosidic substituents. The modified hydroxypropyl β-cyclodextrin is regarded as safe for parenteral use and approved by FDA for oral, intra-dermal, subcutaneous and parental use. 80% to 90% of injected hydroxypropyl β-cyclodextrin (HPβCD) is excreted unchanged through renal route with the T½ ≈ 1.5 hrs. With this background we had labeled HPβCD and explored its role in infection imaging

Experimental

In-siloco Docking studies: Based on the above hypothesis, Ligand-protein interaction of HPβCD and MBP was studied by using autodock 4 program on auto dock tool (ADT) in Linux operating system. Crystallographic coordinates of MBP have been taken from the Brookhaven Protein Data Bank (pdb) (Reference -1DMB). The HPβCD was drawn by Chemdraw software and three

dimensional structure was taken as ligand. Hydrogen atoms were added and Gasteigen charges were assigned. Since MBP is a monomeric protein, the grid box was selected covering the whole protein. The docking was done for 10 genetic algorithms. Minimum energy state (most stable) ligand-protein complex was chosen. Ligand–protein contacts were derived with the help of LPC software.

Radiolabeling and Quality control: Hydroxypropyl-β-cyclodextrin was labeled with technetium-99m (^{99m}Tc-HPβCD). The quality control of ^{99m}Tc-HPβCD was done by Instant thin layer chromatography (ITLC) and chemical characterization by ^{1}H-NMR. The route of excretion of ^{99m}Tc-HPβCD was assessed in albino wistar rats. Initially, ^{99m}Tc HPβCD was injected in human subjects with clinically confirmed infected knee joints.

Results

Docking analysis: The structures after docking were analyzed on the basis of energy levels. The lowest energy co-ordinates were taken for ligand protein interaction on LPC server. Obtained structure demonstrated hydrophobic, hydrophilic and van der Waals interactions between ligand and amino acid residues of MBP (Figure). HPβCD was located at the base of the cleft on MBP.

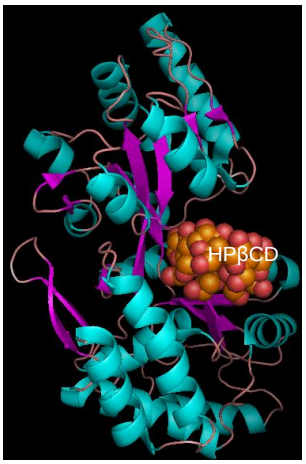
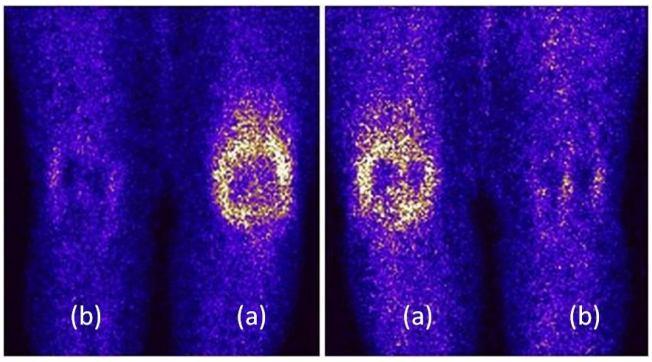


Figure . The structure of HPβCD (spheres) with MBP (cartoon) derived from docking experiment.

Other experimental data: The ^{1}H NMR studies revealed the binding of ^{99m}Tc at C-8/H-8 position of HPβCD. The excretion of ^{99m}Tc HPβCD showed renal route of excretion. There was marked increased uptake of ^{99m}Tc-HPβCD in the patient's left knee with proven infection (Figure). However, there is no uptake in the right knee due to the lack of infection. The differentiation between septic and aseptic loosening was also evident on single photon emission tomography (SPECT).



^{99m}Tc-HPβCD images in human subjects with B/L Knee joint replacement (a) with prosthesis infection, (b) with no prosthesis infection

Discussion

Maltose binding protein has two globular domains. Shraff et al (1995) demonstrated that βCD bind to MBP. The βCD was at the base of the cleft either with direct βCD-protein hydrogen bond or water-mediated hydrogen bonds. In our studies minimum energy coordinates of HPβCD-MBP complex were taken to study interactions of MBP and HPβCD by using LPC server. The established interactions were hydrogen bonds, van der Waals interactions and hydrophilic interactions. As HPβCD-MBP interactions are similar to βCD-MBP as shown by our docking experiment, we can also assume that some amino acid side chains (MBP) and HPβCD interactions are mediated with water molecules. The interactions between HPβCD and MBP stabilize HPβCD in the cleft of MBP and are responsible for imaging bacterial infection when labeled with radionuclide (^{99m}Tc).

Cyclodextrin (^{99m}Tc-HPβCD) can easily be labeled with diagnostic isotope (^{99m}Tc) and the retention of ^{99m}Tc-HPβCD in the knee joint with prosthesis infection was due to the ligand protein intetraction between ^{99m}Tc-HPβCD and MBP or similar receptor proteins having maltodextrin binding sites on membranes of infection causing microbes. Docking results are complementing our clinical findings and are helpful for defining the mechanism of localization of newly developed radiopharmaceutical, ^{99m}Tc hydroxypropyl-β-cyclodextrin, for infection imaging.

Conclusion

Docking data confirmed the interaction between HPβCD and bacterial maltose binding protein (MBP) and supported our experimental and clinical data for ^{99m}Tc HPβCD as a molecular probe for infection imaging.

Application of ccpdb in designing of protein-ligand mutation sites

Harinder Singh & Gajendra P. S. Raghava

Bioinformatics Centre, CSIR- Institute of Microbial Technology, Chandigarh, India.

ccPDB (Compilation and Creation of datasets from PDB) is a comprehensive platform which provides gamut of facilities from creating different datasets of proteins available in PDB, to study their interactions with different ligands.

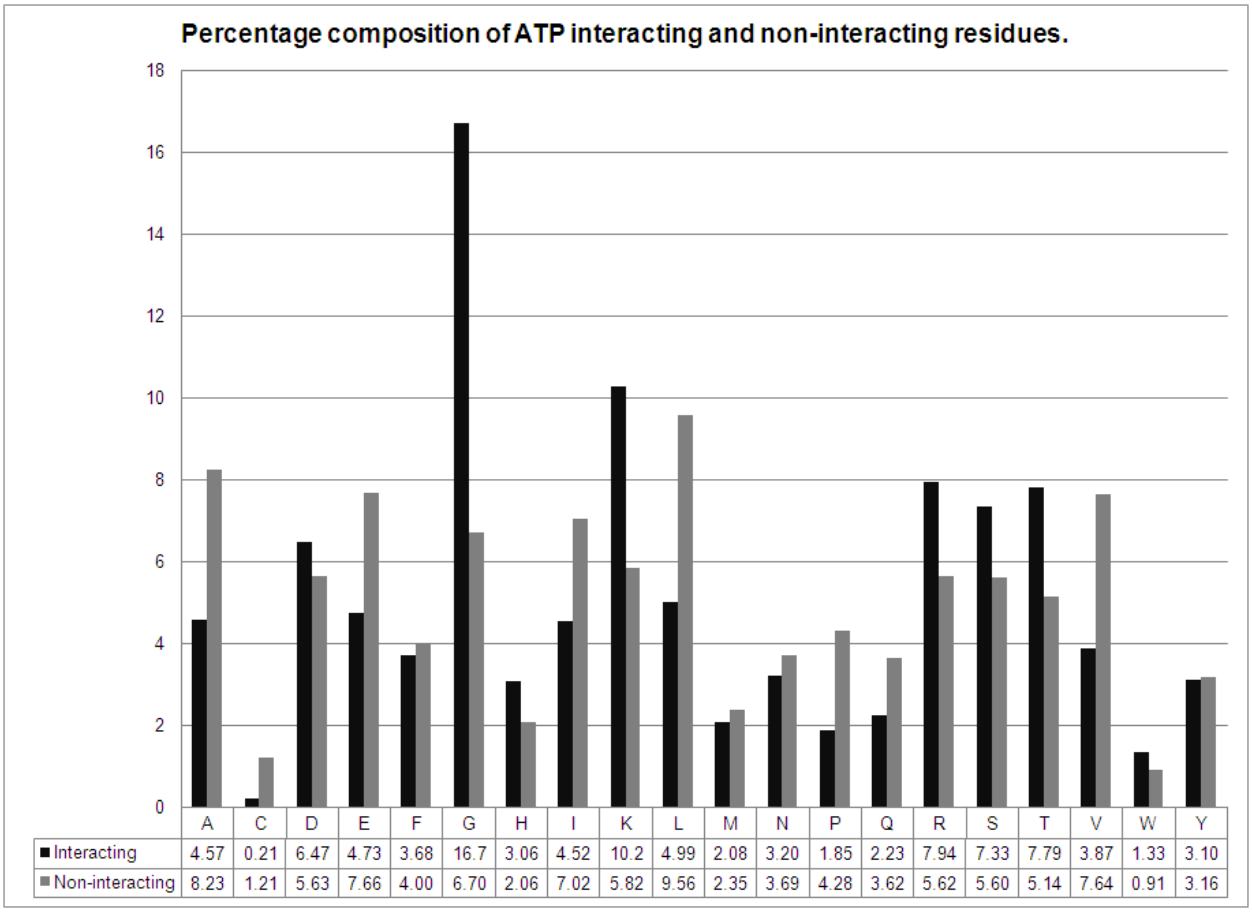
Currently, ccPDB contains 11002 ligands. Each ligand has preference of interaction with some amino acid (strong binders) more than other amino acids (weak binders). Based on ligand-amino acid interaction data, it's possible to mutate interaction sites of a protein to increase or decrease its ligand interaction affinity. For a given protein-ligand complex, the binding efficiency of ligand can be improved by mutating weak binders with strong binders.

ccPDB has a module ligand interacting residues which provides ligand interaction details for a given set of PDB. The module allows users to assign interacting residue for each ligand based on interaction distance, contact surface area, bond type between ligand and residues. The data obtained can be use to interfere strong and weak binders. The beauty of this module is that is allows users to customize options for finding strong and weak binders from a given set of proteins.

A Case study: In the present study, we have analyzed latest set of PDB chain interacting with an ATP ligand. A total of 109 ATP-interacting PDB chains were found in PDB. Only those residues are counted in which interaction distance is less than 4.0 A°. We found a total 16669 ATP-interaction and 377047 non ATP-interaction. Percentage of interacting and non-interacting residues were calculated, which are shown in Figure 1. Analysis of results showed that, glycine had maximum number of interactions with ATP followed by lysine. Cysteine showed the least number of interactions, whereas arginine, serine and threonine residues show considerable number of interactions. Class wise trend shows that Basic amino acids show more number of interactions than basic acidic acids. Alanine, glutamate, isoleucine, leucine, proline, glutamine and valine are more dominating in non-interacting residues. Moreover, in terms of bond type,

hydrogen bonding (13157) was most prominent followed by hydrophobic-hydrophilic contacts (8715), hydrophobic-hydrophobic contact (3671) and aromatic-aromatic contact (1152). No aromatic-aromatic, hydrophobic-hydrophobic contacts are found. A total 6339 interactions are due to only hydrogen bonds and 1138 interactions are due to hydrophobic-hydrophilic contacts. In short, we found glycine, lysine, arginine, serine, threonine as strong binders and cysteine, glutamine, methionine, proline as weak binders. We are also, analyzing FMN, FAD, NAP ligands, of which data will be shown in poster.

In conclusion, this type of studies will be helpful in find strong and weak binders for a particular ligand. It will be helpful in designing novel protein or to mutate pre-existing ones either to enhance ligand-residue interaction or decline ligand-residue interaction.



	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
Interacting	4.57	0.21	6.47	4.73	3.68	16.7	3.06	4.52	10.2	4.99	2.08	3.20	1.85	2.23	7.94	7.33	7.79	3.87	1.33	3.10
Non-interacting	8.23	1.21	5.63	7.66	4.00	6.70	2.06	7.02	5.82	9.56	2.35	3.69	4.28	3.62	5.62	5.60	5.14	7.64	0.91	3.16

Figure 1: Percentage composition of ATP interacting and non-interacting residues in PDB.

Development of predictive models for lipid peroxidation inhibitory activity of cinnamic acid and caffeic acid derivatives

Indrani Mitra[1], Achintya Saha[2] and Kunal Roy[1*]

[1]Drug Theoretics and Cheminformatics Laboratory, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India

[2]Department of Chemical Technology, University of Calcutta, Kolkata, India.

Introduction

The design and development of antioxidant molecules have lately gained a great deal of focus which is attributed to their immense biomedicinal importance in combating the free radical associated health hazards. However, when produced in large excess they participate in a series of unwanted side reactions resulting in cell damage. The free radicals are regulated by a balance between tissue oxidant and antioxidant activity. An imbalance in the oxidant-antioxidant reaction either due to excess free radical formation or insufficient removal by antioxidants leads to oxidative stress. Thus to balance the overall antioxidant status of the human body, dietary antioxidant supplementation becomes essential. Viewing the immense biomedicinal importance of the antioxidant molecules, design and synthesis of such chemical entities with improved potency have gained remarkable attention over the last few decades. Such designing of active chemical compounds employing a time effective approach involves the incorporation of QSAR methodology. Development of QSAR models enables preliminary screening of databases for selection of optimally active molecules in a cost effective manner. The present work deals with derivatives of cinnamic acid and caffeic acids, which were modeled for their ability to inhibit lipid peroxidation within the biological system. Three different types of models were developed: (a) descriptor based QSAR models, (b) 3D pharmacophore models and (c) HQSAR (hologram QSAR) models.

Present Work and Methods

The dataset used for the present work comprised of derivatives cinnamic and caffeic acids (1-3). The essential molecular attributes contributing to the activity profile of the molecules were determined based on different QSAR models developed using the genetic function approximation (GFA) technique. The pharmacophore and crucial molecular fragments of the compounds were also determined based on 3D pharmacophore mapping and HQSAR techniques.

To assess the quality of the generated pharmacophore hypotheses, cost functions were calculated and Fischer validation test was performed at 95% confidence level followed by prediction of test set compounds. HQSAR model was derived based on various combinations of fragment distinction and fragment generation parameters for each hologram length and the best PLS model was selected based on the maximum value of Q^2. All the models developed were validated using the internal validation external validation and randomization techniques.

Results

Among all models, the descriptor-based model (Q^2=0.841, $\overline{r_{m}^{2}{}_{(LOO)}}$=0.745 and $\Delta r_m^2{}_{(LOO)}$=0.115) developed using the GFA-spline technique yielded the most satisfactory results [R^2_{pred} =0.710, $\overline{r_{m}^{2}{}_{(test)}}$ = 0.585 and $\Delta r_m^2{}_{(test)}$=0.139]. The model describes that an increase in activity is achieved with a decrease in the nucleophilic character of the molecules. Additionally, reduced positive charge or negative charge on the C_4 atom adversely affects the activity profile of the molecules. Additionally, an increase in the number of atom centered fragments bearing hydrogen atom attached to a heteroatom (-OH) further enhances to the activity profile of the molecules.

The three feature (HBA, HBA, HYD) pharmacophore developed in the present work was selected based on the values of the correlation coefficient (0.756) and the cost functions (null cost = 325.656, total cost = 212.261).The hydrophobic (HYD) feature maintains a distance of 6.233Å from one of the hydrogen bond acceptor (HBA) groups while the HBA features are separated from each other by 7.702Å (Fig. 4a). The HBA features (green) facilitate formation of the hydrogen bond with the free radicals while the HYD feature (cyan) interact with the free radicals through the development an area of transient electron deficiency. Mapping of the most active compound with the developed pharmacophore revealed that the ketonic (=O fragment for the amide/carboxylic acid group) and the ethereal (for the substituent at C_6) position oxygen atoms are well corroborated with the two hydrogen bond acceptor features that appeared in the 3D pharmacophore model developed in this work. Besides these, mapping of the parent phenyl moiety with the hydrophobic feature implies its significance in imparting the necessary biological activity to the molecules. Fischer validation performed for the developed model implied the robustness of the developed model. Acceptable values for all the external predictive parameter (R^2_{pred}=0.568, $\overline{r_{m}^{2}{}_{(test)}}$ = 0.516 and $\Delta r_m^2{}_{(test)}$=0.166) for the selected hypothesis indicates that the pharmacophore obtained reflects statistical significance and bears improved external

predictive potential.

The best HQSAR model was obtained with specific fragment contribution (atoms, bonds and hydrogen bond donor & acceptor), fragment size (5-9), hologram length (307) and optimum component number (4). The model (Q^2=0.586) thus obtained was validated externally yielding acceptable values for the external validation parameters [R^2_{pred} =0.736, $\overline{r_{m}^{2}{}_{(test)}}$=0.605 and $\Delta r_m^2{}_{(test)}$=0.168]. The green coloured atoms in the contribution map for the most active compound indicated that the maximum contributing fragments included the ethylene (-CH=CH-) linkage of and the ketonic (=O) fragment of the acid/amide derivatives. This indicated that these fragments are essential for optimum activity of the molecules. The ketonic fragment being in conjugation with the 9, 10-double bond (-CH=CH-) enhances the electron transfer ability and the free radical scavenging action of the molecules through electron delocalization. The fragments contributing moderately (yellow) to the activity profile constituted the C_5 atom of the parent moiety and the secondary carbon atom of the propoxy linkage attached at the C_2 position of the parent phenyl moiety. On the contrary, red colouration of the carbon atom, *ortho* to one of the t-butyl groups, in phenyl ring signified its worst impact towards activity profile of the compounds. The HQSAR model thus denoted that lack of the essential fragments in the parent moiety may result in deterioration of the anti-lipid peroxidative activity of the molecules.

Conclusion

The three different analyses performed in the present work refer to the essential molecular fragments for optimum activity profile of the molecules. All the models infer that the ketonic oxygen fragment of the acid/amide functional group attached to the ethylene carbon which is in turn linked to the C_4 position of the phenyl ring play the prime role in governing the lipid peroxidation inhibitory activity of the molecules. Additionally, the C_4 atom and the ethylene chain also constitute the essential structural features for the molecules. Thus the QSAR models can be efficiently utilized for prediction of lipid peroxidation inhibitory activity of untested molecules belonging to the class of cinnamic acid derivatives. Moreover, the 3D pharmacophore model and the HQSAR model developed in the present work serve as efficient 3D query tools that may be utilized for screening of large databases and selecting the molecules bearing the necessary structural features for exhibiting optimal activity.

***In silico* evolutionary analysis of *c-myc* and its role in oncogenesis causing Burkitt's lymphoma (8.-14)**

Sandeep Kumar, Gaurav Anant1, Baru Ambica2, Louis Balan3

B.M.S.Block, Panjab University& Department of Hematology, PGI Chandigarh, India.

The *in silico* analysis facilitates the analysis of both the known and the unknown genes and proteins. It facilitates quicker and relevant analysis of the data being consolidated by wetlab experimentation.The presented study deals with understanding functional aspects of the proto oncogene thereby identifying the point mutations in the oncogene c-myc. The identification of point mutations plays a significant role in understanding the nature of the tumour caused. Furthermore, the functional understanding can be improved by analyzing and evaluating the domain level differences in the proto oncogene and oncogene. Along with the analysis at the functional levels, *c-myc* gene can be studied for its homology and conserved domains. The aim of this study is to completely understand the oncogenesis caused due to malfunctioning of the proto oncogene and poor association with various transcription factors. To achieve this purpose the mutations and translocations are verified at the sequence level and homology searches are performed for the corresponding nucleoprotein "*c-Myc*".

Background Information

c-myc is the cellular gene, homolog of avian viral mylecytomatosis. It belongs to the Myc gene family composed of retro virus associated DNA sequences. Targeted homozygous deletions of the murine c-*myc* gene results in embryonic lethality, suggesting that it is critical for development.

Composition of C-Myc- The human c-myc gene is located at 8q24 on the long arm of chromosome 8. The sequence of c-myc gene in humans was recognized by using BAC cloning vector. Human cellular oncogene c-myc is composed of three exons and is transcribed from two initiation sites separated by 175-base-pair DNA in HeLa cells. For both resulting mRNA species, exon 1 composes the 5' untranslated region and the initiator methionine is located 16 base pairs down-stream from the 5' splice acceptor of exon 2.The c-myc gene being a transcription factor occupies the centre stage in the G1 phase when the cell gears itself for DNA duplication. Malfunctioning of c-myc leads to alterations in the amount of the tumour suppressor gene p53. Positive functioning of c-myc gives an indication to the cell for proceeding towards the next

event. c-myc helps in the smooth transition of the cell from the mid G1 to S phase. Interaction of c-myc with the cell-cycle inhibitors promotes controlled cell division.

Burkitt s lymphoma (8.-14) - Burkitt's lymphoma is a solid tumor of B lymphocytes, the lymphocytes that the immune system uses to make antibodies. The genes for making antibodies are located on chromosomes 14 (the heavy [H] chains), 2 (kappa light chains), and 22 (lambda light chains).

Material and methods:-

- Analysing the mutations in the *c-myc* proto-oncogene thereby leading to oncogenesis causing Burkitt's lymphoma (8.-14). The analysis aids identifying the mutation, pointing out the missing domains and thereby determining the impact on the gene function
- Establishing the phylogeny of c-Myc protein in multiple organisms. This is done with an aim to understand the evolution of c-Myc nucleoprotein in different organisms ranging from amphibians to higher mammals.

The Study aims at analysing the gene in all the possible aspects. The point mutations and translocations are located using "BL2SEQ". The proto-oncogene protein is used for similarity searching using "BLASTP". The proteins from different organisms with lowest e-value and highest score are used for multiple sequence alignment using "CLUSTAL-W" for analysing the conserved region and understanding the evolutionary relationships via the phylogenetic tree being predicted. The location of the protein action is validated using the tool "PSORT". The nature of the protein is a crucial factor for its physiochemical property determination and is predicted using "ProtScale". Overall all the functional, physiochemical and evolutionary aspects of the c-myc are identified using the *in silico* tools like GENBANK, OMIM, and EXPASY Server.

Conclusion

It is currently regarded as a central transcriptional oncogenic switch that regulates a large variety of cellular functions through altering gene expression. Recent studies suggest that c-Myc is able to activate the cell cycle machinery and its safeguards. Intriguingly, its ability to activate glycolsis suggests that in addition to triggering the cell cycle, c-Myc also sustains the fuel necessary to run the cell cycle machinery. Indeed, its ability to enhance the activities of specific

enzymes involved in DNA metabolism and other metabolic pathways further suggests that it is a key molecular integrator of cell cycle machinery and cellular metabolism. The future of the study of c-Myc target genes lies in the use of arrayed gene expression analysis to determine the common and divergent patterns of c-Myc target gene expression in a variety of physiological and neoplastic conditions. The benefits from such advances in technology, however, will require the expertise of biologists who are able to tease out the roles of the target genes in producing the multitude of c-Myc-mediated phenotypes.

Development of Predictive *In Silico* Classification and Regression Models for *Vibrio Fischeri* Toxicity of Ionic Liquids: Green Solvents for the future

Rudra Narayan Das and Kunal Roy*

Drug Theoretics and Cheminformatics Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, INDIA

Introduction

Ionic liquids (ILs) are considered as less hazardous chemicals in various industrial operations than the conventional volatile organic solvents for their remarkable physical properties in terms of lower vapor pressure, better solvation of chemicals, non-flammability, higher thermal and chemical stabilities, and reusability. Although ILs are lesser harmful chemicals, further assessment of their toxicity is necessary to make them more environmentally friendly '*green solvents*'. Development of predictive mathematical models for ionic liquids can help in designing better derivatives with reduced toxicity profile. The present work has been aimed at developing predictive classification and quantitative structure-toxicity relationship (QSTR) models. Such models provide suitable structural informations of ILs responsible for its toxic manifestation and thereby will guide in the design and development of lesser toxic analogues with its improved property of interest.

Present work and methods

In this study, we have used a dataset of 147 ionic liquids comprising of 30 anions and 64 cations with reported *Vibrio fischeri* toxicity. Various two-dimensional chemical descriptors including the extended topochemical atom (ETA) indices developed by the present authors' group along with indicator variables for cations and anions were used in this present study. Dataset division has been performed using *k*-means clustering technique. Linear discriminant analysis has been used to develop the classification model (toluene served as the reference toxic chemical, pEC_{50}=3.670 mol/L), whereas multiple linear regression method accompanied with stepwise and genetic function approximation (GFA) algorithm have been employed in the regression model. Dragon (version 6.0), cerius 2 (version 4.10), MINITAB, and STATISTICA are the different chemometric tools used in this work. All the developed models have been subjected to statistical validation using multiple validation strategies and the results have been satisfactory.

Results

The discriminant equation with 12 independent variables showed an encouraging Wilk's λ value of 0.298. The model showed highly appreciable sensitivity and specificity values of 87.10% and 96.20% for the training set (N=110), and 60% and 86.36% for the test set (N=37) compounds respectively. We have also performed ROC analysis and the area under the curve (AUC) values has been nearer to unity ($AUROC_{training}$=0.976, $AUROC_{test}$=0.852) corresponding to good quality of achieved discrimination. The classification model principally showed importance of branching (MSD), aromaticity (SaasC), lipophilicity (MLOGP) and molecular bulk (Σα) in explaining differences in toxicity of ILs to *Vibrio fischeri*.

The regression models have been developed and judged according to OECD guidelines. The best regression model was developed using GFA spline technique and has been found to be statistically reliable from internal, external and overall validation tests. We have additionally determined various r_m^2 metrics, developed by the present authors' group [5], for internal, external as well as overall validation. The regression model was characterized by acceptable values (greater than 0.5) of R^2 (0.694), Q^2 (0.651), R^2_{pred} (0.739) as well as other parameters. The applicability domain of the regression model was checked by leverage method leaving no test set chemicals molecule as outlier. Figure 1 shows a plot between the observed and calculated (training set)/predicted (test set) toxicity values of the selected ionic liquids to *V. fischeri* with lesser degree of scattering corresponding to good prediction of toxicity. Different descriptors present in the regression model chiefly explain the effect of branching (shown by GMTI, $^5\chi^v$, and $[\eta']^{local}$) and the solvation entropy and dispersion interactions (shown by $^4\chi^s$).

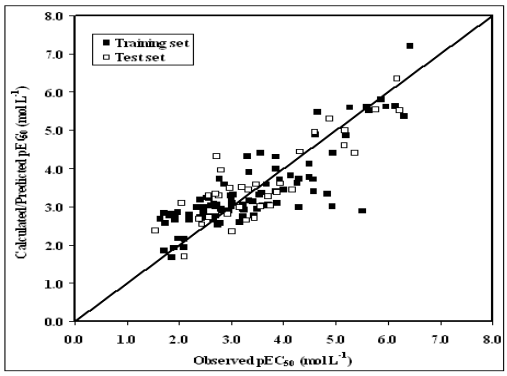


Fig. 1. Observed vs calculated (training set)/predicted (test set) toxicity plot for the regression model

Conclusions

Toxicity due to chemicals is an integrated part of modern developments in industrialization. Room temperature ionic liquids have proved to be good alternative solvents to exploit a wide variety of industrial operations especially in various separation processes and numerous chemical reactions as '*green solvents*'. In this present study, attempts have been made to derive suitable mathematical relationships between the structure and *Vibrio fischeri* toxicity of ionic liquids employing lesser cost of research. The classification model can be useful as a preliminary model for the grouping of toxic or non-toxic ionic liquids; whereas the regression model can be useful for direct prediction of toxicity values. Both the models chiefly account for important contributions of branching, lipophilicity and molecular bulk of the compounds. Good statistical quality of results in both the analyses confirms that these models can be used for the classification and prediction of *V. fischeri* toxicity of the ionic liquids and thereby aiding the design of more "greener ionic liquids".

Re-scoring docking results using MM/PBSA improves prediction of binding energy

Rashmi Kumari and Andrew M Lynn

School of Computational and Integrative Sciences, Jawaharlal Nehru University, India

Introduction

Drug discovery and development process is often time consuming and requires large resources like time, money and human beings. Estimated time and cost for bringing a drug in market vary from 7-12 years and around $ 1.2 billion. Therefore, new approaches are needed for streamline drug discovery and development process for saving these resources. Computer-aided drug design (CADD) is one of such evolutionary technologies. Molecular Docking (part of the CADD), one of the widely used methods of virtual screening computationally screens thousands to millions of organic molecules against the active site of target macromolecule, looking for those with complementary fits and then do scoring for the assessment of ligand binding strength. The technique is often used to find novel ligands for drug discovery. Many approximations are made to screen many molecules in a timely fashion, often resulting in low "hit-rates". These include using only one conformation of the protein, neglecting the internal energies of the docking molecules, using simplified models of ligand solvation energies, typically ignoring protein desolvation, and ignoring most entropic terms entirely. While these methods have been improved respect to the accuracy and efficiency of the available algorithms, drawbacks and limitations still exist. For example, docking techniques still lack reliable simulation of the flexibility of both ligands and receptor. Another major drawback concerns the application of scoring functions that largely fail to estimate ligand binding energies in reasonable agreement with experiment. As a result, false-positive and false-negative hits still populate docking screening results performed with standard methods.

Owing to the above limitations, it is generally agreed that docking results need to be post-processed with more accurate tools to refine docking orientations, filter out poor structures from the docking ensembles, and rank potential ligands to give better agreement with experimental results. An alternative approach involves generation of ensemble of structures using molecular dynamics simulation (MD simulation) and a higher level of theory to evaluate binding free energy from ensemble. Molecular Mechanics Poisson Boltzmann Surface Area (MM/PBSA) is currently thought to be quite effective at incorporating implicit solvation into the estimation of

ligand binding free energies. This more physically realistic method has improved models for solvation, electrostatic interactions and conformational change compared to most docking programs. The term MM/PBSA stands for Molecular Mechanics Poisson Boltzmann Surface Area. The overall objective of the MM-PBSA method is to calculate the free energy difference between two states which most often represent the bound and unbound state of two solvated molecules or alternatively to compare the free energy of two different solvated conformations of the same molecule.

$\Delta G_{bind,\,solv} = \Delta G_{bind,\,vacumm} + \Delta G_{solv,\,complex} - (\Delta G_{solv,\,protein} + \Delta G_{solv,\,ligand})$

In the above equation, on right hand side $\Delta G_{solv,complex}$, $\Delta G_{solv,ligand}$ and $\Delta G_{solv,protein}$ are solvation free energy of complex, ligand and receptor. $\Delta G_{bind,vacuum}$ is free energy of binding in vacuum.

MM/PBSA is already implemented, optimized and validated in the AMBER package. MD simulations through AMBER are still computationally expensive and high throughput rescoring of the docked complexes is time consuming. Additionally, the MD simulation module of AMBER package is not available as an open source program. These two major drawbacks motivated this work with the objectives of implementation, validation and optimization of MM/PBSA through open source program. GROMACS and APBS are used for the MD simulations and PB calculation respectively.

Objectives

- Implementation and automation of MM/PBSA in GROMACS
- Validation of method by comparison with experimental results
- Optimization of the protocol
- Re-scoring of top-ranked docked complexes obtained from virtual screening

Implementation and Automation: The docked complexes obtain from virtual screening has undergone following three steps for getting the relative binding energy and is automated through in-house written perl script:

- Input files are prepared for simulation by separating the complexes into protein and ligand and then processing it using GROMACS and Ambertools respectively.
- MD simulation of 10ns is performed for each complex after minimization and successive release of position restrain allowing flexibility to both protein and ligand.
- Ensemble of conformations obtained from trajectory of MD simulation, has been further used for calculation of relative binding free energy.

Validation and Optimization: Large datasets are needed to be validated to show high throughput capability of this implementation and also, there are lot of scope for the optimization of the entire protocol and the method. Therefore, the whole method is implemented using a perl script written in-house and validated on the HIV Protease I with inhibitors that have a broad K_i range. Further optimization of the method has been performed by changing parameters of polar and non-polar solvation energy calculation, concerning various literatures to get high correlation between calculated and experimental results. After optimization, a good correlation coefficient of 0.832 is obtained between calculated relative binding energy and the experimental $logK_i$ values as shown in the figure below.

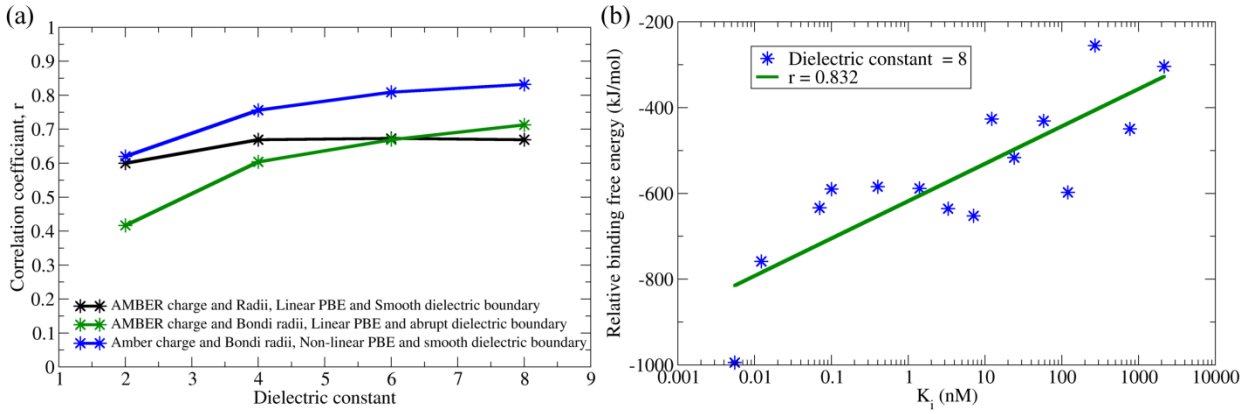


Figure: (a). Correlation between experimental binding free energies and MM/PBSA calculations, changing different parameters of polar solvation at different dielectric constants. (b) Correlation between MM/PBSA predicted and experimental binding free energies at dielectric 8.

This optimized method is used to re-score the docked complexes of DHDP (DiHydroDiPicolinate) reductase, a key enzyme of Diaminopimelate pathway of *M. tuberculosis*. Docked complexes of some other known *M. tuberculosis* target are also re-scored which clearly discriminate between active and in-active compounds from docking hit-list. Re-scoring of top ranking inhibitors from docking experiments by MM/PBSA can be used as a filter to more accurately select molecules for experimental validation.

Conclusion: Implementation of this method with GROMACS will speed up MD simulations by many times and it will open a platform for the high throughput re-scoring of protein-ligand complexes by MM/PBSA, which would be a major leap in the drug discovery process. After the re-scoring, there will be fair chance that screened ligand will show biological activity and synthesis of those ligand may become lead compound in the drug discovery process.

Simplified Receptor Based Pharmacophore Approach to Retrieve Potent Ptp-Lar Inhibitors using Apoenzyme

Dara Ajay and M. Elizabeth Sobhia*

Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), S.A.S. Nagar, Punjab, INDIA.

The design of biological active compounds from the apoenzyme is still a challenging task. Herein a simple yet efficient technique is reported to generate a receptor based pharmacophoresolely using a ligand-free protein crystal structure. Human leukocyte antigen-related phosphatase (PTP-LAR) is an apoenzyme and a receptor like transmembrane phosphatase that has emerged as a drug target for diabetes, obesity and cancer. The prior knowledge of the active residues responsible for the mechanism of action of the protein was used to generate the LUDI interaction map. Then, the complement negative image of the binding site was used to generate the pharmacophore features. A unique strategy was followed to design a pharmacophore query maintaining crucial interactions with all the active residues, essential for the enzyme inhibition. The same query was used to screen several databases consisting of the Specs, IBS, MiniMaybridge, NCI and an in-house PTP inhibitor databases. In order to overcome the common bioavailability problem associated with phosphatases, the hits obtained were filtered by Lipinski's Rule of Five, SADMET properties and validated by docking studies in Glide and GOLD. These docking studies not only suggest the essential ligand binding interactions but also the binding patterns necessary for the LAR inhibition. The ligand pharmacophore mapping studies further validated the screened protocol and supported that the final screened molecules, presumably, showed potent inhibitory activity. Subsequently, these molecules were subjected to Derek toxicity predictions and nine new molecules with different scaffold were obtained as non-toxic PTP-LAR inhibitors. The present prospective strategy is a powerful technique to identify potent inhibitors using the protein 3D structure alone and is a valid alternative to other structure-based and random docking approaches.

Methodology:As it is difficult to generate a pharmacophore query from a homology model or the apoenzymes alone, without having any prior information on the inhibitors or its complexes. The aim of the current study is to fill this gap by using the knowledge based approach to generate

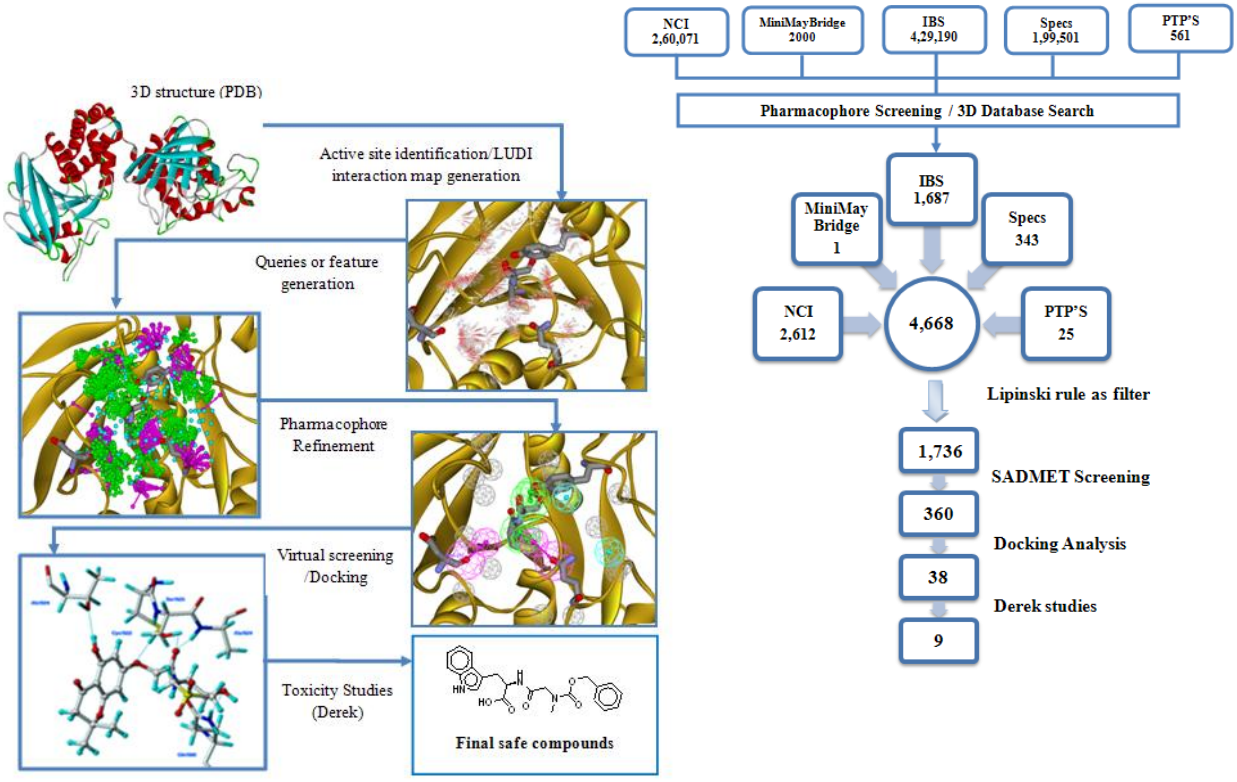
pharmacophore model from the active residues responsible for the mechanism of action of the protein. To design and optimize potent LAR inhibitors we followed combinatorial approaches which include receptor based pharmacophore generation, 3D database search, SADMET screening, and molecular docking followed by Derek toxicity prediction (Fig)

Results and Discussion: The results from the LUDI interaction protocol were subjected to two different strategies to generate a rational and valid pharmacophore model. This was used to screening compounds from 5 various 3D Databases, the hits obtained were filtered by Lipinski's Rule of Five, SADMET properties and validated by docking studies in Glide and GOLD. The 3D information provided by these docking studies is meaningful, as essential interactions required for the inhibition are retained by the screened molecules. Finally, toxicity prediction by Derek yielded nine molecules out of 38 ligands that are free from toxicological endpoints

Conclusions: The pharmacophore concept has proven to be extremely successful, not only in rationalizing the structure-activity relationships, but also by its large impact in developing appropriate 3D-tools for efficient virtual screening. In the present study, knowledge-based approach was used to design a simple and efficient technique to generate a receptor-based pharmacophore as a part of structure-based drug design strategy. Our method can thus be used independently to identify potent hits molecules without using any prior information from the ligand or its complexes. The LUDI interaction map and the complementary pharmacophore features of the active site were used to preferentially target the features of the active residues alone and generate 3D queries. Based on this concept a pharmacophore model was designed with two hydrogen-bond acceptor (A), two hydrogen-bond donor (D), and two hydrophobic (H) features. The docking analysis also showed a minimum of four interactions being maintained out of the six pharmacophore features selected, highlighting cogent evidence to confirm the mentioned interactions to be indeed crucial for ligand-protein interactions. The final nine screened molecules were superimposed onto the pharmacophore model of strategy-II. The results indicate that out of the 9 screened molecules, 8 molecules were mapped unto the pharmacophore fitting method. This showed that the screened molecules have the characteristic pharmacophore features as designed by strategy-II. The ligand pharmacophore mapping studies further validated the screened protocol and supported that the final screened molecules presumably showed potent

inhibitory activity. Although we could not achieve large-scale benchmarks, the interactions obtained from these docking studies ameliorate the accuracy of the model. However, the general protocol presented in this study is sensitive and specific enough to prioritize virtual hits of interest using an apoenzyme exclusively.

Finally, we have successfully identified new potent inhibitors using a unique sequential in silico strategies comprising of the receptor based pharmacophore, SADMET based virtual screening, and docking study followed by toxicity study using Derek. These methods allowed us to screen nine compounds of different scaffolds as PTP-LAR inhibitors from a databases containing 8,91,323 molecules. The compounds obtained from the NCI and PTP databases were found to be promising and free from toxic endpoints. The compounds PTP-13811 and NCI-361664 could be further subjected to in vitro analysis in the search for better leads that could improve the treatment of obesity and diabetic disease.



POSTER PRESENTATION

Analysis of Functionally Diverse Toxic Peptides

Arun Sharma[1] and GPS Raghava[2]

[1] Project Assistant, Bioinformatics Centre, Institute of Microbial Technology, Chandigarh, India.

[2] Head, Bioinformatics Centre, Institute of Microbial Technology, Chandigarh, India

Peptides are well known for their diverse therapeutic properties. A number of peptides having antibacterial, anticancer, cell penetrating, tumor homing, antiviral, antihypertensive, immunemodulatory, antithrombotic and antioxidant properties have been reported in literature and very well complied in peptide specific databases such as CPPsite, APD2, ANTIMIC, PhytAMP, CAMP etc. In this study, we have used datasets of NTXpred and analyzed selected functionally diverse toxic peptides. NTXpred classified diverse neurotoxins into five functionally different classes. In order to differentiate peptides from proteins, cut-off length of 100 residues was chosen during present study. Separately we analyzed each functional class for its amino acid composition against composition of equal number of non-toxic peptides. In comparison with non-toxic peptides from Swiss-Prot database Cys, Gly, Lys, Asn and Tyr residues are preferred in class I peptides (involved in blocking of ion channels). The class II peptides involved in blocking of acetylcholine receptors showed relative abundance of Cys, Asn, Pro, Arg, Thr residues. The class III of peptides, inhibiting acetylcholine release via phospholipase A2 activity, was found with Cys, Gly, Lys, Asn, Gln, Thr and Tyr residues abundancy. Asn, Thr and Tyr residues were found relatively more in toxic peptides in contrast to non-toxic peptides. No peptide was retrieved with length below 100 residues, from class IV peptides, involved in Inhibition of acetylcholine release via metalloproteolytic activity. Finally analyzed class V toxic peptides (facilitating acetylcholine release), showed the most abundance of Cys, Asn, Ser and Tyr residues. In conclusion, most of toxic peptides showed abundance of Cys and Asn residues. Instead, the Cys residue abundance can be attributed to disulfide bond formation and stabilizing 3-dimensional structures of these peptides. This analysis of functionally diverse peptides will be helpful in understanding the preference of certain amino acid residues among neurotoxic peptides over non-toxic peptides. It may prove a milestone during designing of therapeutically useful peptides having optimal toxicity increasing their usage in pharmaceutical industry.

Pharmacophore-based virtual screening and docking studies on HDAC Inhibitors

Monika Sharma, Astha Gupta, Madhu Chopra*
Dr. B. R. Ambedkar Center for Biomedical Research,University of Delhi, Delhi, India.

Histone Deacetylase (HDAC) has emerged as an effective anticancer target. Currently many HDAC inhibitors are in clinical trials but ongoing research is to find more potent inhibitors with improved pharmacokinetic properties and therapeutic index. As the role of various HDAC subtypes have been found to be related to many other cellular functions and with many other diseases, so the discovery and development of Type –Specific HDAC inhibitors is of research interest in pharmaceutical and academic labs. Ligand Based Pharmacophore Modeling is an important method for the identification of features in ligands that can modulate the activity of a particular drug target. We summarize in this paper, the development of pharmacophore model from a dataset of inhibitors for HDAC by using the Discovery Studio 2.5 with chemically diverse series of compounds. A training set consisting of 22 compounds was carefully selected considering structural diversity and wide coverage of activity range. The most predictive pharmacophore model (hypothesis 1)consisting of four features, namely 3 hydrogen bond acceptor and one Hydrophobic Aromatic one aromatic feature, had a correlation (r) of 0.95. The entropy (configuration cost) value of the hypotheses was 16.41, within the allowed range. The difference between the null hypothesis and the fixed cost and between the null hypothesis and the total cost of the best hypothesis (hypothesis 1) was 103.44 and 81.7 respectively. The model was validated using the Fischer's cross validation test. A test set consisting of a different series of structurally diverse 25 compounds was built to validate the predictive ability of the pharmacophore. This validation approach provides confidence in the utility of the predictive pharmacophore model developed in this work as a 3D query tool in the virtual screening of drug like molecules to retrieve new chemical entities as potent HDAC inhibitors. The validated pharmacophore model (Hypo-1) was used as a 3D query for virtual screening to retrieve potential inhibitors from the Maybridge and druglike subset of Zinc database. The hit compounds were subsequently subjected to molecular docking studies using CDOCKER.

In-silico Predictive Mutagenicity Model Generation using Supervised Learning Approaches

Anurag Passi[1]*, Abhik Seal[2$]*, U.C. Abdul Jaleel[3], David Wild[1], OSDD Consortium[2]

[1] Open Source Drug Discovery, Council of Scientific and Industrial Research, India

[2] Indiana University Bloomington School of Informatics and Computing, Bloomington, USA

[3]Department of Cheminformatics, Malabar Christian College, Kerala, India.

Background

Screening of chemical compounds is a time consuming and resource expensive practice. To speed up the drug discovery process, there is always a need to shortlist compounds most relevant to the study so as to have a specific chemical space to work with. The aim of this work was to develop *in-silico* predictive mutagenicity models than can classify random sets of compounds into mutagens or non-mutagens. This work is based on the idea that molecular descriptors quantify the structural information of a molecule, which in turn, defines its activity. The Bursi mutagenicity data set and a Benchmark data set were taken as inputs in this work. A third data set (Set 3) was prepared by joining the previous two sets. Classifier algorithms such as Naïve Bayes, Random Forest, J48 and SMO with 10 fold cross-validation and default parameters were used for model generation on these data sets. It was observed that models built using Set 3 performed better than those developed from the Benchmark data set. Moreover, for each data set it was observed that Random Forest outperformed other classifiers for all the data sets, especially for Set 3 with 89.27% accuracy, 89% precision and ROC of 95.3%. To validate the developed models two external data sets, AID1189 and AID1194, with mutagenicity data were tested showing 62% accuracy with 67% precision and 65% ROC area and 91% accuracy, 91% precision with 96.3% ROC area respectively. Moreover, we tested the Set 3 Random Forest model on the approved drugs from DrugBank and metabolites from the Zinc Database with True Positives rate almost 85% showing the robustness of the model. Thus the predictive models developed would be able to predict and classify the sets of compounds based on their structural features acting as filters to screen the most relevant hits in the drug discovery process.

Microwave Assisted Synthesis of Flouro Chloro Benzimidazolo Substituted Thiazolidinone Derivatives for Antimicrobial Activities

Singhal L*, Patil S, Hippargi SM, Meharoon D

Department of Pharmaceutical Chemistry, Kle universitys College of Pharmacy Rajajinagar Bangalore, India.

Abstract

Flouro chloro benzimidazolo substituted thiazolidinone derivatives synthesized by reacting 3-chloro, 4-flouro ortho phenylenediamine with para amino benzoic acid and the aspartic acid respectively followed by different aromatic aldehyde and thioglycolic acid in presence of aluminium chloride. The compound shows absorption bands ranging from 3433- 3320 cm^{-1} for N-H, 3149-3034 cm^{-1} for C-H aromatic stretching and 1521-1342 cm^{-1} for NO_2 functional group. In ^{1}HNMR the presence of methylene proton and methyl protons between δ 2.49 ppm and δ 3.31 ppm respectively was observed respectively. For aromatic protons multiplets were observed between δ 6.8-7.25 ppm and N-H δ 6.8ppm.

Result

We have synthesize series of 10 derivatives of substituted Benzothiazoles with different amines by reacting m-nitroanilines with potassium thiocynate, 2-amino-5-nitro(1,3)benzothiazole with sodium nitrite yield diazonium salt, which on coupling gives substituted anilines and napthols to yield Benzothiazoles been reported as anti-microbial agents.

Conclusion

The derivatives screened were compared with the standard amoxicillin(100µg/ml). In anti-microbial activity sp103, sp 203, sp 206, sp 207 show better activity in gram(-)ve bacteria E.coli, and sp 201 and sp 207 show good activity in gram(+)ve bacteria carried by agar diffusion method and the radius of zone of inhibition was recovered.

A New Python Script for Qsar Study

Vadiraj Kurdekar, Hemant R. Jadhav

Department of Pharmacy, Birla Institute of Technology and Sciences Pilani, Pilani Campus, India.

Quantitative Structure Activity Relationship (QSAR) is the stastical correlation of physicochemical properties of the structure with biological activity. QSAR study involves two main steps, first is the generation of descriptors and second is building and validating the models, so that it can be efficiently used for predicting activity of unknown compounds. Many QSAR models published in the literature are not utilized for designing new drugs because these models are developed using costly commercial software and it prevents many researchers form testing and adopting published models. If computational models were generated by open source software then it can be more easily shared in scientific community.

We have developed a python program for QSAR studies using Chemistry Development Kit (CDK, a open source molecular descriptors generator) and various python modules. Initially descriptors have to be generated using CDK. Then using python modules like Numpy, Scipy, Python-Sklearn and matplotlib; linear algebra calculations, statistical values, machine/statistical learning, data mining and plotting, respectively can be done using our software.

Our python code, by using above modules will generate coefficient of determination (R^2) matrix, Multiple Linear Regression R^2, Adjusted R^2, F statistics, p-value, Leave One Out (LOO) R^2, Leave Many Out (LMO) R^2, Y-scrambling and Bootstrap R^2. We used this software to study some samples and was found to perform satisfactorily.

Evaluation of Eukaryotic Gene Prediction Programms

Author: Divya Singhal*, Pooja Sharma, Moksha Shandilya, Hardeep Kaur, Kanika Malhotra, Shivalika Sharma

Deptt. of Biotechnology, Ambala College of Engg. & Applied Research, Ambala, India.

Gene finding typically refers to the area of computational biology that is concern with algorithmically identifying stretches of sequence, usually genomic DNA, that are biologically functional. This specially includes protein coding genes but may also include other functional elements such as RNA gene and regulatory regions. Gene finding is one of the first and most steps in understanding the genome of specie once it has been sequenced. Gene prediction software's are bioinformatics tools to predict the gene structure of a given sequence in Fasta format. Gene prediction involves determining the number and location of exons (initial, intermediate or terminal), number and location of introns, CDS region, location of promoter and terminal regions (PolyA). In this study, various windows based online gene prediction software's were compared against genebank sequences for 5 different sequences. Softwares used were: HMMgene, EMBOSS, FGENESH, GENMARK and GENSCAN. Results were analyzed by calculating specificity and sensitivity at nucleotide and exon level. Correlation coefficient, average conditional probability, approximate correlations were calculated and compared to determine most efficient software for use. FgeneSH software found to be best eukaryotic gene prediction software.

Evolutionary relatedness among viral miRNAs

Sujit Kumar Pujhari[1], and Akhlash P Singh[2]

Department of Virology1, PGIMER, Chandigarh, Departments of Biochemistry [2], GGDSD College, Chandigarh, India.

MicroRNAs (miRNAs) are small, noncoding RNA molecules that direct post-transcriptional suppression of gene expression. miRNAs are found to be encoded by the simpler prokaryotic to complex eukaryotic genome. Interestingly viruses that do not have an independent existence outside the permissive living host system also encodes for the miRNA. Viruses use these miRNAs as an effective modulator to sustain inside the host; especially to escape the host immune system, modulate cell cycle and other machinery of the host to replicate, turn latency related genes and many other. In the present study miRNA sequences in the miRBase were browsed. The viral miRNA sequences were aligned and a phylogenetic tree was constructed. There are 240 viral miRNA sequences from 23 viruses; of them 94 miRNA sequences were from 10 human viruses. All viruses encoding miRNA have DNA as their genetic material or have a DNA phase in their lifecycle eg. HIV. This study demonstrates the evolutionary relatedness among viral miRNA that were placed in three distinct clusters.

Generation of Ligand-Based Pharmacophore Model And Virtual Creening for Identification of Novel Mbta Inhibitors With Potent Antitubercular Activity

Lakshmi Maganti, Nanda Ghoshal*

Structural Biology and Bioinformatics Division, CSIR - Indian Institute of Chemical Biology, Kolkata, India.

Background: One of the main goals in drug discovery is to identify new chemical entities that have high likelihood of binding to the target protein to obtain the desired biological response. To this end, pharmacophore modeling and 3D database searching are now recognized as integral components of lead discovery and lead optimization. MbtA is a key enzyme involved in siderophore biosynthetic pathway in Mycobacterium tuberculosis. It is an aryl acid adenylation enzyme and initiates biosynthesis of the aryl-capped siderophores, which are critical for virulence and growth of the bacteria. Only a few inhibitors for MbtA are currently known. To discover more potent and diverse inhibitors we have conducted pharmacophore-based virtual screening.

Results: Virtual screening of the commercial databases was done by using chemical feature based pharmacophore model, which has been generated from known Mbt A inhibitors. The top ranked hypothesis contained five features: three hydrogen-bond acceptors, one hydrogen-bond donor and one hydrophobic aromatic. The initial *in silico* screening of a library of 1.6 million commercially available compounds against pharmacophoric query and using sequential filter of drug-likeness, yielded a hit list of 67 compounds. The obtained hits were further evaluated by molecular docking and 16 compounds were short-listed based on docking scores and the binding site interactions. Resulting hits were subsequently subjected to ADMET screening to enrich the retrieved hits.

Conclusion: The pharmacophore model developed in this study would be used as a query to browse through different databases and retrieve compounds with pharmacophoric features known to be critical for inhibitory activity. The discovery of potent and selective MbtA inhibitors by ligand based 3D pharmacophoric virtual screening indicates the efficiency of this kind of approach in drug discovery program, which is economical and saves considerable amount of labor and time.

Metabolic Modeling of Autoimmune Skin Disorder Targeting Psoriasis through Systems Biology Approach

Shikha Agnihotry[1], Shivangi Singh Parihar[1], Manish Kumar Gupta[1*], Manisha Lalwani[1]

[1]Department of Bioinformatics, University Institute of Engineering and Technology, Chhatrapati Shahu Ji Maharaj University , Kanpur, U.P. India.

Background:

Psoriasis is a chronic inflammatory autoimmune disorder associated with many genetic and environmental factors. Psoriasis is characterized by well-defined red scaly plaques on scalp, knees and elbows. Systems Biology approach is anticipated to enhance our understanding of complex behavior of cells, through the integration of molecular components and their interactions. In this study, we present an exhaustive map of the abnormal psoriasis pathway curating reactions from 49 original papers. The metabolic pathway and signaling pathway responsible for the occurrence of psoriasis plaques is constructed through Cell Designer (ver. 4.1). The simulation results yields in finding the novel targets for the treatments of disease. The simulation result shows that the concentration of TNF-α remains constant with time, concentration of IL-2 falls with time and that of IFN-γ rises with time. The current therapeutic drugs mainly target TNF-α and IL-2 which are produced in later stages of psoriasis.

Result:

IFN- γ can be considered as a novel target for the drug discovery for psoriasis and curcumin can act as a potential drug candidate for psoriasis based on its docking analysis and least energy-score value.

Reversible Inhibitors to Combat Tuberculosis: Homology Modeling, Virtual Screening And Comparative Docking Studies

Sanjukta Patra*, Parameswaran Saravanan

Department of Biotechnology, Indian Institute of Technology Guwahati, Assam, India.

Cell wall and lipid metabolism plays a vital role in the survival and infection of *Mycobacterium tuberculosis*. Increase in the drug resistant tuberculosis worsen the existing scenario and urge the need of new druggable targets and new drugs. Targeting Rv3802c, an essential cell wall lipase which participates in cell wall remodeling, can open up a new arsenal to fight the dreadful pathogen. Based on the experimental studies on Rv3802c and its role in the cellular integrity from *M. smegmatis*, Rv3802c is considered to be good drug target for treating tuberculosis. Our current study concentrates on finding potent reversible inhibitors against Rv3802c, which lack adversities of irreversible inhibitors. Bioinformatics analysis identified Ser175, Asp268 and His299 as the catalytic residues of Rv3802c. 3D structure of Rv3802c is predicted for the first time which provides insight in identifying the substrate binding sites and identified Thr83 and Gln176 as the oxyanion hole residues. 80% similarity search of reversible inhibitor [PDB id: ZYH] complexed with human MGL [PDB id: 3PE6] is considered as the initial data set for virtual screening with AutoDock. Followed by virtual screening on 448 molecules, comparative docking studies of Rv3802c with its closest human structural homolog has been carried out to identify potential mycobacterial specific inhibitors. ZINC22795517 identified as a potent mycobacterial specific reversible inhibitor with difference in predicted binding free energy of -2.17. Rv3802c is a promising drug target and ZINC22795517 could be a mycobacterial specific potential inhibitor and also a step towards enriching the drug candidates to combat tuberculosis.

Analysis of Cell Penetrating Peptides

Ankur Gautam, Harinder Singh, Rahul Kumar, Kumardeep Chaudhary and G.P.S. Raghava
Bioinformatics Centre, Institute of Microbial Technology (CSIR), Chandigarh, India.

Cell-penetrating peptides (CPPs) are short peptides, show no cell type specificity and are able to translocate into eukaryotic cells in a receptor independent manner without causing significant membrane damage. In addition, they have enormous capability to deliver various conjugated cargoes such as proteins, siRNA and drugs inside the cell with high efficiency. Therefore, they are considered as a versatile delivery vehicle. We have collected 843 CPPs from literature and analyzed them. We have also developed a database- "CPPsite".

We have observed that most of the CPPs are consists of 9-35 amino acids and are cationic in nature. In order to evaluate the abundance of particular type of residues in CPPs, we have computed average amino acid composition of these peptides and compared with swissprot sequences (randomly generated sequences of 5-30 amino acids). Analysis revealed that Arg, Lys, Ala, Leu and Pro are more abundant in CPPs, while Asp, Glu, Gly, Val, Ser, and Thr are more abundant in swissprot sequences. We have also generated sequence logos of 10 N-terminal and C-terminal residues using online software. Although overall CPPs are dominated by positively charged residues, certain residues for example Gly, Leu, Ile, Ala, Trp and Ser are also preferred at specific positions in CPPs. Since structure of CPP play important role in their interaction with membrane and subsequently their internalization, we have predicted the structure of all the CPPs and secondary structure composition of each peptide has been calculated. Most of the peptides have less than 20% helix and beta-strand content. Only limited peptides have helix or beta-strand content more than 80%. In case of turn, it is observed that peptides having turns up to 60% are equally distributed. In case of coil, a large number of peptides have coil content in the rage of 20-80%. This is expected as it is difficult for a small peptide to maintain a regular secondary structure. Analysis of CPPs is important to understand their properties. In future, we anticipate more detailed analysis, which will be helpful for designing novel CPPs.

Distribution of Chi1 Dihedral Angles in Peptides and Proteins: A Comparative Analysis

Sandeep Singh[1], Harinder Singh[1] and G.P.S. Raghava[1]

[1]Bioinformatics Center, CSIR-Institute of Microbial Technology, Chandigarh 160036, India.

Side chain prediction plays an important role in the accuracy of protein structure prediction. The distribution of Chi1 torsion angles are not random but occupy only certain range of values. In this work we analysed distribution of chi1 dihedral angle of all amino acids (except ALA and GLY) in peptides and proteins separately. 1819/7995 peptide/protein chains were obtained from PDB respectively after filtering data at resolution <=2, redundancy 30% and peptide/protein length <= 22/80 respectively. The distribution of Chi1 torsion angle for each amino acid in both the datasets shows a Gaussian distribution within selected range of angles in the Chi1 conformational space making a clear cut 4 regions (except PRO which has one region from 42˚ to -38˚) where majority of Chi1 values fall. These regions are defined as: R1 from 179˚ to 161˚, R2 from 77˚ to 51˚, R3 from -42˚ to -96˚ and R4 from -152˚ to -179˚. A step further, calculation of Chi1 percentage of each residue in each region was done in both the datasets for comparison. R2 region in cysteine occupies less chi1 distribution in peptides as compared to proteins but in case of phenylalanine the trend is reversed. In peptides, tryptophan has more chi1 distribution in R3 and less chi1 distribution in R4 as compared to proteins. We are also working on distribution of chi2, chi3, chi4 and chi5 in peptides and proteins as well as dependence of distribution of chi1 torsion angles based on different phi-psi ranges.

Table1: Percentage of chi1in each region in peptides and proteins for C, F, H and W residues having at-least 5% difference in any region.

In Peptides					In Proteins				
Residue	R1	R2	R3	R4	Residue	R1	R2	R3	R4
C	10.48	8.54	66.91	14.06	C	11.44	15.78	55.06	17.71
F	16.34	14.90	51.65	17.09	F	16.33	9.33	55.96	18.36
H	12.83	8.81	51.40	26.94	H	11.99	10.90	55.88	21.21
W	15.79	13.87	60.32	10.00	W	15.61	12.77	52.44	19.16

Benchmarking Of All Genotypic Hiv Coreceptor Usage Prediction Methods

Ravi Kumar & Gajendra P.S. Raghava

Bioinformatics Centre, CSIR – Institute of Microbial Technology, Chandigarh, India.

HIV-1 cell entry commonly uses CD4 as primary receptor and one of the chemokine receptors CCR5 or CXCR4 as coreceptors. Knowledge of coreceptor usage is important for monitoring the disease progression as well as for supporting therapy with the novel drug class of coreceptor antagonists. All simple heuristics e.g. 11/25 rule as well as statistical learning methods proposed to date predict coreceptor usage mostly based on sequence features of the V3 loop while avoiding expensive phenotypic assays. 11/25 charge rule was the first of the genotypic methods which determined the HIV coreceptor usage based on the presence of basic (positively charged) amino acids at 11^{th} or 25^{th} position of V3 loop confers the CXCR4 coreceptor usage. According to Pillai S et al., Four classifiers (Charge rule, SVM, C4.5 and PART) were able to discriminate between virus capable of using CXCR4 as a coreceptor versus those which were incapable with accuracies of 87.45%, 90.86%, 89.51% and 89.37% respectively. Jensen M et al. used PSSM matrix to score V3 sequence of HIV-1 subtype B and C, the higher the score, the more similar the given V3 sequence is to an average actual X4 sequence. Random Forest as a machine learning technique was utilized by Xu S et al., based on 37 random features of V3 peptide (35 amino acids with net charge and polarity) developing a prediction model which achieved 95.1% accuracy for coreceptor usage. The clinical and host data information is inculcated by Sing T el al., and compared the performance of 11/25 rule, SVM, PSSM based approach in HIV coreceptor tropism prediction. Sander O et al used the structural descriptors of V3 peptide and found that prediction model improved using both structural and sequence information of V3 peptide. In another experiment, SVM with distant segment kernel was used by Boisvert S et al and achieved high accuracies of 96.35% (R5), 94.80% (X4) and 95.15% (R5X4) with this approach.

Most of the studies regarding the HIV coreceptor usage have been concentrated only to V3 sequences but Prosperi MCF et al. investigated the associations of the whole HIV-1 envelope genetic features and clinical markers with viral tropism. Masso M et al. used random forest, SVM, boosted decision tree, and neural network machine learning algorithm to develop an efficient structure based model. Thelian A et al. reported the impact of the V2 loop on coreceptor usage prediction. In an experiment, the author Dimonte S et al., access the signatures (mutations) in V3 and gp41 region which were statistically correlated with coreceptor usage.

Analysis of antibody class-specific B-cell epitopes

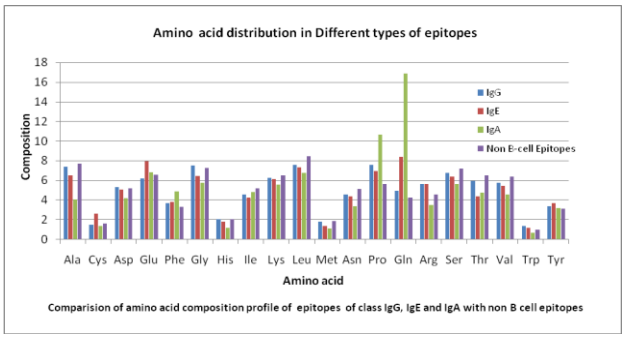
Sudheer Gupta, Dr G.P.S. Raghava

Bioinformatics Centre, Institute of Microbial Technology, Chandigarh, India.

The interaction between antigen and antibody triggers the particular kind of effector function which leads to the elimination of the antigen or pathogen from body. Tongren et al have elaborated that antigen or epitope can dictate the outcoming antibody isotype and thus effector function, therefore it is very important to explore the antigen or epitope features which are responsible for antibody class determination. We analysed the experimentaly validated B-cell epitopes from IEDB which have been shown to generate different classes of antibody. The final data contained 7598, 2343 and 403 epitope sequences for IgG, IgE and IgA classes respectively. The purpose of the study is to see the general and residual characteristics of B-cell epitopes belonging to various classes of antibody.

Results:

After observing the length distribution we found that most of the IgG specific epitope lie in the range of 4-20 residue while for IgE it is 4-15 and for IgA it is 4-10 residues. The amino acid composition profile shows that four epitope classes differ in composition profile and certain residues are preferred or not proffered for every class like the composition of Gly, Gln and Arg in IgG epitopes. In case of IgE epitopes, Ala, Cys, Glu, Gln and Pro residues show difference. IgA epitopes show change in composition of Ala, Phe, Pro, Gln, Arg and Thr amino acids.



Comparision of amino acid composition profile of epitopes of class IgG, IgE and IgA with non B cell epitopes

Conclusion: The results of this study prove the existence of difference between the B-cell epitopes belonging to different antibody class. This study increases the understanding of class-specific properties of B-cell epitopes which may help in generation of particular class of antibody or increasing therapeutic efficiency of an antibody.

Data-mining Approach for Identification of *M. tuberculosis* Inhibitors Using PubChem BioAssays

Deepak Singla[1,2] , Rupinder Tewari[2], Ashwani Kumar[1], Open Source Drug Discovery Consortium and G.P.S Raghava[1]

[1] Bioinformatics Centre, Institute of Microbial Technology, Chandigarh, India.

[2] Centre for Microbial Biotechnology, Panjab University, Chandigarh, India.

Background

The innovation of new therapeutic molecule has long been recognized as time-consuming and labor-intensive, costing on average about $800 million as well as 10–12 years. Lack of publicly available biomedical assay data may represent another barrier in the drug discovery process. Fortunately, this is changing since more public resources like Protein Data Bank (PDB), PubChem, DrugBank and KEGG are emerging, and offering new opportunities to chemical biology researchers for drug development. At present, PubChem bioAssay hold total 28 cell-based as well as target based *Mycobacterium tuberculosis (M.tb)* H37Rv bioassays. These contains screening data for lakhs of compounds classified into inhibitors and non-inhibitors as a results of high-throughput screening effort. Our major goal is to identified potential multi-targeted non-toxic inhibitors of *M.tb.*

Results

In this study, we have chosen four assays two for inhibitors searching and two for toxicity. Firstly, we searched for inhibitors that are active in both replicative as well as non-replicative phase. Secondly, we searched for compounds that are inactive in bioassays involving human protein as a target. Finally, we have removed all the compounds that are toxic to THP-1 and Vero E6 cell lines. Thus, we got only three compounds which are active non-toxic, inhibitors of m.tb.

Conclusion

Based on this study, we conclude that these compounds may act as lead molecules and further optimization of these may help in finding a better compound than existing ones against drug resistant mycobacterium. We also assume that this approach can also be applied to find out lead molecules against other neglected diseases.

Analysis and Identification of Drugable Molecules

Sandeep Kumar[1], Deepak Singla[1,2], GPS Raghava[1]

[1] Bioinoformatics centre, Institute of Microbial Technology, Chandigarh, India.

[2]Centre of Microbial Biotechnology, Panjab University, Chandigarh, India.

Background: Predicting druglike molecules is always crucial for drug discovery process. In past, different rules (like Lipinski rule of five, Oprea rules) have been derived for identifying druglike compounds. However these rules are not universally applicable. Thus classifying the approved vs. experimental drugs is the need of hour. When a lot of time and money is at the stake, only a fraction of molecules finally get approved. Thus, to rank molecules based upon drugability properties is a major issue. This study is an attempt to analyze the differences in approved from experimental drugs present in DrugBank database.

Results: Our study shows that a better predictive model can be developed using 2D fingerprints calculated from freely available softwares.

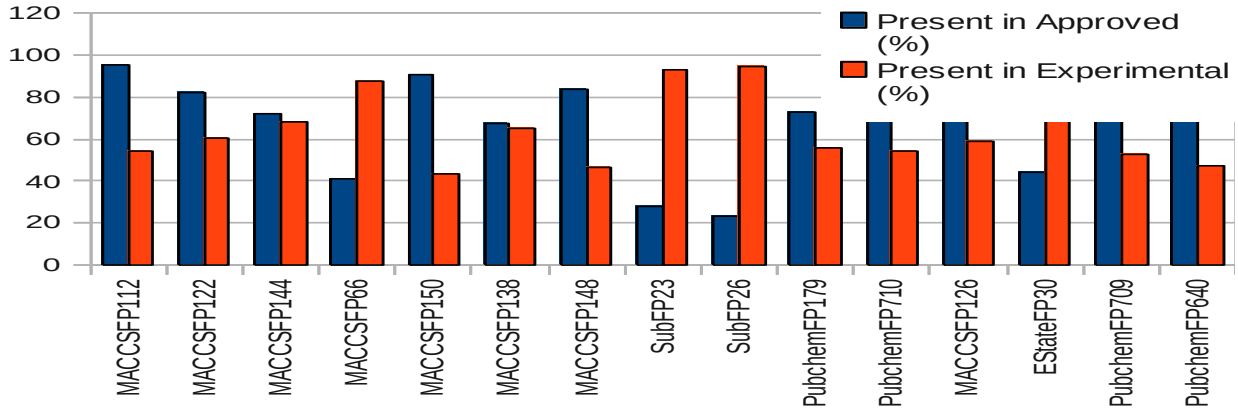


Fig 1. Top 15 descriptors versus their presence in approved and experimental drugs

Conclusion: This study demonstrates that binary descriptors calculated using open source PaDEL software can be used to discriminate approved drugs from experimental drugs with high accuracy. This discriminantion will help in developing a tool for searching the drugable molecules out of huge chemical spaces.

Statistical analysis of secondary structure of peptides

Kumardeep Chaudhary, Harinder Singh & Gajendra P. S. Raghava

Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, India.

Protein structure prediction is of utmost importance in the field of structural biology. One of the renowned protein secondary structure prediction methods is Chou-Fasman method which utilizes information from PDB data and predicts structure of protein on the basis of relative frequencies of amino acids present in alpha helix, beta sheet and turns regions. Though it has been shown in literature that methods developed so far for protein structure prediction fail in case of peptide structure prediction, till date there is no method which addresses the problem of structure prediction of peptides in particular.

In recent years, a number of anticancer, anti bacterial and antiviral peptides have been found. As peptides are very short in length, a single modification can lead to significant change in peptide structure. Since present methods developed for protein structure prediction are inefficient for peptide structure prediction, there is a need to develop novel methods for peptide structure prediction. PEPstr and PEPFOLD are among top performing methods for peptide structure prediction. Both these methods use data derived from proteins and not from peptides. Thus, it is imperative to develop a new algorithm based on peptides derived data.

In this study, we have analyzed secondary structure of peptides and observed that most of residues in peptides fall in irregular secondary structure region (*e.g.* 61% of residues were found in coil region). It was also observed that only 8% of residues form beta sheets in peptides. These observations demarcate secondary structure composition of peptides from that of proteins. The helical nature for methionine and threonine was found more in proteins as compared to peptides. Cysteine and valine were having remarkably high content of beta sheet structure in peptides than in proteins. Beta sheet structure was preferred for leucine in case of proteins as compared to peptides. Aspartic acid and asparagine were preferred in coil region when present in proteins as compared to peptides. It was observed that amino acid propensities are significantly different in peptides and proteins. These observations instigate to develop separate algorithm for prediction of secondary structure in peptides. In summary, we analyzed secondary structure of peptides.

Applications of Open Source Workflow Management Tools in Drug Discovery

Nilkanth Patel,[a] KhaledTumbi,[a] NaeemAttari,[a] Abhay T. Sangamwar*,[a]

[a] Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), India.

Workflow management tools are widely used to integrate various databases, web servers, software and services to extract desired outcome by transfer of information between different fields. As the huge amount of 'observational' and 'experimental' data gathered from enormous sources, it is very tedious to evaluate or process every dataset manually or by using conventional methods. In the past decade, a steep upsurge was observed in utilization of workflow management tools in various fields of drug discovery and development. The various workflow management tools are available for the drug discovery process, which includes Pipeline Pilot, VIBE, Biolib, etc. from the commercial domain; and KNIME, TAVERNA, GRID, Kepler, etc. from public domain. These tools allow users to construct a protocol for processing and analyzing the complex data. The protocols can be defined, managed and executed using a variety of software, which involved metadata as input, transformation and output in several phases. In Present study, the KNIME (Konstanz Information Miner) workflow tool has been analyzed for the optimum use. KNIME workflow with its Schrodinger extensions has been used for molecular docking analysis on diverse protein structures. These extensions provide nodes for each step in molecular docking procedure, which can be connected together to final results. The results were found to be same as that of results from tedious manual procedures. It is worthwhile to note manual procedures may lead to errors and also time consuming. The automation significantly reduces time, manual interferences and thus eliminates various sources for errors. However, we should not completely rely on an automated workflow system due to diversity and exception in different metadata, and scientific workflows are still in emerging phase and much is required to increase their reliability for rational drug design and development.

Computer Aided Drug Design (Cadd) Approach to Design Novel and Potent Pknb Inhibitors as Anti-Tubercular Agents

R.Venkata Krishnan[a], Abhay .T. Sangamwar*[,a]

[a]Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), Punjab, India.

Tuberculosis remains one of the most devastating diseases in the world, with more than 13 million people suffering from an active tuberculosis infection and 1.8 million resulting deaths in 2008 alone. The emergence of multi-drug resistant strains has highlighted the need for new drugs to treat tuberculosis. PknB is an essential serine/threonine kinase of Mycobacterium tuberculosis, which plays an important role in a number of signalling pathways involved in cell division and metabolism. The recent finding has demonstrated that the Ser/Threonine protein kinase (PknB) is essential for sustaining Mycobacterial growth and support the development of protein kinase inhibitors as new potential anti-tuberculosis drugs. In the last decade, it has been reported that inhibition of PknB serine/threonine kinase results in inhibiting evolutionarily-conserved steps in central metabolic processes. In the present study, we have employed structure based pharmacophore and *de novo* design approach to design novel and potent inhibitors for PknB receptor. Docking studies were performed to explore binding affinity and hydrogen bond interaction between the ligand and active site of PknB receptor. Designed molecules were screened for ADMET and Lipinski's rule of five. Finally 10 different molecules were obtained as potential molecules that could be used as PknB inhibitors.

Bioinformatics Profiling of Mtb Esxg Protein in Host

Abid Qureshi, Nishant Thakur & Manoj Kumar

Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, India.

Analysis of the *Mycobacterium tuberculosis* (MTB) genomes have identified 4,013 protein genes out of which 52 are believed to be secretory, including members of the ESAT-6 secretion system (esx family). It appears that the various esx proteins play different roles in virulence and have been linked to granuloma formation, cell-to-cell spread of the mycobacteria and escape from arrested phagosomes etc. Recently it was shown that esxG, a putative member of the esx family, is required for optimal growth of MTB and has been associated with essential processes such as iron and zinc acquisition. In this study, we have used motif and other search methods including post-translational modification (PTM) events in the host to predict salient roles of this important protein in the host environment. Amino acid motifs play vital roles in carrying out most signaling aspects of the cell, and also control protein degradation and its location. Scanning signaling motifs from online databases MOTIF Search, Mega Motif base, Motif Scan, TOPS etc. for the host proteins and using phi blast we mapped some putative stretches with predicted function on the esxG sequence. From these predictions it is apparent that this protein may be interfering in the host cell cycle, apoptosis, interaction/binding with adaptor proteins involved in phosphorylation and other PTM events which are indispensable for the pathogen and hence it may be a novel drug target for MTB.

Svm based Prediction of *O-GlcNAc* sites in proteins

Nishant Thakur, Abid Qureshi & Manoj Kumar*.

Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, India.

O-GlcNAcylation is a post-translational modification (PTM), analogous to phosphorylation that regulates the activity, the stability or the subcellular localisation of target proteins, including kinases, phosphatases, transcription factors, and cytoskeletal proteins. In this type of PTM O-linked N-acetylglucosamine (O-GlcNAc) transferase catalyzes the addition of a single N-acetylglucosamine in O-glycosidic linkage to serine or threonine residues of intracellular proteins. O-GlcNAcylation plays critical role in many cellular processes and chronic diseases such as diabetes, neurodegeneration and cancer. Identification of new *O-GlcNAc* sites will help in finding new drug targets for various diseases. Importance of O-GlcNAcylation actuated us to develop O-GlcNAc site prediction algorithm using SVM. Training Dataset was collected from updated version of Phosphositeplus database with total 612 O-GlacNAc sites. We used various sequence features like amino acid composition, dipeptide composition, proteins disorderness, PSSM, binary profile and we achieved overall accuracy 74% during five fold cross validation and AUC 0.78.

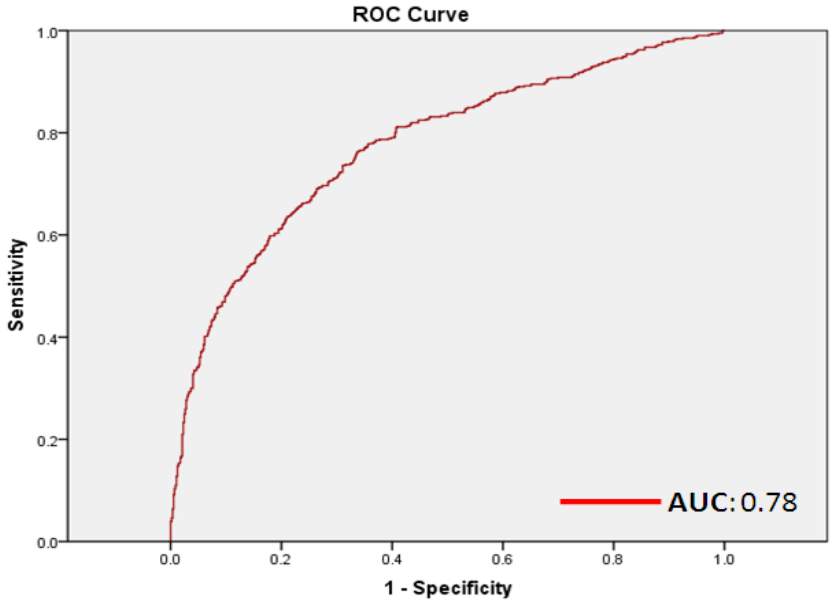


Figure. 1: Area under curve for O-GlcNAc algorithm.

Structure Based *De Novo* Design Of Ispd Inhibitors As Anti-Tubercular Agents

Rohith Anand Varikoti,[a] Gaurao V. Dhoke,[a] Abhay .T. Sangamwar*,[a]

[a]Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), India.

Tuberculosis (TB) remains one of the leading infectious diseases to humans infected with Mycobacterium tuberculosis. According to the WHO, there are 8.8 million new TB cases estimated per annum. The total number of new cases of TB is still growing slowly. In addition to this, 3.2% of the new cases of TB per annum show multidrug-resistance to isoniazid and rifampin, the two first-line drugs. In this perspective, isoprenoid synthesis is being studied and identified as a novel drug target. In Mycobacterium tuberculosis, isoprenoids plays an important role, which utilizes the methylerythritol phosphate (MEP) pathway for the biosynthesis of isopentenyl diphosphate and its isomer, dimethylallyl diphosphate, precursors of all isoprenoid compounds. As this pathway is absent in humans and causes disruption of the responsible genes in Escherichia coli, it emerged as a new target for TB. In the present study, we have employed LigBuilder and LUDI to design potent inhibitors for IspD receptor. To explore binding affinity and hydrogen bond interaction between the ligand and active site of IspD receptor, docking studies were performed. ADMET and Lipinski's rule of five filters were used to screen designed molecules. Finally, ten compounds were selected and subsequently submitted for the *in vitro* studies as IspD inhibitors.

Minimal Mutations that Define Lewis Y Reactive Antibodies

Somdutta Saha[1], Anastas Pashov[3], Ramachandran Murali[2], Thomas Kieber-Emmons[3]

[1]Bioinformatics Graduate Program, University of Arkansas at Little Rock/University of Arkansas,[2]Department of Biological Sciences, Cedars-Sinai Medical Center, Los Angeles, CA. [3]Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Arkansas, USA.

High-resolution structures reveal how a germline antibody can recognize a range of clinically relevant carbohydrate epitopes.Such analysis suggest that the expansion from a germ line antibody recognizing a carbohydrate epitopeinvolves linking the carbohydrate binding site with a subset of complementarity-determining regions (CDRs) of limited flexibility positioned to recognize the remainder of the epitopes. Here we show that antibodies reactive with the neolactoseries antigen Lewis Y (LeY) follow this strategy. Analyses of four murine and one humanized anti-Lewis Y (LeY) antibodies were performed to define the minimal mutations that can dramatically expand the reactive repertoire for the LeY antigen. Molecular dynamics calculations of these antibody-LeY complexes suggest that for the LeY reactive antibodies BR55-2, B3, 1S3K, mu3S193, the single germline precursor is sufficient to account for polyspecific recognition of the LeY determinate whereas hypermutationsaccounted for the fine specificity of the epitopes displayed on the LeY sugar determinate. Structural studies revealed that more often substitution mutations leading to antibody diversity occur in the loop regions of the CDRs than in the framework regions, which more predominantly have silent mutations. The optimization of the binding by somatic mutations fixes particular binding conformations formed by a limited flexible variable region of the precursor antibody attained upon antigen binding which directly impacts antigen recognition. The use of a predominate germline sequence to contact antigen suggests that selection is occurring as part of the germline diversity and not somatic hypermutations within the antigen combining site to increase antigen affinity. This observation emphasizes the fine specificity of these antibodies despite recognizing the same tumor associated antigen. This contributes to polyreactivity while CDR3 is reserved for fine tuning of epitope recognition of a particular determinate.

Classification of Kinases and Their Inhibitors Based On Their Interaction Using Clustering

Jagat Singh Chauhan & Gajendra P.S. Raghava*

Bioinformatics Centre, CSIR – Institute of Microbial Technology, Chandigarh, India.

The human protein kinases are the largest known families of human proteome it contains >500 genes. Protein kinases play vital role in signal transduction through phosphorylation mechanism. They catalyse the transfer of phosphate from ATP to a hydroxyl group of Serine, Threonine or Tyrosine of target proteins. Any deregulation in kinase causes various diseases such as cancer, diabetes, inflammation, cardiovascular disease and neurological disorders so they are emerging class of drug target in drug discovery processes. Imatinib, Nilotinib and Dasatinib drug were designed based on ATP binding sites of kinase target (Manley et al 2002). To date, 11 kinase inhibitors are approved by FDA as cancer treatments and 80 kinase inhibitors are in clinical trial. The protein kinases can be inhibited by compounds competing with ATP to ATP binding pocket with gatekeeper residues. To address this significant issue, computational techniques would be helpful to support experimental efforts directed toward designing of selective kinase inhibitors.

In this study we used a set of 38 kinase inhibitors and their experimentally validated dissociation constant (Kd) with 317 kinases were extracted from literature (Fabian 2009). Finally we selected 38 kinase target with more than 15 inhibitors have significant bioactivity values. In order to understand selectivity and specificity of kinases, we have first classified kinase inhibitors based on their inhibition efficacy of different kinases by using clustering software CLUTO. Here we found that certain inhibitor, for example EKB-569 and Roscovitine and similiarly SB-431542 and Staurosporine show highly similiarity. Further, we also classify different kinases based on their inhibition from various chemicals, where any single kinase is inhibited by same class of chemical inhibitor, so here we clusterd the highly similar kinase in separate group. Finally we have observed that KIT and RET kinase, similarly EPHA6 and TNIK kinase shows highly similiarty. Finally we clustered the highly correlated kinase in same group, so we can designed specific inhibitor against specific kinase of highly similar kinase group, thus this study will be helpful in selecting the highly efficient and specific kinase inhibitors against kinase.

Isolation and characterization of antimycobacterial compounds from aloe vera (l.) Burm. F. Leaf

Pratap Chandran, R., Dibu Divakaran, Kannan, M., Sujeesh, B. V., Lekshmy, R. Nair., Aswathi Chandran.

Department of Biotechnology and Research, K.V.M. College of Engineering and Information Technology, India.

Emergence of multi-drug resistant (MDR) and extensively-drug resistant (XDR) strains of *Mycobacterium tuberculosis* has further complicated the problem of tuberculosis (TB) control. Medicinal plants offer a hope for developing alternate medicines for the treatment of TB. The aim of the present study was to evaluate antimycobacterial activity of extracts of *Aloe vera* using the different solvents. Accordingly only Chloroform Hot, Methanol Hot and Methanol Cold extracts of *Aloe vera* plant exhibited effective antimycobacterial activity. GC MS analysis of these three extracts has been carried out to detect the volatile compounds and revealed many compounds in each extracts. In the present study only methanolic cold extract has been subjected further for the isolation of bioactive compound using Silica gel column chromatography (60-120 mesh) and initially yielded about 25 fractions. Among the 25 fractions only fraction 5 exhibited highest antimycobacterial activity of about 40mm zone formation. Further isolation of fraction 5 using Silica gel (230-400 mesh) yielded 2 sub fractions (Fr_m.1 and Fr_m.2) of which the first fraction (Fr_m.1) showed an activity with a zone of 25mm. Purity of the final bioactive fraction yielded about 62% in HPLC. Extensively purification using preparative HPLC and characterization using NMR spectroscopy as single bioactive compound is in process. Thus the present study gives an insight to take *Aloe vera* for further bioactive compound characterization against *Mycobacterium tuberculosis* in the battle of identifying a novel drug to combat the mortality due to extensively drug resistant Tuberculosis.

Lipophilic Proteins Analysis Of Aminoglycoside Resistant Isolates of Mycobacterium *Tuberculosis*

Divakar Sharma, Manjulata, Bhavnesh Kumar, Prashant Sharma, Krishnamurthy Venkatesan & Deepa Bisht.

Department of Biochemistry, National JALMA Institute for Leprosy and Other Mycobacterial Diseases (ICMR), Tajganj, Agra, India.

Tuberculosis is a great threat to mankind since ancient times. This gets further worsen with the development of drug resistance and co-infection with HIV which is the main problem to deal with. Globally, 1.7 million deaths occurred due to this disease. BCG vaccination in uninfected people and chemotherapy using drug combinations is the only existing tool against tuberculosis. Aminoglycosides are commonly used drugs in the multi-drug resistant tuberculosis treatment. They inhibit protein synthesis in susceptible bacteria by interacting with several steps of translation. Several explanations have been put forward to explain the mechanism of aminoglycoside resistance but still our knowledge is incomplete. As proteins are the main functional moiety in the cell and they manifest most of the biological processes, these are the attractive targets for developing drugs, immunodiagnostics or therapeutics. The aim of our study was to compare the protein profiles of lipophilic proteins from *Mycobacterium tuberculosis* isolates susceptible and resistant to aminoglycosides by two-dimensional gel electrophoresis. On comparing two dimensional gels some proteins were found to be upregulated in resistant isolates and were identified by MALDI-TOF mass spectrometry and bioinformatic tools. These proteins are likely to be used as markers of aminoglycosides resistance and/or might play an important role in developing better drugs, diagnostics and vaccines.

Numerical Simulation Of The Internal Vibrations Of Oh Group In Amino-Salicylic Acids

Ratnesh Das * & Y. P. Singh

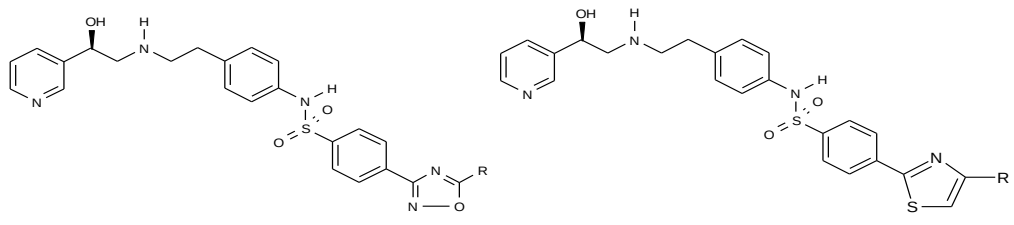
Department of Chemistry, Dr. Harisingh. Gour University, M.P. India.

Our present work reports the IR spectra of amino substituted salicyclic Acids recorded by FTIR spectrometer and also simulated theoretically. The simulation were performed using GF matrixand AM1, PM3, DFT method. In this work following steps were taken: optimizing the geometry,computing the IR spectra and comparing it with experimental spectra. assuming Cs point symmetry, vibrational assignments for the observed frequencies have been proposed. The spectra exhibit distinct features originating from low frequency vibrational modes caused by intra-molecular motion.

Comfa and comsia investigations on substituted thiazole/ alkyl oxadiazole benzenesulfonamide as beta-3 adrenergic receptor agonist ligands for anti-obesity drug development

Sushil K. Kashaw[1], Vikas Mishra[1], Varsha Kashaw[2] and N.K. Jain[1]
[1]Division of Medicinal and Computational Chemistry, Deptt. of Pharm. Sci., Dr. H.S. Gour University, M.P. India [2]Sagar Institute of Pharmaceutical Science, M.P. India.

The last 25 years have seen a great increase in the incidence of obesity, both in the Western world and in developing third world countries. Doubts have been raised on the long-term sustainability of this weight loss. β_3-adrenoceptors agonist capable of increasing metabolic rates by selective activation of these receptors is considered as a potentially effective approach for the treatment of obesity and diabetes in recent years. The main problem with β_3-adrenoceptor agonists is their cross-reactivity at the β_1 and β_2-receptors [1]. Comparative molecular field analysis (CoMFA) has been a useful technique in designing important 3-dimensional properties associated with the optimum binding of ligands to a binding site. CoMFA samples the differences in steric and electrostatic field surrounding a set of compounds and maps this biological activity. Successful CoMFA, Adv CoMFA and COMSIA models have been generated for the mentioned compounds. Common structure based fitting, Rigid body field fit alignment using the steric and electrostatic fields and common feature hypothesis was used for alignment. Two structural series of Human β_3-adrenergic receptor (ARs) agonists were used in the studies. The cross validated r^2 (q^2) was highly significant for all the three alignments but common structure based alignment was the best among these. The best 3D-QSAR model using this alignment were CoMFA (q^2=0.66), Adv CoMFA (q^2=0.569) and CoMSIA (q^2 =0.667). These models showed good correlation between the observed and predicted activity of the training set and test set (r=0.77) and so the studies may be helpful in designing new chemical entities for β_3-adrenoceptor agonists



Identification of Novel Inhibitors for Lysine/Dap Pathway in *Mycobacterium Tuberculosis*

Priyanka Shrivastava[1], Rahul Shubhra Mandal[2], Smriti Sharma[2], Balasubramanian Gopal[3], OSDD Consortium, Srinivasan Ramachandran

[1,2] Functional Genomics Unit, CSIR - Institute of Genomics & Integrative Biology, , Delhi, India

[3] Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India.

The Meso-diaminopimelic acid (DAP), a key component of highly immunogenic mycobacterial cell wall, has been implicated as a potential virulence factor. The DAP pathway could serve as a target for design of new antimycobacterial agents and potential drug targets.

Computational studies have screened various drug like inhibitors against DAP pathway. In order to carry out experimental validation, a set of compounds including pyruvate analogs and pyruvate like compounds were obtained from chemists. The *DapA* and *DapB* genes were cloned, expressed, and purified using Ni^{2+} - NTA chromatography. The identities of the purified proteins were confirmed using Western Blot and MALDI analysis. A coupled assay included DapA and DapB in a reaction using NADPH, which reduces the product of DapA by DapB. The consumption of NADPH in the reaction was measured at 334nm. The compounds were tested for their inhibition. We used alpha keto pimelic acid, a known inhibitor for the assay as reference to examine the inhibition effect of test compounds.

We have tested six compounds, PUB 19751056: 2-(3-nitrophenyl)-2-oxoacetate, PUB 15288093: 2-(2-nitrophenyl)-2-oxoacetic acid, PUB 4451056: 2-phosphonato acetate, PUB 3092: 2-methyl-3,5-dinitrobenzamide, PUB 70269: (E)-3-(5-nitrofuran-2-yl) prop-2-enoic acid. The compounds were dissolved in water except PUB 15288093 and PUB 3092, which were dissolved in DMSO. One of the compounds, methanol soluble EPMC (ethyl-p-methoxycinnamate), was obtained from Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram. We tested these compounds with a gradient of inhibitor and solvent concentrations ranging from 100µM to 200µM. EPMC, PUB 19751056, PUB 15288093 and PUB 3092 were also tested at higher concentrations. PUB 70269 and PUB 4451056 showed significant inhibition at a concentration of 200µM as compared to other compounds. All measurements were performed in duplicate or triplicate. The results were plotted as non linear regression fit. These results hold hope for developing new inhibitors for the DAP pathway.

Understanding Mnanog-Dna Interactions Using Structure Based Approaches

Sangita Kachhap & Balvinder Singh

Bioinformatics center, CSIR – Institute of Microbial Technology, Chandigarh, India.

Nanog is a transcription factor, responsible for embryonic cell pluripotency and binds to DNA through homeodomain. It has been shown that gene knockdown of Nanog promotes cell differentiation. Further, Nanog overexpression in human embryonic stem cell enables their propagation for multiple passages during which cell remains pluripotent. As it has important role on mammalian development, it is necessary to understand its molecular function and how it recognizes its cognate DNA sequence. However, its structure with cognate DNA sequence is not available. It will be interesting to model its interactions with cognate DNA and to understand its selective promoter sequence recognition. EMSA study shows that mNanog homeodomain preferably binds to Tcf3 promoter having sequence TAATGG. It has also been shown to binds DNA sequences such as 5’ TAATTG 3’, 5’ TAATTT 3’ and 5’ TAATGT 3’ in this EMSA study. Out of these four sequences, 5’ TAATTG 3’and 5’ TAATGG 3’ are natural DNA binding site of homeodomain. In this study, interaction between mNanog and DNA has been predicted and modeled based on different structural approaches involving homology and docking techniques. An analysis of different type and kind of interactions between mNanog and different promoter sequences has been carried out. Structural and functional implications shall be presented.

Targeting the agent responsible for invasion: an approach towards cure of malaria.

Pramod Kumar, Nimish Gopal, Abin Ghosh, Pravindra Kumar

Department of Biotechnology, Indian Institute of Technology Roorkee, Uttranchal, India.

Every day we can notice the increasing number of deaths due to the killer disease malaria. Further the challenge has become grim by the reports of the new drug resistant forms these days. It has been reported that 95% of the malaria is caused by *Plasmodium falciparum.* Out of the several approaches to cure the disease, targeting the invasion process of the *P. falciparum* to RBC is of immense importance. For intruding the mammalian erythrocyte the proteolytic enzymes are employed. “Subtilisin” is one the most important enzyme used by *P. falciparum* for initiating the invasion process of RBC surface. Further the subtilisin homologue in malaria is highly specific for its substrate recognition in comparison to the lower class of eukaryotes which reinforces the approach to become successful. In the present study we have used an approach for targeting this subtilisin with a modified drug to act over it with the flash activity of fast action and fast disappearance so the peak phase of the invasion and spreading of the disease can be targeted. We have used the Poly vinly pyrolidone (PVP) to form the complex with known drug Darunavir. We have used bioinformatics approach to find the way of interaction and the binding ability of the compound to the targeted subtilisin. A clear interacting pocket was obtained where the drug docked successfully with the reasonable interacting distance where in the key residues were involved in the activity of the enzyme and having evolution conserved profile, moreover the energy profile of the modelled structure was within the valid parameter range. Here a clear pocket having active site was characterized in complex with the drug compound, where the apo form is having the open status that is transformed to the close loop structure upon binding of the ligand confirms the hypothesis to using a PVP based derived complex. Further the ability of the PVP-complex to make the interaction profile better is characterized by the interaction with polar residues lying in the catalytic pocket. Here the interaction is shifted to the favourable confirmation which is assessed by the docking software Autodock, Hex and Maestro. On the basis of above analysis and further validation with biochemical experiments this approach can become a promising way to fight the malaria.

Computational Identification Of Novel Adhesins Of *Mycobacterium Tuberculosis* H37rv And Their Experimental Characterization

Sanjiv Kumar, Rahul S Mandal, Shahila Parween, Pradip Nahar and Srinivasan Ramachandran*

FGU, CSIR-Institute of Genomics and Integrative Biology, Delhi, India.

Pathogenic bacteria express adhesins on their surface that causes membrane perturbation during initial interaction that mediate infection of eukaryotic host. Aim of the present study was to identify novel adhesins of *M. tuberculosis.* The proteins encoded in genome of *M. tuberculosis* were screened for identifying potential surface adhesins using SPAAN, along with additional criteria of non-homology with human proteins, low molecular weight (<100 kDa) and surface localization of proteins.

The protein Rv3717 has predicted role in cell adhesion and is a cell wall hydrolase. *Rv3717* is a 726 bp gene encoding a 24 kDa protein. The protein has signal peptide with cleavage site between amino acids 24 and 25. CDD database search predicted NAM-LAA-amidase domain. The gene was cloned in pET28a, expressed in *E. coli* C41 (DE3) and protein was purified using Ni-NTA chromatography. To examine cell wall hydrolase activity, zymographs were performed with 1% heat killed bacterial cells. Positive reaction was observed for various bacterial species (i.e. *L. acidophilus, E. coli, B. thuringiensis, B. avium, E. aerogenes* etc.). Protein was also able to hydrolyze heat killed *M. smegmatis*. These results show that Rv3717 is a novel cell wall hydrolase in addition to Rv3915 in *M. tuberculosis*. Preliminary results also show that it could be a probable adhesin with binding affinity to human ECM proteins.

Homology model of Rv3717 was prepared using CPHmodels 3.0. Template used was 1JWQ (N-acetylmuramoyl-L-alanine amidase) from *B. polymyxa*. Sequence identity between template and target structure was 23%. Stereo-chemical quality of model was examined by PROCHECK. Total of 86.7% residues fell in favored region of Ramachandran plot. Verify_3D and Errat scores were 90.50% and 98.54% respectively. Final model was verified using SAVS serve. Alignment of aminoacid sequence of Rv3717 with 1JWQ and 3NE8 (NAM-LAA amidase) showed conserved zinc binding and active domain.

Uridine Monophosphate Kinase as Novel Target For Tuberculosis: From Target To Lead Identification

Vaibhav Jain,[a] Akanksha Arvind,[a] Parameswaran Saravanan[a], P. V. Bharatam*[b] and C. Gopi Mohan*[a]

[a]Department of Pharmacoinformatics, [b]Department of Medicinal Chemistry National Institute of Pharmaceutical Education and Research, S.A.S. Nagar Punjab, India.

ABSTRACT

Mycobacterium tuberculosis (Mtb) is a causative agent of tuberculosis (TB) disease, which has affected approximately 2 billion people worldwide. Due to the emergence of resistance towards the existing drugs, discovery of new anti-TB drugs is an important global healthcare challenge. To address this problem, there is an urgent need to identify new drug targets in Mtb. In the present study, subtractive genomics approach has been employed for the identification of new drug targets against TB. Screening the Mtb proteome using Database of Essential Genes (DEG) and human proteome resulted in the identification of 60 key proteins which have no eukaryotic counterparts. Critical analysis of these proteins using Kyoto Encyclopedia of Genes and Genomes(KEGG) metabolic pathways database revealed uridine monophosphate kinase (UMPK) enzyme as potential drug target for developing novel anti-TB drugs. Homology model of Mtb-UMPK was constructed for the first time on the basis of the crystal structure of *E. Coli*-UMPK, in order to understand its structure-function relationships, and which would in turn facilitate to perform structure-based inhibitor design. Furthermore, structural similarity search was carried out using physiological inhibitor UTP of Mtb-UMPK to virtually screen ZINC database. Retrieved hits were further screened by implementing several filters like ADME and toxicity followed by molecular docking. Finally, on the basis of Glide docking score and the mode of binding, 6 putative leads were identified as potential inhibitors of this enzyme. These novel inhibitors are suggested for *in vitro/in vivo* analysis, and which can potentially act as future drugs for the treatment of TB.

Deciphering key amino acid residues for trimerization of *m. tuberculosis* pii

Sriyans Jain[1], Rahul Shubhra Mandal[1], Swadha Anand[2], Debasisa Mohanty[2], Srinivasan Ramachandran[1]

[1]CSIR - Institute of Genomics and Integrative Biology, New Delhi, India.

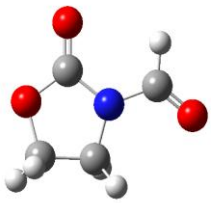
[2]National Institute of Immunology, New Delhi, India

The signal transduction protein PII plays important role in nitrogen assimilation. The homologues of this protein family are present in a wide variety of organisms that include archaea, bacteria and eukaryotes. This protein acts as sensor for intracellular nitrogen through the ratio of nitrogen to carbon represented by glutamine and 2-oxoglutarate respectively. Many species have two to more homologues of PII, while *M. tuberculosis* has only one copy of the gene Rv2919c coding for PII. In *M. tuberculosis*, PII (MtbPII) has been shown to be essential for survival. Its role in nitrogen regulation, essentiality and absence of its homologues in humans makes it an attractive candidate gene for drug targeting. The PII family proteins are known to form a homotrimer. In this study, our goal was to decipher amino acid residues important for its trimer formation. The MtbPII crystal structure (PDB ID – 3BZQ) was analyzed using protein interaction calculator (PIC) and PDBePISA (Protein Interfaces, Surfaces and Assemblies) servers to predict the interacting residues between the subunits. These were further compared to examine their conservation across other species. Using 47 PII orthologous proteins from various species we obtained 3 target pairs (K2-D97, R60-E62, and D71-R107) of which two pairs K2-D97 and R60-E62 were electrostatically conserved across orthologues of MtbPII. We generated a series of single and double mutants using site directed mutagenesis by modifying the residues forming the target pairs and the mutant proteins were purified through Ni^{2+} - NTA chromatography. The homotrimer formation potential of the mutated PII proteins was tested using the glutaraldehyde cross-linking assay. It was observed that mutation of R107 alone or in combination with other residues had no effect on homotrimer formation. The mutation of amino acids R60 and D97 singly had no effect on the protein's trimerization. However, a double mutant with modified R60 and D97 proved refractory to solubilization and purification. These results show that amino acid interactions are mutually compensatory whereas drastic changes with multiple mutations have a biologically unacceptable protein. It is evident that among the three target pairs, two are important for maintaining stable inter-subunit interactions and at least one of them is critical. The residue pair D71-R107 may supplement the critical interactions energetically.

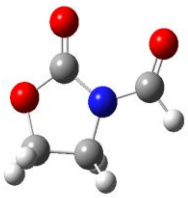
Evaluation of Proton Affinities of Oxazolidin-2-One Derivatives: Components Of Antibiotics And Antibacterial Agents

Ritika Sharma, Damanjit Kaur Department of Chemistry, Guru Nanak Dev University, Amritsar, India.

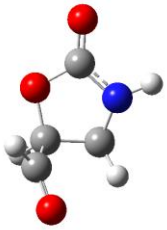
Oxazolidinones are component of a novel class of totally synthetic antimicrobial agents against multidrug resistant gram-positive bacteria and some gram-negative bacteria for example Linezolid is the first antibiotic containing oxazolidinone ring that exhibits an antimicrobial spectrum that encompasses a broad range of susceptible and multidrug resistant gram-positive coci. Mainly the oxazolidinone derivatives substituted at C5 and N4 positions have applications as antibiotics. The antibacterial effect of oxazolidinones is by working as protein synthesis inhibitors, targeting an early step involving the binding of N-formylmethionyl-tRNA to the ribosome. They are the last generation of antibiotics used against gram-positive pathogens, including superbugs such as Methicillin-resistant Staphylococcus aureus. Cycloserine (*4-amino-1,2-oxazolidin-3-one*) is a second line drug against Tuberculosis. Several reactions involving binding of oxazolidinone drugs to the target molecule, involve proton or cation transfers, the understanding of which, at the molecular level, passes through the knowledge of the precise site of protonation, changes of conformation upon protonation, and protonation thermochemistry. Therefore N, C5-oxazolidinones have been molecules of interest for us for calculating proton affinity values. In this research, study of geometrical isomerism, conformational flexibility, relative stabilities of oxazolidin-2-one (OXA), N-/C5-formyl-oxazolidinones, N-/C5-amino-oxazolidinones and their protonated forms have been performed by employing *ab* initio and DFT calculations. The proton affinities (PAs) of the above said molecules have been determined at their potential sites. The preferred protonation site and magnitude of PAs have been analyzed in terms of relative stability difference between the neutral and protonated species, electron delocalization, NPA charges, various electrc



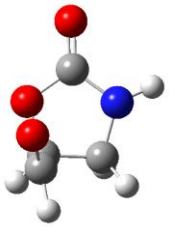
N4-CHO(a)-OXA



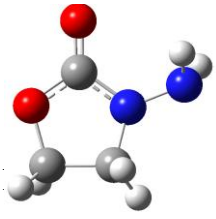
N4-CHO(b)-OXA



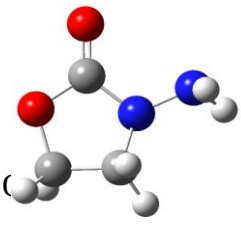
C5-CHO(a)-OXA



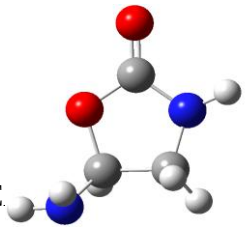
C5-CHO(b)-OXA



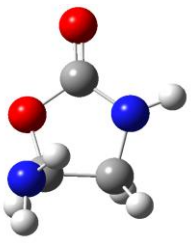
N4-NH_2(a)-OXA



N4-NH_2(b)-OXA



C5-NH_2(a)-OXA



C5-NH_2(b)-OXA

Benzylidene Derivatives to Design Potent Hiv-Integrase Inhibitors

VK Srivastav & M Tiwari

Shri GS Institute of Technology and Science, M.P. India.

Background: AIDS is the most devastating disease throughout the world. Multidrug resistance is the greatest challenge in treatment of AIDS. The absolute requirement of the integrase enzyme for HIV replication, and the fact that it has no host cell counterpart, makes it an attractive drug development target. The introduction of the integrase targeted drug *raltegravir* has validated this enzyme as a very promising HIV/AIDS therapeutic target.

Methods: In the present study a QSAR analysis was performed to correlate the calculated physicochemical properties of compounds with anti HIV-1Integrase activity and docking study was done to predict the binding mode pattern of the compounds in the binding site of integrase enzyme. Designed compounds were compared with Raltegravir and novel leads (under clinical trials) for their binding pattern to integrase enzyme by docking studies.

Results: The QSAR model predicted that parameters of the model showed good contribution for the anti-HIV integrase activity. The best QSAR Model was

pIC_{50} = [13.7809(± 6.3671)] +HOMO [1.3504(± 0.6900)] +BI [4.6464(± 2.4538)] +SC [3.9531(± 1.5949)]

n=18, r=0.8814, r^2=0.7770, r^2adj=0.7292, variance=0.0681, SEE=0.2610, F=16.2618, Q^2=0.6362, S_{PRESS}=0.3334, S_{DEP}=0.2940, r^2_{pred}=0.3125, Chance<0.001

The designed compounds showed better activity than the dataset compounds and possess good scoring function than standard compounds in docking studies. In addition to that these compounds also bind to the resistant residues present in the integrase protein.

Conclusion: QSAR and docking study of the Coumarin derivative proved that this lead has good potential against HIV. The binding pattern to the integrase protein as well as the resistant residues proved its potency against HIV.

Structural Minimization Study of Anatoxin-A: A Potent Neurotoxin

Naresh Kumar[1] and Vikas Pruthi[2]

* Department of Biotechnology, IIT Roorkee, Roorkee, India.

Cyanotoxins are toxins produced by bacteria called cyanobacteria (blue-green algae). Cyanobacteria are found almost everywhere, but particularly in eutropic lakes, rivers, reservoirs and in the ocean where, under certain conditions, they reproduce exponentially to form blooms. Blooming cyanobacteria can produce cyanotoxins in such concentration that they poison and even kill animals and humans. Continuous increase in the number of cases of animal poisoning and human illness by cyanotoxins have altered worldwide attention on one of the most potential cyanotoxin called as Anatoxin-a. Anatoxin-a have been implicated in the development of loss of coordination, muscular fasciculations, convulsions and death by respiratory paralysis. It acts as analogues of acetylcholine and believes to cause the competitive inhibition of the nicotinic acetylcholine receptor. In this *in silico* study our aim is to get more insight on the molecular structure of Anatoxin-a. To characterize Anatoxin-a structure, we performed Density Functional Theory (DFT) calculations by Gaussian (computational chemistry software). We used the DFT employing B3LYP exchange correlation for the structure minimization. The geometry of the Anatoxin-a molecule were fully optimized at B3LYP/6-311G** level both in gas phase and solution phase (Water). We also calculate the chemical shifts for both Proton (^{1}H) and Carbon (^{13}C). Gaussian predicts the energies, molecular structures, and vibrational frequencies of molecular systems along with the numerous molecular properties derived from basic computation types. Gaussian quickly became a popular and widely-used electronic structure program.

Development of Classification and Regression Based Qsar Models to Predict Rodent Carcinogenic Potency Using Oral Slope Factor

Supratik Kar[1], Omar Deeb[2] and Kunal Roy[1*]

[1] Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India.

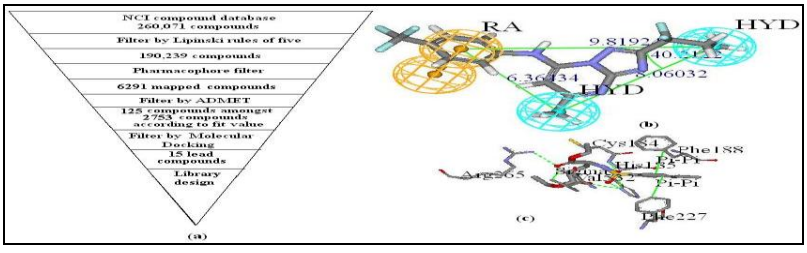
Carcinogenicity is among the toxicological endpoints posing the highest concern for human health. Oral slope factors (OSFs) are used to estimate quantitatively the carcinogenic potency or the risk associated with exposure to the chemical by oral route. Regulatory agencies in food and drug administration and environmental protection are employing quantitative structure activity-relationship (QSAR) models to fill the data gaps related with properties of chemicals affecting the environment and human health. In this background, we have developed quantitative structure-carcinogenicity regression models for rodents based on the carcinogenic potential of 70 chemicals with wide diversity of molecular structures, spanning a large number of chemical classes and biological mechanisms. All the developed models have been assessed according to the OECD principles for the validation of QSAR models. We have also attempted to develop a carcinogenicity classification model based on linear discriminatory analysis (LDA). Our *in silico* studies make it possible to obtain a quantitative interpretation of the structural information of carcinogenicity along with identification of the discriminant functions between lower and higher carcinogenic compounds by LDA. Pharmacological distribution diagrams (PDDs) are used as a visualizing technique for discrimination. The current *in silico* methods have provided: a) some structural alerts to determine the potential carcinogenic modes of action via direct interaction with DNA and other macromolecules, b) it may be possible to design chemicals with lesser carcinogenicity to living ecosystem and proper risk assessment of those chemicals, c) in the absence of OSF toxicity data for a chemical, the developed model can be used to predict the carcinogenicity chemicals, and d) most importantly, the obtained results can be used as a starting point for regulatory decision making and risk assessment in future.

***In Silico* Design Of Dihydroorotate Dehydrogenase (Dhodh) Inhibitors As Potential Antimalarial Agents Using Pharmacophore Mapping, Molecular Docking And Virtual Screening Strategies**

Probir Kumar Ojha, Indrani Mitra, Supratik Kar, Rudra Narayan Das, Kunal Roy*

Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India.

According to the World Health Organization report, malaria secures the fifth position among the top ten causes of worldwide death. Dihydroorotate dehydrogenase (DHODH)] is a potential target for antimalarial drug development. In this work, we have tried to screen and design novel *Pf*DHODH inhibitors from NCI database employing pharmacophore modeling and molecular docking as the prime virtual screening approaches [Fig 1(a)]. 3D pharmacophore models were developed based on a series of triazolopyrimidine derivatives (130 compounds) exhibiting *Pf*DHODH inhibitory activity. The three feature (HYD, HYD, RA) pharmacophore [Fig. 1(b)] developed in the present work was selected based on the values of the correlation coefficient (0.813) and the cost functions (null cost = 609.696, total cost = 292.004, fixed cost=123.778 and configuration cost=11.412). External validation of the developed model was done based on qualitative assessment of the predictions of the test set molecules that were mapped to the developed pharmacophore. Docking studies revealed that the bound ligand DSM1 and the dataset compounds are docked in the same binding pocket of the enzyme showing specific interactions with Arg265, His185, Phe188, Phe227 and Val532 residues [Fig 1(c)]. Thus pharmacophore mapping and docking analysis provided useful insight about the necessary structural attributes of the molecules. Based on these results, the NCI database was screened for selection of hit compounds and further designing of focused library comprising of potent derivatives.



Development of an *In Silico* Predictive Model for Olfactory Threshold of a Series of Pyrazine Derivatives

Pallabi Pal, Indrani Mitra and Kunal Roy*

Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India.

Nowadays, much interest is driven towards the prediction of different organoleptic properties of molecules for their vivid use. Among the various chemical senses, odor constitutes a complex chemical sense by which human beings can identify different materials. So, it is essential to know the olfactory threshold (log T) of different chemicals so as to identify or alter this organoleptic property as per the requirement of the products. Recently, the *in silico* techniques are widely employed for such property prediction procedures which enable to screen the potent molecules selectively in a cost effective manner. In the present work, a quantitative structure property relationship (QSPR) model has been developed based on a series of pyrazine derivatives for their odor threshold. 30% of the compounds were selected as test set based on cluster analysis. Following computation of different categories of descriptors and application of different chemometric tools, a model developed using genetic function approximation (GFA) technique emerged as the best model which can be used for predicting the threshold values of pyrazine derivatives. After validating the models both internally and externally, the GFA-spline model showed encouraging predictive quality. From the deduced model, it can be inferred that increase in branching of the molecules associated with a decrease in their van der Waal's area decreases the odor threshold of the pyrazine derivatives. Again, decreased π electron density which in turn accounts for a decrease in the negative surface area of the molecules helps to reduce the odor threshold of the molecules. On the other hand, increase in hydrophobic surface area of the molecules will lower the odor threshold. Finally, the QSPR model may be used to predict the olfactory threshold values of pyrazine derivatives falling within the domain of applicability of the model.

Qspr with extended topochemical atom (eta) indices. Exploring effects of hydrophobicity, branching and electronic parameters on *log*-cmc values of anionic surfactants

Humayun Kabir and Kunal Roy

Department of Pharmaceutical Technology, Jadavpur University, Kolkata, INDIA.

In this work, quantitative structure-property relationship (QSPR) models have been developed to establish the relationship between the molecular structures and the critical micelle concentration (CMC) of 37 anionic surfactants using extended topochemical atom (ETA) indices along with computed hydrophobicity descriptors. The ETA models have also been compared to those developed with non-ETA topological descriptors along with hydrophobicity terms. The selection of the training and test compounds (n = 28 and 9 respectively) was done by using a principle component analysis (PCA) score plot. The best ETA model could outperform in statistical quality and predictive ability the models developed with non-ETA and combined set (ETA and non-ETA) of descriptors. In this study, it is observed that hydrophobicity plays a major role in the model development for CMC of anionic surfactants while the branching character and electronic nature of the surfactants are also important as evidenced from different ETA parameters.

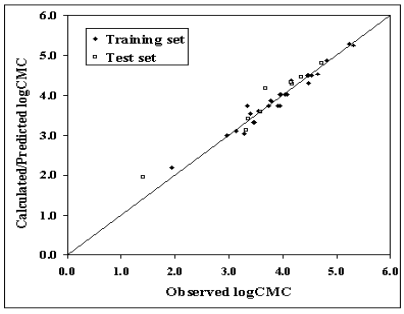


Figure 1: Scatter plot for observed vs calculated/predicted activity data of the training and test sets.

In-Silico Evalution of Ligand against Sur1 Receptor (Diabetes Mellitus)

*Rohit Sahu , Ankit Awasthi & Prof. D.K. Gupta

Center of Bioinformatics, University of Allahabad, Allahabad, India.

Diabetes Mellitus is a metabolic disorder characterized by the hypoglycaemia, glycosuria, hyperlipaemia, negative nitrogen balance & sometimes ketonaemia.

Type II Noninsulin-dependent diabetes mellitus (NIDDM), maturity onset diabetes mellitus:- Cause may be (a) Abnormality in glucose receptor of beta-cell so that they respond at high glucose concentration.(b) Reduced sensitivity of peripheral tissues to insulin, reduction in number of insulin receptor, 'down regulation of insulin receptors'.(c) Excess of hyperglycaemic hormone. Prepare a oral drugs an in-silico (Drug Designing) research has been worked-out which is based on some computational tools & on the principle of protein-protein & protein-ligands binding. Protein-protein and protein-ligand interactions are fundamental as many proteins mediate their biological function through these interactions. Many important applications follow directly from the identification of residues in the interfaces between protein-protein and protein-ligand interactions, such as drug design, protein mimetics engineering, elucidation of molecular pathways, and understanding of disease mechanisms. The identification of interface residues can also guide the docking process to build the structural model of protein-protein complexes.

This dissertation focuses on developing computational approaches for protein-ligand and protein-protein binding site prediction and applying these predictions to improve protein-ligand docking.

The Ligand could be developed with the help of some available software of Bioinformatics like ChemSketch of ACDLabs. The Ligand must follow the Lipinski Rule. After that a Docking procedure is performed with a specified Receptor of a Ligand prepared with the help of such software like Auto-Dock, Virtual Dock.

A result that obtained when compared with the result of prescribed drug is lower, than this proved to proposed a more chance of better Ligand against that particular Protein (SUR1-NIDDM).

Tumor homing peptides-Targetted Approach to Cancer Therapy: Analysis

Pallavi Kapoor, Harinder Singh, Kumardeep chaudhary, Rahul Kumar & Gajendra P. S. Raghava

Bioinformatics Centre-(CSIR), Institute of Microbial Technology, Chandigarh, India.

Homing peptides are those peptides which homes specifically to either normal tissues or pathological conditions after systemic delivery, thus tumor homing peptides are those peptides which homes specifically to the tumor tissues and microenviroment sorrounding the tumor. The high specificity of therapeutically active drugs to the tumor cells is one of the major challenges in drug development. Homing peptides or their mimetic can be used to improve selectivity and specificity.Thus these tumor homing peptides are providing new directions for diagnostic and therapeutic strategies.With such potential of these peptides we collected and analysed tumor homing peptides .These peptides were collected and compiled from published papers , patents and database. Tumor homing peptides have been found to target different types of tumors that include breast, lung, prostate, melanoma, colon, etc. Average amino acid composition of these peptides shows that certain types of residues (Cys, Arg, Gly, Leu and Ser) are more abundant in tumor homing peptides . Tumor homing peptides have significant higher average composition of Arg, Cys and Trp compared to SwissProt sequences, while residues Gly and Leu are abundant in both the cases.Cys has highest composition which is also indicated by peptide cyclic nature.We have also calculated the frequency of poly-Arg (RR, RRR and RRRR) in tumor homing peptides showing that high average composition of Arg in tumor homing peptides is not due to the presence of poly-Arg peptides/motifs and it may be due to the presence of some common motifs like RGD, NGR, etc. which contain Arg. It is also observed that most of the peptides have 7, 9, 10 and 12 residues.Secondary structure composition of each peptide and classified peptides based on secondary structure content shows that most of the peptides have no beta-strand or less than 20% beta-strand. Only few peptides have 40-60% helix content. In case of coil, a large number of peptides have most of the residues forming coil. This is just a primary analysis, and a detailed analysis is required in order to understand the relation between structure and tumor homing capability of the peptides. Collection of tumor homing peptides and analysis will be very helpful to researchers to design novel peptides, which can be used further for developing novel anti-cancer drugs.

Linear Discriminant Analysis for Skin Sensitization Of Diverse Organic Chemicals

Ashis Nandy, Supratik Kar, Rudra Narayan Das and Kunal Roy*

Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India.

Topical pharmaceutical products such as cream, lotion, shampoo may cause sensitization to the normal skin cell resulting in irritation, swelling and redness of the skin and sometimes may even produce aging and carcinoma. Skin sensitization is one of the emerging toxicological endpoints posing the latest concern for human health. Stimulation index and murine local lymph node assay (LLNA) are used to estimate quantitatively the skin sensitization potential and also to classify chemicals as sensitizer or non sensitizer. To overcome these time consuming and expensive approaches, development of *in silico* predictive models have gained large attention over the last few decades. *In silico* methods that have been extensively used to predict the skin sensitivity of new compounds provide an important tool for initial screening of new chemicals. In this background, we have developed LDA (Linear Discriminant Analysis) model for the skin sensitization potential data based on the LLNA of 147 chemicals with wide diversity of molecular structures (1). The developed LDA model is rigorously validated. Our *in silico* studies make it possible to identify the major discriminatory features between lower and higher skin sensitizing compounds. Our developed model and contribution plot (Fig. 1) suggest that skin sensitization of chemicals decreases along with the increasing number of the branching index. On the contrary, high number of rotatable bonds or molecular flexibility positively contributes towards high skin sensitization. Number of sulfonate fragments (thio-/dithio-), number of triple bonds and number of nitrogen atoms in chemicals are also identified major discriminatory features for high and low skin sensitizing chemicals. One may consider all these points to reduce the skin sensitization potential of the newly designed molecules.

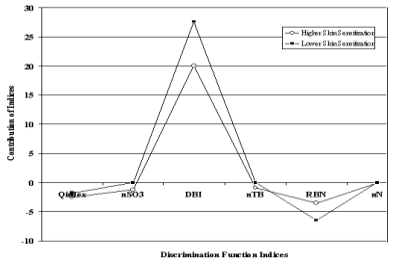


Fig. 1 Contribution plot of the discriminant functions for higher and lower skin sensitizing molecules.

Prevalence and Antibiotic Resistance of Gram-Negative pathogenic Bacteria Species isolated from *Periplaneta americana* and *Blattella germanica* in Varanasi, India.

D Leshan Wannigama [1], Rishabh Dwivedi [2], Alireza Zahraei Ramazani [3],*

[1 2] Molecular Biology Section, Laboratory of Industrial Microbiology, Department of Botany, [3] Department of Zoology, Centre of Advanced Study, Banaras Hindu university, Varanasi, India.

Background: Cockroaches are among the medically important pests found within the human habitations that cause serious public health problems. They may harbor a number of pathogenic bacteria with antibiotic resistance, which were carried either on the cuticle or in the gut. In this study aimed at conceive the prevalence and antibiotic resistance of gram-negative pathogenic bacteria species isolated from *Periplaneta americana* and *Blattella germanica* in Varanasi, India.

Method: In this study, 203 adult cockroaches (*Periplaneta americana* 130 and *Blatella germanica* 73) were collected form 44 households and 52 food-handling establishments in Varanasi city, India by trapping. Medically important gram-negative bacteria were isolated from external surfaces using standard methods. The Kirby-Bauer disc diffusion method was used to determine the profile of susceptibility to antimicrobials.

Result: Among the places analyzed, we found that have 54% cockroaches infestation in households and 77% in food- handling establishments. There was no significant different between the overall bacteria load of the external surface in *Periplaneta americana* (64.04%) and *Blattela germanica* (35.96%). The bacterial species were most frequently in food-handling establishments; however few bacterial species were identified from samples in households. Antibiotic resistance was found in *Pseudomonas aeruginosa, Klebsiella pneumoniae, Escherichia coli, Proteus mirabilis, Citrobacter freundii, Enterobacter aerogenes* species isolated from the cockroaches.

Conclusion: Therefore, we conclude that cockroaches species are uniformly distributed in domestic environment, which can be a possible carrier for transmission of drug-resistant bacteria and food-borne diseases.

In silico identification of proteins and peptides having therapeutic potential

Rahul Kumar, Kumardeep Choudhary and GPS Raghava

Bioinformatics Centre (BIC), @CSIR- Institute of Microbial Technology, Chandigarh, India.

Protein plays the most dynamic role in the body, from giving structural support to catalyzing a complex biochemical reaction, protein play immense diverse role. Because of this proteins always placed into an important position in the therapeutic world. Viewed from the perspective of therapeutics, proteins have several advantages like; they are highly specific to their targets, no or minimum adverse effects over small-molecule drugs, less approval time from FDA for therapeutic proteins make them favourite for pharmaceutical companies. Since the discovery of insulin, there is a vast expansion of therapeutic proteins market. Apart from this success of therapeutic proteins, we still need proteins which provide therapeutic potential for future. In literature, researchers have reported many peptides with anticancer activity and cell penetrating activity. But still, we need peptides and proteins having such important potential. We did data mining of swissprot proteins to find out some more handful number of proteins which can assure therapeutics potential. First, we collect peptides from the literature having cell penetrating activity and anticancer peptides which were collected from the APD2 database. We took out all the swissprot proteins having length less than 500 amino acids and divide them in five categories according to their lengths. Then, we carried out the BLAST, in which cell penetrating peptides was hit against the each class of swissprot proteins mentioned above at e-value of 0.001 and we noted the number of proteins for each class found in the BLAST output. We analysed all these proteins and found new proteins having either cell penetrating peptide sequences or anticancer peptide sequences which were not reported in the literature earlier. These proteins includes Zinc finger protein Krox-20, cycloviolacin O2 which belongs to the cyclotide family of plant proteins, Kalata, Histone 2A and various other proteins. These proteins can be potentially used to find out peptides with cell penetrating and anticancer activities. Cycloviolacin is previously known to have antitumor activity and in our analyses we show it also has cell penetrating potential, so this cyclic peptide is promisable therapeutic molecule. We also find out proteins having both kind of sequences (CPP and ACP) which includes varv peptides A, B & C, vodo peptide, Magainins and various other proteins. This study provides new proteins and peptides which can be used in future to develop cell penetrating peptides, anticancer peptides and proteins.

***Denovo* genome assembly and annotation of Microbes**

Surendra vikram, Shailesh Kumar, and Gajendra Pal Singh Raghava*

Bioinformatics Centre (BIC), @CSIR- Institute of Microbial Technology, Chandigarh, India.

Whole genome sequencing by the second generation sequencing technologies or the Next Generation Sequencing technologies (NGS) provides a new way to study whole genome of an organism. These technologies include Illumina's Solexa technology, ABI's SOLiD and Roche's 454 GS FLX, have a significant impact on the genomics. Reducing cost of whole genome sequencing now mean that it is realistic to use this to explore the full genomic potential of a microbe. These technologies can be helpful for the study of genome evolution and also for the microbial pathogenesis. Genome assembly and annotation of genomes from the data of Next Generation Sequencing technologies (NGS) is quite challenging. All the microbes have unique characteristic features. Here, we are showing the strategies of genome assembly and annotation of microbial genomes.

First, we have used SeqQC (http://genotypic.co.in/SeqQC.html?mnu=1), Fastxtoolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html) and NGS toolkit to filter the raw data produced by the sequencing machines. Second, we have used freely available genome assemblers like SOAPdenovo, Velvet, Ray and Newbler for assembly of short reads produced by sequencing technologies. Whole genome assembly was carried out at different parameters i.e. Hash length (K) and best assembly in terms of N 50 contig length have been taken for further analysis.

Genome annotation of prokaryotes has been carried out with the help of automated pipelines like IMG-ER, RAST, ISGA and Blast2go. Gene prediction and annotation of Eukaryotes was done by MAKER pipeline. Predicted proteins were further aligned to several databases like Non redundant database (NR) of NCBI (www.ncbi.nlm.nih.gov), SwissProt and KASS etc.

Lastly, to provide necessary information about genome assembly (i.e. *Denovo* and reference based) by Next Generation Sequencing we made a web page i.e. CRAG (www.imtech.res.in/raghava/crag). We have also developed useful Perl scripts for the analysis of genome assembly and annotations.

Characterization Of A Broad Spectrum Drug To Wide Range Of Mycobacterium

Visamsetti Amarendra[1], Kandasamy Dhevendaran[2] and Ramachandran Sarojini Santhosh[3]

[1, 3] Genetic Engineering Laboratory, Anusandhan Kendhra, SASTRA University, Thanjavur, India. [2] School of Chemical and Biotechnology, SASTRA University, Thanjavur, India.

Explorations of marine natural products gain interest since last two decades for disease treatment. Huge numbers of drugs have been characterized from the marine sources against infectious diseases. Tuberculosis is one of the most dreadful diseases in India killing 0.33 million every year and developing infections about 0.87 million people. *Mycobacterium* regained its resistance to the available antibiotics making the treatment worse than before. Combined efforts are taking in from novel product exploiters and drug synthesizing chemists to synthesize the products in bulk amount which is a major hurdle to get the compound from their natural habitat. Researchers are finding the molecular regulatory mechanisms involved in pathogenesis of tuberculosis species inside the host system and developing drugs targeting the underlying pathogenic mechanisms. Our group characterized extracts from two marine microbes, *Oceanobacillus iheyensis* KDRSSA and *Kocuria* sp. isolated from Vizhinium Coast marine sponge, *Tedania anhelans* for the presence of anti tubercular drugs. The extracts were showing the presence of diketopiperazines having antagonistic activity against *Mycobacterium* species. The extracts were biologically assayed against *M. fortuitum*, *M. smegmatis*, *M. scrofulaceum*, *M. vaccae*, *S. aureus*, and *P. aeruginosa* and minimum inhibitory concentrations were determined as ~100 µg/ml. Bacterium-bacterium assays showed effective inhibitions against *Mycobacterium* species. Compounds characterization and their chemical synthesis are the major works to be carried out in the future.

Computational screening for new inhibitors of *m. Tuberculosis* mycolyltransferases antigen 85 group of proteins as potential drug targets

Shachi Gahoi[1#], Rahul Shubhra Mandal[1#], Priyanka Shrivastava[1], Sriyans Jain[1], Nikita Ivanisenko[6], Ashish Kumar Singh[1], Muthukurrusi Varieth Raghunandanan[1], Swarna Kanchan[4], Bhupesh Taneja[3], Chhabinath Mandal[5], Vladimir A. Ivanisenko[6], Anil Kumar[2], Rita Kumar[2] [Open Source Drug Discovery Consortium][7], Srinivasan Ramachandran[1*]

[1]G.N. Ramachandran Knowledge Centre for Genome Informatics, [2]Environmental Biotechnology Division, [3]Institute of Genomics and Integrative Biology (CSIR), Delhi, India [4]Biological Science Group, Birla Institute of Technology and Science, Rajasthan, [5]Indian Institute of Chemical Biology (CSIR), 4, Raja S.C., Kolkata, India [6]Institute of Cytology and Genetics, SBRAS, Lavrentyev Avenue 10, Novosibirsk, Russia [7]Council of Scientific and Industrial Research, New Delhi, India

The group of antigen 85 proteins of Mycobacterium tuberculosis is responsible for converting trehalosemonomycolate to trehalosedimycolate, which contributes to cell wall stability. Here we have used a serial enrichment approach to identify new potential inhibitors by searching the libraries of compounds using both 2D atom pair descriptors and binary fingerprints followed by molecular docking. Three different docking softwares, AutoDock, GOLD and LigandFit were used for docking calculations. In addition, we applied the criteria of selecting compounds with binding efficiency close to the starting known inhibitor, and showing potential to form Hydrogen bonds with the active site amino acid residues. The starting inhibitor was ethyl-3-phenoxybenzyl-butylphosphonate, which had IC50 value of 2.0 µM in mycolyltransferase inhibition assay. Our search from more than 34 million compounds from public libraries yielded 49 compounds. Subsequently selection was restricted to compounds conforming to the Lipinski rule of 5 and exhibiting hydrogen bonding to any of the amino acid residues in the active site pocket of all three proteins of antigen 85A, 85B and 85C. Finally we selected those ligands which were ranked top in the table with other known decoys in both the docking results. The compound NIH415032 from Tuberculosis Antimicrobial Acquisition and Coordinating Facility was further examined using molecular dynamics simulations for 10 nanoseconds. These results showed that the binding is stable, although some of the hydrogen bond atom pairs varied through the course of simulation. The NIH415032 has anti-tubercular properties with IC90 at 20 µg/ml. These results will helpful to the medicinal chemists for developing new anti-Tubercular molecules.

Comparative Analysis of *Gatcab* Operon from Pathogenic and Non-Pathogenic Mycobacterium

Balasubramanian Suriyanarayanan[1], Ramachandran Sarojini. Santhosh[2]

[1,2]ASK 302, Genetic Engineering Lab, Anusandhan Kendra,SASTRA's Hub for Research and Innovation SASTRA University,Thanjavur, India

Mycobacterium species is known to cause several dreadful diseases in and around the world. Out of many species in this group *M.tuberculosis*, *M.leprae* are some of the pathogenic organisms causing diseases such as tuberculosis, leprosy etc. These bacteria have become highly resistant to almost all antibiotics and towards the host defence mechanisms (immune responses). At present, various molecular targets has been considered in the process of designing a suitable inhibitor as effective drugs against organisms. Mycobacterium has several strategies in escaping from the host defence. Out of many, the bacterium has the ability to remain dormant inside the macrophages and subsequent prevention of phagosome and lysosome formation which is one of the main reasons. These species are characterised by the presence of *gat*CAB operons. These genes code for the enzyme glutamine amido transferases which mediates transamidation reaction. These are heterodimeric proteins with three subunits namely GatC, GatA and GatB. These proteins are involved in transamidating mis-charged glu- $tRNA^{gln}$ to gln- $tRNA^{gln}$. With the help of this operon, the bacteria might escape from the defence mechanisms of the host. Thus these proteins might act as effective drug targets against Mycobacterial infections. We are cloning and over express gat CAB operon from *M.smegmatis* and *M.tuberculosis*. The proteins are to be purified and a comparative study of the enzyme activity and structural analysis are to be made.

Analyzing Highly Co-Expressed Sub-Network Modules across Multiple Strains Of Mycobacterium Tuberculosis

Bhanwar Lal Puniya, Srikant Prasad Verma and Srinivasan Ramachandran

CSIR-Institute of Genomics and Integrative Biology, Delhi, India

Networks and models analysis can be used as data exploratory method and to generate testable hypotheses. Gene expression variability among Mycobacterial strains is important to understand critical biological processes and genes. To investigate network modules of variably expressed genes we have constructed a weighted gene co-expression network by using log-phase expression data of 12 different strains of *Mycobacterium tuberculosis*.

Our main aim of this study is to explore the following biological questions using network approach. First, what are the network modules of genes, which are variably expressed across multiple strains? Second, are these modules enriched with some biological processes? Third, what is the distribution of known validated or predicted drug targets? What are the genes which form hubs in these network modules?

Our weighted gene co-expression network was built by using 38 samples of different Mycobacterial strains under log phase expression. After selecting biologically relevant co-expression modules, we mapped expression categories on networks to identify variably or differentially expressed genes. We also found distribution of known high confidence drug targets in network clusters.

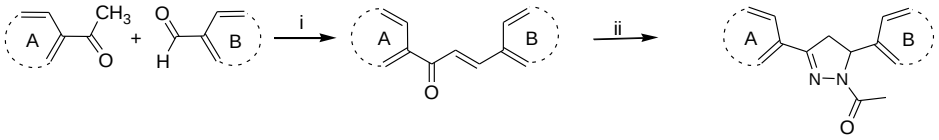
In our study we have identified network clusters, which are significantly enriched with biological processes such as nucleotide binding proteins, tryptophan biosynthesis, oxidative phosphorylation, ribonucleoproteins and transposable elements. Transposable elements showed the highest enrichment of 46 fold. All selected sub-networks are significantly enriched with variably expressed genes. By using these clusters we have extracted the highly variable co-expressed sub-network modules. We have located Rv1611 (trpC) as top hub in network which is already found to be high confidence drug targets.

Non purine based xanthine oxidase inhibitors: synthesis, evaluation and docking studies.

Preet mohinder singh bedi[a], Kunal nepali[a]

[a] Department of pharmaceutical sciences, Guru Nanak Dev University, Amritsar, Punjab, India

Background : Interactions of purine analogue XO inhibitors on activities of purine and pyrimidine metabolism enzyme leading to the hypersensitivity, (Stevens-Johnsons) a syndrome characterized by fever, skin rash, hepatitis, leukocytosis with eosinophilia and worsening renal function induced in some of the patients encouraged the researchers to focus on XO inhibitors with structurally diverse and novel non-purine isosteres. 1-Acetyl-3,5-diaryl-4,5 dihydro(1*H*)pyrazoles have previously been reported to exhibit a variety of biological activities. In the present work, synthesis of non-purine 1-acetyl-3,5-diaryl-4,5-dihydro(1*H*) pyrazoles without ignoring shape and their structure activity relationships as a new class of XO inhibitors for the first time has been carried out. Analogues of 1-acetyl-3,5-diaryl-4,5-dihydro(1H)pyrazoles were rationally designed and synthesised as per the below mentioned scheme:



Reagents and conditions: (i) 5% NaOH, ethanol, rt, 1 h (ii) $NH_2NH_2 \cdot H_2O$, CH_3COOH, reflux, 4h

The molecules were evaluated for in vitro xanthine oxidase inhibitory activity using bovine milk xanthine oxidase (grade 1, ammonium sulfate suspension, Sigma–Aldrich) enzymatic assay as described in the literature.

Result: Among the series of compounds, KL-1 was found to be the most potent (IC_{50} = 5.30 µM) comparable to that of allopurinol (IC_{50} = 8.3 µM). Docking studies were also performed and it was found that the S-isomer fits well in the Xanthine Oxidase Binding cavity.

Conclusion: On the basis of SAR studies, a basic pharmacophore with the structural features required for Xanthine Oxidase inhibitory activity has been proposed.

Gene Ontology Based Annotation using HmmModE

Swati Sinha and Andrew M. Lynn

School of Computational and Integrative Sciences, Jawaharlal Nehru University, India

Profile Hidden Markov Models (HMM) are statistical representations of protein families derived from patterns of sequence conservation in multiple alignments and have been used in identifying remote homologues with considerable success. These conservation patterns arise from fold specific signals, shared across multiple families, and function specific signals unique to the families. The availability of sequences pre-classified according to their function permits the use of negative training sequences to improve the specificity of the HMM, both by optimizing the threshold cutoff and by modifying emission probabilities to minimize the influence of fold-specific signals. A protocol HmmModE was developed earlier to generate family specific HMMs that first constructs a profile HMM from an alignment of the family's sequences and then uses this model to identify sequences belonging to other classes that score above the default threshold (false positives). This protocol uses the program *hmmer,* which was recently upgraded to a newer version with a design goal of combining the power of probabilistic inference with high computational speed. Although faster, the new HMM profile used by the program, besides having a different format wherein all the probability parameters are stored as negative natural log probabilities, has changes like the addition of more options for controlling weights, profile construction, effective sequence number and e-value calibration etc. The program HMM-ModE has been re-written as perl modules that work with *hmmer3*. The newer version of HmmModE protocol has been tested on three different data sets including pre-classifier a protein kinase dataset, GPCR dataset and the entire metabolic enzyme dataset previously described by our group, and compared with the older version using hmmer2. The protocol has been further implemented to create highly specific profiles from pre-classified Gene Ontology sequences from GO database and Non-Supervised Orthologous data from eggNOG which are useful for annotation of sequences of unknown function.

MetaNET: A web-accessible platform for metabolic network simulation and analysis

Shawez Khan, Pankaj Narang and Andrew M. Lynn

School of Computational and Integrative Sciences, Jawaharlal Nehru University, India

We describe an open source, integrated, web accessible platform MetaNET, to perform various tasks of computational systems biology. It is a freely available representing an advancement for the easy application of systems biology. By providing a user-friendly rich interface, the platform allows users to simulate genome-scale metabolic networks under various genetic or environment conditions. It has been integrated with different functionalities of System Biology Research Tool (SBRT) which provides only a basic non-interactive GUI, merely to launch the execution of its processes. MetaNET integrates all SBRT functions with the open source platform Galaxy - a web accessible platform to integrate different command line tools, with work-flow capabilities. Along with all the processes of system biology research tool, we have implemented some new functions using Perl and Java language that is being increasingly used by scientific community. MetaNET currently has more than 60 tools including the optimization of objective functions, single or pairwise gene knock-out or knock down simulations, reaction and catalyst deletion analysis, robustness analysis and prediction of cell phenotype under different genetic or environment conditions using methods of FBA. The present version of the tool incorporates

(i) Optimization of Objective function for wild type strain, Gene Knock Analysis, Reaction deletion Analysis, using the method of Flux Balance Analysis (FBA).

(ii) Computation of flux intervals of all reactions in a stoichiometric network using Flux variability analysis (FVA)

(iii) Chemical species Participation.

(iv) Choke Point reactions analysis to facilitate drug target identification in drug discovery

(v) Graphically visualize results from various tools.

Text Mining Applications in Translational Research

Jeyakumar Natarajan

Data Mining and Text Mining Laboratory, Department of Bioinformatics, School of Life Sciences, Bharathiar University, Coimbatore, INDIA

The linkage between the clinical and laboratory research domains is a key issue in translational research. Most of the current laboratory research outputs are available online in the form of abstracts (e.g. PubMed/MEDLINE) and full text research articles (e.g. PubMED Central, BioMed Central). However, the growth in publications is exponential and at this rate of publication, it is difficult or impossible for biologists to keep up with the underlying hypotheses. For example the biomedical literature database MEDLINE currently contains about 15 million abstracts and an average about 40,000 to 50,000 new abstracts are added every month. This demands the use of automated text processing methods for the organization, the maintenance and the analysis of these data.

Text mining is the application of techniques from machine learning in conjugation with natural language processing, information retrieval and statistical/mathematical approaches to extract useful knowledge from text. The application of biomedical text mining includes named entity tagging (e.g. genes, proteins, enzymes, drugs etc.), entity concept relation extraction (e.g. protein-protein interactions, drug-disease relations etc.) and mining new facts (e.g. biomedical pathways and functions). Integrating the above text mining results to a data warehouse related to a broad range of human diseases will make these data usable/available at the point-of-need for a translational research community.

This presentation will fist highlight text mining basic concepts, methodology, and its various applications in biomedical domain. This will be followed by a brief introduction cum demo of our newly developed open source text mining tools NAGGNER – A hybrid named entity tagger for tagging human protein and gene mentions in biomedical literature and ProNormz – A gene/protein name normalization system available on line at http://www.biomining-bu.in

Potential use of Biodegradable Nanoparticle-Based Anti-Tb Drug Delivery Systems

Vandana Sharma [1]

MCMDAV College for Women, Chandigarh, India

The appearance of extensively drug resistant TB poses new challenges for its prevention, treatment and control. The treatment with available drugs is associated with noncompliance to therapy and only a small fraction of the anti-tubercular drugs reach an alveoli which is the desired site of drug action. This has created an intense need for presenting old drugs, such as those encompassed in the anti-TB regimen, in new forms utilising novel modes of delivery and dosage forms. Drug delivery, which takes into consideration the carrier, the route and the target, has evolved into a strategy of processes designed to enhance the efficacy of therapeutic agents through modified or controlled release. Carrier or delivery systems such as liposomes and microspheres have been developed for the sustained delivery of anti-TB drugs and have demonstrated better chemotherapeutic efficacy when investigated in animal models. Solid lipid nanoparticles (SLNs) are another antimicrobial drug delivery platform which possess good tolerability, ability to incorporate hydrophilic and hydrophobic drugs and enhanced stability of incorporated drugs. The advantage with SLNs is that unlike liposomes, their long term stability as well as drug incorporation efficiency is better, whereas in contrast to polymeric formulations, the risk of residual organic solvents is minimum. Non-ionic surfactant vesicles or niosome can entrap solutes in a manner analogous to liposomes, are relatively more stable *in vitro* and can improve the stability of entrapped drug as compared with stability in conventional dosage forms. Dendrimers are synthetic, high branched polymers which provide enormous surface area to size ratio and allows great reactivity with microorganisms *in vivo*. Various advantages of nanoparticles as drug carriers, including size, surface charge, composition and the presence of ligands, their *in vivo* stability, increased specificity of these drug delivery systems to recognize and bind to target cells, improvement of drug bioavailability and reduction of the dosing frequency and feasibility of the versatile routes of drug administration, including oral and inhalation routes. Present article compiles the new research updates on the potential of nanoparticle based antimicrobial drug delivery systems to improve treatment to bacterial infections, especially in life threatening diseases such as tuberculosis.

Fold and function specific residues prediction from protein sequences using an information theoretic method.

Rehan M, Andrew M. Lynn

Jawaharlal Nehru University, New Delhi. India

According to the neutral theory of molecular evolution, once a protein has evolved to a useful level of functionality, the majority of mutation are selectively neutral at the molecular level and do not affect the function and folding of the protein, whereas those mutations which are deleterious provide selection pressure for residue conservation. Thus, the conservation indicates the importance of residue for structure and function of the protein. Often, the mutations at a conservation site after the gene duplication lead to functional divergence. The residues of a protein at these sites called function specific residues when such residue is mutated causes change or total loss of the protein's function. In general, for a class of protein, sequences are grouped into subtypes contain subtype conserved residues indicative of functional differences among the subtypes. For such a protein class, we present a method for ranking residues importance for functional activity and finally their separation into function specific and fold specific ones. Our method is based on two kinds of Relative Entropy, RE_{fold} and $RE_{function}$. RE_{fold} calculates the conservation over background distribution of amino acids in the whole family whereas the $RE_{function}$ tells about differential conservation among subtypes. The $RE_{function}$ and the RE_{fold} are combined in a way to give a new score, $RE_{context}$ to rank the residues for fold and function. The methodology is implemented using HMM, validated on Major Facilitator Superfamily (MFS), G-Protein Coupled receptors (GPCRs) & AGC Protein Kinases and performing better than other existing methods.

Molecular modeling, docking and dynamic analyses of ndm-1 proteins.

Abhijit Chakraborty[1], Tuhin Kumar Pal[1] and Saikat Chakrabarti[1].

[1]CSIR-Indian Institute of Chemical Biology (IICB), Kolkata, India.

Background:

New Delhi Metallo-β-lactamase (NDM-1) mediated antibiotic resistance has created severe clinical threat for human health. Metallo-β-lactamases (MBL) are divided into four classes where class A, C and D are serine- β-lactamase, and class B is called metallo-β-lactamase, as they require one or two Zinc atoms for their hydrolyzing activity. NDM-1 is a class B MBL that hydrolyzes beta-lactam ring of most of the carbapenem group of antibiotics.

Result:

Sequence analysis suggests that Verona integron-encoded metallo-β-lactamase (VIM) family of MBL as the closest relatives of NDM-1 protein. Three dimensional modeling studies suggest mono-Zinc dependent lactamase activity for a truncated NDM-1 (158 residues) protein while the predicted structure of the longer NDM-1 (270 residues) protein greatly resembles the recently solved, experimentally derived three-dimensional (3D) structure. With respect to VIM-2 proteins, several mutations at the important loop regions (loop1 and loop 2) regulating the substrate recognition and accessibility to the active sites, are observed in NDM-1. Molecular dynamic results shows higher structural flexibility around these loops making it accommodative for multiple ranges of antibiotics. Docking analysis showed a better binding of carbapenem antibiotics by NDM-1 proteins compared to the VIM family proteins.

Conclusion:

Our sequence and structural analyses suggest that the truncated NDM-1 might adopt a mono-Zinc dependent hydrolysis mechanism as it lacks suitable residues for crucial second Zn^{++} coordinating. Molecular dynamic results indicate higher loop flexibility in NDM-1 proteins and molecular docking results suggest better predictive binding of NDM-1 proteins to a broad spectrum of antibiotic ligands compared to VIM-2 protein.

In-Silico Analysis of Transfr RNA Modifications

Bharat Panwar and G. P. S. Raghava

Bioinformatics Centre, Institute of Microbial Technology (CSIR), Chandigarh, India

Background

The post-transcriptional modification diversifies four basic bases, provides flexibility and specifies each transfer RNA (tRNA) for the particular function. The basic cloverleaf structure of tRNA obtained only after certain modification steps. There are great chemical diversity in tRNA modifications and ~50 different types of nucleotides modification are known. The nucleotide modification of anticodon loop affects the codon-anticodon interactions, which leads into the protein translation information. It is important to analyze and understand the base-specific modification patterns of tRNA.

Results

In this study, we have used a non-redundant dataset of 136 tRNA and 40 different type of modifications from MODOMICS database. We analyzed that average ~8 bases of each tRNA modified after transcription. We have created overlapping patterns of modified and unmodified bases and created two sample logo. It was observed that 60.36% modified bases are uridine-derived and an average 26.83% of total uridine in tRNA is modified. Pseudo-uridine, Dihydro-uridine and 5-methyl-uridine are three most common uridine-derived modifications prsent in tRNAs. Pseudo-uridine (Y) contains 40.54% of all Uridine-derived modifications. Dihyro-uridine (D) modification is present in 31.32% of uridin-derived modifictions, it is a major part of the D-loop but also present in the other loops of tRNAs. 5-methyl-uridine pressnt only once in the T-loop of tRNAs and there is a conserved site (GTDC) for the modification. It means the prediction of 5-methyl-uridine in T-loop is simple but which base will convert into the Pseudo-uridine and Dihydro-uridine is still difficult to predict.

Conclusions

The study shows that there are different patterns for different uridine-modification types and these knowledge can also be convert into the automatic *in silico* prediction method to predict and classify these uridine-derived modifications.

A Complex Network Of Micrornas Expressed In Brain And Genes Associated With Amyotrophic Lateral Sclerosis (Als)

Santosh P. Shinde, Neelima Arora and Utpal Bhadra

Functional Genomics and Gene Silencing Group, Centre for Cellular and Molecular Biology-CCMB, Hyderabad, India.

Amyotrophic Lateral Sclerosis (ALS) is a rare neurological disease affecting mainly motor neurons and often leads to paralysis and death in extreme cases. Though the number of patients of ALS is relatively small compared to other neurodegenerative disorders, it holds distinct importance owing to high fatality among victims. Various mechanisms are proposed for the disease but exact mechanism still remains a mystery. Recently various microRNA (small non coding RNA) expressed in a spatially and temporally manner in the brain. The pathological implication of aberrant miRNA expression in ALS remains largely unknown. For exploring the role of microRNAs in ALS disease genes regulation and to identify the complete miRNA- ALS related gene network, miRanda was employed for prediction of target sites of miRNAs expressed in various parts of brain and CNS on 35 genes associated with ALS. MicroRNA target sites were classified according to the position and the location of these sites were used to infer their mode of action in target gene silencing. For verification, similar search was conducted using TargetScan and PicTar for prediction of target sites. Target site hotspots were identified and were recognized as hotspots for multiple miRNAs action, thus, acting as favoured sites of action for repression of gene expression. The complex interplay of genes and miRNAs brought about by multiplicity and cooperativity was explored. This investigation will aid in elucidating the mechanism of action of miRNAs for the considered genes. The intrinsic network of miRNAs expressed in nervous system and genes associated with ALS may provide rapid and effective outcome for therapeutic applications and diagnosis.

Predictive models for anti-tubercular molecules using machine learning on high-throughput biological screening datasets

Vinita Periwal[1], K.R. Jinuraj[2], U.C.A. Jaleel[2], Vinod Scaria[1]

[1]GN Ramachandran Knowledge Center for Genome Informatics, Institute of Genomics and Integrative Biology (CSIR), New Delhi, India, [2]Department of Cheminformatics, Malabar Christian College, Kerala, India

Tuberculosis is a contagious disease caused by *Mycobacterium tuberculosis* (Mtb), affecting more than two billion people around the globe and is one of the major causes of morbidity and mortality in the developing world. Recent reports suggest that Mtb has been developing resistance to the widely used anti-tubercular drugs resulting in the emergence and spread of multi drug-resistant (MDR) and extensively drug-resistant (XDR) strains throughout the world. In view of this global epidemic, there is an urgent need to facilitate fast and efficient lead identification methodologies. Target based screening of large compound libraries has been widely used as a fast and efficient approach for lead identification, but is restricted by the knowledge about the target structure. Whole organism screens on the other hand are target-agnostic and have been now widely employed as an alternative for lead identification but they are limited by the time and cost involved in running the screens for large compound libraries. This could be possibly be circumvented by using computational approaches to prioritize molecules for screening programmes.

We utilized physicochemical properties of compounds to train four supervised classifiers (Naïve Bayes, Random Forest, J48 and SMO) on three publicly available bioassay screens of Mtb inhibitors and validated the robustness of the predictive models using various statistical measures.

This study is a comprehensive analysis of high-throughput bioassay data for anti-tubercular activity and the application of machine learning approaches to create target-agnostic predictive models for anti-tubercular agents.

Association of Snp-63 of Calpain 10 Gene and Risks Associated with Type 2 Diabetes in Chandigarh and its Surrounding Population

Aatish Mahajan *

*SUS College of Research and Technology, Punjabi university, India

Countries like India have shown significant improvement in health care delivery in communicable disease sectors, but face severe problem due to the burden of non-communicable diseases like type 2 diabetes mellitus (T2DM). T2DM is a multifactorial disease with both environmental and genetic factors contributing to its development. Our study included 30 T2DM patients and 30 unrelated non-diabetic controls. The study was carried out to find out the various risk factors associated with T2DM. The present case-control study had an advantage because it was drawn from relatively small uniform environment and may provide new information for diagnostic treatment and prevention of T2DM in Indians. Subject was designated diabetic/ non-diabetic on basis of various biochemical parameters. A questionnaire was filled up by each subject regarding their place of birth, alcohol consumption habits, educational qualifications, physical activity and economic status. The blood samples drawn were used for extraction of genomic DNA for genotyping of *CAPN-10* gene using PCR and then the results were checked by applying statistical methods like *P*- value, C.I., ODDS ratio by using SPSS 10. Two to four fold increased risk of developing T2DM due to Physical inactivity was observed. In genotypic and allelic study a significant association ($P< 0.05$) of SNP-63 of *CAPN 10* gene was found to be associated with 2.5 fold increased risk. Education, alcohol consumption, clinical Information of subjects stratified by sex and disease, coronary heart disease, ocular infections and dental diseases were found significantly ($P< 0.05$) associated with T2DM. The identification of environmental risk factors and diabetes susceptibility genes will allow better understanding of the molecular mechanisms underlying pathogenesis.

Comprehensive *in-silico* analysis of genomic variations in human noncoding Rnas and their functional implications

Deeksha Bhartiya, Saurabh V. Laddha, Jatin Talwar, Saakshi Jalali, Yasha Hasija, Arijit Mukhopadhyay, Vinod Scaria*

GN Ramachandran Knowledge Center for Genome Informatics, CSIR Institute of Genomics and Integrative Biology, Delhi, India

Recent advances in sequencing technologies have made genome-scale mapping of transcriptional potential possible at single-nucleotide resolutions. Though consistent evidence now suggests the existence of a large number of previously uncharacterized transcripts, a vast majority of the transcripts, termed as noncoding RNAs (ncRNAs), surprisingly do not encode for a functional protein. This class now encompasses a wide diversity of subclasses which include the well studied class of microRNAs and others like snoRNAs, piRNAs, and long noncoding RNAs, including long intergenic noncoding RNAs, anti-sense transcripts and processed pseudogenes. Though genomic variations have been extensively studied in protein-coding genes, there remains paucity in understanding the effect of genomic variations in ncRNAs.

To analyze the potential roles of genomic variations in ncRNAs and predict potential biological outcomes of the variations, we have systematically compiled and mapped genomic variations to a manually curated compendium of ncRNA loci. We also developed a systematic computational pipeline which includes global structural ensemble comparison and assessment of potential biological functions to assess the potential functional role of the variations. It also includes assessment of sequence conservation and modules for predicting effects on biogenesis and or function, apart from analysis of the RNA structural diversity ensemble. We performed the analysis for 3 major classes on ncRNAs: miRNAs, snoRNAs and lncRNAs which revealed a subset of variations with potential to alter the structure and function of ncRNAs including potential alteration of the biogenesis of miRNAs. Further mapping of genome-wide association signals and analysis of the minor allele frequencies of the variations in global populations revealed potential mechanisms of variations modulating effect through ncRNAs. The implication of the variations *vis a vis* the regulatory networks suggest potential novel ways variations in ncRNA loci could modulate biological and disease processes.

Identification of common target proteins for neurodegenerative disorders through metabolic pathway approach

Khurshid Ahmad[1]*, Saif Khan[1]

[1]Integral University, Lucknow, India

Neurodegenerative disorder is a condition in which cells of the brain and spinal cord are lost. The brain and spinal cord are composed of neurons that do different functions such as controlling movements, processing sensory information, and making decisions. Aside from a small number of neural stem cells that are created daily, cells of the brain and spinal cord are not readily regenerated en mass, so excessive damage can be devastating. The study of metabolic pathways is becoming increasingly important to exploit an integrated, systems-level approach for optimizing a desired cellular property or phenotype. The most prominent CNS disorders are Alzheimer s Disease, Parkinson s Disease, Huntington s Diseases, amyotrophic lateral sclerosis (ALS) and few others. There are six following neurodegenerative disorders which are most common in human. Alzheimer's disease, Prion disease, Parkinson's disease, ALS (Amyotrophic Lateral Sclerosis), DRPLA (Dentatorubropallidoluysian atrophy) and Huntington's disease. The pathways for these neurodegenerative disorders are available in KEGG (Kyoto Encyclopedia of Genes and Genomes) and in other pathway portals which provide the details of the responsible proteins or genes for that disease. A comparative study of the available disease pathways shows that there are many proteins or genes which are interlinked with more than one disease.

The pathway analysis of above mentioned major neurodegenerative diseases shows that there are following proteins which are interlinked with more than one disease: NGFR, APLP1, GFAP, BCL2, HSPA5, MAPT, FBXW7, CASP8, CASP3, CASP6, CASP7, CASP1, GRB2, GAPD, APBA1, BCL2L1. This study can be more useful in drug designing field by getting a common and suitable drug for more than one neurodegenerative disease. The significance of this work is to put a step forward in constructing a common pathway for major neurodegenerative diseases.

Phylogenetic Comparison of Protein Ser/Thr Phosphorylation among Pathogenic and Non Pathogenic Mycobacterium Species

Gunjan Arora[1], Andaleeb Sajid[1], Amit Kumar Yadav, Puneet Kumar Kadimi, Nupur Verma[1] and Yogendra Singh[1]

CSIR-Institute of Genomics and Integrative Biology, Delhi, India

A major challenge of post-genomic biology is to understand the complex phosphorylation networks of interacting proteins that regulate cellular functions. Advances in Mass spectrometry techniques led to characterization of such networks by phosphoproteomic approaches. The large number of phosphoproteins discovered in *Mycobacterium tuberculosis* present opportunity to understand the evolution and function of such patterns in the Mycobacterium species. The aim of this study is to segregate the conserved phosphoproteins from non-conserved ones. We have tried to identify the Ser/Thr phosphosites which are conserved in pathogen and non pathogen species such as *M. tuberculosis*, *M. bovis* BCG, *M. marinum*, *M. ulcerans*, *M. smegmatis* and *M. avium* paratuberculosis. Evolutionarily cross-species comparison will enhance our understanding of regulatory mechanisms of Mycobacterium species.

Association of Big Two: "Tuberculosis And Malaria" Co-Infection And Need For Computational Tools

Harsimrat Kaur

Dolphin (PG) College of Life Sciences, Chunni Kalan. Punjab, India

Tuberculosis and malaria are endemic in many regions of the world, and due to significant overlap of areas endemic for them; co-infection with *Mycobacterium tuberculosis* and *Plasmodium* sp. is likely to be common. Co-infections come along with epidemiological, immunological and chemotherapeutic implications. Only very few studies have been carried out to determine the effect(s) that *Myobacterium* and *Plasmodium* bear on each other when present concomitantly. The disease associations between the big two "TB and Malaria" during co-infections need detailed deterministic studies in respect of immunodynamics and co-treatment strategies. In our laboratory* we tested *Mycobacterium tuberculosis* and *Plasmodium sp.* co-infection in murine model established by inoculating Swiss and Balb/c mice intravenously with *M. tuberculosis* H37Rv and four weeks later with *P. berghei* (moderately lethal)- or *P. chabaudi chabaudi* AS (non lethal)- or *P. yoelii nigeriensis* (highly lethal)-parasitized erythrocytes, intraperitoneally. In Swiss mice having *M. tuberculosis* H37Rv infection and co-infected with *P. berghei, P. chabaudi chabaudi AS and P. yoelii nigeriensis* the rates of progression of parasitaemia were significantly lower ($p<0.05$) compared to *Plasmodium* sp.-only infected controls. The lower rate of parasitemia in co-infected mice was associated with increased survival. On the contrary, *Plasmodium* sp.-only infected or co-infected Balb/c mice neither differed with respect to rates of parasitaemia progression or mortality rates clearly implicating role of host genetics during TB-malaria co-infection. There is a pressing need to further our understanding in this critical area of TB-malaria co-infection and it seems using computational methods would be a logical choice to sort wide range of assortment of microbes, molecules and mechanisms operative during TB malaria co-infections. Mtisi (2009) operated a realistic deterministic ODE based compartmental model for the transmission of malaria and tuberculosis co-dynamics in a population. Wiwanitkit (2006) used a new gene ontology technology "GoFigure" to delve into molecular functions of HSP70 and ATPBP during episode of TB-malaria co-infection. Clearly, the studies are numbered in this area and vie for more input.

In Silico Analysis of *Sorghum* Oxalate Oxidase

Sneha Munshi, Vijay Dangi

University Institute of Engineering & Technology, M.D. University, Rohtak, Haryana, India

Oxalate oxidase is a multifunctional glycoprotein which catalyzes the oxidation of oxalate. Oxalates are produced naturally during assimilation of carbon in the form of their insoluble calcium salts which lead to chronic, sometimes acute, 'stone diseases' in human. Present study was carried out to investigate the structural and biochemical properties of Sorghum oxalate oxidase by in silico approach. Comparative Modeling was done to predict protein structure based on sequence homology. Sorghum oxalate oxidase sequence has 48% similarity with Barley oxalate oxidase (1FI2). Physiochemical properties of Sorghum oxalate oxidase were calculated with Expasy ProtParam server. 2D structure of Sorghum oxalate oxidase sequence was predicted by PSIPRED server. 3D structure was predicted with Modeller v9.10 using crystal structure of Barley oxalate oxidase as template. Orientations of first six amino acids, which have no similarity with Barley oxalate oxidase, were modeled on the basis of 2D structure as predicted by PSIPRED. DOPE score profile of predicted structure was aligned with DOPE score profile of Barley oxalate oxidase. Best structure was selected on the basis of low DOPE score and similarity of DOPE profile. Regions of selected structure where DOPE score plot varies significantly from the DOPE score of template were further refined by loopmodel class of Modeller. The best structure selected after the loop refinement were further optimized by MD Simulation using NAMD 2.6. The protein model was solvated in a water box and periodic boundary conditions were applied to the system in the canonical ensemble and system was equilibrated for 5ps. Final structure was obtained from trajectory of simulation and evaluated with Structure Assessment tool of SWISS-MODEL server. In Ramachandran plot, 86.7% residues were in core region, 11.6% in allowed region and 1.7% in generously allowed region. Procheck analysis showed no bad contact of atoms and all parameters for main chain and side chain were better. QMEAN showed 84.4 % residues were in secondary structure agreement. The dDFIRE score of predicted model was -368.48. This low value of free energy suggests it as a good conformation.

***In-silico* designing of potent lead molecules against β_2 adrenergic g-protein coupled receptor ($\beta_2$ar)**

Varinder Kumar[1], Parakh Sehgal[1]

[1]Department of Bioinformatics, GGD SD College, Chandigarh INDIA

The drug designing approach involves a large number of chemical compounds that are simulated, synthesized, evaluated and in most cases, discarded. Crystal structures of proteins and potential ligands are the cornerstones of our approach towards *in-silico* ligand designing. Since the first brick to build upon is the protein target, in this study we selected β_2 Adrenergic G- Protein Coupled Receptor (PDB ID: 2RH1, 2.4 Å resolution), a member of the GPCR family, eukaryotic signal transducing transmembrane protein as our receptor protein. These proteins are responsible for a number of diseases like Parkinsons, Chronic asthma, Alzheimer's etc, which make members of this family important drug targets. In PDB file 2RH1 the complex β_2AR-T4L was found bounded to partial inverse agonist, Carazolol ((2-propanol,1-9H-carbazol-4-yloxy)-3-[(-1\-methylethyl)amineamino]). Carazolol occupy the binding cavity and found showing interaction at carbazole moiety adjacent to Phe289, phe290 and Trp286. Ligands were generated by applying structure based approach which would occupy the active site of β_2AR. A total of 4000 ligands were generated using two different seed molecules, and finally 20 elite molecules were chosen on the basis of comparison of physiochemical properties of these with already available drug structures. These elite conformers were checked for their toxicity risk assessment, drug likeliness and other drug like properties like non-tumorogenecity, non- irritability and finally two ligands were found to be having high drug score and low toxicity assessment risks. These molecules were than docked into the receptor binding pocket of β_2AR which resulted into 156 conformers with different binding modes and energies. On the basis of the binding position of the ligand within the active site of the receptor molecule and binding energy, 19 conformers were selected. After analyzing all 19 conformers for their interaction with the active site residues, two elite ligands were choosen which shows hydrogen bonding with active site residues Phe193, Asp192, Trp313 (E_{total} = -215.18) and Ser203, Asn293 (E_{total} = -202.92) respectively as seen in case of carazolol. These potential lead molecules show close similarities with already available successful drug compounds i.e. molecular weight, $\log_p$, H-donor & acceptor, drug score etc. We proposed potential lead molecules against the receptor binding ite of β_2AR.

Phenyl-Propanoids from *Alpinia Galanga* as Modulator of Ethidium Bromide Susceptibility against *Mycobacterium Smegmatis* Mc^{2}155

Sonika Pahwa[1], Somendu K. Roy[2], Hemraj S. Nandanwar[1] and Sanjay M. Jachak[2]

[1] Institute of Microbial Technology, Chandigarh, [2] Department of Natural Products, National Institute of Pharmaceutical Education and Research, Punjab, India

Background: Efflux pump inhibitors (EPIs) are expected to block the efflux of antibiotics from the site of action and provide complete antimicrobial effect. Emerging evidence indicates that secondary metabolites from plants as well as microbes combat efflux-mediated microbial resistance either by modulating the anti-mycobacterial activity of known anti-TB drugs or exhibiting EPI activity. So, in search of new modulators for anti-mycobacterial activity, lipophilic extracts of *Alpinia galanga* were investigated. Previous report revealed high antimicrobial activity of lipophilic extract of *A. galanga* against Gram-positive and anaerobic bacteria. Here in, we investigated Minimum Inhibitory Concentration (MIC) of eight isolated pure compounds and their modulating effect on ethidium bromide (EtBr) against *M. smegmatis* mc^{2}155.

Results: Eight phenlypropanoids viz. 1´-*S*-1´-Acetoxycavicol acetate (1, ACA), *trans*-*p*-coumaryl diacetate (2), 1´-*S*-1´-acetoxyeugenol acetate (3), 1´-*S*-1´-hydroxychavicol acetate (4), 4-(*E*)-3´-hydroxyprop-1´-enylphenyl acetate (5), *trans*-*p*-hydroxycinnamaldehyde (6), *trans*-*p*-hydroxylcinnamaldehyde acetate (7) and *trans*-*p*-coumaryl alcohol (8) [Fig. 1]; were isolated from *A. galanga* lipophilic extracts where 1 and 3 exhibited MIC values of ≤ 50 µg/ml. ACA was found to be potent modulator and potentiated the EtBr by 64 fold at 2.5 µg/ml.

Conclusion: The change in modulation activity may be due to methoxy group at C-5 position of phenyl ring in ACA. Further, *in-silico* designing of structural variants by QSAR and possible synthetic derivatives may add potential drug candidate to be used as promising modulator and EPI for TB drugs, in future.

Acetyl-11-Keto -B -Boswellic Acid; a Potent Antibacterial and Anti-Biofilm Agent for Oral Cavity Pathogens

[1]Alsaba Fida Raja* and [2]Furqan Ali

Infectious lab PK/PD, Indian Institute of Integrative Medicine (CSIR), [1]Sanatnagar, Srinagar India [2]Jammu, India.

Background: Several microorganisms inhabit the human oral cavity, and there is always a risk of infection with bacterial pathogens associated with the oral cavity. *Streptococcus mutans* is a key contributor to the formation of biofilms associated with dental caries disease. The enzyme glucosyltransferase produced by *S. mutans* is the factor behind the production of metabolic organic acids, demineralization of the tooth surface and lead to dental caries.

Boswellic acids, major constituents of the gum resin derived from the plant *Boswellia serrata*, comprises of β-boswellic acids as the main triterpenic acid along with 11- keto- β-boswellic acids and their acetates. Boswellic acids have been extensively studied for a number of activities including anti inflammatory, antitumor, immunomodulatory and inflammatory bowel diseases. The present study describes the antimicrobial activities of boswellic acid molecules against oral cavity pathogens. Acetyl-11-keto-β-boswellic acid (AKBA), which exhibited the most potent antibacterial activity, was further evaluated in time kill studies, postantibiotic effect (PAE), effect of AKBA on glucosyl transeferase and biofilm susceptibility assay against oral cavity pathogens.

Results: AKBA was the most active of the four boswellic acids against all bacterial pathogens. AKBA exhibited MIC ranging from 2-4 μg/ml. It exhibited concentration dependent killing of *Streptococcus mutans* up to 8 × MIC. AKBA demonstrated PAE of 5.7 ± 0.1 h at 2 × MIC. Our study revealed that AKBA inhibited the activity of GTFs isolated from *S. mutans* resulting in the decreased formulation of glucans. At 8 μg/ml AKBA prevented ~50% glucan formation by GFTs. Furthermore, AKBA effectively inhibited and eradicated the formation of biofilm generated by *S. mutans* and *Actinomyces viscosus*, with 50% $MBIC_{50}$ and $MBRC_{50}$ from 16-32 μg/ml.

Conclusions: AKBA can be useful compound for the development of antibacterial agent against oral pathogens and it has great potential as oral care applications in mouthwash for preventing and treating oral infections. AKBA structure can be further exploited to evolve potential lead compounds in the discovery of new anti-Gram-positive and anti-biofilm agents.

Docking-Based 3d-Qsar Analyses Of Triclosan Analogs as Pfenr Inhibitors

Priyanka Shah[1] and Mohammad Imran Siddiqi[1]

[1]Molecular and Structural Biology Division, Central Drug Research Institute, Lucknow, India

The final step of type II FAS pathways, catalyzed by enoyl-ACP reductase (PfENR) enzyme in Plasmodium is a validated drug target. Statistically significant 3D-QSAR models of 53 diverse compounds were developed against PfENR to understand chemical-biological interactions governing their activities with cross-validated coefficients (q^2) of 0.64 and 0.54 and conventional coefficients (r^2) of 0.96 and 0.97, for CoMFA and CoMSIA analysis respectively. Both the models were validated by an external test set of nine compounds giving satisfactory prediction (r^2pred) of 0.534 and 0.765 for CoMFA and CoMSIA models, respectively. Final three dimensional conformations used in the QSAR studies were selected on the basis of minimum binding energy of protein-ligand complexes for each ligand. Furthermore, to compliment the QSAR studies, in silico studies were performed on native and mutant protein structures to investigate the possible role of Ala217, Met281 and Phe368 in solvent accessibility and stability of protein.

CDK-Taverna 2: an open-source cheminformatics workflow environment

Kalai Vanii Jayaseelan[1], Andreas Truszkowski[2], Stefan Neumann[3], Egon L Willighagen[4], Achim Zielesny[2]and Christoph Steinbeck[1]

[1]Chemoinformatics and Metabolism, European Bioinformatics Institute (EBI), Cambridge, UK

[2]Institute for Bioinformatics and Cheminformatics, University of Applied Sciences of Gelsenkirchen, Recklinghausen, Germany

[3]GNWI - Gesellschaft fuer naturwissenschaftliche Informatik mbH, Oer-Erkenschwick, Germany

[4]Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

The computational processing and analysis of small molecules is at heart of cheminformatics and structural bioinformatics and their application in e.g. metabolomics or drug discovery. Pipelining or workflow tools allow for the Lego(TM)-like, graphical assembly of I/O modules and algorithms into a complex workflow, which can be easily deployed, modified and tested without the hassle of implementing it into a monolithic application. Here we present,CDK-Taverna 2.0 [1] an improvised version of CDK-Taverna 1.0 [5], a free open-source cheminformatics workflow solution combining different open-source projects such as the Chemistry Development Kit (CDK) [2,3], the open-source workflow environment, Taverna [4] and the Waikato Environment for Knowledge Analysis (WEKA) [5].The CDK-Taverna 2.0 plug-in provides 192 drag and drop components for input and output (I/O) of various chemical file and line notation formats, substructure filtering, aromaticity detection, atom typing, reaction enumeration, natural-product likeness calculation [7], molecular descriptor calculation and data analysis.Example workflows are published on myexperiment.org, an open-source community for workflow sharing, and are available for free download.

Useful links:

http://www.myexperiment.org/groups/105.html

http://cdk-taverna-2.ts-concepts.de/wiki/index.php?title=Main Page

Synthesis Of A-Cyano-N-(5-Methylisoxazol-3-Yl) Substituted Cinnamides And Evaluation For *In Vitro* Antioxidant Properties

Madhavi Kuchana*

Institute of Pharmaceutical Technology, Sri PadmavatiMahilaVisvavidyalayam, (Women's University) Tirupati, A.P. India

A series of α-cyano-N-(5-methylisoxazol-3-yl) substitutedcinnamides were prepared by using 2-cyano-N-(5-methylisoxazol-3-yl)acetamide and substituted benzaldehydes as starting materials. The starting compound, 2-cyano-N-(5-methylisoxazol-3-yl) acetamide was synthesized by cyanoacetylation with mixture of cyanoacetic acid and acetic anhydride in anhydrous dioxane containing equimolar quantity of pyridine. The tile compounds were formed by the knovenagel condensation of active methylene group of 2-cyano-N-(5-methylisoxazol-3-yl) acetamide with substituted benzaldehydes in toluene containing small amounts of glacial acetic acid and piperidine. Twelve final compounds were prepared using above procedure. All the compounds were purified by recrystallization from suitable solvent and characterized by spectral and analytical methods.

The title compounds were evaluated forantioxidant properties by*in vitro* models like reduction of 1,1-diphenyl-2-picrylhydrazyl (DPPH) and Scavenging of Nitric Oxide (NO) at 100 µM concentration. Among the series, compounds with phenolic hydroxyl group on α-cyanocinnamide moiety were found to possess good anti-oxidant properties.

In-Silico Binding Energy Studies Of Natural Antimicrobial Peptides (Amp) Against *Mycobacterium Tuberculosis* H37ra

[1]Kumar Jitender, [2]Namdeo Rajnikant

[1]Beant College of engineering and technology, Punjab, [2]IBI Biosolutions Pvt. Ltd. Hryana, India

Tuberculosis (TB), one of the oldest known human diseases, and still one of the unbeaten major causes of mortality due to total drug resistance (TDR). With the two million new cases of tuberculosis per year worldwide and emergence of new strains of *Mycobacterium tuberculosis* characterized by TDR or increased virulence have made clear the pressing need for the evolution of newer and more powerful drugs.

Many natural Anti Microbial Peptide's have already explored by various independent researchers in vivo across the globe and have proved their potential role against different micobial infection. Specifically; studies with Defensins and Protegrins over Mtb H37Ra showed 86.3 to 99.0% reduction in CFU (Colony Forming Unit) by causing lesions on the surface of H37Ra.

Different reported (in total 10) AMP's were collected from NCBI sequence repository and were searched for 3D structure availability. Eight (8) surface proteins of Mtb H37Ra were also selected related to surface genes which are involved in the synthesis and formation of lipid layer of bacterium. Unavailable 3D structure's were modeled using Modeller version 9.9 and DOPE analysis were carried out to detect best stable structure of different AMP's and surface protein of Mtb H37Ra.

80 sets of partial mono flexible Protein-Protein docking were then performed over Autodock4.2. Binding affinity analysis suggests the inhibition of PROBABLE CONSERVED MCE ASSOCIATED MEMBRANE PROTEIN with Neutrophil Defensin-1 Preprotein AMP (DEFA-1). Results of this insilico affinity analysis seems promising and suggesting a new approach of using Neutrophil Defensin-1 Preprotein AMP (DEFA-1) to fight with drug resistance problem of tuberculosis bacterium H37Ra.

Effect of Montelukast and its interaction with the protective effect of rofecoxib and caffeic acid in kainic acid–induced cognitive dysfunction model

Anil Kumar*, Atish Prakash and Deeksha Pahwa

Pharmacology Division, University Institute of Pharmaceutical Sciences, Panjab University, Chandigarh, India

There is an evolving consensus that mild cognitive impairment (MCI) serves as a prodrome to Alzheimer's disease. Recent studies suggested the involvement of oxidative stress, COX-2 and LOX up-regulation in MCI. Further, newer drugs like cysteinyl leukotriene inhibitors have shown neuroprotective effect in animal models of ischemia. Thus, the present study purports to explore the potential role of montelukast (a cysteinyl leukotriene inhibitor) in concert with rofecoxib (COX-2 inhibitor) and caffeic acid (a 5-LOX inhibitor and potent antioxidant) against kainic acid induced cognitive dysfunction in rat. In the experimental protocol, kainic acid (0.4μ/2μl ACSF) was given intrahippocampally (CA3 region) to induce a condition similar to MCI since only the amnestic cognitive domain was affected. Memory performance along with the locomotor activity was measured on the days 10-14 and 1,7 and 14 respectively. The brains were isolated on the 14^{th} day; oxidative stress parameters and mitochondrial enzyme complexes were then estimated. All drug groups consisting of montelukast (0.5 and 1 mg/kg), rofecoxib (5 and 10 mg/kg) and caffeic acid (5 and 10 mg/kg) showed significant improvement in memory performance as compared to that in control (kainic acid treated). Further, two weeks of the drug treatments reversed all the biochemical indices and showed favourable mitochondrial protection. Combination treatments of montelukast with rofecoxib or caffeic acid showed significant synergism of their protective effect which was significant as compared to their effect per se. Thus the present study is unique in that it shows the positive modulation of cysteinyl leukotriene receptor inhibition on COX and LOX pathways in the control of existing neuroinflammation.

Prediction of Inhibitors against Dihydrodipicolinate Reductase using Multiple Docking Algorithms

Romika Kumari and Andrew M. Lynn

School of Computational and Integrative Sciences, Jawaharlal Nehru University, India

Mycobacterium tuberculosis is a global health threat, infecting about one third of the human population and resulting in an annual casualty of 2 million people worldwide. A majority of drugs in current clinical use for many diseases have been designed without the knowledge of the targets, perhaps because standard methodologies to identify such targets in a highthroughput fashion do not really exist. The metabolome of Mycobacterium *tuberculosis* was filtered to identify drug target candidates. Some enzymes in the metabolic pathways are standard drug targets. The enzymes of the diaminopimelate biosynthetic pathway are investigated as good potential drug targets. In mycobacteria, the diaminopimelate pathway is used to synthesize lysine and diaminopimelate from L-aspartate. Diaminopimelate is an essential component of the cell wall peptidoglycan in both Gram-positive and Gram-negative bacteria and disruption of diaminopimelate biosynthesis in mycobacteria is known to result in cell death. Enzyme dihydrodipicolinate reductase is an important enzyme of this pathway and it can be used as a potential drug target.

We describe the results of virtual screening of library of inhibitors procured by OSDD using four programs – Glide, Gold, Autodock 4.2 and Dock v 6.0 v and combine the results using an ensemble score. The energies for known inhibitors are used as positive controls.

Insilico Analysis of Influenza A Viral Genome

Brijesh Singh Yadav[1], Anup Kainthola[1], Nidhi Srivastav[1], D.P. Vashishtha[1], A.B. Bhatt[1]

[1]Department of Botany and Microbiology, H.N.B.Garhwal University, Srinagar, Uttarakhand, India

Influenza virus belong to *orthomyxoviridae* family is a special kind of virus whose sequence is available in segments (total 8 segments) not in genome. The influenza A virus genome is contained on 8 single non-paired RNA strands that code for 10 proteins. The segmented nature of genome allows for the exchange of entire genes between different viral strains when they cohabitate the same cell. In this study gene sequences of five different strains have been investigated. The first step following data retrieval is the execution of a multiple sequence alignment, and phylogenetic analysis is performed. In second step modeling of H5N1 strain of influenza A virus was modelled and validated. Afterword protein neuraminidase of avian influenza virus was considered because of its pathogenicity and docking was performed to find the suitable ligand. This observation it might be concluded that in the course of evolution, the genes underwent suitable modifications in strains H1N1, H9N2, H5N1 and H3N2 as compared to H2N2. This proves that H2N2 is less pandemic as compared to others, which are main cause of pandemic bird flu now-a-days. The 3-D structure prediction and Docking analysis further help in investigation of structural point of view.

Post Transcriptional Gene Silencing Of Sporadic Alzheimer's Disease-A Systems Biology Approach Towards Molecular Medicine

Monga[1].Isha and Sahrawat[2].Tammanna R.

[1]Monga- M.Sc (system biology & bioinformatics), Panjab University.Chandigarh, India

[2] Sahrawat- Asst. Professor, Centre for System Biology and Bioinformatics, Panjab University, Chandigarh, India

Neurodegenerative disorders (ND's) are major class of disorders for which thus far any effective small molecule drug therapy has failed to emerge (Saido, 1998). Reason behind the failure is that they target only single gene in cellular pathways. Moreover, those studies are only on rare monogenic disease variants [*APP, PSEN-1* and *PSEN-2*, *BACE-1*], which correspond to only 25 percent familial form of total, therefore not informative to the remained 75% sporadic, more common forms of neurodegenerative disease, where combinations of many genes and non-genetic determinants are thought to play crucial roles.

Thus finding the target gene for Sporadic forms of ND, simultaneous study of all genes was necessary, which was difficult using present drug strategies, who target only single gene. Therefore solution to this problem of studying genes simultaneously as a network to find the target gene was presented by systems biology approach which not only helps in studying all disease associated genes simultaneously to discover biological target, but also fed it for its down regulation by RNA Interference. The overall workflow of present research work entitled *"Post Transcriptional Gene Silencing of Sporadic Alzheimer's Disease : A Systems Biology Approach Towards Molecular Medicine"* was to understand these complex biological pathways to identify novel crucial gene with the help of tools of Systems biology,their down regulation by RNA Interference. To achieve this goal, AlzGene and GeneCards databases were screened and Virtual Library of the most significant genes responsible for AD was prepared. Various systems biology and bioinformatics tools such as STRING, CYTOSCAPE, CABIN and MCODE,KEGG, si-RNA Target finder, RNA fold, RNA cofold and DNA Pattern Finder were used to short list the most critical gene and specifically silence it using homology dependent RNA Interference. Thus "Biological pathway based si-RNA screens will not only provide an approach for interrogating biological pathways, but also decrease target validation time lines and the cost associated with target development before clinical trials.

Ligand Based Virtual Screening by Ensemble Classification

Nisha Chandran[1] , Andrew M Lynn[1], Max Kuhn[2]

[1] Jawaharlal Nehru University, New Delhi, India

[2] Non-Clinical Statistics, Pfizer Global R&D, Connecticut , USA

Introduction

The prodigious amount of biological data, which is being generated through combinatorial synthesis and High Throughput Screening, has laid the groundwork for application of Data-Mining approaches in biological computational analysis. These approaches have hence become an integral part of Cheminformatics and Pharmaceutical research, thereby also complementing the Structure Based and Ligand Based approaches in Drug Discovery in a rational manner. The study described below used data mining approaches provided by the package CARET (Classification and Regression Training), available through the Open Source Statistical software:R.

Results

CARET as of 15th January 2012 provides over 70 classification and regression models. The initial input data-set used for the study was that of MDRR (MultiDrug Resistance Reversal) agents. These are agents whose original response variable is a ratio measuring the ability of a compound to reverse a leukemia cell's resistance to adriamycin. The standard dataset present with the caret package is a set of 528 compounds (298 actives and 230 inactives) containing 342 descriptors.

The initial assessment of the separability of the classes was done using PCA (Principal Component Analysis). 342 predictors were separated into 3 components which accounted for 44.6%, 10.5% and 6.3% of the variability in the original predictors.

There were 43 predictors that were removed due to severely unbalanced distributions that could negatively affect the model fit. Mean centering and Scaling to unit variance were used in the Pre-processing methods. After pre-processing, 297 predictors remained for model building.

The Ensemble score was higher than the individual classifiers at 0.865.

abcdeFBA: An Evolvable Open-Source Alternative for Constraint Based Modeling in R

Abhilash Gangadharan[1], Neha Rohatgi[1], Andrew M. Lynn1

[1]Jawaharlal Nehru University, New Delhi, India

Constraint based modeling currently provides the only tractable means for simulation of large genome scale metabolic models. Using simplifying assumptions of steady state kinetics and maximal rates constrained by stoichiometry, it has circumvented the limitations of traditional kinetic modeling techniques using ODE's which rely on the specification of experimentally hard-to-determine kinetic parameters. The technique by itself is more than twenty years old but has recently seen an evolution from small models of metabolism used in metabolic engineering to genome scale models constructed tour de force. Flux Balance Analysis has enabled the prediction of phenotypes and the design of mutant strains for over-production of bio-products in metabolic engineering successfully.

We introduce a free, intuitive, open source package integrated with R – called abcdeFBA available on CRAN at http://cran.r-project.org/web/packages/abcdeFBA/index.html . It is hoped that A Biologist Can Do Everything of FBA with this package. Most of the common forms of analysis covered in literature related to Constraint Based Modeling have been covered in this package notably the methods in COBRA ToolBox. The most useful feature is the abstraction of some of the more complex forms of simulation into monolithic functions which take in parameters and "intelligently" perform the simulations and return viewer-friendly results. Apart from all this abcdeFBA is perfectly extendable with the addition of new functions using existing functions or even functions from other packages not utilized as a part of abcdeFBA.

Aminoglycoside induced nephrotoxicity: molecular modeling studies of calreticulin-gentamicin complex.

Pratibha Sharma, G. Hariprasad, Manoj Kumar, Punit Kaur and A. Srinivasan

Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India.

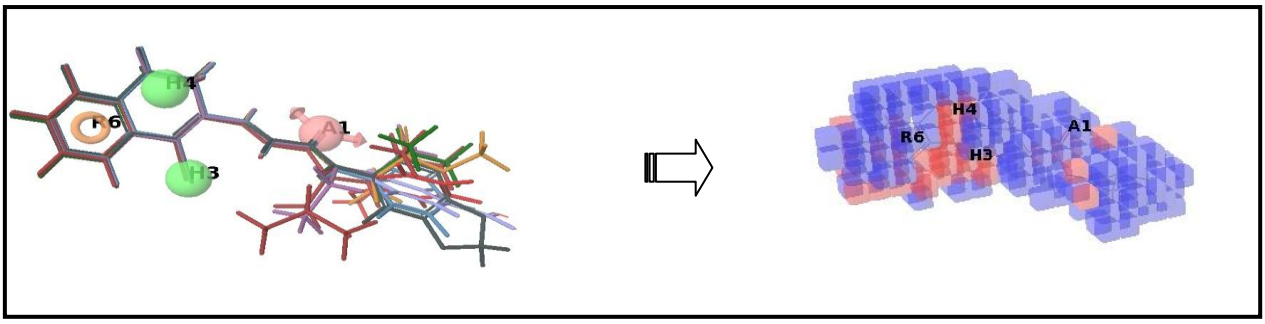
Gentamicin is a member of aminoglycoside group of broad spectrum antibiotics. It impairs protein synthesis by binding to A site of the 30S subunit of bacterial ribosomes. One of the main side effects of this drug is nephrotoxicity. The drug is known to bind to calreticulin, a chaperone essential for the folding of glycosylated proteins. We provide a detailed structural insight of the calreticulin-gentamicin complex by molecular modeling and the binding of the drug in the presence of explicit solvent was analyzed by molecular dynamics simulation. The gentamicin molecule binds to the lectin site of the calreticulin and lies in the concave channel formed by the long beta sheets. It makes interactions with residues Tyr109, Asp125, Asp135, Asp317 and Trp319 which are crucial for the chaperone function of the calreticulin. The superimposing of the modeled complex with the only available crystal structure complex of calreticulin with a tetrasaccharide (Glc (1) Man (3)) shows interesting features. First, the rings of the gentamicin occupy the positions of glucose and the first two mannose sugars of the tetrasaccharide molecule. Second, the oxygen atoms of the glycosidic linkage of these two ligands have a positional deviation of 1.3 Å. The predicted binding constant of 16.9 μM is in accordance with the previous kinetic study experiments. The details therefore, strongly implicate gentamicin as a competitive inhibitor of sugar binding with calreticulin.

Insilico* models for virtual screening, identification and lead optimization of inhibitors of putative multidrug resistant efflux pump Rv1258c of *Mycobacterium tuberculosis

Rukmankesh, Amit Nargotra, Nitin Kalia, Inshad Ali Khan, Surrinder Koul, Ram Vishwakarma

Indian Institute of Integrative Medicine (Council of Scientific and Industrial Research), Jammu, India

Drug resistance in bacteria, including *Mycobacterium tuberculosis*, has been associated with membrane-located efflux pumps that prevent cytosolic accumulation of drugs. Rv1258c is one of the efflux proteins which is an active transporter localized in the cytoplasmic membrane of the cells. Hence, there is a need to design inhibitors against this efflux pump, not only to restore the activity of the antibiotic that is being effluxed out but also to prevent the emergence of drug resistance. In the present study, structure based and ligand based approaches have been used for designing potent Rv1258c inhibitors. Since 3D-structure of Rv1258c is not available, it was modeled using several online servers, and based on the in-house experimental data available, the probable structure with the binding site has been predicted. The final model was selected based on the protein structural analysis and validation server. The predicted binding site was further analysed for fragment binding studies in order to design several virtual compounds around a selected scaffold for lead optimization and identification of potent Rv1258c inhibitors. Further a robust four point pharmacophore model was also developed based on the Rv1258c inhibitory data from in-house wet lab studies (unpublished data). This four point pharmacophore model was further used to develop an atom based QSAR model, which can be used effectively for lead optimization.



Author Index

Notes

Notes

Notes

Notes

Notes

Notes

Notes