JOURNAL OF BACTERIOLOGY, Feb. 1997, p. 684–688 Vol. 179, No. 3
0021-9193/97/$04.00+0


# Binding of a Novel Host Factor to the pT181 Plasmid Replication Enhancer

QING ZHANG, SELMA SOARES DE OLIVEIRA,† ROBERTO COLANGELI, AND MARIA LAURA GENNARO*

*Public Health Research Institute, New York, New York 10016*

Received 6 September 1996/Accepted 17 November 1996

**Replication enhancers are *cis*-acting genetic elements that stimulate the activity of origins of DNA replication. The enhancer found in plasmid pT181 of *Staphylococcus aureus*, called *cmp*, functions at a distance of 1 kb from the origin of DNA replication to stimulate the interaction between the replication initiation protein and the origin. DNA encoding *cmp*-binding activity was isolated by screening an expression library of *S. aureus* DNA in *Escherichia coli*, and a novel gene, designated *cbf1*, was identified. The *cbf1* locus codes for a polypeptide of 313 amino acid residues (*cmp*-binding factor 1 [CBF1]; $M_r$ = 35,778). In its COOH-terminal region, the protein sequence contains the helix-turn-helix motif common to many DNA binding proteins that usually bend DNA. The specificity of CBF1 binding for *cmp* was demonstrated by affinity chromatography using *cmp* DNA and by competition binding studies. DNase I footprinting analysis of the CBF1-*cmp* complexes revealed DNase I-hypersensitive sites in phase with the helical periodicity of DNA, implying that CBF1 increases distortion of the intrinsically bent *cmp* DNA.**

Replication enhancers are *cis*-acting genetic elements that stimulate the function of an origin of replication (*ori*) from a physically separate, often distant site on the same DNA molecule (5, 6, 14, 17). Our work has identified the replication enhancer in plasmid pT181, *cmp*, a 100-bp genetic element located 1 kb from *ori* (4, 9). Replication enhancement by *cmp* was first demonstrated in plasmid incompatibility studies, where deletion of *cmp* drastically lowered the ability of pT181 to coexist with an incompatible $cmp^+$ plasmid (4, 5). Further genetic studies (3) showed that second-site suppressor mutations of $Cmp^-$ map in *repC* and produce mutant RepC proteins that fail to discriminate between $cmp^+$ and *cmp* mutant plasmids. Taken together, plasmid incompatibility and suppression studies show that *cmp* strengthens the interaction of the linked origin with RepC.

A question crucial to our understanding of replication enhancement by *cmp* is whether *cmp* acts as a protein binding site. The 100-bp *cmp* sequence is characterized by pentameric sequences sharing a loose consensus sequence and phased oligo(dT) tracts that confer an overall bent conformation to the *cmp* DNA (5, 9). Intrinsically bent DNA and the pentameric repeats suggest that *cmp* acts as a protein binding site, where the degree of curvature in the DNA is often enhanced upon binding of specific proteins (7). Because *cmp* does not interact with RepC or with other plasmid-encoded proteins (4, 5a, 11), we set out to determine whether *cmp* functions as a binding site for a protein(s) encoded by the host chromosome. In this work, we describe the isolation of *cbf1*, a novel chromosomal gene of *Staphylococcus aureus*. The *cbf1* locus codes for *cmp*-binding factor 1 (CBF1), a polypeptide of 313 amino acid residues. The CBF1 protein, expressed at a high level in *Escherichia coli*, binds *cmp* with high specificity and enhances cleavage by DNase I at sites that are in phase with the helical periodicity of DNA.

* Corresponding author. Mailing address: Public Health Research Institute, 455 First Ave., New York, NY 10016. Phone: (212) 578-0844. Fax: (212) 578-0804. E-mail: gennaro@phri.nyu.edu.

† Present address: Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

## MATERIALS AND METHODS

**Library construction and screening.** A genomic DNA library of *S. aureus* NCTC 8325-4 was constructed in the bacteriophage λ ZAP II (Stratagene) according to the protocols supplied by the manufacturer. In this vector, DNA inserts in the correct orientation and reading frame are expressed as β-galactosidase fusion proteins after induction of cultures with isopropyl-thio-β-D-galactoside (IPTG). By an established procedure (16), plaques were screened with a radiolabelled 100-bp DNA probe containing the *cmp* sequence (9) to identify clones expressing a specific *cmp*-binding protein(s).

**Purification of CBF1.** The following protocol was used to purify the CBF1 protein from *E. coli* XL1-Blue MRF′ harboring the recombinant plasmid pCBF1. Overnight cultures of *E. coli* expressing CBF1 were diluted 1:100 in 250 ml of Luria broth containing ampicillin (50 μg/ml) and were grown with vigorous shaking to an optical density at 600 nm of 0.5. IPTG was added to the culture to a final concentration of 0.4 mM, and the induced culture was grown for an additional 3 h. Cells were harvested, washed in a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 5 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), and 10% glycerol, and resuspended in 5 ml of the same buffer. Following addition of KCl to a final concentration of 0.1 M, the cell suspension was treated with lysozyme (4 mg/ml) for 30 min at 4°C and then subjected to one freeze-thaw cycle. After centrifugation at 33,000 rpm (Beckman SW50.1 rotor) for 60 min, the soluble fraction was dialyzed overnight at 4°C against a buffer containing 50 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 1 mM DTT, 10 mM $MgSO_4$, 0.1 M KCl, 10% glycerol, and 0.1 U of DNase I (Sigma). The dialyzed protein was precipitated by addition of ammonium sulfate (44.2 g/100 ml) and collected by centrifugation at 16,000 rpm for 10 min. The precipitate, which was suspended in 0.5 ml of a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 1 mM PMSF, 0.1 M KCl, and 10% glycerol, was dialyzed overnight against the same buffer and stored in aliquots of 50 to 100 μl at −70°C.

CBF1 was further purified either by chromatography on heparin-agarose or by DNA affinity chromatography. In the first method, the protein (10 mg) from the previous step was applied to a column of HiTrap heparin (Pharmacia) equilibrated with a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 1 mM PMSF, 0.1 M KCl, and 10% glycerol. The column was washed with 5 volumes of the same buffer, and the protein was eluted with 3 volumes of buffer containing 1 M KCl. The peak fractions containing *cmp*-binding activity were combined, dialyzed against a buffer containing 0.1 M KCl, and stored at −70°C. Under these conditions, *cmp*-binding activity was stable for 3 to 4 months. In the second method, CBF1 was purified by specific DNA affinity chromatography after ammonium sulfate fractionation. Approximately 50 pmol of biotinylated *cmp* DNA (100 bp) was coupled to 1 mg of streptavidin-coupled beads (Dynabeads M-280; Dynal) in a buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 1 M NaCl for 30 min at 25°C with gentle rotation and then washed twice with the same buffer. The buffer used in the subsequent steps contained 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, and 10% glycerol; the concentration of KCl was as indicated below. Prior to protein binding, the DNA-coupled beads were washed three times with a 0.1 M KCl wash buffer and then incubated with 5 to 10 μg of extracts containing CBF1 for 60 min at 4°C with

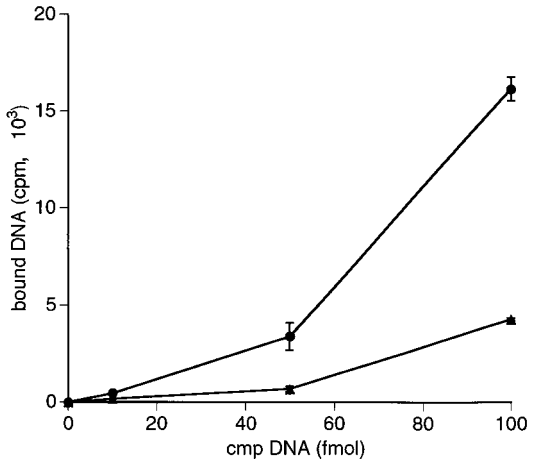


FIG. 1. Expression of *cmp*-binding activity by pCBF1. Binding of CBF1 to radiolabelled *cmp* DNA was determined by a nitrocellulose filter binding assay with *E. coli* cultures containing plasmid pCBF1 grown in the presence and absence of IPTG. ●, excess CBF1 (35 ng; 1 pmol) purified by heparin-agarose from cells grown in the presence of IPTG; ▲, mock-purified protein (35 ng) present in corresponding heparin-agarose fractions from cells that were grown in the absence of IPTG. Increasing amounts of 5′-end-labelled 100-bp *cmp* DNA (as indicated) and protein were incubated for 30 min at 4°C in 25 μl of a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 50 mM KCl, 1 mM DTT, 10% glycerol, and 1 mg of salmon sperm DNA/ml. At the end of the binding reaction, the mixtures (25 μl) were immediately applied with gentle suction (approximately 2 ml/min) to nitrocellulose filters (25-mm diameter, type BA85, 0.45-μm pore size; Millipore) that had been presoaked in binding buffer for 30 min. Filters were washed three times with 500 μl of buffer and dried, and the radioactivity was measured in a liquid scintillation counter.

gentle rotation. After binding, the beads were washed extensively with a 50 mM KCl wash buffer at 4°C. The CBF1 protein was eluted with a 1 M KCl buffer for 30 min at 4°C with gentle rotation.

**$NH_2$-terminal protein sequencing.** The $NH_2$-terminal sequences of the two proteins encoded by *cbf1* were determined after the proteins were separated on 10% polyacrylamide gels containing sodium dodecyl sulfate (SDS) and transferred to polyvinylidene difluoride (PDVF) membranes (Immobilon P; Millipore) with a semidry transfer unit (SemiPhor; Hoefer Scientific Instruments) according to the standard protocol (13). The $NH_2$-terminal amino acid sequences were determined by M. A. Gawinowicz at the Protein Chemistry Core Facility, Columbia University, New York, N.Y., with an Applied Biosystems 477-A protein sequencer.

**Nucleotide sequence accession number.** The GenBank accession number of the *cbf1* nucleotide sequence is U21636.

## RESULTS

**Expression, cloning, and characterization of the *cbf1* structural gene.** An expression library of *S. aureus* DNA was constructed in the bacteriophage λ ZAP II vector and screened with a radiolabelled 100-bp *cmp* DNA probe to identify clones expressing a specific *cmp*-binding protein(s). In a screen of approximately $5 \times 10^5$ plaques, six positive phage clones were obtained. Plasmids pCBF1 to pCBF6, excised in vivo, were analyzed by restriction endonuclease digestion. The six inserts of *S. aureus* DNA, ranging from 6 to 8 kb in length, displayed four different but overlapping restriction patterns (data not shown). One of the clones, pCBF1, expressed *cmp*-binding activity in an IPTG-inducible manner as determined by measuring *cmp*-binding activity in cells grown in the presence and in the absence of IPTG by a filter binding assay (Fig. 1). Because the gene encoding the *cmp*-binding activity in pCBF1 is expressed under *lac* control, the nucleotide sequence of the region distal to *lacZ* in the vector was determined, and an open reading frame (ORF) of 939 nucleotides was found (Fig. 2). The ORF, which is expressed as a *lacZ* operon fusion, codes for a protein of 313 amino acid residues with a calculated $M_r$ of 35,778. Comparison of both the nucleotide and deduced amino acid sequences with sequences in the GenBank database revealed no homology with any known gene or protein. A search for motifs present in DNA binding proteins revealed the presence of a putative helix-turn-helix (HTH) element (amino acids 211 to 233) located in the carboxyl-terminal region of the deduced polypeptide sequence (Fig. 2).

The nucleotide sequence of the ORF in pCBF1 was found in all other positive clones obtained in the library screen by Southern blotting with a PCR-generated DNA fragment containing a sequence internal to the ORF as a probe (data not

```
1   ATTGTTCCATCAAAAGAAGTGATTACATTAAACAAAATAGAGGAAGGCGGGAAACGATGAGAAATATAGAGAATCTAAATCCCGGAGATTCAGTTGATCACT 102
      C  S  I  K  R  S  D  Y  I  K  Q  N  R  G  R  R  E  T  M  R  N  I  E  N  L  N  P  G  D  S  V  D  H  F
                                                              ↑
103 TTTTCTTAGTGCATAAAGCTACACAGGGTGTAACAGCACAAGGTAAAGATTATATGACATTACATTTGCAAGATAAAAGTGGTGAAATTGAAGCGAAATTTT 204
      F  L  V  H  K  A  T  Q  G  V  T  A  Q  G  K  D  Y  M  T  L  H  L  Q  D  K  S  G  E  I  E  A  K  F  W

205 GGACGGCTACAAAAAATGATATGGCAACAATCAAGCCTGAAGAAATTGTACATGTTAAAGGTGACATCATAAACTATCGCGGAAATAAACAGATGAAAGTCA 306
      T  A  T  K  N  D  M  A  T  I  K  P  E  E  I  V  H  V  K  G  D  I  I  N  Y  R  G  N  K  Q  M  K  V  N
                                                                                          ↑
307 ACCAAATTAGACTAGCGACAACTGAAGATCAATTAAAAACAGAACAATTTGTAGATGGTGCACCTTTATCACCGGCAGAAATACAAGAAGAGATTTCTCATT 408
      Q  I  R  L  A  T  T  E  D  Q  L  K  T  E  Q  F  V  D  G  A  P  L  S  P  A  E  I  Q  E  E  I  S  H  Y

409 ATTTGCTAGATATTGAAAATGCTAATTTACAACGTATCACACGTCATTTATTGAAAAAATATCAAGAACGATTTTACACATATCCAGCTGCTAGTTCTCATC 510
      L  L  D  I  E  N  A  N  L  Q  R  I  T  R  H  L  L  K  K  Y  Q  E  R  F  Y  T  Y  P  A  A  S  S  H  H

511 ATCATAACTTTGCGAGTGGCTTAAGCTATCATGTATTAACGATGTTACGTATTGCAAAATCAATTTGTGACATTTATCCATTGTTAAACAAAAGTTTGTTAT 612
      H  N  F  A  S  G  L  S  Y  H  V  L  T  M  L  R  I  A  K  S  I  C  D  I  Y  P  L  L  N  K  S  L  L  Y

613 ATAGTGGTATTATTTTGCATGATATTGGTAAAGTTAGAGAATTGAGTGGTCCTGTTGCGACGTCGTATACAGTCGAAGGTAACTTATTAGGACACATCTCGA 714
      S  G  I  I  L  H  D  I  G  K  V  R  E  L  S  G  P  V  A  T  S  Y  T  V  E  G  N  L  L  G  H  I  S  I

715 TTGCGAGTGATGAAGTAGTTGAAGCAGCTCGTGAATTGAACATTGAAGGAGAAGAAATCATGTTGTTAAGACATATGATTTTATCTCATCATGGTAAGTTAG 816
      A  S  D  E  V  V  E  A  A  R  E  L  N  I  E  G  E  E  I  M  L  L  R  H  M  I  L  S  H  H  G  K  L  E

817 AGTATGGTTCTCCAAAACTGCCATACTTAAAAGAAGCAGAAATTTTATGCTATATCGATAATATCGATGCTAGAATGAATATGTTTGAAAAGGCATATAAAA 918
      Y  G  S  P  K  L  P  Y  L  K  E  A  E  I  L  C  Y  I  D  N  I  D  A  R  M  N  M  F  E  K  A  Y  K  K

919 AAACTGACAAGGGTCAGTTTACAGATAAAATATTTGGTCTTGAAAATCGTAGATTCTACAATCCTGAATCACTCGATTAAG 999
      T  D  K  G  Q  F  T  D  K  I  F  G  L  E  N  R  R  F  Y  N  P  E  S  L  D  *
```

FIG. 2. Nucleotide sequence of *cbf1*. The sequence contains a 939-nt ORF, which is expressed as a *lacZ* operon fusion. We call this ORF *cbf1*. In *E. coli*, the *cbf1* gene expresses two proteins of 35 and 25 kDa. Translation start sites, as determined from the $NH_2$-terminal sequences of the two proteins, are indicated by arrows. Putative ribosome binding sites are underlined. Boldface letters denote a proposed HTH element, which was obtained by computer analysis of the deduced amino acid sequence of CBF1 and a consensus HTH sequence. The consensus HTH sequence (QTELAXKIGVXQQTIGRIFN, where X indicates any amino acid) and its simplified version (DADIADHIAIADDAIAHIID, where D = acidic, polar; A = small, neutral; I = hydrophobic; and H = positively charged) were generated from the known HTH motifs of eight DNA binding proteins that bend DNA (8), using the Wisconsin sequence analysis package by Genetics Computer Group, Inc.

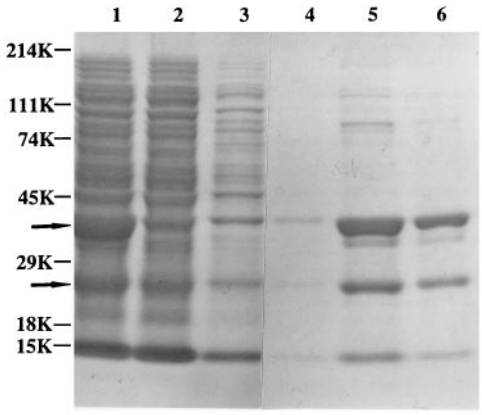


FIG. 3. DNA affinity purification of the CBF1 protein. pCBF1 expressed two IPTG-inducible proteins (indicated by arrows in lane 1) whose apparent molecular masses by SDS-PAGE are approximately 35 and 25 kDa. Extracts containing CBF1 were purified by DNA affinity chromatography. Samples obtained after each step of the procedure were assayed for protein content by SDS-PAGE. Lanes: 1, crude cell extract; 2, flowthrough; 3 and 4, first and fourth low-salt washes, respectively; 5 and 6, high-salt-elution product. The DNA affinity-purified protein shown in the last two lanes contained both the 35 and 25-kDa proteins.

shown). Thus, using the expression cloning strategy described here, we identified in the *S. aureus* chromosome a single, novel gene coding for a protein that binds the *cmp* enhancer. We call this gene *cbf1* (encoding *cmp*-binding factor 1).

**Characterization of the CBF1 protein.** We began to characterize the IPTG-inducible *cmp*-binding activity expressed by the plasmid pCBF1 by measuring the concentration and the molecular weight of the CBF1 protein in cell extracts obtained from *E. coli* transformed with pCBF1. Addition of IPTG to the growth medium elicited increased synthesis of a protein of 35 kDa. This value is in good agreement with the calculated molecular weight of the protein encoded by *cbf1*. Surprisingly, a second protein of about 25 kDa was also expressed in large amounts (Fig. 3, lane 1). To assess the *cmp*-binding activity of the purified proteins, we developed a protocol for specific DNA affinity chromatography with streptavidin-coupled magnetic beads (Dynabeads M-280; Dynal) and biotinylated *cmp* DNA in the presence of excess salmon sperm DNA (see the legend to Fig. 3). After elution with salt at a high concentration, the affinity-purified CBF1, which was about 90% homogeneous as judged by SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 3, lanes 5 and 6), contained the 35- and 25-kDa proteins in a 2:1 molar ratio (estimated by densitometry), indicating that both proteins, either independently or in the form of a heterodimer, bind to the *cmp* DNA matrix with high specificity. In control experiments, beads coupled to nonspecific DNAs were $\leq$10% as efficient as those coupled with *cmp* DNA in retaining the CBF1 protein (data not shown).

The *cbf1* gene, extending from nucleotides (nt) 60 to 998 (Fig. 2), was subcloned in an *E. coli* expression vector and shown to encode both 35- and 25-kDa proteins and to express *cmp*-binding activity that was indistinguishable from that expressed by pCBF1 (data not shown). Further, the $NH_2$-terminal sequence of each protein, determined after separation of the two proteins by SDS-PAGE and transfer to PDVF membranes, shows unambiguously that the 35-kDa protein initiates at the first AUG codon (nt 57 to 59) of the *cbf1* gene whereas the 25-kDa protein initiates in the same reading frame from an internal AUG codon (nt 297 to 299) (Fig. 2). The molecular weights of the two proteins determined by SDS-PAGE are identical with the molecular weights calculated from the deduced amino acid sequence. Taken together, the results of the subcloning experiments and the $NH_2$-terminal sequence analysis show that both the 35- and 25-kDa proteins are encoded by the *cbf1* gene.

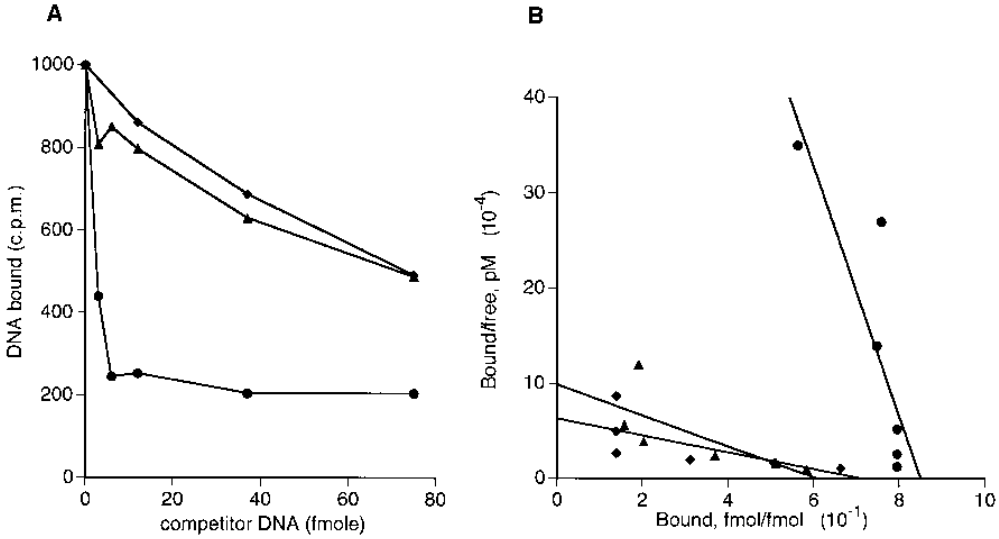


FIG. 4. Analysis of CBF1 binding to *cmp* DNA. (A) Binding of CBF1 to radiolabelled *cmp* DNA was determined in the absence and in the presence of increasing amounts of nonlabelled competitor DNAs, using a nitrocellulose filter binding assay according to the procedure described in the legend to Fig. 1. Binding mixtures contained in a total volume of 25 μl of binding buffer, 1 fmol of 5′-end-labelled 100-bp *cmp* DNA and 10 fmol of partially purified CBF1 (freshly diluted). Filter-bound DNA was corrected for background retention and normalized to the value of 1,000 cpm in the control (no unlabelled competitor DNA). Each data point represents the average of at least three independent experiments. The level of specific retention (80%) measured even at high concentrations of added unlabelled *cmp* DNA is similar to levels observed in other filter binding studies (10, 15). (B) The binding data are expressed as bound DNA (*x* axis) (femtomoles of unlabelled DNAs bound per femtomole of CBF1 monomer) and bound/free DNA (*y* axis) (bound DNA/concentration [picomolar] of total [labelled plus unlabelled] DNA) and were analyzed by linear least-squares fitting to the Scatchard equation. ●, pT181 *cmp* DNA (126-bp fragment); ◆, pBluescript (132-bp *Ssp*I fragment); ▲, pT181 (119-bp *Alu*I fragment, distantly located from *cmp*).

**Binding of CBF1 protein to the *cmp* site.** The specificity of CBF1 binding to *cmp* DNA was investigated by competition binding studies. Binding by CBF1 of $^{32}$P-labelled *cmp* DNA was quantitated with a filter binding assay by measuring the incorporation of radiolabelled DNA into protein-bound DNA in the absence and in the presence of increasing amounts of unlabelled competitor DNAs. The binding data (Fig. 4A) showed that 50% dilution of labelled *cmp* in the CBF1-*cmp* complex required only an ~2-fold excess of nonradiolabelled *cmp* DNA. The same level of isotope dilution was obtained only after addition of a 75-fold excess of two nonspecific DNAs (a 119-bp *Alu*I fragment from pT181 and a 132-bp *Ssp*I fragment from the pBluescript plasmid of *E. coli*). In control experiments (data not shown), single-stranded DNAs derived from *cmp* or from nonspecific DNAs did not compete with *cmp* for CBF1 binding. Taken together, the results of these experiments indicate that CBF1 binds specifically to double-stranded, but not single-stranded, *cmp* DNA. The binding of CBF1 with *cmp* was further analyzed by linear least-squares fitting of the data to the Scatchard equation (Fig. 4B). Assuming the molecular weight of CBF1 to be that of the monomer, the $K_d$ of the CBF1-*cmp* complex was calculated to be $(5 \pm 1) \times 10^{-12}$ $M^{-1}$. This value is similar to the values reported for the *lac* repressor-operator complex (12) and for the λ cI repressor-operator complex (2). As calculated by Scatchard analysis (Fig. 4B), CBF1 binding with two nonspecific DNAs was only 2 and 5% of that with *cmp* DNA, confirming the specificity of CBF1 binding to *cmp* DNA.

Mapping of the CBF1 binding site(s) in *cmp* DNA was initiated by measuring the dilution of radiolabelled full-length *cmp* DNA (100 bp) in the CBF1-*cmp* complex caused by the addition of unlabelled overlapping segments (50 bp) of *cmp* DNA. Analysis of the binding data (not shown) indicated that 50% isotope dilution in the CBF1-*cmp* complex required a concentration of each 50-mer 10 to 30 times higher than that of full-length *cmp* regardless of whether the 50-mer was derived from the left half, central portion, or right half of *cmp*. These findings indicate that CBF1 binds with higher affinity to the full-length, 100-bp *cmp* sequence than it does to any 50-bp region internal to *cmp*.

**DNase I footprinting analysis of CBF1-*cmp* complexes.** Binding of CBF1 to double-stranded *cmp* DNA end labelled with $^{32}$P was analyzed by DNase I footprinting (Fig. 5). Little, if any, protection from DNase I attack was observed upon binding of CBF1 to *cmp* DNA. This result may be explained by a loosening of the CBF1-*cmp* interactions owing to $Mg^{2+}$ ions present in the DNase I buffer or by less-than-full occupancy by CBF1 when multiple CBF1 binding sites are present within *cmp*. The most prominent change in *cmp* DNA in the presence of CBF1 was the appearance of multiple DNase I-hypersensitive sites. Even though two hypersensitive sites located in the top strand of *cmp* (1201 and 1207 [Fig. 5A]) gave a weak signal, multiple strong hypersensitive sites were observed in the bottom strand of *cmp* (ca. 1236 to 1238, ca. 1244 to 1246, ca. 1253 to 1254, ca. 1262 to 1264, and 1270 to 1276 [Fig. 5B]). These sites were located approximately every 10 nt, a distance corresponding to the periodicity of the DNA helix. The number of hypersensitive sites generated by binding of CBF1 to *cmp* DNA together with the periodicity of the sites strongly suggests that CBF1 causes distortion of the *cmp* DNA.

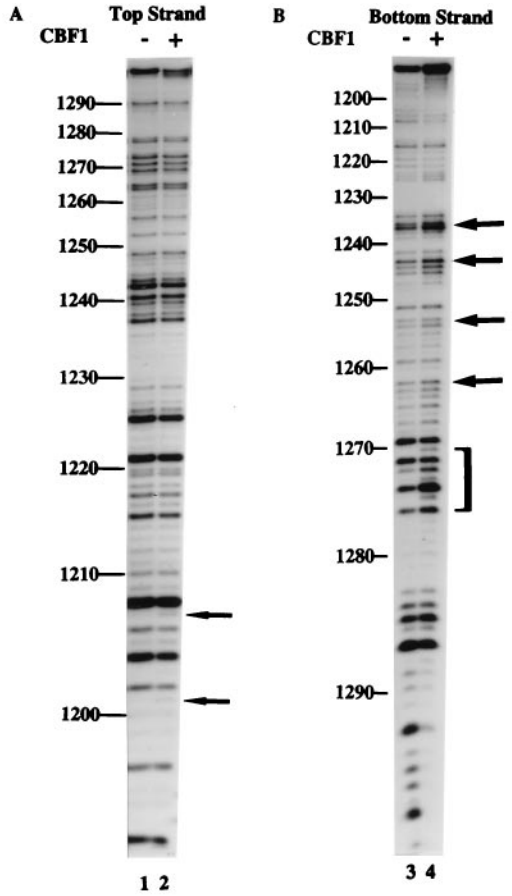


FIG. 5. DNase I footprinting assay. Binding reactions were performed with 5′-end-labelled (100-bp) *cmp* DNA and excess CBF1 purified by heparin-agarose under conditions similar to those described for the filter binding assay (see the legend to Fig. 4). By following an established protocol (1a), the CBF1-*cmp* complex was digested with 0.1 U of DNase I (Sigma) for 2 min at 25°C in a 200-μl volume containing 10 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 1 mM $CaCl_2$, 2 mM DTT, 50 μg of bovine serum albumin/ml, 2 μg of calf thymus DNA/ml, and 0.1 M KCl. DNase I digestion was quenched by addition of 700 μl of stop solution (645 μl of 100% ethanol, 5 μl of tRNA at 1 mg/ml, 50 μl of saturated $NH_4$ acetate). The samples were precipitated, rinsed in ethanol, resuspended in loading buffer, and resolved on 8% polyacrylamide-urea gels. The gels were dried and autoradiographed with a Dupont-Cronex intensifying screen at −70°C. Lanes 1 and 3, DNase I controls (no CBF1 added); lanes 2 and 4, hypersensitive sites (indicated by arrows and brackets) detected on both DNA strands when CBF1 was added to the reaction. Differences in banding patterns in lanes with and without CBF1 were highly reproducible in replicate experiments and were not due to differences in the amounts of DNA applied to the gel. In contrast, the protection of about 10 nt seen at one end of the *cmp* DNA in panel B was not reproducible.

## DISCUSSION

In the present communication, we have shown that *S. aureus* cells contain DNA encoding a specific *cmp*-binding protein that we call CBF1. An expression library of *S. aureus* DNA was probed with *cmp* DNA (100 bp), and six positive clones carrying overlapping inserts of *S. aureus* DNA were isolated. The chromosomal gene encoding the *cmp*-binding activity expressed by these clones was characterized by nucleotide sequence analysis. Comparison of nucleotide and deduced amino acid sequences with sequences in the database revealed no homology with known genes and proteins, indicating that the gene designated *cbf1* that we have isolated in the present studies is novel.

The *cbf1* gene produces two proteins of 35 and 25 kDa, which are translated in the same frame. The 25-kDa species initiates with an AUG codon (nt 297 to 299) which is preceded by a short sequence (GGAAA) that may function as a ribosomal binding site (Fig. 2). Thus, it is likely that the 25-kDa

protein is a product of translation reinitiation, either real or induced by *cbf1* overexpression.

Several lines of evidence indicate that CBF1 protein binds to the *cmp* enhancer in a specific manner. First, CBF1 can be purified by DNA affinity chromatography with beads coupled to *cmp* DNA. Second, CBF1 binds to *cmp* DNA with high specificity (the calculated dissociation constant for the CBF1-*cmp* complex is $5 \times 10^{-12}$ $M^{-1}$). CBF1 binding to *cmp* DNA can be effectively competed by *cmp* DNA but not by other DNAs, irrespective of their G+C content. Third, DNase I footprinting analysis defined specific alterations (DNase I-hypersensitive sites) in the *cmp* DNA upon its association with the CBF1 protein.

Data presented in this paper allow us to delineate some important features of the CBF1-*cmp* complex. The appearance of phased, DNase I-hypersensitive sites strongly suggests that binding of CBF1 causes further distortion of the already intrinsically bent (9) *cmp* DNA. The presence of a common HTH motif in CBF1 and in eight proteins known to bend DNA to some extent (8) supports this idea. Our studies also show that different 50-bp segments of *cmp* bind CBF1 with similar affinities and that all of the segments bind CBF1 less tightly than the entire *cmp* region, suggesting that the CBF1 recognition region spans the entire length of *cmp*. Multiple contact points between CBF1 and *cmp* are consistent with the presence of 11 pentameric repeats in the *cmp* DNA (9) and with the observation that combinations of mutations in *cmp* DNA, which have little effect by themselves, abolish *cmp* activity (1). These considerations, together with the lack of cooperative binding (as inferred by the linearity of the Scatchard plot in Fig. 4), suggest that a multimer of CBF1 establishes multiple contacts with and distorts the 100-bp *cmp* DNA.

Having established that CBF1 binds specifically to the *cmp* DNA, the next important question is whether CBF1 is involved in DNA replication enhancement by *cmp*. Genetic and biochemical studies are currently under way to establish a biological function for CBF1 in plasmid DNA replication. We previously suggested that replication enhancement by *cmp* could ensure a rapid response to copy number fluctuations related to changes in cell growth conditions (5). If shown to be involved in enhancement of plasmid DNA replication, a chromosomally encoded factor such as CBF1 may constitute a regulatory bridge that couples plasmid replication with the cell's physiological state.

## ACKNOWLEDGMENTS

We thank Brian Monk and David Perlin for advice on the preparation of protein samples for $NH_2$-terminal sequencing, Mary Anne Gawinowicz (Protein Chemistry Core Facility, Columbia University) for $NH_2$-terminal sequencing, and Loren Day and Karl Drlica for critical reading of the manuscript. We are especially indebted to Carol Lusty for help with the analysis of binding data and for advice throughout the course of this work.

S.S.O. was supported by a CAPES/Fulbright Fellowship (Brazil/USA Interchange). R.C. was the recipient of a training fellowship from Universitá La Sapienza, Rome, Italy. This work was supported by National Science Foundation grant MCB 9419027 to M.L.G.

## REFERENCES

1. **Bogdanos, N., D. Colombo, and M. L. Gennaro.** Unpublished observations.
1a. **Brenowitz, M., D. F. Senear, and R. E. Kingston.** 1987. DNaseI footprint analysis of protein-DNA binding, p. 12.4.1–12.4.16. *In* F. M. Ausubel et al. (ed.), Current protocols in molecular biology. John Wiley & Sons, New York, N.Y.
2. **Chadwick, P., V. Pirrotta, R. Steinberg, N. Hopkins, and M. Ptashne.** 1970. The λ and 434 phage repressors. Cold Spring Harbor Symp. Quant. Biol. **35:**283–294.
3. **Gennaro, M. L.** 1993. Genetic evidence for replication enhancement from a distance. Proc. Natl. Acad. Sci. USA **90:**5529–5533.
4. **Gennaro, M. L., and R. P. Novick.** 1986. *cmp*, a *cis*-acting plasmid locus that increases interaction between replication origin and initiator protein. J. Bacteriol. **168:**160–166.
5. **Gennaro, M. L., and R. P. Novick.** 1988. An enhancer of DNA replication. J. Bacteriol. **170:**5709–5717.
5a. **Gennaro, M. L.** Unpublished data.
6. **Greenstein, D., and K. Horiuchi.** 1990. Replication enhancer-independent mutation increase the cooperativity with which the initiator protein binds its origin. J. Mol. Biol. **211:**91–101.
7. **Hagerman, P. J.** 1990. Sequence-directed curvature of DNA. Annu. Rev. Biochem. **59:**755–781.
8. **Harrison, S. C., and A. K. Aggarwal.** 1990. DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem. **59:**933–969.
9. **Henriquez, V., V. Milisavljevic, J. D. Khan, and M. L. Gennaro.** 1993. Sequence and structure of *cmp*, the replication enhancer of the *Staphylococcus aureus* plasmid pT181. Gene **134:**93–98.
10. **Johnson, A. D., C. O. Pabo, and R. T. Sauer.** 1980. Bacteriophage λ repressor and cro protein: interactions with operator DNA. Methods Enzymol. **65:** 839–856.
11. **Koepsel, R. R., R. W. Murray, W. D. Rosenblum, and S. A. Khan.** 1985. The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. Proc. Natl. Acad. Sci. USA **82:** 6845–6849.
12. **Lin, S.-Y., and A. D. Riggs.** 1975. The general affinity of *lac* repressor for E. coli DNA: implications for gene regulation in procaryotes and eucaryotes. Cell **4:**107–111.
13. **Matsudaira, P.** 1989. A practical guide to protein and peptide purification for microsequencing. Academic Press, Inc., San Diego, Calif.
14. **Mukherjee, S., H. Erickson, and D. Bastia.** 1988. Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. Cell **52:**375–383.
15. **Riggs, A. D., H. Suzuki, and S. Burgeois.** 1970. *lac* repressor-operator interaction. I. Equilibrium studies. J. Mol. Biol. **48:**67–83.
16. **Singh, H., J. H. LeBowitz, A. S. Baldwin, Jr., and P. A. Sharp.** 1988. Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell **52:**415–423.
17. **Walker, S. S., S. C. Francesconi, and S. Eisenberg.** 1990. A DNA replication enhancer in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **87:**4665–4669.