



УДК 547.963.32.07

СИНТЕЗ ПРОМОТОРНЫХ ЛИНКЕРОВ ДЛЯ ЭКСПРЕССИИ
ГЕНОВ В *E. COLI**Коробко В. Г., Добрынин В. Н., Северцова И. В.,
Быстров Н. С., Колосов М. Н.**Институт биоорганической химии им. М. М. Шемякина
Академии наук СССР, Москва*

При конструировании и клонировании искусственных генов одной из важных проблем является их экспрессия в используемом хозяине, обычно в *E. coli*. Чтобы ген мог функционировать, он должен быть снабжен промотором и нуклеотидной последовательностью, транскрипт которой способен образовать инициаторный комплекс с рибосомой. До сих пор для этого применялись лишь фрагменты природных ДНК, например регуляторная область гена *lacZ E. coli* [1]. Между тем было бы выгодно во многих отношениях располагать синтетическими ДНК, специально приспособленными для этих целей. Поэтому мы осуществили синтез двухцепочечных ДНК, которые содержат нуклеотидную последовательность «идеального» промотора *E. coli* [2] и имеют разные концы — выступающий и тупой — для соединения с вектором и клонируемым геном.

Для получения этих промоторных линкеров мы химически синтезировали ряд олигодезоксинуклеотидов и с помощью Т4 ДНК-лигазы соединили их в короткие дуплексы, изображенные на рис. 1. Синтез олигонуклеотидов (VIII) — (XI) был выполнен фосфотриэфирным методом как в работе [3], олигонуклеотиды (I) — (VII) были синтезированы нами ранее [4], а (VIII') и (VIII'') получены из промежуточных продуктов синтеза тетрадекануклеотида (VIII).

Лигированием ундекануклеотида (I) с 5'-фосфорилированными олигонуклеотидами (pII) — (pIV) был получен дуплекс А, а из 5'-фосфорилированных олигонуклеотидов (pV) — (pVII) и нефосфорилированного декануклеотида (VIII') был синтезирован дуплекс В, энзиматическое сшивание которого с А привело к 49-членному промоторному линкеру (рис. 2, Л). Образование дуплекса В и его сшивание с А сопровождалось побочной реакцией самолигирования по 5'-выступающим концам pT-T-T-A-A-G с образованием димерного (40-членного) дуплекса, в центральной части которого имеются две неканонические пары G-T. Чтобы в дальнейшем избежать этой побочной реакции, дуплекс А был заменен более длинным (В), полученным из предыдущего путем присоединения к нему додекануклеотида (pV). Лигированием олигонуклеотидов (pVI) с (VIII) на тетрадекануклеотидной матрице (pVII) был синтезирован дуплекс Г, в котором верхняя цепь была достроена до тупого 3'-конца с помощью dATP, dCTP и dTTP под действием ДНК-полимеразы I *E. coli*. Образовавшийся дуплекс Д был соединен с В в 52-членный линкер (рис. 2, М), который содержит всю

тодом нуклеотидных карт, а двухцепочечных полинуклеотидов — методом Максама — Гилберта [5]. Конечные продукты синтеза, дуплексы *L*, *M* и *H*, были разделены на комплементарные цепи двухмерным электрофорезом (сначала мы ацетилцеллюлозе при pH 3,5, а затем в денатурирующем 20% полиакриламидном геле), и нуклеотидная последовательность каждой из цепей была определена методом [6] с модификациями, описанными в статьях [7, 8].

Синтезированные нами двухцепочечные ДНК *L* — *H* (рис. 2) структурно соответствуют «идеальному» промотору *I* Шерера и др. [2], причем ДНК *H* дополнительно содержит протяженный участок (последовательность Шайн — Дальгарно [9]), который комплементарен 3'-концу 16S рРНК *E. coli* (*K*) и находится на таком же расстоянии от иницирующего кодона, как в *lacZ* и других генах *E. coli* и колифагов [10]. Благодаря этому ДНК *H* должна обеспечивать в *E. coli* эффективную экспрессию структурных генов, полученных обратной транскрипцией или химическим синтезом и не имеющих участков инициации ни транскрипции, ни трансляции. Следует отметить, что нуклеотидная последовательность этой синтетической ДНК спланирована таким образом, что участки с разными биохимическими функциями (узнавание и связывание РНК-полимеразы, инициация транскрипции, комплексообразование с рибосомой и инициация биосинтеза белка) разделены сайтами разных эндонуклеаз рестрикции. Это позволяет относительно легко варьировать нуклеотидную последовательность и длину каждого участка независимо от других с целью изучения зависимости между структурой и функцией, а также для получения максимальной экспрессии гена.

ЛИТЕРАТУРА

- Goeddel D. V., Heyneker H. L., Kozumi T., Arentzen R., Itakura K., Yansura D. G., Ross M. J., Miozzari G., Crea R., Seeburg P. H. (1979) *Nature*, **281**, 544–548.
- Scherer G. E. F., Walkinshaw M. D., Arnott S. (1978) *Nucl. Acids Res.*, **5**, 3759–3773.
- Коробко В. Г., Добрынин В. Н., Болдырева Е. Ф., Северцова И. В., Чернов Б. К., Колосов М. Н., Городецкий С. И., Слюсаренко А. Г., Копелинская Т. В., Лисенков А. Ф., Дубинин Н. П. (1979) *Биоорганич. химия*, **12**, 1802–1915.
- Добрынин В. Н., Коробко В. Г., Северцова И. В., Колосов М. Н. (1980) *Биоорганич. химия*, **6**, 783–785.
- Maxam A. M., Gilbert W. (1977) *Proc. Nat. Acad. Sci. USA*, **74**, 560–564.
- Maxam A. M., Gilbert W. (1980) in: *Methods in Enzymology*, vol. 65, *Nucleic Acids*, Part I (Grossman L., Moldave K., eds), pp. 499–560, Acad. Press, N. Y.
- Коробко В. Г., Грачев С. А. (1977) *Биоорганич. химия*, **3**, 1420–1422.
- Коробко В. Г., Грачев С. А., Колосов М. Н. (1978) *Биоорганич. химия*, **4**, 1281–1283.
- Shine J., Dalgarno L. (1974). *Proc. Nat. Acad. Sci. USA*, **71**, 1342–1346.
- Stetz J. A. (1979) in: *Biological Regulation and Development*, vol. 1, *Gene Expression* (Goldberger R. F., ed.), pp. 219–277, Plenum Press, N. Y.

Поступило в редакцию
4.VII.1980

SYNTHESIS OF PROMOTER LINKERS FOR THE EXPRESSION OF GENES IN *ESCHERICHIA COLI*

IKOROBKO V. G., DOBRYNIN V. N., SEVERTSOVA I. V., BYSTROV N. S., KOLOSOV M. N.

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

Three double-stranded DNAs have been synthesized by T4 DNA ligase catalyzed joining of oligodeoxynucleotides prepared by the improved phosphotriester method. One of them is 52-membered and contains an ideal promoter sequence of Scherer et al. [2], whereas the other one is 49-membered and contains all but three nucleotides (-1 through +2) of that sequence. The third DNA (a 67 membered) comprises both the promoter sequence and a ribosome binding site composed of an initiator triplet and an extended Shine-Dalgarno sequence which are separated by a *Bcl* I site while the RNA polymerase recognition and binding sites are flanked with *Eco* RI-*Hind* III and *Hind* III-*Kpn* I sequences, respectively. The three DNAs are potential promoter linkers for genes to be cloned in *E. coli*, especially the third RNA is designed to promote the expression of structural genes which are prepared by reverse transcription or by chemical synthesis and lack initiation regions for both transcription and translation.