



ПИСЬМА РЕДАКТОРУ

УДК 547.963.32.02

НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НАЧАЛА ГЕНА *groB* *ESCHERICHIA COLI*

Гуревич А. И., Азаков А. Э., Колосов М. Н.

Институт биоорганической химии им. М. М. Шемякина
Академии наук СССР, Москва

Из трансдуцирующего ффага λ rif^d47 [1] мы выделили *EcoRI*-фрагмент бактериальной ДНК длиной около 1100 нуклеотидных пар (н.п.), который содержит конец гена *rplL*, начало гена *groB* и регуляторный участок между ними. Ранее нами было выяснено положение этого фрагмента (обозначенного *EcoRI-K*) на физической карте λ rif^d47 [2] и исследована структура двух его *HpaII*-субфрагментов [3, 4]. Продолжая изучение оперона *groBC*, мы определили полную нуклеотидную последовательность фрагмента *EcoRI-K*, включая начальную часть структурного гена *groB*, кодирующую первые 185 аминокислотных остатков β -белка РНК-полимеразы.

Выясненная нами карта расщепления фрагмента *EcoRI-K* рестриктазами и схема его секвенирования приведены на рис. 1. Как видно из этого рисунка, большая часть нуклеотидной последовательности была определена по обеим цепям ДНК. Анализ проводили методом Максама — Гилберта [5] в модификации Коробко и др. [6, 7]; 5'- и 3'-концевую ³²P-метку вводили как описано ранее [5, 3], комплементарные цепи разделяли по методу [5]. При определении нуклеотидной последовательности некоторых участков ДНК, например терминатора гена *rplL*, встретились труд-

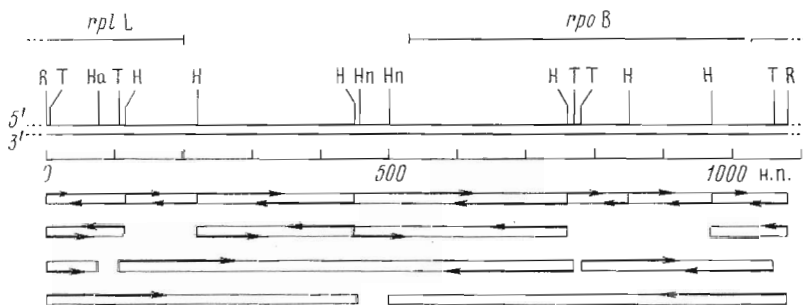


Рис. 1. Рестриктивная карта фрагмента *EcoRI-K* генов *rplL-groB* *E. coli* и схема определения его нуклеотидной последовательности. Указано расположение структурных генов *rplL* и *groB* и сайтов действия эндонуклеаз *EcoRI* (R), *HpaII* (H), *HaeIII* (Ha), *HindII* (Hn) и *TaqI* (T). Горизонтальные стрелки показывают длину установленных последовательностей и направление секвенирования в каждой из двух цепей соответствующего субфрагмента (полярность верхней цепи ДНК везде 5'→3')

...GluPheAspValIleLeuLysAlaAlaGlyAlaAsnLysValAlaValIleLysAlaValArgGly
EcoRI TaqI 50

...GAATTCGACGTAATTCCTGAAAGCTGCTGGCGCTAACAAAGTTGCTGTTATCAAAGCAGTACGTGGC
...CTTAAAGCTGCATTAAGACTTTCGACGACCGGATTGTTTCAACGACAATAGTTTCGTGCATGCACCG

AlaThrGlyLeuGlyLeuLysGluAlaLysAspLeuValGluSerAlaProAlaAlaLeuLysGluGly
HaeIII 100 TaqI HpaII

GCAACTGGCCTGGGTCTGAAAGAAGCTAAAGACTGGTCAATCTGCACCGGCTGCTCTGAAAGAAGGC
CGTTGACCGGACCCAGACTTTCTTCGATTCTTGACCAGCTTAGACGTGGCCGACGAGACTTTCTTCCG

ValSerLysAspAspAlaGluAlaLeuLysLysAlaLeuGluGluAlaGlyAlaGluValGluValLys
I50 200

GTGAGCAAAGACGACGCAGAAAGCACGAAAAAGCTCTGGAAGAAGCTGGCGCTGAAGTTGAAGTTAAAC
CACTCGTTTCTGCTGCGTCTTCGTGACTTTTTTCGAGACCTTCTTCGACCGGACTTCAACTTCAATTT

HpaII 250
TAAGCCAACCTTCCGGTTCAGCCTGAGAAATCAGGCTGATGGCTGGTACTTTTTCAGTCAACAGCTT
CTTCGGTTGGGAAGGCCAACGTCGGACTCTTTCAGTCCGACTACCGACCACTGAAAAATCAGTGGTCCAA

300
TTTTTGCCTGTAAAGCGCCAGTAGCGTTTCACACTGTTTACTACTGCTGTGCCTTTCATGCTTGTTC
AAAAACCGGACATTCGCGGTTCATCGCAAAGTGTGACAAACTGATGACGACACGGAAAGTTACGAAACAA

350 400
TCTATCGACGACTTAATACTATCGCAGACAGGACGTCGGTTCTGTGTAATCGCAATGAATGGTTAAAGC
AGATAGCTGCTGAATTATATGACGCTGTCTCGCAGGCAAGACACATTTAGCGTTACTTTACCAAATTCG

HpaII HindII
GTGATAGCAACAGGCATTTCGGAAAGTGTTCATTTTCCGGTCAACAAAATAGTGTGCACAAAAGTGC
CACTATCGTTGTCGGTAACGCCTTTCACAAGGTAAGGCCAGTGTGTTTATCACAACGTTTGTGACAG

MetValTyrSerTyrThrGlu
500 HindII
CGCTCAATGGACAGATGGGTTCGACTTGTCAAGCGAGCTCAGGAACCCATATGGTTTACTCCTATACCGAG
GCGAGTTACCTGTCTACCCAGCTGAACAGTCTGCTCGACTCCTTGGGATACCAATGAGGATATGGCTC

LysLysArgIleArgLysAspPheGlyLysArgProGlnValLeuAspValProTyrLeuLeuSerIle
550 600
AAAAAACGTATTCGTAAGGATTTTGGTAAACGTCACAAAGTCTGGATGTACCTTATCTCCTTTCTATC
TTTTTGCATAAGCATTCCTAAAACCATTTGCAGGTGTTCAAGACCTACATGGAATAGAGGAAAGATAG

GlnLeuAspSerPheGlnLysPheIleGluGluAspProGluGlyGlnTyrGlyLeuGluAlaAlaPhe
650
CAGCTTGACTCGTTTCAGAAATTTATCGAGCAAGATCCTGAAAGGCGAGTATGGTCTGGAAGCTGCTTTC
GTGCAACTGAGCAAAGTCTTAAATAGCTCGTTCTAGGACTTCCCGTCATACCAGACCTTCGACGAAAG

ArgSerValPheProIleGlnAlaThrAlaValIleProSerCysAsnAsnValThrThrProTrpArg
700 750
CGTTCGGTATTCCCGATTACAGCTACAGCGGTAATTCCGAGCTGCAATAACGTCACTACGCTTGGCGA
GCAAGGCATAAGGGCTAAGTCCGATGTCGCCATTAAAGGCTCGACGTTATTGCAGTGATGCGGAACCGCT

β -белка РНК-полимеразы 185-членную N-концевую аминокислотную последовательность, изображенную на рис. 2. Спейсер между генами *rplL* и *groV* занимает почти треть фрагмента *EcoRI*-K (около 320 н.п.). В нем недалеко от конца гена *rplL* имеются два обращенных повтора: 8-членный (225–232, 237–244) и обнаруженный ранее [4] 10-членный (248–257, 262–271), сразу за которым находится последовательность TTTTТТТG; по нашему мнению, именно этот участок является аттенуатором оперона *rplL-groV*. В средней части спейсера есть еще один (правда, несовершенный) обращенный повтор: СААТGСРyTGТТ (332–342, 420–430); его функциональная роль и значение для вторичной структуры ДНК (и соответствующей мРНК) не известны, но интересно, что находящиеся в этой области сайты TCGA (347–350, 500–503) не расщепляются рестриктазой *TaqI*.

После завершения настоящей работы в печати появилась статья Номуры и сотр. [11], в которой опубликована нуклеотидная последовательность большого сегмента (3072 н.п.) ДНК *E. coli*, выделенного из трансдуцирующего фага λ rif^r18 и содержащего участок 1–628 исследованного нами фрагмента *EcoRI*-K. Наши результаты совпадают с полученными группой Номуры, за исключением следующих четырех. В гене *rplL* американскими учеными найдены нуклеотиды 100-Т и 105-А, а нами — 100-С и 105-С; эти различия относятся к третьей букве триплета и не меняют его смысла в отношении кодируемой аминокислоты. Следует отметить, что и мы на основании химической дегградации вначале идентифицировали нуклеотид 105 как А, но затем обнаружили, что непосредственно перед ним ДНК расщепляется рестриктазой *TaqI* и, следовательно, этот нуклеотид должен быть С. В межгенной области, по данным Поста и соавт. [11], вместо 272-Т содержится СТ_n, а между нуклеотидами 500 и 501 находится триплектид СGT, отсутствующий в формуле на рис. 2. Мы повторно секвенировали соответствующие участки ДНК, но не смогли получить подтверждения этих данных.

ЛИТЕРАТУРА

1. Миндлин С. З., Ильина Т. С., Горленко Ж. М., Хачикян Н. А., Ковалев Ю. Н. (1976) Генетика, 12, № 12, 116–130.
2. Гуревич А. И., Аваков А. Э., Киселева О. А., Колосов М. Н. (1978) Биоорган. химия, 4, 628–638.
3. Гуревич А. И., Аваков А. Э. (1979) Биоорган. химия, 5, 301–304.
4. Гуревич А. И., Аваков А. Э., Колосов М. Н. (1979) Биоорган. химия, 5, 779–784.
5. Maxam A. M., Gilbert W. (1977) Proc. Nat. Acad. Sci. USA, 74, 560–564.
6. Коробко В. Г., Грачев С. А. (1977) Биоорган. химия, 3, 1420–1422.
7. Коробко В. Г., Грачев С. А., Колосов М. Н. (1978) Биоорган. химия, 4, 1281–1283.
8. Terhorst C., Müller M., Laursen R., Wittmann-Liebold B. (1973) Eur. J. Biochem., 34, 138–152.
9. Shine J., Dalgarno L. (1974) Proc. Nat. Acad. Sci. USA, 71, 1342–1346.
10. Fujiki H., Zurek G. (1975) FEBS Lett., 55, 242–244.
11. Post L. E., Strycharz G. D., Nomura M., Lewis H., Dennis P. P. (1979) Proc. Nat. Acad. Sci. USA, 76, 1697–1701.

Поступило в редакцию
2.VII.1979

THE NUCLEOTIDE SEQUENCE AT THE PROXIMAL END OF *rpoB* GENE OF *ESCHERICHIA COLI*

GUREVICH A. I., AVAKOV A. E., KOLOSOV M. N.

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

A *EcoRI* fragment of *E. coli* DNA isolated from transducing phage λr_{ij}^{d47} has been sequenced by a modified Maxam – Gilbert method. The fragment is 1080 nucleotide long and contains the distal and proximal ends of genes *rplL* and *rpoB*, respectively, separated by a spacer of ca. 320 b.p. The terminal 204 nucleotide sequence determined in *rplL* was found to completely agree with the published structure of the ribosomal protein L7/L12. The start of *rpoB* is assumed to be at the ATG in position 528–530, because it is preceded by a putative ribosome binding site GAGG and followed by no termination codons till the end of the DNA fragment studied while several nonsense codons are present in the two other reading frames. The starting 555 nucleotide sequence thus determined in *rpoB* codes for 185 amino acids of the β -protein of RNA polymerase, the first three amino acids being identical to those previously reported as the N-terminal. In the *rplL-rpoB* spacer, two inverted repeats are present near the end of *rplL*, an 8-membered and a 10-membered followed by T₇G, which supposedly constitute an attenuator of the *rplL-rpoBC* operon. There is also a nonperfect 11-membered inverted repeat at a longer distance whose function, if any, remains to be known.

Технический редактор *Е. С. Кузьмишкина*

Сдано в набор 20.08.79 Подписано к печати 02.10.79 Т-13673 Формат бумаги 70×108^{1/16}
Высокая печать Усл. печ. л. 13,3+1 вкл. Уч.-изд. л. 13,9 Бум. л. 4^{3/4} Тираж 870 экз. Зак. 2173

Издательство «Наука». 103717 ГСП, Москва, К-12, Подсосенский пер., 21
2-я типография издательства «Наука». 121099, Москва, Шубинский пер., 10