JOURNAL OF BACTERIOLOGY, Mar. 1993, p. 1860–1864
0021-9193/93/061860-05$02.00/0

Vol. 175, No. 6

Characterization of a Plasmid-Encoded Urease Gene Cluster Found in Members of the Family *Enterobacteriaceae*

SARAH E. F. D'ORAZIO AND CARLEEN M. COLLINS*

Department of Microbiology and Immunology, University of Miami School of Medicine, P.O. Box 016960, Miami, Florida 33101

Received 29 September 1992/Accepted 8 January 1993

Plasmid-encoded urease gene clusters found in uropathogenic isolates of *Escherichia coli*, *Providencia stuartii*, and *Salmonella cubana* demonstrated DNA homology, similar positions of restriction endonuclease cleavage sites, and manners of urease expression and therefore represent the same locus. DNA sequence analysis indicated that the plasmid-encoded urease genes are closely related to the *Proteus mirabilis* urease genes.

Urease acts as a virulence determinant for uropathogenic bacteria by catalyzing the hydrolysis of urea to ammonia and carbon dioxide, thus producing alkaline conditions in the urinary tract. Increased urine pH can lead to formation of struvite stones that harbor the infecting organism (15, 17), enhanced attachment of bacteria to the renal epithelium (23), direct renal tissue damage (16), and inactivation of certain antibiotics (7). As a result, urease-producing uropathogens are better able to colonize and survive in the urinary tract than are isogenic urease-negative variants (8, 10).

Urease gene clusters have been isolated from several members of the family *Enterobacteriaceae*, including *Proteus mirabilis* (12), *Proteus vulgaris* (20), *Providencia stuartii* (18, 22), *Morganella morganii* (9), *Escherichia coli* (4), *Klebsiella pneumoniae* (5), and *Klebsiella aerogenes* (21). Each of the clusters that has been DNA sequenced thus far contains *ureA*, *ureB*, and *ureC*, which encode the urease structural subunits. Four additional genes, *ureD*, *ureE*, *ureF*, and *ureG*, have been found in *K. aerogenes* (21), *K. pneumoniae* (5), and *E. coli* (4). The exact function of the polypeptides encoded by *ureD*, *ureE*, *ureF*, and *ureG* has yet to be determined, although these polypeptides are thought to be involved in the acquisition of nickel by urease (13, 21). Urease is a nickel metalloenzyme, and a mutation in either *ureD*, *ureE*, *ureF*, or *ureG* results in formation of a nickel-deficient, inactive urease apoenzyme.

Clinical isolates of urease-positive *E. coli* contain one of two distinct urease gene clusters (3). The urease genes of approximately one-half of the urease-positive clinical isolates are thought to be chromosomally encoded; the urease genes of the remaining clinical isolates are located on large plasmids. The chromosomally encoded and the plasmid-encoded urease loci differ in their restriction endonuclease maps and in the sizes of the polypeptides they encode. The two gene clusters do not demonstrate DNA homology under stringent hybridization conditions. Recent DNA sequence analysis also indicates that these two loci are distinct (6).

Previously, we demonstrated that the urease genes of uropathogenic, urease-positive *E. coli* 1440, were found on a 160-kb plasmid. The urease genes of *E. coli* 1440 were isolated and inserted into pUC19, resulting in recombinant plasmid pURE14.8. Under stringent hybridization conditions, a urease gene probe derived from pURE14.8 showed DNA homology with large plasmids from one-half of the urease-positive *E. coli* isolates tested and all of the urease-producing *Salmonella* and *P. stuartii* strains tested (two *Salmonella* and five *P. stuartii* strains were examined). The various urease-encoding plasmids differed in size (from 82 to approximately 230 kb) and coded for distinct biochemical traits, suggesting that they were not identical plasmids. However, Southern analysis indicated that the urease loci associated with these plasmids were similar.

Urease loci from *E. coli*, *P. stuartii*, and *S. cubana* have similar restriction endonuclease maps. Southern hybridization analysis was performed to assess the degree of relatedness between the various plasmid-encoded urease loci. Total DNAs from the ureolytic strains *E. coli* 1440, *P. stuartii* JH3, and *S. cubana* 3237-71 were digested with *Eco*RI and probed with a 3.4-kb ^{32}P-labeled *Xba*I-*Cla*I fragment of pURE14.8. The plasmid-encoded urease gene cluster was found on identical 14.0-kb *Eco*RI fragments in all three species (Fig. 1A).

To demonstrate further the similarity between the plasmid-encoded urease gene clusters, restriction endonuclease cleavage site profiles were generated. Total cell DNA from each of the three species was digested with various combinations of restriction enzymes and probed with one of three different ^{32}P-labeled DNA fragments of pURE14.8: a 0.5-kb *Pvu*II-*Stu*I fragment that contains portions of *ureD* and *ureA*, a 0.8-kb *Cla*I-*Hin*dIII fragment that contains portions of *ureF* and *ureG*, and a 6.7-kb *Hin*dIII fragment that spans the entire gene cluster. Representative Southern blots are shown in Fig. 1B and C. The results from the Southern hybridizations enabled us to construct restriction maps for the *P. stuartii* and *S. cubana* urease loci. These maps were compared with the previously determined restriction map of the *E. coli* 1440 urease gene cluster (Fig. 2) (3). In regard to the restriction endonuclease cleavage sites determined, the three loci were indistinguishable. The maps presented here are in close agreement with the previously published restriction endonuclease map of the *P. stuartii* BE2467 urease gene cluster (22).

Expression of urease from the plasmid-encoded locus is induced by urea. Expression of urease activity from *E. coli* 1440 is induced in the presence of urea (3). *P. stuartii* JH3 and *S. cubana* 3237-71 were examined for similarity in induction of urease activity. Urease gene expression was increased at least 40-fold in the presence of urea for each of the above strains (Table 1). In contrast, *E. coli* 1021, which

* Corresponding author.

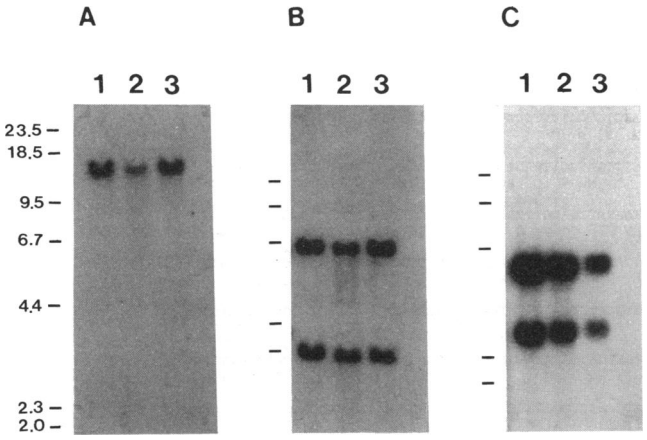

FIG. 1. The plasmid-encoded urease gene clusters share DNA homology. Southern hybridizations were performed under stringent conditions. Total cell DNAs from *E. coli* 1440 (lane 1), *S. cubana* 3237-71 (lane 2), and *P. stuartii* JH3 (lane 3) were probed with ^{32}P-labeled fragments of pURE14.8. (A) *Eco*RI-digested DNA probed with the 3.4-kb *Xba*I-*Cla*I fragment. (B) *Hin*dIII-*Nru*I-digested DNA probed with the 6.7-kb *Hin*dIII fragment. (C) *Hin*dIII-*Ava*I-digested DNA probed with the 6.7-kb *Hin*dIII fragment. Molecular size markers (in kilobases) are indicated on the left for panel A. The dashes to the left of panels B and C represent markers of 9.5, 6.7, 4.4, 2.3, and 2.0 kb, respectively, from the top.

has a chromosomally encoded urease locus, was constitutive for urease expression.

It appears that the urease genes found in these distinct bacterial species represent the same urease locus, on the basis of the DNA homology detected between the loci, the similarities in the restriction maps, and the similar manners of urease expression (induction in the presence of urea). We refer to this locus as the plasmid-encoded urease gene cluster.

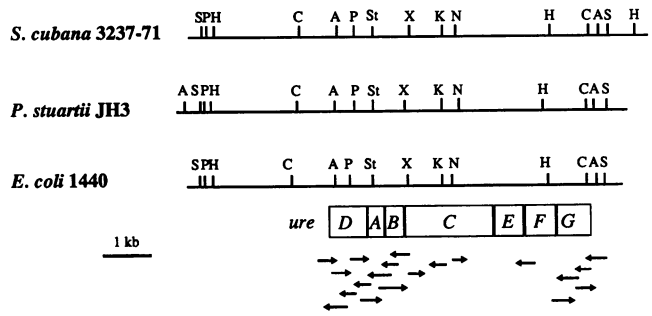

FIG. 2. The plasmid-encoded urease gene clusters of members of the family *Enterobacteriaceae* have similar restriction endonuclease maps. Physical maps of the urease loci of *E. coli* 1440, *S. cubana* 3237-71, and *P. stuartii* JH3 are shown. The map of the *E. coli* 1440 urease locus was derived from recombinant plasmid pURE14.8, which was constructed by ligating the 8.0-kb *Spe*I fragment to the *Xba*I site of pUC19 (3). Southern blot analysis of total cellular DNA probed with ^{32}P-labeled DNA fragments of pURE14.8 was used to determine the restriction maps of the other urease gene clusters. Open boxes represent the relative positions of the plasmid-encoded urease genes. The strategy used for DNA sequence analysis of the plasmid-encoded urease locus is shown below the maps. The arrows indicate the template and strand of pURE14.8 DNA sequenced. A, *Ava*I; C, *Cla*I; H, *Hin*dIII; K, *Kpn*I; N, *Nru*I; P, *Pvu*II; S, *Spe*I; St, *Stu*I; X, *Xba*I.

TABLE 1. Members of the family *Enterobacteriaceae* that contain urease-encoding plasmids express urease activity in a urea-inducible manner

Strain[a]	Urease activity[b]	
	Without urea	With urea
E. coli DH5α	0.1 ± 0.02	0.1 ± 0.02
E. coli 1021	6.3 ± 0.5	6.7 ± 1.1
E. coli 1440	0.1 ± 0.01	5.0 ± 0.2
S. cubana 3237-71	0.1 ± 0.02	7.9 ± 1.0
P. stuartii JH3	0.5 ± 0.1	19.7 ± 3.5

[a] All of the strains used, except *E. coli* DH5α (Bethesda Research Laboratories, Gaithersburg, Md.), were urease-positive clinical isolates that have been described previously (2).

[b] Urease activity is expressed as moles of ammonia produced per milligram of protein per minute (10^3). Overnight cultures grown in Luria broth were diluted to an optical density at 550 nm of 0.05 in Christensen's broth (1) supplemented with 10 mM $MgSO_4$ and 10 nM $NiCl_2$, with or without a final concentration of 2% urea. Indophenol production of mid-logarithmic-phase cells was determined as described previously (3). The values shown represent averages ± the standard errors of duplicate samples from three individual experiments.

DNA sequence analysis of the plasmid-encoded urease gene cluster. To characterize the plasmid-encoded urease locus further, the DNA sequences of both strands of 4.2 kb of pURE14.8 were determined (Fig. 2). The DNA sequences obtained were compared with previously published DNA sequences of urease gene clusters from other bacterial species. Three open reading frames (ORFs) that displayed 76 to 83% DNA identity to *P. mirabilis ureD*, *ureA*, and *ureB* (Fig. 3A and Table 2) were found within the plasmid-encoded locus, and these ORFs have been named accordingly. Downstream of *ureB*, the first 90 bp of an ORF with 82% DNA identity to *P. mirabilis ureC* was found (bp 16455 to 1735 [Fig. 3A]); this region was termed *ureC*. (Additionally, partial DNA sequences of two noncontiguous fragments located downstream of the translational start of this ORF showed 82 to 83% nucleotide identity to *P. mirabilis ureC*.)

Partial DNA sequences obtained from restriction fragments of pURE14.8 located downstream of *ureC* provided evidence for the presence of *ureE* and *ureF*. The sequences demonstrated significant nucleotide identity to *P. mirabilis ureE* (83%) and *ureF* (78%). In addition, an ORF found immediately downstream of *ureF* had 68% DNA identity to *K. aerogenes ureG*. The nucleotide sequence of the plasmid-encoded *ureG* gene is presented in Fig. 3B.

The relative positions of *ureD*, *ureA*, *ureB*, *ureC*, *ureE*, *ureF*, and *ureG* were aligned as shown in Fig. 2. Tn*5* insertions mapping to the regions now termed *ureD*, *ureA*, *ureB*, *ureC*, and *ureG* are known to eliminate urease activity of *E. coli* DH5α, supporting the notion that these ORFs represent genes (3).

Features of the nucleotide sequence. No ORF of any significant length was found on the strand opposite that which codes for UreD, UreA, UreB, and UreG. Translation of each of these polypeptides begins with a characteristic ATG start codon, and there are putative *E. coli* consensus-like ribosome-binding sites (Shine-Dalgarno sequences) (25) immediately upstream of each (*ureD*, bp 142 to 148; *ureA*, bp 998 to 1004; *ureB*, bp 1309 to 1313; *ureG*, bp 181 to 189) (Fig. 3A and B). The predicted molecular masses of plasmid-encoded polypeptides UreD, UreA, UreB, and UreG were 31.5, 11.1, 12.1, and 22.4 kDa, respectively. A possible Rho-independent transcriptional termination site was found immediately downstream of *ureG* (Fig. 3B).

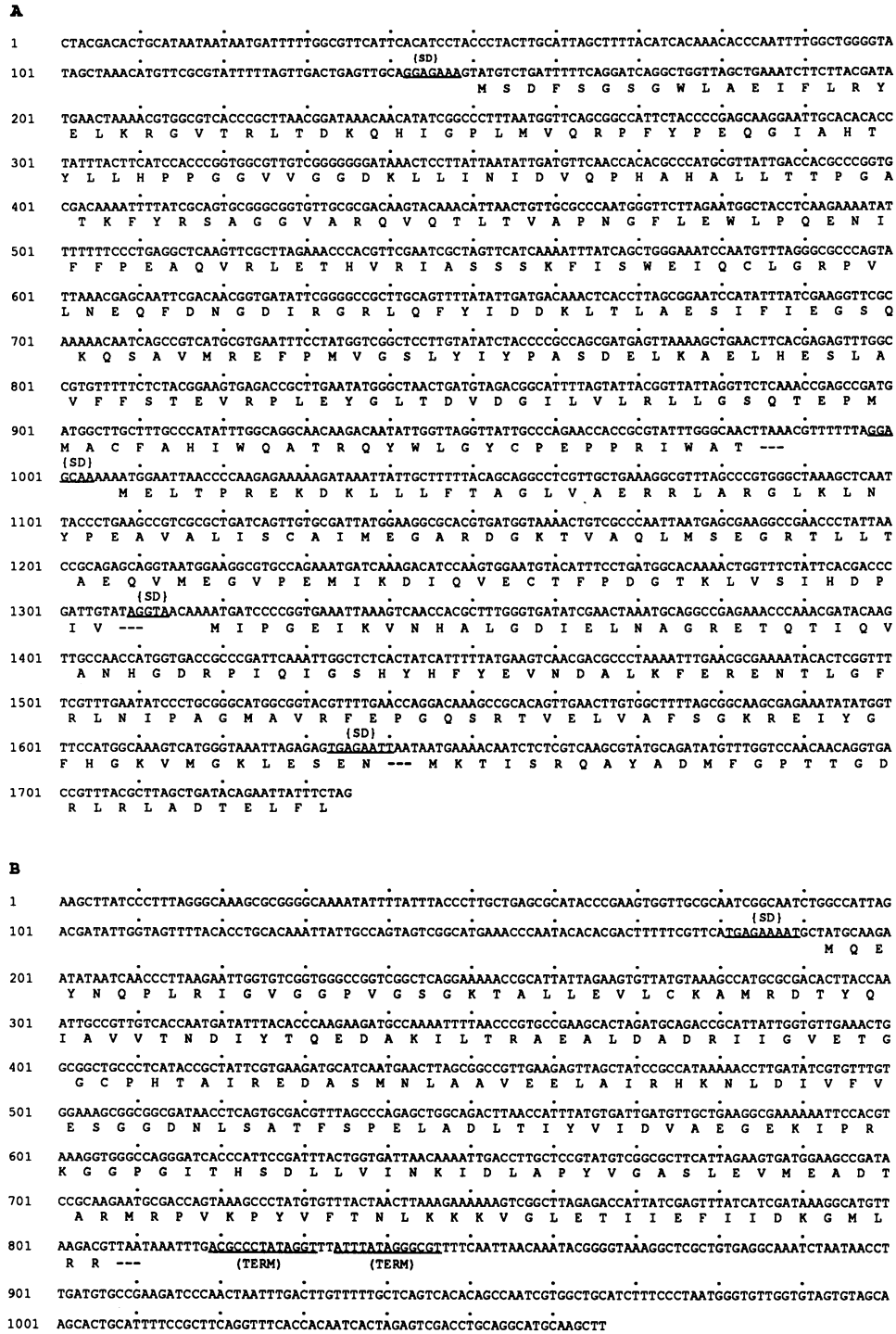

FIG. 3. Nucleotide and deduced amino acid sequences of plasmid-encoded *ureD*, *ureA*, *ureB*, and *ureG*. The numbers to the left indicate nucleotide positions. (A) DNA sequences of *ureD* (bp 151 to 984), *ureA* (bp 1008 to 1307), *ureB* (bp 1319 to 1639), and the 5′ terminus of *ureC* (bp 1645 to 1735). The predicted amino acid sequence of each gene product is shown in standard one-letter code below the DNA sequences. (B) Nucleotide and deduced amino acid sequences of *ureG* (bp 193 to 807). Nucleotides 1 to 183 display a significant degree of homology to the *P. mirabilis* and *K. aerogenes ureF* genes, as described in the text. Putative ribosome-binding sites (Shine-Dalgarno [SD] sites) and a possible Rho-independent transcriptional termination site (TERM) are underlined. DNA sequences were determined by the dideoxy-chain termination method of Sanger et al. (24) with Sequenase (U.S. Biochemicals) in accordance with the manufacturer's instructions, as described previously (4).

TABLE 2. Identity of the plasmid-encoded urease genes with *P. mirabilis* and *K. aerogenes* urease genes[a]

Plasmid-encoded gene	% DNA identity[b]		% Amino acid identity[c]	
	P. mirabilis	*K. aerogenes*	*P. mirabilis*	*K. aerogenes*
ureD	83	50	73	40
ureA	81	69	94	83
ureB	76	63	87	63
ureC	82[d]	62[d]	100[d]	76[d]
ureE	83[d]	60[d]	95[d]	64[d]
ureF	78[d]	57[d]	55[d]	55[d]
ureG	NA[e]	68	NA	80

[a] The DNA-protein sequence analysis software programs of Intelligenetics, Inc., release 5.4, were used to determine ORFs and deduced amino acid sequences and to search the GenBank data base of sequences.

[b] Identity of the indicated plasmid-encoded gene with the corresponding gene of *P. mirabilis* (12) or *K. aerogenes* (21).

[c] Identity of the predicted amino acid sequence of the indicated plasmid-encoded gene with the amino acid sequence of the corresponding gene product of *P. mirabilis* or *K. aerogenes*.

[d] Partial sequence data.

[e] NA, data not available.

Comparison of the plasmid-encoded urease gene products with the urease gene products of other bacterial species. Among members of the family *Enterobacteriaceae*, the urease structural genes are highly conserved. The predicted amino acid sequences of the urease structural subunits of the plasmid-encoded locus and the urease loci of *P. mirabilis* and *K. aerogenes* share a significant degree of identity. The deduced amino acid sequence of the plasmid-encoded UreA demonstrated 94% identity to UreA of *P. vulgaris* and *P. mirabilis* and 83% identity to *K. aerogenes* UreA (Table 2). In a comparison of the sequence of the plasmid-encoded locus and that of the *Proteus* species, five of the six amino acid differences were conservative substitutions. Similar identities were found between the plasmid-encoded UreB and UreB proteins of *P. vulgaris* (88%), *P. mirabilis* (87%), and *K. aerogenes* (63%). The amino-terminal portion of UreC, predicted from the first 90 bp of *ureC*, showed 100% identity to the deduced amino acid sequences of the *P. vulgaris* and *P. mirabilis* UreC and 76% homology to *K. aerogenes* UreC.

The predicted amino acid sequence of the plasmid-encoded UreG protein displayed 80% identity to *K. aerogenes* UreG. The region just upstream of *ureG* had DNA homology to *P. mirabilis ureF* (78%) and *K. aerogenes ureF* (57%). Although the plasmid-encoded *ureF* sequence displays a higher degree of homology to *P. mirabilis*, the carboxy terminus of the predicted ORF for UreF in the plasmid-encoded locus extends approximately 20 amino acids beyond the end of *P. mirabilis ureF*. The size of the predicted UreF polypeptide was more similar to that of *K. aerogenes* UreF, which has a molecular mass of 24.7 kDa, than to that of *P. mirabilis* UreF, which has a predicted molecular mass of 22.5 kDa.

The plasmid-encoded UreD had 73% amino acid identity to *P. mirabilis* UreD. However, the plasmid-encoded UreD shares only 40 to 41% identity with the predicted amino acid sequences of the *K. aerogenes* and *K. pneumoniae* (5) UreD proteins. This gene product seems to be less highly conserved among members of the family *Enterobacteriaceae* than either the structural subunits or the other accessory polypeptides.

In this report, we demonstrate that all of the members of the family *Enterobacteriaceae* examined to date that carry large urease-encoding plasmids contain identical urease gene clusters. Approximately 30% of clinical *P. stuartii* isolates (14, 19), but only a small fraction of *E. coli* and *Salmonella* clinical isolates, are urease positive and contain this locus. Thus, the plasmid-encoded urease gene cluster is found predominantly in *P. stuartii*. It is interesting that the urease-positive *Salmonella* isolates examined here were obtained from urine samples of patients with urinary tract infections. Presumably, these strains acquired the plasmids by conjugation with a neighboring *E. coli* or *P. stuartii* strain. The urease-encoding plasmids in the *Salmonella* isolates were not the *Salmonella* virulence plasmid, as determined by hybridization with a DNA probe specific for the virulence plasmid (data not shown).

The plasmid-encoded urease gene cluster is more closely related to the *P. mirabilis* urease locus than to the *K. aerogenes* urease genes. Like the plasmid-encoded locus, the urease activity of *P. mirabilis* is induced in the presence of urea (11), suggesting a possible evolutionary relationship between the two loci.

Nucleotide sequence accession numbers. The sequences presented here have been deposited in the GenBank data base under accession no. L03307 (*ureD*, *ureA*, *ureB*, and *ureC*) and L03308 (*ureF* and *ureG*).

We thank Clifford Mintz and Dorothy Pierson for critical reading of the manuscript.

This work was supported by Public Health Service grant AI27907 to C.M.C. S.E.F.D. is a predoctoral trainee supported by grant AI07346 from the National Institutes of Health.

REFERENCES

1. **Christensen, W. B.** 1946. Urea decomposition as a means of differentiating *Proteus* and paracolon cultures from each other and from *Salmonella* and *Shigella* types. J. Bacteriol. **52**:461.
2. **Collins, C. M., and S. Falkow.** 1988. Genetic analysis of an *Escherichia coli* urease locus: evidence of DNA rearrangement. J. Bacteriol. **170**:1041–1045.
3. **Collins, C. M., and S. Falkow.** 1990. Genetic analysis of *Escherichia coli* urease genes: evidence for two distinct loci. J. Bacteriol. **172**:7138–7144.
4. **Collins, C. M., and D. M. Gutman.** 1992. Insertional inactivation of an *Escherichia coli* urease gene by IS*3411*. J. Bacteriol. **174**:883–888.
5. **Collins, C. M., D. M. Gutman, and H. Laman.** Identification of a nitrogen regulated promoter controlling expression of *Klebsiella pneumoniae* urease genes. Mol. Microbiol., in press.
6. **D'Orazio, S. E. F., D. M. Gutman, and C. M. Collins.** Unpublished data.
7. **Forland, M.** 1984. The management of urinary tract infections, p. 222–233. *In* W. N. Suki and S. G. Massry (ed.), Therapy of renal diseases and related disorders. Martinus Nijhoff Publishers, Boston.
8. **Gatermann, S., J. John, and R. Marre.** 1989. *Staphylococcus saprophyticus* urease: characterization and contribution to uropathogenicity in unobstructed urinary tract infection of rats. Infect. Immun. **57**:110–116.
9. **Hu, L.-T., E. B. Nicholson, B. D. Jones, M. J. Lynch, and H. T. Mobley.** 1990. *Morganella morganii* urease: purification, characterization, and isolation of gene sequences. J. Bacteriol. **172**:3073–3080.
10. **Jones, B. D., V. Lockatell, D. E. Johnson, J. W. Warren, and H. T. Mobley.** 1990. Construction of a urease-negative mutant of *Proteus mirabilis*: analysis of virulence in a mouse model of ascending urinary tract infection. Infect. Immun. **58**:1120–1123.
11. **Jones, B. D., and H. T. Mobley.** 1988. *Proteus mirabilis* urease: genetic organization, regulation, and expression of structural genes. J. Bacteriol. **170**:3342–3349.
12. **Jones, B. D., and H. T. Mobley.** 1989. *Proteus mirabilis* urease:

nucleotide sequence determination and comparison with jack bean urease. J. Bacteriol. **171**:6414–6422.

13. **Lee, M. H., S. B. Mulrooney, M. J. Renner, Y. Markowicz, and R. P. Hausinger.** 1992. *Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis. J. Bacteriol. **174**:4324–4330.
14. **Lennette, E. H., A. Balows, W. J. Hausler, Jr., and H. J. Shadomy (ed.).** 1985. Manual of clinical microbiology, 4th ed., p. 267. American Society for Microbiology, Washington, D.C.
15. **Lerner, S. P., M. J. Gleeson, and D. P. Griffith.** 1989. Infection stones. J. Urol. **141**:753–758.
16. **MacLaren, D. M., and P. G. H. Peerbooms.** 1986. Urinary infections by urea-splitting micro-organisms, p. 183–195. *In* A. W. Asscher and W. Brumfitt (ed.), Microbial diseases in nephrology. John Wiley & Sons, Inc., New York.
17. **McLean, R. J. C., J. C. Nickel, K.-J. Cheng, and J. W. Costerton.** 1988. The ecology and pathogenicity of urease-producing bacteria in the urinary tract. Crit. Rev. Microbiol. **16**:37–79.
18. **Mobley, H. L. T., B. D. Jones, and A. E. Jerse.** 1986. Cloning of urease gene sequences from *Providencia stuartii*. Infect. Immun. **54**:161–169.
19. **Mobley, H. L. T., and J. W. Warren.** 1987. Urease-positive bacteriuria and obstruction of long-term urinary catheters. J. Clin. Microbiol. **25**:2216–2217.
20. **Morsdorf, G., and H. Kaltwasser.** 1990. Cloning of the genes encoding urease from *Proteus vulgaris* and sequencing of the structural genes. FEMS Microbiol. Lett. **66**:67–74.
21. **Mulrooney, S. B., and R. P. Hausinger.** 1990. Sequence of the *Klebsiella aerogenes* urease genes and evidence for accessory proteins facilitating nickel incorporation. J. Bacteriol. **172**:5837–5843.
22. **Mulrooney, S. B., M. J. Lynch, H. T. Mobley, and R. P. Hausinger.** 1988. Purification, characterization, and genetic organization of recombinant *Providencia stuartii* urease expressed by *Escherichia coli*. J. Bacteriol. **170**:2202–2207.
23. **Parsons, C. L., C. Stauffer, S. G. Mulholland, and D. P. Griffith.** 1984. Effect of ammonium on bacterial adherence to bladder transitional epithelium. J. Urol. **132**:365–366.
24. **Sanger, F. S., S. Nicklen, and A. K. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **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.