International Research Journal of Biochemistry and Bioinformatics Vol. 1(3) pp.066-070, April 2011
Available online http://www.interesjournals.org/IRJBB

Full Length Research Paper

Graph theoretic approach for metabolism disruption and developing a drug targeting methodology for the cure of tuberculosis.

Veeky Baths* Utpal Roy*

*Department of Biological Sciences, Birla Institute of Technology and Science (BITS) Pilani K K BIRLA Goa Campus, Goa 403 726, India

Accepted 20 March, 2011

Various networks, such as transcriptional, gene regulatory, metabolic or protein-protein interaction networks of various organisms have been studied, which provide insights into metabolism and regulation. Here, we have attempted to construct the metabolic network of *Mycobacterium tuberculosis* wherein the fatty acid biosynthesis pathway was chosen for analysis of the potential drug targeting. The metabolic network was constructed based on the KEGG LIGAND database, followed by graph spectral analysis of the network to identify hubs as well as sub-clustering of the reactions. Analysis of the eigen values and spectrum of the normalized laplacian matrix of the reaction pathway indicate the enzyme catalyzing ACP – Acetyl ACP may considered as a potential drug target.

Keywords: Acyl carrier protein; Eigen values; *Mycobacterium tuberculosis;* mycolic acid pathway; spectral graph analysis

INTRODUCTION

Mycobacterium tuberculosis is a pathogenic bacterium and the causative agent of most cases of tuberculosis. The physiology of *Mycobacterium tuberculosis* is highly aerobic and requires high levels of oxygen. Primarily a pathogen of the mammalian respiratory system, *Mycobacterium* infects the lungs and is the causative agent of tuberculosis. Given the intrinsic robustness mechanisms in the bacterial cell, it is no surprise that the bacterium finds new ways of overcoming the problem and developing resistance or tolerance to these drugs through activation of alternate pathways, or through manipulation of the drug or its bioavailability of the target itself. Several front-line drugs used for treating tuberculosis actually inhibit mycolic acid synthesis. Understanding the biochemical pathway that synthesizes these compounds is therefore of great interest. Availability of the genome sequence and various computational methods enable us to study pathways as whole functional units, rather than having to infer from the study of individual proteins [Raman et al. 2005, Cole et al. 1998]. The mycolic acid pathway (MAP) has been studied with great interest, and a large amount of biochemical and genetic information is available in the literature [Sassetti et al. 2003]. It is possible to exploit these large volumes of data to construct an *in silico* model of the pathway, which can then be simulated and analyzed. Constructing such models forms an important step in understanding the underlying molecular mechanisms of disease, and facilitates rational approaches to drug design. Several computational methods have emerged in recent years to simulate biochemical models, which aid in the systems approach to understanding pathways, processes, and whole cell metabolism [Raman et al. 2005, Verkhedkar et al. 2007, Bollobas et al. 2002].
Here, we present a comprehensive identification of the strategic point and components of the mycolic acid pathway and represent it mathematically based on reaction stoichiometry.

*Corresponding author Email: veeky_baths@yahoo.co.in

Spectral Graph Theory

Node-degree and the Adjacency Matrix

For an undirected graph G, we shall write deg (u) for the degree of a node u in V (G). This is simply the total number of edges at u. For the graphs we shall consider, this is equal to the number of neighbors of u,
Deg (u) = |N (u)|

Then the adjacency matrix, A, of G is given by
A_{ij} = 1 if v_iv_j belongs to E(G)
= 0 if v_iv_j doesn't belong to E(G)
Thus, the adjacency matrix of an undirected graph is symmetric while this need not be the case for a directed graph [Baths et al. 2009, Patra and Vishveshwara 2000, Ma and Zeng 2003b, Ma and Zeng 2003a].

Diagonal Matrix

In linear algebra, a diagonal matrix is a square matrix in which the entries outside the main diagonal (↘) are all zero. The diagonal entries themselves may or may not be zero [Baths et al. 2009, Patra and Vishveshwara 2000, Ma and Zeng 2009, Ma and Zeng 2003a]. Thus, the matrix D = ($d_{i,j}$) with n columns and n rows is diagonal if:
D_{ij} = 0, if i != j, for all i,j = {1 to n}

Laplacian Matrix

Given a graph *G* with *n* vertices (without loops or multiple edges), its Laplacian matrix

$$L := (\ell_{i,j})_{n\times n}$$

is defined as:

$$\ell_{i,j} := \begin{cases} \deg(v_i) & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise.} \end{cases}$$

which indicates the difference in the degree matrix and the adjacency matrix of the graph. In the case of directed graphs, either the in degree or the out degree might be used, depending on the application [Baths et al. 2009, Patra and Vishveshwara 2000, Ma and Zeng 2009, Ma and Zeng 2003a, Ma and Zeng 2003a]

Eigen Values

Given a linear transformation *A*, a non-zero vector **x** is defined to be an *eigen vector* of the transformation if it satisfies the eigen value equation $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. In this situation, the scalar λ is called an *eigen value* of *A* corresponding to the eigenvector **x** [Bollobas and Riordan 2002, Hu 2005, Ma and Zeng 2003a, Que´mard 1995].

MATERIALS AND METHODS

Graph spectral Analysis

Graph spectral analysis was performed to find and analysis spectra (Eigen values and Eigen vector components) of nodes in the graph. Such an analysis provided information on the overall structure and topology of the graph. To obtain Eigen value spectra of the graph, the adjacency matrix is converted to a Laplacian matrix L, by the equation: **L=D-A**, where, **D** being the degree matrix of the graph, is the diagonal matrix in which the ith element of the diagonal is equal to the number of connections that the i^{th} node makes in the graph digitalization of the Laplacian matrix yielding the spectra of the graph comprising the Eigen values and corresponding Eigen vectors.
Graph spectral theory serves as a tool useful for analysing the topological structure and organisation of large complex networks. This technique yields information about the sub-clustering of nodes in the network and identifies the cluster centres by a single numeric computation. Analysis of sub-clusters of the mycobacterial reaction networks detected by this method suggests that modularity of metabolic networks is possibly less well-defined at the level of biochemical reactions; clusters have been discernedwell from metabolite networks [Ravasz et al. 2002].

Bacterial strain used

Mycobacterium tuberculosis genomic H37rv strain
The genome contains 250 genes involved in fatty acid metabolism, with 39 of these involved in the polyketide metabolism generating the waxy coat. Such a large number of conserved genes show the evolutionary importance of the waxy coat to pathogen survival [Cole et al. 1998].]. Two clustered gene families that encode acidic glycine rich proteins take up 10% of the coding capacity. These proteins have a conserved N-terminal motif, deletion of which impairs growth in macrophages and granulomas [Cole et al. 1998].

Softwares used

VisANT was used for analyzing networks of the pathways. Given user-defined sets of interactions or groupings between genes or proteins, VisANT provides supporting function and annotation data for different genomes from the Gene Ontology and KEGG databases [Van Helden et al. 2002]. MATLAB was used for calculating Eigen values.

Database used

KEGG (Kyoto Encyclopedia of Genes and Genomes)
We have used the KEGG database to reconstruct the reaction networks of *Mycobacterium tuberculosis genomic H37rv strain.* A list of metabolic pathways and their constituent biochemical reactions were downloaded as flat files. These files contain

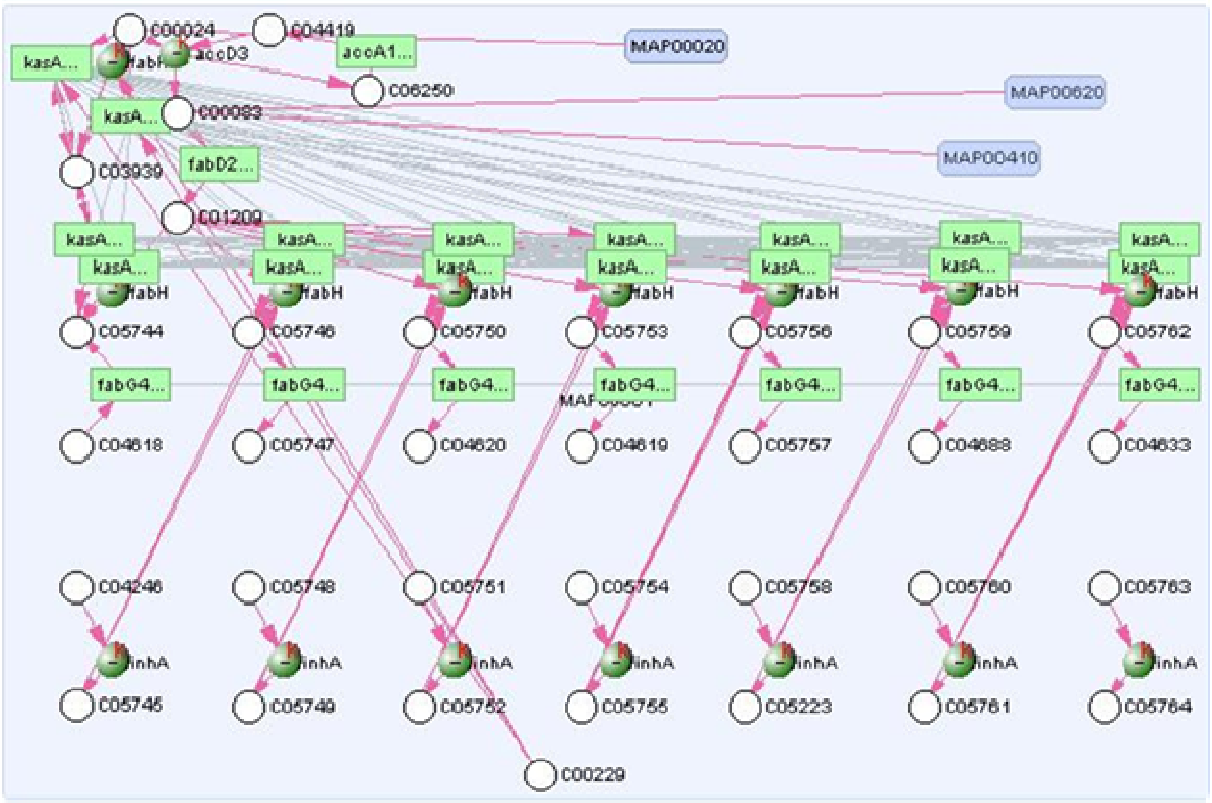

Figure 1. Fatty acid synthesis pathway of *Mycobacterium tuberculosis* as generated by VisANT.

information about reactants, products, and reversibility and steady state stoichiometries of biochemical reactions [Papin et al. 2004].

Metabolic Network Reconstruction:

The metabolic network MAP00061 of the strain was reconstructed as follows. Each metabolite of the network is a node, and reactions are the edges. Enzymes that catalyze these reactions are the potential drug targets. All edges have an equal weight of 1. In order to make the network amenable to network analysis, it is represented in the form of adjacency matrix or reaction-interaction matrix (RIM), which is an nxn matrix; n being the number of nodes (biochemical reactions) in the graph .The elements of Aij of the RIM have values according to the following rules:

$\mathbf{A_{ij}}$ = 1 if V_iV_j belong to the set of edges

= 0 if V_iV_j do not belong to the set of edges

To construct the RIM, the set of reactions in the flat file representing the metabolome was first represented as stoichiometric matrix S(mxn), with every metabolite being represented by a row and every reaction by a column [Golub et al. 2000, Kremer et al. 2002, Marrakchi et al. 2002].

RESULTS AND DISCUSSION

Tuberculosis continues to be a major health challenge, warranting the need for newer strategies for therapeutic intervention and newer approaches to discover them. Some issues that need to be addressed specifically are to increase efficiency rates of bacterial clearance, so as to minimise both treatment time and persistence. One way of achieving that could be by significant disruption of mycobacterial metabolism. This, however, would have to be done efficiently using minimal points of attack for any practical application in drug discovery.

The metabolic network for the chosen pathway was reconstructed as described earlier in the previous section and analyzed using VisANT software. [http://visant.bu.edu/] VisANT creates a network of the given pathway where each metabolite was treated as node and reaction as an edge. The adjacency and the diagonal matrix for the network were constructed as described earlier using Figure 1. It was formed in a 53x53 matrix where each cell (i,j) represents a link between the i^{th} and j^{th} metabolite. If (i,j) = 1; it means that upon enzyme action, the i^{th} metabolite leads to the production of the j^{th} metabolite.

Followed by the adjacency matrix and diagonal matrix reconstruction, Laplacian matrix was found using: L = D – A. L was then normalized in MATLAB using: **L' =** [1/sqrt (D)]*L*[1/sqrt(D)]. The maximum Eigen value calculated by the MATLAB is the Spectral Radius

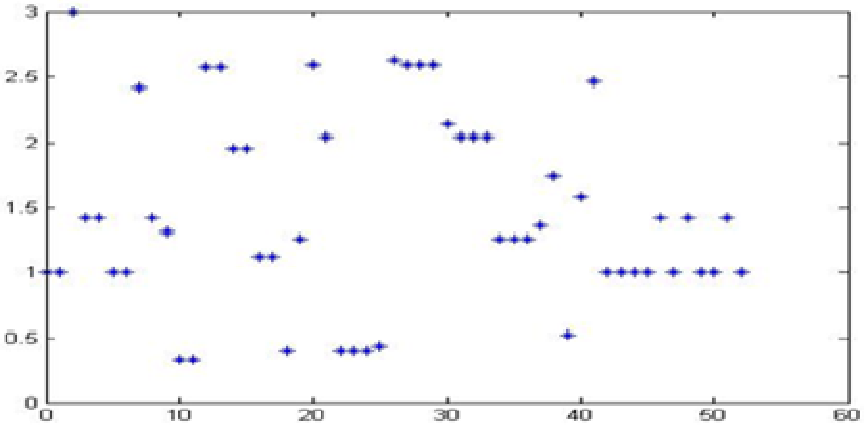

Figure 2. Eigen Values on a scatter plot. We can clearly see that node 3 i.e. the metabolite ACP has the highest Eigen value and defines the spectral radius.

which is equal to Maximum Eigen value 3.

According to previous studies carried out [5], it was particularly interesting to note that the first reaction of the FAS-I cycle (Fatty acid synthase) of the mycolic acid pathway unique to *mycobacteria* ranks as the 38th unique hub in mycobacterium. Thus, the enzyme acyl carrier protein-fatty acid synthase (ACP-FAS) involved in this reaction can be explored as a potential drug target against mycobacteria**.** Independent studies by carried out earlier have also identified the FAS enzyme as one of the putative anti-tubercular drug targets [Raman et al. 2005]

The metabolite – ACP identified by our study of the Fatty Acid Pathway of Mycobacterium, using the spectral radius, undergoes a reaction catalyzed by the enzyme ACP –FAS which could be the potential drug target. The results are in accordance with the studies previously carried out [Ravasz et al. 2002].

CONCLUSION

Upon reconstructing the metabolic network for FAS (Fatty Acid Synthase), and on further carrying out the spectral analysis for the same, the enzyme acyl carrier protein-fatty acid synthase (ACP-FAS) [Ravasz et al. 2002] is identified as a potential drug target against *Mycobacterium tuberculosis* and this has been demonstrated in Figure 2, We can clearly see that node 3 i.e. the metabolite ACP has the highest Eigen value and defines the spectral radius. Such essential metabolites can be good targets for drug designing, and they can serve as strategic point to combat tuberculosis which has been identified in previous studies. Locating the protein which affects the maximum number of proteins in a given network will help for finding the intervention strategies against tuberculosis.

REFERENCES

Raman K, Rajagopalan P, Chandra N(2005) Flux Balance Analysis of Mycolic Acid Pathway: Targets for Anti-Tubercular Drugs. PLoS Comput Biol 1(5): e46. doi:10.1371/journal.pcbi.0010046

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.

Sassetti CM, Boyd DH, Rubin EJ (2003). Genes required for mycobacterium growth defined by high-density mutagenesis. Mol. Microbiol. 48: 77–84.

Verkhedkar KD, Raman K, Chandra NR, Vishveshwara S(2007) Metabolome Based Reaction Graphs of *M. tuberculosis* and *M. leprae*: A Comparative Network Analysis. PLoS ONE 2(9): e881. doi:10.1371/journal.pone.0000881

Bollobas B, Riordan O (2002). Handbook of Graphs and Networks, chapter "Mathematical results on scale-free graphs". Wiley.

Baths V, Rohit K, Praneeth GVR, Roy, U. (2009) Graph Theoretic Approach on Metabolomic Networks of Mycobacterial Strains for Potential Drug Targets. Research Journal of Microbiology. Page No.: 132-137

Patra SM, Vishveshwara S (2000) Backbone cluster identification in proteins by a graph theoretical method. Biophys Chem , 14;84(1):13-25.

Ma HW, Zeng AP(2003b) The connectivity structure, gaint strong component and centrality of metabolomic networks. Bioinformatics. 1423-1430.

Vishweshwara S, Bindra KV, Kanna N (2002) Protien structure : Insights from graph theory .J th Comp Chem I , 187-211

Hu Z (2005). VisANT: data integrating visual framework for biological networks and modules. Nucleic Acid Res. (2005) 33:W352–W357.

http://www.genome.ad.jp/kegg

Ma, H.W., Zeng, A.P. (2003a) Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics 19: 270–277.

Van Helden J, Wernisch L, Gilbert D, Wodak S(2002) Graph-based

analysis of metabolic networks. In: HW Mewes, H Seidel, B Weiss, eds. Bioinformatics and Genome Analysis, volume 38 of Ernst Schering Research Foundation Workshop Springer Verlag. pp 245–274.

Papin JA, Reed JL, Palsson BO(2004). Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. Trends Biochem Sci 29: 641–647.

Golub Gene F, van der Vorst , Henk A (2000), "Eigenvalue computation in the 20th century", Journal of Computational and Applied Mathematics 123: 35–65, doi:10.1016/S0377-0427(00)00413-1 .

Kremer L, Dover L, Carre` re S, Nampoothiri K, Lesjean S(2002) Mycolic acid biosynthesis and enzymic characterization of the betaketoacyl- ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis*. Biochem J 364: 423–430.

Marrakchi H, Ducasse S, Labesse G, Montrozier H, Margeat E (2002) MabA (FabG1), a Mycobacterium tuberculosis protein involved in the long chain fatty acid elongation system FAS-II. Microbiology 148: 951–960.

Que´mard A, Sacchettini JC, Dessen A, Vilcheze C, Bittman R(1995) Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. Biochemistry 34: 8235–8241.

Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Baraba´ si AL (2002) Hierarchical organization of modularity in metabolic networks. Science 297: 1551–1555.

http://visant.bu.edu/