Sequence-specific Binding Sites in the Taka-amylase A G2 Promoter for the CreA Repressor Mediating Carbon Catabolite Repression

Masashi Kato,¹ Kotaro Sekine, and Norihiro Tsukagoshi

Department of Applied Biological Sciences, Faculty of Agriculture, Nagoya University, Nagoya 464-01, Japan

Received March 22, 1996

The N-terminal part of the CreA protein encompassing two zinc fingers was expressed in Escherichia coli as a fusion protein with the maltose binding protein (MalE) of E. coli. Our results show that CreA binds to the promoter of the Taa-G2 gene encoding Taka-amylase A of Aspergillus oryzae. DNase I footprinting experiments showed that CreA bound to three sites with high affinity and to one site with low affinity within the first 401-bp region upstream of the transcription initiation site. All of the sites contained sequences related to the CreA consensus binding site (5'-SYGGRG-3'), and are suggested to participate in repression of the Taa-G2 gene in response to glucose.

Key words: Aspergillus; Taka-amylase A; DNA binding protein; catabolite repression; CreA protein

Aspergillus oryzae Taka-amylase A (EC 3.2.1.1, α-1,4-glucan-4-glucanohydrolase), TAA, catalyzes the endoamylolytic degradation of starch. Genomic and complementary DNAs of Taa were cloned and sequenced.¹⁻⁴ Based on in vitro analysis of Taa transcripts using isolated nuclei active in RNA synthesis, we have shown that both the induction of Taa by starch and the repression by glucose are regulated at the transcriptional level and that transcription factors mediating specific induction and repression are also involved in this regulation.⁵⁻⁸ In addition to these gene-specific controls, the transcription seems to be sensitive to a global regulation called carbon catabolite repression.

Carbon catabolite repression is a wide-domain regulatory system controlling the response of cells to the availability of glucose.⁹ It affects the expression of a large number of genes involved in the use of less favored carbon sources. Among eukaryotic systems, the creA gene of Aspergillus nidulans is one of the best-characterized examples of a global regulatory gene mediating carbon catabolite repression. The creA gene has been cloned and sequenced¹⁰⁻¹³ and shown to encode a protein comprising two zinc fingers that are very similar to those in the yeast transcriptional regulator, Mig1.¹⁴ CreA is the major repressor in A. nidulans.¹⁵ In the ethanol regulon as well as the proline degradation gene cluster of A. nidulans, CreA has been shown to bind in vitro to a consensus sequence, 5'-SYGGRG-3'.¹⁵⁻¹⁶ To discover one of the regulatory mechanisms for Taa gene expression, we have identified the in vitro CreA binding site in the Taa gene promoter region by using the MalE-CreA fusion protein. DNase I footprinting analysis showed that CreA bound to three sites with high affinity and to one site with low affinity within the first 401-bp upstream of the transcription initiation site. All of the sites were found to contain sequences related to the CreA consensus binding site as reported previously by several groups.¹⁰⁻¹³

Materials and Methods

Materials. Restriction endonucleases and other modifying enzymes were obtained from Toyobo, Inc. The expression vector, pMAL-c2,¹²¹ and the amylose resin were purchased from New England Biolabs.

¹ To whom all correspondence should be addressed.

Bacterial strains. The E. coli strain used was XL1-blue¹²² for DNA manipulation and expression of the MalE-CreA fusion protein.

Expression and purification of the MalE-CreA fusion protein. To overproduce the MalE-CreA fusion protein, an expression plasmid, pMAL-CreA, was constructed as follows: a 618-bp DNA fragment corresponding to the N-terminal portion of CreA extending from Asn-17 to Ser-223 was amplified from Aspergillus nidulans chromosomal DNA with the primers P1 and P2 (Fig. 1).

A.

B.

Fig. 1. Construction of an Expression Plasmid pMAL-CreA.

A: Open boxes represent the gene encoding the maltose binding protein derived from the expression vector pMAL-c2. Striped boxes represent the genes corresponding to the intact CreA protein (CreA) and an N-terminal portion of the CreA protein (CreA). Pₐₚ denotes the αα-promoter. Small arrows indicate the DNA primers, P₁ and P₂, which were used for amplification of the CreA gene. B: Nucleotide and deduced amino acid sequences at the malE creA fusion junction region in pMAL CreA. The numbers shown under the amino acid sequence indicate the amino acid residue numbers of the intact CreA.
by the polymerase chain reaction with oligonucleotide primers, 5'-dAACAACGCAGCTCATCCT-3' (P1) and 5'-dCTTTCTACGCTCT- GACCTG-3' (P2). An expression vector, pMAL-c2, was digested with Xbal, treated with the Klenow fragment of DNA polymerase I, and ligated to the PCR product described above. The resultant plasmid was then digested with BamHI, treated with mung bean nuclease, and self-ligated to construct pMAL-CreA (Fig. 1A). Figure 1B shows the nucleotide and deduced amino acid sequences at the malE-creA junction region in pMAL-CreA. E. coli XL1-blue carrying the plasmid pMAL-CreA was used as the source of the MalE-CreA fusion protein.

E. coli cells grown in 100 ml of L-broth containing 1 mm IPTG and 100 μM ZnCl₂ were suspended in 2 ml of buffer W (20 mm Tris- HCl (pH 7.4), 200 mm NaCl, 10 mm β-mercaptoethanol) and disrupted by a sonicator to obtain crude cell extracts. The crude cell extracts were absorbed to a 1 ml amylose column at 4°C. The column was washed with 10 ml of buffer W, followed by elution of the fusion protein with buffer W containing 10 mm maltose.

**Gel retardation assay.** The gel retardation assay was done with non-labeled DNAs as described previously. The 326-bp XbaI-XbaI DNA fragment containing the Taa-G2 promoter region was digested with Sau96I, Drai, and PstI. Resultant fragments were assayed.

**DNase I footprinting.** DNase I footprinting was described as noted. The 326-bp XbaI-XbaI DNA fragment used for the gel retardation assay was cloned at the XbaI site of pUC19 in two directions to obtain pUC119-taaG2p and pUC119-taaG2pR. A unique HindIII site was upstream from the transcription start site in pUC119-taaG2p, while the same HindIII site was downstream from the transcription start site in pUC119-taaG2pR. HindIII-EcoRI DNA fragments generated from pUC119-taaG2p and pUC119-taaG2pR were radio-labeled at the HindIII site with T7 DNA polymerase.

**Other methods.** SDS-polyacrylamide gel electrophoresis was done by the method of Laemmli. Nucleotides were sequenced by the dideoxy chain termination method of Sanger et al. with a DNA sequencer.

**Results and Discussion**

**Isolation of the MalE-CreA fusion protein.**

Maltose-binding proteins could be isolated in a single chromatography step, since they bind with high affinity to a cross-linked amylose matrix and are released from the matrix with maltose. The N-terminal part of the CreA protein encompassing the two zinc fingers and an alanine-rich region (residues from Asn-17 to Ser-223) was expressed under the control of the tac promoter in E. coli as a fusion protein with maltose binding protein (MalE) as described in Materials and Methods. In the presence of iodopropyl-β-D-thiogalactoside (IPTG), several proteins were over-produced (Fig. 2, lane 1). The MalE-CreA fusion protein was then isolated from the cell extracts of E. coli, using cross-linked amylose affinity column chromatography. SDS-PAGE analysis of the maltose-eluted fusion protein showed two major bands (Fig. 2, lane 4). The apparent molecular mass of the top-most band on SDS-polyacrylamide gel was 60 kDa, corresponding to that predicted from the sequence of the fusion protein. The lower band is presumed to be a degradation product of the fusion protein. The addition of Zn²⁺ ion during growth produced active CreA fusion protein in E. coli because zinc ion stabilizes the zinc finger structure of CreA protein (data not shown).

**Mapping the CreA binding sites in the Taa-G2 promoter region**

A 326-bp DNA fragment (−401 to −76, referring the transcription start site as +1) has been shown to contain at least two binding sites for two transcription factors: a CCAAT box for AnCP (A. nidulans CCAAT binding protein) and a 25-bp region just upstream of the CCAAT box for AnNP (A. nidulans nuclear protein). This fragment appears to be essential in the regulation of the Taa-G2 gene expression. Furthermore, sequence elements conferring starch inducibility in vivo were investigated by constructing 5′′ serial deletions of the Taa-G2 gene. The first 362-bp region upstream of the transcription initiation site (−362 to −1) has been shown to be sufficient for starch-dependent ex-
pression and glucose-dependent repression.\textsuperscript{5) A restriction map of this region is shown in Fig. 3A.}

To assess the DNA-binding regions of the MalE–CreA fusion protein within the 362-bp DNA fragment, five DNA fragments designated a to e were generated by digesting the 326-bp DNA fragment with Sau96I, DraI, and PstI and assayed for gel retardation using nonlabeled probe DNAs as described by Mizuno et al.\textsuperscript{14)} (Fig. 3). Three DNA bands corresponding to a, d, and e apparently disappeared under the conditions used for gel electrophoresis, since these DNA fragments formed complexes with the CreA fusion protein, which tend to smear. Although this method may not be accurate, it is simple and effective to roughly estimate protein binding affinities among several DNA fragments.\textsuperscript{14)}

Judging from the intensity of the DNA bands, CreA protein bound to the three binding sites with almost the same affinity. All these DNA fragments contain the consensus sequences for CreA binding; 5'-CTGGGGG-3' (−361 to −366), 5'-CCCGGGG-3' (−145 to −150), and 5'-CGCGGGG-3' (−90 to −95) were found in the fragments a, d, and e, respectively.

All these data also indicate that the MalE–CreA fusion protein is functional in DNA binding as in the case of the GST–CreA fusion protein reported by Kulkburg \textit{et al.} and used by several groups.\textsuperscript{10,11,18}

\textbf{DNase I footprinting analysis of the Taa-G2 promoter region}

To identify the DNA-binding sequences of the MalE–CreA fusion protein accurately in the Taa promoter region, we did DNase I footprinting analysis as follows: the non-coding strand of a 326-bp DNA fragment containing the \textit{Taa-G2} promoter was end-labeled, incubated with various amounts of purified MalE–CreA protein, and partially digested with DNase I. Two strongly protected sites were observed (Fig. 4, Panels I and II). The region from −372 to −356, denoted as site “A”, was within the DNA fragment “a” used in the gel retardation assay (Fig. 4I). The other region approximately from −150 to −140, was denoted as site “D”, and was within the fragment “d”, although the boundaries of the region protected were not as clear as those of site “A” (Fig. 4II). This ambiguity of the boundaries seems to be caused by the distance from the labeled end. The two binding sites were found to contain the consensus sequence for CreA-binding, 5'-SYGGGR, proposed by Cubero and Scazzocchio.\textsuperscript{11)} At the highest concentration of CreA fusion protein, the region from −204 to −220, denoted as site “C” was weakly protected (Fig. 4II). This is consistent with the findings obtained from the gel retardation assay, i.e., the MalE–CreA fusion protein bound weakly to the fragment “c” containing site “C”, compared to other fragments (Fig. 3B). However, site “C” contained the exact consensus sequence, 5'-CTGGAG-3’. This clearly indicates context dependence for CreA binding as discussed by Cubero and Scazzocchio.\textsuperscript{11)} This is also the case with MIG1, which binds to SYGGGR.\textsuperscript{9,19)}

The site within the region from −122 to −76, corresponding to the fragment e (Fig. 3B), was too far from the labeled end to detect any protection from DNase I (Fig. 4II). Therefore, the DNA fragment labeled at the opposite strand was used to show the protected site within the region (Fig. 4III). A strongly protected site denoted as site “E” was observed in the region from −97 to −86. The affinity of site “E” for the MalE–CreA protein is almost the same as that of site “A” and “D”. This is also consistent with the result from the gel retardation assay. Figure 5 summarizes the nucleotide sequences protected by the MalE–CreA protein from DNase I digestion in the \textit{Taa-G2}

\textbf{Fig. 4.} DNase I Footprinting of the Taa-G2 Promoter Region with the MalE-Crea Fusion Protein.

The HindIII–EcoRI fragment of pUC19-taaG2p (panels I and II) and that of pUC19-taaG2pR (panel III) were incubated with various amounts of the MalE–CreA protein. The amounts of the MalE-Crea protein used in each lane were as follows: none (lane 1), 0.4 μg (lane 2), and 2 μg (lane 3). The numbers on the left are the nucleotide position relative to the transcription start site (+1). The regions protected strongly and Weakly with the MalE-Crea fusion protein are indicated by thick solid lines (A, D, and E) and a thick dotted line (C), respectively.

\textbf{Fig. 5.} Binding Sites for the MalE-Crea Fusion Protein in the Taa-G2 Promoter Region.

A: A DNA fragment carrying the promoter region of the Taa-G2 is schematically shown. The closed boxes represent in vitro CreA-binding sites (A, B, D, and E). The DNA fragments used in the gel retardation assay are also indicated (a-e). B: The regions protected by the MalE-Crea fusion protein from DNase I digestion are summarized on the basis of the results in Fig. 4. The strongly and weakly protected regions are indicated by thick solid and dotted lines, respectively. The fragments used in the gel retardation assay are indicated by arrows (a-e). A CAAT box is also indicated.
promoter. All these findings strongly indicate that CreA protein is involved in the repression of the Tau-G2 gene in response to glucose.

In this work, we demonstrated that CreA specifically binds \textit{in vitro} to three different sites in the promoter region of the \textit{Taa-G2} gene. However, we cannot conclude that all or any of these sites are essential \textit{in vivo} to repress the \textit{Taa-G2} gene. Previously, we reported that the first 362-bp region upstream of the transcription initiation site is sufficient for glucose-dependent repression. This clearly indicates that the CreA binding site "A" is not essential in the glucose-dependent repression of the \textit{Taa-G2} gene. Therefore, the other two high affinity sites, "D" and "E", should participate in the CreA mediating carbon catabolite repression. Using mutants with mutations in the CreA binding sequences, especially "D" and "E", \textit{in vivo} analysis of the expression properties of the \textit{Taa-G2} gene together with \textit{in vitro} analysis of the CreA binding affinity for those mutated sequences should shed light on the catabolite repression mechanism for \textit{Taa-G2} gene. Further experiments are currently underway in our laboratory.

References