



УДК 547.963.32.02

НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ПРОМОТОРНОГО
УЧАСТКА A_1 — A_3 ФАГА T7**Коробко В. Г., Чувпило С. А., Грачев С. А., Колосов М. Н.***Институт биоорганической химии им. М. М. Шемякина
Академии наук СССР, Москва***Грачев М. А., Зайчиков Е. Ф., Плетнев А. Г.***Новосибирский институт органической химии
Сибирского отделения Академии наук, Новосибирск*

Транскрипция ранних генов фага T7 инициируется РНК-полимеразой *E. coli* преимущественно на промоторном участке A, расположенном вблизи левого конца фагового генома [1, 2]. Этот участок содержит три сильных промотора — A_1 , A_2 и A_3 [1], эффективность которых по-разному зависит от внешних условий [3]. Для двух из них, A_2 и A_3 , была установлена первичная структура [4, 5], причем было обнаружено, что у ряда промоторов имеется сходная гептануклеотидная последовательность (так называемый бокс Прибнова), центр которой приблизительно на 10 нуклеотидов предшествует началу транскрипции [4—6]. Недавно нами была выяснена первичная структура фрагмента *HpaII*-100 ДНК T7 [7], который содержит промотор A_2 и, согласно рестриктивной карте ранней области фага [8], отстоит на 513 нуклеотидов от начала генома. В развитие этой работы мы исследовали смежные фрагменты ДНК T7 и установили изображенную на рисунке нуклеотидную последовательность всего инициаторного участка A ранних генов фага T7, включая первый главный промотор A_1 .

Для определения нуклеотидной последовательности, расположенной за промотором A_2 , мы исследовали фрагмент *HpaII/AluI*-190, который в ДНК T7 примыкает к фрагменту *HpaII*-100 справа и, по данным [9], содержит промоторную область A_3 . Он был получен из более крупного концевой фрагмента *HaeIII*-1300 [7] расщеплением эндонуклеазами *AluI* и *HpaII* и после введения 5'- 32 P-метки (действием щелочной фосфатазы, а затем $[\gamma$ - 32 P]АТР (> 1000 Ки/ммоль) и Т4-полинуклеотидкиназы) был выделен электрофорезом в 7% полиакриламидном геле (ПАГ). Этот фрагмент денатурировали 0,3 н. КОН, цепи разделяли в 6% ПАГ по методу [10] и анализировали, как описано ранее [11, 12]. При этом было идентифицировано 86 нуклеотидов в быстродвижущейся цепи и 94 в менее подвижной цепи (на рисунке соответственно нуклеотиды 629—714 в *l*-цепи и 695—787 в *r*-цепи), что дало 20-нуклеотидное перекрытие. Структуру 5'-концевых участков обеих цепей определяли двухмерным разделением продуктов химической деградации по методу [13]; при этом были идентифицированы нуклеотиды 616—632 в *l*-цепи и 789—799 в *r*-цепи. Установленная таким образом последовательность включает две опубликованные ранее (689—740 [4, 5] и 681—730 [9]) и полностью с ними согласуется.

(l) 5' ... GTG³⁴⁰ TA . TTTTAA³⁵⁰ CA³⁶⁰TTGTCTT³⁷⁰ ATTAATACAA³⁸⁰ CTCACTATA³⁹⁰ GGAGAGACAA-
 (r) 3' ... CACT AT . AAAATTA GTAACAGAAA TAATTATGTT GAGTGATATT CCTCTCTGTT-

400 410 420 430 440 450
 CTTAAAGAG⁴⁰⁰ CTTAAAGAT⁴¹⁰ TAATTTAAA⁴²⁰ TTTATCAAAA⁴³⁰ AGAGTATTG⁴⁴⁰ CTTAAAGTC⁴⁵⁰-
 GAATTTCTCT GAATTTTCTA ATTAAATTTT AAATAGTTTT TCTCATAACT GAATTTCAGA-

A₁

460 470 480 490 500 510
 AACCTATA⁴⁶⁰ G⁴⁷⁰ ATACT⁴⁸⁰ TACAG⁴⁹⁰ CCATCGAGAG⁵⁰⁰ GGACACGGCG⁵¹⁰ AACAGCCATC⁵¹⁰ CC ... GTCAA-⁸
 TTGGATATCC⁴⁶⁰ TATGAA⁴⁷⁰ TGTC⁴⁸⁰ GGTAGCTCTC⁴⁹⁰ CCTGTGCCG⁵⁰⁰ TTGTCGGTAG⁵¹⁰ GG ... CAGTT-
 pppAUCGAGAG GG ...

HpaII 520 530 540 550 560 570
 CCGGATAAG⁵²⁰ AGACAGCCTG⁵³⁰ ATAAGTCGC⁵⁴⁰ CGAAAAACAG⁵⁵⁰ GTATTGACAA⁵⁶⁰ CATGAAGTAA-
 GG⁵⁷⁰CTATTCA TCTGTCCGAC TATTACGCGT GCTTTTTGTC CATAACTGTT GACTTTCATT-

A₂

580 590 600 610 HpaII 620 630
 CATGCAGTAA⁵⁸⁰ GATACAAATC⁵⁹⁰ GCTAGGTAAC⁶⁰⁰ ACTAGCAGCG⁶¹⁰ TCAAC⁶²⁰ CCGGGG⁶³⁰ CGCACAGTGC-
 GTACGTC⁵⁸⁰ ATT⁵⁹⁰ CTAT⁶⁰⁰ GTTTAG⁶¹⁰ CGATCCATTG⁶²⁰ TGATCGTCGC⁶³⁰ AGTTGGG⁶³⁰ CCC GCGTGTCCAG-
 pppGCUAGGUAAC ...

640 650 660 670 680 690
 CTCTAGGTG⁶⁴⁰ CTTAAGCGCA⁶⁵⁰ CCACGGCAC⁶⁶⁰ TAAGGTGAA⁶⁷⁰ CAAAACGGTT⁶⁸⁰ GACAACATGA-
 GAGATCCACT⁶⁹⁰ GAATTCGCGT GGTGCCGTGT ATTCCACTTT GTTTTGCCAA CTGTTGACT-

A₃

700 710 720 730 740 750
 AGTAAACGG⁷⁰⁰ T⁷¹⁰ ACGAT⁷²⁰ GTACC⁷³⁰ ACATGAAACG⁷⁴⁰ ACAGTGAGTC⁷⁵⁰ ACCACACTGA⁷⁵⁰ AAGGTGATGC-
 TCAATTTGCCA⁷⁰⁰ TGCTGCA⁷¹⁰ ATGG⁷²⁰ TGTACTTTGC⁷³⁰ TGCACTCAG⁷⁴⁰ TGGTGTGACT⁷⁵⁰ TTCCACTACG-
 pppAUGAAACG AC ...

760 770 780 790 800
 GGTCTAACGA⁷⁶⁰ AACCTGACCT⁷⁷⁰ AAGACGCTCT⁷⁸⁰ TTAACAA . CT⁷⁹⁰ GGTAAGAGC⁸⁰⁰ T ...
 CCAGATTGCT⁷⁶⁰ TTGGACTGGA⁷⁷⁰ TTCTGCGAGA⁷⁸⁰ AATTGTT . GA⁷⁹⁰ CCATTTCTCG⁸⁰⁰ A ...
 AluI

Нуклеотидная последовательность промоторного участка A₁ — A₃ фага T7

Для выяснения нуклеотидной последовательности, предшествующей промотору A₂, т. е. примыкающей к фрагменту *HpaII*-400 слева, мы использовали полученный ранее [7] терминальный фрагмент *HpaII*-500. Его денатурировали нагреванием и цепи разделяли в 5% ПАГ; введение 5'-метки и структурный анализ проводили так же, как в случае фрагмента *HpaII/AluI*-190. Было найдено, что у менее подвижной цепи 5'-концевая последовательность совпадает с установленной для левого конца ДНК фага T7 [7] и, следовательно, представляет собой *l*-цепь. Мы определили 174-членную 5'-концевую последовательность быстро движущейся цепи (на рисунке нуклеотиды 337—510 *r*-цепи), которая должна содержать промотор A₁, поскольку он картирован [9] в районе нуклеотида 460, считая от левого конца фаговой ДНК. Чтобы точно локализовать этот промотор, мы исследовали мРНК, считываемую с него РНК-полимеразой *E. coli* при использовании [γ -³²P]АТФ в качестве единственного меченого из четырех нуклеозидтрифосфатов; при этом матрицей служила ДНК делеционного мутанта T7D111, у которого отсутствуют промоторы A₂ и A₃ [8]. Продукт транскрипции был выделен электрофорезом в 3% ПАГ и подвергнут частичному расщеплению пиперидином при 20° с

последующим электрофорезом на ацетилцеллюлозе и гомохроматографией. В результате были определены 10 первых нуклеотидов этой мРНК ($^{32}\text{pppAUCGAGAGGG}$), что позволило установить стартовую точку транскрипции с промотора A_1 (нуклеотид 473) и постулировать для бокса Прибнова положение 460—466.

Из результатов этой работы следует, что промоторы A_1 , A_2 и A_3 находятся на примерно равном расстоянии один от другого (около 120 н. п., что соответствует 12 виткам спирали ДНК) и каждый содержит гексануклеотид PuTTGAC в l -цепи на расстоянии 30—35 нуклеотидов перед стартовой точкой, т. е. в предполагаемом участке узнавания РНК-полимеразы [6]. Промоторы A_2 и A_3 содержат идентичную 17-членную последовательность на значительном удалении от старта транскрипции (нуклеотиды 554—570 и 679—695). По структуре боксов Прибнова промотор A_1 очень похож на промотор P_L фага λ [6, 14] и сильно отличается от A_2 и A_3 , которые сходны между собой; очевидно, это является причиной наблюдаемого различия [3] в функциональных свойствах ранних промоторов фага T7.

ЛИТЕРАТУРА

1. Minkley E. G., Pribnow D. (1973) *J. Mol. Biol.*, **77**, 255—277.
2. Dunn J. J., Studier F. W. (1973) *Proc. Nat. Acad. Sci. USA*, **70**, 1559—1563.
3. Dausse J.—P., Sentenac A., Fromageot P. (1976) *Eur. J. Biochem.*, **65**, 387—393.
4. Pribnow D. (1975) *Proc. Nat. Acad. Sci. USA*, **72**, 784—788.
5. Pribnow D. (1975) *J. Mol. Biol.*, **99**, 419—443.
6. Gilbert W. (1976) in: *RNA Polymerase* (Losick R., Chamberlin M., eds), pp. 193—205, Cold Spring Harbor Lab.
7. Коробко В. Г., Чувпило С. А., Колосов М. Н., Грачев М. А., Зайчиков Е. Ф., Плетнев А. Г. (1978) *Биоорганич. химия*, **4**, 1132—1134.
8. Gordon R. L., Humphries P., McConnell D. J. (1978) *Mol. Gen. Genet.*, **162**, 329—339.
9. Hsieh Tao-shin, Wang J. C. (1976) *Biochemistry*, **15**, 5776—5783.
10. Maxam A. M., Gilbert W. (1977) *Proc. Nat. Acad. Sci. USA*, **74**, 560—564.
11. Коробко В. Г., Грачев С. А. (1977) *Биоорганич. химия*, **3**, 1420—1422.
12. Коробко В. Г., Грачев С. А., Колосов М. Н. (1978) *Биоорганич. химия*, **4**, 1281—1283.
13. Коробко В. Г., Грачев С. А., Петров Н. А. (1977) *Биоорганич. химия*, **3**, 1423—1426.
14. Maniatis T., Ptashne M., Barrell B. G., Donelson J. (1974) *Nature*, **250**, 394—397.

Поступило в редакцию
31.VIII.1978

THE NUCLEOTIDE SEQUENCE OF THE A_1 — A_3 PROMOTER REGION OF PHAGE T7

KOROBKO V. G., CHUVPILO S. A., GRACHEV S. A., KOLOSOV M. N.

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

GRACHEV M. A., ZAYCHIKOV E. F., PLETNYOV A. G.

*Novosibirsk Institute of Organic Chemistry,
Siberian Branch of the Academy of Sciences of the USSR, Novosibirsk*

In the early region of T7 DNA, 170 nucleotides preceding the leftmost *HpaII* site and 182 nucleotides following the next *HpaII* site were determined by a modified Maxam — Gilbert method, whereas the DNA structure between the two restriction sites had been reported. The total sequence comprises ca. 460 base pairs and contains three major promoters A_1 , A_2 , and A_3 , the last two having previously been sequenced. To locate the A_1 initiation site, we determined the 5'-terminal sequence in a RNA transcript synthesized in vitro by *E. coli* RNA polymerase on the T7D111 DNA template lacking the A_2 — A_3 region. It was found that on A_1 , A_2 , and A_3 promoter sites transcription starts at nucleotides 473, 591, and 713, respectively (numbering from the left end of the phage DNA); a 17 nucleotide long repetition is present in positions 554—570 and 679—695; with respect to the Pribnow box, the A_1 promoter is similar to λP_L and very different from both A_2 and A_3 .