The *Bacillus subtilis* NatK–NatR two-component system regulates expression of the *natAB* operon encoding an ABC transporter for sodium ion extrusion

Mitsuo Ogura, Kensuke Tsukahara, Kentaro Hayashi and Teruo Tanaka

Correspondence
Mitsuo Ogura
oguram@scc.u-tokai.ac.jp

Institute of Oceanic Research and Development, Tokai University, 3-20-1 Orido-Shimizu, Shizuoka 424-8610, Japan

Received 23 October 2006
Revised 13 December 2006
Accepted 14 December 2006

A previous microarray analysis suggested that multicopy *yccH*, encoding a function-unknown response regulator, enhances expression of *natAB*, which encodes a two-gene ATP-binding cassette transporter involved in the extrusion of sodium ions. The two-component regulatory system YccG–YccH was therefore renamed NatK–NatR. Here, this observation was confirmed by a *lacZ* fusion analysis using a strain carrying *natA*–*lacZ*. Further, in both *natK* and *natR* mutants, *natA*–*lacZ* expression was completely abolished, indicating that the NatK–NatR system positively regulates the expression of *natAB*. In a gel retardation analysis, NatR bound to the *natA* promoter region. Using purified His-tagged NatR, DNase I footprinting analysis of the *natA* promoter region suggested that a direct repeat of [TTCA(G)CGACA], separated by a 12 bp space, would be recognized by NatR. Deleted and mutagenized promoter regions of *natA* were analysed using a *lacZ* fusion, and it was confirmed that the direct repeat is critical for *natA* activation by NatR.

INTRODUCTION

Bacteria employ elaborate signal transduction mechanisms to respond to environmental fluctuations. The two-component systems are a large family of proteins, and they are typically composed of a sensor kinase and its cognate response regulator (Stock *et al.*, 2000). The catalytic part of the kinase autophosphorylates a histidine residue in response to an environmental signal. The phosphoryl group is then transferred to a conserved aspartate residue on the cognate response regulator, which acts as a transcription factor in most cases.

In the *Bacillus subtilis* genome, 35 response regulators, and 30 operonic kinase-regulator pairs, have been identified (Kunst *et al.*, 1997). Following the completion of genomic sequencing, several newly identified operons that encode two-component systems have been characterized, e.g. the BceRS, CitST, CssSR, DesKR, GlnKL, LiaRS, MalKR, YdbFG, YdfHI, YxdJK and YycFG systems (Aguilar *et al.*, 2001; Asai *et al.*, 2000; Doan *et al.*, 2003; Fukuchi *et al.*, 2000; Hyrylainen *et al.*, 2001; Joseph *et al.*, 2004; Mascher *et al.*, 2004; Ohki *et al.*, 2003; Satomura *et al.*, 2005; Serizawa & Sekiguchi, 2005; Tanaka *et al.*, 2003; Yamamoto *et al.*, 2000). In addition, close relationships have been identified between

some operons encoding two-component systems and genes for ATP-binding cassette (ABC) transporter homologues, e.g. YccGH, YclJK, YfiJK, YvcPQ, YvfTU and YxdJK (Joseph *et al.*, 2002).

Comprehensive DNA microarray analyses of response regulators, using gene amplification in a sensor-deficient strain, have contributed to the identification of the target genes for these systems (Kobayashi *et al.*, 2001; Ogura *et al.*, 2001). In those studies, the NatR (YccH) response regulator was identified as an activator for the *natAB* operon encoding a two-gene ABC transporter involved in Na⁺ extrusion (Cheng *et al.*, 1997; Kobayashi *et al.*, 2001). This ABC transporter is a member of a small subfamily named ABC-2, which includes the two-component daunomycin- and doxorubicin-efflux system in *Streptomyces peucetius* Drr (Guilfoile & Hutchinson, 1991; Reizer *et al.*, 1992). It has been reported that NatAB has a physiological role in Na⁺ resistance at alkaline pH, since a high concentration of cellular Na⁺ is toxic to the cell, and plays a role in ethanol resistance at neutral pH (Cheng *et al.*, 1997; Padan & Schuldiner, 1994).

In this paper, we investigated whether or not NatK–NatR (formerly YccG–YccH) positively regulates the *natAB* operon. Gel retardation and footprint assays revealed direct binding of NatR to the promoter of *natAB*. Furthermore, we investigated the role of a direct repeat sequence, located within the *natA* promoter, in NatR recognition.

Abbreviation: ABC, ATP-binding cassette.

A supplementary figure showing the structure of plasmid pDG-N17 is available with the online version of this paper.

METHODS

Bacterial strains and culture media. All the *B. subtilis* strains used for this study are listed in Table 1. One-step competence medium (MC) was used for the determination of β -galactosidase activity (Ogura & Tanaka, 1996; Ogura *et al.*, 1997). TKM medium was used for growth tests in the presence of NaCl or ethanol (Cheng *et al.*, 1997). *Escherichia coli* JM103 cells for DNA manipulation were grown in Luria–Bertani (LB) medium. Concentrations of antibiotics used have been described previously (Ogura & Tanaka, 1996).

Plasmid construction. Synthetic oligonucleotides were commercially prepared by the Tsukuba Oligo Service (Ibaraki, Japan). The plasmids and oligonucleotides used in this study are listed in Tables 1 and 2, respectively. The vector pMutInIII (Vagner *et al.*,

1998), and a PCR product spanning the 310 bp upstream region from the first nucleotide of the 24th codon of *natA*, prepared using the *natA*-FH and *natA*-RB primers, were digested with *Hind*III and *Bgl*II, and ligated to form pMutIn-*natA*. To construct pDG-N17 (see supplementary Fig. S1, available with the online version of this paper), pDG-His-degU (Shimane & Ogura, 2004) was digested with *Bam*HI, and the large fragment was ligated with a double-stranded oligonucleotide, made by annealing single-stranded NotI-I and NotI-II oligonucleotides, after which the reaction mixture was used to transform *E. coli*. The presence of the 6 $\times$ His tag, followed by the *Bam*HI–NotI sites, on pDG-N17 was confirmed by sequencing using a 377 DNA Sequencer (Perkin-Elmer), and a Dye Terminator Cycle Sequencing Kit (Applied Biosystems). In order to construct pDG-N17-*natR*, which has a histidine tag at the N-terminus of *NatR*, pDG-N17 and a PCR product, prepared using the *NatR*-*Bam* and *NatR*-NotI primers, were digested with *Bam*HI and NotI, and

Table 1. Strains and plasmids used for this study

Strain or plasmid	Genotype or description	Reference or source
<i>B. subtilis</i> strains		
168	<i>trpC2</i>	Laboratory stock
OAM298	<i>trpC2 natA::pMutIn-natA (natA⁺ natA-lacZ Em^r)</i>	This study
OAM269	<i>trpC2 natA::pMutIn-natA (natA⁺ natA-lacZ Em^r) natK::Pm^r</i>	This study
TF120	<i>trpC2 natR::Cm^r</i>	Hayashi <i>et al.</i> (2006)
OAM270	<i>trpC2 natA::pMutIn-natA (natA⁺ natA-lacZ Em^r) natR::Cm^r</i>	This study
NATAd	<i>trpC2 natA::pMutIn2-natA (natA natA-lacZ Em^r)</i>	BSORF (http://bacillus.genome.jp/)
NATRd (YCCHd)	<i>trpC2 natR::pMutIn2-yccH (natR natR-lacZ Em^r)</i>	BSORF
YCCKd	<i>trpC2 yccK::pMutIn2-yccK (yccK yccK-lacZ Em^r)</i>	BSORF
OAM271	<i>trpC2 amyE::natA1-lacZ (Cm^r)</i>	This study
OAM272	<i>trpC2 amyE::natA2-lacZ (Cm^r)</i>	This study
OAM273	<i>trpC2 amyE::natA3-lacZ (Cm^r)</i>	This study
OAM274	<i>trpC2 amyE::natA1M-lacZ (Cm^r)</i>	This study
OAM275	<i>trpC2 amyE::natA2M-lacZ (Cm^r)</i>	This study
OAM359	<i>trpC2 amyE::yccK-lacZ (Cm^r)</i>	This study
Plasmids		
pDG-His-degU	Ampicillin, erythromycin and spectinomycin resistance, T5 promoter, <i>amyE::6 $\times$ His-degU</i>	Shimane & Ogura (2004)
pDG-N17	Ampicillin, erythromycin and spectinomycin resistance, T5 promoter, <i>amyE::6 $\times$ His</i>	This study
pDG-N17- <i>natR</i>	pDG-N17 carrying <i>amyE::6 $\times$ His-natR</i>	This study
pRep4	Kanamycin resistance, <i>lacI</i>	Qiagen
pDG148	Ampicillin and kanamycin resistance, <i>lacI</i> , <i>Pspac</i>	Stragier <i>et al.</i> (1988)
pDG148- <i>natR</i>	pDG148 carrying <i>Pspac-natR</i> , formerly pDG148- <i>yccH</i>	Kobayashi <i>et al.</i> (2001)
pPhl2	Phleomycin resistance	Ogura <i>et al.</i> (1997)
pPhl2- <i>natK</i>	pPhl2 carrying a part of <i>natK</i>	This study
pMutInIII	Ampicillin and erythromycin resistance, <i>lacZ</i>	Vagner <i>et al.</i> (1998)
pMutIn-NatA	pMutInIII carrying a part of <i>natA (natA-lacZ)</i>	This study
pIS284	Ampicillin and chloramphenicol resistance, <i>lacZ</i>	I. Smith*
pIS-yccK	pIS284 carrying a <i>yccK</i> upstream region (–197 to –12 from the initiation codon)	This study
pIS-natA1	pIS284 carrying a <i>natA</i> upstream region (–116 to +27)	This study
pIS-natA2	pIS284 carrying a <i>natA</i> upstream region (–79 to +27)	This study
pIS-natA3	pIS284 carrying a <i>natA</i> upstream region (–57 to +27)	This study
pIS-natA1M	pIS-natA1 carrying a mutagenized <i>natA</i> upstream region	This study
pIS-natA2M	pIS-natA2 carrying a mutagenized <i>natA</i> upstream region	This study

*TB Center, The Public Health Research Institute, Newark, NJ, USA.

Table 2. Oligonucleotides used for this study

Oligonucleotide	Sequence (5'→3')
natA-FH	CGCAAGCTTGGAAAAGACAATTTTCGACG
natA-RB	AGCGGATCCATCTCGCACCGCTTTGACT
natK-E	ATGGAATTCTATCAGGGAGCTGCTGCTT
natK-Sa	ATGGTCGACTCCGATACGGAGCCCTTCAC
natR-F1	TCAGCGCTCTGGCAATGGCG
natR-F2	CGTCGTGACTGGGAAAACGATCACTAAGCAGCAGCTTG
natR-R1	GTTATCCGCTCACAATTCTACCATTATCATTCCCCTAG
natR-R2	CGAAACACTCACATTCGAGG
pUC-F	GTTTTCACAGTCACGACG
pUC-R	GAATTGTGAGCGGATAAC
NotI-I	GATCTGCGGCCGCG
NotI-II	GATCCGCGGCCGCA
NatR-Bam	CTAGGATCCATGGTAAAAGTCGGACTTG
NatR-NotI	ATGACCTAGCGGCCGCTCAAAAATAGTAGTTTGA
NatA-bio	Biotin-CTCGCACCGCTTTGACTACT
NatA-pISB1	CATGGATCCTATCATTCCCCTAGTTACAT
NatA-pISB2	CATGGATCCCATTTCACGACATAAA
NatA-pISB3	CATGGATCCGAACCTCTTCGCGACAACCTT
NatA-pISH	CATAAGCTTGTACAATCTCCCTTATCTT
YccK-pISB	ATTGGATCCATCTCTTTTACTAGCC
YccK-pISH	GATAAGCTTCTTTGAAACCAATTTTGATG
NatA-M1	AACTCTTCGATCAAACTTTATC
NatA-M2	GATAAAGTTTGATCGAAGAGTT
NatA-pISB2-M	CATGGATCCCATTTCATCAATAAAAAGAAC
NatA-up-F	Biotin-TATCATTCCCCTAGTTACAT
NatA-R-trpH	ATGAAGCTTGGCTGCAATCGGTCAGTGT
NatA-RB	Biotin-GTCACAATCTCCCTTATCTT
NatA-up-R	GTCACAATCTCCCTTATCTT

ligated. Sequencing of the cloned fragment in pDG-N17-natR confirmed the absence of PCR amplification errors. The plasmid pPhl2 (Ogura *et al.*, 1997), and a PCR product prepared using the natK-E and natK-Sa primers, were digested with *EcoRI* and *Sall*, and ligated to form pPhl2-natK. PCR products were prepared using the primer pairs NatA-pISB1 and NatA-pISH, NatA-pISB2 and NatA-pISH, NatA-pISB2M and NatA-pISH, NatA-pISB3 and NatA-pISH, and YccK-pISB and YccK-pISH. The fragments were digested with *BamHI* and *HindIII*, and ligated into the similarly digested pIS284, in order to construct pIS-natA1, pIS-natA2, pIS-natA2M, pIS-natA3 and pIS-yccK, respectively. The plasmid pIS-natA1M was constructed by site-directed mutagenesis of pIS-natA1 using the oligonucleotide-based PCR method described previously (NatA-pISB1, NatA-M1, NatA-M2 and NatA-pISH; Ogura & Tanaka, 1996). The sequences cloned into pIS284 were confirmed by sequencing.

Construction of strains. The *natR*-disruption mutant was constructed by homologous recombination with PCR fragments containing the upstream and downstream regions of *natR*, flanking a chloramphenicol-resistance gene. PCR amplification of the upstream and downstream regions of *natR* was conducted using the primers natR-F1 and natR-R1, and natR-F2 and natR-R2, respectively. PCR amplification of the chloramphenicol-resistance gene from pCBB31 was conducted using the primers pUC-F and pUC-R. The three PCR fragments were used as templates for a PCR reaction with the

primers natR-F1 and natR-R2, and the amplified DNA fragment was used to transform *B. subtilis* 168. Disruption of *natR* in the resulting chloramphenicol-resistant strain was confirmed by PCR.

Primer-extension analysis. The isolation of total RNA, and primer-extension analysis, were performed as described previously (Yoshida *et al.*, 2001; Ogura & Tanaka, 1996; respectively). Sequencing was performed using a Cycle sequencing kit (Toyobo), and detection of biotinylated DNA was performed using a chemiluminescence DNA detection kit (Toyobo).

Protein purification. The induction and purification procedures for preparation of protein expressed in *E. coli* have been described previously (Ogura *et al.*, 2003; Shimane & Ogura, 2004). His-tagged NatR was expressed as a soluble protein in *E. coli* M15 cells carrying pRep4, and step-wise elution with imidazole from a Ni-affinity column was used for the purification. Following SDS-PAGE analysis of fractions, the purified protein was dialysed against TEDG buffer (Mukai *et al.*, 1990).

Gel shift and DNase I footprinting assays. We used biotinylated DNA probes for the gel shift assay (Ogura *et al.*, 2003); the DNase I footprinting assay was carried out as described previously (Hayashi *et al.*, 2005).

β -Galactosidase assays. Samples were taken at hourly intervals, and the level of β -galactosidase activity was determined as described previously (Ogura & Tanaka, 1996).

RESULTS

Confirmation of the DNA microarray analysis results

Our DNA microarray analysis of the NatR response regulator, using a multicopy plasmid expressing *natR* in the *natK* mutant, showed three candidate gene targets for NatR located in tandem (*natA*, *natB* and *yccK*; Kobayashi *et al.*, 2001). Northern analysis has indicated that *natA* forms an operon with *natB*, but not with *yccK* located immediately downstream of *natAB* (Cheng *et al.*, 1997). There is no ORF immediate downstream of *yccK* on the same DNA strand, and this strongly suggests that *yccK* is monocistronic. This notion was reinforced by two observations. (1) The presence of a putative rho-independent terminator downstream of *natB* (Fig. 1a; Cheng *et al.*, 1997). (2) The promoter activity of the *yccK* upstream region at an ectopic *amyE* locus; *yccK* shows a temporal expression pattern that is different from that of *natA-lacZ* (data not shown). YccK is predicted to encode a soluble protein with similarity to members of the aldo/keto reductase family (BSORF; <http://bacillus.genome.jp/>).

A multicopy plasmid [pDG148-*natR* (formerly pDG148-*yccH*), Table 1] expressing *natR* was introduced into the *natK* strain carrying a transcriptional *natA-lacZ* or *yccK-lacZ* fusion. Each of the fusions was generated by Campbell-type recombination of derivatives of the plasmid pMutIn (Table 1). Next, β -galactosidase activities were examined in the presence and absence of IPTG. The induction of *natR* by the addition of IPTG resulted in 35- and 5-fold increases in the levels of expression of *natA-lacZ* (Fig. 1b) and *yccK-lacZ* (data not shown), respectively. Thus, the observations obtained with the *lacZ* fusions confirmed the results of DNA

microarray analysis. However, the promoter of *yccK* was not a target of NatR, because induction of *natR* on pDG148-*natR* did not enhance the expression of *yccK-lacZ* at the *amyE* locus (data not shown). Thus, the apparent induction of *yccK* expression detected by microarray analysis was most probably caused by artificial read-through of the *yccK* gene resulting from the increased activity of the *natAB* promoter under the condition of *natR* overexpression. In addition, in the presence of an intact *natK* gene, induction of *natR* on pDG148-*natR* resulted in a ninefold increase in the expression of *natA-lacZ* (Fig. 1b).

natAB is located adjacent to *natRK*, but on the complementary strand (Fig. 1a). The chromosomal structure of the region that contains the transcription control elements for *natRK* and *natAB* raised the possibility that *natR* and *natK* genes might be part of an autoregulatory loop under the control of NatR. To examine this, we introduced pDG148-*natR* into cells carrying *natR-lacZ*, and measured β -galactosidase activities in the presence and absence of IPTG. We did not observe any change in the expression of *natR-lacZ* when the *natR* gene was induced by IPTG (data not shown). This indicates that *natR* expression itself is not regulated by NatR.

Expression of *natA* was abolished in the mutants *natK* and *natR*

To examine the notion that *natR* is a positive regulator of *natAB*, we measured the β -galactosidase activities of the fusions *natA-lacZ* in wild-type, *natK* and *natR* strains. The expression level of *natA-lacZ* in the wild-type cells was significantly high compared with that in cells carrying the same fusion and pDG148-*natR*, without addition of IPTG

Fig. 1. Structure of the *natAB* region, and expression of *natA-lacZ* in wild-type, *natK* and *natR* cells. (a) P, phosphoryl groups; boxes, genes; ovals, NatR proteins; stem-loop structures, rho-independent terminators; bent arrows, promoters; SK, sensor kinase; RR, response regulator. (b, c) Cells were grown in MC medium and β -galactosidase activities were determined. Labels on the x-axis represent growth time in hours relative to the end of vegetative growth (t₀). All strains carried a transcriptional *natA-lacZ* fusion. The result shown is a representative experiment from at least three independent experiments. (b) Open and filled symbols correspond to the presence and absence of 1 mM IPTG, respectively. Circles, OAM298 (wild-type) carrying pDG148-*natR*; squares, OAM269 (*natK*) carrying pDG148-*natR*. (c) Open circles, OAM298 (wild-type); open squares, OAM269 (*natK*); filled circles, OAM270 (*natR*).

(Fig. 1b, c). The reason for this is unknown; however, it is possible that addition of kanamycin affects the expression level of the fusion. The expression of *natA*–*lacZ* was almost abolished in both *natK* and *natR* backgrounds (Fig. 1c). Thus, we concluded that *natK* and *natR* are required for *natA* expression. In addition, the expression of *yccK*–*lacZ* at the *yccK* locus was not affected by the disruption of *natR*, reinforcing the notion that *yccK* is not a member of the NatR regulon (data not shown).

Functional analysis of the *natR* strain

It was expected that introduction of the *natR* disruption mutation would decrease Na⁺ resistance at alkaline pH, and ethanol resistance at neutral pH, because NatR positively regulates the *natAB* operon, whose product is responsible for the resistant phenotypes (Cheng *et al.*, 1997). Thus, we examined the growth properties of the *natR* and *natA* strains in the presence or absence of 200 mM NaCl in TKM medium. The medium composition, pH, concentration of NaCl or ethanol, and culture conditions, including temperature and culture volume, were the same as those used by Cheng *et al.* (1997), with the exception that we used *B. subtilis* 168 as a control strain, and not the BD99 strain,

which is a derivative of 168. As shown in Fig. 2(b, d), the *natA* and *natR* strains showed a delay in growth, and a low cell density at the end of the exponential phase, compared with the control strain (Fig. 2a, c), indicating that NaCl sensitivity at alkaline pH, and ethanol sensitivity at neutral pH, were higher in both strains. These observations confirmed our prediction regarding the phenotypes of the *natR* strain. The *natK* strain showed similar phenotypes to the *natR* strain under both alkaline and neutral conditions (data not shown). We used strain 168, which was more resistant to NaCl at alkaline pH, and ethanol at neutral pH, than BD99. Subtle genetic differences between our strain and the one used by Cheng *et al.* (1997) may explain these differences.

Determination of the transcription start site of *natA*

To characterize *natAB* operon expression, we used primer extension to determine the transcription start site of *natA*. We isolated RNA from the wild-type strain carrying pDG148-*natR* (Table 1), which was grown in MC medium with and without 1 mM IPTG. According to the above results, *natAB* expression should be enhanced by the

Fig. 2. Growth properties of *B. subtilis* 168, *natA* and *natR* strains at alkaline pH in the presence and absence of 200 mM NaCl, and at neutral pH in the presence and absence of 3% ethanol. The results shown are a representative experiment from three independent experiments. Growth was started by diluting an overnight culture 100-fold in TKM medium (pH 8.3 for a and b, and pH 7.0 for c and d). Growth was monitored at hourly intervals by using a Klett spectrometer; cell densities are shown in arbitrary units. Circles, 168 control strain; triangles, NATAd (*natA*); squares, NATRd (*natR*). (a, b) Cells were grown in 50 ml TKM medium at 30 °C (pH 8.3). Filled and open symbols correspond to the absence and presence of 200 mM NaCl, respectively. (c, d) Cells were grown in 50 ml TKM medium at 30 °C (pH 7.0). Filled and open symbols correspond to the absence and presence of 3% ethanol.

addition of IPTG. The transcription start was clearly indicated by a band that was detected using a *natA*-specific primer in the reaction with RNA from the cells grown with 1 mM IPTG, and not from those grown without IPTG (Fig. 3). Using the RNA from the wild-type cells without multicopy *natR*, we observed a faint band of the same size as that in Fig. 3, indicating that *natA* transcription starts at the same nucleotide position, irrespective of the presence of multicopy *natR* (data not shown).

NatR binds to the *natA* promoter

The observations described above led us to hypothesize that NatR directly regulates *natA* expression by binding to the *natA* promoter. We produced a His-tagged NatR protein in *E. coli*, and purified it from a cleared cell lysate. Gel retardation analysis indicated that the recombinant His-tagged NatR bound to the *natA* promoter (Fig. 4a). The binding was specific, because the reaction contained a 500-fold excess amount of poly dI-dC.

Fig. 3. Determination of the 5' end of the *natA* transcript, using primer extension and a *natA*-specific primer. Total RNA was isolated from wild-type cells carrying pDG148-*natR*, and grown in MC medium with and without 1 mM IPTG. The cells were harvested at early stationary phase. The *natA* template for a sequencing ladder was prepared by PCR amplification with NatA-plSB1 and NatA-up-R, and is shown in lanes G, A, T and C. The NatA-bio primer was used for sequencing and primer extension analysis of *natA*. The nucleotide sequence of *natA* is shown below the gel. The asterisk indicates the transcription start site, and the deduced -35 and -10 sequences are in bold; + and - indicate the presence and absence of 1 mM IPTG, respectively.

Footprinting analysis of NatR

To characterize the *natA* promoter sequence(s) bound by NatR, we performed DNase I footprinting analysis. Protection of the regions spanning -72 to -37, and -79 to -37, relative to the transcription start site, was observed for both strands; protection was also observed for the -31 to -17 region on the top strand (Fig. 4b). These results are consistent with the observations made using gel retardation. The commonly protected region of both strands contains a direct repeat of 5'-TTCRCGACA-3', separated by 12 bp. This relatively long space between directly repeated motifs is characteristic of the LytTR family of DNA-binding proteins (see Discussion).

Confirmation that the direct repeat in the *natA* promoter region acts as a positive *cis*-element for NatR

Transcriptional *lacZ* fusions were constructed using three distinct promoter regions, with different 5' ends of *natA*, at the *amyE* locus. The common 3' end of the promoter regions was 1 nt upstream from the initiation codon. Since the fusion expression levels were very low, we introduced multicopy *natR* on pDG148-*natR*, driven by an IPTG-inducible promoter. The β -galactosidase activities of the strains were examined in the presence and absence of IPTG. As expected, a high level of expression of the fusions carrying the direct repeat recognized by NatR, i.e. *natA1-lacZ* (-116 to +27) and *natA2-lacZ* (-79 to +27), was induced by the addition of IPTG (Fig. 5). The deletion of the -116 to -79 region resulted in a 2.5-fold increase in fusion expression when *natR* was overexpressed, suggesting that there is a negative regulatory region in this region. However, the expression of *natA3-lacZ* (-57 to +27), which carries the downstream repeat only, was not induced by IPTG. This indicates that the complete direct repeat is required for expression of *natA*. Further, we mutagenized both of the direct repeat sequences in the fusions, and examined their β -galactosidase activities. The expression of both fusions carrying disruptions in the direct repeat was severely decreased in the presence of IPTG compared with the reference strains. These results demonstrate that the direct repeat, recognized by NatR, is important for the regulation of *natA*.

DISCUSSION

This study demonstrated that the NatK-NatR two-component system is a positive regulator of the *natAB* operon. Multicopy *natR* enhanced expression of both *natA-lacZ* fusions in the natural context, and at the *amyE* locus in the wild-type cells (Figs. 1b, 5). In the several *B. subtilis* two-component regulatory systems, i.e. DegS-DegU, ComP-ComA and PhoP-PhoR, multicopy of the response regulator gene can only induce the expression of the target genes when the cognate sensor kinase gene is absent (Ogura *et al.*, 2001). In the NatK-NatR system, however, the multicopy *natR* gene was able to enhance the expression of

Fig. 4. Gel retardation and footprint assays revealed binding of NatR to the *natA* promoter. (a) Gel retardation assay. Samples were separated using 6% native PAGE. Reactions contained poly dl-dC (1.0 μg per 25 μl). The DNA probe (2 nM) spans the –116 to +52 region of the *natA* promoter, and it was prepared by PCR amplification using the primers NatA-up-F and NatA-R-trpH. The probes were incubated with His-tagged NatR (lanes: 1, no addition; 2, 0.3 μM; 3, 0.6 μM; 4, 1.2 μM); c, DNA–protein complexes; f, a free probe. (b) Footprint assays. The top and bottom strand probes were prepared by PCR amplification with the primer pairs NatA-up-F and NatA-R-trpH, and NatA-plSB1 and NatA-RB, respectively. Probes (8 nM) were incubated with His-tagged NatR (0.5 and 2 μM), and subjected to DNase I cleavage. A sequencing ladder is shown in lanes G, A, T and C. Brackets and filled triangles indicate protected regions and cleavage at DNase-I-hypersensitive sites, respectively. (c) Nucleotide sequences of the *natA* promoter region. Brackets and filled triangles indicate protected regions and cleavage at DNase-I-hypersensitive sites, respectively, and the direct repeat identified is denoted in bold. The numbers at either end of the sequence indicate nucleotide positions relative to the transcription start site.

natA-lacZ, even in the presence of the cognate *natK* gene, with a slightly decreased level compared with that obtained in the absence of *natK* (Fig. 1b). The discrepancy might be due to the nature of the sensor kinase. For example, it is known that some sensor kinases have both kinase and phosphatase activities for the cognate response regulator, while, on the other hand, other kinases lack the phosphatase activity (Stock *et al.*, 2000).

The NatK sensor kinase has three transmembrane domains, and is classified as a group IIIA kinase (Fabret *et al.*, 1999). Generally, the kinases in this group are paired with group IIIA regulators. However, NatR is distinct from the other members of the group because it has characteristics that are specific to the LytTR family; these characteristics include response regulators involved in the biosynthesis of extracellular polysaccharides, fimbriation, expression of

exoproteins (including toxins), and quorum sensing (Nikolskaya & Galperin, 2002). The regulators in this family have been reported to bind to imperfect direct repeat sequence patterns composed of nine nucleotides [T/A] [A/C] [C/A]GTTN[A/G] [T/G] separated by a 12–13 bp spacer. Although the NatR-binding sequence that was identified, i.e. TTCACGACA-N12-TTCGCGACA, does not contain this sequence, it does maintain the 9-12-9 bp structure. This conforms to the idea that regulators within the same protein family may have evolved to recognize different sequence motifs.

In the footprint analysis, NatR protected not only the upstream region, but also the core promoter region of *natA* on the top strand, from the DNase I digestion. Since this region does not contain the NatR-recognition sequence, protection might be generated by the spatial proximity of

Fig. 5. Deletion and mutation experiments using the *natA-lacZ* fusion. Cells carrying the five fusions at the *amyE* locus, and pDG148-*natR*, were grown in MC medium. Sampling from the late exponential phase was hourly, and five samples were taken. The β -galactosidase activities from the transcriptional *natA-lacZ* fusions were determined, and the peak values are shown; + and - indicate the presence and absence of 1 mM IPTG, respectively. The result shown is a representative experiment from at least three independent experiments – the SD did not exceed 15%. The 5' and 3' termini of the *natA* promoter region used are shown relative to the transcription start site. Open boxes indicate the direct repeat sequences, and lower case letters indicate replaced nucleotides.

DNA-bound NatR to the core promoter region. This raised the possibility that NatR may interact with RNA polymerase in the *natA* promoter, leading to activation of *natAB* transcription, although how NatR activates *natAB* expression remains unknown. In the gel retardation assay, at the highest concentration of the protein (1.2 μ M), a significant amount of the probe was not bound by NatR. Similarly, in the footprint analysis, a large amount of NatR (2.0 μ M) was required for the clear footprint pattern. These results suggested that a significant portion of purified protein was not active.

To examine the possibility that other genes might be directly regulated by NatR, we searched the NatR-recognition sequence in the *B. subtilis* genome using GRASP-DNA sequence software (Schilling *et al.*, 2000). However, we did not find a candidate target of NatR.

It has been reported that the expression of *natAB-lacZ* is induced fourfold by the addition of 2% ethanol (Cheng *et al.*, 1997). The expression of the fusion was measured in the *natB* background (Cheng *et al.*, 1997). We also observed the induction of *natA-lacZ* expression by ethanol in *natA* cells; however, there was no ethanol induction in the wild-type cells (unpublished results). It is speculated that the lack of NatAB may perturb cellular ionic homeostasis, which affects the activity of the *natAB* promoter.

The basal expression of the *natA1-lacZ* fusion at the *amyE* locus was very low compared with the *natA-lacZ* fusion in the natural context. There could be several possible reasons

for this: (1) some regulatory *cis*-element might be lost in *natA1-lacZ*, since -116 relative to the transcription start site is the 5' end of the fusion; (2) a local topology of the *natA* promoter region might be changed in the *natA* fusions at the *amyE* locus, leading to low expression; (3) different mRNA structures of the two fusions might result in altered mRNA stability, leading to different expression levels. However, the reason for the discrepancy is yet to be determined.

Response regulators demonstrating close similarities to NatR have been found in many bacterial genomes, including those of *E. coli*, *Bacillus halodurans*, *Shigella flexneri* and *Streptococcus thermophilus*. However, we investigated over 20 genomes, among which only the *Str. thermophilus* genome has conserved the genomic structure as shown in Fig. 1(a) (Bolotin *et al.*, 2004). In the case of *Str. thermophilus* LMG18311, the LytTR-type response regulator (stu0543) is associated with a putative ABC-transporter for extrusion of an unknown substrate.

ACKNOWLEDGEMENTS

We thank I. Smith (TB Center, The Public Health Research Institute, Newark, NJ, USA) for the kind gift of plasmid. We also thank M. Tanaka and Y. Hanawa for technical assistance. This work was supported by a Grant-in-aid for Scientific Research on Priority Areas (C) 'Genome Biology' to T.T., and a Grant-in-aid for Scientific Research (C) to M.O., and (B) to T.T., from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

REFERENCES

- Aguilar, P. S., Hernandez-Arriaga, A. M., Cybulski, L. E., Erazo, A. C. & de Mendoza, D. (2001). Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J* **20**, 1681–1691.
- Asai, K., Baik, S. H., Kasahara, Y., Moriya, S. & Ogasawara, N. (2000). Regulation of the transport system for C₄-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* **146**, 263–271.
- Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S. D., Kulakauskas, S., Lapidus, A., Goltzman, E., Mazur, M. & other authors (2004). Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* **22**, 1554–1558.
- Cheng, J., Guffanti, A. A. & Krulwich, T. A. (1997). A two-gene ABC-type transport system that extrudes Na⁺ in *Bacillus subtilis* is induced by ethanol or protonophore. *Mol Microbiol* **23**, 1107–1120.
- Doan, T., Servant, P., Tojo, S., Yamaguchi, H., Lerondel, G., Yoshida, K., Fujita, Y. & Aymerich, S. (2003). The *Bacillus subtilis* *ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate. *Microbiology* **149**, 2331–2343.
- Fabret, C., Feher, V. A. & Hoch, J. A. (1999). Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *J Bacteriol* **181**, 1975–1983.
- Fukuchi, K., Kasahara, Y., Asai, K., Kobayashi, K., Moriya, S. & Ogasawara, N. (2000). The essential two-component regulatory system encoded by *yycF* and *yycG* modulates expression of the *ftsAZ* operon in *Bacillus subtilis*. *Microbiology* **146**, 1573–1583.

- Guilfoile, P. G. & Hutchinson, C. R. (1991). A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. *Proc Natl Acad Sci U S A* **88**, 8553–8557.
- Hayashi, K., Ohsawa, T., Kobayashi, K., Ogasawara, N. & Ogura, M. (2005). The H₂O₂ stress-responsive regulator PerR positively regulates *srfA* expression in *Bacillus subtilis*. *J Bacteriol* **187**, 6659–6667.
- Hayashi, K., Kensuke, T., Kobayashi, K., Ogasawara, N. & Ogura, M. (2006). *Bacillus subtilis* RghR (YvaN) represses *rapG* and *rapH*, which encode inhibitors of expression of the *srfA* operon. *Mol Microbiol* **59**, 1714–1729.
- Hyrylainen, H. L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M., Sarvas, M., Pragai, Z., Bron, S. & other authors (2001). A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Mol Microbiol* **41**, 1159–1172.
- Joseph, P., Fichant, G., Quentin, Y. & Denizot, F. (2002). Regulatory relationships of two-component and ABC transport systems and clustering of their genes in the *Bacillus/Clostridium* group suggest a functional link between them. *J Mol Microbiol Biotechnol* **5**, 503–513.
- Joseph, P., Guiseppi, A., Sorokin, A. & Denizot, F. (2004). Characterization of the *Bacillus subtilis* YxdJ response regulator as the inducer of expression for the cognate ABC transporter YxdLM. *Microbiology* **150**, 2609–2617.
- Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K., Ogasawara, N., Tanaka, T. & Fujita, Y. (2001). The comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *J Bacteriol* **183**, 7365–7370.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A. & other authors (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256.
- Mascher, T., Zimmer, S. L., Smith, T. A. & Helmann, J. D. (2004). Antibiotic-inducible promoter regulated by the cell envelope stress-sensing two-component system LiaRS of *Bacillus subtilis*. *Antimicrob Agents Chemother* **48**, 2888–2896.
- Mukai, K., Kawata, M. & Tanaka, T. (1990). Isolation and phosphorylation of the *Bacillus subtilis* *degS* and *degU* gene products. *J Biol Chem* **265**, 20000–20006.
- Nikolskaya, A. N. & Galperin, M. Y. (2002). A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res* **30**, 2453–2459.
- Ogura, M. & Tanaka, T. (1996). Transcription of *Bacillus subtilis* *degR* is σ^D -dependent and suppressed by multicopy *proB* through σ^D . *J Bacteriol* **178**, 216–222.
- Ogura, M., Ohshiro, Y., Hirao, S. & Tanaka, T. (1997). A new *Bacillus subtilis* gene, *med*, encodes a positive regulator of *comK*. *J Bacteriol* **179**, 6244–6253.
- Ogura, M., Yamaguchi, H., Yoshida, K., Fujita, Y. & Tanaka, T. (2001). DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. *Nucleic Acids Res* **29**, 3804–3813.
- Ogura, M., Shimane, K., Asai, K., Ogasawara, N. & Tanaka, T. (2003). Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol Microbiol* **49**, 1685–1697.
- Ohki, R., Giyanto, Tatenno, K., Masuyama, W., Moriya, S., Kobayashi, K. & Ogasawara, N. (2003). The BceRS two-component regulatory system induces expression of the bacitracin transporter, BceAB, in *Bacillus subtilis*. *Mol Microbiol* **49**, 1135–1144.
- Padan, E. & Schuldiner, S. (1994). Molecular physiology of Na⁺/H⁺ antiporters, key transporters in circulation of Na⁺ and H⁺ in cells. *Biochim Biophys Acta* **1185**, 129–151.
- Reizer, J., Reizer, A. & Saier, M. H., Jr (1992). A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates. *Prot Sci* **1**, 1326–1332.
- Satomura, T., Shimura, D., Asai, K., Sadaie, Y., Hirooka, K. & Fujita, Y. (2005). Enhancement of glutamine utilization in *Bacillus subtilis* through the GlnK–GlnL two-component regulatory system. *J Bacteriol* **187**, 4813–4821.
- Schilling, C. H., Held, L., Torre, M. & Saier, M. H. (2000). GRASP-DNA: a web application to screen prokaryotic genomes for specific DNA-binding sites and repeat motifs. *J Mol Microbiol Biotechnol* **2**, 495–500.
- Serizawa, M. & Sekiguchi, J. (2005). The *Bacillus subtilis* YdfHI two-component system regulates the transcription of *ydfJ*, a member of the RND superfamily. *Microbiology* **151**, 1769–1778.
- Shimane, K. & Ogura, M. (2004). Mutational analysis of the helix–turn–helix region of *Bacillus subtilis* response regulator DegU, and identification of *cis*-acting sequences for DegU in the *aprE* and *comK* promoters. *J Biochem* **136**, 387–397.
- Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* **69**, 183–215.
- Stragier, P., Bonamy, C. & Karmazyn-Campelli, C. (1988). Processing of a sporulation sigma factor in *Bacillus subtilis*: how morphological structure could control gene expression. *Cell* **52**, 697–704.
- Tanaka, K., Kobayashi, K. & Ogasawara, N. (2003). The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YfIS, and is essential for utilization of malate in minimal medium. *Microbiology* **149**, 2317–2329.
- Vagner, V., Dervyn, E. & Ehrlich, S. D. (1998). A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**, 3097–3104.
- Yamamoto, H., Murata, M. & Sekiguchi, J. (2000). The CitST two-component system regulates the expression of the Mg-citrate transporter in *Bacillus subtilis*. *Mol Microbiol* **37**, 898–912.
- Yoshida, K.-I., Kobayashi, K., Miwa, Y., Kang, C.-M., Matsunaga, M., Yamaguchi, H., Tojo, S., Yamamoto, M., Nishi, R. & other authors (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res* **29**, 683–692.

Edited by: J. M. van Dijk