

(19)



URZĄD
PATENTOWY
RZECZYPOSPOLITEJ
POLSKIEJ

(10)

PL 244825 B1

(12)

Opis patentowy

(21) Numer zgłoszenia: **439656**

(22) Data zgłoszenia: **2021.11.26**

(43) Data publikacji o zgłoszeniu: **2023.05.29 BUP 22/2023**

(45) Data publikacji o udzieleniu patentu: **2024.03.11 WUP 11/2024**

(51) MKP:

C12N 9/58 (2006.01)

C12N 15/57 (2006.01)

C12N 1/19 (2006.01)

C12N 15/80 (2006.01)

C12P 21/00 (2006.01)

(73) Uprawniony z patentu:

BLIRT SPÓŁKA AKCYJNA, Gdańsk, PL

(72) Twórca(-y) wynalazku:

ROBERT BRODZIK, Gdańsk, PL

ARKADIUSZ POPINIGIS, Gdańsk, PL

JUSTYNA LEIBNER-CISZAK, Borkowo, PL

DOMINIK ZIĘTKOWSKI, Gdynia, PL

KRZYSZTOF KUR, Reda, PL

(74) Pełnomocnik:

rzecz. pat. Rafał Witek, Wrocław, PL

(54) Tytuł:

Mutant proteiny K *Tritirachium album* i jej zymogenu, plazmid ekspresyjny, rekombinantowy szczep *Pichia pastoris* i sposób wytwarzania dojrzałej formy mutantu proteiny K

PL 244825 B1

Opis wynalazku

Przedmiotem wynalazku są mutant proteiny *K* *Tritirachium album* i jej zymogenu, plazmid ekspresyjny, rekombinantowy szczep *Pichia pastoris* i sposób wytwarzania dojrzałej formy mutantu proteiny *K*.

Proteinaza *K* (E.C. 3.4.21.64), proteaza serynowa, jest enzymem proteolitycznym syntetyzowanym przez komórki pleśni *Tritirachium album* (*Engyodontium album*) ATCC 22625. Jej analogi znaleźć można także w komórkach bakterii np. *Serratia* sp., *Pseudoalteromonas* sp., *Alteromonas* sp., *Thermus* sp., *Vibrio* sp. i in.

W zależności od źródła, z którego pochodzą, enzymy te różnią się masą cząsteczkową, długością domeny C-terminalnej, wartościami optymalnej temperatury i pH działania, a także zapotrzebowaniem na jony metali niezbędne dla aktywności enzymu lub/i stabilizowania jego struktury. Aktywność enzymów z rodziny proteinaz *K* zazwyczaj wzrasta, gdy w buforze reakcyjnym znajdują się czynniki denaturujące jak SDS czy mocznik. Jony wapnia zabezpieczają natomiast proteinazę *K* przed autoproteolizą.

Enzymy te posiadają zdolność katalizowania reakcji hydrolizy wiązań peptydowych, wykazując przy tym szerokie spektrum substratowe, dlatego wykorzystuje się je głównie w procesach izolacji DNA i innych zastosowaniach biologii molekularnej, jak również w przypadku usuwania zanieczyszczeń białkowych z materiałów niebiałkowych.

Proteinaza *K* jest znanym od dawna enzymem, izolowanym z płynu pohodowlanego pleśni *Tritirachium album*. Proces produkcji jest jednak długotrwały i charakteryzuje się niską wydajnością produkcji na 1 litr płynu pohodowlanego. Podjęto również próby konstrukcji bakteryjnych układów ekspresyjnych umożliwiających produkcję proteiny *K* w postaci ciał inkluzyjnych, lecz problem renaturacji białka uniemożliwił otrzymanie jego znacznych ilości (Gunkel, F. A. and Gassen, H. G. (1989) Eur. J. Biochem. Vol. 179(1), 185–194; Samal, B. B. et al. (1996) Adv. Exp. Med. Biol. Vol. 379, 95–104).

Z patentu EP1360283 znana jest także metoda produkcji proteiny *K* w układach drożdżowych, w postaci fuzji translacyjnej zymogenu z peptydem sygnałnym alfa faktora *Saccharomyces cerevisiae*. W tym przypadku autokatalityczna aktywacja zymogenu do formy aktywnej białka następuje w czasie sekrecji lub w płynie pohodowlanym, lecz może również następować przedwcześnie przez co nie jest procesem ściśle kontrolowanym.

Z patentu PL 213045 znany jest mutant zymogenu proteiny *K* z *Tritirachium album* zawierający sekwencję aminokwasową rozpoznawaną przez proteazę Kex2 umieszczoną pomiędzy propeptydem, a sekwencją białka o aktywności dojrzałej proteiny *K*. Taka modyfikacja zymogenu, (tj. nieaktywnej enzymatycznie formy białka) prowadząca do kontrolowanego odcinania propeptydu (tj. części białka pełniącego rolę wewnętrznego czaperonu) w komórkach drożdży i uzyskania dojrzałej formy proteiny *K* lub jej mutantu (tj. białka o aktywności dojrzałej proteiny *K*) umożliwia wydajną produkcję enzymu w gospodarzu drożdżowym, w szczególności komórkach rekombinantowego szczepu *Pichia pastoris*.

Pomimo znanych rozwiązań, nadal poszukuje się nowych, wydajnych układów pozwalających na efektywną i tanią biosyntezę enzymów z rodziny proteineaz *K*. Od bardzo wielu lat trwają prace nad konstrukcją układu pozwalającego na wydajną biosyntezę i łatwą izolację enzymu posiadającego aktywność proteiny *K*. Celem wynalazku jest dostarczenie takiego układu pozwalającego na wydajną produkcję białka o aktywności proteiny *K* oraz nowych mutantów proteiny *K*, które mogą być wykorzystane dla uzyskania takiego pożądanego układu.

Kolejnym celem wynalazku jest uzyskanie nowego białka o aktywności proteolitycznej zbliżonej do aktywności proteiny *K*, które będzie jednak mogło być łatwo inaktywowane w celu uniknięcia niepożądanej aktywności proteolitycznej mieszaniny reakcyjnej uzyskiwanej po zakończeniu reakcji proteolizy. W przypadku wykorzystywania proteiny *K* jako odczynnika mającego za zadanie usunięcie wszelkich białek, np. w trakcie oczyszczania kwasów nukleinowych, pozostawanie w końcowym produkcie zanieczyszczenia enzymem o aktywność proteiny jest bardzo niepożądane. Dlatego uzyskanie białka poddającego się łatwiejszej inaktywacji niż dzika proteinaza *K* jest bardzo pożądanym.

Nieoczekiwanie określony powyżej złożony efekt techniczny został osiągnięty w niniejszym wynalazku.

Nieoczekiwanie określony powyżej cel został osiągnięty w niniejszym wynalazku.

Przedmiotem wynalazku jest mutant zymogenu proteiny *K* z *Tritirachium album* zawierający na C-końcu mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzałej proteiny *K* w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu.

Korzystnie, mutant zymogenu proteinazy K z *Tritirachium album* według wynalazku posiada sekwencję przedstawioną jako sekwencja nr 9.

Nieoczekiwanie, zaproponowana w wynalazku modyfikacja zymogenu zwiększa wydajność produkcji enzymu w komórkach drożdżowych, zwłaszcza *Pichia pastoris* zgodnie z opisanym przykładem 9.

Kolejnym przedmiotem wynalazku jest mutant dojrzalej proteinazy K z *Tritirachium album* zawierający na C-końcu mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzalej proteinazy K w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu. Korzystnie, posiada on sekwencję przedstawioną jako sekwencję nr 22.

Kolejnym przedmiotem wynalazku jest sekwencja DNA zawierająca sekwencję kodującą mutantą zymogenu proteinazy K według wynalazku jak zdefiniowano powyżej albo jego fragment. Korzystnie, posiada ona sekwencję nr 8.

Kolejnym przedmiotem wynalazku jest sekwencja oligonukleotydów o symbolach ProtK-G-F i ProtK-G-R służących do otrzymywania sekwencji DNA kodującej mutantą zymogenu proteinazy K według wynalazku zdefiniowanego powyżej albo jego fragment wybrana spośród sekwencji: sekwencja nr 6 i 7.

Kolejnym przedmiotem wynalazku jest plazmid ekspresyjny zawierający sekwencję DNA kodującą mutantą zymogenu proteinazy K według wynalazku zdefiniowanego powyżej, zgodnie z opisem w przykładzie 4. Korzystnie, plazmid według wynalazku posiada sekwencję wybraną spośród sekwencji: sekwencja nr 16.

Kolejnym przedmiotem wynalazku jest rekombinantowy szczep drożdżowy transformowany dowolnym plazmidem według wynalazku zdefiniowanym powyżej, zgodnie z opisem w przykładzie 6. Korzystnie szczep drożdży według wynalazku może być otrzymany z jednego ze znanych szczepów drożdży wybranych spośród *Pichia sp.*, *Hansenula sp.*, *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Yarrowia sp.*, lub *Kluyveromyces sp.*, korzystnie wybranego spośród szczepów *Pichia pastoris*.

Kolejnym przedmiotem wynalazku jest sposób otrzymywania białka o aktywności proteinazy K *Tritirachium album* charakteryzujący się tym, że prowadzi się hodowlę komórek drożdży transformowanych plazmidem ekspresyjnym kodującym mutantą zymogenu proteinazy K z *Tritirachium album* zawierającego na C-końcu mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzalej proteinazy K w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu, a następnie uzyskiwane białko izoluje się i oczyszcza z płynu hodowlanego.

Korzystnie prowadzi się hodowlę komórek drożdży według wynalazku zdefiniowanych powyżej, korzystnie w temperaturze około 28 do 30°C, korzystnie na pożywce zawierającej jako źródło węgla glicerol lub metanol, przy czym indukuje się ekspresję zmodyfikowanego genu proteinazy K metanolem. Równie korzystnie, białko o aktywności proteinazy K *Tritirachium album* izoluje się z płynu hodowlanego metodą chromatografii oddziaływań hydrofobowych lub metodą chromatografii na złożu hydroksypatytu, a następnie oczyszcza się prowadząc dializę wobec roztworu węglanu amonu, po czym suszy się ze stanu zamrożenia zgodnie z opisem w przykładzie 9.

Skonstruowano układ ekspresyjny pozwalający na wydajną biosyntezę dojrzalej formy zmodyfikowanej proteinazy K *Tritirachium album* ATCC 22625 w komórkach *Pichia pastoris*. Gen kodujący zymogen zmodyfikowanego białka, przedstawiony w sekwencji nr 8, został namnożony na drodze PCR na matrycy sekwencji genu kodującego dziką wersję białka przedstawionym w sekwencji nr 5, zgodnie z opisem w przykładzie 2 i zsekwencjonowany. Z namnożonej sekwencji usunięto sekwencję nukleotydową kodującą peptyd sygnałowy, przedstawiony w sekwencji 2, a pozostawiono sekwencję DNA kodującą propeptyd dzikiego białka, zgodnie z sekwencją nr 3, stabilizujący zymogen proteinazy K przed przedwczesną aktywacją do formy dojrzalej, aktywnie enzymatycznie białka (Wolfgang Ebeling, *et al*, Proteinase K from *Tritirachium album* Limber, Eur. J Biochem. 47,91–97,1974). Dodatkowo dodano resztę aminokwasową glicyny na C-końcu dojrzalego białka, co nieoczekiwanie, zwiększyło efektywność autolizy dojrzalej formy zmutowanego białka w warunkach stresowych, w podwyższonym stężeniu surfaktanta. Taka modyfikacja dzikiego białka pozwoliła na wydajną inaktywację enzymu w odpowiednich warunkach buforowych, zgodnie z opisem w przykładzie 10, co jest szczególnie korzystną właściwością, umożliwiającą na łatwą inaktywację enzymu, w procesach oczyszczania kwasów nukleinowych, gdzie wymagana jest wysoka jakość i czystość preparatów DNA i RNA pozbawionych zanieczyszczeń białkowych, bez stosowania dodatkowych czynników chemicznych czy fizycznych, które mogą uszkadzać izolowane preparaty kwasów nukleinowych.

Uzyskany w wynalazku szczep ekspresyjny *Pichia pastoris*/pD912-ProtK-G, zgodnie z opisem w przykładzie 6, umożliwia otrzymywanie funkcjonalnej zmodyfikowanej proteiny K *Tritirachium album* w znacznych ilościach, a ponadto pozwala na produkcję zmodyfikowanego enzymu, który nie odbiega aktywnością od enzymu izolowanego ze szczepu dzikiego. Produkowane dojrzałe zmutowane białko Proteiny K, przedstawione w sekwencji 22, nie różni się w znaczącym stopniu aktywnością od białka izolowanego z *Tritirachium album* mimo obecności dodatkowej reszty aminokwasowej glicyny na C-końcu dojrzałego białka.

Realizację wynalazku stanowi także sekwencja DNA plazmidu ekspresyjnego o symbolu pD912-ProtK i sekwencji 21 otrzymana w wyniku amplifikacji DNA, zgodnie z opisem w przykładzie 7, z zastosowaniem starterów o sekwencji 18 i symbolu ProtK-G-out-F oraz o sekwencji 19 i symbolu ProtK-G-out-R, która zawiera fragment DNA o sekwencji 20 kodującej sekwencję aminokwasową identyczną ze zmutowaną wersją proteiny K ale pozbawionej dodatkowej reszty aminokwasowej glicyny na C-końcu dojrzałego białka, zgodnie z sekwencją 17.

Realizację wynalazku stanowi rekombinantowy szczep *Pichia pastoris* o symbolu *Pichia pastoris*/pD912-ProtK, zgodnie z opisem w przykładzie 8, który stanowią komórki drożdży *Pichia pastoris* BG10 zawierający gen kodujący zmodyfikowaną proteinę K *Tritirachium album*, przedstawiony w sekwencji 20.

Realizację wynalazku stanowi także sekwencja DNA plazmidu ekspresyjnego o symbolu pPink_HC-ProtK-WT i sekwencji 11 otrzymana w wyniku syntezy de novo DNA i klonowania do wektora o symbolu pPink_HC i sekwencji 10, zgodnie z opisem w przykładzie 1. Uzyskany plazmid ekspresyjny pPink_HC-ProtK-WT zawiera fragment DNA o sekwencji 5 i symbolu ProtK-WT kodujący sekwencję aminokwasową dzikiej wersji proteiny K *Tritirachium album*. Zawierającej natywny peptyd sygnałowy proteiny K (sekwencja 2) wraz z natywnym propeptydem (sekwencja 3).

Realizację wynalazku stanowi rekombinantowy szczep *Pichia pastoris* o symbolu *Pichia Pink*/pPink_HC-ProtK-WT, zgodnie z opisem w przykładzie 5, który stanowią komórki drożdży *Pichia Pink* (ThermoFisher, nr kat. A11150) zawierający gen kodujący dziką wersję proteiny K *Tritirachium album*, przedstawiony w sekwencji 1.

Realizację wynalazku stanowi również sposób wytwarzania zmodyfikowanej proteiny K *Tritirachium album*, zgodnie z opisem w przykładzie 9, polegający na zewnątrzkomórkowej produkcji enzymu przez komórki rekombinantowego szczepu drożdży w pożywkach płynnych wzbogaconych o jony wapnia (Ca^{2+}) charakteryzujący się tym, że prowadzi się hodowlę komórek szczepu drożdży o symbolu *Pichia Pink*/pPink_HC-ProtK-WT lub *Pichia pastoris*/pD912-ProtK lub *Pichia pastoris*/pD912-ProtK-G w temperaturze 28–30°C, indukuje ekspresję genu kodującego dziką lub zmodyfikowaną proteinę K metanolem, a następnie izoluje się i oczyszcza białko z płynu pohodowlanego na drodze chromatografii oddziaływań hydrofobowych, po czym oczyszczony enzym suszy się ze stanu zamrożenia. Można tutaj również zastosować metodę oczyszczania chromatograficznego na złożu w postaci hydroksypatytu.

Dzięki wykorzystaniu wynalazku uzyskuje się układ ekspresyjny umożliwiający wydajną produkcję zmodyfikowanej proteiny K *Tritirachium album* (ProtK-G) w rekombinantowych komórkach drożdży *Pichia pastoris*/pD912-ProtK-G, która jest o około 10% wyższa w porównaniu z poziomem ekspresji natywnej proteiny (ProtK) prowadzonej w tych samych warunkach, w tym samym układzie ekspresyjnym w rekombinantowym szczepie *Pichia pastoris*/pD912-ProtK i aż o około 30% w porównaniu do ekspresji w dzikiej wersji proteiny K w szczepie *Pichia Pink*/pPink_HC-ProtK-WT. Szczep *Pichia Pink*™ to komercyjny system ekspresyjny oferowany przez firmę Thermo Fisher, USA do wysokowydajnej produkcji białek rekombinantowych w układzie drożdżowym *Pichia Pink*.

W wyniku realizacji wynalazku uzyskano szczep ekspresyjny *Pichia pastoris*/pD912-ProtK-G, który okazał się najbardziej wydajny w produkcji proteiny K *Tritirachium album* w układzie rekombinantowym spośród testowanych systemów w trakcie realizacji wynalazku. Przy zachowaniu 10% różnicy wydajności i przeskalowaniu procesu do produkcyjnej skali 1000L, możliwe jest w ten sposób wytworzenie ponad 260 g białka więcej na proces.

Opis rycin i sekwencji:

Ryc. 1 – przedstawia schemat plazmidu ekspresyjnego o sekwencji 11 i symbolu pPink_HC-ProtK-WT kodującego dziką wersję proteiny K.

Ryc. 2 – przedstawia schemat plazmidu ekspresyjnego o sekwencji 16 (lub 21) i symbolu pD912-ProtK-G (lub pD912-ProtK) kodującego rekombinantową wersję proteiny K ze zmutowanym C-końcem (lub wersję dziką dojrzałego białka).

Ryc. 3 – Porównanie wydajności ekspresji proteiny K *Tritirachium album* w trzech szczepach ekspresyjnych: Pichia Pink/pPink_HC-ProtK-WT, Pichia pastoris/pD912-ProtK-G oraz Pichia pastoris/pD912-ProtK.

Ryc. 4 – Porównanie stabilności białka rekombinantowej wersji ProtK-G oraz natywnej ProtK_WT w warunkach typowych podczas oczyszczania kwasów nukleinowych oraz warunkach stresowych, z podwyższoną ilością surfaktanta do 1%. Na żelach SDS-PAGE pokazano stabilność rekombinantowej wersji ProtK-G oraz natywnej ProtK_WT po 30 min. inkubacji w warunkach typowych dla izolacji DNA i/lub RNA (panel lewy) lub w warunkach stresowych z podwyższonym stężeniem surfaktanta (panel środkowy). Oba białka, ProtK-G i ProtK-WT nie ulegają autolizie w warunkach typowych dla izolacji kwasów nukleinowych i pozostają na podobnym poziomie, natomiast w warunkach stresowych, białko rekombinantowej ProtK-G ulega szybszej i wydajniejszej degradacji w porównaniu do wersji dzikiej, ProtK-WT.

Tabela 1. Porównanie wydajności efektywność fermentacji proteiny K *Tritirachium album* w trzech szczepach ekspresyjnych: Pichia Pink/pPink_HC-ProtK-WT, Pichia pastoris/pD912-ProtK-G oraz Pichia pastoris/pD912-ProtK. Fermentacja drożdży prowadzona była w przy użyciu tego samego protokołu w tych samych warunkach dla wszystkich układów ekspresyjnych. Po zakończeniu hodowli mierzona była ilość biomasy komórek w brzeczce pofermentacyjnej po oddzieleniu pożywki pofermentacyjnej przez wirowanie, przedstawiona w gramach osadu w 1 litrze hodowli. Zawartość proteiny K w pożywce pofermentacyjnej po oddzieleniu komórek mierzona była densytometrycznie oraz metodą HPLC. Porównano względną wydajność trzech układów ekspresyjnych przyjmując jako 100% układ najbardziej wydajny, Pichia pastoris/pD912-ProtK-G.

Sekwencja 1 – przedstawia sekwencję aminokwasową białka dzikiego proteiny K *Tritirachium album* o symbolu **ProtK-WT**

Sekwencja 2 – przedstawia sekwencję aminokwasową peptydu sygnałowego z dzikiego białka proteiny K *Tritirachium album* o symbolu ProtK-WT SP

Sekwencja 3 – przedstawia sekwencję aminokwasową propeptydu dzikiego białka proteiny K *Tritirachium album* o symbolu **ProtK-WT propeptyd**

Sekwencja 4 – przedstawia sekwencję aminokwasową fragmentu aktywnego enzymatycznie białka dojrzałej proteiny K *Tritirachium album* o symbolu **ProtK-WT mature**

Sekwencja 5 – przedstawia sekwencję nukleotydową genu o długości 1173 nukleotydów kodującego wersję białka dzikiej proteiny K *Tritirachium album* o symbolu **ProtK-WT**

Sekwencja 6 – przedstawia sekwencję oligonukleotydu o długości 41 nukleotydów o symbolu **ProtK-G-F** użytego w reakcji PCR do amplifikacji sekwencji 8

Sekwencja 7 – przedstawia sekwencję oligonukleotydu o długości 41 nukleotydów o symbolu **ProtK-G-R** użytego w reakcji PCR do amplifikacji sekwencji 8

Sekwencja 8 – przedstawia sekwencję nukleotydową genu o długości 1139 nukleotydów kodującego fragment propeptydu i peptyd białka dojrzałego wersji rekombinantowej proteiny K *Tritirachium album* o symbolu **pro-ProtK-G**

Sekwencja 9 – przedstawia sekwencję aminokwasową o długości 370 aminokwasów propeptydu i peptydu fragmentu aktywnego enzymatycznie białka dojrzałej rekombinantowej proteiny K *Tritirachium album* o symbolu **pro-ProtK-G**.

Sekwencja 10 – przedstawia sekwencję nukleotydową wektora ekspresyjnego pPink_HC o długości 7667 nukleotydów zakupionego z firmy Thermo Fisher, USA, nr kat. A11152 zawierającego promotora AOX1 *Saccharomyces cerevisiae* indukowalnego przez metanol

Sekwencja 11 – przedstawia sekwencję nukleotydową wektora ekspresyjnego pPink_HC-ProtK-WT o długości 8805 nukleotydów o symbolu **pPink_HC-ProtK-WT** zawierającego sekwencję 5 pod kontrolą promotora AOX1 *Saccharomyces cerevisiae* indukowalnego przez metanol

Sekwencja 12 – przedstawia sekwencję aminokwasową białka rekombinantowej wersji proteiny K *Tritirachium album* o symbolu **ProtK-G** i długości 460 aminokwasów ze zmodyfikowanym C-końcem białka

Sekwencja 13 – przedstawia 90 aminokwasową sekwencję peptydu sygnałowego z białka alpha factor *Saccharomyces cerevisiae* o symbolu **ProtK-G SP**

Sekwencja 14 – przedstawia sekwencję nukleotydową genu o długości 1324 nukleotydów kodującego wersję rekombinantową białka proteiny K *Tritirachium album* o symbolu **ProtK-G**

Sekwencja 15 – przedstawia sekwencję nukleotydową wektora ekspresyjnego pD912 o długości 3837 nukleotydów zakupionego z firmy ATUM, USA, nr kat. 867858d1 zawierającego promotora AOX1 *Saccharomyces cerevisiae* indukowalnego przez metanol

Sekwencja 16 – przedstawia sekwencję nukleotydową wektora ekspresyjnego pD912-ProtK-G o długości 4928 nukleotydów o symbolu **pD912-ProtK-G** zawierającego sekwencję 14 pod kontrolą promotora AOX1 *Saccharomyces cerevisiae* indukowalnego przez metanol

Sekwencja 17 – przedstawia sekwencję aminokwasową białka rekombinantowej wersji proteiny K *Tritirachium album* o symbolu **ProtK** i długości 459 aminokwasów z natywnym C-końcem białka (bez dodanego aminokwasu glicyny w pozycji 460)

Sekwencja 18 – przedstawia sekwencję oligonukleotydu o długości 59 nukleotydów o symbolu **ProtK-G-out-F** użytego w reakcji PCR do amplifikacji sekwencji 20

Sekwencja 19 – przedstawia sekwencję oligonukleotydu o długości 36 nukleotydów o symbolu **ProtK-G-out-** użytego w reakcji PCR do amplifikacji sekwencji 20

Sekwencja 20 – przedstawia sekwencję nukleotydową genu o długości 1324 nukleotydów kodującego wersję rekombinantową białka proteiny K *Tritirachium album* o symbolu **ProtK** (G460stop)

Sekwencja 21 – przedstawia sekwencję nukleotydową wektora ekspresyjnego pD912-ProtK o długości 4928 nukleotydów o symbolu **pD912-ProtK** zawierającego sekwencję 20 pod kontrolą promotora AOX1 *Saccharomyces cerevisiae* indukowalnego przez metanol

Sekwencja 22 – przedstawia sekwencję aminokwasową fragmentu aktywnego enzymatycznie białka dojrzałej rekombinowanej proteiny K *Tritirachium album* o symbolu **ProtK-G mature**

Przykład 1. Otrzymywanie sekwencji DNA genu kodującego dziką proteinazę K *Tritirachium album* o sekwencji 1 i symbolu ProtK-WT.

W celu otrzymania DNA genu kodującego białko dzięki proteiny K o Sekwencji 1 i symbolu ProtK-WT zawierającego peptyd sygnałowy o Sekwencji 2 i symbolu ProtK-WT SP, propeptyd o Sekwencji 3 i symbolu ProtK-WT propeptyd oraz dojrzały aktywny enzymatycznie polipeptyd proteiny K o Sekwencji 4 i symbolu ProtK-WT mature, zaprojektowano sekwencję DNA w oparciu o sekwencję białkową proteiny K dostępną w bazie danych UniProtKB numer P06873 (<https://www.uniprot.org/uniprot/P06873>). Sekwencję DNA kodującą dziką proteinazę K uzyskano poprzez odwrotną translację sekwencji polipeptydowej nr 1 o symbolu ProtK-WT. Zaprojektowana sekwencja DNA została zoptymalizowana pod kątem wydajnej ekspresji w drożdżach *Pichia pastoris*. Dodatkowo, by umożliwić klonowanie do wektora ekspresyjnego, zaprojektowano miejsca rozpoznawane przez enzymy restrykcyjne EcoRI i KpnI, odpowiednio na 5- i 3- końcach. Synteza genu *de novo* została zlecona do firmy GeneART (<https://www.thermofisher.com/pl/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis.html>). Sekwencja DNA uzyskanego genu o symbolu ProtK-WT przedstawiona została jako Sekwencja 5.

Przykład 2. Otrzymywanie sekwencji DNA genu kodującego fragment rekombinantowej proteiny K *Tritirachium album* z dołączonym na C-końcu białka aminokwasu glicyny o sekwencji 8 i symbolu pro-ProtK-G.

W celu otrzymania fragmentu DNA genu o sekwencji 8 i symbolu pro-ProtK-G kodującego fragment rekombinantowej proteiny K *Tritirachium album* z dołączonym na C-końcu białka aminokwasu glicyny o Sekwencji 9 i symbolu pro-ProtK-G, otrzymany w Przykładzie 1 fragment DNA o Sekwencji 5 i symbolu ProtK-WT amplifikuje się w reakcji PCR stosując dwa startery oligonukleotydowe o Sekwencji 6 i symbolu ProtK-G-F oraz o Sekwencji 7 i symbolu ProtK-G-R przy użyciu kitu do amplifikacji zakupionego od firmy New England Biolabs, Inc., nr kat. 50-995-156 w warunkach rekomendowanych przez producenta. Dodatkowo, by umożliwić klonowanie fragmentu DNA o Sekwencji 8 i symbolu pro-ProtK-G do wektora ekspresyjnego, użyte w reakcji PCR, oligonukleotydy zawierają miejsce rozpoznawane przez enzym restrykcyjny SapI. Otrzymany, fragment DNA o Sekwencji 8 i symbolu pro-ProtK-G, wytrąca się etanolem i wykorzystuje się do klonowania do wektora ekspresyjnego.

Przykład 3. Otrzymywanie plazmidu ekspresyjnego o sekwencji 11 i symbolu pPink_HC-ProtK-WT zawierającego gen kodujący sekwencję polipeptydową dzięki proteiny K *Tritirachium album* o sekwencji 1 i symbolu ProtK-WT.

W celu otrzymania plazmidu ekspresyjnego o Sekwencji 11 i symbolu pPink_HC-ProtK-WT, oczyszczony na drodze precypitacji etanolowej fragment DNA o Sekwencji 1 i symbolu ProtK-WT, uzyskany w Przykładzie 1 trawi się enzymami restrykcyjnymi EcoRI i KpnI, a następnie liguje się z DNA wektora plazmidowego o Sekwencji 10 i symbolu pPink_HC zakupionym w Thermo Fisher, USA, nr kat. A11152 trawionym tymi samymi enzymami restrykcyjnymi.

Mieszaniną ligacyjną transformuje się do komórek kompetentnych *Escherichia coli* TOP10F i wysiewa na szalki Petriego z podłożem LA (1% pepton K; 0,5% ekstrakt drożdżowy; 1% NaCl; 1,5% agar) zawierającym 100 µg/ml ampicyliny. W wyniku izolacji plazmidowego DNA z otrzymanych kolonii bakteryjnych uzyskuje się plazmid ekspresyjny o Sekwencji 11 i symbolu pPink_HC-ProtK-WT zawierający gen o sekwencji 5 i symbolu ProtK-WT kodujący dziką wersję proteiny K *Tritirachium album* o sekwencji 1 i symbolu ProtK-WT. Mapa plazmidu o sekwencji 11 i symbolu pPink_HC-ProtK-WT przedstawiona jest na Rycinie 1.

Przykład 4. Otrzymywanie plazmidu ekspresyjnego o sekwencji 16 i symbolu pD912-ProtK-G zawierającego gen o sekwencji 14 i symbolu ProtK-G kodującego sekwencję polipeptydową rekombinantowej proteiny K *Tritirachium album* o sekwencji 12 i symbolu ProtK-G.

Otrzymany w przykładzie 2 fragment DNA o sekwencji 8 i symbolu pro-ProtK-G trawi się enzymem restrykcyjnym SapI, a następnie liguje się z DNA wektora plazmidowego o Sekwencji 15 i symbolu pD912 zakupionym w firmie ATUM, Newark, USA, nr kat. 867858d1, trawionym tym samym enzymem restrykcyjnym. W wyniku ligacji fragmentu DNA o sekwencji 8 i symbolu pro-ProtK-G z DNA wektora pD912 o Sekwencji 15 powstaje połączenie we właściwej ramce odczytu (fuzja translacyjna) sekwencji 8 o symbolu pro-ProtK-G z peptydem sygnałowym znajdującym się na wektorze pD912 o sekwencji 13 i symbolu ProtK-G SP.

Mieszaniną ligacyjną transformuje się do komórek kompetentnych *Escherichia coli* TOP10F i wysiewa na szalki Petriego z podłożem LA (1% pepton K; 0,5% ekstrakt drożdżowy; 1% NaCl; 1,5% agar) zawierającym 25 µg/ml zeozyny. W wyniku izolacji plazmidowego DNA z otrzymanych kolonii bakteryjnych uzyskuje się plazmid ekspresyjny o Sekwencji 16 i symbolu pD912-ProtK-G zawierający gen o sekwencji 14 i symbolu ProtK-G kodujący rekombinantową proteinę K *Tritirachium album* o sekwencji 12 i symbolu ProtK-G. Mapa plazmidu o sekwencji 16 i symbolu pD912-ProtK-G przedstawiona jest na Rycinie 2.

Przykład 5. Otrzymywanie rekombinantowego szczepu *Pichia pastoris* o symbolu Pichia Pink/pPink_HC-ProtK-WT

W celu uzyskania rekombinantowego szczepu Pichia Pink/pPink_HC-ProtK-WT wykonuje się transformację komórek drożdży Pichia Pink liniową formą DNA plazmidu ekspresyjnego o sekwencji 11 i symbolu pPink_HC-ProtK-WT uzyskaną przez uprzednie trawienie DNA plazmidu o sekwencji 11 i symbolu pPink_HC-ProtK-WT enzymem restrykcyjnym SpeI (Thermo Fisher, nr kat. ER1251). Selekcję pozytywnych klonów komórek drożdży Pichia Pink wykonuje się na podłożu selekcyjnym PAD, zgodnie z zaleceniami dostawcy komórek Pichia Pink (ThermoFisher, nr kat. A11150, wersja instrukcji MAN0000717). Obecność wbudowanego rekombinantowego genu ProtK-WT do genomu drożdży o symbolu Pichia Pink/pPink_HC-ProtK-WT potwierdza się przez amplifikację PCR z użyciem oligonukleotydów jak w przykładzie 2 na matrycy genomowego DNA izolowanego z komórek drożdży.

Przykład 6. Otrzymywanie rekombinantowego szczepu *Pichia pastoris* o symbolu Pichia pastoris/pD912-ProtK-G

W celu uzyskania rekombinantowego szczepu Pichia pastoris/pD912-ProtK-G wykonuje się transformację komórek drożdży Pichia pastoris szczep BG10 (ATUM, USA nr kat. PPS-9010) liniową formą DNA plazmidu ekspresyjnego o sekwencji 16 i symbolu pD912-ProtK-G uzyskaną przez uprzednie trawienie DNA plazmidu o sekwencji 16 i symbolu pD912-ProtK-G enzymem restrykcyjnym PmeI (Thermo Fisher, nr kat. ER1342). Selekcję pozytywnych klonów komórek drożdży Pichia pastoris/pD912-ProtK-G wykonuje się na podłożu selekcyjnym YPDS z dodatkiem 200 pg/ml zeocyny a następnie przesiewa się na podłożu YPD z dodatkiem 200 pg/ml zeocyny, zgodnie z zaleceniami dostawcy komórek Pichia pastoris szczep BG10 (ATUM, USA). Obecność wbudowanego rekombinantowego genu ProtK-G do genomu drożdży o symbolu Pichia pastoris/pD912-ProtK-G potwierdza się przez amplifikację PCR z użyciem oligonukleotydów jak w przykładzie 2 na matrycy genomowego DNA izolowanego z komórek drożdży.

Przykład 7. Otrzymywanie plazmidu ekspresyjnego o sekwencji 21 i symbolu pD912-ProtK zawierającego gen o sekwencji 20 i symbolu ProtK kodującego sekwencję polipeptydową rekombinantowej proteiny K *Tritirachium album* o sekwencji 17 i symbolu ProtK.

Otrzymany zgodnie z przykładem 4 plazmid ekspresyjny o sekwencji 16 i symbolu pD912-ProtK-G zawierający gen o sekwencji 14 i symbolu ProtK-G kodujący sekwencję polipeptydową rekombinantowej proteiny K *Tritirachium album* o sekwencji 12 i symbolu ProtK-G używa się jako matrycy w procesie PCR do uzyskania wektora ekspresyjnego o sekwencji 21 i symbolu pD912-ProtK zawierającego gen o sekwencji 20 i symbolu ProtK kodujący sekwencję polipeptydową rekombinantowej proteiny K *Tritirachium album* o sekwencji 17 i symbolu ProtK. Jako startery w procesie PCR stosuje dwa oligonukleotydy, jeden o Sekwencji 18 i symbolu ProtK-G-out-F oraz drugi, o Sekwencji 19 i symbolu ProtK-G-out-R przy użyciu kitu do amplifikacji zakupionego od firmy New England Biolabs, Inc., nr kat. 50-995-156, w warunkach rekomendowanych przez producenta.

Mieszaninę PCR inkubuje się z enzymem restrykcyjnym DpnI (Thermo Fisher, nr kat. ER1701) i transformuje się do komórek kompetentnych *Escherichia coli* DH5(alpha) a następnie wysiewa na szalki Petriego z podłożem LA (1% pepton K; 0,5% ekstrakt drożdżowy; 1% NaCl; 1,5% agar) zawierającym 25 µg/ml zeozyny. W wyniku izolacji plazmidowego DNA z otrzymanych kolonii bakteryjnych uzyskuje się plazmid ekspresyjny o Sekwencji 21 i symbolu pD912-ProtK zawierający gen o sekwencji 20 i symbolu ProtK kodujący rekombinantową proteinę K *Tritirachium album* o sekwencji 17 i symbolu ProtK. Mapa plazmidu o sekwencji 21 i symbolu pD912-ProtK przedstawiona jest na Rycinie 2.

Przykład 8. Otrzymywanie rekombinantowego szczepu *Pichia pastoris* o symbolu *Pichia pastoris*/pD912-ProtK

W celu uzyskania rekombinantowego szczepu *Pichia pastoris*/pD912-ProtK- wykonuje się transformację komórek drożdży *Pichia pastoris* szczep BG10 (ATUM, USA nr kat. PPS-9010) liniową formą DNA plazmidu ekspresyjnego o sekwencji 21 i symbolu pD912-ProtK uzyskaną przez uprzednie trawienie DNA plazmidu o sekwencji 21 i symbolu pD912-ProtK enzymem restrykcyjnym PmeI (Thermo Fisher, nr kat. ER1342). Selekcję pozytywnych klonów komórek drożdży *Pichia pastoris*/pD912-ProtK wykonuje się na podłożu selekcyjnym YPDS z dodatkiem 200 pg/ml zeocyny a następnie przesiewa się na podłoże YPD z dodatkiem 200 pg/ml zeocyny, zgodnie z zaleceniami dostawcy komórek *Pichia pastoris* szczep BG10 (ATUM, USA). Obecność wbudowanego rekombinantowego genu ProtK-G do genomu drożdży o symbolu *Pichia pastoris*/pD912-ProtK potwierdza się przez amplifikację PCR z użyciem oligonukleotydów jak w przykładzie 2 na matrycy genomowego DNA izolowanego z komórek drożdży.

Przykład 9. Otrzymywanie proteiny K *Tritirachium album* oraz porównanie wydajności ekspresji rekombinowanej (ProtK-G) oraz dzikiej wersji (ProtK-WT lub ProtK) proteiny K przy wykorzystaniu odpowiednio komórek *Pichia pastoris*/pD912-ProtK-G, *Pichia Pink*/pPink_HC-ProtK-WT lub komórek *Pichia pastoris*/pD912-ProtK.

Uzyskane, odpowiednio w przykładach 5, 6 lub 8, szczepy ekspresyjne *Pichia Pink*/pPink_HC-ProtK-WT, *Pichia pastoris*/pD912-ProtK-G lub *Pichia pastoris*/pD912-ProtK produkujące odpowiednio rekombinantową (ProtK-G), natywną (ProtK-WT) lub wersję dziką dojrzałego polipeptydu proteiny K, hodzi się w pożywkę MGY przez 18–20 godzin w temperaturze 30°C, do uzyskania gęstości komórek mierzonej spektrofotometrycznie przy długości fal 600 nm, OD₆₀₀ większego niż 2. Następnie uzyskanym inokulum zaszczepia się pożywkę BSM w stalowym reaktorze o objętości roboczej 20 dm³, zawierającą 2% glicerolu jako źródło węgla i prowadzi się hodowlę w 30°C przez 46–48 godzin, dodając glicerolu jako źródła węgla. Po uzyskaniu odpowiedniej gęstości komórek, hodowlę indukuje się do ekspresji genu zmodyfikowanej lub natywnej proteiny K poprzez dodanie metanolu. Hodowlę prowadzi się przez kolejne 3–4 dni (72–96 h) dodając metanol jako źródło węgla i induktora ekspresji genu. Na zakończenie hodowli oraz z próbek pobranych w trakcie hodowli, komórki drożdży oddziela się od płynu pohodowlanego poprzez wirowanie. Zawartość proteiny K w pożywkę pofermentacyjnej, po oddzieleniu komórek, w pożywkę pohodowlanej analizowana jest na żelu SDS-PAGE i/lub metodą HPLC.

Białko z pożywki pohodowlanej, po oddzieleniu biomasy drożdżowej przy pomocy wirowania lub za pomocą filtracji stycznej z użyciem kasety filtracyjnej o porach o wielkości 0,2 mikrometra, oczyszczane jest metodą chromatografii oddziaływań hydrofobowych z wykorzystaniem złoża np. Phenyl Sepharose HS (Cytiva nr kat. 17097399) lub metodą chromatografii na złożu hydroksypatytu, a następnie oczyszcza się prowadząc dializę wobec roztworu węglanu amonu, po czym suszy się ze stanu zamrożenia uzyskując aktywnie enzymatycznie końcowy produkt.

Porównanie poziomu ekspresji w czasie hodowli, po 2, 3 oraz 4 dniach od indukcji w szczepach *Pichia Pink*/pPink_HC-ProtK-WT, *Pichia pastoris*/pD912-ProtK-G i *Pichia pastoris*/pD912-ProtK przedstawiona jest na rycinie 3. Wydajność ekspresji rekombinowanej proteiny K (ProtK-G) w szczepie *Pichia pastoris*/pD912-ProtK-G jest wyższa o około 10% w porównaniu z poziomem ekspresji natywnej proteiny (ProtK) w tym samym układzie ekspresyjnym w szczepie *Pichia pastoris*/pD912-ProtK i aż o około 30% w porównaniu do ekspresji w szczepie *Pichia Pink*/pPink_HC-ProtK-WT. Szczep *Pichia Pink*™ to komercyjny system ekspresyjny oferowany przez firmę Thermo Fisher, USA zoptymalizowany do wysokowydajnej produkcji białek rekombinantowych w układzie drożdżowym *Pichia Pink*. Mimo użycia zoptymalizowanego gospodarza ekspresyjnego, produkcja białka w natywnej sekwencji była o 30% niższa w porównaniu do dzikich komórek *Pichia pastoris*, przy użyciu tego samego protokołu wytwórczego.

Szczep ekspresyjny *Pichia pastoris*/pD912-ProtK-G okazał się najbardziej wydajny do produkcji proteiny K *Tritirachium album*. Przy zachowaniu 10% różnicy wydajności i przeskalowaniu procesu do produkcyjnej skali 1000L, można w ten sposób wytworzyć o 260 g białka więcej z jednej fermentacji.

Przykład 10. Możliwość szybszej dezaktywacji i usuwania proteiny K o zmodyfikowanej sekwencji, ProtK-G względem natywnej proteiny K, ProtK-WT w procesie oczyszczania kwasów nukleinowych.

Proteinaza K z *Tritirachium album*, jak endopeptydaza o szerokim spektrum działania i bardzo wysokiej aktywności specyficznej, ma szerokie zastosowanie podczas izolacji kwasów nukleinowych (DNA i RNA) bez naruszania ich struktury do wytrawiania białek, włączając DNazy i RNazy i inne białka oddziałujące z kwasami nukleinowymi, które zmniejszają czystość końcowych preparatów DNA lub RNA. Proteinaza K zachowuje wysoką aktywność w szerokim spektrum warunków buforowych i jest dodatkowo aktywowana przez dodatek niewielkiej ilości surfaktanta SDS, w zakresie 0,1–0,5%. Efektywne wytrawienie zanieczyszczeń białkowych, włącznie z dodaną w procesie izolacji proteinazą K, jest warunkiem uzyskania wysokiej jakości preparatów DNA lub RNA a wszelkie pozostałości białkowe podczas procesu oczyszczania, negatywnie wpływają na jakość oczyszczanych próbek. Jest to niezwykle istotne w przygotowywaniu próbek używanych do sekwencjonowania, diagnostyki z użyciem qPCR, gdzie zanieczyszczenia białkowe w przygotowanych próbkach powodować mogą zafałszowanie uzyskiwanych wyników.

Dla wykazania możliwości wydajnej dezaktywacji i usuwania proteiny K w procesie oczyszczania DNA i/lub RNA, w wariantach proteiny K będącym przedmiotem niniejszego wynalazku, porównano aktywność obu wariantów (ProtK-G o sekwencji nr 12 i ProtK-WT o sekwencji nr 1) w warunkach typowych podczas procesu oczyszczania kwasów nukleinowych (inkubacja w 10 mM Tris-HCl pH7.5, 37°C) bez czynnika stresującego, jak i w warunkach stresowych, z dodatkiem 1% SDS (inkubacja w 10 mM Tris-HCl pH7.5, 1% SDS, 37°C).

Nieoczekiwanie, okazało się, że wersja rekombinantowa proteiny K, tj. ProtK-G traci aktywność w szybszym tempie w warunkach stresowych w porównaniu do wersji dzikiej, tj. ProtK-WT, co pozwala na łatwiejsze usuwanie pozostałości proteiny K po etapie wytrawiania białek oddziałujących z kwasami nukleinowymi podczas izolacji DNA i/lub RNA, w przeciwieństwie do białka natywnego, ProtK-WT. Wyniki porównania obu białek w warunkach inkubacji standardowych i stresowych przedstawione są na rycinie 4. Zdolność do trawienia wykazywana przez mutanta po inkubacji w warunkach stresowych pozostaje na poziomie zaledwie 1,9%, w porównaniu z 69,7% wyjściowej aktywności. Białko referencyjne, ProtK-WT utrzymuje ponad 30% zdolność do trawienia, po inkubacji w takich samych warunkach z surfaktantem.

Uzyskane wyniki sugerują, że w odpowiednich warunkach (1% surfaktant) możliwa jest prawie całkowita inaktywacja enzymu w wariantie zmutowanym, ProtK-G, według niniejszego zgłoszenia bez użycia dedykowanego odczynnika/inhibitora [Pefabloc® SC] czy drastycznego podwyższania temperatury, jak w przypadku innych rekombinantów proteiny K i produktów dostępnych rynkowo znanych ze stanu techniki wynikających ze zgłoszeń patentowych EP3670655A1 [New England Biolabs Termolabilna Proteinaza K], czy WO2019170809A1 [Arctic Zymes Proteinaza]. Dodatek 1% jonowego związku powierzchniowo czynnego (surfaktant, laurylosiarczanu sodu- SDS) oraz 30 minutowa inkubacja w podwyższonej temperaturze 37°C, powoduje spadek aktywności zmutowanej proteiny K, ProtK-G wg niniejszego wynalazku niemal do zera w przeciwieństwie do wariantu natywnego, ProtK-WT.

Daje to możliwość łatwej inaktywacji enzymu, zwłaszcza w procesach oczyszczania kwasów nukleinowych, gdzie wymagana jest wysoka czystość preparatów DNA i RNA, bez stosowania dodatkowych czynników chemicznych czy fizycznych, które mogą uszkadzać izolowane preparaty kwasów nukleinowych.

SEQUENCE LISTING

<110> Blirt S.A.

<120> Mutant proteiny K Tritirachium album i jej zymogenu,
plazmid
ekspresyjny, rekombinantowy szczep Pichia pastoris i sposób
wytwarzania dojrzalej formy mutantu proteiny K

<130> PK/8738/RW

<160> 22

<170> PatentIn version 3.5

<210> 1

<211> 384

<212> PRT

<213> Tritirachium album

<400> 1

Met Arg Leu Ser Val Leu Leu Ser Leu Leu Pro Ser Ser Leu Gly Ala
1 5 10 15

Pro Ala Val Glu Gln Arg Ser Glu Ala Ala Pro Leu Ile Glu Ala Arg
20 25 30

Gly Glu Met Val Ala Asn Lys Tyr Ile Val Lys Phe Lys Glu Gly Ser
35 40 45

Ala Leu Ser Ala Leu Asp Ala Ala Met Glu Lys Ile Ser Gly Lys Pro
50 55 60

Asp His Val Tyr Lys Asn Val Phe Ser Gly Phe Ala Ala Thr Leu Asp
65 70 75 80

Glu Asn Met Val Arg Val Leu Arg Ala His Pro Asp Val Glu Tyr Ile
85 90 95

Glu Gln Asp Ala Val Val Thr Ile Asn Ala Ala Gln Thr Asn Ala Pro
100 105 110

Trp Gly Leu Ala Arg Ile Ser Ser Thr Ser Pro Gly Thr Ser Thr Tyr
115 120 125

Tyr Tyr Asp Glu Ser Ala Gly Gln Gly Ser Cys Val Tyr Val Ile Asp
 130 135 140

Thr Gly Ile Glu Ala Ser His Pro Glu Phe Glu Gly Arg Ala Gln Met
 145 150 155 160

Val Lys Thr Tyr Tyr Tyr Ser Ser Arg Asp Gly Asn Gly His Gly Thr
 165 170 175

His Cys Ala Gly Thr Val Gly Ser Arg Thr Tyr Gly Val Ala Lys Lys
 180 185 190

Thr Gln Leu Phe Gly Val Lys Val Leu Asp Asp Asn Gly Ser Gly Gln
 195 200 205

Tyr Ser Thr Ile Ile Ala Gly Met Asp Phe Val Ala Ser Asp Lys Asn
 210 215 220

Asn Arg Asn Cys Pro Lys Gly Val Val Ala Ser Leu Ser Leu Gly Gly
 225 230 235 240

Gly Tyr Ser Ser Ser Val Asn Ser Ala Ala Ala Arg Leu Gln Ser Ser
 245 250 255

Gly Val Met Val Ala Val Ala Ala Gly Asn Asn Asn Ala Asp Ala Arg
 260 265 270

Asn Tyr Ser Pro Ala Ser Glu Pro Ser Val Cys Thr Val Gly Ala Ser
 275 280 285

Asp Arg Tyr Asp Arg Arg Ser Ser Phe Ser Asn Tyr Gly Ser Val Leu
 290 295 300

Asp Ile Phe Gly Pro Gly Thr Ser Ile Leu Ser Thr Trp Ile Gly Gly
 305 310 315 320

Ser Thr Arg Ser Ile Ser Gly Thr Ser Met Ala Thr Pro His Val Ala
 325 330 335

Gly Leu Ala Ala Tyr Leu Met Thr Leu Gly Lys Thr Thr Ala Ala Ser
 340 345 350

Ala Cys Arg Tyr Ile Ala Asp Thr Ala Asn Lys Gly Asp Leu Ser Asn
 355 360 365

Ile Pro Phe Gly Thr Val Asn Leu Leu Ala Tyr Asn Asn Tyr Gln Ala
 370 375 380

<210> 2
 <211> 15
 <212> PRT
 <213> Tritirachium album

<400> 2

Met Arg Leu Ser Val Leu Leu Ser Leu Leu Pro Ser Ser Leu Gly
 1 5 10 15

<210> 3
 <211> 90
 <212> PRT
 <213> Tritirachium album

<400> 3

Ala Pro Ala Val Glu Gln Arg Ser Glu Ala Ala Pro Leu Ile Glu Ala
 1 5 10 15

Arg Gly Glu Met Val Ala Asn Lys Tyr Ile Val Lys Phe Lys Glu Gly
 20 25 30

Ser Ala Leu Ser Ala Leu Asp Ala Ala Met Glu Lys Ile Ser Gly Lys
 35 40 45

Pro Asp His Val Tyr Lys Asn Val Phe Ser Gly Phe Ala Ala Thr Leu
 50 55 60

Asp Glu Asn Met Val Arg Val Leu Arg Ala His Pro Asp Val Glu Tyr
 65 70 75 80

Ile Glu Gln Asp Ala Val Val Thr Ile Asn
 85 90

<210> 4
 <211> 279

<212> PRT

<213> Tritirachium album

<400> 4

Ala Ala Gln Thr Asn Ala Pro Trp Gly Leu Ala Arg Ile Ser Ser Thr
 1 5 10 15

Ser Pro Gly Thr Ser Thr Tyr Tyr Tyr Asp Glu Ser Ala Gly Gln Gly
 20 25 30

Ser Cys Val Tyr Val Ile Asp Thr Gly Ile Glu Ala Ser His Pro Glu
 35 40 45

Phe Glu Gly Arg Ala Gln Met Val Lys Thr Tyr Tyr Tyr Ser Ser Arg
 50 55 60

Asp Gly Asn Gly His Gly Thr His Cys Ala Gly Thr Val Gly Ser Arg
 65 70 75 80

Thr Tyr Gly Val Ala Lys Lys Thr Gln Leu Phe Gly Val Lys Val Leu
 85 90 95

Asp Asp Asn Gly Ser Gly Gln Tyr Ser Thr Ile Ile Ala Gly Met Asp
 100 105 110

Phe Val Ala Ser Asp Lys Asn Asn Arg Asn Cys Pro Lys Gly Val Val
 115 120 125

Ala Ser Leu Ser Leu Gly Gly Gly Tyr Ser Ser Ser Val Asn Ser Ala
 130 135 140

Ala Ala Arg Leu Gln Ser Ser Gly Val Met Val Ala Val Ala Ala Gly
 145 150 155 160

Asn Asn Asn Ala Asp Ala Arg Asn Tyr Ser Pro Ala Ser Glu Pro Ser
 165 170 175

Val Cys Thr Val Gly Ala Ser Asp Arg Tyr Asp Arg Arg Ser Ser Phe
 180 185 190

Ser Asn Tyr Gly Ser Val Leu Asp Ile Phe Gly Pro Gly Thr Ser Ile
 195 200 205

Leu Ser Thr Trp Ile Gly Gly Ser Thr Arg Ser Ile Ser Gly Thr Ser
 210 215 220

Met Ala Thr Pro His Val Ala Gly Leu Ala Ala Tyr Leu Met Thr Leu
 225 230 235 240

Gly Lys Thr Thr Ala Ala Ser Ala Cys Arg Tyr Ile Ala Asp Thr Ala
 245 250 255

Asn Lys Gly Asp Leu Ser Asn Ile Pro Phe Gly Thr Val Asn Leu Leu
 260 265 270

Ala Tyr Asn Asn Tyr Gln Ala
 275

<210> 5
 <211> 1173
 <212> DNA
 <213> *Tritirachium album*

<400> 5
 gaattcgaaa cgatgagact ctccgtcctc ctctccctcc tcccctcgag tctgggcgct
 60

ccagctgttg aacaaagatc tgaagctgct ccattgatcg aggctagagg tgaaatggtt
 120

gctaacaagt acatcgttaa gttcaaagag ggttctgctt tgtcagcttt ggacgctgct
 180

atggaaaaga tctctggtaa gccagaccac gtttacaaga acgttttctc cggtttcgct
 240

gctactttgg acgagaacat ggtagagtt ttgagagctc acccagacgt tgagtacatt
 300

gagcaagacg ctgttggtac tatcaacgct gtcagacta acgctccatg gggtttggt
 360

agaatttctt ctacttcccc aggtacttcc atttactact acgacgaatc tgctggtcag
 420

ggttcctgtg ttacgttat cgacactggt attgaggctt cccaccaga atttgagggt
 480

agagctcaga tggtaagac ttactattac tcctccagag atggtaacgg tcacggtact
 540

cattgtgctg gtactgttgg ttccagaact tacgggtgtg ctaagaaaac tcagttgttc
600

gggtgtaagg ttttgacga caacgggtcc ggtcagtact ccactattat tgctggatatg
660

gacttcgttg cttccgacaa gaacaacaga aactgtccaa aggggtgttgt tgcttccttg
720

tctttgggtg gtggttactc ttcttccgtt aactctgctg ctgctagatt gcagtcctcc
780

gggtgttatgg ttgctgttgc tgctggtaac aacaacgctg acgctagaaa ctactctcca
840

gcttctgaac catccgtttg tacagttggt gcttctgaca gatacgacag aagatcctcc
900

ttctccaact acggttccgt tttggacatt ttcggctctg gaacttccat cttgtccact
960

tggattggtg gttccactag atccatttcc ggtacttcta tggctactcc acacgttgct
1020

ggattggctg cttacttgat gactttgggt aagactactg ctgcttctgc ttgtagatat
1080

atcgctgaca cagctaacaa gggtgacttg tccaacatcc cattcggtag tggttaacttg
1140

ttggcttaca acaactacca ggcttaaggt acc
1173

<210> 6

<211> 41

<212> DNA

<213> artificial

<220>

<223> primer ProtK-G-F

<400> 6

caggcgctct tctatggctc cagctgttga acaaagatct g
41

<210> 7

<211> 41

<212> DNA

<213> artificial

<220>

<223> primer ProtK-G-R

<400> 7

gcagcgctct tctaccagcc tggtagttgt tgtaagccaa c
41

<210> 8

<211> 1139

<212> DNA

<213> artificial

<220>

<223> pro-ProtK-G DNA

<400> 8

caggcgctct tctatggctc cagctgttga acaaagatct gaagctgctc cattgatcga
60

ggctagaggt gaaatggttg ctaacaagta catcgtaag ttcaaagagg gttctgcttt
120

gtcagctttg gacgctgcta tggaaaagat ctctggtaag ccagaccacg tttacaagaa
180

cgttttctcc ggtttcgctg ctactttgga cgagaacatg gttagagttt tgagagctca
240

cccagacgtt gagtacattg agcaagacgc tgttggtact atcaacgctg ctcagactaa
300

cgctccatgg ggtttggcta gaatttcttc tacttcccca ggtacttcca cttactacta
360

cgacgaatct gctggtcagg gttcctgtgt ttacgttatc gacactggta ttgaggcttc
420

ccaccagaa tttagaggta gagctcagat ggtaagact tactattact cctccagaga
480

tggtaacggt cacggtactc attgtgctgg tactgttggg tccagaactt acggtgttgc
540

taagaaaact cagttgttcg gtgttaaggt ttggacgac aacggttccg gtcagtactc
600

cactattatt gctggtatgg acttcgttgc ttccgacaag aacaacagaa actgtccaaa
660

gggtgttggt gcttccttgt ctttgggtgg tggttactct tcttccgta actctgctgc
720

tgctagattg cagtcctccg gtgttatggg tgctgttgct gctggtaaca acaacgctga
780

cgctagaac tactctccag ctctgaacc atccgtttgt acagttggtg ctctgacag
840

atacgacaga agatcctcct tctccaacta cggttccggt ttggacattt tcggtcctgg
900

aacttccatc ttgtccactt ggattggtgg ttccactaga tccatttccg gtacttctat
960

ggctactcca cacgttgctg gattggctgc ttacttgatg actttgggta agactactgc
1020

tgcttctgct tgtagatata tcgctgacac agctaacaag ggtgacttgt ccaacatccc
1080

attcgg tact gttacttgt tggcttaca caactaccag gctggtagaa gagcgctgc
1139

<210> 9

<211> 370

<212> PRT

<213> artificial

<220>

<223> pro-ProtK-G protein

<400> 9

Ala	Pro	Ala	Val	Glu	Gln	Arg	Ser	Glu	Ala	Ala	Pro	Leu	Ile	Glu	Ala
1				5					10					15	

Arg	Gly	Glu	Met	Val	Ala	Asn	Lys	Tyr	Ile	Val	Lys	Phe	Lys	Glu	Gly
			20					25					30		

Ser	Ala	Leu	Ser	Ala	Leu	Asp	Ala	Ala	Met	Glu	Lys	Ile	Ser	Gly	Lys
		35					40					45			

Pro	Asp	His	Val	Tyr	Lys	Asn	Val	Phe	Ser	Gly	Phe	Ala	Ala	Thr	Leu
	50					55					60				

Asp	Glu	Asn	Met	Val	Arg	Val	Leu	Arg	Ala	His	Pro	Asp	Val	Glu	Tyr
65					70					75					80

Ile	Glu	Gln	Asp	Ala	Val	Val	Thr	Ile	Asn	Ala	Ala	Gln	Thr	Asn	Ala
				85					90					95	

Pro Trp Gly Leu Ala Arg Ile Ser Ser Thr Ser Pro Gly Thr Ser Thr
 100 105 110

Tyr Tyr Tyr Asp Glu Ser Ala Gly Gln Gly Ser Cys Val Tyr Val Ile
 115 120 125

Asp Thr Gly Ile Glu Ala Ser His Pro Glu Phe Glu Gly Arg Ala Gln
 130 135 140

Met Val Lys Thr Tyr Tyr Tyr Ser Ser Arg Asp Gly Asn Gly His Gly
 145 150 155 160

Thr His Cys Ala Gly Thr Val Gly Ser Arg Thr Tyr Gly Val Ala Lys
 165 170 175

Lys Thr Gln Leu Phe Gly Val Lys Val Leu Asp Asp Asn Gly Ser Gly
 180 185 190

Gln Tyr Ser Thr Ile Ile Ala Gly Met Asp Phe Val Ala Ser Asp Lys
 195 200 205

Asn Asn Arg Asn Cys Pro Lys Gly Val Val Ala Ser Leu Ser Leu Gly
 210 215 220

Gly Gly Tyr Ser Ser Ser Val Asn Ser Ala Ala Ala Arg Leu Gln Ser
 225 230 235 240

Ser Gly Val Met Val Ala Val Ala Ala Gly Asn Asn Asn Ala Asp Ala
 245 250 255

Arg Asn Tyr Ser Pro Ala Ser Glu Pro Ser Val Cys Thr Val Gly Ala
 260 265 270

Ser Asp Arg Tyr Asp Arg Arg Ser Ser Phe Ser Asn Tyr Gly Ser Val
 275 280 285

Leu Asp Ile Phe Gly Pro Gly Thr Ser Ile Leu Ser Thr Trp Ile Gly
 290 295 300

Gly Ser Thr Arg Ser Ile Ser Gly Thr Ser Met Ala Thr Pro His Val
 305 310 315 320

Ala Gly Leu Ala Ala Tyr Leu Met Thr Leu Gly Lys Thr Thr Ala Ala
325 330 335

Ser Ala Cys Arg Tyr Ile Ala Asp Thr Ala Asn Lys Gly Asp Leu Ser
340 345 350

Asn Ile Pro Phe Gly Thr Val Asn Leu Leu Ala Tyr Asn Asn Tyr Gln
355 360 365

Ala Gly
370

```
<210> 10
<211> 7667
<212> DNA
<213> artificial
```

<220>
<223> pPink_HC plasmid

```
<400> 10
agatctaaca tccaaagacg aaagggtgaa tgaaaccttt ttgccatccg acatccacag
60
```

gtccattctc acacataagt gccaaacgca acaggagggg atacactagc agcagaccgt
120

tgcaaacgca ggacctccac tcctcttctc ctcaacaccc acttttgcca tcgaaaaacc
180

agcccagtta ttgggcttga ttggagctcg ctcattccaa ttccttctat taggctacta
240

acaccatgac tttattagcc tgtctatcct ggccccctg gcgaggttca tgtttgttta
300

tttccgaatg caacaagctc cgcattacac ccgaacatca ctccagatga gggctttctg
360

agtgtgggggt caaatagttt catgttcccc aaatggccca aaactgacag tttaaacgct
420

gtcttggaac ctaatatgac aaaagcgtga tctcatccaa gatgaactaa gtttggttcg
480

ttgaaatgct aacggccagt tgggtcaaaaa gaaacttcca aaagtcggca taccgtttgt
540

cttgtttggt attgattgac gaatgctcaa aaataatctc attaatgctt agcgagctct
600

ctctatcgct tctgaacccc ggtgcacctg tgccgaaacg caaatgggga aacacccgct
660

ttttggatga ttatgcattg tctccacatt gtatgcttcc aagattctgg tgggaatact
720

gctgatagcc taacgttcat gatcaaaatt taactgttct aaccctact tgacagcaat
780

atataaacag aaggaagctg cctgtctta aacctttttt tttatcatca ttattagctt
840

actttcataa ttgcgactgg ttccaattga caagcttttg attttaacga cttttaacga
900

caacttgaga agatcaaaaa acaactaatt attcgaaacg gaattccgga ccggcatgcc
960

aaggcctcag gtaccggccg gccatttaaa tacaggcccc ttttcctttg tcgatatcat
1020

gtaattagtt atgtcacgct tacattcacg cctcctccc acatccgctc taaccgaaaa
1080

ggaaggagtt agacaacctg aagtctaggt ccctatttat tttttttaat agttatgtta
1140

gtattaagaa cgttatttat atttcaaatt tttctttttt ttctgtacaa acgcgtgtac
1200

gcatgtaaca ttatactgaa aaccttgctt gagaagggtt tgggacgctc gaaggcttta
1260

atttgcaagc tggatccgcg gccgccttcc aaactctcat ggattctcag gtaataggta
1320

ttctaggagg aggccagcta ggccgaatga ttgttgaggc cgctagcagg ctcaatatca
1380

agaccgtgat tcttgatgat ggtttttcac ctgctaagca cattaatgct gcgcaagacc
1440

acatcgacgg atcattcaaa gatgaggagg ctatcgccaa gttagctgcc aaatgtgatg
1500

ttctcactgt agagattgag catgtcaaca cagatgctct aaagagagtt caagacagaa
1560

ctggaatcaa gatatatcct ttaccagaga caatcgaact aatcaaggat aagtacttgc
1620

aaaaggaaca tttgatcaag cacaacattt cggtgacaaa gtctcagggg atagaatcta
1680

atgaaaaggc gctgcttttg tttggagaag agaattggatt tccatatctg ttgaagtccc
1740

ggactatggc ttatgatgga agaggcaatt ttgtagtgga gtctaaagag gacatcagta
1800

aggcattaga gttcttgaaa gatcgtccat tgtatgccga gaagtttgct ctttttgta
1860

aagaattagc ggtaatgggt gtgagatcac tggaaggcga agtattctcc tacccaaccg
1920

tagaaactgt gcacaaggac aatatctgtc atattgtgta tgctccggcc agagttaatg
1980

acaccatcca aaagaaagct caaatattag ctgaaaacac tgtgaagact ttcccaggcg
2040

ctggaatctt cggagttgag atgttcctat tgtctgatgg agaacttctt gtaaagtaga
2100

ttgctccaag gccccacaat tctggtcact atacaatcga tgcattgtga acatctcagt
2160

tcgaagcaca tgtaagagcc ataactgggc tgccaatgcc actagatttc accaaactat
2220

ctacttccaa caccaacgct attatgctca atgttttggg tgctgaaaaa tctcacgggg
2280

aattagagtt ttgtagaaga gccttagaaa caccgggtgc ttctgtatat ctgtacggaa
2340

agaccacccg attggctcgt aagatgggtc atatcaacat aataggatct tccatgttgg
2400

aagcagaaca aaagttagag tacattctag aagaatcaac ccacttacca tccagtactg
2460

tatcagctga cactaaaccg ttggttgag ttatcatggg ttcagactct gatctacctg
2520

tgatttcgaa aggttgcat attttaaaac agtttggtgt tccattcgaa gttactattg
2580

tctctgctca tagaacacca cagagaatga ccagatatgc ctttgaagcc gctagtagag
2640

gtatcaaggc tatcattgca ggtgctggtg gtgctgctca tcttcaggga atggttgctg
2700

ccatgactcc gttgccagtc attggtgttc ctgtcaaggg ctctacgttg gatggtgtag
2760

actcgctaca ctcgattgtc caaatgccta gaggtgttcc tgtggctacg gttgctatca
2820

acaacgccac caatgccgct ctgttggcca tcaggatttt aggtacaatt gaccacaaat
2880

ggcaaaagga aatgtccaag tatatgaatg caatggagac cgaagtgttg gggaaggcat
2940

ccaacttga atctgaaggg tatgaatcct atttgaagaa tcgtctttga atttagtatt
3000

gttttttaat agatgtatat ataatagtac acgtaactta tctattccat tcataatttt
3060

attttaaggg ttcggtagaa atttgtcctc caaaaagttg gttagagcct ggcagttttg
3120

ataggcatta ttatagattg ggtaatatat accctgcacc tggaggaact ttgcaaagag
3180

cctcatgtgc ggcgcgccag gccataatgg ccaaacggtt tctcaattac tatatactac
3240

taaccattta cctgtagcgt atttcttttc cctcttcgag aaagctcaag ggcattctct
3300

tgactcatga aaaatatctg gatttcttct gacagatcat cacccttgag cccaactctc
3360

tagcctatga gtgtaagtga tagtcatctt gcaacagatt attttggaac gcaactaaca
3420

aagcagatac acccttcagc agaatccttt ctggatattg tgaagaatga tcgccaaagt
3480

cacagtcttg agacagttcc taatctttac cccatttaca agttcatcca atcagacttc
3540

ttaacgcctc atctggctta tatcaagctt accaacagtt cagaaactcc cagtccaagt
3600

ttcttgcttg aaagtgcgaa gaatggtgac accgttgaca ggtacacctt tatgggacat
3660

tccccagaa aaataatcaa gactgggcct ttagaggggtg ctgaagtga ccccttggtg
3720

cttctggaaa aagaactgaa gggcaccaga caagcgcaac ttcttggtat tcctcgtcta
3780

agtgggtggtg ccataggata catctcgtac gattgtatta agtactttga accaaaaact
3840

gaaagaaaac tgaaagatgt tttgcaactt ccggaagcag ctttgatgtt gttcgacacg
3900

atcgtggcctt ttgacaatgt ttatcaaaga ttccaggtaa ttggaaacgt ttctctatcc
3960

gttgatgact cggacgaagc tattcttgag aaatattata agacaagaga agaagtggaa
4020

aagatcagta aagtggatatt tgacaataaa actgttcctt actatgaaca gaaagatatt
4080

attcaaggcc aaacgttcac ctctaattt ggtcaggaag ggtatgaaaa ccatgttcgc
4140

aagctgaaag aacatattct gaaaggagac atcttccaag ctgttccttc tcaaagggta
4200

gccaggccga cctcattgca cccittcaac atctatcgtc atttgagaac tgtcaatcct
4260

tctccataca tgttctatat tgactatcta gacttccaag ttgttggtgc ttcacctgaa
4320

ttactagtta aatccgacaa caacaacaaa atcatcacac atcctattgc tggaactctt
4380

cccagaggta aaactatcga agaggacgac aattatgcta agcaattgaa gtcgtctttg
4440

aaagacaggg ccgagcacgt catgctggta gatttggcca gaaatgatat taaccgtgtg
4500

tgtgagccca ccagtaccac ggttgatcgt ttattgactg tggagagatt ttctcatgtg
4560

atgcatcttg tgtcagaagt cagtggaaca ttgagaccaa acaagactcg cttcgatgct
4620

ttcagatcca ttttcccagc aggaaccgtc tccggtgctc cgaaggtaag agcaatgcaa
4680

ctcataggag aattggaagg agaaaagaga ggtgtttatg cgggggcccgt aggacactgg
4740

tcgtacgatg gaaaatcgat ggacacatgt attgccttaa gaacaatggt cgtcaaggac
4800

ggtgtcgctt accttcaagc cggaggtgga attgtctacg attctgaccc ctatgacgag
4860

tacatcgaaa ccatgaacaa aatgagatcc aacaataaca ccatcttggg ggctgagaaa
4920

atctggaccg ataggttggc cagagacgag aatcaaagtg aatccgaaga aaacgatcaa
4980

tgaacggagg acgtaagtag gaatttatgg ttggccata atggcctagc ttggcgtaat
5040

catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac
5100

gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa ctcacattaa
5160

ttgcgttgcg ctactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat
5220

gaatcgcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc
5280

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg
5340

cggtataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag
5400

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc
5460

gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag
5520

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga
5580

ccctgccgct taccggatac ctgtccgcct ttctcccttc ggggaagcgtg gcgctttctc
5640

atagctcacg ctgtaggtat ctcaattcgg tgtaggtcgt tcgctccaag ctgggctgtg
5700

tgcacgaacc ccccgttcag cccgaccgct gcgccttata cggtaactat cgtcttgagt
5760

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca
5820

gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca
5880

ctagaaggac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag
5940

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca
6000

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg
6060

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa
6120

aaaggatctt cacctagatc cttttaaaatt aaaaatgaag ttttaaatca atctaaagta
6180

tatatgagta aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag
6240

cgatctgtct atttcgttca tccatagtgt cctgactccc cgtcgtgtag ataactacga
6300

tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac
6360

cggtccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc
6420

ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta
6480

gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac
6540

gctcgtcgtt tggatatggc tcattcagct ccggttccca acgatcaagg cgagttacat
6600

gatccccat gttgtgcaaa aaagcggta gtccttcgg tcctccgatc gttgtcagaa
6660

gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg
6720

tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag
6780

aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc
6840

cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct
6900

caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat
6960

cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg
7020

ccgcaaaaaa ggggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc
7080

aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta
7140

tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg
7200

tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct
7260

ttcgtctcgc gcgtttcggg gatgacgggtg aaaacctctg acacatgcag ctcccggaga
7320

cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag
7380

cgggtgttgg cgggtgtcgg ggctggctta actatgcggc atcagagcag attgtactga
7440

gagtgcacca tatgcgggtg gaaataccgc acagatgcgt aaggagaaaa taccgcatca
7500

ggcgccattc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg cgggcctctt
7560

cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt tgggtaacgc
7620

cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattg
7667

<210> 11
<211> 8805
<212> DNA
<213> artificial

<220>
<223> pPink_HC-ProtK-WT plasmid

<400> 11
agatctaaca tccaaagacg aaagggtgaa tgaaaccttt ttgccatccg acatccacag
60

gtccattctc acacataagt gccaaacgca acaggagggg atacactagc agcagaccgt
120

tgcaaacgca ggacctccac tcctcttctc ctcaacaccc acttttgcca tcgaaaaacc
180

agcccagtta ttgggcttga ttggagctcg ctcatccaa ttccttctat taggctacta
240

acaccatgac tttattagcc tgtctatcct ggccccctg gcgaggttca tgtttgttta
300

tttccgaatg caacaagctc cgcattacac ccgaacatca ctccagatga gggctttctg
360

agtgtggggg caaatagttt catgttcccc aaatggccca aaactgacag tttaaagcgt
420

gtcttggAAC ctaatatgac aaaagcgtga tctcatccaa gatgaactaa gtttggttcg
480

ttgaaatgct aacggccagt tggTcaaaaa gaaacttcca aaagtcggca taccgtttgt
540

cttgtttggg attgattgac gaatgctcaa aaataatctc attaatgctt agcgagctct
600

ctctatcgct tctgaacccc ggtgcacctg tgccgaaacg caaatgggga aacacccgct
660

ttttggatga ttatgcattg tctccacatt gtatgcttcc aagattctgg tgggaatact
720

gctgatagcc taacgttcat gatcaaaatt taactgttct aaccctact tgacagcaat
780

atataaacag aaggaagctg cctgtctta aacctttttt tttatcatca ttattagctt
840

actttcataa ttgcgactgg ttccaattga caagcttttg attttaacga cttttaacga
900

caacttgaga agatcaaaaa acaactaatt attcgaaacg gaattcgaaa cgatgagact
960

ctccgtcctc ctctccctcc tcccctcgag tctgggcgct ccagctgttg aacaagatc
1020

tgaagctgct ccattgatcg aggctagagg tgaaatgggt gctaacaagt acatcgtaa
1080

gttcaaagag ggttctgctt tgtcagcttt ggacgctgct atggaaaaga tctctggtaa
1140

gccagaccac gtttacaaga acgttttctc cggtttcgct gctactttgg acgagaacat
1200

ggttagagtt ttgagagctc acccagacgt tgagtacatt gagcaagacg ctgttgttac
1260

tatcaacgct gctcagacta acgctccatg gggtttggct agaatttctt ctacttcccc
1320

aggctacttcc acttactact acgacgaatc tgctgggtcag ggttcctgtg tttacgttat
1380

cgacactggg attgaggcct cccacccaga atttgagggt agagctcaga tggtaagac
1440

ttactattac tcctccagag atggtaacgg tcacgggtact cattgtgctg gtactgttgg
1500

ttccagaact tacggtgttg ctaagaaaac tcagttgttc ggtgttaagg ttttgacga
1560

caacgggttcc ggtcagtact ccactattat tgctgggtatg gacttcgttg cttccgacaa
1620

gaacaacaga aactgtccaa aggggtgttg tgcttccttg tctttgggtg gtggttactc
1680

ttcttccgtt aactctgctg ctgctagatt gcagtcctcc ggtgttatgg ttgctgttgc
1740

tgctggtaac aacaacgctg acgctagaaa ctactctcca gcttctgaac catccgtttg
1800

tacagttggg gcttctgaca gatacgacag aagatcctcc ttctccaact acggttccgt
1860

tttgacatt ttcggtcctg gaacttccat ctgtccact tggattgggt gttccactag
1920

atccatttcc ggtacttcta tggctactcc acacgttgct ggattggctg cttacttgat
1980

gactttgggt aagactactg ctgcttctgc ttgtagatat atcgctgaca cagctaacaa
2040

gggtgacttg tccaacatcc cattcggtac tgtaacttg ttggcttaca acaactacca
2100

ggcttaagg accggccggc catttaata caggccccct ttctttgtc gatatcatgt
2160

aattagttat gtcacgtta cattcacgcc ctctccac atccgctcta accgaaaagg
2220

aaggagttag acaacctgaa gtctagggtc ctatttattt ttttaatat ttatgttagt
2280

attaagaacg ttatttatat ttcaaattt tcttttttt ctgtacaaac gcgtgtacgc
2340

atgtaacatt atactgaaaa ccttgcttga gaaggttttg ggacgctcga aggctttaat
2400

ttgcaagctg gatccgcggc cgccttccaa actctcatgg attctcaggt aataggtatt
2460

ctaggaggag gccagctagg ccgaatgatt gttgaggccg ctagcaggct caatatcaag
2520

accgtgattc ttgatgatgg tttttcacct gctaagcaca ttaatgctgc gcaagaccac
2580

atcgacggat cattcaaaga tgaggaggct atcgccaagt tagctgccaa atgtgatgtt
2640

ctcactgtag agattgagca tgtcaacaca gatgctctaa agagagttca agacagaact
2700

ggaatcaaga tatatccttt accagagaca atcgaactaa tcaaggataa gtacttgcaa
2760

aaggaacatt tgatcaagca caacatttcg gtgacaaagt ctcagggtat agaatctaata
2820

gaaaaggcgc tgcttttgtt tggagaagag aatggatttc catatctgtt gaagtcccgg
2880

actatggctt atgatggaag aggcaatttt gtagtggagt ctaaagagga catcagtaag
2940

gcattagagt tcttgaaaga tcgtccattg tatgccgaga agtttgctcc ttttggtaaa
3000

gaattagcgg taatggttgt gagatcactg gaaggcgaag tattctccta cccaaccgta
3060

gaaactgtgc acaaggacaa tatctgtcat attgtgtatg ctccggccag agttaatgac
3120

accatccaaa agaaagctca aatattagct gaaaacactg tgaagacttt cccaggcgct
3180

ggaatcttcg gagttgagat gttcctattg tctgatggag aacttcttgt aatgagatt
3240

gctccaaggc cccacaattc tggtcactat acaatcgatg catgtgtaac atctcagttc
3300

gaagcacatg taagagccat aactggctcg ccaatgccac tagatttcac caaactatct
3360

attccaaca ccaacgctat tatgctcaat gttttgggtg ctgaaaaatc tcacggggaa
3420

ttagagtttt gtagaagagc cttagaaaca cccggtgctt ctgtatatct gtacggaaag
3480

accacccgat tggctcgtaa gatgggtcat atcaacataa taggatcttc catgttggaa
3540

gcagaacaaa agttagagta cattctagaa gaatcaaccc acttaccatc cagtactgta
3600

tcagctgaca ctaaaccgtt ggttggagtt atcatgggtt cagactctga tctacctgtg
3660

atttcgaaag gttgcgatat tttaaaacag ttgggtgttc cattcgaagt tactattgtc
3720

tctgtcata gaacaccaca gagaatgacc agatatgcct ttgaagccgc tagtagaggt
3780

atcaaggcta tcattgcagg tgctgggtgt gctgtcatc ttccaggaat ggttgctgcc
3840

atgactccgt tgccagtcac ttggtgttct gtcaagggtt ctacgttga ttggtgtagac
3900

tcgctacact cgattgtcca aatgcctaga ggtgttctg ttggtacggt tgctatcaac
3960

aacgccacca atgccgctct gttggccatc aggattttag gtacaattga ccacaaatgg
4020

caaaaggaaa tgtccaagta tatgaatgca atggagaccg aagtgttggg gaaggcatcc
4080

aacttggaa ctgaagggtg tgaatcctat ttgaagaatc gtctttgaat ttagtattgt
4140

tttttaatag atgtatatat aatagtacac gtaacttacc tattccattc ataattttat
4200

tttaaagggt cggtagaaat ttgtcctcca aaaagtgtgt tagagcctgg cagttttgat
4260

aggcattatt atagattggg taatatttac cctgcacctg gaggaacttt gcaaagagcc
4320

tcattgtcgg cgcgccaggc cataatggcc aaacggtttc tcaattacta tatactacta
4380

accatttacc tgtagcgtat ttcttttccc tcttcgcgaa agctcaaggg catcttcttg
4440

actcatgaaa aatatctgga tttcttctga cagatcatca cccttgagcc caactctcta
4500

gcctatgagt gtaagtgata gtcactttgc aacagattat ttggaacgc aactaacaaa
4560

gcagatacac ccttcagcag aatcctttct ggatattgtg aagaatgatc gccaaagtca
4620

cagtcctgag acagttccta atctttaccc catttacaag ttcattccaat cagactttctt
4680

aacgcctcat ctggcttata tcaagcttac caacagttca gaaactccca gtccaagttt
4740

cttgcttgaa agtgcgaaga atggtgacac cgttgacagg tacaccttta tgggacattc
4800

ccccagaaaa ataatacaaga ctgggccttt agagggtgct gaagttgacc ccttggtgct
4860

tctggaaaaa gaactgaagg gcaccagaca agcgcaactt cctggtattc ctctgctaag
4920

tgggtggtgcc ataggataca tctcgtacga ttgtattaag tactttgaac caaaaactga
4980

aagaaaactg aaagatgttt tgcaacttcc ggaagcagct ttgatgttgt tcgacacgat
5040

cgtggctttt gacaatgttt atcaaagatt ccaggttaatt ggaaacgttt ctctatccgt
5100

tgatgactcg gacgaagcta ttcttgagaa atattataag acaagagaag aagtggaaaa
5160

gatcagtaaa gtggtatttg acaataaaac tgttccctac tatgaacaga aagatattat
5220

tcaaggccaa acgttcacct ctaatatagg tcaggaaggg tatgaaaacc atgttcgcaa
5280

gctgaaagaa catattctga aaggagacat cttccaagct gttccctctc aaagggtagc
5340

caggccgacc tcattgcacc ctttcaacat ctatcgatcat ttgagaactg tcaatccttc
5400

tccatacatg ttctatatgg actatctaga cttccaagtt gttggtgctt cacctgaatt
5460

actagttaaa tccgacaaca acaacaaaat catcacacat cctattgctg gaactcttcc
5520

cagaggtaaa actatcgaag aggacgacaa ttatgctaag caattgaagt cgtctttgaa
5580

agacagggcc gagcacgtca tgctggtaga tttggccaga aatgatatta accgtgtgtg
5640

tgagcccacc agtaccacgg ttgatcgttt attgactgtg gagagatttt ctcattgtgat
5700

gcattcttgtg tcagaagtca gtggaacatt gagaccaaac aagactcgct tcgatgcttt
5760

cagatccatt ttcccagcag gaaccgtctc cgggtgctccg aaggtaagag caatgcaact
5820

cataggagaa ttggaaggag aaaagagagg tgtttatgcg ggggccgtag gacactggtc
5880

gtacgatgga aaatcgatgg acacatgtat tgccttaaga acaatggctg tcaaggacgg
5940

tgtcgcttac cttcaagccg gaggtggaat tgtctacgat tctgaccct atgacgagta
6000

catcgaacc atgaacaaaa tgagatccaa caataacacc atcttgagg ctgagaaaat
6060

ctggaccgat aggttggcca gagacgagaa tcaaagtga tccgaagaaa acgatcaatg
6120

aacggaggac gtaagtagga atttatggtt tggccataat ggcctagctt ggcgtaatca
6180

tggatcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga
6240

gccggaagca taaagtgtaa agcctggggg gcctaattgag tgagctaact cacattaatt
6300

gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga
6360

atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttctctgctc
6420

actgactcgc tgcgctcggt cgttcggtg cggcgagcgg tatcagctca ctcaaaggcg
6480

gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc
6540

cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc
6600

ccccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga
6660

ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc
6720

ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat
6780

agctcacgct gtaggtatct cagttcgggtg taggtcgttc gtcceaagct gggctgtgtg
6840

cacgaacccc ccgttcagcc cgaccgtgc gccttatccg gtaactatcg tcttgagtcc
6900

aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga
6960

gcgaggtatg taggcgggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact
7020

agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt
7080

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtgggttttt tgtttgcaag
7140

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg
7200

tctgacgctc agtggaacga aaactcacgt taagggattt tggatcatgag attatcaaaa
7260

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaataaat cttaaagtata
7320

tatgagtaaa cttgggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg
7380

atctgtctat ttcgttcac catagttgcc tgactccccg tcgtgtagat aactacgata
7440

cgggaggggt taccatctgg cccagtgct gcaatgatac cgcgagaccc acgctcaccg
7500

gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgag aagtggtcct
7560

gcaactttat ccgcctccat ccagtctatt aattggtgcc gggaagctag agtaagtagt
7620

tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc
7680

tcgtcgtttg gtagggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga
7740

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt
7800

aagttagccg cagtgttatt actcatgggt atggcagcac tgcataattc tcttactgtc
7860

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa
7920

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca
7980

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca
8040

aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct
8100

tcagcatctt ttactttcac cagcgtttct gggtagcaaa aacaggaag gcaaaatgcc
8160

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cttttttcaa
8220

tattattgaa gcatttatca gggttattgt ctcagagcg gatacatatt tgaatgtatt
8280

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc
8340

taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt
8400

cgtctcgcgc gtttcgggtga tgacggtgaa aacctctgac acatgcagct cccggagacg
8460

gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg
8520

ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat tgtactgaga
8580

gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
8640

cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcgggtcg ggcctcttcg
8700

ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttg ggtaacgcca
8760

gggttttccc agtcacgacg ttgtaaaacg acggccagtg aattg
8805

<210> 12
 <211> 460
 <212> PRT
 <213> artificial

<220>
 <223> ProtK-G protein

<400> 12

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Met Ala Pro Ala Val Glu Gln
 85 90 95

Arg Ser Glu Ala Ala Pro Leu Ile Glu Ala Arg Gly Glu Met Val Ala
 100 105 110

Asn Lys Tyr Ile Val Lys Phe Lys Glu Gly Ser Ala Leu Ser Ala Leu
 115 120 125

Asp Ala Ala Met Glu Lys Ile Ser Gly Lys Pro Asp His Val Tyr Lys
 130 135 140

Asn Val Phe Ser Gly Phe Ala Ala Thr Leu Asp Glu Asn Met Val Arg
 145 150 155 160

Val Leu Arg Ala His Pro Asp Val Glu Tyr Ile Glu Gln Asp Ala Val
 165 170 175

Val Thr Ile Asn Ala Ala Gln Thr Asn Ala Pro Trp Gly Leu Ala Arg
 180 185 190

Ile Ser Ser Thr Ser Pro Gly Thr Ser Thr Tyr Tyr Tyr Asp Glu Ser
 195 200 205

Ala Gly Gln Gly Ser Cys Val Tyr Val Ile Asp Thr Gly Ile Glu Ala
 210 215 220

Ser His Pro Glu Phe Glu Gly Arg Ala Gln Met Val Lys Thr Tyr Tyr
 225 230 235 240

Tyr Ser Ser Arg Asp Gly Asn Gly His Gly Thr His Cys Ala Gly Thr
 245 250 255

Val Gly Ser Arg Thr Tyr Gly Val Ala Lys Lys Thr Gln Leu Phe Gly
 260 265 270

Val Lys Val Leu Asp Asp Asn Gly Ser Gly Gln Tyr Ser Thr Ile Ile
 275 280 285

Ala Gly Met Asp Phe Val Ala Ser Asp Lys Asn Asn Arg Asn Cys Pro
 290 295 300

Lys Gly Val Val Ala Ser Leu Ser Leu Gly Gly Gly Tyr Ser Ser Ser
 305 310 315 320

Val Asn Ser Ala Ala Ala Arg Leu Gln Ser Ser Gly Val Met Val Ala
 325 330 335

Val Ala Ala Gly Asn Asn Asn Ala Asp Ala Arg Asn Tyr Ser Pro Ala
 340 345 350

Ser Glu Pro Ser Val Cys Thr Val Gly Ala Ser Asp Arg Tyr Asp Arg
 355 360 365

Arg Ser Ser Phe Ser Asn Tyr Gly Ser Val Leu Asp Ile Phe Gly Pro
 370 375 380

Gly Thr Ser Ile Leu Ser Thr Trp Ile Gly Gly Ser Thr Arg Ser Ile
 385 390 395 400

Ser Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Leu Ala Ala Tyr
 405 410 415

Leu Met Thr Leu Gly Lys Thr Thr Ala Ala Ser Ala Cys Arg Tyr Ile
 420 425 430

Ala Asp Thr Ala Asn Lys Gly Asp Leu Ser Asn Ile Pro Phe Gly Thr
 435 440 445

Val Asn Leu Leu Ala Tyr Asn Asn Tyr Gln Ala Gly
 450 455 460

<210> 13
 <211> 90
 <212> PRT
 <213> artificial

<220>
 <223> ProtK-G SP protein

<400> 13

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Met
 85 90

<210> 14
 <211> 1380

<212> DNA

<213> artificial

<220>

<223> ProtK-G DNA

<400> 14

atgagattcc catctat ttt caccgctgtc ttgttcgctg ctcctctgc attggctgcc
60

cctgttaaca ctaccactga agacgagact gctcaaattc cagctgaagc agttatcggt
120

tactctgacc ttgaggggtga tttcgacgtc gctgttttgc ctttctctaa ctccactaac
180

aacgggttgt tgttcattaa caccactatc gcttcattg ctgctaagga agaggggtgc
240

tctctcgaga aaagagagggc cgaagctatg gctccagctg ttgaacaaag atctgaagct
300

gctccattga tcgaggctag aggtgaaatg gttgctaaca agtacatcgt taagttcaaa
360

gagggttctg ctttgtcagc tttggacgtc gctatggaaa agatctctgg taagccagac
420

cacgtttaca agaacgtttt ctccggtttc gctgctactt tggacgagaa catgggttaga
480

gttttgagag ctcaccacaga cggtgagtag attgagcaag acgctgttgt tactatcaac
540

gctgctcaga ctaacgctcc atgggggttg gctagaattt cttctacttc cccaggtact
600

tccacttact actacgacga atctgctggt caggggttcct gtgtttacgt tatcgacact
660

ggtattgagg cttcccaccc agaatttgag ggtagagctc agatgggttaa gacttactat
720

tactcctcca gagatggtaa cggtcacggt actcattgtg ctggtactgt tggttccaga
780

acttacgggtg ttgctaagaa aactcagttg ttcgggtgta aggttttgga cgacaacggt
840

tccggtcagt actccactat tattgctggt atggacttcg ttgcttccga caagaacaac
900

agaaactgtc caaaggggtg tgttgcttcc ttgtctttgg gtgggtggtta ctcttcttcc
960

gttaactctg ctgctgctag attgcagtc tccgggtgta tggttgctgt tgctgctggt
1020

aacaacaacg ctgacgctag aaactactct ccagcttctg aaccatccgt ttgtacagtt
1080

ggtgcttctg acagatacga cagaagatcc tccttctcca actacggttc cgttttggac
1140

attttcggtc ctggaacttc catcttgtcc acttggattg gtggttccac tagatccatt
1200

tccggtactt ctatggctac tccacacggt gctggattgg ctgcttactt gatgactttg
1260

ggtaagacta ctgctgcttc tgcttgtaga tatatcgctg acacagctaa caagggtgac
1320

ttgtccaaca tcccattcgg tactgttaac ttgttggtt acaacaacta ccaggctggt
1380

<210> 15

<211> 3837

<212> DNA

<213> artificial

<220>

<223> pD912 plasmid

<400> 15

gctcttcagg ttaaaggggc ggccgctcaa gaggatgtca gaatgccatt tgcctgagag
60

atgcaggctt cttttttgat acttttttat ttgtaacctata tatagtatag gatttttttt
120

gtcattttgt ttcttctcgt acgagcttgc tcctgatcag cctatctcgc agcagatgaa
180

tatcttgtgg taggggtttg ggaaaatcat tcgagtttga tgtttttctt ggtatttccc
240

actcctcttc agagtacaga agattaagtg aaaccttcgt ttgtgcggat ccttcagtaa
300

tgtctttgtt cttttgttgc agtggtgagc cttttgact tcgtgaaagt ttcttttagaa
360

tagttgtttc cagaggccaa acattccacc cgtagtaaag tgcaagcgta ggaagaccaa
420

gactggcata aatcaggat aagtgtcgag cactggcagg tgatcttctg aaagtttcta
480

ctagcagata agatccagta gtcatgcata tggcaacaat gtaccgtgtg gatctaagaa
540

cgcgtcctac taaccttcgc attcgttggc ccagtttggt gttatcgatc aacgtgacaa
600

ggttgtcgat tccgcgtaag catgcatacc caaggacgcc tgttgcaatt ccaagtgagc
660

cagttccaac aatctttgta atattagagc acttcattgt gttgcgcttg aaagtaaaat
720

gcgaacaaat taagagataa tctcgaaacc gcgacttcaa acgccaatat gatgtgcggc
780

acacaataag cgttcatatc cgctgggtga ctttctcgct ttaaaaaatt atccgaaaaa
840

atcttctaga gtgttggtac ttatatactc cggctcgatc aatacgacaa ggtgtaagga
900

ggactaaacc atggctaaac tcacctctgc tgttccagtc ctgactgctc gtgatgttgc
960

tgggtgctgtt gagttctgga ctgatagact cggtttctcc cgtgacttcg tagaggacga
1020

ctttgccggt gttgtacgtg acgacgttac cctgttcac cccgcagttc aggaccaggt
1080

tgtgccagac aacactctgg catgggtatg gggtcgtggc ctggacgaac tgtacgctga
1140

gtggctctgag gtcgtgtcta ccaacttccg tgatgcatct ggtccagcta tgaccgagat
1200

cggatgaacag ccctggggtc gtgagtttgc actgcgtgat ccagctggta actgcgtgca
1260

tttcgtcgca gaagaacagg actaacaatt gacaccttac gattatttag agagtattta
1320

ttagttttat tgtatgtata cggatgtttt attatctatt tatgccctta tattctgtaa
1380

ctatccaaaa gtcctatctt atcaagccag caatctatgt ccgcgaacgt caactaaaaa
1440

taagcttttt atgctgttct ctcttttttt cccttcggta taattatacc ttgcatccac
1500

agattctcct gccaaatttt gcataatcct ttacaacatg gctatatggg agcacttagc
1560

gccctccaaa acccatattg cctacgcatg tataggtggt tttccacaa tttttctct
1620

gtgctctctt ttatttaaag agaagctcta tatcgagaa gcttctgtgg ccgttatatt
1680

cgcccttatt gtgggaccac attgcctgaa ttggtttgcc ccggaagatt ggggaaactt
1740

ggatctgatt accttagctg caggtaccac tgagcgtcag acccgtaga aaagatcaaa
1800

ggatcttctt gagatccttt tttctgcgc gtaatctgct gcttgcaaac aaaaaacca
1860

ccgctaccag cgggtggttg ttgcccggat caagagctac caactctttt tccgaaggta
1920

actggcttca gcagagcgca gataccaaat actgttcttc tagttagacc gtagttaggc
1980

caccatttca agaactctgt agcaccgcct acatacctcg ctctgctaatt cctgttacca
2040

gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta
2100

ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttgagg
2160

cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt
2220

cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc
2280

acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac
2340

ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac
2400

gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc
2460

tttcctgcgg taccagatc caattcccgc ttgactgcc tgaaatctcc atgcctaca
2520

atgatgacat ttggatttgg ttgactcatg ttggtattgt gaaatagacg cagatcggga
2580

acactgaaaa atacacagtt attattcatt taaataacat ccaaagacga aaggttgaat
2640

gaaacctttt tgccatccga catccacagg tccatttctca cacataagtg ccaaacgcaa
2700

caggagggga tacactagca gcagaccgtt gcaaacgcag gacctccact cctcttctcc
2760

tcaacaccca cttttgccat cgaaaaacca gcccagttat tgggcttgat tggagctcgc
2820

tcattccaat tccttctatt aggctactaa caccatgact ttattagcct gtctatcctg
2880

gccccctgg cgaggttcat gtttgtttat ttccgaatgc aacaagctcc gcattacacc
2940

cgaacatcac tccagatgag ggctttctga gtgtggggtc aaatagtffc atgttcccca
3000

aatggcccaa aactgacagt ttaaacgctg tcttgggaacc taatatgaca aaagcgtgat
3060

ctcatccaag atgaactaag tttgggtcgt tgaaatgcta acggccagtt ggtcaaaaag
3120

aaacttccaa aagtcggcat accgtttgtc ttgtttggta ttgattgacg aatgctcaaa
3180

aataatctca ttaatgctta gcgcagtcct tctatcgctt ctgaaccccg gtgcacctgt
3240

gccgaaacgc aaatggggaa acacccgctt ttgggatgat tatgcattgt ctccacattg
3300

tatgcttcca agattctggg ggggaatactg ctgatagcct aacgttcatg atcaaaattt
3360

aactgttcta acccctactt gacagcaata tataaacaga aggaagctgc cctgtcttaa
3420

accttttttt ttatcatcat tattagctta ctttcataat tgcgactggg tccaattgac
3480

aagcttttga ttttaacgac ttttaacgac aacttgagaa gatcaaaaaa caactaatta
3540

ttgaaagaat tccgaaacga tgagattccc atctattttc accgctgtct tgttcgctgc
3600

ctctctgca ttggctgccc ctgttaacac taccactgaa gacgagactg ctcaaattcc
3660

agctgaagca gttatcgggt actctgacct tgagggtgat ttcgacgtcg ctgttttgcc
3720

tttctctaac tccactaaca acggtttggt gttcattaac accactatcg cttccattgc
3780

tgctaaggaa gaggggtgtct ctctcgagaa aagagaggcc gaagctatgt gaagagc
3837

<210> 16

<211> 4928

<212> DNA

<213> artificial

<220>

<223> pD912-ProtK-G plasmid

<400> 16

aacatccaaa gacgaaaggt tgaatgaaac ctttttgcca tccgacatcc acaggtccat
60

tctcacacat aagtgccaaa cgcaacagga ggggatacac tagcagcaga ccgttgcaaa
120

cgcaggacct ccactcctct tctcctcaac acccactttt gccatcgaaa aaccagccca
180

gttattgggc ttgattggag ctcgctcatt ccaattcctt ctattaggct actaacacca
240

tgactttatt agcctgtcta tcctggcccc cctggcgagg ttcattgttg tttatttccg
300

aatgcaacaa gtcctgcatt acacccgaac atcactccag atgagggcct tctgagtgtg
360

gggtcaaata gtttcatgtt ccccaaattg cccaaaactg acagtttaaa cgctgtcttg
420

gaaccttaata tgacaaaagc gtgatctcat ccaagatgaa ctaagtttgg ttcgttgaaa
480

tgctaacggc cagttgggtca aaaagaaact tccaaaagtc ggcataccgt ttgtcttggt
540

tggtattgat tgacgaatgc tcaaaaataa tctcattaat gcttagcgca gtctctctat
600

cgcttctgaa ccccggtgca cctgtgccga aacgcaaatt gggaaacacc cgctttttgg
660

atgattatgc attgtctcca cattgtatgc ttccaagatt ctggtgggaa tactgctgat
720

agcctaacgt tcatgatcaa aatttaactg ttctaacccc tacttgacag caatatataa
780

acagaaggaa gctgccctgt cttaaaccct tttttttatc atcattatta gcttactttc
840

ataattgcga ctggttccaa ttgacaagct ttgatttta acgactttta acgacaactt
900

gagaagatca aaaaacaact aattattgaa agaattccga aacgatgaga ttcccatcta
960

ttttcaccgc tgtcttggtc gctgcctcct ctgcattggc tgcccctggt aacactacca
1020

ctgaagacga gactgctcaa attccagctg aagcagttat cggttactct gaccttgagg
1080

gtgatttcga cgctgctggt ttgcctttct ctaactccac taacaacggt ttgttggtca
1140

ttaacaccac tatcgcttcc attgctgcta aggaagaggg tgtctctctc gagaaaagag
1200

aggccgaagc tatggctcca gctggtgaac aaagatctga agctgctcca ttgatcgagg
1260

ctagaggtga aatgggttgc aacaagtaca tcgttaagtt caaagaggggt tctgctttgt
1320

cagctttgga cgctgctatg gaaaagatct ctggttaagcc agaccacggt tacaagaacg
1380

ttttctccgg tttcgctgct actttggacg agaacatggt tagagttttg agagctcacc
1440

cagacgttga gtacattgag caagacgctg ttgttactat caacgctgct cagactaacg
1500

ctccatgggg ttgggctaga atttcttcta ctccccagg tacttccact tactactacg
1560

acgaatctgc tggtcagggt tcctgtgttt acgttatcga cactgggtatt gaggcttccc
1620

accagaatt tgagggtaga gctcagatgg ttaagactta ctattactcc tccagagatg
1680

gtaacgggtca cggtactcat tgtgctggta ctgttggttc cagaacttac ggtgttgcta
1740

agaaaactca gttgttcggt gttaagggtt tggacgacaa cggttccggt cagtactcca
1800

ctattattgc tggatatggac ttcgttgctt ccgacaagaa caacagaaac tgtccaaagg
1860

gtgttggtgc ttccttgtct ttgggtggtg gttactcttc ttccgttaac tctgctgctg
1920

ctagattgca gtcctccggt gttatggttg ctgttgctgc tggtaacaac aacgctgacg
1980

ctagaaacta ctctccagct tctgaaccat ccgtttgtac agttggtgct tctgacagat
2040

acgacagaag atcctccttc tccaactacg gttccgtttt ggacattttc ggtcctggaa
2100

cttccatctt gtccacttgg attggtggtt ccactagatc catttccggt acttctatgg
2160

ctactccaca cgttgctgga ttggctgctt acttgatgac tttgggtaag actactgctg
2220

cttctgcttg tagatatatc gctgacacag ctaacaaggg tgacttgtcc aacatcccat
2280

tcggtactgt taacttggtg gcttacaaca actaccaggc tggttaaagg ggcggccgct
2340

caagaggatg tcagaatgcc atttgccatga gagatgcagg cttcattttt gatacttttt
2400

tatttgaac ctatatagta taggattttt tttgtcattt tgtttcttct cgtacgagct
2460

tgctcctgat cagcctatct cgcagcagat gaatatcttg tggtaggggt ttgggaaaat
2520

cattcgagtt tgatgttttt ctggtatatt cccactcctc ttcagagtac agaagattaa
2580

gtgaaacctt cgtttgtgcg gatccttcag taatgtcttg tttcttttgt tgcagtgggtg
2640

agccattttg acttcgtgaa agtttcttta gaatagttgt ttccagaggc caaacattcc
2700

acccgtagta aagtgaagc gtaggaagac caagactggc ataaatcagg tataagtgtc
2760

gagcactggc aggtgatctt ctgaaagttt ctactagcag ataagatcca gtagtcatgc
2820

atatggcaac aatgtaccgt gtggatctaa gaacgcgtcc tactaacctt cgcattcgtt
2880

ggtccagttt gttgttatcg atcaacgtga caaggttgtc gattccgcgt aagcatgcat
2940

acccaaggac gcctgttgca attccaagtg agccagttcc aacaatcttt gtaatattag
3000

agcatttcac tgtgttgccg ttgaaagtaa aatgcgaaca aattaagaga taatctcgaa
3060

accgcgactt caaacgccaa tatgatgtgc ggcacacaat aagcgttcat atccgctggg
3120

tgactttctc gctttaaaaa attatccgaa aaaattttct agagtgttgt tactttatac
3180

ttccggctcg tataatacga caagggtgaa ggaggactaa accatggcta aactcacctc
3240

tgctgttcca gtcctgactg ctcgtgatgt tgctgggtgt gttgagttct ggactgatag
3300

actcggtttc tcccgtgact tcgtagagga cgactttgcc ggtgttgtac gtgacgacgt
3360

taccctgttc atctccgcag ttcaggacca ggttgtgcca gacaacactc tggcatgggt
3420

atgggttcgt ggtctggacg aactgtacgc tgagtgggtct gaggtcgtgt ctaccaactt
3480

ccgtgatgca tctgggtccag ctatgaccga gatcggtgaa cagccctggg gtcgtgagtt
3540

tgcactgcgt gatccagctg gtaactgcgt gcatttcgtc gcagaagaac aggactaaca
3600

attgacacct tacgattatt tagagagtat ttattagttt tattgtatgt atacggatgt
3660

tttattatct atttatgccc ttatattctg taactatcca aaagtcctat cttatcaagc
3720

cagcaatcta tgtccgcgaa cgtcaactaa aaataagctt tttatgctgt tctctctttt
3780

tttcccttcg gtataattat accttgcatc cacagattct cctgccaaat ttgcataat
3840

cctttacaac atggctatat gggagcactt agcgcctcc aaaaccata ttgcctacgc
3900

atgtataggt gttttttcca caatatcttc tctgtgctct ctttttatta aagagaagct
3960

ctatatcgga gaagcttctg tggccgttat attcggcctt atcgtgggac cacattgcct
4020

gaattggttt gccccggaag attggggaaa cttggatctg attaccttag ctgcaggtag
4080

cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttctg
4140

cgcgtaatct gctgcttgca aacaaaaaaaa ccaccgctac cagcggtggt ttgtttgccg
4200

gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca
4260

aatactgttc ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcacccg
4320

cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg
4380

tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga
4440

acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac
4500

ctacagcgtg agctatgaga aagcgcacag cttcccgaag ggagaaaggc ggacaggtag
4560

ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc
4620

tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga
4680

tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt ttacggttc
4740

ctggcctttt gctggccttt tgctcacatg ttctttctg cggtaccag atccaattcc
4800

cgctttgact gcctgaaatc tccatgcctt acaatgatga catttggatt tggttgactc
4860

atgttggtag tgtgaaatag acgcagatcg ggaacactga aaaatacaca gttattattc
4920

atttaaat
4928

<210> 17

<211> 459

<212> PRT

<213> Tritirachium album

<400> 17

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Met Ala Pro Ala Val Glu Gln
 85 90 95

Arg Ser Glu Ala Ala Pro Leu Ile Glu Ala Arg Gly Glu Met Val Ala
 100 105 110

Asn Lys Tyr Ile Val Lys Phe Lys Glu Gly Ser Ala Leu Ser Ala Leu
 115 120 125

Asp Ala Ala Met Glu Lys Ile Ser Gly Lys Pro Asp His Val Tyr Lys
 130 135 140

Asn Val Phe Ser Gly Phe Ala Ala Thr Leu Asp Glu Asn Met Val Arg
 145 150 155 160

Val Leu Arg Ala His Pro Asp Val Glu Tyr Ile Glu Gln Asp Ala Val
 165 170 175

Val Thr Ile Asn Ala Ala Gln Thr Asn Ala Pro Trp Gly Leu Ala Arg
 180 185 190

Ile Ser Ser Thr Ser Pro Gly Thr Ser Thr Tyr Tyr Tyr Asp Glu Ser
 195 200 205

Ala Gly Gln Gly Ser Cys Val Tyr Val Ile Asp Thr Gly Ile Glu Ala
210 215 220

Ser His Pro Glu Phe Glu Gly Arg Ala Gln Met Val Lys Thr Tyr Tyr
225 230 235 240

Tyr Ser Ser Arg Asp Gly Asn Gly His Gly Thr His Cys Ala Gly Thr
245 250 255

Val Gly Ser Arg Thr Tyr Gly Val Ala Lys Lys Thr Gln Leu Phe Gly
260 265 270

Val Lys Val Leu Asp Asp Asn Gly Ser Gly Gln Tyr Ser Thr Ile Ile
275 280 285

Ala Gly Met Asp Phe Val Ala Ser Asp Lys Asn Asn Arg Asn Cys Pro
290 295 300

Lys Gly Val Val Ala Ser Leu Ser Leu Gly Gly Gly Tyr Ser Ser Ser
305 310 315 320

Val Asn Ser Ala Ala Ala Arg Leu Gln Ser Ser Gly Val Met Val Ala
325 330 335

Val Ala Ala Gly Asn Asn Asn Ala Asp Ala Arg Asn Tyr Ser Pro Ala
340 345 350

Ser Glu Pro Ser Val Cys Thr Val Gly Ala Ser Asp Arg Tyr Asp Arg
355 360 365

Arg Ser Ser Phe Ser Asn Tyr Gly Ser Val Leu Asp Ile Phe Gly Pro
370 375 380

Gly Thr Ser Ile Leu Ser Thr Trp Ile Gly Gly Ser Thr Arg Ser Ile
385 390 395 400

Ser Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Leu Ala Ala Tyr
405 410 415

Leu Met Thr Leu Gly Lys Thr Thr Ala Ala Ser Ala Cys Arg Tyr Ile
420 425 430

Ala Asp Thr Ala Asn Lys Gly Asp Leu Ser Asn Ile Pro Phe Gly Thr
435 440 445

Val Asn Leu Leu Ala Tyr Asn Asn Tyr Gln Ala
450 455

<210> 18
<211> 59
<212> DNA
<213> artificial

<220>
<223> ProtK-G-out-F primer

<400> 18
gcttacaaca actaccaggc ttgataaagg ggcggccgct caagaggatg tcagaatgc
59

<210> 19
<211> 36
<212> DNA
<213> artificial

<220>
<223> ProtK-G-out-R primer

<400> 19
gccgcccctt tatcaagcct ggtagttggt gtaagc
36

<210> 20
<211> 1380
<212> DNA
<213> Tritirachium album

<400> 20
atgagattcc catctatattt caccgctgtc ttgttcgctg cctcctctgc attggctgcc
60

cctgttaaca ctaccactga agacgagact gctcaaattc cagctgaagc agttatcggt
120

tactctgacc ttgagggtga tttcgacgtc gctgttttgc ctttctctaa ctccactaac
180

aacggtttgt tgttcattaa caccactatc gcttccattg ctgctaagga agaggggtgtc
240

tctctcgaga aaagagagggc cgaagctatg gctccagctg ttgaacaaag atctgaagct
300

gctccattga tcgaggctag aggtgaaatg gttgctaaca agtacatcgt taagttcaaa
360

gagggttctg ctttgtcagc tttggacgct gctatggaaa agatctctgg taagccagac
420

cacgtttaca agaacgtttt ctccggtttc gctgctactt tggacgagaa catggttaga
480

gttttgagag ctcacccaga cgttgagtac attgagcaag acgctgttgt tactatcaac
540

gctgctcaga ctaacgctcc atggggtttg gctagaattt cttctacttc cccagggtact
600

tccacttact actacgacga atctgctggg cagggttcct gtgtttacgt tatcgacact
660

ggtattgagg cttcccaccc agaatttgag ggtagagctc agatgggttaa gacttactat
720

tactcctcca gagatggtaa cggtcacggg atcattgtg ctggtactgt tggttccaga
780

acttacgggtg ttgctaagaa aactcagttg ttcgggtgta aggttttgga cgacaacggg
840

tccggtcagt actccactat tattgctggg atggacttcg ttgcttccga caagaacaac
900

agaaactgtc caaaggggtg tgttgcttcc ttgtctttgg gtggtgggta ctcttcttcc
960

gttaactctg ctgctgctag attgcagtc tccgggtgta tgggtgctgt tgctgctggg
1020

aacaacaacg ctgacgctag aaactactct ccagcttctg aaccatccgt ttgtacagtt
1080

ggtgcttctg acagatacga cagaagatcc tccttctcca actacgggtc cgttttggac
1140

attttcggtc ctggaacttc catcttgtcc acttggattg gtggttccac tagatccatt
1200

tccggctactt ctatggctac tccacacggt gctggattgg ctgcttactt gatgactttg
1260

ggtaagacta ctgctgcttc tgcttgtaga tatatcgctg acacagctaa caaggggtgac
1320

ttgtccaaca tcccattcgg tactgttaac ttgttggctt acaacaacta ccaggcttga
1380

<210> 21

<211> 4928

<212> DNA

<213> artificial

<220>

<223> pD912-ProtK plasmid

<400> 21

aacatccaaa gacgaaagggt tgaatgaaac ctttttgcca tccgacatcc acaggtccat
60

tctcacacat aagtgccaaa cgcaacagga ggggatacac tagcagcaga ccgttgcaaa
120

cgcaggacct ccactcctct tctcctcaac acccactttt gccatcgaaa aaccagccca
180

gttattgggc ttgattggag ctgctcatt ccaattcctt ctattaggct actaacacca
240

tgactttatt agcctgtcta tcctggcccc cctggcgagg ttcattgtttg tttatttccg
300

aatgcaacaa gctccgcatt acacccgaac atcactccag atgagggcct tctgagtgtg
360

gggtcaaata gtttcatgtt ccccaaattg cccaaaactg acagtttaaa cgctgtcttg
420

gaacctaata tgacaaaagc gtgatctcat ccaagatgaa ctaagtttgg ttcgttgaaa
480

tgctaacggc cagttgggtca aaaagaaact tccaaaagtc ggcataccgt ttgtcttggt
540

tggtattgat tgacgaatgc tcaaaaataa tctcattaat gcttagcgca gtctctctat
600

cgcttctgaa ccccggtgca cctgtgccga aacgcaaatt gggaaacacc cgctttttgg
660

atgattatgc attgtctcca cattgtatgc ttccaagatt ctggtgggaa tactgctgat
720

agcctaacgt tcatgatcaa aatttaactg ttctaacccc tacttgacag caatatataa
780

acagaaggaa gctgccctgt cttaaaccctt tttttttatc atcattattia gcttactttc
840

ataattgcga ctggttccaa ttgacaagct tttgatttta acgactttta acgacaactt
900

gagaagatca aaaaacaact aattattgaa agaattccga aacgatgaga ttcccatcta
960

ttttcaccgc tgtcttggtc gctgcctcct ctgcattggc tgcccctggt aacactacca
1020

ctgaagacga gactgctcaa attccagctg aagcagttat cggttactct gaccttgagg
1080

gtgatttcga cgtcgctggt ttgcctttct ctaactccac taacaacggt ttgttggtca
1140

ttaacaccac tatcgcttcc attgctgcta aggaagaggg tgtctctctc gagaaaagag
1200

aggccgaagc tatggctcca gctggtgaac aaagatctga agctgctcca ttgatcgagg
1260

ctagaggtga aatgggtgct aacaagtaca tcgttaagtt caaagaggggt tctgctttgt
1320

cagctttgga cgctgctatg gaaaagatct ctggtaagcc agaccacgtt tacaagaacg
1380

ttttctccgg tttcgctgct actttggacg agaacatggt tagagttttg agagctcacc
1440

cagacgttga gtacattgag caagacgctg ttgttactat caacgctgct cagactaacg
1500

ctccatgggg tttggctaga atttcttcta ctccccagg tacttccact tactactacg
1560

acgaatctgc tggtcaggggt tcctgtgttt acgttatcga cactggtatt gaggcttccc
1620

accagaatt tgagggtaga gtcagatgg ttaagactta ctattactcc tccagagatg
1680

gtaacggtca cggctactcat tgtgctggta ctgttggttc cagaacttac ggtgttgcta
1740

agaaaactca gttgttcggt gttaaggttt tggacgacaa cggttccggt cagtactcca
1800

ctattattgc tggtatggac ttcgttgctt ccgacaagaa caacagaaac tgtccaaagg
1860

gtgttgttgc ttccttgtct ttgggtggtg gtiactcttc ttccgttaac tctgctgctg
1920

ctagattgca gtcctccggt gttatggttg ctgttgctgc tggtaacaac aacgctgacg
1980

ctagaaacta ctctccagct tctgaaccat ccgtttgtac agttggtgct tctgacagat
2040

acgacagaag atcctccttc tccaactacg gttccgtttt ggacattttc ggtcctggaa
2100

cttccatctt gtccacttgg attggtggtt ccactagatc catttccggt acttctatgg
2160

ctactccaca cgttgctgga ttggctgctt acttgatgac tttgggtaag actactgctg
2220

cttctgcttg tagatatatc gctgacacag ctaacaaggg tgacttgtcc aacatcccat
2280

tccgtactgt taacttgttg gcttacaaca actaccaggc ttgataaagg ggcggccgct
2340

caagaggatg tcagaatgcc atttgcctga gagatgcagg cttcattttt gatacttttt
2400

tatttgtaac ctatatagta taggattttt ttgtcattt tgtttcttct cgtacgagct
2460

tgctcctgat cagcctatct cgcagcagat gaatatcttg tggtaggggt ttgggaaaat
2520

cattcgagtt tgatgttttt cttggtatth cccactcctc ttcagagtac agaagattaa
2580

gtgaaacctt cgtttgtgcg gacccctcag taatgtcttg tttcttttgt tgcagtgggtg
2640

agccattttg acttcgtgaa agtttcttta gaatagttgt ttccagaggc caaacattcc
2700

accgtagta aagtgaagc gtaggaagac caagactggc ataaatcagg tataagtgtc
2760

gagcactggc aggtgatctt ctgaaagttt ctactagcag ataagatcca gtagtcatgc
2820

atatggcaac aatgtaccgt gtggatctaa gaacgcgtcc tactaacctt cgcattcgtt
2880

ggtccagttt gttgttatcg atcaacgtga caaggttgtc gattccgcgt aagcatgcat
2940

acccaaggac gcctgttgca attccaagtg agccagttcc aacaatcttt gtaatattag
3000

agcacttcat tgtgttgccg ttgaaagtaa aatgcgaaca aattaagaga taatctcgaa
3060

accgcgactt caaacgccaa tatgatgtgc ggcacacaat aagcgttcat atccgctggg
3120

tgactttctc gctttaaaaa attatccgaa aaaattttct agagtgttgt tactttatac
3180

ttccggctcg tataatacga caagggtgaa ggaggactaa accatggcta aactcacctc
3240

tgctgttcca gtcctgactg ctcgtgatgt tgctgggtgct gttgagttct ggactgatag
3300

actcggtttc tcccgtgact tcgtagagga cgactttgcc ggtgttgtac gtgacgacgt
3360

taccctgttc atctccgcag ttcaggacca ggttgtgcca gacaacactc tggcatgggt
3420

atgggttcgt ggtctggacg aactgtacgc tgagtgggtct gaggtcgtgt ctaccaactt
3480

ccgtgatgca tctgggtccag ctatgaccga gatcggtgaa cagccctggg gtcgtgagtt
3540

tgcactgcgt gatccagctg gtaactgcgt gcatttcgtc gcagaagaac aggactaaca
3600

attgacacct tacgattatt tagagagtat ttattagttt tattgtatgt atacggatgt
3660

tttattatct atttatgccc ttatatctcg taactatcca aaagtcctat cttatcaagc
3720

cagcaatcta tgtccgcgaa cgtcaactaa aaataagctt tttatgctgt tctctctttt
3780

tttcccttcg gtataattat acctgcatc cacagattct cctgccaaat ttgcataat
3840

cctttacaac atggctatat gggagcactt agcggcctcc aaaaccata ttgcctacgc
3900

atgtataggt gttttttcca caatattttc tctgtgctct ctttttatta aagagaagct
3960

ctatatcgga gaagcttctg tggccgttat attcggcctt atcgtgggac cacattgcct
4020

gaattggttt gccccggaag attggggaaa ctgggatctg attaccttag ctgcaggtag
4080

cactgagcgt cagaccccggt agaaaagatc aaaggatctt ctgagatcc ttttttctg
4140

cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg
4200

gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca
4260

aatactgttc ttctagtgtt gccgtagtta ggccaccact tcaagaactc tgtagcacccg
4320

cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg
4380

tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga
4440

acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac
4500

ctacagcgtg agctatgaga aagcgccacg ctccccgaag ggagaaaggc ggacaggtag
4560

ccggttaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc
4620

tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgctg atttttgtga
4680

tgctcgtcag gggggcgag cctatggaaa aacgccagca acgcggcctt tttacggttc
4740

ctggcctttt gctggccttt tgctcacatg ttctttcctg cggtaccag atccaattcc
4800

cgctttgact gcctgaaatc tccatgcct acaatgatga catttgatt tggttgactc
4860

atgttggtat tgtgaaatag acgcagatcg ggaacactga aaaatacaca gttattattc
4920

atttaaat
4928

<210> 22

<211> 280

<212> PRT

<213> Tritirachium album

<400> 22

Ala Ala Gln Thr Asn Ala Pro Trp Gly Leu Ala Arg Ile Ser Ser Thr
 1 5 10 15

Ser Pro Gly Thr Ser Thr Tyr Tyr Tyr Asp Glu Ser Ala Gly Gln Gly
 20 25 30

Ser Cys Val Tyr Val Ile Asp Thr Gly Ile Glu Ala Ser His Pro Glu
 35 40 45

Phe Glu Gly Arg Ala Gln Met Val Lys Thr Tyr Tyr Tyr Ser Ser Arg
 50 55 60

Asp Gly Asn Gly His Gly Thr His Cys Ala Gly Thr Val Gly Ser Arg
 65 70 75 80

Thr Tyr Gly Val Ala Lys Lys Thr Gln Leu Phe Gly Val Lys Val Leu
 85 90 95

Asp Asp Asn Gly Ser Gly Gln Tyr Ser Thr Ile Ile Ala Gly Met Asp
 100 105 110

Phe Val Ala Ser Asp Lys Asn Asn Arg Asn Cys Pro Lys Gly Val Val
 115 120 125

Ala Ser Leu Ser Leu Gly Gly Gly Tyr Ser Ser Ser Val Asn Ser Ala
 130 135 140

Ala Ala Arg Leu Gln Ser Ser Gly Val Met Val Ala Val Ala Ala Gly
 145 150 155 160

Asn Asn Asn Ala Asp Ala Arg Asn Tyr Ser Pro Ala Ser Glu Pro Ser
 165 170 175

Val Cys Thr Val Gly Ala Ser Asp Arg Tyr Asp Arg Arg Ser Ser Phe
 180 185 190

Ser Asn Tyr Gly Ser Val Leu Asp Ile Phe Gly Pro Gly Thr Ser Ile
 195 200 205

Leu Ser Thr Trp Ile Gly Gly Ser Thr Arg Ser Ile Ser Gly Thr Ser
 210 215 220

Met Ala Thr Pro His Val Ala Gly Leu Ala Ala Tyr Leu Met Thr Leu
 225 230 235 240

Gly Lys Thr Thr Ala Ala Ser Ala Cys Arg Tyr Ile Ala Asp Thr Ala
 245 250 255

Asn Lys Gly Asp Leu Ser Asn Ile Pro Phe Gly Thr Val Asn Leu Leu
 260 265 270

Ala Tyr Asn Asn Tyr Gln Ala Gly
 275 280

Zastrzeżenia patentowe

1. Mutant zymogenu proteiny K z *Tritirachium album* zawierający na C-końcu mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzałej proteiny K w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu.
2. Mutant według zastrz. 1, **znamienny tym**, że posiada sekwencję przedstawioną jako sekwencja nr 12.
3. Mutant dojrzałej proteiny K z *Tritirachium album* zawierający na C-końcu mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzałej proteiny K w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu.
4. Mutant według zastrz. 3, **znamienny tym**, że posiada sekwencję przedstawioną jako sekwencja nr 22.
5. Sekwencja DNA zawierająca sekwencję kodującą mutantu zymogenu proteiny K jak określono w zastrz. 1 lub 2 albo jego fragment.
6. Sekwencja DNA według zastrz. 5 **znamienna tym**, że posiada sekwencję nr 8.
7. Sekwencja oligonukleotydu służącego do otrzymywania sekwencji DNA kodującej mutantu zymogenu proteiny K jak określono w zastrz. 1 lub 2 albo jego fragment wybrana spośród sekwencji nr 6 i 7.
8. Plazmid ekspresyjny zawierający sekwencję DNA kodującą mutantu zymogenu proteiny K jak określono w zastrz. 1 lub 2.
9. Plazmid według zastrz. 8, **znamienny tym**, że posiada sekwencję wybraną spośród sekwencji nr 16.
10. Rekombinantowy szczep drożdżowy transformowany plazmidem wybranym z grupy obejmującej plazmid jak określono w zastrz. od 8 do 9.
11. Szczep drożdży według zastrz. 10, **znamienny tym**, że został otrzymany z jednego ze znanych szczepów drożdży wybranych spośród *Pichia* sp., *Hansenula* sp., *Saccharomyces* sp., *Schizosaccharomyces* sp., *Yarrowia* sp., lub *Kluyveromyces* sp., korzystnie wybranego spośród szczepów *Pichia pastoris*.
12. Sposób otrzymywania białka o aktywności proteiny K *Tritirachium album*, **znamienny tym**, że prowadzi się hodowlę komórek drożdży transformowanych plazmidem ekspresyjnym kodującym mutantu zymogenu proteiny K z *Tritirachium album* zawierającego na C-końcu

- mutację zwiększającą wydajność ekspresji w komórkach drożdżowych oraz ułatwiającą dezaktywację dojrzałej proteiny K w warunkach stresowych, przy czym mutacją tą jest dodatkowa reszta Gly znajdująca się na C-końcu, a następnie uzyskiwane białko izoluje się i oczyszcza z płynu hodowlanego.
13. Sposób według zastrz. 12, **znamienny tym**, że prowadzi się hodowlę komórek drożdży jak określono w zastrz. 11, korzystnie w temperaturze około 28 do 30°C, korzystnie na pożywce zawierającej jako źródło węgla glicerol lub metanol, przy czym indukuje się ekspresję zmodyfikowanego genu proteiny K metanolem.
 14. Sposób według zastrz. 12, **znamienny tym**, że białko o aktywności proteiny K *Tritirachium album* izoluje się z płynu pohodowlanego metodą chromatografii oddziaływań hydrofobowych lub metodą chromatografii na złożu hydroksyapatytu, a następnie oczyszcza się prowadząc dializę wobec roztworu węglanu amonu, po czym suszy się ze stanu zamrożenia.

Rysunki

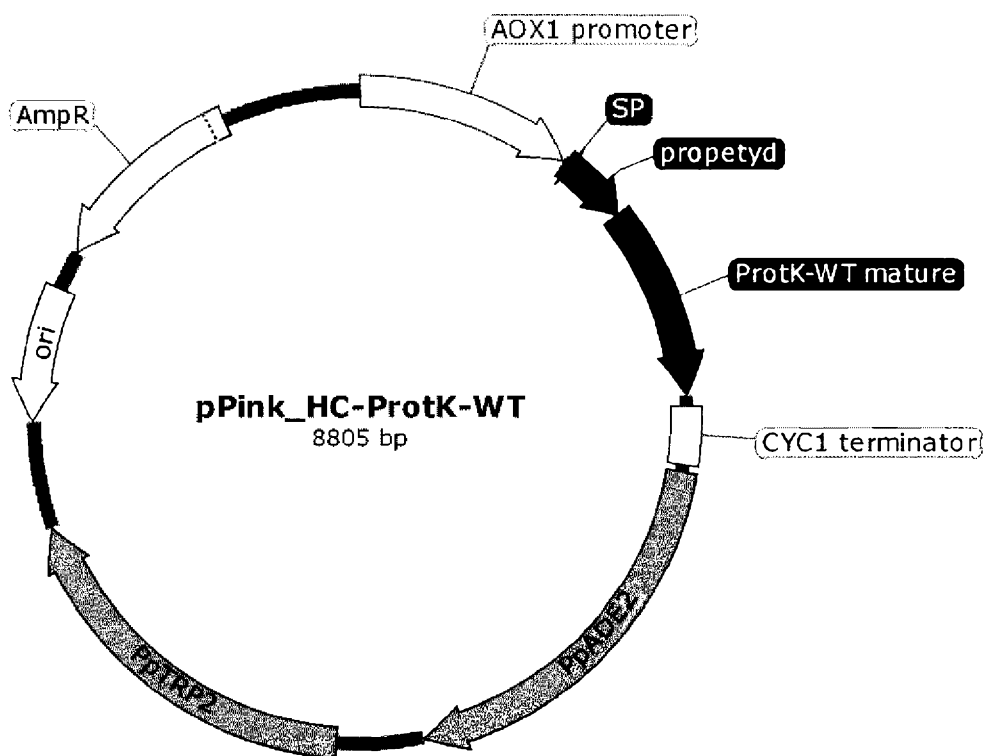


Fig.1

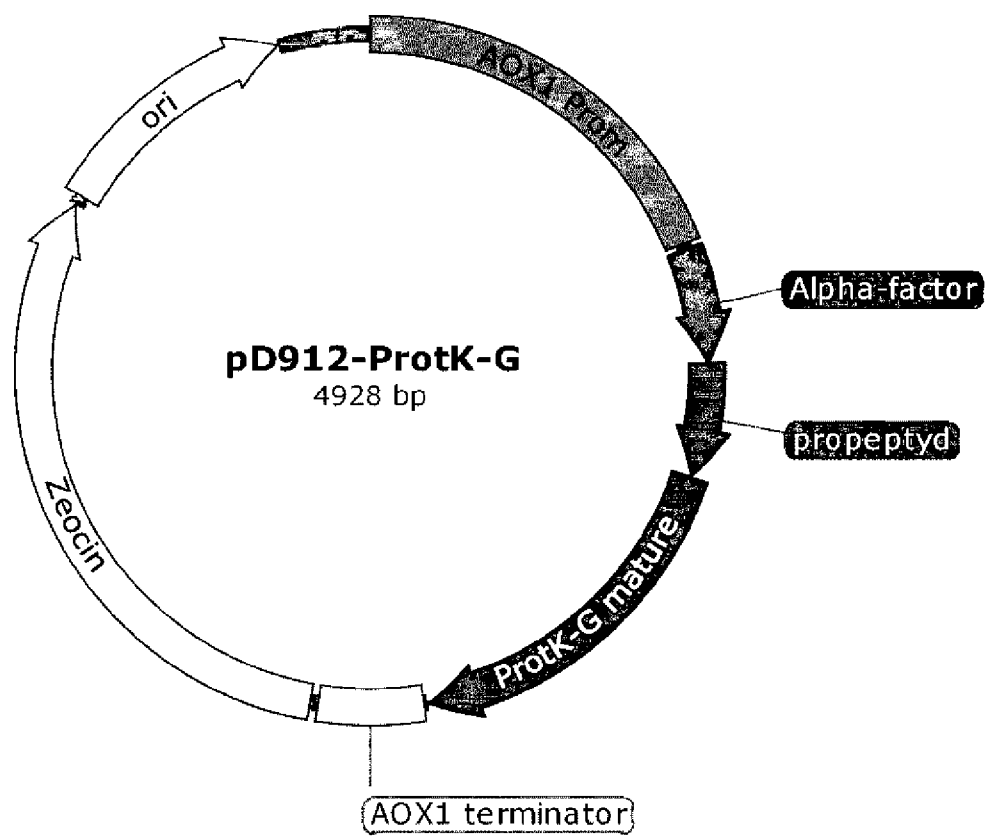
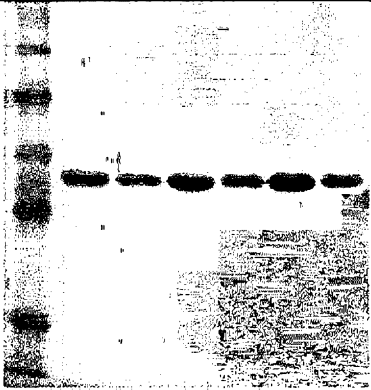


Fig.2

1	2	3	4	5	6	7	Opis próbek
							1. Marker białkowy: 170kDa, 130kDa, 100kDa, 70kDa, 55kDa, 40kDa, 35kDa, 25kDa, 15, 10kDa (Thermo Fisher nr. Kat. SM0671) 2. pD912-ProtK-G 48h OD= 294 3. pPink_HC-ProtK-WT 48h OD=296 4. pD912-ProtK-G 72h OD= 322 5. pPink_HC ProtK-WT 72h OD= 326 6. pD912-ProtK-G 96h OD= 500 7. pPink_HC ProtK-WT 96h OD= 492
8	9	10					
							
			</				

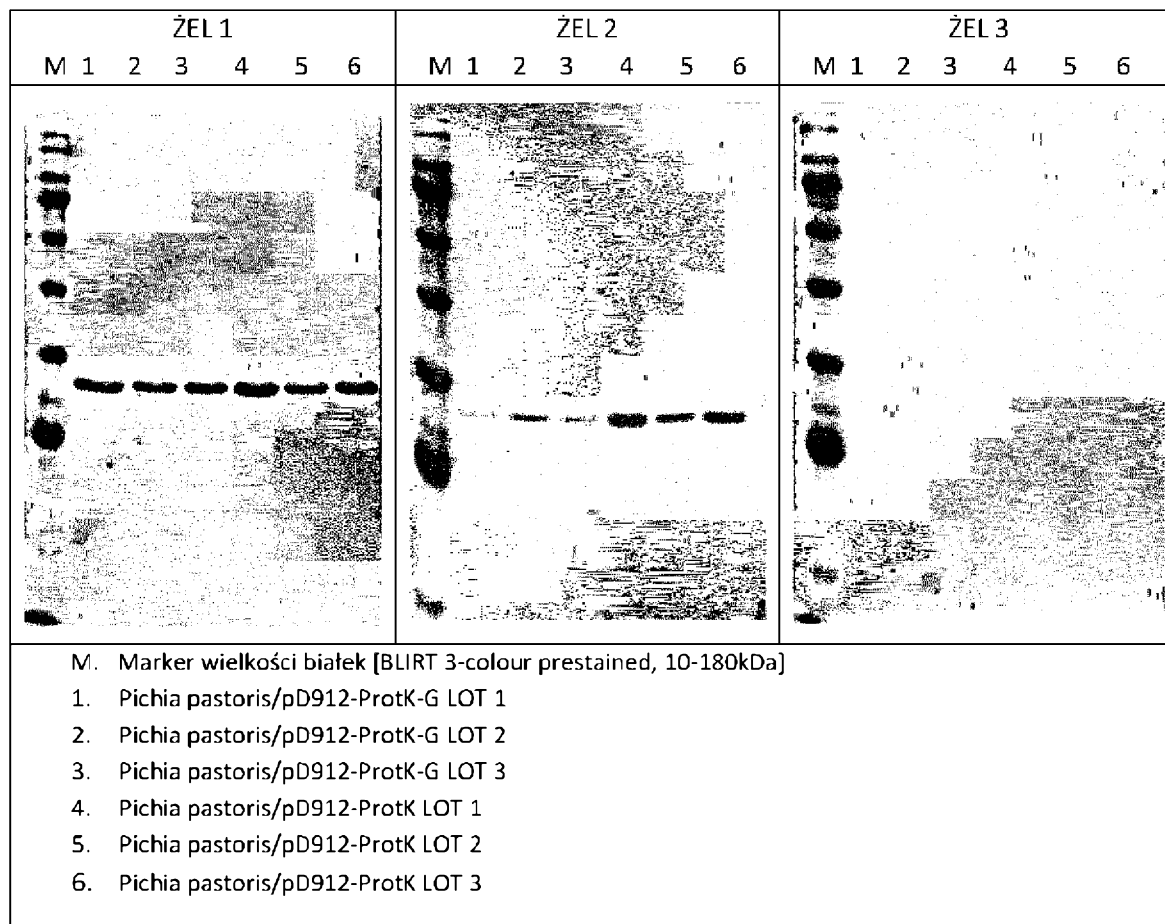


Fig. 4

	pD912-ProtK-G (linie: 2, 4, 6)	pD912-ProtK (linie: 8, 9, 10)	pPink_HC-ProtK-WT (linie: 3, 5, 7)
Ilość drożdży na koniec fermentacji	392,1 g/ L (SD 28)	362,5 g/l (SD 10,5)	437 g/l
Średnia końcowa ilość białka zmierzona metodą HPLC i/lub densytometrycznie	2,14 g/l (SD 0,01)	1,88 g/l (SD 0,01)	1,40 g/l
Ilość białka z procesu [wartość względna]	100%	89,4%	70%

Tabela 1