



УДК 577.214.622

СИНТЕЗ И ЭКСПРЕССИЯ ИСКУССТВЕННОГО ГЕНА
IgG-СВЯЗЫВАЮЩЕГО ФРАГМЕНТА БЕЛКА А
STAPHYLOCOCCUS AUREUS

*Ефимов В. А., Бурякова А. А., Полушин Н. Н.,
Папкова И. И., Дмитракова Е. В., Чахмахчева О. Г.*

*Институт биоорганической химии им. М. М. Шемякина Академии наук СССР,
Москва*

Осуществлен химико-ферментативный синтез гена, который кодирует иммуноглобулинсвязывающий фрагмент белка А стафилококков, состоящий из сигнального пептида и модифицированных доменов *E* и *B* этого белка. На основе синтетического гена получены плазмиды, способные обеспечивать экспрессию этого фрагмента белка А в *E. coli*.

Современная белковая инженерия открывает широкие перспективы для введения заранее запланированных изменений в первичную структуру соединений белковой природы с помощью направленного мутагенеза их генов с целью получения белков и пептидов с новой или повышенной биологической активностью. Кроме того, существуют возможности для микробиологического синтеза разнообразных белков и полипептидов с ранее неизвестной, произвольной структурой [1]. Одним из классических примеров белковой инженерии является получение нестабильных в бактериях гетерологичных полипептидов в виде гибридов с бактериальными белками или их фрагментами. При этом стабильность целевого полипептида достигается либо за счет способности гибридного белка образовывать в цитоплазме клеток нерастворимые тела включения, либо за счет секреции гибрида в периплазматическое пространство клетки или культуральную среду, что обеспечивается сигнальными последовательностями лидеров. Последний способ, по-видимому, наиболее перспективен и обладает рядом преимуществ. Так, в результате эффективной секреции гибридного белка в культуральную среду должна достигаться его значительная очистка от остальных белков клетки, облегчающая дальнейшую процедуру его выделения. Поскольку культуральная среда *E. coli* практически не содержит протеиназ, в процессе ферментации секретируемого продукта обеспечивается его устойчивость. Секреция может также способствовать правильному замыканию дисульфидных мостиков целевых белков, чего обычно не происходит в цитоплазме бактерий [2].

С целью конструирования экспрессирующих векторов, способных обеспечивать секрецию гетерологичных пептидов и белков в межклеточное пространство *E. coli* в виде гибридов с фрагментом белка А стафилококков, нами осуществлены синтез, клонирование и экспрессия гена иммуноглобулинсвязывающего фрагмента этого белка.

Белок А является компонентом клеточной стенки большинства штаммов *Staphylococcus aureus* и обычно ковалентно связан с пептидогликаном. Однако существуют мутанты и клишические штаммы, секретирующие весь производимый ими белок А в межклеточное пространство. Характерная особенность белка А — способность селективно связываться с Fc-районом иммуноглобулина G (IgG), оставляя при этом свободными антигенсвязывающие сайты [3].

В функциональном отношении этот белок может быть разделен на два характерных района: IgG-связывающий и так называемый район X, который ответствен за ассоциацию белка А с пептидогликаном клеточной

стенки *S. aureus* [4]. Структурный ген белка А (*spa*) был клонирован, секвенирован и экспрессирован в гетерологичных системах [5]. Он состоит из трех функционально отличных частей: 5'-конец гена кодирует сигнальный пептид, который способствует транспорту белка через цитоплазматическую мембрану как в *S. aureus*, так и в *E. coli*; средняя часть гена кодирует пять практически идентичных доменов (*E, D, A, B, C*), которые ответственны за связывание с иммуноглобулинами; 3'-конец гена кодирует X-район белка. Следует отметить, что последовательность самого белка также сильно гомологична как в области IgG-связывающего района (состоит из пяти схожих доменов длиной около 58 а.о. каждый), так и в X-районе (состоит из 12 повторяющихся доменов длиной 8 а.о. каждый). Ранее было показано, что небольшой фрагмент гена белка А, кодирующий домены *E* (или его N-концевую часть), *B* и сигнальную последовательность, дает в *E. coli* правильно процессируемый и эффективно секретируемый продукт, способный аффинно связываться с IgG, иммобилизованным на твердом носителе [6]. Были также опробованы векторы, приспособленные для получения пептидов и белков в виде гибридов с белком А и его фрагментами [7], и с их помощью была достигнута экспрессия некоторых белков, в том числе инсулинподобного ростового фактора (IGF) [8].

Последовательность синтезированного нами искусственного гена, на основе которого были получены экспрессирующие векторы, кодирует фрагменты белка А. Она была выведена на основании известной первичной структуры этого белка и генетического кода и составлена из кодонов, наиболее часто используемых в генах *E. coli*. Основываясь на данных работы [6], вместо полной последовательности гена мы синтезировали только его часть, которая состояла из фрагментов, кодирующих сигнальный пептид, модифицированный домен *E* и домен *B*, что должно быть достаточно для эффективного связывания с IgG и процессинга. Модифицированный *E*-домен представлял собой гибрид фрагментов *E*- и *B*-доменов. Ранее было показано [6], что первые аминокислоты домена *E* природного белка А, являющиеся одновременно и первыми аминокислотами всего белка, отличаются от N-концевых аминокислот остальных IgG-связывающих доменов и они необходимы для правильного процессинга сигнального пептида. Поэтому в структуре модифицированного *E*-домена содержалось 12 первых а.о. нативного *E*-домена, а остальные остатки были взяты из последовательности *B*-домена.

Выбор *B*-домена в качестве основного в структуре гибридного белка был связан с тем, что его последовательность наиболее близка к гипотетической консенсус-последовательности IgG-связывающего домена. Кроме того, в нем нет остатков метионина, и, следовательно, он устойчив к действию бромциана. Чтобы исключить чувствительность *B*-домена к обработке гидроксиламином, входящий в состав одного из его α -спиральных участков дипептид Asn-Gly был заменен на дипептид Asn-Ala. Это было сделано на основании того, что остаток Ala следует за Asn в аналогичном месте структуры второго α -спирального участка, вовлеченного в связывание с IgG. Возможность такой замены подтверждалась также данными компьютерно-графического анализа [9].

Синтетический ген фрагмента белка А (*sspa*) был получен нами в трех вариантах, различающихся структурой регуляторной части. Один из них имел длину полинуклеотидной цепи 475 п.о. (ген *a*) и содержал непосредственно перед первыми кодонами структурного гена сайт *S-D*, предшествующий природному гену *spa*. Вторым вариантом (ген *b*, 471 п.о.) вместо этого участка ДНК содержал *S-D*-сайт гена *lacZ*, а третий (ген *c*, 505 п.о.) — *S-D*-сайт и несколько первых кодонов лидерного пептида триптофанового оперона, за которым следовали сигнал терминирования трансляции и природный *S-D*-сайт гена *spa* (рис. 1). Схема сборки гена основывалась на конструировании целевой ДНК из предварительно клонированных субфрагментов-модулей М1—М4. С этой целью в последовательность гена был введен ряд уникальных сайтов эндонуклеаз рестрикции, в частности *KpnI*, *MluI*, *BglII*, *XhoI*, *BamHI*. В конструкцию искусствен-



Рис. 1. Первичная структура фрагментов F1a и F2 искусственного гена IgG-связывающего участка белка А стафилококков и кодируемых ими пептидов. Сверху даны 5'-концевые последовательности вариантов В и С фрагмента F1 (F1c и F1b). Стрелками показаны олигонуклеотиды и модули M1-M4, из которых составлялись эти фрагменты ДНК. Штриховой линией отмечены сайты расщепления эндонуклеазами рестрикции

ного гена была также заложена возможность мультипликации как участка, кодирующего *B*-домен, так и целого гена.

Фрагменты ДНК, включающие регуляторные элементы плюс участок, кодирующий модифицированный *E*-домен (фрагмент F1), и ген *B*-домена (фрагмент F2), собирались параллельно из предварительно клонированных модулей, которые в свою очередь составлялись из синтетических олигонуклеотидов и сшивались с помощью ДНК-лигазы фага T4. F1 получали в трех вариантах (A, B, C), которые различались структурой 5'-концевого участка, и имели на концах сайты эндонуклеаз рестрикции *XhoI* и *BamHI*, а фрагмент F2 — сайты *BglII* и *BamHI*. При этом благодаря идентичности выступающих тетра-нуклеотидных концов у *BglII* и *BamHI*-сайтов создавалась возможность соединения этих фрагментов и мультипликации фрагмента II при полном сохранении рамки считывания и нативной структуры соединяющего гомологичные домены пептида Ala-Pro-Lys-Ala-Asp. Правда, при этом происходила замена следующего через один от этого пептида остатка Asp на Leu, но в принципе эта замена не должна была повлиять на активность синтезируемого химерного белка, поскольку она была значительно удалена от обоих α -спиральных участков А-белка, участвующих во взаимодействии с IgG.

Клонирование отдельных модулей гена *sspa* и их сборка проводились в полилинкерной плазмиде pHS3, специально сконструированной для этих целей введением в плазмиду pHS2 [10] между *XhoI*- и *HindIII*-сайтами синтетического дуплекса длиной 36 п.о., содержащего *ClaI*-, *KpnI*-, *BglII*-, *BclI*- и *MluI*-сайты (рис. 2). Все варианты фрагмента F1 собирались из трех модулей M—M3, которые были ограничены сайтами эндонуклеаз *XhoI* и *KpnI*, *KpnI* и *MluI*, *MluI* и *BamHI*. Фрагмент F2 собирался из двух модулей, первый из которых был ограничен *BglII*- и *MluI*-сайтами, а второй — *MluI*- и *BamHI*-сайтами и был аналогичен последнему модулю из фрагмента F1. В результате были получены плазмиды pSSPAI (a, b, c) и pSSPAII, содержащие фрагменты целевого гена. Последние объединялись, как это показано на рис. 3.

С этой целью ДНК плазмид pSSPAIa, pSSPAIb и pSSPAIc обрабатывали эндонуклеазами *BamHI* и *PstI*. Полученный при этом меньший фрагмент лигировали с большим фрагментом, образующимся при обработке плазмиды pSSPAII нуклеазами *BglII* и *PstI*, и полученными рекомбинантными ДНК трансформировали клетки *E. coli* HB101. Колонии, выросшие на агаризованной LB-среде, содержащей ампициллин, скринировали на присутствие плазмидных ДНК pSSPAIII (a, b, c), содержащих полноразмерный целевой ген, с помощью гибридизации с 32 P-мечеными олигонуклеотидами d(AAAAACATCTACTCTATCCGTAAGTGG) и d(GATCTGAAATTCAACAAAAA), входящими в состав фрагментов F1 и F2 соответственно. Обратные при этом плазмидные ДНК подвергали рестриктному анализу, а структуру синтетической вставки *XhoI/BamHI* подтверждали методом Максама — Гилберта.

Помимо вариантов гена, кодирующих бидоменный фрагмент белка А, нами были получены гены три- и тетрадоменных структур этого белка. С этой целью три варианта плазмиды pSSPAIII расщепляли эндонуклеазами *PstI* и *BamHI* и полученные при этом малые фрагменты соединяли с большим *BglII/PstI*-фрагментом плазмиды pSSPAII. Полученные при этом рекомбинантные ДНК после введения в бактериальные клетки скринировали на присутствие *XhoI/BamHI*-фрагмента длиной около 700 п.о., соответствующего гену тридоменной белковой структуры (плазмиды pSSPAIVa, pSSPAIVb и pSSPAIVc). Аналогично был получен ген, кодирующий тетрадоменный фрагмент белка А (плазмиды pSSPAVa, pSSPAVb и pSSPAVc) (рис. 3).

С целью получения векторов, способных обеспечить экспрессию искусственного гена фрагмента белка А, сконструированные нами варианты гена *sspa* вводились в плазмиду pUCL1 (производную pUC8 [11]) между *XhoI*- и *BamHI*-сайтами (рис. 4). В полученных при этом экспрессирующих плаزمидях pUCL2a, pUCL2b и pUCL2c искусственный ген фрагмента белка А находился под контролем *lac*-промотора, а *BamHI*-,

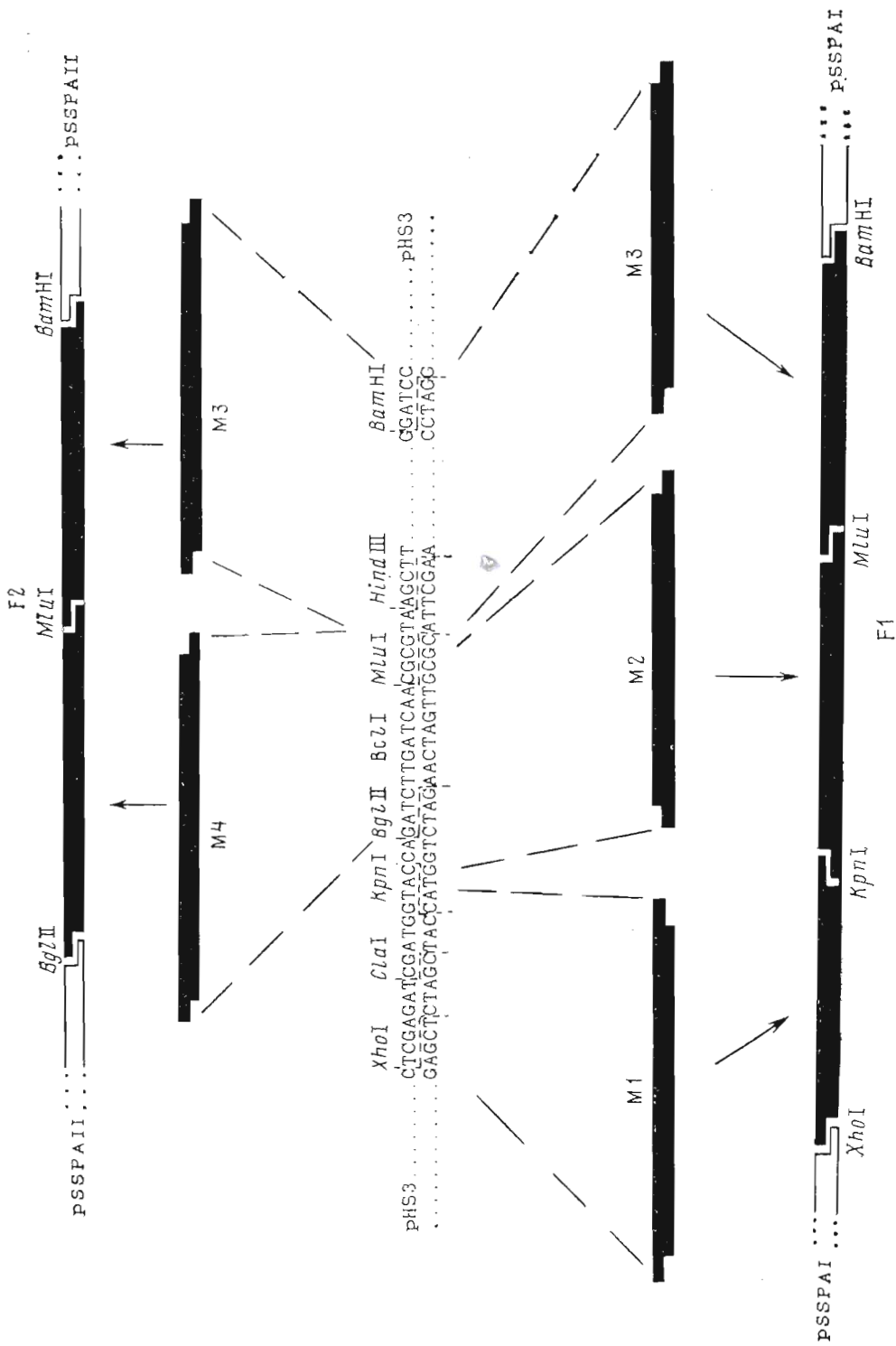


Рис. 2. Схема сборки
 фрагментов F1 и F2
 синтетического гена
 IgG связывающего
 участка белка А из
 модулей M1 - M4. По-
 казана последователь-
 ность полилинкера
 плазмиды pHS3.

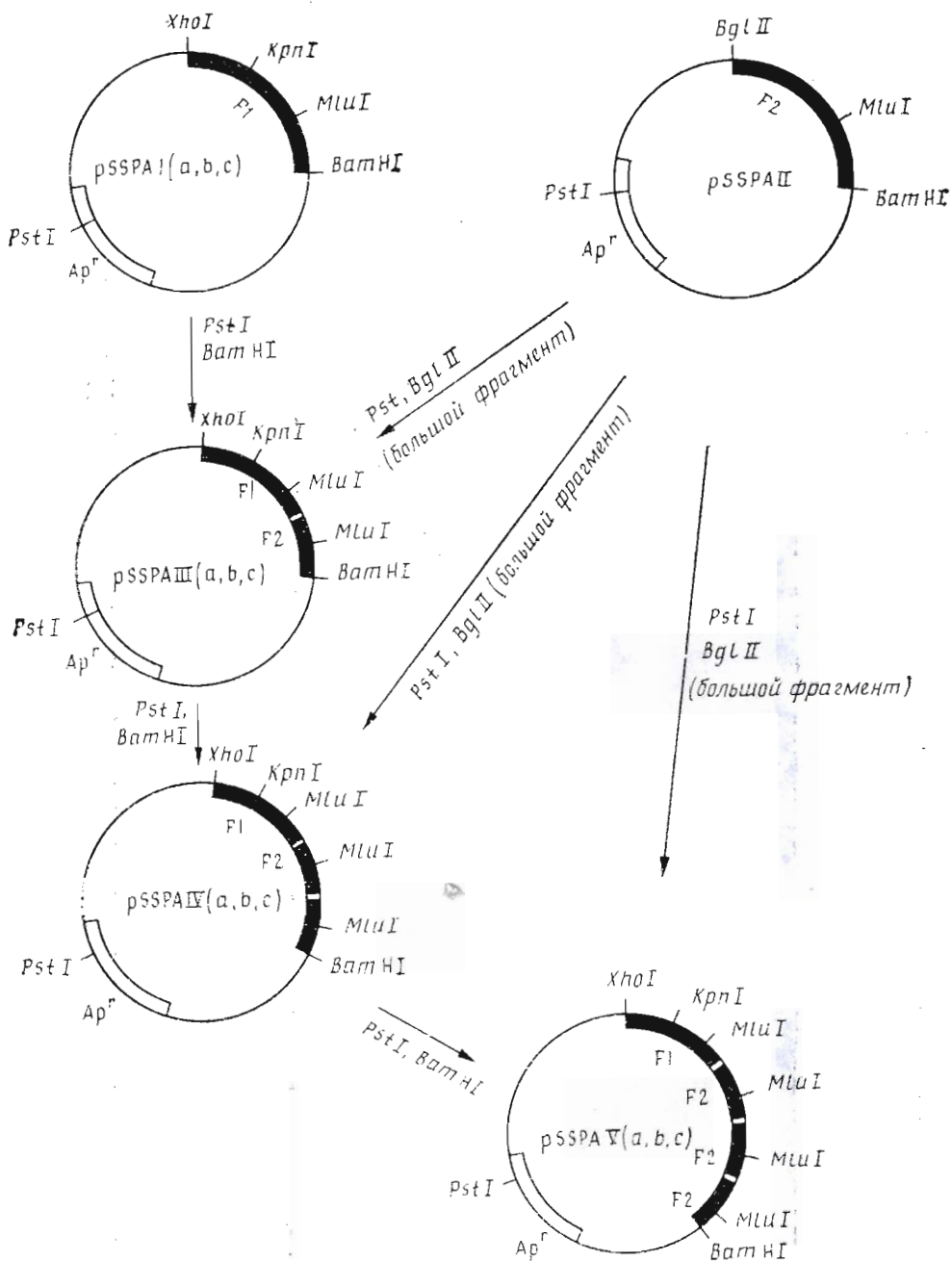


Рис. 3. Схема получения плазмид, несущих гены, кодирующие ди-, три- и тетрадоменные структуры искусственного IgG-связывающего участка белка А. Зачернены фрагменты синтетического гена *ssra*

EcoRI-, *Sall*-, *StuI*- и *Kpn2I*-сайты предназначались для введения в них генов гетерологичных пептидов и белков. С использованием *XhoI*- и *Sall*-сайтов плазмид серии pUCL2 были получены векторы, содержащие несколько копий гена *sspa*. На рис. 4 в качестве примера показано получение тандема генов *sspa*. Необходимо отметить, что аналогичное мультикопирование может быть проведено и после введения в плазмиду серии pUCL2 гена какого-либо другого белка или пептида, продуцируемого в составе гибрида с фрагментом белка А. Подобное увеличение дозы целевого гена в плазмиде должно приводить к повышению уровня его экспрессии, а следовательно, к увеличению выхода целевого белка.

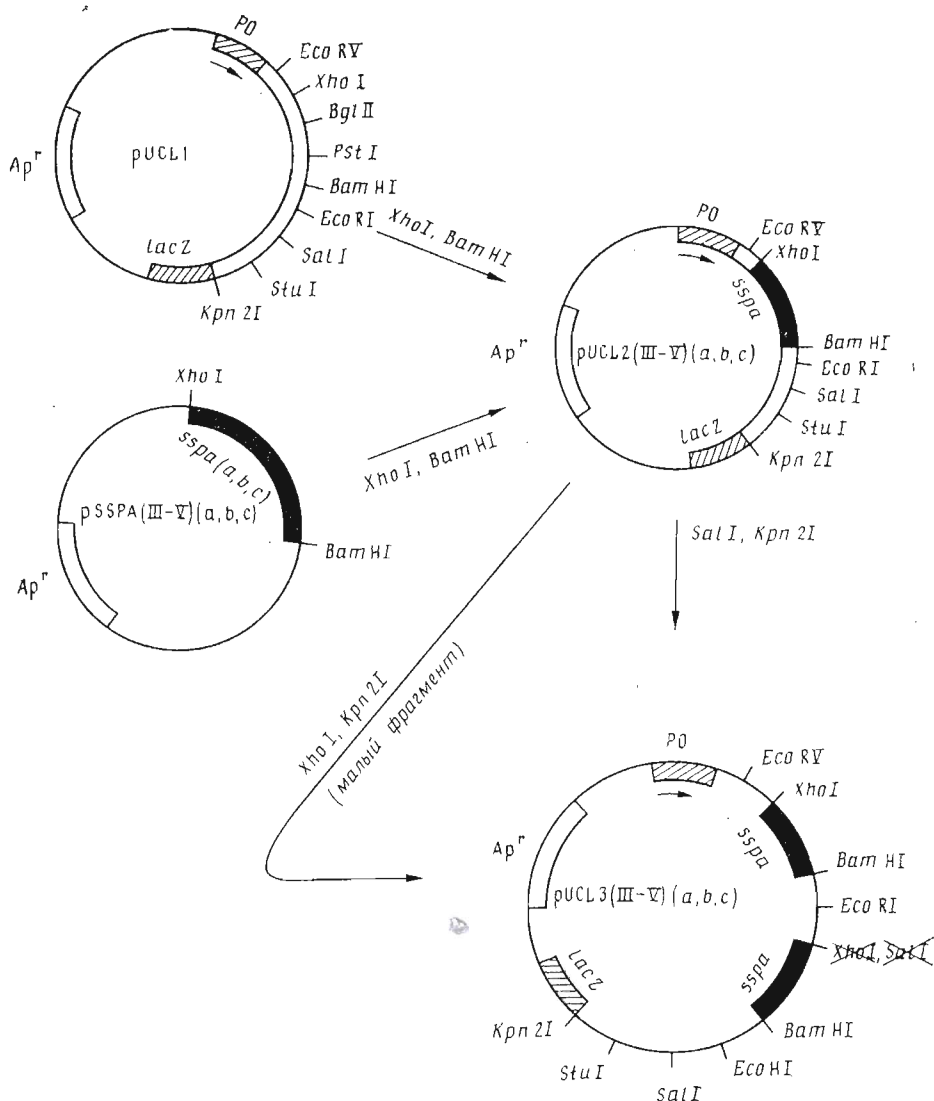


Рис. 4. Получение экспрессирующих плазмид типа pUCLs_{pa} на основе синтетического гена фрагмента белка А, содержащих одну и две копии последнего. Зачернен синтетический ген *sspa*. Штриховкой показаны регуляторный участок PO и фрагмент гена *lacZ*

Для сравнения эффективности экспрессии фрагментов белка А ди-, три- и тетрадоменной структуры в сконструированных нами векторах штаммы, полученные введением этих плазмид в *E. coli* HB101, выращивали в течение 1 сут в LB-бульоне, содержащем ампициллин, сначала при 37° С, а затем температуру повышали до 44° С. Клетки отделяли от культуральной среды центрифугированием, осадок и супернатант тестировали на присутствие IgG-связывающего домена белка А иммуоферментным анализом с использованием конъюгата Fab-фрагмента антител к белку А, присоединенного к β-галактозидазе. Как показали результаты этих экспериментов, все полученные нами конструкции направляли синтез искусственных фрагментов белка А (см. таблицу). Уровень синтеза пептидов, секретирующихся в межклеточное пространство, колебался от 1 до 10 мкг/мл культуральной среды. При этом структура 5'-концевой части гена (*a*, *b* и *c*) оказывала незначительное влияние на выход фрагмента белка А. Как и следовало ожидать, увеличение дозы синтетического гена фрагмента белка А в 2, 3 и 4 раза в плаزمидях серий pUCL3, pUCL4 и pUCL5 приводило к соответствующему увеличению уровня его экспрессии.

Определение фрагментов белка А в клетках штаммов *E. coli*, содержащих плазмиды с искусственными генами этих фрагментов, и в культуральной среде

Плаزمида	Содержание белка А, мкг/мл *	
	клетки **	среда
pUCL1	<0,01	<0,01
pSSPAIII (a, b, c)	0,1	0,5
pUCL2IIIa	1,0	5,0
pUCL2IIIb	1,2	5,5
pUCL2IIIc	1,0	5,0
pUCL2IVa	1,5	6,5
pUCL2IVb	1,6	7,0
pUCL2IVc	1,5	6,8
pUCL2Va	2,0	9,0
pUCL2Vb	2,1	10,0
pUCL2Vc	2,0	10,0
pUCL3IIIa	2,5	15,0
pUCL4IIIa	2,2	20,0
pUCL5IIIa	2,5	26,0

* Определялось методом иммуноферментного анализа как описано в «Экспериментальной части». Даны усредненные значения из трех независимых определений.

** Содержание белка А в клетках определялось после их отделения центрифугированием с последующей деградацией ультразвуком.

Экспериментальная часть

В работе использованы эндонуклеазы рестрикции фирм Boehringer (ФРГ), Reanal (Венгрия) и Pharmacia (Швеция), Т4-ДНК-лигаза, Т4-полинуклеотидкиназа и ДНК-полимераза I *E. coli* (фрагмент Кленова) фирмы Pharmacia.

Химический синтез олигодезоксирибонуклеотидов осуществляли фосфотриэфирным методом с применением О-нуклеофильных катализаторов как описано в работе [12], а также фосфоамидитным методом [13] на синтезаторе Gene Assembler (Pharmacia). Синтетические олигонуклеотиды после удаления защитных групп очищали препаративным гель-электрофорезом и обращенно-фазовой хроматографией на приборе FPLC System (Pharmacia), а их структуры подтверждали методом Максама — Гилберта [14]. 5'-Фосфорилирование олигонуклеотидов осуществляли с помощью Т4-полинуклеотидкиназы и [γ - 32 P]АТР (3000 Ки/ммоль, Amersham, Англия) [14].

Расщепление ДНК эндонуклеазами осуществляли в буфере, содержащем 20 мМ трис-НСl (рН 7,5), 10 мМ MgCl₂, 50 мМ NaCl и 2 мМ дитиотреит, в течение 1 ч при 37°С. Реакции лигирования проводили с 250—500 пмоль олигомеров в реакционной смеси (100 мкл), содержащей 50 мМ трис-НСl (рН 7,5), 10 мМ MgCl₂, 5 мМ дитиотреит, 100 мМ АТР и 20 ед. акт. Т4-ДНК-лигазы, при 10°С в течение 6—12 ч. Продукты лигазных шивок выделяли электрофорезом в 8—10% полиакриламидном геле.

Клонирующее рекомбинантных плазмид и трансформацию компетентных клеток *E. coli*, а также скрининг колоний гибридизацией с мечеными синтетическими олигонуклеотидами осуществляли согласно работе [15]. Плазмидные ДНК выделяли как описано в работе [16]. Нуклеотидные последовательности клонированных ДНК определяли по методу Максама — Гилберта с использованием фрагментов, меченных по 3'-концу с помощью ДНК-полимеразы I *E. coli* (фрагмент Кленова) [14].

Репаративный синтез проводили в буфере, содержащем 50 мМ трис-НСl (рН 7,5), 10 мМ MgCl₂, 2 мМ дитиотреит, или в рестрикционном буфере (см. выше). Во всех случаях использовали 100 мкМ растворы dNTP, 2—4 ед. акт. ДНК-полимеразы *E. coli* и 100 мкл реакционной смеси. Реакцию проводили в течение 30 мин при 20°С и останавливали фенольной экстракцией белка и осаждением ДНК спиртом.

Детектирование и количественная оценка белка А осуществлялись им-

муноферментным анализом (ELISA) с использованием кроличьего IgG. Для этого раствором последнего в 0,1 М бикарбонате натрия (pH 9,5) обрабатывали лунки планшета для иммунологических исследований в течение 1 ч при 37° С. После промывки лунок и обработки их раствором BSA в каждую лунку прибавляли по 90 мкл буфера, содержащего 0,1% твин-20, и по 1—10 мкл клеточных экстрактов или культуральной жидкости для связывания тестируемых фрагментов белка А с IgG и выдерживали 1 ч при 37° С. Затем лунки промывали и обрабатывали Fab-фрагментом кроличьих антител к белку А, конъюгированных с β-галактозидазой (1 ч при 37° С), для связывания с фрагментом белка А. После промывки лунок измеряли активность связанной β-галактозидазы колориметрическим анализом с использованием в качестве субстрата *o*-нитрофенил-β-*D*-галактопиранозида (4 мг/мл).

Авторы благодарят д-ра М. Инганэса (Pharmacia, Швеция) за предоставление Fab-фрагмента антител к белку А.

СПИСОК ЛИТЕРАТУРЫ

1. Efimov V. A., Chakhmakhcheva O. G., Ovchinnikov Yu. A. // *Biotechnology and Genetic Engineering Reviews*/Ed. Russell G. E. Intercept, England, 1986. V. 4. P. 79—116.
2. Nilsson B., Holmgren E., Josephson S., Gatenbeck S., Philipson L., Uhlen M. // *Nucl. Acids Res.* 1985. V. 13. № 4. P. 1151—1162.
3. Uhlen M., Guss B., Nilsson B., Gotz F., Lindberg M. // *J. Bacteriol.* 1984. V. 159. № 2. P. 713—719.
4. Lindmark R., Movitz I., Sjoquist J. // *Eur. J. Biochem.* 1977. V. 74. № 3. P. 623—628.
5. Uhlen M., Guss B., Nilsson B., Gatenbeck S., Philipson L., Lindberg M. // *J. Biol. Chem.* 1984. V. 259. № 3. P. 1695—1702.
6. Moks T., Abrahmsen L., Nilsson B., Hellman U., Sjoquist J., Uhlen M. // *Eur. J. Biochem.* 1986. V. 156. № 3. P. 637—643.
7. Uhlen M., Nilsson B., Guss B., Lindberg M., Gatenbeck S., Philipson L. // *Gene.* 1983. V. 23. № 3. P. 369—378.
8. Moks T., Abrahmsen L., Osterlof B., Josephson S., Ostling M., Enfors S., Persson I., Nilsson B., Uhlen M. // *Biotechnol.* 1987. V. 5. № 2. P. 379—382.
9. Nilsson B., Moks T., Jansson B., Abrahmsen L., Elmblad A., Holmgren E., Henrichson C., Jones T. A., Uhlen M. // *Protein Engng.* 1987. V. 1. № 2. P. 107—113.
10. Чахмахчева О. Г., Бурякова А. А., Мурских О. В., Реввердато С. В., Ефимов В. А., Овчинников Ю. А. // *Биоорг. химия.* 1985. Т. 11. № 11. С. 1533—1546.
11. Messing J., Vieira J. // *Gene.* 1982. V. 19. № 3. P. 259—268.
12. MacBride L. J., Kierzek A., Beausage S. L., Caruthers M. H. // *J. Amer. Chem. Soc.* 1986. V. 108. № 8. P. 2040—2048.
13. Efimov V. A., Buryakova A. A., Dubey I. Y., Polushin N. N., Chakhmakhcheva O. G., Ovchinnikov Yu. A. // *Nucl. Acids Res.* 1986. V. 14. № 16. P. 6525—6540.
14. Maxam A. M., Gilbert W. // *Meth. Enzymol.* 1980. V. 65. P. 499—560.
15. Ефимов В. А., Чахмахчева О. Г. // *Биоорг. химия.* 1982. Т. 8. № 8. С. 2084—2093.
16. Маниагис Т., Фрич Э., Сэмбрун Дж. Молекулярное клонирование. М.: Мир, 1984. С. 101—102, 333—335.

Поступила в редакцию
31.VIII.1988

SYNTHESIS AND EXPRESSION OF AN ARTIFICIAL GENE FOR IgG-BINDING FRAGMENT OF PROTEIN A FROM *STAPHYLOCOCCUS AUREUS*

EFIMOV V. A., BURYAKOVA A. A., POLUSHIN N. N., PASHKOVA I. N.,
DMITRAKOVA E. V., CHAKHMAKHCHEVA O. G.

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

The chemical-enzymatic synthesis of a gene for IgG-binding fragment of the staphylococcal protein A has been carried out. The design of the gene, which consists of signal peptide and modified *E* and *B* domains, and strategy of the synthesis provided possibility of various degrees of polymerization of the gene fragment coding for *B* domain and of the whole gene. Several protein A-like polypeptides composed of the leader sequence, *E* domain and 1 to 4 copies of *B* domain were produced in *E. coli* cells under the *lac* promoter control.