Nucleotide sequence of a *Bacillus subtilis* gene homologous to the *grpE* gene of *E.coli* located immediately upstream of the *dnaK* gene

Michael Wetzstein and Wolfgang Schumann
Lehrstuhl für Genetik, Universität Bayreuth, D-8580 Bayreuth, FRG

Submitted January 30, 1990 EMBL accession no. X51477

Using a *dnaK* probe from *B. megaterium*, kindly provided by P.Setlow (1) we cloned the analogous heat shock gene from *B. subtilis*. Immediately upstream of this gene, nucleotide sequencing revealed an open reading frame coding for a protein of 187 amino acids. This protein shows 50% homology to the *grpE* gene of *E. coli* (2). The nucleotide sequence determined by the dideoxy sequencing method (3) includes the complete sequence for the *grpE* homologous gene and part of the sequence of the *dnaK* gene. Our sequence of the *dnaK*-gene differs from that of the published one (4) at four positions. Start and stop codons of the open reading frame are underlined, a potential sigma-43 promoter is boxed.

ACKNOWLEDGEMENTS

We thank B. Rak for computer work. This research project was supported by a grant from the BMFT.

REFERENCES

1. Sussmann,M.D. and Setlow,P. (1987) *Nucl. Acids Res.* **15**, 3923.
2. Lipinska,B., King,J., Ang,D. and Georgopoulos,C. (1988) *Nucl. Acids Res.* **16**, 7545.
3. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5468.
4. Hearne,C.M. and Ellar,D.J. (1989) *Nucl. Acids Res.* **17**, 8373.

```
   1 CTGCAGTTGC CATGTTTTAC GGCAGTGAGA GCAGAGATAG CGCTGATGTC CGGCGGTGCT
  61 TTTGCCGTTA CGCACCACCC CGTCAGTAGC TGAACAGGAG GGACAGCTGA TAGAAACAGA
 121 AGCCACTGGA GCACCTCAAA AACACCATCA TACACTAAAT CAGTAAGTTG GCAGCATCAC
 181 CAACAAGTCT GTATGATGAA TAAGGGAATT TTGGCAAATT TTATCGAAGG GCAGCACCTG
 241 TCCTTCTCCT TACACTTTGA GGGAGGTGAA CACAATGTCA GAAGAAAAAC AAACCGTTGA
 301 ACAAAACGAA ACAGAAGAGC AAGAAATCAT TGAAGAACAA GCTGCCGCTG ATGAACAGCA
 361 GGAAGAAACA AATGAAAGCG AACTTCTTCA AAACCAAATT AACGAATTGC AAGGTTTGCT
 421 TGAGGAAAAA GAAAACAAAC TTTTGCGTGT TCAAGCAGAC TTTGAAAACT ATAAACGACG
 481 CAGCCGTTTA GAGATGGAAG CGTCCCAAAA ATACCGTTCT CAAAATATCG TGACTGATTT
 541 GCTGCCGGCT CTTGACAGTT TTGAACGAGC GCTTCAGGTT GAAGCCGACA ATGAACAGAC
 601 GAAAAGTTTG CTCCAGGGAA TGGAAATGGT CCACCGTCAG CTCGTAGAAG CCTTGAAAAA
 661 AGAAGGCGTC GAAGCCATCG AAGCTGTAGG GCAGGAATTT GATCCTAATC TGCACCAAGC
 721 TGTTATGCAG GCTGAAGACG AAAACTACGG CTCCAACATT GTTGTTGAGG AAATGCAAAA
 781 AGGCTATAAG CTGAAGGATC GCGTCATTCG CCCTTCCATG GTCAAAGTGA ATCAATAATT
 841 ACATAGCAGG AGGTTATTCA AGTGAGTAAA GTTATCGGAA TCGACTTAGG AACAACAAAC
 901 TCATGTGTGG CAGTGCTTGA AGGCGGCGAG CCTAAAGTTA TTGCTAACGC TGAAGGAAAC
 961 CGCACAACGC CATCAGTTGT TGCATTTAAA AACGGCGAAC GTCAAGTAGG GGAAGTGGCT
1021 AAACGCCAAT CTATTACAAA CCCTAACACA ATTATGTCTA TCAAACGTCA TATGGGTACT
1081 GATTATAAAG TTGAAATTGA AGGAAAGGAT TACACTCCAC AAGAAGTGTC TGCTATCATC
1141 CTTCAACACC TTAAATCATA CGCTGAAAGC TATCTTGGCG AAACAGTATC AAAAGCAGTT
1201 ATCACAGTTC CTGCATACTT TAACGATGCT GAGCGTCAAG CAACAAAAGA CGCTGGTAAA
1261 ATTGCAGGTC TTGAAGTAGA ACGTATCATC AACGAGCCGA CTGCAGCAGC GCTTGCATAC
1321 GGACTTGATA AAACAGATGA AGATCAAACG ATCCTAGTAT ACGACCTTGG CGGCGGTACA
1381 TTCGACGTTT CCATCCTTGA GCTTGGCGAC GGTGTATTCG AAGTTCGTTC AACTGCCGGC
1441 GACAACCGTC TGGGTGGGGA CGATTTTGAC CAAGTTATCA TCGATCATCT GGTGTCTGAA
1501 TTCA
```