BIOINFORMATICS DISCOVERY NOTE Vol. 21 no. 10 2005, pages 2161–2166
doi:10.1093/bioinformatics/bti375

Sequence analysis

Computational prediction and experimental verification of novel IdeR binding sites in the upstream sequences of *Mycobacterium tuberculosis* open reading frames

Prachee Prakash[1], Sailu Yellaboina[2], Akash Ranjan[2] and Seyed E. Hasnain[1,3,*]

[1]Laboratory of Molecular and Cellular Biology, [2]Laboratory of Computational and Functional Genomics, Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500 076, India and [3]Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560064, India

Received on November 18, 2004; revised on January 31, 2005; accepted on March 2, 2005
Advance Access publication March 3, 2005

ABSTRACT

IdeR (iron-dependent regulator) is a key regulator of virulence factors and iron acquisition systems in *Mycobacterium tuberculosis*. Despite the wealth of information available on IdeR-regulated genes of *M.tuberculosis*, there is still an underlying possibility that there are novel genes/pathways that have gone undetected, the identification of which could give new insights into understanding the pathogenesis of *M.tuberculosis*. We describe an *in silico* approach employing the positional relative entropy method to identify potential IdeR binding sites in the upstream sequences of all the 3919 ORFs of *M.tuberculosis*. While many of the predictions made by this approach overlapped with the ones already identified by microarray experiments and binding assays, pointing to the accuracy of our method, a few genes for which there has been no evidence for IdeR regulation were additionally identified. Our results have implications on the iron-dependent regulatory mechanism of *M.tuberculosis* vis-a-vis the activity of urease operon and novel transcription regulators and transporters.

Contact: hasnain@cdfd.org.in

INTRODUCTION

In pathogenic bacteria, many virulence factors and iron acquisition systems are regulated by iron-dependent transcription regulators (Litwin and Calderwood, 1993). One of the key regulators of such systems in *Mycobacterium tuberculosis* is IdeR (iron-dependent regulator), first identified as a homologue of the DtxR (diphtheria toxin repressor) protein of *Corynebacterium diphtheriae* (Schmitt *et al*., 1995). IdeR has been known to govern the expression of a wide variety of genes ranging from those involved in iron acquisition and oxidative stress response to the ones that code for enzymes involved in aromatic amino acid biosynthesis (Gold *et al*., 2001; Rodriguez and Smith, 2003). The success of *M.tuberculosis* in the establishment of an infection is also dependent upon its ability to acquire iron from its neighbouring environment. While low iron is a limiting factor for pathogen growth and survival, even high iron is detrimental as it leads to the formation of highly reactive hydroxyl radicals via the Fenton reaction. Hence, the acquisition of iron by pathogenic bacteria has to be tightly regulated.

*To whom correspondence should be addressed.

IdeR was first identified as the mycobacterial equivalent of DtxR of *C.diphtheriae* (Schmitt *et al*., 1995). DtxR serves as a repressor of the *tox* gene, the structural gene for diphtheria toxin. Apart from this function, DtxR also behaves as a regulatory protein involved in the iron metabolism of the bacterium (Boyd *et al*., 1990; Schmitt and Holmes, 1991). The function of DtxR was found to be similar to the very well-known Fur protein of gram-negative bacteria. Under iron sufficient conditions, DtxR causes the repression of genes involved in iron metabolism by binding to their operator sequences, with a high specificity, thereby blocking transcription (Tao *et al*., 1992). The IdeR of *M.tuberculosis* shares a 59% overall amino acid identity within a 230 amino acid stretch to DtxR. Initial experiments were carried out to determine if IdeR represses transcription of DtxR-regulated promoters. DNA mobility shift assays and a DNA footprinting analysis showed that IdeR binds to the same promoter sequences to which the Corynebacterial DtxR protein binds. Binding was observed to be metal-dependent (Schmitt *et al*., 1995).

In *M.tuberculosis*, under iron sufficient conditions, IdeR binds to the upstream sequences of genes required for growth in low iron conditions thereby preventing their transcription. Under iron limiting conditions, IdeR no longer binds to these promoter regions which are free to allow the binding of RNA polymerase and subsequent transcription of the downstream gene. *ide*R is also an essential gene of *M.tuberculosis* and the encoded protein regulates the expression of genes involved in the metabolism of iron and oxidative stress response (Rodriguez *et al*., 2002). Though the *ide*R null mutant of *M.tuberculosis* cannot be generated without the incorporation of a second copy of the gene, Rodriguez and coworkers obtained a rare mutant of *ide*R in which the lethal effects of *ide*R inactivation were alleviated by a suppressor mutation. This mutant showed a restricted iron assimilation capacity. The authors also studied the transcription profiles of wild-type, *ide*R mutant and *ide*R-complemented mutant *M.tuberculosis* strains using DNA microarray. This resulted in the identification of genes regulated by iron and IdeR. These genes encode proteins involved in siderophore biosynthesis and iron storage, enzymes of aromatic amino acid biosynthesis, putative transporters, members of the PE/PPE family, transcriptional regulators and enzymes involved in lipid metabolism.

The IdeR of *M.tuberculosis* in association with ferrous ions binds to a 19 bp inverted repeat consensus sequence or iron

box [TTAGGTTAGGCTAACCTAA] in the upstream sequences of several genes (Schmitt *et al*., 1995). Gel mobility shift assays, DNA footprinting, Reporter gene assays and DNA microarray are four techniques that have been exploited by a multitude of workers to determine the genes expressed/repressed in *M.tuberculosis* as a function of iron availability (Gold *et al*., 2001; Camacho *et al*., 1999). Genes involved in iron acquisition and storage have been shown to be IdeR-regulated (Dussurget *et al*., 1996; Gold *et al*., 2001). Several other genes not directly involved in siderophore biosynthesis have also been shown to be expressed or repressed as a function of iron stress (Rodriguez and Smith, 2003; Dussurget *et al*., 1999). These reports suggest that IdeR is a global regulator that controls several genes involved in iron metabolism and processes related to iron metabolism. Experimental evidence for iron-mediated regulation of quite a few mycobacterial genes exists. Two divergently transcribed genes, *his*E (a part of the histidine operon) and a PPE gene (Rv2123) have been shown to be IdeR-regulated (Rodriguez *et al*., 1999). Gel shifts and footprinting assays have revealed that IdeR regulates the expression of these genes by binding to the iron boxes in the regulatory region and binding was dependent upon the concentration of iron in the reaction mix. A few other genes involved in the biosynthesis of siderophores (*mbt*A, *mbt*B, *mbt*I), biosynthesis of aromatic amino acids (*trp*E2, *phe*A, *his*E, *his*B) and others like iron storage proteins (bacterioferritins, *bfr*A, *bfr*B) have also been experimentally shown to be part of the IdeR regulon (Gold *et al*., 2001). Functional characterization of genes not apparently involved in iron metabolism would lead to further insights into the relation between iron metabolism and various aspects of mycobacterial physiology. This report describes the use of a computational approach to identify novel genes under the regulatory control of IdeR followed by its experimental verification. Our results, while confirming already known IdeR-regulated genes, have additionally identified new genes.

METHODS

Computational prediction of IdeR binding sites

The complete genome sequence of *M.tuberculosis* H37Rv was downloaded from the NCBI ftp site (ftp://ftp.ncbi.nlm.nih.gov) and IdeR binding sites were collected from the literature (Table 1). Profiles for recognition of IdeR binding sites were calculated by the positional relative entropy method assuming that each position is an independent site (Yellaboina *et al*., 2004a,b). A matrix was developed for the purpose, which was used to scan the upstream sequences of all the genes from −400 to +20 of the translation start site to identify potential IdeR binding. Consensus IdeR binding sites were used to compute probability distributions of four different nucleotides within the binding sites of known sequences as well as throughout the genome. The probability distributions of nucleotides within and outside the regulatory region were used to compute the relative entropy of segments (length 19 bp) along the +20 to −400 regions of all the genes of *M.tuberculosis*. Finally segments were sorted according to the relative entropy, and segments with a high relative entropy were considered as probable iron-dependent repressor binding sites.

Cloning, expression and purification of *M.tuberculosis* IdeR

The pRSETa expression vector (Promega) with an N-terminal 6X His tag was used to clone the ORF Rv2711 of *M.tuberculosis* that encodes IdeR. Briefly, Rv2711 was amplified from *M.tuberculosis* H37Rv DNA using primers with specific restriction enzyme sites (forward primer: ATT***GGATCC***ATGAACGAGTTGGTTGATA; reverse primer: TGT***AAGCTT***GACTTTCTCGACCTTGACC) and the amplicon was cloned into the corresponding sites of the pRSETa expression vector.

Table 1. Known IdeR binding sites in the upstream sequences of *M.tuberculosis* ORFs

IdeR binding sequence	Downstream ORF
CAAGGTAAGGCTAGCCTTA	Rv1519
TTATGTTAGCCTTCCCTTA	Rv3403c
TTAACTTAGGCTTACCTAA	Rv3839
TTAGGCAAGGCTAGCCTTG	Rv1343c
CAAGGCTAGCCTTGCCTAA	Rv1344
TATGGCATGCCTAACCTAA	Rv1347c
TTCGGTAAGGCAACCCTTA	Rv1348
ATAGGTTAGGCTACCCTAG	Rv2122c
CTAGGGTACCCTAACCTAT	Rv2123
AGAGGTAAGGCTAACCTCA	Rv3402c
TTAGTGGAGTCTAACCTAA	Rv1876
GTAGGTTAGGCTACATTTA	Rv2386c
CTAGGAAAGCCTTTCCTGA	Rv3841
TTAGCTTATGCAATGCTAA	Rv0282
TTAGGCTAGGCTTAGTTGC	Rv0451c
TTAGCACAGGCTGCCCTAA	Rv2383c
TTAGGGCAGCCTGTGCTAA	Rv2384

Escherichia coli BL21DE3 cells transformed with the 6X His tagged chimeric construct were grown in 1 l of Luria-Bertani medium supplemented with 100 μg/ml of ampicillin and 10% glycerol. IPTG (0.1 mM) was added to a mid-log phase culture. The cells were kept in an incubator shaker for another 5 h at 27°C and 150 rpm to allow protein expression. After induction, cells were harvested by centrifugation and resuspended in 20 ml of lysis buffer (10 mM Tris HCl, 100 mM NaCl and 10% glycerol, pH 7.5) with 0.1 mM PMSF and disrupted using a sonicator. After a second round of centrifugation for 10 min at 10 000 × *g*, the supernatant was applied to a Ni-NTA affinity column (Qiagen, USA).

Affinity chromatography The supernatant was allowed to bind to a Ni-NTA column (Qiagen) packed in a polypropylene column. The recombinant protein was purified after washing the column with five bed volumes of lysis buffer containing 30 mM imidazole and eluting with 250 mM imidazole. The eluates were analyzed by SDS PAGE and dialyzed against Tris buffer to remove salts and imidazole. The purity of the eluted protein was checked on SDS PAGE followed by Coomassie Blue staining.

Gel retardation assay Binding of IdeR to the predicted iron box was carried out in a 20 μl reaction consisting of 1X buffer (10 mM Tris HCl, 50 mM NaCl, 10% glycerol, 5 μg/ml acetylated BSA, 1 mM DTT, 1 mM PMSF and 50 mM $MgCl_2$), 1% NP40, 1 μg/ml poly dIdC, purified IdeR (1 μM) and 3–5 fmol of ^{32}P labelled probe. The probe consisted of the annealed 19 bp oligo corresponding to the predicted IdeR box end labelled with ^{32}P using T4 polynucleotide kinase enzyme. The reaction was performed in the presence and absence of $CoCl_2$ (200 μM). Unlabelled oligo was used for specific competition. After the addition of the labelled probe, the reaction mixture was incubated for 15 min at 25°C followed by loading on a 4% polyacrylamide gel in 1XTBE buffer. Electrophoresis was carried out at 200 V for 30 min at 4°C. After electrophoresis, the gel was dried and analyzed by autoradiography.

South-western assay The bacterial extract over-expressing *M.tuberculosis* IdeR was separated on an SDS/PAGE gel and the proteins were electrophoretically transferred to a nitrocellulose membrane in a buffer containing 25 mM Tris, 190 mM glycine and 20% methanol for 16 h at 30 mA. The protein on the membrane was renatured by incubating in blocking buffer (2% non-fat dry milk, 1% BSA, 10 mM Hepes NaOH, 0.1 mM EDTA, 200 mM NaCl, 50 mM $MgCl_2$ and 16 μg/ml sonicated sperm DNA). After renaturation, the

Table 2. Candidate IdeR binding sites in the genome of *M.tuberculosis*[a]

IdeR binding site	Position of binding site relative to translation start site	Score	Gene annotation	Rv number	Predicted function
TTAGTGGAGTCTAACCTAA	−226	5.2563	*bfr*A	Rv1876	Bacterioferritin
ATAGGCAAGGCTGCCCTAA	**−151**	**5.19346**	—	**Rv1846c**	**Predicted transcriptional regulator**
TTAGCACAGGCTGCCCTAA	−86	5.16997	*mbt*A	Rv2384	Peptide arylation enzymes
TTAGGGCAGCCTGTGCTAA*	−32	5.15772	*mbt*B	Rv2383c	Peptide arylation enzymes
TTATGTTAGCCTTCCCTTA*	−2	5.14546	—	Rv3403c	Hypothetical protein
CTAGGAAAGCCTTTCCTGA*	−73	5.12055	*bfr*B	Rv3841	Ferritin-like protein
TTAGGCAAGGCTAGCCTTG*	−85	5.09743	—	Rv1343c	Hypothetical protein
TTAACTTAGGCTTACCTAA	−36	5.09181	—	Rv3839	Hypothetical protein
ATAGGTTAGGCTACCCTAG*	−51	5.07767	PPE	Rv2123	PPE
TTAGGTAAGCCTAAGTTAA	−79	5.04482	*phe*A	Rv3838c	Prephenate dehydratase
CTAGGGTAGCCTAACCTAT*	−95	5.04385	*his*I	Rv2122c	Phosphoribosyl-ATP pyrophosphohydrolase
TTAGGGCAGCCTTGCCTAT	**−146**	**5.0165**	—	**Rv1847**	**Hypothetical protein**
CAAGGCTAGCCTTGCCTAA	−292	4.97724	*fad*D33	Rv1345	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II
CAAGGTAAGGCTAGCCTTA*	−50	4.97077	—	Rv1519	Hypothetical protein
CAAGGTAAGGCTAGCCTTA	−345	4.97077	—	Rv1520*	Glycosyltransferases involved in cell wall biogenesis
GTAGGTTAGGCTACATTTA*	−25	4.8669	*trp*E2	Rv2386c	Anthranilate/para-aminobenzoate synthases component I
GCAGGTCAGGCTACCCTTA	−26	4.82224	*mur*B	Rv0482	UDP-N-acetylmuramate dehydrogenase
ATAGGAAAGCCGATCCTTA	−36	4.64865	—	Rv0114	Histidinol phosphatase and related phosphatases
GTAGACCAGGCTCCCCTTG	**−302**	**4.62592**	***fec*B**	**Rv3044**	**ABC-type Fe3+-siderophores transport systems**
TAAGGGTAGCCTGACCTGC	−20	4.61752	—	Rv0481c	Hypothetical protein
TTAGGCTAGGCTTAGTTGC*	−112	4.59032	*mmp*S4	Rv0451c	mmpS4
GCAACTAAGCCTAGCCTAA	−139	4.54925	—	Rv0452	Transcriptional regulator
CTATGTGATACTGACCTGA	−42	4.5466	*glp*Q2	Rv0317c	Glycerophosphoryl diester phosphodiesterase
AGATGCTAGACTTTCCTGA	**−77**	**4.54327**	—	**Rv1404**	**Transcriptional regulator**
TTACGGCAGCCTGTTGTAA	**−35**	**4.53876**	—	**Rv2663**	**Hypothetical protein**
TTAGCTTATGCAATGCTAA*	−50	4.49914	—	Rv0282	Hypothetical protein
TTCGGTAAGGCAACCCTTA*	−213	4.41965	—	Rv1348	Hypothetical protein
TCACTGTAGTCTTAGCTGA	−179	4.39591	—	Rv0698	Hypothetical protein
ATCCGTAAGTCTAAACTTA	−26	4.35929	—	Rv2034	Predicted transcriptional regulators
TTACTGCAATCTCCACTGA	−149	4.33623	*fad*A5	Rv3546	Acetyl-CoA acetyltransferases
TATGGCATGCCTAACCTAA	−31	4.02212	—	Rv1347c	Acetyltransferase
TTACCGCGCACTGCTCTAT	−17	3.51297	—	Rv1344	Acyl carrier protein
TATGGCATGCCTAACCTAA	−50	4.02212	—	Rv1347c	Acetyltransferase
GTAGGTTAGGACAGCCTTT	−102	3.92933	—	Rv0338c	Fe-S oxidoreductases
TAATGGCAGACTGTGATTC	−3	3.89219	*ppi*A	Rv0009	Peptidyl-prolyl cis trans isomerase

[a]Sequences with * represent the experimentally confirmed IdeR binding sites. Sequences in bold represent the experimentally unverified novel IdeR binding sites predicted by the positional relative entropy method used in this study.

membrane was placed in a hybridization bag with binding buffer (blocking buffer with 0.2% non-fat dry milk and 10^6 cpm/ml labelled oligo). Hybridization was performed with constant shaking for 16 h. The membrane was briefly rinsed in blocking buffer without skimmed milk or BSA, dried, covered with plastic wrap and subjected to autoradiography.

RESULTS

Novel IdeR binding sites are present upstream of *fec*B, a periplasmic lipoprotein coding gene and Rv1404, a putative transcriptional regulator

The consensus IdeR binding site collected from the published literature (Table 1) was used to identify similar IdeR binding regions in the −400 to +20 regions of all the 3919 ORFs of *M.tuberculosis*. A complete list of IdeR binding sites with the highest scores as calculated by the positional relative entropy method is shown in Table 2.

To date, the most detailed study on the prediction of IdeR binding sites along with experimental verification has been carried out by Gold *et al*. (2001). Additionally, microarray analysis of genes induced by low iron and in an IdeR mutant strain have also shed light on the iron-dependent regulation of mycobacterial genes (Rodriguez *et al*., 2002). While our method indeed identified novel IdeR binding sites, the possibility of occurrence of additional such sites cannot be ruled out. As a first step towards the analysis of our predictions, the results were compared with the available information on IdeR-regulated genes. Though most of these genes were earlier known to be IdeR-regulated, the present study identified for the first time that a part of the ferric dicitrate type transporter complex, FecB, a periplasmic lipoprotein and Rv1404, a putative transcriptional regulator, are possibly regulated by IdeR (Table 2). The upstream sequence of *fec*B shows the presence of an IdeR box at −302 position. On account of the absence of reports on the details of the iron transport system of *M.tuberculosis*, the ferric dicitrate transporter system does seem to be an important candidate. A new transcription regulator (Rv1404) and a hypothetical protein Rv2663 that were not earlier predicted to be part of the IdeR regulon could also be identified in this study.

IdeR binds to the IdeR box present in the intergenic region between the ORFs Rv1846c and Rv1847c

While many of the IdeR binding sites predicted by this study overlapped with ones predicted by earlier workers, experimental evidence demonstrated by *in vitro* binding experiments and reporter gene assays is available for only a few. These include *his*E, Rv2123 (Rodriguez *et al*., 1999), Rv3402, *mbt*I, *his*G, *mbt*A, *mbt*B, *mbt*I, Rv3402 and *bfr*A (Gold *et al*., 2001), etc. As per our prediction, the IdeR box upstream of the ORF Rv1846c shows one of the highest similarity scores to the IdeR consensus sequence. However, experimental evidence for the same does not exist. Moreover, Rv1846c does not figure in the list of genes induced in an IdeR mutant strain (Rodriguez and Smith, 2003). The binding site between Rv1846 and Rv1847 was also observed to be conserved in other mycobacteria. Hence, it was decided to determine if IdeR binds to this predicted iron box. The ORF Rv2711 that encodes IdeR was cloned in the *Bam*HI and *Hin*dIII sites of the pRSETa vector with an N-terminal Histidine tag and expressed and purified as a recombinant protein in *E.coli* BL21 cells (Fig. 1). Purified recombinant IdeR was used in gel retardation and south-western assays to test if it binds to the predicted IdeR box in the intergenic region between Rv1846c and Rv1847 (Fig. 2A). As evident from the gel shift assay (Fig. 2B), IdeR does

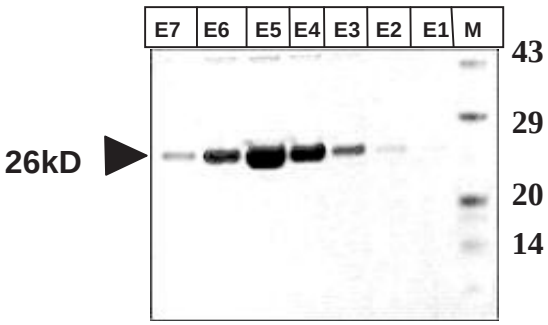

Fig. 1. Purification of the IdeR of *M.tuberculosis* as a recombinant protein in *E.coli*. *M.tuberculosis* IdeR (coded by ORF Rv2711) was cloned in the *Bam*HI/*Hin*dIII sites of the pRSETa expression vector and purified as a 6X Histidine-tagged recombinant protein using affinity chromatography procedures. Purified protein was fractionated on a 10% polyacrylamide gel and stained with Coomassie Brilliant Blue dye. M represents the protein molecular size marker (Broad range, Genei, India), E1–E7 show the successive eluted fractions of the recombinant protein. The arrowhead indicates the position of the pure eluted protein.

bind to the 19 bp IdeR binding site present in the intergenic region between Rv1846c and Rv1847. The binding could be completed out using cold oligos indicating the specificity of binding.

To convincingly demonstrate the binding of IdeR to the above mentioned probable iron box, south-western analysis was carried out. *E.coli* BL21 strain transformed with recombinant plasmid carrying *M.tuberculosis ide*R was grown to mid-log phase and fractionated by electrophoresis on a polyacrylamide gel. The gel was probed with radiolabelled oligonucleotide corresponding to the predicted iron box. While the vector control lysate lane (Fig. 2C, lane 1) did not show any binding, the induced cultures showed a positive binding. These data conclusively demonstrate that IdeR indeed binds to the predicted iron box element present in the divergently transcribed ORFs, Rv1846c and Rv1847, of *M.tuberculosis*.

DISCUSSION

Non-availability of the soluble form of iron is an important form of nutritional stress presented by the host to the bacterium; it is therefore logical to assume that genes responsible for the acquisition of iron are essential for full virulence and establishment of a successful infection. *M.tuberculosis* has an elaborate network of genes for the biosynthesis of siderophores, the iron acquisition systems (Quadri *et al*., 1998). Recent experiments have shown that these genes are regulated by iron-dependent regulatory proteins (Gold *et al*., 2001). Transcriptional control plays a key role in regulating gene expression in response to various environmental conditions. Apart from the production of siderophores as a function of low iron availability, *M.tuberculosis* also produces many other iron-regulated proteins, which are the probable virulence factors of the bacterium (Rodriguez and Smith, 2003).

The ferric dicitrate type transporter complex of *M.tuberculosis* as a probable IdeR-regulated system

While a number of transporter proteins like Rv1463 (an ABC transporter), Rv2459 (a probable drug efflux pump), Rv1348 (a membrane protein similar to Yersiniabactin uptake system), etc., have been

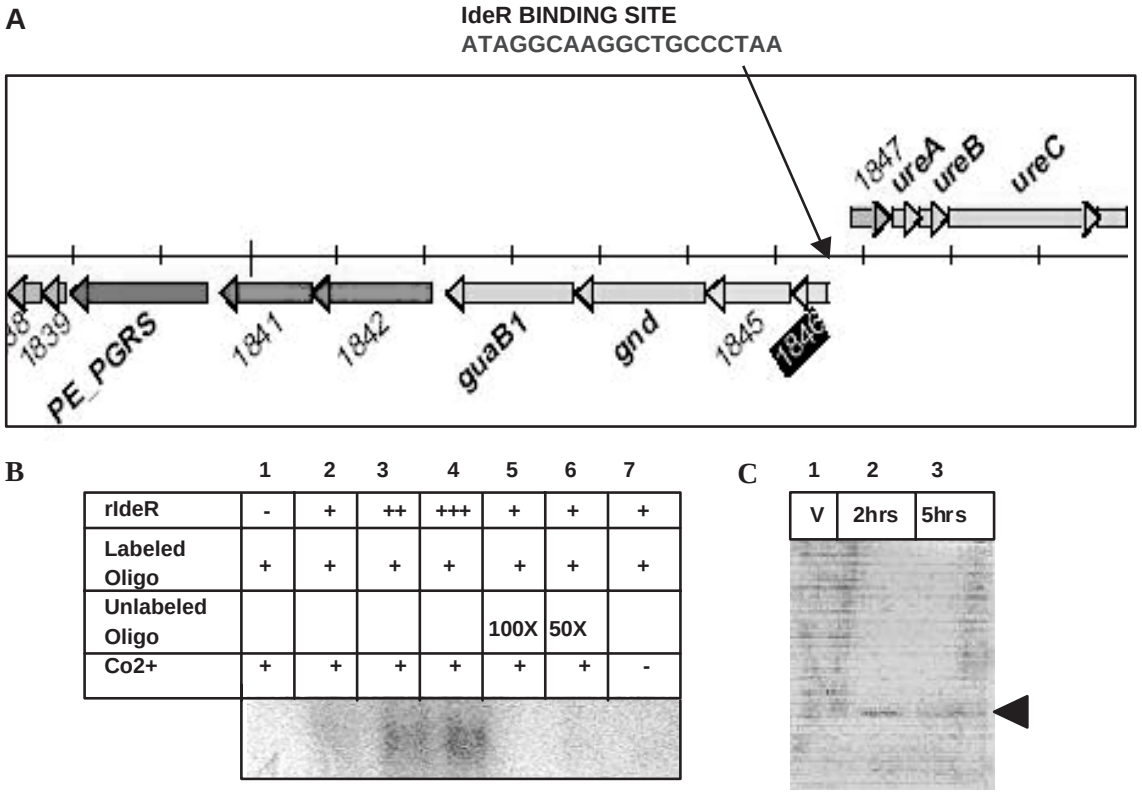

Fig. 2. Recombinant IdeR binds to predicted iron box element. **(A)** Schematic representation of the divergently transcribed ORFs, Rv1846c and Rv1847, with an IdeR binding site in the intergenic region. *ure*A, B and C are genes of the urease operon. **(B)** Autoradiogram of the gel retardation assay demonstrating the binding of IdeR to the predicted iron box shown in (A). Binding was specific as indicated by the disappearance of the band upon addition of cold oligo (lanes 5 and 6). Absence of a band in lane 7 confirms a metal-dependent binding of IdeR to the predicted iron box. **(C)** Autoradiogram of the south western assay demonstrating the binding of *M.tuberculosis* IdeR in *E.coli* BL21 cell lysates (induced for 2 and 5 h) to the predicted iron box shown in (A). The cultures were induced for 2 and 5 h, fractionated on a 10% polyacrylamide gel, transferred to a nitrocellulose membrane, renatured and hybridized with ^{32}P labeled 19 bp oligo deoxyribonucleotide. The arrowhead indicates the position of the band. Specificity of binding was confirmed by the absence of the corresponding band in vector control lane (lane1).

earlier predicted to be IdeR-regulated, the present work suggests for the first time IdeR-dependent regulation of *fec*B of *M.tuberculosis*. FecB has been annotated as a probable Fe[III] dicitrate binding periplasmic lipoprotein. The *fec* operon is very well characterized in *E.coli* and a dyad repeat sequence GAAAATAATTCTTATTTCG present upstream to *fec*A has been proposed to serve as the binding site of the Fur iron repressor protein in *E.coli* (Zimmermann *et al*., 1984; Pressler *et al*., 1988). It is thus likely that FecB of *M.tuberculosis* could also be part of the iron transport complex of the bacterium and the regulation of the gene is brought about by IdeR, the Fur equivalent of *M.tuberculosis*. Additionally, as predicted by the method described above as well as the NCBI pattern search by Gold *et al*. (2001), another membrane protein coded by Rv1348c (similar to the Yersiniabactin uptake system) shows an IdeR box in its upstream sequence. This protein also appears to be an important candidate in the uptake of siderophore-like compounds.

Regulation of a probable MarR equivalent transcriptional regulator, Rv1404 by IdeR

Quite a few transcription regulators are known to be under the regulatory control of IdeR (Rodriguez *et al*., 2002). The results presented above could also identify Rv1404, a novel transcriptional regulator that shares some similarity to the multiple antibiotic resistance regulator (MarR) protein from *E.coli*, as a probable IdeR-regulated gene. If the antibiotic resistance regulator function of Rv1404 is proven, this could provide a clue to iron-dependent regulation of antibiotic resistance in *M.tuberculosis*. Here, it would be worth mentioning that the ORF Rv1846c that is also predicted to have an IdeR box in its upstream sequence also shows some similarity to the penicillase repressor protein of *Bacillus licheniformis*. These findings suggest that IdeR could be a global regulator that activates even other regulatory proteins that take care of the iron-dependent regulation of a broader network of *M.tuberculosis* genes.

Regulation of the urease operon by IdeR

The ORFs Rv1846 and Rv1847 have interesting predicted functions that are important in the context of the pathology of *M.tuberculosis*. While Rv1847 is a hypothetical protein, probably a thioesterase involved in the biosynthesis of aromatic compounds, Rv1846c codes for a transcription regulator with some similarity to the penicillase repressor protein of *B.licheniformis*. Interestingly, Rv1847 also appears to be part of the same operon that codes for genes involved in

the biosynthesis of the urease enzyme. It is known that *M.tuberculosis* survives in the acidic, toxic and hostile environment of the macrophage phagolysosome. One mechanism of survival is to somehow increase the pH of the phagolysosome. In this respect, the activity of the urease operon assumes importance as it could possibly help in neutralization of the acidic pH (Clemens *et al*., 1995). However, the mechanism of regulation of the *M.tuberculosis* urease operon has not yet been described anywhere. As an iron box exists upstream of the urease operon (directly upstream of ORF Rv1847), it was tempting to speculate that urease could also be regulated by IdeR. Additional evidence springs from the fact that in many pathogenic bacteria like *H.pylori*, the urease operon is regulated by ferric uptake regulatory (Fur) proteins (Bijlsma *et al*., 2002).

Along with the prediction of a high score, experimental evidence for binding of the IdeR of *M.tuberculosis* to an iron box element upstream of the urease operon has been provided in this work. Urease has been implicated in the virulence of several other pathogenic microorganisms. In *H.pylori*, *Salmonella typhimurium* and *E.coli*, urease is regulated by the ferric uptake regulator in response to pH (Bijlsma *et al*., 2002; Heimer *et al*., 2002). In the case of *M.tuberculosis*, ammonia generated by the action of urease may be of considerable importance in alkalinizing the microenvironment of the organism and preventing phagosome–lysosome fusion and phagosome acidification. In addition ammonia generated by the action of urease should be available to the organism for assimilation of nitrogen into biomolecules.

In summary, this study enhances the current understanding of the complex network of *M.tuberculosis* genes expressed/repressed as a consequence of iron stress. The study also adds considerably to the understanding of the various mechanisms of survival adopted by the bacterium to survive inside the iron deficient environment presented by the host.

ACKNOWLEDGEMENTS

Research in Hasnain's and Ranjan's laboratories was supported by grants from the Department of Biotechnology, Council of Scientific & Industrial Research (CSIR), Government of India. P.P. and S.Y. are recipients of Senior Research Fellowship from the CSIR.

REFERENCES

Bijlsma,J.J. *et al*. (2002) The *Helicobacter pylori* homologue of the ferric uptake regulator is involved in acid resistance. *Infect. Immun.*, **70**, 606–611.

Boyd,J. *et al*. (1990) Molecular cloning and DNA sequence analysis of a diphtheria tox iron-dependent regulatory element (dtxR) from *Corynebacterium diphtheriae. Proc. Natl Acad. Sci. USA*, **87**, 5968–5972.

Camacho,L.R. *et al*. (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol. Microbiol.*, **34**, 257–267.

Clemens,D.L. *et al*. (1995) Purification, characterization, and genetic analysis of *Mycobacterium tuberculosis* urease, a potentially critical determinant of host-pathogen interaction. *J. Bacteriol.*, **177**, 5644–5652.

Dussurget,O. *et al*. (1996) An ideR mutant of *Mycobacterium smegmatis* has derepressed siderophore production and an altered oxidative-stress response. *Mol. Microbiol.*, **22**, 535–544.

Dussurget,O. *et al*. (1999) Transcriptional control of the iron-responsive *fxb*A gene by the mycobacterial regulator IdeR. *J. Bacteriol.*, **181**, 3402–3408.

Gold,B. *et al*. (2001) The *Mycobacterium tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Mol. Microbiol.*, **42**, 851–865.

Heimer,S.R. *et al*. (2002) Urease of enterohemorrhagic *Escherichia coli*: evidence for regulation by fur and a trans-acting factor. *Infect. Immun.*, **70**, 1027–1031.

Litwin,C.M. and Calderwood,S.B. (1993) Role of iron in regulation of virulence genes. *Clin. Microbiol. Rev.*, **6**, 137–149.

Pressler,U. *et al*. (1988) Genetics of the iron dicitrate transport system of *Escherichia coli. J. Bacteriol.*, **170**, 2716–2724.

Quadri,L.E. *et al*. (1998) Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem. Biol.*, **5**, 631–645.

Rodriguez,G.M. and Smith,I. (2003) Mechanisms of iron regulation in mycobacteria: role in physiology and virulence. *Mol. Microbiol.*, **47**, 1485–1494.

Rodriguez,G.M. *et al*. (1999) Identification and characterization of two divergently transcribed iron regulated genes in *Mycobacterium tuberculosis. Tuber Lung Dis.*, **79**, 287–298.

Rodriguez,G.M. *et al*. (2002) *ide*R, an essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect. Immun.*, **70**, 3371–3381.

Schmitt,M.P. and Holmes,R.K. (1991) Iron-dependent regulation of diphtheria toxin and siderophore expression by the cloned *Corynebacterium diphtheriae* repressor gene *dtx*R in *C.diphtheriae* C7 strains. *Infect. Immun.*, **59**, 1899–1904.

Schmitt,M.P. *et al*. (1995) Characterization of an iron-dependent regulatory protein (IdeR) of *Mycobacterium tuberculosis* as a functional homolog of the diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheriae. Infect. Immun.*, **63**, 4284–4289.

Tao,X. *et al*. (1992) Specific binding of the diphtheria tox regulatory element DtxR to the tox operator requires divalent heavy metal ions and a 9-base-pair interrupted palindromic sequence. *Proc. Natl Acad. Sci. USA*, **89**, 5897–5901.

Yellaboina,S. *et al*. (2004a) Prediction of DtxR regulon: Identification of binding sites and operons controlled by Diphtheria toxin repressor in *Corynebacterium diphtheriae. BMC Microbiol.*, **4**, 38.

Yellaboina,S. *et al*. (2004b) Predictregulon: a webserver for the prediction of the regulatory protein binding sites and operons in prokaryote genomes. *Nucleic Acids Res.*, **32**, W318–W320.

Zimmermann,L. *et al*. (1984) Exogenous induction of the iron dicitrate transport system of *Escherichia coli* K-12. *J. Bacteriol.*, **159**, 271–277.