



УДК 577.217.343'112:575.117.2

**КЛОНИРОВАНИЕ И ХАРАКТЕРИСТИКА ГЕНА
РИБОСОМНОГО БЕЛКА S21 ЧЕЛОВЕКА**© 2000 г. **Е. В. Смирнова[#], Т. В. Ракитина, А. Ю. Евтодиенко, И. А. Костанян, В. М. Липкин***Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/110*

Поступило в редакцию 22.12.99 г. Принято к печати 19.01.2000 г.

Клонирован и охарактеризован полноразмерный функциональный ген, кодирующий рибосомный белок S21 человека. Определена его нуклеотидная последовательность, экзон-интронная организация и точка инициации транскрипции. Ген состоит из 6 экзонов и 5 интронов, имеет размер 1417 п. о. Подобно большинству известных генов рибосомных белков млекопитающих он не содержит в промоторной области канонические TATA- и CAAT-последовательности, а потенциальные участки связывания транскрипционных факторов расположены в нем как до, так и после точки инициации транскрипции. Первый интрон гена *rpS21* находится в 5'-нетранслируемой области. Точка инициации транскрипции определена на расстоянии 53 нт от АТГ-кодона, инициаторный цитидин окружен полипиримидиновым трактом. В 5'-фланкирующей области гена обнаружены два повтора, принадлежащие к *Alu-S-* и *Alu-J-* семействам.

Ключевые слова: рибосомный белок S21; экзон-интронная организация гена.

Эукариотическая рибосома представляет собой комплексную структуру, состоящую из 4 видов РНК и более 80 различных белков [1]. Известно, что каждому рибосомному белку в геноме млекопитающих соответствует целое семейство кодирующих последовательностей [2]. Установлено, что только один из представителей такого семейства, имеющий экзон-интронную структуру, является транскрипционно активным, функциональным геном [3]. Единственным известным исключением служит ген рибосомного белка S4 человека, для которого показано наличие двух высокомолекулярных, интронсодержащих и транскрипционно активных генов, локализованных на X- и Y-хромосомах [4]. Остальные представители семейств генов рибосомных белков являются безынтронными ретропозонподобными псевдогенами [5].

Проводимые в нашей лаборатории комплексные структурно-функциональные исследования эндогенного фактора дифференцировки (HLDF) из человеческой линии промиелоцитарного лейкоза HL-60, индуцированной ретиноевой кислотой, позволили предположить, что мРНК HLDF транскрибируется с гена, принадлежащего к семейству генов рибосомного белка S21 [6, 7].

Рибосомный белок S21 – один из наименее изученных компонентов рибосомы. Известно, что он локализован на поверхности бокового выступа

малой субъединицы [8], предположительно участвует в инициации трансляции и является партнером предшественника ламининсвязывающего белка LBP-40 [9]. Хотя были определены нуклеотидные последовательности кДНК, кодирующих рибосомные белки S21 человека, крысы, дрозофилы и дрожжей, только для дрозофилы был клонирован и структурно охарактеризован ген этого белка [10–12].

Для выяснения механизма синтеза HLDF в клетках HL-60, индуцированных ретиноевой кислотой, было решено изучить семейство генов рибосомного белка S21 и, в первую очередь, найти и охарактеризовать его функциональный ген.

Первоначально для поиска гена рибосомного белка S21 человека мы применили метод ПЦР. Матрицей для реакций служила геномная ДНК человека из клеток промиелоцитарного лейкоза линии HL-60, индуцированных ретиноевой кислотой. В качестве праймеров были выбраны пары олигонуклеотидов P1–P2 и P3–P4 (рис. 1), синтезированные на основании известной структуры кДНК рибосомного белка S21 человека [10]. В обоих случаях анализ продуктов ПЦР в агарозном геле показал наличие двух полос. Мажорные низкомолекулярные полосы соответствовали фрагментам псевдогенов S21. Минорные высокомолекулярные полосы представляли собой перекрывающиеся фрагменты внутренних областей функционального гена и в совокупности содержали 5 экзонов и 4 интрона.

[#] Автор для переписки (тел.: (095) 330-63-47; e-mail: smirnova@ibch.siobc.ras.ru).

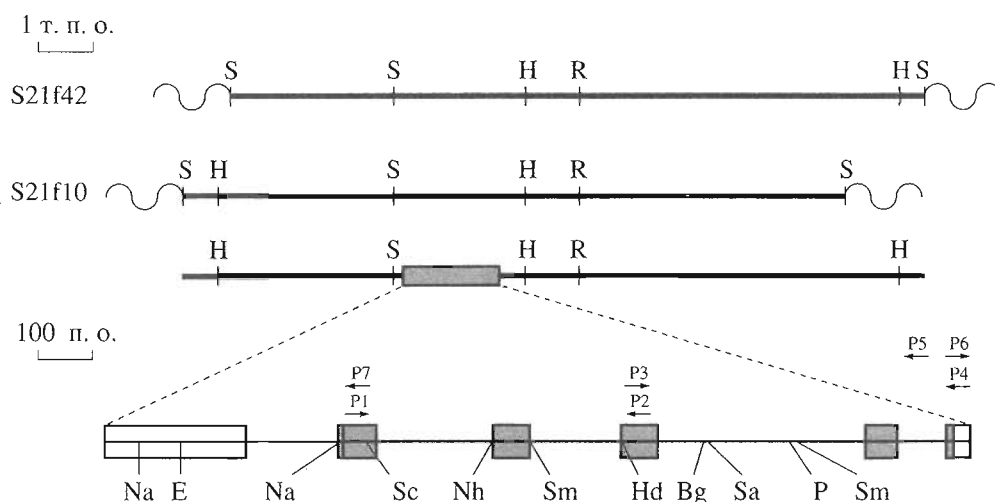


Рис. 1. Физическая карта и геномная организация гена рибосомного белка S21 человека. Серые прямоугольники соответствуют кодирующей области гена; белые – 5'- и 3'-нетранслируемым областям; стрелками обозначены олигонуклеотиды, использованные для ПЦР, обратной ПЦР и наращивания праймера. Эндонуклеазы рестрикции обозначены: Bg – *BgIII*; E – *EcoRV*; H – *HindIII*; Hd – *HindII*; Na – *NaeI*; Nh – *NheI*; S – *Sall*; Sa – *Sau3AI*; Sc – *SacII*; Sm – *SmaI*; P – *PstI*; R – *EcoRI*.

Для поиска 5'- и 3'-фланкирующих областей гена S21 применяли метод обратной ПЦР на матрице геномной ДНК, обработанной эндонуклеазой рестрикции *AluI* и лигированной в кольцо [13]. В качестве праймеров были использованы олигонуклеотиды P5–P6 (рис. 1), синтезированные на основе последовательностей последнего клонированного интрона и 3'-нетранслируемой области кДНК. В результате обратной ПЦР был получен фрагмент длиной 329 п. о., содержащий 3'-нетранслируемую и фланкирующую области гена *rpS21*. В случае 5'-нетранслируемой области гена такой подход оказался неэффективным, возможно из-за того, что ген S21, как и все гены рибосомных белков, является геном “домашнего хозяйства” и находится в GC-богатых островках генома [5], что затрудняет или делает невозможным прохождение ПЦР.

Для поиска 5'-нетранслируемой и промоторной областей гена S21 нами проведен скрининг лейкоцитарной человеческой геномной клонотеки, созданной на основе вектора λ FIX II, представительностью 3×10^6 клонов (Stratagene, USA). В результате гибридизации *in situ* с ранее полученными фрагментами гена *rpS21* найдены два перекрывающихся клон, содержащие полноразмерный ген рибосомного белка S21. Была установлена физическая карта найденных клонов (рис. 1) и определена полная нуклеотидная последовательность гена рибосомного белка S21 и его фланкирующих областей (рис. 3, см. ниже).

Следует отметить, что нуклеотидные последовательности фрагментов гена рибосомного белка S21, полученных методом ПЦР на матрице геномной ДНК из клеток HL-60, индуцированных ретиноевой кислотой, полностью соответствова-

ли нуклеотидной последовательности гена, найденного в лейкоцитарной клонотеке. Это исключает вероятность наличия каких-либо мутаций в гене рибосомного белка S21 из клеток промиелоцитарного лейкоза линии HL-60.

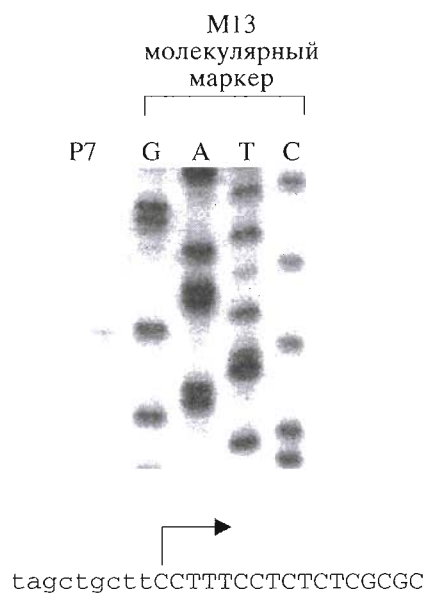


Рис. 2. Определение точки инициации транскрипции методом наращивания праймера. Дорожка P7 представляет собой уникальный 98-нт продукт реакции обратной транскрипции с праймера P7. Дорожки G, A, T, C – маркерный сиквенс. Под фотографией приведена нуклеотидная последовательность гена *rpS21*, окружающая старт инициации транскрипции, отмеченный стрелкой, полипиримидиновый тракт подчеркнут.

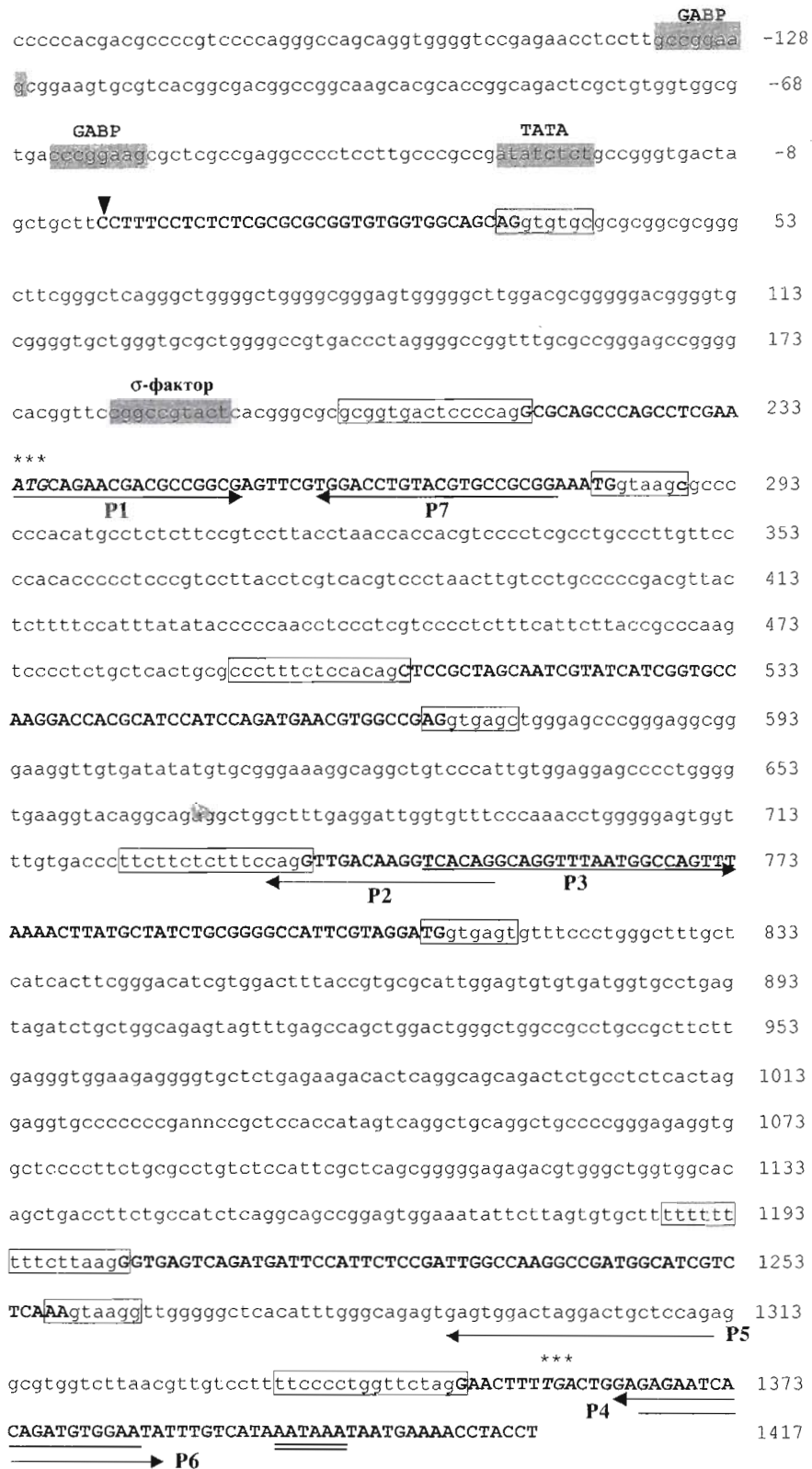


Рис. 3. Нуклеотидная последовательность и экзон-интронная организация гена рибосомного белка S21 человека. Нумерация дана от точки инициации транскрипции, которая обозначена черной стрелкой. Старт- и стоп-кодоны отмечены звездочками. Экзоны написаны полужирным шрифтом прописными буквами, интроны и 5'-фланкирующая область – строчными. Экзон-интронные границы, соответствующие консенсусам: **AG**tragt – 5'-донорный сайт сплайсинга и **u**ncag**G** – 3'-акцепторный сайт сплайсинга, обведены рамками. Стрелками показаны пары праймеров, используемые для ПЦР (P1–P2, P3–P4), обратной ПЦР (P5–P6), и праймер для определения точки инициации транскрипции (P7). Сигнал полиаденилирования подчеркнут двойной чертой. Участки связывания транскрипционных факторов выделены серым и подписаны.

Повторный скрининг данной клонотеки с использованием в качестве гибридизационной пробы фрагмента кДНК HLDF выявил 17 новых положительных клонов. ПЦР-анализ фаговой ДНК этих клонов показал, что все они соответствуют процессированным псевдогенам рибосомного белка S21. Таким образом, ПЦР на матрице геномной ДНК из промиелоцитарной клеточной линии HL-60, индуцированной ретиноевой кислотой, и скрининг лейкоцитарной человеческой геномной клонотеки позволил выявить и охарактеризовать единственный интронсодержащий ген из семейства генов рибосомного белка S21.

Было показано, что полноразмерный ген рибосомного белка S21 человека имеет длину 1417 п. о. и содержит 6 экзонов и 5 интронов. Аналогичный ген из дрозофилы имеет длину 841 п. о. и состоит из 4 экзонов и 3 интронов, причем, положения второго и последнего интронов в генах человека и дрозофилы совпадают, в результате, 3-й экзон гена дрозофилы соответствует совокупности 3-го, 4-го и 5-го экзонов гена человека. Все экзон-интронные границы подчиняются GT/AG-правилу [14]. Все интроны, за исключением первого, расположены внутри кодирующей области гена. Первый интрон расположен в 5'-нетранслируемой области. Старт-кодон находится во втором экзоне, хотя в большинстве известных рибосомных белков инициирующий ATG-кодон и следующие несколько аминокислот кодируются первым экзоном. Сигнал полиаденилирования AATAAA находится в положении 1396–1401 нт (нумерация дана от точки инициации транскрипции).

Точка инициации транскрипции гена рибосомного белка S21 была определена методом наращивания праймера. Этот метод представляет собой обратную транскрипцию на матрице тотальной РНК, где в качестве затравки в реакции синтеза используется праймер, комплементарный последовательности исследуемой мРНК, находящейся на расстоянии 100–200 п. о. от предполагаемой точки инициации транскрипции [15]. В результате реакции обратной транскрипции с праймера P7 (рис. 1) был получен уникальный фрагмент, размером 98 п. о., указывающий на то, что инициирующим остатком является цитидин (рис. 2). Данный цитидин локализован в полипиримидиновом тракте длиной 19 нт, прерываемом одним остатком гуанозина. Для ряда рибосомных белков ранее было показано, что инициаторный цитидин, находящийся в полипиримидиновом окружении, необходим для трансляционного контроля [16].

В промоторной области гена *rpS21* не было обнаружено типичных регуляторных последовательностей TATA и CAAT, хотя в положениях (–27)–(–20) нт был найден AT-богатый участок (рис. 3). Отсутствие TATA/CAAT-консенсусов

обычно требует участия в процессе транскрипции дополнительных ядерных факторов, участки связывания которых располагаются в генах рибосомных белков как до, так и после точки инициации транскрипции [17]. В 5'-фланкирующей области гена *rpS21* нами были обнаружены два потенциальных участка связывания транскрипционного фактора GABP [18], а в первом интроне гена – предполагаемый участок связывания σ -фактора [19] и многочисленные GC-боксы, являющиеся сайтами связывания транскрипционного фактора Sp1 [20] (рис. 3).

Компьютерный анализ с помощью программы Grail Ver.1.3 (<http://avalon.epm.ornl.gov/Grail-bin/EmptyGrailForm>) содержания CpG-нуклеотидов в гене *rpS21* показал, что экзоны 1–3 и интроны 1 и 2 находятся в GC-богатой области (содержание GC-нуклеотидов 68%).

В 5'-фланкирующей области гена в положениях (–2113)–(–1349) нт и (–1084)–(–989) нт были найдены два *Alu*-повтора, принадлежащие к *Alu-S* (89.6% гомологии) и *Alu-J* (81.9% гомологии) семействам [21].

Нуклеотидная последовательность была депонирована в базу данных EMBL/GeneBank (регистрационный номер AJ250907).

Работа была поддержана Международным научно-техническим центром (проект № 463) и Российским фондом фундаментальных исследований (проект № 97-04-49462).

Авторы благодарны сотрудникам ИБХ РАН Н.С. Быстрову за синтез олигонуклеотидов и К.Б. Игнатову за ценные советы.

СПИСОК ЛИТЕРАТУРЫ

1. Wool I.G. // Ann. Rev. Biochem. 1979. V. 48. P. 719–754.
2. Feo S., Davies B., Fried M. // Genomics. 1992. V. 13. P. 201–207.
3. Wagner M., Perry R.P. // Mol. Cell Biol. 1985. V. 5. P. 3560–3576.
4. Zinn A.R., Alagappan R.K., Brown L.G., Wool I., Page D.C. // Mol. Cell. Biol. 1994. V. 14. P. 2485–2492.
5. Dudov K.P., Perry R.P. // Cell. 1984. V. 37. P. 457–468.
6. Костянян И.А., Астапова М.И., Старовойтова Е.В., Драницына С.М., Липкин В.М. // Биоорган. химия. 1995. Т. 21. С. 243–248.
7. Драницына С.М., Костянян И.А., Андреева С.Г., Астапова М.В., Бабиченко И.И., Баева О.В., Богачук А.П., Молотковская И.М., Родионов И.Л., Смирнова Е.В., Липкин В.М. // Биоорган. химия. 2000. Т. 26. С. 340–351.
8. Lutsch G., Stahl J., Kargel H.-J., Noll F., Bielka H. // Eur. J. Cell Biol. 1990. V. 51. P. 140–150.
9. Sato M., Saeki Y., Tanaka K., Kaneda Y. // Biochem. Biophys. Res. Commun. 1999. V. 256. P. 385–390.
10. Bhat K.S., Morrison S.G. // Nucl. Acids Res. 1993. V. 21. P. 2939.

11. Itoh T., Otake E., Matsui K.A. // *Biochemistry*. 1985. V. 24. P. 7418–7423.
12. Torok I., Herrmann-Horle D., Kiss I., Tick G., Speer G., Schmit R., Mechler M.B. // *Mol. Cell. Biol.* 1999. V. 19. P. 2308–2312.
13. Ochman H., Ajioka J.W., Garza D., Hartl D.L. // *Biotechnology*. 1990. V. 8. P. 759–760.
14. Mount S.M. // *Nucl. Acids Res.* 1982. V. 10(2). P. 459–472.
15. Sambrook J., Fritsch E.F., Maniatis T. *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 1989. P. 7.79–7.81.
16. Levy S., Avni D., Hariharan N., Perry R., Meyuhas O. // *Proc. Natl. Acad. Sci. USA*. 1991. V. 88. P. 3319–3323.
17. Hariharan N., Kelley D.E., Perry R.P. // *Genes Dev.* 1989. V. 3. P. 1789–1800.
18. LaMarco K., Thompson C.C., Byers B.P., Walton E.M., McKnight S.L. // *Science*. 1991. V. 253. P. 789–792.
19. Park K., Atchison M.L. // *Proc. Natl. Acad. Sci. USA*. 1991. V. 88. P. 9804–9808.
20. Pugh B.F., Tjian R. // *Cell*. 1990. V. 61. P. 1187–1197.
21. Jurka J., Milosavljevic A. // *J. Mol. Evol.* 1991. V. 32. P. 105–121.

Cloning and Characterization of the Gene for the Human Ribosomal Protein S21

E. V. Smirnova[#], T. V. Rakitina, A. Yu. Evtodienko, I. A. Kostanyan, and V. M. Lipkin

*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, GSP-7 Moscow, 117871 Russia*

A full-size functional gene encoding the human ribosomal protein S21 was cloned and characterized. Its nucleotide sequence, exon–intron organization, and transcription initiation site were determined. The gene comprises 1417 bp and is composed of six exons and five introns. Like most known genes of mammalian ribosomal proteins, it lacks the canonical TATA- and CAAT sequences in the promoter region and harbors potential binding sites for transcription factors both upstream and downstream the transcription initiation site. The first intron of the *rpS21* gene is located in the 5'-untranslated region. The transcription initiation site is at a 53-bp distance from the ATG codon, and the initiation cytidine is surrounded by polypyrimidine tracts. The 5'-flanking region contains two repeats belonging to the *Alu-S* and *Alu-J* families.

Key words: gene exon-intron organization, ribosomal protein S21

[#] To whom correspondence should be addressed; phone: +7 (095) 330-6347; e-mail: smirnova@ibch.siobc.ras.ru.

For the full text of English translation of this article, see *Russian Journal of Bioorganic Chemistry*, 2000, vol. 26, no. 5. It is also available (free) in PDF format through the Web Site <http://www.maik.rssi.ru/>.