



(51) МПК
C07K 14/47 (2006.01)
C07K 14/575 (2006.01)
C07K 14/62 (2006.01)
C07K 19/00 (2006.01)
C12P 21/02 (2006.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
 ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

C07K 14/575 (2023.05); *C07K 14/62* (2023.05); *C07K 14/605* (2023.05); *C07K 19/00* (2023.05); *C12P 21/02* (2023.05); *C12N 15/62* (2023.05); *C12N 15/70* (2023.05); *C07K 2319/35* (2023.05); *C07K 2319/50* (2023.05); *C07K 2319/60* (2023.05)

(21)(22) Заявка: 2021130321, 18.03.2020

(24) Дата начала отсчета срока действия патента:
18.03.2020

Дата регистрации:
04.08.2023

Приоритет(ы):

(30) Конвенционный приоритет:
19.03.2019 CN 201910210102.9

(43) Дата публикации заявки: 19.04.2023 Бюл. № 11

(45) Опубликовано: 04.08.2023 Бюл. № 22

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 19.10.2021

(86) Заявка РСТ:
CN 2020/080036 (18.03.2020)

(87) Публикация заявки РСТ:
WO 2020/187270 (24.09.2020)

Адрес для переписки:
121205, г. Москва, территория инновационного
центра "Сколково", ул. Луговая, д. 4, корп. 2,
ООО "ЦИС "Сколково", Котлов Дмитрий
Владимирович

(72) Автор(ы):

У, Сун (CN),
 ЧЗАН, Чжэньшань (CN),
 ЛЮ, Хуэйлин (CN),
 ЧЭНЬ, Вэй (CN)

(73) Патентообладатель(и):

НИНГБО КУНПЕНГ БИОТЕКХ КО.,
 ЛТД. (CN)

(56) Список документов, цитированных в отчете
о поиске: WO 2013142859 A2, 26.09.2013, с.3-
13. US 20160289291 A1, 06.10.2016. UGRINOV
K. G. et al., Cotranslational folding increases GFP
folding yield, Biophysical journal, 2010, v. 98, n.
7, p.1312-1320. PAKULA A.A. et al., Genetic
analysis of protein stability and function, Annual
review of genetics, 1989, v. 23, n. 1, p.289-310.
SHOELSON S. et al., (см. прод.)

(54) ГИБРИДНЫЙ БЕЛОК, СОДЕРЖАЩИЙ ФРАГМЕНТЫ ФЛУОРЕСЦЕНТНЫХ БЕЛКОВ, И ЕГО ПРИМЕНЕНИЕ

(57) Реферат:

Изобретение относится к области биотехнологии, конкретно к гибридным белкам на основе фрагментов флуоресцентных белков и целевых терапевтических полипептидов, и может быть использовано для экспрессии и получения целевого полипептида. Предложена конструкция гибридного белка, которая содержит фрагменты флуоресцентных белков и целевой

терапевтический полипептид, выбранный из белка-предшественника человеческого инсулина, белка-предшественника инсулина лизпро, белка-предшественника инсулина гларгина, глюкагоноподобного пептида и его производных эксенатида и лираглутида, инсулина, кальцитонина, тедуглутида. Изобретение обеспечивает повышение эффективности

получения указанных целевых терапевтических пр.
полипептидов. 7 н. и 3 з.п. ф-лы, 5 ил., 1 табл., 6

(56) (продолжение):

Identification of a mutant human insulin predicted to contain a serine-for-phenylalanine substitution, Proceedings of the National Academy of Sciences, 1983, v. 80, n. 24, p.7390-7394. ОВЧИННИКОВ Ю. А., Биоорганическая химия, Москва, 1987, 815с., с.42.

RU 2801248 C2

RU 2801248 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
C07K 14/47 (2006.01)
C07K 14/575 (2006.01)
C07K 14/62 (2006.01)
C07K 19/00 (2006.01)
C12P 21/02 (2006.01)

(12) ABSTRACT OF INVENTION

(52) CPC

C07K 14/575 (2023.05); *C07K 14/62* (2023.05); *C07K 14/605* (2023.05); *C07K 19/00* (2023.05); *C12P 21/02* (2023.05); *C12N 15/62* (2023.05); *C12N 15/70* (2023.05); *C07K 2319/35* (2023.05); *C07K 2319/50* (2023.05); *C07K 2319/60* (2023.05)

(21)(22) Application: **2021130321, 18.03.2020**(24) Effective date for property rights:
18.03.2020Registration date:
04.08.2023

Priority:

(30) Convention priority:
19.03.2019 CN 201910210102.9(43) Application published: **19.04.2023** Bull. № 11(45) Date of publication: **04.08.2023** Bull. № 22(85) Commencement of national phase: **19.10.2021**(86) PCT application:
CN 2020/080036 (18.03.2020)(87) PCT publication:
WO 2020/187270 (24.09.2020)

Mail address:

**121205, g. Moskva, territoriya innovatsionnogo
tsentra "Skolkovo", ul. Lugovaya, d. 4, korp.2, OOO
"TSIS "Skolkovo", Kotlov Dmitrij Vladimirovich**

(72) Inventor(s):

**WU, Song (CN),
ZHANG, Zhenshan (CN),
LIU, Huiling (CN),
CHEN, Wei (CN)**

(73) Proprietor(s):

**NINGBO KUNPENG BIOTECH CO., LTD.
(CN)**

(54) HYBRID PROTEIN CONTAINING FRAGMENTS OF FLUORESCENT PROTEINS AND ITS APPLICATION

(57) Abstract:

FIELD: biotechnology.

SUBSTANCE: invention relates to hybrid proteins based on fragments of fluorescent proteins and target therapeutic polypeptides, and can be used to express and obtain the target polypeptide. A fusion protein construct is proposed that contains fragments of fluorescent proteins and a target therapeutic polypeptide selected from human insulin precursor protein, insulin

lispro precursor protein, insulin glargine precursor protein, glucagon-like peptide and its derivatives exenatide and liraglutide, insulin, calcitonin, teduglutide.

EFFECT: invention provides an increase in the efficiency of obtaining said target therapeutic polypeptides.

10 cl, 5 dwg, 1 tbl, 6 ex

Область техники

Настоящее изобретение относится к биотехнологии, в частности, к гибридным белкам, содержащим фрагменты флуоресцентных белков, и к области их применения.

Уровень техники

5 Пептиды являются классом биомолекул, широко используемых в качестве реагентов во многих областях биомедицинских исследований, а также в качестве терапевтических средств при лечении заболеваний, в качестве диагностических средств при обнаружении патогенных бактерий и в качестве биомаркеров. Широко распространены два способа синтезирования пептидов; одним является химический синтез, а другим - рекомбинантная
10 экспрессия. Химический синтез используется для приготовления разнообразных терапевтических пептидов, включая кортикорелин, паратиреоидный гормон (ПТГ), глюкагоноподобный пептид (ГПП-1) и его производные эксенатид и лираглутид, а также энфувиртид, кальцитонин, бивалирудин, зиконотид, семорелин, соматорелин, секретин, тедуглутид и инсулин. Данный способ требует выполнения многостадийной
15 конденсации фрагментов аминокислот для образования пептидов, а также требует проведения неудобных стадий защиты, снятия защиты и очистки. На сегодняшний момент большинство коммерческих пептидов с количеством аминокислотных остатков менее 50 производятся указанным способом. По мере увеличения спроса на пептиды в фармацевтической промышленности и индустрии проведения биомедицинских
20 исследований растет также и цена аминокислотных фрагментов, используемых при химическом синтезе. Как следствие, в будущем поддержание цен на доступном уровне на терапевтические средства на основе пептидов (например, ГПП-1), используемые для ежедневного приема, может оказаться проблематичным. Хотя с технической точки зрения и существует возможность химического синтеза пептидов с количеством
25 аминокислотных остатков менее 50, однако низкая полезная продуктивность этого процесса и большое количество образующихся при этом органических отходов делают такой процесс синтеза неэкономичным. В настоящее время большинство пептидов с количеством аминокислотных остатков более 50 производятся способом рекомбинантной экспрессии в клетках-хозяевах таких организмов, как, например,
30 бактерии, дрожжи, насекомые и млекопитающие. В течение многих лет обычным способом обеспечения экспрессии полипептидов являлось использование гибридных белков. Тем не менее, использование доступных на данный момент времени способов использования гибридных белков для обеспечения экспрессии пептидов сопряжено со множеством технических проблем, особенно при получении пептидов, продуцирующих
35 менее 50 аминокислотных остатков. Например, известный из предыдущего уровня техники гибридный белок характеризуется большим молекулярным весом, сильной гидрофобностью и трудностью проведения сепарации. Целевой белок характеризуется низким значением удельной плотности, низким значением коэффициента слияния, устойчивой структурой, трудностью усвоения и значительным мертвым объемом в
40 загрузочной ионной колонке и гидрофобной колонке.

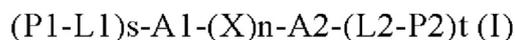
Таким образом, существует насущная потребность в разработке нового гибридного белка для обеспечения экспрессии целевого пептида, для преодоления ограничений, связанных с использованием существующих гибридных белков и, в то же самое время, для увеличения полезной продуктивности вставляемых аминокислотных белков,
45 например, повышения выработки искусственных аминокислотных белков, таких как трет-бутоксикарбонил-лизиновые белки.

Сущность изобретения

Целью настоящего изобретения является создание нового гибридного белка для

обеспечения экспрессии целевого пептида и, в то же самое время, увеличения полезной продуктивности вставляемых искусственных аминокислотных белков, например, повышения выработки искусственных аминокислотных белков, таких как трет-бутоксикарбонил-лизиновые белки.

5 Первый объект настоящего изобретения касается гибридного белка, имеющего структуру, показанную в Формуле I:



где

10 «-» обозначает пептидную связь или линкерный пептид,
каждый P1 независимо является первым целевым пептидом;
каждый P2 независимо является вторым целевым пептидом;
каждый L1 независимо является отсутствующим или первым линкерным пептидом;
каждый L2 независимо является отсутствующим или вторым линкерным пептидом;
15 A1 является отсутствующим или сигнальным пептидом;
A2 является отсутствующим или сигнальным пептидом;
каждый X независимо является одиночным β-складчатым блоком флуоресцентного белка;

n является положительным целым числом от 1 до 8;

20 s может принимать значения 0, 1 или 2;

t может принимать значения 0, 1 или 2; и

дополнительным условием является неравенство s и t нулю одновременно.

В еще одном предпочтительном варианте осуществления s = 0 и t = 1.

В еще одном предпочтительном варианте осуществления s = 1 и t = 0.

25 В еще одном предпочтительном варианте осуществления значение n принадлежит диапазону 1-6, предпочтительно, диапазону 2-4.

В еще одном предпочтительном варианте осуществления A1 является сигнальным пептидом, а A2 является отсутствующим пептидом.

30 В еще одном предпочтительном варианте осуществления β-складчатый блок выбирается из группы, состоящей из β-складчатых блоков u1, u2, u3, u4, u5, u6, u7, u8, u9, u10 и u11 флуоресцентного белка.

В еще одном предпочтительном варианте осуществления длина каждого X составляет 10-14 аминокислот.

35 В еще одном предпочтительном варианте осуществления каждый X является различным.

В еще одном предпочтительном варианте осуществления каждый X является тем же самым.

40 В еще одном предпочтительном варианте осуществления флуоресцентный белок выбирается из группы, состоящей из зеленого флуоресцентного белка (ЗФБ), желтого флуоресцентного белка (ЖФБ), синего флуоресцентного белка (СФБ), гена голубого флуоресцентного белка (ГФБ) и любой из комбинаций указанных элементов.

В еще одном предпочтительном варианте осуществления в качестве флуоресцентного белка используется зеленый флуоресцентный белок (ЗФБ).

45 В еще одном предпочтительном варианте осуществления ЗФБ имеет последовательность аминокислот, показанную на SEQ ID NO.: 13.

В еще одном предпочтительном варианте осуществления между любыми двумя (3-складчатыми блоками отсутствует гибкое соединение.

В еще одном предпочтительном варианте осуществления (X)_n выполняет роль элемента, способствующего экспрессии белка.

В еще одном предпочтительном варианте осуществления общая длина L_n элемента (X) $_n$ составляет 3,7-30% общей длины L_0 флуоресцентного белка, предпочтительно, 7,5-20,5%, более предпочтительно, 7,5-15,5%).

5 В еще одном предпочтительном варианте осуществления X выбирается из группы, состоящей из

элемента u_j , который имеет последовательность аминокислот, показанную на SEQ ID NO.: j;

где j может принимать значения 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или 11.

10 В еще одном предпочтительном варианте осуществления X выбирается из группы, состоящей из:

X	Последовательность аминокислот
u1	VPILVELDGDVNG (SEQ ID NO.: 1)

15	u2	HKFSVRGEGEGDAT (SEQ ID NO.: 2)
	u3	KLTLKFICTT (SEQ ID NO.: 3)
	u4	YVQERTISFKD (SEQ ID NO.: 4)
20	u5	TYKTRAEVKFEGD (SEQ ID NO.: 5)
	u6	TLVNRIELKGIDF (SEQ ID NO.: 6)
	u7	HNVYITADKQ (SEQ ID NO.: 7)
25	u8	GIKANFKIRHNVED (SEQ ID NO.: 8)
	u9	VQLADHYQQNTPIG (SEQ ID NO.: 9)
	u10	HYLSTQSVLSKD (SEQ ID NO.: 10)
30	u11	HMVLLFVTAAGI (SEQ ID NO.: 11)

В еще одном предпочтительном варианте осуществления X, кроме того, включает в себя последовательность аминокислот, образованную путем взаимного замещения элементов R, K и H в любой из последовательностей, как показано на SEQ ID NO.: 1-11; и/или

35 X, кроме того, включает в себя последовательность аминокислот, образованную путем взаимного замещения элементов P и Q в последовательности, как показано на любом из SEQ ID NO.: 1-11; и/или

40 X, кроме того, включает в себя последовательность аминокислот, образованную путем взаимного замещения элементов T и S в последовательности, как показано на любом из SEQ ID NO.: 1-11.

В еще одном предпочтительном варианте осуществления X выбирается из группы, состоящей из

(A) полипептида, имеющего последовательность аминокислот, показанную на любом из SEQ ID NO.: 1-11;

45 (B) полипептида, имеющего гомологию $\geq 80\%$ (предпочтительно, гомологию $\geq 85\%$, более предпочтительно, $\geq 90\%$, более предпочтительно, $\geq 95\%$, наиболее предпочтительно, $\geq 97\%$) с последовательностью аминокислот, показанной на любом из SEQ ID NO.: 1-11, с сохранением характеристик;

(С) производного полипептида, образованного путем замещения, удаления или добавления аминокислотных остатков 1-5 к последовательности аминокислот, показанной на любом из SEQ ID NO: 1-11, с сохранением характеристик.

В еще одном предпочтительном варианте осуществления (X)_n может быть u8, u9, u2-u3, u4-u5, u8-u9, u1-u2-u3, u2-u3-u4, u3-u4-u5, u5-u6-u7, u8-u9-u10, u9-u10-u11, u3-u5-u7, u3-u4-u6, u4-u7-u10, u6-u8-u10, u1-u2-u3-u4, u2-u3-u4-u5, u3-u4-u3-u4, u3-u5-u7-u9, u5-u6-u7-u8, u1-u3-u7-u9, u2-u2-u7-u8, u7-u2-u5-u11, u3-u4-u7-u10, u1-I-u2, u1-I-u5, u2-I-u4, u3-I-u8, u5-I-u6 или u10-I-u11.

В еще одном предпочтительном варианте осуществления A1 или A2 имеет последовательность аминокислот, показанную на SEQ ID NO.: 12 (MVSKGEELFTGV).

В еще одном предпочтительном варианте осуществления A1-(X)_n имеет последовательность аминокислот, показанную на SEQ ID NO.: 14, 15, 16, 17, 22, 23, 24, 26, 27, 28, 29 или 30.

В еще одном предпочтительном варианте осуществления первый линкерный пептид содержит первый участок рестрикции (например, участок рестрикции TEV).

В еще одном предпочтительном варианте осуществления второй линкерный пептид содержит второй участок рестрикции.

В еще одном предпочтительном варианте осуществления гибкий линкер I содержит третий участок рестрикции.

В еще одном предпочтительном варианте осуществления третий участок рестрикции является участком рестрикции EK (как показано в последовательности DDDDK, SEQ ID NO.: 25).

В еще одном предпочтительном варианте осуществления (X)_n может являться одним из следующих элементов: u1-EK-u2, u1-EK-u5, u2-EK-u4, u3-EK-u8, u5-EK-u6 или u10-EK-u11, где EK представляет собой участок рестрикции фермента EK.

В еще одном предпочтительном варианте осуществления первый, второй и третий участки рестрикции отличаются друг от друга.

В еще одном предпочтительном варианте осуществления два или три из состава первого, второго и третьего участков рестрикции являются одинаковыми.

В еще одном предпочтительном варианте осуществления первый, второй и третий участки рестрикции отсутствуют в P1 и P2.

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид, кроме того, содержат участки рестрикции, отличающиеся от первого, второго и третьего участков рестрикции.

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид содержат участок рестрикции трипсина, предпочтительно, первый линкерный пептид и/или второй линкерный пептид содержат, по меньшей мере, либо аргинин (R), либо лизин (K).

В еще одном предпочтительном варианте осуществления N-концевой аминокислотой первого линкерного пептида является Arg или Lys.

В еще одном предпочтительном варианте осуществления C-концевой аминокислотой второго линкерного пептида является Arg или Lys.

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид, кроме того, содержат последовательность распознавания протеазы вируса гравировки табака.

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид содержат последовательность аминокислот, показанную на SEQ ID NO.: 18.

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид, кроме того, содержат последовательность меток, содействующую экспрессии и/или очистке.

5 В еще одном предпочтительном варианте осуществления последовательность меток представляет собой гистидиновый таг, предпочтительно, таг размерностью 6×HIS.

В еще одном предпочтительном варианте осуществления как P1, так и P2 независимо выбираются из следующего: белок-предшественник человеческого инсулина, белок-предшественник инсулина лизпро, белок-предшественник инсулина гларгин, паратиреоидный гормон, кортикорелин, кальцитонин, бивалирудин, глюкагоноподобный пептид и его производные эксенатид и лираглутид, а также зиконотид, серморелин, соматорелин, секретин, тедуглутид, гирудин, гормон роста, фактор роста, фактор, высвобождающий гормон роста, адренкортикотропный гормон, рилизинг-фактор, деслорелин, десмопрессин, элкатонин, глюкагон, лейпрорелин, лютропин-рилизинг-гормон, соматостатин, тиреотропин-рилизинг-гормон, трипторелин, вазоактивный пептид кишечника, интерферон, паратиреоидный гормон, пептид ВНЗ, амилоидный пептид или фрагменты указанных выше пептидов, или одна из комбинаций (предпочтительно, указанных выше полипептидов, однако не ограничиваясь ими).

В еще одном предпочтительном варианте осуществления как P1, так и P2 независимо являются белком искусственной аминокислоты.

20 В еще одном предпочтительном варианте осуществления как P1, так и P2 независимо имеют последовательность длиной 10-200 аминокислот, предпочтительно, последовательность длиной 10-80 аминокислот.

В еще одном предпочтительном варианте осуществления как P1, так и P2 независимо являются проинсулином или инсулином, предпочтительно, человеческим проинсулином или человеческим инсулином.

В еще одном предпочтительном варианте осуществления инсулин включает в себя инсулин медленного действия или инсулин быстрого действия.

В еще одном предпочтительном варианте осуществления лизин в позиции 29 проинсулина является производным алкинилоксикарбонил-лизина, производным трет-бутоксикарбонил-лизинового белка (трет-бутоксикарбонил-L-лизина), ацилированного алифатического лизина или комбинацией указанных элементов.

В еще одном предпочтительном варианте осуществления проинсулин имеет последовательность аминокислот, показанную на SEQ ID NO: 19 или 20.

В еще одном предпочтительном варианте осуществления целевой пептид P2 располагается у C-конца (X)_n (или соединяется с ним).

В еще одном предпочтительном варианте осуществления гибридный белок перерабатывается для образования первого целевого пептида и/или второго целевого пептида.

40 В еще одном предпочтительном варианте осуществления первый целевой пептид и второй целевой пептид могут быть одинаковыми или различными.

В еще одном предпочтительном варианте осуществления гибридный белок перерабатывается для образования первого целевого пептида и/или второго целевого пептида, а (X)_n разделяется с образованием короткого пептида длиной L_x, намного меньшей длины L_p первого целевого пептида и/или второго целевого пептида.

45 В еще одном предпочтительном варианте осуществления каждый L_x является аминокислотой 1-25.

В еще одном предпочтительном варианте осуществления отношение длины L_x к длине L_p составляет от 1/2 до 1/10, предпочтительно, от 1/3 to 1/8.

В еще одном предпочтительном варианте осуществления разность длины L_p и длины L превышает 1,3KD.

В еще одном предпочтительном варианте осуществления гибридный белок имеет структуру, выбираемую из группы, состоящей из: A1-u8-L2-P2, A1-u9-L2-P2, A1-u2-u3-L2-P2, A1-u4-u5-L2-P2, A1-u8-u9-L2-P2, A1-u3-u5-u7-L2-P2, A1-u1-u2-u3-L2-P2, A1-u1-u2-u3-u4-L2-P2, A1-u3-u4-u6-L2-P2, A1-u4-u7-u10-L2-P2, A1-u3-u4-u7-u10-L2-P2 или A1-u5-EK-u6-L2-P2.

В еще одном предпочтительном варианте осуществления гибридный белок имеет последовательность аминокислот, показанную на SEQ ID NO.: 21.

Второй объект настоящего изобретения касается изолированного полинуклеотидного кодирования гибридного белка в соответствии с первым объектом настоящего изобретения.

Третий объект настоящего изобретения касается вектора, включающего в себя полинуклеотид в соответствии со вторым объектом настоящего изобретения.

В еще одном предпочтительном варианте осуществления вектор выбирается из группы, состоящей из ДНК, РНК, плазмиды, лентивирусного вектора, аденовирусного вектора, ретровирусного вектора, транспозона или комбинации вышеуказанных элементов.

В еще одном предпочтительном варианте осуществления вектор является плазмидой; предпочтительно, вектор является вектором pBAD-HisA и/или вектором pEvol-pBpF.

Четвертый объект настоящего изобретения касается клетки-хозяина, содержащей вектор согласно третьему объекту настоящего изобретения, или полинуклеотида согласно второму объекту настоящего изобретения, интегрированному в хромосому или же обеспечивающему экспрессию гибридного белка согласно первому объекту настоящего изобретения.

В еще одном предпочтительном варианте осуществления клетка-хозяин является кишечной палочкой, сенной палочкой, дрожжевой клеткой, клеткой насекомого, клеткой млекопитающего или же комбинацией вышеуказанных элементов.

В еще одном предпочтительном варианте осуществления ни одна из клеток-хозяев не содержит протеаз, соответствующих первому, второму и третьему участкам рестрикции.

Пятый объект настоящего изобретения касается способа приготовления белка, содержащего следующие этапы:

(а) производят культивирование клетки-хозяина согласно четвертому объекту настоящего изобретения, получая, таким образом, гибридный белок согласно первому аспекту настоящего изобретения.

В еще одном предпочтительном варианте осуществления способ, кроме того, содержит следующий этап: (b) очищают гибридный белок, полученный на этапе (а).

В еще одном предпочтительном варианте осуществления способ, кроме того, включает в себя процедуру протеолитической переработки гибридного белка согласно первому аспекту настоящего изобретения, для высвобождения целевого пептида из гибридного белка.

В еще одном предпочтительном варианте осуществления при выполнении этапа (с) перерабатывают гибридный белок согласно первому объекту настоящего изобретения с использованием протеазы для получения переработанного продукта; и, необязательно,

(d) изолируют или очищают целевой пептид от переработанного продукта.

В еще одном предпочтительном варианте осуществления этап очистки включает в себя дальнейшую очистку целевого пептида путем гель-фильтрации или использования

процесса высокоэффективной жидкостной хроматографии.

Шестой объект настоящего изобретения касается использования гибридного белка согласно первому объекту настоящего изобретения, или полинуклеотида согласно второму объекту настоящего изобретения, или вектора согласно третьему объекту настоящего изобретения, или клетки-хозяина согласно четвертому объекту настоящего изобретения, для обеспечения экспрессии и подготовки целевого пептида.

Следует понимать, что в пределах объема настоящего изобретения все вышеупомянутые технические особенности настоящего изобретения и все технические характеристики, подробно описанные ниже (например, варианты осуществления), могут быть объединены друг с другом для образования нового или предпочтительного технического решения. Вследствие ограниченного объема указанное в настоящем документе не повторяется.

Описание чертежей

На Фигуре 1 показана структура созревшего инсулина, корректно образующего дисульфидные связи.

На Фигуре 2 показана плазмидная карта экспрессии гибридного белка рекомбинантного человеческого инсулина.

На Фигуре 3 показано схематическое изображение гибридного белка рекомбинантного человеческого инсулина.

На Фигуре 4 показана структура трет-бутоксикарбонил-лизина.

На Фигуре 5 продемонстрирована экспрессия рекомбинантного гибридного белка. Линия 1, Экспрессия гибридного белка A1-u4-u5-TEV-R-MiniINS с молекулярной массой 11,2 кД; Линия 2, Экспрессия гибридного белка A1-u4-u5-TEV-R-GLP1 с молекулярной массой 8,7 кД; Линия 3, Экспрессия гибридного белка A1-u8-TEV-R-MiniINS с молекулярной массой 10,0 кД; линия 4, экспрессия гибридного белка A1-u3-u5-u7-TEV-R-MiniINS с молекулярной массой 12,2 кД; линия 5, Экспрессия гибридного белка A1-и5-EK-u6-TEV-R-MiniINS с молекулярной массой 12,0 кД; М представляет собой стандарт молекулярной массы белка.

Подробное описание

После обширных и всесторонних исследований изобретатели впервые получили новый гибридный белок, пригодный для обеспечения экспрессии целевого пептида. Наличие (X)_n или A1-(X)_n в гибридном белке настоящего изобретения обеспечивает улучшенные параметры процесса складывания и экспрессии целевого гибридного пептида; при этом повышается полезная продуктивность и растворимость гибридного белка и снижается степень его межмолекулярного взаимодействия, так что целевой гибридный пептид может складываться при высоких значениях концентраций, что обеспечивает промышленную значимость процесса. Кроме того, элементы (X)_n или A1-(X)_n гибридного белка могут быть расщеплены на множество коротких пептидов намного меньшей длины, по сравнению с целевым пептидом, что обеспечивает более эффективное отделение его от целевого пептида, повышая удобство процедуры очистки целевого пептида. Руководствуясь вышеизложенным, изобретатели представили настоящее изобретение.

Термины

Перед представлением описания настоящего изобретения необходимо указать, что настоящее изобретение не ограничивается описанными конкретными способами и экспериментальными условиями, т.к. такие способы и условия могут варьироваться. Необходимо также понимать, что термины, использованные в данном документе, предназначены лишь для описания конкретных вариантов осуществления, а не для

ограничения объема настоящего изобретения, который ограничивается лишь прилагаемой формулой изобретения.

Если не указано иное, все технические и научные термины, используемые в данном документе, имеют значение, согласующееся с обычным трактованием его специалистами в той области техники, к которой относится настоящее изобретение.

В контексте данного документа, при использовании в отношении конкретно приводимого значения, термин «около» означает, что значение может отклоняться от приведенного значения не более чем на 1%. Например, в контексте данного документа выражение «около 100» включает в себя все значения от 99 до 101 (например, 99,1; 99,2; 99,3; 99,4 и т.д.).

В контексте данного документа термин «содержащий» или «включающий» может быть неограничивающим, частично ограничивающим и ограничивающим. Другими словами, такой термин также включает в себя значения «в значительной степени состоящий из» или «состоящий из».

Трехбуквенные и однобуквенные коды аминокислот, используемые в настоящем изобретении, соответствуют описанию, приведенному в публикации J. biool. chem, 243, с. 3558 (1968).

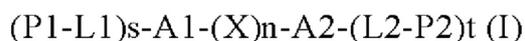
В контексте данного документа использование термина «необязательный» или «необязательно» означает, что событие или ситуация, описываемые далее, могут произойти или возникнуть, но не обязаны произойти или возникнуть.

Термин «идентичность последовательности», используемый в настоящем изобретении, относится к степени идентичности двух нуклеиновых кислот или двух последовательностей аминокислот при их оптимальном выравнивании и сравнении с допустимыми мутациями, такими как замещение, вставка или делеция. Степень идентичности при сравнении последовательности, описанной в настоящем изобретении, с ее идентичной последовательностью может достигать, по меньшей мере, до 85%, 90% или 95%; предпочтительно, по меньшей мере до 95%. Неограничивающие примеры включают в себя 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%.

Гибридный белок

В контексте данного документа термины «гибридный белок настоящего изобретения», «рекомбинантный гибридный белок» и «полипептид» относятся к гибриднему белку согласно первому объекту настоящего изобретения.

Так, гибридный белок настоящего изобретения включает в себя структуру, показанную в Формуле I:



где

«-» обозначает пептидную связь или линкерный пептид,
 каждый P1 независимо является первым целевым пептидом;
 каждый P2 независимо является вторым целевым пептидом;
 каждый L1 независимо является отсутствующим или первым линкерным пептидом;
 каждый L2 независимо является отсутствующим или вторым линкерным пептидом;
 A1 является отсутствующим или сигнальным пептидом;
 A2 является отсутствующим или сигнальным пептидом;
 каждый X независимо является одиночным β-складчатым блоком флуоресцентного белка;

n является положительным целым числом от 1 до 8;

s может принимать значения 0, 1 или 2;

t может принимать значения 0, 1 или 2; и
дополнительным условием является неравенство s и t нулю одновременно.

В еще одном предпочтительном варианте осуществления используемый флуоресцентный белок представляет собой ЗФБ, причем предпочтительно в состав аминокислотной последовательности ЗФБ входит аминокислотная последовательность (241 AA), как показано на SEQ ID NO.: 13.

MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTTLKFIC
TTGKLPVPWPTLVTTLTYGVCFSRYPDHMKRHDFFKSAMPEGYVQERTISF
KDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYI
TADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLST
QSVLSKDPNEKRDHMLLEFVTAAGITHGMDELYAGS (SEQ ID NO.: 13)

Среди них, в указанной выше последовательности подчеркнутые части представляют собой β-складчатые блоки u1, u2, u3, u4, u5, u6, u7, u8, u9, u10 и u11 флуоресцентного белка последовательности.

В еще одном предпочтительном варианте осуществления (X)_n может являться одним из следующих элементов: u8, u9, u2-u3, u4-u5, u8-u9, u1-u2-u3, u2-u3-u4, u3-u4-u5, u5-u6-u7, u8-u9-u10, u9-u10-u11, u3-u5-u7, u3-u4-u6, u4-u7-10 u10, u6-u8-u10, u1-u2-u3-u4, u2-u3-u4-u5, u3-u4-u3-u4, u3-u5-u7-u9, u5-u6-u7-u8, u1-u3-u7-u9, u2-u2-u7-u8, u7-u2-u5-u11, u3-u4-u7-u10, u1-I-u2, u1-I-u5, u2-I-u4, u3-I-u8, u5-I-u6 или u10-I-u11.

В еще одном предпочтительном варианте осуществления каждый X(u_j) из состава (X)_n имеет или не имеет гибкое соединение I, предпочтительно, гибкий линкер I.

В еще одном предпочтительном варианте осуществления гибкий линкер I содержит третий участок рестрикции.

В еще одном предпочтительном варианте осуществления гибкий линкер I содержит участок рестрикции фермента EK (DDDDK).

В еще одном предпочтительном варианте осуществления гибридный белок перерабатывается для образования первого целевого пептида и/или второго целевого пептида, а (X)_n разрезается с образованием короткого пептида с длиной L_x, являющейся намного меньшей длины L_p первого целевого пептида и/или второго целевого пептида.

В еще одном предпочтительном варианте осуществления A1-(X)_n представляет собой A1-u8, A1-u9, A1-u2-u3, A1-u4-u5, A1-u8-u9, A1-u3-u5-u7, A1-u1-u2-u3, A1-u1-u2-u3-u4, A1-u3-u4-u6, A1-u4-u7-u10, A1-u3-u4-u7-u10 или A1-u5-EK-u6.

В еще одном предпочтительном варианте осуществления A1-(X)_n имеет 30 последовательность аминокислот, показанную на SEQ ID NO.: 14, 15, 16, 17, 22, 23, 24, 26, 27, 28, 29 или 30.

Последовательность A1-u8:

MVSKGEELFTGVGIKANFKIRHNVED (SEQ ID NO.: 23)

Последовательность A1-u9:

MVSKGEELFTGVVQLADHYQQNTPIG (SEQ ID NO.: 17)

Последовательность A1-u2-u3:

MVSKGEELFTGVHFKFSVRGEGEGDATKLTTLKFICTT (SEQ ID NO.: 24)

Последовательность A1-u4-u5:

MVSKGEELFTGVVYVQERTISFKDTYKTRAEVKFEGD (SEQ ID NO.: 16)

Последовательность A1-u3-u5-u7:

MVSKGEELFTGVKLTTLKFICTTTYKTRAEVKFEGDHNVYITADKQ (SEQ ID NO.: 15)

Последовательность u8-u9 (A1 отсутствует):

GIKANFKIRHNVEDVQLADHYQQNTPIG (SEQ ID NO.: 14)

Последовательность A1-u5-EK-u6:

MVSKGEELFTGVTYKTRAEVKFEGDDDDDKTLVNRIELKGIDF (SEQ ID NO.: 22)

Последовательность A1-u1-u2-u3:

5 MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATKLTCLKFICTT (SEQ ID NO.: 26)

Последовательность A1-u1-u2-u3-u4:

MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATKLTCLKFICTT
YVQERTISFKD (SEQ ID NO.: 27)

10 Последовательность A1-u3-u4-u6:

MVSKGEELFTGVKLTCLKFICTTYVQERTISFKD TLVNRIELKGIDF (SEQ ID NO.: 28)

Последовательность A1-u4-u7-u10:

MVSKGEELFTGVYVQERTISFKDHNVYITADKQHLYSTQSVLSKD (SEQ ID NO.: 29)

Последовательность A1-u3-u4-u7-u10:

15 MVSKGEELFTGVKLTCLKFICTTYVQERTISFKDHNVYITADKQHLYSTQS VLSKD (SEQ ID NO.: 30)

В еще одном предпочтительном варианте осуществления первый линкерный пептид и/или второй линкерный пептид содержат последовательность аминокислот, показанную на SEQ ID NO.: 18 (ENLYFQGR).

20 В еще одном предпочтительном варианте осуществления проинсулин имеет последовательность аминокислот, показанную на SEQ ID NO: 19 или 20.

FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGGQVELGGGP

GAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN (SEQ ID NO.: 19)

25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN (SEQ ID NO.: 20)

В еще одном предпочтительном варианте осуществления гибридный белок имеет последовательность аминокислот, показанную на SEQ ID NO.: 21.

30 MVSKGEELFTGVYVQERTISFKDTYKTRAEVKFEGDENLYFQGRFVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN (SEQ ID NO.: 21)

35 В контексте данного документа термин «гибридный белок» также включает в себя варианты, проявляющие указанную выше активность. В состав указанных вариантов входят следующие (список, однако, не является исчерпывающим): 1-3 (обычно 1-2, более предпочтительно 1) аминокислотных делеций, вставок и/или замещений, а также добавление или удаление одной или нескольких (как правило, 3-х или менее, предпочтительно, 2-х или менее, более предпочтительно, 1-й или менее) аминокислот у С-конца и/или N-конца. В качестве примера, в этой области при замещении аминокислот с близкими или идентичными свойствами функция белка обычно не изменяется. В качестве еще одного примера, добавление или удаление одной или нескольких аминокислот у С-конца и/или N-конца обычно не изменяет структуру и функцию белка. Кроме того, указанный термин также включает в себя полипептид
45 настоящего изобретения в мономерной и мультимерной формах. Указанный термин также включает в себя линейные и нелинейные полипептиды (например, циклические пептиды).

Настоящее изобретение также включает в себя активные фрагменты, производные

и аналоги упомянутого выше гибридного белка. В контексте данного документа термины «фрагмент», «производная» и «аналог» относятся к полипептиду, который в значительной степени сохраняет функцию или активность гибридного белка настоящего изобретения. Фрагменты, производные или аналоги полипептида настоящего изобретения могут представлять собой (i) полипептид, в котором замещены один или несколько консервативных или неконсервативных аминокислотных остатков (предпочтительно, консервативных аминокислотных остатков), или (ii) полипептид с заместительной группой в одном или нескольких аминокислотных остатках, или (iii) полипептид, образованный путем слияния полипептида с другим составом (например, составом, продлевающим время полужизни полипептида, такой как полиэтиленгликоль), или (iv) полипептид, образованный путем слияния дополнительной аминокислотной последовательности с последовательностью этого полипептида (гибридный белок, образованный путем слияния с лидерной последовательностью, секреторной последовательностью или последовательностью тага bHis). Согласно принципам, содержащимся в данном документе, указанные фрагменты, производные и аналоги подпадают под сферу действия, хорошо известную специалистам в данной области.

Использование термина «предпочтительный тип» в применении к активной производной означает, что, в сравнении с аминокислотной последовательностью настоящего изобретения максимум 3, предпочтительно, максимум 2 и, более предпочтительно, максимум 1 аминокислота в процессе образования полипептида замещается аминокислотами с близкими или аналогичными свойствами. Лучшим способом производства указанных консервативных вариантов полипептидов, согласно Таблице А, является замещение аминокислот.

Таблица А

Исходные остатки	Репрезентативное замещение	Предпочтительное замещение
Ala (A)	Val; Leu; Ile	Val

	Arg (R)	Lys; Gln; Asn	Lys
	Asn (N)	Gln; His; Lys; Arg	Gln
5	Asp (D)	Glu	Glu
	Cys (C)	Ser	Ser
	Gln (Q)	Asn	Asn
10	Glu (E)	Asp	Asp
	Gly (G)	Pro; Ala	Ala
	His (H)	Asn; Gln; Lys; Arg	Arg
15	Ile (I)	Leu; Val; Met; Ala; Phe	Leu
	Leu (L)	Ile; Val; Met; Ala; Phe	Ile
	Lys (K)	Arg; Gln; Asn	Arg
20	Met (M)	Leu; Phe; Ile	Leu
	Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
	Pro (P)	Ala	Ala
25	Ser (S)	Thr	Thr
	Thr (T)	Ser	Ser
	Trp (W)	Tyr; Phe	Tyr
30	Tyr (Y)	Trp; Phe; Thr; Ser	Phe
	Val (V)	Ile; Leu; Met; Phe; Ala	Leu

Настоящее изобретение также представляет аналоги гибридного белка настоящего изобретения. Отличие между указанными аналогами и полипептидом настоящего изобретения может заключаться в наличии иной аминокислотной последовательности, может также заключаться в наличии иной модифицированной формы, не влияющей на последовательность, или же в наличии обоих указанных различий. В состав аналогов также входят аналоги, имеющие остатки, отличающиеся от естественных L-аминокислот (например, D-аминокислоты), и аналоги, имеющие аминокислоты, не встречающиеся в естественном виде, или же синтетические аминокислоты (например, β - и γ -аминокислоты). Необходимо понимать, что полипептид настоящего изобретения может являться не только одним из репрезентативных полипептидов, приведенных в качестве примера выше.

Кроме того, гибридный белок настоящего изобретения также может быть модифицирован. Модифицированные (как правило, без изменения первичной структуры) формы включают в себя: химически полученные (например, в процессе ацетилирования или карбоксилирования) *in vivo* или *in vitro* формы полипептидов. Процедуры модифицирования также включают в себя гликозилирование, используемое, например, для получения полипептидов в процессе гликозилирующей модификации во время

синтеза и переработки полипептида или же во время выполнения дальнейших шагов переработки. Такое модифицирование может быть достигнуто путем обеспечения воздействия на полипептид фермента, выполняющего гликозилирующую роль (например, гликозилазы или дегликозилазы млекопитающих). Модифицированные формы также включают в себя последовательности с аминокислотными остатками (например, фосфотирозином, фосфосерином, фосфотреонином). В их состав также входят полипептиды, подвергшиеся модификации для улучшения своих антипротеолитических свойств или оптимизации своих свойств растворимости.

Термин «полинуклеотид, кодирующий гибридный белок настоящего изобретения» может включать в себя полинуклеотид, кодирующий гибридный белок настоящего изобретения, или же полинуклеотид, включающий в себя также дополнительные кодирующие и/или некодирующие последовательности.

Настоящее изобретение также относится к вариантам вышеупомянутых полинуклеотидов, которые кодируют фрагменты, аналоги и производные полипептидов или гибридных белков, имеющих аминокислотную последовательность, аналогичную рассматриваемой в настоящем изобретении. Указанные нуклеотидные варианты включают в себя варианты замещения, варианты делеции и варианты вставки. Как известно из уровня техники, аллельный вариант представляет собой альтернативную форму полинуклеотида. Он может представлять собой замещение, делецию или вставку одного или нескольких нуклеотидов, но не изменяет функцию закодированного гибридного белка существенным образом.

Настоящее изобретение также относится к полинуклеотидам, гибридирующимся с вышеупомянутыми последовательностями и характеризующимся, по меньшей мере, 50%, предпочтительно, по меньшей мере, 70% и, более предпочтительно, по меньшей мере 80% идентичностью указанных двух последовательностей. Настоящее изобретение, в частности, относится к полинуклеотидам, гибридирующимся с полинуклеотидом настоящего изобретения при соблюдении строгих условий (или в жестких условиях). В настоящем изобретении термин «строгие условия» относится к следующему: (1) гибридизации и элюированию при низких значениях ионной силы и высоких температурах, например, 0,2×SSC, 0,1% SDS, 60°C; или (2) добавлению денатуранта в процессе гибридизации, например, 50% (по объему) формамида, 0,1% телячьей сыворотки / 0,1% фикола, 42°C, и т.д.; или (3) гибридизации, происходящей лишь тогда, когда степень идентичности двух последовательностей достигает значения 90% или более и, более предпочтительно, 95% или более.

Предпочтительно, обеспечивается наличие гибридного белка и полинуклеотидов настоящего изобретения в изолированной форме и, более предпочтительно, очищенных до достижения однородности.

Полноразмерная последовательность полинуклеотида настоящего изобретения может, как правило, быть получена с использованием способа ПЦР-амплификации, способа рекомбинации или способа искусственного синтезирования. При использовании способа ПЦР-амплификации праймеры могут разрабатываться на основе соответствующей нуклеотидной последовательности, раскрываемой в настоящем изобретении, особенно, последовательности открытой рамки считывания, а также с использованием коммерчески доступной библиотеки кДНК, или же библиотеки кДНК, подготовленной с использованием традиционного способа, известного специалистам в данной области техники, в качестве шаблона для усиления соответствующей последовательности. Если последовательность является длинной, часто оказывается необходимым выполнение двух или более процедур ПЦР-амплификации; затем

амплифицированные фрагменты сращиваются в правильном порядке.

Как только была получена соответствующая последовательность, для получения этой соответствующей последовательности в больших количествах может быть использован способ рекомбинации. Как правило, это достигается путем клонирования последовательности в вектор, а затем перенесения ее в клетку и изолирования от пролифелированных клеток-хозяев с использованием традиционных методов.

Кроме того, для синтезирования родственных последовательностей могут также использоваться способы искусственного синтезирования, особенно в случае малой длины фрагмента. (Как правило, путем синтезирования множественных малых фрагментов с последующим получением фрагментов с очень длинными последовательностями.)

На данный момент времени процедура кодирования белка (или же его фрагмента или производной) ДНК-последовательностью, согласно настоящему изобретению, реализуется полностью с использованием химического синтеза. Затем ДНК-последовательность может вводиться в различные существующие ДНК-молекулы (или же, например, векторы) и клетки, известные из уровня техники.

Предпочтительно, для получения полинуклеотида настоящего изобретения должен использоваться способ, заключающийся в применении ПЦР-технологии для амплификации ДНК/РНК. Предпочтительно, должен использоваться способ быстрой амплификации концов кДНК (RACE-cDNA), особенно в тех случаях, когда получение полноразмерных кДНК из библиотеки затруднено, а надлежащий отбор праймеров, используемых для ПЦР, может осуществляться в соответствии с информацией о последовательности, раскрытой в настоящем изобретении; указанные праймеры могут синтезироваться с применением традиционных способов. Амплифицированные фрагменты ДНК/РНК могут разделяться и очищаться с использованием традиционных методов, например, гель-электрофореза.

Вектор экспрессии

Настоящее изобретение также относится к вектору, содержащему полинуклеотид настоящего изобретения, клетку-хозяина, полученную способом генной инженерии с использованием вектора настоящего изобретения, или же последовательность кодирования гибридного белка настоящего изобретения, а также к способу получения полипептида настоящего изобретения с использованием рекомбинантной технологии.

В случае применения традиционной рекомбинантной ДНК-технологии полинуклеотидная последовательность настоящего изобретения может использоваться для обеспечения экспрессии или получения рекомбинантного гибридного белка. В общем случае, присутствуют следующие этапы:

- (1) использование полинуклеотида (или варианта) настоящего изобретения для кодирования гибридного белка настоящего изобретения, или же использование рекомбинантного вектора экспрессии, содержащего полинуклеотид, для трансформирования или преобразования подходящей клетки-хозяина;
- (2) культивирование клетки-хозяина в подходящей среде;
- (3) изолирование и очистка белка от культуральной среды или клеток.

В настоящем изобретении полинуклеотидная последовательность, кодирующая гибридный белок, может вставляться в рекомбинантный вектор экспрессии. Термин «рекомбинантный вектор экспрессии» относится к бактериальным плазмидам, бактериофагам, дрожжевым плазмидам, вирусам растительных клеток, вирусам клеток млекопитающих, таким как аденовирусы, ретровирусы, или же к другим векторам, известным из данной области техники. Могут использоваться любая плаزمиды и любой

вектор, при условии, что удастся их воспроизвести и стабилизировать в хозяине. Важной характеристикой вектора экспрессии является то, что он, как правило, содержит точку начала репликации, промотор, маркерный ген и элементы управления трансляцией.

Для конструирования вектора экспрессии, содержащего ДНК-последовательность, кодирующую гибридный белок настоящего изобретения, и соответствующие сигналы управления транскрипцией/трансляцией, могут использоваться способы, хорошо известные специалистам в данной области техники. Указанные способы включают в себя рекомбинантную ДНК-технология *in vitro*, технологию синтеза ДНК и рекомбинантную технологию *in vivo*. ДНК-последовательность может быть эффективно связана с соответствующим промотором в векторе экспрессии для направления синтеза мРНК. Типичными примерами указанных промоторов являются следующие: Лактозный промотор или триптофан промотор кишечной палочки; лямбда-фаг РL промотор; эукариотические промоторы, включая предранний промотор ЦМВ, промотор тимидинкиназы ВПГ, ранний и поздний промотор вируса SV40, ретровирусные LTR-последовательности и некоторые другие известные промоторы, способные контролировать экспрессию генов в прокариотических или эукариотических клетках или вирусах. Вектор экспрессии также включает в себя участок связывания рибосомы для инициирования трансляции, и терминатор транскрипции.

Кроме того, вектор экспрессии, предпочтительно, содержит один или несколько селективируемых маркерных генов для обеспечения фенотипических признаков для выбора трансформированных клеток-хозяев, таких как дигидрофолатредуктаза для эукариотической клеточной культуры, ген резистентности к неомицину, и зеленый флуоресцентный белок (ЗФБ) или же ген резистентности *E. Coli*. к тетрациклину или ампициллину.

Вектор, содержащий вышеупомянутую соответствующую ДНК-последовательность и соответствующий промотор или управляющую последовательность, может использоваться для трансформирования соответствующей клетки-хозяина таким способом, что она окажется способной обеспечивать экспрессию белка.

Клетка-хозяин может быть прокариотической клеткой, например, бактериальной клеткой, или эукариотической клеткой, например, дрожжевой клеткой, или же высшей эукариотической клеткой, например, клеткой млекопитающего. Типичными примерами являются следующие: Кишечная палочка, стрептомицеты; бактериальные клетки палочки мышинного тифа; грибковые клетки, такие как дрожжи, и клетки растений (например, клетки женьшеня).

Если экспрессия полинуклеотида настоящего изобретения происходит в высших эукариотических клетках, то при вставке в вектор энхансерной последовательности транскрипция будет усилена. Энхансеры представляют собой факторы ДНК, состоящие, как правило, из 10-300 базовых пар и действующие в *cis*-положении на промоторы для усиления транскрипции генов. Примеры включают в себя энхансер SV40 (состоящий из 100-270 базовых пар) на поздней стороне точки начала репликации, полиомный энхансер на поздней стороне точки начала репликации и аденовирусные энхансеры (включая подобные им).

Специалисты, обладающие обычными знаниями в данной области техники, знакомы с процедурой выбора надлежащих векторов, промоторов, энхансеров и клеток-хозяев.

Трансформирование клеток-хозяев с рекомбинантной ДНК может осуществляться с использованием традиционных методик, хорошо известных специалистам в этой области техники. Если хозяин является прокариотом (например, кишечная палочка), то сбор компетентных клеток, способных к абсорбции ДНК, может осуществляться

после фазы экспоненциального роста, а обработка их - по методу CaCl_2 . Используемые при этом этапы хорошо известны из уровня техники. Другой способ заключается в использовании MgCl_2 . При необходимости процесс трансформирования может

5 осуществляться при помощи электропорации. Если хозяин является эукариотом, могут быть выбраны следующие способы трансфекции ДНК: способ совместного осаждения фосфата кальция, традиционные механические методы, например, микроинъекция, электропорация, упаковка в липосомы и т.д.

Получаемые трансформанты могут культивироваться с использованием традиционных методов для обеспечения экспрессии полипептида, кодируемого геном настоящего изобретения. В зависимости от используемой клетки-хозяина может использоваться одна из традиционных культуральных сред. Процесс культивирования осуществляется в условиях, подходящих для выращивания клеток-хозяев. После того как в процессе роста клеток-хозяев будет достигнуто необходимое значение плотности клеток, выбранный промотор индуцируется с использованием соответствующего метода

10 (например, температурной конверсии или химической индукции), и клетки культивируются в течение определенного периода времени.

При использовании указанного выше метода рекомбинантный полипептид может экспрессироваться в клетке или на клеточной мембране или же секретироваться из клетки. При необходимости могут использоваться физические, химические и другие

20 характеристики рекомбинантного белка для его отделения и очистки с использованием различных методов разделения. Указанные методы хорошо известны специалистам в этой области техники. Примеры указанных методов включают в себя следующее (список, однако, не является исчерпывающим): традиционная ренатурационная обработка, обработка осадителем белка (способ высаливания), центрифугирование, разрушение

25 бактерий посредством осмоса, ультраобработка, ультрацентрифугирование, молекулярно-ситовая хроматография (гель-фильтрация), адсорбционная хроматография, ионообменная хроматография, высокоэффективная жидкостная хроматография (ВЭЖХ) и различные другие методы жидкостной хроматографии, а также комбинации указанных методов.

30 Синтезирование инсулина

Инсулин является четко выраженным пептидом с известной аминокислотной последовательностью и структурными характеристиками. Он является белком, насчитывающим 51 аминокислотный остаток, с двумя аминокислотными цепями. Гормон содержит две независимые пептидные цепи, цепь А (21 аминокислота) и цепь

35 В (30 аминокислот). Имеются 6 цистеиновых остатков в двух аминокислотных цепях, и каждая из цепей имеет два цистеиновых остатка, соединенных друг с другом посредством дисульфидных связей. С точки зрения статистики существует 15 возможных вариантов образования дисульфидных связей в молекуле человеческого инсулина. Однако в человеческом инсулине, проявляющем биологическую активность,

40 присутствует лишь один из указанных 15 возможных вариантов, и указанные дисульфидные связи являются следующими: 1) А6-А11; 2) А7-В7; 3) А20-В19. Проинсулин является биологическим предшественником инсулина. Он представляет собой одноцепной пептид, образующийся путем связывания цепи А и цепи В пептидом С. Две пептидные цепи инсулина соединяются дисульфидными связями (Фигура 1).

Инсулин представляет собой белковый гормон, выделяемый островными клетками поджелудочной железы, стимулируемыми эндогенными веществами, такими как глюкоза. Первым секретом островных клеток поджелудочной железы является проинсулин, длинноцепочечный полипептид, состоящий из 84 аминокислот. Проинсулин отщепляется

от средней части (цепи С) проинсулина под действием особых ферментов PC1 и PC2, осуществляющих преобразование протеазы в проинсулин, а также под действием карбоксипептидазы E (CPE), в то время как карбоксильная часть (цепь А) и аминная часть (цепь В) проинсулина соединяются дисульфидными связями с образованием
5 инсулина. Зрелый инсулин содержится в секреторных пузырьках островных клеток поджелудочной железы и координируется при помощи ионов цинка для образования гексамеров. При внешней стимуляции инсулин высвобождается в кровь вместе с секреторными пузырьками и оказывает свое физиологическое воздействие. Кровь
10 пациентов с диабетом 1-го типа теряет способность к регулированию уровня глюкозы вследствие утери островными клетками поджелудочной железы способности к образованию инсулина.

На данный момент времени существуют два способа получения различных типов коммерческого рекомбинантного человеческого инсулина - способ «комбинирования цепей» и «проинсулиновый» способ. При использовании способа «комбинирования
15 цепей» две пептидные цепи, составляющие инсулин - цепь А и цепь В - синтезируются отдельно с применением биологического рекомбинирования, а затем цепь А и цепь В смешиваются для образования дисульфидных связей для формирования биологически активного человеческого инсулина. Тем не менее, эффективность процесса прямого смешивания двух пептидных цепей для образования биологически активного
20 человеческого инсулина является относительно невысокой, и окончательный выход вещества составляет лишь около 7%. Данный метод постепенно вытесняется вторым, «проинсулиновым» способом. При использовании «проинсулинового» способа проинсулин, состоящий из инсулиновых цепей В, С и А, сначала экспрессируется в *E. coli* или дрожжах, а затем, после очистки, ренатурируется *in vitro*. Ренатурированный
25 проинсулин затем подвергается гидролизу и перерабатывается с использованием трипсина и карбоксипептидазы В, для получения человеческого инсулина, характеризующегося естественной активностью. При использовании «проинсулинового» способа трипсин специфически распознает лизин и аргинин в белке и расщепляет пептидную связь у С-конца лизина и аргинина. По конформационным причинам аргинин
30 в позиции В22 проинсулина не гидролизуется трипсином. Однако трипсин распознает и гидролизует лизин в позиции В29 инсулина, образуя, как следствие, инсулин в качестве побочного продукта (инсулин DesB30), в котором исключен треонин в позиции В30. Для снижения выработки инсулина DesB30 должны строго контролироваться количество трипсина и время реакции. Тем не менее, определенное количество инсулина
35 DesB30 будет все же произведено. Вследствие того, что инсулин DesB30 и обычный инсулин отличаются лишь наличием или отсутствием треонина, их разделение является весьма трудной задачей. Для промышленного разделения инсулина используется технология высокоэффективной жидкостной хроматографии, проводимой в крупных масштабах; однако в процессе разделения обычного инсулина и инсулина DesB30 с
40 использованием указанного способа образуется большое количество промышленных отходов, что существенно повышает стоимость производства существующего рекомбинантного человеческого инсулина.

Настоящее изобретение представляет способ получения человеческого инсулина с корректно соединенными цистеиновыми мостиками. В процессе своей реализации
45 данный метод требует меньшего количества этапов, благодаря чему можно добиться большей производительности процесса выработки человеческого инсулина.

Химическим способом был синтезирован фрагмент ДНК; при этом структура кодирования 2 β -TEV-R-MiniINS фрагмента ДНК состояла из аминокислотной

последовательности, показанной на SEQ ID NO.: 21. Указанный фрагмент ДНК клонировался в вектор бактериальной экспрессии, регулируемый промотором araBAD. Вектор экспрессии, содержащий 2 β -TEV-R-MiniINS, преобразовывался в штамм Top10 E. coli, а рекомбинантные клетки культивировались в среде Лурия-Бертани, содержащей примесные элементы. Гибридный белок 2 β -TEV-R-MiniINS извлекался из внутриклеточных включений и складывался при определенных условиях, что приводило к тому, что образованная дисульфидная связь в сложенном гибридном белке 2 β -TEV-R-MiniINS была идентична дисульфидной связи в сложенном человеческом проинсулине, т.е., формировались дисульфидные связи A6-A11, A7-B7 и A20-B19. Разделенный корректно сложенный гибридный белок 2 β -TEV-R-MiniINS перерабатывался с использованием трипсина и карбоксипептидазы В для образования правильно сложенного трет-бутоксикарбонильного человеческого инсулина. Трет-бутоксикарбонильный человеческий инсулин очищался до степени чистоты 90% с использованием процесса гидрофобной хроматографии. Затем осуществлялась депротекция трет-бутоксикарбонильной кислоты для образования корректно сложенного человеческого инсулина, который затем очищался в процессе ВЭЖХ в режиме обращения фазы. Для определения свойств полученного таким способом чистого человеческого инсулина использовались процессы N-терминального анализа последовательностей, определения молекулярной массы и пептидного картирования.

По сравнению с предшествующим уровнем техники основными преимуществами настоящего изобретения являются следующие:

- 1) Элемент гибридного белка настоящего изобретения, способствующий экспрессии белка, может также способствовать складыванию гибридного белка.
- 2) Элемент гибридного белка настоящего изобретения, способствующий экспрессии белка, может повышать степень растворимости растворимого гибридного белка и снижать степень его межмолекулярного взаимодействия, так что появляется возможность складывания гибридного белка при высоких значениях концентраций, что обеспечивает промышленную значимость процесса.
- 3) В процессе подготовки целевого пептида отсутствует необходимость в проведении бромцианового расщепления, окислительного гидролиза сульфитов и связанных с ними этапов очистки.
- 4) В процессе подготовки целевых пептидов отсутствует необходимость в использовании тиолов или гидрофобных адсорбирующих смол в высоких концентрациях.
- 5) Обеспечивается защита целевых пептидов от внутриклеточной деструкции микробных хозяев.
- 6) Целевой белок характеризуется высокими значениями удельной плотности (повышенным коэффициентом слияния). Элементы (X)_n или A1-(X)_n гибридного белка могут перерабатываться в малые фрагменты с участием протеазы. В сравнении с целевым белком обеспечивается широкое различие значений молекулярной массы и существенная простота разделения.
- 7) Гибридный белок настоящего изобретения может способствовать экспрессии целевого пептида; существенно повышаются уровень экспрессии и выход целевого пептида.
- 8) Гибридный белок настоящего изобретения весьма подходит для обеспечения экспрессии целевых пептидов с участием искусственных аминокислот и, очевидно, может способствовать складыванию целевых пептидов с участием искусственных аминокислот.

Настоящее изобретение будет дополнительно разъяснено ниже с демонстрацией

конкретных вариантов осуществления. Следует понимать, что указанные варианты осуществления используются только для иллюстрации настоящего изобретения, а не для ограничения объема настоящего изобретения. Экспериментальные способы без указания конкретных условий их осуществления в следующих примерах обычно реализуются в общепринятых условиях, например, в условиях, описанных в работе Сэмбрука (Sambrook) с соавторами «Молекулярное клонирование»: Лабораторное руководство (Нью-Йорк: Cold Spring Harbor Laboratory Press, 1989), или же в соответствии с условиями, предлагаемыми производителем. Если не указано иное, проценты и части представляют собой массовые проценты и массовые части.

Указанные варианты осуществления используют в качестве примера целевого пептида человеческий инсулин, однако данный пример не является ограничивающим.

Пример 1. Конструирование вектора экспрессии A1-u4-u5-TEV-R-MiniINS

Структура экспрессии A1-u4-u5-TEV-R-MiniINS содержит ген, кодирующий белок человеческого инсулина, который слит с С-концом структуры A1-u4-u5. Линкерный пептид, соединяющий A1-u4-u5 и инсулиновый белок MiniINS, представляет собой октапептид Glu-Asn-Leu-Tyr-Phe-Gln-Gly-Arg. Указанный октапептид может гидролизироваться трипсином у карбоксильного конца Arg, и может также гидролизироваться TEV-протеазой между Gln и Gly. ДНК-последовательность октапептида является кодон-оптимизированной, что позволяет реализовать высокоуровневую экспрессию функционального белка в *E. coli*.

Все использованные фрагменты гибридного рекомбинантного белка синтезировались при помощи Gen Script и загружались в вектор pUC57. Структура A1-u4-u5-TEV-R-MiniINS удалялась из синтетического вектора pUC57-A1-u4-u5-TEV-R-MiniINS с использованием рестрикционных ферментов NcoI и XhoI; в то же самое время вектор экспрессии pBAD/His A(KanaR) вырезался с использованием тех же самых рестрикционных ферментов NcoI и XhoI. Переработанные продукты разделялись сначала с использованием процедуры электрофореза в агарозном геле, а затем - набора для восстановления ДНК в агарозном геле для экстракции; наконец, два ДНК-фрагмента соединялись с использованием ДНК-лигазы T4. Продукт лигирования трансформировался в клетки *E. coli* Top 10, а затем трансформированные клетки культивировались на агаровой среде Луриа-Бертани, содержащей канамицин в концентрации 50 мкг/мл (дрожжевой пептон 10 г/л, порошок дрожжевого экстракта 5 г/л, NaCl 10 г/л, агар 1,5%), в течение ночи. Были выбраны 3 живых колонии; в течение ночи производилось их культивирование в 5 мл жидкой среды Луриа-Бертани (дрожжевой пептон -10 г/л, порошок дрожжевого экстракта - 5 г/л, NaCl - 10 т/л), содержащей 50 мкг/мл канамицина, и с использованием для экстракции плазмиды мини-набора Plasmid. Затем экстрагированная плазмида подвергалась секвенированию. Полученной в конце всех этих манипуляций плазмиде было присвоено название pBAD-A1-u4-u5-TEV-R-MiniINS. Карта плазмиды и схема гибридного белка показаны на Фиг. 2 и 3, соответственно.

Пример 2. Конструирование структуры экспрессии A1-u4-u5-TEV-R-GLP1

Порядок конструирования и синтеза структуры экспрессии pUC57-A1-u4-u5-TEV-R-GLP1 описан в Примере 1. В то же самое время для вырезания вектора экспрессии pBAD/His A(KanaR) использовались рестрикционные ферменты NcoI и XhoI, и переработанные продукты разделялись с использованием процедуры электрофореза в агарозном геле, а затем - с использованием набора для восстановления ДНК в агарозном геле для экстракции; наконец, два ДНК-фрагмента соединялись с использованием ДНК-лигазы T4. Продукт лигирования химически трансформировался

в клетки *E. coli* Top10, а затем трансформированные клетки культивировались на агаровой среде Лурия-Бертани, содержащей канамицин в концентрации 50 мкг/мл (дрожжевой пептон - 10 г/л, порошок дрожжевого экстракта - 5 г/л, NaCl - 10 г/л, агар - 1,5%) в течение ночи. Были выбраны 3 живых колонии; в течение ночи производилось их культивирование в 5 мл жидкой среды Лурия-Бертани (дрожжевой пептон - 10 г/л, порошок дрожжевого экстракта - 5 г/л, NaCl - 10 г/л), содержащей 50 мкг/мл канамицина, и с использованием для экстракции плазмиды мини-набора Plasmid. Затем экстрагированная плаزمида подвергалась секвенированию с использованием олигонуклеотидного праймера для секвенирования 5'-ATGCCATAGCATTTTATCC-3' для подтверждения корректности вставки. Полученной плазмиде было присвоено название pBAD-A 1-u4-u5-TEV-R-GLP1.

Пример 3. Экспрессия, разделение и очистка гибридного белка A1-u4-u5-TEV-R-MiniINS

Для обеспечения экспрессии фрагмента слияния A1-u4-u5-TEV-R-MiniINS, содержащего аминокислотную последовательность SEQ ID NO.: 21. Плазмиды pBAD-A1-u4-u5-TEV-R-MiniINS, подтвержденная посредством секвенирования, и плазмиды аминоксил-тРНК-синтетазы для пирролизина pEvol-pylRs-pylT (где плазмиды аминоксил-тРНК-синтетазы для пирролизина pEvol-pylRs-pylT использовались для обеспечения экспрессии аминоксил-тРНК-синтетазы и тРНК, как показано в Примере 1 Заявки на патент №2011103886241) совместно трансформировались в штамм Top10 *E. coli*. Трансформационный раствор размещался на ночь на агаровой среде Лурия-Бертани, содержащей канамицин в концентрации 25 мкг/мл и хлорамфеникол в концентрации 17 мкг/мл. Выбиралась одиночная колония; в течение ночи производилось ее культивирование в жидкой среде Лурия-Бертани, содержащей канамицин в концентрации 25 мкг/мл и хлорамфеникол в концентрации 17 мкг/мл. Затем ночная культура вносилась в 100 мл среды Лурия-Бертани (жидкая среда Лурия-Бертани: дрожжевой пептон - 12 г/л, порошок дрожжевого экстракта - 24 г/л, глицерол - 4 мл/л, партнерский белок киназы 4%, пеноподавляющее средство 0,3%. Раствор партнерского белка киназы: KH_2PO_4 23,1 г/л, безводный K_2HPO_4 125,4 г/л), содержащий канамицин в концентрации 25 мкг/мл и хлорамфеникол в концентрации 17 мкг/мл; инкубирование производилось при 37°C до достижения уровня значений 2~4 оптической плотности OD_{600} . Затем к среде добавлялся 25% раствор арабинозы до достижения окончательной концентрации 0,25%, и добавлялось 0,1 М трет-бутоксикарбонил-лизинового раствора (трет-бутоксикарбонил-лизина, где структура трет-бутоксикарбонила показана на Фиг. 4) до достижения окончательной концентрации 5 нМ для индуцирования экспрессии гибридного белка. Раствор культуры непрерывно культивировался в течение 16-20 часов, а затем осуществлялась процедура сбора центрифугированием (10000 об/мин, 5 мин, 4°C).

Экспрессия гибридного белка A1-u4-u5-TEV-R-MiniINS осуществлялась в форме нерастворимых «внутриклеточных включений». Для высвобождения внутриклеточных включений клетки *E. coli* разрушались с использованием гомогенизатора высокого давления. Нуклеиновые кислоты, клеточный детрит и растворимые белки удалялись посредством центрифугирования при 10000g. Внутриклеточные включения, содержащие гибридный белок A1-u4-u5-TEV-R-MiniINS, вымывались чистой водой; образующийся в результате осадок внутриклеточных включений использовался в качестве сырья для складывания. Окончательный уровень экспрессии гибридного белка составлял 14 г/л в пересчете на ферментативный бульон. Схема ДСН-ПААГ-электрофореза гибридного белка показана на Фиг. 5. Из Фиг. 5 можно видеть, что целевой белок, экспрессируемый

гибридным белком, может экспрессироваться полностью без разрыва, и гибридный белок, вставляемый в целевой искусственный аминокислотный белок, также может экспрессироваться в целости, без разрыва.

Для повторного складывания гибридного белка внутриклеточные включения растворяются в 7,5 М раствора мочевины (рН 10,5), содержащего 2-10 мМ меркаптоэтанола; так что общая концентрация белка после растворения составляла 10-25 мг/мл. Образец разбавлялся 5-10 раз; процесс традиционного складывания проводился в течение 16-30 часов при температуре 4-8°C и значении рН от 10,5 до 11,7. При температуре 18-25°C значение рН поддерживалось на уровне 8,0-9,5; для перерабатывания гибридного белка в течение 10-20 часов использовались трипсин и карбоксипептидаза В; затем добавлялся 0,45 М раствор сульфата аммония для прекращения реакции ферментативного гидролиза. Результаты ВЭЖХ анализа в режиме обращения фазы показывают, что выход продукта на указанном этапе ферментативного гидролиза составляет более 90%. Аналог инсулина, получаемый после переработки с участием трипсина и карбоксипептидазы В, носит название «трет-бутоксикарбонильный человеческий инсулин». Процедура ферментативного гидролиза трет-бутоксикарбонильного человеческого инсулина в указанных выше условиях является невозможной. Образец осветлялся в процессе мембранной фильтрации, с использованием 0,45 мМ раствора сульфата аммония в качестве буфера; предварительная очистка трет-бутоксикарбонильного человеческого инсулина осуществлялась с использованием процесса гидрофобной хроматографии. Степень чистоты электрофореза в полиакриламидном геле с использованием натрия додецилсульфата достигает 90%, а окончательный выход трет-бутоксикарбонильного человеческого инсулина на 1 литр ферментативного бульона составляет около 2,1 г; получаемый трет-бутоксикарбонильный человеческий инсулин анализировался с использованием метода масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF); было найдено, что фактическое значение молекулярной массы соответствовало теоретическому значению 5907,7 Да. Сбор образцов осуществлялся с применением метода гидрофобной хроматографии; для обеспечения протекания реакции снятия защиты трет-бутоксикарбонильного человеческого инсулина добавлялась соляная кислота. Для доведения значения рН до уровня 2,8 ~ 3,2 (для прекращения реакции) добавлялся раствор гидроксида натрия. После проведения двух этапов обращенно-фазовой хроматографии высокого давления выход рекомбинантного человеческого инсулина составлял более 85%. Окончательный выход рекомбинантного человеческого инсулина на 1 литр ферментативного бульона составляет около 700 мг.

Пример 4. Конструирование структуры экспрессии pBAD-A1-u8-TEV-R-MiniINS и экспрессия, разделение и очистка гибридного белка A1-и8-TEV-R- MiniINS

Используется тот же самый экспериментальный способ, что и в Примере 1-3; осуществляется замена A1-u4-u5 (SEQ ID NO.: 16) в A1-u4-u5-TEV-R-MiniINS последовательностью A1-u8 (SEQ ID NO.: 23). Была сконструирована рекомбинантная плазида, названная pBAD-A1-u8-TEV-R-MiniINS; экспрессия гибридного белка A1-u8-TEV-R-MiniINS обеспечивалась согласно указанному способу (Фиг. 5). Окончательный выход трет-бутоксикарбонильного человеческого инсулина составляет около 1,9 г/л; получаемый трет-бутоксикарбонильный человеческий инсулин анализировался с использованием метода масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF). Результаты показывают соответствие фактического значения молекулярной массы теоретическому.

Пример 5. Конструирование структуры экспрессии pBAD-A1-u3-u5-u7-TEV-R-MiniINS

и экспрессия, разделение и очистка гибридного белка A1-u3-u5-u7-TEV-R-MiniINS

Используется тот же самый экспериментальный способ, что и в Примере 1-3; осуществляется замена A1-u4-u5 (SEQ ID NO.: 16) в A1-u4-u5-TEV-R-MiniINS последовательностью A1-u3-u5-u7 (SEQ ID NO.: 15). Была сконструирована рекомбинантная плазмида, названная pBAD-A1-u3-u5-u7-TEV-R-MiniINS; экспрессия гибридного белка A1-u3-u5-u7-TEV-R-MiniINS обеспечивалась согласно указанному способу. Окончательный выход трет-бутоксикарбонильного человеческого инсулина составляет около 2,0 г/л; получаемый трет-бутоксикарбонильный человеческий инсулин анализировался с использованием метода масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF). Результаты показывают соответствие фактического значения молекулярной массы теоретическому.

Пример 6. Конструирование структуры экспрессии pBAD-A1-u5-EK-u6-TEV-R-MiniINS и экспрессия, разделение и очистка гибридного белка A1-u5-EK-u6-TEV-R-MiniINS

Используется тот же самый экспериментальный способ, что и в Примере 1-3; осуществляется замена A1-u4-u5 (SEQ ID NO.: 16) в A1-u4-u5-TEV-R-MiniINS последовательностью A1-u5-EK-u6 (SEQ ID NO.: 22), где EK представляет собой участок рестрикции фермента EK DDDDK (SEQ ID NO.: 25). Была сконструирована рекомбинантная плазмида, названная pBAD-A1-u5-EK-u6-TEV-R-MiniINS; экспрессия гибридного белка A1-u5-EK-u6-TEV-R-MiniINS обеспечивалась согласно указанному способу. Окончательный выход трет-бутоксикарбонильного человеческого инсулина составляет около 1,9 г/л; получаемый трет-бутоксикарбонильный человеческий инсулин анализировался с использованием метода масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF). Результаты показывают соответствие фактического значения молекулярной массы теоретическому.

Все документы, упомянутые в настоящем изобретении, включены в настоящий документ в виде ссылок, как если бы они были включены при помощи ссылок по отдельности. Кроме того, необходимо понимать, что после ознакомления с вышеуказанными идеями настоящего изобретения специалисты в этой области техники могут вносить в изобретение различные изменения и модификации, а также что указанные эквиваленты равным образом подпадают под сферу действия формулы изобретения, прилагаемой к данной заявке.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> НИНГВО КУНПЕНГ ВИОТЕКХ КО., ЛТД.

<120> ГИБРИДНЫЙ БЕЛОК, СОДЕРЖАЩИЙ ФРАГМЕНТЫ ФЛУОРЕСЦЕНТНЫХ БЕЛКОВ, И ЕГО ПРИМЕНЕНИЕ

<130> P2020-0086

<150> CN201910210102.9

<151> 2019-03-19

<160> 31

<170> PatentIn version 3.5

<210> 1

<211> 13

<212> ПРТ

<213> искусственная последовательность

<220>

<223> u1

<400> 1

Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly

1

5

10

<210> 2
 <211> 14
 <212> ПРТ
 <213> искусственная последовательность
 5 <220>
 <223> u2
 <400> 2
 His Lys Phe Ser Val Arg Gly Glu Gly Glu Gly Asp Ala Thr
 1 5 10
 10 <210> 3
 <211> 10
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 15 <223> u3
 <400> 3
 Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr
 1 5 10
 20 <210> 4
 <211> 11
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> u4
 25 <400> 4
 Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp
 1 5 10
 <210> 5
 <211> 13
 30 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> u5
 <400> 5
 35 Thr Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp
 1 5 10
 <210> 6
 <211> 13
 <212> ПРТ
 40 <213> искусственная последовательность
 <220>
 <223> u6
 <400> 6
 Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe
 45 1 5 10
 <210> 7
 <211> 10
 <212> ПРТ

<213> искусственная последовательность
 <220>
 <223> u7
 <400> 7
 5 His Asn Val Tyr Ile Thr Ala Asp Lys Gln
 1 5 10
 <210> 8
 <211> 14
 <212> ПРТ
 10 <213> искусственная последовательность
 <220>
 <223> u8
 <400> 8
 Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Val Glu Asp
 15 1 5 10
 <210> 9
 <211> 14
 <212> ПРТ
 <213> искусственная последовательность
 20 <220>
 <223> u9
 <400> 9
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly
 1 5 10
 25 <210> 10
 <211> 12
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 30 <223> u10
 <400> 10
 His Tyr Leu Ser Thr Gln Ser Val Leu Ser Lys Asp
 1 5 10
 <210> 11
 35 <211> 13
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> u11
 40 <400> 11
 His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile
 1 5 10
 <210> 12
 <211> 12
 45 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> A1 или A2

<400> 12
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
 1 5 10
 <210> 13
 5 <211> 241
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> GFP
 10 <400> 13
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly
 20 25
 15 Glu Gly Glu Gly Asp Ala Thr Asn Gly Lys Leu Thr Leu Lys Phe Ile 45
 35 40
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 20 65 70 75
 Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90
 Arg Thr Ile Ser Phe Lys Asp Asp Gly Thr Tyr Lys Thr Arg Ala Glu
 100 105 11
 25 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Phe Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn
 30 145 150 155
 Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Val Glu Asp Gly Ser
 165 170 175
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 19
 35 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Val Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Ala Gly
 40 225 230 235
 Ser
 <210> 14
 <211> 28
 <212> ПРТ
 45 <213> искусственная последовательность
 <220>
 <223> u8-u9
 <400> 14

RU 2 801 248 C2

Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Val Glu Asp Val Gln
1 5 10
Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly
20 25

5 <210> 15
<211> 45
<212> ПРТ
<213> искусственная последовательность
<220>

10 <223> A1-u3-u5-u7
<400> 15
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Lys Leu Thr Leu
1 5 10
Lys Phe Ile Cys Thr Thr Thr Tyr Lys Thr Arg Ala Glu Val Lys Phe
15 20 25
Glu Gly Asp His Asn Val Tyr Ile Thr Ala Asp Lys Gln
35 40

<210> 16
<211> 36

20 <212> ПРТ
<213> искусственная последовательность
<220>
<223> A1-u4-u5
<400> 16

25 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Tyr Val Gln Glu
1 5 10
Arg Thr Ile Ser Phe Lys Asp Thr Tyr Lys Thr Arg Ala Glu Val Lys
20 25
Phe Glu Gly Asp
35

30 <210> 17
<211> 26
<212> ПРТ
<213> искусственная последовательность
<220>

35 <223> A1-u9
<400> 17
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Gln Leu Ala
1 5 10
40 Asp His Tyr Gln Gln Asn Thr Pro Ile Gly
20 25

<210> 18
<211> 8
<212> ПРТ

45 <213> искусственная последовательность
<220>
<223> первый линкерный пептид и/или второй линкерный пептид
<400> 18

Glu Asn Leu Tyr Phe Gln Gly Arg
 1 5
 <210> 19
 <211> 86
 5 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> проинсулин
 <400> 19
 10 Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
 20 25
 Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro
 15 35 40 45
 Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys
 50 55 60
 Arg Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln
 65 70 75
 20 Leu Glu Asn Tyr Cys Asn
 85
 <210> 20
 <211> 52
 <212> ПРТ
 25 <213> искусственная последовательность
 <220>
 <223> проинсулин
 <400> 20
 30 Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Gly
 20 25
 Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu
 35 40 45
 35 Asn Tyr Cys Asn
 50
 <210> 21
 <211> 96
 <212> ПРТ
 40 <213> искусственная последовательность
 <220>
 <223> гибридный белок
 <400> 21
 45 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Tyr Val Gln Glu
 1 5 10
 Arg Thr Ile Ser Phe Lys Asp Thr Tyr Lys Thr Arg Ala Glu Val Lys
 20 25
 Phe Glu Gly Asp Glu Asn Leu Tyr Phe Gln Gly Arg Phe Val Asn Gln

<220>
 <223> участок рестрикции фермента ЕК
 <400> 25
 Asp Asp Asp Asp Lys
 5 1 5
 <210> 26
 <211> 49
 <212> ПРТ
 <213> искусственная последовательность
 10 <220>
 <223> A1-u1-u2-u3
 <400> 26
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10
 15 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly
 20 25
 Glu Gly Glu Gly Asp Ala Thr Lys Leu Thr Leu Lys Phe Ile Cys Thr
 35 40 45
 Thr
 20 <210> 27
 <211> 60
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 25 <223> A1-u1-u2-u3-u4
 <400> 27
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly
 30 20 25
 Glu Gly Glu Gly Asp Ala Thr Lys Leu Thr Leu Lys Phe Ile Cys Thr
 35 40 45
 Thr Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp
 50 55 60
 35 <210> 28
 <211> 46
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 40 <223> A1-u3-u4-u6
 <400> 28
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Lys Leu Thr Leu
 1 5 10
 Lys Phe Ile Cys Thr Thr Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys
 45 20 25 3
 Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe
 35 40 45
 <210> 29

<211> 45
 <212> ПРТ
 <213> искусственная последовательность
 <220>
 5 <223> A1-u4-u7-u10
 <400> 29
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Tyr Val Gln Glu
 1 5 10
 Arg Thr Ile Ser Phe Lys Asp His Asn Val Tyr Ile Thr Ala Asp Lys
 10 20 25 30
 Gln His Tyr Leu Ser Thr Gln Ser Val Leu Ser Lys Asp
 35 40 45
 <210> 30
 <211> 55
 15 <212> ПРТ
 <213> искусственная последовательность
 <220>
 <223> A1-u3-u4-u7-u10
 <400> 30
 20 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Lys Leu Thr Leu
 1 5 10
 Lys Phe Ile Cys Thr Thr Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys
 20 25 30
 Asp His Asn Val Tyr Ile Thr Ala Asp Lys Gln His Tyr Leu Ser Thr
 25 35 40 45
 Gln Ser Val Leu Ser Lys Asp
 50 55
 <210> 31
 <211> 20
 30 <212> ДНК
 <213> искусственная последовательность
 <220>
 <223> Олигонуклеозиды
 <400> 31
 35 atgccatagc atttttatcc 20

(57) Формула изобретения

1. Гибридный белок для экспрессии и получения целевого полипептида, содержащий структуру, показанную в Формуле I:

40 (P1-L1)_s-A1-(X)_n-A2-(L2-P2)_t (I)

где

«-» обозначает пептидную связь или линкерный пептид,

каждый P1 независимо является первым целевым пептидом;

каждый P2 независимо является вторым целевым пептидом;

45 каждый L1 независимо является отсутствующим или первым линкерным пептидом;

каждый L2 независимо является отсутствующим или вторым линкерным пептидом;

A1 является отсутствующим или сигнальным пептидом;

A2 является отсутствующим или сигнальным пептидом;

каждый X независимо является одиночным β-складчатым блоком флуоресцентного белка;

n является положительным целым числом от 1 до 8;

s может принимать значения 0, 1 или 2;

5 t может принимать значения 0, 1 или 2; и

дополнительным условием является неравенство s и t нулю одновременно;

и (X)n является u8, u9, u2-u3, u4-u5, u8-u9, u1-u2-u3, u2-u3-u4, u3-u4-u5, u5-u6-u7, u8-u9-u10, u9-u10-u11, u3-u5-u7, u3-u4-u6, u4-u7-u10, u6-u8-u10, u1-u2-u3-u4, u2-u3-u4-u5, u3-u4-u3-u4, u3-u5-u7-u9, u5-u6-u7-u8, u1-u3-u7-u9, u2-u2-u7-u8, u7-u2-u5-u11, u3-u4-u7-u10, u1-I-u2, u1-I-u5, u2-I-u4, u3-I-u8, u5-I-u6 или u10-I-u11;

I - гибкий линкер, содержащий сайт рестрикции;

где аминокислотные последовательности u1-u11 показаны ниже:

15	u1	VPILVELDGDVNG (SEQ ID NO.: 1)
	u2	HKFSVRGEGEGDAT (SEQ ID NO.: 2)
	u3	KLTLKFICTT (SEQ ID NO.: 3)
	u4	YVQERTISFKD (SEQ ID NO.: 4)
	u5	TYKTRAEVKFEGD (SEQ ID NO.: 5)
	u6	TLVNRIELKGIDF (SEQ ID NO.: 6)
	u7	HNVYITADKQ (SEQ ID NO.: 7)
20	u8	GIKANFKIRHNVED (SEQ ID NO.: 8)
	u9	VQLADHYQQNTPIG (SEQ ID NO.: 9)
	u10	HYLSTQSVLSKD (SEQ ID NO.: 10)
	u11	HMVLLFEVTAAGI (SEQ ID NO.: 11);

P1 и P2 каждый независимо выбирают из: белка-предшественника человеческого инсулина, белка-предшественника инсулина лизпро, белка-предшественника инсулина гларгина, глюкагоноподобного пептида и его производных эксенатида и лираглутида, инсулина, кальцитонина, тедуглутида.

2. Гибридный белок по п.1, где первый целевой пептид и/или второй целевой пептид представляют собой белок с искусственной аминокислотой.

3. Гибридный белок по п.1, где A1-(X)n имеет последовательность аминокислот, показанную в SEQ ID NO.: 14, 15, 16, 17, 22, 23, 24, 26, 27, 28, 29 или 30.

4. Гибридный белок по п.1, где между любыми двумя β-складчатыми блоками присутствует гибкий линкер I, и гибкий линкер I содержит участок рестрикции.

5. Изолированный полинуклеотид, кодирующий гибридный белок по п.1.

35 6. Вектор для экспрессии гибридного белка по п.1, который включает полинуклеотид по п.5.

7. Клетка-хозяин для экспрессии гибридного белка по п.1, содержащая вектор по п.6.

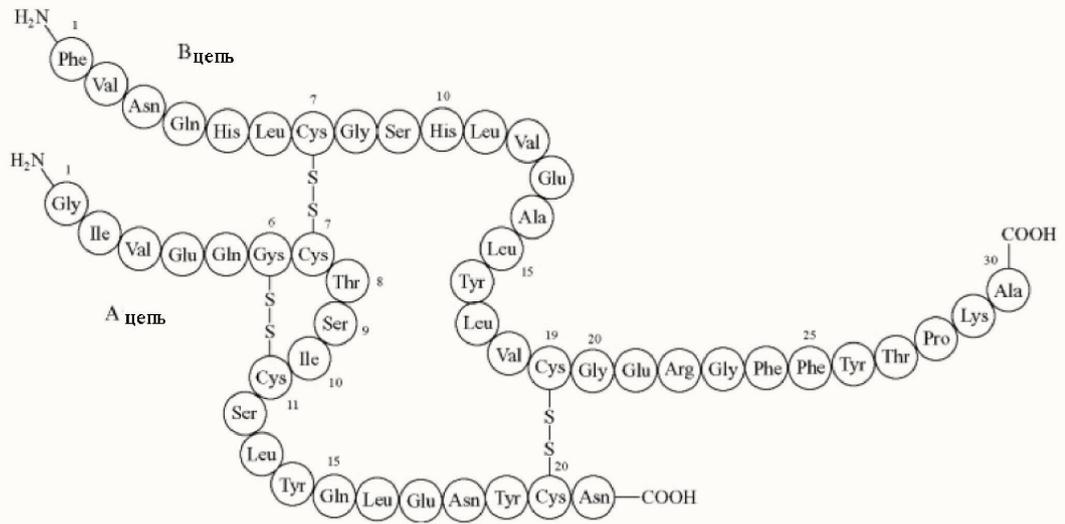
40 8. Клетка-хозяин для экспрессии гибридного белка по п.1, в которой экзогенный полинуклеотид по п.5 интегрирован в хромосому клетки-хозяина.

9. Способ получения гибридного белка по п.1, содержащий следующие этапы:

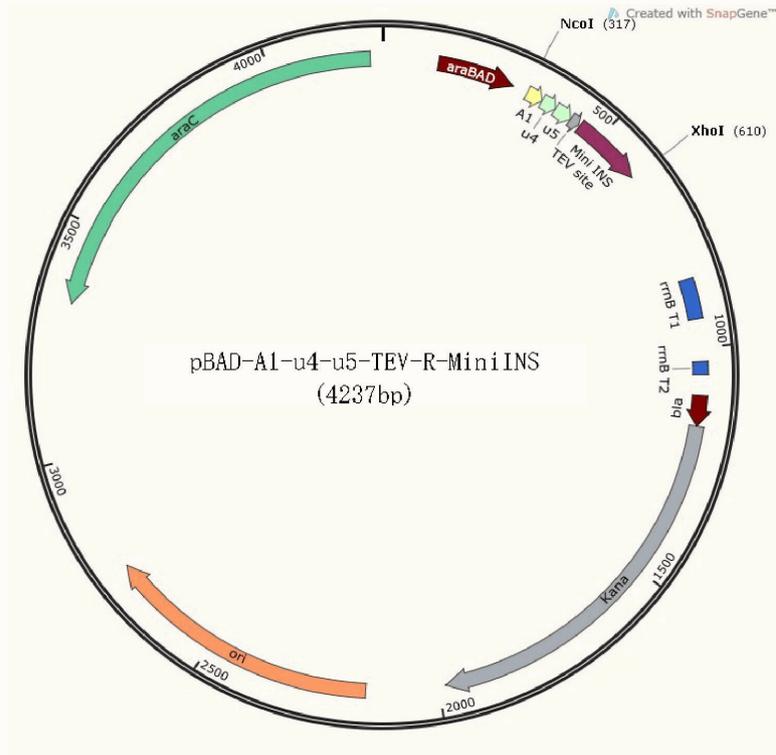
(а) в условиях, подходящих для получения гибридного белка по п.1, культивируют клетку-хозяина по любому из пп.7-8, получая, таким образом, гибридный белок по п.1.

45 10. Применение клетки-хозяина по любому из пп.7-8 для экспрессии и получения целевого полипептида по п.1.

1

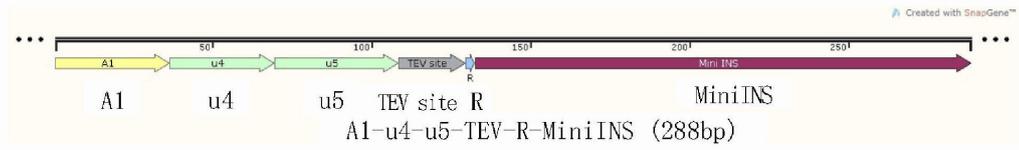


Фигура 1

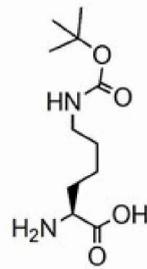


Фигура 2

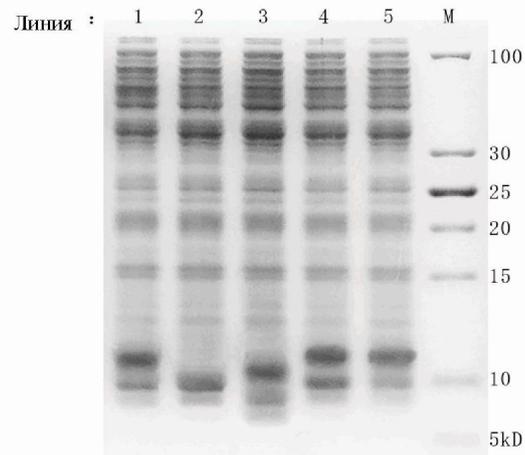
2



Фигура 3



Фигура 4



Фигура 5