Biochemical and Biophysical Research Communications **263,** 58–62 (1999)
Article ID bbrc.1999.1316, available online at http://www.idealibrary.com on IDEAL®

Overlapping Sequences with High Homology to Functional Proteins Coexist on Complementary Strands of DNA in the Rumen Bacterium *Prevotella albensis*

Nicola D. Walker, Neil R. McEwan,[1] and R. John Wallace
Rowett Research Institute, Greenburn Road, Bucksburn, Aberdeen AB21 9SB, Scotland

Received August 9, 1999

The potential for two complementary fragments of DNA from a clone from the ruminal bacterium *Prevotella albensis* to encode sequences with homology to at least part of functional proteins is described. One strand contains a sequence with high homology to *dnaK*, a member of the *hsp70* family, and the other strand contains a sequence with some homology to glutamate dehydrogenase genes. Overlapping of these two genes on opposite strands has been reported in eukaryotic species, and is now reported for the first time in a bacterial species. Further investigation of previously described *dnaK* genes demonstrates that it is more widespread than might be anticipated, with all thirty other *dnaK* genes investigated also retaining long sequences encoding at least part of a sequence with high homology to a glutamate dehydrogenase gene.

The existence of open reading frames (ORFs) on the antisense strand of a message encoding a functional protein is well documented. Initially it was thought that the role of antisense messages was associated with genomes which adopt a minimalistic approach to replication. These genomes manage to encode messages on both strands of the DNA, or in an overlapping but different reading frame, of the same strand of DNA. This phenomenon is most widely described in extremely small genomes such as those of viruses or small plasmids.

More recently, the use of long antisense reading frames has been described as being more common than first expected (1, 2, 3). Merino *et al.* (1) demonstrate that overlapping antisense ORFs were found in every genome investigated in their experimental work. They attribute this to a non-random phenomenon, which is primarily dependent on codon usage, but also to an extent the GC content of the genome and the size of the gene with the antisense counterpart.

An example of two genes being encoding on opposite strands is illustrated by a 70 kDa heat shock protein and an NAD-specific glutamate dehydrogenase being found on opposite strands of the DNA in the freshwater mould *Achlya klebsiana* (4, 5). In this organism both genes have been shown to be functional on the basis of enzymatic studies. In a similar manner, a long, antiparallel, coupled open reading frame (LAC ORF) has been described in the antisense strand of a 70K heat shock gene from *Drosophila auraria* (6) and an ORF has been described on the antisense strand in *Schizosaccharomyces pombe* (7). In prokaryotes, Silke (3) reports a long antisense open reading frame (aORF) on the antisense strand of the *dnaK* homologue of *E. coli* but does not indicate the frequency of this sequence in other bacterial species, nor does he assign a function to this ORF. He also suggests that it is one of the few *hsp70*s which contains an aORFs.

This work identifies the presence of a sequence with some homology to a glutamate dehydrogenase (GDH) gene lying on the complementary strand of a DNA sequence encoding *dnaK*, a member of the heat shock protein family which have a molecular weight of around 70K. Further investigation demonstrates that a number of other bacterial species also encode sequences with some homology to GDH on their antisense strand. In many cases the antisense sequence contains a number of stop codons, and so is unlikely to be functional. The implications of this discovery are discussed from the perspective of gene evolution.

MATERIALS AND METHODS

Growth of Prevotella albensis and isolation of DNA. *P. albensis* M384 was grown in Hobson's medium (8). DNA was isolated from *P. albensis* M384 using a microwaving technique (9). Cells were pelleted by centrifugation at 13000g for 10 min. The pellet was resuspended in 66.7 mM Tris.HCl (pH 8.0), 6.7 mM EDTA, 3.3% (w/v) SDS and incubated at 65°C for 2 h. The cells were lysed further by microwaving for 2 min on ice at 430W. DNA was collected by phenol:

[1] To whom correspondence should be addressed. Fax: + 44 1224 715349. E-mail: n.mcewan@rri.sari.ac.uk.

0006-291X/99 $30.00

```
CTTCAACAAGCTATCACTAACCCTACAAAGACCATCTACTCTATCAAGCGTTTCATGGGT   60
  L  Q  Q  A  I  T  N  P  T  K  T  I  Y  S  I  K  R  F  M  G    20

GAGACTTATGAGCAGTCAAAGAAAGAGGCTGAGTCAATGCCTTACACCGTAATCAATGAG  120
  E  T  Y  E  Q  S  K  K  E  A  E  S  M  P  Y  T  V  I  N  E    40

AATGGTTACCCACGTGTACAGATCGATGACCGCAAATATACTCCACAGGAGATTTCTGCA  180
  N  G  Y  P  R  V  Q  I  D  D  R  K  Y  T  P  Q  E  I  S  A    60

ATGGTTCTTCAGAAGATGAAGAAGACCGCAGAGGATTATCTTGGTCAGGAAGTTACCGAT  240
  M  V  L  Q  K  M  K  K  T  A  E  D  Y  L  G  Q  E  V  T  D    80

GCTGTCATCACAGTACCTGCTTACTTCTCTGATTCTCAGCGTCAGGCTACTAAGGAGGCT  300
  A  V  I  T  V  P  A  Y  F  S  D  S  Q  R  Q  A  T  K  E  A   100

GGTCAGATTGCTGGTCTTAACGTTCAGCGTATCGTAAACGAGCCAACTGCTGCTGCACTT  360
  G  Q  I  A  G  L  N  V  Q  R  I  V  N  E  P  T  A  A  A  L   120

GCTTATGGTGTAGACAAGAACGACAAAGATATGAACATCGCAGTATTCGACCCTTGGTGG  420
  A  Y  G  V  D  K  N  D  K  D  M  N  I  A  V  F  D  P  W  W   140

TGGTACTTTCGTAATCTCCATTTAACCTTCGGTGGTGGTGTATTCGAGGTTCTTTCTACT  480
  W  Y  F  R  N  L  H  L  T  F  G  G  G  V  F  E  V  L  S  T   160

AATGGTGATACTCACCTCGGTGGTGATGACTTCGACCGTGTTTTCATCAACTGGTTGATC  540
  N  G  D  T  H  L  G  G  D  D  F  D  R  V  F  I  N  W  L  I   180

GATGGTTTCAAGGCTGACGAGGGTATCGACTTGAGCAAAGACCCTATGGCAATGCAGCGT  600
  D  G  F  K  A  D  E  G  I  D  L  S  K  D  P  M  A  M  Q  R   200

TTGAAGGAAGCTGCAGAGAAGGCTAAGATCGAGTTGAGTTCTTCTACTTCTACTGAGATC  660
  L  K  E  A  A  E  K  A  K  I  E  L  S  S  S  T  S  T  E  I   220

AACTTGCCATATATCACAGCAGAGGGTGGTGTACCAAAGCACTTGGTTAAGACTTTGACT  720
  N  L  P  Y  I  T  A  E  G  G  V  P  K  H  L  V  K  T  L  T   240

CGTTCACAGTTCGAGCAGTTGGCACACGACCTGATCCAGGCTTGTCTTGTACCATGCCAG  780
  R  S  Q  F  E  Q  L  A  H  D  L  I  Q  A  C  L  V  P  C  Q   260

AATGCTATCCGTGACGCAAAACTTTCAACTTCAGATATCGACGAAGTTATCCTCGTAGGT  840
  N  A  I  R  D  A  K  L  S  T  S  D  I  D  E  V  I  L  V  G   280

GGTAGCTCTCGTATCCCTGCCGTTCAGACTTTGGTAAAGAACTACTTCGGCAAAGAGCCT  900
  G  S  S  R  I  P  A  V  Q  T  L  V  K  N  Y  F  G  K  E  P   300

TCTAAGGGTGTTAACCCAGACGAGGTTGTTGCAGTTGGTGCAGCTATCCAGGGTGCCGTT  960
  S  K  G  V  N  P  D  E  V  V  A  V  G  A  A  I  Q  G  A  V   320

CTTAACAAGGAAGAGGGTGTTGGCAACATCGTATTGCTCGACGTTACCCCATTGACACTT 1020
  L  N  K  E  E  G  V  G  N  I  V  L  L  D  V  T  P  L  T  L   340

GGTATCGAGACTATGGGTTGTGGACGACGA                                1050
  G  I  E  T  M  G  C  G  R  R                                  350
```

FIG. 1. The DNA sequence of the PCR product and its derived amino acid sequence in frame one of the sense strand.

chloroform extraction and ethanol precipitation (10). DNA was resuspended in sterile distilled water.

PCR amplification of DNA and cloning of PCR products. PCR was performed using the primers ATGACNCCNCARYTNCARCARGCNAT and RTCRTCRTCNACNACCCA (where R = A or G, Y = C or T and N = any one of the four nucleotides). These primers were designed on the basis of amino acid sequences which are conserved in proteins involved in nitrogen regulation in a number of bacterial species. 35 cycles of DNA amplification were performed using a melt temperature of 94°C for 30 s, an annealing temperature of 42°C for 2 min and an amplification temperature of 68°C for 6 min. The reaction cocktail used followed the instructions of the manufacturers for the enzyme eLONGase (Life Technologies, Paisley, Scotland).

Amplification products were checked by electrophoresis prior to cloning. A number of different bands in excess of 1 kb were observed on an agarose gel, suggesting the primers had recognised a number of sites within the genome. Amplified products were cloned into the TA Cloning Kit (Invitrogen), following the manufacturers specifications.

Sequencing of DNA and characterisation of clones. Plasmid DNA was isolated using the method of Birnboim and Doly (11) and purified using a Wizard column (Promega) following the manufacturer's instructions. Plasmid DNA was sequenced in both directions on an ABI Prism 377XL DNA sequencer, using an ABI Prism BigDye terminator sequencing ready reaction kit, following the manufacturer's specifications. Due to the clone being longer than the limits of accuracy for the DNA sequencer, a second round of sequencing was necessary to verify the DNA sequence at either end of the clone. This was performed using two complementary primers which are within the DNA sequence; CCTTGGTGGTGGTACTTTCG and CGAAAGTACCACCACCAAGG. Sequences were translated using the Protein Machine translation facility (URL: http://www.ebi.ac.uk/contrib/tommaso/translate.html). Homology to putative ORFs was established using BLASTP searches (URL: http://www.blast.genome.ad.jp/).

RESULTS AND DISCUSSION

The clone was found to contain a large long ORF (Fig. 1) which showed high homology to the hsp70 protein encoded by the gene *dnaK* ($P = 9.4 \times 10^{-168}$) This sequence has been deposited in the EBI database with Accession Number AJ243536. In addition to its

```
TCGTCGTCCACAACCCATAGTCTCGATACCAAGTGTCAATGGGGTAACGTCGAGCAATAC   60
  S  S  S  T  T  H  S  L  D  T  K  C  Q  W  G  N  V  E  Q  Y    20

GATGTTGCCAACACCCTCTTCCTTGTTAAGAACGGCACCCTGGATAGCTGCACCAACTGC  120
  D  V  A  N  T  L  F  L  V  K  N  G  T  L  D  S  C  T  N  C    40

AACAACCTCGTCTGGGTTAACACCCTTAGAAGGCTCTTTGCCGAAGTAGTTCTTTACCAA  180
  N  N  L  V  W  V  N  T  L  R  R  L  F  A  E  V  V  L  Y  Q    60

AGTCTGAACGGCAGGGATACGAGAGCTACCACCTACGAGGATAACTTCGTCGATATCTGA  240
  S  L  N  G  R  D  T  R  A  T  T  Y  E  D  N  F  V  D  I  *    80

AGTTGAAAGTTTTGCGTCACGGATAGCATTCTGGCATGGTACAAGACAAGCCTGGATCAG  300
  S  *  K  F  C  V  T  D  S  I  L  A  W  Y  K  T  S  L  D  Q   100

GTCGTGTGCCAACTGCTCGAACTGTGAACGAGTCAAAGTCTTAACCAAGTGCTTTGGTAC  360
  V  V  C  Q  L  L  E  L  *  T  S  Q  S  L  N  Q  V  L  W  Y   120

ACCACCCTCTGCTGTGATATATGGCAAGTTGATCTCAGTAGAAGTAGAAGAACTCAACTC  420
  T  T  L  C  C  D  I  W  Q  V  D  L  S  R  S  R  R  T  Q  L   140

GATCTTAGCCTTCTCTGCAGCTTCCTTCAAACGCTGCATTGCCATAGGGTCTTTGCTCAA  480
  D  L  S  L  L  C  S  F  L  Q  T  L  H  C  H  R  V  F  A  Q   160

GTCGATACCCTCGTCAGCCTTGAAACCATCGATCAACCAGTTGATGAAAACACGGTCGAA  540
  V  D  T  L  V  S  L  E  T  I  D  Q  P  V  D  E  N  T  V  E   180

GTCATCACCACCGAGGTGAGTATCACCATTAGTAGAAAGAACCTCGAATACACCACCACC  600
  V  I  T  T  E  V  S  I  T  I  S  R  K  N  L  E  Y  T  T  T   200

GAAGGTTAAATGGAGATTACGAAAGTACCACCACCAAGGGTCGAATACTGCGATGTTCAT  660
  E  G  *  M  E  I  T  K  V  P  P  P  R  V  E  Y  C  D  V  H   220

ATCTTTGTCGTTCTTGTCTACACCATAAGCAAGTGCAGCAGCAGTTGGCTCGTTTACGAT  720
  I  F  V  V  L  V  Y  T  I  S  K  C  S  S  S  W  L  V  Y  D   240

ACGCTGAACGTTAAGACCAGCAATCTGACCAGCCTCCTTAGTAGCCTGACGCTGAGAATC  780
  T  L  N  V  K  T  S  N  L  T  S  L  L  S  S  L  T  L  R  I   260

AGAGAAGTAAGCAGGTACTGTGATGACAGCATCGGTAACTTCCTGACCAAGATAATCCTC  840
  R  E  V  S  R  Y  C  D  D  S  I  G  N  F  L  T  K  I  I  L   280

TGCGGTCTTCTTCATCTTCTGAAGAACCATTGCAGAAATCTCCTGTGGAGTATATTTGCG  900
  C  G  L  L  H  L  L  K  N  H  C  R  N  L  L  W  S  I  F  A   300

GTCATCGATCTGTACACGTGGGTAACCATTCTCATTGATTACGGTGTAAGGCATTGACTC  960
  V  I  D  L  Y  T  W  V  T  I  L  I  D  Y  G  V  R  H  *  L   320

AGCCTCTTTCTTTGACTGCTCATAAGTCTCACCCATGAAACGCTTGATAGAGTAGATGGT 1020
  S  L  F  L  *  L  L  I  S  L  T  H  E  T  L  D  R  V  D  G   340

CTTTGTAGGGTTAGTGATAGCTTGTTGAAG                                1050
  L  C  R  V  S  D  S  L  L  K                                  350
```

FIG. 2. The DNA sequence of the PCR product and its derived amino acid sequence in frame one of the antisense strand.

TABLE 1

Species	Accession No.	Score	Best P value	Reference
Agrobacterium tumefaciens	X87113	676	3.2e-63	12
Bacillus sphaericus	Y17157	139	4.2e-15	13
Bacillus subtilis	M84964	139	4.2e-15	14
Bradyrhizobium japonicum	Y09633	808	2.8e-77	15
*Brucella ovis**	M95799	648*	6.7e-60*	16
Buchnera sp.	D88673	179	1.8e-16	17
Burkholderia pseudomallei	AF016711	572	3.2e-52	18
Caulobacter crescentus	M55224	680	9.2e-64	19
Clostridium acetobutylicum	M74569	55	3.0e-06	20
Clostridium perfringens	X62915	54	3.0e-06	21
Haemophilus influenzae	U32803	106	1.0e-21	22
Legionella pneumophila	D89498	121	3.0e-26	23
Leptospira interrogans	AF007813	64	8.0e-10	24
Methanosarcina mazei	X60265	106	2.0e-22	25
Methylovorus sp.	AF106835	192	9.0e-48	26
Mycobacterium tuberculosis	Z95324	118	7.0e-26	27
Myxococcus xanthus	U83800	140	6.0e-32	28
Nitrosomonas europaea	AB018706	130	7.0e-29	29
Porphyra purpurea	X62240	64	4.0e-09	30
Porphyromonas gingivalis	AB015879	73	3.0e-12	31
Pseudomonas cepacia	L36603	62	1.0e-08	32
Ralstonia eutropha	AJ001727	124	2.0e-27	33
Rhizobium leguminosarum	Y14649	156	3.0e-37	34
Rhizobium meliloti	L36602	156	3.0e-37	32
Rhodobacter capsulatus	U57637	213	2.0e-54	35
Rhodopseudomonas sp.	D78133	196	3.0e-49	36
Salmonella typhimurium plasmid	U58360	206	3.0e-52	37
Staphylococcus aureus	D30690	63	2.0e-08	38
Synechococcus sp.	D28551	208	2.0e-52	39
Synechocystis sp.	M57518	115	3.0e-24	40
Treponema pallidum	AE001203	151	2.0e-35	41
Vibrio harveyi	AF055368	56	2.0e-06	42

* Duplicate gene in different reading frames (100%).

similarity to *dnaK*, it was observed that a fragment of the DNA on the antisense strand had the potential to encode amino acids corresponding to the gene for glutamate dehydrogenase ($P = 3.3 \times 10^{-24}$). This sequence on the antisense strand was interrupted at several places by stop codons, suggesting that although it had large domains which possess the capacity to encode amino acids found in glutamate dehydrogenase, it does not encode a functional gene. This is an interesting observation in the light of the work of LeJohn *et al.* on *Achlya klebsiana* (4, 5), where both sequences were described as having the capacity to encode proteins (Fig. 2).

It should also be noted that the two sequences are also encoded by DNA in the opposite but same frames on the two strands (*i.e.* in the first reading frame of both the sense and antisense strands). This is in keeping with the prediction of Silke (3), where he suggests that duel-stranded potential of encoding two messages would be maximised in the case of sequences which have a tendency to preferentially use codons with the consensus RNY. This bias may be measured as follows:

$$\mathrm{RNY} = \frac{\mathrm{Thr}^{\mathrm{ACC}} + \mathrm{Thr}^{\mathrm{ACT}} + \mathrm{Ser}^{\mathrm{AGC}} + \mathrm{Ser}^{\mathrm{AGT}} + \mathrm{Val}^{\mathrm{GTC}} + \mathrm{Val}^{\mathrm{GTT}} + \mathrm{Ala}^{\mathrm{GCC}} + \mathrm{Ala}^{\mathrm{GCT}}}{\mathrm{Thr}^{\mathrm{ACN}} + \mathrm{Ser}^{\mathrm{AGC}} + \mathrm{Ser}^{\mathrm{AGT}} + \mathrm{Arg}^{\mathrm{AGA}} + \mathrm{Arg}^{\mathrm{AGG}} + \mathrm{Val}^{\mathrm{GTN}} + \mathrm{Ala}^{\mathrm{GCN}}}$$

In the case of this sequence, the RNY index value is 0.68, and hence shows a bias deviating from 50% for the 70 codons using ACN, AGN, GTN or GCN.

To our knowledge, only the work of LeJohn *et al.* previously described *dnaK* and *gdh* existing on opposite strands of DNA, although others have described the presence of an ORF on the strand complementary to the *dnaK* gene. In prokaryotes, this appears to be the first time that the homology to a *gdh* gene has been recorded in a genome.

It was unknown how widespread a phenomenon domains having the capacity encoding parts of the *dnaK* and *gdh* genes was likely to be, and so we extended the investigation to the top 30 hits generated on a BLASTP search for the *dnaK* sequence. Details of sequences used in this work, and the level of homology their amino acids in their complementary strand showed to

gdh sequences is shown in Table 1. In this work, most sequences on the complementary strand contained stop codons. Thus, the investigation is being carried out on what would appear to be at best a relic of a former gene, rather than a functional gene *per se*.

Clearly, a number of other organisms also possess at least partial homology to a glutamate dehydrogenase gene on the strand complementary to the *dnaK* gene. In some cases, the length of the putative ORF is extremely long e.g. *Pseudomonas cepacia* (679 amino acids), *Ralstonia eutropha* (694 amino acids), *Rhodobacter capsulatus* (686 amino acids), *Bradyrhizobium japonicum* (793 amino acids), *Brucella ovis* (1134 amino acids – two copies) and *Burkholderia pseudomallei* (651 amino acids – entire sequence of *dnaK* gene).

Given the length of these sequences on the antisense strand, it might seem likely that they would be functional. This would be in keeping with the proposal that if an antisense non-stop reading frame (aNRF) cannot be easily explained by codon usage, then it is likely to be present for the purposes of translation (3). In this case, there are a number of sequences which are interrupted by the presence of stop codons. Clearly it is unlikely that these are being translated, and so it would appear more likely that they are arising due to a relic from evolution. Alternatively, it may be possible that the one sequence which has been shown to have both heat shock and *gdh* functions (4, 5) is actually the more highly evolved sequence, by virtue of it being able to maintain two genes at the expense of only one portion of double stranded DNA. In the absence of additional evidence, it is not clear which of these two hypotheses is correct–although the former does seem more likely for organisms which have a functional *gdh* gene elsewhere in the genome.

As already stated, we were unable to identify any publications which document coexistence of the *hsp70* and *gdh* genes on parallel strands in bacteria. Furthermore, a number of publications dealing with the *hsp70* genes describe ORFs solely on the sense strand (i.e. on the strand encoding *hsp70*), including genes such as *dnaJ* and *grpE*, and do not report any sequence on the complementary strand. Only one of these papers identifies an ORF on the opposite strand, present in *E. coli* (3)–and this paper does not assign a function to this ORF.

In many cases, the partial homology to the *gdh* gene is easy to miss. Firstly, the longer ORFs on this strand often does not contain a likely start codon 3′ of the previous stop codon. As such, they would be discounted as being functional. Secondly, a BLASTX search using the *P. albensis* M384 DNA sequence, with a cut-off value of 500 sequences gives proteins with a probability of $P = 8.0 \times 10^{-48}$ or less. Since the value of the best sequence with *gdh*-like homology is $P = 3.3 \times 10^{-24}$, none of the *gdh*-like sequences on the complementary strand is identified at this cut-off level. A combination of these two circumstances makes it probable that the apparent relic of a *gdh* gene would remain undetected in the majority of cases.

In conclusion, we have demonstrated that there are sufficiently long aNRFs in a number of prokaryotic species, which often lack a readily identifiable start codon, but which still retain sufficient homology to segments of the protein encoding gdh in the water mould *Achlya klebsiana*. The frequent absence of obvious start codons, and the regularity of stop codons within this strand leads us to postulate that the observation is the result of an evolutionary relic, rather a functional *gdh* gene.

ACKNOWLEDGMENT

The Rowett Research Institute receives funding from the Scottish Office of Agriculture, Fisheries, and Environment Department.

REFERENCES

1. Merino, E., Balbas, P., Puente, J. L., and Bolivar, F. (1994) *Nucl. Acids Res.* **22,** 1903–1908.
2. Yomo, T., and Urabe, I. (1994) *J. Mol. Evol.* **38,** 113–120.
3. Silke, J. (1997) *Gene* **194,** 143–155.
4. LeJohn, H. B., Cameron, L. E., Yang, B., MacBeath, G., Barker, D. S., and Williams, S. A. (1994) *J. Biol. Chem.* **269,** 4513–4522.
5. LeJohn, H. B., Cameron, L. E., Yang, B., and Rennie, S. L. (1994) *J. Biol. Chem.* **269,** 4523–4531.
6. Konstantopoulou, I., Ouzounis, C. A., Drosopoulou, E., Yiangou, M., Sideras, P., Sander, C., and Scourar, Z. G. (1995) *J. Mol. Evol.* **41,** 414–420.
7. Usui, T., Yoshida, M., Kasahara, K., Honda, A., Beppu, T., and Horinouchi, S. (1997) *Gene* **189,** 43–47.
8. Hobson, P. N. (1969) *in* Methods in Microbiology (Norris, J. R., and Ribbons, D. W., Eds.), Vol. 3B, pp. 133–145, Academic Press, London.
9. McEwan, N. R., Wheeler, C. T., and Milner, J. J. (1994) *Soil Biol. Biochem.* **26,** 541–545.
10. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.
11. Birnboim, H. C., and Doly, J. (1979) *Nucl. Acids Res.* **7,** 1513–1523.
12. Segal, G., and Ron, E. Z. (1995) *J. Bacteriol.* **177,** 5952–5958.
13. Ahmad, S., Selvapandiyan, A., Gasbarri, M., and Bhatnagar, R. K. (1998) Direct Submission to GenBank.
14. Wetzstein, M., Voelker, U., Dedio, J., Loebau, S., Zuber, U., Schiesswohl, M., Herget, C., Hecker, M., and Schumann, W. (1992) *J. Bacteriol.* **174,** 3300–3310.
15. Minder, A. C., Narberhaus, F., Babst, M., Hennecke, H., and Fischer, H. M. (1997) *Mol. Gen. Genet.* **254,** 195–206.
16. Cellier, M. F., Teyssier, J., Nicolas, M., Liautard, J. P., Marti, J., and Sri Widada, J. (1992) *J. Bacteriol.* **174,** 8036–8042.
17. Sato, S., and Ishikawa, H. (1997) *J. Biochem.* **122,** 41–48.
18. See, L. H., Yap, E. H., and Yap, E. P. H. (1997) Direct Submission to GenBank.
19. Gomes, S. L., Gober, J. W., and Shapiro, L. (1990) *J. Bacteriol.* **172,** 3051–3059.

20. Narberhaus, F., Giebeler, K., and Bahl, H. (1992) *J. Bacteriol.* **174,** 3290–3299.
21. Galley, K. A., Singh, B., and Gupta, R. S. (1992) *Biochim. Biophys. Acta* **1130,** 203–208.
22. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G. G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science* **269,** 496–512.
23. Amemura-Maekawa, J., and Watanabe, H. (1997) *Gene* **197,** 165–168.
24. Ballard, S. A., Go, M., Segers, R. P. A. M., and Adler, B. (1998) *Gene* **216,** 21–29.
25. Macario, A. J., Dugan, C. B., and Conway de Macario, E. (1991) *Gene* **108,** 133–137.
26. Eom, C. Y., and Kim, Y. M. (1998) Direct Submission to GenBank.
27. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squares, R., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393,** 537–544.
28. Weimer, R. M., Creighton, C., Stassinopoulos, A., Youderian, P., and Hartzell, P. L. (1998) *J. Bacteriol.* **180,** 5357–5368.
29. Iizumi, T., and Nakamura, K. (1997) *Appl. Environ. Microbiol.* **63,** 1777–1784.
30. Reith, M., and Munholland, J. (1991) *FEBS Lett.* **294,** 116–120.
31. Yoshida, A., Nakano, Y., Yamashita, Y., Oho, T., Shibata, Y., Ohishi, M., and Koga, T. (1999) *FEBS Lett.* **446,** 287–291.
32. Falah, M., and Gupta, R. S. (1994) *J. Bacteriol.* **176,** 7748–7753 (1994).
33. Talbi, S., and van der Lelie, D. (1997) Direct Submission to GenBank.
34. Simpkins, S. A., Johnston, A. W. B., and James, R. (1997) Direct Submission to GenBank.
35. Nickel, C. M., Vandekerckhove, J., Beyer, P., and Tadros, M. H. (1997) *Gene* **192,** 251–259.
36. Momma, K., Inui, M., Yamagata, H., and Yukawa, H. (1997) *Biochim. Biophys. Acta* **1350,** 235–239.
37. Stephen, R. J., and Hinton, J. C. D. (1996) Direct Submission to GenBank.
38. Ohta, T., Saito, K., Kuroda, M., Honda, K., Hirata, H., and Hayashi, H. (1994) *J. Bacteriol.* **176,** 4779–4783.
39. Nimura, K., Yoshikawa, H., and Takahashi, H. (1994) Direct Submission to GenBank.
40. Chitnis, P. R., and Nelson, N. (1991) *J. Biologic. Chem.* **266,** 58–65.
41. Fraser, C. M., Norris, S. J., Weinstock, G. M., White, O., Sutton, G. G., Dodson, R., Gwinn, M., Hickey, E. K., Clayton, R., Ketchum, K. A., Sodergren, E., Hardham, J. M., McLeod, M. P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J. K., Chidambaram, M., Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M. D., Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Watthey, L., Weidman, J., Smith, H. O., and Venter, J. C. (1998) *Science* **281,** 375–388.
42. Klein, G., Walczak, R., Krasnowska, E., Blaszczak, A., and Lipinska, B. (1995) *Mol. Microbiol.* **16,** 801–811.