

Research article

Open Access

## Prediction of DtxR regulon: Identification of binding sites and operons controlled by Diphtheria toxin repressor in *Corynebacterium diphtheriae*

Sailu Yellaboina<sup>1</sup>, Sarita Ranjan<sup>1</sup>, Prachee Chakhaiyar<sup>2</sup>,  
Seyed Ehtesham Hasnain<sup>2</sup> and Akash Ranjan\*<sup>1</sup>

Address: <sup>1</sup>Computational and Functional Genomics Group, Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500076, INDIA and <sup>2</sup>Laboratory of Cellular and Molecular Biology, Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500076, INDIA

Email: Sailu Yellaboina - sailu@cdfd.org.in; Sarita Ranjan - sarita@cdfd.org.in; Prachee Chakhaiyar - prachee@cdfd.org.in; Seyed Ehtesham Hasnain - ehtesham@cdfd.org.in; Akash Ranjan\* - akash@cdfd.org.in

\* Corresponding author

Published: 24 September 2004

BMC Microbiology 2004, 4:38 doi:10.1186/1471-2180-4-38

This article is available from: <http://www.biomedcentral.com/1471-2180/4/38>

Received: 08 April 2004

Accepted: 24 September 2004

© 2004 Yellaboina et al; licensee BioMed Central Ltd.

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

### Abstract

**Background:** The diphtheria toxin repressor, DtxR, of *Corynebacterium diphtheriae* has been shown to be an iron-activated transcription regulator that controls not only the expression of diphtheria toxin but also of iron uptake genes. This study aims to identify putative binding sites and operons controlled by DtxR to understand the role of DtxR in patho-physiology of *Corynebacterium diphtheriae*.

**Result:** Positional Shannon relative entropy method was used to build the DtxR-binding site recognition profile and the later was used to identify putative regulatory sites of DtxR within *C. diphtheriae* genome. In addition, DtxR-regulated operons were also identified taking into account the predicted DtxR regulatory sites and genome annotation. Few of the predicted motifs were experimentally validated by electrophoretic mobility shift assay. The analysis identifies motifs upstream to the novel iron-regulated genes that code for Formamidopyrimidine-DNA glycosylase (FpG), an enzyme involved in DNA-repair and starvation inducible DNA-binding protein (Dps) which is involved in iron storage and oxidative stress defense. In addition, we have found the DtxR motifs upstream to the genes that code for sortase which catalyzes anchoring of host-interacting proteins to the cell wall of pathogenic bacteria and the proteins of secretory system which could be involved in translocation of various iron-regulated virulence factors including diphtheria toxin.

**Conclusions:** We have used an *in silico* approach to identify the putative binding sites and genes controlled by DtxR in *Corynebacterium diphtheriae*. Our analysis shows that DtxR could provide a molecular link between Fe<sup>+2</sup>-induced Fenton's reaction and protection of DNA from oxidative damage. DtxR-regulated Dps prevents lethal combination of Fe<sup>+2</sup> and H<sub>2</sub>O<sub>2</sub> and also protects DNA by nonspecific DNA-binding. In addition DtxR could play an important role in host interaction and virulence by regulating the levels of sortase, a potential vaccine candidate and proteins of secretory system.

## Background

Iron is an important inorganic component of a cell. Iron is required as co-factor for various essential enzymes and proteins some of which are involved in electron transport (Cytochromes), redox reactions (oxidoreductases) and regulation of gene expression (fumarate-nitrate reduction regulatory protein, iron-binding protein) [1]. However a higher level of intracellular iron can catalyze formation of hydroxyl radicals and reactive oxygen species through Fenton's reaction which could be lethal to the cell [2]. Hence, a careful regulation of iron-requiring enzymes/proteins and iron uptake proteins/enzymes is required for the survival of bacteria.

Inorganic iron is also known to influence virulence in many pathogenic bacteria such as *Corynebacterium diphtheriae*, *Escherichia coli*, and *Bordetella bronchiseptica* [3-5]. The diphtheria toxin repressor DtxR is known as an iron-activated global transcription regulator that represses the transcription of various iron-dependent genes in *C. diphtheriae* [6,7]. Eight DtxR-binding sites in upstream sequences of operons/genes named as *tox*, *hmuO*, *irp1*, *irp2*, *irp3*, *irp4*, *irp5* and *irp6* have been identified by DNA footprinting methods [6]. The product of *tox* gene is diphtheria toxin which catalyzes the NAD-dependent ADP ribosylation of eukaryotic aminoacyl-transferase-II, thereby causing inhibition of protein synthesis and subsequent death of the host. The *hmuO* gene, which encodes a haem oxygenase, oxidizes the haem to release free iron. The operons *irp1* and *irp6* encode the products with homology to ABC-type ferric-siderophore transport systems. The gene *irp3* encodes a homologue of AraC-type transcriptional activators. The products of *irp2*, *irp4* and *irp5* do not show any homology to the other known proteins. In addition, *C. diphtheriae* with inactive DtxR has been shown to be sensitive to killing by exposure to high iron conditions or hydrogen peroxide than the wild type [8].

This work uses an *in silico* method to identify additional DtxR-binding sites and target genes to understand the role of DtxR in virulence and patho-physiology of *C. diphtheriae*.

## Results

### **In silico identification of putative DtxR-binding sites**

Experimentally characterized DtxR-binding motifs were collected from the literature (Table 1). These binding sites were used to identify additional putative DtxR-binding sites along with associated operons in *C. diphtheriae* NCTC13129 genome (see materials and methods). Table 2 shows the predicted DtxR-binding sites with score 3.7438 or more. We could identify five (*tox*, *irp4*, *irp5*, *irp6* and *hmuO*) of the eight known DtxR-binding sites, in sequenced *C. diphtheriae* NCTC13129 genome. We could not find *irp1* and *irp2* motifs as the corresponding genes (*irp1*, *irp2*) are not present in the sequenced strain NCTC13129 [9]. The regulator binding sites of *irp3*, *irp4* and *irp6* genes in the strain NCTC13129 shows one base change from the binding sites reported in strain C7 [6]. Binding site of *irp3* gene (TTAGCTGACACCCACCCAT) although exists in strain NCTC13129, but not there in the predicted sites, because it is located within the coding region of *irp3* ORF. The predicted ORF of *irp3* in the sequenced strain NCTC13129 has different start position and is larger than what was previously reported in strain C7 [9,10].

In addition, we have identified binding sites in upstream sequences of eight genes recently reported to be regulated by DtxR [7]. However, our prediction differs from the previous report for five (*secY*, *deoR*, *chtA*, *frgA*, *sidA*) of the seven sites which were identified by BLAST search (Table 2). Our prediction agreed with the previous report that the genes such as *recA* (DIP1450) and *yujA* (DIP1735) are not under a direct DtxR regulation as we could not detect any motif upstream to these gene with scores above the cutoff value [7].

**Table 1: Known DtxR-binding sites from *C. diphtheriae***

Binding site	Gene	Product	Reference
TTAGGATAGCTTACCTAA	<i>tox</i>	Diphtheria toxin	[25]
TTAGGTTAGCCAAACCTTT	<i>irp1</i>	Periplasmic protein of siderophore transport system	[26]
GCAGGGTAGCCTAACCTAA	<i>irp2</i>	Hypothetical protein	[26]
TTAGGTGAGACGCACCCAT	<i>irp3</i>	AraC-type transcription regulator	[10]
ATTACTAACGCTAACCTAA	<i>irp4</i>	Hypothetical protein	[10]
CTAGGATTGCCTACACTTA	<i>irp5</i>	Hypothetical protein	[10]
TTCCCTTGCCTAGCCTAA	<i>irp6</i>	Periplasmic protein of siderophore transport system	[6]
TGAGGGGAACCTAACCTAA	<i>hmuO</i>	Haem oxygenase	[27]

**Table 2: Predicted DtxR-binding sites in *C. diphtheriae***

Score	Position	Site	Gene	Synonym	Product
4.45904	-80	TGAGGGGAACCTAACCTAA	<i>hmuO</i>	DIP1669**	heme oxygenase
4.39003	-52	TTAGGATAGCTTACCTAA	<i>Tox</i>	DIP0222**	Diphtheria toxin precursor
4.25877	-60	ATAGGCTACACTTACCTAA	-	DIP0624	Putative membrane protein
4.21068	-168	TTGGATTAGCCTACCTAA	-	DIP2162**	ABC-type peptide transport system periplasmic component
4.2033	-21	TTAGGGTAGCTTGCCTAA	<i>iucA</i>	DIP0586	Putative siderophore biosynthesis related protein
4.17632	-78	ATAGGCATGCCTAACCTCA	-	DIP2330	Putative membrane protein
4.07921	-130	TTAGGTCAGGGTACCTAA	-	DIP0370	Putative succinate dehydrogenase cytochrome B subunit
4.03559	-30	TTAGCTAACCTTGCCTAT	<i>arsR</i>	DIP0415	Putative ArsR family regulatory protein
4.01967	-239	TTAGGGTAGGCTAATCAA	<i>sidA</i> *	DIP2161	nonribosomal peptide synthase
3.99985	-74	TTTTCTTGCCTAGCCTAA	<i>irp6A</i>	DIP0108**	Ferrisiderophore receptor Irp6A
3.99195	-241	TTAGGCACCCCTAACCTAG	-	DIP0539	Putative sugar ABC transport system ATP-binding protein
3.98554	-72	TTAGCTTAGCCCTAGCTAA	-	DIP0169	Putative secreted protein
3.9296	-26	CTAGGATTGCCTACACTTA	<i>lrp5</i>	DIP0894**	Conserved hypothetical protein
3.9073	-93	GTTGGGTTGCCAACCTAC	-	DIP2106	Putative ABC transport system, ATP-binding subunit
3.89763	-86	ATAGGTTAGGTTAACCTTG	<i>chtA</i> *	DIP1520	Putative membrane protein
3.89676	-130	TTGTGTTAGCCTAGGCTAA	<i>secA</i>	DIP0699	Translocase protein
3.89169	-26	TTGGGGTGGCCTATCCTTA	-	DIP2304	Putative DNA-repair glycosylase
3.88042	-172	TTAGGTAAGTGTAGCCTAT	<i>htaA</i> *	DIP0625	Putative membrane protein
3.86534	-69	ATTACTAATGCTAACCTAA	<i>lrp4</i>	DIP2356**	Putative conserved membrane protein
3.85539	-173	TTAGGGTGGGCTAACCTGC	<i>deoR</i> *	DIP1296	Putative DNA-binding protein
3.84889	-75	TTAGGGAACCTTGCCTTA	<i>piuB</i> *	DIP0124	Putative membrane protein
3.83816	-121	TTAGCTAGGGCTAAGCTAA	-	DIP0168	Putative glycosyl transferase
3.83576	-219	GTAACAAAGGCAAGCCTAA	<i>xerD</i>	DIP1510	Putative integrase/recombinase
3.8224	-216	ATAGGCAAGGTTAAGCTAA	-	DIP0417	Putative membrane protein
3.81905	-47	GTTGGACAGGTTACCTAA	<i>frgA</i> *	DIP1061	Putative iron-siderophore uptake system permease
3.8148	-37	TGTGGCACACCAACCTAA	-	DIP2272	possible sortase-like protein
3.76235	-136	TTGGGGTTGCCCTCCTAA	-	DIP0142	Hypothetical protein
3.76233	-268	CTAGGTTAGGGTGCCTAA	<i>secY</i> *	DIP0540	preprotein translocase SecY subunit
3.74673	-110	TAAACATAGCCAACCAA	<i>nrdF1</i>	DIP1865	ribonucleotide reductase beta-chain I
3.7438	-81	TAAGGATAGGCCACCCCAA	<i>Dps</i>	DIP2303	Starvation inducible DNA-binding protein

Note: \*\*Indicate the gene synonym with experimentally identified binding site in *C. diphtheriae* [6]. \* Indicates the genes known to be regulated by DtxR [7]. The binding sites in Italics were verified by EMSA. The gene pairs, DIP0624-DIP0625, DIP2161-DIP2162, DIP0168-DIP0169, DIP0539-DIP0540 and DIP2303-DIP2304 are divergently transcribed and contain common regulatory regions.

### Experimental validation of predicted binding sites

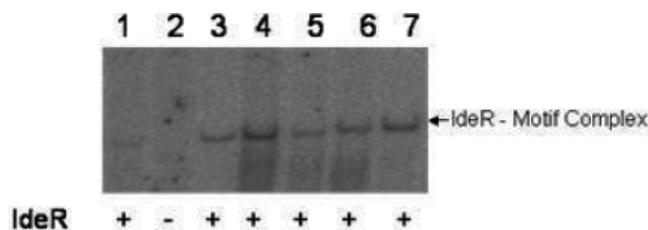
Since our approach to identify DtxR-regulated genes is purely computational in nature, we decided to test the validity of our predictions. A sample of predicted regulator binding motifs (Table 2) (upstream to ORFs: DIP2161, DIP0699, DIP0586, DIP2304, DIP2272) were experimentally verified by EMSA using IdeR, an orthologue of DtxR from *M. tuberculosis*. DtxR and IdeR are iron-dependent regulators. A pair wise sequence comparison of the two proteins shows a high (58%) overall sequence identity (similarity 72%) which increases further to 92% identity and 100% similarity in DNA recognition domain. In addition, the structural comparison of two regulators also shows a very similar 3D organization, suggesting that the IdeR regulator would be able to recognize the DtxR motif [11].

Synthetic double stranded oligonucleotides corresponding to DNA-binding sites were labeled with  $^{32}\text{P}$  and mixed

with purified IdeR in presence of manganese ions and was assayed for the formation of DNA-protein complex using EMSA. Manganese was used as the divalent metal in the binding reactions on account of its redox stability compared with ferrous ion. Electrophoretic mobility of all five double stranded oligonucleotides tested was retarded by IdeR (Figure 1). However a synthetic motif (TTTTCAT-GACGTCTTCTAA) used as a negative control did not show any complex formation. These results indicate that the predicted DtxR-binding sites can indeed bind to DtxR.

### Identification and annotation of DtxR-regulated genes *C. diphtheriae* genome

In addition to the binding site prediction, we have also identified co-regulated genes (operons) downstream to the predicted DtxR-binding site (Table 3). Function of the proteins encoded by the putative genes in Table 2 and Table 3 was predicted by RPS-BLAST search against conserved domain database [12].

**Figure 1**

IdeR binds the predicted DtxR-binding DNA fragments. 30 pmoles of IdeR was added to  $^{32}\text{P}$ -labelled DNA probes in the presence of 200  $\mu\text{M}$   $\text{Mn}^{2+}$ , and complexes were resolved on a 7% Tris-borate polyacrylamide gel containing 150  $\mu\text{M}$   $\text{Mn}^{2+}$ . Lane 1: Control gel retardation using Radiolabeled DNA motif without DtxR-binding site. Lane 2: Radiolabeled DIP2161 motif without IdeR. Lane 3: Radiolabeled DIP2161 motif with IdeR. Lane 4: Radiolabeled DIP0699 motif with IdeR. Lane 5: Radiolabeled DIP0586 motif with IdeR. Lane 6: Radiolabeled DIP2304 motif with IdeR. Lane 7: Radiolabeled DIP2272 motif with IdeR.

## Discussion

Our analysis identified putative DtxR motifs upstream to various operons/genes which could be involved in siderophore biosynthesis, ABC-type transport systems, iron storage, oxidative stress defense and iron-sulfur cluster biosynthesis. In addition, we have also identified the motifs upstream of operons that could be involved in anchoring of host-interacting proteins to the cell wall and secretion of various virulence factors. Important functions of some of these DtxR-regulated genes and their role in *C. diphtheriae* physiology are discussed here.

### Regulation of siderophore biosynthesis and ABC-type transport systems

Predicted member of the DtxR regulon, the gene DIP0586, codes for the IucA/IucC family of enzymes that catalyze discrete step in the biosynthesis of the aerobactin [13]. In addition to known DtxR-regulated siderophore transport genes (irp1, irp6), DtxR could also regulate other ABC-type transport systems similar to Manganese/Zinc, peptide/Nickel and multidrug subfamilies of ABC transporters. The peptide/nickel transport system (DIP2162-DIP2165) has been suggested to be recently acquired by pathogenic *C. diphtheriae* [9].

### Regulation of iron storage and oxidative stress defense

We predict that DtxR could regulate divergently transcribed genes DIP2303 and DIP2304 whose products are similar to starvation inducible DNA-binding protein (Dps) and Formamidopyrimidine-DNA glycosylase (Fpg), respectively. Dps in *Escherichia coli* is induced in response to oxidative or nutritional stress and protects DNA from oxidative stress damage by nonspecific binding

[14]. Dps also catalyzes oxidation of ferrous iron to ferric iron by hydrogen peroxide ( $2\text{Fe}^{2+} + \text{H}_2\text{O}_2 + 2\text{H}_2\text{O} \rightarrow 2\text{Fe}^{3+}\text{OOH}_{(\text{core})} + 4\text{H}^+$ ) which in turn prevents hydroxyl radical formation by Fenton's reaction ( $\text{Fe}^{2+} + \text{H}_2\text{O}_2 \rightarrow \text{Fe}^{3+} + \text{HO}^- + \text{HO}^-$ ) and thereby prevents subsequent DNA damage [15]. The enzyme, formamidopyrimidine-DNA glycosylase is a primary participant in the repair of 8-oxo-guanine, an abundant oxidative DNA lesion [16]. The gene DIP1510 which codes for the site-specific recombinase XerD could also be regulated by DtxR. The *xerD* gene in *E. coli* belongs to the oxidative stress regulon [17].

### Regulation of proteins involved in iron-sulfur cluster biosynthesis and iron-sulfur cluster containing proteins

We predict that the operon DIP1288-DIP1296, which is similar to the *suf* operon of *E. coli*, could be regulated by DtxR. The *suf* operon in bacteria encodes the genes for Fe-S cluster assembly machinery [18]. In addition, genes encoding the iron-sulfur containing proteins such as succinate dehydrogenase (Sdh), cytochrome oxidase (CtaD) and Ribonucleotide reductase (NrdF1) in *C. diphtheriae* also show DtxR motif in their upstream sequences.

### Regulation of sortases

We predict that DtxR could regulate the recently acquired pathogenic island DIP2271-DIP2272, encoding the sortase *srtA* and hypothetical protein, respectively [9]. Sortases are membrane-bound trans-peptidases that catalyze the anchoring of surface proteins to the cell wall peptidoglycan [9]. Such systems are often used by gram-positive pathogens to anchor host-interacting proteins to the bacterial surface [19].

### Regulation of protein translation and translocation system

DtxR could regulate two operons that contain genes DIP0699 (*secA*) and DIP0540 (*secY*) that code for the protein translocation system. The *secY*-containing operon, which is similar to the streptomycine operon *spc* from *B. subtilis* and other bacteria, involves the genes required for protein translation and translocation [20]. The operon contains additional sialidase gene (DIP0543) in comparison to non pathogenic *Corynebacterium* species. Activity of sialidase has been linked to virulence in several other microbial pathogens and may enhance fimbriae mediated adhesion in *Corynebacterium diphtheriae* by unmasking receptors on mammalian cells [9].

The Sec system can both translocate proteins across the cytoplasmic membrane and insert integral membrane proteins into it. The former proteins but not the latter possess N-terminal, cleavable, targeting signal sequences that are required to direct the proteins to the Sec system. Some of the DtxR-regulated genes including diphtheria toxin (Table 4) show predicted signal sequences by SignalP 3.0 [21] and hence they may play an important role in host interaction and virulence of *Corynebacterium diphtheriae* [9].

**Table 3: Predicted DtxR-regulated operons in *C. diphtheriae***

Synonym	Gene	Orthologue	Product
DIP2158		COG1131	ABC-type transport system permease and ATPase component
DIP2159		COG1131	ABC-type transport system permease and ATPase component
DIP2160	-	COG3321	Polyketide synthase modules and related proteins
DIP2161*	-	COG1020	Non-ribosomal peptide synthetase modules and related proteins
DIP0586	<i>iucA</i>	Pfam04183	Catalyse discrete steps in biosynthesis of the siderophore aerobactin
DIP0587	-	-	Putative membrane protein
DIP0588	-	-	Putative membrane protein
DIP1059	<i>fepC</i>	COG1120	ABC-type cobalamin/Fe3+-siderophores transport systems
DIP1060	<i>fepG</i>	COG4779	ABC-type enterobactin transport system
DIP1061*	<i>fepD</i>	COG0609	ABC-type Fe3+-siderophore transport system
DIP2162	<i>ddpA</i>	COG0747	ABC-type peptide transport system periplasmic component
DIP2163	<i>ddpB</i>	COG0601	ABC-type peptide/nickel transport systems permease components
DIP2164	<i>ddpC</i>	COG1173	ABC-type peptide/nickel transport systems permease components
DIP2165	<i>ddpD</i>	COG0444	ABC-type peptide/nickel transport systems ATPase component
DIP0169	<i>lral</i>	COG0803	ABC-type metal ion transport system, periplasmic component
DIP0170	<i>znuC</i>	COG1121	ABC-type Mn/Zn transport systems, ATPase component
DIP0171	<i>znuB</i>	COG1108	ABC-type Mn2+/Zn2+ transport systems, permease components
DIP0172	<i>znuB</i>	COG1108	ABC-type Mn2+/Zn2+ transport systems, permease components
DIP0173	<i>lral</i>	COG0803	ABC-type metal ion transport system, periplasmic component
DIP2106	<i>mdrB</i>	COG1131	ABC-type multidrug transport system, ATPase and permease component
DIP2107	<i>mdrB</i>	COG1131	ABC-type multidrug transport system, ATPase and permease component
DIP0625	<i>htaa</i>	Pfam04213	Haemin transporter associated protein
DIP0626	<i>hmuT</i>	COG4558	ABC-type haemin transport system
DIP0627	<i>hmuU</i>	COG0609	ABC-type Fe3+-siderophore transport system
DIP0628	<i>hmuV</i>	COG4559	ABC-type haemin transport system
DIP0629*	<i>htaa</i>	Pfam04213	Haemin transporter associated protein
DIP1519*	<i>htaa</i>	pfam04213	Haemin transporter associated protein
DIP1520*	<i>htaa</i>	pfam04213	Haemin transporter associated protein
DIP2303	<i>dps</i>	COG0783	Starvation inducible DNA-binding protein
DIP2304	-	COG0266	Formamidopyrimidine-DNA glycosylase
DIP2305	-	COG0063	Predicted sugar kinase
DIP1510	<i>xerD</i>	COG4974	Site-specific recombinase
DIP1288	-	-	Conserved hypothetical protein
DIP1289	<i>uup</i>	COG0488	ATPase components of ABC transporters with duplicated ATPase domains
DIP1290	-	COG2151	Predicted metal-sulfur cluster biosynthetic enzyme
DIP1291	<i>iscU</i>	COG0822	NifU homolog involved in Fe-S cluster formation
DIP1292	<i>csd</i>	COG0520	Selenocysteine lyase
DIP1293	<i>sufC</i>	COG0396	ABC-type transport system involved in Fe-S cluster assembly
DIP1294	-	COG0719	ABC-type transport system involved in Fe-S cluster assembly
DIP1295	<i>sufB</i>	COG0719	ABC-type transport system involved in Fe-S cluster assembly
DIP1296*	<i>deoR</i>	COG2345	DeoR family transcriptional regulator
DIP0370	-	-	Putative succinate dehydrogenase (cytochrome b)
DIP0371	-	COG1053	Succinate dehydrogenase/fumarate reductase
DIP0372	-	COG0479	Succinate dehydrogenase/fumarate reductase
DIP0373	-	-	Putative membrane protein
DIP0374	-	-	Putative membrane protein

**Table 3: Predicted DtxR-regulated operons in *C. diphtheriae* (Continued)**

DIP0375	-	-	Putative membrane protein
DIP0376	-	-	Putative membrane protein
DIP0377	-	-	Putative membrane protein
DIP1864	<i>ctaD</i>	COG0843	Heme/copper-type cytochrome/quinol oxidases
DIP1865	<i>nrdF1</i>	COG0208	Ribonucleotide reductase
DIP2330	-	-	Putative membrane protein
DIP2331	-	COG1012	NAD-dependent aldehyde dehydrogenases
DIP0124*	-	Pfam03929	Uncharacterized iron-regulated membrane protein (DUF337)
DIP0622	-	-	Putative membrane protein
DIP0623	<i>metA</i>	COG2021	Homoserine acetyltransferase
DIP0624	-	-	Putative membrane protein
DIP0415	-	Pfam01022	Bacterial regulatory protein
DIP0539	-	COG3839	ABC-type sugar transport systems
DIP0168	-	-	Putative glycosyl transferase
DIP0417	-	-	Putative membrane protein
DIP0142	-	-	Hypothetical protein
DIP0143	-	-	-
DIP0144	<i>tra8</i>	COG2826	Transposase and inactivated derivatives
DIP2271	-	-	Putative membrane protein
DIP2272	-	COG3764	Sortase (surface protein transpeptidase)
DIP0699	<i>secA</i>	COG0653	Preprotein translocase subunit SecA (ATPase)
DIP0700	-	-	Hypothetical protein
DIP0540*	<i>secY</i>	Pfam00344	Eubacterial secY protein
DIP0541	<i>Adk</i>	COG0563	Adenylate kinase and related kinases
DIP0542	<i>mapA</i>		Methionine aminopeptidase
DIP0543	-	-	Sialidases or neuraminidases;
DIP0544	<i>erfK</i>	Pfam03734	This family of proteins contains a conserved histidine and cysteine
DIP0545	<i>infA</i>	COG0361	Translation initiation factor I (IF-I)
DIP0546	<i>rpsM</i>	COG0099	Ribosomal protein S13
DIP0547	<i>rpsK</i>	COG0100	Ribosomal protein S11
DIP0548	<i>rpsD</i>	COG0522	Ribosomal protein S4 and related proteins
DIP0549	<i>rpoA</i>	COG0202	DNA-directed RNA polymerase
DIP0550	<i>rplQ</i>	COG0203	Ribosomal protein L17
DIP0551	<i>truA</i>	COG0101	Pseudouridylate synthase

Note: \* Indicate the genes reported be regulated by DtxR. Genes listed together belongs to same operon.

## Conclusions

The bioinformatics method used to predict the targets of DtxR in *C. diphtheriae* NCTC13129 genome is promising, as some of the predicted targets were experimentally verified. The approach identified novel DtxR-regulated genes, which could play an important role in physiology of *C. diphtheriae* NCTC13129. DtxR, generally known as a repressor of diphtheriae toxin and iron siderophore/transport genes, can also regulate other metal ion transport genes, iron storage, oxidative stress, DNA-repair, biosyn-

thesis of iron-sulfur cluster, Fe-S-cluster containing proteins, and even protein sortase and translocation systems.

## Methods

### Source of genome sequence

The complete genome sequence of *C. diphtheriae* was downloaded from NCBI ftp site [22], and the DtxR-binding sites identified by experimental methods were collected from literature [6,10,25-27].

**Table 4: DtxR-regulated genes containing the potential signal sequence**

Gene	Product
DIP0222	Diphtheria toxin
DIP0109	IRP6B
DIP2356	IRP4
DIP2162	ABC-type peptide transport system periplasmic component
DIP0172	Putative membrane protein
DIP2107	Putative integral membrane transport protein
DIP0625	Haemin transporter associated protein
DIP0626	ABC-type haemin transport system
DIP0627	ABC-type haemin transport system
DIP1519	Haemin transporter associated protein
DIP0629	Haemin transporter associated protein
DIP1520	Haemin transporter associated protein
DIP2330	Putative membrane protein
DIP0543	Sialidases or neuraminidases

#### Prediction of DtxR-binding sites

DtxR-binding site recognition profile was calculated by positional Shannon relative entropy method [23,24]. The positional relative entropy  $Q_i$  at position  $i$  in a binding site is defined as

$$Q_i = \sum_{b=A,T,G,C} f_{b,i} \log_{10} \frac{f_{b,i}}{q_b}$$

where  $b$  refers to each of the possible base (A, T, G, C),  $f_{b,i}$  is observed frequency of each base at position  $i$  and  $q_b$  is the frequency of base  $b$  in the genome sequence. The contribution of each base to the positional Shannon's relative entropy is calculated by multiplying positional frequency of each base with positional relative entropy. The binding site profile thus generated was used to scan upstream sequences of all the genes of the *Corynebacterium diphtheriae* genome. The score of each site is calculated as the sum of the respective positional Shannon relative entropy of each of the four possible bases. A maximally scoring site is selected from the upstream sequence of each gene. The lowest score among the input binding sites is considered as cut-off score. The sites scoring higher than the cut-off value are reported as potential binding sites conforming to the consensus sequence.

#### Prediction of operons

Co-directionally transcribed genes, downstream to the predicted binding site were selected as potential co-regulated genes (operons) according to one of the following criteria (a) Co-directionally transcribed orthologous gene pairs, conserved in at least 4 genomes; (b) genes belong to the same cluster of orthologous gene function category and the intergenic distance is less than 200 base pairs; (c) the first three letters in gene names are identical (gene names for

putative genes were assigned from COG database); (d) intergenic distance is less than 90 base pairs [24].

#### Functional assignment of genes

The function of predicted genes was inferred using the RPS-BLAST search against conserved domain database [12]. These genes were further classified according to their function.

#### Expression and purification of IdeR

The iron-dependent regulator IdeR from *M. tuberculosis* was expressed from a recombinant pRSET vector containing the IdeR gene fused to a six His affinity tag (P. Chakhiyar unpublished). The expressed protein was first purified using Ni-NTA Metal Chelate Affinity chromatography; later it was desalted and concentrated using Centricon Ultra filtration device. The concentration of the recombinant protein was estimated using Bradford method.

#### Electrophoretic mobility shift assay

Double-stranded oligonucleotides containing the predicted binding motif (19 bp long) were end labeled with T4 polynucleotide kinase and [ $\gamma^{32}$ P]-ATP and were incubated with the recombinant purified IdeR protein in a binding reaction mixture. The binding reaction mixture (20- $\mu$ l total volume) contain the DNA-binding buffer (20 mM Tris-HCl [pH 8.0], 2 mM DTT, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 50% glycerol, 5  $\mu$ g of bovine serum albumin per ml), 10  $\mu$ g of poly(dI-dC) per ml (for nonspecific binding) and 200  $\mu$ M MnCl<sub>2</sub>. The reaction mixture was incubated at room temperature for 30 min. Approximately 2  $\mu$ l of the tracking dye (50% sucrose, 0.6% bromophenol blue) was added to the reaction mixture at the end of incubation and was loaded onto 7% polyacrylamide gel containing 150  $\mu$ M MnCl<sub>2</sub> in 1  $\times$  Tris-borate-EDTA buffer. The gel was electrophoresed at 200 V for 2 hours. Subsequently the gel was dried and exposed to Fuji Storage Phosphor Image Plates for 16 hours. The image plates were subsequently scanned in Fuji Storage Phosphor Imaging workstation.

#### List of abbreviations

DtxR – Diphtheria toxin repressor; IdeR – Iron-dependent regulator; Dps – DNA-binding protein from starved cells; RPS-BLAST – Reversed Position Specific – Basic Local Alignment Search Tool; EMSA – Electrophoretic Mobility Shift Assay

#### Authors' contributions

SY: carried out the computation, data analysis, and manuscript preparation. SR: Carried out the EMSA and drafted the manuscript. PC: provided the cloned IdeR construct, drafted the manuscript. SH: Manuscript preparation and coordination. AR: Design of the study and coordination. All authors read and approved the final manuscript.

## Acknowledgements

This work is partially supported by CSIR NMITLI Grant to AR. SR is supported by CSIR NMITLI Grant. YS and PC is supported by CSIR Research Fellowships.

## References

1. Castagnetto JM, Hennessy SW, Roberts VA, Getzoff ED, Tainer JA, Pique ME: **MDB: the Metalloprotein Database and Browser at The Scripps Research Institute.** *Nucleic Acids Res* 2002, **30**:379-382.
2. Urbanski NK, Beresewicz A: **Generation of \*OH initiated by interaction of Fe<sup>2+</sup> and Cu<sup>+</sup> with dioxygen; comparison with the Fenton chemistry.** *Acta Biochim Pol* 2000, **47**:951-962.
3. Tao X, Schiering N, Zeng HY, Ringe D, Murphy JR: **Iron, DtxR, and the regulation of diphtheria toxin expression.** *Mol Microbiol* 1994, **14**:191-197.
4. Russo TA, Carlino UB, Johnson JR: **Identification of a new iron-regulated virulence gene, ireA, in an extraintestinal pathogenic isolate of *Escherichia coli*.** *Infect Immun* 2001, **69**:6209-6216.
5. Register KB, Ducey TF, Brockmeier SL, Dyer DW: **Reduced virulence of a *Bordetella bronchiseptica* siderophore mutant in neonatal swine.** *Infect Immun* 2001, **69**:2137-2143.
6. Qian Y, Lee JH, Holmes RK: **Identification of a DtxR-regulated operon that is essential for siderophore-dependent iron uptake in *Corynebacterium diphtheriae*.** *J Bacteriol* 2002, **184**:4846-4856.
7. Kunkle CA, Schmitt MP: **Analysis of the *Corynebacterium diphtheriae* DtxR Regulon: Identification of a putative siderophore synthesis and transport system that is similar to the *Yersinia* high-pathogenicity island-encoded yersiniabactin synthesis and uptake system.** *J Bacteriol* 2003, **185**:6826-6840.
8. Oram DM, Avdalovic A, Holmes RK: **Construction and characterization of transposon insertion mutations in *Corynebacterium diphtheriae* that affect expression of the diphtheria toxin repressor (DtxR).** *J Bacteriol* 2002, **184**:5723-5732.
9. Cerdeno-Tarraga AM, Efstratiou A, Dover LG, Holden MT, Pallen M, Bentley SD, Besra GS, Churcher C, James KD, De Zoysa A, Chillingworth T, Cronin A, Dowd L, Feltwell T, Hamlin N, Holroyd S, Jagels K, Moule S, Quail MA, Rabbinowitsch E, Rutherford KM, Thomson NR, Unwin L, Whitehead S, Barrell BG, Parkhill J: **The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129.** *Nucleic Acids Res* 2003, **31**:6516-6523.
10. Lee JH, Wang T, Ault K, Liu J, Schmitt MP, Holmes RK: **Identification and characterization of three new promoter/operators from *Corynebacterium diphtheriae* that are regulated by the diphtheria toxin repressor (DtxR) and iron.** *Infect Immun* 1997, **65**:4273-4280.
11. Feese MD, Ingason BP, Goranson-Siekierke J, Holmes RK, Hol WG: **Crystal structure of the iron-dependent regulator from *Mycobacterium tuberculosis* at 2.0-Å resolution reveals the Src homology domain 3-like fold and metal binding function of the third domain.** *J Biol Chem* 2001, **276**:5959-66.
12. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madaj T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH: **CDD: a curated Entrez database of conserved domain alignments.** *Nucleic Acids Res* 2003, **31**:383-387.
13. de Lorenzo V, Neilands JB: **Characterization of iucA and iucC genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*.** *J Bacteriol* 1986, **167**:350-355.
14. Martinez A, Kolter R: **Protection of DNA during oxidative stress by the non specific DNA-binding protein Dps.** *J Bacteriol* 1997, **179**:5188-5194.
15. Zhao G, Ceci P, Ilari A, Giangiacomo L, Laue TM, Chiancone E, Chasteen ND: **Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells. A ferritin-like DNA-binding protein of *Escherichia coli*.** *J Biol Chem* 2002, **277**:27689-27696.
16. Zaika EI, Perlow RA, Matz E, Broyde S, Gilboa R, Grollman AP, Zharkov DO: **Substrate discrimination by formamidopyrimidine-DNA glycosylase: a mutational analysis.** *J Biol Chem* 2004, **279**:4849-4861.
17. Gaudu P, Weiss B: **Flavodoxin mutants of *Escherichia coli* K-12.** *J Bacteriol* 2000, **182**:1788-1793.
18. Outten FW, Wood MJ, Munoz FM, Storz G: **The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in *Escherichia coli*.** *J Biol Chem* 2003, **278**:45713-45719.
19. Ton-That H, Schneewind O: **Assembly of pili on the surface of *Corynebacterium diphtheriae*.** *Mol Microbiol* 2003, **50**:1429-1438.
20. Suh JW, Boylan SA, Oh SH, Price CW: **Genetic and transcriptional organization of the *Bacillus subtilis* spc-alpha region.** *Gene* 1996, **169**:17-23.
21. Jannick DB, Henrik N, Gunnar VH, Søren B: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
22. NCBI FTP site [[http://ftp.ncbi.nih.gov/genomes/Bacteria/Corynebacterium\\_diphtheriae](http://ftp.ncbi.nih.gov/genomes/Bacteria/Corynebacterium_diphtheriae)]
23. Shannon CE: **A mathematical theory of communication.** *Bell System Technical Journal* 1948:379-423. 623-656
24. Yellaboina S, Seshadri J, Kumar MS, Ranjan A: **PredictRegulon: A webserver for the prediction of the regulatory protein binding sites and operons in prokaryote genomes.** *Nucleic Acids Res* 2004, **32**:W318-W320.
25. Tao X, Murphy JR: **Binding of the metalloregulatory protein DtxR to the diphtheria tox operator requires a divalent heavy metal ion and protects the palindromic sequence from DNase I digestion.** *J Biol Chem* 1992, **267**:21761-21764.
26. Schmitt MP, Holmes RK: **Cloning, sequence, and footprint analysis of two promoter/operators from *Corynebacterium diphtheriae* that are regulated by the diphtheria toxin repressor (DtxR) and iron.** *J Bacteriol* 1994, **176**:1141-1149.
27. Schmitt MP: **Transcription of the *Corynebacterium diphtheriae* hmuO gene is regulated by iron and heme.** *Infect Immun* 1997, **65**:4634-4641.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

