



ПИСЬМА РЕДАКТОРУ

УДК 547.963.32.02+547.96.02

ПЕРВИЧНАЯ СТРУКТУРА ФРАГМЕНТА *EcoRI*-F ГЕНОВ *groB*, *C*
И СООТВЕТСТВУЮЩИХ ОБЛАСТЕЙ β - И β' -СУБЪЕДИНИЦ
РНК-ПОЛИМЕРАЗЫ *E. COLI*Монастырская Г. С., Губанов В. В., Гурьев С. О.,
Динкин В. М., Свердлов Е. Д.Институт биоорганической химии им. М. М. Шмякина
Академии наук СССР, Москва

В процессе параллельного исследования структуры *groB*, *C*-оперона *E. coli* и аминокислотной последовательности β - и β' -субъединиц РНК-полимеразы ранее мы определили структуру *EcoRI*-С-фрагмента ДНК, содержащего большую среднюю часть оперона *groB*, и соответствующую ей аминокислотную последовательность β -субъединицы [1].

В данном сообщении приводятся результаты определения первичной структуры *EcoRI*-F-фрагмента ДНК (см. в [1] рис. 1), содержащего участки, кодирующие С-концевую последовательность β -субъединицы, N-концевую последовательность β' -субъединицы и межцистровную область.

Общая стратегия исследования была аналогична примененной нами в работе [1]. Схема установления нуклеотидной последовательности *EcoRI*-F дана на рис. 1. Практически вся структура определена по двум цепям. Установленная последовательность, составляющая 1207 пар оснований, представлена на рис. 2. Ранее [2] мы предположили, что трипептический пептид β -Т-IX-1-2, полученный из β -субъединицы, для которого была установлена аминокислотная последовательность Ser-Leu-Gly-Phe-Asx-Phe-Glx (Leu, Glx, Asx, Glx) (см. координаты 568-600), является С-концевым в молекуле. При гидролизе β -субъединицы протеиназой из *St. aureus* из этой же области полипептидной цепи были выделены пептиды Phe-Arg-Ser-Leu-Gly-Phe-Asn-Phe-Glu и Leu-Glu-Asp-Glu (562-588 и 589-600). Нуклеотидные последовательности, соответствующие данным пептидам, находятся непосредственно перед терминирующим кодоном ТАА, подтверждая, что эти пептиды являются С-концевыми в β -субъединице. Таким образом, С-концевая последовательность β -субъединицы, кодируемая фрагментом *EcoRI*-F, содержит 200 аминокислотных остатков (рис. 2).

В нетранспирируемой области на расстоянии 65 пар оснований от терминирующего кодона обнаруживается последовательность GGAG (669-672), комплементарная 3'-концу 16S РНК [3, 4] в сочетании с редко иницирующим кодоном GTG (680-682) [5-7]. Это сочетание может являться сигналом начала трансляции β' -субъединицы.

Недавно были детально исследованы структуры информационных РНК прокариот, предшествующих точке инициации трансляции [8], и сделан

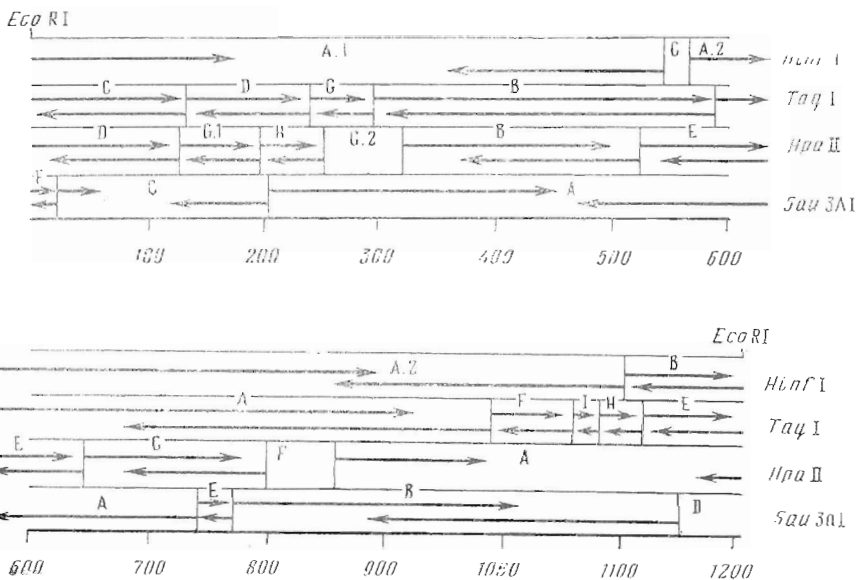


Рис. 1. Детальная карта расщепления фрагмента *EcoRI-F* рестрикционными эндонуклеазами *HinfI*, *TagI*, *HpaII* и *Sau3AI* и схема установления его последовательности. Субфрагменты, получающиеся при действии рестриктаз, изображены прямоугольниками. Стрелки обозначают длину установленной последовательности элементарных цепей данного субфрагмента

вывод, что в пределах 21–23 оснований от иницирующего кодона может находиться nonsense-кодон UAA или UGA и никогда не обнаруживается третий стоп-кодон UAG. В установленной нами последовательности на расстоянии 22 н.п. от кодона GTG расположен стоп-кодон TAA (658–660) и не обнаружен TAG, что может служить дополнительным подтверждением локализации начальной точки трансляции β' -субъединицы в указанном месте.

С помощью автоматического секвенатора для β' -субъединицы РНК-полимеразы нами была установлена N-концевая аминокислотная последовательность Met-Lys-Asp-Leu-Leu-Lys-Phe-Leu-, которая находится в соответствии с найденной ранее [9].

Если принять за начальную точку трансляции кодон GTG (680–682), то аминокислотная последовательность, гипотетически выведенная из нуклеотидной последовательности *EcoRI-F*-фрагмента, совпадает с найденной. Следовательно, кодон GTG является иницирующим для β' -субъединицы и соответствующая фрагменту *EcoRI-F* N-концевая аминокислотная последовательность β' -субъединицы содержит 176 аминокислотных остатков.

Интерцистронная область (604–679) между генами β - и β' -субъединиц составляет 76 пар оснований. Это существенно меньше, чем длина интерцистронной области между геном *rplL* белка L7/L12 и геном *rpoV* β -субъединицы РНК-полимеразы [10]. В настоящее время предполагают, что в интерцистронной области *rplL+rpoV* находится некий дополнительный регуляторный элемент, возможно аттенюатор [11–13]. С другой стороны, длина интерцистронной области между геном *rpoA*, кодирующим α -субъединицу, и ближайшим рибосомным геном всего 25 пар оснований [14], что делает мало вероятным существование контрольных элементов транскрипции в этом районе.

При анализе установленной нами последовательности в интерцистронной области *rpoV-rpoC* в положениях 623–635 и 638–650 был обнаружен несовершенный обратный повтор. В настоящее время нет данных, указывающих на функциональные свойства интерцистронной области

	10	20	30	40	50	60
I-60	<u>GAATTCATCCAGCGTGGGTACGATCTGGGCGGTGACGTTTCGTCAGAAAGTTACCTGAGT</u> <u>GluPheIleGlnArgAlaTyrAspLeuGlyAlaAspValArgGlnLysValAspLeuSer</u>					
61-120	<u>ACCTTCAGCGATGAAGAAGTTATGCGTCTGGCTGAAAACCTGCGCAAAGGTATGCCAATC</u> <u>ThrPheSerAspGluGluValMetArgLeuAlaGluAsnLeuArgLysGlyMetProIle</u>					
121-180	<u>GCAACGCCGCTGTTTCGACGGTGCAGAAAGAAGCAGAAATTAAGAGCTGCTGAAACTTGGC</u> <u>AlaThrProValPheAspGlyAlaLysGluAlaGluIleLysGluLeuLeuLysLeuGly</u>					
181-240	<u>GACCTGCCGACTTCCGGTCACATCCGCCTGTACGATGGTTCGCACTGGTGAACAGTTCGAG</u> <u>AspLeuProThrSerGlyGlnIleArgLeuTyrAspGlyArgThrGlyGluGlnPheGlu</u>					
241-300	<u>CGTCCGGTAAACCGTTGGTTACATGTACATGCTGAAACTGAACCAC*[*]TGGTCGACGACAAG</u> <u>ArgProValThrValGlyTyrMetTyrMetLeuLysLeuAsnHisLeuValAspAspLys</u>					
301-360	<u>ATGCCAGCGCGTTCCACCGTTCCTACAGCCT*[*]TGGTACTCAGCAGCGCTGGGTGGTAAG</u> <u>MetHisAlaArgSerThrGlySerTyrSerLeuValThrGlnGlnProLeuGlyGlyLys</u>					
361-420	<u>GCACAGTTCGGTGGTCAGCGTTTCGGGCAGATCGAAGTGTGGCGCTGGAAGCATACGGC</u> <u>AlaGlnPheGlyGlyGlnArgPheGlyGluMetGluValTrpAlaLeuGluAlaTyrGly</u>					
421-480	<u>GCAGCATACCCCTGCAGGAAATGCTCACCGTTAAGTCTGATGACGTGAACGGTTCGTACC</u> <u>AlaAlaTyrThrLeuGlnGluMetLeuThrValLysSerAspAspValAsnGlyArgThr</u>					
481-540	<u>AAGATGTATAAAAACATCGTGGACGGCAACCATCAGATGGAGCCGGGCATCCAGAATCC</u> <u>LysMetTyrLysAsnIleValAspGlyAsnHisGlnMetGluProGlyMetProGluSer</u>					
541-600	<u>TTCAACGTATTGTTGAAAGAGATTGCTTCGCTGGGTATCAACATCGAACTGGAAGACGAC</u> <u>PheAsnValLeuLeuLysGluIleArgSerLeuGlyIleAsnIleGluLeuGluAspGlu</u>					
601-660	<u>TAATTCTCGCTCAAACAGGTCACTGCTGTCGGGTTAAACCCGGCAGCGGATTGTGCTAA</u> <u>ter</u>					
661-721	<u>CTCCGACGGGAGCAAATCCCTGAAAGATTTATTAAGTTTCTGAAAGCCGAGACTAAAACC</u> <u>fMetLysAspLeuLeuLysPheLeuLysAlaGlnThrLysThr</u>					
722-781	<u>GAAGAGTTTGATGCGATCAAATGCTCTGGCTTCGCCAGACATGATCCGTTTCATGGTCT</u> <u>GluGluPheAspAlaIleLysIleAlaLeuAlaSerProAspMetIleArgSerTrpSer</u>					
782-841	<u>TTCGGTGAAGTTAAAAAGCCGGAAACCATCAACTACCGTACGTTCAAACGAGAACGTGAC</u> <u>PheGlyGluValLysLysProGluThrIleAsnTyrArgThrPheLysProGluArgAsp</u>					
842-901	<u>GGCCTTTTCTGCGCCCGTATCTTTGGGCCGGTAAAGATTACGACTGCCTGTGCGGTAAG</u> <u>GlyLeuPheCysAlaArgIlePheGlyProValLysAspTyrGluCysLeuCysGlyLys</u>					
902-961	<u>TACAAGCGCCTGAAACACCGTGGCGTCATCTGTGAGAAGTGGCGGTTGAAGTGACCCAG</u> <u>TyrLysArgLeuLysHisArgGlyValIleCysGluLysCysGlyValGluValThrGln</u>					
962-1021	<u>ACTAAAGTACGCCGTGAGCGTATGGGCCACATCGAACTGGCTTCCCGACTGCGCACATC</u> <u>ThrLysValArgArgGluArgMetGlyHisIleGluLeuAlaSerProThrAlaHisIle</u>					
1022-1081	<u>TGCTTCCTGAAATCGCTGCCGTCCCGTATCGGTCTGCTGCTCGATATGCCGCTGCGCGAT</u> <u>TrpPheLeuLysSerLeuProSerArgIleGlyLeuLeuLeuAspMetProLeuArgAsp</u>					
1082-1141	<u>ATCGAACCGCTACTGTACTTTGAATCCTATGTGGTTATCGAAGCGGATGACCAACCTG</u> <u>IleGluArgValLeuTyrPheGluSerTyrValValIleGluGlyGlyMetThrAsnLeu</u>					
1142-1201	<u>GAACGTCAGCAGATCCTGACTGAAGAGCAGTATCTGGACCGCTGGAAGAGTTCGGTGAC</u> <u>GluArgGlnGlnIleLeuThrGluGluGlnTyrLeuAspAlaLeuGluGluPheGlyAsp</u>					
1202-1207	<u>GAATTC</u> <u>GluPhe</u>					

Рис. 2. Нуклеотидная последовательность фрагмента *EcoRI*-F и аминокислотные последовательности С-концевой части β-субъединицы и N-концевой части β'-субъединицы РНК-полимеразы *E. coli*. Приведена последовательность комплементарной цепи ДНК, адекватная последовательности мРНК. Ориентировка фрагмента *EcoRI*-F соответствует приведенной на рис. 1. Подчеркнуты аминокислотные последовательности, структура которых установлена анализом пептидов, полученных при расщеплении белка трипсином и прогеиназой из *Staphylococcus aureus*. Двойной чертой подчеркнута N-концевая последовательность β'-субъединицы. В области 623-650 в рамки взяты последовательности межцистронной области, несущие несовершенный обратный повтор. С* — остатки 5-метилцитидина

rpoB-rpoC, однако ее размер допускает существование какого-либо регуляторного элемента транскрипции или трансляции. Ответ на вопрос о функции этого промежутка может дать анализ свойств мутантов с изменениями в этой области.

ЛИТЕРАТУРА

1. Овчинников Ю. А., Свердлов Е. Д., Липкин В. М., Монастырская Г. С., Чертов О. Ю., Губанов В. В., Гурьев С. О., Модянов Н. Н., Гринкевич В. А., Макарова И. А., Марченко Т. В., Половникова И. Н. (1980) *Биоорган. химия*, **6**, 655–665.
2. Липкин В. М., Марченко Т. В., Хохряков В. С., Половникова И. Н., Потопенко Н. А., Модянов Н. Н., Овчинников Ю. А. (1980) *Биоорган. химия*, **6**, 332–342.
3. Shine J., Dalgarno L. (1974) *Proc. Nat. Acad. Sci. USA*, **71**, 1342–1346.
4. Steitz J. A., Jakes K. (1975) *Proc. Nat. Acad. Sci. USA*, **72**, 4734–4738.
5. De Wachter R., Vandenberghe A., Merregaert J., Contreras R., Fiers W. (1971) *Proc. Nat. Acad. Sci. USA*, **68**, 585–589.
6. Schaller H., Beck E., Takanami M. (1978) in *Single-stranded DNA phages* (Denhardt D. T., Dressler D., Ray D. S., eds), Cold Spring Harbor Laboratory, N. Y.
7. Dickson R. C., Abelson J., Barnes W. M., Reznikoff W. S. (1975) *Science*, **187**, 27–35.
8. Atkins J. F. (1979) *Nucl. Acids Res.*, **7**, 1035–1041.
9. Fujiki H., Zurek G. (1975) *FEBS Lett.*, **55**, 242–244.
10. Post L. E., Strycharz G. D., Nomura M., Lewis H., Dennis P. P. (1979) *Proc. Nat. Acad. Sci. USA*, **76**, 1697–1701.
11. Yamamoto M., Nomura M. (1978) *Proc. Nat. Acad. Sci. USA*, **75**, 3891–3895.
12. Linn T., Scaife J. (1978) *Nature*, **276**, 33–37.
13. Dennis P. P., Fiil N. P. (1979) *J. Biol. Chem.*, **254**, 7540–7547.
14. Post L. E., Nomura M. (1979) *J. Biol. Chem.*, **254**, 10 604–10 606.

Поступило в редакцию
17.III.1980

PRIMARY STRUCTURE OF *EcoRI*-F FRAGMENT OF *rpoB*, *C* GENES AND CORRESPONDING FRAGMENTS OF β - AND β' -SUBUNITS OF RNA POLYMERASE FROM *E. COLI*

MONASTYRSKAYA G. S., GUBANOV V. V., GURYEV S. O.,
LIPKIN V. M., SVERDLOV E. D.

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

In the course of the parallel study of primary structures of β - and β' -subunits of DNA-dependent RNA polymerase from *E. coli* and its structural genes the complete nucleotide sequence (1207 base pairs) of *EcoRI*-F fragment of genes *rpoB*, *C* was established. The amino acid sequences of C-terminal part of β -subunit and N-terminal part of β' -subunit contain 200 and 176 amino acid residues, respectively.