



1984

ПИСЬМА РЕДАКТОРУ

УДК 577.113.5

ПЕРВИЧНАЯ СТРУКТУРА УЧАСТКА ДНК *E. coli*, ПРЕДШЕСТВУЮЩЕГО ГЕНАМ ТРИПТОФАНОВОГО ОПЕРОНА

Губанов В. В., Лебедев Ю. Б., Монастырская Г. С.,
Рубцов П. М*., Скрыбин К. Г*.

Институт биоорганической химии им. М. М. Шемякина Академии наук СССР, Москва;
* Институт молекулярной биологии Академии наук СССР, Москва

Промотор триптофанового оперона *E. coli* часто используют для экспрессии различных генов, вводимых методами генной инженерии в эту бактерию [1, 2]. При этом возникает задача вырезания этого промотора вместе с его оператором для последующего присоединения к требуемому гену. Для этой цели крайне желательно знать первичную структуру промотора и прилегающих областей. В настоящее время известна структура небольшого участка ДНК *E. coli*, представляющая собой собственно промотор, и структура всего триптофанового оперона [3]. Однако сведения о структуре участка, прилегающего к промотору с другой стороны, отсутствуют. В данной работе установлена первичная структура этого участка, что создает возможность наиболее рационального использования рестрикционных эндонуклеаз и методов химического синтеза при конструировании различных векторов. Клонирование этого фрагмента было осуществ-

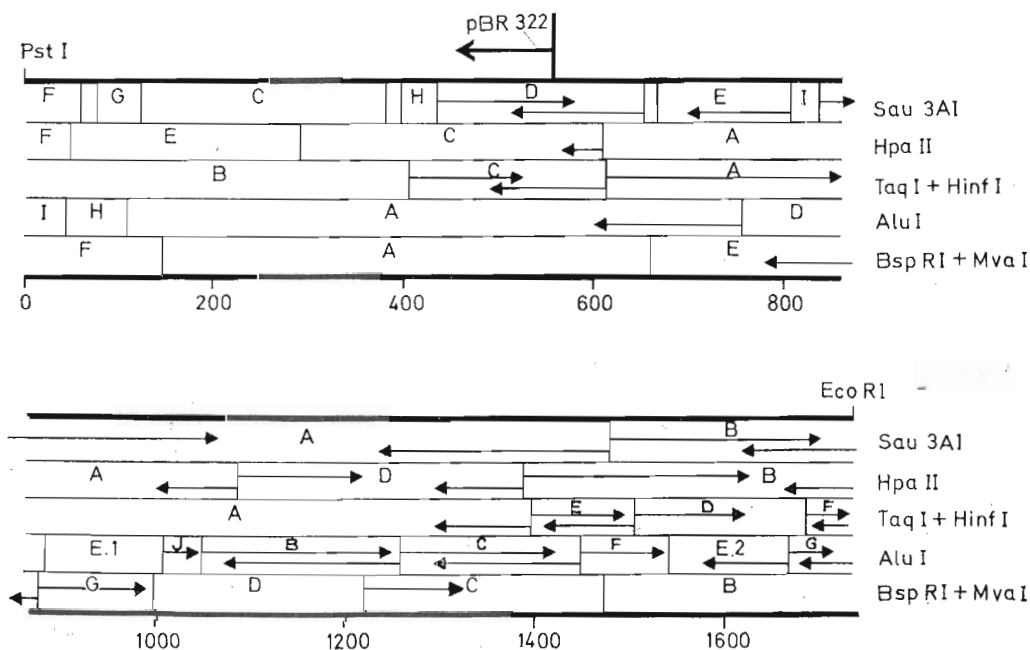


Рис. 1. Детальная карта расщепления фрагмента *Pst*I - *Eco*RI рестрикционными эндонуклеазами и схема установления его первичной структуры. Субфрагменты, образующиеся при действии рестриктаз, изображены прямоугольниками. Стрелки обозначают длину установленной последовательности комплементарных цепей субфрагмента

← pBR322 →

526-600 CACGGAAATGTTGAATACTCATACTCTTCSTTTTTCCAAATACGCTCCATCGCGTTTTTATCTTCAATTTTACAG
601-675 CCGAGCGCATAGCCGGAATCAGTTGGATAAACAATCACCCCGCCTTTACGCACGATCTCCACCCGCTGGTTGATC
676-750 AGACGTTGCTGTGGGTTATCAGGATGAATATAAAAAAAGCTGGCTCATAATCCCTCTCTGTGGTGTCTCGCGCT
751-825 GTTCCCATAGCTGCCAGACGCCTTCAACGCCTGCGGGCAGCCAGAGTTTACGACCCAGTTCGATCCACGGGCATG
826-900 GCTGATGAAAAATCAGATCCTTGTGATGCCATAAATGATGCTGACGCGCAAGGGCCGACGTTGGGTACGTTTCAT
901-975 TGGGCGATGTGCGCACTGCGCGACTTCCATCGCGCTCACCGTGGTGTTCGGCAAAATGCGCAACCAATCTTTTCA
976-1050 GCCATTTAGCASTAAGATTGTACCGCCAGGATGAGCTAGCACCCGCTTACCGCCAGAATGATGAATGACATCAA
1051-1125 TAGCTTGTCTATTGTACACCACTGTGGCGGAACGTATCCGGTTTTCCCGCGCCAGATACCTTTTAAAGACAT
1126-1200 CCGCATTGAACTGGCTTTGCCGCACTCAACAAGGAAACGCGCAAAATGACCGCGCTCACTGCGCCCTGCGC
1201-1275 CAGTCGTTGTGCGCCTTCCAGCGCCAGGATTTGCGCTTTTTTCAAGCTTTTCGGCAATCAGTGCGCCGCTGA
1275-1350 TTGCCCGCTTCTGTCTGCTGCGGAGGAACTCACACATAGCGGATGAGTAATATCAATATTAGCCCGACAATA
1351-1425 TGAATTTTCATGATTTTCCAGACCGTGGAAATTTCCACGCGCGGAAATAGATTCAACGCCAGTCCCGAACGTGAA
1426-1500 ATTTCCCTCTGTGCTGCGCGATTGACAGCTGTGGTGTCTATGGTCCGTGATCGCCAGGTGCGCGACCGCATCTCG
1501-1575 ACTGCACGGTGCACCAATGCTTCTGGCGTCAAGGCACTCCCGTCTGGATAATGTTTTTTCGCCGACATCATAACGGTCT
1576-1650 AGTGCATAATTCGTCTGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTCGCCGACATCATAACGGTCTCT
1651-1725 GGCAATAATTCGAAATGAGCTGTGACAAATTAATCATCGAACTAGTTAACTAGTACGCAAGTTCACGTAATAAAG
1726-1740 GG... . . . GAATTC

Рис. 2. Нуклеотидная последовательность фрагмента ДНК *E. coli*, предшествующего генам триптофанового оперона. Стрелкой отмечен конец плазмидной части фрагмента. Нумерация нуклеотидов фрагмента *Pst*I—*Eco*RI дана от начала участка узнавания *Pst*I (координата — 3607 в обозначениях работы [8])

лено в плазмиде *pBR322*; после реконструкции рекомбинантная плаزمид была использована для экспрессии генов интерферона $\alpha 2$ [2].

Для вырезания исследуемого фрагмента из исходной плазмиды были взяты эндонуклеазы рестрикции *Eco*RI и *Pst*I. При этом захватывалась часть плазмидной ДНК (~500 п.о.), что, однако, не мешало исследованию.

Общая стратегия определения структуры данного фрагмента ДНК аналогична примененной нами в работах [4, 5] (рис. 1). Фрагмент расщепляли поочередно одной из эндонуклеаз рестрикции: *Sau*3AI, *Hpa*II, *Alu*I или смесями *Taq*I+*Hinf*I и *Bsp*I+*Mva*I. Последовательности комплементарных цепей выделенных субфрагментов определяли модифицированным нами [6] методом Максама — Гилберта [7] после введения 32 P-концевой фосфатной группы с помощью [γ - 32 P]АТФ и полинуклеотидкиназы фага Т4. Структура субфрагментов из плазмидной области не определялась. Большая часть остальной структуры определена по двум цепям. Установленная последовательность фрагмента ДНК, включающего промотор триптофанового оперона и предшествующий ему участок ДНК *E. coli*, содержит 1179 п.о. (без плазмидной части фрагмента). Она представлена на рис. 2.

По данным работы [9], ближайшим к локусу триптофанового оперона со стороны промотора является локус *opp*, ответственный за транспорт олигонуклеотидов в клетку. Оба локуса расположены на 27-й мин 100-минутной генетической карты *E. coli* (сцепленность при половой рекомбинации — 98% [10]). Эти сведения послужили предпосылкой для поиска структурных генов в определенной нами последовательности. С помощью компьютера был осуществлен поиск терминирующих кодонов. При этом оказалось, что максимальная длина участков с «открытой рамкой» считывания в одном направлении (от 5'- к 3'-концу на рис. 2) составляет 276 п.о., а в противоположном — 291 п.о., что практически исключает наличие в этом районе структурных генов. Такая возможность сохраняется лишь для концевой участка фрагмента, прилегающего к сайту *Pst*I; максимальная длина последовательности с «открытой рамкой» считывания составляет здесь 267 п.о.

Таким образом, участок ДНК, предшествующий генам триптофанового оперона, представляет собой значительную по размерам интерцистронную область и предназначен для осуществления регуляторных функций.

1. Goeddel D. V., Heyneker H. L., Hozumi T., Arentzen R., Itakura K., Yansura D. G., Ross M. J., Miozzari G., Crea R., Seeburg P. H. *Nature*, 1979, v. 281, № 5, p. 544–548.
2. Овчинников Ю. А., Свердлов Е. Д., Царев С. А., Ходкова Е. Д., Монастырская Г. С., Саломатина И. С., Ефимов В. А., Чихмагичева О. Г., Соловьев В. Д., Кузнецов В. П., Жданов В. М., Новохацкий А. С., Асперов П. Д. Докл. АН СССР, 1982, т. 265, № 1, с. 238–242.
3. Bennett G. N., Yanofsky C. J. *Mol. Biol.*, 1978, v. 121, № 2, p. 179–192.
4. Ovchinnikov Yu. A., Monastyrskaya G. S., Gubanov V. V., Guryev S. O., Chertov O. Yu., Modyanov N. N., Grinkevich V. A., Makarova I. A., Marchenko T. V., Polovnikova I. N., Lipkin V. M., Sverdlov E. D. *Eur. J. Biochem.*, 1981, v. 116, p. 621–629.
5. Ovchinnikov Yu. A., Monastyrskaya G. S., Gubanov V. V., Guryev S. O., Salomatina I. S., Shuvaeva T. M., Lipkin V. M., Sverdlov E. D. *Nucl. Acid Res.*, 1982, v. 10, № 13, p. 4035–4044.
6. Ovchinnikov Yu. A., Guryev S. O., Kravtsov A. S., Monastyrskaya G. S., Skryabin K. G., Sverdlov E. D., Zakharyev B. M., Bayev A. A. *Gene*, 1979, v. 6, № 3, p. 235–249.
7. Maxam A., Gilbert W. *Proc. Nat. Acad. Sci. USA*, 1977, v. 74, № 2, p. 560–564.
8. Sutcliffe J. G. *Nucl. Acids Res.*, 1978, v. 5, № 8, p. 2721–2728.
9. Bachmann B. J., Low K. B., Taylor A. L. *Bacteriol. Rev.*, 1978, v. 40, № 1, p. 116–167.
10. Baras Z., Gilvary Ch. *J. Biol. Chem.*, 1974, v. 249, № 1, p. 143–148.

Поступило в редакцию
30.XI.1983

THE PRIMARY STRUCTURE OF THE *E. COLI* DNA FRAGMENT,
PRECEEDING THE TRYPTOPHAN OPERON GENES

GUBANOV V. V., LEBEDEV Yu. B., MONASTYRSKAYA G. S.,
RUBTSOV P. M*, SKRYABIN K. G*.

*M. M. Shemyakin Institute of Bioorganic Chemistry Academy
of Sciences of the USSR, Moscow;*

* *Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow*

The nucleotide sequence of 1179 b.p. preceding the *trp* operon genes has been established. There are no open reading frames large enough to code for proteins containing more than 97 amino acid residues. In all cases the coding sequences do not contain the initiation codons. The determined sequence is concluded to represent an intercistronic region.