Kolkata International School cum Conference on Systems Biology

KOLSYSBIO

December 29'th 2012 to January 3'rd 2013

Venue: Auditorium, Saha Institute of Nuclear Physics, Kolkata

Abstract Book

Saha Institute of Nuclear Physics
Kolkata, INDIA

Abstracts of Poster Presentations

KOLKATA INTERNATIONAL SCHOOL CUM CONFERENCE ON SYSTEMS BIOLOGY (KOLSYSBIO)

December 29, 2012 to January 3, 2013

CONVENORS:

Pradeep K Mohanty, SINP
Soumen Roy, Bose Institute

P01: Comparative network analysis

Sumeet Agarwal
Department of Electrical Engineering,
Indian Institute of Technology Delhi
Hauz Khas, New Delhi - 110016

Many real-world systems are naturally represented as networks, and a variety of measures exist that characterise their structure. However, studies of networks typically employ only small, partly arbitrarily selected subsets of these, and the lack of a comparison makes it unclear which structural diagnostics are redundant or complementary. We present a highly comparative study of networks and network features, analysing a wide variety of networks, derived both from empirical observation and from mathematical models. We make use of a total of over four hundred network metrics or summary statistics thereof. We demonstrate how our approach can be used to organise and classify networks, as well as to obtain insights into how network structure relates to functionally relevant characteristics in a variety of settings. These include detecting structural features of metabolic networks that correlate with biological evolution, and constructing summary statistics that allow for efficient fitting of evolutionary models to observed protein interaction networks, via a novel Approximate Bayesian algorithm. Our methodology provides a general-purpose data-driven approach to aid in the study and understanding of networked systems.

P02: Systems approaches for identification of essential genes: *Mycobacterium tuberculosis* as case study

Priyanka Baloni, Soma Ghosh, Sumanta Mukherjee, Praveen Anand and Nagasuma Chandra
Department of Biochemistry, Indian Institute of Science

A systems approach is necessary to identify essential genes in an organism. Essential genes are defined as the minimal set of genes required for the survival of an organism. In the present work, we report an integrated approach to identify a set of essential genes in the causative agent of tuberculosis, Mycobacterium tuberculosis (M.tb). Different levels of studies have been incorporated to overcome the limitations of individual methods. As a first criterion, genes which show consistently high expression values in the 39 wild-type samples are short-listed as important for the survival of the organism. Next, at the metabolic level, FBA combined with gene expression data is used to identify genes that result in lethal knock-outs. Following that, a weighted and directed protein-protein interaction network is obtained from various sources and important controlling points identified using broken path analysis. The genes are then checked for phyletic retention with the basic assumption that genes that are conserved across the species are essential for the survival of the organism. The genes identified from the different methods are further combined using set-theory to obtain a final list of 229 essential genes. These are further validated using the transposon mutagenesis dataset, showing an overall agreement of 73for the 229 genes conserved amino acid residue of its protein products are identified and mapped on the binding sites of the protein. It is seen that majority of the conserved residues fall on the binding site. Thus, a general and robust methodology has been developed for the identification of essential genes in M.tb. It should be noted that the method can be easily extrapolated to other organisms.

P03: Computational study of differential gene expression in *Bordetella pertussis*

Arnab Bandyopadhyay and Suman K Banik
Department of Chemistry, Bose Institute
93/1 A P C Road, Kolkata 700 009, India

Bacterial pathogens undergo profound physiological changes when they infect their host and require coordinated regulation of gene expression in response to the stress encountered during infection. *Bordetella pertussis*, a human pathogen and causative agent for the disease whooping cough, expresses a number of virulence factors that enable the bacteria to colonize in the respiratory tract of the host, intoxicate the host and cause disease. A number of environmental factors, such as temperature and chemical agents, are known to control the transcription of the virulence related genes, e.g., filamentous haemagglutinin, pertussis toxin, adenylate cyclase etc., in *B. pertussis*. The expression of virulence factors is coordinately regulated by the *bvg* locus, which encodes the proteins BvgA and BvgS, the sensor and the response regulator domain, respectively, of *B. pertussis*. By sequence homology these fall into the 'two component' family of bacterial signal transduction machinery. In the first step of activation, *bvg* locus activates its own autoregulated promoter and the promoter of the adherence factor filamentous haemagglutinin. The second step occurs several hours later and consist of the activation of adenylate cyclase and pertussis toxin genes.

The aforesaid phenomena can be explained in terms of reaction kinetics, as in a living cell several sub-cellular events occur as a manifestation of biochemical reactions. Considering all the biochemical reactions follow mass action kinetics we have developed a computational model for signal transduction and differential gene regulation in *B. pertussis*, that explains the existing experimental results.

P04: Key to network controllability

Soumya Jyoti Banerjee and Soumen roy
Bose Institute, 93/1 APC Roy Road,Kolkata 700009 India

Controlling complex systems has always been a challenge. Of late, there has been significant interest in controllability of complex networks. Recent studies has claimed that a network's controllability is to a great extent encoded by the underlying network's degree distribution $P(k_{in}, k_{out})$ and degree correlations. Is the controllability of a network decided almost completely by the immediate neighborhood of a node, while, even slightly distant nodes play no role at all? Motivated by the above question, we tried to define controllability metric in a more general way, using distance based measures like closeness centrality and betweenness centrality of a network. We want to show that controllability of a network may not only require the purely local connectivity measures like in-degree and out-degree. Degree reflects information about the immediate neighborhood of a node. However, closeness centrality encodes both local and global information and signifies a node's potential to choose good control paths passing through it. Again, a node with high betweenness centrality could be rather distant from the node in question but could connect it to a matched or unmatched node. Thus, we are immediately led to investigate the important role that closeness centrality and betweenness centrality should play in deciding controllability.

P05: Mutations alter protein interactions leading to loss and gain of biological functions

Mahashweta Basu, Nitai P. Bhattacharyya and P. K. Mohanty
Saha Institute of Nuclear Physics
1/AF-Bidhannagar, Kolkata-700064

We tested the hypothesis that loss/gain of interactions of mutant HTT which causes Huntingtons disease (HD), or mutant TP53 involved in various cancers, result in loss or gain of biological function(s) in respective pathological conditions. We have constructed protein-protein interaction network (PPIN) with the proteins interacting with wild type and mutant proteins and modularized the networks. Functional role of modules are assessed by Gene Ontology (GO) enrichment analysis for the biological processes (BPs). Such analyses reveal the plausible candidate BPs of the modules. The unique GO terms enriched significantly (Hyp ¡ 0.0001) with proteins in modules of mutant PPIN indicate gain of BPs due to mutation whereas the enriched GO terms in wild type PPIN represents the loss. Several other enriched GO terms common in the modules of mutant and wild type networks indicate loss as well as gain of biological functions in pathological conditions. It turns out that most of BPs assigned to the modules of these protein networks are already known to be altered in HD or cancers. We argue that gain of BPs, and thus the biological functions, are due to new interacting partners acquired by mutant proteins. The methodology we adopted here could be applied to genetic diseases where the mutation alters the ability of the protein to interact with other proteins.

P06: Formal methods for analysis of biological systems

Sukanya Basu and Supratik Chakraborty
Dept of Computer Science and Engineering, IIT Bombay

Recently, there has been an explosion of interest in applying mathematical techniques to understand system-level dynamics of biological systems, and even to predict outcomes of virtual biological experiments.There are several open questions that need be answered in order to develop a meaningful abstraction-refinement framework for biological systems. The first question concerns the discovery of an appropriate abstraction function and the corresponding concretization function. In the context of static analysis and formal verification of computer systems, abstraction and concretization functions have been traditionally been defined as pairs in a Galois connection between a concrete state space and an abstract state space. One needs to re-examine these notions in the context of biological networks. In addition, one needs to devise techniques for automatic discovery of abstractions. As an initial proposal, we wish to use ideas from predicate abstraction of programs to abstract the information represented in a biological network using a set of carefully chosen and biologically relevant predicates. An n-dimensional predicate is a mapping from an n-dimensional domain to the range True, False. Refinement of an abstraction obtained using a set of predicates has traditionally been done (in the context of program verification) by making available additional predicates for use in abstracting the concrete behaviour. In the context of biological networks, this approach can be used as well. However, the nature of biological systems offers additional opportunities for refinement. Yet another important aspect to consider is the choice of the underlying reasoning engine. How does one represent the abstraction, and analyze it to obtain meaningful sub-structures? In this context, the use of satisfiability and/or constraint solvers appears to be a promising direction to explore. We have some very preliminary evidence that this might be a fruitful line to pursue. Needless to say, this requires a significant amount of additional investigation.

We propose to develop an iterative abstraction-refinement framework for analyzing large biological networks. The goal would be to identify all sub-structures/ information content from the network that are relevant to a accuracy by a biologist. We should be able to provide guarantees that no relevant sub-structure has been missed by our analysis, and strive to ensure that as few non-relevant sub-structures as possible are returned by the analysis.

P07: Construction of a colon cancer differentiation model to identify new targets for colorectal cancer

Nirmalya Dasgupta, Bhupesh Kumar Thakur, Atri Ta, Santasabuj Das
National Institute of Cholera and Enteric Diseases,
P-33 C.I.T. Rd., Scheme XM, Beliaghata, Kolkata 700010, India.

Colorectal cancer (CRC) is the third most common form of cancer and the second leading cause of death among cancers worldwide. Sequential accumulation of mutations in specific genes, such as adenomatous polyposis coli (APC), Kirsten-ras ($K-ras$) and p53 drives the transition from healthy colonic epithelia through increasingly dysplastic adenoma to colorectal cancer. Human colon cancer-derived cell line HT-29 expresses mutant APC and p53 proteins, the commonest combination of mutations (27.1%) in colon cancers. However, HT-29 cells can be differentiated, which ultimately leads to apoptosis of the cells by butyrate, a bacterial metabolite present in the intestinal lumen. To achieve differentiation, butyrate regulates approximately 1500 genes at the transcriptional level. We constructed a Butyrate Network with these 1500 genes. Analysis of the network divided the hubs into two categories with respect to their presence in the p53-APC interactome. Substantial differences in the biochemical properties and degree of the hubs were observed between the two categories. Butyrate network was more vulnerable to the removal of the hub proteins, which were shared by the p53-APC interactome. We identified Syk, a non-receptor tyrosine kinase as a butyrate regulated hub protein present in the p53-APC interactome. Expression of Syk was associated with Breast and head and neck carcinomas. However, its role in colon carcinomas is yet to be investigated. We found that Syk may play an important role in colon cancer growth and metastasis. This association needs to be further investigated through clinical studies.

P08: Rule-based modeling of host-pathogen interactions: a focus on tuberculosis

Soma Ghosh, Prasad KVS, Saraswathi Vishveshwara, Nagasuma Chandra
Molecular Biophysics Unit, Indian Institute of Science;
Computer Science and Engineering, Chalmers University of Technology;
Department of Biochemistry, Indian Institute of Science

Iron is an essential element in most biological cells, required for the function of many proteins in diverse processes such as oxygen transport, repression, detoxification and DNA synthesis. Intricate mechanisms have evolved to tightly regulate iron metabolism both at the cellular and organismal levels. Iron overload as well as iron insufficiency are both known to lead to pathological conditions such as hemochromatosis and anemia respectively. *Mycobacterium tuberculosis (M.tb)* has formulated various means of attacking the host system to establish itself within the host system. One such crucial strategy is the exploitation of the iron resources of the host system. Both host and pathogen compete with each other to obtain and maintain the required concentration of iron through complex molecular interactions. The extent of complexity, makes it important to obtain a systems perspective of the interplay between host and pathogen with respect to iron homeostasis. We have reconstructed a systems model comprising of 66 components and 78 protein-protein or protein-metabolite interactions, captured as a set of 194 rules. A rule-based modelling approach, Kappa is used to simulate the system separately under infection and non-infection conditions. Various perturbations including knock-outs and dual perturbation are carried out to identify bottle necks in the network that can lead to decrease in the pathogenesis. The model is able to re-establish the importance of iron dependent regulator (ideR) in M.tb and transferrin (Tf) in the host. Perturbations, where iron storage is increased appear to enhance nutritional immunity and the analysis indicates its adverse effects on the host, while decreasing the rate of iron uptake by Tf is shown to be beneficial for the host. Simulation and perturbation studies help in identifying Tf as a possible drug target. Regulating mycobactin concentration is also identified as a possible strategy to control bacterial growth. The model reported in this study presents a comprehensive framework to study iron homeostasis and provide significant insights into iron homeostasis and potential drug targets for combating TB infection.

P09: Gene set control analysis (GSCA) predicts haematopoietic control mechanisms from genome-wide transcription factor binding data.

Dr. Anagha Joshi
The Roslin Institute, University of Edinburgh, UK

Transcription factors are key regulators of both normal and malignant haematopoiesis. Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) has become the method of choice to interrogate the genome-wide impact of transcription factors. Here we have collected and integrated 142 publicly available ChIP-Seq datasets for both normal and leukaemic murine blood cell types, and introduce the new bioinformatic tool Gene Set Control Analysis (GSCA). GSCA predicts likely upstream regulators for lists of genes based on statistical significance of binding event enrichment within the gene loci of a user-supplied gene set. We show that GSCA analysis of lineage-restricted gene sets reveals expected as well as previously unrecognised candidate upstream regulators. Moreover, application of GSCA to leukaemic gene sets allowed us to predict the reactivation of blood stem cell control mechanisms as a likely contributor to LMO2 driven leukaemia as well as to clarify the recent debate on the role of Myc in leukaemia stem cell transcriptional programmes. GSCA therefore provides a valuable new addition to analysing gene sets of interest, complementary to Gene Ontology and Gene Set Enrichment analyses. To facilitate access to the wider research community, we have implemented GSCA as freely accessible web tool (http://bioinformatics.cscr.cam.ac.uk/GSCA/GSCA.html).

P10: Structural and functional analysis of hypothetical proteins in Mycobacterium tuberculosis using bioinformatics tools

Lakshmi Maganti, Nanda Ghoshal
Structural Biology and Bioinformatics Division,
CSIR - Indian Institute of Chemical Biology, Kolkata- 700032.

Genome sequencing projects have led to an explosion of large amount of gene products of which many are hypothetical proteins with unknown function. Analyzing and annotating the functions of hypothetical proteins are important in Mycobacterium tuberculosis (Mtb), the causative agent of tuberculosis (TB), which remains as a serious public health threat. Genome sequencing of Mycobacterium tuberculosis holds promise for the development of new drug candidates and search for protein coding regions in the genes. However, up to 50% of genes within a genome are often labeled unknown, uncharacterized or hypothetical. The biological functions of proteins encoded by these genes are not known limiting our understanding of virulence and pathogenicity of these organisms. In this study, six hypothetical proteins of Mycobacterium tuberculosis were retrieved from swissprot and analyzed for their structural and functional characteristics by using various bioinformatics tools like CDD-BLAST, INTERPROSCAN, PFAM, STRING. The analyses revealed that some of them possessed functionally important domains, protein-protein interacting partners and belong to families, which are Ppx-GPPA phosphate, Exonuclease, CblQ-cobalt transport protein, peptidase S13, Glycerate kinase, DNA primase S. This suggests that those hypothetical proteins may have the functions of the respective families to which they belong. In this way, searching functional conserved domains and structure prediction of uncharacterized (hypothetical) proteins will make them the future drug targets by identifying their existence in the metabolic pathways of Mycobacterium tuberculosis life cycle. The structure prediction of those proteins along with identification of the binding sites have been done, which may be useful for identifying novel drug candidates against these hypothetical proteins to assist in the control of tuberculosis.

P11: Histone acetyl transferase-1 from Leishmania donovani (LdHAT1) is regulated by S-phase kinase LdCyc1-CRK3

Anup Kumar Maity and Partha Saha
Crystallography and Molecular Biology Division,
Saha Institute of Nuclear Physics
1/AF-Bidhannagar, Kolkata-700064

Protein phosphorylation is a key regulatory post-translational modification and cyclin dependent protein kinases (Cdks) play a central role in the regulation of cell cycle progression and cell division by reversibly phosphorylating different proteins in eukaryotes. Several Cdk related kinases (CRKs) and cyclins have been identified in the early branching pathogenic kinetoplastida parasites Leishmania and an S-phase cell cycle kinase LdCyc1-CRK3 from Leishmania donovani was characterized from our laboratory. The cyclin subunit LdCyc1 contains a conserved MRAIL motif, which is responsible for interaction with proteins harboring RXL type cyclin binding (Cy) motif. Hence, a screening for substrates was carried out based on the presence of Cy motif as well as Cdk phosphorylation site (S/T-P-X-R/K) in the targets resulting in the identification of three substrates of LdCyc1-CRK3. All the three substrates are phosphorylated by LdCyc1-CRK3 in a Cy-motif dependent manner. Among the three identified substrates, one is a unique protein with no known homologues. Another one contains Ku-70 related conserved domain and the third one, which we term as LdHAT1, is similar to MYST family histone acetyl transferase (HAT). In eukaryotes, histone acetylation is one of the major posttranslational modifications that remodels chromatin to regulate DNA replication and repair, gene expression, cell proliferation and terminal differentiation. We have shown that LdHAT1 acetylates histone H4 at K10 residue, and interestingly, the HAT activity is down-regulated after its phosphorylation by LdCyc1-CRK3. Therefore, the S-phase specific kinase LdCyc1-CRK3 may play an important role in the regulation of LdHAT1 activity through phosphorylation to control the cell cycle and life cycle specific chromatin dynamics in L. donovani.

P12: Characterization of functional domain in natively unfolded human proteins

Mritunjoy Maity and Nakul C. Maiti
Structural Biology and Bioinformatics Division,
Indian Institute of Chemical Biology
Council of Scientific and Industrial Research,
4, Raja S.C. Mullick Road, Kolkata 700032

Studying unstructured proteins are of growing importance as it was recently discovered that many proteins without any globular fold involve in cell signalling and pathology of neurological disorder and other diseases. We computationally defined the compositional aspects of the binding/functional region of a particular class of human proteins which poses little or no globular fold. Many of these proteins possessed multiple binding regions (BR) and the sequence length of the BRs varied from region to region. Number of residues in binding region found to increase with the protein sequence length and distributed throughout protein sequence. The regions were distributed throughout the protein sequence and not localized. Grand average of hydropathy of most of the protein was negative, however many binding regions were with + ve values. Majority of residues in the binding region showed extended/random coil conformation possibly making them adaptable to different target molecules.

P13: Interplay between Fur and HNS in controlling virulence gene expression in Salmonella typhimurium

Mahendra Kumar Prajapat and Supreet Saini

Chemical Engineering, Indian Institute of Technology Gandhinagar, VGEC Campus, Chandkheda, Ahmdeabad, Gujarat 382424, India; Chemical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra - 400067, India

Salmonella enterica is a pathogen responsible for a large number of enteric diseases in a wide-range of hosts. Fur (Ferric uptake regulator) and HNS (histone-like small nucleoid-associated protein) are two of the global regulators involved in controlling gene expression during the infection cycle of the *Salmonella*. Fur is a Fe^{2+} dependent transcriptional activator or repressor of gene expression. HNS is a DNA binding transcriptional regulator which negatively regulates gene expression by shielding DNA from transcription factors and RNA polymerase. In this work, we demonstrate computationally that Fur and HNS have disproportionately high density of binding sites in the Pathogenicity Islands on the *Salmonella* chromosome. Moreover, the frequency of binding sites for the two proteins is correlated throughout the genome of the organism. These results indicate a complex interplay between Fur and HNS in regulating cellular global behavior.

P14: Potent anticancer activity of small peptides and their physical properties

Biswadip Banerji, Sumit Kumar Pramanik,
Uttam Pal, Nakul Chandra Maiti
Department of Chemistry and
Department of Structural Biology and Bioinformatics;
CSIR-Indian Institute of Chemical Biology
4, Raja S.C. Mullick Road, Kolkata - 700032, India

Cancer is a severe threat to human society. In the scientific community worldwide cancer remains a big challenge as there are no remedies as of now. Cancer is quite complicated as it involves multiple signaling pathways and it may be caused by genetic disorders. Various natural products and synthetic molecules have been designed to prevent cell proliferation. Peptide-based anticancer drugs, however, are not explored properly. Though peptides have their inherent proteolytic instability; they could still act as anticancer agents. Here we wish to report small peptide which can act as anticancer molecule. Potent anticancer activities were confirmed by MTT assay (a laboratory test and a standard colorimetric assay, which measures changes in colour, for measuring cellular proliferation) and phase contrast images. The IC50 value of these compounds ranges in the low-micromolar level.

P15: Investigating gene co-expression modules in optimally growing Mycobacterium tuberculosis

Bhanwar Lal Puniya, Srikant Prasad Verma and Srinivasan Ramachandran
CSIR-Institute of Genomics and Integrative Biology

Biological systems are robust, which allow them to survive in changing environmental conditions. These robust traits have tendency to be selected in evolution. Modularity in the cellular organization of dynamics of expression is a one of the property of robust biological system. In this work we have constructed weighted gene co-expression network from gene expression data of mid log phase of optimally growing different strains of Mycobacterium tuberculosis. We have used 1068 genes, which have statistically significant variation in gene expression at least between one pair of strains. We have identified 7 co-expressed modules out of which 4 modules are in good agreement with biological enrichment analysis. Turquoise module was enriched with fatty acid biosynthesis, DNA replication and non membrane bound intracellular organelles. The blue module was enriched with two component system. Blue module has NADH dehydrogenase and Cell membrane. The Black module was containing mostly Transposons and transposition related genes. These modules were used for investigating highly weighted connections. These highly weighted connections were filtered by using weight threshold of 0.08. We have compared these connected gene pairs with other sources such as STRING database to identify the genomic neighborhood, gene fusion, conserved co-expression and phylogenetic co-occurrences. To identify the common transcriptional regulation between genes we have compared these with recently published Transcriptional Regulatory Network. We have found 202 connections out of total 1642 connections were in agreement. We have used these filtered network modules for identification of coordinated biological functions and hubs which could be attractive candidates for drug target identification. We identified Rv1611 (trpC), Rv0545c (pitA) and Rv3404c as top hubs. We found 39 high confidence drug targets mapped on these modules. Top 3 hubs were already predicted as high confidence drug targets. We have identified 7 novel candidates for drug targets which have degree greater than average degree of mapped high confidence drug targets and lacks homology with human proteins. Using this information we have proposed Rv0996, pknA, fgd2, lpqL, murC and phos1 as novel candidates for drug targets.

P16: Use of structural bioinformatics to deduce novel functions of proteins

Amit Kumar Singh Gautam, Prasanna lab, ACTREC

The primary structure of protein defines its structure, function and localization. Short amino acid sequences play an important role in signaling and carrying localization or retention signals. We hypothesize that solvent exposed short sequences may play central role in protein-protein interaction and can bring together proteins of very different functions in a system wide network. We have tested this principle under different biological conditions-
1) In order to globally profile natural substrate of human endo-proteases, we have designed a novel method. We used minimal protease recognition sequence (4 residues) to fish out the putative substrates and filter them based on whether the recognition sequence is exposed to solvent (and hence accessible for protease) and whether protease and substrate are in the same cellular compartment. We have developed an online server to predict natural endo-protease substrate with high confidence (PNSAS). As a proof of principle, we identified several novel substrates for a serine protease matriptase, elevated level of which has been associated with several diseases including cancer. One of the novel putative matriptase substrate was desmogleine-2 (Dsg-2).
2) We designed tetra amino acid peptide from C-terminus of proteins and screened for binding with PSMD9. PSMD9 is a non-ATPases shuttling proteasome subunit containing PDZ like domain. The function of PSMD9 is not well understood; with the help of Cterminal short peptides we identified some novel binding partners of PSMD9. By pull down and immunoprecipitation experiments we showed that protein containing short sequence also interact with PSMD9 and mutation or deletion of these residues abrogate binding. Using classical binding studies, homology based modeling, structure based analysis of binding sites and site directed mutagenesis, we have provided the first glimpse of the probable structure of PDZ-like domain of PSMD9, its binding pockets and their role in protein-protein recognition.
3) Oncogene PSMD10 interacts with one of ATPases subunit of proteasome. The EEVD sequence was essential for this interaction. Since PSMD10 is an oncogene it will rewire cellular networks. It, therefore, becomes very important to identify the binding partners. We have identified several novel PSMD10 binding partners by searching the EEXD sequences in proteome and filtering them based on solvent exposure. Many of these interactions have been proved by immunoprecipitation and mutagenesis.

P17: Bringing Down the House: A Non-linear Pay-off Driven Growth Model for Remediation of Antibiotic Resistant Infection by Introduction of Cheats

Sanhita Ray and Anjan Kr. Dasgupta
Department of Biochemistry
University of Calcutta
Kolkata 700019

Microbial cheats may serve as a useful weapon in the war against antibiotic resistant infections. Cheaters that do not produce public goods (e.g. beta-lactamase) required for conferring antibiotic resistance may destabilize a resistant population when introduced into the same niche. However, the conditions under which such remediation would occur can only be understood with the help of realistic dynamical models. We propose a model wherein a pay-off term having non-linear dependence on worker frequency decreases the death rate attributable to antibiotics. Simulations and analytical results show: if cheaters have a critical advantage in terms of growth rate, even low antibiotic dosages are sufficient for bringing down a resistant infection. We also deal with the question of how transference of resistance genes to cheaters may affect results.

P18: Towards cataloguing peptide-mediated protein-protein interactions

Debasree Sarkar, Tanmoy Jana and Sudipto Saha
Bose Institute
Kolkata-700054

Protein Protein Interactions (PPIs) are known to be regulated by linear motifs (LMs), which are short elements embedded within large protein sequences. These short LMs are often present in the disordered regions in the eukaryotic proteins and are responsible for many protein-protein interactions. LMs plays a vital role in many cell signalling pathways including Wnt , MAPK and TGF in forming peptide mediated PPIs and complexes. To better understand the role of the LMs in pathways involved in oncoproteins including p53, RAS, MYC and APC, we focus on compiling and curating LMs using text mining followed by manual curation. We are developing a database with an aim to curate the LMs to avoid false positives by using filtering logic such as surface accessibility, structure, cellular compartment and orthogonal search in different species. Targeting protein networks using LMs have a potential in developing drug targets.

P19: Stochastic optimization based study of dimerization kinetics

Srijeeta Talukder, Shrabani Sen, Ralf Metzler, Suman K Banik, Pinaki Chaudhury ; Department of Chemistry, University of Calcutta; Institute for Physics & Astronomy, University of Potsdam; Department of Chemistry, Bose Institute

The evaluation of correct individual step based pathways (both the nature of the reaction as well as the correct magnitude of the rate constant) present in a multi-step reaction scheme is central in establishing a complete reaction model in chemical and biological reaction networks. The conventional way to study reaction kinetics is to write down the mean field rate equations for the process, integrate them and follow the variation in the concentration of each species involved as a function of time. However this strategy is too simplistic and fails in situations where the number of reacting species is small (e.g. biochemical reaction in living cell), since for small number of particles fluctuations in the species population become relevant. In the conventional rate equations approach, it is assumed that the process is both continuous and deterministic. Reaction kinetics at such low concentrations are intrinsically discrete and stochastic. The stochastic simulation algorithm (SSA) is an elegant formulation to incorporate these effects and predict cor- rect results in a complicated multi-step reaction network. The rate constants associated with each individual step in a multi-step reaction scheme might not always be known a priori, or there might be a range of values of the rate constants, for which predictions for the overall reaction are compatible with experimental data. The correct prediction of all individual rate constants is not always an easy task and involves an optimization process. If an optimization scheme can be linked to SSA, then it should be possible to evaluate a correct set of reaction parameters, quantifying the complete kinetic behavior of a reaction network. We investigate the potential of numerical algorithms to decipher the kinetic parameters involved in a multi-step chemical reaction. To this end we study a dimerization kinetics of protein as a model system. We follow the dimerization kinetics using a stochastic simulation algorithm and combine it with three different optimization techniques (Genetic Algorithm, Simulated Annealing and Parallel Tempering) to obtain the rate constants involved in each reaction step. We find good convergence of the numerical scheme to the rate constants of the process. We also perform a sensitivity test on the reaction kinetic parameters to see the relative effects of the parameters for the associated profile of the monomer/dimer distribution.

P20: Drug targets in Leishmania spp

Barnali Waugh and Rahul Banerjee
Crystallography and Molecular Biology Division
Saha Institute of Nuclear Physics
1/AF, BidhanNagar, Kolkata-64

Leishmania spp. are the causative agents of a wide spectrum of diseases referred to as leishmaniasis. Of these, visceral leishmaniasis can prove to be fatal unless treated properly in time. Several thousands of patients are afflicted by these parasites especially amongst poorer sections of the populace in tropical/subtropical regions of third world countries. Till date, there is no vaccine for this disease and only a few drugs are available which have toxic side effects, in addition strains resistant to these drugs have also appeared.
Experimental approach:
Several proteins from the glycolytic pathway are unique to the parasite with respect to human and thus could qualify as valid drug targets. Specifically the protein phosphofructokinase has been validated as a drug target and its inhibition would be lethal to the parasite. The clones of four enzymes from the glycolytic pathway or related to it have been obtained (phosphofructokinase [PFK], fructose1,6-bisphosphatse[FBP],phosphoglyceratekinase[PGK], phosphoglucomutase[PGM]) and with the exception of PFK have been successfully expressed. Of these four proteins, PGK has been purified to homogeneity. Biophysical studies on these proteins are envisaged in the near future. In addition, crystallization trials on Aldose reductase(ALR)) from L. donovani , yielded crystals which were found to diffract poorly. Presently the crystallization trials for improving the quality of the crystal are being carried out.
Computational approach:
Since the enzyme PFK is unique to the parasite and is a validated drug target a pharmacophore is being designed based on the structure of its inhibitors as reported in the literature. Initially two models of PFK were constructed using the crystal structures of PFK (in apo-form: 2HIG) and the same protein ligated to ATP (3F5M) from T. Brucei as templates. About thirty confirmed inhibitors with relevant IC50 values were selected from the literature, and their atomic models constructed using locally developed software, and subjected to several cycles of energy minimization. These molecules will now be docked onto the enzyme subsequently leading to the design of an effective pharmacophore.