JOURNAL OF BACTERIOLOGY, Aug. 1992, p. 5171–5175
0021-9193/92/155171-05$02.00/0

Vol. 174, No. 15

# A General Method for Cloning *recA* Genes of Gram-Positive Bacteria by Polymerase Chain Reaction

PATRICK DUWAT,* S. DUSKO EHRLICH, AND ALEXANDRA GRUSS

*Laboratoire de Génétique Microbienne, Institut de Biotechnologie, Institut National de la Recherche Agronomique, Domaine de Vilvert, 78352 Jouy en Josas Cedex, France*

Received 3 March 1992/Accepted 25 May 1992

**An internal fragment of the *recA* gene from eight gram-positive organisms has been amplified by using degenerate primers in a polymerase chain reaction. The internal 348- or 360-bp *recA* DNA segments from *Bacillus subtilis*, *Clostridium acetobutylicum*, *Lactobacillus bulgaricus*, *Lactobacillus helveticus*, *Leuconostoc mesanteroides*, *Listeria monocytogenes*, *Staphylococcus aureus*, and *Streptococcus salivarus* subsp. *thermophilus* were amplified, cloned, and sequenced. The G+C contents of the DNA from these species range from 28 to 52%. The sequences of the bacterial *recA* genes show strong relatedness. This method is particularly useful for the recovery of the *recA* genes of gram-positive bacteria and avoids the difficulties of using a genetic complementation test for cloning.**

The *recA* gene of *Escherichia coli*, which codes for a protein of 352 amino acids, has been extensively characterized and is essential for numerous cell functions, such as homologous recombination, induction of the SOS pathway, and bacteriophage activation (11). The *recA* gene has been cloned from more than 40 organisms (9, 11). Nevertheless, only five *recA* genes have been cloned from the gram-positive bacteria *Acholeplasma laidlawii* (4), *Bacillus subtilis* (8), *Staphylococcus aureus* (cited in reference 11), *Mycobacterium tuberculosis* (2), and *Lactococcus lactis* (3). The underrepresentation of cloned *recA* genes from the gram-positive group may be due to difficulties in gene expression (2) or to the instability of the cloned gene in a heterologous organism such as *E. coli* (15). In this study, we describe a new method based on the polymerase chain reaction (PCR) (12) to clone an internal DNA segment of the *recA* gene (here called $recA_{int}$) of numerous organisms. This method enables the recuperation of entire *recA* genes without a complementation step and can be applied to gram-positive bacteria which seem to be refractive to cloning by other means.

The amino acid sequences of 18 previously characterized *recA* genes were aligned, and highly conserved stretches of the protein were chosen as templates in order to design degenerate primers (Fig. 1). Oligonucleotides with 192-fold, 128-fold (coding-strand primers a and b, respectively), and 384-fold (complementary-strand primer c) degeneracies were chosen. The relatively low G+C contents of the organisms involved (28 to 52%) were considered in the designing of the first 10 nucleotides (5′ end) but not in that of the remainder of the primer, since perfect homology on the 3′ end is desirable for good elongation. The two primer pairs, a plus c

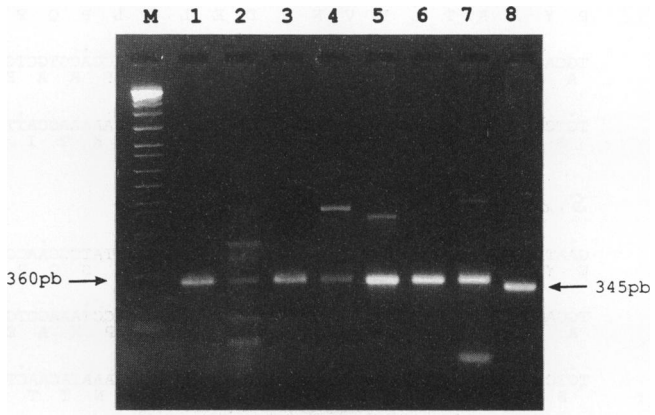


FIG. 1. Description of the degenerate primers used in the PCRs. The degenerate oligonucleotide primer pairs used for the PCR amplification of the $recA_{int}$ fragment are a and c or b and c. The amino acid segments upon which the primers are based are shown above the DNA sequences. Amino acid positions are given according to the *E. coli* sequence.




FIG. 2. Amplification by PCR of chromosomal DNA from gram-positive organisms. Chromosomal DNA of the different gram-positive strains was amplified by using PCR with the primer couples a and c (lanes 1 to 7) or b and c (lane 8). Lanes 1 to 8 correspond to 10 μl of the PCR mixture with the chromosomal DNA of *B. subtilis*, *Lactobacillus bulgaricus*, *Leuconostoc mesanteroides*, *Listeria monocytogenes*, *S. aureus*, *L. lactis*, *Lactobacillus helveticus*, and *S. thermophilus*, respectively. Lane M corresponds to the Raoul marker (Appligene). The result of the PCR amplification of *C. acetobutylicum* chromosomal DNA is not shown.

* Corresponding author.

*C.acetobutylicum*

```
AGTTATGCCCAAAAACTTGGAGTTGATGTAGATAGCTTAATAATTTCTCAACCTGATACAGGGGAACAAGGACTTGAAATAGCTGAAGCATTAGTCCGTTCTGG  104
 S  Y  A  Q  K  L  G  V  D  V  D  S  L  I  I  S  Q  P  D  T  G  E  Q  G  L  E  I  A  E  A  L  V  R  S  G   35

AGCTATTGATGTTTTGGTAGTAGACTCAGTTGCGGCTTTAGTACCTAAAGCTGAAATTGAAGGAGAAATGGGAGATTCACATATTGGTCTTCAAGCAAGACTTA  208
 A  I  D  V  L  V  V  D  S  V  A  A  L  V  P  K  A  E  I  E  G  E  M  G  D  S  H  I  G  L  Q  A  R  L  M   70

TGTCACAGGCGTTAAGAAAGCTTGCAGGAACAATAAATAAGACTAATTGTGTTGCTATATTTATTAACCAATTAAGAGAAAAGTTGGGGTTTATGTTTGGTTCT  312
   S  Q  A  L  R  K  L  A  G  T  I  N  K  T  N  C  V  A  I  F  I  N  Q  L  R  E  K  L  G  F  M  F  G  S   103
```

*L.bulgaricus*

```
GCTTACGCTGAAGCCTTGGGCGTGGACATCGACCAATTGATCCTGTCTCAGCCAAACACTGGGGAAGAAGGGCTGCAAATCGCGGACACCTTGATCTCCAGCGG  104
 A  Y  A  E  A  L  G  V  D  I  D  Q  L  I  L  S  Q  P  N  T  G  E  E  G  L  Q  I  A  D  T  L  I  S  S  G   35

GGCCATCGACATCGTCGTGGTCGACTCCGTTGCCGTCCTGGTGCCGCGGGCCGAAATCGAAGGTGAAATGGGTGACTCCCACGTTGGACTCCAGGCCCGCCTGA  208
 A  I  D  I  V  V  V  D  S  V  A  V  L  V  P  R  A  E  I  E  G  E  M  G  D  S  H  V  G  L  Q  A  R  L  M   70

TGAGCCAGGCCTTGCGCAAGCTCTCCGGGACGATTGCCAAGACCAAGACCATTGCCATCTTCATCAACCAGATCCGGGAAAAGGTCGGCGTCATGTTTGGTAAT  312
   S  Q  A  L  R  K  L  S  G  T  I  A  K  T  K  T  I  A  I  F  I  N  Q  I  R  E  K  V  G  V  M  F  G  N   104
```

*L.helveticus*

```
GCATATGCTGAAGCATTAGGTGTAGACATTGATTCACTTATTTTATCTCAACCTAATACAGGTGAAGAAGGTTTACAAATTGCTGATACTTTGATTTCTAGTGG  104
 A  Y  A  E  A  L  G  V  D  I  D  S  L  I  L  S  Q  P  N  T  G  E  E  G  L  Q  I  A  D  T  L  I  S  S  G   35

TGCCATTGATATTGTGGTAGTTGACTCTGTTGCAGCGTTAGTTCCACGTGCTGAAATTGAAGGCGAGATGGGGGATGCTCATGTTGGTCTTCAAGCTCGATTAA  208
 A  I  D  I  V  V  V  D  S  V  A  A  L  V  P  R  A  E  I  E  G  E  M  G  D  A  H  V  G  L  Q  A  R  L  M   70

TGAGTCAGGCTTTGCGTAAATTATCTGGTACTATTTCAAAGACTAAGACCATTGCAATTTTTATTAATCAGATTCGTGAAAAAAGTTGGGATTATGTTTGGTAAC  312
   S  Q  A  L  R  K  L  S  G  T  I  S  K  T  K  T  I  A  I  F  I  N  Q  I  R  E  K  V  G  I  M  F  G  N   104
```

*L.mesantheroides*

```
AAGTACGCTGAGGCATTGGGCGTACAAAAAGATGAATTACTGCTATCACAACCAGATACTGGCGAACAAGGCTTAGAAATTGCTGATGCGCTAGTACAATCAGG  104
 K  Y  A  E  A  L  G  V  Q  K  D  E  L  L  L  S  Q  P  D  T  G  E  Q  G  L  E  I  A  D  A  L  V  Q  S  G   35

TGCGGTCGATATTATTGTTGTTGATTCAGTCGCAGCCCTTGTACCTCGAGCTGAAATTGAAGGTGAAATGGGTGATTCGCATGTTGGTTTGCAAGCACGATTAA  208
 A  V  D  I  I  V  V  D  S  V  A  A  L  V  P  R  A  E  I  E  G  E  M  G  D  S  H  V  G  L  Q  A  R  L  M   70

TGAGTCAAGCGCTTCGTAAATTAGCTGGGACTTTAAATCGCACAGGAACCATTGCTATTTTTATTAACCAAATTCGTGAAAAAAATTGGTGTGATGTTTGGTAAT  312
   S  Q  A  L  R  K  L  A  G  T  L  N  R  T  G  T  I  A  I  F  I  N  Q  I  R  E  K  I  G  V  M  F  G  N   104
```

*L.monocytogenes*

```
CCTTATGCTAAAACCCTAGTTGTAAATATTGATGAATTATTACTACCTCAACCAGATACAGGAGAACAAGCTTTAGAGATTGCTGAAGCTTTAGTTAGAAGTGG  104
 P  Y  A  K  T  L  V  V  N  I  D  E  L  L  L  P  Q  P  D  T  G  E  Q  A  L  E  I  A  E  A  L  V  R  S  G   35

TGCAGTTGATATGTTAGTAATTGACTCCGTTGCAGCACTTGTACCACGTGCTGAAATCGAAGGCGAGATGGGCGATGCTCATGTTGGATTACAAGCACGTTTAA  208
 A  V  D  M  L  V  I  D  S  V  A  A  L  V  P  R  A  E  I  E  G  E  M  G  D  A  H  V  G  L  Q  A  R  L  M   70

TGTCCCAAGCATTGCGTAAACTTTCTGGTGTTATTAATAAATCAAAAACCATTGCTATTTTCATTAACCAAATTCGTGAAAAAGTTGGTGTTATGTTTGGTAAC  312
   S  Q  A  L  R  K  L  S  G  V  I  N  K  S  K  T  I  A  I  F  I  N  Q  I  R  E  K  V  G  V  M  F  G  N   104
```

*S.aureus*

```
GAATATGCTCAAGCATTAGGCGTAGATATCGATAATTTATATTTATCGCAACCGGATCATGGTGAACAAGGTCTTGAAATCGCCGAAGCATTTGTTAGAAGTGG  104
 E  Y  A  Q  A  L  G  V  D  I  D  N  L  Y  L  S  Q  P  D  H  G  E  Q  G  L  E  I  A  E  A  F  V  R  S  G   35

TGCAGTTGATATTGTAGTTGTAGACTCAGTTGCTGCTTTAACACCTAAAGCTGAAATTGAAGGAGAAATGGGAGACACTCACGTTGGTTTACAAGCTCGTTTAA  208
 A  V  D  I  V  V  V  D  S  V  A  A  L  T  P  K  A  E  I  E  G  E  M  G  D  T  H  V  G  L  Q  A  R  L  M   70

TGTCACAAGCGTTACGTAAACTTTCAGGTGCTATTTCTAAATCAAATACAACTGCTATTTTCATCAACCAAATTCGTGAAAAAGTTGGTGTTATGTTCGGTAAT  312
   S  Q  A  L  R  K  L  S  G  A  I  S  K  S  N  T  T  A  I  F  I  N  Q  I  R  E  K  V  G  V  M  F  G  N   104
```

*S.thermophilus*

```
GCGTATGCACGAGCTCTAGGTGTTAATATCGATGAGCTTCTTTTGTCGCAGCCTGATTCTGGTGAGCAAGGTCTCGAAATTGCAGGTAAGCTGATTGACTCTGG  104
 A  Y  A  R  A  L  G  V  N  I  D  E  L  L  L  S  Q  P  D  S  G  E  Q  G  L  E  I  A  G  K  L  I  D  S  G   35

TGCAGTGGATTTAGTTGTTGTTGACTCAGTTGCGGCCTTCGTACCACGTGCAGAAATTGATGGAGATAGTGGTGACAGTCATGTAGGACTTCAAGCGCGTATGA  208
 A  V  D  L  V  V  V  D  S  V  A  A  F  V  P  R  A  E  I  D  G  D  S  G  D  S  H  V  G  L  Q  A  R  M  M   70

TGAGTCAAGCCATGCGTAAACTTTCTGCATCTATTAATAAAACAAAAACGATTGCTATCTTTATTAACCAGTTGCGTGAAAAAGTTGGTATCATGTTTGGTAAC  312
   S  Q  A  M  R  K  L  S  A  S  I  N  K  T  K  T  I  A  I  F  I  N  Q  L  R  E  K  V  G  I  M  F  G  N   104
```

FIG. 3. DNA and amino acid sequences of the $recA_{int}$ fragments. The DNA and deduced amino acid sequences of the seven sequenced $recA_{int}$ fragments are shown. Sequences corresponding to the degenerate primers can be inaccurate and are therefore not shown. Sequencing of both strands was performed, with about 99% sequence overlap. In order to confirm sequence results, two independent clones were analyzed for six of the amplifications (one clone was sequenced for *C. acetobutylicum* and *S. thermophilus* DNA).

and b plus c, gave efficient amplification with the DNA samples described below, although various other pairs were initially tried (data not shown).

The gram-positive bacterial strains used in this study are *B. subtilis* SB202 (*leuA8 tyrA1 aroB2 hisH2*; laboratory collection), *Clostridium acetobutylicum* NCIB8052 (M. Young), *Lactobacillus bulgaricus* IL431 (plasmid free; M. C. Chopin), *Lactobacillus helveticus* IL1235 (plasmid free; M. C. Chopin), *Lactococcus lactis* subsp. *lactis* IL1403 (here called *L. lactis*) (plasmid free; M. C. Chopin), *Leuconostoc mesanteroides* IL1250 (plasmid free; M. C. Chopin), *Listeria monocytogenes* B7 (J. J. Gratadoux), *Staphylococcus aureus* RN450 (R. Novick), and *Streptococcus salivarius* subsp. *thermophilus* IL73 (here called *S. thermophilus*) (plasmid free; M. C. Chopin).

One microgram of chromosomal DNA isolated from each strain (extracted as previously described [3, 7]) was added to a PCR mixture of 100 μl (2.5 U of Perkin-Elmer Cetus *Taq* DNA polymerase in its buffer [1×] with 250 pmol of a primer couple) and submitted to 30 cycles of 1 min of denaturation at 95°C, 2 min of annealing at 55°C, and 3 min of elongation at 72°C in a Genamp PCR System 9600 (Perkin-Elmer). *L. lactis* chromosomal DNA was amplified as a positive control, as we recently showed that this reaction gives a unique $recA_{int}$ segment with the primer pairs used here (3). The presence of a DNA fragment of the expected size (either 348 or 360 bp) was contingent upon the presence of formamide in the reaction mixture (3, 13). PCR products were separated on a 3% agarose gel (eight representative samples are shown in Fig. 2). For each chromosomal DNA preparation, a DNA fragment of the expected size was amplified with at least one of the two primer combinations tried. Good amplification of *S. thermophilus* DNA was observed only with the primer pair b plus c (the expected size fragment of 348 bp was obtained). In some cases, more than one DNA fragment appeared (Fig. 2, lanes 2, 4, 5, and 7). The extra bands are nonspecific and may appear because the motifs chosen for primer synthesis are present in other proteins. The weak intensity of the $recA_{int}$-amplified bands of *Lactobacillus bulgaricus* (Fig. 2, lane 2) and those of *C. acetobutylicum* (data not shown) might be due to the relative extremes in the G+C contents of the DNA of these organisms (52 and 28% [10], respectively) compared with the G+C distribution of the primers.

We verified that the amplified DNA bands correspond to $recA_{int}$. A 10-μl sample of the PCR mixture was treated with T4 DNA polymerase and DNA polymerase I (Boehringer, Mannheim, Germany) and separated on a 3% agarose gel. The bands of 348 or 360 bp (according to the primer pairs used) were purified and cloned into a *Sma*I-linearized pBluescript KS$^+$ plasmid (Stratagene) as previously described (3, 7). The cloned fragments were sequenced on both strands and from two independent clones (only one clone for *C. acetobutylicum* and *S. thermophilus* was sequenced) except for *B. subtilis* PCR-amplified $recA_{int}$, which was verified only by hybridization with a segment of the *Bacillus recA* gene used as a probe (kindly provided by K. Bayles).

In every case, the DNA and deduced amino acid sequences of the $recA_{int}$ segment show good homology with those of *E. coli* $recA_{int}$ (Fig. 3 and 4). The sequence similarity corresponds to positions 92 through 211 (primers a and c) or positions 96 through 211 (primers b and c) of the *E. coli* RecA amino acid sequence, as expected from the choice of the primers. Crystallographic studies of *E. coli* RecA revealed seven amino acids which are essential for protein function and invariant among all known bacterial RecA-like proteins (14). Of these seven amino acids, four are contained in the $RecA_{int}$ region (Asp-144, Ser-145, Asn-193, and Gln-194) and are also invariant in the gram-positive proteins. Thus, the $recA_{int}$ DNA sequences are also highly conserved among the gram-positive species (Fig. 5).

The *recA* gene of *S. aureus* has already been partially sequenced (11). Comparison of the published deduced amino acid sequence with that derived from the PCR-cloned fragment reveals three differences: Ser, Thr, and Gln instead of Arg, His, and Leu at positions 136, 162, and 167. Two amino acids, Ser-136 and Gln-167, are conserved in all of the gram-positive RecA sequences (Fig. 3), which may indicate

```
      102      110       120       130       140       150       160       170       180       190       200
Ecoli   IYARKLGVDIDNLLCSQPDTGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGNLKQSNTLLIFINQIRMKIGVMF.GN
 Bsub   VYAQKLGVNIEELLLSQPDTGEQALEIAEALVRSGAVDIVVVDSVAALVPKAEIEGDMGDSHVGLQARLMSQALRKLSGAINKSKTIAIFINQIREKVGNMFFRE
 Saur   EYAQALGVDIDNLYLSQPDHGEQGLEIAEAFVRSGAVDIVVVDSVAALTPKAEIEGEMGDTHVGLQARLMSQALRKLSGAISKSNTTAIFINQIREKVGVMF.GN
Lbulg   AYAEALGVDIDQLILSQPNTGEEGLQIADTLISSGAIDIVVVDSVAVLVPRAEIEGEMGDSHVGLQARLMSQALRKLSGTIAKTKTIAIFINQIREKVGVMF.GN
Lhelv   AYAEALGVDIDSLILSQPNTGEEGLQIADTLISSGAIDIVVVDSVAALVPRAEIEGEMGDAHVGLQARLMSQALRKLSGTISKTKTIAIFINQIREKVGIMF.GN
 Llac   EYAKALGVNIDELLLSQPDYGEQGLQIAEKLITSGAVDLVVIDSVAALVPKAEIDGEIGDSSVGLQARMMSQAMRKLAGHINKTKTTAIFINQLREKVGVMF.GS
Stherm  AYARALGVNIDELLLSQPDSGEQGLEIAGKLIDSGAVDLVVVDSVAAFVPRAEIDGDSGDSHVGLQARMMSQAMRKLSASINKTKTIAIFINQLREKVGIMF.GN
 Lmono  PYAKTLVVNIDELLLPQPDTGEQALEIAEALVRSGAVDMLVIDSVAALVPRAEIEGEMGDAHVGLQARLMSQALRKLSGVINKSKTIAIFINQIREKVGVMF.GN
Lmesan  KYAEALGVQKDELLLSQPDTGEQGLEIADALVQSGAVDIIVVDSVAALVPRAEIEGEMGDSHVGLQARLMSQALRKLAGTLNRTGTIAIFINQIREKIGVMF.GN
  Cbut  SYAQKLGVDVDSLIISQPDTGEQGLEIAEALVRSGAIDVLVVDSVAALVPKAEIEGEMGDSHIGLQARLMSQALRKLAGTINKTNCVAIFINQLREKLGFMF.GS
        **  * *    *   **  **  * *+     *** *  * ****   * *** *  ***  **+** **** ***          +***** *+* * **

consens  YA  LGV ID L LSQPDTGEQGLEIA  L  SGAVD  VVDSVAALVP AEIEGEMGDSHVGLQARLMSQALRKL G I K  T AIFINQIREKVG MF GN
```

FIG. 4. Comparison of the deduced amino acid sequences of the $recA_{int}$ fragments. The deduced amino acid sequences of the $recA_{int}$ fragments of *B. subtilis* (Bsub), *C. acetobutylicum* (Cbut), *E. coli* (Ecoli), *Lactobacillus bulgaricus* (Lbulg), *Lactobacillus helveticus* (Lhelv), *L. lactis* (Llac), *Leuconostoc mesanteroides* (Lmesan), *Listeria monocytogenes* (Lmono), *S. aureus* (Saur), and *S. thermophilus* (Stherm) were aligned by using the MULTALIN Multiple Sequence Alignments software of F. Corpet (1). Amino acid positions correspond to those of the *E. coli* RecA protein. A star indicates that the amino acids are identical, and a plus means that the amino acids are identical in the gram-negative species and not in *E. coli*. Amino acids in bold letters have been studied functionally in *E. coli* and are discussed in the text. A plus means that the corresponding amino acid is discussed in the text. A consensus amino acid is given only when at least seven amino acids are identical.

| PROT. \ DNA | Ecoli | Bsub | Llac | Cbut | Lbulg | Lhelv | Lmesan | Lmono | Saur | Stherm |
|---|---|---|---|---|---|---|---|---|---|---|
| Ecoli | - | 65 | 60 | 61 | 66 | 60 | 65 | 63 | 65 | 64 |
| Bsub | 73 | - | 72 | 66 | 69 | 68 | 67 | 74 | 74 | 68 |
| Llac | 69 | 79 | - | 68 | 66 | 70 | 67 | 73 | 75 | 78 |
| Cbut | 70 | 78 | 71 | - | 62 | 68 | 68 | 68 | 72 | 63 |
| Lbulg | 70 | 80 | 76 | 75 | - | 77 | 69 | 69 | 64 | 66 |
| Lhelv | 68 | 77 | 75 | 75 | 95 | - | 72 | 77 | 75 | 68 |
| Lmesan | 78 | 81 | 77 | 77 | 82 | 81 | - | 72 | 73 | 69 |
| Lmono | 73 | 85 | 78 | 76 | 77 | 77 | 79 | - | 78 | 71 |
| Saur | 74 | 83 | 79 | 79 | 81 | 81 | 81 | 81 | - | 70 |
| Stherm | 70 | 79 | 86 | 69 | 78 | 78 | 77 | 76 | 75 | - |

FIG. 5. Homology among the $recA_{int}$ sequences. $recA_{int}$ DNA sequences and deduced amino acid sequences from *B. subtilis* (Bsub), *C. acetobutylicum* (Cbut), *E. coli* (Ecoli), *Lactobacillus bulgaricus* (Lbulg), *Lactobacillus helveticus* (Lhelv), *L. lactis* (Llac), *Leuconostoc mesanteroides* (Lmesan), *Listeria monocytogenes* (Lmono), *S. aureus* (Saur), and *S. thermophilus* (Stherm) were compared. Results are given as percentages of strict sequence identity.

that our sequence is correct. The third amino acid, Thr-162, is not conserved in the gram-positive RecA proteins. It is known that PCR can introduce mutations into the DNA sequence (5); however, our previous *recA* cloning confirmed that PCR-amplified DNA and cloned chromosomal DNA had identical sequences (3). This sequence discrepancy is possibly due to strain differences.

It is of interest that the gram-positive $RecA_{int}$ fragments show more homology to each other than to gram-negative *E. coli* $RecA_{int}$ (Fig. 3). For example *E. coli* and *Lactobacillus bulgaricus* have similar G+C contents (51 and 52%) and show 70% $RecA_{int}$ sequence identity; the DNA of *C. acetobutylicum* is only 28% G+C, yet it shares 75% identity with $RecA_{int}$ of *Lactobacillus bulgaricus*. These data suggest that *recA* existed before the divergence between gram-negative and gram-positive bacteria, as previously proposed (9).

The eight $recA_{int}$ sequences provided here can be used to define the elements important for gram-positive RecA function. Several amino acids are known to be necessary for *E. coli* RecA activity by either mutational or structural analyses. The *E. coli* amino acids Leu-126, Gly-157, Gly-160, Arg-169, and Gly-200 (11) show identity with their gram-positive counterparts (with the exception of the RecA protein of *Listeria monocytogenes*, which has a proline residue at position 137, and that of *A. laidlawii*, which has a serine residue at position 160), and amino acids Ser-117, Glu-158, and Ala-179 (11) are conserved. However, two other amino acid positions differ between the gram groups. Thr-121 is conserved among the gram-negative but not the gram-positive RecA proteins. Although it appears to be essential (11), its role is unknown. Conversely, Cys-116 is known to be involved in the ATPase function of the *E. coli* RecA protein (6); nevertheless, it is poorly conserved in the gram-negative group, and among the gram-positive RecA proteins, a leucine residue (Ile in *C. acetobutylicum*) is always found in this position.

Additional amino acid positions, which are not characterized in *E. coli* RecA, may have functional importance in the gram-positive RecA proteins. Four such amino acids (Ala-129, Gln-167, Ala-189, and Glu-197) are identical in all of the gram-positive proteins but differ from those in *E. coli*. The first three are not conserved among the gram-negative RecA proteins. Glu-197 is replaced by a hydrophobic amino acid in all gram-negative RecA proteins. The presence of conserved but distinct amino acids for each group may reflect functions of these amino acids in RecA of gram-positive and/or gram-negative bacteria.

We have described a simple method to obtain an internal part of a *recA* gene from gram-positive bacteria. We are confident that this method can be used to isolate other gram-positive $recA_{int}$ gene segments or could be adapted to identify $recA_{int}$ of a gram-negative species. The isolation of an entire *recA* gene is also facilitated by this method. The amplified DNA fragment can be used as a probe for precise mapping of the entire *recA* gene. The appropriate DNA segment can thereby be selected for cloning and subsequently identified in colony hybridization by using the $recA_{int}$ DNA as the probe. This method has been successfully used in our laboratory for the cloning of the *L. lactis* subsp. *lactis* ML3 *recA* gene (3).

**Nucleotide sequence accession number.** The sequences of the *recA* PCR products of the following bacteria have been submitted to GenBank and assigned the indicated accession numbers: *C. acetobutylicum*, M94057; *Lactobacillus bulgaricus*, M94058; *Lactobacillus helveticus*, M94059; *Leuconostoc mesanteroides*, M94060; *S. aureus*, M94061; and *S. thermophilus*, M94062.

**Addendum.** A paper describing a similar method was published by Dybvig et al. (J. Bacteriol. **174**:2729–2732, 1992) after this paper was submitted.

We are grateful to P. Dabert, E. Maguin, and P. Langella for frequent discussion of this work.

Our research was supported in part by a grant from the European Economic Community (BRIDGE grant CT91-0263).

## REFERENCES

1. **Corpet, F.** 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16**:10881–10890.
2. **Davis, E. O., S. G. Sedwick, and J. M. Colston.** 1991. Novel structure of the *recA* locus of *Mycobacterium tuberculosis* implies processing of the gene product. J. Bacteriol. **173**:5653–5662.
3. **Duwat, P., S. D. Ehrlich, and A. Gruss.** PCR cloning and sequencing of the *L. lactis* subsp. *lactis recA* gene. Appl. Environ. Microbiol., in press.
4. **Dybvig, K., and A. Woodard.** 1992. Cloning and DNA sequence of a mycoplasmal *recA* gene. J. Bacteriol. **174**:778–784.
5. **Keohavong, X., and W. G. Thilly.** 1989. Fidelity of DNA polymerases in DNA amplification. Proc. Natl. Acad. Sci. USA **86**:9253–9257.
6. **Kuramitsu, S., K. Hamagushi, H. Tachibana, T. Horii, Y.**

**Hogawa, and H. Hogawa.** 1984. Cysteinyl residues of *Escherichia coli* RecA protein. Biochemistry **23**:2363–2367.
7. **Maniatis, T. E., E. F. Fritsch, and J. Sambrook.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Marrero, R., and R. E. Yasbin.** 1988. Cloning of the *Bacillus subtilis* $recE^+$ gene and functional expression of $recE^+$ in *B. subtilis*. J. Bacteriol. **170**:335–344.
9. **Miller, R. V., and T. A. Kokjohn.** 1990. General microbiology of *recA*: environmental and evolutionary significance. Annu. Rev. Microbiol. **44**:365–394.
10. **Owen, R. J.** 1982. National collection of type cultures, Central Public Health Laboratory, London.
11. **Roca, A. I., and M. M. Cox.** 1990. The RecA protein—structure and function. Crit. Rev. Biochem. Mol. Biol. **25**:415–456.
12. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Hullis, and H. A. Ehrlich.** 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. Science **239**:487–491.
13. **Sarkar, G., S. Kapelner, and S. S. Sommer.** 1990. Formamide can dramatically improve the specificity of PCR. Nucleic Acids Res. **18**:7465.
14. **Story, R. M., and T. A. Steitz.** 1992. Structure of the RecA protein-ADP complex. Nature (London) **355**:374–376.
15. **Yasbin, R. E., M. Stranathan, and K. W. Bayles.** 1991. The *rec*E(A)$^+$ gene of *B. subtilis* and its gene product: further characterization of this universal protein. Biochimie **73**:245–250.