Interaction of a putative transcriptional regulatory protein and the thermo-inducible *cts*-52 mutant repressor in the *Bacillus subtilis* phage ϕ105 genome

Annie Y. Chan[1] and Boon. L. Lim[1*]

[1] Department of Zoology, The University of Hong Kong, Pokfulam Road, Hong Kong, China

*Author of correspondence:

Department of Zoology, The University of Hong Kong, Pokfulam Road, Hong Kong, China.

Tel: +852 2299-0826, Fax: +852 2559-9114

Email: bllim@hkucc.hku.hk

Keywords: *Bacillus subtilis*; bacteriophage ϕ105; *cts*-52 repressor; RT-PCR; surface plasmon resonance; *imm*F region

Abbreviations: a.a., amino acids; GST, glutathione S-transferase; IPTG, isopropyl-β-*D*-thiogalactopyranoside; kb, kilobase; K_D, equilibrium dissociation constant; mtcϕ105, *cts*-52 mutant repressor; ORF, open-reading frame; O_R, operator; SPR, Surface plasmon resonance; wtcϕ105, wild-type repressor

Abstract

A 144 amino acid *cts*-52 mutant repressor (mtcϕ105) located in the *Eco*RI-F immunity region (*imm*F) of *Bacillus subtilis* phage ϕ105 is involved in the control mechanism of a thermo-inducible expression system. Adjacent to the repressor gene, an open-reading frame, designated *ORF4*, encodes a polypeptide of 90 amino acid residues, which shares a 37% homology with the amino acid sequence of the repressor. Based on protein sequence alignment, a DNA-binding αhelix-βturn-αhelix (HTH) motif was identified in the N-terminal region (residues 18–37) of the repressor as well as in the polypeptide of ORF4 (residues 22–41). *In vivo* expression of the mutant repressor and ORF4 were confirmed by real-time reverse transcriptase polymerase chain reaction (RT-PCR) and Western blot analysis. To study their DNA binding properties, the wild-type repressor (wtcϕ105) and the mutant repressor mtcϕ105, which has a Thr17 to Ile substitution, were overexpressed in *Escherichia coli* and purified for affinity assays. Their affinities towards six operator sites at various temperatures were elucidated by surface plasmon resonance (SPR). Our data showed that a temperature shift does not influence the wtcϕ105-operators' binding affinity, while the binding of mtcϕ105 to the operators was temperature sensitive. This explains how thermo-induction triggers the release of the mutant repressor and renders heterologous gene expression. Interestingly, mtcϕ105 and ORF4 demonstrated

a large affinity discrepancy towards individual operators at different temperatures. mRNA levels monitored by real-time RT-PCR indicated a suppression of *mtcϕ105* expression, but a stimulation of *ORF4* transcription after thermo-induction. Our data suggested that ORF4 might be a counter protein to the phage repressor in the modulation of the two divergent-oriented promoters P_M and P_R within the *imm*F region.

Introduction

Immediately after temperate bacteriophage infection, lysogeny is tightly regulated, by repressing the lytic promoter. Switching from lysogeny to lytic development, or prophage induction, occurs after the host SOS system is activated in response to DNA damage, for example, after ultraviolet (UV) radiation[1,2]. In the *Bacillus subtilis* phage ϕ105, a thermo-inducible prophage mutant that carries a *cts* mutation was discovered two decades ago[3]. Dhaese and colleagues[4] demonstrated that an open-reading frame (ORF), designated *cϕ105*, was responsible for the maintenance of lysogeny (Figure 1). It suppressed phage DNA replication and adjacent bacterial genes transcription before thermo-induction. Using chloramphenicol acetyl transferase (CAT) activity assays, transcriptional control in the immunity region was found to rely on two divergently orientated promoters, P_M (Left) and P_R (Right)[5,6]. Similarly to coliphage λcI, wtcϕ105 is both a negative and positive regulator of transcription. It stimulates P_M, the promoter for its own gene, but represses P_R, which probably signals the onset of a lytic pathway[6]. The repressor was shown specifically to recognise six operator sites that are identified within the 3.2 kb immunity region fragment[6–11]. Similarly to the recently reported operator sites identified in two mycobacteriophages, L5[12] and Bxb1[12,13], the 14 bp-core operator sequence 5'-GACGGAAATACAAG-3' of wtcϕ105 is lacking in two-fold rotational symmetry[10,11].

A thermo-inducible expression system based on phage ϕ105 has been established in *Bacillus subtilis*[14,15]. Unlike other *Bacillus* expression systems that employed plasmid-based, constitutive promoters, this prophage expression system is capable of driving recombinant protein expression in a thermo-inducible manner. The complete suppression of heterologous protein expression before thermo-induction is desirable for expressing toxic products. The genetic switch results from a mutation (*cts*-52) on the wtcϕ105 that renders the expression system thermo-inducible[16–18]. As shown in previous studies[19,20], induction of heterologous gene expression in the ϕ105MU331 system was achieved by a temperature shift from 37°C to 50°C. The aim of our study was to examine whether transcriptional control is closely regulated by altered affinities between the mutant repressor and the six operators at different temperatures. To address the effect of the *cts*-52 mutation[18], site-directed polymerase chain reaction (PCR) mutagenesis was performed to generate a wild-type copy of *wtcϕ105*[8,21,22].

A model for the role of cooperativity in the regulation of genetic switches is most characterised in the λ phage system. Pairs of cI dimers interact cooperatively to occupy adjacent operator sites at O_R and at O_L. These cI tetramers repress the lytic promoter P_R and activate its own transcription from P_{RM}. Transcription of cI is in turn negatively regulated by the Cro protein, which primarily represses P_{RM}[23–25]. A more complicated immunity mechanism has recently been characterised in *Salmonella*

phage P22; this involves three immunity systems[26–29], showing that some phages utilise more than one immunity repressor to safeguard the control of immunity. In the ϕ105 genome, an ORF adjacent to the repressor gene, denoted *ORF4*, which is situated downstream from P_R[10,11], has never been studied (Figure 1). Sequence analysis showed that ORF4 is potentially another transcriptional regulatory protein. ORF4 shares a 37% sequence identity with mtcϕ105 and both proteins carry a typical DNA-binding αhelix-βturn-αhelix (HTH) motif at their N-termini. To examine whether ORF4 is also a transcriptional regulator that acts upon P_M and P_R, it was overexpressed and its binding affinities towards the operators were compared with that of the repressor. The mRNA levels of *mtcϕ105* and *ORF4* were also monitored by real-time RT-PCR so as to determine the influence of temperature shift on *mtcϕ105* and *ORF4* transcription. Based on these findings, a model of the impact of thermo-induction on the transcriptional regulation within the *imm*F region is proposed.

Results

Sequence analysis of mtcϕ105 and ORF4

Based on the results of nucleotide sequencing, the *cts*-52 mutation designated by previous studies[16,17] was found to be a T to C transition that is equivalent to the *cts*-23 mutation[3,4,18]. This point mutation not only leads to an amino acid (a.a.) substitution from Thr17 to Ile (Figure 2A), but also renders the repressor thermo-inducible. Analysis of the ϕ105 genome sequence found that *ORF4* coding sequence is located adjacent to the gene of the cϕ105 repressor and shares a 37% a.a. sequence identity. A DNA-binding HTH motif was identified in the N-terminal regions (residues 18–37) of the 144 a.a *cts*-52 mutant repressor as well as of the 90 a.a. polypeptide of ORF4 (residues 22–41). Such similarities indicated that ORF4 is potentially another DNA-binding protein within the *imm*F region and it is likely to interact with the same six operator sites as the repressor..

After the a.a. sequence of ORF4 was deduced by DNA sequencing, the sequence was blasted in GenBank and showed a high homology to the consensus sequence of a large DNA-binding protein family, termed HTH3. λcI, λCro and P22 are the prominent representatives in this family. The alignment of the protein sequences of the DNA-binding proteins (Figure 2B) is based, in part, on sequence homology and

also on structural homology among them. The region of maximal sequence homology is found within a 20-residue sequence that aligns with the α2 and α3 helices of λcI, λCro and P22[27,28,30].

The ϕ105 genome was searched for sequences that are closely related to the 14 bp-core region. A consensus sequence template was determined using the core sequences in O_R1, O_R2, O_R3, O_R4, O_R5 and O_R6. At least 28 regions were found to align with the consensus sequence GWCGKRAATWCMAK (W=A or T, K=G or T, R=A or G, M=A or C), including the O_R1–O_R6 (data not shown). In the thermo-inducible expression system[14,15], the gene of interest is cloned downstream of a strong holin gene promoter[14]. However, no such consensus sequence was identified at the –35 or –10 RNA recognition regions of the holin gene promoter. A direct control of heterologous gene expression in the phage expression system is therefore unlikely.

Identification of a novel protein designated ORF4

To determine whether the *ORF4* is a functional gene or a pseudogene, a RT-PCR was performed. Total RNA was extracted from wild-type *B. subtilis* strain 168 and the lysogen *IA304(ϕ105MU331).* Contamination of genomic DNA was prevented by DNaseI treatment. As shown in Figure 3, PCR products of both ORF4 and the

repressor were amplified from strain *IA304(ϕ105MU331)*, but not from *B. subtilis* strain 168. The identity of the PCR product was verified by Southern hybridisation (data not shown). Samples that had not been reverse transcribed showed no detectable amplification (lane 3 and 7), indicating the absence of contaminating DNA.

Western blot analysis further verified the *in vivo* expression of mtcϕ105 repressor and ORF4 coding sequence in *IA304(ϕ105MU331)* (Figure 4). Data showed that the sizes of mtcϕ105 and ORF4 proteins were approximately 16 kDa and 13 kDa, respectively.

mtcϕ105 and *ORF4* transcription levels in the course of thermo-induction

To analyze the effect of thermo-induction on mtcϕ105 and ORF4 transcription, total RNA was extracted at four time points: before induction, 10 minutes, 30 minutes and 60 minutes after induction. The expression levels of each target sequence were compared using the $2^{-\Delta\Delta C}{}_T$ method[31]. PCR efficiency of the target coding sequences (*mtcϕ105* and *ORF4*) and internal control (16SrRNA) was shown to be similar using real-time RT-PCR and TaqMan detection (Figure 5). Thus, relative quantification by the $2^{-\Delta\Delta C}{}_T$ method was validated in these experiments. According to the data collected, the transcript level of *mtcϕ105* decreased moderately after thermo-induction and then gradually recovered (Figure 6). Following a 60 minute incubation, the expression

level increased by three-fold. In the case of *ORF4*, expression was dramatically increased by 65-fold immediately after thermo-induction, followed by a moderate drop to approximately 30-fold. These observations suggested that the two ORFs were functional and expressed in *IA304(ϕ105MU331)*, but that the amounts of transcript were not equivalent at various time points. In general, this indicated the P_M and P_R activities are influenced by a temperature shift.

Kinetic analysis of protein and operator sites interaction

To assess the interaction of the two proteins with P_M and P_R, real-time protein-DNA interaction is recorded continuously in terms of resonance units over a function of time, leading to a sensorgram which allows us to study the association and dissociation of protein-DNA complexes within the same experiment[32]. To address the specificity of protein–operator interaction, two control experiments, using double-stranded oligonucleotides without a consensus sequence and a sensor flow cell containing no DNA as negative controls were performed. The control experiments showed no non-specific binding between the reference cell and the non-specific oligonucleotides (data not shown). Prior to the calculation of the equilibrium dissociation constant, all resonance units in the kinetic measurements were subtracted with the reference data.

To evaluate the accuracy of the surface plasmon resonance (SPR) technique, equilibrium dissociation constants of repressor–operator site interactions obtained from SPR were compared with the results generated from *in vitro* repressor-binding competition experiments[11]. Both experiments provided similar results. At room temperature (22.5°C), mtcϕ105 showed the strongest binding to O_R3 ($K_D \sim 10^{-11}$M), followed by O_R1, O_R2, then O_R6 (Table 1). In contrast, ORF4 showed a very weak interaction ($K_D \sim 10^{-7}$M) towards O_R3 that overlapped with its own gene. Both mtcϕ105 and ORF4 showed low intrinsic affinity towards O_R4 and O_R5 at 22.5°C. This could be explained by the fast association and dissociation rates of the complexes. Surprisingly, although O_R1, O_R2 and O_R3 bear an identical consensus sequence, they differ in protein-binding affinity. Seemingly, the flanking regions of these operators to some extent lead to discrepancies in DNA conformation, thereby affecting the binding affinities towards mtcϕ105 and ORF4. In short, we conclude that both mtcϕ105 and ORF4 are DNA-binding proteins that bind to multiple but non-identical 14-bp operator sites within the *imm*F region. Specific binding to all operator sites was obtained, both with mtcϕ105 and ORF4.

In order to examine the effect of temperature on these DNA–protein interactions, SPR analyses were performed at 37°C, with or without prior incubation of the proteins at 50°C (Table 1). Obviously, a temperature shift leads to an alternation in binding

behaviour between the proteins and operators. At 37°C, both proteins showed an increase in affinity towards O_R4 and O_R5. The affinities of mtcϕ105 towards O_R2, O_R3, O_R4 and O_R6 were strong and similar ($K_D \sim 10^{-10}$M). The interactions between mtcϕ105 and O_R1 or O_R5 were ten-fold ($K_D \sim 10^{-9}$M) weaker. In the case of ORF4, a weak association with O_R3 ($K_D \sim 10^{-8}$M) was recorded, while it was found to have comparable affinities ($K_D \sim 10^{-10}$M) towards the other five operator sites.

To determine whether thermo-induction alters these affinities, protein samples were incubated at 50°C for 10 minutes and then reincubated at 37°C before analysis. SPR analysis was then carried out at 37°C. Our data indicated that there was more than a ten-fold decrease in affinity in the interaction of mtcϕ105 towards O_R3, O_R4, O_R5 and O_R6 when compared with that obtained without heat treatment. Under the same conditions, the binding profile of ORF4 was also altered. ORF4 showed a very strong affinity towards O_R1 ($K_D \sim 10^{-13}$M) after heat incubation, i.e. a 10,000-fold increase. Additionally, the binding between ORF4 and O_R3 seemed to be stronger than that measured before induction ($K_D \sim 10^{-9}$M).

Mutational analysis of thermo-inducible repressor

A wild-type repressor was prepared by site-directed PCR mutagenesis. Interactions between the wild-type repressor and the six operator sites were assayed at 37°C,

before and after heat treatment (Table 2). At 37°C, the binding behaviour of the wild-type repressor was similar to that of the mutant repressor. However, unlike the mutant repressor, the binding affinities of the wild-type repressor towards the six operators were not affected by heating. Even after heat treatment, no significant alternation in binding was observed.

Discussion

Most of the studies on bacteriophage ϕ105 focused on the regulatory role of *cϕ105* in maintaining lysogeny but ignored the existence of *ORF4* in the close proximity. Our data suggested that ORF4 is a functional polypeptide and its expression is regulated by the repressor.

The repressor and ORF4 isolated from *IA304(ϕ105MU331)* are likely to belong to the large family of HTH DNA-binding proteins, designated as HTH3. Both proteins showed homology to the α_3 helix of λcI that presents DNA recognition specificity. Western blot analysis indicated that proteins of the two ORFs were expressed *in vivo* and were able specifically to bind to the six individual operators. It is believed that ORF4 is potentially another DNA-binding protein within the *imm*F region of the ϕ105 genome.

To determine whether both proteins bind to various operators at different affinities and to examine whether their affinities are affected by a temperature shift, protein-DNA interaction was measured *in vitro* using SPR technology. The SPR technique has an advantage over traditional electrophoretic mobility shift assays, since it can also measure the association and dissociation rates of the protein–DNA interaction. Therefore, this technique prevents underestimation of an affinity constant

as a result of a high dissociation rate of the complexes, and vice versa.

In the *ϕ105MU331* expression system, cells bearing a heterologous gene construct were grown at 37°C and then thermo-induced at 50°C. Subsequently, heterologous protein was detected after 30 minutes' incubation at 37°C. To imitate this physical change, purified proteins were incubated at the desired temperature prior to SPR. The activities of the P_M and P_R promoter were also monitored by their gene products, *mtcϕ105* and *ORF4* mRNAs, using real-time RT-PCR analysis. Based on these findings, a model of P_M and P_R modulation within the *imm*F region is proposed and illustrated in Figure 7. Before thermo-induction, mtcϕ105 could bind to all six operators with a similar level of affinities. It is likely that its binding to O_R1, O_R2 and O_R4 could stimulate $P_{M,}$ while its interaction with O_R5 and O_R6 could suppress P_R, as proven by the CAT activity assay[6]. Moreover, binding of mtcϕ105 to O_R3 indefinitely hinders ORF4 expression (Figure 7A). The above assumptions are confirmed by the transcript levels of mtcϕ105 (1.05×10^5 copies/0.1 μg total RNA) and ORF4 (2.29×10^2 copies/0.1 μg total RNA) at the lysogeny state (data not shown).

After thermo-induction, the affinities of mtcϕ105 towards O_R4–6 were moderately reduced (30–60 fold). Release of mtcϕ105 from O_R3, O_R5 and O_R6 derepressed P_R activity and permitted *ORF4* transcription. Thus, a dramatic increase in *ORF4* transcript level was observed (>60 fold). Thermo-induction also induced a drastic

increase (>1000 fold) in the affinity of ORF4 towards O_R1 (Figure 7B). A moderate fall in *mtcϕ105* transcript level was observed after thermo-induction (Figure 6), which could be either attributed to the strong binding of ORF4 to O_R1, or the loss of binding of mtcϕ105 to O_R4, or both.

Our data further confirmed that wtcϕ105 is insensitive to thermo-induction, which is also supported by Rutberg L and colleagues[5]. Hence, the *cts*-52 mutation on *mtcϕ105* is the sole critical protein responsible for the modulation of the genetic switch in the *ϕ105MU331* expression system.

In the *ϕ105MU331* expression system, the expression of the heterologous gene was controlled by the promoter of the holin gene, which is located 10.6 kb from the *imm*F region. Although thermo-induction was found to turn on the promoter activity of the holin gene[19], it was deemed unlikely that mtcϕ105 nor ORF4 had any direct interaction with this promoter, as no consensus operator sequence was identified at its upstream region. It is supposed that the binding of mtcϕ105 to certain operator sequences in the ϕ105 genome will block either the expression of ORF4 or other *trans* factor necessary for triggering heterologous gene expression. To further elucidate the sequence of events, a microarray displaying all 51 ORFs (GenBank accession number: AB016282) in the bacteriophage genome in both the forward and reverse directions should provide more information.

In summary, we conclude that ORF4 is a putative DNA-binding protein. It is a novel transcriptional regulator discovered in bacteriophage ϕ105. Both mtcϕ105 and ORF4 respond to thermo-induction, which in turn directs the expression of the heterologous gene in the *ϕ105MU331* system. It is likely that mtcϕ105 and ORF4 in bacteriophage ϕ105 resemble the genetic switch of the lambda phage cI and Cro system. These two proteins may be tightly counterbalanced to respond to environmental stress.

Materials and Methods

Bacterial strains

Details of the bacterial strains and phage employed in this study are listed in Table 3. *E. coli* strain JM109 (Stratagene) was used for plasmid construction and *E. coli* strain BL21(DE3) (Stratagene) was used as an expression host for the pGEX vectors (Amersham Biosciences).

Wild-type *B. subtilis* strain 168 was used in this study as a control. The detection of *mtcϕ105* and *ORF4* was performed in *B. subtilis* strain *IA304* carrying a *ϕ105MU331* prophage designated *IA304(ϕ105MU331)*[14]. This ϕ105 derivative prophage carries an *ind* mutation, which restores almost normal competence, and a *cts*-52 mutation, which confers thermo-induction.

One-step RT-PCR and Southern Blot analysis

To detect the presence of *mtcϕ105* and *ORF4* mRNA transcripts in *B. subtilis IA304 (ϕ105MU331),* one-step RT-PCR and Southern hybridisation were performed. Single colonies of *B. subtilis* strain 168 and *IA304 (ϕ105MU331)* were cultivated for 12 hours at 37°C and 280 rpm in 5 ml brain heart infusion (BHY) medium (3.7%(w/v) (Oxoid), 0.5% (w/v) yeast extract (Oxoid) and 5 µg/ml erythromycin). One millilitre of the overnight culture was sub-cultured the next day in 15 ml BHY medium without

antibiotic until the OD_{600} reached 3.0. The culture was then heat induced in a 50°C water bath for 4 minutes, with vigorous shaking, followed by re-incubation at 37°C. Bacterial cells were collected at four time points: before induction, and 10 minutes, 30 minutes and 60 minutes after induction.

Total RNA was extracted by the RNease Mini Kit (QIAgen) and treated with DNase I (DNA-*free* kit, Ambion) to remove contaminating DNA. After removal of DNase I, the integrity of the RNA samples was monitored by gel electrophoresis and OD_{260} measurement. Fifty nanograms of total RNA were added as the PCR template in individual PCR reactions using the OneStep RT-PCR System (QIAgen). Additionally, two negative controls were prepared, either in the absence of RNA templates or by omitting the reverse transcription step. Two pairs of gene-specific primers, ORF1BamS/ORF1EcoA and ORF4BamS/ORF4EcoA (Table 4), were employed for *mtcϕ105* and *ORF4* amplification, respectively. Each primer was added at a final concentration of 0.6 μM. The RNA samples were reverse-transcribed at 50°C for 30 minutes, followed by incubation at 95°C for 15 minutes before amplification. Forty amplification cycles were then carried out by denaturation of cDNA at 94°C for 30 seconds, annealing at 50°C for 40 seconds and polymerization at 72°C for 1 minute. A final extension step was performed at 72°C for 10 minutes. The same volume of PCR products was loaded onto 1% (w/v) formaldehyde gel and visualised under UV

transillumination. To avoid cross-contamination, filtered tips and separate areas were used for RNA extraction and RT-PCR.

The identity of the PCR product was verified by non-radioactive Southern hybridization by applying gene-specific DNA probes labelled with digoxigenin (Roche). An *mtcϕ105*-specific probe was synthesized by PCR using the primer pair ORF1PS/ORF1PA, whereas the *ORF4*-specific probe was synthesized using the primer pair ORF4PS/ORF4PA. All working procedures were followed as recommended by the manufacturer. Chemiluminescent detection was conducted using CSPD (Roche) as substrate.

Real-time RT-PCR

We assessed the relative quantity of *mtcϕ105* and *ORF4* mRNA transcripts at various time points using real-time RT-PCR (Bio-Rad iCycler iQ Detection System) and the TagMan MGB probes (Applied Biosystems) (Table 4). One hundred nanograms of total RNA was added to 25 μl of TaqMan EZ RT-PCR Core Reagent (Applied Biosystems) in each reaction. No Template Controls (NTC) were prepared in parallel reactions to test for contamination. All the reaction mixtures were transferred to a 96-well PCR plate (Bio-Rad) and subjected to initial AmpErase UNG (uracil N-glycosylase) treatment at 50°C for 2 minutes, followed by reverse transcription at

60°C for 30 min. Deactivation of UNG was carried out at 95°C for 5 minutes and a PCR amplification protocol of 40 cycles of denaturation at 94°C for 20 seconds and annealing at 60°C for 1 minute was performed.

We adopted a relative quantification approach in which the expression levels of the target genes were compared to the data obtained at a time point before thermo-induction. The mean fold change in expression of the target gene at each time point was calculated using the $2^{-\Delta\Delta C}{}_T$ method[31], where $\Delta\Delta C_T = (C_{T,\ Target} - C_{T,\ 16SrRNA})_{Time\ x} - (C_{T,\ Target} - C_{T,\ 16SrRNA})_{Time\ 0}$. The amount of *mtcϕ105* and *ORF4* mRNA transcripts was normalised to the expression of *B. subtilis 168* 16SrRNA in each sample. To validate the $\Delta\Delta C_T$ calculation, the amplification efficiency of the target and internal control were assessed by preparing a serial dilution of total RNA. For each dilution, samples were amplified using primers and fluorogenic probes for *mtcϕ105*, *ORF4* and 16SrRNA. The average C_T was determined and a plot of C_T versus the log total RNA dilution was analysed. When the slope is close to zero, the PCR efficiency of the target and internal control is similar.[31].

DNA manipulation and plasmid construction

In order to overexpress mtcϕ105 and ORF4 proteins, plamids pCϕ105 and pORF4 were constructed. Gene-specific primer pairs composed of specific restriction sites at the 5'ends, ORF1BamS/ORF1EcoA and ORF4BamS/ORF4EcoA, (Table 4), were

used for *mtcϕ105* and *ORF4* amplification, respectively. Genomic DNA from *B. subtilis IA304 (ϕ105MU331)*, employed as a template, was firstly denatured at 94°C for 3 minutes, followed by 30 amplification cycles of denaturation at 94°C for 40 seconds, annealing at 50°C for 40 seconds and polymerization at 75°C for 1 minute. A final extension was carried out at 75°C for 5 minutes. All amplification steps were performed using *Pfu* polymerase (Invitrogen) to enhance fidelity. After restriction enzyme digestion, the PCR products were ligated to the *Bam*HI and *Eco*RI sites of the expression vector pGEX-2T (Amersham Biosciences), at a position downstream of the glutathione-S-transferase (GST)-tag. The recombinant plasmids designated pCϕ105 and pORF4 were transformed into JM109 by electroporation. Positive clones were selected on Luria-Bertani (LB) agar plates supplemented with 100 μg/ml ampicillin and PCR screened. All plasmids were sequenced using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Subsequently, the clones were transformed into the expression *E. coli* host BL21(DE3).

Site-directed PCR mutagenesis

Site-directed PCR mutagenesis was performed in a three-step PCR reaction. Template DNA used for PCR mutagenesis was pCϕ105, and the primers are listed in Table 4. The first PCR was carried out with primer ORF1BamHS and primer wtA. The second PCR was performed using primer wtS and primer ORF1EcoA. The DNA fragments

bearing the wild-type *cϕ105* sequences were annealed in a 1:1 ratio by heating at 95°C for 10 minutes and then allowed to cool down at room temperature for 1 hour. A final PCR was performed using primer pairs ORF1BamS/ORF1EcoA. Finally, the PCR product was digested and ligated to pGEX-2T (Amersham Biosciences) as described above.

Protein overexpression and purification in *E. coli*

LB broth supplemented with 2% (w/v) glucose and 100 μg/ml ampicillin was used to grow BL21(DE3) strain. Induction was achieved by the addition of 0.1 mM IPTG (Sigma) when the cells reached the log phase. The cells were then incubated at 25°C for 3 hours. Expression of both fusion proteins, GST-mtcϕ105 and GST-ORF4, was confirmed by SDS-PAGE and Western blot analysis using anti-GST antibodies (Amersham Biosciences).

The fusion proteins were purified using Glutathione Sepharose4B matrix (Amersham Biosciences). After sample loading and column washing, thrombin was loaded for the cleavage of the GST-tag. Target proteins were then eluted. The protein identities were verified by using the Protein Sequencer G1000A System (Hewlett Packard).

SDS-PAGE and Western blot analysis of mt*cϕ105* and *ORF4* expression in *B.*

subtilis IA304 (ϕ105MU331)

Twenty μg of protein extracts from various time points were examined on 15% SDS-PAGE, followed by Western Blot analysis. Purified mtcϕ105 and ORF4 polypeptides were applied as positive controls. Protein concentration was determined by a Bio-Rad dye-binding assay, using IgG as the standard.

Kinetic analysis of protein-DNA interaction

To study the interaction of mtcϕ105 and ORF4 with the six operator sequences, BIAcore 2000 (BIAcore AB) was used. By monitoring the adsorption of biomolecules on a sensor chip, real-time protein–DNA interaction was measured. This optical technique measures changes in refractive index in the vicinity of the surface. Such changes are directly proportional to the change in adsorbed mass and are expressed in resonance units (RU); 1000 RU corresponds to a surface concentration of ~1 ng/mm^2.

Prior to DNA immobilisation on the sensor chip matrix, biotinylated oligonucleotides were annealed to their complementary oligonucleotides (Table 4), by being heated at 95°C for 10 minutes and allowed to cool at room temperature for 1 hour. The double-stranded oligonucleotides individually possessed of the six operator-consensus sequences were denoted, in order, as O_R1, O_R2, O_R3, O_R4, O_R5 and O_R6. Addition of a spacer at the N or C terminus was based on the sequence of the parent DNA

sequences to avoid steric hindrance during protein binding. A pair of duplex oligonucleotides without the consensus sequence was designed to test for non-specific binding.

Biotinylated double-stranded oligonucleotides were immobilised on a streptavidin-coated SA chip (Amersham Biosciences). The procedure was followed according to the method described by Blaesing and colleagues[33]. About 100 RU DNA was immobilised on the sensor chip matrix. This relatively low concentration of DNA avoids mass transport limited conditions that may interfere with kinetic measurements. In each sensor chip, a blank surface without DNA was included as a reference. After oligonucleotide immobilisation, the sensor chip surface was subsequently saturated with biotin.

In order to examine the binding between mtcϕ105 and ORF4 with the six operator sequences, purified proteins were diluted to a concentration range from 0.8 nM to 40 nM, so as to perform a precise kinetic measurement. Analyses were performed at a flow rate of 100 µl/min in HBS-EP buffer (0.01 M HEPES, pH 7.4, 0.15M NaCl, 3 mM EDTA, 0.005% polysorbate 20 (v/v)). To simulate thermo-induction, protein samples were incubated at 50°C for 10 minutes and then at 37°C for 10 minutes before analysis. At every measurement, 250 µl protein was injected. Refractive deviations of the different sensor chip surfaces were normalised at the beginning of

each measurement. For regeneration, the sensor chip surface was briefly pulsed by 10 mM glycine-HCl, pH 3.0, to remove bound protein. Data analysis was performed with BIAevaluation 2.1 software (BIAcore). All data were handled after subtraction of the reference data obtained from a sensor flow cell containing no DNA (data not shown). These data were successfully fitted to a 1:1 Langmuir binding model.

Statistics

The results of the quantitative PCR data were analysed by One Way ANOVA, using SPSS version 11.0, and ranked by Turkey test into different groups. The relationship between the level of expression corresponding to each studied coding sequence and the time point after thermo-induction was evaluated by simple linear regression. The significance of the model was assessed by analysis of variance of the regression. Statistical significance was set at a *p*-value of less than 0.05.

Acknowledgements

We would like to thank Dr. Thomas Leung for providing the *Bacillus* strains and Dr. Van-Kaer for his advice on protein expression. We are also grateful to Miss Elisa Lau for her assistance in the operation of the BIAcore biosensor.

References

1. Rubinstein, C.P., Coso, O.A., Ruzal, S., and Sanchez-Rivas, C. 1993. Anti-SOS effects induced in Bacillus subtilis by a phi 105 mutant prophage. *Arch.Microbiol.* 160:486-491.

2. Rubinstein, C.P., Guerchicoff, A.and Sanchez-Rivas, C. 1998. Normal induction of the SOS response in Bacillus subtilis is prevented by the mutant repressor from phage phi 105cts23. *FEMS Microbiol. Lett*. 167:315-320.

3.Armentrout, R.W. and Rutberg, L. 1971. Heat induction of prophage phi 105 in Bacillus subtilis: replication of the bacterial and bacteriophage genomes. *J. Virol*. 8:455-468.

4. Dhaese, P., Hussey, C., and Van-Montagu, M. 1984. Thermo-inducible gene expression in Bacillus subtilis using transcriptional regulatory elements from temperate phage phi 105. *Gene* 32:181-194.

5. Rutberg, L. 1973. Heat induction of prophage phi 105 in Bacillus subtilis: bacteriophage-induced bidirectional replication of the bacterial chromosome. *J. Virol*. 12 :9-12.

6. Van-Kaer, L., Van-Montagu, M., and Dhaese, P. 1987. Transcriptional control in the EcoRI-F immunity region of Bacillus subtilis phage phi 105. Identification and unusual structure of the operator. *J. Mol.Biol.* 197:55-67.

7. Bugaichuk, U.D., Deadman, M., Errington, J. and Savva, D. 1984. Restriction enzyme analysis of Bacillus subtilis bacteriophage phi 105 DNA. *J. Gen.Microbiol*. 130:2165-2167.

8. Cully, D.F., and Garro, A.J. 1985. Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage phi 105: identification of the repressor gene and its mRNA and protein products. *Gene* 38:153-164.

9. Scher, B.M., Law, M.F., and Garro, A.J., 1978. Correlated genetic and EcoRI cleavage map of Bacillus subtilis bacteriophage phi105 DNA. *J. Virol*. 28:395-402.

10. Van-Kaer, L., Gansemans, Y., Van-Montagu, M., and Dhaese, P. 1988. Interaction of the Bacillus subtilis phage phi 105 repressor DNA: a genetic analysis. *EMBO J.* 7:859-866.

11. Van-Kaer, L., Van-Montagu, M., and Dhaese, P.1989. Purification and in vitro DNA-binding specificity of the Bacillus subtilis phage phi 105 repressor. *J.Biol.Chem.* 264:14784-14791.

12. Mediavilla, J., Jain, S., Kriakov, J., Ford, M.E., Duda, R.L., Jacobs, W.R., Hendrix, R.W., and Hatfull, G.F. 2000. Genome organization and characterization of mycobacteriophage Bxb1. *Mol. Microbiol*. 38:955-970.

13. Jain, S., and Hatfull, G.F. 2000. Transcriptional regulation and immunity in mycobacteriophage Bxb1. *Mol. Microbiol.* 38:971-985.

14. Leung, Y.C., and Errington, J. 1995. Characterization of an insertion in the phage phi 105 genome that blocks host Bacillus subtilis lysis and provides strong expression of heterologous genes. *Gene* 154:1-6.

15. Thornewell, S.J., East, A.K., and Errington, J. 1993. An efficient expression and secretion system based on Bacillus subtilis phage phi 105 and its use for the production of B. cereus beta-lactamase I. *Gene* 133:47-53.

16. Errington, J., and Jones, D. 1987. Cloning in Bacillus subtilis by transfection with bacteriophage vector phi 105J27: isolation and preliminary characterization of transducing phages for 23 sporulation loci. *J. Gen. Microbiol*. 133:493-502.

17. Jones, D., and Errington, J. 1987. Construction of improved bacteriophage phi 105 vectors for cloning by transfection in Bacillus subtilis. *J. Gen. Microbiol.* 133:483-492.

18. Osburne, M.S., Craig, R.J., and Rothstein, D.M. 1985. Thermoinducible transcription system for Bacillus subtilis that utilizes control elements from temperate phage phi 105. *J. Bacteriol*. 163:1101-1108.

19. Chan, A.Y., Chan, M.M., Lo, H.M., Leung, Y.C., and Lim, B.L. 2002. A dual protein expression system in *Bacillus subtilis*. *Protein Expr. Purif*. 26:337-342.

20. Tye, A.J., Siu, F.K.Y., Leung, T.Y.C., and Lim, B.L. 2002. Molecular cloning and the biochemical characterization of two novel phytases from *B. subtilis* 168 and *B. licheniformis*. *Appl. Microbiol. Biotechnol*. 59:190-197.

21. Armentrout, R.W., and Rutberg, L. 1970. Mapping of prophage and mature deoxyribonucleic acid from temperate Bacillus bacteriophage phi 105 by marker rescue. *J. Virol*. 6:760-767.

22. Dhaese, P., Seurinck, J., De-Smet, B., and Van-Montagu, M. 1985. Nucleotide sequence and mutational analysis of an immunity repressor gene from Bacillus subtilis temperate phage phi 105. *Nucleic Acids Res.* 13:5441-5455.

23. Backman, K., Hymayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B., Sauer, R.T., and Ptachne, M. 1976. Autoregulation and function of a repressor in bacteriophage lambda: interactions of a regulatory protein with sequences in DNA meditate intricate patterns of gene regulation. *Science* 194:156-161.

24. Ptashne, M. 1992. The master elements of control. In: Mark Ptashne, Editors.. A genetic switch: phage lambda and higher organisms. Blackwell Scientific Publications & Cell Press, Cambridge, MA, USA, p. 13.

25. Meyer, B.J., Maurer, R., and Ptashne, M. 1980. Gene regulation at the right operator (O-R) of bacteriophage lambda: 2. O-R1,O-R2 and O-R3: their roles in mediating the effects of repressor and cro. *J. Mol..Biol.* 139:163-194.

26. Berggrun, A., and Sauer, R.T. 2001. Contributions of distinct quaternary contacts to cooperative operator binding by Mnt repressor. *Proc. Natl. Acad. Sci. U.S.A.* 98:2301-2305.

27. De-Anda, J., Poteete, A.R., and Sauer, R.T. 1983. P22 c2 repressor. Domain structure and function. *J. Biol. Chem*. 258:10536-10542.

28. Poteete, A.R. and Ptashne, M. 1982. Control of transcription by the bacteriophage P22 repressor. *J. Mol. Biol*. 157:21-48.

29.Vander-Byl, C., and Kropinski, A.M. 2000. Sequence of the genome of Salmonella bacteriophage P22. *J. Bacteriol*. 182:6472-6481.

30. Lee, S.J., Shirakawa, M., Akutsu, H., Kyogoku, Y., Shiraishi, M., Kitano, K., Shin, M., Ohtsuka, E., and Ikehara, M. 1987. Base sequence-specific interactions of

operator DNA fragments with the lambda-cro repressor coupled with changes in their conformations. *EMBO J.* 6:1129-1135.

31. Livak, K.J., and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402-408.

32. Persson B., Buckle M., and Stockley P.G. 2000. Kinetics of DNA interactions surface plasmon resonance spectroscopy,. In: Travers A. and Buckle M., Editors. DNA-protein interactions. Oxford University Press, New York, NY, USA, pp. 257-279.

33. Blaesing, F., Weigel, C., Welzeck, M., and Messer, W. 2000. Analysis of the DNA-binding domain of Escherichia coli DnaA protein. *Mol. Microbiol*. 36:557-569.

Figure legends

Figure 1. Orientation of the six operator sites and location of the two ORFs (*cϕ105* and *ORF4*) within the imm*F* region[6]. Generally, O_R1, O_R2 and O_R4 sites are located closer to P_M. O_R1 and O_R2 are positioned upstream from the RNA polymerase recognition site (-35), while O_R4 overlaps with this sequence. O_R5 and O_R6 are positioned near to P_R and the latter overlaps with the RNA polymerase recognition site (–10). O_R3 is located within the coding sequence *ORF4*.

Figure 2. Protein alignment. (A) Comparison of mtCϕ105 and ORF4 polypepetides. Identical residues are shaded by grey boxes. The Thr17→ Ile mutation that renders mtcϕ105 thermo-inducible is squared. (B) Alignment of the αhelix-βturn-αhelix motif. Conserved residues are underlined and the asterisk indicates the position of “turn”. The ninth glycine residue is an amino acid favourable for forming a short loop between the two helices.

Figure 3. RT-PCR analysis. *mtcϕ105* (lane 4) and *ORF4* transcripts (lane 8) were detected in *B. subtilis IA304(ϕ105MU331)* but not in *B. subtilis* strain 168 (lanes 2 and 6). Two negative controls, without the addition of RNA templates (lanes 1 and 5)

or by omitting the reverse transcription step (lanes 3 and 7), are shown.

Figure 4. Western blot analysis of the *mtcϕ105* and *ORF4* proteins after thermo-induction. The protein bands of 16 kDa and 13 kDa were mtCϕ105 and ORF4, respectively.

Figure 5. Validation of the $2^{-\Delta\Delta C}{}_{T}$ method in relative quantification of *mtcϕ105* and *ORF4* transcript levels. The data were fitted using least-squares linear regression analysis (n=4). When the slope is close to zero, the PCR efficiency of the target and internal control is similar.

Figure 6. Fold-changes in the mRNA levels of *mtcϕ105* and *ORF4*. The data were normalised to the expression of 16SrRNA at various time points. Horizontal bars represent the geometric means for four replicates at each time point (n=4).

Figure 7. A model on the modulation of P_M and P_R by mtcϕ105 and ORF4 at the *imm*F region of the *Bacillus* phage ϕ105 genome. Interaction of mtcϕ105 and ORF4 with individual operators are illustrated (A) before and (B) after thermo-induction. (* represents a competition; represents promoter activation; , represents promoter repression.)

Table 1. Equilibrium dissociation constants of mtcϕ105 and ORF4 towards the six operator sites.

Table 2. Differences in the equilibrium dissociation constants between the mutant repressor and wild-type repressor.

Table 3. Bacterial strains and bacteriophage employed in this study.

Table 4. Oligonucleotides for PCR, RT-PCR, DIG-probe, TaqMan® probe synthesis and surface plasmon resonance.

Table 1

Protein	mtcϕ105 K_D(nM)			ORF4 K_D(nM)		
Oligonucleotide	22.5°C [a]	37°C [a]	50°C → 37°C [a]	22.5°C [a]	37°C [a]	50°C → 37°C [a]
O_R1	0.60	1.01	0.30	0.09	0.25	0.00018
O_R2	0.89	0.26	0.24	0.17	0.11	0.87
O_R3	0.03	0.36	21.9	719	42.4	0.57
O_R4	N.D[b]	0.64	24.2	N.D[b]	0.44	0.92
O_R5	N.D[b]	1.23	38.3	N.D[b]	0.98	0.30
O_R6	3.29	0.52	26.1	3.43	0.18	0.29

[a] Real-time interaction was monitored at 22.5°C, 37°C or after incubation at 50°C prior to analysis at 37°C. Results were obtained from two independent measurements, with protein concentrations ranging from 0.8 nM to 40 nM. These data successfully fitted into the 1:1 Langmuir binding model.

[b]Not determined because of fast association and dissociation rate.

Table 2

Protein	Wild-type cϕ105 K_D (nM)	Mutant cϕ105 K_D(nM)	Wild-type cϕ105 K_D (nM)	Mutant cϕ105 K_D(nM)
Oligonucleotides	37°C [a]		50°C → 37°C [a]	
O_R1	1.18	1.01	1.04	0.30
O_R2	0.28	0.26	0.13	0.24
O_R3	0.52	0.36	0.28	21.9
O_R4	0.59	0.64	0.28	24.2
O_R5	1.73	1.23	1.72	38.3
O_R6	0.26	0.52	0.30	26.1

[a] Real-time interaction was monitored at 37°C or after incubation at 50°C prior to analysis at 37°C. Results were obtained from two independent measurements, with protein concentrations ranging from 0.8 nM to 40 nM. These data successfully fitted into the 1:1 Langmuir binding model.

Table 3

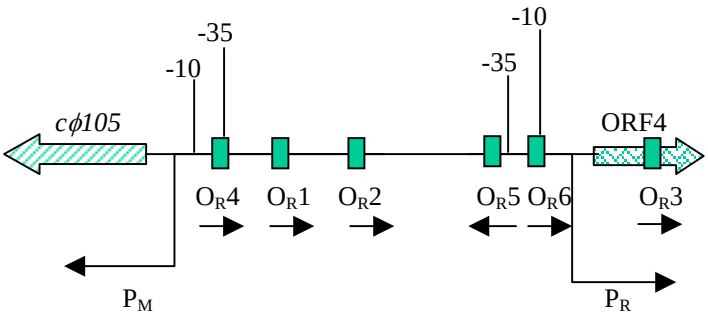
Strains	Genotype	Reference/source
Bacteria		
B. subtilis		
168	*trpC2*	ATCC 27370[a]
IA304	*trpC2 metB52 xin-1 SPβ(S)*	Leung *et al.*,1995
E. coli		
JM109	*e14-(McrA-) recA1 endA1 gyrA96 thi-1 hsdR17(rK- mK+) supE44 relA1 D(lac-proAB) [F' traD36 proAB lacI^{q}ZDM15]*	Stratagene
BL21(DE3)	*E. coli B F- dcm ompT hsdS(r_B- m_B-) gal l (DE3)*	Stratagene
Bacteriophage		
ϕ105MU331	*ind-125 cts-52Ω(lacZ'[ClaI]-ermC-cat'[NcoI])331Δ(DI:1t)*	Leung *et al.*,1995

[a] ATCC= American Type Culture Collection

Table 4

Primer Name	Sequence (5' to 3')	Orientation	Remarks
Gene-specific primers for PCR and RT-PCR			
ORF1BamS	AAGGATCCATGACTGTAGGGCAAAGAA	Sense	
ORF1EcoA	AGAATTCCTATTCTTGATCGTCATTTCT	Antisense	
ORF4BamS	AAGGATCCATGCTGGATGGGAAAAAGCT	Sense	
ORF4EcoA	AGAATTCTCATAAAGCCTGTCTTCTAC	Antisense	
16SF	TCCGCAATGGACGAAAGTCT	Sense	
16SA	ACGATCCGAAAACCTTCATCA	Antisense	
Primers for PCR DIG-DNA probe preparation			
ORF1PS	AACTGGCTGAAAAAGCCAATCT	Sense	
ORF1PA	AGTCCTTTCTTATTTCCTCCAT	Antisense	
ORF4PS	CACTTGAAACAGACAGAAATGG	Sense	
ORF4PA	TCCTCAACTACTTGTATTTCCG	Antisense	
Primers for site-directed PCR mutagenesis			
wtS	GGAACGTAAATTAACCCAAGTCGAA	Sense	Base mutation underlined
wtA	GTTGCACTTGGGTTAATTTACGTTCC	Antisense	Base mutation underlined
Sequences of TaqMan MGB probe applied			
FAM'ORF1	FAM-GCGCGTTGGGCAT[1]	Sense	FAM fluorescent reporter dye labelled at 5' end
FAM'ORF4	FAM-GATATCGAAAACGGCA[1]	Sense	FAM fluorescent reporter dye labelled at 5' end
VIC'16SrRNA	VIC-ACGGAGCAACGCCGCGTGA[1]	Sense	VIC fluorescent reporter dye labelled at 5' end
Oligonucleotides for surface plasmon resonance			
B'Or1S	5'(B)-aaaaaaGACGGAAATACAAGtatttt[2]	Sense	O_R1, Biotinylated at 5' end
B'Or2S	5'(B)-taaattGACGGAAATACAAGataaat[2]	Sense	O_R2, Biotinylated at 5' end
B'Or3S	5'(B)-aaaaatGACGGAAATACAAGtagttg[2]	Sense	O_R3, Biotinylated at 5' end
B'Or4S	5'(B)-ccgaatGTCGGAAATACAATactaaa[2]	Sense	O_R4, Biotinylated at 5' end
B'Or5S	5'(B)-aaaatgGACG-AAATTCAAGaaattt[2]	Sense	O_R5, Biotinylated at 5' end
B'Or6S	5'(B)-caaaatGTCGTGAATACCATacaatt[2]	Sense	O_R6. Biotinylated at 5' end
B'Non-boxS	5'(B)-caattgtacgtcaaagagatgaagca[2]	Sense	No consensus sequence, Biotinylated at 5' end

[1]All Taqman MGB probes are 3' labelled with non-fluorescent minor groove binder as the quencher.
[2]Complementary oligonucleotides without labelling were prepared for double-strand annealing

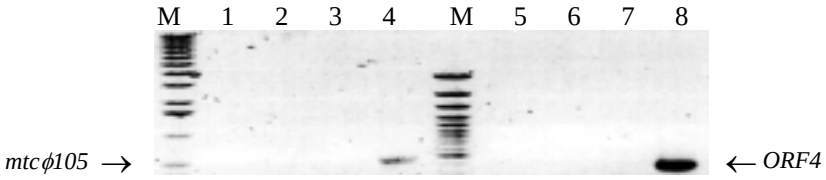

-35
-10
-10
-35
cϕ105
ORF4
OR4
OR1
OR2
OR5
OR6
OR3
PM
PR

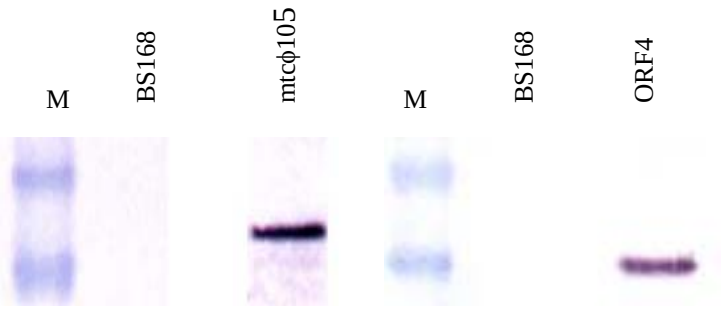
A

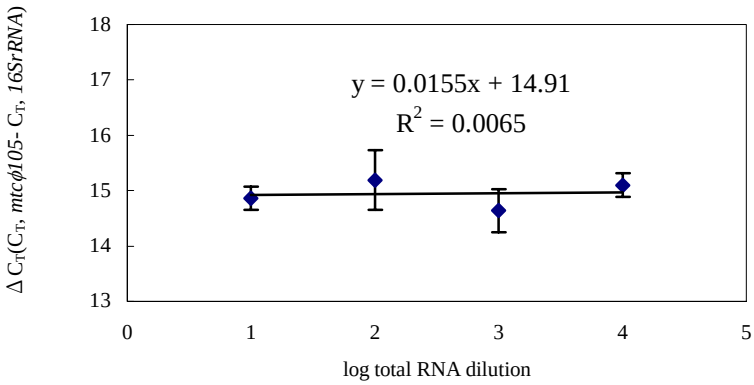
```
mtcϕ105    1 ----MTVGQRIKAIRKERKLIQVQLAEKANLSRSYLADIERDRYNPSLSTLEAVAGALGI
ORF4       1 MLDGKKLGALIKDKRKEKHLKQTEMAKALGMSRTYLSDIENGRYLPSTKTLSRIAILINL

mtcϕ105   57 QVSAIVGEETLIKEEQAEYNSKEEKDIAKRMEEIRKDLEKSDGLSFSGEPMSQEAVEFLM
ORF4      61 DLNVLKMTEIQVVEE--G----------------------GYDRAAGTCRRQAL----

mtcϕ105  117 EAMEHIVRQTQRINKKYTPKKYRNDDQE
ORF4         ----------------------------
```

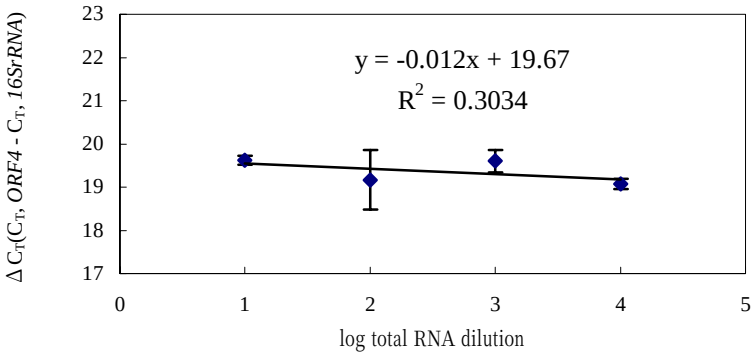
B

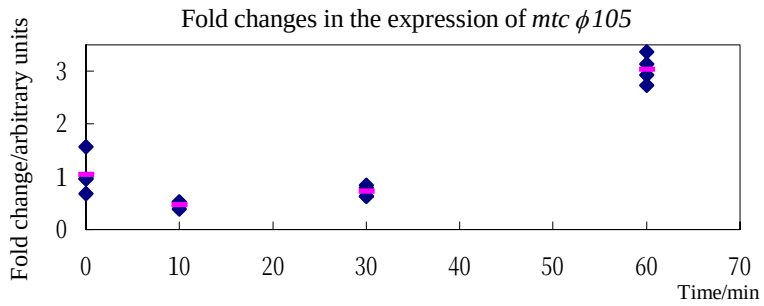
```
mtcϕ105     1QVQLAEKANLSRSYLADIER20
ORF4        1QTEMAKALGMSRTYLSDIEN20
P22         1QAALGKMVGVSNVAISQWER20
λCro        1QTKTAKDLGVYQSAINKAIH20
λCI         1QESVADKMGMGQSGVGALFN20
                      *

          --α2helix--βturn----α3helix--
```

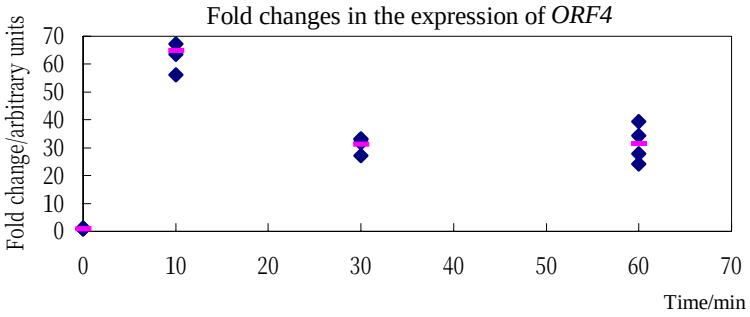

M 1 2 3 4 M 5 6 7 8
mtcϕ105 →
← ORF4


M
BS168
mtcϕ105
M
BS168
ORF4

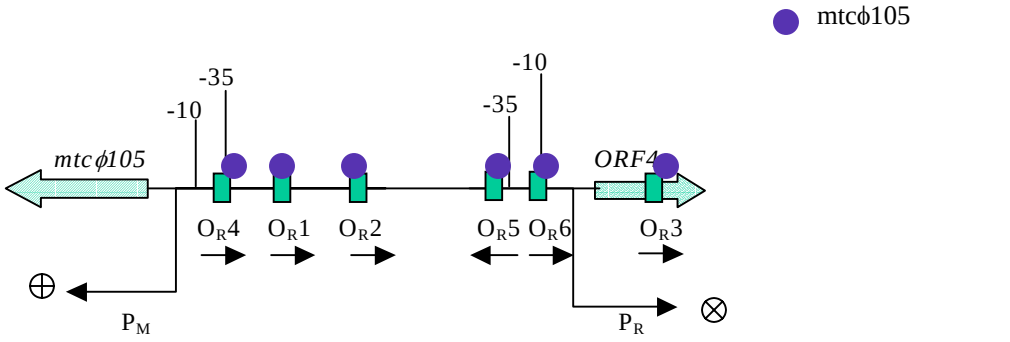

y = 0.0155x + 14.91
R2 = 0.0065
ΔCT(CT, mtcφ105- CT, 16SrRNA)
log total RNA dilution


y = -0.012x + 19.67
R2 = 0.3034
ΔCT(CT, ORF4 - CT, 16SrRNA)
log total RNA dilution


Fold changes in the expression of *mtc ϕ105*
Fold change/arbitrary units
0
1
2
3
0
10
20
30
40
50
60
70
Time/min


Fold changes in the expression of *ORF4*
Fold change/arbitrary units
0
10
20
30
40
50
60
70
0
10
20
30
40
50
60
70
Time/min

A Before thermo-induction



B After thermo-induction