



## ПИСЬМА РЕДАКТОРУ

УДК 547.963.32.02

СТРУКТУРА ЦЕНТРАЛЬНОЙ ЧАСТИ ОПЕРОНА *groBC* *E. coli*.  
НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ГЕНА  
β-СУБЪЕДИНИЦЫ РНК-ПОЛИМЕРАЗЫ

Гуревич А. И., Игошин А. В., Колосов М. Н.

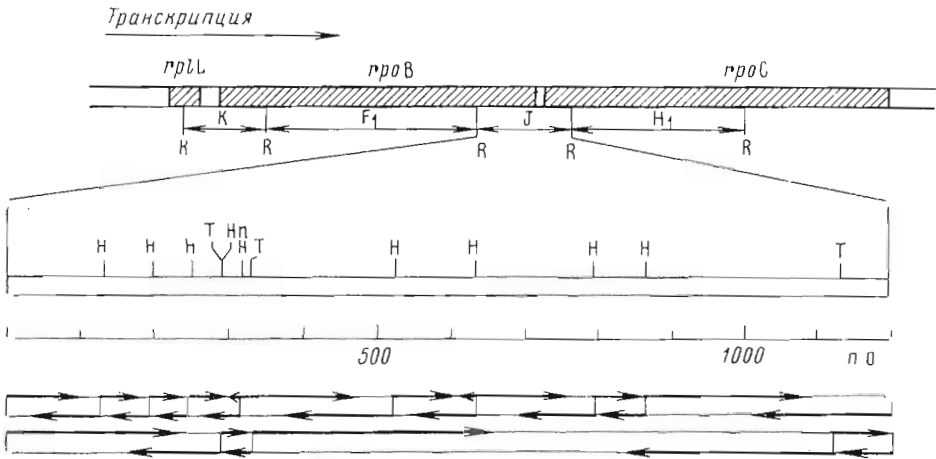
Институт биоорганической химии им. М. М. Шемякина  
Академии наук СССР, Москва

При изучении оперона *groBC* нами в 1979 г. было выяснено [1] строение *EcoRI*-фрагмента ДНК *E. coli* величиной около 1100 пар оснований (п. о.), который содержит аттенюатор этого оперона и начальную 555-нуклеотидную последовательность структурного гена *groV*, кодирующую β-белок РНК-полимеразы. В 1980 г. Ю. А. Овчинников с сотр. [2] опубликовали нуклеотидную последовательность *EcoRI*-фрагмента, смежного с предыдущим и заключающего в себе большую часть структурного гена *groV* (2862 п. о. между двумя сайтами *EcoRI*). Настоящая работа посвящена определению структуры соседнего *EcoRI*-фрагмента, в котором содержится центральная часть оперона *groBC*, т. е. находятся конец гена *groV* и начало гена *groC*.

Исследуемый фрагмент ДНК *E. coli* был нами ранее получен из трансдуцирующего фага  $\lambda rif^{d47}$  и обозначен *EcoRI-J* [3]. При его расщеплении эндонуклеазами *HpaII*, *HindII* и *TaqI*\* была составлена физическая карта, показанная на схеме 1. Полярность фрагмента была установлена по расположению в нем сайта эндонуклеазы *SalI* (= *HindII*), который был картирован ранее в опероне *groBC* в составе ДНК фага  $\lambda rif^{d18}$  [5]. Определение нуклеотидной последовательности мы проводили по методу Максама — Гилберта [7] с модификациями, разработанными В. Г. Коробко и др. [8, 9]. Радиоактивную 5'- и 3'-концевую метку вводили как описано ранее [7, 10], комплементарные цепи разделяли по методу [7] после термической денатурации в 75–100% диметилсульфоксиде. Как видно из схемы 1, большая часть нуклеотидной последовательности фрагмента *EcoRI-J* была определена по обеим цепям ДНК. При этом каждый ее отрезок анализировали не менее 3 раз и, где было возможно, использовали субфрагменты, полученные с помощью разных эндонуклеаз.

\* Эта эндонуклеаза была нами выделена из штамма *Thermus aquaticus*, полученного от А. Г. Слюсаренко (Институт общей генетики АН СССР), и ошибочно принята [1] за рестриктазу *TaqI*. При проверке ее специфичности мы нашли, что она является изошизомером эндонуклеазы *EcoRII*, т. е. расщепляет ДНК перед пента-нуклеотидной последовательностью  $CC(\frac{T}{A})GG$ . В связи с этим в опубликованную ранее нуклеотидную последовательность фрагмента *EcoRI-K* [1] в сайтах *TaqI* и *EcoRII* нужно внести следующие изменения (данные А. Э. Авакова): 105-А вместо 105-С, 767-Т вместо 767-С, 774-С вместо 774-Г, 1060-С вместо 1060-А, 1057-С отсутствует.

Схема 1



Рестриктазная карта фрагмента *EcoRI*-J генов *groV* – *groC* *E. coli* и схема определения его нуклеотидной последовательности. Указано направление транскрипции, расположение структурных генов (заштрихованы) и сайтов эндонуклеаз *EcoRI* (R), *HpaII* (H), *HindIII* (Hn) и *TaqX\** (T). *EcoRI*-фрагменты обозначены в соответствии с работой [3]. Горизонтальные стрелки показывают длину установленных последовательностей и направление секвенирования в каждой из двух нитей соответствующего субфрагмента (полярность верхней цепи ДНК везде 5'→3')

Установленная нами нуклеотидная последовательность изображена на схеме 2, где показаны также границы генов *groV* и *groC* и постулированные на основании этих данных С- и N-концевые аминокислотные последовательности β- и β'-белка РНК-полимеразы. Очевидно, что ген *groV* оканчивается в положении 599 перед тринуклеотидом ТАА (здесь и далее указаны нуклеотиды верхней цепи), поскольку в двух других рамках считывания терминирующие триплеты встречаются уже в начале фрагмента *EcoRI*-J. Что касается гена *groC*, то его стартовой точкой должен быть нуклеотид 673-G, так как начинающаяся с него триплетная последовательность GTG.AAA.GAT.ТТА, во-первых, не содержит терминирующих кодонов до конца фрагмента *EcoRI*-J (в отличие от обеих других рамок считывания) и, во-вторых, соответствует опубликованной для β'-субъединицы РНК-полимеразы N-концевой структуре Met-Lys-Asp-Leu- [11].

Интересно, что в гене *groC* вместо обычного иницирующего кодона АТГ содержится триплет GTG, который является значительно менее эффективным инициатором трансляции [12]. Вероятно, это обстоятельство имеет существенное значение для координации биосинтеза β- и β'-субъединиц РНК-полимеразы генами *groV* и *groC* и связано с тем, что второй аминокислотой в β'-белке является лизин, оба кодона которого начинаются с нуклеотида А. Действительно, перед обоими генами, *groV* и *groC*, имеются сходные тетрапуриновые последовательности (соответственно GAGG и GGAG), комплементарные 3'-концевой части 16S рРНК (так называемые последовательности Шайн – Дальгарно [13]) и, если бы *groC* начинался с АТГА, то, по данным работы [4], его мРНК образовала бы инициаторный комплекс с рибосомой втрое эффективнее, чем мРНК гена *groV*, начинающегося с тетраплексида АТGG. Замена в гене *groC* иницирующего кодона АТГ на GTG должна выравнивать уровень трансляции обоих генов, приводя к синтезу β- и β'-субъединиц РНК-полимеразы в приблизительно равных количествах.

\* См. примечание на стр. 1580.

С х е м а 2

MetValTyrSerTyr... ..PheGluPhe...

HindII EcoRI

...GTCGACTGTCAGCGAGCTGAGGAACCCATGGGTTTACTCC1AT...TTCGAATTC...  
 ...CAGCTGAACAGTCGCTCGACTCCTTGGGATACCAAATGAGGATA..(K)..AAGCTTAAG..(F1)...

GluPheIleGlnArgAlaTyrAspLeuGlyAlaAspValArgGlnLysValAspLeuSerThrPheSerAsp  
 EcoRI 50

GAATTCATCCAGCGTGCCTACGATCTGGGCGCTGACGTTGTCAGAAAGTTGACCTGAGTACCTTCAGCGAT  
 CTTAAGTAGGTCGCACGCATGCTAGACCCGCGACTGCAAGCAGTCTTTCAACTGGACTCATGGAAGTCGCTA

GluGluValMetArgLeuAlaGluAsnLeuArgLysGlyMetProIleAlaThrProValPheAspGlyAla  
100 HpaII

GAAGAAGTTATGCGTCTGGCTGAAAACCTGCGCAAAGGTATGCCAATCGCAACGCCGGTGTTCGACGGTGCG  
 CTTCTTCAATACGCACGCCGACTTTTGGACGCGTTTCCATACGGTTAGCGTTGCGGCCACAAGCTGCCACGC

LysGluAlaGluIleLysGluLeuLeuLysLeuGlyAspLeuProThrSerGlyGlnIleArgLeuTyrAsp  
150 HpaII 200

AAAGAAGCAGAAATTAAGAGCTGCTGAAACTTGGCGACCTGCCGACTTCCGGTCAGATCCGCCCTGTACGAT  
 TTTCTTCTGCTTTAATTTCTCGACGACTTTGAACCGCTGGACGGCTGAAGGCCAGTCTAGGCGGCATGCTA

GlyArgThrGlyGluGlnPheGluArgProValThrValGlyTyrMetTyrMetLeuLysLeuAsnHisLeu  
HpaII 250 TaqX

GGTCGCACTGGTGAACAGTTCGAGCGTCCGGTAAACCGTTGGTTACATGTACATGCTGAAACTGAACCACCTG  
 CCAGCGTGACCACCTGTCAAGCTCGCAGGCCATTGGCAACCAATGTACATGTACGACTTTGACTTGGTGGAC

ValAspAspLysMetHisAlaArgSerThrGlySerTyrSerLeuValThrGlnGlnProLeuGlyGlyLys  
 HindII 300 HpaII TaqX 350

GTCGACGACAAAATGCACGCGCTTCCACCGGTTCTTACAGCCTGGTTACCCAGCAGCCGCTGGGTGGTAAG  
 CAGCTGCTGTTTTACGTGCGCGCAAGGTGGCCAAGAATGTCGGACCAATGGTTCGTGCGCGACCCACCATTG

AlaGlnPheGlyGlyGlnArgPheGlyGluMetGluValTrpAlaLeuGluAlaTyrGlyAlaAlaTyrThr  
400

GCACAGTTCCGGTGGTCAGCGTTTCGGGGAGATGGAAGTGTGGGCGCTGGAAGCATACGGCCGAGCATACACC  
 CGTGTC AAGCCACAGTCGCAAAGCCCTCTACCTTACACCCGCGACCTTCGTATGCCGCGTCGTATGTGG

LeuGlnGluMetLeuThrValLysSerAspAspValAsnGlyArgThrLysMetTyrLysAsnIleValAsp  
450 500

CTGCAGGAAATGCTCACCCTTAAGTCTGATGACGTGAACGGTCGTACCAAGATGTATAAAAACATCGTGGAC  
 GACGCTCCTTTACGAGTGGCAATTCAGACTACTGC ACTTGCCAGCATGGTTCACATATTTTGTAGCACCTG

GlyAsnHisGlnMetGluProGlyMetProGluSerPheAsnValLeuLeuLysGluIleArgSerLeuGly  
HpaII 550

GGCAACCATCAGATGGAGCCGGGATGCCAGAATCCTTCAACGTATTTGTTGAAAGAGATTTCGTTCCGTTGGGT  
 CCGTTGGTAGTCTACCTCGGCCGTCAGGCTTAGGAAGTTGCATAACAACCTTCTCTAAGCAAGCGACCCA

Нуклеотидная последовательность генов *rpoB* — *rpoC* *E. coli* и соответствующие аминокислотные последовательности β- и β'-белка РНК-полимеразы. Для краткости эущены опубликованные ранее нуклеотидные последовательности *EcoRI*-фрагментов К

Схема 2 (продолжение)

IleAsnIleGluLeuGluAspGlnTER

600

Hpa II

ATCAACATCGAACTGGAAGACCAGTAATTCTCGCTCAAACAGGTCAC TGCTCTTTAAACCGGCAGCGGATTGT  
TAGTTGTAGCTTGACCTTCTGCTCATTAAAGAGCGAGTTTGTCCAGTGACGAGAAATTTGGCCGTGCCTAACA

fMetLysAspLeuLeuLysPheLeuLysAlaGlnThrGluThrGluGlu

550

700

GCTAAC TCCGACGGGAGCAAATCCGTGAAAAGATTTATTAAGTTTCTGAAAGCGCAGACTGAAACCGAAGAG  
CGATTGAGGCTGCCCCGCTTAGGCACCTTCTAAATAATTTCAAAGACTTTCGCGCTGACTTTGGCTTCTC

PheAspTrnIleLysValAlaIleuValSerProAspMetIleArgSerTrpSerPheGlyGluValLysLys

750

TTTGATACGAICAAAGTTGCTCTGGTTTCGCCAGACATGAICCGTTCCGTGGTCTTTCGGTGAAGTTAAAAAG  
AAACTATGCTAGTTTCAACGAGACCAAAGCGGTCTGTACTAGGCAAGCACCAGAAAGCCACTTCAATTTTTCT

ProGluThrIleAsnTyrArgThrPheLysProGluArgAspGlyLeuPheCysAlaArgIlePheGly

Hpa II 800

850

CCGGAACCATCAACTACCGTACGTTCAAACCGAAGCGTGACGGCTTTTCTGCGCCGATCTTTGGG  
GGCTTTGGTAGTTGATGGCATGCAAGTTTGGTCTTGCACTGCCGAAAAGACGGGGCATAGAAACCC

ProValLysAspHisGluArgProArgGlyLysTyrLysArgLeuLysHisArgGlyValIleCysGluLys

Hpa II

900

CCGGTAAAAGATCACGAGCGCCCGCGCGGCAAGTACAAGCGCTGAAACACCGTGCGTCACTCTGCGAGAAG  
GGCATTCTTAGTCTCGCGGGCGCGGTTTCATGTTGCGGGACTTTGTGGCACCGCAGTAGACGCTCTTC

CysGlyValGluAlaThrGlnThrLysValArgArgGluArgMetSerHisIleGluLeuAlaSerProThr

950

1000

TGGGGGTTGAAGCGACCCAGACCAAAGTACGGCGTGAGCGTATGAGCCACATCSAACTGGCTTCCCAGACT  
ACGGCGCAACTTCGCTGGGTCTGGTTTCATGCGGCACCTCGCATACTCGGTGTAGCTTGACCAGAGGGGCTGA

AlaHisMetTrpPheLeuLysSerLeuProSerArgIleGlyLeuLeuLeuAspMetProLeuArgAspIle

1050

GCGCACATGTGGTTCTGAAATCGCTGCCGTCCCCTATCGGTCTGCTGCTCGATATGCCGCTGCCGATATC  
CGGTGTACACCAAGGACTTAGCGACGGCAGGGCATAGCCAGACGACGAGCTATACGGCGACGCGCTATAG

GluArgValLeuTyrPheGluSerTyrValValIleGluGlyGlyValThrAsnLeuGluArgGlnGlnIle

1100

TaqX

GAACGCGTACTGTACTTTGAATCCTATGTGGTTATCGAAGGCGGTGAACCAACCTGGAACGTCAGCAGATC  
CTTGCCGATGACATGAAACTTAGGATACACCAATAGCTTCCGCCACATTTGGTTGGACCTTGACGTGCTGAC

LeuThrGluGluGlnTyrLeuAspAlaLeuGluGluPheGlyAspGluPhe...

1150

EcoRI 1200

CTGACTGAAGAGCAGTATCTGGACGCGCTGGAAGAGTTCGGTGACGAATTC...  
GACTGACTTCTCGTCATAGACCTGCGCGACCTTCTCAAGCCACTGCTTAAG...

[1] и F<sub>1</sub> [2]. Подчеркнуты участки, комплементарные 3'-концевой части 16S рРНК, и иницирующие кодоны в обоих генах. Указаны экспериментально найденные сайты расщепления рестриктазами

## ЛИТЕРАТУРА

1. Гуревич А. И., Аваков А. Э., Колосов М. Н. (1979) Биоорган. химия, 5, 1735–1739.
2. Овчинников Ю. А., Свердлов Е. Д., Липкин В. М., Монастырская Г. С., Чертов О. Ю., Губанов В. В., Гурьев С. О., Модянов Н. Н., Грипкевич В. А., Макарова И. А., Марченко Т. В., Половникова И. Н. (1980) Биоорган. химия, 6, 655–665.
3. Гуревич А. И., Аваков А. Э., Киселева О. А., Колосов М. Н. (1978) Биоорган. химия, 4, 628–638.
4. Tanaguchi T., Weissmann C. (1978) J. Mol. Biol., 118, 533–565.
5. Boros I., Kiss A., Venetianer P. (1979) Nucl. Acids Res., 6, 1817–1830.
6. Collins J. (1979) Mol. Gen. Genet., 173, 217–220.
7. Maxam A. M., Gilbert W. (1977) Proc. Nat. Acad. Sci. USA, 74, 560–564.
8. Коробко В. Г., Грачев С. А. (1977) Биоорган. химия, 3, 1420–1422.
9. Коробко В. Г., Грачев С. А., Колосов М. Н. (1978) Биоорган. химия, 4, 1281–1283.
10. Гуревич А. И., Аваков А. Э. (1979) Биоорган. химия, 5, 301–304.
11. Fujiki H., Zurek G. (1975) FEBS Lett., 55, 242–244.
12. Watson J. D. (1976) Molecular Biology of the Gene, p. 361, Benjamin Inc., London.
13. Shine J., Dalgarno L. (1974) Proc. Nat. Acad. Sci. USA, 71, 1342–1346.

Поступило в редакцию  
9.IV.1980

После доработки  
24.VI.1980

### STRUCTURE OF A CENTRAL PART OF *E. COLI* OPERON *rpoBC*. NUCLEOTIDE SEQUENCE OF THE GENE FOR $\beta$ SUBUNIT OF RNA POLYMERASE

GUREVICH A. I., IGOSHIN A. V., KOLOSOV M. N.

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

We have analysed, by the Maxam-Gilbert method, an *EcoRI* fragment (of ca. 1200 b.p.) of phage  $\lambda$ rif<sup>d</sup>47 DNA containing a central part of *rpoBC* operon of *E. coli*. The sequences were thus determined at the distal end of gene *rpoB* (600 b. p.) and at the proximal end gene *rpoC* (528 b.p.), coding for 200 C-terminal and 176 N-terminal amino acids of RNA polymerase proteins  $\beta$  and  $\beta'$  respectively.