JOURNAL OF BACTERIOLOGY, Oct. 1980, p. 444-446
0021-9193/80/10-0444/03$02.00/0

Vol. 144, No. 1

Nucleotide Sequence Homology Between the Heat-Labile Enterotoxin Gene of *Escherichia coli* and *Vibrio cholerae* Deoxyribonucleic Acid

STEVE L. MOSELEY* AND STANLEY FALKOW

Department of Microbiology and Immunology, SC-42, University of Washington, Seattle, Washington 98195

Isolated deoxyribonucleic acid fragments encoding the heat-labile enterotoxin of *Escherichia coli* were used to probe for homologous sequences in restricted whole-cell deoxyribonucleic acid from *Vibrio cholerae*. Significant sequence homology between the heat-labile enterotoxin gene and *V. cholerae* deoxyribonucleic acid was demonstrated, and apparent differences were observed in the organization of the cholera toxin gene among different strains of *V. cholerae*.

Vibrio cholerae and some strains of *Escherichia coli* cause diarrheal disease in humans by similar mechanisms: colonization of the small intestine and production of enterotoxin. It has been well established that one of the enterotoxins of *E. coli*, the heat-labile toxin (LT), shares immunological determinants with cholera toxin (8) and has the same mechanism of action, the activation of adenyl cyclase (7). Recent immunological (2), biochemical (3), and genetic (4) data strongly suggest that LT and cholera toxin possess similar subunit structure consisting of an enzymatically active protein (A subunit) and a binding protein (B subunit) which mediates the adsorption of the toxin molecule to the target cell. Such striking similarities between the two toxins could conceivably result from the convergent evolution of two different proteins, but a common evoluntary origin is a more appealing explanation. We report here the detection of significant nucleotide sequence homology between DNA encoding LT and *V. cholerae* DNA.

The gene encoding LT has been isolated from porcine strain P307 and characterized by Dallas et al. (5). The gene encodes two proteins: a 25,500-dalton protein possessing adenyl cyclase-activating activity (LT-A) and an 11,500-dalton protein possessing adsorption activity for Y-1 adrenal cells (LT-B). The regions of the gene encoding each of these subunits can be isolated on specific restriction endonuclease-generated fragments (Fig. 1). We employed a 1,275-base pair *Hinc*II fragment encoding the LT-A subunit and a 590-base pair *Eco*RI-*Hin*dIII fragment encoding the LT-B subunit as probes for sequence homology with *V. cholerae* DNA.

Cellular DNA was prepared by the method of Brenner et al. (1) from *V. cholerae* classical biotype strains 569B and 14035 and El Tor biotype strain 30167 and from *E. coli* strains C600 and C600 (P307). The DNA was digested to completion with *Eco*RI, and the resulting fragments were separated by agarose gel electrophoresis. The fragments were transferred to nitrocellulose by the method of Southern (14) and hybridized to radiolabeled LT probes. The nitrocellulose was then washed and exposed to X-ray film for detection of probe DNA duplexed with homologous sequences on the nitrocellulose. The degree of homology required for stable duplex formation depends on the stringency of the hybridization conditions, i.e., the temperature at which the hybridization is carried out and the ionic strength of the hybridization solution. Highly stringent conditions will allow stable duplex formation between sequences of very close homology, but such conditions are not necessarily appropriate for examining the relationship between DNAs whose sequences may have diverged from a common evolutionary origin. We therefore used varying degrees of stringency in hybridizations to detect homology between the LT gene and *V. cholerae* DNA. The most stringent hybridizations were carried out in 5 × SSC (1 × SSC: 0.15 M sodium chloride, 0.015 M sodium citrate) at 37°C in 50% formamide. Formamide lowers the T_m of DNA duplex molecules approximately 0.7°C for each percent formamide in the hybridization solution (12). Assuming a guanine plus cytosine content of 40 to 45% based on preliminary sequence data of the LT gene (Dallas and Falkow, in preparation), the calculated T_m of completely homologous sequences under these conditions would be approximately 65 to 68°C. Assuming that each percent nucleotide mismatch lowers the T_m of the duplex by 1.4°C (10), these conditions would allow a maximum mismatch of approximately 20%. Reduced-stringency hybridizations were carried out under the same conditions of temperature and ionic strength in reduced formamide concentrations. In 25% formamide, a maxi-

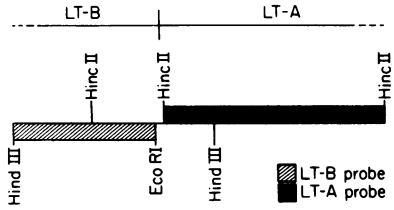

FIG. 1. *LT encoding region of plasmid EWD299 (5). Probes were prepared by polyacrylamide gel electrophoresis of restriction endonuclease digests of EWD299, followed by electroelution of the appropriate fragments from the polyacrylamide. Isolated fragments were ^{32}P-labeled in vitro by nick translation (11) to a specific activity of 2.5 × 10^7 to 5 × 10^7 cpm/μg of DNA.*

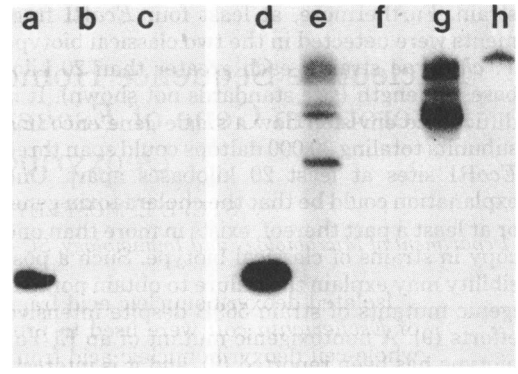

FIG. 2. *Autoradiograph showing hybridization of the LT-A probe to EcoRI fragments of V. cholerae and E. coli DNA. Cellular DNA was prepared by the method of Brenner et al. (1) and digested with EcoRI. Approximately 1 μg of each DNA digest was separated by electrophoresis on a 0.7% agarose gel. The fragments were denatured and transferred to nitrocellulose by the method of Southern (14). The nitrocellulose was incubated at 37°C for 3 h in the following hybridization solution: 25% or 50% formamide, 5× SSC, 1 mM EDTA, 0.1% sodium dodecyl sulfate, and 1× Denhardt's solution (6). The nitrocellulose was then incubated for 36 h in fresh hybridization solution containing 8 × 10^5 cpm of ^{32}P-labeled probe DNA (denatured) per ml and 100 μg of sheared, denatured calf thymus DNA per ml. The nitrocellulose was then washed in 5× SSC–0.1% sodium dodecyl sulfate for 45 min at 54.5°C or 65°C. The nitrocellulose was then rinsed in 2× SSC at room temperature, allowed to dry, and exposed to X-ray film for 24 h. Tracks (a) through (c) were hybridized in 50% formamide and washed at 65°C. Tracks (d) through (h) were hybridized in 25% formamide and washed at 54.5°C. (a) C600 (P307); (b) 569B; (c) C600; (d) C600 (P307); (e) 569B; (f) C600; (g) 14035; (h) 30167.*

mum mismatch of approximately 33% would be allowed under these conditions. In 20% formamide, the maximum mismatch in stable duplex molecules would be approximately 36%.

The results of the hybridizations with the LT-A probe are shown in Fig. 2. Under stringent conditions (50% formamide), homology with the LT-A probe was detected only in the *E. coli* C600 (P307) DNA. This strain carries the Ent plasmid from which the LT probes were derived. Under conditions of reduced stringency, however, the LT-A probe detected homologous sequences in the *V. cholerae* DNA as well as the C600 (P307) DNA. Fragments identical in size to those detected by the LT-A probe were also detected by the LT-B probe in the *V. cholerae* DNAs if the hybridization was carried out in 20% formamide (data not shown). Therefore, some sequence homology exists between the LT gene and *V. cholerae* DNA. However, since the stringency must be reduced for stable duplex formation between the LT and *V. cholerae* DNA, significant sequence divergence must have occurred since a presumed exchange of the enterotoxin gene between the two organisms. These data do not permit a precise determination of the degree of sequence homology between the LT and cholera toxin genes. Predictions of percent mismatch from hybridization analyses are influenced by the distribution of mismatched base pairs, the nucleotide composition of each mismatch, and the guanine plus cytosine composition of homologous sequences neighboring the mismatched base pairs (15). Therefore, direct sequence analyses have not always borne out the percent homology predicted from hybridization analyses (13, 15), and it is likely that the influence of local guanine plus cytosine content differences within the hybrid molecules would be increased with the small probes used in the present study. With these limitations in mind, it is still possible to obtain a rough estimate of the degree of homology between the LT and cholera toxin genes from these data. A percent homology on the order of 70% can be predicted from the hybridization conditions required for detection of homology between the LT probes and *V. cholerae* DNA. The fact that a somewhat lower stringency was required for detection of homology with the LT-B probe than with the LT-A probe (20% formamide for LT-B versus 25% for LT-A) could indicate differences in degrees of homology of the two probes with the cholera toxin gene. It is as likely, however, that differences in guanine plus cytosine content and distribution between the two probes are responsible for this observation.

The hybridization data obtained here raise several interesting questions concerning organization of the cholera toxin gene. Different patterns of restriction fragments were detected by the LT probes in each of the strains examined. This suggests that the sequence environment of the cholera toxin gene differs from strain to

strain. Furthermore, at least four *Eco*RI fragments were detected in the two classical biotype *V. cholerae* strains, each greater than 20 kilobases in length (size standards not shown). It is difficult to envision how a single gene encoding subunits totaling 30,000 daltons could span three *Eco*RI sites at least 20 kilobases apart. One explanation could be that the cholera toxin gene, or at least a part thereof, exists in more than one copy in strains of classical biotype. Such a possibility may explain the failure to obtain nontoxigenic mutants of strain 569B despite intensive efforts (9). A nontoxigenic mutant of an El Tor biotype has been reported (9), and it is interesting to note that only a single *Eco*RI fragment is detected by the LT probes in the El Tor strain studied here, as well as other El Tor strains we have examined. Finally, identically sized fragments were detected by both the LT-A and LT-B probes in a given strain. This suggests that portions of the cholera toxin gene encoding the A subunit are contiguous with portions encoding the B subunit, as in the case of *E. coli*.

Demonstration of nucleotide sequence homology between the LT gene and *V. cholerae* DNA establishes the evolutionary relatedness of the two enterotoxins. This homology will permit the use of LT gene fragments to probe the structure and organization of the cholera toxin gene. The LT probe should also prove useful in determining the genetic potential of toxin production by strains of vibrios of questionable toxigenicity, such as strains of atypical biotypes and environmental isolates.

This work was supported by a Public Health Service grant from the National Institutes of Health, AI-10885-08, and S.M. was supported by a Public Health Service training grant from the National Institutes of Health, National Research Service Award AI-07149-02.

LITERATURE CITED

1. **Brenner, D. J., G. R. Fanning, K. E. Johnson, R. V. Citarella, and S. Falkow.** 1969. Polynucleotide relationships among members of *Enterobacteriaceae*. J. Bacteriol. **98**:637–650.
2. **Clements, J. D., and R. A. Finkelstein.** 1978. Demonstration of shared and unique immunological determinants in enterotoxins from *Vibrio cholerae* and *Escherichia coli*. Infect. Immun. **22**:709–713.
3. **Clements, J. D., and R. A. Finkelstein.** 1979. Isolation and characterization of homogenous heat-labile enterotoxins with high specific activity from *Escherichia coli*. Infect. Immun. **24**:760–769.
4. **Dallas, W. S., and S. Falkow.** 1979. The molecular nature of heat-labile enterotoxin (LT) of *Escherichia coli*. Nature (London) **277**:406–407.
5. **Dallas, W. S., D. M. Gill, and S. Falkow.** 1979. Cistrons encoding *Escherichia coli* heat-labile toxin. J. Bacteriol. **139**:850–858.
6. **Denhardt, D. T.** 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641–646.
7. **Evans, D. J., L. C. Chen, D. T. Curlin, and D. G. Evans.** 1972. Stimulation of adenyl cyclase by *Escherichia coli* enterotoxin. Nature (London) New Biol. **236**:137–138.
8. **Gyles, C. L., and D. A. Barnum.** 1969. A heat labile enterotoxin from strains of *Escherichia coli* enteropathogenic for pigs. J. Infect. Dis. **120**:419–426.
9. **Honda, T., and R. A. Finkelstein.** 1979. Selection and characteristics of a *Vibrio cholerae* mutant lacking the A (ADP-ribosylating) portion of the cholera enterotoxin. Proc. Natl. Acad. Sci. U.S.A. **76**:2052–2056.
10. **Hyman, R. W., I. Brunovskis, and W. C. Summers.** 1973. DNA base sequence homology between coliphages T7 and ϕII and between T3 and ϕII as determined by heteroduplex mapping in the electron microscope. J. Mol. Biol. **77**:189–196.
11. **Maniatis, T., A. Jeffrey, and A. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage lambda. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
12. **McConaughy, B. L., C. D. Laird, and B. J. McCarthy.** 1969. Nucleic acid reassociation in formamide. Biochemistry **8**:3289–3295.
13. **Nichols, B. P., and C. Yanofsky.** 1979. Nucleotide sequences of *trpA* of *Salmonella typhimurium* and *Escherichia coli*: an evolutionary comparison. Proc. Natl. Acad. Sci. U.S.A. **76**:5244–5248.
14. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
15. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA sequence: extent of homology with simian virus 40 DNA. Proc. Natl. Acad. Sci. U.S.A. **76**:1179–1183.