In *silico* genome-wide screening for TnrA-regulated genes of *Bacillus clausii*

Abbas Farazmand[1,2], Bagher Yakhchali[1*], Parvin Shariati[1], Zarrin Minuchehr[1], Hamideh Ofoghi[2]

[1]Department of Industrial and Environmental Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), P.O. Box 14965-161, Tehran, I.R. Iran [2]Department of Biotechnology, Iranian Research Organization for Science and Technology (IROST), P.O. Box 15815-3538, Tehran, I.R. Iran

Abstract
Bacillus clausii TnrA transcription factor is required for global nitrogen regulation. In order to obtain an overview of gene regulation by TnrA in *B. clausii* KSM-K16, the entire genome of *B. clausii* was screened for the consensus sequence, 5'-TGTNAN7TNACA-3' known as the TnrA box, and 13 transcription units were found containing a putative TnrA box. The TnrA targets identified in this study were *tnrA*, *glnA*, *nrgA*, *nasFDEB*, *puc* genes, *licT*, the two operons of the oligopeptide ABC transporter, *lytR*, transcriptional regulator of the Lrp/AsnC family, sodium-dependent transporter of SNF family, *hyu* genes and a biochemically uncharacterized protein.
Keywords**:** Gene regulation; TnrA; *Bacillus clausii*; nitrogen metabolism

The MerR family is a group of transcriptional activators, which regulates gene expression and controls transcription in response to diverse physiological signals, such as nitrogen availability. The *B. clausii* TnrA transcription factor, a member of the MerR family has been found to contain 100 amino acids. Also, the *Bacillus subtilis* TnrA, as one of the best understood members of the MerR family, composed of 110 amino acids. When nitrogen sources are in excess, the glutamine synthetase (GS), a key enzyme in nitrogen metabolism, becomes subject to feedback inhibition by glutamine and adenosine monophosphate (AMP). The feedback-inhibited GS forms a complex with TnrA via its C-terminal domain, thereby preventing TnrA from interacting with specific operators and regulating gene expression (Wray *et al*., 2000).

In *B. subtilis*, *Bacillus licheniformis*, *Geobacillus Kaustophilus* and *Oceanobacillus iheyensis*, the two transcription factors TnrA and GlnR control many genes for utilization of glutamine and other nitrogen-containing compounds. *Bacillus halodurans* lacks GlnR but possesses a single TnrA regulator of nitrogen assimilation and two paralogs of *glnA*, both with TnrA-binding sites.

Bacillus clausii is known to produce a commercially important extracellular alkaline serine protease (AprE). This enzyme, produced during the stationery phase of growth, is temporally controlled and subjected to regulation by a large number of positive and negative regulators that inevitably allow for the effective use of the enzyme in the cell's surrounding environment (Ghaemi Oskouiea *et al*., 2008; Saeki *et al*., 2007). Elucidating of the molecular mechanisms of the metabolism and gene regulatory networks could thus be used to design metabolic engineering strategies for maximizing alkaline serine protease production in *B. clausii*.

The aim of this work was to distinguish and analyze the potentially TnrA regulons of *B. clausii*, involved in metabolism of nitrogen, and thus have an insight into the nature of the regulation of this metabolic system and reveal the similarities and differences of the associated transcriptional regulatory networks,

*Correspondence to: **Bagher Yakhchali, Ph.D.***
Tel: +98 21 44580353; Fax: +98 21 44580399
E-mail: bahar@nigeb.ac.ir

present in *B. clausii*, *B. subtilis* and other bacilli. The complete genome sequence of *B. clausii KSM-K16*, was obtained from GenBank (accession number, AP006627.1) (http://www.ncbi.nlm.nih.gov) (Benson *et al*., 2005). The nucleotide sequences of the promoter and the coding region of the *tnrA* belonging to *B. clausii EHY L2* deposited previously in GenBank were also used in this study (accession number, HM488959).

A genome-wide analysis for TnrA-regulated genes of *B. clausii* associated with a TnrA box was performed. It was found that there were 13 such transcription units containing a putative TnrA box, including the global regulator of nitrogen regulation (*tnrA*), the glutamine synthetase (*glnA*) gene, two oligopeptide ABC transporter operons, the β-glucoside *bgl* operon antiterminator (*licT*), the transcriptional attenuator of autolytic activity (*lytR*), the transcriptional regulator of the Lrp/AsnC family, the sodium-dependent transporter (SNF family), the assimilatory nitrite reductase operon (*nasFDEB*), the genes for purine catabolism (*puc*), the operon for hydantoin utilization (*hyu*), the ammonium transporter (*nrgA*), and a biochemically uncharacterized protein. Investigation of the genes that have been reported as the *B. subtilis* TnrA regulon showed that the glutamine ABC transporter (*glnQ*, *glnH* and *glnMP*), glutamate synthase (*gltAB*), sodium: alanine symporter (*alsT*), gamma-glutamyltranspeptidase (*ywrD*) and L-asparaginase (*ansZ*) genes of *B. clausii* do not have a reliable TnrA site in their regulatory regions. Furthermore, bioinformatic analysis showed that the respective orthologs of urease (*ureABC*), the GlnR transcription factor (*glnR*), proline transporter (*yodF*) and gamma-aminobutyrate (GabA) permease (*gabP*) are not present in *B. clausii*.

Alignment of TnrA sites in the promoters was used to compare the *tnrA* promoter sequences in *B. clausii KSM-K16* and *EHY L2*. The *tnrA* genes of the two *B. clausii* strains, *KSM-K16* and *EHY L2*, contain two TnrA sites in their promoter region, with a 25 bp interspace (Fig. 1).

The *B. clausii* TnrA protein is smaller than most MerR family members. It contains 100 amino acids and two domains (Fig. 2). A conserved N-terminal DNA binding domain is located between residues 5 and 76. Based on the crystal structure of the multidrug-binding transcription regulator BmrR of *B. subtilis* and secondary structure prediction by PSIPRED, this domain was shown to contain a β-strand, a helix-turn-helix motif formed by helices 1 and 2, and a second wing formed by helices 3 and 4. A conserved 15-amino-acid C-terminal region was also found, which like other TnrA orthologs, functions as a signal transduction domain. In fact a similar domain in the TnrA

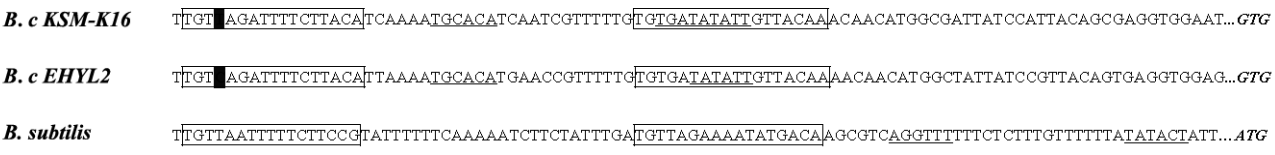

Figure 1. Comparison of the nucleotide sequence of the *tnrA* promoter region in *B. clausii KSM-K16, B. clausii EHY L2* and *B. subtilis*. The proposed -35 and -10 promoter elements are underlined. The TnrA sites are shown in boxes. The translational start site is indicated by bold italic letters.

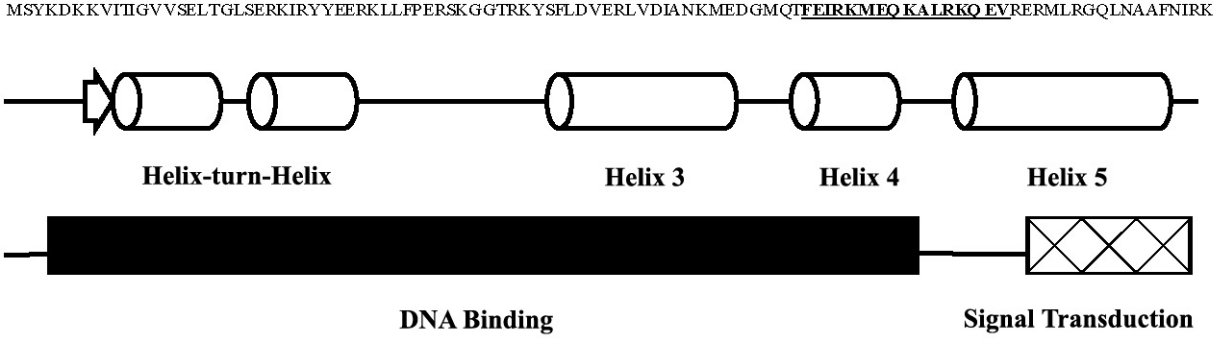

Figure 2. A graphical view of the secondary structure and domains of the *B. clausii* TnrA protein, where α-helices and β-strands are represented by cylinders and arrows. The coiled-coil domain is indicated by underlined bold letters and interface amino acids are underlined.

of *B. subtilis* has also been reported to be involved in signal transduction (Newberry *et al*., 2008; Wray *et al*., 2001; Jones, 1999).

Using the COILS program (Lupas *et al*., 1991), the C-terminal region of TnrA was predicted to contain coiled-coil structures, arising from the association of amino acid residues (68 to 83) with other similar C-terminal regions of the TnrA. Furthermore, Ile-70, Met-73, Ala-77 and Lys-80 were recognized as interface residues on α-helices 4 and 5 of the TnrA protein (Fig. 2).

Comparison of the nucleotide and deduced amino acid sequences showed that there was strong homology between the *tnrA* genes of *B. clausii KSM-K16* and *B. clausii EHY L2* (97% and 100% identities at the nucleotide and amino acid levels, respectively). BLAST analysis was used to detect paralogous genes (Altschul *et al*., 1997). *B. clausii* contains two paralogs of the gene encoding the GS, GlnA1 (*ABC3940*) and GlnA2 (*ABC2179*). The *glnA1* gene, contains a TnrA site, 87 bp upstream of the translation start site. This TnrA site is located downstream of the -10 region of the promoter. Comparison of the deduced amino acid sequences of the *B. clausii* GS with other bacteria revealed that the *B. halodurans* GS has a high degree of similarity (90%) with that of *B. clausii*. However, the *glnA2* gene does not contain the TnrA site at its regulatory region. *B. halodurans* like *B. clausii* only has the *tnrA* and a monocistronic *glnA* operon, but is devoid of *glnR*. *B. halodurans* possesses two paralogs of *gln*A, both with TnrA-binding sites (Fig. 3) (Doroshchuk *et al*., 2006).

The *nrgA* gene, whose product is an ammonium transporter, is present in the genome of *B. clausii* and has a conserved TnrA site in its regulatory region (Fig. 4). Similarly, the *nrgA (amtB)* of many other bacteria

glnA1 B. clausii CATCCAATTGGATGGTTTTGCCATTATTATCTGAATATTAGTGTGTCAGATAATCTGACA...***ATG***

glnA (BH2360) B. halodurans TTTGTTAGATTAATTACTATGTCAGATTATCTTACATATATTTGACAGATCGATACCCTT...***ATG***

glnA (BH3867) B. halodurans ATTCACTCTTTTTTTATTCTATTGTAGATTACC…GTTTGTCAGATAATCTAACAAACCCA…***ATG***

glnRA B. subtilis ATTTGATGTTAAGAATCCTTACATCGTATTGACACATAATATAACATCACCTATAATGA...***ATG***

Figure 3. Comparison of the nucleotide sequences of the GS promoter regions in *B. clausii KSM-K16, B. halodurans* and *B. subtilis*. The proposed -35 and -10 promoter elements are underlined. The TnrA sites are shown in boxes. The translational start site is indicated by bold italic letters.

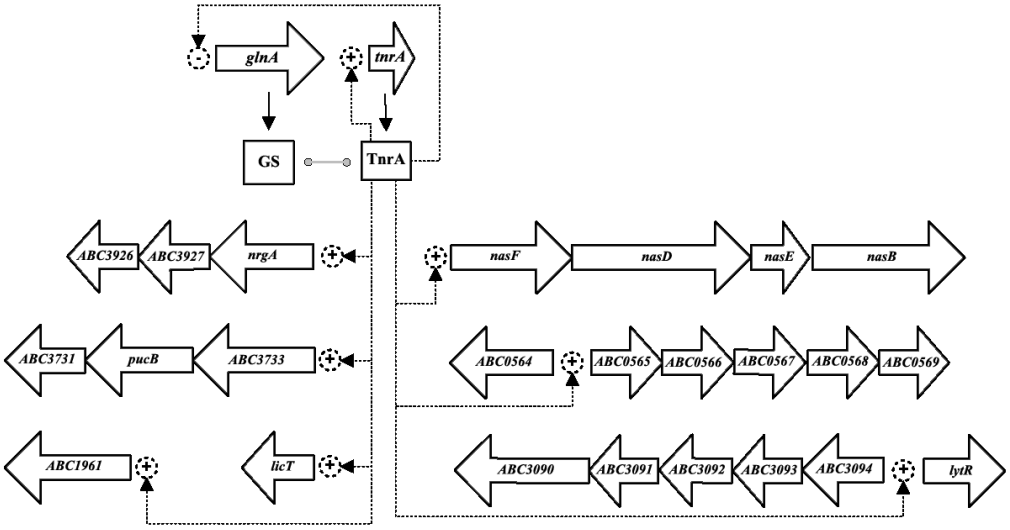

Figure 4. Regulatory network of nitrogen metabolism in *B. clausii*. Arrows indicate the relative length and organization of genes on the chromosome. Dotted black arrows illustrate the regulatory interactions of the respective regulator of nitrogen metabolism (orange) including its predicted function, in activating (plus) or repressing (minus) the target gene(s). Gray lines with circles indicate post-translational interactions.

such as *B. subtilis*, *B. cereus*, *B. licheniformis*, *Oceanobacillus. iheyensis* and *G. Kaustophilus* contain one binding site for TnrA (Doroshchuk *et al*., 2006). Bioinformatics analysis showed another set of genes associated with the TnrA box, known as the *nas* genes, which are involved in nitrate and nitrite assimilation, and are organized as two transcription units, *nasFDEB* and *nasCA*. The *nas* operon with the conserved TnrA site, occurs in the genomes of *B. clausii*, *B. subtilis*, *B. licheniformis*, *B. cereus* and *B. halodurans* (Doroshchuk *et al*., 2006). The *puc* genes involved in purine catabolism containing a TnrA site. In *B. subtilis*, the *puc* operon contains two binding sites for TnrA (Fig. 4).

This study also identified the hydantoin utilization operon (*hyu*), as a new member of the potentially TnrA regulons, that contins a reliable TnrA site in the regulatory region of its promoter. The hydantoin operon, which among the *Bacillus* species appears to be unique to *B. clausii*, consists of genes encoding aspartate/glutamate hydantoin racemase; it is a hypothetical conserved protein, which is thought to be involved in hydantoin and pyrimidine utilization under nitrogen-limited conditions. This research also identified two other operons with TnrA sites, which encode the oligopeptide ABC transporter in *B. clausii*. It has already been shown that orthologs of these genes (*opp ABCDF*) found in *B. subtilis* and *B. licheniformis* are under the control of TnrA (Yoshida *et al*., 2003; Doroshchuk *et al*., 2006). The *licT* gene, whose product is a β-glucoside *bgl* operon transcriptional antiterminator protein of the BglG family, is also present in the genome of *B. clausii* and has a conserved TnrA site in its regulatory region. However, the ortholog of this gene is not regulated by TnrA in *B. subtilis* (Doroshchuk *et al*., 2006; Yoshida *et al*., 2003).

Hence, *B. clausii* TnrA responds by regulating not only genes for nitrogen utilization but also those involved in amino acid (*ABC3095*) and carbon (*licT*) metabolism (Lindner *et al*., 2002). The TnrA-regulated gene *lytR*, a transcriptional attenuator of autolytic activity, was found only in *B. clausii*. By contrast the *B. subtilis* LytR, described as a transcriptional attenuator of itself and the *lytABC* operon, is not regulated by TnrA (Doroshchuk *et al*., 2006; Yoshida *et al*., 2003). Hence, it could be that in *B. clausii*, the TnrA regulator is involved indirectly in the inhibition of autolysis under nitrogen-limited conditions. The *B. clausii ABC3095* gene of the Lrp/AsnC family of transcriptional regulators also has a potential TnrA site. The ortholog of this gene has been found not to be regulated by TnrA in other bacilli (Doroshchuk *et al*., 2006; Yoshida *et al*., 2003). Members of the Lrp family are small DNA-binding proteins with molecular masses of approximately 15 kDa, which are important regulatory systems of amino acid metabolism and related processes (Brinkman *et al*., 2003). The TnrA-regulated sodium-dependent transporter gene (SNF family) was found only in *B. clausii*.The ortholog of this gene is not regulated by TnrA in *B. halodurans, Bacillus Pseudofirmus* and *O. iheyensis* (Doroshchuk *et al*., 2006).

It is worth nothing that most of the *B. clausii* TnrA target genes are very likely involved in the utilization of nitrogen sources, such as ammonium (*nrgA*), nitrite, nitrate (*nasFDEB* and *nasA*) and oligopeptides, thus emphasizing the physiological role of TnrA regulation. Furthermore, the *B. clausii* TnrA may also respond by regulating genes involved in amino acid (*ABC3095*) and carbon (*licT*) metabolism (Fig. 4).

Comparison of potentially TnrA regulons in *B. clausii* with those which have been previously studied in *B. subtilis*, *B. licheniformis*, *B. halodurans*, *O. iheyensis* and *G. kaustophilus* reveals that in all these starins the TnrA sites of the *glnA*, *tnrA* and *nrgA* promoters are conserved (Doroshchuk *et al*., 2006; Yoshida *et al*., 2003). Comparison of TnrA regulons, which contain the TnrA sites of *B. clausii* and *B. subtilis* revealed that the general transcription factor, TnrA (*tnrA*), the glutamine synthetase gene (*glnA*), oligopeptide ABC transporter operons, the assimilatory nitrate and nitrite reductase operon (*nas*), the genes of purine catabolism (*puc*) and ammonium transport (*nrg*A), are conserved in both *B. clausii* and *B. subtilis* bacteria (Doroshchuk *et al*., 2006; Yoshida *et al*., 2003)

It is important to note that the alkaline protease of B. *clausii*, similar to that of *B. subtilis*, is also expressed in abundance under nitrogen-limited conditions (Abe *et al*., 2009). So it may be possible that the *aprE* (coding for alkaline protease) of *B. clausii* is also under nitrogen regulation through the GlnA-TnrA pathway. On this basis, a nitrogen-replete status in the cell may be a situation where TnrA is captured by complex formation with feedback-inhibited GlnA. Therefore, we propose to construct a potent *B. clausii* for the production of alkaline serine proteases by the disruption of *glnA* or truncation of the C-terminal region of *tnrA*, which will lead to the release of TnrA from the feedback-inhibited GlnA, thus mimicking a nitrogen limited situation.

References

Abe S, Yasumura A, Tanaka T (2009). Regulation of Bacillus subtilis aprE expression by glnA through inhibition of scoC and sigma(D)-dependent degR expression. *J Bacteriol*. 191: 3050-3058.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005). GenBank. *Nucleic Acids Res*. 33: D34-38.

Brinkman AB, Ettema TJ, de Vos WM, van der Oost J (2003). The Lrp family of transcriptional regulators. *Mol Microbiol*. 48: 287-294.

Doroshchuk NA, Gel'fand MS, Rodionov DA (2006). [Regulation of nitrogen metabolism in gram-positive bacteria]. *Mol Biol* (*Mosk*). 40: 919-926.

Ghaemi Oskouiea F, Tabandehb F, Yakhchali B, Eftekhara F (2008). Response surface optimization of medium composition for alkaline protease production by Bacillus clausii. *Biochem Eng J*. 39: 37-42.

Jones DT (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*. 292: 195-202.

Lindner C, Hecker M, Le Coq D, Deutscher J (2002). Bacillus subtilis mutant LicT antiterminators exhibiting enzyme I- and HPr-independent antitermination affect catabolite repression of the bglPH operon. *J Bacteriol*. 184: 4819-4828.

Lupas A, Van Dyke M, Stock J (1991). Predicting coiled coils from protein sequences. *Science* 252: 1162-1164.

Newberry KJ, Huffman JL, Miller MC, Vazquez-Laslop N, Neyfakh AA, Brennan RG (2008). Structures of BmrR-drug complexes reveal a rigid multidrug binding pocket and transcription activation through tyrosine expulsion. *J Biol Chem*. 283: 26795-26804.

Saeki K, Ozaki K, Kobayashi T, Ito S (2007). Detergent alkaline proteases: enzymatic properties, genes, and crystal structures. *J Biosci Bioeng*. 103: 501-508.

Wray LV, Jr, J. M. Zalieckas, S. H. Fisher, (2000). Purification and *in vitro* activities of the Bacillus subtilis TnrA transcription factor. *J Mol Biol*. 300: 29-40.

Wray LVJr, Zalieckas JM,Fisher SH (2001). *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell* 107: 427-435.

Yoshida K, Yamaguchi H, Kinehara M, Ohki YH, Nakaura Y, Fujita Y (2003). Identification of additional TnrA-regulated genes of Bacillus subtilis associated with a TnrA box. *Mol Microbiol*. 49: 157-165.