



Urząd Patentowy
Rzeczypospolitej Polskiej

(21) Numer zgłoszenia: **415888**

(51) Int.Cl.
C07K 14/62 (2006.01)
C12P 21/06 (2006.01)

(22) Data zgłoszenia: **22.01.2016**

(54)

Sposób wytwarzania insuliny i jej pochodnych

(43) Zgłoszenie ogłoszono:

31.07.2017 BUP 16/17

(45) O udzieleniu patentu ogłoszono:

02.11.2021 WUP 31/21

(73) Uprawniony z patentu:

**SIEĆ BADAWCZA ŁUKASIEWICZ – INSTYTUT
CHEMII PRZEMYSŁOWEJ IMIENIA PROFESORA
IGNACEGO MOŚCICKIEGO, Warszawa, PL**

(72) Twórca(y) wynalazku:

**PIOTR BOROWICZ, Warszawa, PL
ANDRZEJ PŁUCIENNICZAK, Warszawa, PL
GRAŻYNA PŁUCIENNICZAK, Warszawa, PL
IWONA SOKOŁOWSKA, Warszawa, PL
AGNIESZKA ROMANIK-CHRUŚCIELEWSKA,
Józefów, PL
NATALIA ŁUKASIEWICZ, Warszawa, PL
MARCIN ZIELIŃSKI, Warszawa, PL
JAROSŁAW ANTOSIK, Skierniewice, PL
AGNIESZKA SOBOLEWSKA, Warszawa, PL
DIANA MIKIEWICZ, Warszawa, PL
ANNA WÓJTOWICZ-KRAWIEC, Warszawa, PL
PIOTR BARAN, Warszawa, PL
ANNA BIERCZYŃSKA-KRZYSIK, Warszawa, PL
MICHAŁ ODROWĄŻ-SYPNIEWSKI, Warszawa, PL
BOŻENA TEJCHMAN-MAŁECKA, Warszawa, PL
DOROTA STADNIK, Warszawa, PL
JACEK STADNIK, Warszawa, PL
WERONIKA SURMACZ-CHWEDORUK,
Nowe Iganie, PL
JOANNA ZIELIŃSKA, Warszawa, PL**

(74) Pełnomocnik:

rzecz. pat. Jarosław Kulikowski

Opis wynalazku

Przedmiotem wynalazku jest sposób wytwarzania insuliny i jej pochodnych.

Cukrzyca (diabetes mellitus), choroba metaboliczna (grupa chorób metabolicznych) o wieloczynnikowej etiologii, jest jedną z najgroźniejszych chorób społecznych i cywilizacyjnych, nazywana jest też czasem epidemią XXI wieku. Wiążąc się z poważnymi powikłaniami, nieleczona cukrzyca może prowadzić do śmierci pacjenta. Jeżeli przewlekła hiperglikemia wynika z niedostatecznego wydzielania insuliny przez organizm – cukrzyca typu 1 – leczenie wymaga stałego stosowania insuliny i/lub jej analogów. Liczba chorych na cukrzycę wynosi 5–6% populacji, a roczny wzrost liczby diabetyków wynosi ok. 4–6% zależności od regionu świata. Zachorowalność na tę chorobę wciąż wzrasta: jest uznana za chorobę cywilizacyjną – do rozwoju cukrzycy w dużym stopniu przyczynia się obecny styl życia – z tego też powodu do tej pory nie udaje się zahamować rozwoju tej choroby. Również mimo postępu medycyny, insulina oraz jej analogi szybko działające i długo działające stanowią podstawę leczenia cukrzycy typu 1 (10% całkowitej liczby przypadków choroby). Stosowana w leczeniu cukrzycy w dużych ilościach insulina i jej rozmaite pochodne są wytwarzane w dużej skali przemysłowej i z tej przyczyny metody jej wytwarzania są stale ulepszone w celu podwyższenia wydajności procesu i obniżenia kosztów produkcji. Ekonomika procesu produkcji jest również istotna dla zwiększającej się liczby wytwórców biopodobnych insulin [m. in.: S. Gough, *Practical Diabetes* 2013, 30 (4), 146–147a, G. Dranitsaris, E. Amir, K. Dorwart, *Drugs* 2011, 71 (12)1527–1536].

W pierwszym okresie po odkryciu insuliny, jako lek dla ludzi stosowana była insulina zwierzęca, wydzielana z trzustek wołowych lub wieprzowych. Humanizowanie tych insulin przez wymianę różnicujących ich aminokwasów metodą syntetyczną nie znalazło zastosowania w skali przemysłowej. Dopiero metody biosyntetyczne rozwiązały problem dostępu chorych do insuliny ludzkiej, a pierwszym procesem było biotechnologiczne oddzielne wytworzenie rekombinowanych łańcuchów A i B insuliny ludzkiej, a następnie ich połączenie w hormon taki, jaki jest wytwarzany endogennie w ludzkiej trzustce (D. V. Goeddel et al., *Proc. Nat. Acad. Sci.*, 76: 106–110, 1979). Od końca ubiegłego stulecia, biosyntetyczna insulina ludzka i jej analogi, wytwarzane przez transformację jednołańcuchowego białka fuzyjnego, stanowią podstawowe źródło tego hormonu jako egzogennej leku. Jest on wytwarzany w rekombinowanych systemach bakteryjnych, przede wszystkim w komórkach *Escherichia coli* (patent polski nr PL 127 843) lub drożdży (L. Thim et al., *Proc. Nat. Acad. Sci. USA*, 83: 6766–6770, 1986). Metoda wytwarzania tych leków jest ogólnie opisana w licznych opisach patentowych oraz publikacjach naukowych i polega na nadekspresji genu kodującego preproinsulinę, czyli polipeptyd hybrydowy, składający się z białka liderowego oraz proinsuliny, tj. łańcucha B insuliny (lub jej pochodnej), peptydu łącznikowego i łańcucha A insuliny (lub jej pochodnej). Po wyizolowaniu białka fuzyjnego, kolejnym etapem wytwarzania jest odcięcie białka liderowego i peptydu łącznikowego, po czym wydziela się czysty hormon. Jednym z przykładów takiego procesu jest sposób wytwarzania insuliny ludzkiej opisany w patencie polskim nr PL 180 818, a metody te są stale udoskonalane (np. polskie opisy patentowe: PL167 810, PL168 114, PL177 002, PL178 466, PL180 968, PL183 284, PL191 901, PL196,626, PL198 190, PL203 195, PL203 2546, PL210 437 i zgłoszenia patentowe P.309 882, P.310 007, P.320 644, P.356 005, P.374 949, P.385 586 i P.391 975).

Biosyntetyczna insulina ludzka i jej analogi są podstawowymi lekami w terapii cukrzycy typu 1 oraz mają istotne znaczenie w leczeniu typu 2 z problemami innej kontroli poziomu glukozy. Choć leczenie insuliną ludzką od 30 lat przynosi w większości przypadków zadowalające wyniki, powstały liczne jej analogi pozwalające na lepszą metaboliczną kontrolę cukrzycy dzięki ich szybszemu lub przedłużonemu działaniu. Od kilku już lat sprzedaż analogów insuliny przekracza połowę światowego rynku insulin, jednocześnie ochrona patentowa tych leków zaczyna wygasać. Ta sytuacja pozwala na wejście na rynki medyczne nowych producentów tych insulin jako leków biopodobnych, zgodnie z regulacjami Unii Europejskiej i USA. Z tych powodów istotne stają się wydajne metody ich biosyntezy, zapewniające wysoką ekspresję pożądanego białka.

W patencie polskim nr PL 180 818 opisano wydajny sposób wytwarzania rekombinowanej insuliny ludzkiej przy użyciu szczepu bakterii *E. coli* pozbawionego represora *cytR* (opisanego w europejskim zgłoszeniu patentowym nr EP.0303972) transformowanego plazmidem z operonem *deo*, zawierającym region kodujący białko hybrydowe (prekursor) obejmujące peptyd liderowy, o długości 62 aminokwasów pochodzący z N-końca modyfikowanej ludzkiej dysmutazy nadtlenkowej (CuZnSOD) poprzedzony od końca N aminokwasem Met i zakończony od końca C resztą Arg łączącą go z łańcuchem B insuliny. Łańcuch B insuliny połączony jest z łańcuchem A krótkim peptydem łącznikowym składającym się ze

znanych łączników Lys-Arg (europejski opis patentowy nr EP195 691) lub Arg (europejski opis patentowy nr EP347 781).

Celem wynalazku było uzyskanie wydajnego sposobu otrzymywania insuliny i jej analogów o kilkanaście procent przewyższającego wydajność procesu z zastosowaniem peptydu C jako białka łącznikowego, a także białek hybrydowych znajdujących zastosowanie w tym sposobie.

Przedmiotem wynalazku jest sposób wytwarzania insuliny ludzkiej obejmujący mikrobiologiczną ekspresję rekombinowanego białka zawierającego insulinę, charakteryzujący się tym, że obejmuje:

- (a) otrzymywanie polipeptydu hybrydowego zawierającego proinsulinę poprzez ekspresję tego polipeptydu hybrydowego w komórce szczepu bakterii *Escherichia coli* pozbawionej represorów *deo i*/lub genu *deo (cytR)* zawierającej DNA kodujący polipeptyd hybrydowy, przy czym ten DNA jest obecny w wektorze pod kontrolą promotora *deoP₁P₂* natomiast polipeptyd hybrydowy zawiera łańcuch B insuliny kowalencyjnie związany z łańcuchem A dipeptydem Arg-Arg,
- (b) odzyskiwanie polipeptydu hybrydowego,
- (c) ufałdowywanie i tworzenie wiązań dwusiarczkowych w polipeptydzie hybrydowym otrzymanym w (b), unikając uprzedniego poddawania polipeptydu hybrydowego siarczynolizie,
- (d) poddawanie uzyskanego w (c) ufałdowanego, z utworzonymi wiązaniami dwusiarczkowymi polipeptydu hybrydowego trawieniu enzymatycznemu w celu otrzymania insuliny,
- (e) oczyszczania tak wytworzonej insuliny.

Korzystnie, peptydem hybrydowym jest peptyd posiadający sekwencję aminokwasową o wzorze ogólnym:



gdzie :

A oznacza polipeptyd łańcucha A insuliny lub jej analogu, korzystnie o sekwencji wybranej spośród Sekw. Nr Id.: 1–4,

B oznacza polipeptyd łańcucha B insuliny lub jej analogu, korzystnie o sekwencji wybranej spośród Sekw. Nr Id.: 5–6,

n oznacza 0 lub 1,

X oznacza polipeptyd białka liderowego, korzystnie o sekwencji wybranej spośród SOD o Sekw. Nr Id.: 12 lub UBI o Sekw. Nr Id.: 13.

Korzystnie, peptyd hybrydowy ma sekwencję aminokwasową wybraną spośród: Sekw. Nr Id.: 8–26.

Korzystnie, w etapie (a) ekspresję prowadzi się w szczepie *E. coli* o genotypie: *F⁻ ara Δ (pro lac) rpsL thi cyt R rec A λ⁻*, pozbawionym aktywnego białka represora *cyt R*.

Korzystnie, w etapie (a) jako wektor stosuje się plazmid *pIBA* o sekwencji Sekw Nr Id: 27.

Korzystnie, etap (a) obejmuje fermentację w obecności glukozy, glicerolu lub galaktozy.

Korzystnie, etap (b) obejmuje:

- (i) niszczenie ściany komórkowej komórki bakteryjnej lub jej fragmentów w celu utworzenia lizatu, korzystnie poprzez działanie lizozymu i Tritonu X100 a następnie sonifikację lub dezintegrację;
- (ii) izolowanie wewnątrzkomórkowego precypitatu ciałek inkluzyjnych z lizatu przez wirowanie; i
- (iii) rozpuszczanie precypitatu ciałek inkluzyjnych.

Korzystnie, etap (c) obejmuje inkubację polipeptydu hybrydowego, korzystnie w obecności kwasu askorbinowego, zwłaszcza w stężeniu około 2 moli na mol grup SH obecnych w mieszaninie, w temperaturze od 4 do 37°C przez około od 1 do 30 godzin, korzystnie około 5 godzin, w pH od 8,5 do 12, korzystnie w pH od 11,0 do 11,25.

Korzystnie, etap (d) obejmuje ponadto:

- (i) doprowadzanie pH do około 8,8–9,0; i
- (ii) trawienie polipeptydu hybrydowego trypsyną i następnie ewentualnie karboksypeptydazą B w temperaturze 16–37°C w ciągu 30 minut do 16 godzin.

Korzystnie, etap (e) obejmuje ponadto oczyszczanie roztworu przez niskociśnieniową chromatografią cieczową na Sefarozie Q przed trawieniem trypsyną.

Korzystnie, etap (e) obejmuje ponadto oczyszczanie roztworu niskociśnieniową chromatografią na DEAE-Sepharose przed trawieniem karboksypeptydazą B.

Korzystnie, etap (e) obejmuje ponadto oczyszczanie uzyskanej insuliny wysokosprawną chromatografią cieczową, korzystnie do czystości co najmniej 98%, a następnie krystalizację, filtrację i suszenie.

Korzystnie, w etapie (e) oczyszczaną insuliną jest białko wybrane z grupy do której należy: insulina ludzka, insulina ludzka Ala^{A22}Lys^{B33}Arg^{B31}, insulina ludzka Ser^{A22}Lys^{B33}Arg^{B31}, insulina ludzka Ser^{A22}Lys^{B33}Arg^{B31}, insulina ludzka Lys^{B28}Pro^{B29} oraz insulina ludzka Asp³²⁸.

Szczegółowy opis wynalazku

W ramach poszukiwań wysokosprawnych systemów ekspresji insulin, zaplanowano badania nad wpływem rodzaju peptydu łącznikowego łączącego łańcuchy A i B (proinsuliny) na wydajność insuliny i jej analogów w systemie genowym opisanym w patencie polskim nr PL 180 818.

Zgodnie z wynalazkiem jako rekombinowaną preproinsulinę przyjmuje się połączenie proinsuliny z dodatkowym polipeptydem liderowym, np. ubikwityną albo dysmutazą ponadtlennkową (SOD), albo ich fragmentami. W świetle uznanej terminologii, przez rekombinowaną proinsulinę rozumie się łańcuch polipeptydowy, w którym łańcuchy A i B insuliny ludzkiej lub jej analogu są połączone dowolnym (poli)peptydem (łącznikowym) umożliwiającym ich ufałdowanie poprzez utworzenie 2 połączeń disiarczkowych między łańcuchami A i B oraz trzeciego – w obrębie łańcucha A. Zgodnie z wynalazkiem, przez insulinę rozumie się rekombinowaną insulinę ludzką lub jej rekombinowany analog o działaniu hipoglikemizującym.

Zgodnie z wynikami przeprowadzonych badań, według wynalazku specjalnie wydajny sposób wytwarzania insuliny i jej analogów obejmujący mikrobiologiczną nadekspresję genu kodującego rekombinowany polipeptyd hybrydowy, zawierający proinsulinę, w komórkach szczepów bakterii *Escherichia coli* pozbawionych represora *cytR* transformowanego odpowiednim wektorem z wstawionym właściwym genem pod kontrolą promotora *deoP1P2*, polega na tym, że wytwarzana proinsulina składa się z łańcucha B insuliny kowalencyjnie związanego z łańcuchem A dipeptydem Arg-Arg jako peptydem łącznikowym. W celu przeprowadzenia badań, na podstawie opisu patentu nr PL 180 818 wytworzono plazmid *pIBA*, zawierający region kodujący promotor *deoP1P2* oraz fragment genu ludzkiej dysmutazy ponadtlennkowej (SOD, 62 aa) lub ubikwityny (UBI) opisany w patentach US 8158382 oraz US 8956848 oraz przez A. Wojtowicz i in., Microbial Cell Fact. 2005, 4: 17 i Microb Cell Fact. 2014; 13(1): 113. W tej samej ramce odczytu znajduje się następnie fragment kodujący odpowiednią proinsulinę – ewentualnie modyfikowany łańcuch B, peptyd łącznikowy i ewentualnie modyfikowany łańcuch A. Elementy regulacyjne plazmidu pochodzą z wektora pBR322 (ATCC 31344). Promotor *deoP1P2* pochodzi z bakteryjnego operonu *deo* złożonego z czterech ściśle powiązanych genów regulujących katabolizm nukleotydów i dezoksynukleotydów w bakteriiach *Escherichia coli*; jego sekwencję zamplifikowano z chromosomalnego DNA szczepu *E. coli* K12 w oparciu o sekwencję nukleotydową z bazy danych genów [GenBank: AP009048]. W oparciu o szczep bakterii *E. coli* CSH50 (ATCC:39111; Cheng i wsp., Gene; 14 Suppl. 1–2: 121–130, 1981), otrzymano szczep *E. coli* IBA, o genotypie: *F⁻ ara Δ (pro lac) rpsL thi cyt R rec A λ -*, pozbawiony aktywnego białka represora *cyt R* poprzez wprowadzenie mutacji do jego genu, stosując metodę opisaną w europejskim zgłoszeniu patentowym nr PL 0303972. Po transformacji szczepu gospodarza *E. coli* IBA plazmidem *pIBA* uzyskano opisany w patencie polskim 180 818 system ekspresji kodowanego białka fuzyjnego w wektorze *pIBAINS* pod kontrolą promotora *deoP1P2*, który jest ujemnie kontrolowany przez białko chromosomalnego represora, produktu genu *cytR* (K. Hammer-Jespersen i inni, Molec. Genet. 1975; 137:327–335). W konstrukcji plazmidów zastosowano elementy-fragmenty DNA kodujące różne białka liderowe i różne pochodne insuliny, przy czym łańcuch A (pro)insuliny był połączony z łańcuchem B insuliny różnymi peptydami łącznikowymi – których przykłady podano poniżej.

Wydajność wytwarzania insuliny i jej analogów w tym systemie ekspresji i z różnymi peptydami łącznikowymi porównano z ekspresją tych hormonów z użyciem innych szczepów *E. coli* i wektora będącego pochodną plazmidu *pGAL* (Bank genów AY 424310), badając przede wszystkim zależność wydajności od rodzaju peptydu łącznikowego.

Zastosowanymi peptydami łącznikowymi były sekwencje:

1. Lys-Arg
2. Arg
3. Arg-Arg
4. Peptyd C

Według wynalazku przebadaną insuliną i jej pochodnymi- analogami były:

rekombinowaną insuliną ludzką (opis patentu polskiego nr PL128 599),
insulina ludzka Lys^{B28}Pro^{B29} (insulina lispro, opis patentu polskiego nr PL180 968),
insulina ludzka Gly^{A22}Arg^{B31}; (opis patentu polskiego nr PL219335 lub US 08618048; insulina GR),
insulina ludzka Ser^{A22}Arg^{B31} (opis polskiego zgłoszenia patentowego nr P.219335, insulina SR),

insulina ludzka Ser^{A22}Lys^{B3}Arg^{B31} (opis polskiego zgłoszenia patentowego nr P.399287, insulina SK3R) oraz

preproinsulina (prekursor) G: białko liderowe – łańcuch B – łącznik – łańcuch A: dezAsnA21GlyA21.

Jako szczep gospodarza *E. coli* do badań wydajności wykorzystano *E. coli* IBA; istotą jego właściwości jest nieaktywny gen białka represyjnego *cytR*, co ujawniono i opisano w literaturze fachowej (m. in. Miller J H. Experiments in molecular genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 1972), czy w opisie patentu USA 4 480 038 lub patentu polskiego 180 818 dla szczepów *E. coli* zdeponowanych pod numerami dostępu ATCC 69361 i 69363. Dla celów porównawczych stosowano również inne, szeroko stosowane w badaniach szczepy – *E. coli* DH5, *E. coli* DH5 α i *E. coli* HB101.

Jako plazmidy do ekspresji białek do badań wydajności wykorzystano wektor pIBA – promotor *deoP1P2*, oporność: tetracyklina, terminator transkrypcji *terA* Trp, ale dla celów porównawczych stosowano również inne wektory – pochodną pGAL1 (Bank genów AY 424310) – promotor *pms*, oporność: ampicylina, terminator transkrypcji ST1; pochodną pGAL1 – promotor *pms*, oporność: tetracyklina, terminator transkrypcji *terA* Trp i pochodną pGAL1 – promotor *deoP1P2*, oporność: tetracyklina, terminator transkrypcji ST1. Jako białko liderowe do badań wydajności stosowano przede wszystkim modyfikowany fragment SOD, jak w opisie patentu polskiego nr PL 180 818 oraz modyfikowaną ubikwitynę (UBI), opisaną w opisie patentowym US 8158382 oraz US 8956848. W dalszej części niniejszego opisu terminy SOD i UBI oznaczają związki opisane odpowiednio w tych opisach patentowych. Sposób wytwarzania każdej insuliny, będącej przedmiotem niniejszego zgłoszenia, realizuje się klasyczną metodą inżynierii genetycznej i biotechnologii. Polega ona na tym, że konstruuje się odpowiedni szczep, który w procesie biosyntezy (fermentacji) wytwarza żądany polipeptyd hybrydowy (odpowiednią preproinsulinę), którą następnie transformuje się w żądany produkt, oczyszcza się i izoluje. System ten jest uniwersalny dla wytworzenia rekombinowanej insuliny ludzkiej i jej badanych – niewiele różniących się – analogów. Dla tego celu skonstruowane zostały modyfikacje genu rekombinowanej preproinsuliny ludzkiej przy wykorzystaniu technik genetycznych takich jak np. reakcja miejscowo – specyficznej mutagenyzy. Reakcję mutagenyzy punktowej przeprowadzono przy użyciu zestawu firmy Stratagene (nr kat. 200518-5); jako matrycę wykorzystano DNA plazmidowe plazmidu pIBA oraz pGAL. Jako matrycę można również wykorzystywać dowolny inny DNA zawierający odpowiednią sekwencję kodującą rekombinowaną ludzką proinsulinę lub preproinsulinę.

Dla każdej insuliny wytworzono konstrukty plazmidowe – wektory – kodujące badaną (pre)proinsulinę z (poli)peptydem łącznikowym, przy czym w konstrukcji plazmidów zastosowano elementy – fragmenty DNA kodujące różne proinsuliny – składające się z różnych peptydów łącznikowych i różnych pochodnych insuliny.

Wektorami tymi transformowano komórki kompetentne odpowiedniego szczepu *Escherichia coli*, jak np. IBA, DH5 α , DH5, czy HB101, przy czym możliwe jest wykorzystanie komórek innych szczepów *E. coli* lub komórek innych mikroorganizmów albo innych znanych linii komórkowych nadających się do ekspresji białek rekombinowanych. Plazmid zawierający określoną modyfikację genu rekombinowanej preproinsuliny ludzkiej izolowano i sekwencjonowano w celu sprawdzenia poprawności transformacji i sekwencji nukleotydowej. Pojedyncze klony transformowanych szczepów hodowano standardowo na odpowiednim podłożu z dodatkiem antybiotyku selekcyjnego, przygotowując materiał bakteryjny do badawczego banku szczepów, a próby w stosunku 1:1 hodowli bakteryjnej i 40% glicerolu zdeponowano w temp. -70°C. Wszystkie otrzymane szczepy banków badawczych poddano wyczerpującym badaniom mikrobiologicznym, biochemicznym i fizykochemicznym, które potwierdziły ich stabilność i zgodność ich właściwości z wymaganiami właściwymi dla banków przemysłowych. Przebieg dalszej biosyntezy oparto na opisie patentu polskiego nr PL 180 818. Po biosyntezie, wytwarzane w szczepach *E. coli* warianty rekombinowanej preproinsuliny izolowano po rozbiciu komórek w postaci ciałek inkluzyjnych, które oddzielano. Z ciałek inkluzyjnych izolowano polipeptyd hybrydowy z insuliną lub jej analogiem (odpowiednią preproinsulinę), a następnie poddawano go ufałdowaniu (foldingowi, renaturacji). Użytkany roztwór ufałdowanego białka hybrydowego poddawano kontrolowanemu działaniu trypsyny, analogicznie jak w przypadku szeregu metod znanych uprzednio i opisanych (np. przez Kemmlera i wsp. w J. Biol. Chem., Vol. 246, str. 6786–6791 (1971) lub w patentach US 6686177, lub US 6100376), albo proteazy deubikwitynującej, jak opisano w patentach US 8158382 oraz US 8956848 oraz przez A. Wojtowicz i in., Microbial Cell Fact. 2005, 4: 17 i Microb Cell Fact. 2014; 13(1): 113, albo obu tych enzymów.

W części przypadków (insulina ludzka, insulina Lispro) działaniem karboksypeptydazy B usuwano dodatkowe aminokwasy peptydu łącznikowego; w pozostałych przypadkach etap ten pomijano. Otrzymaną insulinę i jej analogi poddawano procesowi oczyszczania przy zastosowaniu znanych metod, przede wszystkim chromatografii niskociśnieniowej, ultrafiltracji i HPLC. Z oczyszczonego w dostatecznym stopniu – do wymagań farmakopealnych – roztworu insuliny lub jej analogu, produkt krystalizowano i suszono.

W celu porównania wydajności ekspresji dla każdego gospodarza i wariantu systemu opracowano i zoptymalizowano biosyntezę, izolację i oczyszczanie produktu każdego ze szczepów. Biosyntezę szczepów *Escherichia coli* pozbawionych represora *cytR* transformowanego odpowiednim wektorem z wstawionym właściwym genem pod kontrolą promotora *deoP1P2* prowadzono zgodnie z opisem polskiego patentu nr PL 180 818, w fermentorach o pojemności 7,5 dm³, stosując ok. 3–5% materiału szczepiennego i minimalną, przemysłową pożywkę do inokulum i hodowli produkcyjnej oraz regulując indukcję żądanego produktu za pomocą glukozy, zgodnie z przykładem 2 polskiego patentu nr PL 180 818. Biosyntezę innych szczepów *E. coli* prowadzono zgodnie z ogólnymi sposobami ich hodowli. Hodowlę inokulum prowadzono przez ok. 10 godzin, do czasu zakończenia logarytmicznej fazy wzrostu bakterii, a hodowlę produkcyjną przez ok. 16 godzin. Dalszy proces wydzielania, transformacji i oczyszczania produktu prowadzono również zgodnie z opisem polskiego patentu nr PL 180 818 – zoptymalizowanym przykładem 2 i 3. W szczególności, po zakończonej biosyntezie produkcyjnej wytworzoną brzeczkę chłodzono do ok. 10°C, odwirowywano komórki *E. coli*, poddawano je działaniu lizozymu i Tritonu X100, a następnie w temperaturze 5–10°C poddawano sonifikacji ultradźwiękami lub dezintegracji ciśnieniowej. Ciała inkluzyjne ponownie odwirowywano, przemywano, po czym poddawano renaturacji, jak ogólnie opisano w przykładzie 3A patentu PL 180 818. Za wyjątkiem białka otrzymywanego według Przykładu 6, dalszy zoptymalizowany proces polegał na poddaniu roztworu niskociśnieniowej chromatografii cieczowej z użyciem DEAE Sepharozy po cytrakonylacji i trawieniu trypsyną lub proteazą deubikwinującą i trypsyną. Po decytrakonylacji i strąceniu oraz rozpuszczeniu białka, roztwór poddawano kolejnej niskociśnieniowej chromatografii cieczowej z zastosowaniem złoża Sepharozy Q, po której następowało ewentualne trawienie peptydu łącznikowego karboksypeptydazą B. Po końcowej wysokosprawnej chromatografii cieczowej na złożu Kromasil C8-18, oczyszczony do czystości minimum 98% produkt poddawano krystalizacji, filtracji i suszeniu. Szczegółowy opis całego procesu, całkowicie porównywalnego dla określonych insulin, podano w przykładach. Proces prowadzono w skali pilotowej, stosując metody i aparaturę wykorzystywane w przemyśle. Po zoptymalizowaniu procesu, dla celów badania wydajności procesu przeprowadzono trzy kolejne szarże każdej insuliny, a przedstawione wyniki są średnią z każdej z nich.

Do porównania wydajności ekspresji, przy ustalonym procesie produkcyjnym, z ujednoliconej kontroli procesowej wybrano badania analityczne pięciu operacji:

- (1) Gęstości optycznej (OD₆₀₀) zakończonej hodowli produkcyjnej;
- (2) Masy hodowli (wyizolowanych bakterii) – wyrażanej w g w przeliczeniu na 1 dm³ hodowli produkcyjnej (wilgotnej i/lub suchej);
- (3) Masy wyizolowanych ciał inkluzyjnych, wyrażanej w g na 1 dm³ hodowli produkcyjnej (wilgotnej i/lub suchej albo jako suchej masy zawiesiny w buforze);
- (4) Białka całkowitego (po foldingu prekursora), oznaczanego metodą spektrofotometryczną i wyrażanego w AU/dm³ lub HPLC i wyrażanego w % ufałdowanego prekursora w białku całkowitym;
- (5) Masę oczyszczonego końcowego produktu wyrażaną w mg na 1 dm³ hodowli produkcyjnej.

W przypadku insuliny lispro podano wyniki badań analitycznych większości etapów – białka całkowitego oznaczanego metodą spektrofotometryczną i wyrażanego w AU/dm³ i metodą Bradforda i wyrażanego w g na 1 dm³ hodowli produkcyjnej.

Ogólny schemat blokowy wytwarzania insuliny i jej pochodnych przedstawiono na schemacie 1.

Wynalazek ujawnia, że wydajność końcowa biosyntetycznej insuliny ludzkiej i jej analogów, przy zoptymalizowanym procesie technologicznym biosyntezy, wydzielania, transformacji i oczyszczania, zależąca od stopnia nadekspresji genu kodującego preproinsulinę w komórkach modyfikowanych bakterii *Escherichia coli* (z nieaktywnymi represorami *cytR* negatywnie regulującymi operon *deo*) transformowanych określonym plazmidem (z promotorem *deoP1P2*), nieoczekiwanie zależy przede wszystkim od zastosowanego łącznika peptydowego, łączącego łańcuch A i B insuliny.

Zgodnie z wynalazkiem nieoczekiwanie ustalono, iż największą wydajność w tym systemie uzyskuje się w przypadku, gdy rolę łącznika peptydowego pełni dipeptyd Arg-Arg.

Dzięki wynalazkowi podano specyficzne warunki dla wytwarzania insuliny ludzkiej oraz jej analogów metodą inżynierii genetycznej, przy zastosowaniu mikrobiologicznej nadekspresji odpowiedniej preproinsuliny w komórkach szczepów bakterii *Escherichia coli* pozbawionych represorów *deo* i/lub genu *deo*, zawierających plazmid DNA kodujący odpowiedni polipeptyd hybrydowy w wektorze pod kontrolą promotora *deoP1P2* z optymalną wydajnością – znacząco podwyższoną w porównaniu z opublikowanymi danymi.

W opisanych poniżej przykładach realizacji wynalazku strukturę, wydajność i czystość wszystkich produktów wyznaczano rutynowo znanymi metodami HPLC i spektrofotometrii UV; do pełnego scharakteryzowania wytworzonych insulin zastosowano również liczne nowoczesne metody i techniki analityczne, między innymi MS, CIEF, LC-MS/MS, GC, mapowanie peptydowe, NMR, krystalografię i proszkową dyfrakcję rentgenowską, CD, FTIR, ELISA czy PCR oraz sekwencjonowanie aminokwasów i nukleotydów.

Dla lepszego wyjaśnienia istoty wynalazku niniejszy opis został uzupełniony o szczegółowe omówienie jego przykładowych realizacji, obejmujące również rysunki, przedstawiające ogólną strukturę (Fig. 1) i sekwencję (Fig. 2) plazmidu *pIBA* zarówno gotowego do insercji genu białka hybrydowego jak i przykładowej realizacji w postaci plazmidu zawierającego gen kodujący białko hybrydowe insuliny (Fig. 3 i Fig. 4).

W przykładach nie podano szczegółowych opisów konwencjonalnych metod bioinżynierii zastosowanych do konstrukcji określonych genów, plazmidów, wektorów ze wstawionymi genami kodującymi badane preproinsuliny czy też sposobów transformacji szczepów bakteryjnych tymi wektorami, jak również metod badania stabilności tych szczepów. Metody te są powszechnie znane specjalistom w tej dziedzinie i stosowane w laboratoriach biotechnologicznych oraz opisane w licznych publikacjach, z których kluczowe podano w niniejszym opisie. Produkty końcowe wytwarzane metodą według wynalazku są znanymi insulinami, większość szczegółowych bioinżynierskich metod konstrukcji tych rekombinowanych szczepów bakteryjnych opisano w polskim opisie patentowym nr PL 180 968, P. 385 586 i P. 399 287.

Nie opisano również szczegółowo standardowych metod analitycznych stosowanych do określania zawartości i stężenia białek i insulin na poszczególnych etapach wytwarzania, gdyż metody te są również dobrze znane odpowiednim specjalistom i rutynowo stosowane w oznaczaniu białek.

Podane przykłady, nie ograniczając zakresu wynalazku, opisują szczegółowo zastosowaną metodykę badawczą, pozwalającą w pełni kontrolowanym sposobem na dokonanie porównania uzyskanych reprezentatywnych wyników badań zależności wydajności procesu od struktury peptydu łącznikowego.

Przykłady

I. Ogólna procedura prowadzenia procesu

Etap 1. Bank badawczy – hodowla podstawowa

Hodowle podstawowe szczepu *E. coli* IBA, DH5 DH5alfa i HB101 zawierające jako plazmidy do ekspresji białek wektor *pIBA* i pochodną *pGAL1* (Bank genów AY 424310) prowadzono zgodnie z wymaganiami (ICH Topic Q 5 D: Quality of Biotechnological Products: Derivation and Characterisation of Cell Substrates Used for Production of Biotechnological/Biological Products oraz ICH Topic Q 5 B: Quality of Biotechnological Products:

Analysis of the Expression Construct in Cell Lines Used for Production of r-DNA Derived Protein Products) na podłożu podstawowym :

Podłoże podstawowe:

- | | |
|--|--|
| – Hydrolizat kazeiny | 20 g |
| – Ekstrakt drożdżowy | 10 g |
| – NaCl | 5,0 g |
| – Woda do objętości | 1 dm ³ |
| – Antybiotyk (tetracyklina lub ampicylina) do końcowego stężenia | 12,5 µg/cm ³ TET lub 100 µg/cm ³ Amp |

Kolbę z podłożem podstawowym z antybiotykiem szczepiono pojedynczą kolonią bakteryjną uzyskaną po transformacji i wytrząsano (220 obr./min.) w 37°C do gęstości optycznej OD₆₀₀ ok. 1.

Otrzymaną hodowlę mieszano z podłożem do zamrażania (1:1) i zamrażano w temperaturze -70°C w porcjach po 1 ml.

Podłoże do zamrażania:

– K ₂ HPO ₄	6,3 g
– KH ₂ PO ₄	1,8 g
– Cytrynian sodu mx 2H ₂ O	0,57 g
– MgSO ₄ x 7H ₂ O	0,09 g
– (NH ₄) ₂ SO ₄	0,9 g
– Glicerol	44,0 g
– Woda do objętości	450 cm ³

Dla uzyskanej hodowli bakteryjnej wykonywano następujące badania:

- pomiar gęstości optycznej hodowli OD przy długości fali 600 nm;
- sprawdzenie jednorodności hodowli bakteryjnej wykonując preparat przyżyciowy. Oceniano czystość hodowli bakteryjnej, cechy morfologiczne – wygląd, kształt komórek. Obserwacja preparatu w mikroskopie Olympus CX 40;
- wyznaczano najbardziej prawdopodobną liczbę jednostek tworzących kolonie na podłożu TSA i TSA z konkretnym antybiotykiem – cfu/ml;
- sprawdzano czystość mikrobiologiczną hodowli na podłożach Sabouraud (do identyfikacji zanieczyszczeń grzybami) i TSA (do identyfikacji zanieczyszczeń innymi bakteriami).

Etap 2. Inokulum

Inokulum przygotowuje się w kolbach o pojemności 250–300 ml i wypełnieniu 50 ml pożywki do inokulum. Do każdej z kolb dodawane są sterylnie roztwory:

- Antybiotyk (tetracyklina lub ampicylina, 12,5 mg/ml) 50 µl
- Glukoza 50% 500 µl

Do hodowli szczepów *E. coli* IBA dodaje się dodatkowo:

- Prolina (100 mg/ml) 300 µl
- Tiamina (50 mg/ml) 2–5 µl.

Po zaszczepieniu kolb hodowlą podstawową (bank roboczy), inokulum inkubuje się w temperaturze 30–37°C przez 16–18 godzin do osiągnięcia OD₆₀₀ ok. 0,7–1,0 na wytrząsarce przy 180–200 rpm. Pożywka do inokulum (skład na 1 dm³)

– K ₂ HPO ₄	9 g
– KH ₂ PO ₄	1 g
– NaCl	5 g
– NH ₄ Cl	1 g
– MgSO ₄ ·7H ₂ O	0,4 g
– Cytrynian żelazowo-amonowy	0,01 g
– Ekstrakt drożdżowy	0,2 g
– Roztwór pierwiastków śladowych	1 ml
▪ MnSO ₄ x H ₂ O	1 g/dm ³
▪ ZnSO ₄ x 7 H ₂ O	2,78
▪ CoCl ₂ x 7 H ₂ O	2 g/dm ³
▪ Na ₂ MoO ₄ x 2 H ₂ O	2 g/dm ³
▪ CaCl ₂ x 2 H ₂ O	3 g/dm ³
▪ CuSO ₄ x 5 H ₂ O	1,85 g/dm ³
▪ H ₃ BO ₄	0,5 g/dm ³
▪ HCl (32%)	100 ml/dm ³

Etap 3. Hodowla produkcyjna (Fermentor BioFlo 310 firmy NewBrunswick 7,5 dm³ (objętość startowa 4 dm³))

Parametry wyjściowe hodowli:

- Temperatura: 37°C
- pH: 7,1±0,01
- DO: 30–35%
- Obroty: 250 rpm
- Przepływ powietrza: 5 dm³/min.

Do sterylnego fermentora z 4 dm³ pożywki produkcyjnej i antypieniaczem Antifoam 204 (0,5–1 ml, SIGMA) dodaje się sterylnie roztwory:

- 6 g MgSO₄ x 7H₂O w 80 cm³ wody
- 0,2 g CaCl₂ x 2H₂O w 60 cm³ wody
- glukoza 50%, 4 cm³

Do hodowli szczepów *E. coli* IBA dodaje się dodatkowo:

- Prolina (100 mg/ml) 24 cm³
- Tiamina (50 mg/ml) 200–500 µl.

Pożywka produkcyjna:

- | | |
|---|------------------------|
| ○ K ₂ HPO ₄ | 8 g/dm ³ |
| ○ KH ₂ PO ₄ | 2 g/dm ³ |
| ○ NH ₄ Cl | 3 g/dm ³ |
| ○ Cytrynian żelazowo-amonowy | 0,15 g/dm ³ |
| ○ K ₂ SO ₄ | 0,9 g/dm ³ |
| ○ Cytrynian sodu | 3,0 g/dm ³ |
| ○ Roztwór pierwiastków śladowych | 3 cm ³ |
| ▪ MnSO ₄ x H ₂ O | 1 g/dm ³ |
| ▪ ZnSO ₄ x 7 H ₂ O | 2,78 g/dm ³ |
| ▪ CoCl ₂ x 7 H ₂ O | 2 g/dm ³ |
| ▪ Na ₂ MoO ₄ x 2 H ₂ O | 2 g/dm ³ |
| ▪ CaCl ₂ x 2 H ₂ O | 3 g/dm ³ |
| ▪ CuSO ₄ x 5 H ₂ O | 1,85 g/dm ³ |
| ▪ H ₃ BO ₄ | 0,5 g/dm ³ |
| ▪ HCl (32%) | 100 ml/dm ³ |

Pożywkę w fermentorze zaszczenia się inokulum i hoduje w odpowiednich warunkach. Jako źródło węgla i energii do wytwarzania biomasy stosuje się stałe dozowanie do fermentora roztworu 40% glukozy (w przypadku szczepów *E. coli* IBA z dodatkiem proliny 10–30 mg/ml) tak, aby jej stężenie wynosiło 70–180 mg/dm³.

pH (7,1±0,1) kontroluje się przy pomocy 16% amoniaku przez kaskadę. Poziom tlenu w pożywce utrzymuje się na poziomie nie niższym niż 30–35%. Parametr ten kontrolowany jest w zależności od odczytu z elektrody DO przez zwiększenie obrotów mieszadła z 250 do 1000 rpm i przepływu powietrza z 5 do 10 dm³/min. Po osiągnięciu maksymalnych obrotów mieszadła oraz maksymalnego przepływu powietrza (po ok. 7–10 godz.) zatrzymuje się dodawanie glukozy; wartość OD₆₀₀ wynosi (w przypadku szczepów *E. coli* IBA/pIBA) ok. 25–35.

Po spadku stężenia glukozy, jej dodawanie utrzymywane jest tak, aby stężenie wynosiło ok. 35–50 mg/dm³, a pH było utrzymywane na poziomie 7,1 ± 0,1. Ta faza biosyntezy prowadzona jest do momentu osiągnięcia przez komórki stacjonarnej fazy wzrostu (nie obserwuje się już wzrostu gęstości optycznej hodowli) i trwa ok. 7–8 godzin. Na koniec fazy produkcji rekombinowanego białka OD₆₀₀ wynosi 50–60 w przypadku szczepów *E. coli* IBA/pIBA. Następnie hodowla schładzana jest do temperatury poniżej 20°C, korzystnie ok. 10°C.

Etap 4. Wirowanie

Biomasę odwirowuje się na wirówce stacjonarnej przy 6 000–8 000 rpm w czasie 15 min. i w temperaturze +4°C. Brzeczkę pofermentacyjną można odwirować na wirówce przepływowej.

Etap 5. Izolacja ciałek inkluzyjnych

Biomasę zawiesza się w 1 dm³ buforu 50 mM TRIS, 0,5 M NaCl i 5 mM beta-merkaptotetanolu, o pH 7,5, dodaje się lizozym do stężenia 0,43 mg/cm³ i 0,2M EDTA do stężenia 0,5%, po czym inkubuje się przez 30 minut w temperaturze pokojowej i poddaje dezintegracji przez sonikację lub ciśnienie.

Etap 6 – wariant 1: Sonikacja

Materiał chłodzi się do ok. 10°C, dzieli się na 4 części, dodaje Triton do 1% i każdą porcję sonikuje się przez 30 minut przy amplitudzie 33%. Po sonikacji próbę odwirowuje się przy 8000 rpm przez 15 min w temperaturze pokojowej. Następnie osad zawiesza się w 1 dm³ buforu 50 mM TRIS, 0,5 M NaCl i 5 mM beta-merkaptotetanolu, o pH 7,5 i ponownie sonikuje przez 10 min. Ciała inkluzyjne odwirowuje się przy 8000 rpm przez 15 min w temperaturze pokojowej.

Etap 6 – wariant 2: Dezintegracja ciśnieniowa

Materiał chłodzi się do ok. 10°C, poddaje się dezintegracji ciśnieniowej (Panda+1000 firmy GEA Niro Soavi) przy ciśnieniu 800 bar z szybkością zapewniającą całkowite rozbicie komórek (kontrola mikroskopowa) i chłodząc zawiesinę tak, aby temperatura nie przekroczyła ok. 10°C.

Etap 7. Rozpuszczanie ciał inkluzyjnych

Ciała inkluzyjne rozpuszcza się w odpowiedniej ilości 12 mM buforu wodorowęglanowego i 0,2 mM EDTA do uzyskania wartości absorbancji poniżej 9 przy długości fali 280 nm. Roztwór doprowadza się do pH do wartości $11,5 \pm 0,5$ za pomocą 2 M NaOH i miesza przez 30 min. w temp. pokojowej. Następnie koryguje się pH do $10,8 \pm 0,4$ za pomocą 2 M HCl. Roztwór klaruje się przez wirowanie przy 12 500 rpm przez 15 min w temp. 4°C.

Etap 8. Renaturacja

Po rozpuszczeniu ciał inkluzyjnych białko renaturuje się przez napowietrzanie – silne mieszanie – przez ok. 16–18 h w temperaturze pokojowej. Następnie koryguje się pH do wartości 9,0 za pomocą 2 M HCl.

Etap 9. Cytrakonylacja

Do roztworu białka dodaje się porcjami bezwodnika cytrokonilowego w ilości obliczonej według wzoru: objętość bezwodnika (cm^3) = $0,15 \times$ objętość roztworu [dm^3] $\times A_{280}$ oraz 2 M NaOH do utrzymania pH o wartości w przedziale 8,7–9,3. Po dodaniu całej ilości bezwodnika roztwór miesza się przez 1 godzinę, po czym dodaje się roztwór 2 M etanolaminy w trzykrotnej objętości bezwodnika i miesza przez 30 min.

Etap 10. Trypsynowanie

Po cytrakonylacji pH roztworu koryguje się do wartości 8,8 za pomocą 2 M HCl i rozcieńcza wodą tak, aby przewodnictwo wyniosło poniżej 5 mS. Następnie dodaje się 1% roztwór trypsyny w ilości obliczonej według wzoru: objętość roztworu trypsyny = $A_{280} \times$ objętość roztworu [dm^3] / 95) i miesza przez 16–18 godzin w temperaturze pokojowej. Reakcję hamuje się roztworem aprotyniny o stężeniu 1 mg/ml, dodawanego w ilości dwudziestojednokrotnie mniejszej od objętości trypsyny.

Etap 11. Odcięcie białka liderowego – ubikwityny

Do roztworu białka o temperaturze 10°C dodaje się 20 mM kwas fosforowy do uzyskania pH o wartości 7,5, po czym w temperaturze 37°C dodaje się 19,2 cm^3 roztworu proteazy UBP1ΔC lub UBP1ΔC2 o stężeniu 1 mg/ cm^3 , i miesza się przez 1 godzinę. Po ochłodzeniu do 4°C roztwór klaruje się przez wirowanie przy 12 000 rpm przez 15 minut.

Etap 12. Chromatografia niskociśnieniowa na złożu DEAE Sepharosa

Na kolumnę wypełnioną złożem DEAE (200 cm^3) i zrównoważoną 0,5 M TRIS o pH $8,6 \pm 0,2$, a następnie 20 mM TRIS o pH $8,6 \pm 0,2$ nanosi się roztwór o pH $8,6 \pm 0,2$. Po naniesieniu kolumnę płucze się buforami o pH $8,6 \pm 0,2$: 20 mM TRIS z NaCl do przewodności $5\text{mS} \pm 1\text{mS}$, a następnie 20 mM TRIS z NaCl do przewodności $2\text{mS} \pm 1\text{mS}$ i $20\% \pm 5\%$ izopropanolu. Białko insuliny eluuje się buforem o pH $8,6 \pm 0,2$, składającym się z 20 mM TRIS z NaCl do przewodności $5 \pm 1\text{mS}$ i z dodatkiem $25\% \pm 5\%$ izopropanolu.

Etap 13. Decytrakonylacja

Frakcję główną wyeluowaną z kolumny rozcieńcza się 2 razy, ochładza do 4°C i zakwasza 0,1 M HCl do pH o wartości $2,9 \pm 0,2$, po czym miesza się 10 godzin.

Etap 14. Strącanie insuliny chlorkiem cynku

Do rozcieńczonej frakcji głównej dodaje się taką objętość 1 M chlorku cynku, która odpowiada 1/50 objętości frakcji głównej, koryguje się pH do wartości $5,5 \pm 0,5$ i miesza przez 1 godzinę. Następnie zawiesinę odwirowuje się przy 9000 rpm przez 15 min w 4°C. Otrzymany osad zawiesza się w wodzie.

Etap 15. Chromatografia niskociśnieniowa na złożu Q Sepharosa

Zawiesinę insuliny uzupełnia się wodą do objętości wyliczonej z wzoru: objętość końcowa = całkowita ilość białka otrzymanego po DEAE [$\text{AU } A_{280}$]/7, dodaje TRIS o pH 8,6 do stężenia 30 mM i 0,2 M roztwór EDTA w ilości wyliczonej z wzoru: objętość [cm^3] = całkowita ilość białka otrzymanego po DEAE [$\text{AU } A_{280}$] / 190, po czym doprowadza się pH do wartości 8,6. Na kolumnę wypełnioną złożem Q (80 cm^3) i zrównoważoną 0,5 M TRIS o pH 8,66, a następnie 20 mM TRIS o pH 8,6, nanosi się roztwór rozpuszczonej soli cynkowej insuliny o pH 8,6. Po naniesieniu kolumnę płucze się buforem 20 mM TRIS z $20\% \pm 5\%$ izopropanolu o pH 8,6, a następnie eluuje się białko insuliny buforem o pH 8,6, składającym się z 20 mM TRIS z NaCl do przewodności $3\text{mS} \pm 2\text{mS}$ i $25\% \pm 5\%$ izopropanolu.

Etap 16. Reakcja z karboksypeptydazą B

Frację główną wyeluuowaną z kolumny rozcieńcza się 2 razy i dodaje roztwór karboksypeptydazy B o stężeniu 2,5 mg/cm³ w ilości obliczonej z wzoru: objętość [μm³] = całkowita ilość białka otrzymanego po Q [AU A₂₈₀] / 15. pH roztworu doprowadza się do wartości 8,8, po czym roztwór miesza się przez 16–18 godzin w temperaturze pokojowej. Następnie pH doprowadza się do 3 i roztwór poddaje się oczyszczaniu przez chromatografię wysokociśnieniową.

Etap 17. Chromatografia wysokociśnieniowa RP-HPLC.

Rozdział prowadzi się na kolumnie Kromasil C8 lub C18 z zastosowaniem gradientu acetonitrylu:

Faza A	Faza B
95%	5%
75%	25%

Faza ruchoma A: 2000 cm³ 0,2 M roztworu siarczanu sodu z 440 cm³ acetonitrylu, o pH 2,3 i o temperaturze 25–30°C;

Faza ruchoma B: 1250 cm³ 0,2 M roztworu siarczanu sodu z 1250 cm³ acetonitrylu, o pH 2,3 i o temperaturze 25–30°C;

Roztwór po reakcji z karboksypeptydazą B lub frakcję główną po oczyszczaniu na złożu Q Sepharosa po dwukrotnym rozcieńczeniu i doprowadzeniu pH do wartości 3, oczyszcza się porcjami, niosząc na kolumnę jednorazowo objętość roztworu o zawartości białka całkowitego 200 AU, po czym zbiera frakcję główną tak, aby czystość produktu (HPLC) wynosiła co najmniej 98%.

Etap 18. Strącanie insuliny chlorkiem cynku

Połączone frakcje główne po chromatografii wysokociśnieniowej rozcieńcza się 2 razy i dodaje 1 M ZnCl₂ w proporcji 3 cm³ chlorku cynku na 150 cm³ próbki, po czym pH doprowadza się do wartości 6 ± 1 i miesza przez 1 godzinę. Następnie zawiesinę wiruje się przy 9000 rpm przez 15 min w 4°C, a osad zawiesza się w 10–15 cm³ wody.

Etap 19. Chromatografia na złożu Sephadex G-25 i liofilizacja

Zawiesinę osadu doprowadza się do pH 3, filtruje przez filtr 0,22 μm i nanosi na kolumnę wypełnioną nośnikiem Sephadex G 25 (120 cm³). Przed naniesieniem kolumnę równoważy się 5 mM octanem amonu o pH 4, a białko wymywa się 5 mM octanem amonu o pH 4. Tak otrzymaną substancję poddaje się liofilizacji.

II. Przykłady szczegółowe**Przykład 1.** Insulina ludzka LysB28ProB29 (insulina lispro)

Zgodnie z powyżej podaną ogólną procedurą prowadzenia procesu (punkt I), przeprowadzono wytwarzanie biosyntetycznej insuliny lispro, stosując etapy 1–6, 6b, 7–10 i 12–19 oraz dla systemu ekspresji 1.2. i 1.5. dodatkowo etap II (Tabela 1. Insulina lispro). Proces badano w 20 wariantach 8 systemów ekspresji białka. Z otrzymanych danych wynika, że sprawdzające się w warunkach laboratoryjnych szczepy bakterii *Escherichia coli* DH5α, DH5 i HB101 z plazmidami z promotorami *deoP1P2* i *pms* nie są odpowiednie dla hodowli produkcyjnej z pożywką minimalną (system 1.3., 1.4., 1.5., 1.6. i 1.7.). Szczepy te charakteryzują się niskim wzrostem na minimalnej pożywce technologicznej, wykazując znacznie niższą wartość gęstości optycznej w fermentorach produkcyjnych. Podobnie małą przydatność produkcyjną wykazuje system szczepu IBA z uszkodzonym represorem *cytR* i z plazmidem *pIGAL* (system 1.8.), charakteryzujący się stosunkowo lepszym wzrostem, lecz niską zawartością ciał inkluzyjnych (obserwacje mikroskopowe).

Najlepszą wydajnością ekspresji charakteryzuje się system szczepu IBA/pIBA, opisany w patencie polskim nr PL 180 818, który zbadano w 4 wariantach peptydu łącznikowego i 2 wariantach białka liderowego (system 1.1. i 1.2.). Wszystkie charakteryzują się dobrymi, porównywalnymi parametrami wzrostu w pożywce minimalnej, lecz różnią się masą ciał inkluzyjnych i względną wydajnością renowacji peptydu fuzyjnego oraz – w związku z tym – końcową wydajnością produktu.

Najwyższą wydajność insuliny lispro uzyskano wykorzystując system ekspresji *E. coli* IBA/pIBA wytwarzający łańcuch A połączony z łańcuchem B insuliny lispro dipeptydem ArgArg (Tabela 1, L.p. przykładu 1.1.1. i 1.2.1.), przewyższający wydajność pozostałych systemów co najmniej o 13–15%.

Przykład 2. Insulina GlyA22ArgB31 (insulina GR)

Zgodnie z powyżej podaną ogólną procedurą prowadzenia procesu (punkt I), przeprowadzono wytwarzanie biosyntetycznej insuliny GR, stosując etapy 1–6, 6b, 7–10, 12–15 i 17–19 oraz dla systemu ekspresji 2.2. dodatkowo etap 11 (Tabela 2. Insulina GR). Proces badano w 12 wariantach 5 systemów ekspresji białka (nie stosowano wariantu z łącznikiem LysArg, jako nieprowadzącego do

insuliny GR). Otrzymane dane potwierdzają wnioski z przykładu 1, dotyczące wyższości systemu ekspresji opisanego w patencie polskim nr PL 180 818 nad innymi badanymi oraz wykazują podobne różnice w wydajności procesu w zależności od rodzaju peptydu łącznikowego łączącego łańcuch A i B tego analogu insuliny. W szczególności najwyższą wydajność insuliny GR uzyskano wykorzystując system ekspresji *E. coli* IBA/pIBA, wytwarzający łańcuch A połączony z łańcuchem B insuliny GR dipeptydem ArgArg (Tabela 2, L.p. przykładu 2.1.1. i 2.2.1.) i przewyższający wydajność pozostałych systemów co najmniej o 16%–22%. Z przedstawionych rezultatów wynika przewaga technologiczna łącznika ArgArg bez względu na rodzaj białka liderowego.

Przykład 3. Insulina ludzka.

Zgodnie z powyżej podaną ogólną procedurą prowadzenia procesu, przeprowadzono wytwarzanie biosyntetycznej insuliny ludzkiej w 8 wariantach 2 systemów ekspresji białka. W badaniach prowadzono etapy 1–6, 6b, 7–10 i 12–19 oraz dla systemu ekspresji 1.2. i 1.5. dodatkowo etap 11. Z otrzymanych danych wynika (Tabela 3. Insulina ludzka), że najlepszą wydajnością ekspresji charakteryzuje się system szczepów IBA/pIBA, wykazujących się dobrymi, porównywalnymi parametrami wzrostu w przemysłowej pożywce minimalnej i różniących się masą wytworzonych ciał inkluzyjnych oraz względną wydajnością renaturacji. Ponownie najwyższą wydajność insuliny ludzkiej uzyskano wykorzystując system ekspresji *E. coli* IBA/pIBA wytwarzający łańcuch A połączony z łańcuchem B insuliny ludzkiej dipeptydem ArgArg (Tabela 3, L.p. przykładu 3.1.1. i 3.2.1. Badania analityczne, prowadzone w poszczególnych etapach procesu wykazują, że wyższą wydajność w tym wariantach systemu ekspresji – o 14%–16% – osiąga się po etapie wytwarzania ciał inkluzyjnych i renaturacji, przy czym utrzymuje się ona przez wszystkie pozostałe etapy procesu.

Przykład 4. Insulina ludzka SerA22ArgB31 (insulina SR).

Badania wydajności procesu wytwarzania insuliny SR przeprowadzono zgodnie z powyżej podaną ogólną procedurą (punkt I), realizując etapy 1–6, 6b, 7–10, 12–15 i 17–19 oraz dla systemu ekspresji 4.2. dodatkowo etap 11 (Tabela 2. Insulina GR). Proces badano w 12 wariantach 5 systemów ekspresji białka (nie stosowano wariantu z łącznikiem LysArg, jako nieprowadzącego do insuliny SR). Otrzymane dane potwierdzają poprzednie wyniki, stale wykazując wyższą wydajność systemu ekspresji *E. coli* IBA/pIBA, wytwarzającego łańcuch A połączony z łańcuchem B insuliny SR dipeptydem ArgArg (Tabela 4, L.p. przykładu 4.1.1. i 4.2.1.) w porównaniu z innymi peptydami łącznikowymi tego systemu – co najmniej o 15%.

Przykład 5. Insulina SerA22LizB3ArgB31 (insulina SK3R)

Zgodnie z powyżej podaną ogólną procedurą prowadzenia procesu, przeprowadzono wytwarzanie biosyntetycznej insuliny SK3R w 6 wariantach 2 systemów ekspresji białka. W badaniach prowadzono etapy 1–6, 6b, 7–10, 12–15 i 17–19 oraz dla systemu ekspresji 5.2. dodatkowo etap 11. Z otrzymanych danych wynika (Tabela 5. Insulina SK3R), że najlepszą wydajnością ekspresji charakteryzuje się system szczepów IBA/pIBA, wykazujących się dobrymi, porównywalnymi parametrami wzrostu w przemysłowej pożywce minimalnej i różniących się masą wytworzonych ciał inkluzyjnych oraz względną wydajnością renaturacji. Ponownie najwyższą wydajność insuliny ludzkiej uzyskano wykorzystując system ekspresji *E. coli* IBA/pIBA wytwarzający łańcuch A połączony z łańcuchem B insuliny ludzkiej dipeptydem ArgArg (Tabela 5, L.p. przykładu 5.1.1. i 5.2.1. Wyczerpujące badania analityczne, prowadzone w poszczególnych etapach procesu wykazują, że wyższa wydajność w tych wariantach systemu ekspresji utrzymuje się przez wszystkie etapy procesu po etapie wytwarzania ciał inkluzyjnych i renaturacji, wydajność ta jest wyższa od uzyskanej z innych wariantów systemu ekspresji przynajmniej o 14%–16%.

Przykład 6. Preproinsulina (prekursor) G: białko liderowe – łańcuch B – łącznik – łańcuch A:dezAsnA21GlyA21.

Uniwersalność przewagi łącznika ArgArg nad innymi w ekspresji preproinsuliny wykazano również badając wydajność wytwarzania białka składającego się z białka liderowego poprzedzającego łańcuch B, który był połączony badanymi łącznikami z modyfikowanym łańcuchem A insuliny (A:dezAsnA21GlyA21). Badania zrealizowano przeprowadzając krytyczne etapy 1–8 ogólnej procedury prowadzenia procesu w 8 wariantach 2 systemów ekspresji. Wyniki badań, przedstawione w Tabeli 6. Preproinsulina (prekursor) G: białko liderowe – łańcuch B – łącznik – łańcuch A:dezAsnA21GlyA21 wykazują, że w przypadku łącznika ArgArg wydajność procesu jest wyższa o 10%, niż w pozostałych przypadkach.

Na podstawie końcowej wydajności produktu, potwierdzonej zawartością białka przed oczyszczeniem, nieoczekiwanie okazało się, że najlepsze wyniki w wytwarzaniu rekombinowanej insuliny ludzkiej i jej analogów osiąga się w systemie ekspresji składającym się z plazmidu ekspresyjnego z promotorem

deoP1P2 w szczepach bakterii *Escherichia coli* pozbawionych genu represora *cytR* (patent polski nr PL 180 818), jeżeli w plazmidzie kodującym białko fuzyjne łańcuch A insuliny jest połączony z łańcuchem B za pomocą dipeptydu ArgArg. Co więcej, różnica w wydajności między wariantem systemu z dipeptydem ArgArg a LysArg wynosi 13%–22% i jest porównywalna z opisanym w patencie polskim nr PL 180 818 postępowaniem w wydajności systemu z dipeptydem LysArg i Arg w porównaniu z systemem z peptydem C, wynoszącym odpowiednio 23% i 17%.

Konstrukcja genetyczna plazmidu pIBA i pIBA/INS [pIBA/INSLys(31B)Arg(32B)]

Plazmid pIBA/IMS o wielkości 4354 par zasad jest zbudowany z następujących sekwencji regulatorowych oraz genów:

- od 7 pz do 942 pz zlokalizowany jest region regulatorowy z promotorami *deoP1P2*
- od 946 pz do 1137 pz znajduje się sekwencja kodująca fragment ludzkiego genu SOD,
- od 1138 pz do 1299 pz zawarta jest sekwencja kodująca łańcuchy A, B oraz łańcuch C (w postaci łącznika Lys-Arg) rekombinowanej ludzkiej insuliny INS, sekwencja nukleotydowa insuliny została zmieniona zgodnie z częstością występowania kodonów. w *E. coli*,
- od 1312 pz do 1338 pz do zlokalizowany jest region kodujący terminator transkrypcji Ter *trpA*,
- od 1508 pz do 2698 pz znajduje się gen oporności na tetracyklinę Tet,

Strukturę plazmidu pIBA przedstawiono schematycznie na Fig. 1, a jego sekwencję nukleotydową i aminokwasową; na Fig. 2. Strukturę plazmidu pIBA/INS zawierającego przykładowy gen kodujący białko rekombinowanej insuliny ludzkiej, w której łańcuch A jest połączony z łańcuchem B dipeptydem LysArg przedstawiono schematycznie na Fig. 3, a jego sekwencję nukleotydową i aminokwasową na Fig. 4.

SEQUENCE LISTING

<110> Instytut Biotechnologii i Antybiotyków

<120> Sposób wytwarzania insuliny i jej pochodnych oraz peptyd
hybrydowy stosowany w tym sposobie

<130> PK/3515/RW

<160> 29

<170> PatentIn version 3.5

<210> 1

<211> 21

<212> PRT

<213> artificial

<220>

<223> łańcuch A

<400> 1

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
1 5 10 15

Glu Asn Tyr Cys Asn
20

<210> 2

<211> 22

<212> PRT

<213> artificial

<220>

<223> łańcuch A

<400> 2

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
1 5 10 15

Glu Asn Tyr Cys Asn Gly
20

<210> 3

<211> 22

<212> PRT

<213> artificial

<220>

<223> łańcuch A

<400> 3

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	

Glu	Asn	Tyr	Cys	Asn	Ser
				20	

<210> 4

<211> 21

<212> PRT

<213> artificial

<220>

<223> łańcuch A

<400> 4

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	

Glu	Asn	Tyr	Cys	Gly
				20

<210> 5

<211> 30

<212> PRT

<213> artificial

<220>

<223> łańcuch B

<400> 5

Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5					10					15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr
			20					25					30

<210> 6

<211> 30

<212> PRT

<213> artificial

<220>

<223> łańcuch B

<400> 6

Phe	Val	Lys	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5				10						15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr
			20				25						30

<210> 7

<211> 30

<212> PRT

<213> artificial

<220>

<223> łańcuch B

<400> 7

Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5				10						15	

Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Lys	Pro	Thr
			20				25						30

<210> 8

<211> 64

<212> PRT

<213> artificial

<220>

<223> SOD

<400> 8

Met	Ala	Thr	Lys	Ala	Val	Ser	Val	Leu	Lys	Gly	Asp	Gly	Pro	Val	Gln
1				5				10						15	

Gly	Ile	Ile	Asn	Phe	Glu	Gln	Lys	Glu	Ser	Asn	Gly	Pro	Val	Lys	Val
			20				25					30			

Trp	Gly	Ser	Ile	Lys	Gly	Leu	Thr	Glu	Gly	Leu	His	Gly	Phe	His	Val
		35				40					45				

His	Glu	Phe	Gly	Asp	Asn	Thr	Ala	Gly	Ser	Thr	Ser	Ala	Gly	Pro	Arg
	50					55					60				

<210> 9
 <211> 76
 <212> PRT
 <213> artificial

<220>
 <223> ubi

<400> 9

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
 1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
 20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
 50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
 65 70 75

<210> 10
 <211> 117
 <212> PRT
 <213> artificial

<220>
 <223> X-B-R-R-A (SOD)

<400> 10

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
 1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
 20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
 35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
50 55 60

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
100 105 110

Glu Asn Tyr Cys Asn
115

<210> 11
<211> 117
<212> PRT
<213> artificial

<220>
<223> X-B-R-R-A (SOD) 2

<400> 11

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
50 55 60

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Lys Pro Thr Arg Arg
85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
 100 105 110

Glu Asn Tyr Cys Asn
 115

<210> 12
 <211> 118
 <212> PRT
 <213> artificial

<220>
 <223> X-B-R-R-A (SOD)

<400> 12

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
 1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
 20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
 35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
 50 55 60

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
 85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
 100 105 110

Glu Asn Tyr Cys Asn Gly
 115

<210> 13
 <211> 118
 <212> PRT

<213> artificial

<220>

<223> X-B-R-R-A (SOD)

<400> 13

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
50 55 60

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
100 105 110

Glu Asn Tyr Cys Asn Ser
115

<210> 14

<211> 118

<212> PRT

<213> artificial

<220>

<223> X-B-R-R-A (SOD)

<400> 14

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
50 55 60

Phe Val Lys Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
100 105 110

Glu Asn Tyr Cys Asn Ser
115

```
<210> 15
<211> 117
<212> PRT
<213> artificial
```

<220>
<223> X-B-R-R-A (SOD)

<400> 15

Met Ala Thr Lys Ala Val Ser Val Leu Lys Gly Asp Gly Pro Val Gln
1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Ser Thr Ser Ala Gly Pro Arg
50 55 60

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
65 70 75 80

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
85 90 95

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
100 105 110

Glu Asn Tyr Cys Gly
115

<210> 16
<211> 129
<212> PRT
<213> artificial

<220>
<223> X-B-R-R-A (UBI)

<400> 16

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Phe Val Asn Gln
65 70 75 80

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
85 90 95

Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Gly Ile Val Glu
100 105 110

Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
115 120 125

Asn

<210> 17
 <211> 129
 <212> PRT
 <213> artificial

 <220>
 <223> X-B-R-R-A (UBI)

<400> 17

Met	Gln	Ile	Phe	Val	Lys	Thr	Leu	Thr	Gly	Lys	Thr	Ile	Thr	Leu	Glu
1				5					10					15	

Val	Glu	Ser	Ser	Asp	Thr	Ile	Asp	Asn	Val	Lys	Ser	Lys	Ile	Gln	Asp
			20					25					30		

Lys	Glu	Gly	Ile	Pro	Pro	Asp	Gln	Gln	Arg	Leu	Ile	Phe	Ala	Gly	Lys
		35					40					45			

Gln	Leu	Glu	Asp	Gly	Arg	Thr	Leu	Ser	Asp	Tyr	Asn	Ile	Gln	Lys	Glu
	50					55					60				

Ser	Thr	Leu	His	Leu	Val	Leu	Arg	Leu	Arg	Gly	Gly	Phe	Val	Asn	Gln
65					70					75					80

His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr	Leu	Val	Cys	Gly
			85						90					95	

Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Lys	Pro	Thr	Arg	Arg	Gly	Ile	Val	Glu
			100					105					110		

Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu	Glu	Asn	Tyr	Cys
		115					120					125			

Asn

<210> 18

<211> 130

<212> PRT

<213> artificial

<220>

<223> X-B-R-R-A (UBI)

<400> 18

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Phe Val Asn Gln
65 70 75 80

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
85 90 95

Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Gly Ile Val Glu
100 105 110

Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
115 120 125

Asn Gly
130

<210> 19

<211> 130

<212> PRT

<213> artificial

<220>

<223> X-B-R-R-A (UBI)

<400> 19

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Phe Val Asn Gln
65 70 75 80

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
85 90 95

Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Gly Ile Val Glu
100 105 110

Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
115 120 125

Asn Ser
130

<210> 20
<211> 130
<212> PRT
<213> artificial

<220>
<223> X-B-R-R-A (UBI)

<400> 20

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Phe Val Lys Gln
65 70 75 80

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
85 90 95

Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Gly Ile Val Glu
100 105 110

Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
115 120 125

Asn Ser
130

<210> 21
<211> 129
<212> PRT
<213> artificial

<220>
<223> X-B-R-R-A (UBI)

<400> 21

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ser Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Phe Val Asn Gln
65 70 75 80

His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
85 90 95

Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Gly Ile Val Glu
100 105 110

Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
115 120 125

Gly

<210> 22
<211> 53
<212> PRT
<213> artificial

<220>
<223> B-R-R-A

<400> 22

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
35 40 45

Glu Asn Tyr Cys Asn
50

<210> 23
<211> 53
<212> PRT
<213> artificial

<220>
<223> B-R-R-A

<400> 23

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Lys Pro Thr Arg Arg
20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
35 40 45

Glu Asn Tyr Cys Asn
50

<210> 24

<211> 54

<212> PRT

<213> artificial

<220>

<223> B-R-R-A 3

<400> 24

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
35 40 45

Glu Asn Tyr Cys Asn Gly
50

<210> 25

<211> 54

<212> PRT

<213> artificial

<220>

<223> B-R-R-A

<400> 25

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
 20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
 35 40 45

Glu Asn Tyr Cys Asn Ser
 50

<210> 26
 <211> 54
 <212> PRT
 <213> artificial

<220>
 <223> B-R-R-A

<400> 26

Phe Val Lys Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
 20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
 35 40 45

Glu Asn Tyr Cys Asn Ser
 50

<210> 27
 <211> 53
 <212> PRT
 <213> artificial

<220>
 <223> B-R-R-A

<400> 27

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg
 20 25 30

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
 35 40 45

Glu Asn Tyr Cys Gly
 50

<210> 28
 <211> 4003
 <212> DNA
 <213> artificial

<220>
 <223> pIBA

<400> 28
 gaattcggac cgttggcgat gtgcgggttg ctacattcac agatgttctt cgccacttcc
 60

agcagcaggt catcaggggt gatttcagga tcgtagataa aggtcaggtt cgggtgaaacc
 120

tgtttcaact ctgcatctgc acgtaagatc gcgcgggtaa tgggcgaatc agacgggccg
 180

atattggcgt gcataaaggc gtctggcagg gttctgtcga ggtaacgcc aaaacgtttt
 240

attcgaacat cgatctcgtc ttgtgttaga attaattcta acatacgggt gcaacaacgc
 300

atccagttgc cccaggtaga ccggcatcga tgtgaccgac ggtacgtggt ggtaaagaat
 360

ggtcagcaga gagagtgcgt catcaagatc ttctggcgct tccagctcca gccattcgga
 420

accgttcgcc agaaaacggg cgtaatcggg taagacatag cgcgggtttgt acggcgcgatg
 480

accttcaaac atatcgaga ttacaccttc atccagcgcg cggcgggctt cggcaggaag
 540

ctgtgggtaa ggcagattgt tttctgcttc cagtgccaga aaatggcgct tctgctccgg
 600

gctaagcact gggctggtga caatttgctg gcaacgttgt tgcagtgcac tttcatgaga
660

agtgggcac ttcttttctt tttatgccga aggtgatgag ccattgtaag aagtttcgtg
720

atgttcactt tgatcctgat gcgtttgccg ccaactgacg atccatttga aagtgaatta
780

tttgaaccag atcgcatcac agtgatgcaa acttgtaagt agatttcctt aattgtgatg
840

tgtatcgaag tgtgttgagg agtagatggt agaatactaa caaactcgca aggtgaattt
900

tattggcgac aagcctagggt ttgtttaact ttaaggagaa atcatatgtc tagaatcgat
960

gcccgtttaa tgagcgggct tttttttctc gaggaagctc aagaaaccat tattatcatg
1020

acattaacct ataaaaatag gcgtatcacg aggccctttc gtcttcaaga cctctcatgt
1080

ttgacagctt atcatcgata agctttaatg cggtagttta tcacagttaa attgctaacg
1140

cagtcaggca ccgtgtatga aatctaacaa tgcgctcatc gtcacctcgc gcaccgtcac
1200

cctggatgct gtaggcatag gcttggttat gccggtaact ccgggcctct tgcgggatat
1260

cgtccattcc gacagcatcg ccagtcacta tggcgtgctg ctacgctat atgcgttgat
1320

gcaatttcta tgcgcacccg ttctcggagc actgtccgac cgctttggcc gccgccagc
1380

cctgctcgtc tcgctacttg gagccactat cgactacgag atcatggcga ccacaccgt
1440

cctgtggatc ctctacgccc gacgcatcgt ggcgggcac accggcgcca caggtgcggt
1500

tgctggcgcc tatatcgccc acatcaccca tggggaagat cgggctcgcc acttcgggct
1560

catgagcgtc tgtttcggcg tgggtatggt ggcaggcccc gtggccgggg gactgttggg
1620

cgccatctcc ttgcatgcac cattccttgc ggcggcggtg ctcaacggcc tcaacctact
1680

actgggctgc ttctaatgc aggagtcgca taaggagag cgtcgaccga tgcccttgag
1740

agccttcaac ccagtcagct ccttcgggtg ggcgcggggc atgactatcg tcgccgcaact
1800

tatgactgtc ttctttatca tgcaactcgt aggacagggtg ccggcagcgc tctgggccat
1860

tttcggcgag gaccgctttc gctggagcgc gacgatgatc ggctgtcgc ttgcggtatt
1920

cggaatcttg cagccctcg ctcaagcctt cgtaactggt cccgccacca aacgtttcgg
1980

cgagaagcag gccattatcg ccggcatggc ggccgacgcg ctgggctacg tcttgctggc
2040

gttcgcgacg cgaggctgga tggccttccc cattatgatt ctctcgcctt ccggcggcat
2100

cgggatgcc gcgttgacgg ccattgctgc caggcaggta gatgacgacc atcagggaca
2160

gcttcaagga tcgctcgcg ctcttaccag cctaacttcg atcattggac cgctgatcgt
2220

cacggcgatt tatgccgctt cggcgagcac atggaacggg ttggcatgga ttgtaggcgc
2280

cgccctatac cttgtctgcc tccccgcgtt ggcgcgggt gcattggagcc gggccacctc
2340

gacctgaatg gaagccggcg gcacctcgt aacggattca ccaactccaag aattggagcc
2400

aatcaattct tgcggagaac tgtgaatgcg caaaccaacc cttggcagaa catatccatc
2460

gcgtccgcca tctccagcag ccgcacgcgg cgcattctcg gcagcgttgg gtccctggccc
2520

gggactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca
2580

gtgctgccat aaccatgagt gataaactg cggccaactt acctctgaca acgatcggag
2640

gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc
2700

gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg
2760

cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc
2820

ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg
2880

cccttcgggc tggctggttt attgctgata aatctggagc cggtgagcgt gggctctcgcg
2940

gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga
3000

cggggagtcg ggcaactatg gatgaacgaa atagacagat cgetgagata ggtgcctcac
3060

tgattaagca ttggtaactg tcagaccaag tttactcata tatacttttag attgatttaa
3120

aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca
3180

aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag
3240

gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac
3300

cgtaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa
3360

ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagttaggcc
3420

accacttcaa gaactctgta gcaccgcta catacctcgc tctgctaata ctgttaccag
3480

tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac
3540

cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc
3600

gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc
3660

ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca
3720

cgagggagct tccaggggga aacgcctggc atctttatag tctgtcggg tttcgccacc
3780

tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg geggagccta tggaaaaacg
3840

ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct cacatgttct
3900

ttcctgcgtt atccccgat tctgtggata accgtattac cgcctttgag tgagctgata
3960

cgcctcgccg cagccgaacg accgagcgca gcgagtcagt gag
4003

<210> 29

<211> 4354

<212> DNA

<213> artificial

<220>

<223> pIBA/INSLys (31B) Arg (32B)

<400> 29

gaattcggac cgttggcgat gtgcggtttg ctacattcac agatgttctt cgccacttcc
60

agcagcaggt catcaggggt gatttcagga tcgtagataa aggtcaggtt cggtgaaacc
120

tgtttcaact ctgcatctgc acgtaagatc gcgcgggtaa tgggcgaatc agacgggccg
180

atattggcgt gcataaaggc gtctggcagg gttctgtcga ggtaacgcca gaaacgtttt
240

attcgaacat cgatctcgtc ttgtgttaga attaattcta acatacgggt gcaacaacgc
300

atccagttgc cccaggtaga ccggcatcga tgtgaccgac ggtacgtggt ggtaaagaat
360

ggtcagcaga gagagtgcgt catcaagatc tttcgcgcct tccagctcca gccattcgga
420

accgttcgcc agaaaacggg cgtaatcggg taagacatag cgcgggtttgt acggcgcgatg
480

accttcaaac atatcgaga ttacaccttc atccagcgcg cggcgggctt cggcaggaag
540

ctgtgggtaa ggcagattgt tttctgcttc cagtgccaga aaatggcgct tctgctccgg
600

gctaagcact gggctggtga caatttgctg gcaacgtttgt tgcagtgcac tttcatgaga
660

agtgggcac tttttttcct tttatgccga aggtgatgcg ccattgtaag aagtttcgtg
720

atgttcactt tgatcctgat gcgtttgcc aactgacgc attcatttga aagtgaatta
780

tttgaaccag atcgcattac agtgatgcaa acttgtaagt agatttcctt aattgtgatg
840

tgtatcgaag tgtgttgagg agtagatggt agaatactaa caaactcgca aggtgaattt
900

tattggcgac aagcctaggt ttgtttaact ttaaggagaa atcatatggc tactaaagct
960

gtttctgttt taaaagggtga tgggtccagtt caaggaatta ttaattttga aaaaaagaa
1020

agtaatggac cagttaaagt atggggaagt attaaaggac ttactgaagg cctgcatgga
1080

ttccatgttc atgagtttgg agataatata gcaggcagta ctagtgcagg tctcgtttt
1140

gtcaaccagc acctgtgtgg ttctcacctg gttgaagcac tgtacctggt atgtggcgaa
1200

cgtgggtttct tctacactcc taaaacaaag cgcggcatcg ttgaacagtg ctgtacctct
1260

atctgttccc tgtaccaact ggagaactac tgcaattaat ctagaatcga tgcccgtta
1320

atgagcgggc ttttttttct cgaggacgtc taagaaacca ttattatcat gacattaacc
1380

tataaaaata ggcgtatcac gaggcccttt cgtcttcaag acctctcatg ttgacagct
1440

tatcatcgat aagctttaat gcggtagttt atcacagtta aattgctaac gcagtcaggc
1500

accgtgtatg aaatctaaca atgcgctcat cgtcatcctc ggcaccgtca ccctggatgc
1560

tgtaggcata ggcttgggtta tgccggtact gccgggcctc ttgcgggata tcgtccattc
1620

cgacagcatc gccagtcact atggcgtgct gctagcgcta tatgcgttga tgcaatttct
1680

atgcgcaccc gttctcggag cactgtccga ccgctttggc cgcgcgccag tctgtctgc
1740

ttcgtactt ggaqccacta tcgactacgc gatcatggcg accacacccg tctgtggat
1800

cctctacgcc ggacgcatcg tggccggcat caccggcgcc acaggtgcgg ttgctggcgc
1860

ctatatcgcc gacatcaccg atggggaaga tggggtcgc cacttcgggc tcatgagcgc
1920

ttgtttcggc gtgggtatgg tggcaggccc cgtggccggg ggactgttgg ggcacatctc
1980

cttgcatgca ccattccttg cggcggcggt gctcaacggc ctcaacctac tactgggctg
2040

cttcctaatag caggagtgcg ataagggaga ggcgcgaccg atgcccttga gagccttcaa
2100

cccagtcagc tccttcgggt gggcgggggg catgactatc gtcgccgcac ttatgactgt
2160

cttctttatc atgcaactcg taggacaggt gccggcagcg ctctgggcca ttttcggcga
2220

ggaccgcctt cgctggagcg cgacgatgat cggcctgtcg cttgcggtat tcggaatctt
2280

gcacgccctc gctcaagcct tcgtcactgg tcccgccacc aaacgtttcg gcgagaagca
2340

ggccattatc gccggcatgg cggccgacgc gctgggctac gtcttgctgg cgttcgcgac
2400

gcgaggtcgg atggccttcc ccattatgat tcttctcgct tccggcggca tcgggatgcc
2460

cgcgttgacg gccatgctgt ccaggcaggt agatgacgac catcaggac agcttcaagg
2520

atcgctcgcg gctcttacca gcctaacttc gatcattgga ccgctgatcg tcacggcgat
2580

ttatgccgcc tcggcgagca catggaacgg gttggcatgg atgtaggcg ccgccctata
2640

ccttgtctgc ctccccggt tgcgtcgcgg tgcattggag cgggccacct cgacctgaat
2700

ggaagccggc ggcacctcgc taacggattc accactccaa gaattggagc caatcaattc
2760

ttgcggagaa ctgtgaatgc gcaaaccaac ccttggcaga acatatccat cgcgtccgcc
2820

atctccagca gccgcacgcg gcgcattctg ggcagcgttg ggcctggcc cgggactcac
2880

cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc agtgctgcc
2940

taaccatgag tgataaact gcggccaact tacttctgac aacgatcgga ggaccgaagg
3000

agctaaccgc ttttttgac aacatggggg atcatgtaac tcgccttgat cgttggaac
3060

cggagctgaa tgaagccata ccaaacgacg agcgtgacac cagcatgcct gcagcaatgg
3120

caacaacggt gcgcaacta ttaactggcg aactacttac tctagcttcc cggcaacaat
3180

taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg gcccttcgg
3240

ctggctggtt tattgctgat aaatctggag cgggtgagcg tgggtctcgc ggtatcattg
3300

cagcaactgg gccagatggt aagccctccc gtatcgtagt tatctacacg acggggagtc
3360

aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca ctgattaagc
3420

attggtaact gtcagaccaa gttactcat atatacttta gattgattta aaacttcatt
3480

tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt
3540

aacgtgagtt ttcgttccac tgagcgtcag acccgtaga aaagatcaaa gcatcttctt
3600

gagatccttt tttctcgcg gtaatctgct gcttgcaaac aaaaaacca ccgctaccag
3660

cgggtggtttg tttgcggat caagagctac caactctttt tcogaaggta actggcttca
3720

gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc caccattca
3780

agaactctgt agcaccgcct acatacctcg ctctgctaata cctgttacca gtggctgctg
3840

ccagtggcga taagtctgt cttaccgggt tggactcaag acgatagtta ccggataagg
3900

cgcagcggtc qggctgaacg qggggttcgt gcacacagcc cagcttgag cgaacqacct
3960

acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt cccgaaggga
4020

gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc
4080

ttccaggggg aaacgcctgg tatctttata gtctgtcgg gtttcgccac ctctgacttg
4140

agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg
4200

cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc tttcctgcgt
4260

tatccccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgcc
4320

gcagccgaac gaccgagcgc agcgagtcag tgag
4354

Zastrzeżenia patentowe

1. Sposób wytwarzania insuliny ludzkiej obejmujący mikrobiologiczną ekspresję rekombinowanego białka zawierającego insulinę, **znamienny tym**, że obejmuje:
 - (a) otrzymywanie polipeptydu hybrydowego zawierającego proinsulinę poprzez ekspresję tego polipeptydu hybrydowego w komórce szczepu bakterii *Escherichia coli* pozbawionej represorów *deo* i/lub genu *deo* (*cytR*) zawierającej DNA kodujący polipeptyd hybrydowy, przy czym ten DNA jest obecny w wektorze pod kontrolą promotora *deoP₁P₂* natomiast polipeptyd hybrydowy zawiera łańcuch B insuliny kowalencyjnie związany z łańcuchem A dipeptydem Arg-Arg,
 - (b) odzyskiwanie polipeptydu hybrydowego,
 - (c) ufałdowanie i tworzenie wiązań dwusiarczkowych w polipeptydzie hybrydowym otrzymanym w (b), unikając uprzedniego poddawania polipeptydu hybrydowego siarczynolizie,
 - (d) poddawanie uzyskanego w (c) ufałdowanego, z utworzonymi wiązaniami dwusiarczkowymi polipeptydu hybrydowego trawieniu enzymatycznemu w celu otrzymania insuliny,
 - (e) oczyszczania tak wytworzonej insuliny.
2. Sposób według zastrz. 1, **znamienny tym**, że peptydem hybrydowym jest peptyd posiadający sekwencję aminokwasową o wzorze ogólnym:



gdzie :

A oznacza polipeptyd łańcucha A insuliny lub jej analogu, korzystnie o sekwencji wybranej spośród Sekw. Nr Id.: 1–4,

B oznacza polipeptyd łańcucha B insuliny lub jej analogu, korzystnie o sekwencji wybranej spośród Sekw. Nr Id.: 5–6,

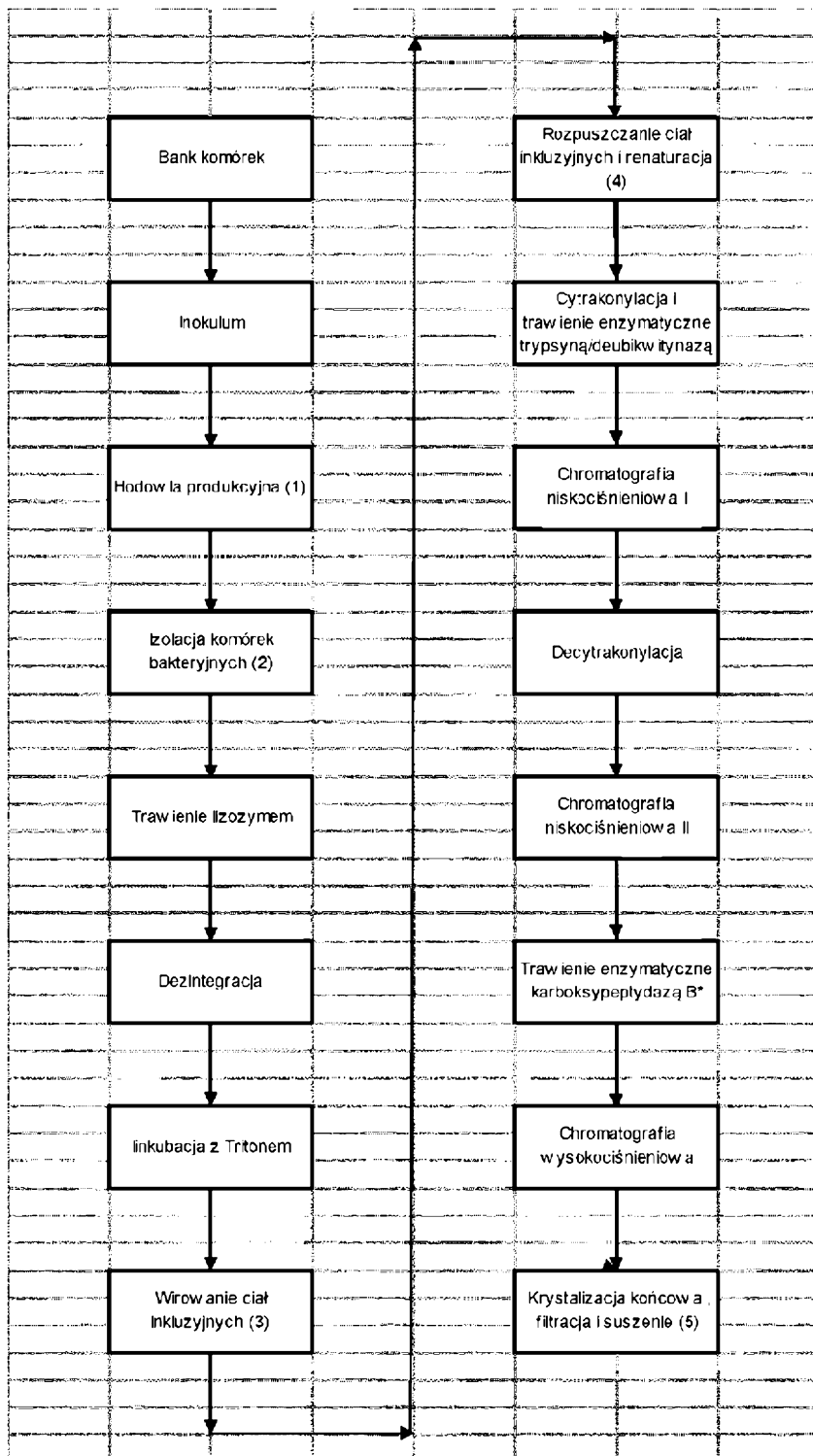
n oznacza 0 lub 1,

X oznacza polipeptyd białka liderowego, korzystnie o sekwencji wybranej spośród SOD o Sekw. Nr Id.:12 lub UBI o Sekw. Nr Id.:13.

3. Sposób według zastrz. 1, **znamienny tym**, że peptyd hybrydowy ma sekwencję aminokwasową wybraną spośród: Sekw. Nr Id.: 8–26.
4. Sposób według zastrz. 1, **znamienny tym**, że w etapie (a) ekspresję prowadzi się w szczepie *E. coli* o genotypie:
 $F^- \text{ ara } \Delta \text{ (pro lac) rpsL thi cyt R rec A } \lambda^-$,
 pozbawionym aktywnego receptora *cyt R*.

5. Sposób według zastrz. 1, **znamienny tym**, że w etapie (a) jako wektor stosuje się plazmid *p/BA* o sekwencji Sekw Nr Id: 27.
6. Sposób według zastrz. 1, **znamienny tym**, że etap (a) obejmuje fermentację w obecności glukozy, glicerolu lub galaktozy.
7. Sposób według zastrz. 1, **znamienny tym**, że etap (b) obejmuje:
 - (i) niszczenie ściany komórkowej komórki bakteryjnej lub jej fragmentów w celu utworzenia lizatu, korzystnie poprzez działanie lizozymu i Tritonu X100 a następnie sonifikację lub dezintegrację;
 - (ii) izolowanie wewnątrzkomórkowego precypitatu ciałek inkluzyjnych z lizatu przez wirowanie; i
 - (iii) rozpuszczanie precypitatu ciałek inkluzyjnych.
8. Sposób według zastrz. 1, **znamienny tym**, że etap (c) obejmuje inkubację polipeptydu hybrydowego, korzystnie w obecności kwasu askorbinowego, zwłaszcza w stężeniu około 2 moli na mol grup SH obecnych w mieszaninie, w temperaturze od 4 do 37°C przez około od 1 do 30 godzin, korzystnie około 5 godzin, w pH od 8,5 do 12, korzystnie w pH od 11,0 do 11,25.
9. Sposób według zastrz. 1, **znamienny tym**, że etap (d) obejmuje ponadto:
 - (i) doprowadzanie pH do około 8,8–9,0; i
 - (ii) trawienie polipeptydu hybrydowego trypsyną i następnie ewentualnie karboksypeptydazą B w temperaturze 16–37°C w ciągu 30 minut do 16 godzin.
10. Sposób według zastrz. 1, **znamienny tym**, że etap (e) obejmuje ponadto oczyszczanie roztworu przez niskociśnieniową chromatografią cieczową na Sefarozie Q przed trawieniem trypsyną.
11. Sposób według zastrz. 1, **znamienny tym**, że etap (e) obejmuje ponadto oczyszczanie roztworu niskociśnieniową chromatografią na DEAE-Sepharose przed trawieniem karboksypeptydazą B.
12. Sposób według zastrz. 1, **znamienny tym**, że etap (e) obejmuje ponadto oczyszczanie uzyskanej insuliny wysokosprawną chromatografią cieczową, korzystnie do czystości co najmniej 98%, a następnie krystalizację, filtrację i suszenie.
13. Sposób według zastrz. 1, **znamienny tym**, że w etapie (e) oczyszczaną insuliną jest białko wybrane z grupy do której należy: insulina ludzka, insulina ludzka Ala^{A22}Lys^{B3}Arg^{B31}, insulina ludzka Ser^{A22}Lys^{B3}Arg^{B31} insulina ludzka Ser^{A22}Lys^{B3}Arg^{B31}, insulina ludzka Lys^{B28}Pro^{B29} oraz insulina ludzka Asp^{B28}.

Schemat 1



- (1) Kontrola gęstości optycznej OD_{600} (Optical Density)
 (2) Masa hodowli (bakterii) po odwirowaniu (lub po odwirowaniu i wysuszeniu)
 (3) Masa czystych ciał inkluzyjnych po odwirowaniu (lub po odwirowaniu i wysuszeniu)
 (4) Białko całkowite (po foldingu, metoda spektrofotometryczna lub HPLC)
 (5) Wydajność końcowa produktu
 w przypadku prowadzenia operacji

Tabela 1

1. insulina ludzka LysB28ProB29											
System E.coli/wektor charakterystyka	Łącznik	{1} OD	{2} Masa hodowli		{3} Masa ciał inkluzyjnych		{4} Względny udział uformowanego prekursora w białku całkowitym		{5} Wydajność końcowa		Lp. przykładu
			g/dm ³	% najwyższej wartości	g/dm ³	% najwyższej wartości	% (HPLC)	% najwyższej wartości	mg/dm ³	% najwyższej wartości	
1.1. IBA/pIBA białko liderowe - SOD, terminator transkrypcji - terA Trp	ArgArg	55	55,0	100%	13,8	100%	7,0	100%	39	100%	1.1.1.
	LysArg	54	55,0	100%	12,8	93%	5,7	81%	33	85%	1.1.2.
	Arg	51	54,4	99%	12,2	89%	5,7	82%	29	74%	1.1.3.
	Peptyd C	53	55,0	100%	12,8	93%	5,9	85%	23	59%	1.1.4.
1.2. IBA/pIBA białko liderowe - UBI, terminator transkrypcji - terA Trp	ArgArg	51	51,2	98%	12,9	100%	7,0	100%	35	100%	1.2.1.
	LysArg	51	51,2	98%	11,6	90%	5,7	81%	31	87%	1.2.2.
	Arg	48	51,1	98%	11,9	92%	5,7	82%	27	75%	1.2.3.
	Peptyd C	48	52,0	100%	12,9	100%	6,6	94%	21	60%	1.2.4.
1.3. DH5α/pIGAL białko liderowe - SOD, oporność: ampicylina, terminator transkrypcji - ST1	ArgArg	16	5,3		0,4		0,4		6**	-	1.3.1.
	LysArg	15	4,2		0,3		0,4		.*	-	1.3.2.
1.4. HB101/pIGAL białko liderowe - UBI, oporność: ampicylina, terminator transkrypcji - ST1	ArgArg	12	3,1		0,2		~0,2		.*		1.4.1.
	LysArg	12	3,1		0,2		~0,2		.*		1.4.2.
1.5. DH5/pIGAL białko liderowe - UBI, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	28	4,8		0,5		0,5		11**		1.5.1
	LysArg	29	5,0		0,5		0,5		9**		1.5.2
1.6. DH5/pIGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - terA Trp	ArgArg	32	5,3		0,7		0,6		17**		1.6.1.
	LysArg	31	5,2		0,6		0,6		16**		1.6.2.
1.7. DH5/pIBA białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - terA Trp	ArgArg	22	3,3		.*		.*		.*		1.7.1.
	LysArg	20	3,2		.*		.*		.*		1.7.2.
1.8. IBA/pIGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	34	7,8		0,6		0,4		.*		1.8.1.
	LysArg	35	7,8		0,6		0,4		.*		1.8.2.

* - proces przerwano na etapie oczyszczania ze względu na niską zawartość produktu.

** - niska i niereprezentatywna wydajność wynika również z niedostosowania skali procesu oczyszczania do niskiej zawartości produktu.

Tabela 2

2. insulina ludzka GlyA22ArgB31 (insulina GR)											
System E.coli/wektor charakterystyka	Łącznik	(1) OD	(2) Sucha masa bakteryjna		(3) Sucha masa ciał inkluzyjnych		(4) Białko całkowite (prekursor) (A _{2aa})		(5) Wydajność końcowa		Lp. przykładu
			g/dm ³	% najwyższej wartości	g/dm ³	% najwyższej wartości	AU/dm ³	% najwyższej wartości	mg/dm ³	% najwyższej wartości	
2.1. IBA/pIBA białko liderowe - SOD, terminator transkrypcji - terA Trp	ArgArg	62	17,6	100%	8,8	100%	7860	100%	40	100%	2.1.1.
	Arg	59	17,4	99%	8,0	91%	6890	88%	31	78%	2.1.2.
	Peptyd C	62	17,5	99%	7,8	89%	6686	85%	24	61%	2.1.3.
2.2. IBA/pIBA białko liderowe - UBI, terminator transkrypcji - terA Trp	ArgArg	68	17,9	99%	9,0	100%	7891	100%	40	100%	2.2.1.
	Arg	67	17,3	96%	8,1	90%	6714	85%	34	84%	2.2.2.
	Peptyd C	65	18,0	100%	8,3	92%	6890	88%	25	63%	2.2.3.
2.3. DHS/pIGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	32	5,3		0,8				.*		2.3.1.
	Arg	32	5,3		0,7				.*		2.3.2.
2.4. DHS/pIBA białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji -	ArgArg	26	4		.*				.*		2.4.1.
	Arg	28	4		.*				.*		2.4.2.
2.5. IBA/pIGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	39	8,3		0,9				.*		2.5.1.
	Arg	39	8,4		0,9				.*		2.5.2.

Tabela 3

System E.coli/wektor charakterystyka	Łącznik	(1) OD 600	(2) Masa bakteryjna		(2) Sucha masa bakteryjna		(3) Masa ciał inkluzyjnych		(3) Sucha masa ciał inkluzyjnych		(4) Białko całkowite (prekursor), A280		Białko całkowite (po biotrans- formacji), A280		Białko całkowite (po DEAE), A280		Białko całkowite (po OI), A280		Białko całkowite (insulina do CPB), A280		Białko całkowite: (insulina po HPLC), metoda Bradforda, A555		(5) Wydajność końcowa	Lp. przy kładu	
			g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	mg/dm ³	% naj- wyższej wartości					
System E.coli/wektor charakterystyka	Łącznik	(1) OD 600	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	g/dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	AU/ dm ³	% naj- wyższej wartości	mg/dm ³	% naj- wyższej wartości	mg/dm ³	% naj- wyższej wartości	(5) Wydajność końcowa	Lp. przy kładu	
białko liderowe - SOD, terminator transkrypcji - terA	ArgArg	59	54,7	100%	26,9	100%	29,3	100%	14,3	100%	10070	100%	10133	100%	1411	100%	571	100%	471	100%	104	100%	82	100%	3.1.1
	LysArg	61	55,0	100%	27,0	100%	27,5	94%	13,7	96%	8459	84%	8410	83%	1157	82%	468	82%	396	84%	87	85%	68	83%	3.1.2
	Arg	59	53,6	98%	26,3	98%	24,8	90%	13,2	92%	8358	83%	8511	84%	1160	82%	465	82%	391	83%	85	83%	67	83%	3.1.3
Irr	PeptidC	53	54,7	100%	26,9	100%	22,9	92%	13,3	93%	7553	75%	7701	76%	1086	77%	457	80%	354	75%	76	74%	59	72%	3.1.4
	ArgArg	50	55,0	100%	27,0	100%	30,1	100%	14,3	100%	10012	100%	9039	100%	1128	100%	456	100%	368	100%	88	100%	72	100%	3.2.1
	LysArg	48	55,0	100%	27,0	100%	28,3	94%	13,0	91%	8210	82%	7683	85%	936	83%	378	83%	316	86%	73	83%	61	85%	3.2.2
Irr	Arg	45	54,7	100%	26,9	100%	26,8	89%	13,3	93%	8203	82%	7593	84%	925	82%	374	82%	305	83%	76	85%	61	84%	3.2.3
	PeptidC	44	52,5	96%	25,8	96%	27,1	91%	13,6	95%	6888	69%	6598	73%	790	78%	379	71%	265	77%	64	74%	54	74%	3.2.4

Tabela 4

4. insulina ludzka SerA22ArgB31 (insulina SR)											
System E.coli/wektor charakterystyka	Łącznik	(1) OD	(2) Sucha masa bakteryjna		(3) Sucha masa ciał inkluzyjnych		(4) Białko całkowite (prekursor) [A280]		(5) Wydajność końcowa		Lp. przykładu
			g/dm ³	% najwyższej wartości	g/dm ³	% najwyższej wartości	AU/dm ³	% najwyższej wartości	mg/dm ³	% najwyższej wartości	
4.1. IBA/pIBA białko liderowe - SOD, terminator transkrypcji - terA Trp	ArgArg	64	17,4	100%	8,8	100%	7860	100%	48	100%	4.1.1.
	Arg	68	17,5	99%	8,1	92%	6890	88%	41	85%	4.1.2.
	Peptyd C	64	17,3	99%	8,2	93%	6686	85%	36	75%	4.1.3.
4.2. IBA/pIBA białko liderowe - UBI, terminator transkrypcji - terA Trp	ArgArg	68	17,9	99%	9,0	100%	7891	100%	46	100%	4.2.1.
	Arg	67	17,3	99%	8,1	90%	6714	85%	39	85%	4.2.2.
	Peptyd C	65	18,0	100%	8,3	92%	6890	88%	35	75%	4.2.3.
4.3. DH5/pGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	32	5,3						.*		4.3.1
	Arg	32	5,3						.*		4.3.2
4.4. DH5/pIBA białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - terA Trp	ArgArg	26	4						.*		4.4.1
	Arg	28	4						.*		4.4.2
4.5. IBA/pGAL białko liderowe - SOD, oporność: tetracyklina, terminator transkrypcji - ST1	ArgArg	39	8,3						.*		4.5.1
	Arg	39	8,4						.*		4.5.2

Tabela 5

5. insulina ludzka SerA22LysB3ArgB31 (SK3R)																			
System E.coli/wektor charakterystyka	Łącznik	(1) OD 600	(2) Masa bakteryjna		(3) Masa ciał inkluzyjnych		(4) Białko całkowite [prekursor], A280		Białko całkowite (po biotransformacji), A280		Białko całkowite (po DEAE), A280		Białko całkowite (po Q), A280		Białko całkowite (insulina po HPLC), metoda Bradforda, A595		(5) Wydajność końcowa		lp. przykłądu
															g/dm3	% naj-wyższej wartości			
5.1. IBA/pIBA białko liderowe - SOD, terminator transkrypcji - terA Trp	ArgArg	57	52,3	100%	23,0	100%	7857	100%	7602	100%	1050	100%	326	100%	93	100%	78	100%	5.1.1.
	Arg	56	52,3	100%	21,2	92%	6757	86%	6386	84%	861	82%	274	84%	77	83%	66	84%	5.1.2.
	Peptyd C	59	51,7	99%	21,6	94%	6600	84%	5778	76%	809	77%	261	80%	68	74%	56	72%	5.1.3.
5.2. IBA/pIBA białko liderowe - UBI, terminator transkrypcji - terA Trp	ArgArg	55	51,8	97%	23,5	100%	7720	100%	7587	100%	1059	100%	332	100%	88	100%	72	100%	5.2.1.
	Arg	55	52,0	98%	22,1	94%	6330	82%	6373	84%	868	82%	272	82%	75	85%	62	86%	5.2.2.
	Peptyd C	56	53,3	100%	22,3	95%	6639	86%	5918	78%	773	73%	249	75%	64	73%	55	76%	5.2.3.

Tabela 6

6. Preproinsulina (prekursor) G: białko liderowe – łańcuch B – łącznik – łańcuch A:dezAsnA21GlyA21									
System E.coli/wektor charakterystyka	łącznik	{1} OD	(2) Sucha masa bakteryjna		(3) Sucha masa ciał inkluzyjnych		(4) Białko całkowite (prekursor) [A280]		Lp. przykładu
			g/dm ³	% najwyższej wartości	g/dm ³	% najwyższej wartości	AU/dm ³	% najwyższej wartości	
6.1. IBA/pIBA białko liderowe - SOD, terminator transkrypcji - terA Trp	ArgArg	58	17,4	99%	8,4	100%	6960	100%	6.1.1.
	LysArg	60	17,3	99%	7,9	94%	6264	90%	6.1.2.
	Arg	60	17,5	100%	7,8	93%	5916	85%	6.1.3.
	Peptyd C	55	17,3	99%	8,1	96%	6334	91%	6.1.4.
6.2. IBA/pIBA białko liderowe - UBI, terminator transkrypcji - terA Trp	ArgArg	55	17,0	100%	8,7	100%	7012	100%	6.2.1.
	LysArg	55	16,9	99%	8,1	93%	6381	91%	6.2.2.
	Arg	57	17,0	100%	8,0	92%	5890	84%	6.2.3.
	Peptyd C	54	17,0	100%	8,4	97%	6451	92%	6.2.4.

Rysunki

mapa wektora pIBA

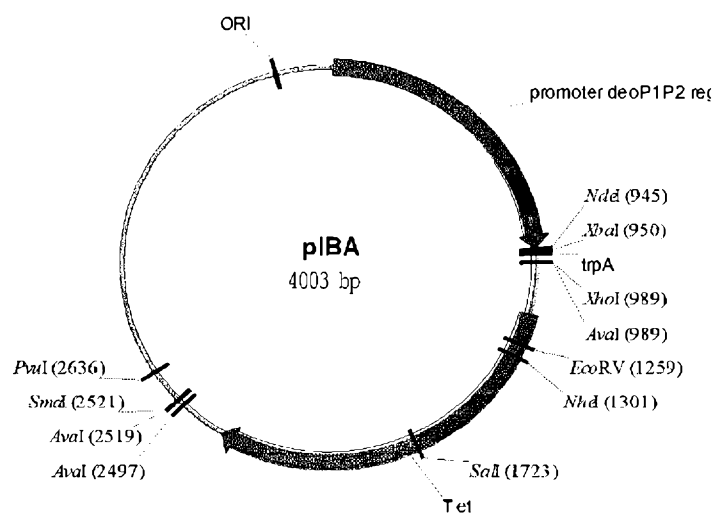


Fig. 1

Fig. 2

sekwencja nukleotydowa wektora pIBA

```

1 GAATTCGGAC CGTTGGCGAT GTGCGGTTTG CTACATTCAC AGATGTTCTT
  CTTAAGCCTG GCAACCGCTA CACGCCAAAC GATGTAAGTG TCTACAAGAA
51 CGCCACTTCC AGCAGCAGGT CATCAGGGGT GATTTCAGGA TCGTAGATAA
  GCGGTGAAGG TCGTCGTCCA GTAGTCCCCA CTAAAGTCCT AGCATCTATT
101 AGGTCAGGTT CGGTGAAACC TGCTTCAACT CTGCATCTGC ACGTAAGATC
  TCCAGTCCAA GCCACTTTGG ACGAAGTTGA GACGTAGACG TGCATTCTAG
151 GCGCGGGTAA TGGGCGAATC AGACGGGCGG ATATGGGCGT GCATAAAGGC
  CGCGCCCATT ACCCGCTTAG TCTGCCCGGC TATAACCGCA CGTATTTCCG
201 GTCTGGCAGG GTTCTGTCTGA GGTAAACGCCA GAAACGTTTT ATTCTGAACAT
  CAGACCGTCC CAAGACAGCT CCATTGCGGT CTTTGCAAAA TAAGCTTGTA
251 CGATCTCGTC TTGTGTTAGA ATTAATTCTA ACATACGGTT GCAACAACGC
  GCTAGAGCAG AACACAATCT TAATTAAGAT TGTATGCCAA CGTTGTTGGC
301 ATCCAGTTGC CCCAGGTAGA CCGGCATCGA TGTGACCGAC GGTACGTGGT
  TAGTCAACG GGGTCCATCT GGCCGTAGCT ACGTGGCTG CCATGCACCA
351 GGTAAAGAAAT GGTGAGCAGA GAGAGTGGGT CATCAAGATC TTTGCGCGCT
  CCATTTCCTTA CCAGTCGTCT CTCTCACGCA GTAGTCTTAG AAAGCGCGGA
401 TCCAGCTCCA GCCATTGCGA ACCGTTGCGC AGAAAACGGG CGTAATCGGG
  AGGTCGAGGT CGGTAAGCCT TGGCAAGCGG TCTTTTGCCC GCATTAGCCC
451 TAAGACATAG CGCGGTTTGT ACGGCGCATG ACCTTCAAAC ATATCGCAGA
  ATTCTGTATC GCGCCAAACA TGCCGCGTAC TGGAAAGTTG TATAGCGTCT
501 TTACACCTTC ATCCAGCGCG CGGCGGGCTT CGGCAGGAAG CTGTGGGTAA
  AATGTGGAAG TAGGTCGCGC GCCGCCCGAA GCCGTCCTTC GACACCCATT
551 GGCAGATTGT TTTCTGCTTC CAGTGCCAGA AAATGGCGCT TCTGCTCCGG
  CCGTCTAACA AAAGACGAAG GTCACGGTCT TTTACCGCGA AGACGAGGCC
601 GCTAAGCACT GGGCTGGTGA CAATTGCTG GCAACGTTGT TGCAGTGCAT
  CGATTCTGTA CCCGACCACT GTTAAACGAC CGTTGCAACA ACGTCACGTA
651 TTTTCATGAGA AGTGGGCATC TTCTTTTCCT TTTATGCCGA AGGTGATGCG
  AAAGTACTCT TCACCCGTAG AAGAAAAGGA AAATACGGCT TCCACTACGC
701 CCATTGTAAG AAGTTTCGTG ATGTTCACTT TGATCCTGAT GCGTTTGCCA
  GGTAAACATT TTCAAAGCAC TACAAGTGAA ACTAGGACTA CGCAAACGGT
751 CCACTGACGC ATTCATTTGA AAGTGAATTA TTTGAACCAG ATCGCATTAC
  GGTGACTGCG TAAGTAAACT TTCACCTAAT AAATTTGGTC TAGCGTAATG
801 AGTGATGCAA ACTTGTAAGT AGATTTCCCT AATGTGTATG TGTATCGAAG
  TCACTACGTT TGAACATTCA TCTAAAGGAA TTAACACTAC ACATAGCTTC
851 TGTGTTGCGG AGTAGATGTT AGAATACTAA CAAACTCGCA AGGTGAATTT
  ACACAACGCC TCATCTACAA TCTTATGATT GTTTGAGCGT TCCACTTAAA
901 TATTTGGCGAC AAGCCTAGGT TTGTTTAACT TTAAGGAGAA ATCATATGTC
  ATAACCGCTG TTCGGATCCA ACAAATTTGA AATTCCTCTT TAGTATACAG
951 TAGAATCGAT GCCCGCTTAA TGAGCGGGCT TTTTTCCTCT GAGGACGCT
  ATCTTAGCTA CGGGCGAATT ACTCGCCCGA AAAAAAAGAG CTCCTGCAGA
1001 AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG
  TTCTTTGGTA ATAATAGTAC TGTAAATTGA TATTTTATC CGCATAGTGC
1051 AGGCCCTTTC GTCTTCAAGA CCTCTCATGT TTGACAGCTT ATCATCGATA
  TCCGGGAAAG CAGAAGTTCT GGAGAGTACA AACTGTCGAA TAGTAGCTAT
1101 AGCTTTAATG CGGTAGTTTA TCACAGTTAA ATTGCTAACG CAGTCAGGCA
  TCGAAATTAC GCCATCAAAT AGTGTCAATT TAACGATTGC GTCAGTCCGT
1151 CCGTGTATGA AATCTAACAA TGCGCTCATC GTCATCCTCG GCACCGTCAC
  GGCACATACT TTAGATTGTT ACGCGAGTAG CAGTAGGAGC CGTGGCAGTG
1201 CCTGGATGCT GTAGGCATAG GCTTGGTTAT GCCGCTACTG CCGGCCTCT
  GGACCTACGA CATCCGTATC CGAACCAATA CGGCCATGAC GGGCCGGAGA
1251 TGGGGGATAT CGTCCATTCC GACAGCATCG CCAGTCACTA TGGCGTGTG
  ACGCCCTATA GCAGGTAAGG CTGTCGTAGC GGTGAGTGAT ACCGCACGAC
1301 CTAGCGCTAT ATGCGTTGAT GCAATTTCTA TGCGCACCCG TTCTCGGAGC
  GATCGCGATA TACGCAACTA CGTTAAAGAT ACGCGTGGGC AAGAGCCTCG
1351 ACTGTCCGAC CGCTTTGGCC GCCGCCAGT CCTGCTCGCT TCGCTACTTG
  TGCAGGGCTG GCGAAACCGG CGGCGGGTCA GGACGAGCGA AGCGATGAAC
1401 GAGCCACTAT CCACTACGCG ATCATGGCGA CCACACCCGT CCTGTGGATC
  CTCGGTGATA GCTGATGCGC TAGTACCGCT GGTGTGGGCA GGACACCTAG
1451 CTCTACGCCG GACGCATCGT GGCCGGCATC ACCGGCGCCA CAGGTGCGGT
  GAGATGCGGC CTGCGTAGCA CCGGCCGTAG TGGCCGCGGT GTCCACGCCA
1501 TGCTGGCGCC TATATCGCCG ACATCACCAG TGGGGAAGAT CCGGCTCGCC
  ACGACCGCGG ATATAGCGGC TGTAGTGGCT ACCCTTCTTA GCGCGAGCGG
1551 ACTTCGGGCT CATGAGCGCT TGTTCGGCG TGGGTATGGT GGCAGGCCCC

```

TGAAGCCCGA GTACTCGCGA ACAAAGCCGC ACCCATACCA CCGTCCGGGG
 1601 GTGGCCGGGG GACTGTTGGG CGCCATCTCC TTGCATGCAC CATTCCTTGC
 CACCGGCCCC CTGACAACCC GCGGTAGAGG AACGTACGTG GTAAGGAACG
 1651 GGGCGCGGTG CTCAACGGCC TCAACCTACT ACTGGGCTGC TTCCTAATGC
 CCGCCGCCAC GAGTTGCCGG AGTTGGATGA TGACCCGACG AAGGATTACG
 1701 AGGAGTCGCA TAAGGGAGAG CGTCGACCGA TGCCCTTGAG AGCCTTCAAC
 TCCTCAGCGT ATTCCCTCTC GCAGCTGGCT ACGGGAACCTC TCGGAAGTTG
 1751 CCAGTCAGCT CCTTCCGGTG GGGCGGGGGC ATGACTATCG TCGCCGCACCT
 GGTCACTCGA GGAAGGCCAC CCGCGCCCCG TACTGATAGC AGCGGCGTGA
 1801 TATGACTGTC TTCTTTATCA TGCAACTCGT AGGACAGGTG CCGGCAGCGC
 ATACTGACAG AAGAAATAGT ACGTTGAGCA TCCTGTCCAC GGCCGTGCGG
 1851 TCTGGGCCAT TTTGGGCGAG GACCGCTTTC GCTGGAGCGC GACGATGATC
 AGACCCGGTA AAAGCCGCTC CTGGCGAAAG CGACCTCGCG CTGCTACTAG
 1901 GGCCTGTGCG TTGCGGTATT CGGAATCTTG CACGCCCTCG CTCGAGCCTT
 CCGGACAGCG AACGCCATAA GCCTTAGAAC GTGCGGGAGC GAGTTCCGGA
 1951 CGTCACTGGT CCGGCCACCA AACGTTTCGG CGAGAAGCAG GCCATTATCG
 GCAGTGACCA GGGCGGTGGT TTGCAAAGCC GCTCTTCGTC CGGTAATAGC
 2001 CCGGCATGGC GGGCGACGCG CTGGGCTACG TCTTGTGGG GTTCCGACG
 GGCCGTACCG CCGGCTGCGC GACCCGATGC AGAACGACCG CAAGCGCTGC
 2051 CGAGGCTGGA TGGCCTTCCC CATTATGATT CTTCTCGCTT CCGCGGGCAT
 GCTCCGACCT ACCGGAAGGG GTAATACTAA GAAGAGCGAA GGCCGCCGTA
 2101 CGGGATGCCC GCGTTGCAGG CCATGCTGTC CAGGCAGGTA GATGACGACC
 GCCCTACGGG CGCAACGTCC GGTACGACAG GTCCGTCCAT CTACTCTGG
 2151 ATCAGGGACA GCTTCAAGGA TCGCTCGCGG CTCTTACCAG CCTAACTTCG
 TAGTCCCTGT CGAAGTTCCCT AGCGAGCGCC GAGAATGGTC GGATTGAAGC
 2201 ATCATTGGAC CGCTGATCGT CACGGCGATT TATGCCGCTT CCGCGAGCAC
 TAGTAACCTG GCGACTAGCA GTGCCGCTAA ATACGGCGGA GCCCTGCTG
 2251 ATGGAACGGG TTGGCATGGA TTGTAGGCGC CGCCCTATAC CTTGTCTGCC
 TACCTTGCCC AACCTTACCT AACATCCGCG GCGGGATATG GAACAGACGG
 2301 TCCCCGCGTT GCGTCGCGGT GCATGGAGCC GGGCCACCTC GACCTGAATG
 AGGGGCGCAA CGCAGCGCCA CGTACCTCGG CCGGTGGAG CTGGACTTAC
 2351 GAAGCCGCGC GCACCTCGCT AACGGATTCA CCACTCCAAG AATTGGAGCC
 CTTGCGCGCG CGTGGAGCGA TTGCCTAAGT GGTGAGGTTT TTAACCTCGG
 2401 AATCAATTCT TGCGGAGAAC TGTGAATGCG CAAACCAACC CTTGGCAGAA
 TTAGTTAAGA ACGCCTCTTG ACACCTACGC GTTTGGTTGG GAACCGCTT
 2451 CATATCCATC GCGTCCGCCA TCTCCAGCAG CCGCACGCGG CGCATCTCGG
 GTATAGGTAG CGCAGGCGGT AGAGGTCGTC GCGGTGCGCC GCGTAGAGCC
 2501 GCAGCGTTGG GTCTTGGCCC GGGACTCACC AGTCACAGAA AAGCATCTTA
 CGTCGCAACC CAGGACCGGG CCTGAGTGG TCAGTGTCTT TTCGTAGAA
 2551 CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT
 GCCTACCGTA CTGTCAATTCT CTTAATACGT CACGACGGTA TTGGTACTCA
 2601 GATAACACTG CCGGCAACTT ACTTCTGACA ACGATCGGAG GACCGAAGGA
 CTATTGTGAC GCGGGTTGAA TGAAGACTGT TGCTAGCCTC CTGGCTTCTT
 2651 GCTAACCGCT TTTTTCACAC ACATGGGGGA TCATGTAACCT CGCCTTGATC
 CGATTGGCGA AAAAACGTGT TGTACCCCTT AGTACATTGA GCGGAACCTAG
 2701 GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA GCGTGACACC
 CAACCTTGG CCTCGACTTA CTTCCGTATG GTTTGCTGCT CGCACTGTGG
 2751 ACGATGCCGT CAGCAATGGC AACACGTTG CGCAAACCTAT TAACCTGGCGA
 TGCTACGGAC GTCGTTACCG TTGTTGCAAC GCGTTTGATA ATTGACCGCT
 2801 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCGG
 TGATGAATGA GATCGAAGGG CCGTTGTAA TTAATCTGACC TACCTCCGCC
 2851 ATAAAGTTGC AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT
 TATTTCAACG TCCTGGTGAA GACGCGAGCC GGGAAGGCCG ACCGACCAA
 2901 ATTGCTGATA AATCTGGAGC CCGTGAGCGT GGGTCTCGCG GTATCATTGC
 TAACGACTAT TTAGACCTCG GCCACTCGCA CCCAGAGCGC CATAGTAACG
 2951 AGCACTGGGG CCAGATGGTA AGCCCTCCCG TATCGTAGTT ATCTACGCA
 TCGTGACCCC GGTCTACCAT TCGGGAGGGC ATAGCATCAA TAGATGTGCT
 3001 CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT CGCTGAGATA
 GCGGCTCAGT CCGTTGATAC CTACTTGCTT TATCTGTCTA GCGACTCTAT
 3051 GGTGCCCTAC TGATTAAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA
 CCACGGAGTG ACTAATTCGT AACCATGAC AGTCTGGTTC AAATGAGTAT
 3101 TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG
 ATATGAAATC TAACATAAAT TTGAAGTAAA AATTAAATTT TCCTAGATCC
 3151 TGAAGATCCT TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT
 ACTTCTAGGA AAAACTATTA GAGTACTGGT TTTAGGGAAT TGCCTCAAA
 3201 TCGTTCCACT GAGCGTCAGA CCGCGTAGAA AAGATCAAAG GATCTTCTTG
 AGCAAGGTGA CTCGAGTCT GGGGCATCTT TTCTAGTTTC CTAGAAGAAC
 3251 AGATCCTTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAAACCAC

```

TCTAGGAAAA AAAGACGCGC ATTAGACGAC GAACGTTTGT TTTTGTGGTG
3301 CGCTACCAGC GGTGGTTTGT TTGCCGGATC AAGAGCTACC AACTCTTTT
GCGATGGTCG CCACCAAACA AACGGCCTAG TTCTCGATGG TTGAGAAAAA
3351 CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT
GGCTTCCATT GACCGAAGTC GTCTCGCGTC TATGGTTTAT GACAGGAAGA
3401 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACCTCTGA GCACCGCCTA
TCACATCGGC ATCAATCCGG TGGTGAAGTT CTTGAGACAT CGTGGCGGAT
3451 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT
GTATGGAGCG AGACGATTAG GACAATGGTC ACCGACGACG GTCACCGCTA
3501 AAGTCGTGTC TTACCGGGTT GGA CTCAAGA CGATAGTTAC CGGATAAGGC
TTACGACACAG AATGGCCCCA CCTGAGTTCT GCTATCAATG GCCTATTCCG
3551 GCAGCGGTCG GCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC
CGTCGCCAGC CCGACTTGCC CCCCAGCAC GTGTGTCGGG TCGAACCTCG
3601 GAACGACCTA CACCGAAGTG AGATACCTAC AGCGTGAGCT ATGAGAAAGC
CTTGCTGGAT GTGGCTTGAC TCTATGGATG TCGCACTCGA TACTCTTTCG
3651 GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG
CGGTGCGAAG GGCTTCCCTC TTTCCGCGTG TCCATAGGCC ATTCGCCGTC
3701 GGTGCGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT
CCAGCCTTGT CCTCTCGCGT GCTCCCTCGA AGGTCCCCCT TTGCGGACCA
3751 ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT
TAGAAATATC AGGACAGCCC AAAGCGGTGG AGACTGAACT CGCAGCTAAA
3801 TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC
AACACTACGA GCAGTCCCCC CGCCTCGGAT ACCTTTTTCG GGTGTTGCG
3851 GGCTTTTFTA CGGTTCCTGG CCTTTTGCTG GCCTTTTGCT CACATGTTCT
CCGAAAAAAT GCCAAGGACC GGAACGAC CGGAAACGA GTGTACAAGA
3901 TTCCTGCGTT ATCCCTGAT TCTGTGGATA ACCGTATTAC CGCCTTTGAG
AAGGACGCAA TAGGGGACTA AGACACCTAT TGGCATAATG GCGGAAACTC
3951 TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGASTCAGT
ACTCGACTAT GGCGAGCGGC GTCGGCTTGC TGGCTCGCGT CGCTCAGTCA
4001 GAG
CTC

```

Fig. 2

Struktura plazmidu pIBA/INSLys(31B)Arg(32B) kodującego białko rekombinowanej lizarginsuliny INS.

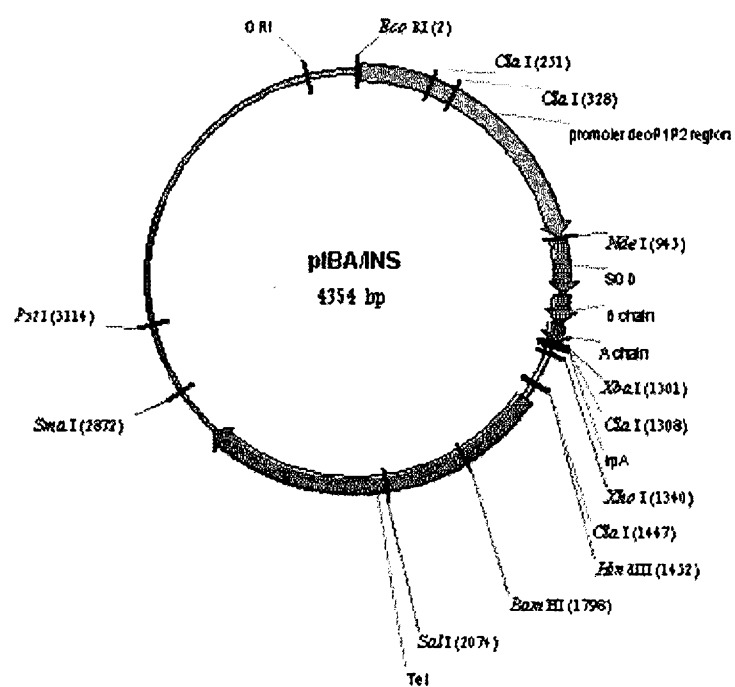


Fig. 3

Fig. 4

Sekuencja nukleotydowa i aminokwasowa plazmidu pIBA/INSLys(31B)Arg(32B).

1	GAATTCGGAC	CGTTGGCGAT	GTGCGGTTTG	CTACATTCAC	AGATGTTCTT
	CTTAAGCCTG	GCAACCGCTA	CACGCCAAAC	GATGTAAGTG	TCTACAAGAA
51	CGCCACTTCC	AGCAGCAGGT	CATCAGGGGT	GATTTTCAGGA	TCGTAGATAA
	GCGGTGAAGG	TCGTGCTCCA	GTAGTCCCCA	CTAAAGTCCT	AGCATCTATT
101	AGGTCAAGTT	CGGTGAAACC	TGCTTCAACT	CTGCATCTGC	ACGTAAGATC
	TCCAGTCCAA	GCCACTTTTG	ACGAAGTTGA	GACGTAGACG	TGCATTCTAG
151	GCGCGGGTAA	TGGGCGAATC	AGACGGGCCG	ATATTGGCGT	GCATAAAGGC
	CGCGCCCAT	ACCCGCTTAG	TCTGCCCGGC	TATAACCGCA	CGTATTTCCG
201	GTCTGGCAGG	GTTCTGTCTGA	GGTAACGCCA	GAAACGTTTT	ATTCTGAACAT
	CAGACCGTCC	CAAGACAGCT	CCATTGCGGT	CTTTGCAAAA	TAAGCTTGTA
251	CGATCTCGTC	TTGTGTTAGA	ATTAATTCTA	ACATACGGTT	GCAACAACGC
	GCTAGAGCAG	AACACAATCT	TAATTAAGAT	TGTATGCCAA	CGTTGTTGCG
301	ATCCAGTTGC	CCCAGGTAGA	CCGGCATCGA	TGTGACCGAC	GGTACGTGGT
	TAGGTCAACG	GGGTCCATCT	GGCCGTAGCT	ACACTGGCTG	CCATGCACCA
351	GGTAAAGAAT	GGTCAGCAGA	GAGAGTGCGT	CATCAAGATC	TTTCGCGCCT
	CCATTTCTTA	CCAGTCGTCT	CTCTCACGCA	GTAGTTCTAG	AAAGCGCGGA
401	TCCAGCTCCA	GCCATTGCGA	ACCGTTCGCC	AGAAAACGGG	CGTAATCGGG
	AGGTGAGGTT	CGGTAAGCCT	TGGCAAGCGG	TCTTTTGCCC	GCATTAGCCC
451	TAAGACATAG	CGCGGTTTGT	ACGGCGCATG	ACCTTCAAAC	ATATCGCAGA
	ATTCTGTATC	GCGCCAAACA	TGCCGCGTAC	TGGAAGTTTG	TATAGCGTCT
501	TTACACCTTC	ATCCAGCGCG	CGGCGGGCTT	CGGCAGGAAG	CTGTGGGTAA
	AATGTGGAAG	TAGGTCGCGC	GCCGCCCGAA	GCCGTCTTTC	GACACCCATT
551	GGCAGATTGT	TTTCTGCTTC	CAGTGCCAGA	AAATGGCGCT	TCTGCTCCGG
	CCGTCTAACA	AAAGACGAAG	GTCACGGTCT	TTTACCGCGA	AGACGAGGCC
601	GCTAAGCACT	GGGCTGGTGA	CAATTGCTG	GCAACGTTGT	TGCAGTGCAT
	CGATTTCGTGA	CCCGACCACT	GTTAAACGAC	CGTTGCAACA	ACGTCACGTA
651	TTTCATGAGA	AGTGGGCATC	TTCTTTTCCT	TTTATGCCGA	AGGTGATGCG
	AAAGTACTCT	TCACCCGTAG	AAGAAAAGGA	AAATACGGCT	TCCACTACGC
701	CCATTGTAAG	AAGTTTCGTG	ATGTTCACTT	TGATCCTGAT	GCGTTTGCCA
	GGTAACATT	TTCAAAGCAC	TACAAGTGAA	ACTAGGACTA	CGCAACCGGT
751	CCACTGACGC	ATTCAATTTGA	AAGTGAATTA	TTTGAACCCAG	ATCGCATTAC
	GGTGACTGCG	TAAGTAAACT	TTCACTTAAT	AACTTGGTTC	TAGCGTAATG
801	AGTGATGCAA	ACTTGTAAGT	AGATTTCCTT	AATTGTGATG	TGTATCGAAG
	TCACTACGTT	TGAACATTCA	TCTAAAGGAA	TTAACACTAC	ACATAGCTTC
851	TGTGTTGCGG	AGTAGATGTT	AGAATACTAA	CAAACCTCGCA	AGGTGAATTT
	ACACAACGCC	TCATCTACAA	TCTTATGATT	GTTTGAGCGT	TCCACTTAAA
					MetAla
901	TATTGGCGAC	AAGCCTAGGT	TTGTTTAACT	TTAAGGAGAA	ATCATATGGC
	ATAACCGCTG	TTCGATCCA	AACAAATTGA	AATTCCTCTT	TAGTATACCG
951	·TACTAAAGCT	·GTTTCTGTTT	·TAAAAGGTGA	·TGGTCCAGTT	·CAAGGAATTA
	·ATGATTTTCCA	·CAAAGACAAA	·ATTTTCCACT	·ACCAGGTCAA	·GTTCCCTAAT
1001	·IASnPheGlu	·GlnLysGlu	·SerAsnGlyPro	·ValLysVal	·TrpGlySer
	·TTAATTTTGA	·ACAAAAAGAA	·AGTAATGGAC	·CAGTTAAAGT	·ATGGGGAGGT
	·AATTAAAACT	·TGTTTTTCTT	·TCATTACCTG	·GTCAATTTC	·TACCCCTTCA
1051	·IleLysGlyLeu	·ThrGluGly	·LeuHisGly	·PheHisValHis	·GluPheGly
	·ATTAAAGGAC	·TTACTGAAGG	·CCTGCATGGA	·TTCCATGTTT	·ATGAGTTTGG
	·TAATTTTCTG	·AATGACTTCC	·GGACGTACCT	·AAGGTACAAG	·TACTCAAACC
1101	·AspAsnThr	·AlaGlySerThr	·SerAlaGly	·ProArgPhe	·ValAsnGlnHis
	·AGATAATACA	·GCAGGCAGTA	·CTAGTGCAGG	·TCCTCGTTTT	·GTCAACCAGC
	·TCTATTATGT	·CGTCCGTCAT	·GATCACGTCC	·AGGAGCAAAA	·CAGTTGGTCG
1151	·HLeuCysGly	·SerHisLeu	·ValGluAlaLeu	·TyrLeuVal	·CysGlyGlu
	·ACCTGTGTGG	·TTCTCACCTG	·GTTGAAGCAC	·TGTACCTGGT	·ATGTGGCGAA
	·TGGACACACC	·AAGAGTGGAC	·CAACTTCGTG	·ACATGGACCA	·TACACCGCTT
1201	·ArgGlyPhePhe	·TyrThrPro	·LysThrLys	·ArgGlyIleVal	·GluGlnCys
	·CGTGGTTTCT	·TCTACACTCC	·TAAACAAAG	·CGCGGCATCG	·TTGAACAGTG
	·GCACCAAAGA	·AGATGTGAGG	·ATTTTGTTC	·GCGCCGTAGC	·AACTTGTCAC
1251	·CysThrSer	·IleCysSerLeu	·TyrGlnLeu	·GluAsnTyr	·CysAsn***
	·CTGTACCTCT	·ATCTGTTCCC	·TGTACCAACT	·GGAGAACTAC	·TGCAATTAAT
	·GACATGGAGA	·TAGACAAGGG	·ACATGGTTGA	·CCTCTTGATG	·ACGTTAATTA
1301	·CTAGAATCGA	·TGCCCGCTTA	·ATGAGCGGGC	·TTTTTTTTCT	·CGAGGACGTC
	·GATCTTAGCT	·ACGGGCGAAT	·TACTCGCCCG	·AAAAAAAGA	·GCTCCTGCAG

1351 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GCGGTATCAC
 ATTCTTTGGT AATAATAGTA CTGTAATTGG ATATTTTAT CCGCATAGTG
 1401 GAGGCCCTTT CGTCTTCAAG ACCTCTCATG TTTGACAGCT TATCATCGAT
 CTCGGGGA A GCAGAAAGTTC TGGAGAGTAC AAACGTCTGA ATAGTAGCTA
 1451 AAGCTTTAAT GCGGTAGTTT ATCACAGTTA AATTGCTAAC GCAGTCAGGC
 TTCGAAATTA CGCCATCAAA TAGTGTC AAT TAAACGATTG CGTCAGTCCG
 1501 ACCGTGTATG AAATCTAACA ATGCGCTCAT CGTCATCCTC GGCACCGTCA
 TGGCACATAC TTTAGATTGT TACGCGAGTA GCAGTAGGAG CCGTGGCAGT
 1551 CCCTGGATGC TGTAGGCATA GGCTTGGTTA TGCCGGTACT GCCGGGCTC
 GGGACCTACG ACATCCGTAT CCGAACCAAT ACGGCCATGA CGGCCCGGAG
 1601 TTGCGGGATA TCGTCCATTC CGACAGCATC GCCAGTCACT ATGGCGTGCT
 AACGCCCTAT AGCAGGTAAG GCTGTCTAGT CCGTCAGTGA TACCGCACGA
 1651 GCTAGCGCTA TATGCGTTGA TGCAATTTCT ATGCGCACCC GTTCTCGGAG
 CGATCGCGAT ATACGCAACT ACGTTAAAGA TACGCGTGGG CAAGACCTC
 1701 CACTGTCCGA CCGCTTTGGC CGCCGCCAG TCCTGCTCGC TTCGTACTT
 GTGACAGGCT GCGCAAACCG GCGGCGGGT AGGACGAGCG AAGCGATGAA
 1751 GGAGCCACTA TCGACTACGC GATCATGGCG ACCACACCCG TCCTGTGGAT
 CCTCGTGAT AGCTGATGCG CTAGTACCGC TGGTGTGGG AGGACACTA
 1801 CCTCTACGCC GGACGCATCG TGCCCGCAT CACCGCGGCC ACAGGTGCGG
 GGAGATGCGG CCTGCGTAGC ACCGGCCGTA GTGGCCGCGG TGTCACGCC
 1851 TTGCTGGCGC CTATATCGCC GACATCACCG ATGGGGAAGA TCGGCTCGC
 AACGACCGCG GATATAGCGG CTGTAGTGGC TACCCCTTCT AGCCCGAGCG
 1901 CACTTCGGG TCATGAGCGC TTGTTTCGGC GTGGGTATGG TGGCAGGCC
 GTGAAGCCCG AGTACTCGCG AACAAAGCCG CACCCATACC ACCGTCCGGG
 1951 CGTGGCCGGG GGACTGTTGG GCGCCATCTC CTTGCATGCA CCATTCTTG
 GCACCGGCCC CCTGACAACC CGCGGTAGAG GAACGTACGT GGTAAAGGAA
 2001 CGGCGGCGGT GCTCAACGGC CTCAACCTAC TACTGGGCTG CTTCTCAATG
 GCCGCCGCCA CGAGTTGCCG GAGTTGGATG ATGACCCGAC GAAGGATTAC
 2051 CAGGAGTCGC ATAAAGGAGA GCGTCGACCG ATGCCCTTGA GAGCCTTCAA
 GTCCTCAGCG TATTCCTCT CTGAGCTGGC TACGGGAACCT CTCGGAAGTT
 2101 CCCAGTCAGC TCCTTCGGT GGGCGCGGG CATGACTATC GTCCCGCAC
 GGGTCAGTCG AGGAAGGCCA CCCGCCGCC GACTGATAG CAGCGGCGTG
 2151 TTATGACTGT CTTCTTTATC ATGCAACTCG TAGGACAGGT GCCCGCAGCG
 AATACTGACA GAAGAAATAG TACGTTGAGC ATCCTGTCCA CGGCCGTCG
 2201 CTCTGGGCTA TTTTCGGCA GGACCGCTT CGCTGGAGCG GCACGATGAT
 GAGACCGGT AAAAGCCGCT CCTGGCGAAA GCGACCTCGC GCTGCTACTA
 2251 CGGCCTGTCT CTTGCGGTAT TCGGAATCTT GCACGCCCTC GCTCAAGCCT
 GCCGGACAGC GAACGCCATA AGCCTTAGAA CGTGCGGGAG CGAGTTCCGA
 2301 TCGTCACTGG TCCCGCCACC AAACGTTTCG GCGAGAAGCA GGCATATC
 AGCAGTGACC AGGGCGGTGG TTTGCAAAGC CGCTCTTCGT CCGGTAATAG
 2351 GCCGGCATGG CGGCCGACGC GCTGGGCTAC GTCTTGCTGG CGTTCCGCAC
 CGGCCGTACC GCCGGCTGCG CGACCCGATG CAGAACGACC GCAAGCGCTG
 2401 GCGAGGCTGG ATGGCCTTCC CCATTATGAT TCTTCTCGCT TCCGGCGGCT
 CGCTCCGACC TACCGGAAGG GGTAACTA AGAAGAGCGA AGGCCCGCT
 2451 TCGGGATGCC CGCGTGCAG GCCATGCTGT CCAGGCAGST AGATGACGAC
 AGCCCTACGG GCGCAACGTC CGGTACGACA GGTCCGTCCA TCTACTGCTG
 2501 CATCAGGGAC AGCTTCAAGG ATCGCTCGCG GCTCTTACCA GCCTAACCTC
 GTAGTCCCTG TCGAAGTTCC TAGCGAGCGC CGAGAATGGT CGGATTGAAG
 2551 GATCATTGGA CCGCTGATCG TCACGGCGAT TTATGCCGCC TCGGCGAGCA
 CTAGTAACCT GGCAGACTAGC AGTGCCGCTA AATACGGCGG AGCCGCTCGT
 2601 CATGGAACGG GTTGSCATGG ATTGTAGGCG CCGCCCTATA CCTTGCTGCT
 GTACCTTGCC CAACCGTACC TAACATCCGC GCGGGGATAT GGAACAGACG
 2651 CTCCCCGCGT TCGCTCGCGG TGCATGGAGC CGGGCCACCT CGACCTGAAT
 GAGGGGCGCA ACGCAGCGCC ACGTACCTCG GCCCGGTGGA GCTGGACTTA
 2701 GGAAGCCGGG GGCACCTCGC TAACGGATTG ACCACTCCAA GAATTGGAGC
 CCTTCGGCCG CCGTGGAGCG ATTGCCTAAG TGGTGAGGTT CTTAACCTCG
 2751 CAATCAATTC TTGCGGAGAA CTGTGAATGC GCAAACCAAC CCTTGGCAGA
 GTTAGTTAAG AACGCCCTT GACACTTACG CGTTTGGTTG GGAACCGTCT
 2801 ACATATCCAT CGCGTCCGCC ATCTCCAGCA GCCGCACGCG GCGCATCTCG
 TGTATAGGTA GCGCAGGCGG TAGAGTCTGT CCGCGTGC GCCTAGAGC
 2851 GGCAGCGTTG GGTCTGGCC CGGGACTCAC CAGTCACAGA AAAGCATCTT
 CCGTCGCAAC CCAGGACCGG GCCCTGAGTG GTCAGTGTCT TTTCTAGAA
 2901 ACGGATGGCA TGACAGTAAG AGAATTATGC AGTGCTGCCA TAACCATGAG
 TGCTTACCGT ACTGTATTG TCTTAATACG TCACGACGGT ATTGGTACTC
 2951 TGATAACACT GCGGCCAACT TACTTCTGAC AACGATCGGA GGACCGAAGG
 ACTATTGTGA CGCCGGTTGA ATGAAGACTG TTGCTAGCCT CCTGGCTTCC
 3001 AGCTAACCGC TTTTTCGAC AACATGGGG ATCATGTAAC TCGCCTTGAT
 TCGATTGGCG AAAAAACGTG TTGTACCCCT TAGTACATTG AGCGGAACCTA

```

3051 CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG AGCGTGACAC
GCAACCCTTG GCCTCGACTT ACTTCGGTAT GGTTCGCTGC TCGCACTGTG
3101 CACGATGCCT GCAGCAATGG CAACAACGTT GCGCAAATA TTAACCTGGC
GTGCTACGGA CGTCGTTACC GTTGTTCGAA CGCGTTTGAT AATTGACCGC
3151 AACTACTTAC TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG
TTGATGAATG AGATCGAAGG GCCGTTGTTA ATTATCTGAC CTACCTCCGC
3201 GATAAAGTTG CAGGACCACT TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT
CTATTTCAAC GTCCGTGGTA AGACGCGAGC CGGGAAGGCC GACCGACCAA
3251 TATTGCTGAT AAATCTGGAG CCGGTGAGCG TGGGTCTCGC GGTATCATTG
ATAACGACTA TTTAGACCTC GGCCACTCGC ACCCAGACG CCATAGTAAC
3301 CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT TATCTACACG
GTCGTGACCC CGGTCTACCA TTCGGGAGGG CATAGCATCA ATAGATGTGC
3351 ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA TCGCTGAGAT
TGCCCTCAG TCCGTTGATA CCTACTTGCT TTATCTGTCT AGCGACTCTA
3401 AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA GTTTACTCAT
TCCACGAGT GACTAATTCG TAACCATTGA CAGTCTGTTT CAAATGAGTA
3451 ATATACTTTA GATTGATTTA AAACCTTCAT TTTAATTTAA AAGGATCTAG
TATAAGAAAT CTAACATAAT TTTGAAGTAA AAATTAATTT TCTTAGATC
3501 GTGAAGATCC TTTTGTGATA TCTCATGACC AAAATCCCTT AACGTGAGTT
CACTTCTAGG AAAAATAATT ACAGTACTGG TTTTAGGGAA TTGCACCTCA
3551 TTCGTTCCAC TGAGCGTCAG ACCCGTAGA AAAGATCAA GGATCTTCTT
AAGCAAGGTG ACTCGCAGTC TGGGGCATCT TTTCTAGTTT CCTAGAAGAA
3601 GAGATCCTTT TTTCTGCGC GTAATCTGCT GCTTGCAAAC AAAAAACCA
CTCTAGGAAA AAAAGACGCG CATTAGACGA CGAACGTTTG TTTTTTGGT
3651 CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTTTTTT
GGCGATGGTC GCCACCAAAC AAACGGCCTA GTTCTCGATG GTTGAGAAAA
3701 TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTCTTTC
AGGCTTCCAT TGACCGAAGT CGTCTCGCGT CTATGGTTTA TGACAGGAAG
3751 TAGTGATAGC GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT
ATCACATCGG CATCAATCCG GTGGTGAAGT TCTTGAGACA TCGTGGCGGA
3801 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA
TGTATGGAGC GAGACGATTA GGACAATGGT CACCGACGAC GGTACCGCT
3851 TAAGTCGTGT CTTACCGGTG TGGACTCAAG ACGATAGTTA CCGGATAAGG
ATTACAGACA GAATGGCCCA ACCTGAGTTC TGCTATCAAT GGCTATTCC
3901 CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC CAGCTTGAG
GCGTCGCCAG CCCGACTTGC CCCCCAAGCA CGTGTGTGCG GTCGAACCTC
3951 CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG
GCTTGCTGGA TGTGGCTTGA CTTATGGAT GTGCACTCG ATACTCTTTC
4001 CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA
GCGGTGCGAA GGGCTTCCCT CTTTCGCGCT GTCCATAGGC CATTGCGCGT
4051 GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG
CCCAGCCTTG TCCCTCGCG TGCTCCCTCG AAGGTCCCCC TTTGCGGACC
4101 TATCTTTATA GTCTGTGCG GTTTCGCCAC CTCTGACTTG AGCGTCGATT
ATAGAAATAT CAGGACAGCC CAAAGCGGTG GAGACTGAAC TCGCAGCTAA
4151 TTTGTGATGC TCGTCAGGGG GGCGGAGCCT ATGGAAAAAC GCCAGCAACG
AAACACTACG AGCAGTCCCC CCGCCTCGGA TACCTTTTTC CGGTGCTTGC
4201 CGGCCTTTTT ACGGTTCCCTG GCCTTTTGCT GGCCTTTTGC TCACATGTTT
GCCGGAAAAA TGCCAAGGAC CGGAAAAACG CCGGAAAAAC AGTGACAAG
4251 TTTCTGCGT TATCCCTGA TTCTGTGGAT AACCGTATTA CCGCCTTTGA
AAAGGACGCA ATAGGGGACT AAGACACCTA TTGGCATAAT GCGGAAACT
4301 GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG
CACTCGACTA TGGCGAGCGG CGTCGGCTTG CTGGCTCGCG TCGCTCAGTC
4351 TGAG
ACTC

```

Fig. 4