JOURNAL OF BACTERIOLOGY, Dec. 1987, p. 5429–5433
0021-9193/87/125429-05$02.00/0

Vol. 169, No. 12

# Nucleotide Sequence of the *iap* Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in *Escherichia coli*, and Identification of the Gene Product

YOSHIZUMI ISHINO, HIDEO SHINAGAWA, KOZO MAKINO, MITSUKO AMEMURA, AND ATSUO NAKATA*

*Department of Experimental Chemotherapy, The Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565, Japan*

Received 1 May 1987/Accepted 22 August 1987

**The *iap* gene in *Escherichia coli* is responsible for the isozyme conversion of alkaline phosphatase. We analyzed the 1,664-nucleotide sequence of a chromosomal DNA segment that contained the *iap* gene and its flanking regions. The predicted *iap* product contained 345 amino acids with an estimated molecular weight of 37,919. The 24-amino-acid sequence at the amino terminus showed features characteristic of a signal peptide. Two proteins of different sizes were identified by the maxicell method, one corresponding to the Iap protein and the other corresponding to the processed product without the signal peptide. Neither the isozyme-converting activity nor labeled Iap proteins were detected in the osmotic-shock fluid of cells carrying a multicopy *iap* plasmid. The Iap protein seems to be associated with the membrane.**

Three isozymes of alkaline phosphatase from *Escherichia coli* have been identified as the prevalent species by polyacrylamide or starch gel electrophoresis when the enzyme is extracted from cells grown in Tris-glucose medium containing Casamino Acids or arginine (17, 20, 25). Based on the differences in the molecular structure of the three isozymes, it was suggested (6, 26) that isozyme formation is a post translational event, and that the amino-terminal arginine residues of the two polypeptides constituting isozyme 1 are removed proteolytically one by one, resulting in the production of isozymes 2 and 3. The conversion of these isozymes is mediated (presumably catalyzed) by the *iap* gene product (13, 14, 16, 17; for a review, see reference 15). Thus, the *iap* gene probably codes for a proteolytic enzyme, some kind of aminopeptidase.

We analyzed the nucleotide sequence of the *iap* gene to identify the primary structure of the putative protease and identified the protein encoded by the gene.

## MATERIALS AND METHODS

**Bacterial strains and the plasmids.** The *E. coli* strains used were AN234 (HfrC *phoT9 iap-1*) (17) to select the *iap*$^+$ plasmids, FE15 (F$^-$ *thr leu phoA thi rpsL*) (16) to prepare cell extracts used for isozyme conversions, JM103 [Δ(*pro-lac*) *supE thi*/F′ *traD36 proAB lacI*$^q$ Δ*lacZM15*] (12) as a host for bacteriophage M13, and CSR603 (*recA1 uvrA phr-1*) as a host for labeling the plasmid-coded proteins by the maxicell method (22).

Plasmid pSN143, which contains the *iap* gene of *E. coli*, was described by Nakata et al. (13). Plasmid vectors pUC8 and pUC9 and the bacteriophage M13mp19 were purchased from Pharmacia Biotechnology, Tokyo.

**Media.** The media used for the routine preparation of M13 phage and for the maxicell method were as previously described (1).

**DNA manipulation.** Plasmid and phage M13 replicative-form DNA were prepared by the method of Birnboim and Doly (2). Restriction endonuclease digestion, agarose, and polyacrylamide gel electrophoresis, in vitro ligation of DNA fragments with phage T4 DNA ligase, transformation, and transfection were done as described elsewhere (8, 9).

**Nucleotide sequencing.** The M13 phage was manipulated as described by Messing et al. (12). The 1.7-kilobase *Eco*RI-*Ava*I fragment containing the *iap* gene on pSN143 (13) was sequenced by the overlapping deletion method (5) as modified previously (1) with the dideoxy-nucleotide chain-termination method (23, 24) after cloning first into plasmid pUC9 in both orientations and then into M13mp19.

**Enzymes and radioisotopes.** The restriction endonucleases, T4 DNA ligase, T4 DNA polymerase, *Hin*dIII linker fragments, and an M13 nucleotide sequencing kit including DNA polymerase (Klenow fragment) were obtained from Takara Shuzo Co., Ltd., Kyoto, Japan. All enzymes were used as directed by the supplier. [α-$^{32}$P]dCTP and [$^{35}$S]methionine were purchased from Amersham Japan, Tokyo.

**Identification of the proteins encoded by plasmids.** The *iap* gene product was identified by the maxicell method (22). CSR603 cells containing plasmids were irradiated with UV light and incubated overnight in the presence of cycloserine (200 μg/ml). The cells were labeled with [$^{35}$S]methionine for 120 min after they had been starved of amino acids for 120 min. Proteins in the cell lysates were separated by electrophoresis on a 11.7% sodium dodecyl sulfate–polyacrylamide gel and visualized by fluorography.

**Other methods.** The procedure for the cold osmotic shock treatment of cells was as described previously (11). The isozyme conversion by cell extracts was done as described before (16).

## RESULTS

**DNA sequencing of *iap* gene.** A 1.7-kilobase *E. coli* chromosomal DNA fragment flanked with *Eco*RI and *Ava*I restriction sites and containing the *iap*$^+$ gene (13) was sequenced (Fig. 1). The *iap*$^+$ gene and its flanking regions were analyzed with deletions of various lengths from one end of the inserted DNA segments as templates. The DNA sequenced covered this region at least twice on both strands with overlapping junctions.

The largest translational open reading frame contains 1,038 bases from the ATG codon at nucleotide 332 to the

* Corresponding author.

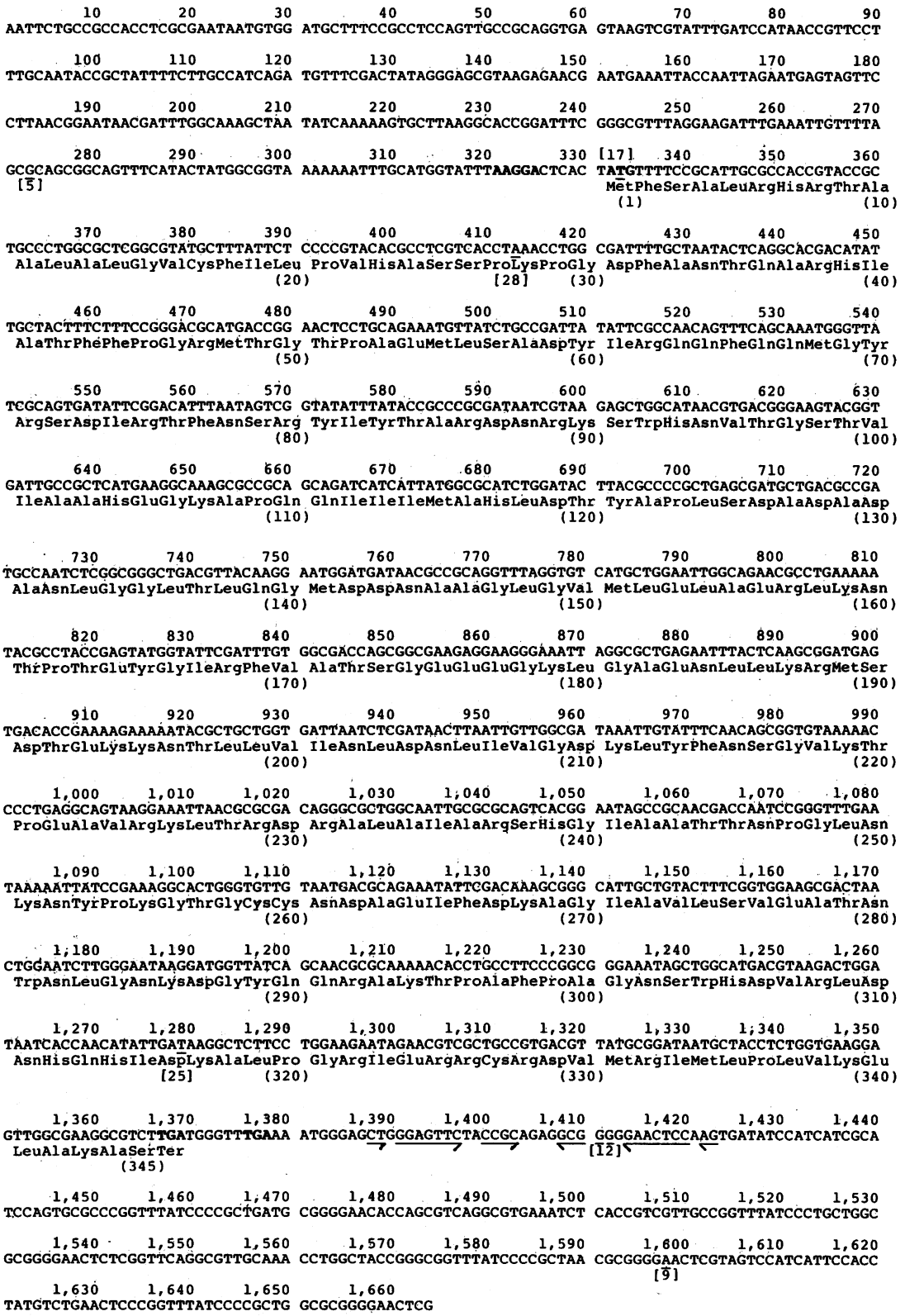

FIG. 1. Nucleotide sequence of the *iap* gene and flanking regions and amino acid sequences of putative Iap protein deduced from the nucleotide sequence. Nucleotides are numbered with the second nucleotide of the *Eco*RI endonuclease recognition sequence taken as 1. The putative translational initiation and termination codons and putative ribosome-binding site are in boldface type. The nucleotide with a transcript that may form a stable stem-and-loop structure in the 3′-flanking region of *iap* gene is shown by arrows. The endpoints of deletions in the 5′- and 3′-end regions of the chromosomal DNA inserts of the phage clones which were used for reconstruction of *iap* plasmids are shown by bracketed numbers and underlined. The reconstruction was done by using the *Pst*I restriction site at nucleotide 486. Each reconstructed DNA fragment was recloned into either pUC8 or pUC9 downstream of the *lac* promoter.

TGA codon at nucleotide 1367 (Fig. 1). This open reading frame can code for a protein of 345 amino acids with a molecular weight of 37,919.

To check the coding region of *iap*, we constructed a plasmid containing a minimum coding region of the *iap* gene. We selected six phage clones used for DNA sequencing. They contained deletions extending to or across the 5′ or 3′ end of the putative coding region as shown in Fig. 1, and we reconstructed the plasmids with deletions in both the 5′ and 3′ ends of the chromosomal DNA fragment. The reconstructed plasmids were introduced into an Iap$^-$ strain. Only two clones that contained the 5′-end fragment derived from the no. 5 clone and the 3′-end fragment derived either from the no. 9 clone (pIN9509) or the no. 12 clone (pIN9512) transformed the *iap* mutant to Iap$^+$. The 5′ end of the coding region was located between nucleotides 273 and 333, and the 3′ end was between nucleotides 1338 and 1412. The complementation tests agreed well with the predicted coding region of the *iap* gene as shown in Fig. 1.

The putative ribosome-binding site, AAGGA, is six nucleotides before the translational initiation codon, ATG (3, 7, 27).

**Isozyme conversion by cell extracts.** Isozyme conversion should occur in the periplasmic space where alkaline phosphatase is localized. Since the predicted amino-terminal 24 amino acid residues of the Iap protein had features characteristic of a signal peptide, experiments were done to determine the localization of the *iap* gene product. This was done in strain FE15 bacteria that contained the pIN9512 plasmid. Isozyme-converting activity was detectable in sonicated extracts prepared from whole cells or osmotically shocked cells of strain FE15 with the plasmid (Fig. 2, lanes 2 and 4). Since no activity was detectable in the shocked fluid (Fig. 2, lane 3), the Iap protein is probably not found in the periplasm.

**Identification of the *iap* gene product.** To identify the *iap* gene product, the plasmid-encoded proteins were labeled with [$^{35}$S]methionine by the maxicell method. Two proteins with approximate molecular weights of 38,000 and 41,000 were observed in the maxicells carrying the *iap*$^+$ plasmids (Fig. 3, lanes 2 and 7). These proteins were not detected in the cells carrying the *iap* plasmid (Fig. 3, lanes 1, 4, and 5).

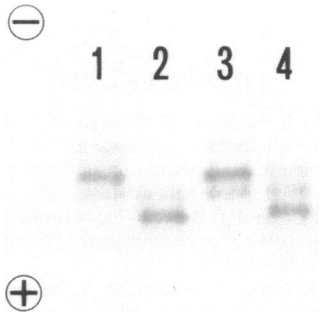


FIG. 2. Isozyme conversion by cell extracts. The reaction mixture (0.2 ml) contained 50 mM Tris hydrochloride (pH 7.5), 20 mM $MgCl_2$, 0.02% azide, 10 μl of alkaline phosphatase isozyme 1 (2.8 enzyme units), and extracts of FE15 cells containing the pIN9512 (*iap*$^+$) plasmid. The reaction mixture was incubated at 37°C overnight. After incubation, it was heated at 80°C for 15 min, and then 50 μl of glycerol containing phenol red (1%) was added to each sample. A 20-μl sample was electrophoresed on a 7.5% polyacrylamide gel. The gel was stained with a mixture of naphthol-AS-MX-phosphate and Fast Blue RR salt (Sigma Chemical Co., St. Louis, Mo.) (10). Lanes contained isozyme 1 with the following: 1, without cell extract (control); 2, with the supernatant of low-speed centrifugation (8,000 × *g*, 15 min) of sonicated whole cells; 3, with the supernatant of low-speed centrifugation of osmotically shocked cells (11); 4, with the supernatant of low-speed centrifugation of the sonicated cells from which the periplasmic fraction had been removed by osmotic shock.

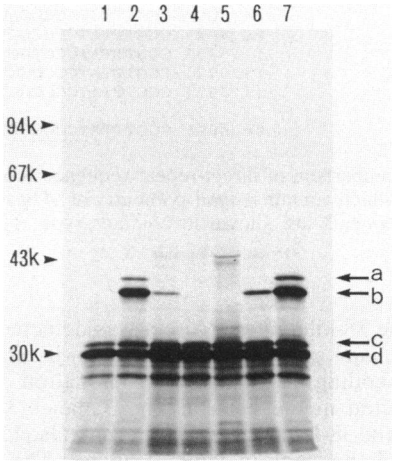


FIG. 3. Identification of the *iap* gene product in maxicells (22). [$^{35}$S]methionine-labeled proteins produced in CSR603 carrying plasmids were separated by electrophoresis on a 11.7% sodium dodecyl sulfate–polyacrylamide gel and visualized by fluorography. Cells carried plasmids pUC9 (lane 1), pIN9054 (*iap*$^+$) (lane 2), pIN8005 (*iap*$^+$) (lane 3), pIN8017 (*iap*) (lane 4), pIN9025 (*iap*) (lane 5), pIN9012 (*iap*$^+$) (lane 6), and pIN9512 (*iap*$^+$) (lane 7). The *Eco*RI-*Hin*dIII fragments of the clones (nos. 5 and 17, Fig. 1) with deletions at the 5′-end region were recloned into pUC8, giving pIN8005 and pIN8017, and those of the clones (nos. 12 and 25, Fig. 1) with deletions at the 3′ end were recloned into pUC9, giving pIN9012 and pIN9025. a and b, *iap* gene product and its presumptive processed product, respectively; c and d, precursor of β-lactamase and β-lactamase, respectively.

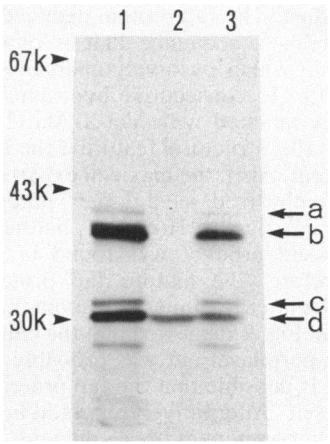


FIG. 4. Location of the Iap protein in the maxicells. Cells of CSR603 carrying pIN9054 labeled with [$^{35}$S]methionine were mixed with nonlabeled FE15 cells carrying the same plasmid and then treated by cold osmotic shock. After centrifugation at low speed, the supernatant and cell pellets obtained were treated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and fluorography. Lanes: 1, total cell fraction; 2, osmotic shock fluid; 3, cell fraction after osmotic shock treatment. a and b, *iap* gene product and its presumptive processed product, respectively; c and d, precursor of β-lactamase and β-lactamase, respectively.

```
TGAAAATGGGAGGGAGTTCTACCGCAGAGGCGGGGGAACTCCAAGTGATATCCATCATCGCATCCAGTGCGCC (1,451)
    (1,452) CGGTTTATCCCCGCTGATGCGGGGAACACCAGCGTCAGGCGTGAAATCTCACCGTCGTTGC (1,512)
    (1,513) CGGTTTATCCCTGCTGGCGCGGGGAACTCTCGGTTCAGGCGTTGCAAACCTGGCTACCGGG (1,573)
    (1,574) CGGTTTATCCCCGCTAACGCGGGGAACTCGTAGTCCATCATTCCACCTATGTCTGAACTCC (1,634)
    (1,635) CGGTTTATCCCCGCTGGCGCGGGGAACTCG (1,664)

consensus: CGGTTTATCCCCGCT(GG/AA)CGCGGGGAACTC
```

FIG. 5. Comparison of direct-repeat sequences consisting of 61 base pairs in the 3'-end flanking region of *iap*. The 29 highly conserved nucleotides, which contain a dyad symmetry of 14 base pairs (underlined), are shown at the bottom. Homologous nucleotides found in at least two DNA segments are shown in boldface type. The second translational termination codon is boxed. The nucleotide numbers are in parentheses.

Although the 38-kilodalton protein was detected in the cells carrying the *iap*$^+$ plasmid with a deletion in either the 5' or 3' end noncoding region, the 41-kilodalton protein was barely detected in the cells in Fig. 3, lanes 3 and 6. The amounts of the 38-kilodalton protein were also lower in these cells than in the ones in lanes 2 and 7. The 41- and 38-kilodalton proteins are probably the nascent Iap protein and its processed product, respectively, since both were detected only in cells carrying the *iap*$^+$ gene and since their sizes corresponded roughly to the *iap* gene product and the processed product, respectively, predicted from the DNA sequence.

To examine whether the processed Iap protein was secreted into the periplasm, we fractionated the maxicells that had been treated by osmotic shock into a supernatant and pellet fraction. Neither the 41-kilodalton protein nor the 38-kilodalton protein was detected in the periplasmic fraction (Fig. 4, lane 2); both were found in the pellet fraction (lane 3). In this experiment, the β-lactamase precursor and its processed product served as the internal controls. The processed product but not the precursor was detected in the periplasmic fraction (lane 2); the precursor was found in the pellet fraction (lane 3). Therefore, it is likely that both the Iap protein and its processed product are in the membrane fraction.

## DISCUSSION

***iap* gene product.** The Iap protein deduced from the DNA sequence contains a sequence that is characteristic of a signal peptide in which positively charged amino acids are followed by 10 to 15 consecutive hydrophobic amino acids and which is terminated with Val-X-Ala (19). It would be consistent with this structural feature if the Iap protein were a secreted protein. Also, the maxicell experiments suggested that it was synthesized as a larger precursor and then processed proteolytically. However, neither the precursor nor the processed product was found in the periplasmic fraction. Therefore, the mature Iap protein is probably associated with either the inner or the outer membrane. The results shown in Fig. 2 suggested that the conversion activity was not in the periplasm but was probably associated with membranes. It is possible that the Iap protein is the conversion enzyme itself. Alternatively, it could be an activator of the enzyme or a component of the enzyme.

**Structural features of the noncoding regions.** Several candidates as the promoter of *iap* were found in the upstream region of the gene. Among them, either TTGAaA at nucleotide 257 for the −35 sequence in combination with TtTcAT for the −10 sequence at nucleotide 283 or TTaACg at nucleotide 182 for the −35 sequence with TAatAT at nucleotide 208 for the −10 sequence may be the promoter for *iap* (3, 4, 21).

A translational termination codon, TGA, was found at both nucleotide 1367 and nucleotide 1376, in the same reading frame.

With nine nucleotides spacing from the second terminator codon, a nucleotide sequence with a transcript that may form a stable stem-and-loop structure was found. Transcription may end at this region.

An unusual structure was found in the 3'-end flanking region of *iap* (Fig. 5). Five highly homologous sequences of 29 nucleotides were arranged as direct repeats with 32 nucleotides as spacing. The first sequence was included in the putative transcriptional termination site and had less homology than the others. Well-conserved nucleotide sequences containing a dyad symmetry, named REP sequences, have been found in *E. coli* and *Salmonella typhimurium* (28) and may act to stabilize mRNA (18). A dyad symmetry with 14 nucleotide pairs was also found in the middle of these sequences (underlining, Fig. 5), but no homology was found between these sequences and the REP sequence. So far, no sequence homologous to these has been found elsewhere in procaryotes, and the biological significance of these sequences is not known.

## ACKNOWLEDGMENTS

We are grateful to K. Hashino for technical assistance.

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan.

## LITERATURE CITED

1. **Amemura, M., K. Makino, H. Shinagawa, A. Kobayashi, and A. Nakata.** 1985. Nucleotide sequence of the genes involved in phosphate transport and regulation of the phosphate regulon in *Escherichia coli*. J. Mol. Biol. **184**:241–250.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1980. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
4. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
5. **Hong, G. F.** 1982. A systematic DNA sequencing strategy. J. Mol. Biol. **158**:539–549.
6. **Kelley, P. M., P. A. Neumann, K. Shriefer, F. Cancedda, M. J. Schlesinger, and R. A. Bradshaw.** 1973. Amino acid sequence of *Escherichia coli* alkaline phosphatase. Amino- and carboxyl-terminal sequences and variations between two isozymes. Biochemistry **12**:3499–3503.
7. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**: 1–45.
8. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K-12. J.

Mol. Biol. **190**:37–44.
9. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 55. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Morita, T., M. Amemura, K. Makino, H. Shinagawa, K. Magota, N. Otsuji, and A. Nakata.** 1983. Hyperproduction of phosphate-binding protein, *phoS*, and pre-*phoS* proteins in *Escherichia coli* carrying a cloned *phoS* gene. Eur. J. Biochem. **130**:427–435.
12. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
13. **Nakata, A., H. Shinagawa, and M. Amemura.** 1982. Cloning of alkaline phosphatae isozyme gene (*iap*) of *Escherichia coli*. Gene **19**:313–319.
14. **Nakata, A., H. Shinagawa, and J. Kawamata.** 1979. Inhibition of alkaline phosphatase isozyme conversion by protease inhibitors in *Escherichia coli* K-12. FEBS Lett. **105**:147–150.
15. **Nakata, A., H. Shinagawa, and F. G. Rothman.** 1987. Molecular mechanism of isozyme formation of alkaline phosphatase in *Escherichia coli*, p. 139–141. *In* A. Torriani, F. Rothman, S. Silver, A. Wright, and E. Yagil (ed.), Phosphate metabolism and cellular regulation in microorganisms. American Society for Microbiology, Washington, D.C.
16. **Nakata, A., H. Shinagawa, and H. Shima.** 1984. Alkaline phosphatase isozyme conversion by cell-free extract of *Escherichia coli*. FEBS Lett. **175**:343–348.
17. **Nakata, A., M. Yamaguchi, K. Izutani, and M. Amemura.** 1978. *Escherichia coli* mutants deficient in the production of alkaline phosphatase isozymes. J. Bacteriol. **134**:287–294.
18. **Newburg, S. F., N. H. Smith, E. C. Robinson, I. E. Hiles, and C. F. Higgins.** 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell **48**:297–310.
19. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39**:615–648.
20. **Piggot, P. J., M. D. Sklar, and L. Gorini.** 1972. Ribosomal alterations controlling alkaline phosphatase isozymes in *Escherichia coli*. J. Bacteriol. **110**:291–299.
21. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
22. **Sancar, A., R. P. Wharton, S. Seltzer, B. M. Kacinski, N. D. Clarke, and W. D. Rupp.** 1981. Identification of the *uvrA* gene product. J. Mol. Biol. **148**:45–62.
23. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **Schlesinger, M. J., and L. Andersen.** 1968. Multiple molecular forms of the alkaline phosphatase of *Escherichia coli*. Ann. N.Y. Acad. Sci. **151**:159–170.
26. **Schlesinger, M. J., W. Bloch, and P. M. Kelley.** 1975. Differences in the structure, function, and formation of two isozymes of *Escherichia coli* alkaline phosphatase, p. 333–342. *In* C. L. Markert (ed.), Isozymes I. Molecular structure. Academic Press, Inc., New York.
27. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
28. **Stern, M. J., G. F.-L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37**:1015–1026.