



УДК 547.963.32

КОНСТРУИРОВАНИЕ СЕМЕЙСТВА ИСКУССТВЕННЫХ ГЕНОВ
ИНСУЛИНА ЧЕЛОВЕКАЕфимов В. А., Бурякова А. А., Пашикова И. Н., Полушин Н. Н.,
Чахламцева О. Г.Институт биоорганической химии им. М. М. Шемякина
Академии наук СССР, Москва

Исходя из синтетического гена проинсулина человека получены гены А- и В-цепей инсулина, С-пептида и его тирозилированного аналога. Кроме того, сконструированы гены препроинсулина, мини-проинсулина, одноцепочечного инсулина человека, а также некоторые модификации этих генов. Использованный при этом набор методических приемов включал олигонуклеотиднаправленный мутагенез и химико-ферментативный синтез ДНК. Все полученные конструкции клонированы в *E. coli* в составе плазмидных векторов.

Инсулин — один из важнейших гормонов поджелудочной железы, который регулирует процессы углеводного обмена. В организме человека он синтезируется в виде предшественника — препроинсулина (110 а.о.), состоящего из сигнального пептида, собственно А- и В-цепей инсулина и С-пептида. Препроинсулин процессируется в клетке в проинсулин (86 а.о.), а затем этот неактивный прогормон в результате протеолитического отщепления С-пептида (35 а.о.) превращается в инсулин (51 а.о.) [1]. Ранее мы описывали синтез и клонирование гена проинсулина человека, последовательность которого была выведена исходя из первичной структуры соответствующего гена крысы заменой 14 его кодонов триплетами, соответствующими аминокислотам, входящим в состав проинсулина человека [2, 3]. Настоящее сообщение посвящено описанию конструирования и клонирования семейства генов, кодирующих инсулин человека, его биосинтетические предшественники и их аналоги.

Полученные нами гены можно разделить на несколько групп. К первой относятся полинуклеотиды, кодирующие отдельные цепи инсулина (А- и В-цепи), С-пептид и его тирозилированный аналог (рис. 1). Наличие остатка Tug в последнем необходимо для введения ¹²⁵I-метки в С-пептид, поскольку в природном человеческом С-пептиде остатки Tug отсутствуют. Вторую группу составляют гены проинсулина человека, отличающиеся от ранее полученного [2] наличием в непосредственной близости к 5'-концу последовательности сайта эндонуклеазы рестрикции *Hpa*I. Один из описываемых нами генов проинсулина отличается от ранее синтезированного также заменой Met-кодона на гексануклеотид, кодирующий дипептид Asn-Gly. Сайт эндонуклеазы *Hpa*I был введен с целью облегчения модификации 5'-концевого района гена проинсулина, а присоединение к N-концу прогормона дипептида Asn-Gly создает возможность его отщепления от лидерного пептида при экспрессии в виде гибридного белка действием гидроксилламина [4] (рис. 2).

В третью группу молекул входят гены, кодирующие природный человеческий препроинсулин и мутантный препроинсулин, представляющий собой гибрид проинсулина с лидерным пептидом основного белка оболочки бактериофага fd (белок гена VIII) [5].

Последняя группа, состоящая из трех генов, кодирует мини-проинсулин, у которого природный С-пептид заменен дипептидом Lys-Arg [6], а также два варианта так называемого одноцепочечного инсулина, у которого А- и В-цепи соединены в один интактный полипептид. Причем один из вариантов полученного нами гена одноцепочечного инсулина кодирует последовательность, лишенную остатка Thr-30 [7] (рис. 3).

¹ MetGlyIleValIGluGlnCysCysThrSerIleCysSerLeuTyrGlnLeuGluAsnTyrCysAsnTerTer
¹⁰ AATTCATGGGCATTGTGGAAACAGTGTGCACCAGCATCTGCTCCCTACCAACTGGAGAACTACTGCCAACTAGTA
²⁰ GTACCCGTAACACACCTTGTCCAGACGTGGTAGACGAGGGAGATGGTTGACCTCTTGATGACGTTGATCATTCGA

¹ MetPheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyrThrProLysThrTerTer
¹⁰ AATTCATGTTTGTCAATCAGCACCTTTGTGGTTCTCACCTGGTGGAGGCTCTGTACTGGTGTGGGAAAGTGGTTCTTCTACACACCCAAAGACCTAGTA
²⁰ GTACAAACAGTTAGTCGTGGAAACACCAAAGAGTGGACCCTCCAGACATGGACACACACCCCTTGCACCAAAGAAGATGTGGGTTCTGGATCATTCSGA

¹ MetArgArgGluAlaGluAspLeuGlnValGlyGlnValGluLeuGlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLysArgTer
¹⁰ AATTCATGCGTGTGAAGCGGAGGACCTGCAAGTGGGACAAGTGGAGCTGGGTGGAGGCCCGGGGCCCTTCAGCCCTTGGCACTGGAGGGTTCCCTGCGAAGACCGTTA
²⁰ GTACGGACGACATTCGCCCTCTGGACGTTTACCCTGTTACCTCGACCCACCTCCGGGGCCCTCAGAAAGTCCGGGAACCGTGCACCTCCCAAGGGACGCTTCGCAATTCCGT

¹ MetTyrArgArgGluAlaGluAspLeuGlnValGlyGlnValGluLeuGlyGlyGlyProGlyAlaGlySerLeuGlnProLeuAlaLeuGluGlySerLeuGlnLysArgTer
¹⁰ AATTCATGACCCTGTGAAGCGGAGGACCTGCAAGTGGGACAAGTGGAGCTGGGTGGAGGCCCGGGGCCCTTCAGCCCTTGGCACTGGAGGGTTCCCTGCGAAGACCGTTA
²⁰ GTASATGGCAGCATTCCGCTCTTGGACGTTTACCCTGTTACCTCGAACCCCTCCGGGGCCCGGGCCCTCAGAAAGTCCGGGAACCGTGCACCTCCCAAGGGACGCTTCGCAATTCCGT

Рис. 1. Нуклеотидные последовательности генов и соответствующие им аминокислотные последовательности А- (а), В-цепи инсулина человека (б), С-пептида (в) и тирозинированного С-пептида (г)

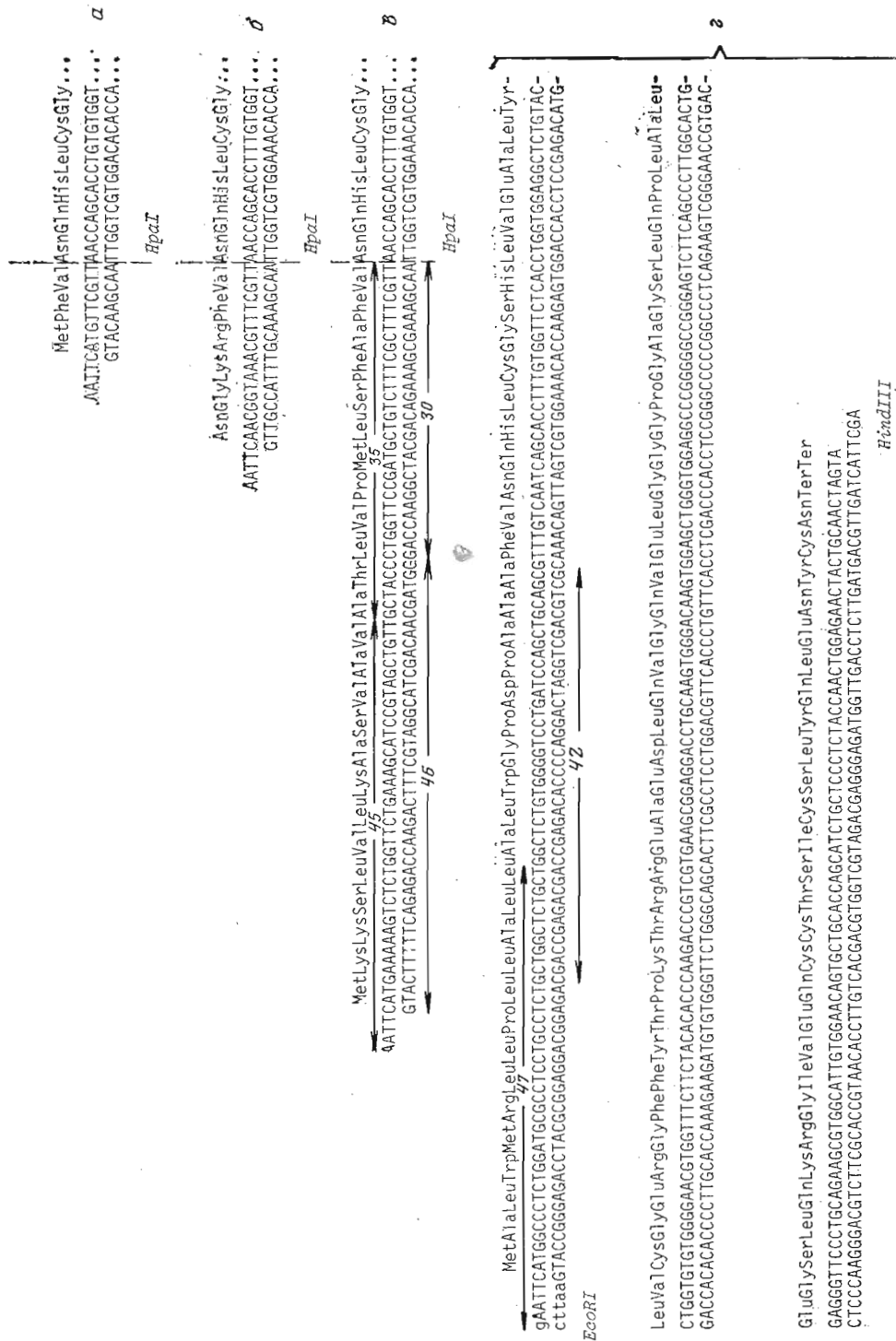


Рис. 2. Нуклеотидные последовательности 5'-концевых участков генов проницулина человека PL'RH (a), PL'RH (b), PL'RF (c) и полная нуклеотидная последовательность гена (PPL'R) (e) и соответствующие им аминокислотные последовательности. Стрелками показаны синтетические олигонуклеотиды, из которых конструировались последовательности, кодирующие лидереые пептиды

α

1 MetPheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyrThrProLysThrLysArgGlyIleValGluGlnCysCysThr- 30- 40
 AATCATGTTTGTCAATCAGCACCTTTGTGGTTCTACCTGGTGGAGGCTCTGTACCTGGTGTGGGAAACAGTGGTTCTTCTACACACCCCAAGCAAGCGTGGCATTTGGAAACAGTGTGCACCC-
 GTACAACACAGTTAGTCTGTGAAACACCAAGAGTGGACCACCTCCGAGACATGGACCACACACACCCCTTGCACCAAGAAGATGTGGTTCTGGTTCCGACCCGTAACACCTTGTCCACGACGTTGG-

50
 SerIleCysSerLeuTyrGlnLeuGluAsnTyrCysAsnTerTer
 AGCATCTGCTCCCTACCAACTGGAGAACTACTGCAACTAGTA
 TCGTAGACGGGAGATGGTTGACCTCTTGATGACCTTGATCATTCGA

β

1 MetPheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyrThrProLysGlyIleValGluGlnCysCysThrSerIleCys- 30- 40
 AATTCATGTTTGTCAATCAGCACCTTTGTGGTTCTACCTGGTGGAGGCTCTGTACCTGGTGTGGGAAACAGTGGTTCTTCTACACACCCCAAGGGCATTTGGAAACAGTGTGCACCCAGCATCTGC-
 GTACAACACAGTTAGTCTGTGAAACACCAAGAGTGGACCACCTCCGAGACATGGACCACACACCCCTTGCACCAAGAAGATGTGGTTCCCGTAACACCTTGTCCAGCGTGGTGGTGGTAGACG-

50
 SerLeuTyrGlnLeuGluAsnTyrCysAsnTerTer
 TCCTCTACCAACTGGAGAACTACTGCAACTAGTA
 AGGGAGATGGTTGACCTCTTGATGACCTTGATCATTCGA

δ

1 MetPheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyrThrProLysThrGlyIleValGluGlnCysCysThrSerIle- 30- 40
 AATTCATGTTTGTCAATCAGCACCTTTGTGGTTCTACCTGGTGGAGGCTCTGTACCTGGTGTGGGAAACAGTGGTTTCGGCCGACACACCCCAAGACCCGGCATTTGGAAACAGTGTGCACCCAGCATC-
 GTACAACACAGTTAGTCTGTGAAACACCAAGAGTGGACCACCTCCGAGACATGGACCACACACCCCTTGCACCAAGAAGCCGCTGTGGTTCCGACCAAGACCTTGTCCACGACGTTGGTGGTAG-

50
 CysSerLeuTyrGlnLeuGluAsnTyrCysAsnTerTer
 TGCTCCCTCTACCAACTGGAGAACTACTGCAACTAGTA
 ACGAGGGAGATGGTTGACCTCTTGATGACCTTGATCATTCGA

Рис. 3. Нуклеотидные по-
 следовательности генов
 и аминокислотные по-
 следовательности мини-
 проинсулина (α), одно-
 цепочечного дезтрео-
 нин-инсулина (β) и од-
 ноцепочечного инсули-
 на (δ)

Последовательности олигонуклеотидов, использованных для проведения направленного мутагенеза гена проинсулина человека

Олигонуклеотид	Пептид, кодируемый мутантным геном
СТАСАСАСССААГАССАААГСГТГГГСАТТГТГ СТТСТАСАСАСССААГГГГСАТТГТГГАААСАГТ	Мини-проинсулин Одноцепочечный дезтреонининсулин
СТАСАСАСССААГАССГГГСАТТГТГГАААСАГТ ААТТСАТГТТГСГТТААССАГСАССТГТГТГГТТС ТСТТСААГААТТСАТГГГГСАТТГТГГАААСАГТ СТАСАСАСССААГАССТАГТААГСТТТААТГГ ТСТТСААГААТТСАТГСГТГСГТГААГСГГАГГ ТТСССТГСАГААГСГТТААГСТТТААТГСГГТ ТСТТСААГААТТСАТГТАССГТГСГТГААГСГГАГГ	Одноцепочечный инсулин Проинсулин (<i>HpaI</i>) А-Цепь инсулина В-Цепь инсулина С-Пептид (N-конец) С-Пептид (C-конец) Тур-С-пептид (N-конец)

Гены А- и В-цепей инсулина, а также С-пептида получали исходя из ранее синтезированного гена проинсулина человека [2] методом олигонуклеотиднаправленного мутагенеза на плазмидных ДНК [2]. Плазмиду рРL'R, несущую исходный ген проинсулина РL'R, расщепляли эндонуклеазой рестрикции *EcoRV*. Полученную при этом линейную ДНК подвергали обработке экзонуклеазой III в контролируемых условиях, которые подбирались таким образом, чтобы образующиеся при действии экзонуклеазы одноцепочечные участки на 5'-концах ДНК включали в себя район предполагаемой мутации и имели протяженность около 500 нуклеотидов. Приготовленная таким образом плазмидная ДНК отжигалась с 5'-фосфорилированным синтетическим олигонуклеотидом, содержащим желаемую мутацию (праймером-мутагеном) (см. таблица). Затем проводилась достройка одноцепочечных участков ДНК с помощью ДНК-полимеразы *E. coli* (фрагмент Кленова) в присутствии четырех дезоксирибонуклеозидтрифосфатов и Т4-ДНК-лигазы. Последняя сшивает разрывы в цепи и обеспечивает замыкание полученного гетеродуплекса в циклическую молекулу. Гетеродуплексной ДНК трансформировали компетентные клетки *E. coli* HB101 и трансформанты высевали на агаризованную LB-среду, содержащую ампициллин. Отбор целевых клонов осуществлялся гибридизацией с соответствующим ³²P-меченым праймером-мутагеном. Из отобранных колоний выделялась плазмидная ДНК, и структура мутантного гена подтверждалась секвенированием (рис. 4).

В случае генов С-пептида и Тур-С-пептида мутагенез проводился одновременно двумя олигонуклеотидами, первый из которых предназначался для удаления участка ДНК, кодирующего В-цепь, а второй — А-цепь инсулина.

Тот же препарат плазмиды рРL'R, обработанной *EcoRV* и экзонуклеазой III, послужил исходным для получения гена мини-проинсулина и двух вариантов гена одноцепочечного инсулина. Аналогично получался вариант гена проинсулина человека, имеющий *HpaI*-сайт. Последовательности праймеров-мутагенов, использованных в этих случаях, приведены в таблице.

Следует отметить, что эффективность предложенного нами метода мутагенеза плазмидных ДНК в вышеописанных примерах составляла в среднем 50%, а в случае мутагенеза одновременно двумя праймерами (С-пептид) — 10–15%.

Для замены Met-кодона в гене проинсулина на гексануклеотид, кодирующий Asn-Gly, плазмиду рРL'RH обрабатывали эндонуклеазами рестрикции *HpaI* и *EcoRI* и полученный при этом большой фрагмент соединяли с помощью Т4-ДНК-лигазы с небольшим синтетическим дуплексом, состоящим из олигонуклеотидов ААТТСААССГГТАААССГТТТГСГТТ и ААССГАААССГТТТАССГТТГ. Полученные рекомбинантные ДНК после трансформации бактериальных клеток отбирали на уровне колоний гибридизацией с ³²P-меченым олигонуклеотидом ААССГАААССГТТТАССГТТГ. Выделенную из отобранных клонов плазмидную ДНК анализируют

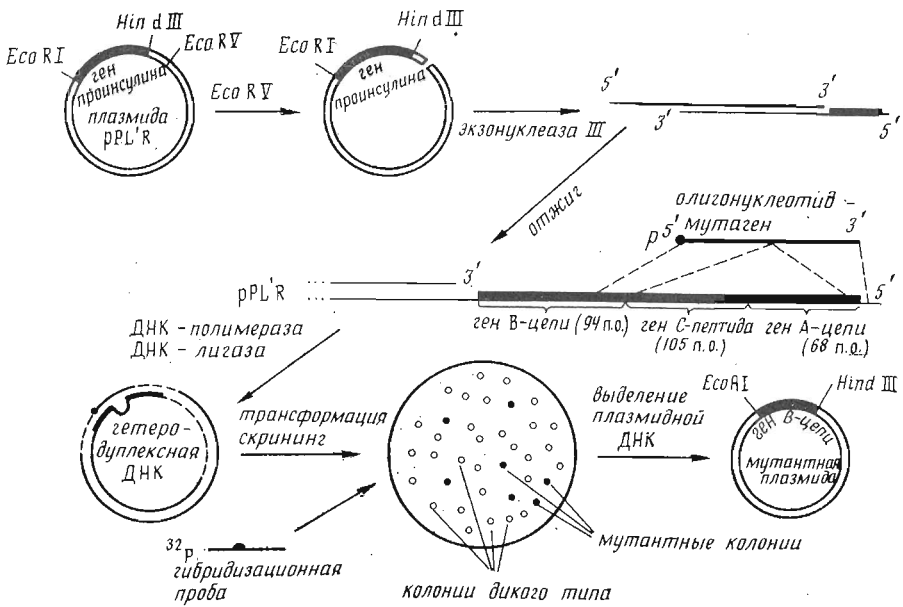


Рис. 4. Схема проведения олигонуклеотиднаправленного мутагенеза гена проинсулина человека на плазмидной ДНК

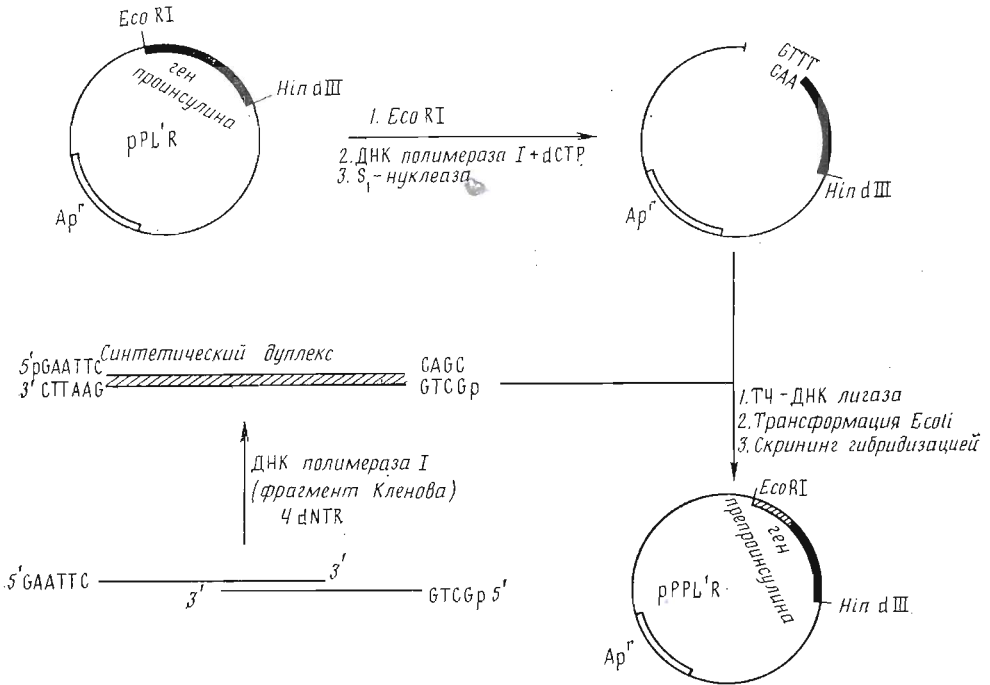


Рис. 5. Схема получения гена препроинсулина человека

вали обработкой эндонуклеазами *EcoRI*, *HpaI* и *HindIII*, а первичную структуру синтетического гена подтверждали секвенированием.

Заменой малого *EcoRI*-*HpaI*-фрагмента плазмиды *rPPL'RN* на 80-звенный синтетический дуплекс, кодирующий преучасток белка оболочки фага fd, была получена плаزمида *rPPL'RF*, в состав которой входит ген гибрида этого лидерного пептида с проинсулином человека. Синтетический дуплекс был получен шивкой четырех олигонуклеотидов с помощью T4-ДНК-лигазы (рис. 2).

Последовательность гена препроинсулина человека конструировалась

исходя также из плазмиды рРL'R введением в нее синтетического дуплекса длиной 77 п.о. Последний получался ДНК-полимеразной достройкой частичного дуплекса, образованного 3'-концевыми участками 42- и 47-звенного олигонуклеотидов. ДНК плазмиды рРL'R расщепляли *EcoRI*, а затем обрабатывали ДНК-полимеразой I (фрагмент Кленова) в присутствии только dCTP. Образовавшиеся при этом выступающие 5'-концы удаляли обработкой S₁-нуклеазой. Полученный таким образом вектор сшивали по тупым концам с синтетическим 77-звенным дуплексом, описанным выше (рис. 5). Полученной рекомбинантной ДНК трансформировали клетки *E. coli*; трансформанты высевали на среду, содержащую ампициллин, и выросшие колонии скринировали на присутствие желаемой конструкции гибридизацией с ³²P-меченым олигонуклеотидом CTGCAGCG·TTTGTCAA, структура которого соответствовала месту стыка гена проинсулина с фрагментом ДНК, кодирующим преучасток. Из отобранных колоний выделяли плазмидную ДНК, которую затем анализировали рестриктным анализом, а структуру гена препроинсулина человека подтверждали секвенированием с помощью метода химических модификаций.

Экспериментальная часть

В работе использованы эндонуклеазы рестрикции, ДНК-полимераза I *E. coli* (фрагмент Кленова), экзонуклеаза III *E. coli* (P-L Biochemicals, США), S₁-нуклеаза *Aspergillus oryzae* (Sigma), [γ -³²P]АТР (Amersham, Англия), ДНК-полимераза, ДНК-лигаза и полинуклеотидкиназа фага T4 (Pharmacia, Швеция).

Бактерии выращивали на LB-среде [9] с добавлением 25–30 мкг/мл ампициллина. Приготовление компетентных клеток *E. coli* и трансформацию бактериальных клеток проводили по описанной методике [10]. Плазмидную ДНК выделяли по методу [11].

Химический синтез олигодезоксирибонуклеотидов осуществляли на синтезаторе фирмы Pharmacia фосфоритриэфирным методом с использованием O-нуклеофильных катализаторов [12] или фосфитамидным методом [13].

Эндонуклеазами рестрикции ДНК расщепляли в буфере, содержащем 20 мМ трис-HCl (pH 7,5), 10 мМ NaCl и 1 мМ дитиотреит.

Для приготовления рекомбинантных плазмид линейную форму исходной плазмиды, полученную обработкой кольцевой формы (2–5 мкг) рестриктазой, смешивали с 50–100 пмоль соответствующего синтетического дуплекса в 100 мкл буфера, содержащего 0,05 М трис-HCl (pH 7,5), 10 мМ MgCl₂, 1 мМ дитиотреит, 500 мкМ АТР, и сшивали в присутствии 10–15 ед. акт. ДНК-лигазы в течение 5–10 ч при 10° С. Аликвоты реакционной смеси использовали для трансформации компетентных клеток *E. coli*.

Подбор условий расщепления одной из цепей ДНК экзонуклеазами проводили следующим образом. После расщепления 20 мкг ДНК *EcoRV* в 300 мкл буфера смесь разделяли на равные части. К одной из них прибавляли 40 ед. акт. экзонуклеазы III. Инкубацию проводили при 25° С. Через определенные промежутки времени отбирали аликвоты по 30 мкл, к ним прибавляли 1 М ацетат натрия (pH 4,5) до концентрации 50 мМ, NaCl — до 200 мМ, ZnCl₂ — до 2 мМ и 1 мкл (4 ед. акт.) S₁-нуклеазы *A. oryzae*. После 15 мин инкубации при 20° С реакцию останавливали прибавлением EDTA. Образцы наносили на 1,5% агарозный гель и при 60 В проводили электрофорез в течение 3 ч. В качестве свидетелей использовали гидролизаты ДНК фага λ рестриктазой *HindIII*, ДНК ϕ X174 — *HpaI* и ДНК pBR322 — *PstI*, *EcoRI* и *BamHI* одновременно.

Для проведения мутагенеза *in vitro* 1–2 мкг двухцепочечной ДНК с выступающими 5'-концевыми участками необходимой длины смешивали с 40–200 пмоль 5'-фосфорилированного олигонуклеотида-мутагена в 50 мкл буфера, содержащего 50 мМ трис-HCl (pH 7,5), 10 мМ MgCl₂, 20 мМ NaCl. Смесь нагревали до 50–55° С и через 5 мин медленно охлаждали до 10–12° С. Затем прибавляли дитиотреит до концентрации 1–2 мМ, четыре dNTP до концентрации 300 мкМ, АТР — до 100 мкМ, 30–40 ед. акт.

ДНК-полимеразы I (фрагмент Кленова) и 20 ед. акт. ДНК-лигазы. Через 5–6 ч аликвоты этой реакционной смеси объемом 1–5 мкл использовали непосредственно для трансформации компетентных клеток *E. coli*.

Скрининг колоний гибридизацией с ^{32}P -меченым олигонуклеотидом-мутагеном проводили на нитроцеллюлозных фильтрах. Температуру гибридизации выбирали на основании данных, полученных в работе [14]. Фильтры промывали 30–45 мин с использованием 2–3 смен промывочного буфера. Температура при этом была такой же, как и при гибридизации. Фильтры экспонировали с рентгеновской пленкой в течение 12–20 ч.

СПИСОК ЛИТЕРАТУРЫ

1. Oyer P. E., Peterson J. D., Steiner D. F. // *J. Biol. Chem.* 1971. V. 246. № 5. P. 1375–1386.
2. Овчинников Ю. А., Ефимов В. А., Чазмахчева О. Г. // Докл. АН СССР. 1983. Т. 270. № 3. С. 743–747.
3. Ovchinnikov Yu. A., Efimov V. A., Ivanova I. N., Reverdatto S. V., Skiba N. P., Chakhmakheva O. G. // *Gene*. 1984. V. 31. № 1. P. 65–78.
4. Allen G. // *Sequencing of proteins and peptides/Eds Work T. S., Burden R. H.* Amsterdam – New York – Oxford: North-Holland Publ. Comp., 1981. P. 49–50.
5. Schaller H., Beck E., Takanami M. // *The single-stranded DNA phages* N. Y.: Cold Spring Harbor Laboratory, 1978. P. 139–163.
6. Thim L., Hansen M. T., Norris K., Hoeck I., Boel E., Forstrom J., Ammerer G., Fiil N. P. // *Proc. Nat. Acad. Sci. USA*. 1986. V. 83. № 18. P. 6766–6770.
7. Markussen J., Hongaard P., Rubei U., Sorensen A. R., Sorensen E. // *Protein Engineering*. 1987. V. 1. № 3. P. 205–213.
8. Ефимов В. А., Мирских О. В., Чазмахчева О. Г., Овчинников Ю. А. // *Биоорганическая химия*. 1985. Т. 11. № 5. С. 621–627.
9. Миллер Д. Эксперименты в молекулярной генетике. М.: Мир, 1976. С. 395.
10. Маниагис Т., Фрич Э., Самбрук Дж. Молекулярное клонирование. М.: Мир, 1986 с.
11. Birnboim H. C., Doly J. // *Nucl. Acids Res.* 1979. V. 7. № 6. P. 1513–1523.
12. Efimov V. A., Buryakova A. A., Dubey I. Y., Polushin N. N., Chakhmakheva O. G., Ovchinnikov Yu. A. // *Nucl. Acids Res.* 1986. V. 19. № 16. P. 6525–6540.
13. McBride L. J., Kiczek A., Beausage S. L., Caruthers M. J. // *J. Amer. Chem. Soc.* 1986. V. 108. № 8. P. 2040–2048.
14. Ефимов В. А., Чазмахчева О. Г. // *Биоорганическая химия*. 1982. Т. 8. № 8. С. 2084–2093.

Поступила в редакцию
10.1.1989

CONSTRUCTION OF A FAMILY OF ARTIFICIAL GENES FOR HUMAN INSULIN

EFIMOV V. A., BURYAKOVA A. A., PASHKOVA I. N., POLUSHIN N. N.,
CHAKHMAKHEVA O. G.

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

Using oligonucleotide-directed mutagenesis and chemical-enzymatic DNA synthesis, genes for A and B insulin chains, C-peptide and Tyr-C-peptide have been constructed starting from synthetic gene for human proinsulin synthesized earlier. The genes for human preproinsulin, mini-proinsulin, single-chain insulin and their modifications were also synthesized. The constructions obtained were cloned in plasmid vectors.