LdrP, a cAMP receptor protein/FNR family transcriptional regulator, serves as a positive regulator for the light-inducible gene cluster in the megaplasmid of *Thermus thermophilus*

Hideaki Takano,1 Yoshihiro Agari,2 Kenta Hagiwara,1 Ren Watanabe,1 Ryuta Yamazaki,1 Teruhiko Beppu,1 Akeo Shinkai2,3 and Kenji Ueda1

1Life Science Research Center, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa, Kanagawa 252-0880, Japan
2RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan
3RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

LdrP (TT_P0055) (LitR-dependent regulatory protein) is one of the four cAMP receptor protein (CRP)/FNR family transcriptional regulators retained by the extremely thermophilic bacterium *Thermus thermophilus*. Previously, we reported that LdrP served as a positive regulator for the light-induced transcription of *crtB*, a carotenoid biosynthesis gene encoded on the megaplasmid of this organism. Here, we showed that LdrP also functions as an activator of the expression of genes clustered around the *crtB* gene under the control of LitR, an adenosyl B12-bound light-sensitive regulator. Transcriptome analysis revealed the existence of 19 LitR-dependent genes on the megaplasmid. S1 nuclease protection assay confirmed that the promoters preceding TT_P0044 (P44), TT_P0049 (P49) and TT_P0070 (P70) were activated upon illumination in the WT strain. An ldrP mutant lost the ability to activate P44, P49 and P70, whilst disruption of litR resulted in constitutive transcription from these promoters irrespective of illumination, indicating that these genes were photo-dependently regulated by LdrP and LitR. An in vitro transcription experiment demonstrated that LdrP directly activated mRNA synthesis from P44 and P70 by the *Thermus* RNA polymerase holocomplex. The present evidence indicated that LdrP was the positive regulator essential for the transcription of the *T. thermophilus* light-inducible cluster encoded on the megaplasmid.

INTRODUCTION

LitR (light-induced transcription regulator) is a MerR family transcriptional regulatory protein that contains an N-terminal helix–turn–helix domain and a C-terminal cobalamin-binding domain. Originally, we discovered that *Streptomyces coelicolor* A3(2), a Gram-positive soil bacterium renowned for its ability to produce a variety of secondary metabolites, produced carotenoid (Crt) in response to illumination. Our study revealed that LitR was involved in the light-induced transcription of an extracytoplasmic function sigma factor litS that directed the transcription of *crt* operons (Takano et al., 2005). Interestingly, LitR homologues are widely distributed in phylogenetically divergent genera of non-phototrophic bacteria (Takano et al., 2005, 2006a, b, 2011, Ortiz-Guerrero et al., 2011), frequently flanked by genes for Crt biosynthesis and DNA photolyase (Phr).

Previously, we revealed that a LitR homologue played a role as a central regulator of light-inducible Crt production in *Thermus thermophilus* HB27 (Takano et al., 2011). In this organism, the light-dependent transcriptional regulation of *crtB* (the phytoene synthase gene) is regulated primarily by two transcriptional regulators: LitR and TT_P0055. TT_P0055 encoded downstream of the *litR* gene [hereafter designated LdrP (LitR-dependent regulatory protein)] belongs to the cAMP receptor protein (CRP)/FNR family. Genetic and biochemical evidence indicated that LitR was bound by cobalamin and associated with the intergenic promoter region between *litR* and *crtB*, repressing the bidirectional transcription.

Abbreviations: CDD, Conserved Domain Database; CRP, cAMP receptor protein; GST, glutathione S-transferase; OHB12, hydroxocobalamin.

Original microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/), accession number GSE31463.

Two supplementary tables are available with the online Supplementary Material.
transcription of *litR* and *crtB*. It is probable that the inactivation of LitR caused by a photo-dependent mechanism induces the expression of the LdrP protein, which serves as a transcriptional activator for the *crtB* operon, resulting in expression of Crt biosynthesis and DNA repair system genes under light conditions (Takano et al., 2011).

Pérez-Marín et al. (2008) showed that CarH of *Mycococcus xanthus*, a homologue of LitR, was involved in light-inducible Crt production. Ortiz-Guerrero et al. (2011) reported that a chimeric protein, CTR2, which is composed of the N-terminal DNA-binding domain of CarH and the C-terminal cobalamin-binding domain of LitR of *T. thermophilus* HB8, was a coenzyme B12 (5'-deoxyadenosylocobalamin, AdoB12)-binding transcriptional regulator with light-sensitive DNA-binding activity. The photolysis of the AdoB12-bound CTR2 leads to the conversion of AdoB12 into hydroxocobalamin (OHB12), resulting in the loss of its DNA-binding activity. It was also reported that the full-length native protein TiCarH (a LitR homologue with a single-residue substitution) from *T. thermophilus* HB8 bound to its operator DNA in an AdoB12- and light-dependent manner (Ortiz-Guerrero et al., 2011, Diez et al., 2013).

We previously revealed that *litR* and *crtB* were under the regulation of LitR in *T. thermophilus* (Takano et al., 2011); however, it was not clear whether the transcription of the *crt* gene cluster as a whole and of other genes were regulated in a LitR-dependent manner. We expected that the identification of LitR-dependent genes would provide a new insight into the protection from light stress in this bacterium. In order to identify the LitR-dependent genes and elucidate the regulatory mechanism, we performed a transcriptome analysis and examined the regulatory mechanism. The evidence obtained in this study indicated that 19 genes encoded on the megaplasmid, including those in the *crt* gene cluster and its flanking genes, were under the regulation of LitR. In addition, LdrP, which is photo-dependently controlled by LitR, served as a transcriptional activator of the operons of these genes.

**METHODS**

**Bacterial strains, plasmids and culture media.** The WT strain of *T. thermophilus* used in this study was HB8 (Oshima, 1974), together with HB27 TH104 (proC) (pTT8) (Hoshino et al., 1994). *Escherichia coli* JM109 and BL21(DE3)pLysS (Takara-Shuzo) were used as hosts for DNA manipulation and protein expression, respectively. pUC19 (Takara-Shuzo) was used for general DNA manipulation. pT7Blue and pMD19 (Takara-Shuzo) were used for TA cloning of PCR-generated DNA fragments. pGEX-6P-2 (GE Healthcare Bio-Sciences) was used for the overexpression of *T. thermophilus* HB27 LitR and LdrP proteins in BL21(DE3)pLysS. Enzymes used for DNA manipulation were purchased from Takara-Shuzo. The conditions for the culture and genetic manipulation for *E. coli* and *Thermus* spp. were as described by Maniatis et al. (1982) and Koyama et al. (1986). *T. thermophilus* was grown at 60 °C in TM medium (containing (1–1): yeast extract (Difco), 2 g; polypeptone, 4 g; NaCl, 1 g; and Castenholz basal salt solution, 10 ml); all chemicals therein were purchased from Wako Pure Chemicals unless otherwise indicated. Castenholz basal salt solution (containing (1–1): nitritriacetic acid, 1 g; CaSO4.2H2O, 0.6 g; MgSO4.7H2O, 1 g; NaCl, 0.08 g; KNO3, 1.03 g; Na2HPO4, 1.11 g; aqueous 0.03 % FeCl3 solution, 10 ml; and Nitsch’s trace elements, 10 ml [pH 8.2]) and Nitsch’s trace elements [containing (1–1): H2SO4, 5 ml; MgSO4.5H2O, 2.2 g; ZnSO4.7H2O, 0.5 g; H3BO3, 0.5 g; CuSO4.0.016 g; Na2MoO4.2H2O, 0.025 g; and CoCl2.6H2O, 0.046 g] were used to prepare TM broth. E. coli was grown in Luria-Bertani (LB) medium (Maniatis et al., 1982) and 1.0–1.5 % agar (Kokusan) was added to prepare solid media. To enable the selection of transformants of *E. coli* and *T. thermophilus*, ampicillin, kanamycin and hygromycin B were added at 50 μg ml⁻¹.

**Gene disruption.** Kanamycin-resistant mutants of *T. thermophilus* HB8 were generated by the standard homologous recombination technique, using disruption plasmids. Each disruption plasmid contained a promoterless thermostable kanamycin resistance (*htk*) gene cassette to avoid polar effects (Hoseki et al., 1999). To construct the disruption plasmid for *litR* (TTHB100) of *T. thermophilus* HB8, two flanking fragments were amplified by PCR using the primer sets P01/P02 and P03/P04 (oligonucleotide primers used in this study are shown in Table S1, available in the online Supplementary Material) and cloned onto pUC19 by three-fragment ligation. Each resulting plasmid was digested with BamHI and ligated with a promoterless *htk* cassette amplified by PCR using the primer sets P05/P06 to generate the disruption plasmid. Disruption plasmids were linearized by digestion with Drai and introduced into *T. thermophilus* WT cells (Hashimoto et al., 2001). Subsequently, kanamycin-resistant mutants were screened and true recombination was verified by PCR using the appropriate primers.

**DNA microarray.** The *T. thermophilus* HB8 strain and the *litR* mutant was cultured at 60 °C for 8 h in TM broth under dark conditions. The crude RNA was extracted from each cell and after the cDNA was synthesized, it was fragmented and labelled with biotin-dideoxyUTP, as described previously (Agari et al., 2008). The 3'-terminally labelled cDNA was hybridized to a TTHB8401a520105F GeneChip (Affymetrix), and the array was then washed, stained and scanned as described previously (Agari et al., 2008). The raw intensities for three independent cultures of the WT strain and the *litR* mutant were each summarized to 2266 ORFs, using the GeneChip Operating Software, version 1.2 (Affymetrix). The datasets were then normalized through the following normalization steps using the Subio Platform, including shifting of low signals <1.0 to 1.0, log-based transformation of the data and global normalization [normalized to 75th percentile (third quartile)]. We excluded several genes with detection levels labelled 'Absent' (Pepper et al., 2007) from the data for three independent cultures of the WT strain and the *litR* mutant. The remaining data of 2061 ORFs were used for the following analysis. The *t*-test *P* values of the observed differences in the normalized intensities between the WT and *litR* mutants were calculated using the Subio Platform and then their false-discovery rates (*q* value) (Storey & Tibshirani, 2003) were calculated from these values using R software (http://www.R-project.org). The microarray data discussed in this study are accessible through GEO accession number GSE31463 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=phupzkcmkwjyipp&acc=GSE31463).

**S1 nuclease mapping.** S1 nuclease mapping was performed using a previously described method (Takano et al., 2011). The transcriptional activities of the promoters preceding *litH* (*LitHt*), *crtB* (*PcrtB*), TT_P0044 (P44), TT_P0049 (P49) and TT_P0070 (P70) of *T. thermophilus* HB8 were evaluated by S1 protection analysis. Hybridization probes were first generated by PCR using the primers P07/P08 (P44), P09/P10 (P49) and P11/P12 (P70) (see Table S1), and cloned onto pT7Blue by TA cloning. For the low- and high-resolution analysis of P44, P49 and P70, probes were prepared by PCR using the primer sets M13-7/18 (P44), M13-7/18 (P49) and M13-7/18 (P70) (primers labelled at their 5' end with [32P]ATP using T4 polynucleotide kinase are denoted with asterisks), respectively. Probes for *litR* and *crtB* were prepared as described previously (Takano et al., 2011). The primers M13-7/18 and...
M13-RV (M13 sequencing primers) were purchased from Takara-Shuzo. Radioactivity was detected by exposing dried gels to a Fuji imaging plate (Fuji Film) and images were scanned with a Typhoon 9410 image analyser (GE Healthcare). Marker 10 (pBR322/Mspl digest; Nippon Gene) labelled with [%-32P]ATP using T4 polynucleotide kinase was used as a standard to estimate the transcript sizes in the low-resolution assay. To determine the transcription start sites in the high-resolution analysis, Maxam–Gilbert sequencing ladders (A+G and C+T reactions) derived from the [%-32P]-labelled probe DNA were used as the reference. The quality of the RNA was verified via a control assay for sigA, encoding the major sigma factor (Nishiyama et al., 1999). The probe for sigA was amplified by PCR using the primers P13/P14* (see Table S1).

**In vitro run-off transcription.** The in vitro run-off transcriptional assay and the preparation of LdrP recombinant protein of *T. thermophilus* HB827 were performed according to methods described previously (Nishiyama et al., 1999; Shinkai et al., 2007; Takano et al., 2011). The DNA templates containing the transcriptional start sites of P44 were generated by PCR, using the primers P07/P08 for Template –186 (276 bp), P15/P08 for Template –50 (142 bp), P16/P08 for Template –51 (143 bp), P17/P08 for Template –52 (144 bp), P18/P08 for Template –53 (145 bp), P19/P08 for Template –54 (146 bp), P20/P08 for Template –55 (147 bp), P21/P08 for Template –56 (148 bp), and P22/P08 for Template –57 (149 bp). For P70, the following primer sets were used: P23/P12 for Template –42 (120 bp), P24/P12 for Template –47 (125 bp), P25/P12 for Template –52 (130 bp), P26/P12 for Template –57 (135 bp) and P27/P12 for Template –62 (140 bp) (see Table S1). The DNA templates (172 bp) containing mutated PcrIB were generated by PCR, using the primers P31/P30 for WT, P32/P30 for G to T substitution at position –50, P33/P30 for T to G at –49, P34/P30 for G to A at –48, P35/P30 for G to T at –47, P36/P30 for G to C at –46, P37/P30 for C to A at –44, P38/P30 for C to T at –43, P39/P30 for G to A at –41, P40/P30 for G to T at –39, P41/P30 for C to A at –38, and P42/P30 for A to C at –35. The DNA templates (157 bp) containing mutated P44 were generated by PCR, using the primers P43/P44 for WT, P45/P44 for A to T substitution at position –49, P46/P44 for T to G at –48, P47/P44 for G to A at –47, P48/P44 for T to G at –46, P49/P44 for G to T at –44, P50/P44 for C to A at –43, P51/P44 for C to A at –42, P52/P44 for G to A at –40, P53/P44 for G to T at –38, P54/P44 for C to A at –37 and P55/P44 for A to C at –34. A total of 0.5 pmol template DNA was mixed with 2 pmol commercial RNA polymerase holoenzyme of *E. coli* (AR Brown), 100 nmol ribonucleotides, including [%-32P]ATP, using T4 polynucleotide kinase was used to generate 2.5–5.0 ng [%-32P]-labelled probe DNA (10000–20000 c.p.m.) was mixed with 30–120 pmol recombinant LdrP prepared as above. The mixture was incubated at 55 °C for 30 min in 50 µl binding buffer (containing 10 mM Tris/HCl (pH 7.0), 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 10% (v/v) glycerol, 1 µg poly(dI–dC) and 50 µg BSA ml⁻¹) under light or dark conditions in a darkroom equipped with white light fluorescent lamps (Toshiba; 15 W). In the light condition, samples were illuminated at –2.4 µmol s⁻¹ m⁻² for 10 min at room temperature. Specific DNA–protein complexes were separated from free probe on non-denaturing polyacrylamide gel containing 6% acrylamide. The gels were dried and radioactive signals were detected by exposing dried gels to a Fuji imaging plate (Fuji Film). Images were scanned with a Typhoon 9410 image analyser (GE Healthcare).

**RESULTS**

**Identification of LitR-dependent genes in *T. thermophilus* HB8**

This study was performed using two strains, *T. thermophilus* HB8 and HB27, whose complete genome sequences have been determined (Brüggemann & Chen, 2006; Henne et al., 2004). Each genome consists of a 1.85 Mb chromosome and a large (0.26 Mb) plasmid, pT27T (Takayama et al., 2004). HB8, but not HB27, harbours two (9.32 and 81 kb) plasmids (Ohtani et al., 2012). The genome and megaplasmid sequences of HB8 and HB27 exhibit marked similarity to each other. Despite the availability of their genomic information, a protocol and DNA chip for the DNA microarray analysis had been established only for *T. thermophilus* HB8 (Agari et al., 2008) and not for the HB27 strain when we started this study. For this reason, we decided to use *T. thermophilus* HB8 for the identification of the LitR-dependent genes, even though we used the HB27 strain for our previous analysis of LitR and LdrP (Takano et al., 2011). We used *T. thermophilus* HB27 for further in vivo and in vitro analysis of LdrP- and LitR-mediated transcriptional regulation.

We first carried out a transcriptome analysis using a *T. thermophilus* HB8 GeneChip custom array to identify genes regulated by LitR. Of the 2266 genes analysed, we selected the down- and upregulated genes with q values <0.05 in the litR mutant compared with the WT under dark conditions. The analysis identified the LitR-dependent genes characterized by 2.0- to 27.5-fold higher transcriptional activity in the litR mutant than in the WT strain under dark conditions—many of these genes were located in the neighbourhood of litR (Table 1) (the other LitR-dependent genes are shown in Table S2). This result suggested that the expression of these genes was negatively regulated by LitR.

As shown in Fig. 1, the LitR-dependent genes clustered around *ldrP* form five putative operons: TTHB089–090–091, TTHB094–095–096, litR operon (TTHB097–098–*ldrP-litR*), *crtB*-phr–TTHB103–104–109–110 and TTHB112–113 (see also Table 1). TTHB089–090–091 had an operon structure and encoded a hypothetical...
**Table 1.** LitR-dependent genes located in the flanking region of the *litR* operon of *T. thermophilus*

<table>
<thead>
<tr>
<th>Coding sequence ID</th>
<th>Expression ΔlitR/WT dark (q value)*</th>
<th>Domain name</th>
<th>E value</th>
<th>Annotation for product†</th>
<th>Possible cellular role</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB8</td>
<td>HB27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTHB088</td>
<td>TT_P0043</td>
<td>cl00446</td>
<td>1.67e–14</td>
<td>Zinc-dependent hydrolase</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB089</td>
<td>TT_P0044</td>
<td>–</td>
<td>–</td>
<td>Hypothetical protein</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB090</td>
<td>TT_P0045</td>
<td>cl12011</td>
<td>6.63e–11</td>
<td>Methyltransferase domain</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB091</td>
<td>TT_P0046</td>
<td>Abhydrolase_6</td>
<td>3.3e–05</td>
<td>z/β Hydrolase fold</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB092</td>
<td>TT_P0047</td>
<td>cl14644</td>
<td>6.18e–13</td>
<td>Pyridine nucleotide disulphide oxidoreductase</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB094</td>
<td>TT_P0049</td>
<td>adh_short</td>
<td>3.2e–21</td>
<td>Putative short-chain oxidoreductase</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB095</td>
<td>TT_P0050</td>
<td>cl00801</td>
<td>7.88e–14</td>
<td>Flavin reductase-like domain</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB096</td>
<td>TT_P0051</td>
<td>cl01001</td>
<td>2.19e–10</td>
<td>Yce1-like domain</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB097</td>
<td>TT_P0052</td>
<td>–</td>
<td>–</td>
<td>Hypothetical protein</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB098</td>
<td>TT_P0054</td>
<td>cd05271</td>
<td>2.46e–45</td>
<td>NAD-dependent epimerase/ dehydratase family</td>
<td>Unknown</td>
</tr>
<tr>
<td>TTHB099</td>
<td>TT_P0055</td>
<td>cl00407</td>
<td>1.29e–06</td>
<td>CRP/FNR family regulator (LdrP)</td>
<td>Activator for light-inducible genes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cl00088</td>
<td>5.32e–05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTHB101</td>
<td>TT_P0057</td>
<td>cd00683</td>
<td>4.03e–64</td>
<td>Phytoene synthase (CrtB)</td>
<td>Crt biosynthesis</td>
</tr>
<tr>
<td>TTHB102</td>
<td>TT_P0058</td>
<td>cl04074</td>
<td>3.12e–67</td>
<td>DNA photolyase (Phr)</td>
<td>DNA repair</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cl02999</td>
<td>3.52e–25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTHB103</td>
<td>TT_P0059</td>
<td>cl12078</td>
<td>2.22e–26</td>
<td>Cytochrome P450 (Cyp175a1)</td>
<td>Crt biosynthesis</td>
</tr>
<tr>
<td>TTHB104</td>
<td>TT_P0060</td>
<td>cl11899</td>
<td>4.75e–09</td>
<td>Phytoene synthase-related protein</td>
<td>Crt biosynthesis</td>
</tr>
<tr>
<td>TTHB109</td>
<td>TT_P0066</td>
<td>cl14644</td>
<td>2.52e–13</td>
<td>Phytoene dehydrogenase</td>
<td>Crt biosynthesis</td>
</tr>
<tr>
<td>TTHB110</td>
<td>TT_P0067</td>
<td>cd02811</td>
<td>5.24e–82</td>
<td>Isopentenyl-diphosphate δ-isomerase (IDI-2)</td>
<td>Crt biosynthesis</td>
</tr>
<tr>
<td>TTHB112</td>
<td>TT_P0069</td>
<td>cl00733</td>
<td>2.37e–29</td>
<td>Conserved hypothetical protein</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cl01284</td>
<td>1.38e–22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTHB113</td>
<td>TT_P0070</td>
<td>–</td>
<td>–</td>
<td>Hypothetical protein</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

*Intensity of the ΔlitR strain normalized to that of the WT strain cultured under dark conditions.

†A product that was not found in the Conserved Domain Database (CDD) search was annotated as a ‘Hypothetical protein’. The possible cellular role of the gene product is shown together with the name of the most closely related domain, E value and annotation of the gene product, which were obtained from CDD or Pfam searches.
protein (TTHB089), a methyltransferase (TTHB090) and the α/β-hydrolase fold protein (TTHB091), respectively. The litR operon was flanked by the TTHB094–095–096 operon, which contained the coding sequences for short-chain dehydrogenases/reductases (TTHB095) and YceI-like protein from E. coli (TTHB096). TTHB112–113 was located downstream of the crt biosynthesis gene cluster as a divergent operon and encoded proteins with no homology to proteins of known function.

We suppose that the six coding sequences in the ldrP region (TTHB093, TTHB105–106–107–108 and TTHB111) are also litR-dependent although they were not included in the list (Table 1). TTHB105–106–107–108 and TTHB111 appear to be included in crtB operon because each intergenic region is very short or even does not exist. TTHB105 and TTHB106 are hypothesized to be Crt biosynthesis enzymes. It is not clear why the analysis did not show the low quality of the corresponding probes on the DNA chip.

Transcriptional analysis by S1 nuclease mapping in *Thermus thermophilus* HB27

We performed a high-resolution analysis of S1 nuclease mapping to determine the transcriptional start sites of TT_P0044, TT_P0049 and TT_P0070 of the HB27 strain, which are orthologous to the first genes of the operons identified by transcriptome analysis in the HB8 strain (Fig. 1, Table 1). As shown in Fig. 2(a), the transcriptional start sites for TT_P0044 and TT_P0070 were 15 and 2 bp upstream from the ATG and GTG translational initiation codons of TT_P0044 and TT_P0070, respectively. The transcription of TT_P0049 was started from the A of ATG of the translation initiation codon of TT_P0049. Based on the position of the transcriptional start sites, the translation initiation codons of TT_P0049 and TT_P0070 were located 207 and 9 bp, respectively, downstream from the respective positions assigned in the genome sequence database (Fig. 2a). The evidence suggested that both TT_P0049 and TT_P0070 were translated by a leaderless mechanism (Moll et al., 2002). We previously observed a similar situation with respect to the transcriptional start site of litR and crtB in this organism (Takano et al., 2011). These promoters were designated P44 (TT_P0044), P49 (TT_P0049) and P70 (TT_P0070). The putative −35 and −10 sequences were CCCATT...TAGGCT (P44), CAAACA...TAAACGT (P49) and GGTACA...TAGAAT (P70); the −10 hexamer of these promoters was similar to the consensus of σ^70-dependent promoters of *Thermus thermophilus* (TGTACA...TANCCCT) described by Sevostyanova et al. (2007) (Fig. 3b).

The transcription level of the light-inducible promoters P44, P49 and P70 was analysed by low-resolution S1 nuclease
mapping with total RNA prepared from cells grown in liquid culture under dark or light conditions (Fig. 2b). The photoinducible promoters PcrtB and PlitR (Takano et al., 2011) were used as positive controls. The sigA gene encoding σ70 was used as a constitutively expressed internal control to measure the amount and purity of the purified RNA. In the WT strain, the activity of P44, P49 and P70 were induced by light similar to PcrtB and PlitR, whilst the activity of PsigA was not significantly affected by illumination. In the ldrP mutant, transcription of P44, P49 and P70 was abolished under both light and dark conditions, whilst PlitR activity was light-dependent at a level comparable with that of the WT strain. In the litR mutant, P44, P49 and P70 were all active irrespective of illumination. The result for the litR mutant is in good agreement with the data of DNA microarray analysis that the transcription of PcrtB, P44, P49 and P70 was upregulated by the deletion of litR under dark conditions. These results indicated that LdrP protein served as an activator essential for the transcription initiated from the promoters P44, P49 and P70, and that the light-dependent expression of LdrP due to the function of LitR was fundamental for the light-induced transcription from the three promoters.

In vitro run-off transcription

To clarify whether LdrP directly activated transcription from the light-inducible promoters by RNA polymerase holoenzyme, we performed an in vitro run-off transcriptional assay as previously carried out with respect to PcrtB (Fig. 3a). An RNA polymerase holoenzyme complex and a recombinant protein of LdrP successfully generated transcripts whose length corresponded to those of P44 (92 bases) and P70 (73 and 74 bases), respectively. The transcription from P70 occurred at a low level even in the absence of LdrP. This result was consistent with that from S1 nuclease mapping (Fig. 2b): the measurement of signal intensity at 40 h indicated that the fold change of transcription (light/dark) at P70 (8.6) was lower than at P44 (53.3) and P49 (12.2). The low-level transcription at P70 was also observed in the ldrP mutant irrespective of illumination (Fig. 2b). These results indicated that the RNA polymerase holoenzyme inefficiently generated the transcript from P70 without LdrP.

We performed the same transcriptional assay with trimmed templates: -50 to -57 for P44 and -42 to -62 for P70 (the number indicates the position of the 5’ end of the antisense
Fig. 3. In vitro run-off transcription assay. (a) T. thermophilus RNA polymerase holoenzyme was added to the reaction mixture containing the promoter DNA fragments in the presence (+) or absence (−) of 2 pmol LdrP recombinant protein to generate transcripts specific for P44 (TT_P0044) and P70 (TT_P0070). Closed triangles denote the positions of the expected transcripts for P44 and P70. The template number indicates the position of the 5' end of each template with respect to the transcriptional start site.
strand of the template DNA when the transcriptional start site is numbered at +1). The assay showed that LdrP-dependent transcription occurred with templates 253 to 257 for P44 and 251 and 261 for P70 (Fig. 3a). When template 242 and 247 for P70 was used, the addition of LdrP partially inhibited the transcription initiated from each template compared with the control lane. This weak inhibition by LdrP may have been due to its incomplete binding to the template. These results indicated that the LdrP-dependent transcriptional initiation at P44 and P70 required the region encompassing positions 253 to +92 and 252 to +74, respectively (Fig. 3a, b).

However, RNA polymerase holoenzyme and LdrP did not generate a specific transcript on the probe DNA for the P49 region in vitro, despite the fact that the in vivo transcriptional analysis showed the abolishment of light-inducible transcription of P49 by the ldrP mutation and that a similar sequence to the LdrP consensus sequence was found in P49 (Fig. 3b). This result suggested that transcriptional initiation at P49 required an additional transcription factor. In all promoters tested in this assay, the addition of cAMP did not affect the efficiency of the transcriptional activation by LdrP (data not shown).

Fig. 3(b) shows a predicted consensus sequence, 5′-(A/G)(T/C)(T/G)(T/G)NGCCN(T/G)NG(G/C)NNA-3′, for the LdrP recognition site, which was proposed on the basis of the nucleotide sequence alignment of the three LdrP-dependent promoters, PcrtB, P44 and P70. The sequence for P49 was not considered for the prediction due to the aforementioned possibility that other transcription factors may be involved in its regulation. In E. coli, the CRP–cAMP complex binds to

---

**Fig. 4.** Gel-shift assay of AdoB12–LitR. The amounts of AdoB12–LitR used were 0 (lane 1), 20 (lane 2), 40 (lane 3), 80 (lane 4) and 160 pmol (lane 5). The various amounts of AdoB12–LitR were mixed with the probes for (a) PlitR, and (b) P44, P49, P70 and PsigA, and applied to a non-denaturing polyacrylamide gel. The AdoB12–LitR protein was incubated in the dark (10 min) or irradiated with white light (10 min) prior to the addition of 32P-labelled probes.
the consensus sequence 5′-AAATGTGATCTAGATCACA-3′ (Ebright et al., 1989). The CRP-binding sites of *E. coli* are grouped into two classes based on the position of the CRP-binding site: class I and class II are centred at positions −61.5 and −41.5, respectively, with respect to the transcription start site (Busby & Ebright, 1999; Lawson et al., 2004). The predicted LdpP consensus sequence did not exhibit a distinctive similarity to that of *E. coli* CRP, although the position of the LdpP consensus [−35 to −50 (*PcrB*), −34 to −49 (*P44*) and −37 to −52 (*P70*)] was similar to that of *E. coli* CRP class II. We could not confirm that the proposed consensus was the LdpP-binding sequence because direct binding of LdpP to P44, P70 and P49 was not observed (see Discussion). To examine whether the predicted consensus was involved in LdpP transcriptional activation, we performed an *in vitro* run-off assay using the mutated *PcrB* and P44 promoters as templates, which contained mutations at conserved nucleotides around the −40 regions. As shown in Fig. 3(c), similar nucleotide mutations caused low-level LdpP-mediated transcriptional activation of both promoters; mutation of the underlined nucleotides in 5′-(A/G)(T/C)(T/G)TNGCCN(T/G)NG(G/C)NNA-3′ reduced mRNA synthesis from *PcrB* and P44 promoters. These results suggested that LdpP recognized a part of the predicted LdpP consensus sequence to regulate the targeted promoters.

**In vitro DNA-binding activity of AdoB12-bound LitR (AdoB12–LitR)**

To examine whether the transcription of P44, P49 and P70 was regulated only by LdpP or LitR was also involved directly in the regulation, we prepared AdoB12–LitR of *T. thermophilus* HB27 and performed a gel-shift assay. As shown in Fig. 4(a), AdoB12–LitR showed specific binding to the intergenic promoter region of *litR* and *crtB* under dark conditions, whilst AdoB12–LitR exposed to white light for 10 min did not bind to the promoter region. In contrast, AdoB12–LitR did not bind to the promoter regions of P44, P49, P70 and PsigA (control) under dark conditions (Fig. 4b). These results indicated that AdoB12–LitR of *T. thermophilus* HB27 was also a light-sensitive DNA-binding protein, as reported in *TtCarH* of HB8 (Ortiz-Guerrero et al., 2011), and suggested that only LdpP affected the transcription of P44, P49 and P70.

**DISCUSSION**

Our transcriptome analysis revealed the existence of 19 LitR-dependent genes encoded on the megaplasmid of *T. thermophilus*. A correlation between many of the newly identified genes and light has not been reported in non-phototrophic bacteria, which indicates that they are novel light-inducible genes. Our data indicate that LdpP is a positive regulator essential for transcriptional initiation of the gene cluster and that LitR confers photo-dependency to LdpP-dependent transcription. The finding that 45 out of 251 genes encoded on the megaplasmid, including cobalamin biosynthesis genes, are associated with light indicates that the megaplasmid of *T. thermophilus* plays an important role in the protection of the cell from light stress.

Fig. 5 shows the current working model for the transcriptional control of the light-inducible gene cluster by LitR and LdpP. This model is supported by the results of our previous study (Takano et al., 2011) and the present study, and other recent findings (Ortiz-Guerrero et al., 2011). The evidence strongly reinforces our view that the members of the CarH/LitR family of proteins [we previously described the family to be CarA/
The light-inducible genes found in this study play a role in protecting cells from \(^1\)O\(_2\).

The transcriptome analysis also identified the genes showing a lower transcriptional level in the litR mutant than that in the WT (Table S2). The downregulated genes existed on the chromosome and included functionally unknown genes forming an operon-like structure (TTHA0771–0774). This result suggests that LitR acts as a positive regulator to control the expression of these chromosomally encoded genes. Whether LitR binds directly to the promoter region of these genes is not known, and because the difference in the expression level of these downregulated genes between the litR mutant and the WT strain is relatively small, the direct involvement of LitR in their regulation is not clear.

The light-inducible gene cluster, including LitR and LdrP, is also conserved in some species of *Deinococcus-Thermus* extremophiles (Fig. 1). Many of these genes, including the *litR*-ldrP operon, are distributed in the chromosome of *Meiothermus ruber* (Tindall et al., 2010), *Meiothermus silvanus* (Sikorski et al., 2010) and *Deinococcus maricopenis* (http://www.genome.ad.jp) – species that were isolated either from hot springs or from soil. This gene conservation suggests that the LitR-dependent genetic regulation also occurs in these bacteria. Meanwhile, the other genome-sequenced species belonging to the phylum *Deinococcus-Thermus*, including *Deinococcus radiodurans*, *Deinococcus geothermarum*, *Oceanithermus profundus* and *Marinithermus hydrothermalis*, do not retain the *litR* and the neighbouring region (http://www.genome.ad.jp). *O. profundus* and *Marinithermus hydrothermalis* were isolated from deep-sea hydrothermal vents, where the environment is dark. This implies that the gene cluster has been acquired due to an advantage specifically developed within some group of the bacterial phylum, which suggests that this regulatory system is more required for their survival in extreme environments.

The biochemical characterization of individual proteins encoded by the light-inducible gene cluster presented here will provide the necessary information to understand the biological responses and adaptation to illumination in this group of bacteria.

**ACKNOWLEDGEMENTS**

This study was supported by a Grant-in-Aid for the Encouragement of Young Scientists (23780093) to H. T. and Scientific Research (C) (22510208) to A.S. This study was also supported by the High-Tech Research Center Project of the Ministry of Education, Culture, Sports, Science and Technology, Japan, the Noda Institute for Scientific Research, the Foundation NAGASE Science Technology Development and Charitable Trust Araki Medical and Biochemical Research Memorial Fund.

**REFERENCES**


Edited by: Y. Ohnishi