JBC Papers in Press. Published on May 24, 2019 as Manuscript RA118.005312
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.RA118.005312

Unraveling the role of the transcriptional regulator VirS in low pH-induced responses of *Mycobacterium tuberculosis* and identification of VirS inhibitors

Swati Singh[1], Nikita Goswami[1], Anil K. Tyagi[1,2*] and Garima Khare[1*]

[1]Department of Biochemistry, University of Delhi South Campus, Benito Juarez road, New Delhi 110021, India.

[2] Vice Chancellor, Guru Gobind Singh Indraprastha University, Sector 16-C, Dwarka, New Delhi 110078, India.

Running title: Role of VirS in acidic stress & its inhibitor identification.

*To whom correspondence should be addressed: Dr. Garima Khare: Department of Biochemistry, University of Delhi South Campus, Benito Juarez Road, New Delhi-110021, India. Email: garima1822@yahoo.co.in, Tel: 91-11-24115209, Fax: 91-11-24115270 and Prof. Anil K. Tyagi: Department of Biochemistry, University of Delhi South Campus, Benito Juarez Road, New Delhi-110021, India. Email: aniltyagi@south.du.ac.in, Tel: 91-11-24110970, Fax: 91-11-24115270

Keywords: *Mycobacterium tuberculosis*, gene microarray, site-directed mutagenesis, transcription factor, tuberculosis, virtual screening, inhibitor, VirS, immune evasion, AraC/XylS type transcription factor.

Abstract

The ability of *Mycobacterium tuberculosis* to respond and adapt to various stresses such as oxygen/nitrogen radicals and low pH inside macrophages is critical for the persistence of this human pathogen inside its host. We have previously shown that an AraC/XylS-type transcriptional regulator, VirS, which is induced in low pH, is involved in remodeling the architecture of the bacterial cell envelope. However, how VirS influences gene expression to coordinate these pH responses remains unclear. Here, using a genetic biosensor of cytoplasmic pH, we demonstrate that VirS is required for the intracellular pH maintenance in response to acidic stress and inside acidified macrophages. Furthermore, we observed that VirS plays an important role in blocking phagosomal–lysosomal fusions. Transcriptomics experiments revealed that VirS affects the expression of genes encoding metabolic enzymes, cell-wall envelope proteins, efflux pumps, ion transporters, detoxification enzymes, and transcriptional regulators expressed under low-pH stress. Employing electrophoretic mobility-shift assays, DNA footprinting, and *in silico* analysis, we identified a DNA sequence to which VirS binds and key residues in VirS required for its interaction with DNA. A significant role of VirS in *M. tuberculosis* survival in adverse conditions suggested it as a potential anti-mycobacterial drug target. To that end, we identified VirS inhibitors in a virtual screen; the top hit compounds inhibited its DNA-binding activity and also *M. tuberculosis* growth *in vitro* and inside macrophages. Our findings establish that VirS mediates *M. tuberculosis* responses to acidic stress and identify VirS-inhibiting compounds that may form the basis for developing more effective anti-mycobacterial agents.

Introduction

Mycobacterium tuberculosis (*M. tuberculosis*) is known to resist the acidic stress encountered in macrophages and multiply in these hostile conditions, however, the mechanisms for its survival in acidic conditions are poorly understood. There are a few genes that have been implicated in

acid resistance in *M. tuberculosis* which include a serine protease Rv3671c, OmpATb operon and a putative magnesium transporter MgtC (1-3). A transposon mutant of Rv3671c was shown to be impaired in the maintenance of intrabacterial pH under acidic conditions suggesting its involvement in the acid resistance (1). OmpATb, an acid responsive porin was shown to be involved in the adaptation to acidic environment by mediating secretion of ammonia, however, its role was considered as redundant as knocking out of OmpATb did not influence the virulence of the pathogen in mice (2,4). *M. tuberculosis* lacking MgtC, a putative magnesium transporter, was found to be attenuated for growth under mild acidic conditions at low Mg^{2+} (3). Apart from *M. tuberculosis*, acid resistance has also been shown to be important for other bacteria such as *Helicobacter pylori*, which colonizes in human stomach having acidic environment, *Streptococcus pneumoniae*, pathogenic strains of *E. coli* and *Salmonella enterica.* Studies have shown the involvement of a few proteins in the acid resistance of these bacteria that include urease and ExbD in *H. pylori*, Gad proteins and F_0F_1 ATPase in *E. coli* and Mg^{2+} transporter in *Salmonella* (5-9).

VirS (Rv3082c) of *M. tuberculosis* belongs to the AraC family of transcriptional regulators (10,11). VirS is present divergently upstream of an acid inducible operon termed as *mymA* operon, which comprises of seven genes (Rv3083-Rv3089)(12). The transcription of *mymA* operon under acidic stress has been earlier shown to be regulated by VirS which itself is regulated by acidic pH (12). Studies demonstrated that *virS* mutant of *M. tuberculosis* exhibited altered cell wall structure, altered mycolic acid content, defective intramacrophage survival and reduced hematogenous dissemination *in vivo* (13). Importantly, *virS* expression was induced during chronic and reactivation phases of murine tuberculosis, implicating VirS in persistence and reactivation of tuberculosis (14). Despite these findings, mechanisms of how VirS exerts its influence on gene expression to elicit *M. tuberculosis's* response under acid stress remain uncharacterized.

Here, our study has delineated the contribution of VirS in acid stress and how it mediates its influence on gene expression to coordinate pH responses in *M. tuberculosis*. Our study shows the requirement for VirS in controlling cytoplasmic pH balance and phagosomal maturation during acidic stress. Further, structure guided mutational studies revealed key residues of VirS required for its interaction with DNA and structure-based virtual screening identified some inhibitors of this interaction with DNA which could be a subject of lead optimization for future drug designing against *M. tuberculosis*.

Results:

VirS is required to survive under acidic stress and to block phagosomal maturation

VirS has been previously reported to be upregulated under acidic conditions (12). To understand the involvement of VirS in acidic responses of *M. tuberculosis*, an *in vitro* growth study was carried out by growing *M. tuberculosis* Erdman (wild type), *virS* mutant and *virS* complemented strain under varying pH conditions (pH 4.5, 5.0, 5.5 and 6.6) in MB7H9 medium and the survival of the cells under these conditions was evaluated. The growth of all three strains was comparable at pH 6.6 and 5.5. However, the growth of the *virS* mutant was significantly reduced at pH 5.0 and 4.5, with a pronounced defect at pH 4.5 as compared to parental *M. tuberculosis* and *virS* complemented strain (Figure 1). We also performed survival studies of wild type, *virS* mutant and *virS* complemented strain at acidic conditions of pH 4.5 in 7H9 medium containing non-hydrolysable Tyloxapol (7H9-4.5-Ty) as the dispersing agent instead of Tween 80 to negate the possibility of growth defect due to hydrolysis of Tween 80 under acidic conditions to free fatty acids, which can be toxic to the cells (1). We monitored survival of these strains after 6 and 9 days of incubation in 7H9-4.5-Ty medium by colony forming units (CFU) enumeration. It was observed that *virS* mutant displayed growth defect in acidified medium post 6 and 9 days of incubation, whereas parental and complemented strain did not exhibit any marked effect on their growth under acidic conditions suggesting the involvement of VirS in the survival of the bacteria under acidic conditions (Figure S1). Moreover, it was earlier reported that *virS* mutant showed survival defects specifically in

immune-activated macrophages (13). *M. tuberculosis* is known to survive under acidic conditions in macrophages by arresting the phagosome-lysosome fusion. Hence, we evaluated the role of VirS in arresting phagosomal maturation in *M. tuberculosis*. For this, we employed FITC labelled *M. tuberculosis*, *virS* mutant and *virS* complemented strain and studied the localization of the pathogen in the lysosomal acidic compartments by using LysoTracker Red dye in THP-1 macrophages. In resting macrophages, all the three strains exhibited similar colocalization of phagosomes with the LysoTracker rich compartments (data not shown). However, in the case of activated macrophages, parental *M. tuberculosis* cells exhibited ~9% colocalization with the Lysotracker stained acidified compartments and largely remained in the non-acidified phagosomes whereas in contrast to the wild type strain, *virS* mutant displayed significantly greater colocalization (~32%) with the acidified compartments (Figure 2A and 2B). The complemented strain also exhibited ~ 6% colocalization which is similar to that observed in the case of wild type. These observations suggest the involvement of VirS in the arrest of phagosomal maturation in the activated macrophages, providing survival advantage to the pathogen under harsher conditions.

Involvement of VirS in maintaining the intrabacterial pH of *M. tuberculosis* in acidic conditions

Activated macrophages are known to have acidic environment and *M. tuberculosis* robustly persists in this acidic milieu, indicating the role of mycobacterial factors in intramycobacterial pH homeostasis. Hence, we next sought to determine the role of VirS in maintaining the pH homeostasis by measuring the intrabacterial pH (pH_{IB}) of various strains by using a pH sensitive ratiometric GFP at various intervals under normal and acidic conditions. The pH_{IB} was calculated by measuring the ratio of the fluorescence intensities at excitation wavelengths of 395 nm and 475 nm with an emission wavelength of 510 nm (1). A standard curve of the ratio of the intensities (excitation wavelengths - 395 nm and 475nm, emission wavelength - 510 nm) at varying pH conditions was plotted to determine the unknown pH_{IB} values (Figure S2A). It was observed that under normal buffer conditions (pH-7.4), the pH_{IB} was maintained to near-neutral in the parental (pH_{IB}-~7.0) as well as in the complemented strain even after 8 and 24 hours of incubation (Figure 3A). However, *virS* mutant consistently maintained marginally lower intrabacterial pH (pH_{IB} – 6.75) even in normal buffer conditions as compared to the parental and complemented strains (Figure 3A). At pH 4.5, the initial pH_{IB} of all the three strains were similar to that observed in normal conditions (Figure 3B). However, after 8 and 24 hours of incubation in pH-4.5 buffer, the intrabacterial pH of the mutant dropped to ~6.4 while the pH_{IB} of parental and complemented strains were still maintained at near neutral conditions of ~6.9 (Figure 3B). This suggested that both the parental and complemented strains maintain pH homeostasis upon acidic stress, whereas the *virS* mutant was relatively defective. Taken together, these observations suggest the importance of VirS in maintaining the intrabacterial pH under acidic conditions.

We next sought to determine the intrabacterial pH (pH_{IB}) of mycobacteria residing in phagosomes of IFN-γ activated macrophages. The unknown pH_{IB} were determined by using standard curve of ratio of intensities at excitation wavelengths-405 nm and 476 nm, emission wavelength - 535 nm at varying pH conditions (Figure S2B). It was observed that wild type and *virS* complemented strains mostly maintained their pH_{IB} at pH 6.5-7.25 after 24 hours postinfection. However, majority of the *virS* mutant mostly maintained its pH_{IB} at pH 6.26 – 7.0 after 24 hours postinfection (Figure 3C) with a small fraction of mutant cells having pH_{IB} in the range of <5.5-6.2, suggesting at 24 hrs post infection, a few mutant cells were unable to maintain their intrabacterial pH at near neutral. This effect was more pronounced after 48hrs post infection in which we found that the pH_{IB} of parental and *virS* complemented strain mostly had a pH_{IB} of 6.51-7.25 whereas *virS* mutant mostly had a pH_{IB} of 5.5 or lower in activated macrophages suggesting that after 48 hours of post infection, the *virS* mutant strain was unable to maintain its intrabacterial pH (Figure 3D). This inability of *virS* mutant to maintain its intrabacterial pH in activated macrophages indicates its role in the maintenance of pH homeostasis upon acidic stress and could be a probable reason for its poor survival under acidic

conditions.

Identification of genes regulated by VirS (VirS-regulon) under acidic condition by microarray studies

To understand the role of VirS in regulating acidic pH mediated gene expression of *M. tuberculosis*, we carried out microarray studies by isolating RNA from *M. tuberculosis* Erdman, *virS* mutant and *virS* complemented strain grown in MB7H9 medium at pH 6.6 as well as at pH 4.5. We shortlisted the VirS dependent genes induced or repressed at acidic pH relative to neutral pH with >2.0 fold difference in the expression level when compared to *M. tuberculosis* Erdman and *virS* mutant. We then further shortlisted these genes on the basis of their similar expression profile at acidic pH in *M. tuberculosis* Erdman and *virS* complemented strain in comparison to *virS* mutant suggesting that indeed the disruption of *virS* was involved in the change of expression of such genes at acidic pH. We also carried out the classification of all these induced and repressed genes into various classes based on their function and homology, however, the hypothetical proteins, unannotated genes and genes of unknown functions were not considered for further analysis. All these analyses revealed that VirS induces expression of 54 genes and represses 39 genes (>2.0 fold difference; p value < 0.05) at acidic pH (Figure 4A and 4B). The major classes of genes (after removing genes belonging to hypothetical proteins, genes of unknown functions and unannotated genes) found to be induced and repressed were related to cellular metabolism followed by cell wall associated genes involved in fatty acid oxidation and biosynthesis, which is in agreement with the role of VirS in the remodelling of cell envelope architecture as reported earlier (13). The other classes of genes induced by VirS at acidic pH were related to PE/PPE family, efflux pumps & ion transporters, detoxification and transcriptional regulation. Additionally, genes related to energy pathways and respiration and genes related to information pathways were also found to be regulated by VirS at acidic pH. Figure 4C represents the heat map showing the expression of various genes that are differentially regulated by VirS in wild type *M. tuberculosis*, *M.tbΔvirS* and *M.tbΔvirSComp* in response to acidic pH relative to neutral pH. Table 1 and 2 describe a few of the genes from the most represented categories induced or repressed by VirS at acidic pH. A detailed description of the genes and their speculated role in counteracting acidic responses is provided in the discussion section. The complete list of the genes that showed upregulation and downregulation with >2.0 fold difference are listed as supplementary excel sheet (Supporting excel sheet X1).

Identification of VirS binding region of DNA

Based on the above observations, VirS appears to be a promising drug target regulating a number of important downstream targets. Hence, we further attempted to characterize VirS by evaluating its DNA binding activity, its DNA binding site, crucial residues of VirS required for its activity and identify its inhibitors. Towards this, we cloned *virS* gene in pET43.1a vector as described in experimental procedures for the expression of fusion protein NusA-VirS, since our earlier studies showed VirS to be an insoluble protein when expressed in *E. coli* (unpublished work) and NusA tag is considered as an effective solubility tag. NusA-VirS was indeed found to be localized in the soluble fraction, however, NusA-VirS fusion protein failed to bind to the resin probably due to the non- exposure of the His tag, hence, cell lysates were used for further studies (data not shown). The DNA binding activity of NusA-VirS fusion was determined by employing EMSA by using a DNA fragment of 325bp upstream region of VirS (the intergenic region between *virS* and *mymA* genes) as VirS regulates its own expression and activates *mymA* operon (12). The fusion protein displayed a shift in the mobility of DNA suggesting the binding of VirS to the DNA fragment employed in the assay (Figure 5A). To rule out the possibility of involvement of NusA tag in the observed shift in the DNA, the lysate with overexpressed NusA was employed as a control which did not display any shift in the mobility of DNA indicating that the observed shift in the mobility was due to the binding of VirS to the DNA (Figure 5A). Further, DNase I protection assay was performed to determine the VirS binding region of DNA. The lysate of overexpressed fusion protein *i.e.* NusA-VirS was used for determining the binding region of VirS which displayed a protected region (Figure 5B). The lysate of only overexpressed NusA protein

was used as a negative control which did not display any protection. It was found that VirS protected a 46 bp sequence that was determined to be as 5'-CTGTCGGAAAATGATAAAAGCG-TGTCGCAAAGTGTCAATACGTGGC-3' (Figure 5B). Based on this VirS binding DNA sequence, we determined a VirS binding consensus logo by using the software Weblogo taking the upstream regions of a few identified downstream targets of VirS for alignment (16) (Figure 5C). For this, we performed multiple sequence alignment of this DNA sequence with the DNA sequence of *micF* and *mar* promoters used in the crystal structure determination of Rob and MarA, respectively, which provided us with a 28bp sequence exhibiting alignment and conservation in DNA binding regions of these three proteins. This 28bp region was employed to search for homology in the upstream regions of a few of the VirS regulated genes namely *icl, ctpG, ahpC, infC, sirA, pdc, Rv1405c, adh, fadD9,* and *fadE35* identified by microarray. These genes were randomly selected to represent a variety of different functional categories. We observed that all these genes carried the 28bp sequence required for binding of VirS although with a few variations in the sequences (Figure 5C).

We further explored the regulation of few of the genes identified by microarray studies by VirS. The genes selected were namely *Rv2428 (ahpD), Rv1706 (ppe23), Rv1994c (cmtR), Rv1992c (ctpG), Rv2354c (ppe38)* and *Rv2590 (fadD9).* For determining whether these genes are directly or indirectly regulated by VirS, we selected ~400-600bp upstream region of these genes that comprised their respective promoters as predicted by online software NNPP (http://www.fruitfly.org/seq–tools/promoter.html). These upstream regions were separately amplified by employing Cy5 PCR Labelling Kit (Jena Biosciences, Germany). The labelled product was then evaluated for their ability to bind VirS by employing EMSA as described in the experimental procedures. It was observed that the upstream regions of all of the selected genes displayed a shift in the mobility suggesting their direct interaction with VirS (Figure 5D). As a negative control to negate a general binding of VirS to DNA, we used Cy5 labelled 32 bp intergenic region between *MbtA* and *MbtB* that contains IdeR (iron dependent transcription regulator) binding region, which did not display any shift in its mobility in the presence of VirS indicating the specificity of VirS to its own downstream targets (Figure 5E).

Homology modelling of VirS

We generated a three dimensional comparative homology model of VirS for performing virtual screening studies and also for carrying out identification of the crucial residues involved in its DNA binding activity. Since, the crystal structure of *M. tuberculosis* VirS is not available; we generated the three dimensional structure of VirS by employing comparative homology modelling as described in experimental procedures. It was found by bioinformatic analysis that C-terminus of VirS (250-340 amino acid) contained the DNA binding region and that was homologous to the C-terminus of the transcriptional regulators of AraC/XylS family. Hence, this C-terminus region was employed for multiple sequence alignment by ClustalW (16) and for model generation by Modeller 9.0 (17) (Figure 6A). The generated 3-D homology model of VirS contained two helix turn helix motifs (Figure 6B). The co-ordinates of DNA bound to MarA structure were used to predict the DNA contact sites of VirS (Figure 6C). The predicted structure of VirS was also validated by using Rampage (18), Errat (19), Verify 3D (20,21), Pro-Q (22) and Prosa (23,24) softwares (Table S1). The results obtained from these programs implied that the generated structure of VirS was in a good configuration.

Identification of the critical residues of VirS involved in its DNA binding activity

On the basis of multiple sequence alignment and the predicted structure of VirS, 15 residues were selected that might affect the DNA binding activity of VirS and structure guided mutational studies were performed (Figure 7A and Table 3). These residues were substituted with alanine for mutational studies since alanine is a non-polar and small amino acid that takes care of charge as well as the size of the side chain.

For carrying out site-directed mutagenesis study, we subcloned the *virS* gene from pET43.1a-*virS* construct to another vector pBEn-SET3a as described in experimental procedures as the size of

pBEn-3a-*virS* construct (6.8kb) is smaller than the pET43.1a-*virS* (8.3kb) and thus more appropriate for SDM studies. 3a-VirS fusion protein was purified and assessed for its DNA binding activity with the VirS binding region of DNA identified by DNA footprint assay containing 46bp flanked by 10bp at each side (66bp). For the ease of work, the ends of this 66bp region were end labelled with fluorophore Cy5. The Cy5 labelled DNA was then employed for the determination of DNA binding activity of VirS by EMSA studies. The purified protein was found to be functionally active and displayed the gel shift which also verified the identified binding region of DNA (data not shown). 0.5 µg of protein was selected for the further work.

Fifteen VirS mutants were generated by using site-directed mutagenesis as described in experimental procedures. The sequences of the primers used for mutagenesis are given in Table S2. All the mutations were confirmed by DNA sequencing and the mutant proteins were separately expressed and purified successfully by using Ni-NTA affinity chromatography (data not shown). Mutants Y61A, R73A and R80A showed poor binding to Ni-NTA resin due to probable changes in the protein folding as a result of mutation. Thus, their lysates were prepared and employed for EMSA studies and lysate of wild type 3a-VirS was used as a control for comparison. The mutants were evaluated for their DNA binding activity by using EMSA (Figure 7B).

Mutants R16A, Y61A and R69A displayed more than 90% reduction in the DNA binding activity of VirS suggesting that the corresponding residues are critical for the DNA binding (Figure 7B). Mutant R20A displayed 84% reduction in the activity indicating the importance of R20 in DNA binding. Mutants R28A and R73A showed moderate effect on the activity as they displayed reduction of 34% and 30%, respectively (Figure 7B). Figure 7C depicts the bar graph representing the percent DNA binding activity exhibited by the mutants when compared to the wild type VirS protein. Mutants H30A, E34A, R38A and S70A did not display any influence on the DNA binding activity suggesting that these residues are not important for the DNA binding activity of VirS. In conclusion, four residues R16, R20, Y61 and R69 were identified that displayed the loss of DNA binding activity by more than 84% in comparison to the wild type VirS. These residues of VirS might directly make contacts with the cognate DNA and can be used as specific target sites for rational design of inhibitors against VirS for anti-tubercular drug development.

Identification of the inhibitory molecules against VirS by structure-based virtual screening and their evaluation against DNA binding activity of VirS and *M. tuberculosis* growth

Amongst the homologous proteins of VirS, only three structures *i.e.* MarA (PDB ID-1BL0), Rob (PDB ID-1D5Y) and AdpA (PDB ID-3W6V) were available in ligand bound form (25-27). The DNA contact sites of VirS were predicted by superimposition with DNA bound MarA structure (Figure 8A). On superimposing the 3D structure of VirS with structure of MarA, two DNA contact sites in VirS were observed (Figure 8B). Hence, virtual screening was carried out at both the DNA contact sites of VirS by using software Autodock4.2 (28,29). The grid covering the DNA contact site 1 of VirS is represented in figure 8C and 8D. The residues *i.e.* Glu5, Gln19, Arg28 and His30, were surrounding the grid and found to be covered in the generated grid (Figure 8E). Similarly, another grid was generated against the DNA contact site 2 of VirS (Figure 8F) and its zoomed representation is shown in Figure 8G. The residues found to surround and covered in the grid were Ala56, Ser53, Tyr61, Ser62, Glu63 and Gln64 (Figure 8H). A filtered library of NCI Open Database was employed for virtual screening against the DNA contact sites of VirS (30). A total of 114 compounds, 57 compounds each from docking at site 1 and site 2, were procured from NCI-DTP, numbered as V1A-V57A (site 1) and V1B-V57B (site 2) and were evaluated for their ability to inhibit the DNA binding activity of VirS at a concentration of 100 µg/ml (Figure 9). 26 compounds *i.e.* 10 from site 1 and 16 from site 2 displayed more than 60% inhibition with 15 compounds (4 from site 1 and 11 from site 2) exhibiting greater than 80% inhibition at 100 µg/ml (Figure 9A, 9B and Table S3). These 26 compounds were further subjected to dose response studies for the determination of their IC_{50} values (Table S3). Compound V7B with an IC_{50} value of 1.8 µg/ml was found to be the most potent one whereas compound V29B displayed an IC_{50} value

of 16.38 μg/ml. Rest of the compounds showed IC_{50} value in the range of 17 μg/ml to 75 μg/ml (Table S3).

These 26 compounds were also evaluated for their ability to inhibit the growth of *M. tuberculosis* in broth culture by subjecting them to Resazurin Microtiter Assay (REMA) (31-33). Rifampicin was used as a positive control in the assay, which exhibited MIC_{90} of 0.25 μg/ml and DMSO was used as a vehicle control. 14 compounds displayed growth inhibition of *M. tuberculosis* with varying MIC_{90} values (Table 4 and Figure S3). The most potent inhibition against *M. tuberculosis* was exhibited by compound V15B with MIC_{90} value of 5 μg/ml. Compound V7B, which was the most potent compound in *in vitro* assay, displayed MIC_{90} of 40 μg/ml. To summarize, we identified 14 compounds (listed in table 4) that exhibited growth inhibition against *M. tuberculosis* as well as inhibited DNA binding activity of VirS in biochemical assays.

We then further evaluated the cytotoxicity of 8 compounds (V4A, V6A, V8A, V44A, V55A, V7B, V15B and V56B) which exhibited an $MIC_{90} \leq 40$ μg/ml and were also active against the DNA binding activity of VirS. Various mammalian cell-lines were employed and IC_{50} values of the compounds were determined as described in experimental procedures. The most potent inhibitor of VirS V7B (IC_{50}-1.8μg/ml, MIC_{90}-40 μg/ml) was found to be non-cytotoxic till a concentration of 100 μg/ml in all the cell-lines (Table S4). Another potent inhibitor of *M. tuberculosis* growth *i.e.* V15B (MIC_{90}-5μg/ml and IC_{50}-31.21μg/ml) was also found to be non-cytotoxic till a concentration of 150 μg/ml in all the cell lines (Table S4). Compounds V4A and V8A also displayed negligible cytotoxicity in all the employed cell-lines till 200 μg/ml and150 μg/ml, respectively. Rest of the compounds displayed moderate cytotoxicity in atleast one or two of the three employed cell-lines (Table S4).

Four compounds (V4A, V8A, V7B and V15B) which displayed inhibition of VirS DNA binding activity as well as growth of the pathogen and did not display any significant cytotoxicity were further evaluated for their potency to inhibit the growth of the pathogen inside macrophages as described in experimental procedures. It was found that compounds V7B and V15B inhibited the intracellular growth of the pathogen at ≥80 μg/ml and ≥40 μg/ml, respectively (Figure 10). To rule out the possibility of this inhibition due to cytotoxic effects of the compounds, uninfected THP-1 cells were also subjected to the same concentrations of the compounds. V7B and V15B were found to be non-cytotoxic to THP-1 cells till 320 μg/ml and 100 μg/ml, respectively. Compounds V4A and V8A failed to inhibit the intracellular growth of the pathogen even at a concentration of 100 μg/ml (data not shown). Thus, these results suggested that the compounds V7B and V15B also possess the ability to inhibit the intracellular growth of the pathogen, however, further optimization of these compounds is required to improve their potency.

Identification of the binding mode of the hit compounds

Based on the $IC_{50,}$ MIC_{90}, cytotoxicity and macrophage experiment results, compounds V7B and V15B emerged as potential hits, hence, we predicted their probable mode of binding at the active site of VirS. The compound V7B was docked at the DNA contact site 2 of VirS and surrounded by residues Gln1, Ser53, Ala56, Val57, Tyr61, Ser62, Glu63, Gln64 and Ser65 (Figure 11A). Most of these residues were also covered in the grid generated against the DNA contact site 2 of VirS. Compound V7B showed van der Waals interactions with Ser53, Glu63 and Tyr61 (Figure 11A). It was found that compound V7B interacted with residues Gln64, Ser65 and Val57 by forming hydrogen bonds (Figure 11B). Compound V15B was found to be surrounded by Gln1, Ser53, Ala56, Val57, Gly60, Tyr61, Ser62, Glu63 and Gln64 (Figure 11C). Compound V15B exhibited van der Waals interactions with residues Gln1, Gly60 and Glu63. It also showed hydrogen bonding with Tyr61 at DNA binding site 2 (Figure 11D). V15B also displayed pi-sigma bond formation with Val57. It also showed hydrophobic interactions like pi-alkyl bond formation with Ala56 and Val57 (Figure 11D).

Structure Activity Relationship studies for compounds V7B and V15B.

On the basis of IC_{50}, MIC_{90}, cytotoxicity and macrophages studies, two hit molecules *i.e.*

compound V7B and V15B were shortlisted for further SAR studies. For this, they were subjected to similarity search for identifying their analogs by using NCI database employing the software Open Babel. The molecules were selected on the basis of Tanimoto coefficient of ≥ 0.7. Seventeen molecules similar to compound V7B and ten molecules similar to compound V15B were procured from NCI-DTP and evaluated for their ability to inhibit VirS DNA binding activity to identify the structural features responsible for inhibiting VirS activity. These molecules were named as V7B-1 – V7B-17 and V15B-1 – V15B-10.

The procured molecules were evaluated for their ability to inhibit DNA binding activity of VirS at a concentration of 100 µg/ml and it was found that 4 analogs of V7B (V7B-4, V7B-5, V7B-8 and V7B-9) and one analog of V15B (V15B-2) displayed >60% inhibition of VirS DNA binding activity at 100 µg/ml while rest of them exhibited lesser inhibition (Table S5 & S6). These five analogs were further evaluated for the determination of their IC_{50} values. They were found to display a reduced inhibitory potential against VirS DNA binding activity in comparison to their parent molecules which are tabulated in Table S5 & S6. It was found that none of these molecules displayed an enhanced potential to inhibit VirS, however, their structures were carefully analysed to derive structure activity relationship.

A detailed comparison of the structures of V7B and its analogs revealed the presence of a benzene sulfonamide moiety in all the structures except in V7B-5 and V7B-17. The structure comparison of a few of the analogs of V7B with its parent compound is shown in figure 12. The absence of acetamide moiety in compound V7B-9 (IC_{50} – 2.9 µg/ml) displayed less potency in comparison to its parent compound V7B (IC_{50} – 1.8 µg/ml). Compound V7B-4 contains 2-amino-3-chloro-1,4-napthoquinone moiety attached to benzene sulphonamide, which resulted in poor IC_{50}. Substitution of methane diamine by benzene sulfonamide attached to 2-amino-3-chloro-1,4-napthoquinone moiety in compound V7B-8 also resulted in poor IC_{50} (12.4 µg/ml) in comparison to its parent molecule. Replacement of benzene sulfonamide with sulfonyl benzene and modification of 1,4-dihydronapthalene-1-one to 1,4-dihydronapthalene-1,4-dione in parent compound resulted in reduced potency to inhibit VirS DNA binding activity as observed in compound V7B-5. Thus, the structural comparisons of all these molecules with their activity emphasize the importance of acetamide ring attached to benzene sulphonamide, which might be essential for the potent inhibition of DNA binding activity of VirS. Secondly, it was also found that bulkier modifications to acetamide group attached to benzene sulfonamide may lead to poor inhibition of VirS DNA binding activity as observed in the cases of V7B-10, V7B-13, V7B-15 and V7B-17.

Amongst the procured analogs of compound V15B, only one compound was found to be functionally active against the VirS DNA binding activity whose structure comparison with its parent compound is shown in figure 13. Replacement of isoquinoline ring from the parent molecule (V15B) with benzene ring, pyridine ring and methyl pyridine displayed poor inhibition of VirS DNA binding activity in comparison to its parent molecule as observed in compound V15B-1, V15B-3, V15B-5 and V15B-6. Further modification of isoquinoline ring to 1, methyl-3,4-dihydroisoquinoline-4-one and its attachment to 2-(cyclohexa-1,3-dien-1-yl) pyridine moiety in compound V15B-2 showed potential to inhibit DNA binding activity of VirS, however, it displayed a slightly less IC_{50} than the parent molecule. It was also observed that modifications to 2-(4,5 dihydro-1H pyrrol-2-y) pyridine moiety failed to show any inhibition against DNA binding activity of VirS as observed in the case of V15B-8 and V15B-10. Thus, the structural comparisons of all these molecules with their activities emphasize the importance of isoquinoline ring attached to 3-nitroso-2 phenyl-pyrrole as present in the parent molecule, which is essential for the inhibition of DNA binding activity of VirS.

Discussion

In this study, we attempted to dissect the role of VirS in the survival of *M. tuberculosis* under acidic stress and in the maintenance of pH homeostasis of the pathogen. We found that the loss of *virS* reduced *M. tuberculosis's* ability to block phagosomal-lysosomal fusion in the activated macrophages as well as its ability to survive in acidic conditions and its absence also led to a disturbance in the

maintenance of the intrabacterial pH under acidic stress and inside activated macrophages thereby suggesting the role of VirS in evading stressful acidic conditions.

To understand how VirS mediates the regulation of gene expression in acid stress, microarray studies were carried out. Transcriptome profiling has provided insights into the role of VirS in the regulation of various pathways such as those involved in cellular metabolism; cell wall and lipid metabolism; PE/PPE family proteins; energy pathways and respiration; efflux pumps and ion transporters; detoxification; information pathways and transcription regulators in acidic stress.

Firstly, we found a striking influence of VirS on the regulation of a number of genes (induced or repressed) involved in fatty acid oxidation and biosynthesis which can contribute to the remodelling of *M. tuberculosis* cell wall. It has been a general belief that the cell wall of mycobacteria remodels itself when exposed to hostile environments to become impervious to outside harsher conditions (14,34). We observed VirS dependent upregulation of genes involved in fatty acid β-oxidation (*fadB2, fadE35*), lipase and hydrolyzing enzymes (*Rv2672, Rv0867c*) and membrane associated proteins (*Rv2911, Rv0625, Rv0531, Rv2180)*; and downregulation of genes encoding enzymes involved in polyketide synthase (*pks2, pks3* and *pks6, papA1*), membrane proteins (*Rv0514, mmpS3, Rv2869c, Rv0412c, Rv2091c, Rv0677c*), phenolic glycolipids (*fadD22*), biosynthesis and export of siderophores (*mmpS5, mmpL5*) and lipoprotein (*Rv2290*). VirS regulated expression of genes involved in lipid metabolism suggests its pivotal role in remodelling of cell envelope that can facilitate intracellular survival under acidic stress.

Our expression profile also revealed the regulation of metabolic enzymes by VirS at acidic pH. It is reported that there is an induction of genes involved in metabolism via anaplerotic node at acidic pH in *M. tuberculosis*. We observed VirS dependent induction of glyoxylate shunt comprising of isocitrate lyase (*icl*) at acidic pH which replenishes intermediates of TCA cycle and gluconeogenesis and decreases carbon flux through TCA cycle (35). The most upregulated gene by VirS in acid stress is *Rv1405c* which is also reported previously to be induced under acid stress and is implicated in adaptation of *M. tuberculosis* to the intramacrophage environment, however, the substrate for Rv1405c is still not known (36-38). It has been reported that fermentative pathways get upregulated under acidic stress in *B. subtilis* and *B. cereus* (39,40). The genes for alcohol dehydrogenase were found to be induced that catalyze the conversion of pyruvate to ethanol generating CO_2 and dissipating H^+ to deal with low intracellular pH or restoration of NAD+/NADH balance (41). Consistent with this fact, we found VirS dependent induction of pyruvate decarboxylase (*pdc)* and alcohol dehydrogenases (*adhE1, Rv1530*) at acidic pH which may reduce acidity through NAD(P)H that transfers electrons to electron transport chain and pumps protons out of the cell (40). We also found a strong influence of VirS on the expression of genes related to cellular metabolism. Genes encoding amino acids feeding into TCA cycle (*argJ*), nitrate reductase (*narX*), for the synthesis of heme and porphyrin (*hemB*) were found to be downregulated by VirS at acidic pH. Genes for sulfur reduction are essential for the synthesis of cysteine, methionine, mycothiol, sulfolipids and these reduced sulfur products contribute to *M. tuberculosis* pathogenesis and we found that genes for sulfur reduction (*sirA, cysH*) were also downregulated in a VirS dependent manner. This data suggests that changes in the central metabolism by VirS enable *M. tuberculosis* to adapt to the acidic intraphagosomal environment.

In the context of maintenance of intrabacterial pH, we observed VirS dependent induction of efflux pumps and ion transporters (*kdpC, ctpG*). It is reported that a decrease in pH leads to reduction in membrane potential and K^+ uptake which induces the genes encoding K^+ transporters (42). Induction of potassium transporters is necessary for the maintenance of internal pH homeostasis and promote survival in the acidic environment (42). Kdp is a K^+ uptake system known to be induced in acidic environment which may be required for bacterial adaptation at low pH levels (42). As Kdp is a K^+/H^+ antiporter, a maintenance of cytoplasmic pH is expected by pumping protons out of the cell. The metal transporter CtpG that transports Cd^{2+} across the mycobacterial cell membrane is also regulated by VirS (43). Such transporters are

previously reported to be critical for ion homeostasis and bacterial survival (43,44). Thus, regulation of these cell wall transporters suggests the role of VirS in response to phagosomal acidification and in the maintenance of intrabacterial pH.

Transcriptome profiling also revealed that VirS further activates other transcriptional regulators (*Rv1994c, Rv3765c*), thus, it might be acting as a master regulator under these conditions inducing the acidic responses of the cell. This is in agreement to the earlier study by Talaat *et. al.* which implicated that VirS could be the master regulator of reactivation of tuberculosis (15). CmtR (Rv1994c) is AsrR-SmtB repressor that senses cadmium and modulate the expression of genes related to the efflux and detoxification of this metal ion (45). The regulation of CmtR and CtpG by VirS indicates its importance in the efflux of the cadmium ions at acidic pH to regulate ion homeostasis.

A downregulation in the VirS dependent gene expression was observed in ribosome biogenesis pathway (*rpsk*) and protein translation (*infC*). Since, the genes involved in ribosome biogenesis and protein translation are the major consumer of cellular energy, their downregulation is also expected (46,47). Repression of these genes indicates a slowing down of the bacterial growth at acidic stress which is mediated by VirS. It is well understood that there is a strong connection between acid stress and oxidative stress (40,48) and our transcriptome profiling revealed VirS dependent upregulation of genes involved in protection from oxidative stress *i.e. ahpC, ahpD, Rv3177,* which may aid bacterial survival at acidic pH. We also observed a decreased expression of aerobic respiration (*nuok*) and such repression at acidic pH is also documented earlier (49). We also found VirS dependent downregulation of cytochrome assembly ATP-binding protein (*cydD)* at acidic pH which promotes proton export for the maintenance of intrabacterial pH (40).

Based on the analysis of our transcriptional profile, we propose a model for the survival of *M. tuberculosis* under acidic conditions mediated by VirS, although, further investigations are needed to confirm these proposed mechanisms and the involvement of the downstream genes of VirS in pH homeostasis (Figure 14).

The importance of VirS for the survival of bacterial cells under acidic stress and its role in regulating multiple downstream targets signifies it as an attractive drug target for the development of new anti-TB drugs. Towards this, we characterized it further by evaluating the DNA binding activity of VirS and by determining the DNA sequence to which VirS binds by employing DNA footprinting. Further, we generated a consensus logo by WebLogo to map the conservation of nucleotides in the employed sequences (Figure 5C) (16). Additionally, we performed EMSA with the upstream regions of a few downstream targets of VirS, which were found to exhibit direct interaction with VirS as evident by a shift in the EMSA studies (Figure 5D).

Site-directed mutational studies revealed four residues *i.e.* R16, R20, Y61 and R69 as the critical residues for binding of VirS to its cognate DNA. Residue Y61 was found to be a conserved residue in other homologous proteins. R20 of VirS corresponds to R46 in MarA which penetrates into the major groove of DNA and makes hydrogen bonds with bases. Similarly, R69 of VirS is homologous to R96 of MarA which makes hydrogen bonding with the base of DNA. Similarly, R16 of VirS corresponds to W42 in MarA; side chains of W42 are involved in the van der Waals contacts with the DNA. The essentiality of the identified critical residues for DNA binding activity suggests their importance in the interaction of VirS with the DNA. Hence, these residues could be utilized for rational designing of inhibitors to target the DNA binding activity of VirS.

Since, VirS appears as an attractive target for the discovery of new inhibitors against *M. tuberculosis*, we performed structure-based virtual screening for the identification of small molecules against the DNA binding activity of VirS, which may serve as hit molecules for novel anti-TB drug development. We identified two hit molecules *i.e.* V7B and V15B which exhibited potent inhibition of VirS activity, inhibited the growth of *M. tuberculosis in vitro* as well as inside macrophages. However, IC_{50} and MIC values of both of these compounds did not correlate well and the most probable reason for such disparity could be attributed to the chemical nature

of the potent enzyme inhibitors leading to their poor penetration into *M. tuberculosis* cells resulting in weak MIC values. Also, the target specificity of these compounds inside *M. tuberculosis* still needs to be validated. Hence, these two compounds can be further explored to increase their potency and could be a subject of lead optimization.

Thus, we have explored the promising drug target VirS by identifying a few crucial residues for its activity that could help in designing of rational inhibitors as well as by identifying few moieties that inhibit VirS activity as well as *M. tuberculosis* growth. This is the first report of exploring VirS as a target for drug development against *M. tuberculosis.*

Experimental Procedures

Oligos for polymerase chain reaction and for site directed mutagenesis, FITC, IFN-γ, Poly dI-dC, Tyloxapol and complete mini EDTA-free protease inhibitor cocktail tablets were purchased from Sigma-Aldrich Co. (St. Louis, Missouri, USA). DpnI and Pfu polymerase were from NEB (Ipswich, Massachusetts, USA). Cy5 labeled double stranded DNA was customized from IDT (Coralville, Iowa, USA). Nickel nitrilotriacetic acid (Ni-NTA) was obtained from IBA (IBA life sciences, Goettingen, Germany). RPMI-1640, DMEM medium, LysoTracker Red DND-99, Prolong Gold antifade reagent with DAPI, Antibiotic-Antimycotic and FBS (fetal bovine serum) were obtained from Thermo Fisher Scientific (Waltham, Massachusetts, USA). Difco Middlebrook (MB) 7H9 media, 7H11 agar and ADC were obtained from Becton Dickinson and Company (Franklin Lakes, New Jersey, USA). The compounds were obtained from the Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health (Bethesda, Maryland, USA).

Bacterial strains and growth conditions

M. tuberculosis strains were grown in MB7H9 media supplemented with 1X ADC, 0.2% glycerol and 0.2% tween 80 with a constant shaking at 200rpm or on 7H11 agar plates containing 10% oleic acid-albumin-dextrose-catalase (OADC) and 0.2% glycerol at 37ºC. The pH of MB7H9 media was adjusted by using HCl and instead of Tween80, 0.02% Tyloxapol was used, wherever required. Hygromycin B and kanamycin were used at a concentration of 50 μg/ml and 25 μg/ml, respectively for mycobacteria. Phosphate-citrate buffers were prepared from 200 mM sodium phosphate and 100 mM citric acid. *Escherichia coli* (*E. coli*) strains XL-1 Blue and BL21(λDE3) were grown in Luria Bertani (LB) broth or on LB agar. Ampicillin was used at a concentration of 50 μg/ml for *E. coli*.

Growth kinetics and colocalization studies by confocal microscopy

virS mutant of *M. tuberculosis* and *virS* complemented strain were earlier generated in our laboratory (12,13). The *M. tuberculosis* Erdman, *virS* mutant and *virS* complemented strain were grown in MB7H9 media adjusted to a pH of 6.6, 5.5, 5.0 and 4.5 at 37ºC with a constant shaking at 200 rpm. The growth was monitored daily by measuring the $O.D._{600nm}$. In case of CFU studies, the cells were appropriately diluted and plated at various time points on MB 7H11 agar plates. The plates were kept at 37ºC for 3-4 weeks for the appearance of colonies. The phagosomal maturation arrest of *M. tuberculosis* was studied as described previously (50). For colocalization studies in activated macrophages, PMA activated macrophages were washed once with plain RPMI-1640 media and cells were further incubated for 24 hours in RPMI-1640 medium containing 10% FBS and 10 ng/ml of IFN-γ before infection. After infection and lysotracker red treatment, coverslips were mounted by using Prolong Gold antifade reagent and visualized by Leica TCS Sp5 confocal laser scanning microscope (Leica Microsystems, Mannheim, Germany). The percent colocalization of FITC labelled mycobacteria with Lysotracker Red was calculated by analyzing ~100 phagosomes.

Intrabacterial pH measurement Assays

The plasmid encoding ratiometric pH-GFP (a kind gift from Dr. Sabine Ehrt, Department of Microbiology and Immunology, Weill Cornell Medical College, New York, USA; received from Dr. Amit Singh, Department of Microbiology and Cell Biology, Indian Institute of Science, Bengaluru, India) was electroporated into

M. tuberculosis Erdman, *virS* mutant and complemented strain and the recombinants were selected on the basis of fluorescence of GFP on 7H11-OADC-hygromycin B plates. The selected recombinant clones were grown in MB7H9 medium in the presence of hygromycin till $O.D._{600nm}$ of 0.3-0.5. The cultures were harvested and pellets were resuspended in phosphate-citrate buffer of pH 7.4 and 4.5 supplemented with ADC, 0.2% tween 80 and hygromycin B. The intrabacterial pH measurements were carried out after incubation of the bacteria in the respective buffers for different time intervals. The bacterial cultures at different time points were harvested and concentrated to 20 times by centrifugation to increase GFP signal. Fluorescence was analyzed in a Cary Varian spectrofluorometer with excitation wavelengths of 395 nm and 475 nm at an emission wavelength of 510 nm. The intrabacterial pH was derived by interpolating the ratio of 395/475 fluorescence intensity on a standard curve. The standard curve for intrabacterial pH measurement was generated by incubating the lysate containing GFP with the phosphate-citrate buffers of pH ranging from 5.5 to 8.5 in a 96 well plate (1). The plate was read at excitation wavelengths of 395nm and 475nm and an emission wavelength of 510nm.

For pH_{IB} measurements of bacteria residing inside macrophages, 5 X 10^5 PMA differentiated THP-1 cells were seeded in poly-L-lysine coated glass bottom culture dishes of 35mm thickness. PMA activated macrophages were washed once with plain RPMI-1640 media and cells were further incubated for 24 hours in RPMI-1640 medium containing 10% FBS and 10 ng/ml of IFN-γ before infection. Macrophages were infected with *M. tuberculosis* at an MOI of 1:5, (Macrophage:Bacteria) for 4 hours followed by amikacin treatment of one hour. After different time points post infection, macrophages were analysed under confocal microscopy to determine the intensities of GFP (exciters D405 and D476, emitter 535). All the images were acquired within an experiment under identical conditions. A bacterial group was defined as atleast one bacterium but may consist of 2-5 bacilli. The bacterial groups were analysed for their emission intensities when excited at 405nm and 476nm and the ratios of intensities at 405:476 were calculated. pH_{IB} were derived by interpolating the 405:476 intensity ratios on a standard curve.

Microarray Studies

M. tuberculosis Erdman, *virS* mutant and *virS* complemented strain were grown till an $O.D._{600nm}$ of 0.3-0.5 and cells were harvested by centrifugation at 4500rpm for 10 minutes. Harvested cells were exposed to MB7H9 media adjusted to pH 6.6 and pH 4.5 and grown at 37ºC at a constant shaking of 200 rpm for 1 hour. Total RNA was isolated from all the mycobacterial strains (grown at pH 6.6 and pH 4.5) and was processed and hybridized to *M. tuberculosis* whole genome Gene Expression Profiling microarray G2509F (AMADID: G2509F_43893, Agilent Technologies). DNA microarrays and data analysis were performed at University of Delhi South Campus Microarray Centre (UDSCMAC). Slides were scanned on a microarray scanner (Agilent Technologies) and analyzed using the GeneSpring software.

Cloning and expression of *virS*

The *virS* gene was PCR amplified by employing *M. tuberculosis* H37Rv genomic DNA as the template and by using the primers F-*virS* containing *spe*I, thrombin site, *Eco*RI followed by 6X his tag and R-*virS* containing *Hind*III site. The sequence of the primers is given in Table S2. The amplified PCR product after digestion with *spe*I and *Hind*III was cloned into pET43.1a vector digested with the same enzymes to create pET43.1a-*virS* (NusA-VirS) for the synthesis of N-terminal His tagged VirS. To subclone *virS* from pET43.1a-*virS* construct into vector pBEn-SET3a, the positive clone of pET43.1a-*virS* was digested with *Eco*RI and *Hind*III to obtain a fallout of ~1.0kb corresponding to the size of *virS*. The pBEn-SET3a vector was also digested with the same enzymes. The digested vector pBEn-SET3a and digested *virS* were ligated to obtain the recombinant clone pBEn-*virS* (3a-VirS). Positive clones were confirmed by restriction digestion. For expression, *E. coli* BL21(λDE3) cells transformed with pET43.1a-*virS* and pBEn-*virS*, were grown at 37ºC in Luria Bertani media containing 50 μg/ml ampicillin till the $O.D._{600nm}$ of 0.8. The cultures *i.e.* NusA-VirS and 3a-VirS, were then induced with 1mM isopropyl-1-thio-β-D-galactopyranoside (IPTG)

and were allowed to grow for 16 hours at 16ºC and 18ºC, respectively. The cells were harvested by centrifugation at 4ºC, 6000g for 10 minutes.

Purification of 3a-VirS

E. coli BL21(λDE3) cells were transformed with pBEn-*virS* and the cells were grown and induced as described above. For purification, the cells from the induced culture were harvested and resuspended in lysis buffer containing 1X PBS pH 7.4, 10 mM imidazole, 5 mM β-mercaptoethanol, 1 mM PMSF (phenylmethylsulfonylfluoride) and complete mini EDTA-free protease inhibitor cocktail tablet and lysed by sonication followed by centrifugation to remove cell debris (15000g, 45 minutes, 4ºC). Clarified lysate was loaded onto a Ni-NTA resin pre-equilibrated with lysis buffer. Washing of the column was performed initially with buffer containing 1X PBS, 10 mM imidazole followed by one wash each with the same buffer containing 20 mM and 50 mM imidazole. The protein was eluted with buffer containing 250 mM imidazole and the purity of the protein was analysed by SDS-PAGE by using a 12.5% polyacrylamide gel. The fractions containing VirS were pooled, and 10% glycerol was added. The pooled protein was dialysed overnight against 20 mM tris pH-8.0 and stored at -80ºC.

Gel Shift Assay for NusA-VirS fusion protein

DNA probes for gel mobility shift assays were amplified by PCR from vector pSG10 comprising of the intergenic region between *virS* and *mymA*, which is hypothesized to be bound by VirS (12). Two primers (UP6 and DW2) were utilized for the amplification of a 325bp intergenic region. (Table S2). Primer DW2 was end labelled with [γ-^{32}P] ATP by using T4 polynucleotide kinase. The PCR product was run on 1.5% agarose gel and the amplified band was cut and gel eluted by using GFX columns and quantitated by β counter (TRI-CARB 2900TR Liquid Scintillation Counter, Perkin Elmer). Lysates of *E. coli* cells overexpressing either the NusA-VirS fusion protein or only NusA protein were prepared. The reaction mixture consisted of 1,20,000 cpm of labelled DNA, binding mix and different amounts *i.e.* 13µg, 26µg and 52µg of the lysate. The binding mix consisted of 1X binding buffer, 1mg/ml BSA, 1mM DTT, 4% PEG-8000 and 1µg of Poly dI-dC. The samples were incubated for 30 minutes at 37ºC. Samples were then electrophoresed on 3.5% polyacrylamide gel, dried and autoradiographed.

DNA Footprinting Assay

Binding reaction mixture was prepared as described above in a total volume of 30 µl that consisted of 126,000 cpm of labelled DNA probe and 52 µg of lysate and incubated at 37ºC for 30 minutes. After incubation, the binding reaction mix was subjected to DNase I digestion (0.2 U) for 1 minute and digestion was terminated by the addition of 180 µl of stop solution (300 mM sodium acetate pH-7.0, 10 mM EDTA pH-8.0, 0.3 mg/ml yeast t-RNA, 0.01% SDS). Samples were extracted with phenol-chloroform and precipitation of nucleic acid was performed with ethanol. DNA was resuspended in water and formamide dye was added. G+A ladder was also prepared by using the same probe and formamide dye was mixed. The samples and G+A ladder were pre-heated at 90ºC for 10 minutes before loading. These samples were separated on 6% polyacrylamide sequencing gel and electrophoresed (in 0.5X TBE buffer and 0.5M sodium acetate in lower chamber and 0.5X TBE in upper chamber) at 1050V. Gel was dried and exposed for 14 hours followed by autoradiography.

Homology modelling of VirS

Amino acid sequence similarity search was performed by using NCBI-BLAST server in PDB structure database. Based on the BLAST results, first few homologs of VirS were identified namely ToxT of *Vibrio cholera* PDB ID-3GBG (51), gadX of *E. coli* PDB ID-3MKL (unpublished data), AdpA of *Streptomyces griseus* PDB ID-3W6V (27), unknown protein of *Listeria innocula* PDB ID-3OOU (unpublished data), AraC of *E. coli* PDBID-2K9S (52), MarA of *E. coli* PDB ID-1BL0 (25), yesn of *Fusobacterium nucleatum* PDB ID-3LSG (unpublished data), Ternary complex of MarA, DNA and C-terminus of RNA polymerase of *E. coli* PDB ID-1XS9(53), Rob of *E. coli* PDB ID-1D5Y (26), transcription regulator of *Chromobacterium violaceum* PDB ID-3OIO (unpublished data) and XylR of *E. coli* PDB ID-4FE4 (54). The sequences of all these homologs were aligned with C-terminal region (250-340 amino acids) of *M. tuberculosis* VirS sequence by using the software ClustalW. The result of sequence

alignments was then submitted to the software Modeller 9.0 to generate a 3-D structure of VirS and the generated homology model was validated by employing Rampage, Errat, Verify 3D, Pro-Q and Prosa softwares.

Site-directed mutagenesis (SDM)

Amino acid substitutions were generated by site directed mutagenesis kit provided by Stratagene (La Jolla, California, USA) according to the manufacturer's instructions. The *E. coli* XL1 Blue cells were transformed to get the colonies on ampillicin plates. These colonies were grown in LB media and DNA was isolated. The mutations in the DNA were confirmed by DNA sequencing. The primers used in mutagenesis are listed in Table S2.

Electrophoretic mobility shift assay (EMSA) for purified 3a-VirS

0.5 µg of purified 3a-VirS protein was taken in a reaction mixture that contained 1X binding buffer (10 mM Tris acetate pH 8.0, 50 mM KCl, 5% glycerol), 4% PEG, 1µg Poly dI-dC, 100 µg/ml BSA, 1mM DTT and 0.8 picomoles of 5` Cy 5 fluorophore labelled DNA and incubated for 30 minutes at 37°C. This labelled double stranded DNA represents the sequence to which VirS binds as identified by DNA footprinting study with 10bp flanking regions on both the sides (66 bp dsDNA) which was labelled with Cy 5 fluorophore at the 5` ends. After incubation, samples were loaded onto a 5% polyacrylamide gel and electrophoresed in 1X TAE buffer for overnight at 32V. Next day, fluorescence was detected by phosphorimager (model-FLA-9000, FUJIFILM Corporation, Minato-ku, Tokyo, Japan).

Regulation of downstream genes by VirS using EMSA.

Upstream regions of the genes were PCR amplified by using Cy5 PCR Labelling Kit (Jena Biosciences, Germany). The amplified products were gel eluted and quantified. 40 fmoles of labelled product was incubated with 1µg of purified VirS in the binding buffer (composition same as mentioned above) for 30 minutes at 37°C. Samples were electrophoresed onto a 5% polyacrylamide gel in 1X TAE buffer for overnight**.** The bands were visualized using phosphorimager.

Virtual Screening

The 3-D structure generated by Modeller 9.0 was superimposed with the crystal structure of MarA (PDB ID-1BL0) and the co-ordinates of DNA bound to MarA structure were used to predict the DNA contact sites of VirS. Virtual Screening was carried out against both the DNA contact sites by using small molecule library of NCI Open Database consisting of 260,071 compounds. These compounds were shortlisted by using FAF server that resulted in 95,748 compounds which were used for screening against the DNA contact sites of VirS by using Autodock4.2 (30). Top compounds were procured from NCI-DTP depending upon the availability for conducting inhibition studies. 57 compounds each on the basis of docking at site 1 and site 2, were procured from NCI-DTP. Compounds obtained from docking at site 1 were denoted as V1A-V57A. Compound V7A was not proceeded further due to its toxicity. Compounds obtained from docking at site 2 were denoted as V1B-V57B.

Assay for inhibition of DNA binding activity of VirS

All the compounds were dissolved in DMSO to prepare a stock solution of 2 mg/ml and were screened for their ability to inhibit the DNA binding activity of VirS at a concentration of 100 µg/ml. The reaction mixture was same as described above. Binding buffer, purified protein (3a-VirS) and the individual test compounds were incubated at 37°C for 30 minutes followed by the addition of rest of the components and further incubated for 30 minutes at 37°C. Rest of the steps were same as mentioned above. Sample containing DMSO was used as a negative control for the assay. IC_{50} is defined as the concentration of the compound which exhibits 50% inhibition of the activity.

Assay for inhibition of growth of *M. tuberculosis*

The inhibition of *M. tuberculosis* growth was evaluated by employing REMA as described previously (30,55). Blue colour depicts no growth while conversion to pink colour depicts viable cells. Fluorescence was measured at an emission

wavelength of 590 nm with an excitation wavelength of 530 nm and percent inhibition was calculated. Rifampicin was employed as a positive control whereas wells with DMSO were used as a negative control. MIC_{90} is defined as the concentration of compound at which growth of the pathogen was inhibited by 90%.

Evaluation of cytotoxicity of the compounds

Compounds were evaluated for their cytotoxicity against three mammalian cell lines namely THP-1 (human monocytic cell line), HEK (human embryonic kidney cell line) and HeLa (cervical cell line) by employing the protocol as described previously (30,55). IC_{50} is defined as the concentration of the compound at which 50 % growth of the cell line is inhibited.

Evaluation of inhibitory potential of the compounds against intracellular survival of *M. tuberculosis* inside macrophages

The ability of the compounds to inhibit intracellular growth of the pathogen was evaluated as described previously (56).

Acknowledgements

We thank University of Delhi South Campus Microarray Center (UDSCMAC), New Delhi and Bionivid Technologies Pvt Ltd, Bengaluru for conducting microarray experiments and data analysis. We thank Dr. Sabine Ehrt, Department of Microbiology and Immunology, Weill Cornell Medical College, New York and Dr. Amit Singh, Department of Microbiology and Cell Biology, Indian Institute of Science, Bengaluru, India for providing us the plasmid encoding ratiometric pH-GFP. We are thankful to Dr. Amit Singh for the critical reading of the manuscript. Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA is highly acknowledged for providing us the compounds. We thank Central Instrumentation Facility (CIF), UDSC for providing the grid facility for docking studies. CIF is also acknowledged for confocal microscopy studies and FLA-9000 phosphorimager. Priti Singh and Tannu Priya Gosain are acknowledged for their excellent technical help. DNA sequencing facility, UDSC is acknowledged for DNA sequencing. Distributed information subcenter, DBT, UDSC is acknowledged for providing computers and internet services. Swati Singh is thankful to Indian Council of Medical Research, India for providing fellowship. This work was supported by financial grant from the Department of Biotechnology, Government of India (BT/01/COE/05/06-I), J.C. Bose fellowship from Department of Science and Technology, Government of India awarded to Prof. Anil K. Tyagi (No. SR/S2/JCB-39/2009(G)) and Innovative Young Biotechnologist Award project from Department of Biotechnology, Government of India awarded to Dr. Garima Khare (BT/07/IYBA/2013-5).

Conflict of Interest

Authors declare that they have no conflict of interest.

Author Contributions

SS, GK and AKT designed the experiments. SS and NG performed the experiments. SS wrote the manuscript; GK and AKT proofread, corrected and edited the manuscript. GK and AKT provided overall supervision of the work.

References

1. Vandal, O. H., Pierini, L. M., Schnappinger, D., Nathan, C. F., and Ehrt, S. (2008) A membrane protein preserves intrabacterial pH in intraphagosomal *Mycobacterium tuberculosis*. *Nat. Med.* **14**, 849-854
2. Song, H., Huff, J., Janik, K., Walter, K., Keller, C., Ehlers, S., Bossmann, S. H., and Niederweis, M. (2011) Expression of the ompATb operon accelerates ammonia secretion and adaptation of *Mycobacterium tuberculosis* to acidic environments. *Mol. Microbiol.* **80**, 900-918
3. Buchmeier, N., Blanc-Potard, A., Ehrt, S., Piddington, D., Riley, L., and Groisman, E. A. (2000) A parallel intraphagosomal survival strategy shared by *Mycobacterium tuberculosis* and *Salmonella enterica*. *Mol. Microbiol.* **35**, 1375-1382
4. Raynaud, C., Papavinasasundaram, K. G., Speight, R. A., Springer, B., Sander, P., Bottger, E. C., Colston, M. J., and Draper, P. (2002) The functions of OmpATb, a pore-forming protein of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **46**, 191-201
5. Athmann, C., Zeng, N., Kang, T., Marcus, E. A., Scott, D. R., Rektorschek, M., Buhmann, A., Melchers, K., and Sachs, G. (2000) Local pH elevation mediated by the intrabacterial urease of *Helicobacter pylori* cocultured with gastric cells. *J. Clin. Invest.* **106**, 339-347
6. Marcus, E. A., Sachs, G., and Scott, D. R. (2013) The role of ExbD in periplasmic pH homeostasis in *Helicobacter pylori*. *Helicobacter.* **18**, 363-372

7. Stincone, A., Daudi, N., Rahman, A. S., Antczak, P., Henderson, I., Cole, J., Johnson, M. D., Lund, P., and Falciani, F. (2011) A systems biology approach sheds new light on *Escherichia coli* acid resistance. *Nucleic Acids Res.* **39**, 7512-7528
8. Martin-Galiano, A. J., Ferrandiz, M. J., and de la Campa, A. G. (2001) The promoter of the operon encoding the F0F1 ATPase of *Streptococcus pneumoniae* is inducible by pH. *Mol. Microbiol.* **41**, 1327-1338
9. Krulwich, T. A., Sachs, G., and Padan, E. (2011) Molecular aspects of bacterial pH sensing and homeostasis. *Nat. Rev. Microbiol.* **9**, 330-343
10. Gupta, S., Jain, S., and Tyagi, A. K. (1999) Analysis, expression and prevalence of the *Mycobacterium tuberculosis* homolog of bacterial virulence regulating proteins. *FEMS Microbiol Lett.* **172**, 137-143
11. Gupta, S., and Tyagi, A. K. (1993) Sequence of a newly identified *Mycobacterium tuberculosis* gene encoding a protein with sequence homology to virulence-regulating proteins. *Gene.* **126**, 157-158
12. Singh, A., Jain, S., Gupta, S., Das, T., and Tyagi, A. K. (2003) mymA operon of *Mycobacterium tuberculosis*: its regulation and importance in the cell envelope. *FEMS Microbiol. Lett.* **227**, 53-63
13. Singh, A., Gupta, R., Vishwakarma, R. A., Narayanan, P. R., Paramasivan, C. N., Ramanathan, V. D., and Tyagi, A. K. (2005) Requirement of the mymA operon for appropriate cell wall ultrastructure and persistence of *Mycobacterium tuberculosis* in the spleens of guinea pigs. *J. Bacteriol.* **187**, 4173-4186
14. Talaat, A. M., Ward, S. K., Wu, C. W., Rondon, E., Tavano, C., Bannantine, J. P., Lyons, R., and Johnston, S. A. (2007) Mycobacterial bacilli are metabolically active during chronic tuberculosis in murine lungs: insights from genome-wide transcriptional profiling. *J. Bacteriol.* **189**, 4265-4274
15. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188-1190
16. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680
17. Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2007) Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci.* **Chapter 2**, Unit 2.9
18. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* **50**, 437-450
19. Wallner, B., and Elofsson, A. (2003) Can correct protein models be identified? *Protein Sci.* **12**, 1073-1086
20. Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science.* **253**, 164-170
21. Luthy, R., Bowie, J. U., and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature.* **356**, 83-85
22. Colovos, C., and Yeates, T. O. (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511-1519
23. Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins.* **17**, 355-362
24. Wiederstein, M., and Sippl, M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407-410
25. Rhee, S., Martin, R. G., Rosner, J. L., and Davies, D. R. (1998) A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc. Natl. Acad. Sci. U S A.* **95**, 10413-10418
26. Kwon, H. J., Bennik, M. H., Demple, B., and Ellenberger, T. (2000) Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat. Struct. Biol.* **7**, 424-430

27. Yao, M. D., Ohtsuka, J., Nagata, K., Miyazono, K., Zhi, Y., Ohnishi, Y., and Tanokura, M. (2013) Complex structure of the DNA-binding domain of AdpA, the global transcription factor in *Streptomyces griseus*, and a target duplex DNA reveals the structural basis of its tolerant DNA sequence specificity. *J. Biol. Chem.* **288**, 31019-31029
28. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **30**, 2785-2791
29. AutoDock website. Available: http://autodock.scripps.edu/downloads/autodock-4.0-registration/autodock-4-2-download-page. Accessed March 24, 2018.
30. Rohilla, A., Khare, G., and Tyagi, A. K. (2017) Virtual Screening, pharmacophore development and structure based similarity search to identify inhibitors against IdeR, a transcription factor of *Mycobacterium tuberculosis*. *Sci. Rep.* **7,** 4653
31. Palomino, J. C., Martin, A., Camacho, M., Guerra, H., Swings, J., and Portaels, F. (2002) Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in *Mycobacterium tuberculosis*, *Antimicrob. Agents Chemother*. **46**, 2720-2722.
32. Martin, A., Camacho, M., Portaels, F., and Palomino, J. C. (2003) Resazurin microtiter assay plate testing of *Mycobacterium tuberculosis* susceptibilities to second-line drugs: rapid, simple, and inexpensive method. *Antimicrob. Agents Chemother*. **47**, 3616-3619
33. Collins, L., and Franzblau, S. G. (1997) Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrob. Agents Chemother*. **41**, 1004-1009
34. Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D., Nathan, C., and Schoolnik, G. K. (2003) Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment. *J. Exp. Med.* **198**, 693-704
35. Baker, J. J., Johnson, B. K., and Abramovitch, R. B. (2014) Slow growth of *Mycobacterium tuberculosis* at acidic pH is regulated by phoPR and host-associated carbon sources. *Mol. Microbiol.* **94**, 56-69
36. Healy, C., Golby, P., MacHugh, D. E., and Gordon, S. V. (2015) The MarR family transcription factor Rv1404 coordinates adaptation of *Mycobacterium tuberculosis* to acid stress via controlled expression of Rv1405c, a virulence-associated methyltransferase. *Tuberculosis (Edinb).* **97**, 154-162
37. Rohde, K. H., Abramovitch, R. B., and Russell, D. G. (2007) *Mycobacterium tuberculosis* invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe.* **2**, 352-364
38. Golby, P., Hatch, K. A., Bacon, J., Cooney, R., Riley, P., Allnutt, J., Hinds, J., Nunez, J., Marsh, P. D., Hewinson, R. G., and Gordon, S. V. (2007) Comparative transcriptomics reveals key gene expression differences between the human and bovine pathogens of the *Mycobacterium tuberculosis* complex. *Microbiology.* **153**, 3323-3336
39. Desriac, N. m., Broussolle, V., Postollec, F., Mathot, A. G., Sohier, D., Coroller, L., and Leguerinel, I. (2013) *Bacillus cereus* cell response upon exposure to acid environment: toward the identification of potential biomarkers. *Front. Microbiol.* **4,** 284
40. Wilks, J. C., Kitko, R. D., Cleeton, S. H., Lee, G. E., Ugwu, C. S., Jones, B. D., BonDurant, S. S., and Slonczewski, J. L. (2009) Acid and base stress and transcriptomic responses in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **75**, 981-990
41. Mols, M., van Kranenburg, R., van Melis, C. C., Moezelaar, R., and Abee, T. (2010) Analysis of acid-stressed *Bacillus cereus* reveals a major oxidative response and inactivation-associated radical formation. *Environ. Microbiol.* **12**, 873-885
42. Cholo, M. C., van Rensburg, E. J., Osman, A. G., and Anderson, R. (2015) Expression of the genes encoding the Trk and Kdp potassium transport systems of *Mycobacterium tuberculosis* during growth *in vitro*. *Biomed. Res. Int.* **2015**, 608682

43. Lopez, M., Quitian, L. V., Calderon, M. N., and Soto, C. Y. (2017) The P-type ATPase CtpG preferentially transports Cd(2+) across the *Mycobacterium tuberculosis* plasma membrane. *Arch. Microbiol.* **200**, 483-492
44. Raimunda, D., Long, J. E., Sassetti, C. M., and Arguello, J. M. (2012) Role in metal homeostasis of CtpD, a Co(2)(+) transporting P(1B4)-ATPase of *Mycobacterium smegmatis*. *Mol. Microbiol.* **84**, 1139-1149
45. Chauhan, S., Kumar, A., Singhal, A., Tyagi, J. S., and Krishna Prasad, H. (2009) CmtR, a cadmium-sensing ArsR-SmtB repressor, cooperatively interacts with multiple operator sites to autorepress its transcription in *Mycobacterium tuberculosis*. *FEBS J.* **276**, 3428-3439
46. Lemke, J. J., Sanchez-Vazquez, P., Burgos, H. L., Hedberg, G., Ross, W., and Gourse, R. L. (2011) Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci. U S A.* **108**, 5712-5717
47. Nomura, M., Gourse, R., and Baughman, G. (1984) Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* **53**, 75-117
48. Maurer, L. M., Yohannes, E., Bondurant, S. S., Radmacher, M., and Slonczewski, J. L. (2005) pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12. *J. Bacteriol.* **187**, 304-319
49. Fisher, M. A., Plikaytis, B. B., and Shinnick, T. M. (2002) Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J. Bacteriol.* **184**, 4025-4032
50. Khare, G., Reddy, P. V., Sidhwani, P., and Tyagi, A. K. (2013) KefB inhibits phagosomal acidification but its role is unrelated to *M. tuberculosis* survival in host. *Sci. Rep.* **3**, 3527
51. Lowden, M. J., Skorupski, K., Pellegrini, M., Chiorazzo, M. G., Taylor, R. K., and Kull, F. J. (2010) Structure of *Vibrio cholerae* ToxT reveals a mechanism for fatty acid regulation of virulence genes. *Proc. Natl. Acad. Sci. U S A.* **107**, 2860-2865
52. Rodgers, M. E., and Schleif, R. (2009) Solution structure of the DNA binding domain of AraC protein. *Proteins.* **77**, 202-208
53. Dangi, B., Gronenborn, A. M., Rosner, J. L., and Martin, R. G. (2004) Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Mol. Microbiol.* **54**, 45-59
54. Ni, L., Tonthat, N. K., Chinnam, N., and Schumacher, M. A. (2013) Structures of the *Escherichia coli* transcription activator and regulator of diauxie, XylR: an AraC DNA-binding family member with a LacI/GalR ligand-binding domain. *Nucleic Acids Res.* **41**, 1998-2008
55. Singh, S., Khare, G., Bahal, R. K., Ghosh, P. C., and Tyagi, A. K. (2018) Identification of *Mycobacterium tuberculosis* BioA inhibitors by using structure-based virtual screening. *Drug Des. Devel. Ther.* **12**, 1065-1079
56. Khare, G., Kumar, P., and Tyagi, A. K. (2013) Whole-cell screening-based identification of inhibitors against the intraphagosomal survival of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **57**, 6372-6377
57. Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., and Wishart, D. S. (2016) Heatmapper: web enabled heatmapping for all. *Nucleic acid Res.* **44**, W147-W153.
58. The Pymol Molecular Graphic System, Version 1.1.eval, Schrodinger, LLC.
59. *Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes, 2016.*

Tables

Table 1- List of VirS dependent genes identified by microarray studies which were found to be upregulated with >2.0 fold at acidic pH. Only a few from the most represented categories are listed here.

Gene	Rv No.	Fold Regulation	Description
Cell wall Associated			
fadB2	Rv0468	8.64	3-hydroxyacyl-CoA dehydrogenase
fadE35	Rv3797	3.88	Acyl-CoA dehydrogenase
Rv2672	Rv2672	2.85	Protease and lipase
Efflux Pumps and Transporters			
ctpG	Rv1992c	2.87	Cation transport ATPase
kdpC	Rv1031	2.69	Potassium-transporting ATPase C chain
Metabolic Enzymes			
Rv1405c	Rv1405c	54.07	similar to phosphatidylethanolamine N-methyltransferase
icl	Rv0467	11.28	isocitrate lyase
pdc	Rv0853c	3.34	pyruvate decarboxylase
adhE1	Rv0162c	2.81	Alcohol dehydrogenase
Detoxification			
ahpD	Rv2429	6.04	member of AhpC/TSA family
ahpC	Rv2428	5.38	alkyl hydroperoxide reductase
Rv3177	Rv3177	4.69	probable non-heme haloperoxidase
Transcriptional regulators			
Rv1994c	Rv1994c	3.22	transcriptional regulator (MerR family)

Table 2- List of VirS dependent genes identified by microarray studies which were found to be downregulated with >2.0 fold at acidic pH. Only a few from the most represented categories are listed here.

Gene	Rv No.	Fold Regulation	Description
Cell wall Associated			
papA1	Rv3824c	5.43	PKS associated protein
fadD9	Rv2590	4.79	Acyl-CoA synthase
pks2	Rv3825c	3.58	Polyketide synthase
mmpS5	Rv0677c	3.59	Membrane protein
mmpL5	Rv0676c	3.23	Membrane protein
Metabolic Enzymes			
lipF	Rv3487c	7.42	Esterase
sirA	Rv2391	2.81	Sulphite reductase
cysH	Rv2392	2.63	3-phosphoadenylsulfate (PAPS) reductase
Information Pathways			
rpsK	Rv3459c	2.74	30S ribosomal protein S11
infC	Rv1641	2.61	Initiation factor IF-3
Energy pathways and respiration			
nuoK	Rv3155	2.81	NADH dehydrogenase chain K
cydD	Rv1621c	2.76	ABC transporter
narX	Rv1736	2.32	Nitrate reductase

Table 3- VirS residues selected for site-directed mutational studies.

S. NO.	Amino Acid Position	Amino acid	Amino acid position in C-terminus of VirS	Mutation
1	254	Glu(E)	5	E5A
2	265	Arg(R)	16	R16A
3	268	Gln(Q)	19	Q19A
4	269	Arg(R)	20	R20A
5	277	Arg(R)	28	R28A
6	279	His(H)	30	H30A
7	283	Glu(E)	34	E34A
8	287	Arg(R)	38	R38A
9	310	Tyr(Y)	61	Y61A
10	318	Arg(R)	69	R69A
11	319	Ser(S)	70	S70A
12	321	Arg(R)	72	R72A
13	322	Arg(R)	73	R73A
14	329	Arg(R)	80	R80A
15	330	Gln(Q)	81	Q81A

Table 4: List of the compounds that displayed *M. tuberculosis* growth inhibition.

S. No.	COMPOUND	MIC_{90} (µg/ml) exhibited by compounds against *M. tuberculosis* growth in *in vitro* broth culture	IC_{50} (µg/ml) exhibited by compounds in EMSA
1	V1A	160	72.01
2	V4A	40	58.33
3	V6A	25	34.59
4	V8A	20	33.63
5	V44A	12.5	24.91
6	V48A	50	27.86
7	V55A	20	74.86
8	V7B	40	1.8
9	V8B	55% inhibition at 200 µg/ml	17.44
10	V11B	61% inhibition at 160 µg/ml	50.45
11	V15B	5	31.21
12	V20B	80	38.44
13	V22B	61% inhibition at 80 µg/ml	72.44
14	V56B	25	37.19

Figure Legends

Figure 1: **Involvement of VirS in *M. tuberculosis* survival at acidic pH.** *In vitro* growth curves of *M. tuberculosis* Erdman (red), *virS* mutant (green) and *virS* complemented (blue) strain under varying pH conditions. The X axis represents number of days while Y axis represents $O.D._{600\ nm}$.. The values are represented as the mean (±SD) of at least two independent experiments. (*p<0.05, **p <0.01, ***p<0.001, two-way ANOVA).

Figure 2: **Role of VirS in phagosomal maturation.** Influence of disruption of *virS* on the ability of *M. tuberculosis* to arrest phagosomal acidification was evaluated in THP-1 macrophages activated with IFN-γ. Macrophages were infected with FITC labelled *M. tuberculosis* strains (green) separately. Blue color represents DAPI staining of the cells. The cells were then subjected to LysoTracker Red staining (red colour) followed by observation under confocal microscope. (B) Bar diagram represents percent co-localization of phagosomes containing FITC labelled bacteria with Lysotracker Red. Data represents the mean (±SD) of 2 independent experiments carried out in triplicates (**p <0.01, ***p<0.001, One-way ANOVA).

Figure 3: *virS* disruption causes change in intrabacterial pH of the pathogen. The intrabacterial pH of the wild type, *virS* mutant and the *virS* complemented strain was measured in near neutral conditions *i.e.* pH 7.4 (A) and in acidified medium *i.e.* pH 4.5 (B). Intrabacterial pH was measured at different time intervals *i.e.* 0, 8 and 24 hours after exposure to acidic pH. Disruption of VirS caused a reduction in intrabacterial pH at pH4.5, which was maintained in the wild type and complemented strain. Data represents the mean (±SD) of atleast 2 independent experiments carried out in duplicates (***p<0.001, two-way ANOVA). The intrabacterial pH of the wild type, *virS* mutant and *virS* complemented strain was measured inside IFN-γ activated macrophages after 24 hours post infection (C) and 48 hours post infection (D). *virS* mutant failed to maintain its intrabacterial pH inside activated macrophages whereas wild type and *virS* complemented strain were able to maintain their $pH_{IB.}$ There were ~115 bacterial groups examined and data represent two independent experiments.

Figure 4: VirS dependent gene regulation at acidic pH. Microarray studies revealed VirS mediated differential gene regulation at acidic pH. Upregulation of VirS at acidic pH causes change in the expression of many genes that were found to be induced and repressed at acidic pH. These genes were categorized into various classes based on their description given by Tuberculist. (A) Pie chart represents the categorization of various genes which are induced by VirS at acidic pH. (B) Pie chart represents the categorization of various genes which are repressed by VirS at acidic pH. The percent distribution of genes is represented in both the pie charts. (C) Heat map showing the expression of VirS regulated genes in *M. tuberculosis Erdman*, *M.tbΔvirS* and *M.tbΔvirSComp* in response to acidic pH relative to neutral pH (>2 fold difference, p <0.05). The figure here depicts complete linkage and was generated by employing the software Heatmapper (57).

Figure 5: Determination of DNA binding activity of VirS and its binding region. (A) Determination of DNA binding activity of VirS by employing EMSA. The DNA binding activity of lysate overexpressing fusion protein NusA-VirS was analyzed with 325bp amplified P^{32} radiolabelled DNA fragment by EMSA. The lysate overexpressing only NusA tag and samples containing DNA (no protein) were utilized as negative controls. Lane 1, 2, 3 represent 13 µg, 26 µg and 52 µg of *E. coli* lysate overexpressing NusA-VirS protein, Lane 4, 5, 6 represent 13 µg, 26 µg and 52 µg of *E. coli* lysate overexpressing NusA protein, Lane 7 represents Free DNA. (B) VirS binding region of DNA was determined by DNase I footprinting assay. Lanes 1&9 – G+A DNA ladder, Lane 2 – Free DNA, lanes 3-5 – DNA bound to NusA-VirS, lanes 6-8 – DNA in the presence of *E. coli* lysate overexpressing NusA protein. Lanes 3-5 shows the protected region bound by VirS. (C) A consensus DNA sequence for binding of VirS was identified. The protected region identified by footprint analysis was used to define the VirS binding sites in its downstream targets. The consensus sequence logo was generated by the software WebLogo in which height of the stack indicates the degree of conservation at that position in the consensus and height of individual letter within a stack

represents the relative frequency of that nucleotide at that position. (D) The upstream regions of few of the downstream targets of VirS (*ahpD, ppe23, cmtR, ctpG, ppe38 and fadD9*) were selected and evaluated for their binding to 1 µg of purified VirS by employing EMSA. All the genes showed a positive binding with VirS suggesting their direct interaction with VirS. (E) As a negative control, an intergenic region between *MbtA* and *MbtB* containing IdeR binding region was employed, which did not display any shift in the mobility of the DNA in the presence of VirS. As a positive control, VirS and its binding sequence were used, which showed shift in the mobility of its DNA.

Figure 6: Prediction of 3-D model of VirS. (A) Multiple sequence alignment of amino acid sequences of VirS and its homologous proteins was generated by employing ClustalW. This alignment file was then used to generate 3-D structure of VirS by using the software Modeller 9.0. (B) The 3-D structure of VirS was generated by comparative homology modeling by using the software Modeller 9.0. The predicted structure of VirS contains two HTH motifs. (C) Structure of VirS with the predicted DNA binding sites. Figure 6B and 6C were generated by software Pymol (58).

Figure 7: Determination of crucial residues of VirS. (A) 15 amino acid residues of VirS selected for mutational studies are highlighted in the homology model of VirS. (B) The effect of mutations on DNA binding activity of VirS was determined by employing EMSA. The wild type 3a-VirS protein was used as a positive control and only DNA was used as negative control. Dashed line depicts that lanes from two different polyacrylamide gels are brought together. DNA binding activity was calculated by comparing the activity in wild type 3a-VirS loaded in the same gel. 0.5 µg of wild type 3a-VirS and mutants was subjected to the evaluation of the effect on DNA binding activity. 20 µg of cell lysates from wild type 3a-VirS and mutants Y61A, R73A and R80A were used to evaluate the effect of mutation on the DNA binding activity of VirS. (C) Bar graph displays percent activity exhibited by various mutants when compared to wild type VirS. Data represents the mean (±SD) of atleast2 independent experiments (*$p<0.05$, **$p<0.01$, ***$p<0.001$, two-way ANOVA).

Figure 8: Target based virtual screening against DNA binding region of VirS. (A) The figure represents the predicted 3D structure of VirS interacting with DNA. (B) The structure of VirS depicts the presence of two DNA contact sites. (C) Virtual screening was carried out against these two DNA contact sites by using Autodock4.2. Surface representation of VirS structure is shown here with grid covering the DNA contact site 1 of VirS. (D) Zoomed representation showing the grid covering the DNA contact site and nearby residues of VirS. (E) The figure represents the residues which are near to the DNA and are covered in the grid. (F) The grid generation against the second DNA contact site of VirS is shown here. (G) Zoomed representation of the grid around DNA contact site of VirS. (H) Residues surrounding the second DNA contact site, covered in the grid are shown here. The figures are prepared by using the software Autodock4.2.

Figure 9: Evaluation of the compounds for their inhibitory potential against the DNA binding activity of VirS. Bar graph displaying percent inhibition exhibited by compounds at 100 µg/ml, obtained on the basis of docking at site 1 (A) and obtained on the basis of docking at site 2 (B). DMSO was used as a negative control in EMSA which displayed no inhibition of DNA binding activity of VirS.

Figure 10: Evaluation of the inhibitory potential of compounds against intracellular survival of the pathogen. Compounds V7B and V15B were evaluated for their inhibitory potential against intracellular survival of *M. tuberculosis* inside macrophages. (A) Compound V7B displayed inhibition of growth of the pathogen at 80 µg/ml and above. Cytotoxicity of compound V7B was determined by resazurin dye. Lower panel depicts the cytotoxicity exhibited by compound V7B to the macrophages. Except at the highest concentration employed, compound V7B did not display any cytotoxicity to the macrophages. (B) Compound V15B displayed inhibition of growth of the pathogen at 40 µg/ml and above. Lower panel depicts the cytotoxicity exhibited by compound V15B to the macrophages. Compound V15B did not display any cytotoxicity to the macrophages.

Figure 11: Binding mode of the hit molecules. The predicted binding mode of the compounds V7B (A) and V15B (C) to the VirS are shown here. Figure 11B and 11D represent the ligand interaction diagrams of compounds V7B and V15B, respectively showing distance measurement with the neighboring residues of VirS. The ligand interaction diagrams are generated by using software Accelrys Discovery Studio (59).

Figure 12. Structure comparison of compound V7B and its analogs.

Figure 13. Comparison of structure of compound V15B and its analogs (ND-not determined).

Figure 14: The proposed model for VirS mediated downstream actions at acidic pH. Exposure to acidic pH causes up-regulation of VirS which in turn regulates a number of genes probably involved in combating acidic stress faced by the pathogen. Firstly, VirS regulates genes related to cellular metabolism which may help bacteria to adapt acidic stress. Secondly, it regulates the genes related to various cell wall components which might help in remodelling of the cell envelope to make it impervious to acidic stress. It also upregulates the expression of efflux pumps and ion transporters which may help in the maintenance of intrabacterial pH by exporting excess of protons outside the cell. VirS also regulates genes related to detoxification, information pathway and other transcriptional regulators which may help bacteria to resist and adapt to the acidic stress.

Figures

Figure 1

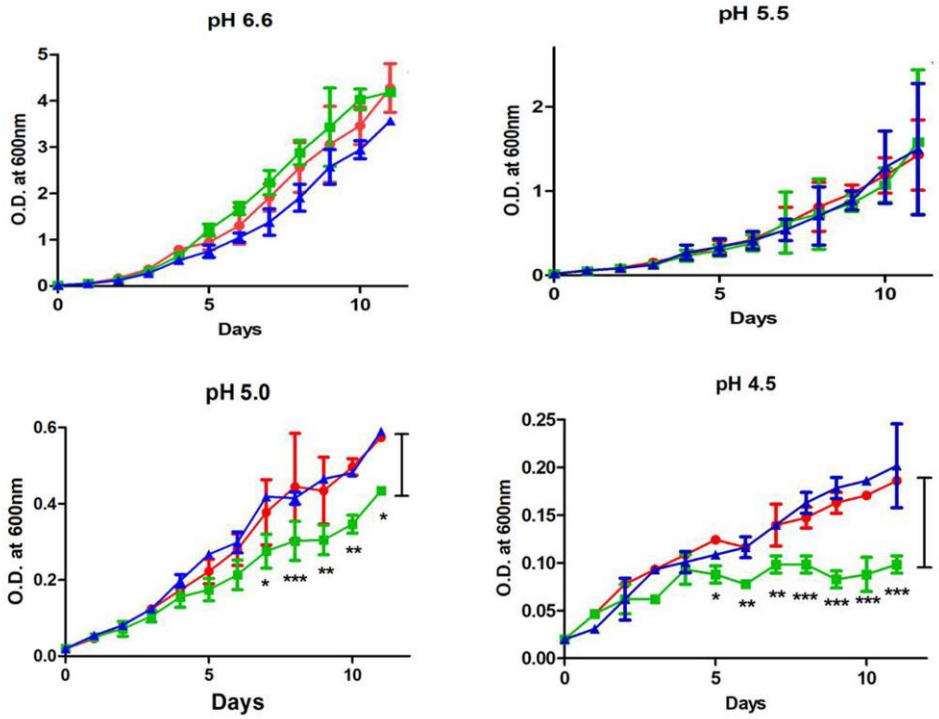


Figure 2

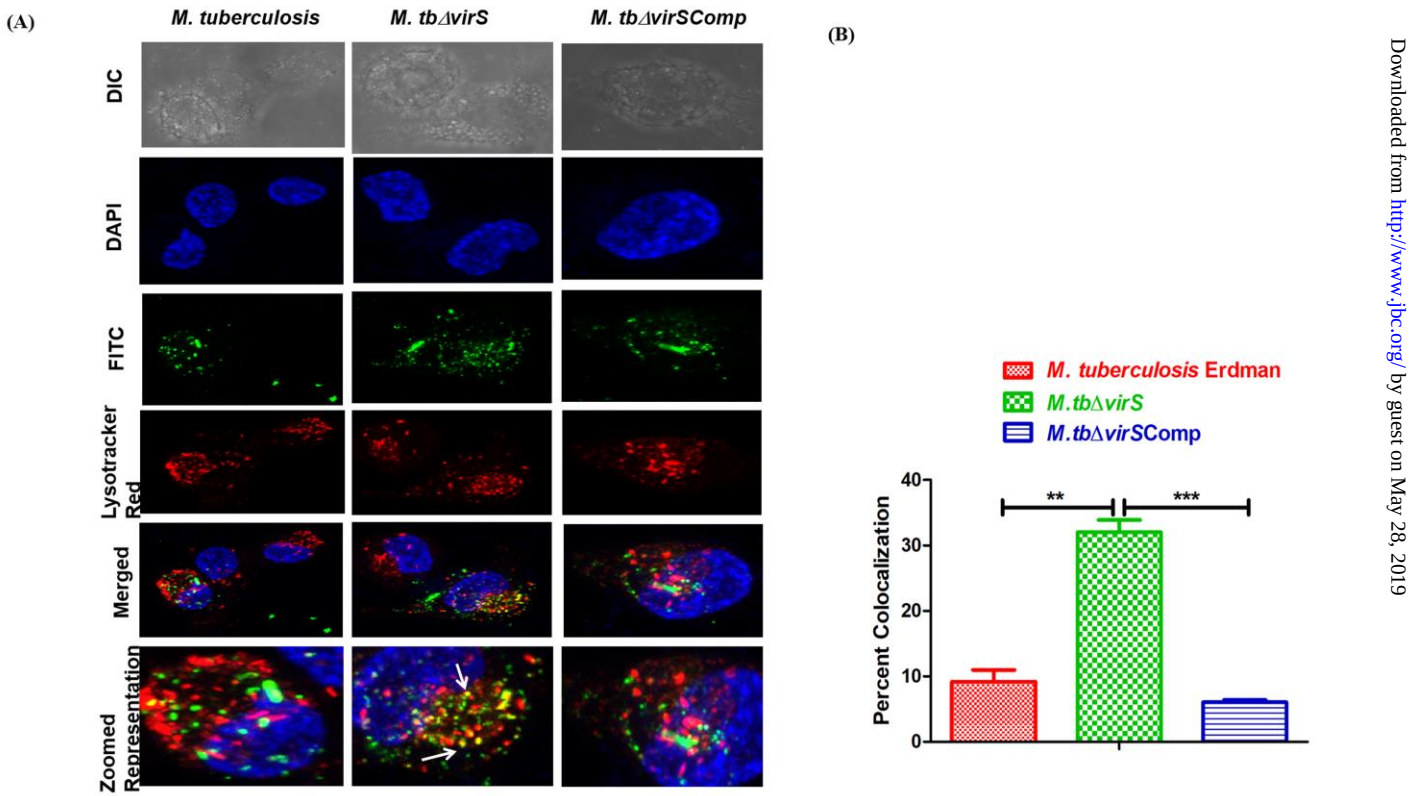


Figure 3

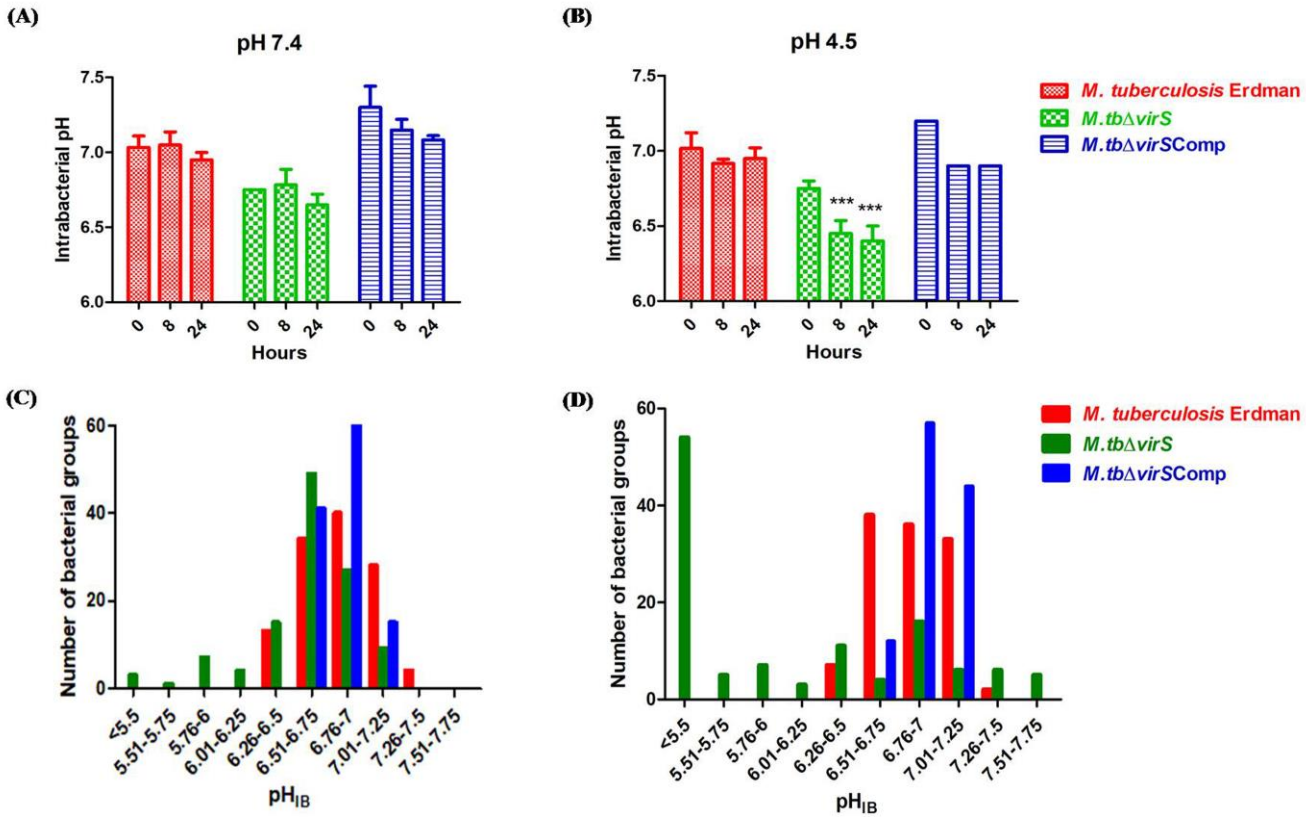


Figure 4

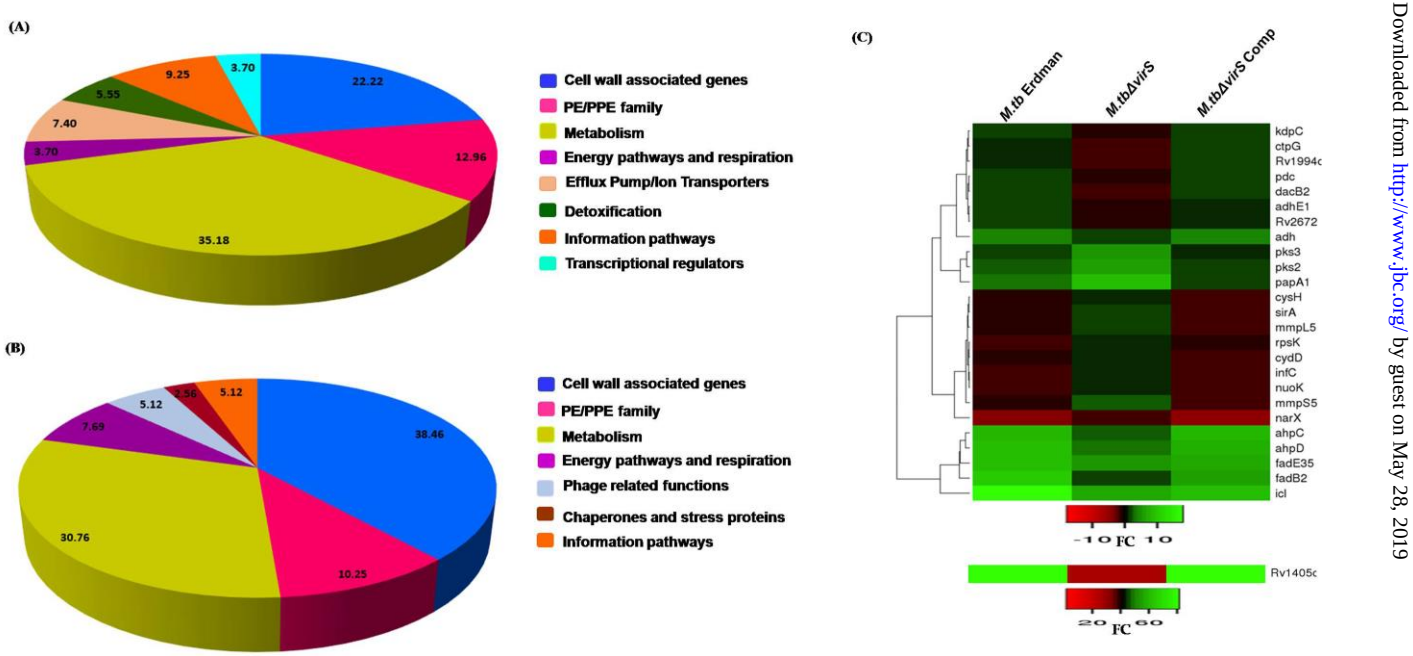


Figure 5

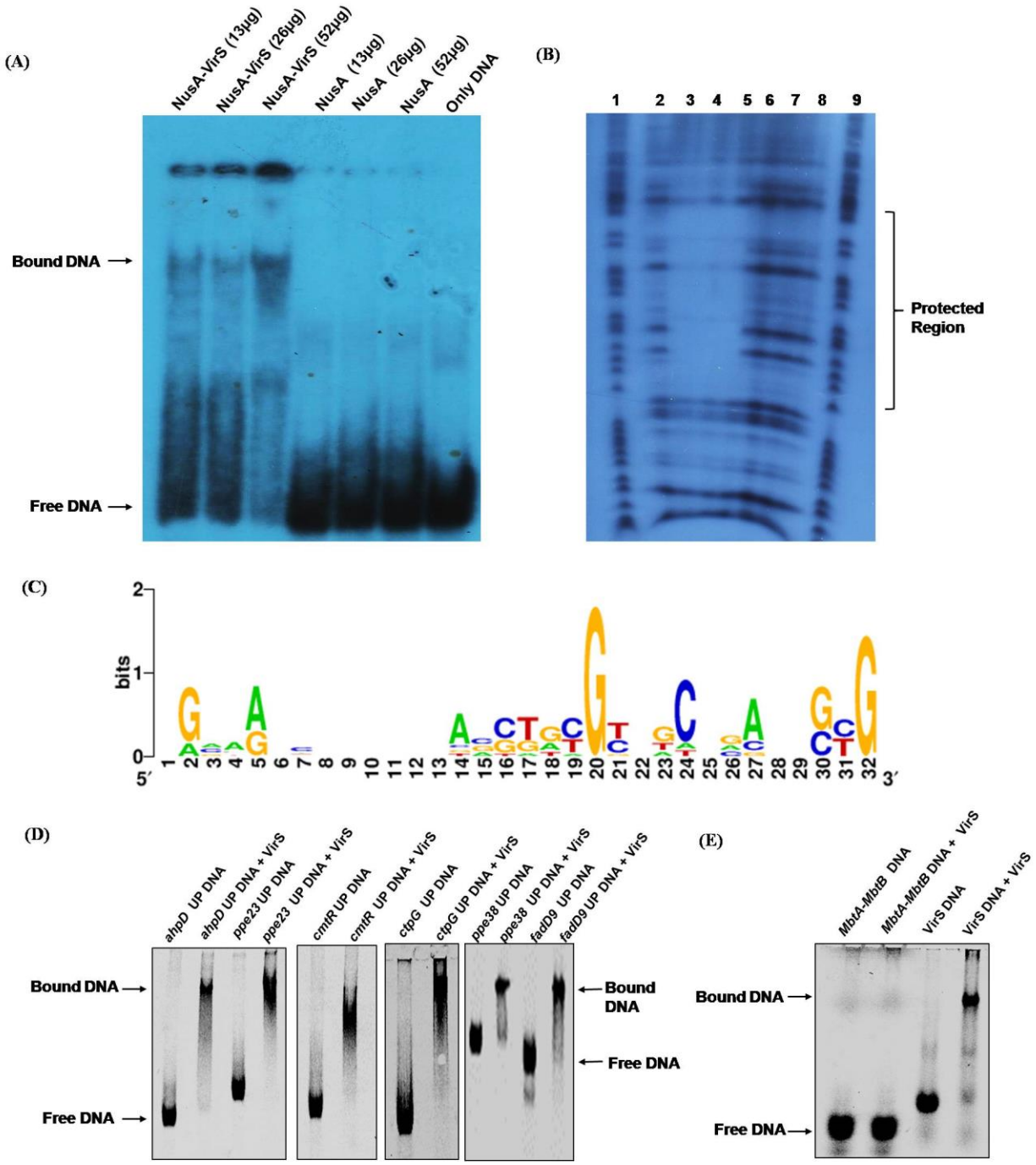


Figure 6

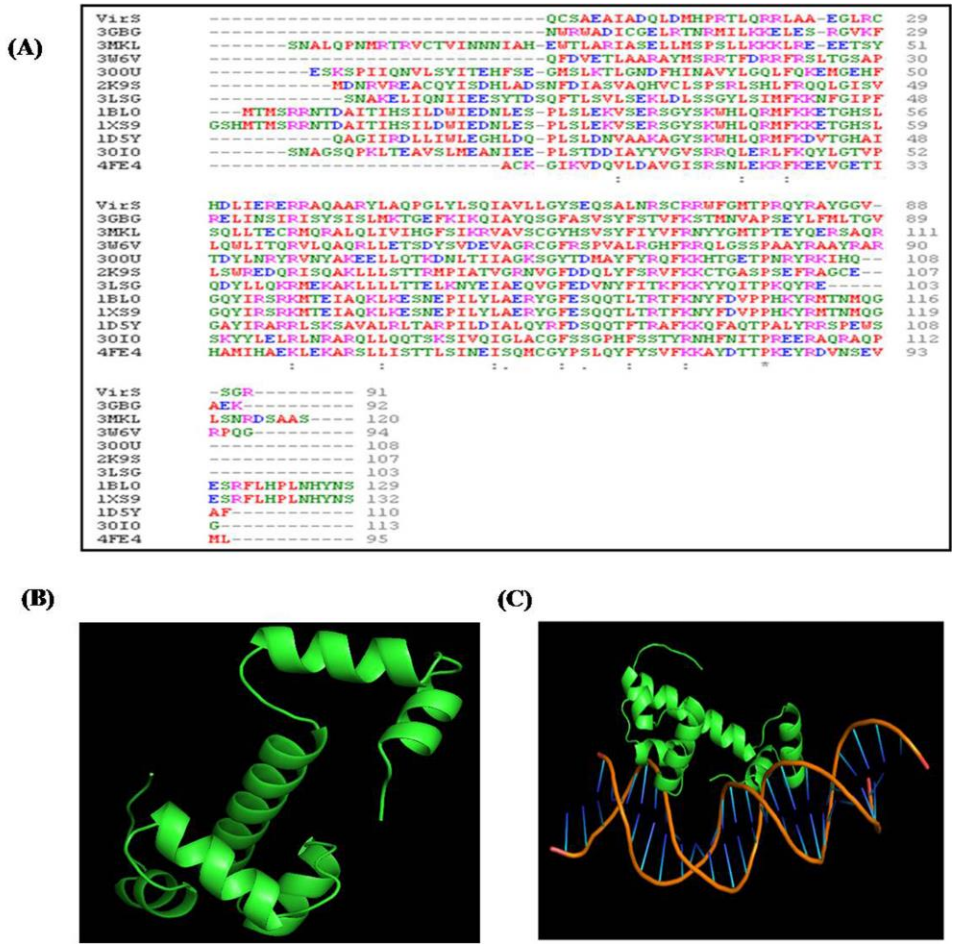


Figure 7

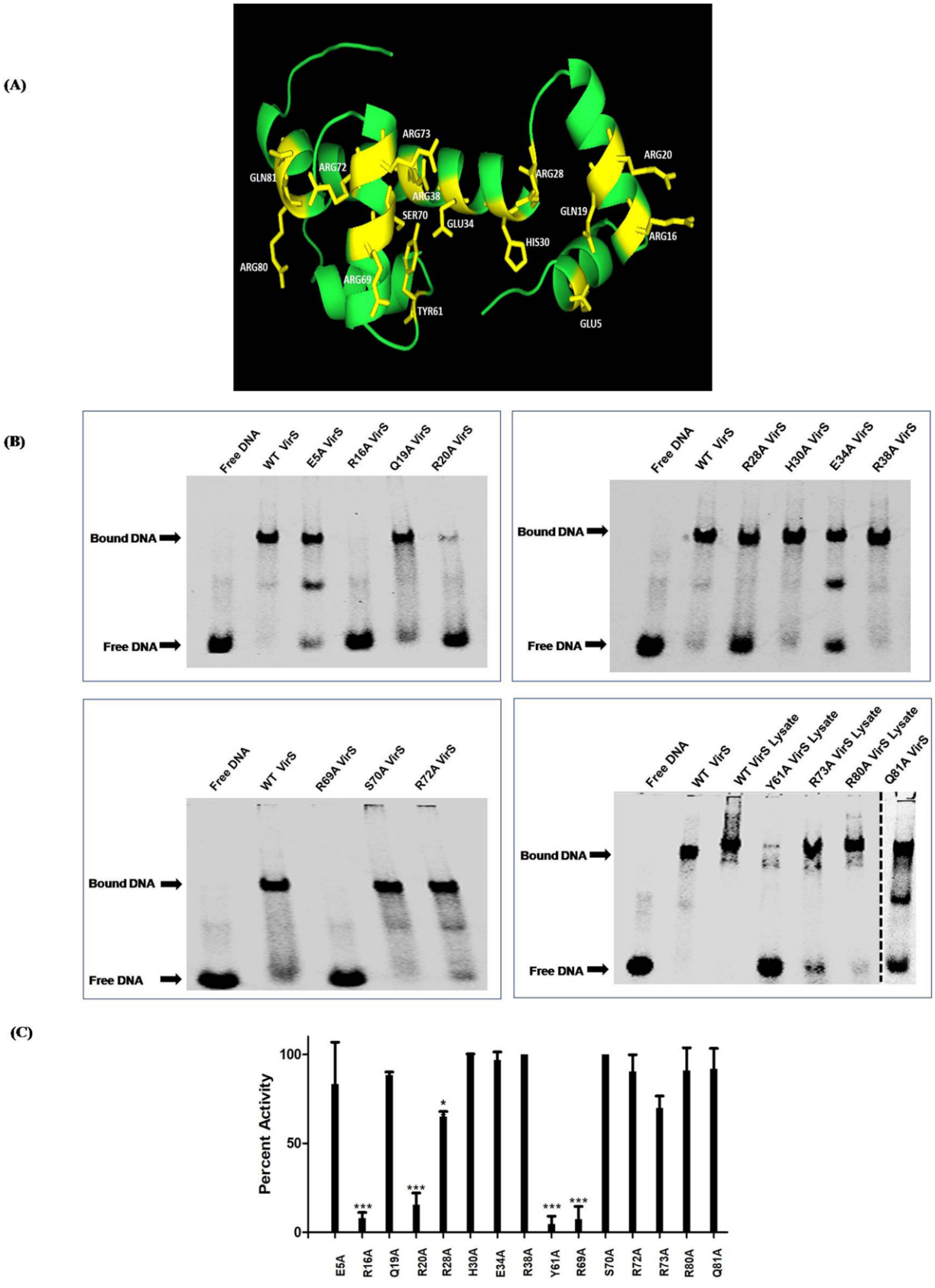


Figure 8

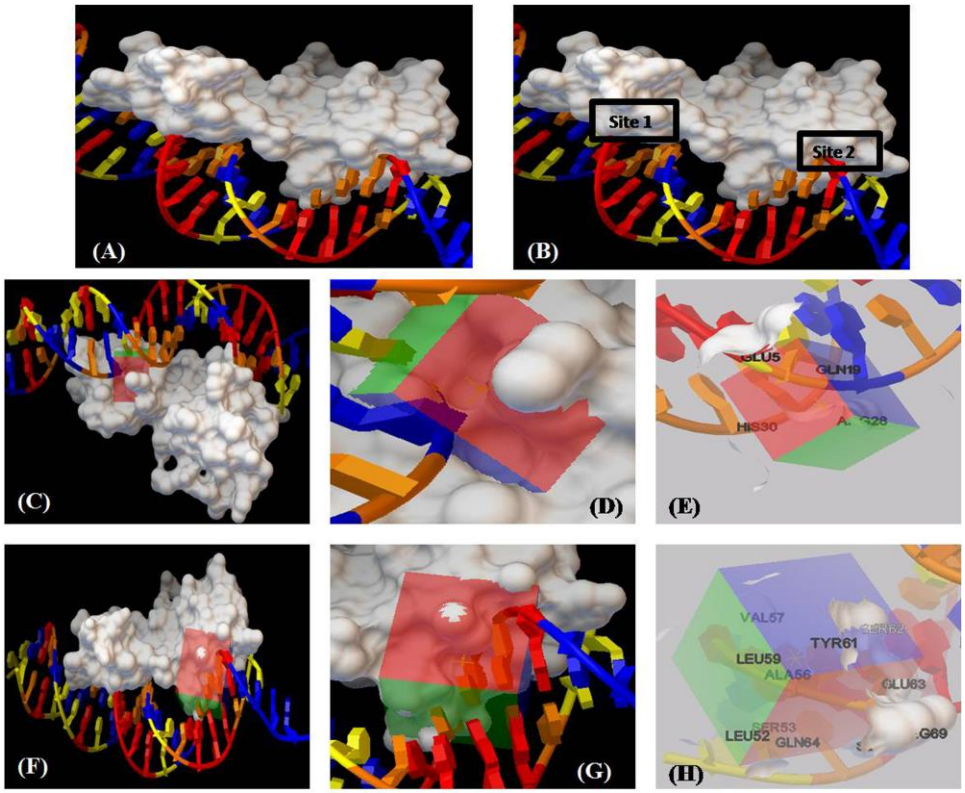


Figure 9

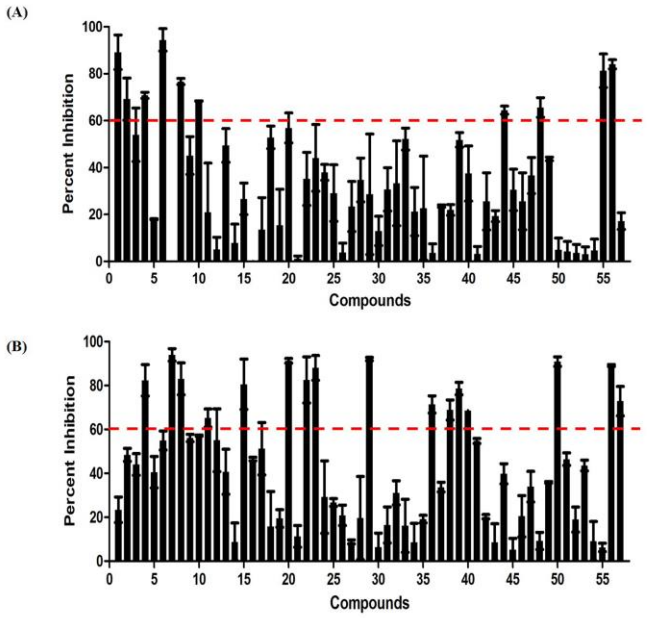


Figure 10

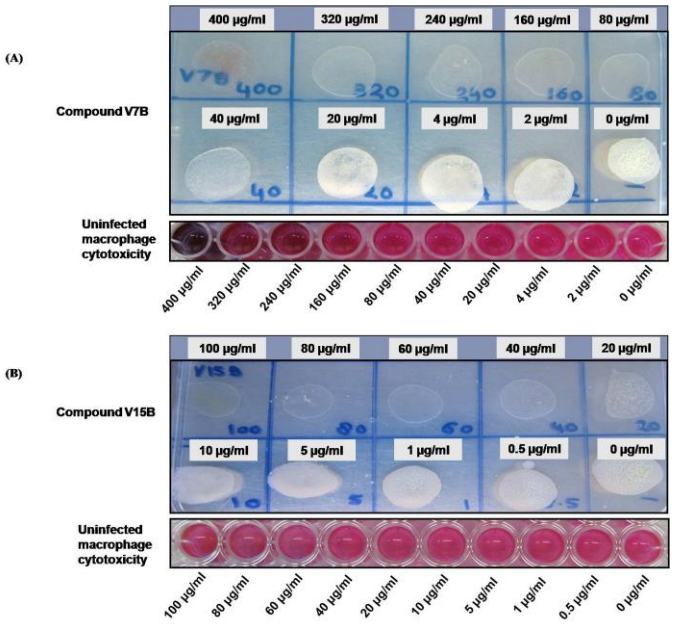


Figure 11

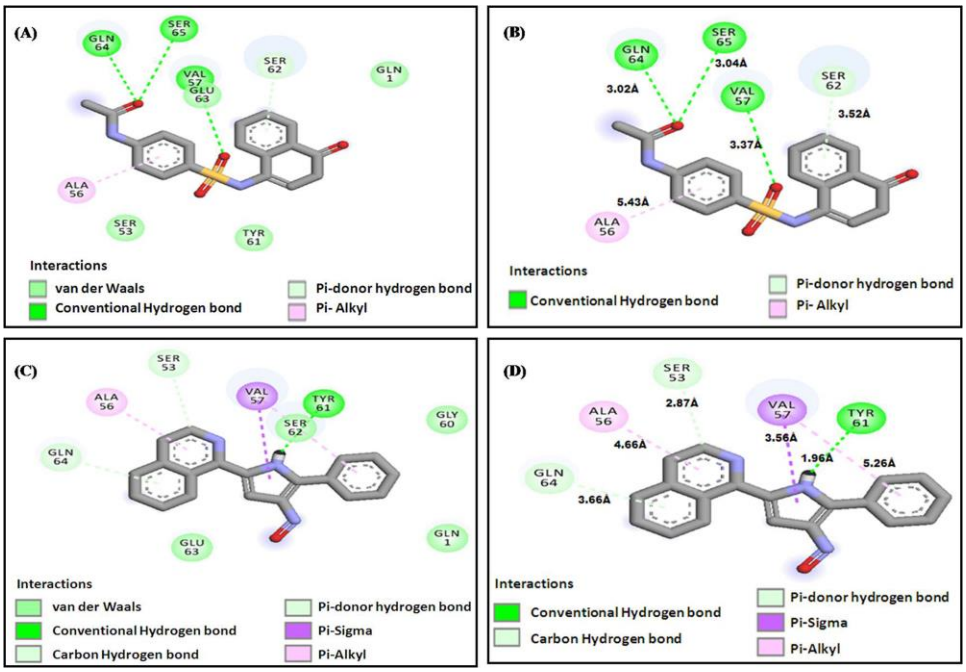


Figure 12

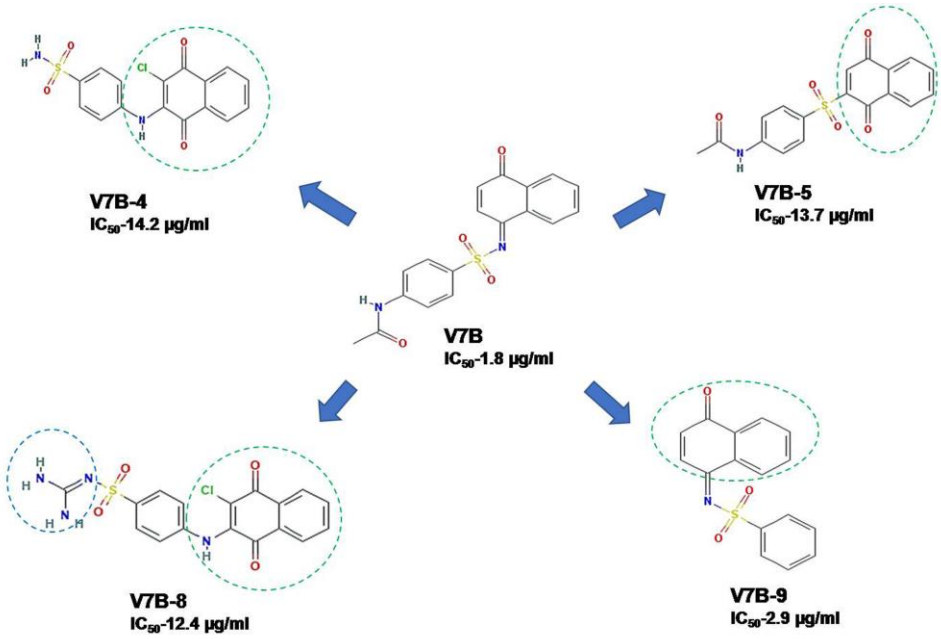


Figure 13

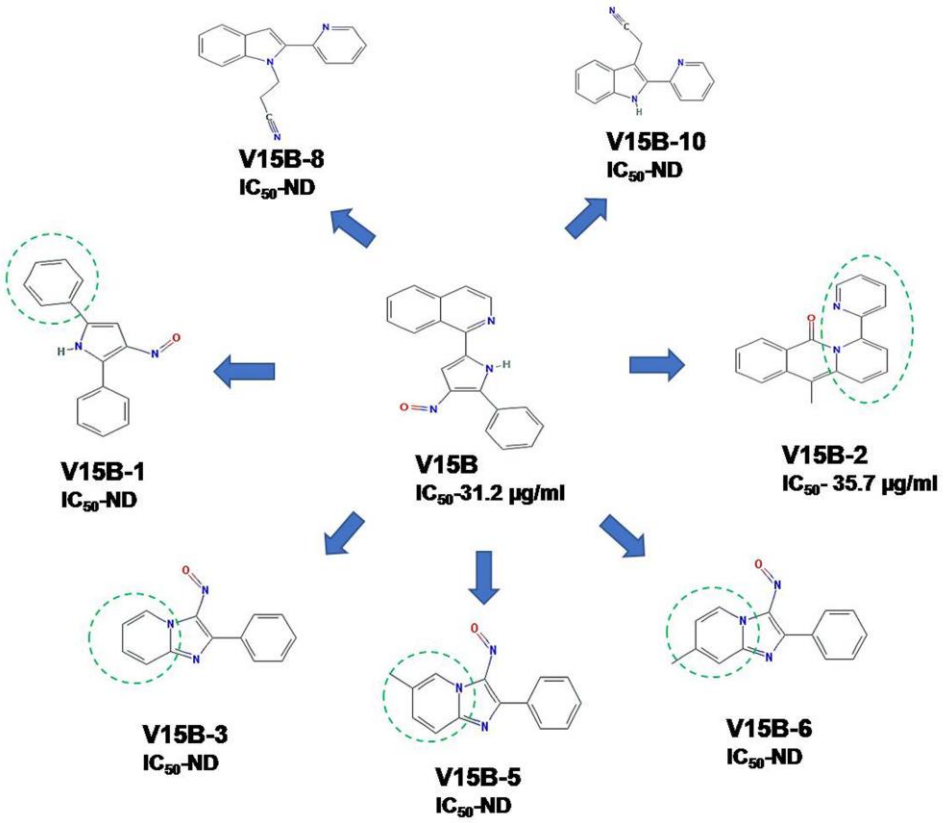
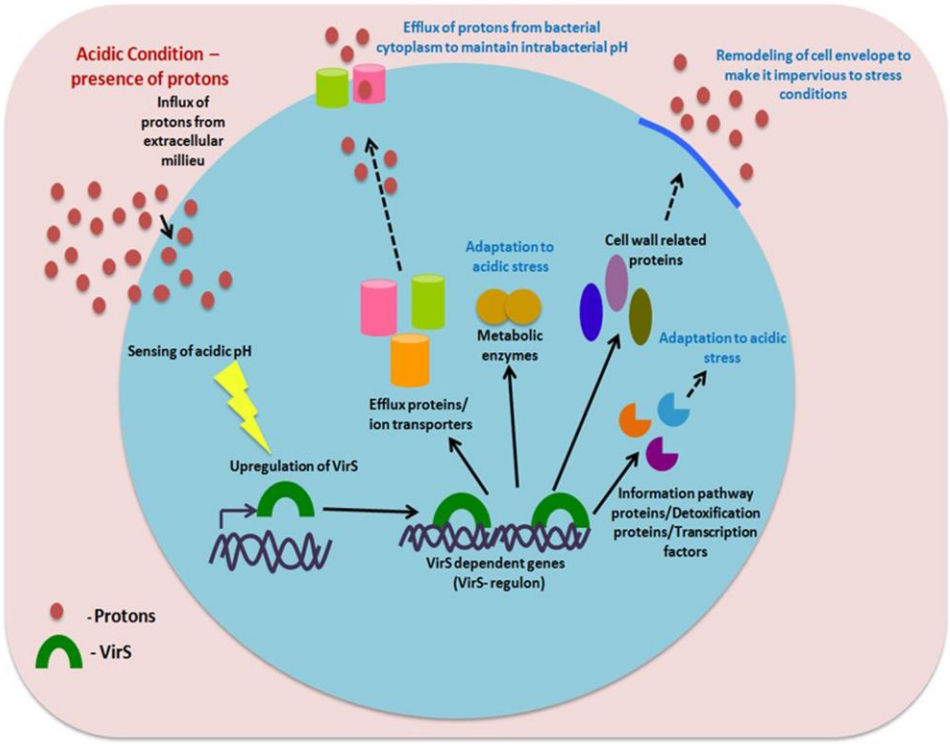


Figure 14



Unraveling the role of the transcriptional regulator VirS in low pH-induced responses of *Mycobacterium tuberculosis* and identification of VirS inhibitors

Swati Singh, Nikita Goswami, Anil K. Tyagi and Garima Khare

J. Biol. Chem. published online May 24, 2019
