



УДК 577.217.525

## ЗАВИСИМОСТЬ УРОВНЯ ЭКСПРЕССИИ ГЕНА В *E. coli* ОТ СТРУКТУРЫ УЧАСТКА ИНИЦИАЦИИ ТРАНСЛЯЦИИ (TIR)

### I. ПЕРВИЧНАЯ СТРУКТУРА TIR

© 1995 г. А. И. Гуревич\*, Р. С. Есипов, Т. А. Качалина,  
А. Л. Каюшин, М. Д. Коростелева

Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,  
117871, Москва, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 16.03.94 г.

Проведено сравнительное изучение уровней экспрессии двух генов – интерлейкина-3 (*il3*) и эпидермального фактора роста с лидерным пептидом *OmpF* (*lompegf*) – в специально сконструированных плаزمиде, содержащих различные структуры усилителей трансляции. Показано, что кроме известных ранее участков связывания района инициации трансляции (TIR) мРНК с 16S рРНК (SD, UB1 и DB) существенную роль в усилении инициации трансляции и соответственно уровня экспрессии генов играет возможность дополнительного связывания TIR с рРНК в участке UB2 (область –30 – –60 н. от иницирующего кодона).

*Ключевые слова:* нуклеиновые кислоты, мРНК, рРНК, инициаторный комплекс, участок инициации трансляции, экспрессия генов.

Взаимосвязь структуры мРНК в участке инициации трансляции (translation initiation region, TIR) и эффективности инициации трансляции (см. обзор [1]), определяющей уровень экспрессии генов, давно привлекает внимание исследователей. Важное практическое значение этой проблемы – создание рекомбинантных ДНК с целью гетерологичной экспрессии в *E. coli*. Подробно изучены основные принципы построения TIR.

– TIR включает около 50 нуклеотидов, покрываемых рибосомой так, что иницирующий кодон AUG делит его на две примерно равные ветви – нетранслируемую (mTIR) и транслируемую (MTIR) [2].

– В ветви mTIR расположена последовательность Шайна-Дальгарно (SD), комплементарная части 3'-концевой последовательности 16S рРНК [3], хотя наличие SD не является строго обязательным и ее отсутствие может компенсироваться другими взаимодействиями мРНК с рибосомой [4 - 6]. Длина участка ветви mTIR между кодоном

AUG и последовательностью SD обычно составляет 6 - 12 н., и для эффективной инициации трансляции существует оптимальное соотношение между этой длиной и структурой триплета, предшествующего инициаторному кодону (например, наилучшим является расстояние в 7 н. при структуре –1-триплета ATA, AGA или GTA) [7].

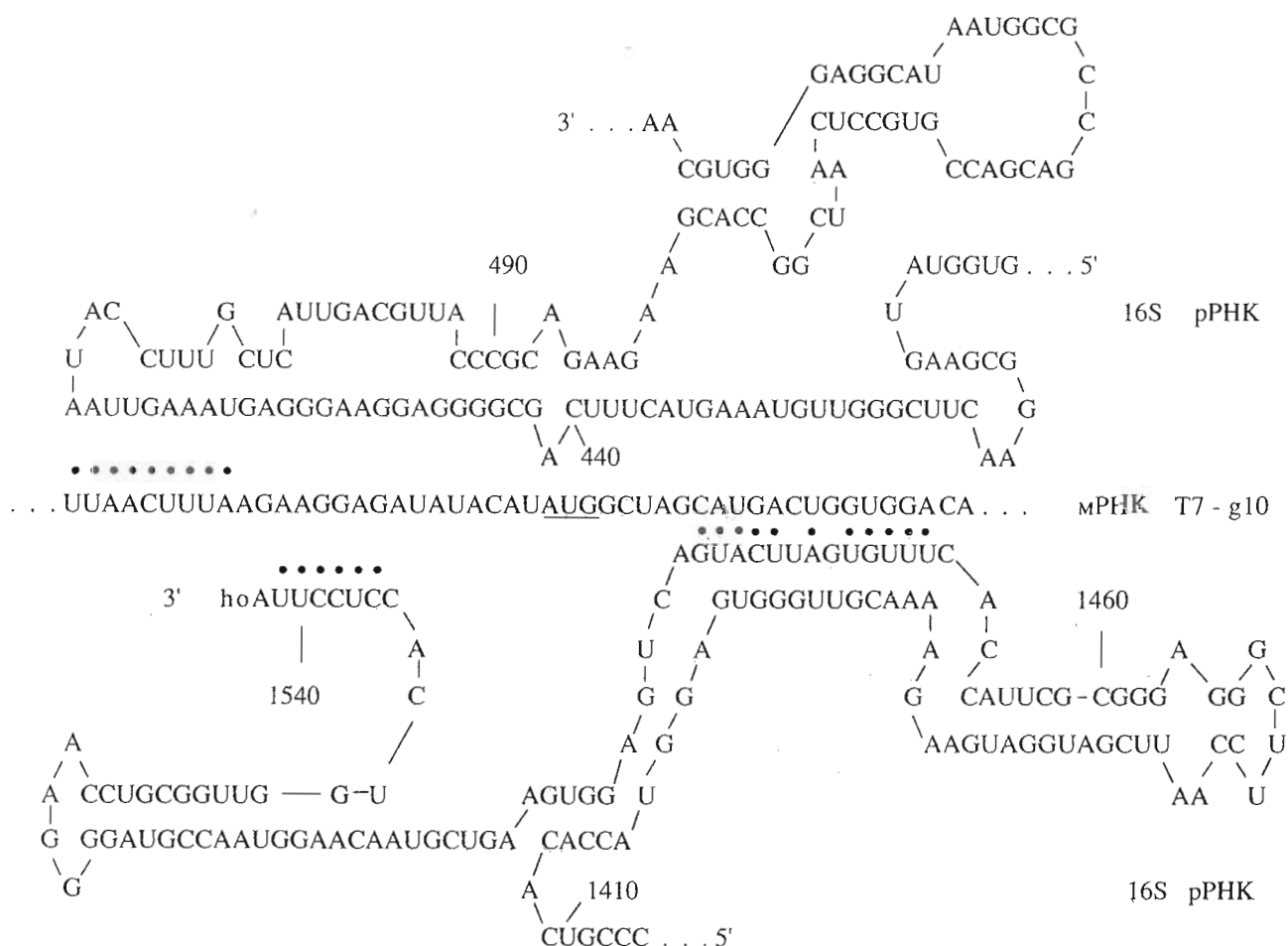
– Ветви mTIR и MTIR могут содержать элементарные последовательности, способные образовать шпильку, с кодоном AUG в экспонированной петле; однако участки вблизи AUG не должны входить в состав стабильных дуплексов. Эффективность инициации трансляции обратно пропорциональна стандартной свободной энергии образования вторичной структуры мРНК в районе TIR,  $\Delta G^0 = -RT \ln[\text{нетранслируемая форма}]/[\text{транслируемая форма}]$ . Если  $\Delta G^0$  ниже порогового значения (–10 – –12.5 ккал/моль), то трансляция блокируется. Следует отметить, что, если  $\Delta G^0$  выше определенного значения, пропорциональность не соблюдается [8 - 10].

– Специфичность нуклеотидной последовательности до и после SD вряд ли играет существенную роль в эффективности инициации трансляции. Скорее имеет значение удельный вес остатков А и U [11, 12].

– Эффективность инициации трансляции значительно увеличивается при наличии в мРНК кроме ближайшей к AUG последовательности SD (SD1) также второго участка связывания рибосомы (SD2 + AUG2), отделенного от SD1 А/У-богатой

Сокращения: TIR, mTIR и MTIR – участок инициации трансляции, его нетранслируемая и транслируемая ветви; SD – последовательность Шайна-Дальгарно; ASD – anti-SD, 3'-концевая последовательность 16S рРНК; UB – upstream box, часть mTIR, комплементарная участку AUB (anti-upstream box) в 16S рРНК; DB – downstream box, часть MTIR, комплементарная участку ADB (anti-downstream box) в 16S рРНК; TE – translational enhancer, усилитель трансляции; MCS – mini-cistron sequence, мини-цистрон; н. – нуклеотидное звено.

\* Автор для переписки.



**Рис. 1.** Возможное строение инициаторного комплекса с участием мРНК гена 10 фага T7 [25] и 16S рРНК. Строение AUB в петле 437 - 497 – согласно работе [23]; строение ADB в петле 1410 - 1540 – согласно работе [24]. Вторичная структура 16S рРНК – согласно работе [26]. Точками отмечены комплементарно взаимодействующие пары нуклеотидов мРНК и 16S рРНК. Подчеркнут кодон AUG.

областью, образующей мини-цистрон (mini-cistron sequence, MCS). MCS сводит к минимуму возможность образования локальной вторичной структуры и тем самым облегчает связывание рибосомы [13, 14].

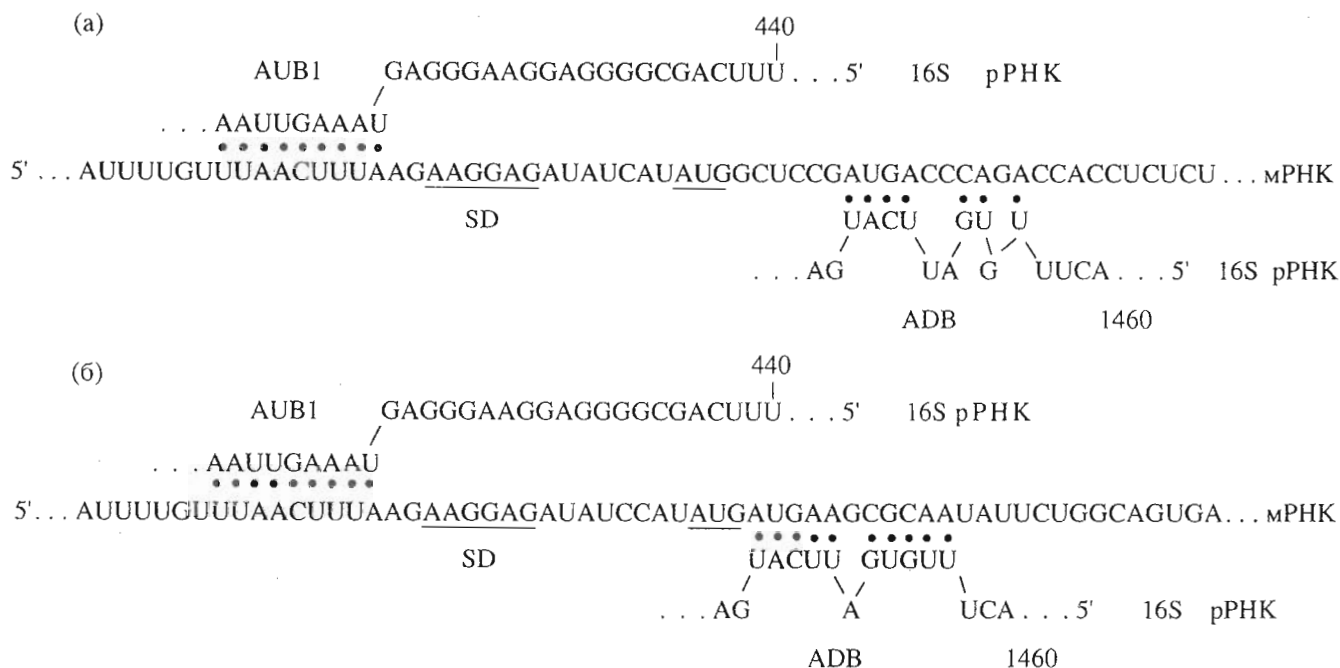
– Существуют другие, предшествующие SD, элементы мРНК, играющие роль специфических усилителей трансляции (translational enhancer, TE) [15 - 21], которые, возможно, препятствуют образованию локальной вторичной структуры [22] или (например, в случае лидерной мРНК гена 10 фага T7) способны комплементарно взаимодействовать с участком 16S рРНК (нуклеотиды 468 - 476, anti-upstream box, AUB) [23].

– Наконец, было найдено, что TIR эффективно транслируемой мРНК гена 0.3 фага T7 в ветви MTIR (нуклеотиды +15 - +26) комплементарна нуклеотидам 1471 - 1482 (anti-downstream box, ADB) 16S рРНК и что подобные, комплементарные тому же району 16S рРНК, участки имеются в других эффективных TIR [24].

Таким образом, эффективность инициации трансляции определяется структурой TIR, позволяющей в результате взаимодействия с 30S-субъединицей правильно ориентировать мРНК, чтобы иницирующий кодон был сближен с антикодонным fMet-тРНК в инициаторном комплексе. Возможное расположение TIR эффективно транслируемой мРНК гена 10 фага T7 относительно 16S рРНК в 30S субъединице рибосом в инициаторном комплексе приведено на рис. 1.

Следует, однако, отметить, что большинство приведенных выше наблюдений было сделано при изучении отдельных или немногочисленных структурных типов mTIR или MTIR, поэтому выводы о консенсусных структурах UB и DB или других TE нуждаются в уточнении, а вся концепция – в дополнительных доказательствах.

С целью расширить представления о взаимодействиях мРНК в инициаторном комплексе и их связи с уровнем экспрессии генов мы предприняли настоящее исследование. Учитывая, что экспери-



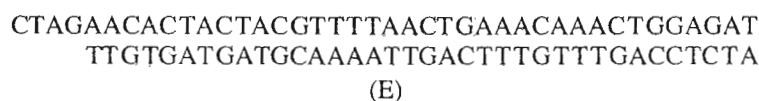
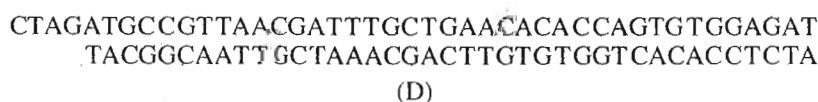
**Рис. 2.** Структуры TIR в мРНК генов *il3* и *lompegf* в плаزمидах рTE2il3 (а) и рTE2lompegf (б). Подчеркнуты кодоны AUG и последовательности SD, показана возможность образования комплементарных дуплексов с 16S рPHK в участках AUB1 и ADB.

ментально определяемый уровень экспрессии генов в *E. coli* неизбежно связан с деградацией белкового продукта, и для того, чтобы избежать в какой-то мере такую деградацию, в качестве объектов мы выбрали два гена – искусственный ген человеческого интерлейкина-3 (hIL3), образующего при экспрессии в *E. coli* тельца включения [15], и искусственный ген человеческого эпидермального фактора роста (hEGF) с лидерным пептидом OmpF – белка, секретируемого в периплазму *E. coli* [27]. Эти гены представляют два типа MTIR в соответствующих мРНК. На рис. 2 приведены структуры MTIR этих генов.

В качестве объектов сравнительного изучения различных mTIR мы решили использовать струк-

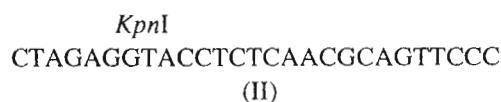
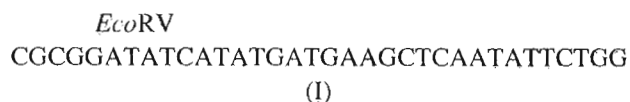
туры трех ТЕ: ТЕ гена *l0* фага T7 (TEX) [15, 20], ТЕ гена *atpE* *E. coli* (TEE) [18, 19, 29] и ТЕ гена *D* фага λ (TED) [21] (см. рис. 3).

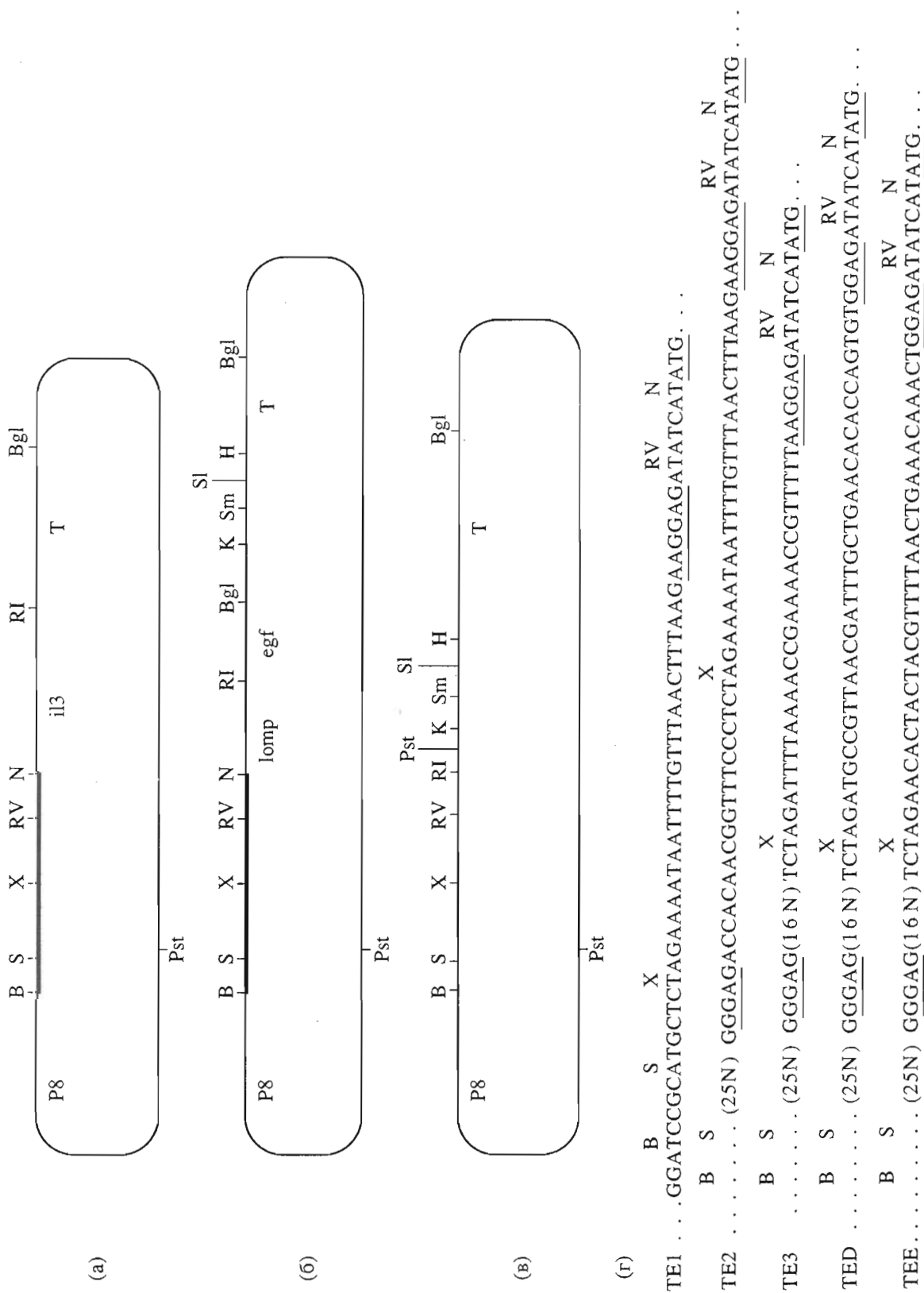
Ранее мы уже изучили экспрессию гена *il3* в трех конструкциях с TEX (рTE1il3, рTE2il3 и рTE3il3) в сравнении с плазмидой рFPCP10il3 (р8il3), лишенной структуры TEX [15], лишь качественно отметив, что структура с TE1 увеличивает экспрессию гена *il3* значительно слабее, чем структуры с TE2 и TE3, которые содержат второй сайт SD. Поэтому далее в конструкциях для сравнительного изучения мы варьировали участок ТЕ в плазмиде рTE2il3 между сайтами *Xba*I и *Eco*RV. В результате такой замены участка TE2 синтетическими дуплексами (D) и (E)



из плазмиды рTE2il3 были получены соответственно плазмиды рTEDil3 и рTEEil3. В качестве вектора для конструирования плазмид с геном *lompegf* мы использовали плазмиду рTE4 (рис. 3в), отличающуюся от ранее описанной

плазмиды рTE2 [15] наличием полилинкера (рис. 3). Ген *lompegf* мы выделили из плазмиды рVThEGF [28] путем амплификации с праймерами, содержащими сайты рестриктаз *Eco*RV (I) и *Kpn*I (II):





**Рис. 3.** Структуры плазмид, экспрессирующих гены белков hIL3 (i13) (а) и OmpF-hEGF (*lomp-egf*) (б), и векторной плазмиды rPE4 (в). Указаны сайты рестриктаз (B – *Bam*HI, Bgl – *Bgl*II, RI – *Eco*RI, RV – *Eco*RV, H – *Hind*III, K – *Kpn*I, N – *Nde*I, Pst – *Pst*I, S – *Sph*I, Sm – *Sma*I, X – *Xba*I), участки промотора P8 и терминатора транскрипции T; участок TE отмечен жирной чертой. (г) – нуклеотидные последовательности различных участков TE в экспрессионных плаزمидях, подчеркнуты последовательности SD и иницирующий кодон.

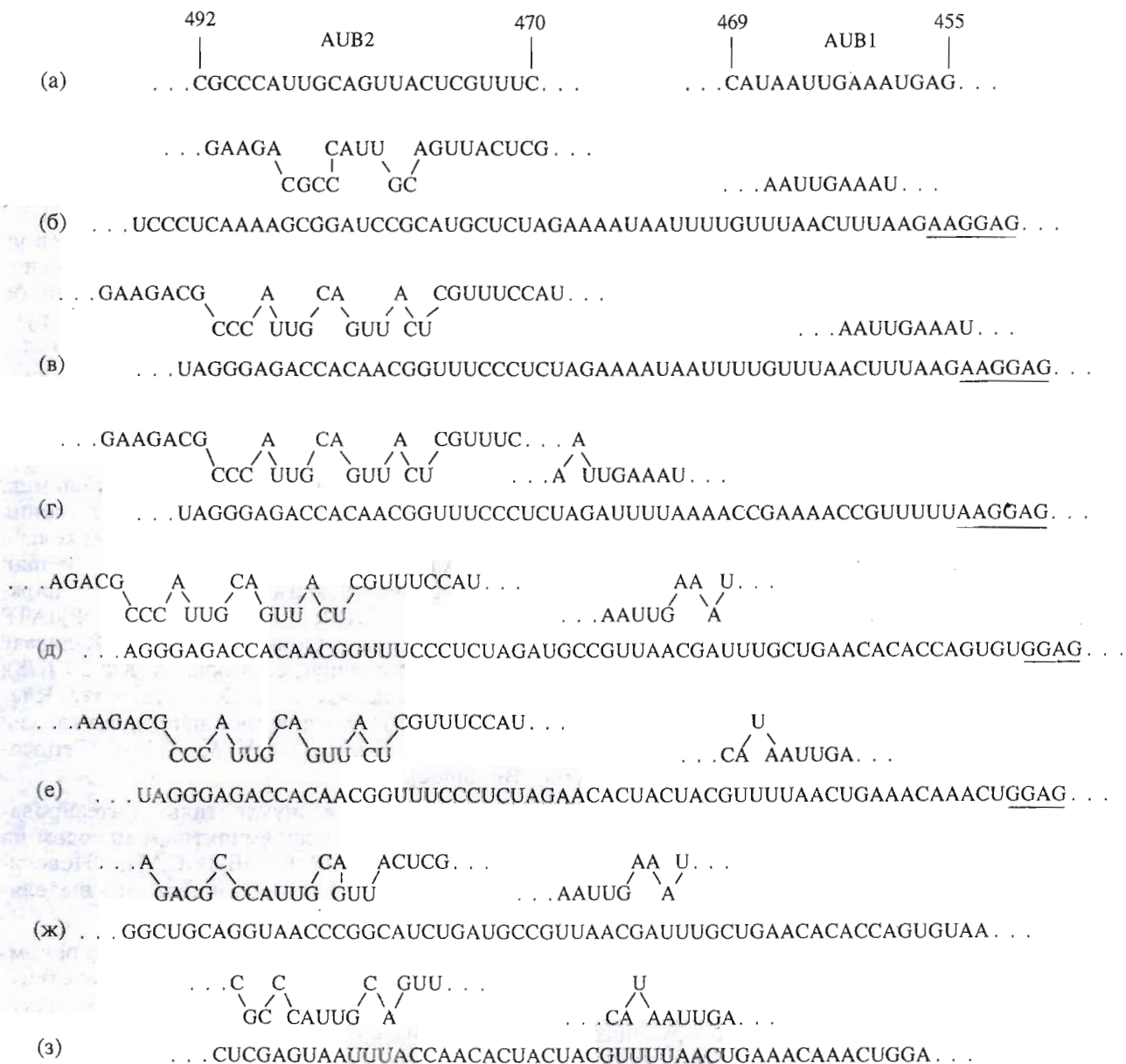


Рис. 4. Структуры mTIR мРНК генов и взаимодействие их с участками AUB1 и AUB2 16S рРНК (указаны номера нуклеотидов в последовательности 16S рРНК) (а). Подчеркнуты последовательности SD. Приведены данные для мРНК плазмид рТЕ1il3 (б), рТЕ2il3 (в), рТЕ3il3 (г), рТЕDi3 и рТЕDlompgef (д); е – рТЕEil3 и рТЕElompgef (е); мРНК гена D фага λ (ж); мРНК гена *arpE* E. coli (з).

После гидролиза этими рестриктазами ген *lompegf* клонировали по соответствующим сайтам вектора рТЕ4 и получили плазмиду рТЕ2lompgef. Рекомбинацией *PstI/EcoRV*-фрагмента этой плазмиды с *PstI/EcoRV*-фрагментами плазмид рТЕDi3 и рТЕEil3 были получены конструкции рТЕDlompgef и рТЕElompgef.

Сконструированными плазмидами трансформировались штаммы E. coli HB101 (*recA*). SG20050 (*lon*), TG1 (*recA*+/*F'*, *lacI*<sup>q</sup>). При этом учитывалось различное соотношение скоростей про-

цессов биосинтеза и деградации белка в различных штаммах [29]. Клетки продуцентов анализировали после культивирования в стандартных условиях путем гель-электрофореза тотальных лизатов с последующим сканированием гелей, прокрашенных кумасси R-250.

Из приведенных в таблице данных следует, что максимальный уровень экспрессии генов *il3* и *lompegf* достигается в плаزمиде, участки TIR которых характеризуются не только последовательностями вблизи иницирующего кодона,

Биосинтез белков на мРНК с различными структурами mTIR

Экспрессионная плаزمида	Ген	Содержание (%) белкового продукта в тотальном лизате штамма-продуцента <i>E. coli</i>		
		HB101	SG20050	TG1
p8i13	<i>il3</i>	<1	<1	<1
pTE1i13	»	2.8	3.4	3.4
pTE2i13	»	24.7	27.9	12.0
pTE3i13	»	7.0	8.0	4.6
pTEDi13	»	5.7	7.9	3.8
pTEEi13	»	15.9	16.2	3.8
pTE2lompegf	<i>lompegf</i>	3.7	2.4	2.8
pTEDlompegf	»	4.2	3.1	2.4
pTEElompegf	»	4.4	2.9	2.4

включающими UB, SD и DB (позиции -30 - +30), но и более отдаленным участком (-30 - -60), который влияет на уровень экспрессии. Мы предполагаем, что последний, возможно, взаимодействует с другим участком той же области 16S рРНК (позиции 437 - 497, см. рис. 1), с которой взаимодействует UB (UB1). Сопоставление соответствующих структур mTIR изученных нами мРНК позволяет идентифицировать предполагаемый участок UB2 (см. рис. 4).

Участок UB2 можно обнаружить в ряде природных генов с высоким уровнем экспрессии, например ген *D* фага  $\lambda$  [30] (рис. 4ж) и ген *atpE* *E. coli* [19] (рис. 4з).

Приведенные выше экспериментальные данные (см. таблицу) хорошо укладываются в общую закономерность. Так, наименьший уровень экспрессии гена *il3* наблюдается в случае плазмиды p8i13, в которой нет участков, кодирующих последовательность UB1 и UB2 (см. выше), а также плазмиды pTE1i13, в которой при наличии хорошего связывания рибосомы с SD и UB1 эффективность взаимодействия UB2 с рРНК мала. Наивысший уровень экспрессии для гена *il3* достигается с плазмидой pTE2i13, в которой эффективность взаимодействия SD, UB1 и UB2 с рРНК оптимальна. Немного ниже уровень экспрессии для гена *il3* с плазмидами pTE3i13 и pTEEi13, в которых связывание участков SD и UB1 с рибосомой слабее, расстояние между структурами UB1 и UB2 уменьшено, но сохраняется тот же, что и в pTE2i13, участок UB2; еще ниже уровень экспрессии этого гена в плазмиде pTEDi13, где существенно слабее связывание участка UB1 с рибосомой. Однако и в этих случаях уровень экспрессии выше, чем с плазмидой pTE1i13. Что же касается плазмид с геном *lompegf*, то с конструкцией, предполагающей конститутивный синтез белка в отсутствие эле-

ментов UB1 и UB2, экспрессия гена была ничтожно мала; в трех других приведенных в таблице конструкциях (pTE2lompegf, pTEDlompegf и pTEElompegf) практически не наблюдалось различий в уровнях экспрессии. Все эти конструкции содержат одинаковый участок UB2 и различаются участком UB1.

Таким образом, возможность дополнительного связывания TIR с 16S рРНК в участке AUB2 приводит к увеличению уровня экспрессии генов, что можно объяснить увеличением эффективности инициации трансляции. Очевидно, что такое связывание зависит также от вторичной структуры TIR, рассмотрению которой посвящено следующее сообщение.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали трис, акриламид, N,N'-метиленбисакриламид, персульфат аммония (Merck); N,N,N',N'-тетраметилэтилендиамин (Reanal); агарозу, АТФ, dNTP, бромистый этидий (Sigma); мочевины, ос. ч. ("Реахим"); агар, триптон, дрожжевой экстракт (Difco); [ $\gamma$ - $^{32}$ P]АТФ, [ $\alpha$ - $^{32}$ P]dАТФ (2000 Ки/ммоль, Обнинск); Т4-ДНК-лигазу (КФ 6.5.1.1), Т4-полинуклеотидкиназу (КФ 2.7.1.78); Taq-ДНК-полимеразу и ДНК-полимеразу Кленова (КФ 2.7.7.7); рестриктазы эндонуклеазы (КФ 32.1.23.x) *EcoRI*, *EcoRV*, *KpnI*, *XbaI* (Fermentas, Вильнюс).

Защищенные олигонуклеотиды синтезировали традиционным фосфоамидитным методом на синтезаторе ASM-102U (БИОСАН, Новосибирск) и после полного деблокирования выделяли, как описано в работе [31].

Условия экспериментов по получению рекомбинантных плазмид, трансформации компетентных клеток, клонированию, а также гель-электрофорезу и выделению ДНК из гелей приведены в работе [32].

Гидролиз рестриктазами проводили в универсальном буфере KGB [33].

Для анализа белков 1 мл бактериальной культуры центрифугировали, растворяли в лизирующем буфере [34], нагревали 5 мин при 100°C и белки разделяли в 15% ПААГ в присутствии додецилсульфата натрия [34]. Гели прокрашивали кумасси R-250 и сканировали на хроматосканере CS-930 (Shimadzu, Япония).

## СПИСОК ЛИТЕРАТУРЫ

1. McCarthy E.G., Guallerzi C. // Trends Genetics. 1990. V. 6. № 3. P. 78 - 85.
2. Steitz J.A. // Biological Regulation and Development. V. 1: Gene Expression / Ed. R.F. Goldberger. N.Y.: Plenum Press, 1979. P. 349 - 399.

3. Shine J., Dalgarno L. // Nature. 1975. V. 254. № 5495. P. 34 - 38.
4. Walz A., Pirotta V., Ineichen K. // Nature. 1976. V. 262. № 5570. P. 665 - 669.
5. Gallie D.R., Kado C.I. // Proc. Nat. Acad. Sci. USA. 1989. V. 86. № 1. P. 129 - 132.
6. Loechel S., Inamine J.M., Hu P.-Ch. // Nucl. Acids Res. 1991. V. 19. № 24. P. 6905 - 6911.
7. Curry K.A., Tomich C.-S.C. // DNA. 1988. V.7. № 3. P. 173 - 179.
8. De Smith M.H., van Duin J. // Proc. Nat. Acad. Sci. USA. 1990. V. 87. № 19. P. 7668 - 7672.
9. De Smith M.H., van Duin J. // Progr. Nucl. Acids Res. Mol. Biol. 1990. V. 38. № 1. P. 1 - 35.
10. Bucheler U.S., Werner D., Schirmer R.H. // Nucl. Acids Res. 1992. V. 20. № 12. P. 3127 - 3133.
11. Fatscher H.P., Geisen R.M., Fuchs E. // Eur. J. Biochem. 1988. V. 175. № 3. P. 461 - 465.
12. Olsen M.K., Rockenbach S.K., Curry K.A., Tomich C.-S.C. // J. Biotechnol. 1989. V. 9. № 2. P. 179 - 190.
13. Cheynet V., Verrier B., Mallet F. // Prot. Express. Purif. 1993. V. 4. № 5. P. 367 - 372.
14. Schoner B.E., Belagale R.M., Schoner R.G. // Proc. Nat. Acad. Sci. USA. 1986. V. 83. № 22. P. 8506 - 8510.
15. Гуревич А.И., Скапцова Н.В., Луценко С.В., Смирнов В.А., Куркин А.Н., Ажаев А.В. // Биоорган. химия. 1991. Т. 17. № 5. С. 647 - 652.
16. Царева Н.В., Музыченко М.Л., Бони И.В. // Биоорган. химия. 1993. Т. 19. № 10. С. 968 - 976.
17. Mahajna J., Oppenheim A.B., Rattray A., Gottesman M. // J. Bacteriol. 1986. V. 165. № 1. P. 167 - 174.
18. McCarthy J.E.D., Schairer H.U., Sebald W. // EMBO J. 1985. V. 4. № 2. P. 519 - 526.
19. McCarthy J.E.D., Sebald W., Lammers R. // Gene. 1986. V. 41. № 2. P. 201 - 216.
20. Olins P.O., Devine C.S., Rangwala S.H., Kavka K.S. // Gene. 1988. V. 73. № 1. P. 227 - 235.
21. Suisse M., Altuvia S., Koby S., Giladi H., Oppenheim A.B. // Mol. Gen. Genet. 1988. V. 214. № 3. P. 570 - 573.
22. Schauder B., McCarthy J.E.D. // Gene. 1989. V. 78. № 1. P. 59 - 72.
23. Olins P.O., Rangwala S.H. // J. Biol. Chem. 1989. V. 264. № 29. P. 16973 - 16976.
24. Sprengart M.L., Fatscher H.P., Fuchs E. // Nucl. Acids Res. 1990. V. 18. № 7. P. 1719 - 1723.
25. Dunn J.J., Studier F.W. // J. Mol. Biol. 1983. V. 166. № 4. P. 477 - 535.
26. Stern S., Powers T., Changchien L.-M., Noller H.F. // Science. 1989. V. 244. № 4906. P. 783 - 790.
27. Schauder B., Blocker H., Frank R., McCarthy J.E.G. // Gene. 1987. V. 52. № 2/3. P. 279 - 283.
28. Батчикова Н.В., Альтман И.Б., Луценко С.В., Смирнов В.А., Назимов И.В., Эшкинд Л.Г., Синягина Е.А., Ажаев А.В. // Биоорган. химия. 1992. Т. 18. № 6. С. 766 - 776.
29. Луценко С.В., Гуревич А.И., Птицын Л.Р., Рязанова Л.А., Смирнов В.А. // Биоорган. химия. 1992. Т. 18. № 3. С. 391 - 397.
30. Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B. // J. Mol. Biol. 1982. V. 162. № 4. P. 729 - 773.
31. Гуревич А.И., Качалина Т.А., Каюшин А.Л., Коростелева М.Д., Мирошников А.И. // Биоорган. химия. 1993. Т. 19. № 6. С. 629 - 632.
32. Гуревич А.И., Некрасова О.В. // Биоорган. химия. 1988. Т. 14. № 2. С. 149 - 152.
33. McClelland M., Hanish J., Nelson M., Patel Y. // Nucl. Acids Res. 1988. V. 16. № 1. P. 364.
34. Laemmli U.K. // Nature. 1970. V. 227. № 5229. P. 680 - 685.

## Dependence of the Gene Expression Level in *E. coli* from the Translation Initiation Region (TIR) Structure

### I. The Primary Structure of TIR

A. I. Gurevich, R. S. Esipov, T. A. Kachalina, A. L. Kayushin, and M. D. Korosteleva

*Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,  
117871 GSP-7 Moscow V-437, Ul. Mikurho-Maklaya, 16/10, Russia.*

**Abstract** – The comparison of expression levels of two genes – interleukin-3 (*il3*) and epidermal growth factor connected to lider peptide of *OmpF* (*lompegf*) – was carried out using specially constructed plasmids contained various structures of translational enhancers. It was shown that besides already known binding sites from mRNA translation initiation region (TIR) to 16S rRNA (SD, UB1 and DB), there is additional binding site of TIR disposed from –30 to –60 nt upstream of start codon AUG (UB2) having significant increasing effect on translation initiation and correspondingly on expression level.

*Key words:* nucleic acids, mRNA, rRNA, initiatory complex, translation initiation region, gene expression.