

Research article

Open Access

Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*

Tao Cui¹, Lei Zhang¹, Xizhou Wang^{1,2} and Zheng-Guo He*¹

Address: ¹National Key Laboratory of Agricultural Microbiology, Center for Proteomics Research, College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, PR China and ²Department of Medicine, Lishui University, Lishui 323000, PR China

Email: Tao Cui - wascm@webmail.hzau.edu.cn; Lei Zhang - zhanglei21e@yahoo.com.cn; Xizhou Wang - wangx@163.com; Zheng-Guo He* - hezhengguo@hotmail.com

* Corresponding author

Published: 19 March 2009

Received: 11 November 2008

BMC Genomics 2009, 10:118 doi:10.1186/1471-2164-10-118

Accepted: 19 March 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/118>

© 2009 Cui et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Analysis of the pathogen interactome is a powerful approach for dissecting potential signal transduction and virulence pathways. It also offers opportunities for exploring new drug targets.

Results: In this study, a protein-protein interaction (PPI) network of *Mycobacterium tuberculosis* H37Rv was constructed using a homogenous protein mapping method, which has shown molecular chaperones, ribosomal proteins and ABC transporters to be highly interconnected proteins. A further analysis of this network unraveled the function of hypothetical proteins as well as a potential signaling pathway. A hypothetical protein, Rv2752c, which was linked to a metal cation-transporting ATPase, was characterized as a metal-beta-lactamase, through domain analysis in combination with an *in vitro* activity experiment. A second hypothetical protein, Rv1354c, and an unknown protein kinase, PknK, interacted with a similar group of inner membrane-associated ABC transporters in the PPI network. The interactions of Rv1354 with these proteins were also confirmed by a further bacterial two-hybrid analysis. According to protein domain structures, the unique *M. tuberculosis* Rv1354c gene was proposed, for the first time, to be responsible for the turnover of cyclic-di-GMP, a second messenger molecule in this bacterium. A further structure-based inhibitors screening for Rv1354c was also performed *in silico*.

Conclusion: We constructed a comprehensive protein-protein interaction network for *M. tuberculosis* consisting of 738 proteins and 5639 interaction pairs. Our analysis unraveled the function of hypothetical proteins as well as a potential signaling pathway. The group of ABC transporters, PknK, and Rv1354c were proposed to constitute a potential membrane-associated signaling pathway that cooperatively responds to environmental stresses in *M. tuberculosis*. The study therefore provides valuable clues in exploring new signaling proteins, virulence pathways, and drug targets.

Background

The intracellular pathogen *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), is responsible for

nearly two million human deaths worldwide every year. In addition, one-third of the world's population is currently infected with the TB bacillus. The situation has

worsened in recent years with the emergence of multi-drug resistant TB and co-infection with HIV [1]. Understanding the signal transduction behind virulence and the infection mechanisms of *M. tuberculosis* is therefore critical for the identification of new drug targets and the development of new drugs.

The *M. tuberculosis* bacterium confronts a highly hostile environment during infection, including restricted access to nutrients and reduced oxygen tension [2]. Its ability to infect under these conditions has been suggested that it may use unique pathogenic mechanisms to facilitate integrated responses to the multiple stresses it encounters within the phagosome. However, little is known regarding the specific bacterial components involved in this process. Moreover, the molecular mechanism involved in sensing of extracellular signals for inducing its metabolic adaptation still remains unclear.

Protein-protein interactions (PPI) play significant roles in many biological processes, such as in the assembly of molecular complexes or in signal transduction. At a more applied level, protein interaction networks also provide tools for exploration of novel drug targets [3]. However, characterization of large-scale protein interaction networks from most organisms is still problematic due to the expense and lengthy time requirements of these types of studies. However, interaction data based on a high throughput experimental technique from model organisms can be utilized in cases where orthologs of interacting proteins can be clearly identified [4]. Walhout et al. [5] introduced the notion of the 'interolog', orthologous pairs of interacting proteins in different organisms. This idea was further extended to include paralogous interactions [6]. Orthologous interactions, or protein pairs with interacting orthologs, have been used to construct entire interactomes for organisms for which actual experimental interactions are sparse [7-9].

In this study, we constructed an *M. tuberculosis* PPI network consisting of 738 proteins and 5639 interaction pairs using a homogenous protein mapping method. Analysis of this network has provided a number of valuable clues for exploration of new virulence pathways and drug targets.

Results

A comprehensive PPI network for *M. tuberculosis*

Based on protein interaction information obtained from high throughput experiments, we have used a computational method called homologous protein mapping (HPM) to predict *M. tuberculosis* protein interactions at the proteomic scale. The rationale is that, for any pair of interactive proteins validated experimentally in the Database of Interaction Protein (DIP) [10], two proteins will be pre-

dicted to have a functional linkage if they demonstrate a higher homology. By this method, interaction evidence from model organisms was used when orthologs of interacting proteins could be clearly identified in *M. tuberculosis*. The schematic diagram shown in Fig. 1A is a simplified model for a protein that corresponds to just one other single protein. Using the HPM method, a virtual protein interaction network of *M. tuberculosis* H37Rv was constructed, consisting of 738 proteins and 5639 non-redundant interaction pairs (see Additional file 1).

Several computational methods have recently been developed to predict protein interactions [3,11,12]. Bowers and his co-workers combined these computational algorithms

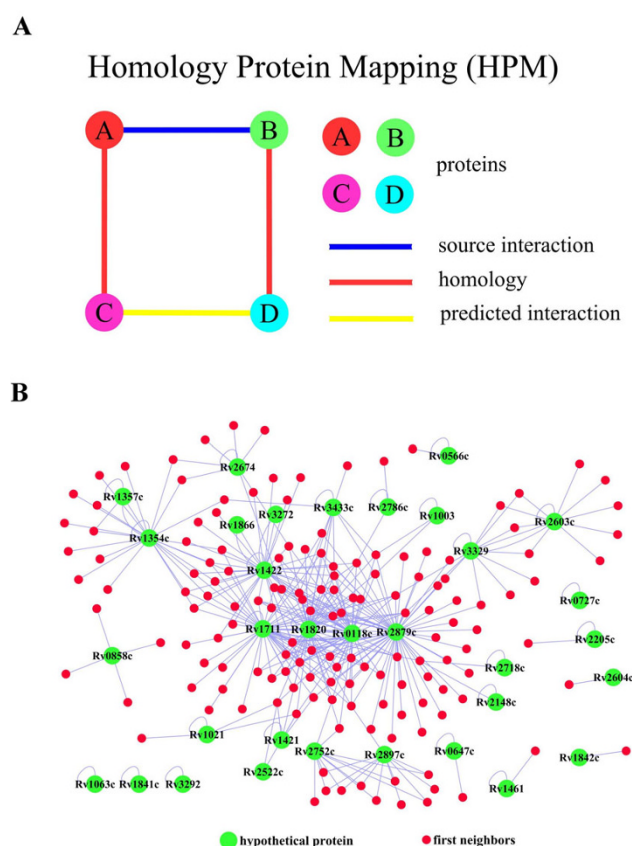


Figure 1
(A) Schematic diagram of the homologous protein mapping (HPM) method. Two proteins are predicted to have functional linkage (linked with the yellow line) if they have higher homology (linked with the red line) with a pair of interactive proteins that have been validated experimentally (linked with the blue line). Circles represent different proteins. **(B) A local PPI network of hypothetical proteins.** Thirty-three proteins (green dots) that were annotated as hypothetical proteins in genome annotation appeared in the predictive protein interaction network generated in this study (see Additional file 3).

to predict the protein interaction information for a multitude of microorganisms and released all of these data on the ProLinks database [13]. To evaluate the reliability of the PPI network, HPM data were compared with other computational methods (Table 1). The HPM data had a relatively high overlap with data produced by the Rosetta Stone method (RS) [11], although it had a very low overlap with that from the Gene Cluster method (GC) [14]. This phenomenon was also observed between other methods (Table 1), thereby demonstrating that there might be complementarities between the methods.

In protein interaction networks, the term degree represents the number of proteins that interact with a given target protein. The most highly connected proteins are usually the most important [15] and are considered to participate extensively in cellular processes. Among the twenty most highly connected proteins, most were found to be molecular chaperones, ribosomal proteins and ABC transporters (see Additional file 2). Molecular chaperones or ribosomal proteins are especially critical for the organism because of their extensive functional importance; for example, molecular chaperones help proteins to fold exactly and degrade those that fold incorrectly. This is extremely important for intracellular pathogens especially when they encounter environmental stresses imposed by host tissues. It is essential for *M. tuberculosis* to be able to resist damage from various pressures exerted by the host cells, which makes molecular chaperones especially critical for the survival of *M. tuberculosis* in its host.

In this respect, it is notable that molecular chaperones of *M. tuberculosis* have been reported to differ from those of other species [16]. For example, *M. tuberculosis* has two groEL genes, located in distinct regions of its genome, while most eubacteria contain only a single copy of this gene. The protein with the highest degree of connection to both groEL genes was Rv0685 (Tuf) and this protein showed increased expression inside macrophages in the BCG strain [17]. This also implicates molecular chaperones as essential for the survival of *M. tuberculosis* in its host, suggesting that these may be potential new drug targets.

In contrast to what is known regarding molecular chaperones, the functional importance of ABC transporters has not well been addressed, although these proteins account for about 2.5% of the genome of *M. tuberculosis* [18,19]. Their high connections to other network proteins imply that these transporter proteins might have important, but as yet unknown, functions in virulence signal transduction.

Rv2752c is a metallo-beta-lactamase

One important application of the PPI is aid in predicting the function of proteins forming the network. Thirty-three proteins have been annotated as hypothetical proteins in the *M. tuberculosis* genome from NCBI; these also appeared in the PPI networks (Fig. 1B) (also see Additional file 3). Their contexts within the network provide a basis by which their potential function and possible roles in cellular processes can be dissected. For instance, Rv2752c appeared to be linked to several metal cation-transporting D-type ATPases, which suggested that it might also be a metal-related protein (Fig. 2A). Using the CDD database [20], a metallo-beta-lactamase fold was found in Rv2752 that contained five sequence motifs. A lactamase B domain (pfam00753), which is common to all metallo-beta-lactamase, was found in the first four motifs of Rv2752 (Fig. 2B). The fifth motif, RMMBL (pfam07521), existed as a partial metallo-beta-lactamase and appeared to be specific to the function of the protein. Therefore, the existence of this RMMBL domain in Rv2752c implied that it was a typical metallo-beta-lactamase.

Previous research has demonstrated that the expression of metal-β-lactamases is one of the mechanisms by which bacterial drug resistances is imparted [21]. For *M. tuberculosis*, the activity of its metallo-beta-lactamase was examined using the purified protein following expression of Rv2752c in *E. coli*. The activity of the metallo-beta-lactamase was detected using a color reaction experiment and no activity was observed when the enzyme or metal ion were removed from the reactions (Fig. 2C).

Rv2752c and Rv2897c appeared to share the same context in local networks (Fig. 2A). Both were linked to several

Table 1: Overlaps between various computational predictive methods.

	GC(2625)	GN(4351)	PP(3894)	RS(7533)	HPM(6091)
GC(2625)		483	83	148	57
GN(4351)	483		567	323	336
PP(3894)	83	567		690	562
RS(7533)	148	323	690		985

GC, Operon/Gene Cluster method; GN, Conserved Gene methods; PP, Phylogenetic Profile method; RS, Rosetta Stone method; HPM, homogenous protein mapping method. HPM datasets was obtained from our work and other datasets obtained from ProLinks database [13]. Numbers in small brackets denote interaction pair number in each dataset.

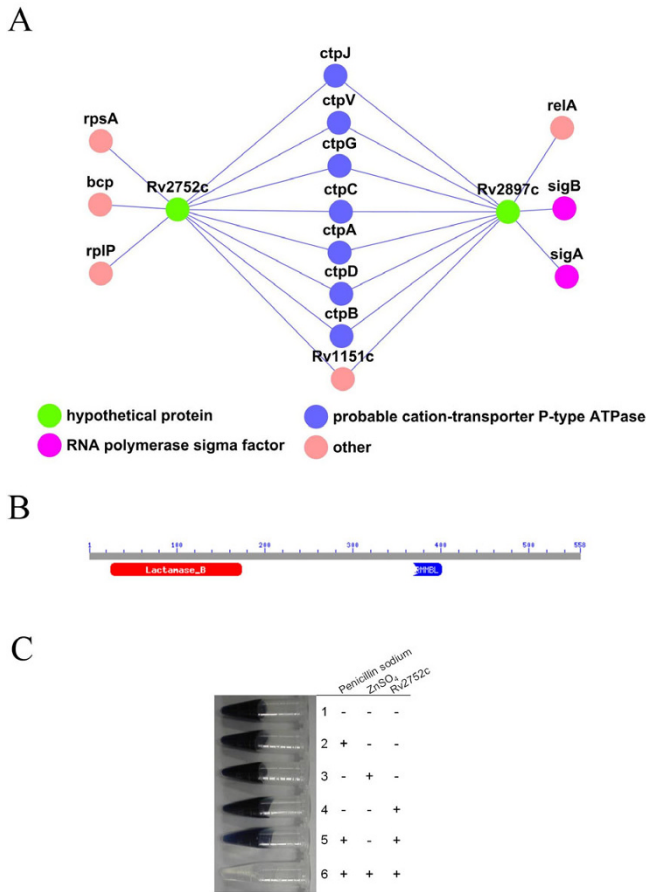


Figure 2
A local PPI network of Rv2752c and its functional assay. (A) A local PPI network of Rv2752c. (B) The domain structures of Rv2752c. (C) The activity assay of metallo-beta-lactamase for Rv2752c. The starch indicator solution and the iodine reagent were prepared. A solution containing 10,000 U of penicillin G per ml of phosphate buffer was freshly prepared and dispensed into small tubes. When a starch indicator was added to the mix with an iodine reagent, a blue color immediately developed due to the reaction of the iodine with the starch. The 0.6 mL reaction mixture contained 100 mM Tris-Cl (pH 7.5), 0.2 M NaCl, 100 ug/mL beta-lactamase, 1 mg/mL penicillin, 1% (m/v) starch, and I₂/KI and with or without 10 mM ZnSO₄. The reaction mixture was further rotated for up to six minutes at room temperature. Rapid decolorization occurred if the penicillin was hydrolyzed by beta-lactamase, which indicated positive beta-lactamase activity.

metal cation-transporting D-type ATPases. A Mg-chelatase domain (pfam01078) was also found in Rv2879c, which leads us to propose that the Rv2879c gene codes for another metallo-beta-lactamase.

M. tuberculosis Ser/Thr protein kinases

The completion of the *M. tuberculosis* genome-sequencing project has now provided a complete list of 11 eukaryotic-

like Ser/Thr protein kinases that form the *M. tuberculosis* STPK family [22]. These protein kinases appear to have important roles during infection [23,24]. Nine of these were found in the PPI networks of the current study (Fig. 3).

Ser/Thr protein kinases of *M. tuberculosis* participate in cell signaling processes such as cell division [23], fatty acid metabolism [24] and response to nitric oxide stress [25]. In the current study, two protein kinases, PknK and PknL, interacted with two cell division proteins, FtsE and FtsH, respectively, in the PPI network (Fig. 3). In addition, four network kinases, PknA, PknB, PknJ and PknL, were found to interact with a group of FadE proteins, indicating a potential regulatory role in fatty acid metabolism. Additionally, interactions were found between kinases such as PknA and PknB, and PknB and PknL, suggesting that co-regulation may also be occurring in the Ser/Thr protein kinases signal system.

PknK, although its function has not yet been established, appeared to extensively interact with the inner membrane-associated ATPase subunits of the network ABC transporters (Fig. 3). Unlike most protein kinases, PknK does not have transmembrane domains although it features a long regulatory structure in its C-terminus [26]. A previous study on protein localization showed that PknK might be associated with the inner membrane [27]. If so, the PknK may be able to participate in signal transductions in the nearby inner membrane by interacting with ABC transporters during infection.

Rv1354c is a potential signal protein associated with ABC transporters

Rv1354c is one of thirty-three hypothetical proteins included in the PPI network (Fig. 1B). Network searching revealed that Rv1354c interacted with essentially the same group of ABC transporter ATPase subunits as did PknK (Fig. 4A). Subcellular localization studies have revealed that Rv1354c is also associated with the inner membrane of *M. tuberculosis* [27]. This is consistent with the fact that no transmembrane region can be found in the protein when analyzed by using TMHMM <http://www.cbs.dtu.dk/services/TMHMM/> or SOSUI [28]. Thus, it implies that Rv1354c may localize near to the inner membrane through interactions with inner membrane proteins, such as the ATPase subunits of the ABC transporters. To confirm the interactions between Rv1354c and these proteins, a bacterial two-hybrid assay was conducted. Most of the contransformants grew well on the screening medium and no self-activation was observed (Fig. 4B). Therefore, Rv1354c was apparently interacting with the ABC transporters included in the PPI network (Fig. 4A).

Further domain analysis revealed that Rv1354c contained two typical domains of GGDEF and EAL (Fig. 5B). Both of

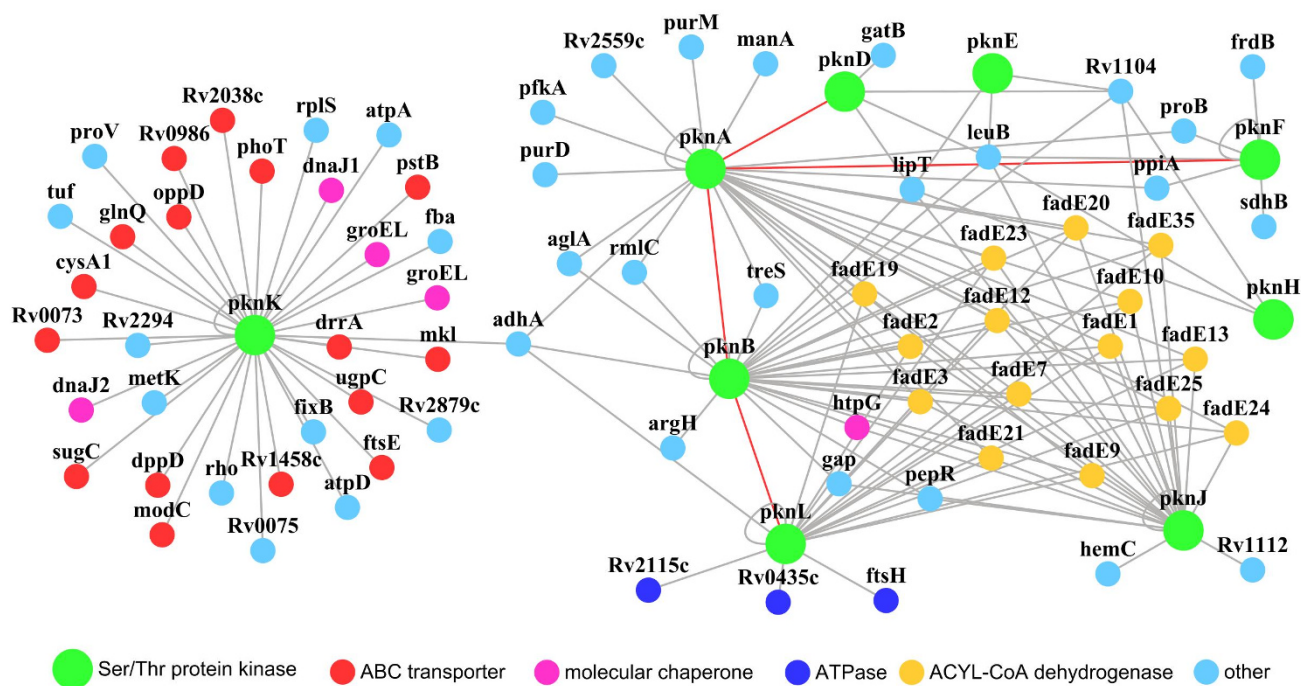


Figure 3
A PPI network of eukaryotic-like Ser/Thr protein kinases. The Protein interaction network was visualized using Cytoscape [42]. Nodes represent individual proteins and are linked with edges if they interacted with one another. Classes of proteins are colored differently as indicated.

these are involved in the turnover of cyclic-di-GMP, a multi-functional second messenger molecule in bacteria. The GGDEF domain exhibited diguanylate cyclase (DGC) activity, which is involved in the synthesis of cyclic-di-GMP from two GTPs. The EAL domain had the phosphodiesterase (PDE) activity required for the hydrolysis of cyclic-di-GMP [29]. Additionally, a sense domain GAF (E-value 1e-04) was also found in the N-terminus of Rv1354c. It has been proposed that the GAF structure has a role in sensing NO [30], a signal molecule involved in mediating the antimicrobial activity of macrophages [31].

An Rv1354 homologous protein from the *M. smegmatis* has been shown recently to have both *in vitro* diguanylate cyclase (DGC) and phosphodiesterase (PDE) activity [32]. Therefore, Rv1354c might be a unique signal protein and possibly is an essential component of the cyclic-di-GMP signaling system in *M. tuberculosis*.

Structure-based inhibitor design for Rv1354c

The cyclic-di-GMP signal system is ubiquitous to bacteria but is absent in archaea and eukaryotes, making this an ideal system to target for potential therapeutic drugs. It also provides a basis for a selective inhibitor design. The Rv1354c appears to be a unique gene in *M. tuberculosis* that contains GAF-GGDEF-EAL domains and interacts

with membrane-associated ABC transporters. Thus, Rv1354c may play an important role in sensing extracellular signals and in regulating the transcription of the *M. tuberculosis* genes. Additionally, Rv1354c is localized near to the inner membrane [27], a location that renders it relatively accessible to drugs. Taken together, the data indicate that Rv1354c may be an ideal target for the design of an anti-tuberculosis drug.

The biochemical and genetic functions of Rv1354c in *M. tuberculosis* have not yet been identified experimentally. For this reason, a structure-based inhibitor screening was performed, which targeted the GGDEF domain to inhibit its cyclic-di-GMP synthesis activity. At present, the 3D structure of Rv1354c is not known. Using the automated comparative protein modeling web server SWISS-MODEL [33] and the PleD protein (PDB ID: 2V0N) from *Caulobacter vibrioides* as a template (Identities = 41%), a structural modeling of the GGDEF domain of Rv1354c (Fig. 5) was performed. The 2V0N is the structure of the PleD in complex with cyclic-di-GMP and GTPαS, which provides critical information for defining the active site. In addition, a docking-based virtual screening using DOCK6.1 [34] was designed. The Table S4 lists the top 10 hits (see Additional file 4). This provides the basis for development of new anti-tuberculosis drugs.

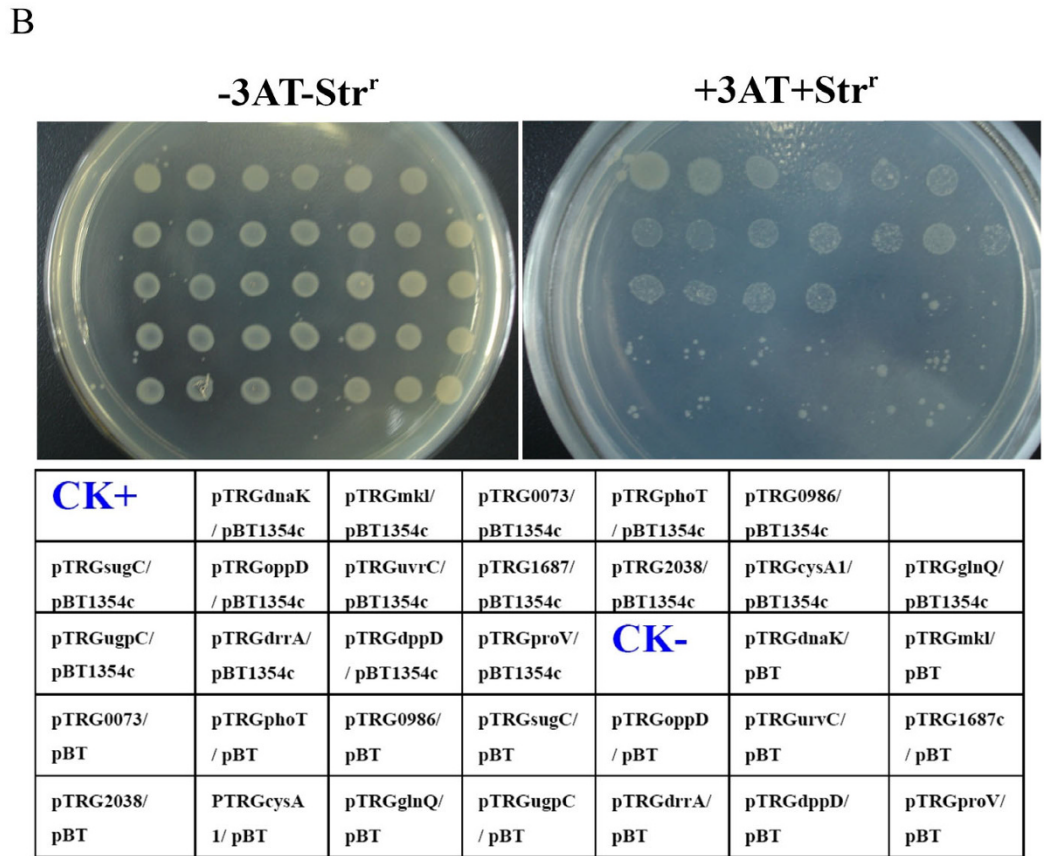
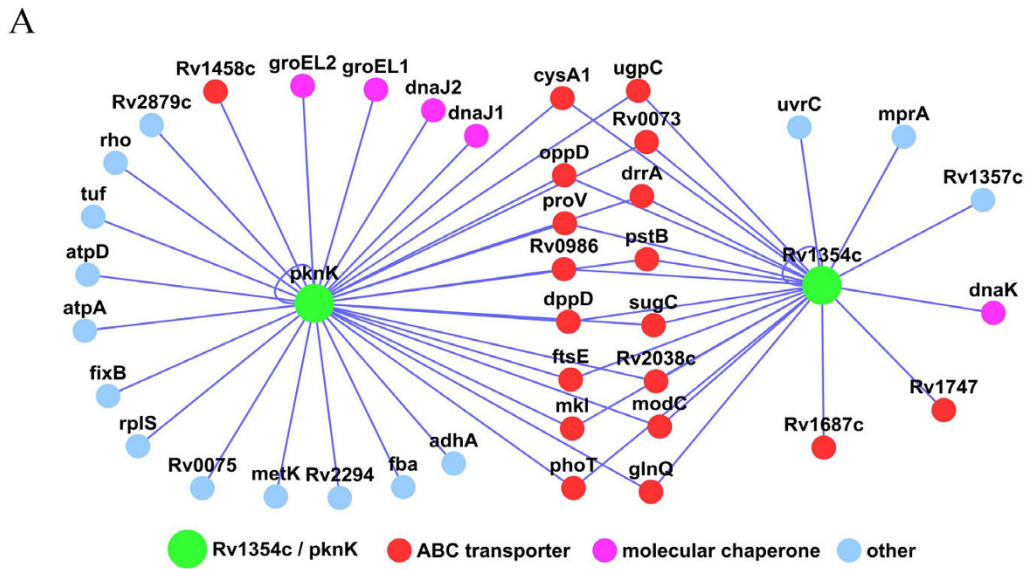
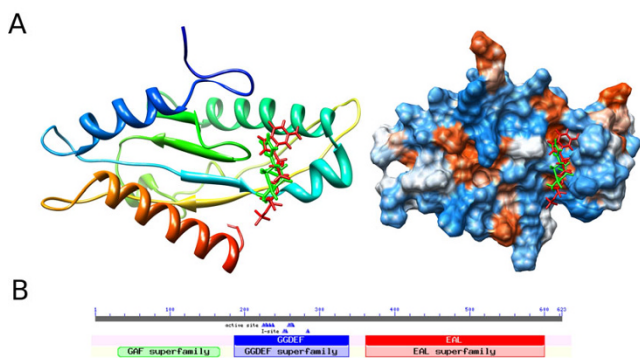


Figure 4
(A) A PPI network of both PknK and Rv1354c with the same group of ABC transporters. Nodes represent individual proteins and are linked with edges if they interacted with one another. Classes of proteins are colored differently as indicated. **(B)** Bacterial analysis of protein-protein interactions. Left panel, plate minus streptomycin (str) and 5 mM 3-amino-1,2,4-triazole (3-AT). Right panel, plate plus str and 5 mM 3-AT. An outline of the plates represents pairs of cotransformant strains. CK⁺, a cotransformant containing pBT-LGF2 and pTRG-Gal I I^P as a positive control. CK⁻, cotransformant containing pBT and pTRG as a negative control.

**Figure 5**

(A) Structure-based inhibitor design of Rv1354c. The structure of the GGDEF domain was obtained using the homology modeling method. The binding conformation of the GGDEF domain with GTP α S (red) and ZINC04632254 (green) from the docking is indicated in a different color (More compounds are listed in Additional file 4). **(B)** Schematic diagram of the Rv1354 domain structure.

Discussion

M. tuberculosis has been a major killer throughout history. Currently, it is still responsible for the deaths of about two million people each year. This unusual pathogenicity suggests that this mycobacterium may use unique mechanisms to facilitate an integrated response to the multiple stresses encountered upon entry into the phagosome, as well as for triggering some as-yet-to-be-identified switches to regulate the different phases of infection. Little is known regarding the molecular mechanisms of how extracellular signals may be sensed by *M. tuberculosis* during the infection process. The analysis of the pathogen interactome is a powerful approach for dissecting potential virulence pathways in order to explore new drug targets. In this study, we have described an approach, HPM, which allows the interactome of *M. tuberculosis* to be obtained based on the experimentally verified interactions from DIP (see Additional file 5). We constructed a comprehensive protein-protein interaction network for *M. tuberculosis*, which, upon further analysis, unraveled the potential function of hypothetical proteins and a probable signaling pathway. However, these methods, because they rely on sequence homology, can still give false positive results. To reduce this risk, we have comprehensively evaluated each query protein and its interaction using the e-value criteria in combination with the coverage and the sequence identity of the query to the hit protein (see Additional file 6).

In recent years, cyclic-di-GMP has been reported as a second messenger [29,35,36]. This signal system is involved

in the regulation of a number of complex physiological processes in numerous pathogenic bacteria, including biofilm formation and virulence factor production [29,35,36]. However, the role of this signal system in *M. tuberculosis* is yet to be characterized. In this study, we found that Rv1354c is a new and unique signal protein that contains GAF-GGDEF-EAL domains and interacts with membrane-associated ABC transporters. Given its characteristics, Rv1354c may play an important role in sensing extracellular signals and in regulating the transcription of *M. tuberculosis* genes. This study's analysis strongly suggests that Rv1354c should be considered an essential component of the cyclic-di-GMP signaling system in *M. tuberculosis*. Based on the evidence from the current study, membrane-associated ABC transporters, a protein kinase signal, and a cyclic-di-GMP signaling system may be coordinating to form an integrated signaling pathway that allows *M. tuberculosis* to respond to extracellular stresses.

With the emergence of multi-drug resistant TB and coinfection with HIV, the incidence of tuberculosis has been increasing substantially worldwide over the past decade. However, no tuberculosis-specific drug has been discovered for more than 40 years. Analysis of the pathogen protein-protein interaction is a powerful approach for dissecting potential signal transduction and virulence pathways [37,38]. It also offers opportunities for exploring new drug targets. Two previous studies proposed the existence of a protein interaction network in *M. tuberculosis* based on the subcellular protein profiling [27] and gene expression profiling [39]. A very recent study discussed the potential pathways to drug resistance based on PPI analysis [40]. The current study suggests that the signal proteins of ABC transporters, PknK and Rv1354c may have previously unknown functions and probably essential roles for the *in vivo* growth of *M. tuberculosis*. As potential targets for drug development, these new proteins are attractive because they are localized near the inner membrane. In particular, the cyclic-di-GMP signaling transduction protein is apparently absent in human and mammalian cells and is essential for the virulence and stress resistance of *M. tuberculosis*. In this study, we found for the first time that Rv1354c contains multiple important signaling structures including GAF, GGDEF, and EAL domains, and that it interacts with membrane-associated ABC transporters. Thus, as a unique cyclic-di-GMP signaling transduction gene in *M. tuberculosis*, Rv1354c may play an important role in sensing extracellular signals that facilitate integrated responses to multiple stresses encountered by *M. tuberculosis* within the phagosome. Rv1354c therefore represents an ideal target for design of an anti-tuberculosis drug and further research should now focus

on active site and docking-based inhibitors as potential new chemical weapons against the persistent pathogen *M. tuberculosis*.

Conclusion

In this study, we comprehensively analyzed a PPI network that contained 738 proteins and 5639 non-redundant interaction pairs. A further analysis of this network unraveled the function of hypothetical proteins as well as a potential signaling pathway. A hypothetical protein Rv2752c was characterized as a metal-beta-lactamase through a domain analysis and *in vitro* activity experiment. Rv1354 was proposed to be responsible for the turnover of cyclic-di-GMP in the bacterium. Its interactions with a group of inner membrane-associated ABC transporters were confirmed by a bacterial two-hybrid analysis. Our analysis suggests that the coordination of a group of ATPase subunits of the ABC transporters, PknK and Rv1354c constitute a potential membrane-associated signaling pathway and all represent potential drug targets. The PPI network developed in this study, coupled with the findings of potential signal proteins, provides a basis for understanding the potential virulence and infection mechanisms of *M. tuberculosis*. They also suggest areas of focus in the search for new drug targets.

Methods

Bacterial strains, plasmids, enzymes, primers, and chemicals

Escherichia coli BL21 cells, purchased from Novagen, were used as the host strains to express *M. tuberculosis* proteins. The pBT, pTRG vectors and *Escherichia coli* host strains were purchased from Stratagene. The pET28a was purchased from Novagen. Restriction enzymes, T4 ligase, modification enzymes, Pyrobest DNA polymerase, deoxy-nucleoside triphosphates (dNTPs) and all antibiotics were obtained from TaKaRa Biotech. All reagents for the two-hybrid assay were purchased from Stratagene, while all PCR primers were synthesized by Invitrogen (Table S1). DNA purification kits were purchased from Watson Biotechnologies, and the Ni-NTA (Ni²⁺-nitrilotriacetate) agarose columns were obtained from Qiagen.

Homologous protein mapping

To integrate accumulative protein interaction information obtained from the high throughput experiment, a computational method called homologous protein mapping (HPM) was used to predict *M. tuberculosis* protein interactions at the proteomic scale. With a pair of interactive proteins validated experimentally in the Database of Interaction Protein (DIP) [10], two proteins were predicted to have a functional linkage if they demonstrated higher homology. The schematic diagram of the HPM method is a simplified model in which a protein corre-

sponds to just a single protein (Fig. 1A). However, when the homologous protein is actually defined, one protein may have a high homology in relation to more than one protein. In this case, a permutation strategy to create all possible interaction pairs is then used. For example, given that A and B have been validated experimentally, if protein A is homologous to C and D, then protein B is homologous to E and F. Consequently, the final mapping result will be composed of interaction pairs of CE, CF, DE and DF (Fig. 1A).

Construction of a PPI network with HPM

The information collected in DIP as source data were selected because they integrated many types of evidence on protein-protein interactions including those from several model organisms [10].

The homology of proteins was assigned using the sequence similarity searching tool, *blast* [41]. The latest *blast* package (2.2.16) and amino acid sequence of *M. tuberculosis* were downloaded from the FTP Web site of NCBI. Using these amino acid sequence as input, a local blast database was subsequently constructed using the program *formatdb* released in the *blast* package. The amino acid sequence of proteins in the DIP database was downloaded from the DIP Web site (update to 20071007). Subsequently, all of the DIP amino acid sequence was used in the query. Then, the previous blast database was employed as target database to perform a large-scale blast search using *blastp* (E-value 1e-20). The preliminary blast results were further processed with our custom script for a convenient mapping procedure in the next step. Consequently, the linkage between these two sets of proteins was procured. Thereafter, the mapping process was completed according to the strategy described above. The Protein interaction network was visualized using Cytoscape [42].

Expression, purification and β -lactamase activity analysis of the Rv2752 protein

The vectors expressing mycobacterial protein were constructed using the pET28a vector and primers described in Additional file 7. *Escherichia coli* BL21 (Novagen) was used as the host strain to express and purify proteins as described in a previously published procedures [43]. Purified proteins were greater than 99% pure as determined by SDS-PAGE. The beta-lactamase assay was also performed as previously described [44]. The starch indicator solution and the iodine reagent were prepared. A solution containing 10,000 U of penicillin G per ml of phosphate buffer was freshly prepared and dispensed in small tubes. When the starch indicator was added to the mix with the iodine reagent, a blue color immediately developed due to the reaction of the iodine with the starch. The 0.6 mL reaction

mixture contained 100 mM Tris-Cl (pH 7.5), 0.2 M NaCl, 100 µg/mL beta-lactamase, 1 mg/mL penicillin, 1% (m/v) starch, and I₂/KI and with or without 10 mM ZnSO₄. The reaction mixture was further rotated for up to six minutes at room temperature. Rapid decolorization occurred if the penicillin was hydrolyzed by beta-lactamase, which indicated positive beta-lactamase activity.

Bacterial two-hybrid analysis of protein-protein interactions

The BacterioMatch II Two-Hybrid System Library Construction Kit (Stratagene) was used to detect protein-protein interactions between mycobacterial proteins. Bacterial two-hybrid analysis was carried out according to the manufacturer's directions. The pBT and pTRG vectors containing mycobacterial genes were generated. All primers used for PCR amplification are described in Additional file 7. Positive growth cotransformants were selected on the Selective Screening Medium plate containing 5 mM 3-AT (Stratagene), 8 µg/ml streptomycin, 15 µg/ml tetracycline, 34 µg/ml chloramphenicol, and 50 µg/ml kanamycin. A cotransformant containing pBT-LGF2 and pTRG-Gal11P (Stratagene) was used as a positive control for expected growth on the Selective Screening Medium. A cotransformant containing empty vector pBT and pTRG was also used as a negative control.

Homology modeling and docking based virtual screening

According to protein interaction information and domain structures, we propose that Rv1354c is a suitable target for new anti-tuberculosis drugs, and a docking-based inhibitor screening was performed *in silico*. The structure of the GGDEF domain of Rv1354c was modeled computationally using the automated comparative protein modeling web server SWISS-MODEL [33]. Likewise, the docking-based virtual screening was performed using DOCK [34]. The compound library was obtained from the ZINC that supplied database molecules as 3D formats that were ready-to-dock [45].

Abbreviations

PPI: Protein-Protein Interaction; HPM: Homologous Protein Mapping; DIP: Database of Interaction Protein; TB: Tuberculosis; DGC: Diguanilate Cyclase; PDE: Phosphodiesterase.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

TC performed data analysis and wrote the manuscript. LZ and XW performed the experiments. ZGH performed the data analysis, helped in the experimental design, supervised the experiments, and wrote the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

A protein-protein interaction network of M. tuberculosis H37Rv. The data provided represent the protein interaction network of M. tuberculosis H37Rv generated from HPM method which composed with 793 individual proteins and 6091 interaction pairs.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S1.doc]

Additional file 2

Top 20 highly connected Nodes in the network. The data provided represent the top 20 highly connected Nodes in the network.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S2.doc]

Additional file 3

Hypothetical proteins in the network. The data present the hypothetical proteins in the network.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S3.doc]

Additional file 4

Top 10 hits compounds in structure based virtual screen. The data provided top 10 hits compounds in structure based virtual screen.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S4.doc]

Additional file 5

The interactome of Mycobacterium tuberculosis obtained based on the experimentally verified interactions from DIP. The data present all protein-protein interactions of Mycobacterium tuberculosis obtained based on the experimentally verified interactions from DIP.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S5.xls]

Additional file 6

Identity, coverage and E-value of each query protein of Mycobacterium tuberculosis to the hit protein in DIP. The data described the identity, coverage and E-value of each query protein of Mycobacterium tuberculosis to the hit protein in DIP.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S6.xls]

Additional file 7

Primers and plasmids used in the study. The data present the primers and plasmids used in the study.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-10-118-S7.doc]

Acknowledgements

This work was supported by 973 Program (2006CB504402), the National Natural Science Foundation of China (30670016) and the National Special

Key Project of China on Major Infectious Diseases (2008–2010). We thank Dr. Yi Wang for offering recombinant vectors.

References

- Johnson R, Streicher EM, Louw GE, Warren RM, vanHelden PD, Victor TC: **Drug Resistance in *Mycobacterium tuberculosis***. *Curr Issues Mol Biol* 2006, **8**:97-112.
- James PE, Grinberg OY, Michaels G, Swartz HM: **Intraphagosomal oxygen in stimulated macrophages**. *J Cell Physiol* 1995, **163**:241-247.
- Strong M, Eisenberg D: **The protein network as a tool for finding novel drug targets**. *Prog Drug Res* 2007, **191**:193-215.
- vonMering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P: **STRING: known and predicted protein-protein associations, integrated and transferred across organisms**. *Nucleic Acids Res* 2005:D433-7.
- Walhout AJ, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, Thierry-Mieg N, Vidal M: **Protein interaction mapping in *C. elegans* using proteins involved in vulval development**. *Science* 2000, **287**:116-122.
- Deane CM, Salwiński Ł, Xenarios I, Eisenberg D: **Protein interactions: two methods for assessment of the reliability of high throughput observations**. *Mol Cell Proteomics* 2002, **1**:349-356.
- Huang TW, Lin CY, Kao CY: **Reconstruction of human protein interolog network using evolutionary conserved network**. *BMC Bioinformatics* 2007, **8**:152-152.
- Jonsson PF, Bates PA: **Global topological features of cancer proteins in the human interactome**. *Bioinformatics* 2006, **22**:2291-2297.
- Saeed R, Deane C: **An assessment of the uses of homologous interactions**. *Bioinformatics* 2008, **24**:689-95.
- Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D: **The Database of Interacting Proteins: 2004 update**. *Nucleic Acids Res* 2004:D449-51.
- Enright AJ, Iliopoulos I, Kyripides NC, Ouzounis CA: **Protein interaction maps for complete genomes based on gene fusion events**. *Nature* 1999, **402**:86-90.
- Dandekar T, Snel B, Huynen M, Bork P: **Conservation of gene order: a fingerprint of proteins that physically interact**. *Trends Biochem Sci* 1998, **23**:324-8.
- Bowers PM, Pellegrini M, Thompson MJ, Fierro J, Yeates TO, Eisenberg D: **Prolinks: a database of protein functional linkages derived from coevolution**. *Genome Biol* 2004, **5**:R35.
- Strong M, Mallick P, Pellegrini M, Thompson MJ, Eisenberg D: **Inference of protein function and protein linkages in *Mycobacterium tuberculosis* based on prokaryotic genome organization: a combined computational approach**. *Genome Biol* 2003, **4**:R59.
- Jeong H, Mason SP, Barabási AL, Oltvai ZN: **Lethality and centrality in protein networks**. *Nature* 2001, **411**:41-2.
- Qamra R, Mandé SC, Coates AR, Henderson B: **The unusual chaperonins of *Mycobacterium tuberculosis***. *Tuberculosis (Edinb)* 2005, **85**:385-94.
- Monahan IM, Betts J, Banerjee DK, Butcher PD: **Differential expression of mycobacterial proteins following phagocytosis by macrophages**. *Microbiology* 2001, **147**:459-71.
- Braibant M, Gilot P, Content J: **The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis***. *FEMS Microbiol Rev* 2000, **24**:449-67.
- Pasca MR, Guglielame P, Arcesi F, Bellinzoni M, DeRossi E, Riccardi G: **Rv2686c-Rv2687c-Rv2688c, an ABC fluoroquinolone efflux pump in *Mycobacterium tuberculosis***. *Antimicrob Agents Chemother* 2004, **48**:3175-8.
- Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, et al.: **CDD: a conserved domain database for interactive domain family analysis**. *Nucleic Acids Res* 2007:D237-40.
- Livermore DM, Woodford N: **The beta-lactamase threat in Enterobacteriaceae, Pseudomonas and Acinetobacter**. *Trends Microbiol* 2006, **14**:413-20.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, et al.: **Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence**. *Nature* 1998, **393**:537-44.
- Kang CM, Abbott DW, Park ST, Dascher CC, Cantley LC, Husson RN: **The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape**. *Genes & Dev* 2005, **19**:1692-1704.
- Molle V, Brown AK, Besral GS, Cozzone AC, Kremer L: **The Condensing Activities of the *Mycobacterium tuberculosis* Type II Fatty Acid Synthase Are Differentially Regulated by Phosphorylation**. *J Biol Chem* 2006, **281**:30094-103.
- Kumar DJ, Jacobs VWR, Narayanan S: **Protein kinase E of *Mycobacterium tuberculosis* has a role in the nitric oxide stress response and apoptosis in a human macrophage model of infection**. *Cell Microbiol* 2008, **10**:365-374.
- Av-Gay Y, Everett M: **The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis***. *Trends Microbiol* 2000, **8**:238-44.
- Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X: ***Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling**. *Mol Biol Cell* 2005, **16**:396-404.
- Hirokawa T, Boon-Chieng S, Mitaku S: **SOSUI: classification and secondary structure prediction system for membrane proteins**. *Bioinformatics* 1998, **14**:378-9.
- Tamayo R, Pratt JT, Camilli A: **Roles of cyclic diguanylate in the regulation of bacterial pathogenesis**. *Annu Rev Microbiol* 2007, **61**:131-48.
- Sardiwal S, Kendall SL, Movahedzadeh F, Rison SC, Stoker NG, Djordjevic S: **A GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem**. *J Mol Biol* 2005, **353**:929-36.
- Boscá L, Zeini M, Travésá PG, Hortela S: **Nitric oxide and cell viability in inflammatory cells: a role for NO in macrophage function and fate**. *Toxicology* 2005, **208**:249-258.
- Kumar M, Chatterji D: **Cyclic di-GMP: a second messenger required for long-term survival, but not for biofilm formation, in *Mycobacterium smegmatis***. *Microbiology* 2008, **154**:2942-55.
- Arnold K, Bordoli L, Kopp J, Schwede T: **The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling**. *Bioinformatics* 2006, **22**:195-201.
- Ewing TJ, Makino S, Skillman AG, Kuntz ID: **DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases**. *J Comput Aided Mol Des* 2001, **15**:411-28.
- Jenal U, Malone J: **Mechanisms of cyclic-di-GMP signaling in bacteria**. *Annu Rev Genet* 2006, **40**:385-407.
- Römling U, Amikam D: **Cyclic di-GMP as a second messenger**. *Curr Opin Microbiol* 2006, **9**:218-28.
- O'Hare H, Juillerat A, Dianisková P, Johnsson K: **A split-protein sensor for studying protein-protein interaction in mycobacteria**. *J Microbiol Methods* 2008, **73**:79-84.
- Singh A, Mai D, Kumar A, Steyn AJ: **Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association**. *Proc Natl Acad Sci USA* 2006, **103**:11346-51.
- Rachman H, Strong M, Schaible U, Schuchhardt J, Hagens K, Mollenkopf H, Eisenberg D, Kaufmann SH: ***Mycobacterium tuberculosis* gene expression profiling within the context of protein networks**. *Microbes Infect* 2006, **8**:747-57.
- Raman K, Chandra N: ***Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance**. *BMC Microbiol* 2008, **8**:234.
- Altschul SF, Lipman DJ: **Protein database searches for multiple alignments**. *Proc Natl Acad Sci USA* 1990, **87**:5509-13.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res* 2003, **13**:2498-504.
- He ZG, Rezende LF, Willcox S, Griffith JD, Richardson CC: **The carboxyl-terminal domain of bacteriophage T7 single-stranded DNA-binding protein modulates DNA binding and interaction with T7 DNA polymerase**. *J Biol Chem* 2003, **278**:29538-29545.
- Catlin BV: **Iodometric Detection of *Haemophilus influenzae* beta-Lactamase: Rapid presumptive test for ampicillin resistance**. *Antimicrob Agents Chemother* 1975, **7**(3):265-270.
- Irwin JJ, Shoichet BK: **ZINC – a free database of commercially available compounds for virtual screening**. *J Chem Inf Model* 2005, **45**:177-82.