BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 6(9)

Identification of membrane associated drug targets in *Borrelia burgdorferi* ZS7- subtractive genomics approach

Shivkumar Madagi[1], Vijayakumari Mali Patil[1], Saremy Sadegh[2], Abhishek Kumar Singh[3], Bhavana Garwal[3], Atreyi Banerjee[4], Usha Talambedu[5], Biplab Bhattacharjee[6]*

[1]DBT_BIF, Karnataka State Women University, Bijapur, India; [2]Brindavan College, Bangalore, India; [3]Amity university, Lucknow, India; [4]Sri Ramswaroop college of Engineering, Lucknow, India; [5]Maharani Lakshmi Ammani College for Women, Bangalore - 560012, India; [6]Institute of Computational Biology, Bangalore-560002, India; Biplab Bhattacharjee - Email: biplabbhattacharjee2010@gmail.com; Phone: +91 -9945532797; *Corresponding Author

Received July 11, 2011; Accepted July 13, 2011; Published July 19, 2011

Abstract:
Lyme disease is an infectious disease caused by a spirochete *Borrelia burgdorferi* ZS7. This spirochete is most often spread by ticks. Single antibiotic therapy is sufficient for containment of the early stage progression of the disease but combinational therapy is more preferred in later stages. Research is in progress for the development of drugs against the pathogen, but till date no vaccines have been developed to effect the late stage infections. There is a rapid rise in the cases of antibiotic-resistant population which is more than 10% of the total infected individuals. In such condition vaccine becomes the sole alternative for prevention. Therefore effective treatment includes antibiotic combination and combination of antigenic surfaces (for vaccine preparation). Thus, a comprehensive list of drug targets unique to the microorganisms is often necessary. Availability of *Borrelia burgdorferi* ZS7 proteome has enabled *insilico* analysis of protein sequences for the identification of drug targets and vaccine targets. In this study, 272 essential proteins were identified out of which 42 proteins were unique to the microorganism. The study identified 15 membrane localized drug targets. Amongst these 15, molecular modeling and structure validation of the five membrane localized drug target proteins could only be achieved because of the low sequence identity of the remaining proteins with RCSB structures. These 3D structures can be further characterized by *invitro* and *invivo* studies for the development of novel vaccine epitopes and novel antibiotic therapy against *Borrelia burgdorferi*.

Keywords: *Borrelia burgdorferi*, Lyme disease, *Insilico*, Homology modeling, subtractive genomics.

Background:
Despite rapid advances in the diagnosis of bacterial infections and the availability of effective antibiotics, Lyme disease, the most common vector borne disease continues to represent a substantial public health problem for most countries like United States and Europe **[1]**. *Borrelia burgdorferi*, the causative organism is a gram negative, fastidious spirochaete, an obligate internal parasite that is vectored between the vertebrate hosts by hard bodied ticks of lxodes ricinis complex **[2]**. It remains in zoonotic cycle for its survival in nature since neither the vertebrates nor the vectors can transmit B.burgdorferi from one generation to another **[3]**. In most cases, infection is marked by erythema migrans. Other signs of early infection are nonspecific and include fever, myalgias, arthralgias, and fatigue **[4]**. It is a helical-shaped bacterium with outer and inner membrane and flexible cell wall. The outer membrane proteins are key components in transmission of bacteria **[5]**. Common antibiotic therapy includes tetracycline for the early manifestation and high doses of erythromycin and penicillin for children and pregnant women **[6]**. Prevention and cure can be accomplished by vaccine preparation involving antigenic surfaces that trigger humoral immune response which has long been found to be protective against *Borrelia burgdorferi* infections **[7]**. In the present study, *insilico* subtractive genomics approach has been carried out to identify non homologous essential membrane proteins of *Borrelia burgdorferi* ZS7. These membrane associated proteins could be effective candidates for new vaccine epitope design and for screening of potential anti-borellioses drugs.

Methodology:
The essential proteins of *Borrelia burgdorferi* ZS7 were identified using subtractive genomics approach and further analysed for identification of potential drug targets **[8]**. The flow chart describing the detailed methodology for identification of pathogen specific essential proteins using subtractive genomics approach is shown in **Figure 1(a)**. The complete proteome of *Borrelia burgdorferi* ZS7 was retrieved from NCBI **[9]**. Each protein sequence of the bacteria was subjected to BLAST-P, specifically against DEG (Database of Essential Genes) available at http://tubic.tju.edu.cn/deg/ for the screening of essential genes **[10]**. Cut off for E-value, bit score and percentage of identity at amino acid level were considered <E-10, >100 and >35% respectively. Matches of length <100 amino acids were not considered for further studies.

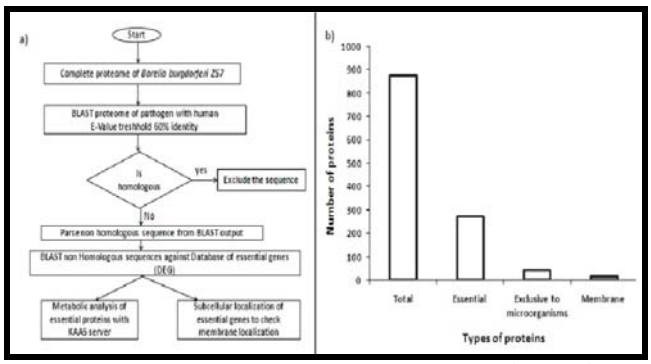


Figure 1(a): Flow chart describing the detailed methodology for identification of pathogen specific essential proteins using subtractive genomics approach; **(b)** Comparative Bar Chart between total number of proteins and membrane associated essential proteins.

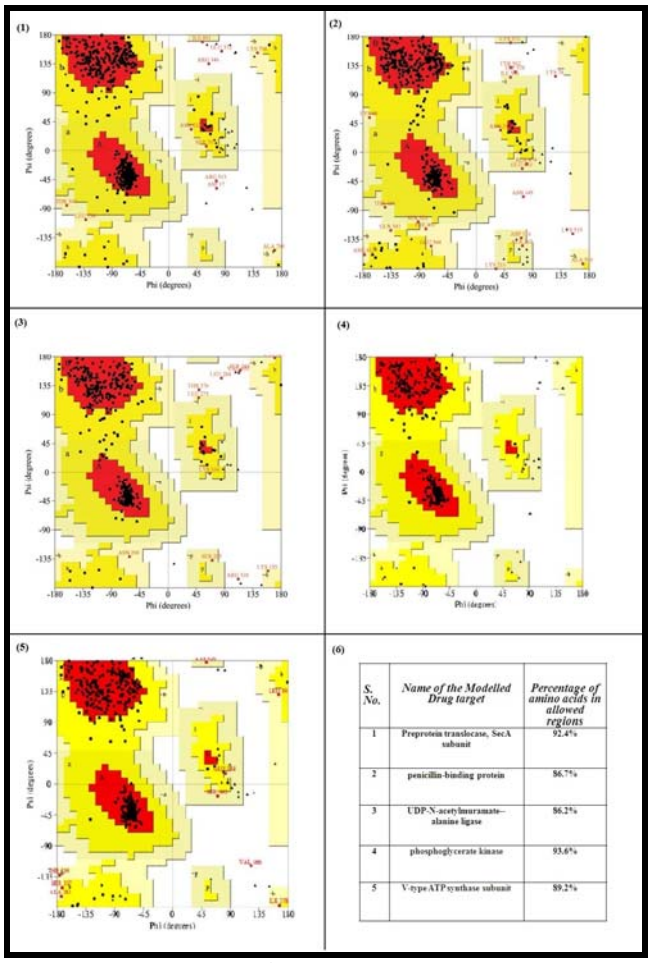
Selection of non homologue essential gene:
The essential proteins were subjected to human BLASTp in NCBI server. Sequences that did not show similarity with any human sequences were selected.

Identification of membrane localized candidate drug targets:
Sub-cellular localization of each identified non homologue essential gene was determined by Specialized Sub Cellular Localization Server v2.5 (PA-SUB) **[11]**.

Metabolic Pathway Analysis:
Metabolic pathway analysis of the essential proteins of *Borrelia burgdorferi* ZS7 was done by KEGG Automatic Annotation Server (KAAS) **[12]**. Comparative analysis of the metabolic pathways of the host and pathogen was performed by using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database for identification of pathogen specific metabolic pathway associated proteins.

Homology modeling for 3D structure of putative drug targets:
Modeling of the target proteins were performed using MODELLER Version 9.9. A template search has been performed through BLAST and PSI-BLAST programs **[13]**. Global alignment method was used for comparison between the target-template sequences **[14]**. Gaps with variable gap penalty function are included for structural loops and core regions, in order to get maximum correspondence between the sequences. Alignment file for MODELLER was prepared by CLUSTALW **[15]**. Fold recognition was done through GenThreader and LOMETS server for fold assignment **[16]**. Energy minimization of generated 3D-model was done through GROMACS (OPLS force field) by using Steepest Descent and Conjugate Gradient Algorithms **[17]**. Parameters like covalent bond distances and angles, stereochemical validation, atom nomenclature were validated using PROCHECK and overall quality factor of non-bonded interactions between different atoms types were measured by ERRAT program **[18]**. RMSD (root-mean-square deviation) and RMSF (Root Mean Square Fluctuation) was calculated for modeled structures. Functionally important residues (Active-site) were identified through comparative result of POCKETFINDER and SURFACE RACER 4.0. The generated 3D model of target proteins was checked by Ramachandran plot through PROCHECK program.

Result and Discussion:
A subtractive genomics approach utilizes the whole proteome of host and pathogen to identify proteins exclusively present in the pathogen by deducing the homologous proteins **[19]**. Based on the similar approach as mentioned in materials and methods, 272 essential genes in *Borrelia burgdorferi* genome were found to be non homologous to human (putative drug targets). **Figure 1(b)** shows the comparative analysis between the total number of proteins to the number of essential & membrane bound proteins respectively. These non human homologous targets identified were further analysed for sub-cellular localization prediction which detected 15 membrane localized proteins. Membrane localized proteins represent largest group (70%) of effective drug targets in any organism **[20]** and can also act as potential epitopes for vaccine design. **Table 1 (see Supplementary material)** gives the details of function, cellular localization and KEGG biochemical pathways of membrane localized proteins of *Borrelia burgdorferi* ZS7. Modeling was a tedious task due to very low sequence similarity and coverage and only five models could be obtained. Three dimensional model of drug targets were generated through identified templates along with fold fitting. Fold recognization was done through GenThreader and LOMETS server for fold assignment. Helices have dominance over other secondary structure. Ramachandran plot of all the five proteins analysed using PROCHECK with the description of the allowed regions are shown in **Figure 2**. Our results yielded kinases namely phosphoglycerate kinase, chemotaxis histidine kinase, that are likely to be involved in cell cycle control, differentiation and response to stress during their complex life cycle. Kinases have been reported to be the drug targets **[21]**. In our study three of the proteins are penicillin binding proteins, role of these proteins as drug targets have also been studied **[22]**. Preprotein translocase plays an important role in extracellular transport and also as drug target **[23]**. ATP-dependent protease La is involved in catalysis of damaged or senescent proteins and short-lived regulatory proteins and therefore serves as drug target. Although most of the proteins mentioned are not validated as drug targets they are putative by virtue of their essentiality using bioinformatics tools and deserves further studies.



S. No.	Name of the Modelled Drug target	Percentage of amino acids in allowed regions
1	Preprotein translocase, SecA subunit	92.4%
2	penicillin-binding protein	86.7%
3	UDP-N-acetylmuramate--alanine ligase	86.2%
4	phosphoglycerate kinase	93.6%
5	V-type ATP synthase subunit	89.2%

Figure 2: Ramachandran plot of modeled drug targets using PROCHECK (1) Preprotein translocase; (2) SecA subunit penicillin-binding protein; (3) UDP-N-acetylmuramate--alanine ligase; (4) phosphoglycerate kinase; (5) V-type ATP synthase subunit; (6) Percentage of amino acids in allowed regions of the target proteins.

Conclusion:
The large scale genome sequencing projects have increased the availability of completely sequenced genomic and proteomic data in public domain. Use of the DEG database is more efficient than conventional methods for identification of essential genes and facilitates the exploratory identification of the most relevant drug targets in the pathogen. The present study has thus led to the identification of several proteins that can be targeted for effective drug

design against *Borrelia burgdorferi* ZS7. The drugs developed against these will be specific to the pathogen, and therefore will be less or non toxic to the host. Since the number of essential genes in the metabolic pathways of *Borrelia burgdorferi* ZS7, has been identified in the present study, these can be further characterized and their role in the survival of the bacteria can be verified. Homology modelling of these targets will help identify the best possible sites that can be targeted for drug design by simulation modelling. Virtual screening against these novel targets might be useful in the discovery of novel therapeutic compounds against *Borrelia burgdorferi* ZS7.

Acknowledgement:
We acknowledge the Department of Bioinformatics, Karnataka State Women University, Bijapur, DBT- BIF facility under BTIS (Biotechnology information system), DBT (Department of biotechnology), Ministry of Science and Technology, Government of India, for providing all the facilities for carrying out the research work.

References:

[1] Johns R *et al*. *J Med Entomol*. 2001 **38**: 99 [PMID: 11268698]
[2] Burgdorfer W *et al*. *Rev Infect Dis*. 1989 **11**(Suppl 6): S1442 [PMID: 2682956]
[3] Magnarelli LA *et al*. *J Infect Dis*. 1987 **156**: 234 [PMID: 3598218]
[4] Elliott DJ *et al*. *Teratology*. 2001 **64**: 276 [PMID: 11745834]
[5] Livey I *et al*. *Clin Infect Dis*. 2011 **52**(Suppl 3): s266 [PMID: 21217174]
[6] Wormser GP *et al*. *Expert Rev Anti Infect Ther*. 2011 **9**: 245 [PMID: 21342072]
[7] Smith BG *et al*. *J Am Acad Orthop Surg*. 2011 **19**: 91 [PMID: 21292932]
[8] Sakharkar KR *et al*. *In silco Biol*. 2004 **4**: 355 [PMID: 15724285]
[9] www.ncbi.nlm.nih.gov/
[10] Zhang R *et al*. *Nucleic Acids Res*. 2004 **32**: D271 [PMID: 14681410]
[11] Lu Z *et al*. *Bioinformatics* 2004 **20**: 547 [PMID: 14990451]
[12] Moriya Y *et al*. *Nucleic Acids Res*. 2007 **35**: W182 [PMID: 17526522]
[13] Chen X & Reynolds CH. *J Chem Inf Comput Sci*. 2002 **42**: 1407 [PMID: 12444738]
[14] Lipinski CA. *J Pharmacol Toxicol Methods*. 2000 **44**: 235 [PMID: 11274893]
[15] Altschul SF *et al*. *Nucleic Acids Res*. 1997 **25**: 3389 [PMID: 9254694]
[16] Sharma N *et al*. *The Internet Journal of Medical Informatics* 2009 **5**(1)
[17] Li KB *et al*. *Bioinformatics* 2003 **19**: 1585 [PMID: 12912844]
[18] Laskowski RA *et al*. *J Biomol NMR*. 1996 **8**: 477 [PMID: 9008363]
[19] Dutta A *et al*. *Insilico Biol*. 2006 **6**: 43 [PMID: 16789912]
[20] Lundstrom K. *J cell Mol Med*. 2007 **11**: 224 [PMID: 17488474]
[21] Naula C *et al*. *Biochim Biophys Acta*. 2005 **1754**: 151 [PMID: 16198642]
[22] Von Rechenberg M *et al*. *Proteomics*. 2005 **5**: 1764 [PMID: 15761956]
[23] Wickner W *et al*. *J Biol Chem*. 1996 **271**: 29514 [PMID: 8939874]

Edited by P Kangueane
Citation: Madagi *et al.* Bioinformation 6(9): 356-359 (2011)

BIOINFORMATION open access

Supplementary material:

Table1: Function, Cellular localization and Biochemical pathways of membrane localized proteins in *Borelia burgdorferi*

S.No	Accession No	Name of the protein	Cellular localization	Metabolic pathways in pathogen	Metabolic pathways in pathogen and not in Human
1	YP_002374587.1	Phosphoglycerate kinase	Cytoplasmic, Outer Membrane	Glycolysis/ Gluconeogenesis,Carbon fixation in photosynthetic organisms	Carbon fixation in photosynthetic organisms
2	YP_002374621.1	V-type ATP synthase beta chain (V-type ATPase subunitB	Cytoplasmic, Outer Membrane	Oxidative phosphorylation, Methane metabolism	Methane metabolism
3	YP_002374664.1	Penicillin-binding protein	Outer Membrane	Peptidoglycan biosynthesis	Peptidoglycan biosynthesis
4	YP_002374680.1	Preprotein translocase, SecA subunit	Cytoplasmic, Outer Membrane	Protein export, Bacterial secretion system	Bacterial secretion system
5	YP_002374728.1	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6- diaminopimelate ligase	Cytoplasmic, Outer Membrane	Lysine biosynthesis, Peptidoglycan biosynthesis	Peptidoglycan biosynthesis
6	YP_002374811.1	Flagellar protein export ATPase FliI	Cytoplasmic, Outer Membrane	Flagellar assembly	Flagellar assembly
7	YP_002375118.1	ATP-dependent protease La	Outer Membrane	Cellcycle Caulobacter	Cellcycle Caulobacter
8	YP_002375152.1	Protein-export membrane protein SecD	Outer Membrane	Protein export, Bacterial secretion system	Bacterial secretion system
9	YP_002375194.1	Signal recognition particle protein	Outer Membrane	Protein export,Bacterial secretion system	Bacterial secretion system
10	YP_002375216.1	Penicillin-binding protein	Outer Membrane	Peptidoglycan biosythesis	Peptidoglycan biosynthesis
11	YP_002375231.1	Penicillin-binding protein	Outer Membrane	Peptidoglycan biosynthesis	Peptidoglycan biosynthesis
12	YP_002375315.1	UDP-N-acetylmuramate--alanine ligase	Cytoplasmic, Outer Membrane	D-glutamine and D-glutamatemetabolism, Peptidoglycan biosynthesis	Cellcycle Caulobacter
13	YP_002374928.1	Chemotaxis protein methyltransferase	Cytoplasmic, Outer Membrane	Two-component system, Bacterial chemotaxis	Two-component system, Bacterial chemotaxis
14	YP_002375074.1	Chemotaxis histidine kinase	Outer Membrane	Two-component system, Bacterial chemotaxis	Two-component system, Bacterial chemotaxis
15	YP_002375181.1	Methyl-accepting chemotaxis protein	Outer Membrane	Two-component system, Bacterial chemotaxis	Two-component system, Bacterial chemotaxis