



Urząd Patentowy
Rzeczypospolitej Polskiej

(21) Numer zgłoszenia: **418712**

(22) Data zgłoszenia: **15.09.2016**

(51) Int.Cl.

C12N 7/00 (2006.01)

C07K 14/01 (2006.01)

C07K 14/005 (2006.01)

C12N 15/70 (2006.01)

C12Q 1/68 (2018.01)

(54) **Polinukleotyd kodujący białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA, wektor ekspresyjny, komórka gospodarza bakteryjnego oraz białko bakteriofaga TP-84 i jego zastosowanie**

(43) Zgłoszenie ogłoszono:

26.03.2018 BUP 07/18

(45) O udzieleniu patentu ogłoszono:

06.06.2022 WUP 23/22

(73) Uprawniony z patentu:

**BIOVENTURES INSTITUTE
SPÓŁKA Z OGRANICZONĄ
ODPOWIEDZIALNOŚCIĄ, Poznań, PL**

(72) Twórca(y) wynalazku:

**PIOTR SKOWRON, Gdańsk, PL
ANDREW KROPINSKI, Guelph, CA
AGNIESZKA ŻYLICZ-STACHULA, Gdańsk, PL
ŁUKASZ JANUS, Poznań, PL
KASJAN SZEMIAKO, Gdańsk, PL
NATALIA MACIEJEWSKA, Płońsk, PL
MAŁGORZATA SKOWRON, Gdynia, PL
MAŁGORZATA BUKREJEWSKA, Grajewo, PL
AGATA ZWARA, Kościerzyna, PL
JOANNA ŁOŚ, Gdańsk, PL
MARCIN PRZEMYSŁAW ŁOŚ, Gdańsk, PL**

(74) Pełnomocnik:

rzecz. pat. Rafał Witek

Opis wynalazku

Przedmiotem wynalazku jest kodujący białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA, wektor ekspresyjny, komórka gospodarza bakteryjnego oraz białko bakteriofaga TP-84 i jego zastosowanie.

Bakteriofag TP-84 infekuje niektóre gatunki termofilne bakterii *Bacillus* (nowa nomenklatura: *Geobacillus*). Polinukleotydy według wynalazku mogą być stosowane do modyfikacji genetycznych genomu TP-84, polegających w szczególności na klonowaniu, ekspresji lub modyfikacji genów TP-84, zwłaszcza mających zastosowanie w technologiach biologii i biotechnologii molekularnej, medycynie oraz przemyśle. Termofilne geny bakteriofagowe, kodujące białka kapsydu mogą znaleźć zastosowanie do konstrukcji systemów 'phage display', a geny kodujące helikazy lub białka wiążące DNA mogą służyć do usprawnienia reakcji powielania kwasów nukleinowych, konstrukcji szczepionek, adjuwantów, biosensorów, systemów wychwytu jonów metali lub terapii fagowej. Ponadto ujawniono sposób uzyskiwania klonów wspomnianych białek oraz wektory DNA do realizacji tego sposobu.

Saunders i Campbell, *Biochemistry* 4:2836–2844 (1965) ujawniają charakterystykę DNA bakteriofaga TP-84, infekującego bakterię *Bacillus stearothermophilus* szczep 10. DNA kodujące genom bakteriofaga ma długość 13,6 μ m, masę cząsteczkową 22,4–22,6 mln Da i zawiera 42% par GC.

Saunders i Campbell, *Journal of Bacteriology* 91:340–348 (1966) ujawniają odkrycie i charakterystykę termofilnego bakteriofaga TP-84, infekującego bakterię *Bacillus stearothermophilus* szczep 10. Ujawniono, że TP-84 jest bakteriofagiem litycznym dla 3 spośród 24 testowanych szczepów *Bacillus stearothermophilus* w zakresie temperatur 43°C do 76°C, wykazujących czas eklipsy 22 do 24 minut. Ujawniono, że kasyd TP-84 zawiera wielościenną (heksagonalną) główkę o wymiarach 53 × 30 nm oraz ogonek o długości 130 nm i 3–5 nm szerokości. Ujawniono, że całkowita masa cząsteczkowa TP-84 wynosi ok. 50 mln Daltonów oraz jego genom zawiera 42% par zasad GC.

Egbert i Mitchell, *Journal of Virology* 1:610–616 (1967) ujawnili odkrycie i charakterystykę termofilnego bakteriofaga T ϕ 3 infekującego *Bacillus stearothermophilus* ATCC8005 w temperaturze 60°C. Bakteriofag ten wykazuje termostabilność w zakresie temperatur do 75°C.

Bassel i współpracownicy *Journal of Virology* 7:663–672 (1971) ujawnili wrażliwość bakteriofaga TP-84 na dysocjację pod wpływem czynników chelatujących, w szczególności EDTA oraz fosforany, oraz powstałe w wyniku dysocjacji substruktury bakteriofagowe: główki, ogonki, wolne DNA i cienie bakteriofagowe.

Epstein i Campbell, *Applied Microbiology* 29:219–223 (1975) ujawniają sposób wydajnej produkcji i oczyszczanie termofilnego bakteriofaga TP-84. Ujawniono warunki hodowli gospodarza bakteryjnego oraz infekcji bakteriofagiem, pozwalające na uzyskanie miana bakteriofagów 5×10^{11} /ml, polegające na zastosowaniu wzbogaconego składu pożywki, zapobiegającego produkcji przetrwalników, pH 6,5, temperatury 58°C oraz dodatku chlorku magnezu i prowadzeniu procesu w fermentorze stosującym środowisko aerobowe. Ujawniono sposób oczyszczania bakteriofaga TP-84, polegający na precypitacji glikolem polietylenowym oraz ultrawirowaniem w gradiencie chlorku cezu.

Sharpy i współpracownicy, *Journal of General Microbiology* 132:1709–1722 (1986) ujawniają izolowanie i charakterystykę 24 bakteriofagów, infekujących obligatoryjnie termofilne szczepy bakterii *Bacillus*, izolowane z różnorodnych źródeł: m.in. kompostu, gleby, gnijącego siana. Ujawniono powszechność infekcyjności bakterii *Bacillus* przez bakteriofagi, o różnorodnych rozmiarach, morfologii i biochemii: od cylindrycznych do wielościennych z ogonkiem długości 15–500 nm, a ich genomowe DNA wykazywało różnice wielkości fragmentów podczas analizy restrykcyjnej. Ujawniono, że większość bakteriofagów była stabilna w temperaturze 50°C przez 4–5 godzin, natomiast w temperaturze 70°C w ciągu 2 godzin miana bakteriofagów obniżały się 10^2 – 10^7 -krotnie.

Blanc i współpracownicy, *Int J Syst Bacteriol.* 47:1246–1248 (1997) ujawniają metodę identyfikacji heterotroficznych, termofilnych, tworzących przetrwalniki bakterii, izolowanych z kompostu. Ujawniają sposób profilowania analizą restrykcyjną fragmentów DNA genomowego, kodującego 16S rDNA, powielonego metodą Polymerase Chain Reaction (PCR), pozwalający na określenie i różnicowanie gatunków i szczepów bakterii *Bacillus* sp. Ujawniono wyniki analiz powielenia i sekwencjonowania 1409 par zasad, kodujących 16S rDNA bakteriofaga TP-84 i zdeponowano je w GeneBank pod numerem dostępu AJ002154.

Golec i współpracownicy, *Journal of Microbiological Methods* 84:486–489 (2011) ujawniają metodę przechowywania bakteriofagów. Przedstawiony sposób jest uniwersalny i pozwala na długoterminowe przechowywanie bakteriofagów w postaci zainfekowanych komórek bakteryjnych, zamrozo-

nych w temperaturze -80°C . W tych warunkach utrata żywotności komórek bakteryjnych jest niewielka, a przeżywalność znajdujących się wewnątrz komórek bakteriofagów wysoka.

Lee i współpracownicy, *International Journal of Nanomedicine* 8:3917–3925 (2013) ujawniają w pracy przeglądowej rozwinięcia technologii '*phage display*' w kierunku konstrukcji biosensorów, opartych na zabiegach inżynierii genetycznej, przeprowadzonych na genomach bakteriofagów.

Pouyanfard i współpracownicy, *PLOS One* 7:e49539 (2012) ujawniają konstrukcję zmodyfikowanego genetycznie bakteriofaga T7, eksponującego w systemie '*phage display*' immunodominujące epitopy pochodzące ze szczurzej onkoproteiny HER2/neu. Ujawniono, że tak skonstruowana chimeryczna nanocząstka bakteriofaga T7 wykazuje wysoką immunogenność w modelowych liniach myszy BALB/c, wykazując działanie przeciwnowotworowe zarówno profilaktyczne jak i terapeutyczne, tym samym stanowiąc prototyp szczepionki przeciwnowotworowej.

Nadal problemem jest łatwe uzyskiwanie użytecznych praktycznie konstruktów genetycznych, opartych na genach termofilnych bakteriofagów, w szczególności bakteriofaga TP-84.

Celem wynalazku jest rozwiązanie tego problemu i zaproponowanie sposobu, który pozwalałby na łatwe uzyskiwanie tego typu konstruktów. W szczególności, celem wynalazku jest dostarczenie konstruktów typu '*phage display*', eksponujących na powierzchni kapsydu polipeptydy i białka naturalne oraz uzyskane w wyniku zabiegów inżynierii genetycznej, przykładowo – sztuczne białka poliepito-powe. Konstrukty takie mogą tworzyć nanocząstki biologiczne, stosowane jako szczepionki, adjuwanty, biosensory, układy wychwyty jonów metali, elementy terapii fagowej. Kolejnym celem wynalazku są klonowane i eksprymowane geny, kodujące białka użyteczne w biologii i biotechnologii molekularnej, medycynie i przemyśle, w szczególności helikazy i białka wiążące DNA, usprawniające reakcje powielania kwasów nukleinowych.

Istota wynalazku

Przedmiotem wynalazku jest polinukleotyd kodujący białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA posiadający sekwencję nukleotydową o numerze 2.

Przedmiotem wynalazku jest polinukleotyd kodujący gen rekombinantowy składający się z polinukleotydu kodującego białko bakteriofaga TP-84 określone powyżej oraz polinukleotydu kodującego sześć reszt histydyny, korzystnie posiadający sekwencję nukleotydową o numerze 4.

Przedmiotem wynalazku jest wektor ekspresyjny, znamieny tym, że zawiera origin replikacji, korzystnie pBR322 ori, gen oporności na antybiotyk, korzystnie ampicylinę, promotor transkrypcji, korzystnie BAD operonu arabinozowego, gen represora, korzystnie araC, sygnał inicjacji translacji, sekwencję kodującą sygnał stopu translacji oraz połączony z nimi operacyjnie polinukleotyd kodujący białko określony powyżej.

Korzystnie, wektor ekspresyjny według wynalazku posiada sekwencję o numerze 6.

Przedmiotem wynalazku jest komórka gospodarza bakteryjnego, zwłaszcza *Escherichia coli*, transformowana wektorem ekspresyjnym określonym powyżej.

Przedmiotem wynalazku jest białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA kodowane przez polinukleotyd określony powyżej.

Przedmiotem wynalazku jest białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA posiadające sekwencję aminokwasową wybraną spośród sekwencji o numerach: 3 lub 5.

Przedmiotem wynalazku jest zastosowanie białka bakteriofaga TP-84 posiadającego sekwencję aminokwasową wybraną spośród sekwencji o numerach: 3 lub 5 do przeprowadzenia *in vitro* manipulacji DNA obejmującej wiązanie dwuniciowego DNA i rozplatanie helisy DNA.

Szczegółowy opis wynalazku

Ujawniono polinukleotyd kodujący element funkcjonalny pochodzący z genomu bakteriofaga TP-84. Jako element funkcjonalny należy rozumieć polinukleotyd posiadający sekwencję DNA pochodzącą z genomu bakteriofaga TP-84, który może znaleźć praktyczne zastosowanie, przykładowo w wytwarzaniu termofilnych systemów '*phage display*', zmodyfikowanych nanocząstek biologicznych, szczepionek, adjuwantów, biosensorów, terapeutyków bakteriofagowych. W szczególności takim elementem funkcjonalnym może być polinukleotyd kodujący pełen genom bakteriofaga TP-84 lub polinukleotyd kodujący jej dowolny fragment funkcjonalny, w szczególności taki jak: promotor, terminator lub otwarta ramka odczytu kodująca białko. Fragmenty funkcjonalne zidentyfikowane w genomie bakteriofaga TP-84 zostały wskazane na Fig. 2.

Mogą one znaleźć szereg praktycznych zastosowań. W szczególności mogą być one użyteczne w produkcji termofilnych genów bakteriofagowych, zwłaszcza kodujących białka kapsydu, które nadają się do konstrukcji systemów 'phage display', a także innych białek fagowych, przykładowo helikaz lub białek wiążących DNA.

Korzystnie polinukleotyd według wynalazku charakteryzuje się tym, że koduje białko bakteriofaga TP-84, przy czym stanowi on polinukleotyd o sekwencji o numerze 2.

Ujawniono polinukleotyd kodujący gen rekombinantowy składający się z polinukleotydu kodującego białko bakteriofaga TP-84 zdefiniowanego powyżej oraz sekwencji kodującej sześć reszt histydy, korzystnie posiadający sekwencję wybraną spośród sekwencji o numerach 4 lub 9.

Białka fagowe mogą być wykorzystane między innymi do usprawnienia reakcji powielania kwasów nukleinowych, produkcji szczepionek, adjuwantów, biosensorów, systemów wychwytu jonów metali lub stosowania w terapii fagowej.

Przykładowe realizacje stanowią polinukleotyd posiadający sekwencję genu rekombinantowego kodującego białko TP84_31, o aktywności enzymatycznej helikazy DNA oraz sekwencja aminokwasowa białka TP84_31. W korzystnym wariantcie, ww. sekwencje są połączone z sekwencją nukleotydową kodującą znacznik sześciu reszt histydy, prowadzącą do biosyntezy fuzyjnego białka TP84_31-His₆, ułatwiającego izolowanie produkowanego białka metodą chromatografii metalopowinowactwa.

Ujawniono również wektor ekspresyjny charakteryzujący się tym, że zawiera *origin* replikacji, korzystnie pBR322 *ori*, gen oporności na antybiotyk, korzystnie ampicylinę, promotor transkrypcji, korzystnie BAD operonu arabinozowego, gen represora, korzystnie *araC*, sygnał inicjacji translacji, sekwencję kodującą sygnał stopu translacji oraz połączony z nimi operacyjnie polinukleotyd kodujący białko bakteriofaga TP-84 lub białko fuzyjne, który został zdefiniowany powyżej.

Korzystnie, wektor ekspresyjny posiada sekwencję o numerze 6.

Przykładowymi realizacjami tego aspektu są wektory DNA, zawierające sklonowane geny bakteriofaga TP-84 (TP84_31), kodujące helikazę DNA lub białko wiążące DNA (TP84-35) w układzie ekspresyjnym bakterii *Escherichia coli* nadające się do ekspresji i izolowania tych białek z bakterii.

Ujawniona została również komórka gospodarza bakteryjnego, zwłaszcza *Escherichia coli*, transformowana wektorem ekspresyjnym opisanym powyżej.

Przykładowy sposób uzyskiwania rekombinantowego, fuzyjnego białka TP84_31-His₆ (ssb) charakteryzuje się tym, że połączony z białkiem znacznik sześciu reszt histydy jest wykorzystany do izolowania białka metodą metalopowinowactwa.

Przykładowy sposób uzyskiwania rekombinantowego, fuzyjnego białka TP84_35-His₆ charakteryzuje się tym, że połączony z białkiem znacznik sześciu reszt histydy jest wykorzystany do izolowania białka metodą metalopowinowactwa.

Ujawniono również białko bakteriofaga TP-84 lub jego pochodną kodowane przez polinukleotyd opisany powyżej. Ujawnione białko bakteriofaga TP-84 lub jego pochodna posiadają sekwencję aminokwasową wybraną spośród sekwencji o numerach: 3 i 5.

Uzyskiwane zgodnie z wynalazkiem białka fagowe lub ich pochodne (np. białka fuzyjne z sekwencjami ułatwiającymi ich oczyszczanie), w szczególności białka TP84_31-His₆ oraz TP84_35-His₆ mogą być wykorzystane do usprawnienia metod biologii molekularnej, biotechnologii molekularnej, medycyny oraz przemysłu, korzystnie metod manipulacji DNA, w szczególności powielania kwasów nukleinowych.

Niniejszy opis został również wzbogacony o opisane poniżej przykłady wykonania, nie powinny być one jednak utożsamiane z pełnym zakresem wynalazku, który został omówiony powyżej oraz zdefiniowany w załączonych zastrzeżeniach patentowych.

P r z y k ł a d 1. Namnażanie termofilnego bakteriofaga TP-84 w bakteriach *Bacillus stearothermophilus* szczep 10.

W przykładowej realizacji namnożono bakterie *Bacillus stearothermophilus* szczep 10, korzystnie będący jednym z możliwych gospodarzy dla bakteriofaga TP-84. Do hodowli gospodarza *Bacillus stearothermophilus* szczep 10 oraz infekcji i namnażania bakteriofaga TP-84 stosowano pożywkę T, zawierającą na 1 L pożywki: 20 g Trypticase™, 3 g ekstraktu drożdżowego, 1,47 g CaCl₂·2H₂O, 7 mg FeCl₂·6H₂O, 15 mg MgSO₄·7H₂O, 1 mg MnCl₂·4H₂O, pH 7,4. Do przygotowania inoculum bakteryjnego pożywkę T zestalano poprzez dodanie i rozpuszczenie 2% agaru oraz przygotowanie płytek Petriego, zawierających 30 ml zestalanej pożywki T, inkubowane z wysianymi bakteriami w temperaturze

55°C przez 18–48 godzin. Pobierano pojedyncze kolonie *Bacillus stearothermophilus* szczep 10 w celu namnożenia w płynnej pożywce T. Kultury bakteryjne inkubowano w temperaturze 65°C, stosując intensywne napowietrzanie poprzez wytrząsanie w inkubatorze-wytrząsarce. Do przygotowania *inoculum* bakteriofaga TP-84 pożywkę T zestalano poprzez dodanie i rozpuszczenie 2% agaru oraz przygotowanie płytek Petriego, zawierających 30 ml zestalanej pożywki T. Następnie przygotowano 2,5 ml pożywki T, do której dodano i rozpuszczono 0,65% agaru, chłodzono do 60°C, dodano 0,3 ml zawiesiny bakterii *Bacillus stearothermophilus* szczep 10, znajdujących się w fazie logarytmicznego wzrostu oraz 0,1 ml zawiesiny bakteriofaga TP-84, inkubowano 10 min w temperaturze 55°C i wlewano jako warstwę górną na płytki Petriego, zawierające 30 ml zestalanej pożywki T. Po zestaleniu warstwy górnej płytki inkubowano w temperaturze 55°C przez 12–20 godzin. W celu hodowli bakteriofaga TP-84 w skali preparatywnej, przygotowano 5 kultur *Bacillus stearothermophilus* szczep 10, zaszczipiając porcje po 1 L pożywki T za pomocą 1 ml hodowli bakterii, znajdującej się w logarytmicznej fazie wzrostu. Hodowle intensywne napowietrzano poprzez wytrząsanie w inkubatorze-wytrząsarce w temperaturze 65°C, aż do osiągnięcia miana populacji bakteryjnej ok. 5×10^8 komórek w 1 ml. Bakterie infekowano lizatem bakteriofaga TP-84, stosując wielokrotność infekcji (MOI) 0,1 oraz dodano roztworu glukozy do stężenia 0,5%. Kontynuowano napowietrzanie przez 3–4 godziny, aż do osiągnięcia lizy komórek i przejaśnienia hodowli. Uzyskane lizaty bakteriofagowe typowo charakteryzowały się mianem bakteriofagów ok. 10^{11} PFU (jednostek formowania łysinek fagowych).

P r z y k ł a d 2. Oczyszczanie bakteriofaga TP-84 oraz izolowanie genomowego DNA TP-84.

W przykładowej realizacji namnożone bakteriofagi TP-84 jak wskazano w Przykładzie 1, poddano oczyszczeniu od pozostałości zlizowanych i całych komórek gospodarza *Bacillus stearothermophilus* szczep 10, stosując wieloetapową procedurę: (i) lizat bakteriofaga poddano różnicowemu wirowaniu w wirówce preparatywnej, rotorze kątowym stosując przyspieszenie 16 000 g przez 15 min. Supernatant, zawierający bakteriofagi TP-84 dekantowano znad debris bakteryjnego; (ii) celem usunięcia bakteryjnego DNA i RNA z lizatu, dodawano DNazę I do stężenia 10 ug/ml i RNazę A do stężenia 1 ug/ml, a następnie inkubowano przez 30 min w temperaturze 37°C; (iii) bakteriofagi precypitowano za pomocą $(\text{NH}_4)_2\text{SO}_4$, stopniowo dodanego w ilości 30 g/1 L lizatu bakteriofagowego oraz inkubowano przez 18 godzin w temperaturze 4°C. Precypitat bakteriofagowy odwirowywano w wirówce w rotorze kątowym, stosując przyspieszenie 12 000 g przez 15 minut w temperaturze 4°C, supernatant odrzucano, bakteriofagi TP-84 zawieszano w wodzie destylowanej, dializowano przez 18 godzin wobec wody destylowanej oraz odwirowywano w wirówce w rotorze kątowym, stosując przyspieszenie 16 000 g przez 10 minut, osad odrzucono; (iv) supernatant wirowano stosując przyspieszenie 27 000 g przez 2 godziny, supernatant odrzucono, osad zawieszono w buforze: 15 mM cytrynian trójsodowy pH 7,0; 0,15 M NaCl; (v) do zawiesiny bakteriofaga TP-84 dodawano CsCl do osiągnięcia gęstości roztworu 1,51 g/cm³ oraz ultrawirowano w rotorze SW-39 przy przyspieszeniu 30 000 g przez 24 godziny. Białe pasmo w gradiencie CsCl, zawierające stężoną zawiesinę bakteriofaga TP-84 wybierano poprzez igłę strzykawki, przebijając ściankę probówki. Zawiesinę bakteriofaga dializowano wobec roztworu zawierającego 10 mM Tris-HCl pH 8,0, 0,5 M NaCl, 10 mM MgCl₂ i przechowywano w 4°C lub po dodaniu glicerolu do 20% w temperaturze -80°C.

DNA bakteriofaga TP-84 izolowano po dodaniu EDTA do stężenia 50 mM za pomocą 3-krotnej ekstrakcji fenolem, nasyconym 20 mM buforem fosforanowym pH 7,0 i 2-krotnej ekstrakcji chloroformem. DNA precypitowano 2-ma objętościami alkoholu etylowego w obecności 0,3 M octanu sodu o pH 7,0, odwirowano, przemywano osad DNA 70% alkoholem etylowym, suszono i rozpuszczano w buforze TE o składzie: 10 mM Tris-HCl pH 8,0, 1 mM EDTA.

P r z y k ł a d 3. Sekwencjonowanie genomu i analiza bioinformatyczna genomu bakteriofaga TP-84.

W przykładowej realizacji, izolowane i oczyszczone DNA bakteriofaga TP-84 uzyskane jak wskazano w Przykładzie 2 poddano sekwencjonowaniu genomowego, korzystnie przy użyciu technologii *Next Generation Sequencing* (NGS) na Sekwenatorze Genomowym HiSeq 2000 firmy Illumina w trybie PE250. Wykonano bibliotekę genomową z krótkimi wstawkami 100–600 par zasad, otrzymane odczyty składano i analizowano programami: CutAdapt, CLC Genomic Workbench i PROKKA. Uzyskano 140683713 nukleotydów, 626934 sekwencji w parach, ostateczna złożona sekwencja stanowiła pojedynczy, liniowy segment dwuniciowego DNA o długości 47703 par zasad (Sek. 1). Uzyskaną sekwencję poddano dalszej analizie bioinformatycznej, korzystnie przy użyciu specjalistycznego oprogramowania: MyRAST, Kodon, BLASTP and SnapGene oraz korekcie wizualnej.

Fig. 1 przedstawia mapę całego genomu bakteriofaga TP-84 z uwidocznionymi Otwartymi Ramami Odczytu (ORF) z przypisanymi indywidualnymi nazwami genów, unikatowymi sekwencjami rozpoznawanymi przez endonukleazy restrykcyjne, rejonami promotorów inicjacji transkrypcji (oznaczonymi jako 'promoter'), hipotetycznymi rejonami regulatorowymi (oznaczonymi jako 'ODD', 'regulatory ODD', 'miscellaneous' i 'regulatory region') oraz rejonami regulacji inicjacji translacji – sekwencji wiązania rybosomów (oznaczonymi jako 'RBS'). Fig. 2 przedstawia pełną sekwencję dwuniciowego DNA, kodującego genom bakteriofaga TP-84 wraz z zaznaczonymi sekwencjami nukleotydowymi Otwartych Ram Odczytu (ORF) z przypisanymi indywidualnymi nazwami genów, promotorów transkrypcji, hipotetycznymi rejonami regulatorowymi oraz rejonami inicjacji translacji RBS. Fig. 3 przedstawia tabelę zawierającą listę Otwartych Ram Odczytu, kodujących hipotetyczne białka wraz z ich podstawowymi właściwościami, ustalonymi na podstawie analizy bioinformatycznej. Wyniki zaprezentowane na Fig. 1, 2, 3 oraz Sek. 1 ujawniają, że genom posiada 85 Otwartych Ram Odczytu (ORF), kodujących hipotetyczne białka bakteriofaga TP-84. Znamienne jest, że niemal wszystkie – 83 (Fig. 1, 2, 3), są skierowane w tym samym kierunku, zgodnym z kierunkiem arbitralnie ustalonej nici górnej DNA genomu, co prowadzi do wniosku, że również dominujący przebieg transkrypcji zachodzi w tym samym kierunku. Tym samym jedna z nici DNA genomu bakteriofaga jest silnie dominująca pod względem kodowania produktów mRNA i białek, jak wskazano w Fig. 1, 2, 3. W kierunku przeciwnym skierowane są jedynie 2 niewielkie Otwarte Ramy Odczytu (ORF): TP84_22 (204 pary zasad) i TP84_54 (264 pary zasad), potencjalnie kodujące polipeptydy o nieustalonej funkcji. Znamiennym również jest, że informacja genetyczna jest wysoce skondensowana, Otwarte Ramy Odczytu (ORF) są położone bardzo blisko siebie, a w szeregu wypadkach są wzajemnie zachodzącą. Właściwość ta służy typowo konserwacji kodującej przestrzeni w obrębie genomu, który ma ograniczoną wielkość ze względu na pojemność kapsydu bakteriofaga. Wśród wykrytych Otwartych Ram Odczytu (ORF), większość (54), koduje hipotetyczne białka o nieustalonej funkcji, wśród których mogą znajdować się białka o użytecznych właściwościach. Wśród kodowanych białek o funkcji przypisanej na podstawie homologii do sekwencji białek umieszczonych w publicznych bazach danych znajduje się m.in. (Fig. 3): 5 hipotetycznych białek membranowych (TP84_01, TP84_18, TP84_37, TP84_64, TP84_84) (Fig. 3); 12 hipotetycznych białek replikacji, rekombinacji, regulacji i metabolizmu DNA (TP84_24 – regulator transkrypcji, TP84_25 – białko RecB-podobne, TP84_26 – białko wspomagające hybrydyzację jednoniciowego DNA, TP84_30 – inhibitor replikacyjnej helikazy DNA, TP84_31 – replikacyjna helikaza DNA, TP84_33 – HNH endonukleaza typu 'homing', TP84_35 – białko wiążące jednoniciowe DNA (*single-stranded binding protein*, ssb), TP84_40 syntetaza tymidylanowa, TP84_41 – dUTP difosfataza, TP84_44 – RecU białko rekombinacyjne, TP84_56 – terminaza, mała podjednostka, TP84_57 – terminaza, duża podjednostka) (Fig. 3); 9 hipotetycznych białek zaangażowanych w budowę otoczki białkowej (kapsydu) bakteriofaga (TP84_61 – białko strukturalne kapsydu, TP84_63 – mniejsze białko kapsydu, TP84_65 – proteaza dojrzewania pro-głowy kapsydu, TP84_67 – główne białko kapsydu, TP84_72 – białko składania ogonka, TP84_76 – białko mierzące długość ogonka, TP84_77 – dystalne białko ogonka, TP84_78 – białko ogonka, TP84_80 – dystalne białko ogonka; 1 hipotetyczna glikozylaza zaangażowana w proces sporulacji (TP84_81); 2 hipotetyczne białka lizy ściany i błony komórkowej (TP84_82 – holina, TP84_83 – endolizyna). Ww. grupy białek mogą zostać rozszerzone o niektóre z pozostałych hipotetycznych białek o w tej chwili nieustalonej funkcji. Białka bakteriofaga TP-84 mogą znaleźć szereg zastosowań praktycznych, m.in. do: usprawnienia metod powielania kwasów nukleinowych, klonowania, diagnostyki molekularnej (korzystnie spośród hipotetycznych białek replikacji, rekombinacji, regulacji i metabolizmu DNA), tworzenia systemu '*phage display*', szczepionek, adjuwantów, prezentacji bio- i chemaktywnych motywów aminokwasowych (korzystnie spośród hipotetycznych białek zaangażowanych w budowę otoczki białkowej), dezintegracji komórek bakteryjnych (korzystnie spośród hipotetycznych białek lizy ściany i błony komórkowej).

P r z y k ł a d 4. Klonowanie molekularne genu TP84_31, kodującego helikazę DNA bakteriofaga TP-84.

W przykładowej realizacji wykorzystania sekwencji nukleotydowej genomu bakteriofaga TP-84 (Sekw. 1) jak wskazano w przykładach 1, 2 i 3 powielono w reakcji *Polymerase Chain Reaction* (PCR) korzystnie gen TP84_31 o długości 564 par zasad, kodujący replikacyjną helikazę DNA, składającą się z łańcucha polipeptydowego o długości 187 aminokwasów, o masie cząsteczkowej 21,7 kDa i teoretycznym punkcie izoelektrycznym 5,32. Fig. 4 przedstawia wyniki reakcji PCR, do której zaprojektowano i zastosowano startery reakcji PCR, korzystnie wprowadzające sekwencję nukleotydową,

kodującą znacznik sześciu reszt histydy do genu helikazy TP-84 oraz sekwencję rozpoznawaną przez endonukleazę restrykcyjną NcoI do klonowania do wektora ekspresyjnego pBAD-MycHisA: 5'-GGAGGACCATGGCTCACCATCATCATCATGATTTTGACGATAAAACACGTGAACG-3' (kodon Start ATG i znacznik sześciu reszt histydy podkreślony) oraz 5'-TTGAGATCTTCATGATGCTTCATCGAATCGGCG-3' (kodon Stop w nici dolnej DNA podkreślony). Sekwencję nukleotydową kodującą znacznik sześciu reszt histydy wprowadzono na koniec 5' genu, tworząc fuzję z N-końcem białka TP84_31, jakkolwiek może być ona wprowadzona również na koniec 3', tworząc fuzję z C-końcem białka. Sekw. 2 przedstawia sekwencję nukleotydową genu TP84_31, Sekw. 3 przedstawia sekwencję aminokwasową białka TP84_31. Korzystnie zastosowano następujące parametry reakcji PCR: stężenie starterów – 0,4 μ M, dNTPs – 0,8 mM, bufor reakcyjny 1 $\times$ (10 mM Tris-HCl pH 8,8 w 25°C, 50 mM KCl, 0,08% (v/v) Nonidet P40), MgCl₂ – 2 mM, 2 jednostki Polimerazy TaqNova (BLIRT S.A., Polska), DNA genomowe bakteriofaga TP-84 – 2 ng, początkowa denaturacja 94°C przez 180 sekund, 30 cykli termicznych: 94°C – 30 sekund, 55°C – 30 sekund, 72°C – 30 sekund; końcowe wydłużanie 72°C – 120 sekund. Zaprezentowany w Fig. 4 powielony i zmodyfikowany za pomocą użytych starterów gen TP84_31 o długości 585 par zasad, koduje replikacyjną helikazę DNA TP-84, opatrzoną znacznikiem 6-ciu reszt histydy oraz dodatkową resztą aminokwasu alaniny, znajdującą się za kodonem Start, składającą się z łańcucha polipeptydowego o długości 194 aminokwasów, o masie cząsteczkowej 22,6 kDa i teoretycznym punkcie izoelektrycznym 5,92. Sekw. 4 przedstawia zmodyfikowaną sekwencję nukleotydową genu TP84_31, Sekw. 5 przedstawia zmodyfikowaną sekwencję aminokwasową białka TP84_31. W uzyskanym produkcie PCR gen TP84_31 opatrzone jest flankującymi segmentami DNA, zawierającymi przeznaczone do klonowania, korzystnie do wektora pBAD-MycHisA, sekwencje rozpoznawane przez endonukleazy restrykcyjne NcoI i BglII. Tak uzyskany moduł genetyczny może być przeniesiony drogą klonowania do różnych klas wektorów. Na Fig. 4 przedstawiono kolejno: ścieżka M, wzorzec wielkości DNA – GeneRuler 1kb (Thermo Scientific, USA); ścieżka 1, reakcja PCR na matrycy genomowego DNA TP-84; ścieżka 2, kontrolna reakcja PCR, bez dodanej matrycy. Oczyszczony produkt PCR oraz DNA wektora pBAD-MycHisA zostały strawione endonukleazami restrykcyjnymi NcoI i BglII oraz poddane reakcji kierunkowej ligacji ligazą bakteriofaga T4 w taki sposób, że nastąpiła fuzja z zachowaniem ciągłości Otwartej Ramy Odczytu (ORF) opatrzonego znacznikiem 6-ciu reszt histydy genu TP84_31 z sygnałami inicjacji translacji wektora pBAD-MycHisA. Fig. 5 przedstawia wyniki uzyskane po transformacji mieszaniną ligacyjną bakterii *Escherichia coli* TOP10. Kolonie rekombinantowych bakterii poddano analizie PCR, stosując startery, zastosowane wcześniej do powielenia genu TP84_31 z genomu bakteriofaga TP-84, uzyskując w przypadku 3 klonów powielenie odcinka DNA o oczekiwanej długości genu TP84_31, opatrzonego sekwencjami flankującymi, wprowadzonymi przez startery. Na Fig. 5 przedstawiono kolejno: ścieżka M, wzorzec wielkości DNA – GeneRuler 1 kb (Thermo Scientific); ścieżki K1-K10, analizowane DNA poszczególnych klonów. Klon oznaczony K9, z przypisaną nazwą konstruktu pBADMycHisA-TP-84_helicase poddano sekwencjonowaniu DNA, potwierdzając pełną zgodność uzyskanej sekwencji nukleotydowej z sekwencją zmodyfikowanego genu TP84_31, wskazaną w Sekw. 4. Sekw. 6 przedstawia sekwencję nukleotydową konstruktu pBADMycHisA-TP-84_helicase.

Przykład 5. Ekspresja genetyczna klonu, kodującego białko helikazę DNA – TP84_31-His₆ bakteriofaga TP-84.

Klon pBADMycHisA-TP-84_helicase, zawierający gen TP84_31 wklonowany w wektorze pBAD-MycHisA, jak wskazano w Przykładach 3 i 4, pod kontrolą indukowalnego promotora arabinozowego BAD, którego poziom transkrypcji jest regulowany stężeniem cukru arabinozy, poddano eksperymentowi ekspresji (biosyntezy białka *in vivo*) w rekombinantowych komórkach bakterii *Escherichia coli* TOP10. Rekombinantowe bakterie *Escherichia coli* TOP10 [pBADMycHisA-TP-84_helicase] korzystnie hodowano w pożywce LB, zawierającą na 1 L pożywki: 10 g Tryptonu, 5 g ekstraktu drożdżowego, 10 g NaCl, pH 7,4, w temperaturze 37°C, stosując intensywne napowietrzanie w inkubatorze-wytrząsarce. Po osiągnięciu przez hodowlę bakterii gęstości optycznej OD = 0,45, wykonano indukcję chemiczną promotora BAD i inicjację biosyntezy rekombinantowego białka TP84_31-His₆ poprzez dodanie arabinozy do stężenia 0,02%.

Fig. 6 przedstawia mapę klonu plazmidowego pBADMycHisA-TP-84_helicase, zawierającego gen wskazany w Przykładach 3 i 4, kodujący białko TP84_35-His₆, wklonowane w wektorze pBAD-MycHisA, pod kontrolą indukowalnego promotora arabinozowego BAD. Na Fig. 7 przedstawiono analizę rezultatów ekspresji transformowanego DNA klonu pBADMycHisA-TP-84_helicase, zawie-

rającego fuzyjny gen, kodujący rekombinantowe białko TP84_31-His₆. Wykonana została analiza elektroforetyczna eksprymowanego białka w żelu poliakrylamidowym w warunkach denaturujących; żel barwiono Coomassie Blue. Pobierano próbki hodowli w odstępach czasowych w zakresie 0 do 16 godzin, odwirowywano bakterie i wykonywano elektroforetyczną analizę profilu białkowego z całych komórek. Na Fig. 7 przedstawiono kolejno: ścieżka 1, hodowla przed indukcją; ścieżka 2, hodowla 2 godziny po indukcji; ścieżka 3, 4 godziny po indukcji; ścieżka 4, 16 godzin po indukcji; ścieżka M, wzorce masy cząsteczkowej białek – LMW-SDS Marker Kit (GE Healthcare, USA). Ewidენტna jest masywna biosynteza rekombinantowego białka TP84_31-His₆, widoczna jako przyrost intensywności prążka białkowego na wysokości żelu odpowiadającej ok. 22 kDa, zgodnej z oczekiwaną wielkością ustaloną na podstawie sekwencji aminokwasowej TP84_31-His₆. Ze względu na to, że TP84_31-His₆ jest białkiem fuzyjnym, zawierającym znacznik sześciu reszt histydyny, do jego oczyszczania mają zastosowanie metody izolowania metodą metalochromatografii powinowactwa IMAC na immobilizowanych jonach metali przejściowych, korzystnie Ni²⁺.

P r z y k ł a d 6. Klonowanie molekularne genu TP84_35 bakteriofaga TP-84, kodującego białko wiążące jednoniciowe DNA (ssb).

W przykładowej realizacji wykorzystania sekwencji nukleotydowej genomu bakteriofaga TP-84 (Sekw. 1) powielono w reakcji PCR, korzystnie gen TP84_35 o długości 468 par zasad, kodujący białko wiążące jednoniciowe DNA jak wskazano w Przykładzie 3, składające się z łańcucha polipeptydowego o długości 155 aminokwasów, o masie cząsteczkowej 17,4 kDa i teoretycznym punkcie izoelektrycznym 5,16. Sekw. 7 przedstawia sekwencję nukleotydową genu TP84_35, Sekw. 8 przedstawia sekwencję aminokwasową białka TP84_35. Fig. 8 przedstawia wyniki reakcji PCR, do której zaprojektowano i zastosowano startery reakcji PCR, korzystnie wprowadzające sekwencję nukleotydową, kodującą znacznik sześciu reszt histydyny do genu białka wiążącego jednoniciowe DNA oraz sekwencję rozpoznawaną przez endonukleazę restrykcyjną NcoI do klonowania do wektora ekspresyjnego pBAD-MycHisA: 5'-GGAGGACCATGGCTCACCATCATCATCATCAACAATGTGACGTTAGTGGGAAGATTGACG-3' (kodon Start ATG i znacznik sześciu reszt histydyny podkreślony) oraz 5'-TTGAGATCTCTAAATGGCAGATCGTCATCATTCACATAAATCGG-3' (kodon Stop w nici dolnej DNA podkreślony). Sekwencję nukleotydową kodującą znacznik sześciu reszt histydyny wprowadzono na koniec 5' genu, tworząc fuzję z N-końcem białka TP84_35, jakkolwiek może być ona wprowadzona również na koniec 3', tworząc fuzję z C-końcem białka. Sekw. 9 przedstawia zmodyfikowaną sekwencję nukleotydową genu TP84_35, Sekw. 10 przedstawia zmodyfikowaną sekwencję aminokwasową białka TP84_35. Korzystnie zastosowano następujące parametry reakcji PCR: stężenie starterów – 0,4 μM, dNTPs – 0,8 mM, bufor reakcyjny 1×, MgCl₂ – 2 mM, 2 jednostki Polimerazy TaqNova (BLIRT S.A.) DNA genomowe bakteriofaga TP-84 – 2 ng, początkowa denaturacja 94°C przez 180 sekund, 30 cykli termicznych: 94°C – 30 sekund, 55°C – 30 sekund, 72°C – 30 sekund; końcowe wydłużanie 72°C – 120 sekund. Zaprezentowany w Fig. 8 powielony i zmodyfikowany za pomocą użytych starterów gen TP84_35 o długości 489 par zasad, koduje białko wiążące jednoniciowe DNA, opatrzone znacznikiem 6-ciu reszt histydyny oraz dodatkową resztą aminokwasu alaniny, znajdującą się za kodonem Start (Sekw. 9), składające się z łańcucha polipeptydowego o długości 162 aminokwasów (Sekw. 10), o masie cząsteczkowej 18,3 kDa i teoretycznym punkcie izoelektrycznym 6,28. Na Fig. 8 przedstawiono kolejno: ścieżka M, wzorzec wielkości DNA – GeneRuler 1 kb (Thermo Scientific); ścieżka 1, reakcja PCR na matrycy genomowego DNA TP-84; ścieżka 2, kontrolna reakcja PCR, bez dodanej matrycy. W uzyskanym produkcie PCR gen TP84_35 opatrzony jest flankującymi segmentami DNA, zawierającymi przeznaczone do klonowania, korzystnie do wektora pBAD-MycHisA, sekwencje rozpoznawane przez endonukleazy restrykcyjne NcoI i BglII. Tak uzyskany moduł genetyczny może być przeniesiony drogą klonowania do różnych klas wektorów. Produkt PCR oraz DNA wektora pBAD-MycHisA zostały strawione endonukleazami restrykcyjnymi NcoI i BglII oraz poddane reakcji kierunkowej ligacji ligazą bakteriofaga T4 w taki sposób, że nastąpiła fuzja z zachowaniem ciągłości Otwartej Ramy Odczytu (ORF) opatrzonego znacznikiem 6-ciu reszt histydyny genu TP84_35-His₆ z sygnałami inicjacji translacji wektora pBAD-MycHisA. Fig. 9 przedstawia wyniki uzyskane po transformacji mieszaniną ligacyjną bakterii *Escherichia coli* TOP10. Kolonie rekombinantowych bakterii poddano analizie PCR, stosując startery, zastosowane wcześniej do powielenia genu TP84_35 z genomu bakteriofaga TP-84, uzyskując w przypadkach 5-ciu klonów powielenie odcinka DNA o oczekiwanej długości genu TP84_35-His₆, opatrzonego sekwencjami flankującymi, wprowadzonymi przez startery. Na Fig. 9 przedstawiono kolejno: ścieżka M, wzorzec wielkości DNA – GeneRuler 1 kb

(Thermo Scientific); ścieżki K1-K9, analizowane DNA poszczególnych klonów. Klon oznaczony K7, z przypisaną nazwą konstruktu pBADMyHisA-TP-84_ssb poddano sekwencjonowaniu DNA, potwierdzając pełną zgodność uzyskanej sekwencji nukleotydowej z sekwencją zmodyfikowanego genu TP84_35-His₆, wskazaną w Sekw. 9. Sekw. 11 przedstawia sekwencję nukleotydową konstruktu pBADMyHisA-TP-84_ssb.

P r z y k ł a d 7. Ekspresja genetyczna klonu, kodującego białko TP84_35-His₆ bakteriofaga TP-84, wiążące jednoniciowe DNA (ssb).

Fig. 10 przedstawia mapę klonu plazmidowego pBADMyHisA-TP-84_ssb, zawierający gen wskazany w Przykładach 3 i 6, kodujący białko TP84_35-His₆, wklonowane w wektorze pBAD-MyHisA, pod kontrolą indukowalnego promotora arabinozowego BAD. Transformowane DNA klonu poddano eksperymentowi ekspresji w komórkach bakterii *Escherichia coli* TOP10. Rekombinantowe bakterie *Escherichia coli* TOP10 [pBADMyHisA-TP-84_ssb] korzystnie hodowano w pożywce LB, w temperaturze 37°C, stosując intensywne napowietrzanie w inkubatorze-wytrząsarce. Po osiągnięciu przez hodowlę bakterii gęstości optycznej OD = 0,40, wykonano indukcję chemiczną promotora BAD i inicjację biosyntezy rekombinantowego białka TP84_35-His₆ poprzez dodanie arabinozy do stężenia 0,02%.

Na Fig. 11 została przedstawiona analiza rezultatów ekspresji fuzyjnego genu, kodującego rekombinantowe białko TP84_35-His₆. Wykonana została analiza elektroforetyczna eksprymowanego białka w żelu poliakrylamidowym w warunkach denaturujących, który barwiono Coomassie Blue. Pobierano próbki hodowli w odstępach czasowych w zakresie 0 do 16 godzin, odwirowywano bakterie i wykonywano analizę profilu białkowego z całych komórek za pomocą elektroforezy denaturującej. Na Fig. 11 przedstawiono kolejno: ścieżka M, wzorce masy cząsteczkowej białek – LMW-SDS Marker Kit (GE Healthcare); ścieżka 1, hodowla przed indukcją; ścieżka 2, hodowla 2 godziny po indukcji; ścieżka 3, 4 godziny po indukcji; ścieżka 4, 16 godzin po indukcji. Ewidentna jest masywna biosynteza rekombinantowego białka TP84_35-His₆, widoczna jako przyrost intensywności prążka białkowego na wysokości żelu odpowiadającej ok. 17 kDa, zgodnej z oczekiwaną wielkością ustaloną na podstawie sekwencji aminokwasowej TP84_35-His₆. Ze względu na to, że TP84_35-His₆ jest białkiem fuzyjnym, zawierającym znacznik sześciu reszt histydyny, do jego oczyszczania mają zastosowanie metody izolowania metodą metalochromatografii powinowactwa IMAC na immobilizowanych jonach metali przejściowych, korzystnie Ni²⁺.

P r z y k ł a d 8. Izolowanie rekombinantowego białka TP84_35-His₆ bakteriofaga TP-84, wiążącego jednoniciowe DNA.

W przykładowej realizacji komórki rekombinantowych bakterii *Escherichia coli* TOP10 [pBAD-MyHisA-TP-84_ssb], wskazane w Przykładach 6 i 7, zawierające eksprymowane białko TP84_35-His₆, wiążące jednoniciowe DNA poddano procesowi separacji białka. Ekspresję genetyczną wykonano w 50 ml pożywki LB, komórki po trwającej 16 godzin indukcji promotora BAD odwirowano i zawieszono w 20 ml buforu do dezintegracji ultradźwiękowej: 50 mM Tris-HCl pH 8,0, 0,5 M NaCl, 0,01% Triton X-100, 0,5 mg/ml lizozymu jaja kurzego, 5 mM beta-merkaptioetanol, 1/5 tabletki SIGMAFAST™ Protease Inhibitor Cocktail, EDTA-Free. Zawiesinę, chłodzoną w łaźni lodowej, poddano sonikacji, a pozostałości komórek bakteryjnych odwirowano przy 19 500 g w temperaturze 4°C przez 30 minut, uzyskując 20 ml supernatantu. 5 ml uzyskanego supernatantu, zawierającego białko TP84_35-His₆ naniesiono na kolumnę, zawierającą 2 ml złoża do chromatografii metalopowinowactwa Ni Sepharose 6 Fast Flow (GE Healthcare). Kolumnę płukano 10-ma ml buforu A zawierającego: 50 mM Tris-HCl pH 8,0, 0,5 M NaCl, 5 mM imidazol, a następnie 10-ma ml buforu A zawierającego: 50 mM Tris-HCl pH 8,0, 0,5 M NaCl, 40 mM imidazol. Związane do kolumny białko TP84_35-His₆ eluowano 10-ma ml buforu E zawierającego: 50 mM Tris-HCl pH 8,0, 0,5 M NaCl, 300 mM imidazol. Frakcje zawierające wyeluowane białko TP84_35-His₆ dializowano wobec buforu S zawierającego: 20 mM Tris-HCl pH 8,0, 150 mM NaCl, 1 mM EDTA, 0,2 mM beta-merkaptioetanol, 0,01% Triton X100, 50% glicerol i przechowywano w temperaturze -20°C. Fig. 12 przedstawia analizę elektroforetyczną w warunkach denaturujących kolejnych etapów izolowania rekombinantowego białka TP84_35-His₆: ścieżki M, wzorce masy cząsteczkowej białek – LMW-SDS Marker Kit (GE Healthcare); ścieżka 2, przesącz z kolumny, zawierający niezwiązane białka; ścieżki 2–5, frakcje po płukaniu buforem A; ścieżki 6–9, frakcje po płukaniu buforem B; ścieżki 10–18, frakcje zawierające wyeluowane buforem E rekombinantowe białko TP84_35-His₆.

P r z y k ł a d 9. Funkcjonalna aktywność biologiczna rekombinantowego białka helikazy DNA – TP84_34-His₆ bakteriofaga TP-84.

W przykładowej realizacji komórki rekombinantowych bakterii *Escherichia coli* TOP10 [pBAD-MycHisA-TP-84_helicase], wskazane w Przykładach 3 i 4, 5, zastosowano do otrzymania oczyszczonego, rekombinantowego białka TP84_35-His₆, któremu na podstawie analizy bioinformatycznej wskazanej w Przykładzie 3, przypisano funkcję enzymu helikazy, rozplatającego dwuniciowe DNA. Idea testu funkcjonalnego polega na wykazaniu zdolności termostabilnego białka TP84_34-His₆ do rozplatania zhybrydyzowanych nici DNA, monitorowane zaawansowaną metodą pomiaru zmiany poziomu fluorescencji barwnika interkalującego jedynie do dwuniciowego DNA (Eva Green, Jena Bioscience, Niemcy), mierzoną w urządzeniu metodą real time PCR (model MyGoPRO, IT-IS Life Science Ltd., Irland). Spadek poziomu fluorescencji po dodaniu rekombinantowego białka TP84_35-His₆ jest wyznacznikiem jego aktywności biologicznej.

Fig. 13 przedstawia wyniki analizy kinetyki zmian fluorescencji w urządzeniu real time PCR. Zaprojektowano dwa wzajemnie komplementarne fragmenty jednoniciowego DNA, które zostały zsyntetyzowane chemicznie o sekwencjach: 5'-GCTAGTCGTCGCGTAAGCCAATCGACCCCAAAATAGCGA GCATCAGTGTAGAGAGTTGGA-3' oraz 5'-TCCAACCTCTCTACACTGATGCTCGCTATTTTGGGGTC GATTGGCTTACGCGACGACTAGC-3'. Korzystnie zastosowano następujące parametry reakcji monitorowania zmian fluorescencji, przeprowadzonych w objętościach po 20 µl: stężenie jednoniciowych odcinków DNA – 1,5 µM, dNTPs – 0,1 mM, bufor reakcyjny 1×: 20 mM Tris-HCl pH 7,5, MgCl₂ – 10 mM, BSA – 1 mg/ml, ATP – 2 mM. Zastosowano schemat cykli termicznych: 37°C – 1 minuta, 55°C – 30 minut. Komponenty reakcji były dodawane do mieszaniny reakcyjnej w temperaturze 0°C a następnie umieszczone w urządzeniu do real time PCR i niezwłocznie poddane cyklom termicznym. Na Fig. 13 przedstawiono wykres obrazujący kinetykę hybrydyzacji jednoniciowych odcinków DNA, gdzie w próbie kontrolnej bez dodanego rekombinantowego białka TP84_34-His₆ widoczny jest stopniowy wzrost stężenia form odcinków DNA zhybrydyzowanych do dwuniciowych odcinków, tym samym zwiększając poziom fluorescencji w czasie cykli hybrydyzacyjnych. W próbach zawierających dodane rekombinantowe białko TP84_34-His₆ widoczny jest znacznie niższy poziom wzrostu fluorescencji lub jej spadek, w zależności od ilości dodanego rekombinantowego białka TP84_34-His₆. Krzywe na wykresie w Fig. 13 oznaczono kolorami: purpurowym – reakcja kontrolna bez dodanego rekombinantowego białka TP84_34-His₆; czarnym – reakcja zawierająca 1,65 µg rekombinantowego białka TP84_34-His₆; czerwonym – reakcja zawierająca 1,24 µg rekombinantowego białka TP84_34-His₆; zielonym – reakcja zawierająca 0,825 µg rekombinantowego białka TP84_34-His₆; niebieskim – reakcja zawierająca 0,412 µg rekombinantowego białka TP84_34-His₆. Obserwowana jest negatywna zależność kinetyki hybrydyzacji jednoniciowego DNA, od ilości dodanego rekombinantowego białka TP84_34-His₆. Tym samym wykazano, że rekombinantowe białko TP84_34-His₆ wykazuje aktywność biologiczną helikazy oddziałującej z DNA, wywierając efekt rozplatania helisy dwuniciowego DNA. Właściwość ta może mieć istotne zastosowania praktyczne w technologiach manipulacji DNA, w szczególności w usprawnieniu reakcji powielania DNA.

P r z y k ł a d 10. Funkcjonalna aktywność biologiczna rekombinantowego białka TP84_35-His₆ bakteriofaga TP-84, wiążącego jednoniciowe DNA.

W przykładowej realizacji komórki rekombinantowych bakterii *Escherichia coli* TOP10 [pBAD-MycHisA-TP-84_ssb], wskazane w Przykładach 6, 7, 8 zastosowano do otrzymania oczyszczonego rekombinantowego białka TP84_35-His₆, wiążącego jednoniciowe DNA. Idea testu funkcjonalnego polega na wykazaniu zdolności termostabilnego białka TP84_35-His₆ do usprawniania reakcji PCR poprzez wiązanie do jednoniciowych starterów i tym samym stymulowanie specyficzności hybrydyzacji do matrycowego DNA, eliminując lub zmniejszając artefakty reakcji PCR, obserwowane podczas technik analitycznych reakcji PCR, takie jak pojawianie się obserwowanych na elektroforetycznych żelach agarozowych lub akrylamidowych dodatkowych prążków DNA, 'smugi' DNA, dimerów starterów. Artefakty te są efektem hybrydyzacji starterów do nie w pełni homologicznych sekwencji nukleotydowych w matrycy oraz wzajemnej hybrydyzacji starterów.

Fig. 14 przedstawia wyniki reakcji PCR, do której zaprojektowano i zastosowano startery reakcji PCR: 5'-ATTGGAGGGCAAGTCTGGTG-3' oraz 5'-CCGATCCCTAGTCGGCATAG-3', powielające testowy odcinek 470 par zasad z genomu *Candida albicans*, gdzie reakcja powielania nie skutkuje powstaniem wyłącznie oczekiwanego jednorodnego, pojedynczego prążka DNA. Tym samym zawiera liczne artefakty, widoczne na agarozowym żelu elektroforetycznym w postaci prążków dodatkowych

oraz 'smugi' DNA. Korzystnie zastosowano następujące parametry reakcji PCR przeprowadzonych w objętości 25 μ l: stężenie starterów – 0,4 μ M, dNTPs – 0,1 mM, bufor reakcyjny 1 $\times$, MgCl₂ – 3 mM, 2 jednostki Polimerazy TaqStoffel, DNA genomowe *Candida albicans* – 23 ng, początkowa denaturacja 95°C przez 180 sekund, 30 cykli termicznych: 95°C – 30 sekund, 45°C – 30 sekund, 72°C – 30 sekund; końcowe wydłużanie 72°C – 180 sekund. Czerwona strzałka oznacza oczekiwany produkt PCR o długości 470 par zasad, przedstawiony na Sekw. 12. Na Fig. 14 przedstawiono kolejno: ścieżki M, marker wielkości fragmentów DNA 100 bp Plus (Fermentas, Litwa); ścieżka 1, kontrolna reakcja PCR na matrycy genomowego *Candida albicans* bez dodanego białka TP84_35-His; ścieżka 2, reakcja PCR, przeprowadzono jak w ścieżce 1 z dodatkiem 0,36 ng białka TP84_35-His; ścieżka 3, PCR z 0,72 ng białka TP84_35-His; ścieżka 4, PCR z 1,08 ng białka TP84_35-His; ścieżka 5, PCR z 1,44 ng białka TP84_35-His; ścieżka 6, PCR z 2,16 ng białka TP84_35-His; ścieżka 7, PCR z 2,88 ng białka TP84_35-His; ścieżka 8, PCR z 3,60 ng białka TP84_35-His. Komponenty reakcji były dodawane do mieszaniny reakcyjnej w temperaturze 0°C, a następnie umieszczone w termocyklerze PCR (model TPROFESSIONAL BASIC, Biometra, Niemcy) i niezwłocznie poddane cyklowi termicznemu. Wyniki wykazują stopniowe zmniejszanie się ilości artefaktów w ścieżkach 2–7, aż do całkowitej ich eliminacji w ścieżce 8. Tym samym wykazano, że rekombinantowe białko TP84_35-His wykazuje aktywność biologiczną oddziaływania z DNA, w przeprowadzonym teście objawiającą się dodatkowo jako usprawnienie reakcji PCR. Właściwość ta może mieć istotne zastosowania praktyczne w technologiach manipulacji DNA, w szczególności w usprawnieniu reakcji powielania DNA.

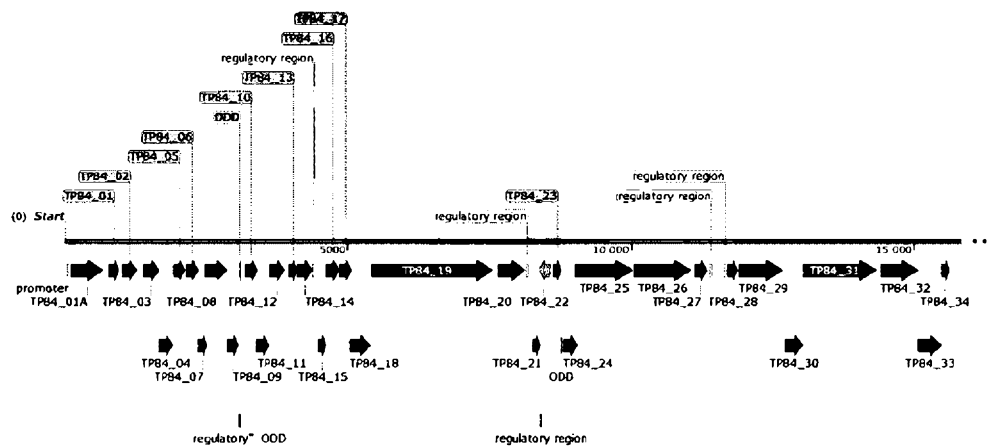
Zastrzeżenia patentowe

1. Polinukleotyd kodujący białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA posiadający sekwencję nukleotydową o numerze 2.
2. Polinukleotyd kodujący gen rekombinantowy składający się z polinukleotydu kodującego białko bakteriofaga TP-84 określone w zastrz. 1 oraz polinukleotydu kodującego sześć reszt histydyny, korzystnie posiadający sekwencję nukleotydową o numerze 4.
3. Wektor ekspresyjny, **znamienny tym**, że zawiera *origin* replikacji, korzystnie pBR322 *ori*, gen oporności na antybiotyk, korzystnie ampicylinę, promotor transkrypcji, korzystnie BAD operonu arabinozowego, gen represora, korzystnie *araC*, sygnał inicjacji translacji, sekwencję kodującą sygnał stopu translacji oraz połączony z nimi operacyjnie polinukleotyd kodujący białko określony w zastrz. 1 albo 2.
4. Wektor ekspresyjny według zastrz. 3, **znamienny tym**, że posiada sekwencję o numerze 6.
5. Komórka gospodarza bakteryjnego, zwłaszcza *Escherichia coli*, transformowana wektorem ekspresyjnym określonym w zastrz. 3 lub 4.
6. Białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA kodowane przez polinukleotyd określony w zastrz. 1 albo 2.
7. Białko bakteriofaga TP-84 wiążące dwuniciowy DNA i wykazujące aktywność rozplatania helisy DNA posiadające sekwencję aminokwasową wybraną spośród sekwencji o numerach: 3 lub 5.
8. Zastosowanie białka bakteriofaga TP-84 posiadającego sekwencję aminokwasową wybraną spośród sekwencji o numerach: 3 lub 5 do przeprowadzenia *in vitro* manipulacji DNA obejmującej wiązanie dwuniciowego DNA i rozplatanie helisy DNA.

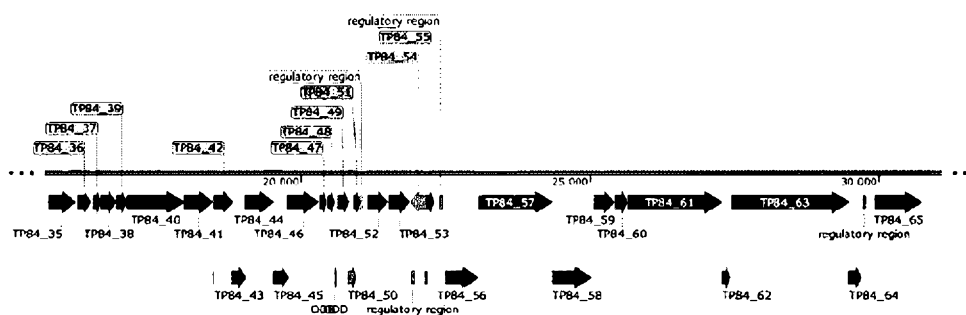
Rysunki

Fig. 1

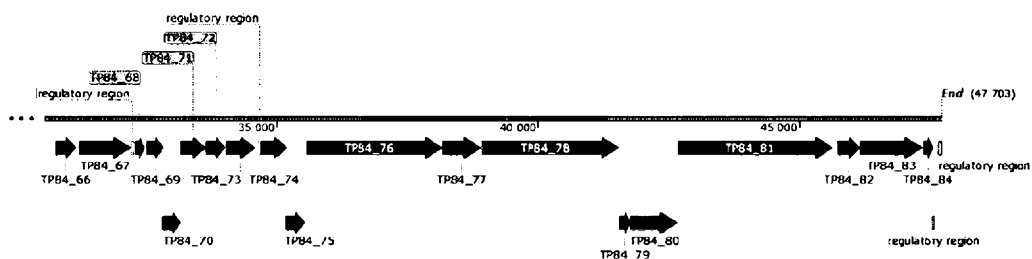
A



B



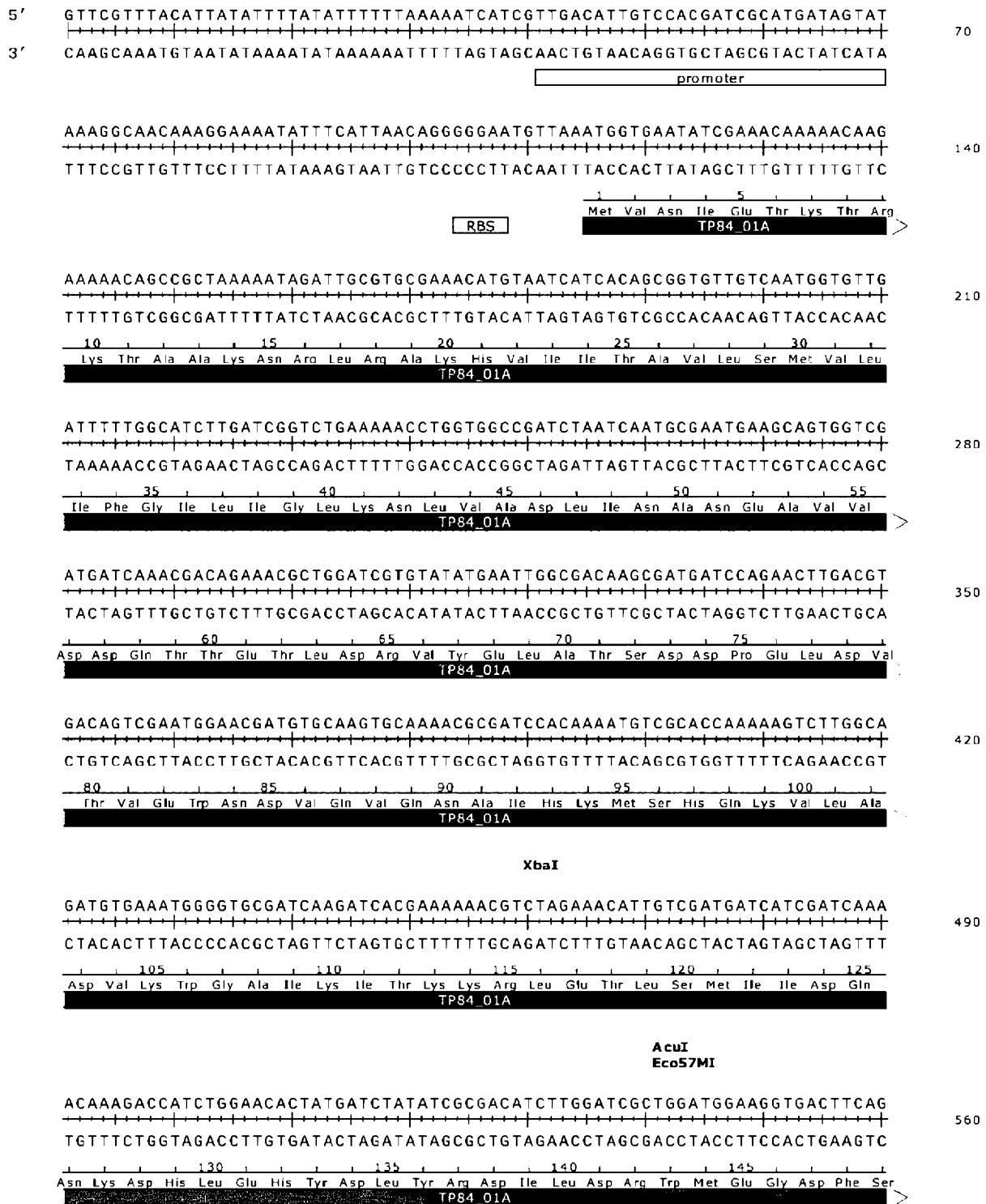
C



Bacillus (Geobacillus) bacteriophage TP-84
47 703 bp

Fig. 2

Start (0)



TCGCGCAGACGAAGATCACAAACGCGATCTGGGAACCTGCAAGGCGGCACGATCGGGAAAGCCTATGGATTA
 630
 AGCGCGTCTGCTTCTAGTGTGCGCTAGACCCTTGACGTTCCGCCGTGCTAGCCCTTTCGGATACCTAAT
 150 155 160 165 170
 Arg Ala Asp Glu Asp His Asn Ala Ile Trp Glu Leu Gln Gly Gly Thr Ile Gly Lys Ala Tyr Gly Leu
 TP84_01A
 CAAACCAAAGAAGAACAAGAATTCATCAAAAAACACATTGCGGTGATGGCGAATAAAAAAATTCGCCAT
 700
 GTTGGTTTCTTCTTCTTGTCTTAAGTAGTTTTTGTGTAAGCCACTACCGCTTATTTTTTAAGCGGTA
 175 180 185 190
 Gln Thr Lys Glu Glu Glu Gln Glu Phe Ile Lys Asn Thr Phe Gly Asp Gly Glu
 TP84_01A
 TTTTCTGTGGACTTGTGACATCGTCAACAAAAAATGGTACATTATAAGTGTAAGGCAATCACGAAAAT
 770
 AAAAAGACACCTGAACAACTGTAGCAGTTGTTTTTACCATGTAATATTCACATTCGGTTAGTGCTTTTA
 RBS
 TCAAGGTGAAGCGATATGAAAAATGGGATCAAGTGGTGATGAAAAAGGCGGAAGCCAACAACTGGGGAT
 840
 AGTTCCACTTCGCTATACTTTTTACCCTAGTTCACCACTACTTTTTCCGCCCTCGGTTGTTTGACCCCTA
 1 5 10 15 20
 Met Lys Arg Tyr Glu Lys Trp Asp Gln Val Val Met Lys Lys Ala Glu Ala Asn Lys Leu Gly Ile
 TP84_01
 CTTTCCAGAACTGCAAGTGATCGAGGATTCGGGCGAATTCCTCACAGTGACGCGACAACGAAAGATTTT
 910
 GAAAGGTCTTGACGTTCACTAGCTCCTAAGCCCGCTTAAGGAGTGTCACATGCCTGTTGCTTTCTAAAA
 25 30 35 40 45
 Phe Pro Glu Leu Gln Val Ile Glu Asp Ser Gly Glu Phe Leu Thr Val Tyr Ala Thr Thr Lys Asp Phe
 TP84_01
 AATACGAAAATGATCATGATCGAACGCGATCAACTGCTTAAAAAGTACCAATAAGATCCGGGCCGCTGAT
 980
 TTATGCTTTTACTAGTACTAGCTTGCGCTAGTTGACGAATTTTTCATGGTTATTCTAGGCCCGGCGACTA
 50 55 60
 Asn Thr Lys Met Ile Met Ile Glu Arg Asp Gln Leu Leu Lys Lys Tyr Gln
 TP84_01
 GGCGGCCCTGGCACAATAAATCAAACAGGGGGAATGAACGATGGAATCAAAACACTGCCGATCAAGTGG
 1050
 CCGCCGGACCGTGTTTATTTAGTTTGCCCCCTTACTTGCTACCTTTAGTTTTGTGACGGCTAGTTCCAC
 1 5 10
 Met Glu Ile Lys Thr Leu Pro Ile Lys Trp
 RBS TP84_02
 TCTGTGCCGACTAATTACATCACAGGGGTTGTGATCGAAACAAGAACGATGATTTATGTGATCCGCAAAA
 1120
 AGACACGGCTGATTAATGTAGTGTCCCCAACACTAGCTTTGTTCTTGCTACTAAATACACTAGGCGTTTT
 15 20 25 30
 Ser Val Pro Thr Asn Tyr Ile Thr Gly Val Val Ile Glu Thr Arg Thr Met Ile Tyr Val Ile Arg Lys
 TP84_02
 ACGGATCTGATGAATTCAGGTCGATCGAATGGATAAAAAACCACAAATGGAATCGATACTGGACAATTGA
 1190
 TGCCTAGACTACTTAAGGTCCAGCTAGCTTACCTATTTTTGGTGTTTACCTTAGCTATGACCTGTTAACT
 35 40 45 50 55
 Asn Gly Ser Asp Glu Phe Gln Val Asp Arg Met Asp Lys Asn His Lys Trp Asn Arg Tyr Trp Thr Ile Asp
 TP84_02

TCGATTCACGATGAAAGGATCAGCCGGCTTGCCGAAAGCGGTGATCGAAGCGATCCAAACGATCAAAGAA
AGCTAAGTGCTACTTTCCTAGTCGGCCGAACGGCTTTCGCCACTAGCTTCGCTAGGTTTGCTAGTTTCTT

1260

60 65 70 75 80
Arg Phe Thr Met Lys Gly Ser Ala Gly Leu Pro Lys Ala Val Ile Glu Ala Ile Gln Thr Ile Lys Glu
TP84_02

AAAGGCATCCTGTGAAAAGCCGAAACGGTCTGGGGATTCTGGACCGTCTGGCCAGATGGGATCTGGTCA
TTTCCGTAGGACACTTTTCGGCTTTGCCAGACCCCTAAGGACCTGGCAGACCGGTCTACCCTAGACCAGT

1330

85
Lys Gly Ile Leu
TP84_02

CTGATGATGGCAACCATCACAGGCCAAAAAATCAAACAGGGGGAATGAAGGATGTTCAAAACTGAAAAA
GACTACTACCGTTGGTAGTGTCGGTTTTTTTAGTTTGTCCCCCTTACTTCTACAAGTTTTGACTTTTT

1400

RBS

1 5
Met Phe Lys Thr Glu Lys
TP84_03

CAATCAATGAAGCGAAAGCCGCATTGAAAAAGGTGTGATCGTCTATTACTTCAATGAAGCGGCGATGA
GTTTAGTTACTTCGCTTTCGGCGTAACTTTTTCCACACTAGCAGATAATGAAGTTACTTCGCCGCTACT

1470

10 15 20 25
Gln Ile Asn Glu Ala Lys Ala Ala Leu Lys Lys Gly Val Ile Val Tyr Tyr Phe Asn Glu Ala Ala Met
TP84_03

ATGCACACATCGGGATCGTGGTCGATACTGATCCAGAATATTATTCGGCGAAGGCGTATCAGTCATGCT
TACGTGTGTAGCCCTAGCACCAGCTATGACTAGGTCTTATAATAAGGCCGCTTCCGCATAGTCAGTACGA

1540

30 35 40 45 50
Asn Ala His Ile Gly Ile Val Val Asp Thr Asp Pro Glu Tyr Tyr Ser Gly Glu Gly Val Ser Val Met Leu
TP84_03

AGTCGATCGAAACTTGAAACCAAAACAAGATGTGTTTCGGCCGGATCATATTGACATCGTGTTTTACCAC
TCAGCTAGCTTTGAACTTTGGTTTTGTTCTACACAAGCCGGCCTAGTATAAGCTGTAGCACAAAATGGTG

1610

55 60 65 70 75
Val Asp Arg Asn Leu Lys Pro Lys Gln Asp Val Phe Gly Arg Ile Ile Phe Asp Ile Val Phe Tyr His
TP84_03

AACGAAATATATCGGATCTGGGAAACAGTCGGCGAATTTTTTGCAGGGGAGGAAGGAAAATGAAACGATT
TTGCTTTATATAGCCTAGACCCTTTGTGAGCCGCTTAAAAAACGTCCCCTCCTTCCTTTTACTTTGCTAA

1680

80 85 90 95
Asn Glu Ile Tyr Arg Ile Trp Glu Thr Val Gly Glu Phe Phe Ala Gly Glu Glu Gly Lys
TP84_03

RBS

1 5
Met Lys Arg Phe
TP84_04

CGAAATCATGTTGAAAGACGGATTGGTGTATGCAGTTTTTCCGATATTGAATCCAAACTATCAAGGCAAC
GCTTTAGTACAACCTTCTGCCTAACCACATACGTCAAAAAGGCTATAACTTAGGTTTGATAGTTCCGTTG

1750

5 10 15 20 25
Glu Ile Met Leu Lys Asp Gly Leu Val Tyr Ala Val Phe Pro Ile Leu Asn Pro Asn Tyr Gln Glv Asn
TP84_04

GGCATTGACGAAATCATCGAAGCTGAAAGTTACGATGACGCAGTCCAAAAAGCGGAAATGAAATATGGCG
CCGTAACCTGCTTTAGTAGCTTCGACTTTCAATGCTACTGCGTCAGGTTTTTCGCCTTTACTTTATACCGC

1820

30 35 40 45 50
Gly Ile Asp Glu Ile Ile Glu Ala Glu Ser Tyr Asp Asp Ala Val Gln Lys Ala Glu Met Lys Tyr Gly
TP84_04

GATGGTATCAATCGCCAGGATCGCAAAACGAATCAACATTCAAAGAGGAAGAAAACGCAAGATATTGCC
CTACCATAGTTAGCGGTCCTAGCGTTTTGCTTAGTTGTAAGTTTCTCCTTCTTTTGC GTTCTATAAGCGG

55 60 65 70
Gly Trp Tyr Gln Ser Pro Gly Ser Gln Asn Glu Ser Thr Phe Lys Glu Glu Glu Asn Ala Arg Tyr Ser Pro
TP84_04

GGAACAAGTTTATGGCATCAAAGGTCTTTTATGAAAAGGGGGATCGATGGTGATCACAAATCCAAATGG
CCTTGTTCAAATACCGTAGTTTCCAGAAAATACTTTCCCCCTAGCTACCACTAGTGTTTAGGTTTTACC

75 80 85
Glu Gln Val Tyr Gly Ile Lys Gly Leu Leu
TP84_04

RBS

1 5 10
Met Lys Arg Gly Ile Asp Gly Asp His Lys Ser Lys Met
TP84_05

ACCGTCACAGACGTTCTATGGGCCACTCTAGGCCTTTTGATCATCGGGGCAATAGTTTATATCATGTCGG
TGGCAGTGCTCTGCAAGATACCGGTGAGATCCGCAAACTAGTAGCCCCGTTATCAAATATAGTACAGCC

15 20 25 30 35
Asp Arg His Arg Arg Ser Met Gly His Ser Arg Arg Phe Asp His Arg Gly Asn Ser Leu Tyr His Val Gly
TP84_05

TCTGAAAAGCCGAAACAGGACCGCTTTAAACGGTCTCTGTCCGGCCAGATGGGATCTGGTCGCTGACGATG
AGACTTTTCGGCTTTGTCTTGGCGAAATTTGCCAGGACAGCGCGGTCTACCCTAGACCAGCGACTGCTAC

40 45 50 55 60
Leu Lys Ser Arg Asn Arg Thr Ala Leu Asn Gly Pro Val Arg Pro Asp Gly Ile Trp Ser Leu Thr Met
TP84_05

GCAACCGTCAGACCAAAATCGAATAGGGGGAATTCAGATGGTGAAACAAATGACAGATGCACAAAGAAA
CGTTGGCAGTCTGGTTTTAGCTTATCCCCCTTAAGTCTACCACTTTGTTTACTGTCTACGTGTTTCTTT

65 70 75 80
Ala Thr Val Arg Pro Lys Ile Glu
TP84_05

RBS

5 10
Met Val Lys Gln Met Thr Asp Ala Gln Arg Lys
TP84_06

GGCGATCGAATGGGCGGTCAAATGGCGGAAAGCCAGATCAAATCAAGTCAAAGACGGATAAAAAGGGCG
CCGCTAGCTTACCGCCAGTTTTACCGCCTTTTCGGTCTAGTTTAGTTTCAGTTTCTGCCTATTTTCCCGC

15 20 25 30
Ala Ile Glu Trp Ala Val Lys Met Ala Glu Ser Gln Ile Lys Ser Ser Gln Arg Arg Ile Lys Arg Ala
TP84_06

GAACGATCACTGGAATTGTTCAAGGGCATGGATGAATTAGATGATAAGCTGAAAAGACGACAAATCAAAC
CTTGCTAGTGACCTTAACAAGTTCCCGTACCTACTTAATCTACTATTCGACTTTTCTGCTGTTTAGTTTG

35 40 45 50 55
Glu Arg Ser Leu Glu Leu Phe Lys Gly Met Asp Glu Leu Asp Asp Lys Leu Lys Arg Arg Gln Ile Lys
TP84_06

ATCAACAAAATCGGATCGAAAGTCTGCAAACTACATCGATGCACTTAAAAGTCTGCTTGATTGACGAAA
TAGTTGTTTTAGCCTAGCTTTCAGACGTTTTGATGTAGCTACGTGAATTTTCAGACGAACCTAACTGCTTT

60 65 70 75
His Gln Gln Asn Arg Ile Glu Ser Leu Gln Asn Tyr Ile Asp Ala Leu Lys Ser Leu Leu Asp
TP84_06

RBS

1
Met Thr Lys
TP84_07

GCCGAAACAGTCCAGGATCATCCCTGGACTGTCTGGCGTGATGGGATCACGTCAGTGACGATGGCAACCG
CGGCTTTGTCAGGTCCTAGTAGGGACCTGACAGACCCCACTACCCTAGTGACGTGACTGCTACCGTTGGC

Ala Glu Thr Val Gln Asp His Pro Trp Thr Val Trp Arg Asp Gly Ile Thr Ser Leu Thr Met Ala Thr
TP84_07

CGTCAGATCAAAAAACACGAAAGGGGATCGAAAAATGCAAGTCAGAAAAATCCACAAAGTCACGTATGAAT
GCAGTCTAGTTTTTGTGCTTTCCCTAGCTTTTACGTTCAAGTCTTTTTAGGTGTTTCAAGTGCATACTTA

Met Gln Val Arg Lys Ile His Lys Val Thr Tyr Glu
TP84_08
RBS
Ala Ser Asp Gln Lys Thr Arg Lys Gly Ile Glu Asn Ala Ser Gln Lys Asn Pro Gln Ser His Val

CAGAAACAATGCCGCTGGAACAAGCGGTGGAAGAAATCGCGCAAAGCACTGGATCGCCTGTGGACCGTGT
GTCTTTGTTACGGCGACCTTGTTCCGCCACCTTCTTTAGCGCGTTTCGTGACCTAGCGGACACCTGGCACA

Ser Glu Thr Met Pro Leu Glu Gln Ala Val Glu Glu Ile Ala Gln Ser Thr Gly Ser Pro Val Asp Arg Val
TP84_08

TCGCGGACTGATCATTCAATTCGGCAAAGTGATGACAGTCCATCATGTCTATCACATCGTGGATGCCAGT
AGCGCCTGACTAGTAAGTTAAGCCGTTTCACTACTGTCAAGGTAGTACAGATAGTGTAGCACCTACGGTCA

Arg Gly Leu Ile Ile Gln Phe Gly Lys Val Met Thr Val His His Val Tyr His Ile Val Asp Ala Ser
TP84_08

GGGATCGGGCTGGTGAAGGATGCAGAAAAATTCCTGGGCCAATCGATGGTGTTCCTCAACTGAAAAGGCGGCG
CCCTAGCCCGACCACTTCTACGTCTTTTAAAGGACCCGGTTAGCTACCACAAAGTTGACTTTCCGCCCGC

Gly Ile Gly Leu Val Lys Asp Ala Glu Lys Phe Leu Gly Gln Ser Met Val Phe Gln Leu Lys Gly Gly
TP84_08

AAACCGTCATTTTGAAGTCATGGAAGTGACACGATGGATGAATGGGGATTATTTTATCAAAGGATATGA
TTTGGCAGTAAAACTTTCAGTACCTTCACTGTGCTACCTACTTACCCCTAATAAAATAGTTTCCTATACT

Glu Thr Val Ile Leu Lys Val Met Glu Val Thr Arg Trp Met Asn Gly Asp Tyr Phe Ile Lys Gly Tyr Asp
TP84_08

TGATGAAGGGATGTGGCGCACGATCGATGTCGATTGCGATTAAAGCATATTTACGCCAGGAAAGGGGATG
ACTACTTCCCTACACCGCGTGCTAGCTACAGCTAAGCTAATTTCTGTATAAAGTGCGGTCCTTTCCCTAC

Asp Glu Gly Met Trp Arg Thr Ile Asp Val Asp Ser Ile Thr Lys Ala Tyr Phe Thr Pro Gly Lys Gly Met
TP84_08

RBS

ATCCGATGAAACCATATGATCCGATCCCGATGGCGATCACATTGAAACCACACCAAACGATCCGAAAAGG
TAGGCTACTTTGGTATACTAGGCTAGGGCTACCGCTAGTGTAACTTTGGTGTGGTTTGCTAGGCTTTTCC

Ile Arg
TP84_08

Met Lys Pro Tyr Asp Pro Ile Pro Met Ala Ile Thr Leu Lys Pro His Gln Thr Ile Arg Lys Gly
TP84_09

25 30 35 40 45
Glu Val Ile Gln Glu Thr Ile Gly Gly Lys Val Val Lys Phe Lys Val Ile Asp Ile Arg Met Val Glu
TP84_09

50 55 60 65
 Ile Arg Glu Thr His Ala Leu Val Gly Val Met Ala Arg Glu Ile Gly Glu Asp Ala Glu
 TP84_09

regulatory" ODD

ODD

regulatory" ODD

Met Asp Lys Ile Pro Lys Leu Val Met Ala Ser

15 20 25 30
 Asp Met Ala Glu Lys Trp Gly Val Asp Arg Arg Val Val Asn Asn Trp Ser Arg Arg Asp Pro Ser Phe
 TP84 10

35 40 45 50 55
Pro Lys Pro Val Met Val Val Gly Asn Gly Lys Tyr Pro Leu Tyr Leu Glu Glu Asp Met Ile Ala Tyr
TP84_10

60 65 70
 Gly Lys Gln Lys Gly Leu Glu Lys Gly Asp Asp Thr Gln Thr
 TP84 10

Met Ile Arg Ser Lys Asp Tyr Trp Glu Gly

Tyr Asn His Ala Gln Lys Glu Ala Glu Lys Arg Trp Glu Gln Ala Arg Ser Glu Leu Thr Ala Arg Phe

TTGAGATCATCGAAGATTCCAAGCGTGACGTGAAGGGGATCGGGCCAAAAGTACACAAAGCACTGGTCTGA
AACTCTAGTAGCTTCTAAGGTCGCACTGCACTTCCCCTAGCCCGGTTTTTCATGTGTTTCGTGACCAGCT
3570
Ile Glu Ile Ile Glu Asp Ser Arg Arg Asp Val Lys Gly Ile Gly Pro Lys Val His Lys Ala Leu Val Glu
TP84_11

ATACGTGATCGAACGAATGATGAATGGGGGGCGGAAATAATGTGGCTAGGATTGAAACACGGCAACAAT
TATGCACTAGCTTGCTTACTACTTACCCCCCGCCTTTATTACACCGATCCTAAGCTTTGTGCCGTTGTTA
3640
Met Trp Leu Gly Phe Glu Thr Arg Gln Gln
TP84_12
RBS
Tyr Val Ile Glu Arg Met Met Asn Gly Gly Arg Lys
TP84_11

ATCTGGACTGGCTTGAATCGTTGCAACCAGGAAACGAAGTGGCAGTGGGATTCGCGATTTCGATATGGAAG
TAGACCTGACCGAACTTAGCAACGTTGGTCCTTTGCTTCACCGTCACCCCTAAGGCGTAAGCTATACCTTC
3710
Tyr Leu Asp Trp Leu Glu Ser Leu Gln Pro Gly Asn Glu Val Ala Val Gly Phe Arg Ile Arg Tyr Gly Arg
TP84_12

GGATGTGCGTTATGAGATCCGAACAGTCACAAGGGTGACAAATCGGCAAATCATCCTGGATGATGGTGCA
CCTACACGCAATACTCTAGGCTTGTCAGTGTTCCCACTGTTTAGCCGTTTAGTAGGACCTACTACCACGT
3780
Asp Val Arg Tyr Glu Ile Arg Thr Val Thr Arg Val Thr Asn Arg Gln Ile Ile Leu Asp Asp Gly Ala
TP84_12

ATCCGCTTCAATCGCGGCACAGGGAAGCAGTGGGGCGGATATGCATGGATCGAGCCAGTCACCGAAAAAA
TAGGCGAAGTTAGCGCCGTGTCCCTTCGTACCCCGCCTATACGTACCTAGCTCGGTCAAGTGGCTTTTTT
3850
Ile Arg Phe Asn Arg Gly Thr Gly Lys Gln Trp Gly Gly Tyr Ala Trp Ile Glu Pro Val Thr Glu Lys
TP84_12

TTCTATCGCAAAAATACGGTAAACAGGCATAATTCGCTGATCCCTGCATACACTGGATTTCATCCAGTGT
AAGATAGCGTTTTTATGCCATTTGTCCGTATTAAGGCGACTAGGGACGTATGTGACCTAAGTAGGTCACA
3920
Ile Leu Ser Gln Lys Tyr Gly Lys Gln Ala
TP84_12

TTGTGGGGATACTCTACACACCGGGGGGATTGAGATGGCATTGAAAGTGTTAAAAAGCAGATGACATGC
AACACCCCTATGAGATGTGTGGCCCCCCTAAGTCTACCGTAACTTTTACCAATTTTTCGTCTACTGTACG
3990
Met Ala Leu Lys Val Val Lys Lys Gln Met Thr Cys
TP84_13
RBS

AAAGAGATCATGATCCTTTATTTCTGCGATGACTGCGGCATCCAGGAATG6TCGGATCGGATCATTCCGA
TTTCTCTAGTACTAGGAAATAAAGACGCTACTGACGCCGTAGGTCCTTACCAGCCTAGCCTAGTAAGGCT
4060
Lys Glu Ile Met Ile Leu Tyr Phe Cys Asp Asp Cys Gly Ile Gln Glu Trp Ser Asp Arg Ile Ile Pro
TP84_13

AAAAGCGATGCGAAAAGTGTGGACTGTTTATGGATCACGAAGAATGGGAAGGGTGATGGTGAATGCATAT
TTTTCGCTACGCTTTTCACACCTGACAAATACCTAGTGCTTCTTACCCTTCCCCTACCACTTACGTATA

Lys Lys Arg Cys Glu Lys Cys Gly Leu Phe Met Asp His Glu Glu Trp Glu Gly
TP84_13 Met His Ile
TP84_14

RBS

CACAAAAAACTGATCCCGGCCGATCAATCCGAAATATATCAAATCCCAAAGCATCACGCACTGGATGTC
GTGTTTTTTTGGCTAGGGCCGGCTAGTTAGGCTTTATATAGTTTAGGGTTTCGTAGTGCCTGACCTACG

Thr Lys Lys Leu Ile Pro Ala Asp Gln Ser Glu Ile Tyr Gln Ile Pro Lys His His Ala Leu Asp Val
TP84_14

CTGGCGAAAGAGATCACCAAAACAACAAGACACATCGAACGGATCAAGATGGACATCCAGGCGGTGAAGG
GACCGCTTTCTCTAGTGTTTGTGTTTCTGTGTAGCTTGCTAGTTCTACCTGTAGGTCCGCCACTTCC

Leu Ala Lys Glu Ile Thr Lys Gln Gln Arg His Ile Glu Arg Ile Lys Met Asp Ile Gln Ala Val Lys
TP84_14

ATCAGCCGAAAAAGGTGGCAGATGGCAAAACAGATCGAGATTACGAAAAAGCGAATCCAGGAATGTCTG
TAGTCGGCTTTTCCACCGTGCTACCGTTTTTGTCTAGCTCTAAGTGCTTTTCGCTTAGGTCCTTACAGC

Asp Gln Pro Lys Lys Val Ala Arg Trp Gln Lys Gln Ile Glu Ile His Glu Lys Arg Ile Gln Glu Cys Arg
TP84_14

CCAGGCGATCAAATGGATCATGGAAAATCTG6TGTAAGGCACTGAAAAAATCAGTGCCTTTTTTATTT
GGTCCGCTAGTTTACCTAGTACCTTTTAGACCACATTCCGTGACTTTTTTTAGTCACGGAAAAAATAAA

Gln Ala Ile Lys Trp Ile Met Glu Asn Leu Val
TP84_14

regulatory region

TTTATGGAATATAGACAGGCACACATGAAATATGTATGTAATGCAAAGCAAATCAAACTAAATGAAAAG
AAATACCTTATATCTGTCCGTGTGTACTTTATACATACATTACGTTTCGTTTAGTTTTGATTTACTTTTC

regulatory region RBS

GGGATGGAATGACATGAAAAAATGAAAAATACACTGTGGTGATCAGAAAATTTTCAGGCAGAAAGGAA
CCCTACCTTACTGTACTTTTTTTACTTTTTTATGTGACACCACTAGTCTTTTAAAAGTCCGTCCTTTCCTT

Met Lys Lys Met Lys Lys Tyr Thr Val Val Ile Arg Lys Phe Ser Gly Arg Lys Glu
TP84_15

RBS

AAATGGGGATTTTTAGCGCAGAACGAACAGATCGCAAAAGCTATTGCAGATGCAATCAGAAAGGGTGGCA
TTTACCCCTAAAAATCGCGTCTTGCTTGCTAGCGTTTTCGATAACGTCTACGTTAGTCTTTCCACCGT

Lys Trp Gly Phe Leu Ala Gln Asn Glu Gln Ile Ala Lys Ala Ile Ala Asp Ala Ile Arg Lys Gly Gly
TP84_15

RBS

AGTAATGTCAATGAAAACGAAAGGGGTGGCATTCAATCTGGCCGATCCCGATCAAAAGGCATTATATGAA
TCATTACAGTTACTTTTGTCTTCCCCACCCTAAGTTAGACCGGCTAGGGCTAGTTTTCCGTAATATACTT

4690

1 5 10 15 20
Met Ser Met Lys Thr Lys Gly Val Ala Phe Asn Leu Ala Asp Pro Asp Gln Lys Ala Leu Tyr Glu
TP84_16

Lys
TP84_15

TATGCCATGACACGCACCAATTTTCAGCGCATACGTCAAACGGCTGATCCAGGCTGATATGCAACGGCAAC
ATACGGTACTGTGCGTGGTTAAAGTCGCGTATGCAGTTTGCCGACTAGGTCCGACTATACGTTGCCGTTG

4760

25 30 35 40 45
Tyr Ala Met Thr Arg Thr Asn Phe Ser Ala Tyr Val Lys Arg Leu Ile Gln Ala Asp Met Gln Arg Gln
TP84_16

AAGCAAAAACCGCAGGGCGACAACATCCTGCGGTCCATTCCGCTGAATATGGCGGCATTAAAGATCCGAAT
TTCGTTTTTGGCGTCCCGCTGTTGTAGGACGCCAGGTAAGGCGACTTATACCGCCGTAATTCTAGGCTTA

4830

50 55 60 65
Gln Ala Lys Thr Ala Gly Arg Gln His Pro Ala Val His Ser Ala Glu Tyr Gly Gly Ile Lys Ile Arg Ile
TP84_16

TGGCGAAAAGTCACTGGGTGAATAGCTTGACGATGATGGCCGCCGTGCCAGTGGTCACAAGCACATATCC
ACCGCTTTTCAGTGACCCACTTATCGAACTGCTACTACCGGCGGCACGGTCACCAAGTTCGTGTATAGG

4900

70 75 80 85 90
Gly Glu Lys Ser Leu Gly Glu Met Thr Met Met Ala Ala Val Pro Val Val Thr Ser Thr Tyr Pro
TP84_16 TP84_17

GATTCCGACAAGGATCGGCATCGATGATCACATCCTTTTCCAAAGTGTCGATCGTGTATAGTATGAAC
CTAAGGCTGTTTCTAGCCGTAGCTACTAGTGTAGGAAAAGGTTTCACAGCTAGCACAAATATCATACTTG

4970

20 25 30 35
Ile Pro Thr Arg Ile Gly Ile Asp Asp His Ile Leu Phe Gln Ser Val Asp Arg Val Tyr Ser Met Asn
TP84_17

AAAAAACGAAAGGGTGATACATGTGTTACACGAAAAAAGGTGGAAACCATTCCATTTTCGCGAATTCATG
TTTTTTGCTTTCCCACTATGTACACAAGTGTGCTTTTTTCCACCTTTGGTAAGGTAAAGCGCTTAAGTAC

5040

RBS Met
TP84_18

40 45 50 55 60
Lys Lys Arg Lys Gly Asp Thr Cys Val His Thr Lys Lys Gly Gly Asn His Ser Ile Ser Arg Ile His
TP84_17

GATGGCAACTGGCGCGAAAAGAAGGCTGAACGCCGCAAAGAAAGCCTGGATGAAGTGATCACCAGACTGG
CTADCGTTGACCGCGCTTTTCTTCCGACTTGCGGCGTTTCTTTCGGACCTACTTCACTAGTGGTCTGACC

5110

5 10 15 20
Asp Gly Asn Trp Arg Glu Lys Lys Ala Glu Arg Arg Lys Glu Ser Leu Asp Glu Val Ile Thr Arg Leu
TP84_18

65 70
Gly Trp Gln Leu Ala Arg Lys Glu Gly
TP84_17

CGGCGGTTGGAACCTGGCGCACATATGGTGATCCATCCGATCGATGCACACGCCGCTGACATGGGCGACAT
GCCGCCAACCTTGACCGCGTGTATACCACTAGGTAGCTAGCTACGTGTGCGGCGACTGTACCCGCTGTA

5180

25 30 35 40 45
Ala Ala Val Gly Thr Gly Ala His Met Val Ile His Pro Ile Asp Ala His Ala Ala Asp Met Gly Asp Ile
TP84_18

CATTATGAAGGCGTGCCAACCGATCATTGATTTATTGCAAGGCATTTTCGTATCCAGTCGCTTTTCATCATG
GTAATACTTCCGCACGGTTGGCTAGTAACATAAATAACGTTCCGTAAAGCATAGGTCAGCGAAAAGTAGTAC

5250

50 55 60 65 70
Ile Met Lys Ala Cys Gln Pro Ile Ile Asp Leu Leu Gln Gly Ile Ser Tyr Pro Val Ala Phe Ile Met
TP84_18

ATCACTGGCGGATTCTGCTGATCATGACTGGGCAGACATCCAGGGGGATGCATTTTCATCAAATGGGCGT
TAGTGACCGCCTAAGGACGACTAGTACTGACCCGTCTGTAGGTCCCCCTACGTAAAGTAGTTTACCCGCA

5320

75 80 85 90
Ile Thr Gly Gly Phe Leu Leu Ile Met Thr Gly Gln Thr Ser Arg Gly Met His Phe Ile Lys Trp Ala
TP84_18

GCCTGGGATATTTAGGATTGCAATTTGCGCCTGCACTGATGCAGATTGTGATCCAGATCGGTCAAAACAT
CGGACCCTATAAATCCTAACGTTAAACGCGGACGTGACTACGTCTAACACTAGGTCTAGCCAGTTTGTGTA

5390

95 100 105 110 115
Cys Leu Gly Tyr Leu Gly Leu Gln Phe Ala Pro Ala Leu Met Gln Ile Val Ile Gln Ile Gly Gln Asn Ile
TP84_18

TCAAGCACAAAGTGGGGGTGGTACGATGAAGCATTTCCGCATTATCCCGGATCAAAAGATTGACAACCGTC
AGTTTCGTGTTACCCCCACCATGCTACTTCGTAAAGCGTAATAGGGCCTAGTTTTCTAACTGTTGGCAG

5460

1 5 10 15
Met Lys His Phe Arg Ile Ile Pro Asp Gln Lys Ile Asp Asn Arg
RBS
120 125
Gln Ala Gln Val Gly Val Val Arg
TP84_18

ATGTGGATCAGATGCTAAAGACGCTGAACATATATAAAGTCCAATACAGCGGTGGAATGGTCGGGGATT
TACACCTAGTCTACGATTCTGCGACTTGATATATATTTTCAGGTTATGTCGCCACCTTACCAGCCCCCTAA

5530

20 25 30 35
His Val Asp Gln Met Leu Lys Thr Leu Asn Ile Tyr Lys Ser Pro Ile Gln Arg Trp Asn Gly Arg Gly Phe
TP84_19

CGATCGAATCCCTTTTATTTTCATACGAAATACACATCCAAAAGGATCACACCACATTCCACATGACGGTG
GCTAGCTTAGGGAAAATAAAGTATGCTTTATGTGTAGGTTTTCTAGTGTTGGTGTAAAGGTGTAAGGTGTAAGTGGCAC

5600

40 45 50 55 60
Asp Arg Ile Pro Phe Ile Ser Tyr Glu Ile His Ile Gln Lys Asp His Thr Thr Phe His Met Thr Val
TP84_19

GCCGATGAAATCGAAACGATCGCGAAAAAAGCACTGGAAACCGCTGGCCGAATGCATCGATCGATCTGG
CGGCTACTTTAGCTTTGCTAGCGCTTTTTTCGTGACCTTTGGCGGACCGGCTTACGTAGCTAGCTAGACC

5670

65 70 75 80 85
Ala Asp Glu Ile Glu Thr Ile Ala Lys Lys Ala Leu Glu Thr Ala Trp Pro Asn Ala Ser Ile Asp Leu
TP84_19

CCGCTGATCCATTCCAGGGATTGCAACCGAACGCAATCAGCCGTCTGGAATTGAAACATCATTCCATGTT
GGCGACTAGGTAAGGTCCTAAGCTTGGCTTGGCTTAGTCGGCAGACCTTAACCTTTGTAGTAAGGTACAA

5740

90 95 100 105
Ala Ala Asp Pro Phe Gln Gly Phe Glu Pro Asn Ala Ile Ser Arg Leu Glu Leu Lys His His Ser Met Phe
TP84_19

TGCCATTGCGGTGGATCGCCGGGAAAATGGATTGATGACAATTTGTTTGAATCGTTGAAGGTGATGGGC
ACGGTAAGCCCACCTAGCGGCCCTTTTACCTAAGTACGTGTTAAACAAGCTTAGCAACTTCCACTACCCG

5810

110 115 120 125 130
Ala Ile Arg Val Asp Arg Arg Glu Asn Gly Phe Met His Asn Leu Phe Glu Ser Leu Lys Val Met Gly
TP84_19

CAGGATGATCGGATCTTCATTGAGATCCTTTGCATCCCTGCATCCCTGGACTGGTACACATCCGCCGCTG
GTCCTACTAGCCTAGAAGTAAGTCTAGGAAACGTAGGGACGTAGGGACCTGACCATGTGTAGGCGGCGAC

5880

135 140 145 150 155
Gln Asp Asp Arg Ile Phe Ile Gln Ile Leu Cys Ile Pro Ala Ser Leu Asp Trp Tyr Thr Ser Ala Ala
TP84_19

ATGCCTACAAACGTTTTAAGGAAGGCAAGATGCCGAAAAAGATTACATCGATGGCAAAGAAGTGGCCAG
TACGGATGTTTGCAAAATTCCTTCCGTTCTACGGCTTTTCTAAGTGTAGCTACCGTTTCTTGACCGGTC

5950

160 165 170 175
Asp Ala Tyr Lys Arg Phe Lys Glu Gly Lys Met Pro Lys Lys Ile His Ile Asp Gly Lys Glu Leu Ala Arg
TP84_19

GGAAGGGATCAAACCTGGTGACACGTGGTGTCTGGGCGTGATCGACACGATTGTGGTCCTAACAGGCGGC
CCTTCCCTAGTTTGACCACTGTGCACCACAGGACCCGCACTAGCTGTGCTAACACCAGGATTGTCCGCCG

6020

180 185 190 195 200
Glu Gly Ile Lys Leu Val Thr Arg Gly Val Leu Gly Val Ile Asp Thr Ile Val Val Leu Thr Gly Gly
TP84_19

ACACCAGAAAAGATCAATCTTGACAATGCAGATCGGGCCATGATGATGCGTGATGGTGGCCTGCGACTGG
TGTGGTCTTTTCTAGTTAGAAGCTGTTACGTCTAGCCCGGTACTACTACGCACTACCACCGGACGCTGACC

6090

205 210 215 220 225
Thr Pro Glu Lys Ile Asn Leu Asp Asn Ala Asp Arg Ala Met Met Met Arg Asp Gly Gly Leu Arg Leu
TP84_19

AAACGATCAACAAGACAAAACACGATGCGTTCGATGTCACGATCCGTGTGGCGATCCAGTCAAAGAATCA
TTTGCTAGTTGTTCTGTTTTGTGCTACGCAAGCTACAGTGCTAGGCACACCGCTAGGTCAGTTTCTTAGT

6160

230 235 240 245
Glu Thr Ile Asn Lys Thr Lys His Asp Ala Phe Asp Val Thr Ile Arg Val Ala Ile Gln Ser Lys Asn His
TP84_19

CGAAAACATCATGAAAATGGTGGCCGCTTCATTCCGTGAATTCGATGGCGACAATCAATTCACATATTAT
GCTTTTGTAGTACTTTTACCACCGGCGAAGTAAGGCACCTAAGCTACCGCTGTTAGTTAAGTGTATAATA

6230

250 255 260 265 270
Glu Asn Ile Met Lys Met Val Ala Ala Ser Phe Arg Glu Phe Asp Gly Asp Asn Gln Phe Thr Tyr Tyr
TP84_19

CCACTGGATGTGAATAAGATGACGATCGGATGGATCGCCGATCGGAAACCAGGGATCAAGCTGTCAAAGG
GGTGACCTACACTTATTCTACTGCTAGCCTACCTAGCGGCTAGCCTTTGGTCCCTAGTTCGACAGTTTCC

6300

275 280 285 290 295
Pro Leu Asp Val Asn Lys Met Thr Ile Gly Trp Ile Ala Asp Arg Lys Pro Gly Ile Lys Leu Ser Lys
TP84_19

ACTATCTATCGACACCAGAACTGGCCAGGATCATTATATGCCGACTGGACCGATCCAGATGAAATGGAA
TGATAGATAGCTGTGGTCTTGACCGGTCTAGTAAGTATACGGCTGACCTGGCTAGGTCTACTTTACCTT

6370

300 305 310 315
Asp Tyr Leu Ser Thr Pro Glu Leu Ala Arg Ile Ile His Met Pro Thr Gly Pro Ile Gln Met Lys Trp Asn
TP84_19

CATCGAACGGATCGAGAAAACACAGACAGCCGTGCCATCCAGGATCACGACTGGTCCTGGCATCTTGATC
GTAGCTTGCTAGCTCTTTTGTGTCTGTGCGCACGGTAGGTCCTAGTGCTGACCAGGACCGTAGAACTAG

6440

320 325 330 335 340
Ile Glu Arg Ile Glu Lys Thr Gln Thr Ala Val Pro Ser Arg Ile Thr Thr Gly Pro Gly Ile Leu Ile
TP84_19

GGTGTGCATACACACAAAAGGAGTGGAACAAACGGTGTTCATGCCGATCGATGACTGGGATGAGTTGTGTC
CCACACGTATGTGTGTTTCCTCACCTTGTTTGCCACAAGTACGGCTAGCTACTGACCCTACTCAACACAG
345 350 355 360 365
Gly Val His Thr His Lys Gly Val Glu Gln Thr Val Phe Met Pro Ile Asp Asp Trp Asp Glu Leu Cys
TP84_19 >

TGCCGCGTGTGGCGATCGGCGGCATGGGCCAGGGCAAGACAAGAGGATTGCGCGCAAACCTGGTTATATCA
ACGGCGCACACCGCTAGCCGCCGTACCCGGTCCCGTTCTGTTCTCCTAAGCCGCGTTTGACCAATATAGT
370 375 380 385
Leu Pro Arg Val Ala Ile Gly Gly Met Gly Gln Gly Lys Thr Arg Gly Phe Gly Ala Asn Trp Leu Tyr Gln
TP84_19 >

AGCGGTCAAAAATGGATTGCGCGGCCTGGCGATCGATCCGGCAAAAAGGGGAAATCGGCAACGAACTGGCG
TCGCCAGTTTTTACCTAAGCCGCCGGACCGCTAGCTAGCCGTTTTCCCTTTAGCCGTTGCTTGACCGC
390 395 400 405 410
Ala Val Lys Asn Gly Phe Gly Gly Leu Ala Ile Asp Pro Ala Lys Gly Glu Ile Gly Asn Glu Leu Ala
TP84_19 >

GCGGTCTGGATGAAGATGAATTCATTGGGATCAACATCGCGCAAAATCCGATCTGCCTGGACTGGTGTG
CGCCAGGACCTACTTCTACTTAAGTAAGCCTAGTTGTAGCGCGTTTTAGGCTAGACGGACCTGACCACAC
415 420 425 430 435
Ala Val Leu Asp Glu Asp Glu Phe Ile Arg Ile Asn Ile Ala Gln Asn Pro Ile Cys Leu Asp Trp Cys
TP84_19 >

AAACAAAATACAGTGAAGTGAAGTGAATTCATTGGGATCAACATCGCGCAAAATCCGATCTGCCTGGACTGGTGTG
TTTGTGTTTATGTCACTTGACCGGTCCCCGTCTGACCGTTTATGCTACTAGTCGAAAAAGTTAAGCTTACG
440 445 450 455
Glu Thr Lys Tyr Ser Glu Leu Ala Arg Gly Arg Leu Ala Asn Ile Thr Met Ile Ser Phe Phe Asn Ser Asn Ala
TP84_19 >

CGATGATGCAGGAGTGCAGACACAGCGATATATCCGGCGATGTTTATGGGGATGCAGGGGAACAACTG
GCTACTACGTCCTCACGTCTGTGTGCTATATAGGCCCGCTACCAATACCCCTACGTCCCTTGTGTTGAC
460 465 470 475 480
Asp Asp Ala Gly Val Gln Thr Gln Arg Tyr Ile Arg Ala Met Val Met Gly Met Gln Gly Asn Lys Leu
TP84_19 >

GATGAATTGATCCGCATGATGAATGATATGGCCTATTTGCAAAAATGCGTGGAAAAGATGCCGCCAGGAT
CTACTTAAGTGGCGTACTACTTACTATACCGGATAAACGTTTTTACGCACCTTTTCTACGGCGGCTCTA
485 490 495 500 505
Asp Glu Leu Ile Arg Met Met Asn Asp Met Ala Tyr Leu Gln Lys Cys Val Glu Lys Met Pro Pro Gly
TP84_19 >

TCCACCGGGCCACACTGGAAGAACTGATCAACTATTCGGATGCCAGACGAATGCAGATCCTATCGCCGAT
AGGTGGCCCCGGTGTGACCTTCTTGACTAGTTGATAAGGCTACGGTCTGCTTACGTCTAGGATAAGCGGCTA
510 515 520 525
Phe His Arg Ala Thr Leu Glu Glu Leu Ile Asn Tyr Ser Asp Ala Arg Arg Met Gln Ile Leu Ser Pro Ile
TP84_19 >

CTTGAATCGATTAGACATGATCCTGGGCGATCCATTCTGGCGAAATGCATGGAATCCGATCGATCCCTG
GAACTTAGCTAATCTGTACTAGGACCCGCTAGGTAAGGACCGCTTTACGTACCTTAGGCTAGCTAGGGAC
530 535 540 545 550
Leu Asn Arg Leu Asp Met Ile Leu Gly Asp Pro Phe Leu Ala Lys Cys Met Glu Ser Asp Arg Ser Leu
TP84_19 >

GACATGGTGGACATTCTGCGGCAACGAAAAGCGATCGTGATCGATGTACCGAAAAAGGATGTCGGTCCAG
CTGTACCACCTGTAAGACGCCGTTGCTTTTCGCTAGCACTAGCTACATGGCTTTTCTACAGCCAGGTC

7140

555 560 565 570 575
Asp Met Val Asp Ile Leu Arg Gln Arg Lys Ala Ile Val Ile Asp Val Pro Lys Lys Asp Val Gly Pro

AAGGCGTGGACATCATTGTGAATCTTCTATCCACCAAAATCGATCTGGCGATGACACTGCGGCCAGACGA
TTCCGCACCTGTAGTAACACTTAGAAGATAGGTGGTTTTAGCTAGACCGCTACTGTGACGCCGGTCTGCT

7210

580 585 590 595
Glu Gly Val Asp Ile Ile Val Asn Leu Leu Ser Thr Lys Ile Asp Leu Ala Met Thr Leu Arg Pro Asp Asp

TGAACAATTTCCGTTTTTCATCGTGTTTCGATGAACCGCACCAATACATGAGAAGCCACACGATCTGGAAA
ACTTGTTAAAGGCAAAAAGTAGCACAAGCTACTTGGCGTGGTTATGTACTCTTCGGTGTGCTAGACCTTT

7280

600 605 610 615 620
Glu Gln Phe Pro Phe Phe Ile Val Phe Asp Glu Pro His Gln Tyr Met Arg Ser His Thr Ile Trp Lys

TCAGCATGTGTGGAATCCAGGAAATGGCGTGTGCGCTATGTCTGGATGTTTCACGAATGGACACAAATCG
AGTCGTACACACCTTAGGTCCTTTACCGCACAGCCGATACAGACCTACAAAGTGCTTACCTGTGTTTAGC

7350

625 630 635 640 645
Ser Ala Cys Val Glu Ser Arg Lys Trp Arg Val Gly Tyr Val Trp Met Phe His Glu Trp Thr Gln Ile

ATGACAAACTGCGGAAGATCATGAAATCCGCGTTGCCGCACTATCACGTTTATCCATCGTCAAAAAACAC
TACTGTTTGACGCCTTCTAGTACTTTAGGCGCAACGGCGTGATAGTGCAAATAGGTAGCAGTTTTTTGTG

7420

650 655 660 665
Asp Asp Lys Leu Arg Lys Ile Met Lys Ser Ala Leu Pro His Tyr His Val Tyr Pro Ser Ser Lys Asn Thr

ATTCTTGGATCTGAAAGAAGAACTGCAACCATTGCAACTGGATGACTTTCTAAAACCTGGAACGCTGGCAC
TAAGGACCTAGACTTTCTTCTTGACGTTGGTAAGCTTGACCTACTGAAAGATTTTGACCTTGCGACCGTG

7490

670 675 680 685 690
Phe Leu Asp Leu Lys Glu Glu Leu Gln Pro Phe Glu Leu Asp Asp Phe Leu Lys Leu Glu Arg Trp His

GCCATCAACGTGATCCGATCCGGCGGCCAAACGATCACACCATTTCATCGCCAGGATGACACCGCCGCCAA
CGGTAGTTGCACTAGGCTAGGCCGCCGGTTTGCTAGTGTTGGTAAGTAGCGGTCCTACTGTGGCGGCGGTT

7550

695 700 705 710 715
Ala Ile Asn Val Ile Arg Ser Gly Gly Gln Thr Ile Thr Pro Phe Ile Ala Arg Met Thr Pro Pro Pro

CAAAAAGAGAAATCTAGTTTTTTGTTGTAATTTGCATAATTTTTGTTACAATTGAATGTGTAAATATTA
GTTTTCTCTTTAGATCAAAAAACAACATTAAACGTATTAAAAACAATGTTAACTTACACATTTTATAAT

7630

720
Thr Lys Arg Glu Ile

CAATTCAAAGAAAGGGTGGACAATAGATGAAAAAGGTATTAAAAGCTATTTTGGCCGGGGCGTTGACATT
GTTAAGTTTCTTTCCACCTGTTATCTACTTTTTCCATAATTTTCGATAAAACCGGCCCGCAACTGTAA

7700

RBS

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

CGGTGTGATGGGAAGTGTGTTTGCAGTATCACCAAAACCGAAGTCCAAGCGGCTGTCGTGGAAGATTGG
GCCACACTACCCCTTGACACAAACGTCATAGTGGTTTTTGGCTTCAGGTTGCGCGACAGCACCTTCTAACC
20 25 30 35
Gly Val Met Gly Thr Val Phe Ala Val Ser Pro Lys Thr Glu Val Gln Ala Ala Val Val Glu Asp Trp
TP84_20

AGCGGATACAAAGAAACAGATCTGCAATTACCTATTGACGCAAGTCTCGCTGACGCTGATCGGGATTATA
TCGCCTATGTTTCTTTGCTAGACGTTAATGGATAACTGCGTTTCAGAGCGACTGCGACTAGCCCTAATAT
40 45 50 55 60
Ser Gly Tyr Lys Glu Thr Asp Leu Gln Leu Pro Ile Asp Ala Ser Leu Ala Asp Ala Asp Arg Asp Tyr
TP84_20

ATTATGGCGACACATACAATGTATTAAAAACGGTCTGAATGGCGTTTTACTCCGAAACACATTAAAACTTA
TAATACCGCTGTGTATGTTACATAATTTTGCAGCTTACCGCAAAATGAGGCTTTGTGTAATTTTAGAT
65 70 75 80 85
Asn Tyr Gly Asp Thr Tyr Asn Val Leu Lys Thr Val Glu Trp Arg Phe Thr Pro Lys His Ile Lys Ile Tyr
TP84_20

TCGTGTGAACGAAGATAAAACGCTACAAAGATATAAGACTATTTACCCGGAACCTATAAATGTTATCGAT
AGCACACTTGCTTCTATTTTGCAGTGTCTCTATATTCTGATAAATGGGCCTTGAATATTTACAATAGCTA
90 95 100 105
Arg Val Asn Glu Asp Lys Thr Leu Gln Arg Tyr Lys Thr Ile Tyr Pro Glu Leu Ile Asn Val Ile Asp
TP84_20

GECATAGAATATCAGTATAGATTCCAGACACCGATCACGACTGGTTTTCCCTGCTGGAAACTATTATGCAG
CCGTATCTTATAGTCATATCTAAGGTCTGTGGCTAGTGCTGACCAAAGGGACGACCTTTGATAATACGTC
110 115 120 125 130
Gly Ile Glu Tyr Gln Tyr Arg Phe Gln Thr Pro Ile Thr Thr Gly Phe Pro Ala Gly Asn Tyr Tyr Ala
TP84_20

ScaI

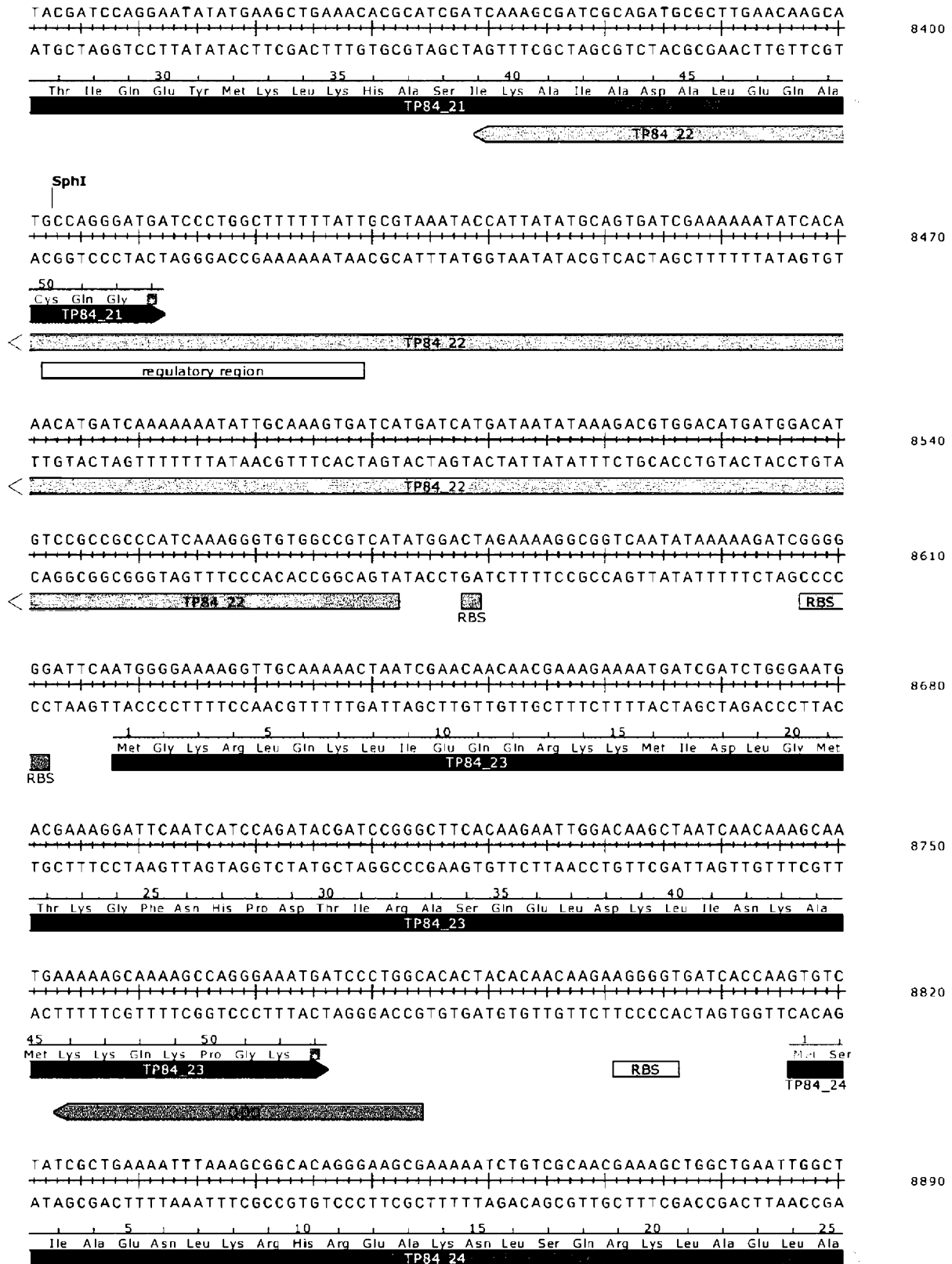
TACTAACATACACATATGAATATGCAGGCGGAAACCTTCGTGACAGAAGCCTATCGCTCGTATCGTTTCAC
ATGATTGTATGTGTATACCTTATACGTCCGCCTTTGAAGCACTGTCTTCGGATAGCGAGCATAGCAAAGTG
135 140 145 150 155
Val Leu Thr Tyr Thr Tyr Glu Tyr Ala Gly Gly Asn Phe Val Thr Glu Ala Tyr Arg Ser Tyr Arg Phe Thr
TP84_20

AATCAATTGATACTGAAACAATTCGAAAGGTCGTTCTTGATCAGACAGGAACGGCCTTTTTTGTGCCC
TTAGTTAACTATGACTTTGTTAAGCTTTCCAGCAAGAACCTAGTCTGTCTTGCCGGAAAAACAACGGG
160 165
Ile Asn
TP84_20

regulatory region

AATAAGTGCCAGTGACACACGTGGTGCAAAAGTGATATGATGGGCGTGAGAAAGGGGGAATGAGTGGTGA
TTATTCACGGTCACTGTGTGCACCACGTTTTCACTATACTACCCGCACTCTTTCCCCCTTACTACCACT
170 175
Met Val
TP84_21

ACAAAATCGATCTGATCAAAACAACCTACTGGAAGCAGGCAAACTGCCGATGCTTGACAAAGACAAAGATTA
TGTTTTAGCTAGACTAGTTTGTGATGACCTTCGTCCGTTTGACGGCTACGAACTGTTTCTGTTTCTAAT
5 10 15 20 25
Asn Lys Ile Asp Leu Ile Lys Gln Leu Leu Glu Ala Gly Lys Leu Pro Met Leu Asp Lys Asp Lys Asp Tyr
TP84_21



GGCGTGCCGCAATCATTGATCAGTCACATCGAAAATGGCGAAAAGAAAAATCCAGGTGTGATCGGCATTA
CCGCACGGCGTTAGTAAC TAGTCAGTGTAGCTTTTACCGCTTTTCTTTT TAGGTCCACACTAGCCGTAAT

30 35 40 45
Gly Val Pro Gln Ser Leu Ile Ser His Ile Glu Asn Gly Glu Lys Lys Asn Pro Gly Val Ile Gly Ile
TP84_24

AAAACTGGCCGATGCACTGGGGATCACAATGGAAGAACTGATCGAAAGCAAGGGGGATGAGGATGAGGA
TTTTTGACCGGCTACGTGACCCCTAGTGTTACCTTCTTGACTAGCTTTCGTTCCCCCTACTCCTACTCCT

50 55 60 65 70
Lys Lys Leu Ala Asp Ala Leu Gly Ile Thr Met Glu Glu Leu Ile Glu Ser Lys Gly Asp Glu Asp Glu Glu
TP84_24
RBS
Met Arg
TP84_25

AATGGGCGGTCTTGACCGATGATCTAAAATTTTGCGAAATCACATTCGATGGCGAATGGTATCGTGTAAG
TTACCCGCCAGAAGCTGGCTACTAGATTTTAAAACGCTTTAGTGTAAGCTACCGCTTACCATAGCACATTC

5 10 15 20 25
Lys Trp Ala Val Leu Thr Asp Asp Leu Lys Phe Cys Glu Ile Thr Phe Asp Gly Glu Trp Tyr Arg Val Arg
TP84_25

75
Met Gly Gly Leu Asp Arg
TP84_24

AATCACCGATGAATTCGGTCAGGAATTAATGACAAACGAATCGATCAGCTTACAAGGGGCGAAAAAACAA
TTAGTGGCTACTTAAGCCAGTCCTTAATTACTGTTTGCTTAGCTAGTCGAATGTTCCCGCTTTTTTGT

30 35 40 45
Ile Thr Asp Glu Phe Gly Gln Glu Leu Met Thr Asn Glu Ser Ile Ser Leu Gln Gly Ala Lys Lys Gln
TP84_25

GTCAAACGATTTCATCGGCACTGECCAAACGTTGAAATGGCAGGAAGTCGAAACCGACAAGATCGAAACAC
CAGTTTGCTAAGTAGCCGTGACCGGTTTGCAACTTTACCGTCCTTCAGCTTTGGCTGTTCTAGCTTTGTG

50 55 60 65 70
Val Lys Arg Phe Ile Gly Thr Gly Gln Thr Leu Lys Trp Gln Glu Val Glu Thr Asp Lys Ile Glu Thr
TP84_25

CATCCAAACTGCATCTGCCATTCGGTTATCTGTCTGTTTCACAAGTTAGAAAATATCTGACATGTCCACG
GTAGGTTTGACGTAGACGGTAAGCCAATAGACAGACAAAGTGTTCAATCTTTTATAGACTGTACAGGTGC

75 80 85 90 95
Pro Ser Lys Leu His Leu Pro Phe Gly Tyr Leu Ser Val Ser Gln Val Arg Lys Tyr Leu Thr Cys Pro Arg
TP84_25

TGCTTATGAATTCAAATATGTCAACAACTGAATGAGCCGATCGGATCGACACTGGTGTGGGCCGGGCA
ACGAATACTTAAGTTTATACAGTTGTTTGACTTACTCGGCTAGCCTAGCTGTGACCACTACCGGCGCCGT

100 105 110 115
Ala Tyr Glu Phe Lys Tyr Val Asn Lys Leu Asn Glu Pro Ile Gly Ser Thr Leu Val Met Gly Arg Ala
TP84_25

TTCCACAAAGGGATGCAGATGGCATCGATCAAAAAGGTGGTCGATGGCGAAATCTTATCCACTGACGATG
AAGGTGTTTCCCTACGTCTACCGTAGCTAGTTTTTCCACCAGCTACCGCTTTAGAATAGGTGACTGCTAC

120 125 130 135 140
Phe His Lys Gly Met Gln Met Ala Ser Ile Lys Lys Val Val Asp Gly Glu Ile Leu Ser Thr Asp Asp
TP84_25

TGCTGGATGTATATTCGGATGCGTTTCGATCAAGAACGTGAAAACAATGATGTGGACTGGGCCGAAGATGA
ACGACCTACATATAAGCCTACGCAAGCTAGTTCTTGCACCTTTTGTACTACACCTGACCCGGCTTCTACT

9520

145 150 155 160 165
Val Leu Asp Val Tyr Ser Asp Ala Phe Asp Gln Glu Arg Glu Asn Asn Asp Val Asp Trp Ala Glu Asp Asp
TP84_25

TCCGGCAAAGGTGAAAGATGATGGCGCGAAGCTGATGCAGAAATACTATGAGGAAATGGGGATCAATGCG
AGGCCGTTTCCACTTTCTACTACCGCGCTTCGACTACGTCTTTATGATACTCCTTACCCTAGTTACGC

9590

170 175 180 185
Pro Ala Lys Val Lys Asp Asp Gly Ala Lys Leu Met Gln Lys Tyr Tyr Glu Glu Met Gly Ile Asn Ala
TP84_25

ATTCCGATGGTTCGATGATCGTGGCCTGCCGCTTGTGCAACGTGAACACGCATTGAGATCGTGCCAGGGC
TAAGGCTACCAGCTACTAGCACCGGACGGCGAACAGCTTGCACTTGTGCGTAAGCTCTAGCACGGTCCCG

9660

190 195 200 205 210
Ile Pro Met Val Asp Asp Arg Gly Leu Pro Leu Val Glu Arg Glu His Ala Phe Glu Ile Val Pro Gly
TP84_25

TGAAAGCGAAAGCAGTCATCGACATCATCGAACAAGACGGATCAATCCGGGATTACAAAACAAGTAAGCG
ACTTTCGCTTTCGTCAGTAGCTGTAGTAGCTTGTCTGCCTAGTTAGGCCCTAATGTTTTGTTTCATTTCGC

9730

215 220 225 230 235
Leu Lys Ala Lys Ala Val Ile Asp Ile Ile Glu Gln Asp Gly Ser Ile Arg Asp Tyr Lys Thr Ser Lys Arg
TP84_25

ATCGCCATCGCAAACGATCATCGATGAAACGATCCAGATGCCAGTCTATGCGCTTGCATATCGTGACATC
TAGCGGTAGCGTTTGCTAGTAGCTACTTTGCTAGGTCTACGGTCAGATACGCGAACGTATAGCACTGTAG

9800

240 245 250 255
Ser Pro Ser Gln Thr Ile Ile Asp Glu Thr Ile Gln Met Pro Val Tyr Ala Leu Ala Tyr Arg Asp Ile
TP84_25

ACTGGCCAGGTGCAAAAAACGGTGCGCTGGATTATGCAGTGAACCTAAAGAAAGAAAAAAGATCATGC
TGACCGGTCCAGCTTTTTTGCCACCCGGACCTAATACGTCACCTGGATTCTTTCTTTTTTCTAGTACG

9870

260 265 270 275 280
Thr Gly Gln Val Glu Lys Thr Val Gly Leu Asp Tyr Ala Val Asn Leu Lys Lys Glu Lys Lys Ile Met
TP84_25

GCCTGGAAACCGATGGGCCAGTCGATGATGGCCGATCGAACGTGTAACAAACCTTCGTGGGTGTGGC
CGGACCTTTGGCTACCCGGTCAGCTACTACCGCGTAGCTTGACATTTTGTGGAAGCACCCACACCG

9940

285 290 295 300 305
Arg Leu Glu Thr Asp Gly Pro Val Asp Asp Gly Arg Ile Glu Arg Val Lys Gln Thr Phe Val Gly Val Ala
TP84_25

AAAGGCGATCAGTGCCGGGATCTTTTACCCGAATGAAGAATCGAACGCCTGCGGCTATTGTTCAATTAAG
TTTCCGCTAGTCACGGCCCTAGAAAAATGGGCTTACTTCTTAGCTTGCGGACGCCGATAACAAGTAAATTC

10 010

310 315 320 325
Lys Ala Ile Ser Ala Gly Ile Phe Tyr Pro Asn Glu Glu Ser Asn Ala Cys Gly Tyr Cys Ser Phe Lys
TP84_25

GACATCTGCAAAAAATCAAAAACATTTTAAGAAAGGGGAATTCAAATGGGATTCGATATGCAATATTTCA
CTGTAGACGTTTTTTAGTTTTTGTAAAATCTTTCCCTTAAGTTTACCCTAAGCTATACGTTATAAAGT

10 080

330 335
Asp Ile Cys Lys Lys Ser Lys Thr Phe
TP84_25

RBS

1 5
Met Gly Phe Asp Met Gln Tyr Phe
TP84_26

GCGAAAAAGAAGCACAGCAATATCAAGATGCAGTGGAAAACTCACTGAACAAAACACGATCCAGGTGGT 10 150
CGCTTTTCTTCGTGTCGTTATAGTTCTACGTACCTTTTGAGTGACTTGTTTTGTGCTAGGTCCACCA
10 15 20 25 30
Ser Glu Lys Glu Ala Gln Gln Tyr Gln Asp Ala Val Glu Lys Leu Thr Glu Gln Asn Thr Ile Gln Val Val
TP84_26

TTCACCTTAACGACATCAAACAAAACTACTGGAAGAAGTGAAGGAATTCGAATTCAGTCCACTGGAATAT 10 220
AAGTGAATTGCTGTAGTTTGTGTTTGTGATGACCTTCTTCACTTCCTTAAGCTTAAGTCAGGTGACCTTATA
35 40 45 50 55
Ser Leu Asn Asp Ile Lys Gln Lys Leu Leu Glu Glu Val Lys Glu Phe Glu Phe Ser Pro Leu Glu Tyr
TP84_26

GTGCCAGGGCCAGATGAAACGCAAAAAACCATCTTTTTAAAGATCGCTAACATCATGGGTGAACTGGACC 10 290
CACGGTCCCGGTCTACTTTGCGTTTTTGGTAGAAAAATTTCTAGCGATTGTAGTACCCACTTGACCTGG
60 65 70 75
Val Pro Gly Pro Asp Glu Thr Gln Lys Thr Ile Phe Leu Lys Ile Ala Asn Ile Met Gly Glu Leu Asp
TP84_26

GAATTCGAAATCGGGATACAACAAGCAATTCGATTACTATTATGTGACTGAATCTGATGTGGTCGGCGC 10 360
CTTAAGGCTTTAGCCCTATGTTGTTTCTTAAGCTAATGATAATACACTGACTTAGACTACCCAGCCGCG
80 85 90 95 100
Arg Ile Pro Lys Ser Gly Tyr Asn Lys Gln Phe Asp Tyr Tyr Val Thr Glu Ser Asp Val Val Gly Ala
TP84_26

GGTGCGGCCGCTGATGGCGAAATATAAACTTGTCTATGATGCCATCGCTGAAAAAATACAAAACAGAAACG 10 430
CCACGCCGGCGACTACCGCTTTATATTGAACAGTACTACGGTAGCGACTTTTTTATGTTTTGTCTTTGC
105 110 115 120 125
Val Arg Pro Leu Met Ala Lys Tyr Lys Leu Val Met Met Pro Ser Leu Lys Lys Tyr Lys Thr Glu Thr
TP84_26

ATCCAGGGCAAATACAACAAGATGATCATGGGAACGGTCGAAATCGAATGGATCGTGCGTGATGCCGAAT 10 500
TAGGTCCCGTTTATGTTGTTCTACTAGTACCCTTGCCAGCTTTAGCTTACCTAGCACGCACTACGCGTTA
130 135 140 145
Ile Gln Gly Lys Tyr Asn Lys Met Ile Met Gly Thr Val Glu Ile Glu Trp Ile Val Arg Asp Ala Glu
TP84_26

CCTATGAACAGATAAAATTCACGATGTTAGGCAAGGGCTGGACAATCTGGAAAAAGACATTTACAAAGC 10 570
GGATACTTGCTATTTTAAAGTGCTACAATCCGTTCCCGGACCTGTTAGACCTTTTCTGTAATGTTTGC
150 155 160 165 170
Ser Tyr Glu Gln Ile Lys Phe Thr Met Leu Gly Lys Gly Leu Asp Asn Leu Glu Lys Asp Ile Tyr Lys Ala
TP84_26

AATCACCGGGAATAAAAAATATGCGCTGATCACACTGTTTATGATCGACAGCGGTGACGATCCAGAACGG 10 640
TTAGTGGCCCTTATTTTATACGCGACTAGTGTGACAAATACTAGCTGTGCGCACTGCTAGGTCTTGCC
175 180 185 190 195
Ile Thr Gly Asn Lys Lys Tyr Ala Leu Ile Thr Leu Phe Met Ile Asp Ser Gly Asp Asp Pro Glu Arg
TP84_26

AACGACACACCGACTGGTGAAGAAGGGCCACAAAACGCACAGAATGGGCCAGAAAACACGAAAAAGACAC 10 710
TTGCTGTGTGGCTGACCACTTCTTCCCGGTGTTTTGCGTGTCTTACCCGGTCTTTTGTGCTTTTCTGTG
200 205 210 215
Asn Asp Thr Pro Thr Gly Glu Glu Gly Pro Gln Asn Ala Gln Asn Gly Pro Glu Asn Thr Lys Lys Thr
TP84_26

AAGGGAATAGTAAACACCAGGGCAACCGAAAAACGGCCAGGAAGTGCCAAAAAATGCGTCACAAGGGCA
10 780
TTCCCTTATCATTTTGTGGTCCCCTTGGCTTTTGGCGGTCTTCACGGTTTTTTACGCAGTGTTCCCCT
220 225 230 235 240
Gln Gly Asn Ser Lys Thr Pro Gly Gln Pro Lys Asn Gly Gln Glu Val Pro Lys Asn Ala Ser Gln Gly Gln
TP84_26

ACCGCAACAACCGCCGCAACAGCCGCCAGTCAAACAACCGACAAAAGGTGATCTATTGACACGGTGGGTG
10 850
TGGCGTTGTTGGCGGCGTTGTGCGGCGGTCAGTTTGTGGCTGTTTTCCACTAGATAACTGTGCCACCCAC
245 250 255 260 265
Pro Gln Gln Pro Pro Gln Gln Pro Pro Val Lys Gln Pro Thr Lys Gly Asp Leu Leu Thr Arg Trp Val
TP84_26

GTCTTGGCAGGGCCGGATCAAAAGAAACAGGAAGTGCGGAAGGCGTTTCGATGAATGGTATGAAAAGAAAA
10 920
CAGGACCGTCCCAGGCTAGTTTTCTTTGTCTTCACGCCTTCCGCAAGCTACTTACCATACTTTTCTTTT
270 275 280 285
Val Leu Ala Gly Pro Asp Gln Lys Lys Gln Glu Val Arg Lys Ala Phe Asp Glu Trp Tyr Glu Lys Lys
TP84_26

AGGCCGAAGGATGGGATCACATGGCCATGATGCAAGCATTGACGAAAAAACTGCATGAAAAGAACCAAAA
10 990
TCCGGCTTCTTACCCTAGTGACCGGTACTACGTTCTGTAAGTCTTTTGGCTACTTTTCTTGGTTTT
290 295 300 305 310
Lys Ala Glu Gly Trp Asp His Met Ala Met Met Gln Ala Leu Thr Lys Lys Leu His Glu Lys Asn Gln Lys
TP84_26

AGAAAAGCAGGCACAGCAACAGCAAGCAGATGGACAAAGTGGACAAGCACAGCAACAGCAAACTGATGAC
11 060
TCTTTTCGTCCGTGTCGTTGTCGTTGCTACCTGTTTCACCTGTTTCGTGTCGTTGCTGTTGACTACTG
315 320 325 330 335
Glu Lys Gln Ala Gln Gln Gln Gln Ala Asp Gly Gln Ser Gly Gln Ala Gln Gln Gln Thr Asp Asp
TP84_26

AAGTCGAAAGAATAACAATCTGGCCAGGATCATCCTGGCCTATCCAAAAAGCAGGGGGGATATGGTGGC
11 130
TTCAGCTTTCTTATTGTTAGACCGGTCCTAGTAGGACCGGATAGGTTTTTTCGTCCCCCTATACCACCG
340
Lys Ser Lys Glu A
TP84_26 RBS TP84_27
Met Val Ala

AAACTGGATCTATCAATACCATGTGTATAACGATCATATCAGTTTTATCGTGTCTATTCCGATGGAATCA
11 200
TTTGACCTAGATAGTTATGGTACACATATTGCTAGTATAGTCAAATAGCACAGATAAGGCTACCTTAGT
5 10 15 20 25
Asn Trp Ile Tyr Gln Tyr His Val Tyr Asn Asp His Ile Ser Phe Ile Val Ser Ile Pro Met Glu Ser
TP84_27

GAAGAATATGCGATCGATTTTGCAAAACGTGTCGCACTGGATTATCTTGATTATCCGCTTTACCAACTGG
11 270
CTTCTTATACGCTAGCTAAAACGTTTGCACAGCGTGACCTAATAGAACTAATAGGCGAAATGGTTGACC
30 35 40 45
Glu Glu Tyr Ala Ile Asp Phe Ala Lys Arg Val Ala Leu Asp Tyr Leu Asp Tyr Pro Leu Tyr Gln Leu
TP84_27

ATGACGTGAAATGTGAATTCATCGTCAGATTTTTTGCGAATCAGTCATCGATGTGAACAACTGCAAAA
11 340
TACTGCACTTTACACTTAAGTAGCAGTCTAAAAAACGCTTAGTCAGTAGCTACACTTGTGTTGACGTTTT
50 55 60 65 70
Asp Asp Val Lys Cys Glu Phe Ile Val Arg Phe Phe Cys Glu Ser Val Ile Asp Val Asn Lys Leu Gln Lys
TP84_27

ATAGCGGCGTGATCAAAGCGATGTGCCGATGATCAAGATCCTGGGAACAAAAGAGGGGTGACAAATCTG
TATCGCGGCACTAGTTTCGCTACACGGCATACTAGTTCTAGGACCTTGTTTTCTCCCACTGTTTAGAC

11 410

TP84_27

regulatory region

ACATCCCTAAATTATAAAAAATACCGTAAAAATAACTTTATAAGATCATAGGGGTGACACTGGTGTCA
TG TAGGGATTTAATATTTTTATGGCATTATTTATTATGAAATATTCTAGTATCCCACTGTGACCACAGT

11 480

regulatory region

CTCCTAATCAAAAAGGGGTGACACTGGTGACATCCCTGGGGTGACACTG6TGTCACTCCGGGGGGGTGAC
GAGGATTAGTTTTTCCCACTGTGACCACGTGAGGGACCCCACTGTGACCACAGTGAGGCCCCCCCACTG

11 550

AAATCTGGCACATAACAATTACTATATATACAAATACTAATTACAAATAGTTATTATATGTCACGATCCT
TTTAGACCGTGATTGTTAATGATATATATGTTTATGATTAAATGTTTATCAATAATATACAGTGCTAGGA

11 620

GTGGATCATGACAATTTTTTCAGTTTTTTGTTTACAGTTTATAACAATAAGCTATAATGAAAGTACCAA
CACCTAGTACTGTTAAAAAGTCAAAAAACAAATGTCAAATATTTGTTATTCGATATTACTTTTCATGTT

11 690

regulatory region

ACAGGAAAGGGGTGCAACGATGGAGAAGAAACAAAATTCAATCCGATCCTTTTCATTTTCAGACCGATGAA
TGTCTTTTCCCACTGTTGCTACCTCTTCTTTGTTTTTAAGTTAGGCTAGGAAAGTAAAGTCTGGCTACTT

11 760

RBS

1 5 10 15
Met Glu Lys Lys Gln Lys Phe Asn Pro Ile Leu Ser Phe Gln Thr Asp Glu
TP84_28

TCGGTGTGCGAATTCATCCAGAAAGTGATGATCAAACGAATGGAATTCACAGATCGAATATCGTGCGTG
AGCCACACGCTTAAGTAGGCTTTTCACTACTAGTTTGCTTACCTTAAGTTGTCTAGCTTATAGCACGCAC

11 830

20 25 30 35 40
Ser Val Cys Glu Phe Ile Gln Lys Val Met Ile Lys Arg Met Glu Phe Asn Arg Ser Asn Ile Val Arg
TP84_28

AGATTTTTATGATTGGCCTGGAAGAATTCAAAAACGTCATCCAGAGATTGACAAGAAGTGAAGTGAATG
TCTAAAAATACTAACCGGACCTTCTTAAGTTTTTTGCAGTAGGTCTCTAACTGTTCTTCACTTCACTTAC

11 900

45 50 55 60
Glu Ile Phe Met Ile Gly Leu Glu Glu Phe Lys Lys Arg His Pro Glu Ile Asp Lys Lys
TP84_28

RBS

GGGGAATGATCTATGTCAGATCAAGAAAAACAAGAATATCTGATGTCGGATGACATGTGGGTGTATGGAT
CCCCTTACTAGATACAGTCTAGTTCTTTTGTCTTATAGACTACAGCCTACTGTACACCCACATACCTA

11 970

RBS

1 5 10 15
Met Ser Asp Gln Glu Lys Gln Glu Tyr Leu Met Ser Asp Asp Met Trp Val Tyr Gly
TP84_29

ATAGTGTCATCGACAATGGTGTGATTTATTCGAAAGAAATCACTGATTCCGAATTCGCACATATTGTGT
TATCACAGTAGCTGTTACCACACTAAATAAGCTTTCTTTAGTGACTAAGGCTTAAGGCGTGTATAACACA

12 040

20 25 30 35 40
Tyr Ser Val Ile Asp Asn Gly Val Ile Tyr Ser Lys Glu Ile Thr Asp Ser Glu Phe Arg Thr Tyr Cys Val
TP84_29

CATCCGATCACTGGTGAATGAACGAAAAGCAGTGGCCTGGCCATCCTATGAACTATTGCAGAATTATCC
GTAGGCTAGTGACCACTTACTTGCTTTTCGTCACCGGACCGGTAGGATACTTTGATAACGTCTTAATAGG

Ile Arg Ser Leu Val Asn Glu Arg Lys Ala Val Ala Trp Pro Ser Tyr Glu Thr Ile Ala Glu Leu Ser
TP84_29

GGGCATTCCAAACGAACTGCGATGCGAAATGTGGCCAGATTGATCGAACTGGATCTGATCGAAAAACGGC
CCCGTAAGGTTTGCTTGACGCTACGCTTTACACCGGTCTAACTAGCTTGACCTAGACTAGCTTTTTGCGG

Gly His Ser Lys Arg Thr Ala Met Arg Asn Val Ala Arg Leu Ile Glu Leu Asp Leu Ile Glu Lys Arg
TP84_29

CACGATCTGGAACGTCAAACGAATTTGTGTTGAAAAAACTGCAAAATTCCAAAGTGTTAAAAACAAACA
GTGCTAGACCTTGACGTTTGCTTAAACACCACTTTTTTGACGTTTAAAGGTTTCACAATTTTTGTTTGT

Pro Arg Ser Gly Thr Ser Asn Glu Phe Val Val Lys Leu Gln Asn Ser Lys Val Leu Lys Asn Lys Gln
TP84_29

AGACATCTGGATTATATAGAGAAATGCCGGGATGATGAACCGAAAAAGCCACTGATCCTGGTGAAAAA
TCTGTAGGACCTAATATATCTCTTACGGCCCTACTACTTGGCTTTTTTCGGTGACTAGGACCACTTTTT

Asp Ile Leu Asp Tyr Ile Glu Lys Cys Arg Asp Asp Glu Pro Lys Lys Ala Thr Asp Pro Gly Glu Lys
TP84_29

GTGGACAAGCCGGAAAAAGCCGATCCGATCCCTTACAAAGAGATCATCGACTATCTAAACGAAAAAGCTG
CACCTGTTTCGGCCTTTTTCGGCTAGGCTAGGGAATGTTTCTCTAGTAGCTGATAGATTGCTTTTTTCGAC

Val Asp Lys Pro Glu Lys Ala Asp Pro Ile Pro Lys Lys Glu Ile Ile Asp Lys Leu Asn Glu Lys Ala
TP84_29

GAACGAAATACAGCCACACTGGATCGGCGAATCAAAAATAATCAAGGCACGATGGAACGAAATGGCCAA
CTTGCTTTATGTCGGTGTGACCTAGCCGCTTAGTTTTGATTAGTTCCGTGCTACCTTGCTTTACCGGTT

Gly Thr Lys Tyr Ser His Thr Gly Ser Ala Asn Gln Lys Leu Ile Lys Ala Arg Trp Asn Glu Met Ala Lys
TP84_29

PmeI

AATAACAAAGACCGTGACTGGATCGTGCGCAGTTTAAACACGTGATCGATGTCAAACAGCACAAATGG
TTATTTGTTTCTGGCACTGACCTAGCACCGCTCAAATTTGTGCACTAGCTACAGTTTGTCTGTTTACC

Ile Asn Lys Asp Arg Asp Trp Ile Val Ala Gln Phe Lys His Val Ile Asp Val Lys Thr Ala Gln Trp
TP84_29

AAGGGTACAGAATGGGAAAAATATCTGCGGCCATCCACGTTATTCGGCAATAAATTTGATCAATACAGAA
TTCCCATGTCTTACCCTTTTTATAGACGCCGGTAGGTGCAATAAGCCGTTATTTAACTAGTTATGTCTT

Lys Gly Thr Glu Trp Glu Lys Tyr Leu Arg Pro Ser Thr Leu Phe Gly Asn Lys Phe Asp Gln Tyr Arg
TP84_29

ACGAATCGCCGAATCACAAGCCAGTCGGTGGACAAAACAGACCAAAAGAGGACATCCGAACAAGACTGGC
TGCTTAGCGGCTTAGTGTTCGGTCAGCCACCTGTTTTGTCTGGTTTTCTCTGTAGGCTTGTCTGACCG

Asn Glu Ser Pro Asn His Lys Pro Val Gly Glu Gln Asn Arg Pro Lys Glu Asp Ile Arg Thr Arg Leu Ala
TP84_29

AGGATTCCTGGCTGATGAAGATTAAAGAGATATAAAAGGGGAAAGGGGCGATCGATCATGACAAAAACAG
TCCTAAGGACCGACTACTTCTAATTTCTCTATATTTTCCCTTTCCCCGCTAGCTAGTACTGTTTTTGTCT

255 260
Gly Phe Leu Ala Asp Glu Asp
TP84_29

RBS

1
Met Thr Lys Thr
TP84_30

AAATGATCACTTTGCTTGAACAGATCGAATCGTTTTTCCAGGCCGCTTCACAGCAGACGAAAAACAGC
TTTACTAGTGAAACGAACTTGTCTAGCTTAGCAAAAAAGGTCCGGCGAAGTGTCGTCTGCTTTTTTGTCTG

5 10 15 20 25
Glu Met Ile Thr Leu Leu Glu Gln Ile Glu Ser Phe Phe Pro Gly Arg Phe Thr Ala Asp Glu Lys Thr Ala
TP84_30

GGCGGTCTGGTGGGAAGTGTTAAAGGATTATGATTTTGAACACTGCAAAAAGAATCTGATCAATCACGTC
CCGCCAGACCACCTTCACAATTTCTAATACTAAACTTGTGACGTTTTTCTTAGACTAGTTAGTGCAG

30 35 40 45 50
Ala Val Trp Trp Glu Val Leu Lys Asp Tyr Asp Phe Glu His Cys Lys Lys Asn Leu Ile Asn His Val
TP84_30

AAGTCATCCGAATTTCCCCAACGATCGCCAATCTGATCGCAGGATCAAAAGACACAAGCCGGAACACTATT
TTCAGTAGGCTTAAAGGGGTTGCTAGCGGTTAGACTAGCGTCTAGTTTTCTGTGTTCCGGCCTTGATAA

55 60 65 70
Lys Ser Ser Glu Phe Pro Pro Thr Ile Ala Asn Leu Ile Ala Gly Ser Lys Asp Thr Ser Arg Asn Tyr
TP84_30

CAGATGATCTGATCGAAAAGATCGAAAATGTACCAGAAATAGATGAAATCGTTGCAGAAATAGAACGGCG
GTCTACTAGACTAGCTTTTCTAGCTTTTACATGGTCTTTATCTACTTTAGCAACGTCTTTATCTTGCCGC

75 80 85 90 95
Ser Asp Asp Leu Ile Glu Lys Ile Glu Asn Val Pro Glu Ile Asp Glu Ile Val Ala Glu Ile Glu Arg Arg
TP84_30

CACAGGAAGAAAGGTGATGTGGTAATGATCTATAATATCGATGCAGAACAAATATTTTCTTGGATGTTTAA
GTGTCCTTCTTTCCACTACACCATTACTAGATATTATAGCTACGCTCTTGTATAAAAGAACCTACAAATT

RBS

5 10 15
Met Ile Tyr Asn Ile Asp Ala Glu Gln Tyr Phe Leu Gly Cys Leu
TP84_31

100 105
Thr Gly Arg Lys Val Met Trp
TP84_30

TGCTTGAACCAGGGCTGGTCAAAGAAACAAAGATAAGACCAGTGTCATCTATCTGAACGACATCGAATCAT
ACGAACTTGGTCCCGACCACTTTCTTTGTTTCTATTCTGGTCACGTAGATAGACTTGCTGTAGCTTAGTA

20 25 30 35
Met Leu Glu Pro Gly Leu Val Lys Glu Thr Lys Ile Arg Pro Val His Leu Ser Glu Arg His Arg Ile Ile
TP84_31

CCTGGATGCGATTCTAAAAGCAGATGAAAAGGGAAACACGCCGGACATCGTGGCCATTGTCACTGAAATG
GGACCTACGCTAAGATTTTCTGCTACTTTTCCCTTTGTGCGGCCGTAGCACCAGGTAACAGTGACTTTAC

40 45 50 55 60
Leu Asp Ala Ile Leu Lys Ala Asp Glu Lys Gly Asn Thr Pro Asp Ile Val Ala Ile Val Thr Glu Met
TP84_31

GGCGATGATAAAGTGGC6GCGATCGGCGGCGTGCTCTATCTTACTGAATTGATCGGTGCAGTGCCTACCA
CCGCTACTATTTACCCGCCGCTAGCCGCCGACAGGATAGAATGACTTAAGTCCACGTCACGGATGGT

65 70 75 80 85
Gly Asp Asp Lys Val Ala Ala Ile Gly Gly Val Ser Tyr Leu Thr Glu Leu Ile Gly Ala Val Pro Thr
TP84_31

CATCAAATGTGAAGATGTATGAACAATTTATTATCGATGCATACAAACTGCGTCTGGCAACCGAAGCGGC
GTAGTTTACACTTCTACATACTTGTTAAATAATAGCTACGTATGTTTGACGCAGACCGTTGGCTTCGCCG

90 95 100 105
Thr Ser Asn Val Lys Met Tyr Glu Gln Phe Ile Ile Asp Ala Tyr Lys Leu Arg Leu Ala Thr Glu Ala Ala
TP84_31

GCAAAGATTTATCGAGGATGTAAGAGGAAGCCGCGATCCGATCCTGGTGAACCAACTGGCACAAAAAATG
CGTTTCTAAATAGCTCCTACATTCTCCTTCGGCGCTAGGCTAGGACCACTTGGTTGACCGTGTTTTTAC

110 115 120 125 130
Gln Arg Phe Ile Glu Asp Val Arg Gly Ser Arg Asp Pro Ile Leu Val Asn Gln Leu Ala Gln Lys Met
TP84_31

GTCGAGATCCAGGAATATGGAAGTATTGTGAAAAGAAGATCCTGGCGTGACATCCTTGTCGATCTGGTG
CAGCTCTAGGTCCTTATACCTTCATAACACTTTTCTTCTAGGACCGCACTGTAGGAACAGCTAGACCACC

135 140 145 150 155
Val Glu Ile Gln Glu Tyr Gly Ser Ile Val Lys Arg Arg Ser Trp Arg Asp Ile Leu Val Asp Leu Val
TP84_31

AAGAATCATATCAAGAAAAAATCGGATTGACTGGAATCAACACTGGCTTTGCTGGATTGAATCGGATGAC
TTCTTAGTATAGTTCTTTTTTAGCCTAACTGACCTTAGTTGTGACCGAAACGACCTAACTTAGCCTACTG

160 165 170 175
Glu Glu Ser Tyr Gln Glu Lys Ile Gly Leu Thr Gly Ile Asn Thr Gly Phe Ala Gly Leu Asn Arg Met Thr
TP84_31

AGACGGTTTGCAAGGGAACAATCTGATCATCATCGGTGCGCGCCATCGATGGGTAAAACGGCATTGCA
TCTGCCAAACGTTCCCTTGTTAGACTAGTAGTACGCACGCGCGGTAGCTACCCATTTTGCCGTAAGCGT

180 185 190 195 200
Asp Gly Leu Gln Gly Asn Asn Leu Ile Ile Ile Gly Ala Arg Pro Ser Met Gly Lys Thr Ala Phe Ala
TP84_31

TTGAACATCGGCACAAATGCTTGACATTAGATGATGCAAGGGTGGACATCTTTTCACTGGAACGCGCG
AACTTGTAGCCGTGTTTACGAACGTGTAATCTACTACGTTCCACCTGTAGAAAAGTGACCTTTGCGGCC

205 210 215 220 225
Leu Asn Ile Gly Thr Asn Ala Cys Thr Leu Asp Asp Ala Arg Val Asp Ile Phe Ser Leu Glu Thr Pro
TP84_31

AAAAACGTCTGGCCAAACGGATCATCGCATCGGTTGGCAATATCAGTGCTGAACGACTGCGGCGAATGGA
TTTTTGACAGACCGGTTTGCTAGTAGCGTAGCCAACCGTTATAGTCACGACTTGCTGACGCGCGTTACCT

230 235 240 245
Glu Lys Arg Leu Ala Lys Arg Ile Ile Ala Ser Val Gly Asn Ile Ser Ala Glu Arg Leu Arg Arg Met Asp
TP84_31

TTTTGACGATAAAACACGTGAACGATTTATTCAAAGGTCGGCCTGGTGGATTGCTTTGACTTGACATT
AAAACGCTATTTTGTGCACTTGCTAAATAAGTTTTCCAGCCGGACCACCTAAGCAAACCTGAACGTGTAA

250 255 260 265 270
Phe Asp Asp Lys Thr Arg Glu Arg Phe Ile Gln Lys Val Gly Leu Val Asp Ser Phe Asp Leu His Ile
TP84_31

CATGATCAATCGACAATCACTGTCGAGGAAATCCGATCGATCGTGGCCGAAAGCAATCGCCAGGCACGGA
GTACTAGTTAGCTGTTAGTGACAGCTCCTTTAGGCTAGCTAGCACCGGCTTTTCGTTAGCGGTCCGTGCCT

275 280 285 290 295
His Asp Gln Ser Thr Ile Thr Val Glu Glu Ile Arg Ser Ile Val Ala Glu Ser Asn Arg Gln Ala Arg
TP84_31

AAGAAGGCAAAAAACATCTGGTCATCATCGACTATCTGCAACTGATCACATATCGTGGGCGGCTACAAAA
14 000
TTCTTCCGTTTTTTGTAGACCAGTAGTAGCTGATAGACGTTGACTAGTGATAGCACCCGGCGATGTTTT
300 305 310 315
Lys Glu Gly Lys Lys His Leu Val Ile Ile Asp Tyr Leu Gln Leu Ile Thr Tyr Arg Gly Pro Leu Gln Asn
TP84_31

CAAGGTCCAACAAATTGGCCATATATCCAGACAATTGAAACTGATGGCCGGGGATTTTGACATCCCGATC
14 070
GTTCCAGGTTGTTTAACCGGTATATAGGTCTGTTAACTTTGACTACCGGCCCTAAAACGTAGGGCTAG
320 325 330 335 340
Lys Val Gln Gln Ile Gly His Ile Ser Arg Gln Leu Lys Leu Met Ala Gly Asp Phe Asp Ile Pro Ile
TP84_31

ATCGCACTGTCGCAATTATCGCGCGGTGTTGAACAACGACAAGACAAACGGCCTGTGATGGCTGATCTGC
14 140
TAGCGTGACAGCGTTAATAGCGCGCCACAACCTGTTGCTGTTCTGTTTGCCGGACACTACCGACTAGACG
345 350 355 360 365
Ile Ala Leu Ser Gln Leu Ser Arg Gly Val Glu Gln Arg Gln Asp Lys Arg Pro Val Met Ala Asp Leu
TP84_31

GTGAATCGGGTGACATCGAACAAGATGCCGACATCATCGCTTTTCTTTACCGGGATGACTATTATGACAA
14 210
CACTTAGCCCACTGTAGCTTGTTCTACGGCTGTAGTAGCGAAAAGAAATGGCCCTACTGATAATACTGTT
370 375 380 385
Arg Glu Ser Gly Asp Ile Glu Gln Asp Ala Asp Ile Ile Ala Phe Leu Tyr Arg Asp Asp Tyr Tyr Asp Lys
TP84_31

AGAATCTGACAAAAAGAACATCACGGAAGTCATCATCGCTAAAAACCGGGATGGCGCAGTTGGTACAGTG
14 280
TCTTAGACTGTTTTCTTGAGTGCCCTTCAGTAGTAGCGATTTTGGCCCTACCGCGTCAACCATGTCAC
390 395 400 405 410
Glu Ser Asp Lys Lys Asn Ile Thr Glu Val Ile Ile Ala Lys Asn Arg Asp Gly Ala Val Gly Thr Val
TP84_31

GAACTGGCGTTTCATTAAAGAATACAACAAATTCCTGGATCTGGAACGCCGATTTCGATGAAGCATCATGAT
14 350
CTTGACCGCAAGTAATTTCTTATGTTGTTTAAAGACCTAGACCTTGCGGCTAAGCTACTTCGTAGTACTA
415 420 425 430 435
Glu Leu Ala Phe Ile Lys Glu Tyr Asn Lys Phe Leu Asp Leu Glu Arg Arg Phe Asp Glu Ala Ser
TP84_31

TCTACATACAGCTTATCCAAACAAGCCCCAAAAATATTCATAAAAAATAAAGAAAGGGAGAATGAAGAATG
14 420
AGATGTATGTCGAATAGGTTTGTTGCGGTTTTTATAAGTATTTTTATTCTTTCCCTCTTACTTCTTAC

RBS Met
TP84_32

CAAGTGATCTCGAAACGATTTCAGATTCAATCAATACTGGGAATTCGTGATGGTCGATGGGCAATTCTACA
14 490
GTTCACTAGAGCTTTGCTAAGTCTAAGTTAGTTATGACCCTTAAGCACTACCAGCTACCCGTTAAGATGT
5 10 15 20
Gln Val Ile Ser Lys Arg Phe Arg Phe Asn Gln Tyr Trp Glu Phe Val Met Val Asp Gly Gln Phe Tyr
TP84_32

TCTTGGCGAAAGACATCCTGGGGTATTTGCGATATAAACGGACCAGTCACAAAGTGATCACCGAATTACT
14 560
AGAACCGCTTTCTGTAGGACCCCATAAAGCCTATATTTGCGCTGGTCAGTGTTTCACTAGTGGCTTAATGA
25 30 35 40 45
Ile Leu Ala Lys Asp Ile Leu Gly Tyr Phe Gly Tyr Lys Arg Thr Ser His Lys Val Ile Thr Glu Leu Leu
TP84_32

CACTAATGATGATCGATTGATCATCAAAGGCAAAGCAACACACGACATGAAGCATGTCGAGAAAGGACAA
GTGATTACTACTAGCTAACTAGTAGTTTCCGTTTCGTTGTGTGCTGTACTTCGTACAGCTCTTTCCTGTT

14 630

50 55 60 65 70
Thr Asn Asp Asp Arg Leu Ile Ile Lys Gly Lys Ala Thr His Asp Met Lys His Val Glu Lys Gly Gln
TP84_32

CATGCGTTATTGATCACAAAGGCCGGGCTGATGACGTTTGTTCACAACCGATTAAAAGATGACAAGCGAG
GTACGCAATAACTAGTGTTTCCGGCCCCGACTACTGCAAACAAATGTTGGCTAATTTTCTACTGTTGCTC

14 700

75 80 85 90
His Ala Leu Leu Ile Thr Lys Ala Gly Leu Met Thr Phe Val Tyr Asn Arg Leu Lys Asp Asp Lys Arg
TP84_32

AAAAAGGGATCGAATTGATCCACTGGATCAACGAAAAAGCAATCCCGGAAATGAAAGAGGAATGGGAAAA
TTTTTCCCTAGCTTAAGTAGGTGACCTAGTTGCTTTTCGTTAGGGCCTTACTTTCTCCTTACCCTTTT

14 770

95 100 105 110 115
Glu Lys Gly Ile Glu Leu Ile His Trp Ile Asn Glu Lys Ala Ile Pro Glu Met Lys Glu Glu Trp Glu Lys
TP84_32

ATTTTTAAAAATCAATGGAAAAGACAGCGTTTTCGAGGAACGGTTTAGCACGGATGATCAAAAACATCCA
TAAAAATTTTGTAGTTACCTTTTCTGTCGCCAAAGCTCCTTGCCAAATCGTGCTACTAGTTTTTGTAGGT

14 840

120 125 130 135 140
Phe Leu Lys Ile Asn Gly Lys Asp Ser Gly Phe Glu Glu Arg Phe Ser Thr Asp Asp Gln Lys His Pro
TP84_32

GATTTATTTGATATGGCACTGGAACAACAAATTGAAGAATTGAAATGCATTCAAGAAGTACAAGAAAAAA
CTAAATAAACTATACCGTGACCTTGTTGTTTAACTTCTTAACTTTACGTAAGTTCTTCATGTTCTTTTTT

14 910

145 150 155 160
Asp Leu Phe Asp Met Ala Leu Glu Gln Gln Ile Glu Glu Leu Lys Cys Ile Gln Glu Val Gln Glu Lys
TP84_32

CGATCCGCCGCTTGCAAGAAAAACTGGCGATGGCCGAAAAGCGCATCGAAGCACTGGAAAAGGCACTGAA
GCTAGGCGGCGAACGTTCTTTTTGACCGCTACCGGCTTTTCGCGTAGCTTCGTGACCTTTTCCGTGACTT

14 980

165 170 175 180 185
Thr Ile Arg Arg Leu Gln Glu Lys Leu Ala Met Ala Glu Lys Arg Ile Glu Ala Leu Glu Lys Ala Leu Lys
TP84_32

ACAAGAGAAGTTAGAAAAAGCGGTCGTGAATGGCGAACGCAAGCTGATCGATGTGGACCGATTGATCAAC
TGTTCTCTTCAATCTTTTTCGCCAGCACTTACCGCTTGCCTTCGACTAGCTACACCTGGCTAACTAGTTG

15 050

190 195 200 205 210
Gln Glu Lys Leu Glu Lys Ala Val Val Asn Gly Glu Arg Lys Leu Ile Asp Val Asp Arg Leu Ile Asn
TP84_32

TATCTGAAAGGTGTGGATGCAAGTGGGCATATTCGATGAAGTGAACGCAGTGCCCAAACCAAAACACAAG
ATAGACTTTCCACACCTACGTTACCCGTATAAGCTACTTCACTTGCGTACGCGGTTTGGTTTTGTGTTT

15 120

215 220
Tyr Leu Lys Gly Val Asp Ala Ser Gly His Ile Arg
TP84_32

RBS

1 5 10 15
Met Gln Val Gly Ile Phe Asp Glu Val Asn Ala Val Pro Lys Pro Lys His Lys
TP84_33

CGGCGGACACCAAAGCGGAAATCCAGGGGCGAATTCAGCCAGGAAACGATTCGGGCGATCTGGGAACTGTG
GCCGCTGTGGTTTCGCTTTAGGTCCCECTTAAGTCGGTCCTTTGCTAAGCCGCTAGACCCTTGACAC

15 190

20 25 30 35 40
Arg Arg Thr Pro Lys Arg Lys Ser Arg Gly Glu Phe Ser Gln Glu Thr Ile Arg Ala Ile Trp Glu Arg
TP84_33

ACGGCTATCGATGCGTGAAATGCGGATCATATCACCTGGAAAAAGTGCCGCACCATGTGATCTATAAAAG
15 260
TGCCGATAGCTACGCACCTTTACGCCTAGTATAGTGGACCTTTTTCACGGCGTGGTACACTAGATATTTTC
45 50 55 60 65
Asp Gly Tyr Arg Cys Val Lys Cys Gly Ser Tyr His Leu Glu Lys Val Pro His His Val Ile Tyr Lys Ser
TP84_33

CCAGGGCGGGTCGTG6CACAAAGCGGAACGGTGTGACTGTTTGCCGTGGATGCCACGACTGGGCACACGGT
15 330
GGTCCC GCCAGCACCGTGTTCGCCTTGCCACACTGACAAACGGCACCTACGGTGCTGACCCGTGTGCCA
70 75 80 85
Gln Gly Gly Arg Gly Thr Lys Arg Asn Gly Val Thr Val Cys Arg Gly Cys His Asp Trp Ala His Gly
TP84_33

TTGAAACCAGGTCCATTTCGGAGAACC GCCAGAAAGATGGCCGCAAATGGTTTGAAAAC TGGGTGGAAACAT
15 400
AACTTTGGTCCAGGTAAAGCCTCTTGCGGCTCTTCTACCGGCGTTTACCAAAC TTTTGACCCACCTTTGTA
90 95 100 105 110
Leu Lys Pro Gly Pro Phe Gly Glu Pro Pro Glu Asp Gly Arg Lys Trp Phe Glu Asn Trp Val Glu Thr
TP84_33

ACTTGGATGAAAACGGTGATCTGATCCGAAAGGGGGATCATCGTGGCAAGATTCGTGACTTTTTCTGGG
15 470
TGAACCTACTTTTGCCACTAGACTAGGCTTTCCCCCTAGTAGCACC GTTCTAAGCACTGAAAAAGGACCC
115 120 125 130 135
Tyr Leu Asp Glu Asn Gly Asp Leu Ile Arg Lys Gly Asp His Arg Gly Lys Ile Arg Asp Phe Phe Leu Gly
TP84_33

AGAATGGGAAACAGGTGATCAGCATCTATGAGGAAGGGATGAAACGAAAGCGGATCATCCAGGAACAAGA
15 540
TCTTACCCTTTGTCCACTAGTGCTAGATACTCCTTCCCTACTTTGCTTTGCGCTAGTAGGTCCTTGTCT
140 145
Glu Trp Glu Thr Gly Asp His Asp Leu
TP84_33

RBS
1 5 10 15
Met Ile Thr Ile Tyr Glu Glu Gly Met Lys Arg Lys Arg Ile Ile Gln Glu Gln Asp
TP84_34

TCGGCCGTTTCTTATCCAGGCGATGCATAAAACAAATGATGGCCGCTATGAGTGGGCCAACGAAAAACCA
15 610
AGCCGGCAAAGAATAGGTCCGCTACGTATTTTGTCTTACTACCGCGATACTCACC CGGTTGCTTTTGGT
20 25 30 35 40
Arg Pro Phe Leu Ile Gln Ala Met His Lys Thr Asn Asp Gly Arg Tyr Glu Trp Ala Asn Glu Lys Pro
TP84_34

AAATACATGAAAGGGGAATAAGAACAATGAACAATGTGACGTTAGTGGGAAGATTGACGAACGATGTGGA
15 680
TTTATGTACTTTCCCTTATTCTTGTTACTTGTTACACTGCAATCACCCCTTCTAACTGCTTGCTACACCT
45 5 10 15
Lys Tyr Met Lys Gly Glu Met Asn Asn Val Thr Leu Val Gly Arg Leu Thr Asn Asp Val Glu
TP84_34 TP84_35

RBS

ATTGAGATACACACAACAAGGGAAAGCAGTGGCCACATTCAACCTGGCAGTGCAAAGAGAATTCAAAAAC
15 750
TAACTCTATGTGTGTGTTCCCTTTTCGTACC GGTGTAAGTTGGACCGTCACGTTTCTCTAAAGTTTGT
20 25 30 35
Leu Arg Tyr Thr Gln Gln Gly Lys Ala Val Ala Thr Phe Asn Leu Ala Val Gln Arg Glu Phe Lys Asn
TP84_35

CAGGATAATGTTTATGAAGTGGACTTTCCACAGATTGTCGTCTGGGGCAAGCCTGCGGAAACGTTGGCCA
15 820
GTCTATTACAAATACTTCACCTGAAAGGTGTCTAACAGCAGACCCCGTTGCGACGCCTTTGCAACCGGT

40 45 50 55 60
Gln Asp Asn Val Tyr Glu Val Asp Phe Pro Gln Ile Val Val Trp Gly Lys Pro Ala Glu Thr Leu Ala
TP84_35

ACTACACCAAAAAAGGCAGTTTAATCGGGATCACAGGTCGATTGCAAACACGGTCCTATGAAAACAGGA
15 890
TGATGTGGTTTTTCCGTCAAATTAGCCCTAGTGTCCAGCTAACGTTTGTGCCAGGATACTTTCTGTCTCT

65 70 75 80 85
Asn Tyr Thr Lys Lys Gly Ser Leu Ile Gly Ile Thr Gly Arg Leu Gln Thr Arg Ser Tyr Glu Arg Gln Asp
TP84_35

TGGCAGTCGTGTTTATGTGACGGAAGTGGTGGCTAACAATGTACGGATCTATCAATGGAAGATAGTGGA
15 960
ACCGTCAGCACAAATACACTGCCTTCACCACCGATTGTTACATGCCTAGATAGTTACCTTTCTATCACCT

90 95 100 105
Gly Ser Arg Val Tyr Val Thr Glu Val Val Ala Asn Asn Val Arg Ile Tyr Gln Trp Lys Asp Ser Gly
TP84_35

CAAGGTGGACAAAGCAGACAGGGCGGATCACAACCACAAGGCCGTGGATCACGTTATCCAGATCCTTATG
16 030
GTTCCACCTGTTTCGTCTGTCCCGCCTAGTGTGGTGTTCGGCACCTAGTGCAATAGGTCTAGGAATAC

110 115 120 125 130
Gln Gly Gly Gln Ser Arg Gln Gly Gly Ser Gln Pro Gln Gly Arg Gly Ser Arg Tyr Pro Asp Pro Tyr
TP84_35

GTCCATATGACGGCGATCCATTTGCAAACGATGGCGAACCGATTTATGTGAATGATGACGATCTGCCATT
16 100
CAGGTATACTGCCGCTAGGTAAACGTTTGCTACCGCTTGGCTAAATACACTTACTACTGCTAGACGGTAA

135 140 145 150 155
Gly Pro Tyr Asp Gly Asp Pro Phe Ala Asn Asp Gly Glu Pro Ile Tyr Val Asn Asp Asp Asp Leu Pro Phe
TP84_35

TTAGGCAGTGAATGAACAGGGGGTGATCATGGTGAAATATTTGTCGGCAGTGATGTTTATTTTATCGCA
16 170
AATCCGTCACCTTACTTGTCCCCCACTAGTACCACTTTATAAACAGCCGTCCTACAAGTAAAAATAGCGT

5 10
Met Val Lys Tyr Phe Val Gly Ser Asp Val His Phe Ile Ala
TP84_35 RBS TP84_36

TTGATCAAAACCTTTTCAAACACGACAAATGTTTTACATCAAACGATCTGATCAAATTTATAGCGGCA
16 240
AAGCTAGTTTTGGAAAAGTTTGCTGCTTTACAAAAGTGTAAGTTGCTAGACTAGTTTAAATATCGCCGT

15 20 25 30 35
Phe Asp Gln Asn Leu Phe Lys His Asp Lys Cys Phe His Ile Lys Arg Ser Asp Gln Ile Tyr Ser Gly
TP84_36

TAACGATTTATAAAGTCGCCGGATCACAATCGCCGACTATCTGCACGACAGGATCATGAACATGTATAT
16 310
ATTGCTAAATATTTACGCGCCTAGTGTGTTAGCGGCTGATAGACGTGCTGTCTAGTACTTGTACATATA

40 45 50 55 60
Ile Thr Ile Tyr Lys Val Ala Gly Ser Gln Ile Ala Asp Tyr Leu His Asp Arg Ile Met Asn Met Tyr Ile
TP84_36

CAAAGGTGACATCAATCTGGAAGTTATGCCGGAAAGCCTACTTTACAAAACCATATAGGCCAGGGATCAT
16 380
GTTTCCACTGTAGTTAGACCTTCAATACGGCCTTTCGGATGAAATGTTTGGTATATCCGGTCCCTAGTA

65 70 75 80
Lys Gly Asp Ile Asn Leu Glu Val Met Pro Glu Ser Leu Leu Tyr Lys Thr Ile
TP84_36

CCTGGCCTGGAAAGGGGTGGACAACATGGACAACGAACGCAAAAAAGGACCAATCGAAGTGGCGTTTCGAT
GGACCGGACCTTTCCCCACCTGTTGTACCTGTTGCTTGCCTTTTCTCTGGTTAGCTTCACCGCAAGCTA

16 450

RBS

RBS

1 5 10 15
Met Asp Asn Glu Arg Lys Lys Gly Pro Ile Glu Val Ala Phe Asp
TP84_37

TGGACATTGGCGATCATTGCACTGGCCACGATCATCGGCCTGGCCTGGAAATACATTTTAAACCATTAT
ACCTGTAACCGCTAGTAACGTGACCGGTGCTAGTAGCCGGACCGGACCTTTATGTAAATTTTGGTAATA

16 520

20 25 30 35
Trp Thr Leu Ala Ile Ile Ala Leu Ala Thr Ile Glu Leu Ala Trp Lys Tyr Ile Leu Lys Pro Leu
TP84_37

TCCTGGGGTGATAAGCATGGATATTCGTGAAGAAATCTATCAAGCGCAAAAGGAACGTGAAGAATTCGAA
AGGACCCCACTATTCGTACCTATAAGCACTCTTTAGATAGTTCGCGTTTTCTTGCCTTCTTAAGCTT

16 590

40
Phe Leu Gly
TP84_37

1 5 10 15
Met Asp Ile Arg Glu Glu Ile Tyr Gln Ala Gln Lys Glu Arg Glu Glu Phe Glu
TP84_38

RBS

AAAGCGCGCGCGCATCGATGCCGGATGAAGTCGAAAGAAAAATCGTGGAACGTGATGAATGATCCACGCA
TTTCGCGCGCGCGGTAGCTACGGCCTACTTCAGCTTTCTTTTATGACACCTTGACTACTTACTAGGTGCGT

16 660

20 25 30 35 40
Lys Ala Arg Ala Ala Ser Met Pro Asp Glu Val Glu Arg Lys Ile Val Glu Leu Met Asn Asp Pro Arg
TP84_38

TTCAAGCGGAAGTGATGGCAATCATTATTCATCTAAAAAACAAAAACGAAAAGACATCGAGATCTTGAC
AAGTTCGCCTTCACTACCGTTAGTAATAAGTAGATTTTTTTGTTTTGCTTTTCTGTAGCTCTAGAAGTG

16 730

45 50 55 60 65
Ile Gln Ala Glu Val Met Ala Ile Ile Ile His Leu Lys Lys Gln Lys Arg Lys Asp Ile Glu Ile Leu Thr
TP84_38

GATCGGTCTGATGGTCACGAACATGGCAGCAAGCGTCATGGAACAAGCGGCGCAAAACCAACAGAAAGGG
CTAGCCAGACTACCACTGCTTGTACCGTGCTTCGCAGTACCTTGTTCCGCGCGTTTGTTTGTCTTTCCC

16 800

70 75 80 85
Ile Gly Leu Met Val Thr Asn Met Ala Arg Ser Val Met Glu Gln Ala Ala Gln Thr Lys Gln Lys Gly
TP84_38

RBS

GAATGATCATGGGTGTGCATTTGCCGAAACACGCCACAGTGACGATCGATGACATCGTGATCGATTTACC
CTTACTAGTACCCACACGTAAACGGCTTTGTGCGGTGTCACTGCTAGCTACTGTAGCACTAGCTAAATGG

16 870

90
Glu
TP84_38

1 5 10 15 20
Met Gly Val His Leu Pro Lys His Ala Thr Val Thr Ile Asp Asp Ile Val Ile Asp Leu Pro
TP84_39

RBS

ATTCGAAATCGAACTGGATTATCAAAAAAGCAAAGCGGAAGATCCAGAAATCACCGAACTGAATCTGCCG
TAAGCTTTAGCTTGACCTAATAGTTTTTTCGTTTCGCCTTCTAGGCTTTAGTGGCTTGACTTAGACGGC

16 940

25 30 35 40
Phe Glu Ile Glu Leu Asp Tyr Gln Lys Ser Lys Ala Glu Asp Pro Glu Ile Thr Glu Leu Asn Leu Pro
TP84_39

AAATCGATCGAAAAAGAATACAACGTCATAAAGATCAAGGAAGGGGAATGAACAATGGAATTCAAACCTGT
TTTAGCTAGCTTTTCTTATGTTGCAGTATTTCTAGTTCCTTCCCCCTACTTGTGTACCTTAAGTTTGACA

45 50 55 60 1 5
Lys Ser Ile Glu Lys Glu Tyr Asn Val Ile Lys Ile Lys Glu Gly Glu Met Glu Phe Lys Leu
TP84_39 TP84_40
RBS

ATAAGGTCGCAGGATATGAAGAAGCGATCATGTCAGTGCAGGATGTCGAAAGGCAAATATTTTCTGGGA
TATTCAGCGCTCTATACTTCTTCGCTAGTACAGTGACGCCTACAGCTTTCGTTTATAAAAAGGACCT

10 15 20 25
Tyr Lys Val Ala Gly Tyr Glu Glu Ala Ile Met Ser Leu Arg Met Ser Lys Gly Lys Tyr Phe Ser Trp Glu
TP84_40

ACGAGCCAAAAAATTCAGCATCTTGTCTATGCGGTGACAGATCACCGCGGATTTCATTGCGCCGCCGCAA
TGCTCGGTTTTTTAAGTCGTAGAACAGATACGCCACTGTCTAGTGGCGCCTAAGTAACGCGCGCGCTT

30 35 40 45 50
Arg Ala Lys Lys Ile Gln His Leu Val Tyr Ala Val Thr Asp His Arg Gly Phe Ile Ala Pro Pro Gln
TP84_40

GTTTATATCAATAACATGCGAAACCTGGCGGCGATCGAGGAAGATATAAAAGTCGATCACCAGGGCAAAA
CAAAATAGTTATTGTACGCTTTGGACCGCCGCTAGCTCCTTCTATATTTTCAGCTAGTGGTCCCGTTTT

55 60 65 70 75
Val Tyr Ile Asn Asn Met Arg Asn Leu Ala Ala Ile Glu Glu Asp Ile Lys Val Asp His Gln Gly Lys
TP84_40

AAGGCGAAATCAGCGGCAACTATATCCGGGATGTGGATGAATTCAAACGATTGTTGGCACTCACACTGAA
TTCCGCTTTAGTCGCGGTTGATATAGGCCCTACACCTACTTAAGTTTGCTAACAACCGTGAGTGTGACTT

80 85 90 95
Lys Gly Glu Ile Ser Gly Asn Tyr Ile Arg Asp Val Asp Glu Phe Lys Arg Leu Leu Ala Leu Thr Leu Asn
TP84_40

CAATGCGATGGGCGAACACAAACACCACACACTGATGAAATACATCGACATTTTATTCTTTACGATCGGA
GTTACGCTACCCGCTTGTGTTTGTGGTGTGTGACTACTTTATGTAGCTGTAAAGTAAGAAATGCTAGCCT

100 105 110 115 120
Asn Ala Met Gly Glu His Lys His His Thr Leu Met Lys Tyr Ile Asp Ile Ser Phe Phe Thr Ile Gly
TP84_40

CTACATAGAGGGGCGCAGGATGATCTGGATGCACATGCCATCGCATTCAATAACCGCATCACACGATATT
GATGTATCTCCCCGCGCTCTACTAGACCTACGTGTACGGTAGCCTAAGTTATTGGCGTAGTGTGCTATAA

125 130 135 140 145
Leu His Arg Gly Ala Gln Asp Asp Leu Asp Ala His Ala Ile Ala Phe Asn Asn Arg Ile Thr Arg Tyr
TP84_40

CAACACGGCTGGCCAACATCCAGGAAACTGTCTTGTCTGAATGGTATCAAGACAAGATCATTCCATTGGA
GTTGTGCCGACCGGTTGTAGGTCCTTTGACAGAACAGACTTACCATAGTTCTGTTCTAGTAAGGTAAGCT

150 155 160 165
Ser Thr Arg Leu Ala Asn Ile Gln Glu Thr Val Leu Ser Glu Trp Tyr Gln Asp Lys Ile Ile Pro Phe Glu
TP84_40

ACATGCACAGCATTTTGTGGGATCGAATGCGCGATGGTGATTGAAACAGACATTGGAAAATTTGAATAC
TGTACGTGTGCTAAAACAACCTAGCTTACCGGCTACCACTAACTTTGTCTGTAACCTTTAAACTTATG

170 175 180 185 190
His Ala Gln His Phe Val Gly Ile Glu Trp Pro Met Val Ile Glu Thr Asp Ile Gly Asn Phe Glu Tyr
TP84_40

ACACCATTTCGGATACATCCATGAAAATTTCAACAGAATCAGCGATGAAAATGGATTGAAAAAAGATGTGA
TGTGGTAAGCCTATGTAGGTACTTTTAAAGTTGTCTTAGTCGCTACTTTTACCTAACTTTTTTCTACACT
195 200 205 210 215
Thr Pro Phe Gly Tyr Ile His Glu Asn Phe Asn Arg Ile Ser Asp Glu Asn Gly Leu Lys Lys Asp Val
TP84_40

AACGCGGTCTGATCCCACTGTCAATGGCATCGAACGCACTATGGAAAATTGATCTGTTTAATCTGCGATA
TTGCGCCAGACTAGGGTGACAGTTACCGTAGCTTGGGTGATACCTTTTAACTAGACAAATTAGACGCTAT
220 225 230 235
Lys Arg Gly Leu Ile Pro Leu Ser Met Ala Ser Asn Ala Leu Trp Lys Ile Asp Leu Phe Asn Leu Arg Tyr
TP84_40

TGTCTATAAAATGCGGTCTGAAGCTGACAAAAGCGAATCCAGAACTGAAACAAGGGATGGAAATGCTGGCG
ACAGATATTTTACGCCAGCTTCGACTGTTTTCGCTTAGGTCTTGACTTTGTTCCTACCTTTACGACCGC
240 245 250 255 260
Val Tyr Lys Met Arg Ser Lys Leu Thr Lys Ala Asn Pro Glu Leu Lys Gln Gly Met Glu Met Leu Ala
TP84_40

GATCAGATCGAAAAACATGTGCCTGTATTCGGTCCATATTTCCGTTATGAATTGACAGACACTGGCGAAT
CTAGTCTAGCTTTTGTACACGGACATAAGCCAGGTATAAAGGCAATACTTAACTGTCTGTGACCGCTTA
265 270 275 280 285
Asp Gln Ile Glu Lys His Val Pro Val Phe Gly Pro Tyr Phe Arg Tyr Glu Leu Thr Asp Thr Gly Glu
TP84_40

GGGAACACATGAACAAAGTCAAGACCGTCACCAGGGAAGAATATGAACTATTGAAAAAGATCAAAAAGCA
CCCTTGTGTACTTGTTCAGTTCTGGCAGTGGTCCCTTCTTATACTTGATAACTTTTTCTAGTTTTTCGT
290 295 300 305
Trp Glu His Met Asn Lys Val Lys Thr Val Thr Arg Glu Glu Tyr Glu Leu Leu Lys Lys Ile Lys Lys Gln
TP84_40

GATGCAACAACAAGGCACAGTCGATGTGGATCTTTACCGTCTGCAACAGGAAGGGAATGAAAAAATGGT
CTACGTTGTTGTTCCGTGTACAGCTACACCTAGAAATGGCAGACGTTGTCCTTCCCTTACTTTTTTACCA
310 315 320 325
Met Gln Gln Gln Gly Thr Val Asp Val Asp Leu Tyr Arg Leu Gln Gln Glu Glu Glu
TP84_40 Met Val
TP84_41

RBS

GAAGGTGGATGAAAGAACGATCAATTTATCAGATCTATTCAAAAACAGAAAATGCTGACGGATTATGTG
CTTCCACCTACTTTCTGCTAGTTAAATAGTCTAGATAAGGTTTTGTCTTTTACGACTGCCTAATACAC
5 10 15 20 25
Lys Val Asp Glu Arg Thr Ile Asn Leu Ser Asp Leu Phe Gln Lys Gln Lys Met Leu Thr Asp Tyr Val
TP84_41

TGGCAAAAACACAACATCAAAAAAGCGGATCTGAAAAAGAAAAAGCGGCGATCATCTGCGAATTGTGGG
ACCGTTTTTGTGTTGTAGTTTTTTCGCCTAGACCTTTTTCTTTTTTCGCCGCTAGTAGACGCTTAACACCC
30 35 40 45
Trp Gln Lys His Asn Ile Lys Lys Ala Asp Leu Glu Lys Glu Lys Ala Ala Ile Ile Cys Glu Leu Trp
TP84_41

AAACCGCAAACGAATTGAAATCCGATGGCTTCAAATATTGGACCGACAAAAAATGCGACAGGGAAAAAAC
TTTGGCGTTTGCTTAACTTTAGGCTACCGAAGTTTATAACCTGGCTGTTTTTTACGCTGTCCCTTTTTTG
50 55 60 65 70
Glu Thr Ala Asn Glu Leu Lys Ser Asp Gly Phe Lys Tyr Trp Thr Asp Lys Lys Cys Asp Arg Glu Lys Thr
TP84_41

GCTGGAAGAAATCGTGGACATGTTGCATTTCTATTTGCAGATCGGGAACATCCTGGGTGTCGTTTATGAA
CGACCTTCTTTAGCACCTGTACAACGTAAGATAAACGTCTAGCCCTTGTAGGACCCACAGCAAATACTT

18 270

75 80 85 90 95
Leu Glu Glu Ile Val Asp Met Leu His Phe Tyr Leu Gln Ile Gly Asn Ile Leu Gly Val Val Tyr Glu
TP84_41

CATTACTGGATCGAAAGACGTGACACGATTCTGGATCAGATCATGGCGATCAACTGGTCATTATTAATGA
GTAATGACCTAGCTTTCTGCACTGTGCTAAGACCTAGTCTAGTACCGCTAGTTGACCAGTAATAATTACT

18 340

100 105 110 115
His Tyr Trp Ile Glu Arg Arg Asp Thr Ile Leu Asp Gln Ile Met Ala Ile Asn Trp Ser Leu Leu Met
TP84_41

TGGATGGGCCGCTGACATGGGCGGTGTCATTTCGCACAATATCGTGGGCTTGTTTCGGATGCTTGGATTCTGA
ACCTACCCGGCGACTGTACCCGCCACAGTAAGCGTGTTATAGCACCCGAACAAGCCTACGAACCTAAGCT

18 410

120 125 130 135 140
Met Asp Gly Pro Leu Thr Trp Ala Val Ser Phe Ala Gln Tyr Arg Gly Leu Val Arg Met Leu Gly Phe Asp
TP84_41

CTGGGATCAAGATATTATCCCGGCATATAACCGAAAAATTCAGGAAAAACATCGCCAGACAGCAGAGGGGA
GACCCTAGTTCTATAATAGGGCCGTATATTGGCTTTTAAGGTCCTTTTGTAGCGGTCTGTCGTCTCCCT

18 480

145 150 155 160 165
Trp Asp Gln Asp Ile Ile Pro Ala Tyr Asn Arg Lys Phe Gln Glu Asn Ile Ala Arg Gln Gln Arg Gly
TP84_41

RBS

TATTAATCGATGTTTATTAATAAATTCATGTCCAGGGATGTATCTGGAATTTAAAAACAAAAACCCGAAAC
ATAATTAGCTACAAATAATTTTAAGGTACAGGTCCCTACATAGACCTTAAATTTTGTTTTGGGCTTTG

18 550

1 5 10 15 20
Tyr Met Phe Ile Lys Ile Pro Cys Pro Gly Met Tyr Leu Glu Phe Lys Asn Lys Asn Pro Lys
TP84_41 TP84_42

RBS

TGTTTGTGGAATATTGCAAAGCGTACATCAAACATAGCCACTATGGCTTGAAGGCGGTGCGGGTGGAAAC
ACAAACACCTTATAACGTTTCGCATGTAGTTGTATCGGTGATACCGAACTTCCGCCAGCCCCACCTTTG

18 620

25 30 35 40
Leu Phe Val Glu Tyr Cys Lys Ala Tyr Ile Lys His Ser His Tyr Gly Leu Lys Ala Val Gly Val Glu Thr
TP84_42

ATCCCGATCCGGCCAAGTCTACATCCTAACAATCCCGGATTATGATGCAAAACGGCCAGTCGCAAAAGGT
TAGGGCTAGGCCGGTTTCAGATGTAGGATTGTTAGGGCCTAATACTACGTTTTGCCGGTCAGCGTTTTCCA

18 690

45 50 55 60 65
Ser Arg Ser Gly Gln Val Tyr Ile Leu Thr Ile Pro Asp Tyr Asp Ala Lys Arg Pro Val Ala Lys Gly
TP84_42

AGAGGAAAAAGCAGACAAGTCATATTAGAACCGATGTATAAAGTGCCGATCGAGAATTTTTTGAAGTGAAG
TCTCCTTTTTTCGTCTGTTTCAGTATAATCTTGCTACATATTTACGGCTAGCTCTTAAAAAACTCACTTC

18 760

70 75 80 85 90
Arg Gly Lys Ser Arg Gln Val Ile Leu Glu Pro Met Tyr Lys Val Pro Ile Glu Asn Phe Leu Ser Glu
TP84_42

AAGAAATCGAATATCTAAAAAGGAAAGGGATGGACAAATATGGCAATAAAATCACGTGAAGCGTACATGA
TTCTTTAGCTTATAGATTTTTCTTTCCCTACCTGTTTATACCGTTATTTTAGTGCACTTCGCATGTACT

95 100 105 110
Glu Glu Ile Glu Tyr Leu Lys Arg Lys Gly Met Asp Lys Tyr Gly Asn Lys Ile Thr
TP84_42

RBS

1 5 10
Met Ala Ile Lys Ser Arg Glu Ala Tyr Met
TP84_43

TCATGAGAAATCCAGAAACGCACATGGAAATCAAATCGATCTGACGAAAGAAAACGCCGGGGATCTGGC
AGTACTCTTTAGGTCTTTGCGTGTACCTTTAGTTTTAGCTAGACTGCTTTCTTTTGGGCCCCCTAGACCG

15 20 25 30
Ile Met Arg Asn Pro Glu Thr His Met Glu Ile Lys Ile Asp Leu Thr Lys Glu Asn Ala Gly Asp Leu Ala
TP84_43

GATCAAAGCGCATGAAAAAGGATTTCGTGCCGATGGATCATGAACTGACACGAAATCAGTACCACAAACAA
CTAGTTTCGCGTACTTTTTCTAAGCACGGCTACCTAGTACTTGACTGTGCTTTAGTCATGGTGTTGTT

35 40 45 50 55
Ile Lys Ala His Glu Lys Gly Phe Val Pro Met Asp His Glu Leu Thr Arg Asn Gln Tyr His Lys Gln
TP84_43

TTCAAATCACTATCGAAGGCGGTCAATCCGAAACCGCGCAAACGGAAAGGAAGCCGCAAAAATGGCAAAA
AAGTTTAGTGATAGCTTCCGCCAGTTAGGCTTTGGCGCGTTTGCTTTCTTCGGCGTTTTTACCGTTTT

RBS

Met Ala Lys
TP84_44

60 65 70 75 80
Phe Lys Ser Leu Ser Lys Ala Val Asn Pro Lys Pro Arg Lys Arg Lys Gly Ser Arg Lys Asn Gly Lys
TP84_43

AGTAAGGGCAAGGAACTGGAAGATCGAATCATCAAGAAAAACAAAGAATATCTGGAACAAGGGATTGCTA
TCATTCGCGTTCTTGACCTTCTAGCTTAGTAGTTCTTTTTGTTTCTTATAGACCTTGTTCCCTAACGAT

5 10 15 20 25
Ser Lys Gly Lys Glu Leu Glu Asp Arg Ile Ile Lys Lys Asn Lys Glu Tyr Leu Glu Gln Gly Ile Ala
TP84_44

Lys
TP84_43

CGATCAAGAAGATCCCGAATAAATGGGCAATCAGACGGAAAGGGCCGCACATTGTCGGGGCGAACCCGAT
GCTAGTTCTTCTAGGGCTTATTTACCCGTTAGTCTGCCTTTCCCGGCGTGTAAACAGCCCCGCTTGGGCTA

30 35 40 45 50
Thr Ile Lys Lys Ile Pro Asn Lys Trp Ala Ile Arg Arg Lys Gly Pro His Ile Val Gly Ala Asn Pro Ile
TP84_44

TCCGTCTGGTCTATGCGATTTTATTGGCACATCCAAACTGGTCAGCGGTGTTTTGATCGTGTTTCGATGCG
AGGCAGACCAGATACGCTAAAAAACCCTGTAGGTTTGACCAGTCGCCAGCAAACTAGCACAAGCTACGC

55 60 65 70
Pro Ser Gly Leu Cys Asp Phe Ile Gly Thr Ser Lys Leu Val Ser Gly Arg Leu Ile Val Phe Asp Ala
TP84_44

AAGGAATGCAAACTGAAAACACGTTTTCCACTGAAAAATATCAAGCCGGAACAAATGGATCACATGCAGG
TTCCTTACGTTTGACTTTTGTCAAAAGGTGACTTTTTATAGTTTCGGCCTTGTTTACCTAGTGACGTCC

75 80 85 90 95
Lys Glu Cys Lys Leu Lys Thr Arg Phe Pro Leu Lys Asn Ile Lys Pro Glu Gln Met Asp His Met Gln
TP84_44

AAACGGTCTGAACACGGCGGCATCGCTTTTCTGATCGTTTCTTTTACTGAACTGGATGAAATGTATTTCTT
TTTGCCAGCTTGTGCCGCCGTAGCGAAAAGACTAGCAAGTAAAATGACTTGACCTACTTTACATAAAGGA
100 105 110 115 120
Glu Thr Val Glu His Gly Gly Ile Ala Phe Leu Ile Val His Phe Thr Glu Leu Asp Glu Met Tyr Phe Leu
TP84_44

ACCATATGAATTCGTGAAACCATATTGGGAAGCGGCTGAAAAGGAAGAAGGGCCATCATCGATTCCGATA
TGGTATACTTAAGCACTTTGGTATAACCCTTCGCCGACTTTTCCTTCTTCCCGGTAGTAGCTAAGGCTAT
125 130 135 140
Pro Tyr Glu Phe Val Lys Pro Tyr Trp Glu Ala Ala Glu Lys Glu Glu Pro Ser Ser Ile Pro Ile
TP84_44

GCGGACATCCGGGAATCTTGACGCGAGATCATCGACATGGATTATTTACGCGATATCCAGATGTTTTACA
CGCCTGTAGGCCCTTAGAACGTGCGTCTAGTAGCTGTACCTAATAAAGTGCCTATAAGGTCTACAAAATGT
145 150 155 160 165
Ala Asp Ile Arg Glu Ser Cys Thr Gln Ile Ile Asp Met Asp Tyr Phe Thr His Ile Gln Met Phe Tyr
TP84_44

RBS

AATGAAACCGAACATCACCGATCGATCCACCTGGACATCTGGGTGGATCGTGGAAGATCCAGACAACCAC
TTACTTTGGCTTGTAGTGGCTAGCTAGGTGGACCTGTAGACCCACCTAGCACCTTCTAGGTCTGTTGGTG

Lys
TP84_44

1 5 10 15 20
Met Lys Pro Asn Ile Thr Asp Arg Ser Thr Thr Ser Gly Trp Ile Val Glu Asp Pro Asp Asn His
TP84_45

ATTTTGATCTGTGTCGATGAATGCGGCAACGTGGTAGGAACATGTGACGCAATGATGCCGCCAGTCATGG
TAAACTAGACACAGCTACTTACGCCGTTGCACCATCCTTGTACACTGCGTTACTACGGCGGTCAGTACC
25 30 35 40 45
Ile Leu Ile Cys Val Asp Glu Cys Gly Asn Val Val Gly Thr Cys Asp Ala Met Met Pro Pro Val Met
TP84_45

TCAAACGGTGGATCGTGAAGGTGGGGGAATTCAAAAAAATCAACATTCGAAACCAAGATGCACTATATACA
AGTTTGCCACCTAGCACTTCCACCCCTTAAGTTTTTTTAGTTGTAAGCTTTGGTCTACGTGATATATGT
50 55 60 65 70
Val Lys Arg Trp Ile Val Lys Val Gly Glu Phe Lys Lys Ile Asn Ile Arg Asn Gln Met His Tyr Ile Gln
TP84_45

GAAATGGAACGAAAGAATGGTGAATGATGATGACGAAAAATATCGATTGGAATAAACAAGCGGCAACCAC
CTTTACCTTGCTTTCTTACCACTTACTACTACTGCTTTTATAGCTAAGCTTATTTGTTGCCGTTGGTG

RBS

1 5 10
Met Thr Lys Asn Ile Asp Ser Asn Lys Gln Ala Ala Thr Thr
TP84_46

75 80 85
Lys Trp Asn Glu Arg Met Val Asn Asp Asp Glu Lys Tyr Arg Phe Glu
TP84_45

GAAACCGCCGAGGGAAAAATATATGATCAGCGGTGGCAGGATTATCAAAGACGGATCGGCTTTTACAAA
CTTTGGCGGCGTCCCTTTTTATATACTAGTCGCCACCGTCCTAATAGTTTCTGCCTAGCCGAAAATGTTT
15 20 25 30 35
Lys Pro Pro Gln Gly Lys Ile Tyr Asp Gln Arg Trp Gln Asp Tyr Gln Arg Arg Ile Gly Phe Tyr Lys
TP84_46

AAGACGGTGGCAAAATACAACGCCATTGTGAAACGGATCGAGGAATTATATAAAATCACGGAACAACTGG
TTCTGCCACCGTTTTATGTTGCGGTAACACTTTGCCTAGCTCCTTAATATATTTTAGTGCCTTGTGACC
19 950
40 45 50 55 60
Lys Thr Val Ala Lys Tyr Asn Ala Ile Val Lys Arg Ile Glu Glu Leu Tyr Lys Ile Thr Glu Gln Leu
TP84_46 >

ATGACAACAATCCCGATGATGTACTGAAAAAGCCGAAATATACGGCATGATCCTGGAAGATGTCGGCGA
TACTGTTGTTAGGGCTACTACATGACTTTTTTCGECTTTATATGCCGTACTAGGACCTTCTACAGCCGCT
20 020
65 70 75 80
Asp Asp Asn Asn Pro Asp Asp Val Leu Lys Lys Ala Glu Ile Tyr Gly Met Ile Leu Glu Asp Val Gly Asp
TP84_46 >

TCTGCATTATATTTTCGATCACTGCCTGGAAATACGCCGATGCACTGAAAAAGGAAGCCTATGCACTGGCG
AGACGTAATATAAAGCTAGTGACGGACCTTTATGCGGCTACGTGACTTTTTCCTTCGGATACGTGACCGC
20 090
85 90 95 100 105
Leu His Tyr Ile Ser Ile Thr Ala Trp Lys Tyr Ala Asp Ala Leu Lys Lys Glu Ala Tyr Ala Leu Ala
TP84_46 >

ATCATCCGGGAAAGGCCAAACGGCCGCACAGTCGAAGCGCACCGGGAAATGGCGGTCTGGAATCGCAAG
TAGTAGGCCCTTTCCGGTTTGCCGGCGTGTCAGCTTCGCGTGCGCCTTTACCGCCAGGACCTTAGCGTCC
20 160
110 115 120 125 130
Ile Ile Arg Glu Arg Pro Asn Gly Arg Thr Val Glu Ala His Arg Glu Met Ala Val Leu Glu Ser Gln
TP84_46 >

AATGGCGATGGAATGGCTGAATGGGAAGGATTGACGAAACGGTGGGAAAACGCAAAGACAACAATTGA
TTACCGCTACCTTTTACCGACTTACCCTTCCTAACTGCTTTGCCACCCTTTTGGCTTTCTGTTGTTAACT
20 230
135 140 145 150
Glu Lrp Arg Lrp Lys Met Ala Glu Lrp Glu Gly Leu Thr Lys Arg Lrp Glu Asn Ala Lys Thr Thr Ile Glu
TP84_46 >

AGAACAGATCAAAATCATGAAGTGAAGATCAAAATGGACACTGGCGAACATGCAACAAGCAGGAATGGCG
TCTTGTCTAGTTTTAGTACTTCACCTTCTAGTTTACCTGTGACCCTTGTACGTTGTTTCGTCCTTACC6C
20 300
155 160 165 170 175
Glu Gln Ile Lys Ile Met Lys Trp Lys Ile Lys Trp Thr Leu Ala Asn Met Gln Gln Ala Gly Met Ala
TP84_46 >

AATCCGAATGCATAACGAAGGGGGGTGTTTCAGTGGTGAAATGTGCTATCTGCGGCTATGAATTGACAATC
TTAGGCTTACGTATTGCTTCCCCCACAAGTCACCACTTTACACGATAGACGCCGATACTTAACTGTTAG
20 370
180 5 10
Asn Pro Asn Ala Met Val Lys Cys Ala Ile Cys Gly Tyr Glu Leu Thr Ile
TP84_46 RBS TP84_47 >

GATGAAGAAATCGAAGTCATGCTGGCCGATGATGATGAATGCCCAAATGCGGCGGCGGATTGACTGTG
CTACTTCTTTAGCTTCAGTACGACCGGCTACTACTACTTACGGGTTTTACGCCGCGGCTTAAGTACAGC
20 440
15 20 25 30 35
Asp Glu Glu Ile Glu Val Met Leu Ala Asp Asp Asp Glu Cys Pro Lys Cys Gly Gly Gly Leu Thr Val
TP84_47 >

AATGAAAAGGAGGTAAAGGAATGAAGCGGCACGGAAAATGCTATTCTGTGCCAGACAACAGGGGAAACGAT
TTACTTTTCTCCATTTCCTTACTTCGCCGTGCTTTTACGATAAGCACGGTCTGTTGTCCCTTTTGCTA
20 510
5 10 15
Met Lys Arg His Gly Lys Cys Tyr Ser Cys Gln Thr Thr Gly Glu Thr Ile
TP84_47 RBS TP84_48 >

ATACAGCCAGACGATGGCCGAATGGTTGTGTATGGAATGCTACAAAGCATTCTGGATCAAATCATCGAT
TATGTCGGTCTGCTACCGGCTTACCAACACATACCTTACGATGTTTCGTAAGACCTAGTTTAGTAGCTA
20 580
Tyr Ser Gln Thr Met Ala Glu Trp Leu Cys Met Glu Cys Tyr Lys Ala Phe Leu Asp Gln Ile Ile Asp
TP84_48

GGATAAAAAGTGGACAAAGTGGACATGTCCACGATGTCCACATCAAAAAAGGGGGATGATCATGGCCAA
CCTATTTTTTACCTGTTTACCTGTACAGGTGCTACAGGTGTAGTTTTTTTCCCCCTACTAGTACCGGTT
20 650
Gly TP84_48 RBS Met Ala Lys TP84_49
ODD

ACTGAATGGATACTGGATCAGCAAACTGATCCGAACGTGACGGTGTATGTCGAAAAGGTCTACAAAAA
TGACTTACCTATGACCTAGTCGTTTGGACTAGGCTTGCACTGCCACATACAGCTTTTCCAGATGTTTTT
20 720
Leu Asn Gly Tyr Trp Ile Ser Lys Thr Asp Pro Asn Val Thr Val Tyr Val Glu Lys Val Tyr Lys Lys
TP84_49

GGATACGTCACCTGGATTCAAATATATGAAGGTTGCAGACGGTGAAATCCATTCCAGATTCAAAGCGACAT
CCTATGCAGTGACCTAAGTTTATATACTTCCAACGCTCTGCCACTTTAGGTAAGGTCTAAGTTTCGCTGTA
20 790
Gly Tyr Val Thr Gly Phe Lys Tyr Met Lys Val Ala Asp Gly Glu Ile His Ser Arg Phe Lys Ala Thr
TP84_49

TCGATGAATTGCTGGATCAATACGAACGAAAAGGATGATAAATATGGATCGATGGGAACGAATGGCGAA
AGCTACTTAACGACCTAGTTATGCTTGCTTTTCTACTATTTTATACCTAGCTACCCTTGCTTACCGCTT
20 860
Phe Asp Glu Leu Leu Asp Gln Tyr Glu Arg Lys Gly TP84_49
RBS TP84_50

CGAACGCGGCCAGGCGATCCAGAAAGCGATCCAAAAGGAAAGGGGGATGCAGAATGGGTGGCAAACCATC
GCTTGCGCCGGTCCGCTAGGTCTTTCGCTAGGTTTCTTTCCCTTACGCTCTTACCCACCGTTTGGTAG
20 930
RBS Met Gly Gly Lys Pro Ser TP84_51
TP84_50

GAAGGGAACGCCGAAAGATATGCGATTGAAACGGAACAGGAAAGCGGCAAGTCTGCCAGTCGGTCCAGGA
CTTCCCTTGCGGCTTTCTATACGCTAACTTTGCCTTGCTCTTTCGCCGTCCAGACGGTCAGCCAGGTCTT
21 000
Lys Gly Thr Pro Lys Asp Met Arg Leu Lys Arg Asn Arg Lys Ala Ala Gly Leu Pro Val Gly Pro Gly
TP84_51
TP84_50

TCGAGATCGAAAAAGACAACAAGTAAACGAAAATAAGGCCAGGGGAAACCTGGTCTTTTTTGCGGACAAA
AGCTCTAGCTTTTTCTGTTGTTTCATTTGCTTTTATTCCGGTCCCTTTGGACCAGAAAAACGCCTGTTT
21 070
Ser Arg Ser Lys Lys Thr Thr Ser Lys Arg Lys TP84_51
regulatory region

ACGGACATTTTAGAAGAATATAACAAAAGCGCATTTCGGCTATGGTAAAATATAGACAAGCCTGGCCAGGT
TGCCTGTAAAATCTTCTTATATTGTTTTCGCGTAAGCCGATACCATTTTATATCTGTTTCGGACCGGTCCA
21 140

AAAAAAGAGGAAAGGGGAAGATCGAATGGCAAAACGACAACACAATTCCAGTACCATTCCGGCGTGGGG
TTTTTCTCCTTTCCCTTCTAGCTTACCGTTTTTGCTGTTGTGTTAAGGTCATGGTAAGGCCGACCCC

21 210

RBS

1 5 10 15
Met Ala Lys Thr Thr Gln Phe Gln Tyr His Ser Gly Val Gly
TP84_52

GATGCAACATCCTACATTCACACATCGGAAGATTACATGAATGTTTATGTGACAGTAGAACTGGGCCGTG
CTACGTTGTAGGATGTAAGTGTGTAGCCTTCTAATGTACTTACAAATACACTGTCATCTTGACCCGGCAC

21 280

20 25 30 35
Asp Ala Thr Ser Tyr Ile His Thr Ser Glu Asp Tyr Met Asn Val Tyr Val Thr Val Glu Leu Gly Arg
TP84_52

TGGGAACATGGGAAACGGAAGCATGGTGCAAATTAGCTTTACAACGATATGAAAACGGCGCATGGAAAAC
ACCTTGTACCCTTTGCCTTCGTACCACGTTTAATCGAAATGTTGCTATACTTTTGCCGCGTACCTTTTG

21 350

40 45 50 55 60
Val Gly Thr Trp Glu Thr Glu Ala Trp Cys Lys Leu Ala Leu Gln Arg Tyr Glu Asn Gly Ala Trp Lys Thr
TP84_52

GATCGCCACAGCGCAAGGATATGCCGCAACAGGGCAAAATCTGAATCGAACATTCAGCAATATCAGTAAT
CTAGCGGTGTCGCGTTCCTATACGGCGTTGTCCCGTTTTAGACTTAGCTTGTAAGTCGTTATAGTCATTA

21 420

65 70 75 80 85
Ile Ala Thr Ala Gln Gly Tyr Ala Ala Thr Gly Gln Asn Leu Asn Arg Thr Phe Ser Asn Ile Ser Asn
TP84_52

GTGATGGAAAAGCCAAATGCGTGTCAAAGTGGATCTTTATGCGAATTCCTCTTATTCGGATTATGTCCAAA
CACTACCTTTTCGGTTACGCACAGTTTACCTAGAAATACGCTTAAGGAGAATAAGGCTAATACAGGTTT

21 490

90 95 100 105
Val Met Glu Lys Pro Met Arg Val Lys Val Asp Leu Tyr Ala Asn Ser Ser Tyr Ser Asp Tyr Val Gln
TP84_52

CTGTCTACAAAAACAGTGGATCAGATAAGGAGTGAATGGGCGACATGGCGACAACCTCTAAAGCGATCG
GACAGATGTGTTTTGTACCTAGTCTATTCCTCACTTACCCGCTGTACCGCTGTTGAAGATTTGCGCTAGC

21 560

110 115
Thr Val Tyr Thr Lys Gln Trp Ile Arg
TP84_52

1 5 10
Met Gly Asp Met Ala Thr Thr Ser Lys Ala Ile
TP84_53

RBS

AATTGAATGATAATTTTACTTATGGTGGTACTGGATACCTTAACACATCCGAAGATTATATGAATGTTTA
TTAACTTACTATTTAAATGAATACCACCATGACCTATGGAATTGTGTAGGCTTCTAATATACTTACAAAT

21 630

15 20 25 30 35
Glu Leu Asn Asp Asn Phe Thr Tyr Gly Gly Thr Gly Tyr Leu Asn Thr Ser Glu Asp Tyr Met Asn Val Tyr
TP84_53

TGCCACCTGGGAAACGCCATATAATACTTATGTTTATGCCAGTATGACACTGCAACGATATGAGGATGGC
ACGGTGGACCCTTTGCGGTATATTATGAATACAAATACGGTCATACTGTGACGTTGCTATACTCCTACCG

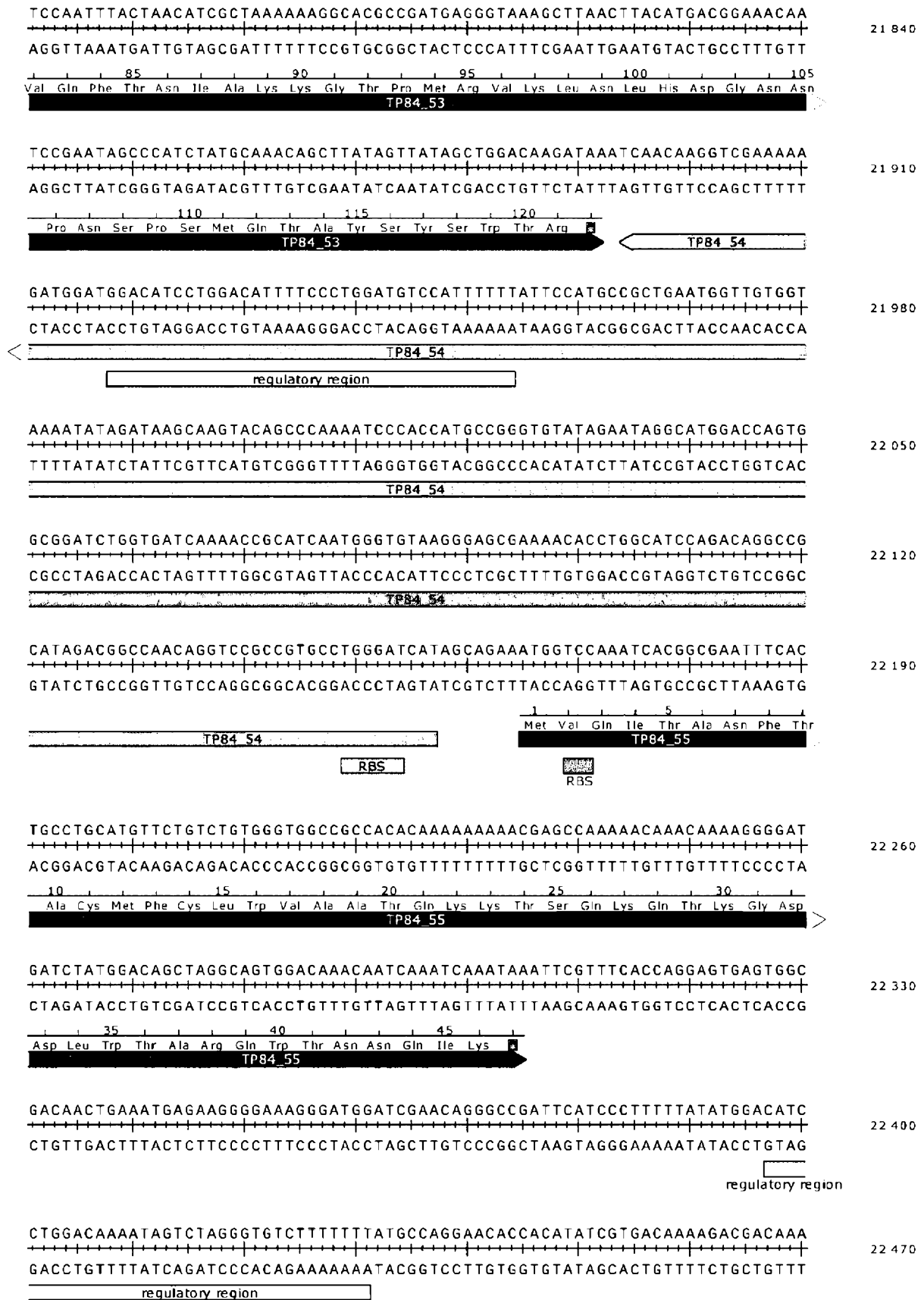
21 700

40 45 50 55
Ala Thr Trp Glu Thr Pro Tyr Asn Thr Tyr Val Tyr Ala Ser Met Thr Leu Gln Arg Tyr Glu Asp Gly
TP84_53

GTGTGGAAAAACATCGAAACGAAAGGTGCTTATGCATATTACACACACCAAGAACGCAAACACGCTAATG
CACACCTTTTGTAGCTTTGCTTTCCACGAATACGTATAATGTGTGTGTTCTTGCGTTTGTGCGATTAC

21 770

60 65 70 75 80
Val Trp Lys Asn Ile Glu Thr Lys Gly Ala Tyr Ala Tyr Tyr Thr His Gln Glu Arg Lys His Ala Asn
TP84_53



AATCTGGTAAAATGAACATAAGGAACACATGATGTTCAATTCGTATCAAAACGATACAACGATGTTTCAT
TTAGACCATTCTTACTTGTATTCTTGTGTACTACAAGTTAAGCATAGTTTTGCTATGTTGCTACAAAGTA

22 540

RBS

Met Phe Asn Ser Tyr Gln Asn Asp Thr Thr Met Phe His
TP84_56

TCGGAAGGGGTGAAACCAAGCGTGGCAAAGAAAAAGAAAGATAAACATGTACCACATGCGGAAAAACAGC
AGCCTTCCCCACTTTGGTTTCGCACCGTTTCTTTTCTTTCTATTTGTACATGGTGTACGCCCTTTTGTCTG

22 610

Ser Glu Gly Val Lys Pro Ser Val Ala Lys Lys Lys Lys Asp Lys His Val Pro His Ala Glu Lys Gln
TP84_56

GAATTGAAACGCCGGAGCAAAAACAGGCGTTTCGAAACTTATTTTCATGATGGGTGAAGATCGATCGTATCG
CTTAACCTTTGCGGCCTCGTTTTTGTCCGCAAGCTTTGAATAAAGTACTACCCACTTCTAGCTAGCATAGC

22 680

Arg Ile Glu Thr Pro Glu Gln Lys Gln Ala Phe Glu Thr Tyr Phe Met Met Gly Glu Asp Arg Ser Tyr Arg
TP84_56

AAAAGTAGCAAAAATGTTGAATAAAGGTGTGACGACAATCTACAACCTGGGCGCAGTGGTTTAACTGGCAA
TTTTGATCGTTTTTACAACCTTATTTCCACACTGCTGTTAGATGTTGACCCGCGTCACCAAATTGACCGTT

22 750

Lys Leu Ala Lys Met Leu Asn Lys Gly Val Thr Thr Ile Tyr Asn Trp Ala Gln Trp Phe Asn Trp Gln
TP84_56

GAACGCATCGAAGAACGTGAAAAACAAGTCCAGAAGATTGTGCAAGAACGCAACAACAAAACGATGGCTG
CTTGCGTAGCTTCTTGCACTTTTTGTTTCAGGTCTTCTAACAGCTTCTTGCGTTGTTGTTTGTCTACCGAC

22 820

Glu Arg Ile Glu Glu Arg Glu Lys Gln Val Gln Lys Ile Val Glu Glu Arg Asn Asn Lys Thr Met Ala
TP84_56

AACTGAAACTGGAACAAGCCAGACTGATCGATGCAACACTGTCACGATTCTGGGAAAAAGTGAAGGCTGG
TTGACTTTGACCTTGTTCGGTCTGACTAGCTACGTTGTGACAGTGCTAAGACCCTTTTTCACTTCCGACC

22 890

Glu Leu Lys Leu Glu Gln Ala Arg Leu Ile Asp Ala Thr Leu Ser Arg Phe Trp Glu Lys Val Lys Ala Gly
TP84_56

AAAAATCGAGATCGAAAGCTGGCAAGATTATGAACGATTATGGCGAATTCTGTTGTGAAATCGGCGGTGAA
TTTTAGCTCTAGCTTTGACCGTTCTAATACTTGCTAATACCGCTTAAGCAACACTTTAGCCGCCACTT

22 960

Lys Ile Glu Ile Glu Ser Trp Gln Asp Tyr Glu Arg Leu Trp Arg Ile Arg Cys Glu Ile Gly Glv Glu
TP84_56

GCGGATCGCAAACAGTCTGGCGCACTCACTGAATTGACGAACATGATCGTGAAAGTGGGTGAACAGATCA
CGCCTAGCGTTTGTGACACCGCGTGAGTGACTTAACTGCTTGTACTAGCACTTTCACCCACTTGTCTAGT

23 030

Ala Asp Arg Lys Gln Ser Gly Ala Leu Thr Glu Leu Thr Asn Met Ile Val Lys Val Gly Glu Gln Ile
TP84_56

TGGCCAAACAACAAGAAAAGGCGGATGATGAAGAATGACAAAATATGTGCCGCCATCTGAAAACCAAGTG
ACCGGTTTGTGTTCTTTTCCGCCTACTACTTCTTACTGTTTTATACACGCGGTAGACTTTTGTGTTTCA

23 100

RBS

Met Thr Lys Tyr Val Pro Pro Ser Glu Asn Gln Val
TP84_57

Met Ala Lys Gln Gln Glu Lys Ala Asp Asp Glu Glu
TP84_56

ATGGCCATTCTGGGATCAAACAAACCATTTAATCTATGGGAAGGATCGATCCGAAGCGGCAAGACATTCT
TACCGGTAAGACCCCTAGTTTGTTGGTAAATTAGATACCCTTCCTAGCTAGGCTTCGCCGTTCTGTAAAG

15 20 25 30 35
Met Ala Ile Leu Gly Ser Asn Lys Pro Phe Asn Leu Trp Glu Gly Ser Ile Arg Ser Gly Lys Thr Phe
TP84_57

GGTCGATGGTCTGGCTATATCACAAGATCCAAACACTGCCGCCAGGGGATGGAATGATCCTGGGCCAAAC
CCAGCTACCAGACCGATATAGTGTCTAGGTTTGTGACGGCGGTCCCCTACCTTACTAGGACCCGGTTTG

40 45 50 55
Trp Ser Met Val Trp Leu Tyr His Lys Ile Gln Thr Leu Pro Pro Gly Asp Gly Met Ile Leu Gly Gln Thr
TP84_57

ACCAGAAACAATCTATCGAAACTTTATGAATGAATTCCTTTTCTGGCTGGATGACGAAGGTGCGAATTAT
TGGTCTTTGTTAGATAGCTTTGAAATACTTACTTAAGGAAAAGACCGACCTACTGCTTCCACGCTTAATA

60 65 70 75 80
Pro Glu Thr Ile Tyr Arg Asn Phe Met Asn Glu Phe Leu Phe Trp Leu Asp Asp Glu Gly Ala Asn Tyr
TP84_57

CGCCATGTGCAAAACAGCCACATCGACATCGAATATATACACAATGGCACGCCAGAGAAACGCCGAATGT
GCGGTACACGTTTTGTCGGTGTAGCTGTAGCTTATATATGTGTACCCTGCGGTCTCTTTGCGGCTTACA

85 90 95 100 105
Arg His Val Gln Asn Ser His Ile Asp Ile Glu Tyr Ile His Asn Gly Thr Pro Glu Lys Arg Arg Met
TP84_57

ATATAGTCGGCGCAAAAGACAAAGGAGCAATCAAGCGGATTCGTGGATCGACACTGATGATTGCATACAT
TATATCAGCCGCGTTTTCTGTTTCTCGTTAGTTCGCTAAGCACCTAGCTGTGACTACTAACGTATGTA

110 115 120 125
Tyr Ile Val Gly Ala Lys Asp Lys Gly Ala Ile Lys Arg Ile Arg Gly Ser Thr Leu Met Ile Ala Tyr Ile
TP84_57

TGACGAATTGACGATGATGCCGCAAGTCGTGTTTGACGAATTGGTGGGCCGCTTGTACATAAAAGAATCG
ACTGCTTAAGTCTACTACGGCGTCCAGCACAACTGCTTAACCACCCGGCGGAACAGTATGTTTCTTAGC

130 135 140 145 150
Asp Glu Leu Thr Met Met Pro Gln Val Val Phe Asp Glu Leu Val Gly Arg Leu Ser Tyr Lys Glu Ser
TP84_57

ATCCTGCTGGCCACAACGAACCCGGACAGTCCACATCACTGGGTGCTGAAACAATATGTCGAACACCCGG
TAGGACGACCGGTGTTGCTTGGGCCTGTCAGGTGTAGTGACCCACGACTTTGTTATACAGCTTGTTGGGCC

155 160 165 170 175
Ile Leu Leu Ala Thr Thr Asn Pro Asp Ser Pro His His Trp Val Leu Lys Gln Tyr Val Glu His Pro
TP84_57

AAAAGAAACAGGACTGGTCCAGATGGCGATTACAAATGGATGACAACCTGGCACTATCGGAAGAATACAA
TTTTCTTTGTCCTGACCAGGCTACCGCTAAGTGTTACCTACTGTTGGACCGTGATAGCCTTCTTATGTT

180 185 190 195
Glu Lys Lys Gln Asp Trp Ser Arg Trp Arg Phe Thr Met Asp Asp Asn Leu Ala Leu Ser Glu Glu Tyr Lys
TP84_57

AGATCGTGTGAAGCGGCAATATTCGGGCATCCCGGCACGCTATCAGCGGATGATCCTGGGTGCGATGGGTG
TCTAGCACACTTCGCCGTTATAAGGCCGTAGGGCCGTGCGATAGTCGCCTACTAGGACCCAGCTACCCAC

200 205 210 215 220
Asp Arg Val Lys Arg Gln Tyr Ser Gly Ile Pro Ala Arg Tyr Gln Arg Met Ile Leu Gly Arg Trp Val
TP84_57

ATGGCCGATGGCTTGAATTTACGATGTGTTTCGATCCGAAAAACACGTCATCACCAGAAAGCAAATGGAAT
TACCGGCTACCGAACTAAATGCTACACAAGCTAGGCTTTTGTGTCAGTAGTGGTCTTTTCGTTTACCTTA
225 230 235 240 245
Met Ala Asp Gly Leu Ile Tyr Asp Val Phe Asp Pro Lys Lys His Val Ile Thr Arg Lys Gln Met Glu
TP84_57

CGATGGTCAAAGGACCGCGGATGAAATGGCACGTGTCAGTGGACTATGGAACGAAAAACCCGACAGTGT
GCTACCAAGTTTCTGCGGCTACTTTACCGTGACAGTCACCTGATACCTTGTCTTTTGGGCTGTCACAA
250 255 260 265
Ser Met Val Lys Gly Pro Pro Met Lys Trp His Val Ser Val Asp Tyr Gly Thr Lys Asn Pro Thr Val Phe
TP84_57

CGGATTGATCGGGCAGTGGCCACATCCGAATCCGACAAGGGATCGAAAAATTTTATATATCTTGGTCAAG
GCCTAACTAGCCCGTCACCGGTGTAGGCTTAGGCTGTTCCCTAGCTTTTATAAATATATAGAACCAGTTC
270 275 280 285 290
Gly Leu Ile Gly Gln Trp Pro His Pro Asn Pro Thr Arg Asp Arg Lys Tyr Leu Tyr Ile Leu Val Lys
TP84_57

GAATATTATTACGATGGCCGCAAAACAGGTGTGTCAAAAACGACATCCGCTTATCTGCAAGATTTTAGAA
CTTATAATAATGCTACCGGCGTTTGTCCACACAGTTTTTGTGTAGGCGAATAGACETTCTAAATCTT
295 300 305 310 315
Glu Tyr Tyr Tyr Asp Gly Arg Lys Thr Gly Val Ser Lys Thr Thr Ser Ala Tyr Leu Gln Asp Phe Arg
TP84_57

AATTTATTGGTCAGAGCGGATAAGCACGATCACGATCGATCCGTCTGCCACGCCGCTGATCGCCGAATT
TTAAATAACCAAGTCTTCGCTATTCTGTGCTAGTGCTAGCTAGGCAGACGGTGCGGCGACTAGCGGCTTAA
320 325 330 335
Lys Phe Ile Gly Gln Lys Arg Ile Ser Thr Ile Thr Ile Asp Pro Ser Ala Thr Pro Leu Ile Ala Glu Phe
TP84_57

CGAAAAATGCAGGATTGAGAGTGACAGAAGCAGACAACGAAGTCCTGGGCGGCATCGCATTGGTGGCGAAT
GCTTTTACGTCCTAACTCTCACTGTCTTCGTCTGTTGCTTCAGGACCCGCGTAGCGTAACCACCGCTTA
340 345 350 355 360
Glu Asn Ala Gly Leu Arg Val Thr Glu Ala Asp Asn Glu Val Leu Gly Gly Ile Ala Leu Val Ala Asn
TP84_57

GCACTGGCCGAAGGATACTTCTATGTGCTGGATGAATGTGAACACACCATTGAGGAATTGGGCTTATACA
CGTGACCGGCTTCCTATGAAGATACACGACCTACTTACACTTGTGTGGTAACTCCTTAAGCCGAATATGT
365 370 375 380 385
Ala Leu Ala Glu Gly Tyr Phe Tyr Val Leu Asp Glu Cys Glu His Thr Ile Glu Glu Phe Gly Leu Tyr
TP84_57

TCTGGGATGAAGCGGCTGGCTTGAAGGAATCGACAGGCCAGTGAAGGAAAACGATCACTGCATGGACAT
AGACCCTACTTCGCCGACCGAACTTTCCTTAGCTGTCCGGTCACTTCCTTTGCTAGTGACGTACCTGTA
390 395 400 405
Ile Trp Asp Glu Ala Ala Gly Leu Lys Gly Ile Asp Arg Pro Val Lys Glu Asn Asp His Cys Met Asp Met
TP84_57

GATCAGATATTTCTTCAAAACACACGCCGATCCATCGAAGCGCGGCGGTGTTCTGGGCGTGTCTGGTTGG
CTAGTCTATAAAGAAGTTTTGTGTGCGGCTAGGTAGCTTCGCGCCGCCACAAGACCCGCACAGACCAACC
410 415 420 425 430
Ile Arg Tyr Phe Phe Lys Thr His Ala Asp Pro Ser Lys Arg Gly Gly Val Leu Gly Val Ser Gly Trp
TP84_57

RBS

Met
TP84_58

		PvuII	
TAAAAACCGAACAAAAACAGATGTCATCGGTCCAAAAATATTGGTGTTTCACCATCAAAATGTCAGCTGA			24 430
ATTTTTGGCTTGTCTACAGTAGCCAGGTTTATAAGCCACAAAGTGGTAGTTTACAGTCGACT			
TP84_57			
Val Lys Thr Glu Gln Lys Gln Met Ser Ser Val Gln Asn Ile Arg Cys Phe Thr Ile Lys Met Ser Ala Asp	5 10 15 20 25		
TP84_58			
TCGTATAGGTCAGCATCCGAAAAATACCGAACAAATCATCATTTTATCCATTGGAAGGGGGAATCATTCATG			24 500
AGCATATCCAGTCGTAGGCTTTTATGGCTTGTTAGTAGTAAAATAGGTAAGCTTCCCCCTTAGTAAGTAC			
Arg Ile Gly Gln His Pro Lys Ile Pro Asn Asn His His Phe Ile His Ser Lys Gly Glu Ser Phe Met	30 35 40 45		
TP84_58			
GGATTGAAAATGTGGATCAAAAAGAAACTGGGCCTGGTCTATCCATCGGATGTGTTAAGGGATAAGCCAG			24 570
CCTAACTTTTACACCTAGTTTTTCTTTGACCCGGACAGATAGGTAGCCTACACAATTCCTATTTCGGTC			
Gly Leu Lys Met Trp Ile Lys Lys Lys Leu Gly Leu Val Tyr Pro Ser Asp Val Leu Arg Asp Lys Pro	50 55 60 65 70		
TP84_58			
GGGGTTTCATTGACACGACAGCGTTCGGCAGCATGCCAGTCAATCAAATCAAACCAAAACAAGGGGGATC			24 640
CCCCAAAGTAACTGTGCTGTCGAAGCCGTGCTACGGTCAGTTAGTTTAGTTTGGTTTTGTTCCCCCTAG			
Gly Gly Phe Ile Asp Thr Thr Ala Phe Gly Thr Met Pro Val Asn Gln Ile Lys Pro Lys Gln Gly Gly Ser	75 80 85 90 95		
TP84_58			
AAACATGGCAGACACATATGCAGAAGCAATGATAAACAAAACAACTATCCAGATGCTGATCCAGAAATG			24 710
TTTGTAACGCTCTGTGTATACGTCTTCGTTACTATTTGTTTTGTTTGATAGGTCTACGACTAGGTCTTTAC			
Asn Met Ala Asp Thr Tyr Ala Glu Ala Met Ile Asn Lys Thr Asn Tyr Pro Asp Ala Asp Pro Glu Met	100 105 110 115		
TP84_58			
ACATTTGAGGTGCAACGATCCACGCTTGACAACTAGCGGAAGAATACAAAAACAATGGACAAAGTGG			24 780
TGTAAACTCCACGTTGCTAGGTGCGAACTGTTTGATCGCCTTCTTATGTTTTTTGTTIACCTGTTTCACC			
Thr Phe Glu Val Gln Arg Ser Thr Leu Asp Lys Leu Ala Glu Glu Tyr Lys Lys Thr Met Asp Lys Val	120 125 130 135 140		
TP84_58			
ACAAAAACGAAAATGTCCAGATGTCCACCAGGATGTGATCAACAAACCACGGCACTATCACCAGGGCGG			24 850
TGTTTTTGCTTTTACAGGTGCTACAGGTGGTCTACACTAGTTGTTTGGTGCCGTGATAGTGGTCCC GCC			
Asp Lys Asn Glu Asn Val His Asp Val His Gln Asp Val Ile Asn Lys Pro Arg His Tyr His Gln Gly Gly	145 150 155 160 165		
TP84_58			
ATTCGATGCGCTTTATGTGATCGAAAGAAAATTCGGTCTGTGTTGTTAAGGGGATTCTACATCGGCAAC			24 920
TAAGCTACGCGAAATACACTAGCTTTCTTTAAGCCAGCACACCACAATTCCTTAAGATGTAGCCGTTG			
Phe Asp Ala Leu Tyr Val Ile Glu Arg Lys Phe Gly Arg Val Val Leu Arg Gly Phe Tyr Ile Gly Asn	170 175 180 185		
TP84_58			

ATCATCAAATACATCCTGCGATTGGAACAGAAAAACGGTGTGGAAGATCTGAAAAAGCACGCTTTTATC 24 990
TAGTAGTTTATGTAGGACGCTAAGCTTGTCTTTTGGCACAGCTTCTAGACTTTTTCGTGCGAAAAATAG
190 195 200 205 210
Ile Ile Lys Tyr Ile Leu Arg Phe Glu Gln Lys Asn Gly Val Glu Asp Leu Lys Lys Ala Arg Phe Tyr
TP84_58

TGGACAAGCTGATCGAACTTGAAGAAGGAAGCGCACCATGATCAAAGGGGGTGATGGATAAAAGCGGCGGA 25 060
ACCTGTTGCGACTAGCTTGAACCTTCTTCTTCTGCGTGGTCTAGTTTCCCCCACTACCTATTTTCGCCGCCT
215 220 225
Leu Asp Lys Leu Ile Glu Leu Glu Gly Ser Ala Pro Asp Gln Arg Gly
TP84_58

AAGGATGGATCACGGTGAAAAAGTGGATCATCGGTGCGGCAATCGGCACGCTGATCGGCTACATGGCCGG 25 130
TTCCTACCTAGTGCCACTTTTTCACCTAGTAGCCACGCCGTTAGCCGTGCGACTAGCCGATGTACCGGCC
5 10 15
Met Lys Lys Trp Ile Ile Gly Ala Ala Ile Gly Thr Leu Ile Gly Tyr Met Ala Gly
RBS TP84_59

AAAATTGATCAGCAAACCTGGACAACATGGACAACCGATCAGACAGATTGAAGAAGAAATCGAAAGTGAT 25 200
TTTTAACTAGTCGTTTGACCTGTTGTACCTGTTTGCCTAGTCTGTCTAACTTCTTCTTTAGCTTTCACTA
20 25 30 35 40
Lys Leu Ile Ser Lys Leu Asp Asn Met Asp Lys Arg Ile Arg Gln Ile Glu Glu Glu Ile Glu Ser Asp
TP84_59

GATGGATGGATCGATGCGAGTCGAAACAGATGACGACATCGCCGTGAAATATCCACCGAAAATGGCATTCA 25 270
CTACCTACCTAGCTACGTACGCTTTGTCTACTGCTGTAGCGGCACCTTTATAGGTGGCTTTTACCGTAAGT
45 50 55 60 65
Asp Gly Trp Ile Asp Ala Val Glu Thr Asp Asp Ile Ala Val Lys Tyr Pro Pro Lys Met Ala Phe
TP84_59

TCTGTCCAAAATGTACCGCACTGGTGATGTTTCATGGCACAGACGAAGCGGCACAGCAAGAGATCGAAGC 25 340
AGACAGGTTTACATGGCGTGACCACTACAAAGTACCGTGCTGCTTCGCCGTGTCGTTCTCTAGCTTCG
70 75 80 85
Ile Cys Pro Lys Cys Thr Ala Leu Val Met Phe His Gly Thr Asp Glu Ala Ala Gln Gln Glu Ile Glu Ala
TP84_59

ATGGCGGCAGTACGAATTTATTTACTGTCCATACTGTGATGAAGAAATAGAAATCGATGACTGGACCAGA 25 410
TACCGCCGTCATGCTTAAATAAATGACAGGTATGACACTACTTCTTTATCTTTAGCTACTGACCTGGTCT
90 95 100 105 110
Trp Arg Gln Tyr Glu Phe Ile Tyr Cys Pro Tyr Cys Asp Glu Glu Ile Glu Ile Asp Asp Trp Thr Arg
TP84_59

CACATCCTGGATCAGTAGAATGAAAGGGGAACAGCATGAACGACAGCAAGATACAACCGATTCCACACAT 25 480
GTGTAGGACCTAGTCATCTTACTTTCCCTTGTGCTACTTGTGCTGCTTCTATGTTGGCTAAGGTGTGTA
115 1 5 10
His Ile Leu Asp Gln Met Asn Asp Ser Lys Ile Gln Pro Ile Pro His Ile
TP84_59 RBS TP84_60

CGTGGTCGCCGAAAACGGTGATGGATTCCATGTGATCCTGGAAGTGGCCATCACCGATCGTGTTTTATG 25 550
GCACCAGCGGCTTTTGGCACTACCTAAGGTACACTAGGACCTTCACCGGTAGTGGCTAGCACAAAAATAC
15 20 25 30 35
Val Val Ala Glu Asn Gly Asp Gly Phe His Val Ile Leu Glu Val Ala Ile Thr Asp Arg Val Phe Met
TP84_60

CGATTACGGTGGACATTGTGGACAACGAGGAAAAAGCATATGAACTGGCACATGAAATCCAGGATCGCT
GCTAAGTGCCACCTGTAACACCTGTTGCTCCTTTTCTGTATACTTGACCGTGTACTTTAGGTCCTAGCGA

40 45 50 55
Arg Phe Thr Val Asp Ile Val Asp Asn Glu Glu Lys Ala Tyr Glu Leu Ala His Glu Ile Gln Asp Arg
TP84_60

TGTCGGTGATCTGGAAGAATAACGAAAGGAGAGGGGGAAGATATGGGCATCAAGTGGACAAAATGGTCAA
ACAGCCACTAGACCTTTCTTATTGCTTTCCTCTCCCTTCTATACCCGTAGTTCACCTGTTTTACCAAGTT

60 65 1 5
Leu Ser Val Ile Trp Lys Glu Met Gly Ile Lys Trp Thr Lys Trp Ser
TP84_60 RBS TP84_61

CAAACGTGATCAAGAAATACCACGGCAACATCCAAAAATATCGCAAATTGTATGACGGTGATCATGCCAA
GTTTGCCTAGTTCCTTTATGGTGCCGTTGTAGGTTTTTATAGCGTTTAACATACTGCCACTAGTACGGTT

10 15 20 25 30
Thr Asn Val Ile Lys Lys Tyr His Gly Asn Ile Gln Lys Tyr Arg Lys Leu Tyr Asp Gly Asp His Ala Lys
TP84_61

ACTGTTCGAACGAGCAAAACGGCTGATCCAGGAAGGGGAAATCACCGATCAGATCATCGAAGGTGCAGAA
TGACAAGCTTGCTCGTTTTGCCGACTAGGTCCTTCCCTTTAGTGGCTAGTCTAGTAGCTTCCACGTCTT

35 40 45 50 55
Leu Phe Glu Arg Ala Lys Arg Leu Ile Gln Glu Gly Glu Ile Thr Asp Gln Ile Ile Glu Gly Ala Glu
TP84_61

GTGGCCAGAAACGTCAAAACACCGTACATCGTGGCCAATGTCTGCAAGATGATCGTGGACATCCCGGCCA
CACCGGTCTTTGCAGTTTTGTGGCATGTAGCACCGGTTACAGACGTTCTACTAGCACCTGTAGGGCCGGT

60 65 70 75
Val Ala Arg Asn Val Lys Thr Pro Tyr Ile Val Ala Asn Val Cys Lys Met Ile Val Asp Ile Pro Ala
TP84_61

TGTTGGTGTCCAGGGCAATTGGTCAAGTGACGACATCGATGTACCAGATGATTTTGTGGAATGGTCAC
ACAACCACAGGTCCCGTTAACCAGTTCAGTCTGTAGCTACAGTGGTCTACTAAAACGACCTTACCAGTG

80 85 90 95 100
Met Leu Val Ser Arg Ala Ile Gly Gln Val Thr Thr Ser Met Ser Pro Asp Asp Phe Ala Gly Met Val Thr
TP84_61

TGATCAAGGTGATGGTACGGTGATCAGTCTATACGAAAAACAAAAAGAACTGATCAAAGGGATCGCCAAA
ACTAGTTCCACTACCATGCCACTAGTCAGATATGCTTTTTGTTTTCTTGACTAGTTTCCCTAGCGGTTT

105 110 115 120 125
Asp Gln Gly Asp Gly Thr Val Ile Ser Leu Tyr Glu Lys Gln Lys Glu Leu Ile Lys Gly Ile Ala Lys
TP84_61

CGGTGCAATCTGCAATTGGAACACAAGACAAACATCATCCATCATCAAAATGGATGGCGGCATTGTAGGAA
GCCAGCTTAGACGTTAAGCTTGTTCTGTTTGTAGTAGGTAGTAGTTTACCTACCGCCGTAACATCCTT

130 135 140 145
Arg Ser Asp Leu Gln Phe Glu His Lys Thr Asn Ile Ile His His Gln Met Asp Gly Gly Ile Val Gly
TP84_61

TGCCATTGATGATGAAAACGGTCTGCGAATTGAATTCAAGTCGCGTGATGTCTATTATCCGCATCCAGA
ACGGTAAGCTACTACTTTTGCAGACGCTTAAGTTCAAGTTCAGCGCACTACAGATAATAGGCGTAGGTCT

150 155 160 165 170
Met Pro Phe Asp Asp Glu Asn Gly Leu Arg Ile Glu Phe Lys Ser Arg Asp Val Tyr Tyr Pro His Pro Asp
TP84_61

TGGCCGTGGATGTGATCTGGTGTATCAACTTGAAATTGAAGATGAAGAAACAGAAGAAGCGATCAAATAT
ACCGGCACCTACACTAGACCACATAGTTGAACCTTAACTTCTACTTCTTTGTCTTCTTCGCTAGTTTATA
175 180 185 190 195
Gly Arg Gly Cys Asp Leu Val Tyr Gln Leu Glu Ile Glu Asp Glu Glu Thr Glu Glu Ala Ile Lys Tyr
TP84:61

TTGCATGTGTACCGGGAAAGAGTGAAGAACAACAACTGGTGACACAACACATGCTTTATAAAATCGGTG
AACGTACACATGGCCCTTTCTCACCTTCTTGTGTTTGGACCACTGTGTTGTGTACGAAATATTTAGCCAC
200 205 210 215
Leu His Val Tyr Arg Glu Arg Val Glu Glu Gln Lys Leu Val Thr Gln His Met Leu Tyr Lys Ile Gly
TP84:61

AATCCGGGATGTTGGAAGAAATCGAGGATGAAGCGGAAGTCAAAGAAATCCTGGGTATTGAAAAACATA
TTAGGCCCTACAACCTTCTTTAGCTCCTACTTCGCCTTCAGTTTCTTTAGGACCCATAACTTTTTGTAT
220 225 230 235 240
Glu Ser Gly Met Leu Glu Glu Ile Glu Asp Glu Ala Glu Val Lys Glu Ile Leu Gly Ile Glu Lys Thr Tyr
TP84:61

CCGGGAATTGGAAGGCCGCGACAAACCATTCTGGTGTACTGGCCAAACAATAAAACATTCACGCATCCG
GGCCCTTAAGCTTCCGGCGCTGTTTGGTAAGCACCATGACCGGTTTGTATTTTGTAAAGTGCCTAGGC
245 250 255 260 265
Arg Glu Phe Glu Gly Arg Asp Lys Pro Phe Val Val Tyr Trp Pro Asn Asn Lys Thr Phe Thr His Pro
TP84:61

CTTGGTCGGTCTGAATTATACAACCTGGCCGGAAAACAGGATGAAATCAACTGGACATTGACACGAAATG
GAACCAAGCCAGACTTAATATGTTGGACCGGCCCTTTGTCTACTTTAGTTGACCTGTAAGTGTGCTTTAC
270 275 280 285
Leu Gly Arg Ser Glu Leu Tyr Asn Leu Ala Gly Lys Gln Asp Glu Ile Asn Trp Thr Leu Thr Arg Asn
TP84:61

CGATTGTTTATGAACGAAACGGCAAGCCGCGAATCGCCGTATCGAAAGAGATATTCCAGGCGTTGCAGGA
GCTAACAAATACCTTGTCTTGGCGTTTGGCGCTTAGCGGCATAGCTTTCTCTATAAGGTCCGCAACGTCT
290 295 300 305 310
Ala Ile Val Tyr Glu Arg Asn Gly Lys Pro Arg Ile Ala Val Ser Lys Glu Ile Phe Gln Ala Leu Glu Asp
TP84:61

CAAAGCGTTTGAACGATATGGCGATGAAAACAAGATCGATCATCGCGATCTGGAATTGTCACGTTTGCAC
GTTTCGCAAACCTTGCTATACCGCTACTTTTGTCTAGCTAGTAGCGCTAGACCTTTAACAGTGCAAACCTG
315 320 325 330 335
Lys Ala Phe Glu Arg Tyr Gly Asp Glu Asn Lys Ile Asp His Arg Asp Leu Glu Ile Val Thr Phe Asp
TP84:61

GAAAACGGAAAAGCGATGGAAGTCATTGAGATCGATGTCAGCAAGATCGGTGACATCAAATGGGTAAAAG
CTTTTGCCTTTTCGCTACCTTCAGTAAGTCTAGCTACAGTCGTTCTAGCCACTGTAGTTTACCCATTTTC
340 345 350 355
Glu Asn Gly Lys Ala Met Glu Val Ile Gln Ile Asp Val Ser Lys Ile Gly Asp Ile Lys Trp Val Lys
TP84:61

ATCTGATGAACTGATGCTGATGGAACACACACATCGGAAAAAGCGGTGATTTTTACCTGGAAGGCAA
TAGACTACTTTGACTACGACTACCTTTGTGTGTGTAGCCCTTTTCGCCAGCTAAAAATGGACCTTCCGTT
360 365 370 375 380
Asp Leu Met Lys Leu Met Leu Met Glu Thr His Thr Ser Glu Lys Ala Val Asp Phe Tyr Leu Glu Gly Asn
TP84:61

CACATCGGCACAGTCTGGGATCGCAAAATTTTATGATCTGTTTCGTATCGATCATGAAGGCGGAACAAATC
GTGTAGCCGTGTCTAGACCTAGCGTTTTAAATACTAGACAAGCATAGCTAGTACTTCCGCCTTGTTTAG

26 880

385 390 395 400 405
Thr Ser Ala Gln Ser Gly Ile Ala Lys Phe Tyr Asp Leu Phe Val Ser Ile Met Lys Ala Glu Gln Ile
TP84_61

GCCACAGAATATGTCCATTTCTTGCAAGAACTATTGCAAACTGCCTGTGGATCGCACACCAGGATGATC
CGGTGTCTTATACAGETAAGAACGTTCTTGATAAGCTTTGACGGACACCTAGCGTGTGGTCTCTACTAG

26 950

410 415 420 425
Ala Thr Glu Tyr Val His Phe Leu Gln Glu Leu Phe Glu Asn Cys Leu Trp Ile Ala His Gln Asp Asp
TP84_61

CAGACATCGTCATCGAAGAACCACGCATTCAGATCAAAGACATGATCCCGATCAGCCGCAGAGAATTGAT
GTCTGTAGCAGTAGCTTCTTGGTGCGTAAGTCTAGTTTCTGTACTAGGGCTAGTCGGCGTCTCTTAACCTA

27 020

430 435 440 445 450
Pro Asp Ile Val Ile Glu Glu Pro Arg Ile Gln Ile Lys Asp Met Ile Pro Ile Ser Arg Arg Glu Leu Ile
TP84_61

CGAACAAGAAAGCACGGCATACAAAAACGGCACACAGTCGCTTGAAACAACAGTACGCAACCAAAATCCG
GCTTGTTCTTTTCGTGCCGTATGTTTTGCGGTGTGTACGGAACCTTTGTTGTATGCGTTGGTTTTAGGC

27 090

455 460 465 470 475
Glu Gln Glu Ser Thr Ala Tyr Lys Asn Gly Thr Gln Ser Leu Glu Thr Thr Val Arg Asn Gln Asn Pro
TP84_61

ACTGCGACAGAAGATTGGATCGAAGATGAACTGGCGGCCATCGAGGAAAGTCAGCAATCGACAGACACAA
TGACGCTGTCTTCTAACCTAGCTTCTACTTGACCGCCGGTAGCTCCTTTCAGTCGTTAGCTGTCTGTGTT

27 160

480 485 490 495
Thr Ala Thr Glu Asp Trp Ile Glu Asp Glu Leu Ala Ala Ile Glu Glu Ser Gln Gln Ser Thr Asp Thr
TP84_61

CATCGATCCTGATGGGGCGGCAGACGTTATCGAATCTATTAGACAATCGAAATCCTAACGGCACGCCGAT
GTAGCTAGGACTACCCCGCCGTCTGCAATAGCTTAGATAATCTGTTAGCTTTAGGATTGCCGTGCGGCTA

27 230

500 505 510 515 520
Thr Ser Ile Leu Met Gly Arg Gln Thr Leu Ser Asn Leu Leu Asp Asn Arg Asn Pro Asn Gly Thr Pro Ile
TP84_61

CGGAGCGGCACAGCAACAGCCACAACAAGGCACACCGCAAACAGGGGGTGGCCAGGCGTGATGATGAATA
GCCTCGCCGTGTGCTTGTGGTGTGTTCCGTGTGGCGTTTGTCCTCCACCGGTCCGCACTACTACTTAT

27 300

525 530 535 540
Gly Ala Ala Gln Gln Gln Pro Gln Gln Gly Thr Pro Gln Thr Gly Gly Gln Ala
TP84_61

RBS

Met Met Met Asn
TP84_62

AAAATGGACAAGTTGGACAAAATGAACGCGAGTGTGATCGTGGACAAAACGGACAAACGAGAAACGGCAAA
TTTTACCTGTTCAACCTGTTTTACTTGCGTCACACTAGCACCTGTTTGCCTGTTTGCTCTTTGCCGTTT

27 370

5 10 15 20 25
Lys Asn Gly Gln Val Gly Gln Asn Glu Arg Ser Val Ile Val Asp Lys Thr Asp Lys Arg Glu Thr Ala Asn
TP84_62

TCTGGTCATATCTTATTGTCTGATCATCGCCGGGATCATCATCAGATTGTCACTCATTTATGAAAAGGG
AGACCAGTATAGAATAACAGACTAGTAGCGGCCCTAGTAGTAGTGCTAACAGTGAGTAAATACTTTTCCC

27 440

30 35 40 45
Leu Val Ile Ser Tyr Cys Leu Ile Ile Ala Gly Ile Ile Thr Ile Val Thr His Leu
TP84_62

RBS

GTGATCCTGGATGGCCAACATGGACCAGATCGAACGACAAATTGTGGAACATGCAGGCAAGCTATTATC
CACTAGGACCTACCGGTTGTACCTGGTCTAGCTTGCTGTTTAAACACCTTGATACGTCGGTTCGATAATAG

Met Ala Asn Met Asp Gln Ile Glu Arg Gln Ile Val Glu Leu Cys Arg Gln Ala Ile Ile
TP84_63

GACATTATCGCATCGGTCGGTGGTGTGATAACTTACTGGATGATGAACGGCGAATGGCACTGGTCAACG
CTGTAATAGCGTAGCCAGCCACCACAGCTATTGAATGACCTACTACTTGCCGCTTACCGTGACCAGTTGC

Asp Ile Ile Ala Ser Val Gly Gly Val Asp Asn Leu Leu Asp Asp Glu Arg Arg Met Ala Leu Val Asn
TP84_63

CCATCGCCGACGAATGCAACAACCTGGGCGCAGACATTTGCGCCATCCTGCCTGCCGAACCTGCAAAAGC
GGTAGCGGCGTGCTTACGTTGTTGACCCGCGTCTGTAAAGCCGGTAGGACGGACGGCTTGAACGTTTTCG

Ala Ile Ala Ala Arg Met Gln Gln Leu Gly Ala Asp Ile Ser Ala Ile Leu Pro Ala Glu Leu Ala Lys Ala
TP84_63

CTATTATCAAGGCATAACCGATGCCGATGCATTCTTGATCAAGAAGGCACGGAAGTCACGGTCCAGGGT
GATAATAGTTCCGTATTGGCTACGGCTACGTAAGAACGTAGTTCTTCCGTGCCTTCAGTGCCAGGTCCCA

Tyr Tyr Gln Gly Ile Thr Asp Ala Asp Ala Phe Leu His Gln Glu Gly Thr Glu Val Thr Val Gln Gly
TP84_63

GTGAATAACAAATCCACCTGGCCGGGATCGAATCGATCGTGTCTGACACAATGATGGATATGCAAGCGG
CACTTATTTGTTTAGGTGGACCGGCCCTAGCTTAGCTAGCACAGACTGTGTTACTACCTATACGTTTCGCC

Val Asn Lys Gln Ile His Leu Ala Gly Ile Glu Ser Ile Val Ser Asp Thr Met Met Asp Met Gln Ala
TP84_63

CAATCCGAACCGCAGTGGCGATGGCCATCGCCGACATCGATCGATCCTGGTCGATGTACGGAACGACAT
GTTAGGCTTGGCGTACCGCTACCGGTAGCGGCTGTAGCTAGCGTAGGACAGCTACATGCCTTGCTGTA

Ala Ile Arg Thr Ala Val Ala Met Ala Ile Ala Asp Ile Asp Arg Ile Leu Val Asp Val Arg Asn Asp Ile
TP84_63

CGCAAAAGGAATGATCCTGGGCGATCCATCACGTGAAATCACAAAACGTGTCATGCAGACATTTGCCAAC
GCGTTTTCTTACTAGGACCCGCTAGGTAGTGCACTTTAGTGTTTTGCACAGTACGTCGTAAACGGTTG

Ala Lys Gly Met Ile Leu Gly Asp Pro Ser Arg Glu Ile Thr Lys Arg Val Met Gln Thr Phe Ala Asn
TP84_63

AGCGGCATGACAGCATTATCAGAAAGATAATAAGCGACTGCCGCTTGATTTTTATGCGATGACTGTCA
TCGCCGTACTGTCGTAAGTAGTGCTTTCTATTATTCGCTGACGGCGAACTAAAAATACGCTACTGACAGT

Ser Gly Met Thr Ala Phe Ile Thr Lys Asp Asn Lys Arg Leu Pro Leu Asp Phe Tyr Ala Met Thr Val
TP84_63

CCAGAACAAAAATGCGCCAGGCACACACAGACGGTGCAGTCAATCGATACAAAGAAAATGGCGTTTATCA
GGTCTTGTGTTTTACGCGGTCCGTGTGTGTCTGCCACGTCAGTTAGCTATGTTCTTTTACCGCAAATAGT

Thr Arg Thr Lys Met Arg Gln Ala His Thr Asp Gly Ala Val Asn Arg Tyr Lys Glu Asn Gly Val Tyr His
TP84_63

CGTCAAGATCAGCGAACACGGTATTACATGCAACAAATGTGCAAGATACCAGGGATTAGTCGTGGCACTT
GCAGTTCTAGTCGCTTGTGCCATAATGTACGTTGTTTACACGTTCTATGGTCCCTAATCAGCACCGTGAA

210 215 220 225 230
Val Lys Ile Ser Glu His Gly Ile Thr Cys Asn Lys Cys Ala Arg Tyr Gln Gly Leu Val Val Ala Leu
TP84_63

GATCCCGAACATGCGGAAGGCTTTCCAGTGGCCGGGATCGATGTGCCACTGCCGCCATATCATCCAAACT
CTAGGGCTTGTACGCCTTCCGAAAGGTCACCGGCCCTAGCTACACGGTGACGGCGGTATAGTAGGTTTGA

235 240 245 250
Asp Pro Glu His Ala Glu Gly Phe Pro Val Ala Gly Ile Asp Val Pro Leu Pro Pro Tyr His Pro Asn
TP84_63

GCCGCCATACAGTGCGGCCATTCTGTGATGGAATATCACAGTCAAAACGACATCCGAAACGAAAAGCGGAA
CGGGGGTATGTACGCGCGGTAAGCACTACCTTATAGTGTACAGTTTGTCTGTAGGCTTTGCTTTTCGCCTT

255 260 265 270 275
Cys Arg His Thr Val Arg Pro Phe Val Met Glu Tyr His Ser Gln Asn Asp Ile Arg Asn Glu Lys Arg Lys
TP84_63

GTGGAAATCGTTCGATCCAGAAGGTGATCCGCGGACCGATGCACAAAAACGACTATACAAGGCCGAACAA
CACCTTTAGCAAGCTAGGTCTTCCACTAGGCGCCTGGCTACGTGTTTTTGTCTGATATGTTCCGGCTTGT

280 285 290 295 300
Trp Lys Ser Phe Asp Pro Glu Gly Asp Pro Arg Thr Asp Ala Gln Lys Arg Leu Tyr Lys Ala Glu Gln
TP84_63

GACATTCGCCGCAAAAGCGCGGCAGGAAATGAAAGAATACATGCTGATGAAAGCCACACTGCCGCCGGATC
CTGTAAGCGGCGTTTCGCGCCGTCCTTTACTTTCTTATGTACGACTACTTTCGGTGTGACGGCGGCTAG

305 310 315 320
Asp Ile Arg Arg Lys Ala Arg Gln Glu Met Lys Glu Tyr Met Leu Met Lys Ala Thr Leu Pro Pro Asp
TP84_63

AAGTGCCGAAAACACTGGCCGCTTATCGCCGGATGAAGCGGAAGAACGATGCACAGTGGCAAAAATTACA
TTCACGGCTTTTGTGACCGGCGAATAGCGGCCCTACTTCGCCTTCTTGCTACGTGTCACCGTTTTTAATGT

325 330 335 340 345
Gln Val Pro Lys Thr Leu Ala Ala Tyr Arg Arg Met Lys Arg Lys Asn Asp Ala Gln Trp Gln Lys Leu Gln
TP84_63

GGCACAATTCAAACAAGCACTGGGCGAAGTGGATCTGCAAGGGCCGCCAGGAACGCCACGAAAGCGG
CCGTGTTAAGTTTGTTCGTGACCCGCTTACCTAGACGTTCCCGGCGGCGGTCTTGCGGTGCTTTTCGCC

350 355 360 365 370
Ala Gln Phe Lys Gln Ala Leu Gly Glu Val Asp Leu Gln Gly Pro Pro Pro Gly Thr Pro Arg Lys Arg
TP84_63

AAAGGGGCAAAAACCTGGTCCTGGATCATCGACACCGAAAACGGACAAAACGGACAAAAACAAAACGTGT
TTTCCCCGTTTTTGACCAGGACCTAGTAGCTGTGGCTTTTGCTGTTTTTGCTGTTTTTTGTTTTGCACA

375 380 385 390
Lys Gly Ala Lys Thr Gly Pro Gly Ser Ser Thr Pro Lys Thr Asp Lys Thr Asp Lys Lys Gln Asn Val
TP84_63

CCACTTTGTCCACACTGCCAAAAGGACCGAAAACCTGACGGAAGCGGATCTGGCGGCACACAAAGAGTTTGA
GGTGAAACAGGTGTGACGGTTTTCTGGCTTTGACTGCCTTCGCTAGACCGCCGTGTGTTTTCTCAAAC

395 400 405 410 415
Ser Thr Leu Ser Thr Leu Pro Lys Gly Pro Lys Leu Thr Glu Ala Asp Leu Ala Ala His Lys Glu Phe Asp
TP84_63

CAACGCATCGATGGCATCGAATTCGCTGATCGCAGACAGACAGCAAGATTCTTACTTGATCAGATCCCT
GTTTGGTAGCTACCGTAGCTTAAGCGACTAGCGTCTGCTGCTGCTTCTAAGAATGAAGTCTAGGGA
420 425 430 435 440
Lys Arg Ile Asp Gly Ile Glu Phe Ala Asp Arg Arg Gln Thr Ala Arg Phe Leu Leu Asp Gln Ile Pro
TP84_63 >

GGCGGCAACCAGATCAAAGTCAGCATCCGAAAAATTGAAGCGAATGGCCACGTGGGAAAATTCAGCGTGG
CCGCCGTTGGTCTAGTTTCAGTCGTAGGCTTTTAACTTCGCTTACCGGTGCACCCTTTTAAATCGCACC
445 450 455 460
Gly Gly Asn Gln Ile Lys Val Ser Ile Arg Lys Ile Glu Ala Asn Gly His Val Gly Lys Phe Ser Val
TP84_63 >

TTGGCGGAAAATTCATAATCGGTGAATATGCATTGTTGAAAGATGATCCGCGGCCGCCGACGATATCGATG
AACCGCCTTTTAAAGTATTAGCCACTTATACGTAACAACTTTCTACTAGCGCGCGCGGCGTCATAGCTAC
465 470 475 480 485
Val Gly Gly Lys Phe Ile Ile Gly Glu Tyr Ala Leu Leu Lys Asp Asp Pro Arg Pro Pro Gln Tyr Arg Trp
TP84_63 >

GAAAACGGTCTTTTCATGAATACTACCATGCACAGATGCATGGCATGGATTATGACGATCATTTTCAGCGT
CTTTTGCCAGAAAGTACTTATGATGGTACGTGTCTACGTACCGTACCTAATACTGCTAGTAAAAGTCGCA
490 495 500 505 510
Lys Thr Val Phe His Glu Tyr Tyr His Ala Gln Met His Gly Met Asp Tyr Asp Asp His Phe Gln Arg
TP84_63

GATAACCAGGAATGGACTATCTGGGAAGAAACAGCCACCGAATGTGCGGCATTCTTCATGGCGCAAAGGG
CTATTGGTCCTTACCTGATAGACCTTCTTTGTGCGGTGGCTTACACGCCGTAAGAAATACCGCGTTTCCC
515 520 525 530
Asp Asn Gln Glu Trp Thr Ile Trp Glu Glu Thr Ala Thr Glu Cys Ala Ala Phe Phe Met Ala Gln Arg
TP84_63 >

CCGGATGGAATGTGTCCACGATCCTACCGTCTTATTCAGAATATCTGATCCGCACACTGCCGATGCTGAA
GGCCTACCTTACACAGGTGCTAGGATGGCAGAATAAGTCTTATAGACTAGGCGTGTGACGGCTACGACTT
535 540 545 550 555
Ala Gly Trp Asn Val Ser Thr Ile Leu Pro Ser Tyr Ser Glu Tyr Leu Ile Arg Thr Leu Pro Met Leu Lys
TP84_63 >

ACAACTGCCAGAATACGCCGATTGTGAAACGATGGTGCATTTTGGTGCAAAGTTTATGAAGTATCGATTT
TGTTGACGGTCTTATGCGGCTAACACTTTGTACCAGCTAAACCACGTTTCAAATACTTCATAGCTAAA
560 565 570 575 580
Gln Leu Pro Glu Tyr Ala Asp Cys Glu Thr Met Val Asp Phe Gly Ala Lys Phe Met Lys Tyr Arg Phe
TP84_63

AATGAAGAACATAAAACGATGGAATGGCGGTGGCTAGTCAGACGTGTGGAACAAAATGCAGATGTGTTGG
TTACTTCTTGATTTTGTACCTTACCGCCACCGATCAGTCTGCACACCTTGTTTTACGTCTACACAAGC
585 590 595 600
Asn Glu Glu His Lys Thr Met Glu Trp Arg Trp Leu Val Arg Arg Val Glu Gln Asn Ala Asp Val Phe
TP84_63

ACATCAACGAATATTTCCGGGAACACTACCTGGAACATATCAAAGCGAACAAAGACCGATATGCCGATAT
TGTAGTTGCTTATAAAGGCCCTTGTGATGGACCTTGATAGTTTCGCTTGTTTCTGGCTATACGECTATA
605 610 615 620 625
Asp Ile Asn Glu Tyr Phe Arg Glu His Tyr Leu Glu His Ile Lys Ala Asn Lys Asp Arg Tyr Ala Asp Met
TP84_63 >

GATTACGATTCACTGTCACAGTCGAACAAATTTACACACGAATCCGTGATCCGGCAGATGATCAAAGAA
CTAAATGCTAAGTGACAGTGTCTAGCTTGTTTAAATGTGTGCTTAGGCACTAGGCCGTCTACTAGTTTCTT

29 400

630 635 640 645 650
Ile Tyr Asp Ser Leu Ser Gln Ser Asn Lys Phe Thr His Glu Ser Val Ile Arg Gln Met Ile Lys Glu
TP84_63

GGCCTGGAATCTGGGATCAGATTCTGGCACGGTGAACCGGGAAATTCGCTTATGTCTGCCAGTCGCGATGA
CCGGACCTTAGACCTAGTCTAAGCCGTGCCACTTGGCCCTTTAAGCGAATACAGACGGTCAGCGCTACT

29 470

655 660 665 670
Gly Leu Glu Ser Gly Ile Arg Phe Gly Thr Val Asn Arg Glu Ile Arg Leu Cys Leu Pro Val Ala Met
TP84_63

ATGAACTGGGGGTGAAGAAGCCATGATCGCGATCTATGATCTGATCGGAAAAATTCAAAACATGGACAAC
TACTTGACCCCACTTCTTCGGTACTAGCGCTAGATACTAGACTAGCCTTTTAAAGTTTTGTACCTGTTG

29 540

675 680
Asn Glu Leu Gly Val Lys Lys Pro
TP84_63

RBS

1 5 10 15
Met Ile Ala Ile Tyr Asp Leu Ile Gly Lys Ile Gln Asn Met Asp Asn
TP84_64

ATGGACAAACTGGATGAAATCATCAGTCGATGGTACTTTCAAACATCCGTATGGACACCACAAGTGAAAG
TACCTGTTTGACCTACTTTAGTAGTCAGCTACCATGAAAGTTGTAGGCATACCTGTGGTGTTCACTTTC

29 610

20 25 30 35
Met Asp Lys Leu Asp Glu Ile Ile Ser Arg Trp Tyr Phe Gln Thr Ser Val Trp Thr Pro Gln Val Lys
TP84_64

AAGAAGTCAAAAAAGCGGCGGTGAAATCCTGCAACTAGGCGAACCGAAAGTGTAACAGTTTATTGTGGA
TTCTTCAGTTTTTTCGCCGCCAGCTTTAGGACGTTGATCCGCTTGGCTTTCACATGGTCAAATAACACCT

29 680

40 45 50 55 60
Glu Glu Val Lys Lys Ala Ala Val Glu Ile Leu Gln Leu Gly Glu Pro Lys Val Tyr Gln Phe Ile Val Asp
TP84_64

TATGTACGGCCTTAAATGTGAACAATGTGGACAACCTGAATGGAAAAATATTCACCTTTTAGCCGAATGT
ATACATGCCGGAATTTTACACTTGTTACACCTGTTGACTTACCTTTTATAAGTGAAAAATCGGCTTACA

29 750

65 70
Met Tyr Gly Leu Lys Met
TP84_64

regulatory region

TTGGTATAATAACGATAAAGACAAGCAATCGGGCCACAAACCGTGTGAGTGGCCCGGCCTGGGATCATTG
AACCATATTATTGCTATTTCTGTTTCGTTAGCCCGGTGTTTGGCACACTACCGGGCCGGACCCTAGTAAC

29 820

regulatory region

GATTGTCTGATCCTGGGTGGTTTCATCAGTGAACGTGTTTTTTTCGGCCTTTTCCTGGTGACGGGCCTT
CTAACAGACTAGGACCCACCAAAGTAGTGCACTTGACAAAAAAGCCGGAAAAGGACCACTGCCCGGAA

29 890

AAACGAAACCAGAGATCGTGCAAGGGGGAAGAAAGTAAATGCAAGACAAACAAATCACGGATGAATTCA
TTTGCTTTGGTCTCTAGCACGTTCCCCCTTCTTTTCAATTTACGTTCTGTTTGTGTTAAGTGCCTACTTAAGT

29 960

RBS

1 5 10
Met Gln Asp Lys Gln Ile Thr Asp Glu Phe
TP84_65

TCCAAAAAGCGAATAAAGCATCCGAATTCGGATTCTGCAATTTTTTGC CGGAGAGGGTGGACAGCAAGG 30 030
AGGTTTTTCGCTTATTTTCGTAGGCTTAAGCCTAAGGACGTTAAAAACGGCCTCTCCACCTGTCGTTCC
15 20 25 30
Ile Gln Lys Ala Asn Lys Ala Ser Glu Phe Gly Phe Leu Gln Phe Phe Ala Gly Glu Gly Gly Gln Gln Gly
TP84_65

CGGCGAACCTACTGATCCAACCGATCCAACCTGATCCGCTTGACGATCCAGAAGGCGGCGAACCAACGAT 30 100
GCCGCTTGGATGACTAGGTTGGCTAGGTTGACTAGGCGAACTGCTAGGTTCTCCGCCGCTTGGTTTGCTA
35 40 45 50 55
Gly Glu Pro Thr Asp Pro Thr Asp Pro Thr Asp Pro Leu Asp Asp Pro Glu Gly Gly Glu Pro Asn Asp
TP84_65

CCGACTGATCCAACCTGATCCAGTCGATGATCAAGATGATAACGATGACGATACTGGGATCACATATGAAA 30 170
GGCTGACTAGGTTGACTAGGTCAGCTACTAGTTCTACTATTGCTACTGCTATGACCCTAGTGATACTTT
60 65 70 75 80
Pro Thr Asp Pro Thr Asp Pro Val Asp Asp Gln Asp Asp Asn Asp Asp Thr Gly Ile Thr Tyr Glu
TP84_65

AAGGCTTGATTGATCGCATTTTTAAAGCAAATAACATCGATTTTGGTCAATTGTTGAAGGACAATCCAGA 30 240
TTCCGAACCTAAGCTAGCTAAATTTTCGTTTATTGTAGCTAAAACCAGTTAACAACCTCCTGTTAGGTTCT
85 90 95 100
Lys Gly Leu Ile Asp Arg Ile Leu Lys Ala Asn Asn Ile Asp Phe Gly Gln Leu Leu Lys Asp Asn Pro Glu
TP84_65

ACTGAAACGGCAATATCAAGCAAGATTCAACAAAAATATGTCCAAACGCCTTGAAAAATACCAGGGTGTC 30 310
TGACTTTGCCGTTATAGTTGCTTCTAAGTTGTTTTTATACAGGTTTGC66AACTTTTATGGTCCCACAG
105 110 115 120 125
Leu Lys Arg Gln Tyr Gln Ala Arg Phe Asn Lys Asn Met Ser Lys Arg Leu Glu Lys Tyr Gln Gly Val
TP84_65

GATGTCGATGAATATTTTGAAGTGAAGAAACGTGCAGAAAGCGGCAATCTCGAAGGTGACGCAAGACTTT 30 380
CTACAGCTACTTATAAACTTGACTTTTTTGACAGCTCTTTCGCCGTTAGAGCTTCCACTGCGTTTCTGAA
130 135 140 145 150
Asp Val Asp Glu Tyr Phe Glu Leu Lys Lys Arg Ala Glu Ser Gly Asn Leu Glu Gly Asp Ala Lys Thr
TP84_65

GGAAAGAAAAATACGAACAATTGAAAGCAGAAATGGAAACAACAACAAAGAAAACAGCCATTCAACAGTA 30 450
CCTTTCTTTTATGCTTGTTAACTTTCGTCTTTACCTTTGTTGTTGTTCTTTTGTGCGTAAGTTGTCAT
155 160 165 170
Trp Lys Glu Lys Tyr Glu Gln Leu Lys Ala Glu Met Glu Thr Thr Thr Lys Lys Thr Ala Ile Gln Gln Tyr
TP84_65

TGCGATCGAAAACGGATTGATTCCGAACAGATCGCATTTCATCACGTCAATGATCGACATGAACAAGCTG 30 520
ACGCTAGCTTTTGCCTAAGCTAAGGCTTGCTAGCGTAAGTAGTGCAGTTACTAGCTGTACTTGTTCGAC
175 180 185 190 195
Ala Ile Glu Asn Gly Phe Asp Ser Glu Gln Ile Ala Phe Ile Thr Ser Met Ile Asp Met Asn Lys Leu
TP84_65

GAACGTGATGACGAAGGTGAATGGATGGGCATTGACGAGGAAGTCGAACGAATCAAAGAAAAATTTCCAC 30 590
CTTGCACTACTGCTTCCACTTACCTACCCGTAAGTGTCTCCTCAGCTTGCTTAGTTTCTTTTAAAGGTG
200 205 210 215 220
Glu Arg Asp Asp Glu Gly Glu Trp Met Gly Ile Asp Glu Glu Val Glu Arg Ile Lys Glu Lys Phe Pro
TP84_65

GGATGTTTGAACACGTCAGGTCAAGTCGGCGGCACACCGAAAAAGAATCGAAATATAATCCTGGCAC
CCTACAAGCTTGGTGCACCTGCCAGTTACGCCGCGGTGTGGCTTTTCTTAGCTTTATATTAGGACCGTG
225 230 235 240
Arg Met Phe Glu Pro Arg Asp Gly Gln Val Gly Gly Thr Pro Lys Lys Glu Ser Lys Tyr Asn Pro Gly Thr
TP84_65

BssSI

GAAAAAGCACAATCAAGACACGAAACCGACTGATCCACGAGAAGTGGGCCGACAAATCGCATTACAGCGG
CTTTTTCGTGTTAGTTCTGTGCTTTGGCTGACTAGGTGCTCTTGACCCGGCTGTTTAGCGTAATGTCGCC
245 250 255 260 265
Lys Lys His Asn Gln Asp Thr Lys Pro Thr Asp Pro Arg Glu Leu Gly Arg Gln Ile Ala Leu Gln Arg
TP84_65

CATAAAACCGCTTGAAACAACAATAAGGAGGTAAAAACACATGAACTTGCAACCGAAAATTGTATCGAT
GTATTTTGGCGAACTTTGTTGTTATTCCTCCATTTTGTGTACTTGAACGTTGGCTTTTAACATAGCTA
270 275 1 5 10
His Lys Asn Arg Leu Lys Gln Gln Met Asn Leu Gln Pro Lys Ile Val Ser Ile
TP84_65 RBS TP84_66

CGCAGGCCAGAAAGAATTCCTGGCCACTACACAAGGCCTTGTCCACAAAGTGGGTGGCGTGACGCTTGAC
GCGTCCGGTCTTTCTTAAGGACCGGTGATGTGTTCCGGAACAGGTGTTTCACCCACCGCACTGCGAACTG
15 20 25 30
Ala Gly Gln Lys Glu Phe Leu Ala Thr Thr Gln Gly Leu Val His Lys Val Gly Gly Val Thr Leu Asp
TP84_66

GCTTCAAAATTCACGCCAGATGAAAAATGGATTTCATCAAGGCCGGATCTGCACTTGCAATTGACAGCAAGCG
CGAAGTTTTAAGTGCGGTCTACTTTTACCTAAGTAGTTCCGGCTAGACGTGAACGTAACGTGTCGTTCCG
35 40 45 50 55
Ala Ser Lys Phe Thr Pro Asp Glu Asn Gly Phe Ile Lys Ala Gly Ser Ala Leu Ala Leu Thr Ala Ser
TP84_66

GAAAAGCTGAACCATTCACGTTTCAACACCTGGCGATCCATCCACTGCAAACGGAACGCCATACATTCT
CTTTTCGACTTGGAAGTTGCAAAGTTGTGGACCGCTAGGTAGGTGACGTTTGCTTGCCTTGCAGTATGTAAGA
60 65 70 75 80
Gly Lys Ala Glu Pro Phe Asn Val Ser Thr Pro Gly Asp Pro Ser Thr Ala Asn Gly Thr Pro Tyr Ile Leu
TP84_66

TGCACACGATGTGCAAATCAAAGACGGCACAACCAACATCGATGCTGTCGCCGGAGTGTTAGAAGCCGCT
ACGTGTGCTACACGTTTAGTTTCTGCCGTGTTGGTTGTAGCTACGACAGCGGCCTCACAATCTTCGGCGA
85 90 95 100
Ala His Asp Val Gln Ile Lys Asp Gly Thr Thr Asn Ile Asp Ala Val Ala Gly Val Leu Glu Ala Ala
TP84_66

TATTTAAAATCGTCAGTGGTCACAACCGCTGAACCTGGTCTGCTGCTGGTGACACAAGATTTTATCGATG
ATAAATTTTAGCAGTCACCAAGTGTGGCGACTTGGACCAGCACAGCACTGTGTTCTAAAATAGCTAC
105 110 115 120 125
Tyr Leu Lys Ser Ser Val Val Thr Thr Ala Glu Pro Gly Arg Val Val Val Thr Gln Asp Phe Ile Asp
TP84_66

CATCGAACGGTCGATTCCACCTTCGATAATGGTGGACAAGCCGGACAATATTCATTTTAAAGGAGGTAAAA
GTAGCTTGCCAGCTAAGGTGGAAGCTATTACCACCTGTTCCGGCTGTTATAAGTAAAATTCCTCCATTTT
130 135
Ala Ser Asn Gly Arg Phe His Leu Arg
TP84_66 RBS

AACGATGCCTTTACATTTAGAACAATTCCAACGTGAAGCATTCCAGGGGTACGTTGAAAACGTGCCGCCA
TTGCTACGGAAATGTAATCTTGTTAAGGTTGCACTTCGTAAGGTCCCATGCAACTTTTGCACGGCGGT
1 5 10 15 20
Met Pro Leu His Leu Glu Gln Phe Gln Arg Glu Ala Phe Gln Gly Tyr Val Glu Asn Val Pro Pro
TP84_67

AAACGTGAATATGCACTGGCAAAATTCATGCCGAATCAGCCAGTCTATGACATTGAATTCACATACAACA
TTTGCACCTTATACGTGACCGTTTTAAGTACGGCTTAGTCGGTCAGATACTGTAACCTTAAGTGTATGTTGT
25 30 35 40 45
Lys Arg Glu Tyr Ala Leu Ala Lys Phe Met Pro Asn Gln Pro Val Tyr Asp Ile Glu Phe Thr Tyr Asn
TP84_67

AscI

TCATCAATGGCGGATATGGACAAATGGCATCCATCACCAGCATGGGATTCTGGCGCGCGGTTGCGTGACAA
AGTAGTTACCGCCTATACCTGTTTACCGTAGGTAGTGGCGTACCCTAAGACCGCGCGGCAACGCACTGTT
50 55 60 65
Ile Ile Asn Gly Gly Tyr Gly Gln Met Ala Ser Ile Thr Ala Trp Asp Ser Gly Ala Pro Leu Arg Asp Lys
TP84_67

AGATGTGATTCAACGCTTGACAGCGCAAATCGCCAAAGTGCAACACGCATATCGCCTTACTGAAAAAGAA
TCTACACTAAGTTGCGAACTGTCGCGTTTAGCGGTTTCACGTTGTGCGTATAGCGGAATGACTTTTTCTT
70 75 80 85 90
Asp Val Ile Gln Arg Leu Thr Ala Gln Ile Ala Lys Val Gln His Ala Tyr Arg Leu Thr Glu Lys Glu
TP84_67

CTGTTAATGTTTACCGGCCACGTATGGATGAAGAACAGCAACAAGTCATTCAAGCGATCTACAACAACA
GACAATTACAAAGTGGCCGGTGTCATACCTACTTCTTGTCGTTGTTTCAGTAAGTTGCTAGATGTTGTTGT
95 100 105 110 115
Leu Leu Met Phe His Arg Pro Arg Met Asp Glu Glu Gln Gln Gln Val Ile Gln Ala Ile Tyr Asn Asn
TP84_67

CAGACAAGCTTGTCTGGGGTGTTCAAGACCGTGAAGAATGGTTGCGCGCTAAAGCTGTTTATGTCGGACA
GTCTGTTTGAACAGACCCACAAGTTCTGGCACTTCTTACCAACGCGCGATTTGACAAATACAGCCTGT
120 125 130 135
Thr Asp Lys Leu Val Trp Gly Val Gln Asp Arg Glu Glu Trp Leu Arg Ala Lys Ala Val Tyr Val Gly Gln
TP84_67

ATTACAATACAGCGAAAACGATGTGCAATTGAACATCGACTTCTTGATTCCGGCAGAAAACAAATTGACT
TAATGTTATGTCGCTTTTGTACACGTTAACTTGTAGCTGAAGAACTAAGGCCGTCTTTTGTTTAACTGA
140 145 150 155 160
Leu Gln Tyr Ser Glu Asn Asp Val Gln Leu Asn Ile Asp Phe Leu Ile Pro Ala Glu Asn Lys Leu Thr
TP84_67

GCCGATGTGGACTGGTCTGATCCGACTGCGCCAGTGATCCAACACTTACAAAGCGCAGTGCAACGATTCA
CGGCTACACCTGACCAGACTAGGCTGACGCGGTCACTAGGTTGTGAATGTTTCGCGTCACGTTGCTAAGT
165 170 175 180 185
Ala Asp Val Asp Trp Ser Asp Pro Thr Ala Pro Val Ile Gln His Leu Gln Ser Ala Val Gln Arg Phe
TP84_67

AAGAAGCAAACAACGGCGAAAAGCCAGTCGAAATGCATATGTCTAGCCGTGTAGAAACATGGTTATTACA
TTCTTCGTTTGTGCGCGCTTTTCGGTCAGCTTTACGTATACAGATCGGCACATCTTTGTACCAATAATGT
190 195 200 205
Lys Glu Ala Asn Asn Gly Glu Lys Pro Val Glu Met His Met Ser Ser Arg Val Glu Thr Trp Leu Leu Gln
TP84_67

AAACGAGCAAGTGAAAGCACACATCTATGGAACACGACTGATCCGCGCATCGTGACAAGCGAACAATTA
 TTTGCTCGTTCACTTTTCGTGTGTAGATACCTTTGTGCTGACTAGCGCGTAGCACTGTTTCGCTTGTAAAT
 210 215 220 225 230
 Asn Glu Gln Val Lys Ala His Ile Tyr Gly Asn Thr Thr Asp Pro Arg Ile Val Thr Ser Glu Gln Leu
 TP84_67

CAACAATTATTCAGCGCGTTGTGCTTCCCGCCATATCGTGTGATCGATGAACAAGTGATTGGCGAAAACG
 GTTGTTAATAAGTCGCGCAACAGCAACGGCGGTATAGCACACTAGCTACTTGTTCCTAACCCTTTTGC
 235 240 245 250 255
 Gln Gln Leu Phe Ser Ala Leu Ser Leu Pro Pro Tyr Arg Val Ile Asp Glu Gln Val Ile Gly Glu Asn
 TP84_67

GCGCGGAAGCACTCATGCCAGAAGATCGTGTCTTACTTGGTGAAGAACTTGGACACACAATGGAAGG
 CGCGCCTTCGTGAGTACGGTCTTCTAGCACAGCAGAATGAACCACTTCTTGAACCTGTGTGTACCTTCC
 260 265 270 275
 Gly Ala Glu Ala Leu Met Pro Glu Asp Arg Val Val Leu Leu Gly Glu Glu Leu Gly His Thr Met Glu Gly
 TP84_67

ACCGACAGTCGAAAACAAC TACAAGCCTGECATTTATGTCATTCCAGAAATCAAAGAAACAAATCCACCA
 TGGCTGTCAAGCTTTTGTGATGTTTCGGACCGTAAATACAGTAAGGTCTTTAGTTTCTTTGTTTAGGTGGT
 280 285 290 295 300
 Pro Thr Val Glu Asn Asn Tyr Lys Pro Gly Ile Tyr Val Ile Pro Glu Ile Lys Glu Thr Asn Pro Pro
 TP84_67

CGCCAGGAAGTATATGTCGGAATCTGTATTTCCGGCATTAGAACGGCCACAAGCAGTCGTTCAATTTAA
 GCGGTCCTTCATATACAGCCTTTTAGACATAAAGGCCGTAATCTTGCCGGTGTTTCGTACAGCAAGTAAATT
 305 310 315 320 325
 Arg Gln Glu Val Tyr Val Gly Lys Ser Val Phe Pro Ala Leu Glu Arg Pro Gln Ala Val Val His Leu
 TP84_67

TCGTTGCACAATCCTAATGACAAAAAAGGGATAGACACATGATCTAGTCTATCCCTTTTTTAAAAAATAC
 AGCAACGTGTTAGGATTACTGTTTTTCCCTATCTGTGTACTAGATCAGATAGGGAAAAAATTTTTTATG
 330
 Ile Val Ala Gln Ser
 TP84_67

regulatory region

ATTCAAAGGGGGATTTTTTCATGCCGAAATACATCGCAAACGTCATTTAGTGACACGCACTGGAATCAA
 TAAGTTTCCCCTAAAAAAGTACGGCTTTATGTAGCGTTTGCAGTAAATCACTGTGCGTGACCTTAGTT
 5 10 15
 Met Pro Lys Tyr Ile Ala Lys Arg His Leu Val Thr Arg Thr Gly Ile Lys
 TP84_68

RBS

AAAACCTGGTGATGTGATCGAATACACGAAAGAACAAGCGCAAAAATTATTGGCCGCAGGATTTCATTGAA
 TTTTGGACCACTACACTAGCTTATGTGCTTTCTTGTTTCGCGTTTTTAATAACCGCGCTCCTAAGTAACTT
 20 25 30 35 40
 Lys Pro Gly Asp Val Ile Glu Tyr Thr Lys Glu Gln Ala Gln Lys Leu Leu Ala Ala Gly Phe Ile Glu
 TP84_68

GAAGCGGAAGAAGAAAAAGAAAACAACAAGCAAAGGAAAACAAAAAGCTGATGAAGCGGACAAAGTGGACA
 CTTGCGCTTCTTTCTTTTGTGTTTCGTTTCTTTTGTTCGACTACTTCGCTGTTTCACCTGT
 45 50 55 60
 Glu Ala Glu Glu Lys Lys Thr Thr Ser Lys Gly Lys Gln Lys Ala Asp Glu Ala Asp Lys Val Asp
 TP84_68

AATAAAATGTCCACTATGTCCACATAAGAGGTGAAGGCCATGACAGTCAAAGAATATTTGGCCAAAATT 32 550
TTATTTTACAGGTGATACAGGTGATTCTCCACTTCGGTACTGTCAGTTTCTTATAAACGGTTTTAA

65
Lys
TP84_68

RBS

1 5 10
Met Thr Val Lys Glu Tyr Leu Ala Lys Ile
TP84_69

CCAGGTGCAGATGAATTCCTGGCACTGCCGGAAGCAGAACAAGAAAAATGGATCTTTATGGCCGAAGAAA 32 620
GGTCCACGTCTACTTAAGGACCGTGACGGCCTTCGTCTTGTTCTTTTACCTAGAAATACCGGCTTCTTT

15 20 25 30
Pro Gly Ala Asp Glu Phe Leu Ala Leu Pro Glu Ala Glu Lys Trp Ile Phe Met Ala Glu Glu
TP84_69

CACTGCGTGATCATTTTCATCCGCGTGTGATCACGCCGCGTGTGGTTGCAAGTCATGTATATGTT 32 690
GTGACGCACTAGTAAAAGTAGGCGCACACTAGTGCGGCGCACACCAACGTGACGTTTCAGTACATATAACA

35 40 45 50 55
Thr Leu Arg Asp His Phe His Pro Arg Val Ile Thr Pro Arg Val Val Ala Leu Gln Val Met Tyr Met Phe
TP84_69

CGAAGGCGAATCCGAAGATTTTGCCATGCTAAAACGACAGGGCGTGGATTCTGTTACGACACAGAAGGACTG 32 760
GCTTCCGCTTAGGCTTCTAAAACGGTACGATTTTGCTGTCCCGCACCTAAGCAAGTCGTGTCTTCCTGAC

60 65 70 75 80
Glu Gly Glu Ser Glu Asp Phe Ala Met Leu Lys Arg Gln Gly Val Asp Ser Phe Ser Thr Glu Gly Leu
TP84_69

TCTGTTTCATTCAAATCTGGATCAATCGCGCCGATGGTCCTGGAATCGTTAATCGAAGCGGATCGGCAA 32 830
AGACAAAGTAAGTTTAGACCTAGTTAGCGCGGCTACCAGGACCTTAGCAATTAGCTTCGCTAGCCGTT

85 90 95 100
Ser Val Ser Phe Lys Ser Gly Ser Ile Ala Pro Met Val Leu Glu Ile Val Asn Arg Ser Gly Ser Ala
TP84_69

RBS

AGGTGGGCCGTTTGATATGAAACCACCGATGAATGATCCAGTCGTGCTGAAACAGCCCCAAAAAGTCGAT 32 900
TCCACCCGGCAAACCTATACTTTGGTGGCTACTTACTAGGTACGACGACTTTGTCGGGTTTTTTCAGCTA

105 110
Lys Val Gly Arg Leu Ile
TP84_69

1 5 10 15
Met Lys Pro Pro Met Asn Asp Pro Val Val Leu Lys Gln Pro Lys Lys Val Asp
TP84_70

GGCAAGATCCAAAAAGACGAAAGAGGATTCGCCGAAATGATCGAAATCCCGGTCATGGCCAGGGTGATCA 32 970
CCGTTCTAGGTTTTTCTGCTTTCTCCTAAGCGGCTTTACTAGCTTTAGGGCCAGTACCGGTCCTCACTAGT

20 25 30 35 40
Gly Lys Ile Gln Lys Asp Glu Arg Gly Phe Ala Glu Met Ile Glu Ile Pro Val Met Ala Arg Val Ile
TP84_70

ACGAAACCACCGAAGTGATCGATGAGAAAGGACAAACACGTCAAGCTGTCATGACTGTCATCCTCCCACC 33 040
TGCTTTGGTGGCTTCACTAGCTACTCTTCTGTTTGTGCAGTTTCGACAGTACTGACAGTAGGAGGGTGG

45 50 55 60 65
Asn Glu Thr Thr Glu Val Ile Asp Glu Lys Gly Gln Thr Arg Gln Ala Val Met Thr Val Ile Leu Pro Pro
TP84_70

ATCTGTGACACCATCCCTGGGCGATGAAATACAAGTCGGATCAGACACCGTGGTGATCCTACAAGTCGAA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAGACACTGTGGTAGGGACCCGCTACTTTATGTTTCAGCCTAGTCTGTGGCACCAGTGGATGTTTCAGCTT

70 75 80 85
Ser Val Thr Pro Ser Leu Gly Asp Glu Ile Gln Val Gly Ser Asp Thr Val Val Ile Leu Gln Val Glu
TP84_70

CCACGAAAGTCCTGGTCTGGCAAAAAGACATATTACTGGGTGACGAACTGTGGCACGTGATCACATCTTT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGTGCTTTCAGGACCAGACCGTTTTTCTGTATAATGACCCACTGCTTGACACCGTGCACTAGTGTAGAAA

1 5
Met Ala Arg Asp His Ile Phe
TP84_71

90 95 100 105
Pro Arg Lys Ser Trp Ser Gly Lys Lys Thr Tyr Tyr Trp Val Thr Asn Cys Gly Thr
TP84_70

TCATTGAGTGGGAAGGTATTGACGAACTGGTCGAGTTATTGGAAAACATGGACAGAAAAACAAGACGGA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTAAGCTCACCTTCCATAACTGCTTGACCAGCTCAATAACCTTTGTACCTGTCTTTTGTCTGCCT

10 15 20 25 30
Ser Phe Glu Trp Glu Gly Ile Asp Glu Leu Val Glu Leu Leu Glu Asn Met Asp Arg Lys Thr Arg Arg
TP84_71

TAATGGTCCAGGAATACACAAAATTCGGCCTATTGGTTGAAGAAGGTGCGAGGGCACTTGCACCGAAAGA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTACCAGTCTTATGTGTTTTAAGCCGGATAACCAACTTCTTCCACGCTCCCGTGAACGTGGCTTTCT

35 40 45 50
Ile Met Val Gln Glu Tyr Thr Lys Phe Gly Leu Leu Val Glu Gly Ala Arg Ala Leu Ala Pro Lys Asp
TP84_71

CGAAGGCAACCTGGAAGATTGATCAACGCAGAAAAAGCGAAATCATCGGTGAAGGCGTGGAAAGTGGAA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTTCCGTTGGACCTTCTAAGCTAGTTGCGTCTTTTTCGCTTTTAGTAGCCACTTCCGCACCTTCACTT

55 60 65 70 75
Glu Gly Asn Leu Glu Asp Ser Ile Asn Ala Glu Lys Ala Lys Ile Ile Gly Glu Gly Val Glu Val Glu
TP84_71

GTCGGTGTGCGGATCGGTCTATGGACTGCGGCGACACGAAGAACCACCACGCATCGGAAAATACCCGAAAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCCACAGCCTAGCCAGATACCTGACGCCGCTGTGCTTCTTGGTGGTGCGTAGCCTTTTATGGECTTTA

80 85 90 95 100
Val Gly Val Gly Ser Val Tyr Gly Leu Arg Arg His Glu Glu Pro Pro Arg Ile Gly Lys Tyr Pro Lys
TP84_71

ATGAGCGCGGCGCAAAATTCCTCGAACTACTATGTGAACGGTCTGGGTGCGCGGACCAGATCAAAACCAGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACTCGCGCGCGTTTTAAGGGCTTGATGATACACTTGCCAGACCCACGCGCTGGTCTAGTTTGGTCC

105 110 115 120
Tyr Glu Arg Gly Ala Lys Phe Pro Asn Tyr Tyr Val Asn Gly Leu Gly Ala Arg Thr Arg Ser Lys Pro Gly
TP84_71

ATGGCGCGGCGAAAAAGCCAGGTGCAAAATATCTGCAAAGGGCCGTGGAATTGGTTGTGGATGATTTTGAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACCGCGCGCTTTTTCGGTCCAGCTTTTATAGACGTTTCCCGGCACCTTAACCAACACCTACTAAAATA

125 130 135 140 145
Trp Arg Gly Glu Lys Pro Gly Arg Lys Tyr Leu Gln Arg Ala Val Glu Leu Val Val Asp Asp Phe Asp
TP84_71

CAAATGAACGAGCGCATCCTTGAACGGATCATGGAAGGCAGGCGATAAACATGATCCAGAAGTGGATCAA 33 670
GTTTACTTGCTCGCGTAGGAACCTTGCCTAGTACCTTCCGTCCGCTATTTGTAAGTCTTCACCTAGTT

150 155 160 1 Met Ile Gln Lys Trp Ile Lys
Gln Met Asn Glu Arg Ile Leu Glu Arg Ile Met Glu Gly Arg Arg
TP84_71 TP84_72
RBS

AGATCAATTGGAAGCGGCGATCCCGAACCTGGAATGGACATATGACTACAAAACAGGCAAGGATCACACT 33 740
TCTAGTTAACCTTCGCCGCTAGGGCTTGGACCTTACCTGTATACTGATGTTTTGTCCGTTCCCTAGTGTGA

10 15 20 25 30
Asp Gln Leu Glu Ala Ala Ile Pro Asn Leu Glu Trp Thr Tyr Asp Tyr Lys Thr Gly Lys Asp His Thr
TP84_72

GGCGTGGTCTATCACGAAACGCCAGGGCAGATCAGCCGGGATGATTTTGAGATCATCACACCGTCATATT 33 810
CCGCACCAGATAGTCTTTGCGGTCCCGTCTAGTCGGCCCTACTAAAACTCTAGTAGTGTGGCAGTATAA

35 40 45 50
Gly Val Val Tyr His Glu Thr Pro Gly Gln Ile Ser Arg Asp Asp Phe Glu Ile Ile Thr Pro Ser Tyr
TP84_72

CGGTGTATATCGAAACATCCGACATGAAAAACGCCGAAAAGTGGGCCTGGATCGTTTATGACACTATGAA 33 880
GCCACATATAGCTTTGTAGGCTGTACTTTTTCGGGCTTTTACCCGGACCTAGCAAATACTGTGATACTT

55 60 65 70 75
Ser Val Tyr Ile Glu Thr Ser Asp Met Lys Asn Ala Glu Lys Trp Ala Trp Ile Val Tyr Asp Thr Met Asn
TP84_72

CAAGCGGCGGCAGGAAGTGGCCACCATCGATGATCGGTCATTCCAGGTCATATTTATCGAAGTCACGACA 33 950
GTTCCGCCCGCTCCTTCACCGGTGGTAGCTACTAGCCAGTAAGGTCCAGTATAAATAGCTTCAGTGCTGT

80 85 90 95 100
Lys Arg Arg Gln Glu Val Ala Thr Ile Asp Asp Arg Ser Phe Gln Val Ile Phe Ile Glu Val Thr Thr
TP84_72

CCACCGATCCTGGTGGGGATCGTTGACAAAAAATGACATACTCAATCAATCTACAAGCGACAATCCGCA 34 020
GGTGGCTAGGACCACCCCTAGCAACTGTTTTTTTACTGTATGAGTTAGTTAGATGTTCCGCTGTTAGGCGT

105 110 115 120
Pro Pro Ile Leu Val Gly Ile Val Asp Lys Lys Met Thr Tyr Ser Ile Asn Leu Gln Ala Thr Ile Arg
TP84_72

AAATATAAGGAGGTAGAAAGGTATGCCAGGGGTAAACATTCCATTCCGGTCTTGCAACGATCACAGTGGGT 34 090
TTTATATTCCTCCATCTTTCCATACGGTCCCATTTGTAAGGTAAGCCAGAACGTTGCTAGTGTACCCCA

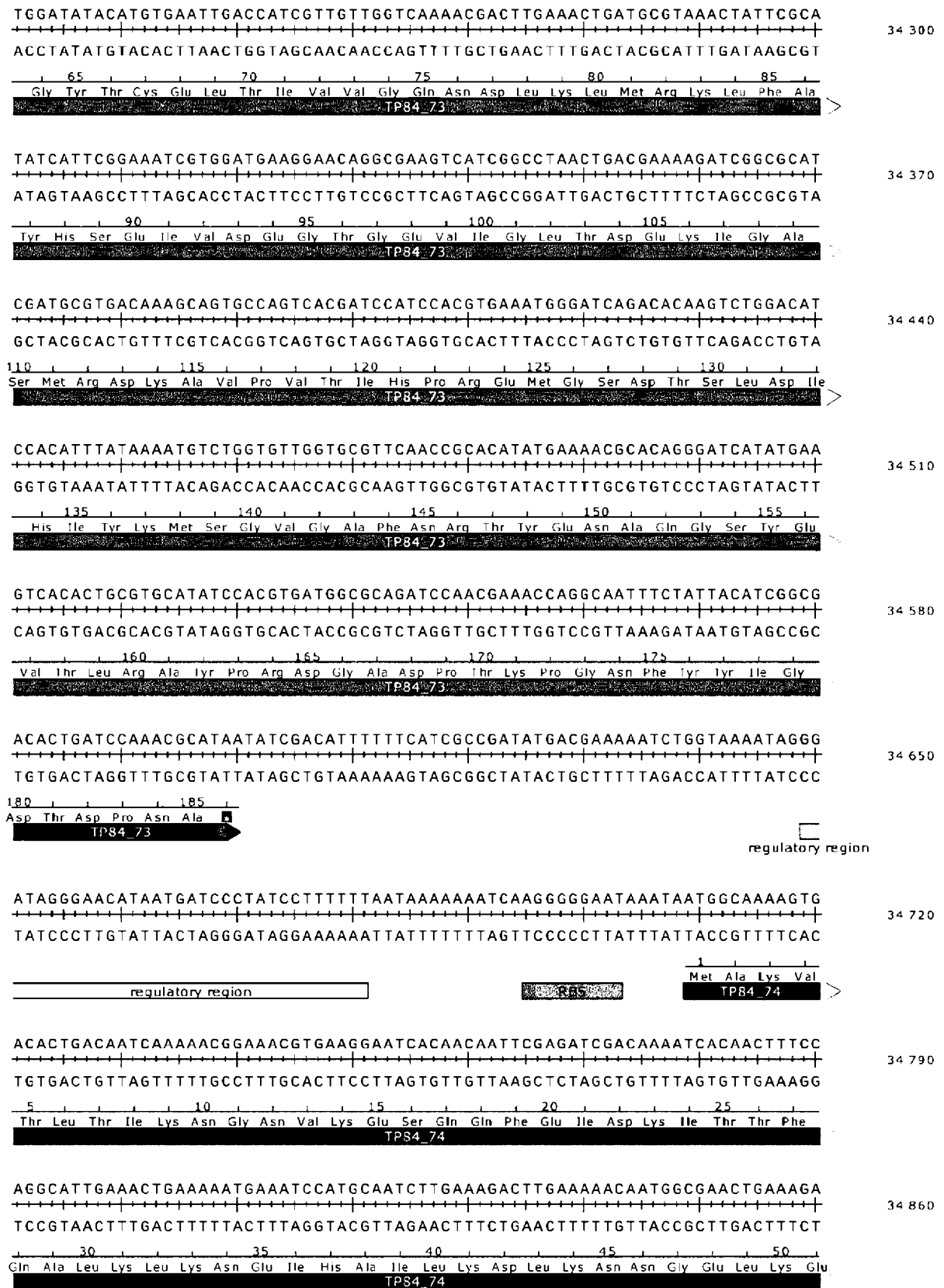
125 1 5 10 15
Lys Ile Met Pro Gly Val Asn Ile Pro Phe Gly Leu Ala Thr Ile Thr Val Gly
TP84_72 TP84_73
RBS

GATCCACAAGATCCGAATAAAATCGTGTTCGATGGCGTGAATTATTTTCAAGCAGAAGGCGGCGAACTTA 34 160
CTAGGTGTTCTAGGCTTATTTTAGCACAAGCTACCGCACTTAATAAAAGTTTCGTCTTCCGCCGCTTGAAT

20 25 30 35
Asp Pro Gln Asp Pro Asn Lys Ile Val Phe Asp Gly Val Asn Tyr Phe Gln Ala Glu Gly Gly Glu Leu
TP84_73

ACATTGAGCCGGAACCTGAACCGATCGAACTGCAAGATTTTGGTGGCAGACCCGTATGATGAACGCATTAA 34 230
TGTAACCTCGGCCTTGAACCTGGCTAGCTTGACGTTCTAAACCACGCTGTGGCATACTACTTGCGTAATT

40 45 50 55 60
Asn Ile Glu Pro Glu Leu Glu Pro Ile Glu Leu Gln Asp Phe Gly Ala Thr Pro Tyr Asp Glu Arg Ile Asn
TP84_73



AGTCATGGAAGGTCTATTTCAGCGGCGGAATTCGATGTGGACAACATGGACATTAAAAACATCACAGCGGAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCAGTACCTTCCAGATAAGTCGCGCGCTTAAGCTACACCTGTTGTACCTGTAATTTTTGTAGTGTGCGCTA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
55 60 65 70
Val Met Glu Gly Leu Phe Ser Gly Glu Phe Asp Val Asp Asn Met Asp Ile Lys Asn Ile Thr Ala Asp
TP84_74

CAACTGGAACAAATGAAAGACGAAAAATTCATCACAAGCCTGGCAGGGGCATTTCGATCGATTGCTTGAAA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTGACCTTGTTTACTTTCTGCTTTTTAAGTAGTGTTGGGACCGTCCCCGTAAGCTAGCTAACGAACTTT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
75 80 85 90 95
Gln Leu Glu Gln Met Lys Asp Glu Lys Phe Ile Thr Ser Leu Ala Gly Ala Phe Asp Arg Leu Leu Glu
TP84_74

CTGTGCCGGAACGTGCGATGAATCTATTGTCCATCATGTCTGGCATCGATCGTGAAGTGTAGAAAAAGC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACACGGCCTTGACCCCTACTTAGATAACAGGTAGTACAGACCGTAGCTAGCACTTCACAATCTTTTTCG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
100 105 110 115 120
Thr Val Pro Glu Arg Ala Met Asn Leu Leu Ser Ile Met Ser Gly Ile Asp Arg Glu Val Leu Glu Lys Ala
TP84_74

GTATCTGGAAGAATTATTCGATGTCTATGATGCTGTTATGGAAGAAAACGACATCATCAAATAATTGAC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATAGACCTTCTTAATAAGCTACAGATACTACGACAATACCTTCTTTTGCTGTAGTAGTTTGATTAAGTG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
125 130 135 140
Tyr Leu Glu Glu Leu Phe Asp Val Tyr Asp Ala Val Met Glu Glu Asn Asp Ile Ile Lys Leu Ile Asp
TP84_74

CGAATGAAACGATCTTTTTTACGACAAAGGGCCAATGGTCGCAGGCACTTCGAACATTTCTGGCCAACA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTTACTTTGCTAGAAAAAAATGCTGTTTCCCGGTTACCAGCGTCCGTAAGCTTGTAAGACCGGTTGT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
145 150 155 160 165
Arg Met Lys Arg Ser Phe Phe Thr Thr Lys Gly Gln Trp Ser Gln Ala Leu Arg Thr Phe Leu Ala Asn
TP84_74

Met Val Ala Gly Thr Ser Asn Ile Ser Gly Gln Gln
RBS TP84_75

AATAAACGATGACACCGATCTATTCGAAATGATCATATACACATTGTCACCAAGTCCTGGGCGGCAGATCC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTATTTGCTACTGTGGCTAGATAAGCTTTACTAGTATATGTGTAACAGTGGTCAGGACCCGCCGTCTAGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Lys
TP84_74

Ile Asn Asp Asp Thr Asp Leu Phe Glu Met Ile Ile Tyr Thr Leu Ser Pro Val Leu Gly Gly Arg Ser
TP84_75

GAAGTCCTGGCCACGCCGTTTGTGAGTGTTTGAATACATCGAACTAGAACGACAGCGGAAAAAGGCGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTTCAGGACCGGTGCGGCAACAGCTCACAAACTTTATGTAGCTTGATCTTGCTGTCGCCCTTTTCCGCC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
40 45 50 55
Glu Val Leu Ala Thr Pro Phe Val Glu Cys Leu Lys Tyr Ile Glu Leu Glu Arg Gln Arg Lys Lys Ala
TP84_75

ATCGATGGAACAGATTTATGGATCTGTTTTATGCACATCCTATGGTGGATAAGAACAAACGGCAAAGATA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAGCTACCTTGCTAAATACCTAGACAAAATACGTGTAGGATACCACCTATTCTTGTGTTGCCGTTTCTAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
60 65 70 75 80
Asp Arg Trp Asn Arg Phe Met Asp Leu Phe Tyr Ala His Pro Met Val Asp Lys Asn Lys Arg Gln Arg Tyr
TP84_75

35 490

35 560

35 630

35 700

35 770

35 840

35 910

35 980

36 050

CGATCGCCGATGCCCAAACAGGTGAATTGGAACGCTTGAAGGAATTCGGTATCACCAAGCAGATGTTGAT
GCTAGCGGCTACGGGTTTGTCCACTTAACCTTGCGAACTTCCTTAAGCCATAGTGGTTCGTCTACAACTA
165 170 175 180
Ala Ile Ala Asp Ala Gln Thr Gly Glu Leu Glu Arg Leu Lys Glu Phe Gly Ile Thr Lys Gln Met Leu Ile
TP84_76

CGACAAGGCCCAAGAACTTTATGGAAAAGAGATCGTCAACGCAAAGGGCCAGATCACTGACATGGAAACA
GCTGTTCCGGGTTCTTGAAATACCTTTTCTCTAGCAGTTGCGTTTCCCGGCTAGTGACTGTACCTTTGT
185 190 195 200 205
Asp Lys Ala Gln Glu Leu Tyr Gly Lys Glu Ile Val Asn Ala Lys Gly Gln Ile Thr Asp Met Glu Thr
TP84_76

ATGAACAAAGCATTATTGCGCATCATGCGTGAACGTTATAGCGGTGGTATGGAATATATATCAAAAACGT
TACTTGTTTCGTAATAAGCGGTAGTACGCACCTGCAATATCGCCACCATACTTATATATAGTTTTTGCA
210 215 220 225 230
Met Asn Lys Ala Leu Phe Ala Ile Met Arg Glu Arg Tyr Ser Gly Gly Met Glu Tyr Ile Ser Lys Thr
TP84_76

TCAATGGTATGTTGGCGAACATCAAAGATTCGATGGGCAACAATCGCCGCCGAAGTAGGAAAACCGATATT
AGTTACCATAACAACGCTTGTAGTTTCTAAGCTACCCGTGTTAGCGGCGGCTTGATCCTTTTGCTATAA
235 240 245 250
Phe Asn Gly Met Leu Ala Asn Ile Lys Asp Ser Met Gly Thr Ile Ala Ala Glu Leu Gly Lys Pro Ile Phe
TP84_76

TGAAAAGTTAAAAGGTCAAATGGAAGATTTAGTGCCGATCATGTCGGCATTACATCGTTCATTGCGGGC
ACTTTTCAATTTTCCAGTTTACCTTCTAAATCACGGCTAGTACAGCCGTAATGTAGCAAGTAAGCGCCG
255 260 265 270 275
Glu Lys Leu Lys Gly Gln Met Glu Asp Leu Val Pro Ile Met Ser Ala Phe Thr Ser Phe Ile Arg Gly
TP84_76

GACATGTCGGGTGCGATGAAAACACTGACGGAAGCATTGCGCGCAAATAAAGCACAGCAGATCATGTCCT
CTGTACAGCCACGCTACTTTTGTGACTGCCTTCGTAAGCCGCGTTTATTTGCTGTCGTCTAGTACAGGA
280 285 290 295 300
Asp Met Ser Gly Ala Met Lys Thr Leu Thr Glu Ala Phe Gly Ala Asn Lys Ala Gln Gln Ile Met Ser
TP84_76

TTTTCCAGACGATCAAAAACGCTGGAATGGGCATAAAAGACTTTTTCTGTGTCACTTGCACCGACAGTGCA
AAAAGGTCTGCTAGTTTTTTCGACCTTACCGGTATTTCTGAAAAAGCACAGTGAACGTGGCTGTACGT
305 310 315 320
Phe Phe Gln Thr Ile Lys Asn Ala Gly Met Gly Ile Lys Asp Phe Phe Val Ser Leu Ala Pro Thr Val Gln
TP84_76

AAACATCGGAACGATCCTGGGGAACATCGCACCGATCATCATCGGTCCACTTGTGATTGCGTTTTAAAGCT
TTTGTAGCCTTGCTAGGACCCCTTGTAGCGTGCTAGTAGTACCAGGTGAACACTAACGCAAATTTTCGA
325 330 335 340 345
Asn Ile Gly Thr Ile Leu Gly Asn Ile Ala Pro Ile Ile Ile Gly Pro Leu Val Ile Ala Phe Lys Ala
TP84_76

ATCGCCGCGGTACTTCCACCAGTTTTAAACACGATCACTGGAATCGTGGCCAAGTTTACTGAATGGGAAG
TAGCGGCGCCATGAAGGTGGTCAAAATTTGTGCTAGTGACCTTAGCACCAGTTCAAATGACTTACCCTTC
350 355 360 365 370
Ile Ala Ala Val Leu Pro Pro Val Leu Asn Thr Ile Thr Gly Ile Val Ala Lys Phe Thr Glu Trp Glu
TP84_76

GATTCATCCCACTGGTGACAGGCCTGGCGGCGGCGATGGCCGTATTCAAGGCGCAGACGGTCGCCACACA
CTAAGTAGGGTGACCACTGTCCGGACCGCCGCCGCTACCGGCATAAGTTCGCGTCTGCCAGCGGTGTGT

375 380 385 390
Gly Phe Ile Pro Leu Val Thr Gly Leu Ala Ala Met Ala Val Phe Lys Ala Gln Thr Val Ala Thr Gln
TP84_76

GATCGTGACAACCGTGGCAACTAGGGCGACACAAATGTGGGCGGCGGCACAAAGATTATTGAACCTGGCG
CTAGCACTGTTGGCACC GTTGATCCCGCTGTGTTTACACCGCCGCCGTGTTTCTAATAACTTGGACCGC

395 400 405 410 415
Ile Val Thr Thr Val Ala Thr Arg Ala Thr Gln Met Trp Ala Ala Ala Gln Arg Leu Leu Asn Leu Ala
TP84_76

ATGTCTATGAATCCGATGACGTTGATCATCGGGCTGATCGTTGGTTTGGTGGCCGCACTGGTTATGGCTT
TACAGATACTTAGGCTACTGCAACTAGTAGCCCGACTAGCAACCAAAACCACGGCGTGACCAATACCGAA

420 425 430 435 440
Met Ser Met Asn Pro Met Thr Leu Ile Ile Gly Leu Ile Val Gly Leu Val Ala Ala Leu Val Met Ala
TP84_76

ATCAACGGTCTGAAACATTCCGCAACATTGTCAATAATGCATGGCAGTCGATCAAATCCGTGGTGTCTGC
TAGTTGCCAGACTTTGTAAGGCGTTGTAACAGTTATTACGTACCGTCAGCTAGTTTAGGCACCACAGACG

445 450 455 460
Tyr Gln Arg Ser Glu Thr Phe Arg Asn Ile Val Asn Asn Ala Trp Gln Ser Ile Lys Ser Val Val Ser Ala
TP84_76

GGTTATCAACTGGTTTGTGACGACAATCCCACAATGGGTGGAAAATGTCAAAAACCTGGTTTGTGGCCTG
CCAATAGTTGACCAAAACACTGCTGTTAGGGTGTACCCACCTTTTACAGTTTTTGACCAACAGCCGGAC

465 470 475 480 485
Val Ile Asn Trp Phe Val Thr Thr Ile Pro Gln Trp Val Glu Asn Val Lys Asn Trp Phe Val Gly Leu
TP84_76

GGTCAACGAATTTCTGAAACGTGGTCAAATATCAAAAACGGCATCGTCAGCATCTGGCAGAATATCATCG
CCAGTTGCTTAAAGACTTTGCACCAAGTTTATAGTTTTTGCCTAGCAGTCGTAGACCGTCTTATAGTAGC

490 495 500 505 510
Gly Gln Arg Ile Ser Glu Thr Trp Ser Asn Ile Lys Asn Gly Ile Val Ser Ile Trp Gln Asn Ile Ile
TP84_76

GAACGGTCTGCCGATTGTCCAGTCTTTTGTCCAGGCTATCGTTGAACGAGCCAAACAGATCGCCACAAA
CTTGCCAGCACGGCTAACAGGTACAGAAAACAGGTCCGATAGCAACTTGCTCGGTTTGTCTAGCGGTGTTT

515 520 525 530
Gly Thr Val Val Pro Ile Val Gln Ser Phe Val Gln Ala Ile Val Glu Arg Ala Lys Gln Ile Ala Thr Asn
TP84_76

TGTGATGAACCTTTATTCGGCCGCTGATCGACTTTTTCAAAAACACCTGGAATAATATCAAATTGCTTGTC
ACACTACTTGAAATAAGCCGGCGACTAGCTGAAAAAGTTTTTGTGGACCTTATTATAGTTTAAACGAACAG

535 540 545 550 555
Val Met Asn Phe Ile Arg Pro Leu Ile Asp Phe Phe Lys Asn Thr Trp Asn Asn Ile Lys Leu Leu Val
TP84_76

CTTCAATTGTTGGTGTGTTTTTGAATCTGCTAGTCGGGAATTTGAAGGTTTGAAAATTTCTCTTTTGG
GAAAGTTAACAACCACAAAAAACTTAGACGATCAGCCCTTAAACTTCCAACTTTTAAAGAGAAAACC

560 565 570 575 580
Leu Ser Ile Val Gly Val Phe Leu Asn Leu Leu Val Gly Asn Phe Glu Gly Leu Lys Ile Ser Leu Leu
TP84_76

GAATCTGGACAGCGATCAAAAACCAGGTCATCAATATCGCCACGACAATCAAGTCAATGGCCGTC AACAT
CTTAGACCTGTCGCTAGTTTTTGGTCCAGTAGTTATAGCGGTGCTGTTAGTTCAAGTTACCGGCAGTTGTA

585 590 595 600
Gly Ile Trp Thr Ala Ile Lys Asn Gln Val Ile Asn Ile Ala Thr Thr Ile Lys Ser Met Ala Val Asn Ile
TP84_76

NmeAIII

TTTCACGGCTTTGAAAAATGGTGTACTTGCAAGTGATCAACGGTTTGAAGTCGCTGGCGATCTCGGCCTGG
AAAGTGCCGAAACTTTTTACCACATGAACGTCAGTTGCCAACTTCAGCGACCGCTAGAGCCGACCC

605 610 615 620 625
Phe Thr Ala Leu Lys Asn Gly Val Leu Ala Val Ile Asn Gly Leu Lys Ser Leu Ala Ile Ser Ala Trp
TP84_76

AACGGTTTGAAGTCGGCGGCGATCTCGGTGTGGACAGGTTTGAAAAAATGGTGTCTGTAATACGGTACAGG
TTGCCAACTTCAGCCGCGCGCTAGAGCCACACCTGTCCAACTTTTTACCACAGCATTTATGCCATGTCC

630 635 640 645 650
Asn Gly Leu Lys Ser Ala Ala Ile Ser Val Trp Thr Gly Leu Lys Asn Gly Val Val Asn Thr Val Gln
TP84_76

CTTTGAAAAACACAGCGATCAACATCGTCAACAGTGTCAAGTCTGGGGTAGTCAATGCGTTTAATTCAGC
GAAACTTTTTGTGTCGCTAGTTGTAGCAGTTGTCACAGTTCAGACCCCATCAGTTACGCAAATTAAGTCG

655 660 665 670
Ala Leu Lys Asn Thr Ala Ile Asn Ile Val Asn Ser Val Lys Ser Gly Val Val Asn Ala Phe Asn Ser Ala
TP84_76

GAAAAACCTGGCAATTTTCAAGCTGGAACGCATTGAAATCTGGTGTATCGAACGCGATCAACAGCGTGAAA
CTTTTTGGACCGTTAAAGTCGGACCTTGCCTAACTTTAGACCACATAGCTTGCCTAGTTGTCGCACTTT

675 680 685 690 695
Lys Asn Leu Ala Ile Ser Ala Trp Asn Ala Leu Lys Ser Gly Val Ser Asn Ala Ile Asn Ser Val Lys
TP84_76

TCCCTGGTCAGCAACATGAAAAACAACATTATCAGCACGATCAAGGGAATCAACTTGTTTCGAAATGGGCA
AGGGACCAGTCGTTGTACTTTTTGTTGTAATAGTCGTGCTAGTTCCCTTAGTTGAACAAGCTTTACCCGT

700 705 710 715 720
Ser Leu Val Ser Asn Met Lys Asn Asn Ile Ile Ser Thr Ile Lys Gly Ile Asn Leu Phe Glu Met Gly
TP84_76

AAAACGTAATCCAGGGATTCATCAAAGGGATCAAATCGATGGTCGGTGGGAAAGCCATCAAAGA
TTTTGCATTAGTGCCCTAAGTAGTTTCCCTAGTTTAACTACCAGCCACGCCAACCTTTTCGGTAGTTTCT

725 730 735 740
Lys Asn Val Ile Gln Gly Phe Ile Lys Gly Ile Lys Ser Met Val Gly Ala Val Gly Lys Ala Ile Lys Glu
TP84_76

AGTCGCATCGAACGTCACGAACAAGATCAAATCGGCACTGGGCATCCATTGCCCATCACGTGTGCTGATG
TCAGCGTAGCTTGCAAGTGCTTGTCTAGTTTATGCCGTGACCCGTAGGTAAGCGGTAGTGCACACGACTAC

745 750 755 760 765
Val Ala Ser Asn Val Thr Asn Lys Ile Lys Ser Ala Leu Gly Ile His Ser Pro Ser Arg Val Leu Met
TP84_76

GAAATCGGCGCATACACTGGTCAAGGATTCGCTATCGGTATTGAGAACATGAAAAAGCGGTCGTCAATG
CTTTAGCCGCGTATGTGACCAAGTTCTAAGCGATAGCCATAACTCTTGTACTTTTTTCGCCAGCAAGTTAC

770 775 780 785 790
Glu Ile Gly Ala Tyr Thr Gly Gln Gly Phe Ala Ile Gly Ile Glu Asn Met Lys Lys Ala Val Val Asn
TP84_76

CAACACAATCCCTG6CCGATGCAACGATCEGAACGATCAGCAGTGCAGAATTGAATCCAACCGAACCACA
38 010
GTTGTGTTAGGGACCGGCTACGTTGCTAGCCTTGCTAGTCGTCACGTCTTAAGTTAGGTTGGCTTGGTGT

795 800 805 810
Ala Thr Gln Ser Leu Ala Asp Ala Thr Ile Gly Thr Ile Ser Ser Ala Glu Leu Asn Pro Thr Glu Pro Gln
TP84_76

AGTTGCGGTGGCGGGCGGCCACAGGAGCGCAGACGAACTATAACGCACCGCTGATGTATGTCGAAAACCAA
38 080
TCAACGCCACCGCCGGGTGTCCTCGCGTCTGCTTGATATTGCGTGGCGACTACATACAGCTTTTGGTT

815 820 825 830 835
Val Ala Val Ala Ala Thr Gly Ala Gln Thr Asn Tyr Asn Ala Pro Leu Met Tyr Val Glu Asn Gln
TP84_76

TACGTGAACGACAATACAGATGTGCGTGACATTTTACATGGTCTGTATAACTTGCAACGCAGATCAGATC
38 150
ATGCACTTGCTGTTATGTCTACACGCACTGTAAAGTGTACCAGACATATTGAACGTTGCGTCTAGTCTAG

840 845 850 855 860
Tyr Val Asn Asp Asn Thr Asp Val Arg Asp Ile Ser His Gly Leu Tyr Asn Leu Gln Arg Arg Ser Asp
TP84_76

GCAAAAAAGGGTGGTGTAGATGCCTATGGGCTATACATTCGGCGGTGCGGCACATTCATGAATTCCTG
38 220
CGTTTTTTCCCACTACATCTACGGATACCCGATATGTAAGCCGCCAGCGTGTAAGTACTTAAGGAC

865
Arg Lys Lys Gly Trp
TP84_76

1 5 10 15
Met Pro Met Gly Tyr Thr Phe Gly Gly Arg His Ile His Glu Phe Leu
TP84_77

RBS

GTCGAAATGACTGGGAAAAATGTGCCGATGACACCGCCGATCAAAAACTATCAGAAGAACAAGGCGGCA
38 290
CAGCTTTACTGACCCTTTTTACACGGCTACTGTGGCGGCTAGTTTTTGGATAGTCTTCTTGTTCGCCGCT

20 25 30 35
Val Glu Met Thr Gly Lys Asn Val Pro Met Thr Pro Pro Ile Lys Asn Leu Ser Glu Glu Gln Gly Gly
TP84_77

TTGACGGCGGATGGGATTTTGGGATTCAATACGAACCGAAAATTATCACCATCGATCATTACATCCTGGC
38 360
AACTGCCGCCCTACCCTAAAACCTAAGTTATGCTTGGCTTTTAATAGTGGTAGCTAGTAATGTAGGACCG

40 45 50 55 60
Ile Asp Gly Gly Trp Asp Phe Gly Ile Gln Tyr Glu Pro Lys Ile Ile Thr Ile Asp His Tyr Ile Leu Ala
TP84_77

GAAAACCAAGGAAGAACGGCAAAACATGATCCGTGAACCTGGCAGGATGGCTGAATCCACGACTGGGTGCA
38 430
CTTTTGGTCCCTTCTTGCCGTTTTGTACTAGGCACTTGACCGTCCTACCGACTTAGGTGCTGACCCACGT

65 70 75 80 85
Lys Thr Arg Glu Glu Arg Gln Asn Met Ile Arg Glu Leu Ala Gly Trp Leu Asn Pro Arg Leu Glu Ala
TP84_77

AGGGAATTGATATTCGATGACGAACCAGACAAAATGTATTATGCAAGATTATCCGAACAATTCGCGCTGG
38 500
TCCCTTAACATAAGCTACTGCTTGGTCTGTTTTACATAATACGTTCTAATAGGCTTGTTAAGCGCGACC

90 95 100 105
Arg Glu Leu Ile Phe Asp Asp Glu Pro Asp Lys Met Tyr Tyr Ala Arg Leu Ser Glu Gln Phe Ala Leu
TP84_77

AAAAGGTTATCGGAACATATAGCGACTTTTCACTGAATTTTATCTGTTATGATCCATTTACTTATTCGGT
38 570
TTTTCCAATAGCCTTGATATCGCTGAAAAGTGACTTAAAGTAGACAATACTAGGTAAATGAATAAGCCA

110 115 120 125 130
Glu Lys Val Ile Gly Thr Tyr Ser Asp Phe Ser Leu Asn Phe Ile Cys Tyr Asp Pro Phe Thr Tyr Ser Val
TP84_77

CCAGGAATACACGCAAAACATCACAGGCAGTGGACAAATCGAACACCTGGGAACACACGTGTCGAAACCG
GGTCCTTATGTGCGTTTGTAGTGTCCGTACCTGTTAGCTTGTGGACCCTTGTGTGCACAGCTTTGGC
135 140 145 150 155
Gln Glu Tyr Thr Gln Asn Ile Thr Gly Ser Gly Gln Ile Glu His Leu Gly Thr His Val Ser Lys Pro
TP84_77

ATCTTGATTGTGGATCATCGCGGCGGATCAGCAACGATCACGAATCAGACACAAGACGGACAAACACAGA
TAGAACTAACACCTAGTAGCGCCGCTAGTCGTTGCTAGTGCTTAGTCTGTGTTCTGCCTGTTTGTGTCT
160 165 170 175
Ile Leu Ile Val Asp His Arg Gly Gly Ser Ala Thr Ile Thr Asn Gln Thr Gln Asp Gly Gln Thr Gln
TP84_77

CGGTGACATTGCGATCGACTACACCGCCAGGTATATTCACAATCGATATGAAGGAAGGAACGGTCAAACCT
GCCACTGTAAGCGTAGCTGATGTGGCGGTCCATATAAGTGTTAGCTATACTTCCTTCCTTGCCAGTTTGA
180 185 190 195 200
Thr Val Thr Phe Ala Ser Thr Thr Pro Pro Gly Ile Phe Thr Ile Asp Met Lys Glu Gly Thr Val Lys Leu
TP84_77

GGGAACGCAAAGCGGCGAAAAATACATCGATTCAATTGAATGGTTTGAATTAGAACAAGGGATGAACACG
CCCTTGCCTTTCGCGCTTTTATGTAGCTAAGTTAACTTACCAAACCTAATCTTGTTCCTACTTGTGC
205 210 215 220 225
Gly Thr Gln Ser Gly Glu Lys Tyr Ile Asp Ser Ile Glu Trp Phe Glu Leu Glu Gln Gly Met Asn Thr
TP84_77

ATCACGCACAGCGGCAACATCAATCAAATTACAGTGAAATATCGGCATACATGGCTATGAAAAAGGGGTG
TAGTGCGTGTGCGCGTTGTAGTTAGTTTAAATGTCACCTTATAGCCGTATGTACCGATACTTTTCCCCAC
230 235 240 245
Ile Thr His Ser Gly Asn Ile Asn Gln Ile Thr Val Lys Tyr Arg His Thr Trp Leu
TP84_77 RBS

ATCGGCATGCGGATTCAAAAAGCGTTATTCACGAATCAGTCAGATTTTATTCAAATAGGATCGAACCAGG
TAGCCGTACCGCTAAGTTTTTCGCAATAAGTGCTTAGTCAGTCTAAAATAAGTTTATCCTAGCTTGGTCC
1 5 10 15 20
Met Ala Ile Gln Lys Ala Leu Phe Thr Asn Gln Ser Asp Phe Ile Gln Ile Gly Ser Asn Gln
TP84_78

TCGGTGTGATCCTTTTCTGTATTATAACGGCGATCAATATCAACAAAACAAAGTGCCATCCAGATGCA
AGCCACAGCTAGGAAAAGACATAATATTGCCGCTAGTTATAGTTGTTTGTTCACCGGTAGGTCTACGT
25 30 35 40 45
Val Gly Val Asp Pro Phe Leu Tyr Tyr Asn Gly Asp Gln Tyr Gln Gln Asn Lys Val Ala Ile Gln Met Gln
TP84_78

ACGATTGCGCAGTCCTGCGGCCATCGGTCCAGGAATCAGAGGGAACGCGCGGCGATGGATGGTGGATGT
TGCTAAGCCGTACAGGACGCCGCTAGCCAGGTCTTAGTCTCCCTTGCGCCGCCGCTACCTACCACCTACA
50 55 60 65
Arg Phe Gly Ser Pro Ala Ala Ile Gly Pro Gly Ile Arg Gly Asn Ala Ala Ala Met Asp Gly Gly Cys
TP84_78

ACCATCTGGCGAAATGTGCCGAATCCGGCGGCTAGAAAACGCATGATCAGTCTATGGATCAAGCCGACAG
TGGTAGACCGCTTTACACGGCTTAGGCCGCCGATCTTTTGGTACTAGTCAGATACCTAGTTTCGGCTGTC
70 75 80 85 90
Thr Ile Trp Arg Asn Val Pro Asn Pro Ala Ala Arg Lys Arg Met Ile Ser Leu Trp Ile Lys Pro Thr
TP84_78

CGGCGGATCTATCACGTTTTTTCAGTGTGATCTGCACAGACCGTGGCGATGGATGGACTGATCATGGGAT
GCGGCCTAGATAGTGCAAAAAAGTCACACTAGACGTGTCTGGCACCCTACCTACCTGACTAGTACCCTA
95 100 105 110 115
Ala Ala Asp Leu Ser Arg Phe Phe Ser Val Ile Cys Thr Asp Arg Gly Asp Gly Trp Thr Asp His Gly Met
TP84_78

GCATCTGGCATTGAACTTCGATAGTGCAAAAGCAAAAAACATTGTTGAATGCGCGGATTATCAGATCAGGA
CGTAGACCGTAACTTGAAGCTATCACGTTTCGTTTTTGTAACTTACGCGCCTAATAGTCTAGTCCT
120 125 130 135
His Leu Ala Leu Asn Phe Asp Ser Ala Lys Gln Lys Thr Leu Leu Asn Ala Arg Ile Ile Arg Ser Gly
TP84_78

TCGGGAACATCACAATGGTACACACATGATCCAGATCAAGGAACGATAGAATTCCAGGCCGATCAATGGT
AGCCCTTGTAGTGTACCATGTGTGTACTAGGTCTAGTTCTTGCTATCTTAAGGTCCGGCTAGTTACCA
140 145 150 155 160
Ser Gly Thr Ser Gln Trp Tyr Thr His Asp Pro Asp Gln Gly Thr Ile Glu Phe Gln Ala Asp Gln Trp
TP84_78

ATCACATTTGCTTCTTTTATGATGCAGATGGCGGCCCTGCCTGGCGGTCGTTATGTTGCAGTGTACGTGAA
TAGTGTAACGAAGAAAATACTACGTCCTACCGCCGGACGGACCGCCAGCAATACAACGTCACATGCACTT
165 170 175 180 185
Tyr His Ile Cys Phe Phe Tyr Asp Ala Asp Gly Gly Leu Pro Gly Gly Arg Tyr Val Ala Val Tyr Val Asn
TP84_78

TAAGACACTGCAACTGCAATCGTTTCGTTGCGCCATTTGATATTACTGCGGCATATACCAGATCGTTCACA
ATTCTGTGACGTTGACGTTAGCAAGCAACGCGGTAAGTATAATGACGCCGTATATGGTCTAGCAAAGTGT
190 195 200 205
Lys Thr Leu Gln Leu Gln Ser Phe Val Ala Pro Phe Asp Ile Thr Ala Ala Tyr Thr Arg Ser Phe Thr
TP84_78

CTTGGCGACATACCAACAACGGATCATATAGCGGCACATATCGATTACGCGGCATGATTGACGAAGTGA
GAACCGCTGTATGTTTGTTCCTAGTATATCGCCGTGTATAGCTAAGTCGCCGTACTAACTGCTTCACT
210 215 220 225 230
Leu Gly Asp Ile Pro Asn Asn Gly Ser Tyr Ser Gly Thr Tyr Arg Phe Ser Gly Met Ile Asp Glu Val
TP84_78

TCCGATTGCACGGTGATGATGTGTGACCATTCAGAAATGCAACAATATTATGACGACATCATGGCCGG
AGGCTAACGTGCCACTACTACACACCTGGTAAGTTCTTTACGTTGTTATAATACTGCTGTAGTACCGGCC
235 240 245 250 255
Ile Arg Leu His Gly Asp Asp Val Trp Thr Ile Gln Glu Met Gln Gln Tyr Tyr Asp Asp Ile Met Ala Gly
TP84_78

AAATTACATCGATGCAGAAACCGAACCAGGAACGATGAAAGTCGGAAAAAATTTTGTGACTGGACAATAT
TTTAATGTAGCTACGTCTTTGGCTTGGTCCTTGCTACTTTACGCCCTTTTAAACACTGACCTGTTATA
260 265 270 275
Asn Tyr Ile Asp Ala Glu Thr Glu Pro Gly Thr Met Lys Val Gly Lys Asn Phe Val Thr Gly Gln Tyr
TP84_78

AATACGAATCTGATGACATGGACATCACCGACAATCGATTTAGGCCAGGATGGATTGATGATTTTGGTC
TTATGCTTAGACTACTGTACCTGTAGTGCGCTGTAGCTAAATCCGGTCCTACCTAAGCTACTAAACCAAG
280 285 290 295 300
Asn Thr Asn Leu Met Thr Trp Thr Ser Pro Thr Ile Asp Leu Gly Gln Asp Gly Phe Asp Asp Phe Gly
TP84_78

305 310 315 320 325
 Arg Val Gln Leu Asn Phe Glu Gln Pro Pro Gly Thr Phe Ile Asn Ile Tyr Thr Arg Thr Ser Asp Asp Gly
 IP84-78

330 335 340 345
 Gln Asn Trp Asp Ala Trp Val Lys Thr Ser Val Asp Gly Thr Ile Asn Ser Ser Asp Lys Arg Tyr Leu
 P84-78

350 355 360 365 370
Gln Ile Lys Ile Glu Phe Gln Thr Thr Asn Gly Ala Ile Thr Pro Lys Ile Met Glu Val Gln Val Leu
TP84-78

375 380 385 390 395
 Asp Tyr Gln Lys Ile Lys Arg Leu Thr Leu Thr Ser Glu Pro Leu Ile Ile Tyr Lys Asp Leu Glu Ser Gly
 IP84/78A

400 405 410 415
 Leu Glu Arg Ile Gly Glu Leu Lys Asn Ala Tyr Asp Val Ile Ile Thr Glu Glu Ile Asn Gly Glu Glu
 TP84:78

420 425 430 435 440
Val Ile Glu Phe Lys Met Ala Ser Asn Asp Pro Lys Arg Ile Glu Leu Gly Ala Glu Pro Val Glu Leu
P84788

445 450 455 460 465
 Ile Ala Arg Ile Gly Asp Lys Gln Phe Ile Ile Arg Asn Ala Ile Asp Lys Arg Asp Glu Asn Gly Lys Lys
 IP84178

Tyr Thr Gln Phe Phe Gly Glu Ala Leu Trp Tyr Glu Leu Arg Asp Ala Lys Val Ile Asn Tyr Glu Gln

490 495 500 505 510
 Val Glu Lys Thr Ala Tyr Glu His Ile Gln Ala Ile Leu Asn Ser Ala Val Val Pro Thr Gly Trp Thr
 TP84 78

TTTACAAAGTCGAATCCGATGGCCGAAAACGAACGATTGCGGGCGAATGGAAATCTGTCCTGGAATTATT
40 530
AAATGTTTCAGCTTAGGCTACCGGCTTTTGCTTGCTAAGCGCCGCTTACCTTTAGACAGGACCTTAATAA

515 520 525 530 535
Ile Tyr Lys Val Glu Ser Asp Gly Arg Lys Arg Thr Ile Arg Gly Glu Trp Lys Ser Val Leu Glu Leu Leu
TP84_78

ACGCGAAGTCGTTGATCAGTTTGGCGGCGAACTGCAATTGACACCATTAAATCGCACAAATCAGCCTGGTG
40 600
TGGGCTTCAGCAACTAGTCAAACCGCCGCTTGACGTTAAGCTGTGGTAATTAGCGTGTTAGTCGGACCAC

540 545 550 555
Arg Glu Val Val Asp Gln Phe Gly Gly Glu Leu Gln Phe Asp Thr Ile Asn Arg Thr Ile Ser Leu Val
TP84_78

AACCGCATCGGTGAAGATAACGGCGTTCGATTCTACTACAATAAGAATCTGAAAACGATCGAACGGTCGG
40 670
TTGGCGTAGCCACTTCTATTGCCGCAAGCTAAGATGATGTTATTCTTAGACTTTTGCTAGCTTGCCAGCC

560 565 570 575 580
Asn Arg Ile Gly Glu Asp Asn Gly Val Arg Phe Tyr Tyr Asn Lys Asn Leu Lys Thr Ile Glu Arg Ser
TP84_78

TGGACACGTACAATCTGATCACACGTCTATATCTATATGGCAAAAACGGCATGACTGTCCAATCTGTCCA
40 740
ACCTGTGCATGTTAGACTAGTGTGCAGATATAGATATACCGTTTTTGCCGTAAGTACAGGTTAGACAGGT

585 590 595 600 605
Val Asp Thr Tyr Asn Leu Ile Thr Arg Leu Tyr Leu Tyr Gly Lys Asn Gly Met Thr Val Gln Ser Val His
P84_78

MluI

TCCACAAGGCCTGGAATACATTGAGGACTTAACATGGGTAAACGCGTTGAATCTGCGAAAATAAGATTGCG
40 810
AGGTGTTCCGGACCTTATGTAACCTGGAATTGTACCCATTGCGCAACTTAGACGCTTTATTCTAAGCG

610 615 620 625
Pro Gln Gly Leu Glu Tyr Ile Glu Asp Leu Thr Trp Val Asn Ala Leu Asn Leu Arg Asn Lys Ile Arg
TP84_78

ATCGGTGTCTGGAAAGATGAACGATACACGATTCCGCAGAACCTTTATGATGATGGCATGAAAATGTTGC
40 880
TAGCCACAGACCTTTCTACTTGCTATGTGCTAAGGCGTCTTGAAATACTACTACCGTACTTTTACAACG

630 635 640 645 650
Ile Gly Val Trp Lys Asp Glu Arg Tyr Thr Ile Pro Gln Asn Leu Tyr Asp Asp Gly Met Lys Met Leu
TP84_78

AAGAAATGGCGAAACCGAATGTGTCTTATGCGATGACTATTGCGGATCTGTCCATGTTGTCCGGCCATGA
40 950
TTCTTTACCGCTTTGGCTTACACAGAATACGCTACTGATAACGCCCTAGACAGGTACAACAGGCCGGTACT

655 660 665 670 675
Gln Glu Met Ala Lys Pro Asn Val Ser Tyr Ala Met Thr Ile Ala Asp Leu Ser Met Leu Ser Gly His Glu
TP84_78

ACACGAATCGATCGGCCTGGGCGACACCGTCTGGGTAGTGGACACCGAACTGATGAATCTGCTTGTGGAA
41 020
TGTGCTTAGCTAGCCGGACCCGCTGTGGCAGACCCATCACCTGTGGCTTGACTACTTAGACGAACACCTT

680 685 690 695
His Glu Ser Ile Gly Leu Gly Asp Thr Val Trp Val Val Asp Thr Glu Leu Met Asn Leu Leu Val Glu
TP84_78

GCGCGCATCGTGGGCGAAAATACAACGTGCGGCAACCGTGAAAAACAGAGTGAATTGAATCAACCGA
41 090
CGCGCGTAGCACGCCGCTTTTATGTTGCACGCCGTTGGCACCTTTTGTCTTCAGCTTAACCTAGTTGGCT

700 705 710 715 720
Ala Arg Ile Val Arg Arg Lys Tyr Asn Val Arg Gln Pro Trp Lys Thr Glu Val Glu Leu Asn Gln Pro
TP84_78

AAAAGGAAGTGGCCGATGCGAATCGGCGTGCTATTGACGATGCAATCGAAACGCTAGTGGAAAGCGATCC
TTTTCTTGACCGGCTACGCTTAGCCGCACGATAACTGCTACGTTAGCTTTGCGATCACCTTTCGCTAGG

725 730 735 740 745
Lys Lys Glu Leu Ala Asp Ala Asn Arg Arg Ala Ile Asp Asp Ala Ile Glu Thr Leu Val Glu Ser Asp Pro
TP84_78

ACTGGACACAAGTGACGTTCAACAGATGACTGTGTTCAATCATTTATTGAATTCACGTGCAGAAGATGGC
TGACCTGTGTTCACTGCAAGTTGTCTACTGACACAAGTTAGTAAATAACTTAAGTGCACGTCTTCTACCG

750 755 760 765
Leu Asp Thr Ser Asp Val Gln Gln Met Thr Val Phe Asn His Leu Leu Asn Ser Arg Ala Glu Asp Gly
TP84_78

ACGAAATACTGGGAAGTGACTGGATCAGACATCAGATCGAACCGTCTGGATTCAAGTGGTAATGCATCTT
TGCTTTATGACCCTTCACTGACCTAGTCTGTAGTGCTAGCTTGGCAGACCTAAGTCACCATTACGTAGAA

770 775 780 785 790
Thr Lys Tyr Trp Glu Val Thr Gly Ser Asp Ile Thr Ile Glu Pro Ser Gly Phe Ser Gly Asn Ala Ser
TP84_78

GGAAAATACAAAGCGGCTATGGAAAAACAAATAAACTGAAACAATCCATTTATGGTGTGTCACATCGTTC
CCTTTTATGTTTCGCGATACCTTTTTGTTTATTTGACTTTGTTAGGTAAATACCACACAGTGTAGCAAG

795 800 805 810 815
Trp Lys Ile Gln Ser Gly Tyr Gly Lys Thr Asn Lys Leu Lys Gln Ser Ile Tyr Gly Val Ser His Arg Ser
TP84_78

GCGGTACACTATCAGTGCATATGTGGCCACAGAAGGAACGATCACCAGGGGAACATCACAGGATGCGTTC
CCGATGTGATAGTCACGTATACACCGGTGCTTCTTCTGCTAGTGCTCCCTTGCTAGTGCTCTACGCAAG

820 825 830 835
Ala Tyr Thr Ile Ser Ala Tyr Val Ala Thr Glu Gly Thr Ile Thr Arg Gly Thr Ser Gln Asp Ala Phe
TP84_78

GTGGGGATCAAGGTGACGATTCCTACACTGAACCAGACAGCGATGGCAAAACATATGAAGAACACTTCC
CACCCCTAGTTCCTACTGCTAAGTGATGTGACTTGGTCTGTCGCTACCGTTTTGTATACTTCTTGTGAAGG

840 845 850 855 860
Val Gly Ile Lys Val Thr Ile His Tyr Thr Glu Pro Asp Ser Asp Gly Lys Thr Tyr Glu Glu His Phe
TP84_78

TGGCGATCCCGGACATCACACAACAAGGCGGTGAAGATAATGGCGGTTAATTCGCTGAATTTTGAACGTA
ACCGCTAGGGCCTGTAGTGTGTTGTTCCGCCACTTCTATTACCGCCAATTAAGCGACTTAAACCTTGACT

865 870 875
Leu Ala Ile Pro Asp Ile Thr Gln Gln Gly Gly Glu Asp Asn Gly Gly
TP84_78

RBS

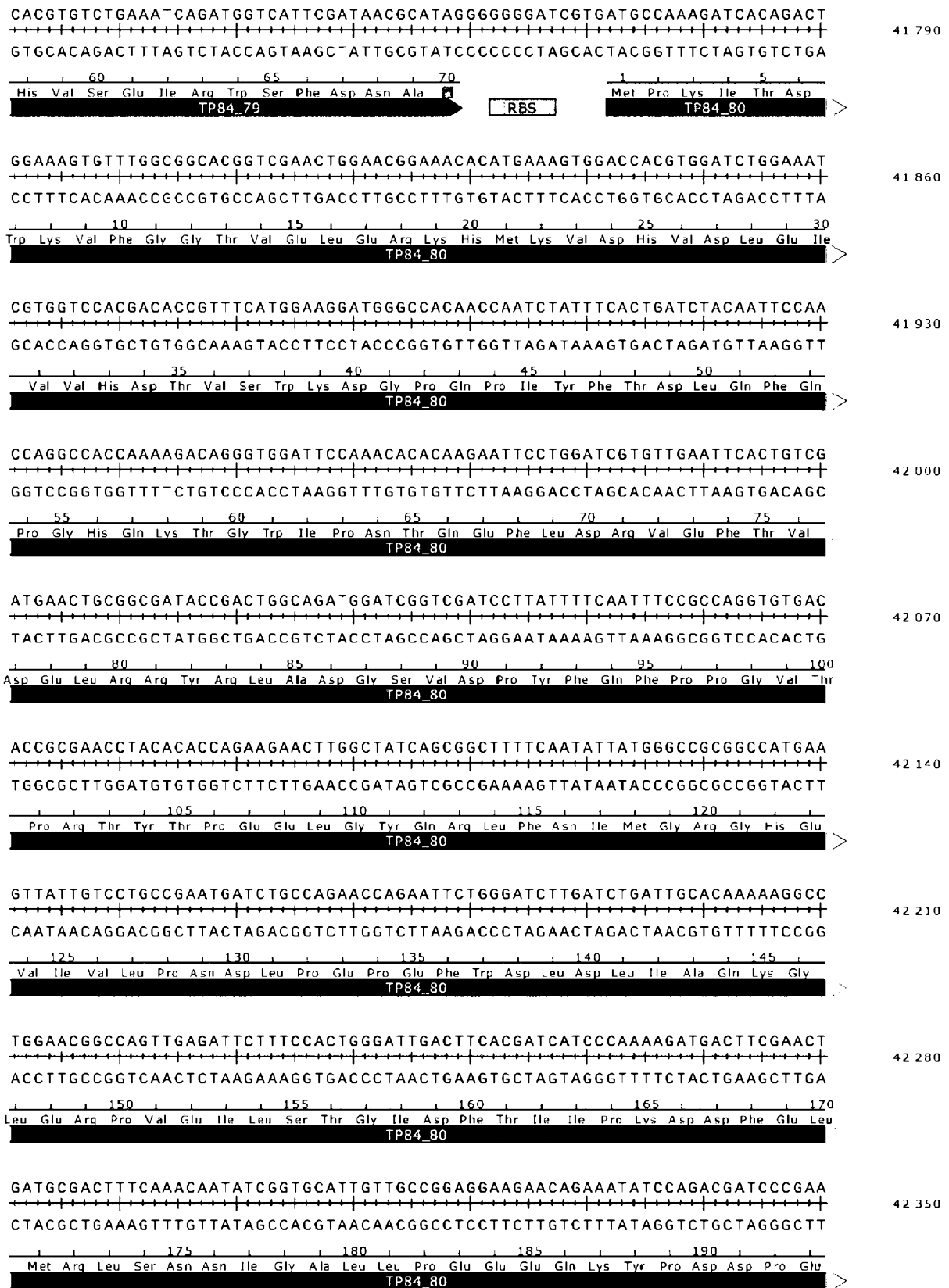
1 5 10
Met Ala Val Asn Ser Leu Asn Phe Glu Leu
TP84_79

TCGTGAAGAAGATAGAAACAGACACATCGCGGAAAGTGGATCACATCGATGTCGAGATCATCGCCAATGA
AGCACTTCTTCTATCTTTGTCTGTGTAGCGCCTTTCACCTAGTGTAGCTACAGCTCTAGTAGCGGTTACT

15 20 25 30
Ile Val Lys Lys Ile Glu Thr Asp Thr Ser Arg Lys Val Asp His Ile Asp Val Glu Ile Ile Ala Asn Asp
TP84_79

CTACACAGGAACGTTTTATTTACGGATATAATGTTTCAATCGGGAACGGTGGCCACATCCTGGGTGGGC
GATGTGTCTTGGCAAAATAAAATGCCTATATTACAAAGTTAGCCCTTGCCACCGGTGTAGGACCCACCCG

35 40 45 50 55
Tyr Thr Gly Thr Phe Tyr Phe Thr Asp Ile Met Phe Gln Ser Gly Thr Val Ala Thr Ser Trp Val Gly
TP84_79



CATCCACTGAATTATCGATATACCAGGGAATTTTGGATCGGAAGCGGCCACGCAGGGGATGTCATCGAGA
GTAGGTGACTTAATAGCTATATGGTCCCTTAAACCTAGCCTTCGCCGGTGCGTCCCCTACAGTAGCTCT

42 420

195 200 205 210 215
His Pro Leu Asn Tyr Arg Tyr Thr Arg Glu Phe Trp Ile Gly Ser Gly His Ala Gly Asp Val Ile Glu
TP84_80

TCAACGCCACCACCATGACGGCCAAAAGTGAATGGCATCACCATCAACACACAAGGGATCAAACAAATCAC
AGTTGCGGTGGTGGTACTGCCGGTTTCACCTACCGTAGTGGTAGTTGTGTGTTCCCTAGTTTGTITAGTG

42 490

220 225 230 235 240
Ile Asn Ala Thr Thr Met Thr Ala Lys Val Asn Gly Ile Thr Ile Asn Thr Gln Gly Ile Lys Gln Ile Thr
TP84_80

GACTGGCAGTGATACAATAAAAAATATACAAAAACAAATTTTCATCTGATGCCGCGTGGATCGGTCCGATTTC
CTGACCGTCACTATGTTATTTTTATATGTTTTGTTTAAAGTAGACTACGGCGCACCTAGCCAGGCTAAG

42 560

245 250 255 260
Thr Gly Ser Asp Thr Ile Lys Ile Tyr Lys Asn Lys Phe His Leu Met Pro Arg Gly Ser Val Arg Phe
TP84_80

CGTGTCGAATTTTATGGCCGGGATGCAAATGGCCGCCTGGCCGACACTGGGATCGGTTATCGCGGCACAG
GCACAGCTTAAATACCGGCCCTACGTTTACCGGCGGACCGGCTGTGACCCTAGCCAATAGCGCCGTGTC

42 630

265 270 275 280 285
Arg Val Glu Phe Tyr Gly Arg Asp Ala Asn Gly Arg Leu Ala Asp Thr Gly Ile Gly Tyr Arg Gly Thr
TP84_80

CAACATTTAAACAATGGACATATGGGGTGGAAAGACTATGACGAAACAAAAACGTGTGATGACATGGTCA
GTTGTAAATTTGTTACCTGTATACCCACCTTTCTGATACTGCTTTGTTTTGCACACTACTGTACCAGT

42 700

1 5 10
Met Thr Lys Gln Lys Arg Val Met Thr Trp Ser
TP84_81
290 295 300
Ala Thr Phe Lys Gln Trp Thr Tyr Gly Val Glu Arg Leu
TP84_80

CTGCGGATTCAAAACGATGACTATCGACACGACATCGAACGTTATGGAAACAAGATCACATCGATCGGAT
GACGCCTAAGTTTTGCTACTGATAGCTGTGCTGTAGCTTGCAATACCTTTGTTCTAGTGTAGCTAGCCTA

42 770

15 20 25 30
Leu Arg Ile Gln Asn Asp Asp Tyr Arg His Asp Ile Glu Arg Tyr Gly Asn Lys Ile Thr Ser Ile Gly
TP84_81

TCCATGAATTCAATGTCGATGCACAGGGCCGGATCGCCATCTATCGTCTAGGGCCACAAAAGGATGCACA
AGGTACTTAAGTTACAGCTACGTGTCCCGGCCCTAGCGGTAGATAGCAGATCCCGGTGTTTTCTACGTGT

42 840

35 40 45 50 55
Phe His Glu Phe Asn Val Asp Ala Gln Gly Arg Ile Ala Ile Tyr Arg Leu Gly Pro Gln Lys Asp Ala Gln
TP84_81

GGGATCGTTCACATCCTTCACAGATGCCACCACATATGATCGTTATGCGCCAGGGCGGACATTATGGCCG
CCCTAGCAAGTGTAGGAAGTGCTACGGTGGTGATACTAGCAATACGCGGTCCCGCCTGTAATACCGGC

42 910

60 65 70 75 80
Gly Ser Phe Thr Ser Phe Thr Asp Ala Thr Thr Tyr Asp Arg Tyr Ala Pro Gly Arg Thr Leu Trp Pro
TP84_81

AATTACATCGAAAAAGACATGAAGCGGTGGCCGCATATCGAATATTACATGCAGTTTGTATATTGGGTC
TTAATGTAGCTTTTCTGTACTTCGCCACCGGCGTATAGCTTATAATGTACGTCAAACAATATAAGCCAG

42 980

85 90 95 100
Asn Tyr Ile Glu Lys Asp Met Lys Arg Trp Pro His Ile Glu Tyr Tyr Met Gln Phe Val Ile Phe Gly
TP84_81

CAGAAACGGTGTACGACTGCTGGATTACAAACGGCGCAAGGATAACTTTATCAACAATCTGAAAACGGT
GTCTTTGCCACAGTGCTGACGACCTAAGTGTTTGCCGCGTCCTATTGAAATAGTTGTTAGACTTTTGCCA
105 110 115 120 125
Pro Glu Thr Val Ser Arg Leu Leu Asp Ser Gln Thr Ala Gln Asp Asn Phe Ile Asn Asn Leu Lys Thr Val
TP84-81

CGTGGCCAGATTCCGAAAGGATCAGAACGGAAATGATCTGGGATACACCGGGATCGAAATCGACTGCGAA
GCACCGGTCTAAGGCTTTCTAGTCTTGCTTTACTAGACCCTATGTGGCCCTAGCTTTAGCTGACGCTT
130 135 140 145 150
Val Ala Arg Phe Arg Lys Asp Gln Asn Gly Asn Asp Leu Gly Tyr Thr Gly Ile Glu Ile Asp Cys Glu
TP84-81

GGATCATTCAAGTGATTGAAATGGGATACACGCGCCGGGGATGATGTGAAGTATATCAATCTATTGAAAC
CCTAGTAAGTCACTAAGCTTTACCCTATGTGCGCGGCCCTACTACACTTCATATAGTTAGATAACTTTG
155 160 165 170
Gly Ser Phe Ser Asp Ser Lys Trp Asp Thr Arg Ala Gly Asp Asp Val Lys Tyr Ile Asn Leu Leu Lys
TP84-81

GGATTAAAAATGAAGTGATCATCGATGCGAATCCAGATTTTAAATTACGGATCAACGCACATGCGATGTG
CCTAATTTTACTTCACTAGTAGCTACGCTTAGGCTAAAATTTAATGCCTAGTTGCGGTGTACGCTACAC
175 180 185 190 195
Arg Ile Lys Asn Glu Val Ile Ile Asp Ala Asn Pro Asp Phe Lys Leu Arg Ile Asn Ala His Ala Met Trp
TP84-81

GGGCGATGGAATCCCGGATTATTATCGATTCCATAACTATAAATTGTTTGGCGAATCCACCGATAAAAAAC
CCCGCTACCTTAGGGCTAATAATAGCTAAGGTATTGATATTTAACAAACGGCTTAGGTGGCTATTTTTG
200 205 210 215 220
Gly Asp Gly Ile Pro Asp Tyr Tyr Arg Phe His Asn Tyr Lys Leu Phe Ala Glu Ser Thr Asp Lys Asn
TP84-81

GGAAATCCGTTGCTGGATGAAGTGCAGATCATGACGTATGACTTTTCTGGTCTGGATCTGCGCCAGGGC
CCTTTAGGCAACGACCTACTTCACGTCTAGTACTGCATACTGAAAAGGACCAGACCTAGACGCGGTCCTCCG
225 230 235 240
Gly Asn Pro Leu Leu Asp Glu Val Gln Ile Met Thr Tyr Asp Phe Ser Trp Ser Gly Ser Ala Pro Gly
TP84-81

CATCGACACCATTGTGGTGGATGCGAAATGTGGCCGAATGGGTGAAACAGTGTGTTTATGATCCATCGGTGAA
GTAGCTGTGGTAACACCACCTACGCTTTACACCGGCTTACCCACTTTGTACAAAACCTAGGTAGCCACTT
245 250 255 260 265
Pro Ser Thr Pro Leu Trp Trp Met Arg Asn Val Ala Glu Trp Val Lys Gln Cys Phe Asp Pro Ser Val Asn
TP84-81

TCCGAACGCAAAATGCACGATTGACAATGTGTATCTGGGCGGCGCAGGGTATGGCCGCGAGATGGCCGATT
AGGCTTGCGITTTACGTGCTAACTGTTACACATAGACCCGCGCGTCCCATACCGGCGTCTACCGGCTAA
270 275 280 285 290
Pro Asn Ala Lys Cys Thr Ile Asp Asn Val Tyr Leu Gly Gly Ala Gly Tyr Gly Arg Arg Trp Pro Ile
TP84-81

CATTGGATGATAACTGGGGATCGACAGTGACATATCGGGATCTCGTTGACTGGCAAAACGGTTATTTGA
GTAAGCCTACTATTGACCCCTAGCTGTCACTGTATAGCCCTAGAGCAACTGACCGTTTTGCCAATAAACT
295 300 305 310
His Ser Asp Asp Asn Trp Gly Ser Thr Val Thr Tyr Arg Asp Leu Val Asp Trp Gln Asn Gly Tyr Leu
TP84-81

TCCACCACCTGGGCGGCGGCAACTGGCAGACCAGGATTTTATCCCTTTAAATGCGTTTAATGATCCAAA
43 680
AGGTGGTGGACCCGCCGCGTTGACCGTCTGGTCCTAAATAGGGAATTTACGCAAATTACTAGGTTT
315 320 325 330 335
Ile His His Leu Gly Gly Gly Gln Leu Ala Asp Gln Asp Phe Ile Pro Leu Asn Ala Phe Asn Asp Pro Asn
TP84_81

CAGTGACAACCAGATCATGCTGATGCATCAATATGACTATTTCAAAGCGCGTCACATGAAGCGATATGAT
43 750
GTCACTGTTGGTCTAGTACGACTACGTAGTTATACTGATAAAGTTTCGCGCAGTGTACTTCGCTATACTA
340 345 350 355 360
Ser Asp Asn Gln Ile Met Leu Met His Gln Tyr Asp Tyr Phe Lys Ala Arg His Met Lys Arg Tyr Asp
TP84_81

TCGATGGGGCAGACGACAGTGCCTGTATCCGAATATAACGGCATCGAATATGCAACGGCCTATTCAAAGA
43 820
AGCTACCCCGTCTGCTGTCACGCACATAGGCTTATATTGCCGTAGCTTATACGTTGCCGGATAAGTTTCT
365 370 375 380
Ser Met Gly Gln Thr Thr Val Arg Val Ser Glu Tyr Asn Gly Ile Glu Tyr Ala Thr Ala Tyr Ser Lys
TP84_81

CGCAACATGCTAAAATTAAGCGGTCTGGCAGGGGTGACAGGGGTGAAAGGTGTCAGCGAACCAGACACAA
43 890
GCGTTGTACGATTTTAATCGCCAGACCGTCCCCACTGTCCCCACTTCCACAATCGCTTGGCTGTGTGTT
385 390 395 400 405
Thr Gln His Ala Lys Ile Ser Gly Leu Ala Gly Val Thr Gly Val Lys Gly Val Ser Glu Pro Thr His Asn
TP84_81

CCTTTCACCATATGATAATCCGAAACGCGGTGAAACAGTCACCATTGATTTTGGGAATGGCCAGTCATAC
43 960
GGAAAGTGGTATACTATTAGGCTTTGCGCCACTTTGTCAGTGGAACATAAAACCCTTACCGGTCAGTATG
410 415 420 425 430
Leu Ser Pro Tyr Asp Asn Pro Lys Arg Gly Glu Thr Val Thr Ile Asp Phe Gly Asn Gly Gln Ser Tyr
TP84_81

ACATTCCAGGCATACCAAACGAAACGTGTGCCGTATGTACCGACAGCAATCAAAGATGCAAACGGAAAAA
44 030
TGTAAGGTCCGTATGGTTTGCTTTGCACACGGCATAACATGGCTGTCGTTAGTTTCTACGTTTGCTTTTT
435 440 445 450
Thr Phe Gln Ala Tyr Gln Thr Lys Arg Val Pro Tyr Val Pro Thr Ala Ile Lys Asp Ala Asn Gly Lys
TP84_81

TAACAGGATATGCGTGTCACTTATCCGAACAGCCAGAAACTGTTTTGACATACACGTTAAATGTGCCGCA
44 100
ATTGTCCTATACGCACAGTCAATAGGCTTGTGGTCTTTGACAAAACGTATGTGCAATTTACACGGCGT
455 460 465 470 475
Ile Thr Gly Tyr Ala Cys Gln Leu Ser Glu Gln Pro Glu Thr Val Leu Thr Tyr Thr Leu Asn Val Pro Gln
TP84_81

AGCTGGAACATACAAGATCGGCGCAATTGTGTCATTTCCATTTTATAACTATACGAAACTAGGTGGACAC
44 170
TCGACCTTGATGTTCTAGCCGCGTTAACACAGTAAAGGTAAAATATTGATATGCTTTGATCCACCTGTG
480 485 490 495 500
Ala Gly Thr Tyr Lys Ile Gly Ala Ile Val Ser Phe Pro Phe Tyr Asn Tyr Thr Lys Leu Gly Gly His
TP84_81

ATAAATGGCGCACCATTACGATCGGCGGCGACAATTTACGAGCTATTATCCGATGATGTTCAAAGCAT
44 240
TATTTACCGCGTGTAAGTGCTAGCCGCGCTGTTAAATGGTCTGATAATAGGCTACTACAAGTTTCGTA
505 510 515 520
Ile Asn Gly Ala Pro Phe Thr Ile Gly Gly Asp Asn Leu Pro Asp Tyr Tyr Pro Met Met Phe Lys Ala
TP84_81

GTACATCTGGGATCTTGGCACACATTTCATTGAACGCCGGAGCAAACACCATCACGATCGAAGGGCCTAT
CAGTGTAGACCCCTAGAACCCTGTGTAAGTAACTTGCGGCCCTCGTTTGTGGTAGTGCTAGCTTCCCGGATA

525 530 535 540 545
Cys His Ile Trp Asp Leu Gly Thr His Ser Leu Asn Ala Gly Ala Asn Thr Ile Thr Ile Glu Gly Pro Met
TP84_81

GTCTGCACATGGCAGCATCATTTTCGGCTTTTTTTCGGTGTCAATCACTGACACTGGAAGTGATCGCGGGA
CAGACGTGTACCGTGCTAGTAAAAGCCGAAAAACGCACAGTTAGTGACTGTGACCTTCACTAGCCGCCT

550 555 560 565 570
Ser Ala His Gly Thr Ile Ile Phe Gly Phe Phe Ala Cys Gln Ser Leu Thr Leu Glu Val Ile Gly Gly
TP84_81

TATATGGACTGCGAATCGACAATATATCCTTACAAAAACGTGATGGATCAGATGCCAGACTGCCTGCAC
ATATACCTGACGCTTAGCTGTTATATAGGAATGTTTTTGCCTACCTAGTCTACGGTCTGACGGACGTG

575 580 585 590
Tyr Met Asp Cys Glu Ser Thr Ile Tyr Pro Tyr Lys Lys Arg Asp Gly Ser Asp Ala Arg Leu Pro Ala
TP84_81

AATTCGCACTGACATCGGAAGTATTGCAACAATCGCCGCGGCCAGTGATCATGTGGGAAGATGTATTCA
TTAAGCGTGACTGTAGCCTTCATAACGTTGTTAGCGCGCCGGTCACTAGTACACCCTTCTACATAAGTC

595 600 605 610 615
Gln Phe Ala Leu Thr Ser Glu Val Leu Gln Gln Ser Pro Arg Pro Val Ile Met Trp Glu Asp Val Phe Arg
TP84_81

ACAGTATTTGAATGATGAATTTGTTCAACAATACGGCCTGGAAGCTACAACCTATTACAGATTGACAGGA
TGTCTATAAAGTTACTACTTAAACAAGTTGTTATGCCGGACCTTCGATGTTGGATAATGTCTAACTGTCT

620 625 630 635 640
Gln Tyr Leu Asn Asp Glu Phe Val Gln Gln Tyr Gly Leu Glu Ala Thr Thr Tyr Tyr Arg Leu Thr Gly
TP84_81

CAGAAAAAAGACTATGGCGGCGGCACGATTGAAGAATGGACAGATGACGTGCTGACTGGATGTTATGCGC
GTCTTTTTTCTGATACCGCGCGCGTGCTAACTTCTTACCTGTCTACTGCACGACTGACCTACAATACGCG

645 650 655 660
Gln Lys Lys Asp Tyr Gly Gly Gly Thr Ile Glu Glu Trp Thr Asp Asp Val Leu Thr Gly Cys Tyr Ala
TP84_81

TAGTGGACAACGGATTACACAAAGGTCCTG6CCAGTGGAACAGACGAAAACGGCACAGCGTGTGCAAG
ATCACCTGTTGCCTAAGTGTGTTCCAGTGACCGGTACCTTTGTCTGCTTTTGCCGTGTGCGCACACGTTT

665 670 675 680 685
Leu Val Asp Asn Gly Phe Thr Gln Gly His Trp Pro Val Glu Thr Asp Glu Asn Gly Thr Ala Cys Ala Arg
TP84_81

ATTCCGGCCTAATGATCCGAAAAACAACGGTGTGACATCTGGACAGTTAGTTTTGAATTACACATATAAA
TAAGGCCGGATTACTAGGCTTTTTGTTGCCACACTGTAGACCTGTCAATCAAACTTAATGTGTATATTT

690 695 700 705 710
Phe Arg Pro Asn Asp Pro Lys Asn Asn Gly Val Thr Ser Gly Gln Leu Val Leu Asn Tyr Thr Tyr Lys
TP84_81

ACCACGAACATTTCTGCGAAGTGCAATTCAAAGTGAAATCGGGTAGACGCGCCGGGATCAGATTCGCAT
TGGTGCTTGTAAGGACGCTTCACGTTAAGTTTCACTTTAGCCCATCTGCGCGGCCCTAGTCTAAGCGTA

715 720 725 730
Thr Thr Asn Ile Ser Cys Glu Val Gln Phe Lys Val Lys Ser Gly Arg Arg Ala Gly Ile Arg Phe Ala
TP84_81

CCACTGGACCAGGGGATGGATATGTCTTTTTAATCGATTATCAAACACAGGAAGCCATGATGTTCTATGA
GGTGACCTGGTCCCCTACCTATACAGAAAAATTAGCTAATAGTTTGTGTCCTTCGGTACTACAAGATACT
44 940

735 740 745 750 755
Ser Thr Gly Pro Gly Asp Gly Tyr Val Phe Leu Ile Asp Tyr Gln Thr Gln Glu Ala Met Met Phe Tyr Glu
TP84_81

AACCGCTGGATCATCGCAACTGGTGCCAGTGATCCCTGGGGGATAGAAGGGCCGACTATGATGAATTG
TTGGCGACCTAGTAGCGTTGACCACCGGTCACGTAGGGACCCCTATCTTCCCGGCTGATACTACTTAAC
45 010

760 765 770 775 780
Thr Ala Gly Ser Ser Gln Leu Val Ala Ser Ala Ser Leu Gly Asp Arg Arg Ala Asp Tyr Asp Glu Leu
TP84_81

ATCACATTGAAAGTCCTGGTCAATAACGGAAAAATGTCGGTGCTATTTGCGCAACGTGATGTTCTTTATGG
TAGTGTAACTTTCAGGACCAGTTATTGCCTTTACAGCCACGATAAAGCCGTTGCACTACAAGAAATACC
45 080

785 790 795 800
Ile Thr Leu Lys Val Leu Val Asn Asn Gly Lys Cys Arg Cys Tyr Phe Gly Asn Val Met Phe Phe Met
TP84_81

ATATGAATCTGCCGCACATGTCGCCAGGCGGCATCGGATTCGTGGCAACAAATTGCGATGCCTATCTATA
TATACTTAGACGGCGTGACAGCGGTCGCCGTAGCCTAAGCACCGTTGTTTAACGCTACGGATAGATAT
45 150

805 810 815 820 825
Asp Met Asn Leu Pro His Met Ser Pro Gly Gly Ile Gly Phe Val Ala Thr Asn Cys Asp Ala Tyr Leu Tyr
TP84_81

TAAACTATCCATCGGAACGACTGAACGATGGGAAACACTTGAACGATTCAAGTGTCATCATCGATGGCCA
ATTTGATAGGTAGCCTTGCTGACTTGCTACCCTTTGTGAAGTTCAGTACAGTAGTAGCTACCGGTC
45 220

830 835 840 845 850
Lys Leu Ser Ile Gly Thr Thr Glu Arg Trp Glu Thr Leu Glu Arg Phe Ser Val Ile Ile Asp Gly Gln
TP84_81

GAATATAAGATGGGCGAAATCAGCAGACCGGGAATCAACCGGGATCAGTGGGGATTTTGTATATTTCTG
CTTATATTCTACCCGCTTTAGTCGTCTGGCCCTTAGTTGGCCCTAGTCACCCCTAAAACTATATAAGAC
45 290

855 860 865 870
Glu Tyr Lys Met Gly Glu Ile Ser Arg Pro Gly Ile Asn Arg Asp Gln Trp Gly Phe Leu Ile Tyr Ser
TP84_81

GATTCAATGAATACAACACCAGGGAAGTCCTGCCGGATGGATCGCGGCCAGAAATCAGTCTGGACTATGT
CTAAGTTACTTATGTTGTGGTCCCTTCAGGACGGCCTACCTAGCGCCGGTCTTTAGTCAGACCTGATACA
45 360

875 880 885 890 895
Gly Phe Asn Glu Tyr Asn Thr Arg Glu Val Leu Pro Asp Gly Ser Arg Pro Glu Ile Ser Leu Asp Tyr Val
TP84_81

ATTTAAACCGATCCATGTGCCATCCTGGACCAGGGAAGGATCAGATCAAGCTGATCGATGCCGGG
TAAATTTGGCTAGGTACACGGTAGGACCTGGCCCTTTTCTTCTAGTGCTAGTTCGACTAGCTACGGCCC
45 430

900 905 910 915 920
Phe Lys Pro Ile His Val Pro Ser Trp Thr Gly Lys Lys Lys Ile Thr Ile Lys Leu Ile Asp Ala Gly
TP84_81

CTATGGTATAAACAGCTATACATCGGCGATGCAAACGGCATGTCGATCGCATACGCAGGGGATGAAGAAT
GATACCATATTTGTGATATGTAGCCGCTACGTTTGGCGTACAGCTAGCGTATGCGTCCCCTACTTCTTA
45 500

925 930 935 940
Leu Trp Tyr Lys Gln Leu Tyr Ile Gly Asp Ala Asn Gly Met Ser Ile Ala Tyr Ala Gly Asp Glu Glu
TP84_81

CGTTTCGACAGGGCGATGAACATCGCGGTGCATGAATATGGATGCAAAGGGATCGGTTTGTGGGTATTAGG
GCAAGCTGTCCCGCTACTTGTAGCGCCACGTACTTATACCTACGTTTCCCTAGCCAAACACCCATAATCC

945 950 955 960 965
Ser Phe Asp Arg Ala Met Asn Ile Ala Val His Glu Tyr Gly Cys Lys Gly Ile Gly Leu Trp Val Leu Gly
TP84_81

TCAAGCAGATCCGCGCATCTTTGAAACATTGCCGGATGTAGTGCCGTGGCATCCAGATCCGAACGAAAAC
AGTTCGTCTAGGCGCGTAGAAACTTTGTAACGGCTACATCACGGCACCGTAGGTCTAGGCTTGTCTTTG

970 975 980 985 990
Gln Ala Asp Pro Arg Ile Phe Glu Thr Leu Pro Asp Val Val Pro Trp His Pro Asp Pro Asn Glu Asn
TP84_81

TAAAAAACACGGCCAAATGATCAGCAAAATATGGTAAATTTAAACAAGAACAGGGGGGTGAATAAAAA
ATTTTTTTGTGCGGTTTACTAGTCGTTTTATACCATTTTAAATTTGTTCTTGTCCCCCACTTATTTTT

TP84_81

TAAAGGAGGTACAACGTTTTGAAAACGCCGCAACAAACCGACACGCTATTCCTGCGATCGCCGGAGGAT
ATTCCTCCATGTTGCAAAACTTTTGGCGGCTTGTGGCTGTGCGATAAGTGACGCTAGCGGCCTCCTA

RBS

1 5 10 15
Met Lys Thr Pro Gln Gln Thr Asp Thr Leu Phe Thr Ala Ile Ala Gly Gly
TP84_82

TTACATCGACAGTGGCTTATTTGATCGGCGGTGTGGACAACCTTTCAAACGCGCTGACGGTGTTCATGAT
AATGTAGCTGTCACCGAATAAACTAGCCGCCACACCTGTTGGAAAGTTTGGCGGACTGCCACAAGTACTA

20 25 30 35 40
Phe Thr Ser Thr Val Ala Tyr Leu Ile Gly Gly Val Asp Asn Leu Ser Asn Ala Leu Thr Val Phe Met Ile
TP84_82

TTTGGATTATTTCACTGGCATCCTATCCGCTTTCTATACGCGCCAGGTGAACAGCTACCTGGCATATAGA
AAACCTAATAAAGTGACCGTAGGATAGGCGAAAGATATGCGCGGTCCACTTGTGATGGACCGTATATCT

45 50 55 60
Leu Asp Tyr Phe Thr Gly Ile Leu Ser Ala Phe Tyr Thr Arg Gln Val Asn Ser Tyr Leu Ala Tyr Arg
TP84_82

GGACTGGCAAAAAAAGCAGGGATGATCGCATTGTTTGTGGCGAATCAATTGGACATCATCACTGGCA
CCTGACCGTTTTTTTCGTCCCTACTAGCGTAAGCAATAACACCGCTTAGTTAACCTGTAGTAGTGACCGT

65 70 75 80 85
Gly Leu Ala Lys Lys Ala Gly Met Ile Ala Phe Val Ile Val Ala Asn Gln Leu Asp Ile Ile Thr Gly
TP84_82

ACACGGAAGGGTTTCTGCGTGATGCCATGATGATGTTTCTGATCGGGATGGAAGGCATTTGATCAAAGA
TGTGCTTCCCAAAGACGCACTACGGTACTACTACAAAGACTAGCCCTACCTTCCGTAAAGCTAGTTTTCT

90 95 100 105 110
Asn Thr Glu Gly Phe Leu Arg Asp Ala Met Met Met Phe Leu Ile Gly Met Glu Gly Ile Ser Ile Lys Glu
TP84_82

GAATGTCGAGAAAATGGGATTCAATGCACCAGGATTCATTGTGGAAGCATTGAAAAAGCTGATGGGGAAT
CTTACAGCTCTTTTACCCTAAGTTACGTGGTCTAAGTAACACCTTCGTAACCTTTTTCGACTACCCCTTA

115 120 125 130
Asn Val Glu Lys Met Gly Phe Asn Ala Pro Gly Phe Ile Val Glu Ala Leu Lys Lys Leu Met Gly Asn
TP84_82

GATAAAAACGATTCATCGAAGGGGGCTTGATAACATGCAAGCAAGATCCGCGAACAATATCAAAGGGATC 46 200
CTATTTTTGCTAAGTAGCTTCCCCGAACATTGTACGTTCTAGGCGCTTGTTATAGTTTCCCTAG

135 140
Asp Lys Asn Asp Ser Ser Lys Gly Ala
TP84_82

1 5 10
Met Gln Ala Arg Ser Ala Asn Asn Ile Lys Gly Ile
TP84_83

RBS

GATGTGTCGCACTGGCAAGGAAAGATTGACTGGGCGAAAGTGAAAGCCGCTGGCATCCAGGTCGCTTATT 46 270
CTACACAGCGTGACCGTTCTTTCTAACTGACCCGCTTTCACTTTCGGCGACCGTAGGTCCAGCGAATAA

15 20 25 30 35
Asp Val Ser His Trp Gln Gly Lys Ile Asp Trp Ala Lys Val Lys Ala Ala Gly Ile Gln Val Ala Tyr
TP84_83

TGAAAGCGACAGAAGGAACGACACATGTGGATAAAATGTTAAAAACAACTATCAGAACGCAAAAAAAGC 46 340
ACTTTCGCTGCTCTTCTTCTGCTGTGTACACCTATTTTACAATTTTGTGTTGATAGTCTTGCGTTTTTTTCG

40 45 50 55
Leu Lys Ala Thr Glu Gly Thr Thr His Val Asp Lys Met Leu Lys Thr Asn Tyr Gln Asn Ala Lys Lys Ala
TP84_83

CGGGATCAAAGTGGGATTTTATCACTTTTTCCGGGCAAAAAACGAACAAAACGCACGTGAACAAGCCAGA 46 410
GCCCTAGTTTACCCTAAAATAGTGAAAAAGGCCCGTTTTTGTCTGTTTTGCGTGCACCTTGTTTCGGTCT

60 65 70 75 80
Gly Ile Lys Val Gly Phe Tyr His Phe Phe Arg Ala Lys Asn Glu Gln Asn Ala Arg Glu Gln Ala Arg
TP84_83

CATTTTCGTGAACACAGTCAAAGGGATGCCGAACGATCTGAAACACGCGCTGGACATCGAAACCACCGAAG 46 480
GTAAGCACTTGTGTCACTTCCCTACGGCTTGCTAGACTTTGTGCGCGACCTGTAGCTTTGGTGGCTTC

85 90 95 100 105
His Phe Val Asn Thr Val Lys Gly Met Pro Asn Asp Leu Lys His Ala Leu Asp Ile Glu Thr Thr Glu
TP84_83

GCCTATCGAATGAAAGCGTTGACGAAATGCGCGATTGCATTCCTGGAAGAAAGTGAAAAAGCTGACTGGTCA 46 550
CGGATAGCTTACTTCGCAACTGCTTTACGCGCTAACGTAAGGACCTTCTTCACTTTTTCGACTGACCAGT

110 115 120 125
Gly Leu Ser Asn Glu Ala Leu Thr Lys Cys Ala Ile Ala Phe Leu Glu Glu Val Lys Lys Leu Thr Gly Gln
TP84_83

AGATCCGATCGTTTACACATACACATCATTGCAAGATCCCGATTAACAGCGGCTATCGCAAAATATCCA 46 620
TCTAGGCTAGCAAAATGTGTATGTGTAGTAAGCGTTCTAGGGCTAATTGTGCGCGATAGCGTTTTATAGGT

130 135 140 145 150
Asp Pro Ile Val Tyr Thr Tyr Thr Ser Phe Ala Arg Ser Arg Leu Thr Ala Ala Ile Ala Lys Tyr Pro
TP84_83

GTGTGGATCGCACATTATGGCGTGACAAGCCTGGCGACAATCCGATCTGGGATCGATGGATCGGATTCC 46 690
CACACCTAGCGTGTAATACCGCACCTGTTCCGACCGCTGTTAGGCTAGACCCTAGCTACCTAGCCTAAGG

155 160 165 170 175
Val Trp Ile Ala His Tyr Gly Val Asp Lys Pro Gly Asp Asn Pro Ile Trp Asp Arg Trp Ile Gly Phe
TP84_83

AGTACACTGACAAAGGGAAGTCAGCGGCATCGCTGGAAATGTGGACATGAATGAATTTACGTCAGACAT 46 760
TCATGTGACTGTTTCCCTTTCAGTCGCCGTAGCGACCTTTACACCTGTACTTAAATGCAGTCTGTA

180 185 190 195
Gln Tyr Thr Asp Lys Gly Lys Val Ser Gly Ile Ala Gly Asn Val Asp Met Asn Glu Phe Thr Ser Asp Ile
TP84_83

TTTTGTCGATGCGGCCAAAGTGGAAACAGCCGAAACAAAAGGTCGATGCTGTCCAATCCACGCCATCGGCC
46 830
AAAACAGCTACGCCGGTTTCACCTTGTCGGCTTTGTTTTCCAGCTACGACAGGTTAGGTGCGGTAGCCGG

200 205 210 215 220
Phe Val Asp Ala Ala Lys Val Glu Gln Pro Lys Gln Lys Val Asp Ala Val Gln Ser Thr Pro Ser Ala
TP84_83

ACTGGCACATATACGATCAAAAGCGGCGACACGTTCTGGGAAGTCTAGAAAGAAAAGTATAACTGGCCACATG
46 900
TGACCGTGTATATGCTAGTTTTCGCCGCTGTGCAAGACCCTTGATCTTCTTTTCATATTGACCGGTGTAC

225 230 235 240 245
Thr Gly Thr Tyr Thr Ile Lys Ser Gly Asp Thr Phe Trp Glu Leu Glu Glu Lys Tyr Asn Trp Pro His
TP84_83

GCACACTGCAAAGATTGAATCCGTCAGTGAATCCGAACGCATTGAAAGTCGGCCAGGTGATAAAAGTACC
46 970
CGTGTGACGTTTCTAACTTAGGCAGTCACTTAGGCTTGCGTAACTTTCAAGCCGGTCCACTATTTTCATGG

250 255 260 265
Gly Thr Leu Gln Arg Leu Asn Pro Ser Val Asn Pro Asn Ala Leu Lys Val Gly Gln Val Ile Lys Val Pro
TP84_83

AAAATCCGAACAGCCGAAACAAAATGCATCATCAGTCACTGGCACATATACGATCAAAAGCGGCGACACG
47 040
TTTTAGGCTTGTGCGGCTTTGTTTTACGTAGTAGTCAGTGACCGTGTATATGCTAGTTTTCGCCGCTGTGC

270 275 280 285 290
Lys Ser Glu Gln Pro Lys Gln Asn Ala Ser Ser Val Thr Gly Thr Tyr Thr Ile Lys Ser Gly Asp Thr
TP84_83

TTCTGGGATCTGGAACAAAAAACGGATGEGCACACGGCACGTTACAAAAATTGAATCCAGGCGTGAATC
47 110
AAGACCCTAGACCTTGTTTTTTTGCCTACCGGTGTGCCGTGCAATGTTTTTAAGTTAGGTCGCACTTAG

295 300 305 310 315
Phe Trp Asp Leu Glu Gln Lys Asn Gly Trp Pro His Gly Thr Leu Gln Lys Leu Asn Pro Gly Val Asn
TP84_83

CGAACAAATTGAAAGTCGGCCAGGTGATAAAAGTACCAAAATCCGAACAAAAGAATGTGCAGCGGACCGT
47 180
GCTTGTTTAACTTTTCAGCCGGTCCACTATTTTCATGGTTTTAGGCTTGTTTTCTTACACGTCGCCTGGCA

320 325 330 335
Pro Asn Lys Leu Lys Val Gly Gln Val Ile Lys Val Pro Lys Ser Glu Gln Lys Asn Val Gln Arg Thr Val
TP84_83

CAAAAACCAAAAAACCGAAGTATCGAACGTATAAAATAAAGAAAGGTGACACGTTCTGGGAAGTTGAA
47 250
GTTTTTGGTGGTTTTTGGCTTGATAGCTTGCAATTTTTATTTCTTTCCACTGTGCAAGACCCTTGAAGTT

340 345 350 355 360
Lys Asn His Gln Lys Pro Asn Tyr Arg Thr Tyr Lys Ile Lys Lys Gly Asp Thr Phe Trp Glu Leu Glu
TP84_83

AAGAAAAATGGCTGGCCACATGGAACGTTGAGAAATTGAATCCAGGCGTGAATCCGGCGAAACTGCAAA
47 320
TTCTTTTACCAGCCGGTGTACCTTGCAACGCTTTTAACTTAGGTCCGCACTTAGGCCGCTTTGACGTTT

365 370 375 380 385
Lys Lys Asn Gly Trp Pro His Gly Thr Leu Gln Lys Leu Asn Pro Gly Val Asn Pro Ala Lys Leu Gln
TP84_83

TCGGCCAAACGATCAAAATTCCAAACTAAAGAGAGGGGAAGCCGCATGAAGAACTTGAAGGCAAAAGACT
47 390
AGCCGGTTTGCTAGTTTTAAGGTTTGATTCTCTCCCTTCGGCGTACTTCTTGAAGTTCCGTTTTCTGA

390 395
Ile Gly Gln Thr Ile Lys Ile Pro Asn
TP84_83

RBS

1 5
Met Lys Asn Leu Lys Ala Lys Asp
TP84_84

GGATCAGCATCATTAAATATTTACTTCCTGTCATCGCACTGGTGATCGCAAAGATCGCAGATGCAGGCGA
CCTAGTCGTAGTAATTTATAAATGAAGGACAGTAGCGTGACCACTAGCGTTTCTAGCGTCTACGTCCGCT

47 460

1015202530
Trp Ile Ser Ile Ile Lys Tyr Leu Leu Pro Val Ile Ala Leu Val Ile Ala Lys Ile Ala Asp Ala Gly Asp

TP84_84

TCCAACCAAGATTCAAAACGCACTGGAAATCTTCCTTACTGCTTTACTGGGATTACTTGCGCGCAATTGGT
AGGTTGGTTCTAAGTTTTGCGTGACCTTTAGAAGGAATGACGAAATGACCCTAATGAACCGCGTTAACCA

47 530

3540455055
Pro Thr Lys Ile Gln Asn Ala Leu Glu Ile Phe Leu Thr Ala Leu Leu Gly Leu Leu Gly Ala Ile Gly

TP84_84

ATCATTAAAAGTCATGACAAAACGACGATCAGTGATATTATATGAAGTGCCGGCCTATGTTCTATTTTC
TAGTAATTTTCAGTACTGTTTTGACTGCTAGTCACTATAATATACTTCACGGCCGGATACAAGATAAAAG

47 600

6065
Ile Ile Lys Ser His Asp Lys Thr Asp Asp Gln

TP84_84

regulatory region

CTGTGAACCTTCGACACGTTTTTTTCGAAATGTGCAAGTGTTCTGCATCGATATTCTTCATTCTCCTATGA
GACACTTGGAAAGCTGTGCAAAAAGCTTTACACGTTTACAAGACGTAGCTATAAGAAGTAAGAGGATACT

47 670

regulatory region

End (47 703)

TCCGCCATCCAAACGGATAGCGGATCTTTTTTT
AGGCGGTAGGTTTGCCTATCGCCTAGAAAAAAA

3'
47 703
5'

regulatory region

Fig. 3 (początek)

Nr	Symbol ORF	Długość ORF (pary zasad)	Orientacja ORF w genomie TP-84	Długość polipeptydu (aminokwasy)	Masa cząsteczkowa polipeptydu (kDa)	Hipotetyczna funkcja na podstawie homologii sekwencji aminokwasowej
1	TP84_01A	573	→	190	21,7	nieustalona
2	TP84_01	189	→	62	7,4	białko związane z błoną cytoplazmatyczną bakterii
3	TP84_02	255	→	84	9,8	nieustalona
4	TP84_03	291	→	96	11,0	nieustalona
5	TP84_04	255	→	84	9,6	nieustalona
6	TP84_05	207	→	68	7,9	nieustalona
7	TP84_06	237	→	78	9,3	nieustalona
8	TP84_07	147	→	48	5,4	nieustalona
9	TP84_08	396	→	131	14,9	nieustalona
10	TP84_09	201	→	66	7,4	nieustalona
11	TP84_10	216	→	71	8,1	nieustalona
12	TP84_11	210	→	69	8,2	nieustalona
13	TP84_12	273	→	90	10,6	nieustalona
14	TP84_13	162	→	53	6,4	nieustalona
15	TP84_14	255	→	84	10,1	nieustalona
16	TP84_15	132	→	43	5,0	nieustalona
17	TP84_16	231	→	76	8,5	nieustalona
18	TP84_17	213	→	70	7,8	nieustalona
19	TP84_18	381	→	126	13,8	białko związane z błoną cytoplazmatyczną bakterii
20	TP84_19	2163	→	720	82,6	nieustalona
21	TP84_20	474	→	157	18,0	nieustalona
22	TP84_21	159	→	52	5,8	nieustalona
23	TP84_22	204	←	67	7,6	nieustalona
24	TP84_23	159	→	52	6,1	nieustalona
25	TP84_24	237	→	78	8,5	regulator transkrypcji (HTH_XRE rodzina)
26	TP84_25	1017	→	338	38,5	białko RecB-podobne
27	TP84_26	1020	→	339	38,8	białko zaangażowane w hybrydyzację jednoniciowego DNA (super-rodzina EFR)
28	TP84_27	222	→	73	8,8	nieustalona
29	TP84_28	183	→	60	7,4	nieustalona
30	TP84_29	783	→	260	30,3	nieustalona
31	TP84_30	318	→	105	12,3	inhibitor helikazy replikacyjnej
32	TP84_31	1305	→	434	49,1	helikaza replikacyjna DNA
33	TP84_32	672	→	223	26,4	nieustalona
34	TP84_33	435	→	144	16,7	HNH endonukleaza <i>homing</i>
35	TP84_34	147	→	48	5,9	nieustalona

36	TP84_35	468	→	155	17,4	białko wiążące jednoniciowe DNA
37	TP84_36	240	→	79	9,3	nieustalona
38	TP84_37	126	→	41	4,6	białko związane z błoną cytoplazmatyczną bakterii
39	TP84_38	270	→	89	10,4	nieustalona
40	TP84_39	183	→	60	6,9	nieustalona
41	TP84_40	987	→	328	38,6	syntaza tymidylanowa
42	TP84_41	501	→	166	20,0	fosfataza dUTP
43	TP84_42	330	→	109	12,7	nieustalona
44	TP84_43	246	→	81	9,4	nieustalona
45	TP84_44	504	→	167	19,3	endonukleaza specyficzna wobec 'Holliday junction'
46	TP84_45	264	→	87	10,4	nieustalona
47	TP84_46	546	→	181	21,4	nieustalona
48	TP84_47	114	→	37	4,0	nieustalona
49	TP84_48	126	→	41	4,8	nieustalona
50	TP84_49	186	→	61	7,2	nieustalona
51	TP84_50	135	→	44	5,0	fosfataza białek
52	TP84_51	123	→	40	4,3	nieustalona
53	TP84_52	354	→	117	13,4	nieustalona
54	TP84_53	366	→	121	14,0	nieustalona
55	TP84_54	264	←	87	9,6	nieustalona
56	TP84_55	141	→	46	5,3	nieustalona
57	TP84_56	567	→	188	22,4	terminaza replikacji DNA, mała podjednostka
58	TP84_57	1299	→	432	50,0	terminaza replikacji DNA, duża podjednostka
59	TP84_58	687	→	228	26,3	nieustalona
60	TP84_59	354	→	117	13,6	nieustalona
61	TP84_60	198	→	65	7,5	nieustalona
62	TP84_61	1629	→	542	62,1	białko składowe kapsydu, łączące główkę i ogonek
63	TP84_62	147	→	48	5,3	nieustalona
64	TP84_63	2046	→	681	78,0	mniejsze białko składowe kapsydu
65	TP84_64	210	→	69	8,2	białko związane z błoną cytoplazmatyczną bakterii
66	TP84_65	828	→	275	31,6	proteaza dojrzewania pro-główki kapsydu
67	TP84_66	408	→	135	14,0	nieustalona
68	TP84_67	993	→	330	37,8	główne białko kapsydu
69	TP84_68	195	→	64	7,2	nieustalona
70	TP84_69	330	→	109	12,2	nieustalona

71	TP84_70	324	→	107	11,9	nieustalona
72	TP84_71	489	→	162	18,7	nieustalona
73	TP84_72	378	→	125	14,6	białko składania ogonka
74	TP84_73	558	→	185	20,3	nieustalona
75	TP84_74	507	→	168	19,3	nieustalona
76	TP84_75	375	→	124	14,7	nieustalona
77	TP84_76	2598	→	865	93,1	białko składania ogonka
78	TP84_77	738	→	245	27,8	cystalne białko ogonka
79	TP84_78	2634	→	877	99,3	białko ogonka
80	TP84_79	210	→	69	7,9	nieustalona
81	TP84_80	900	→	299	34,5	cystalne białko ogonka
82	TP84_81	2976	→	991	112,2	glikozylaza zbliżona do glikozylazy YDHD sporulacji
83	TP84_82	432	→	143	15,5	holina, białko perforujące błonę cytoplazmatyczną bakterii
84	TP84_83	1185	→	394	44,2	endolizyna (lizozym)
85	TP84_84	201	→	66	7,2	białko związane z błoną cytoplazmatyczną bakterii

Fig. 3 (koniec)

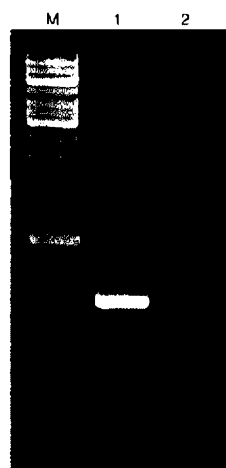


Fig. 4

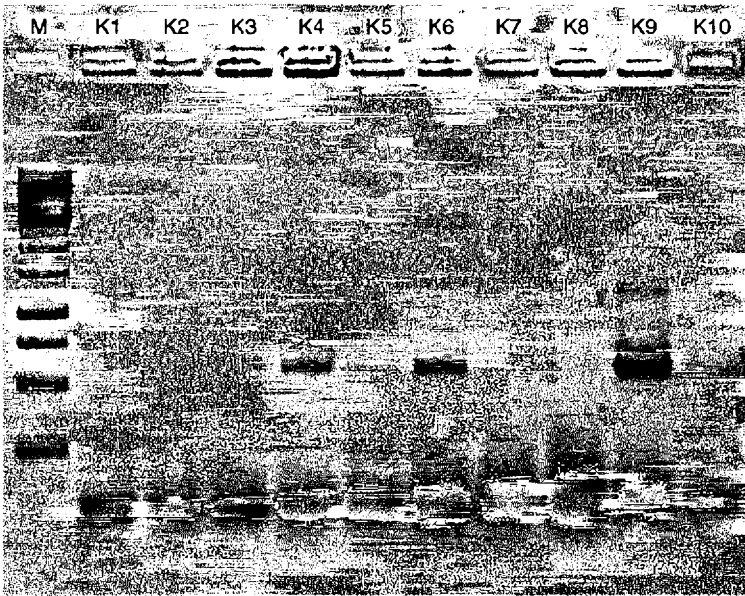


Fig. 5

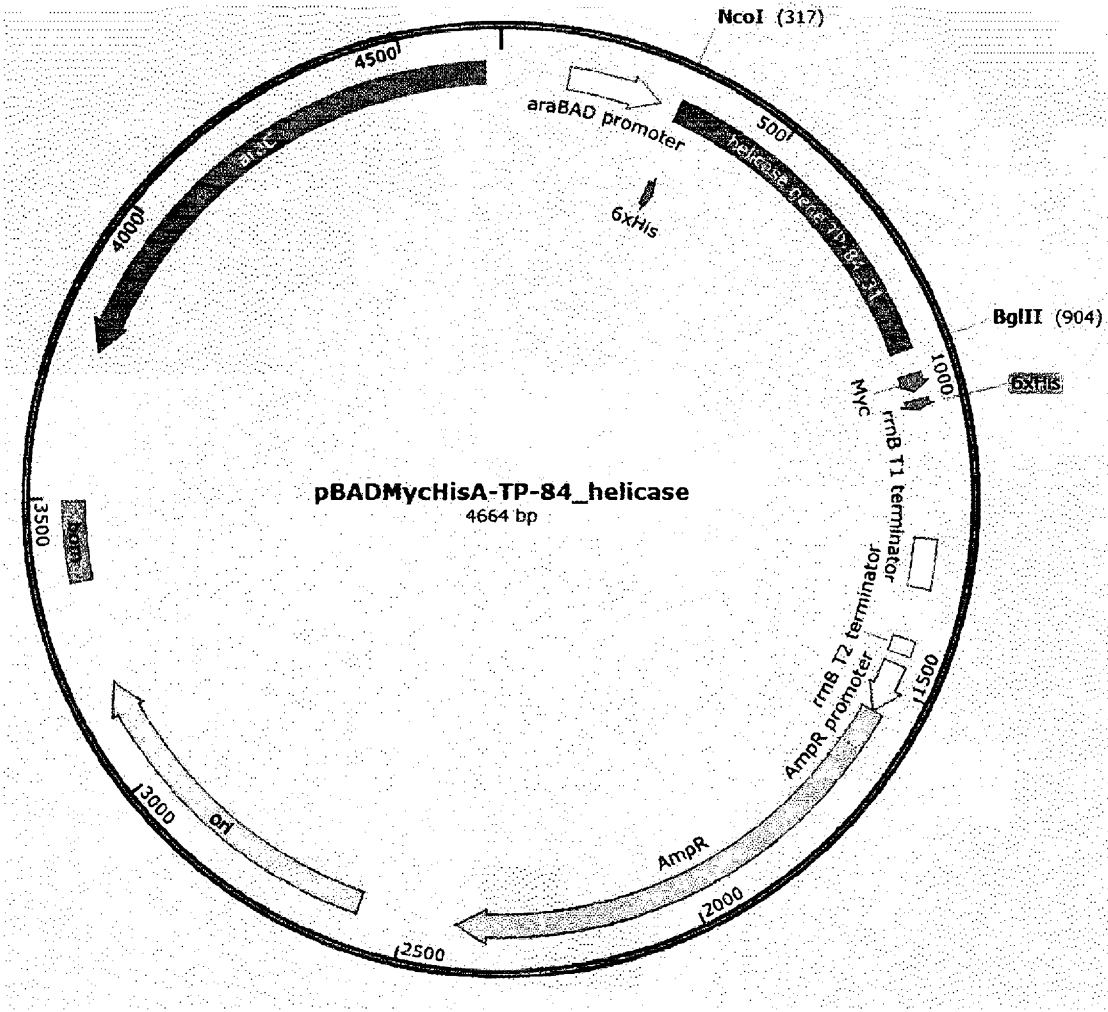


Fig. 6

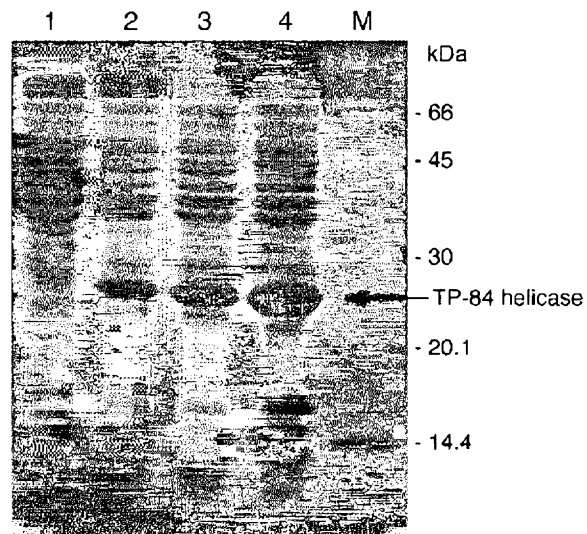


Fig. 7

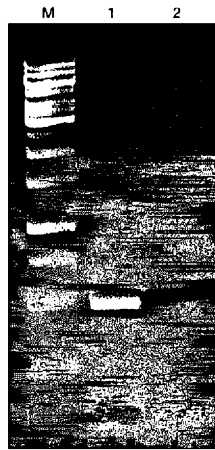


Fig. 8

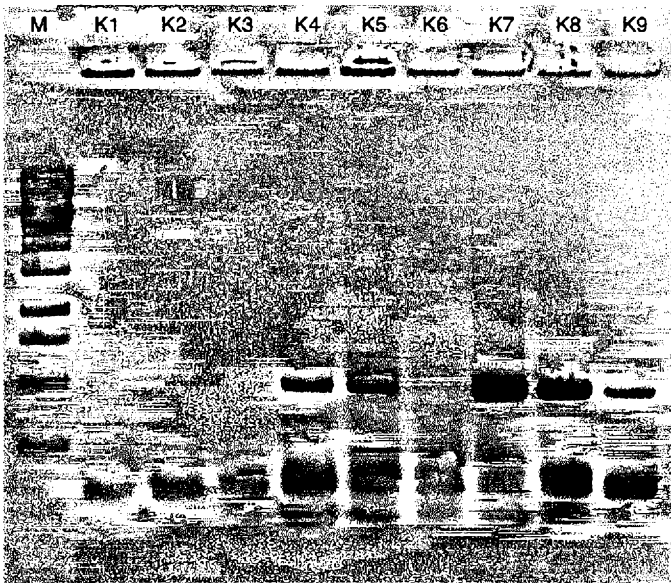


Fig. 9

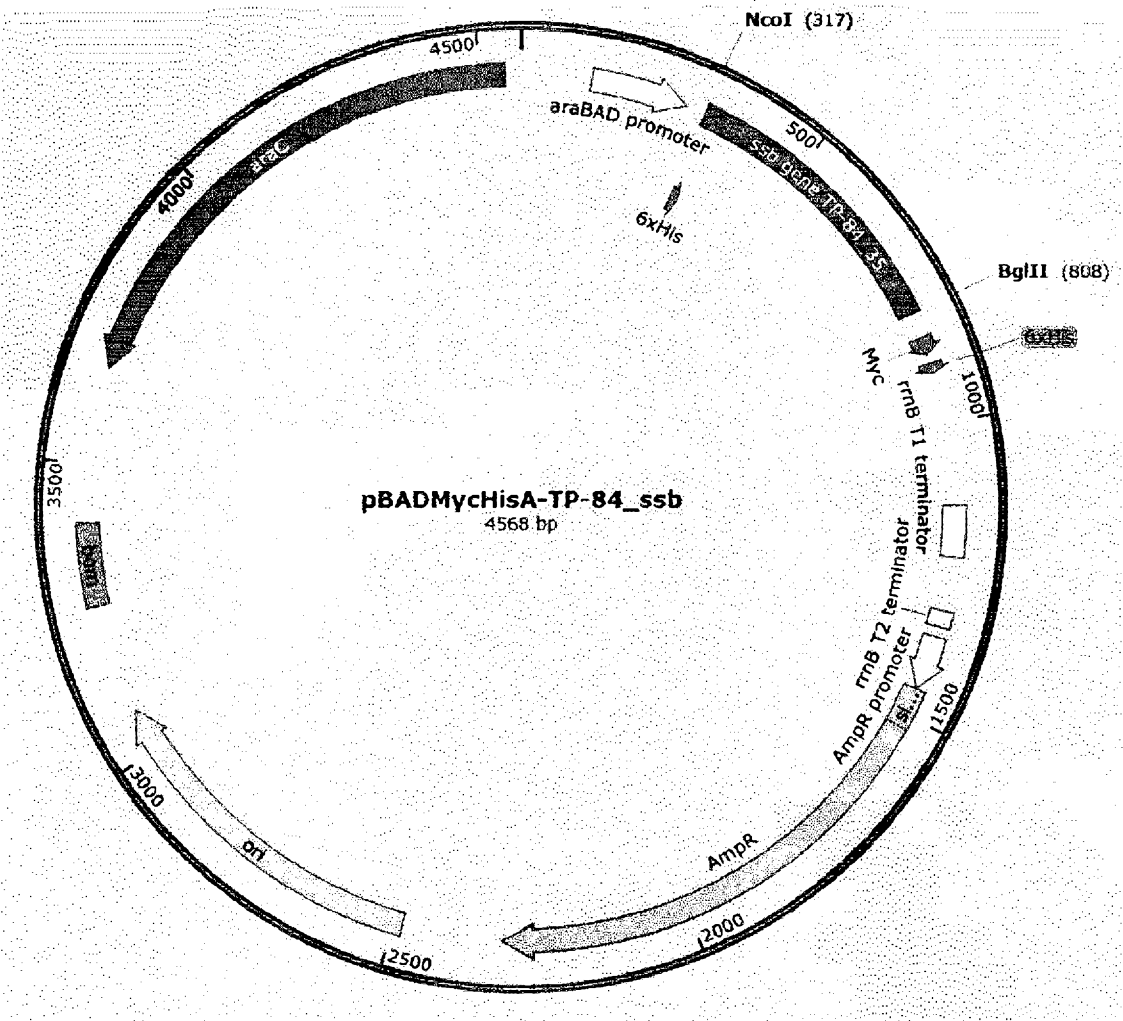


Fig. 10

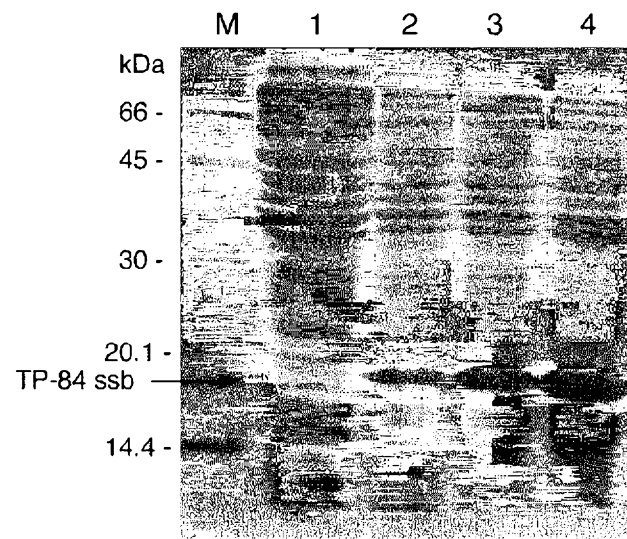


Fig. 11

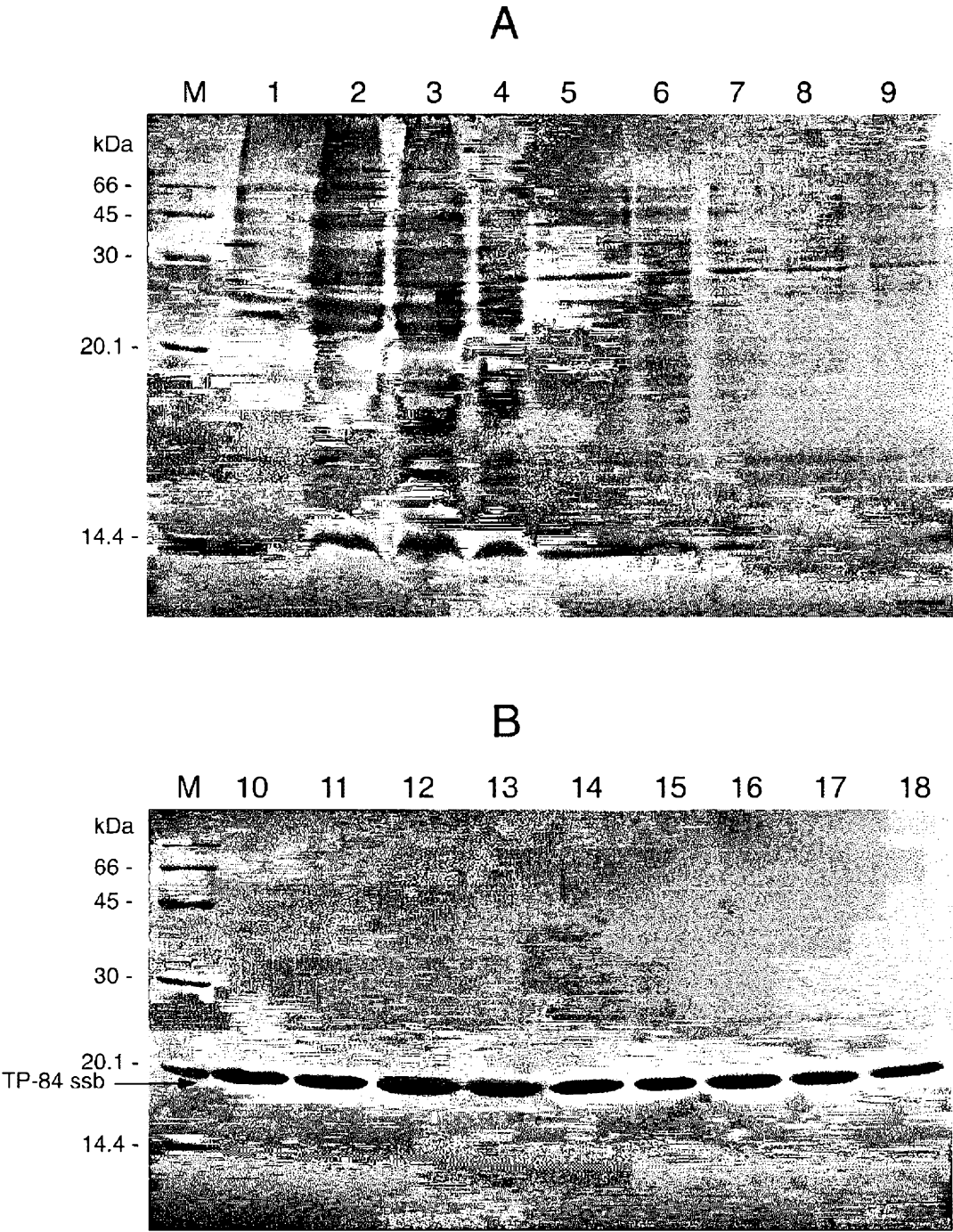


Fig. 12

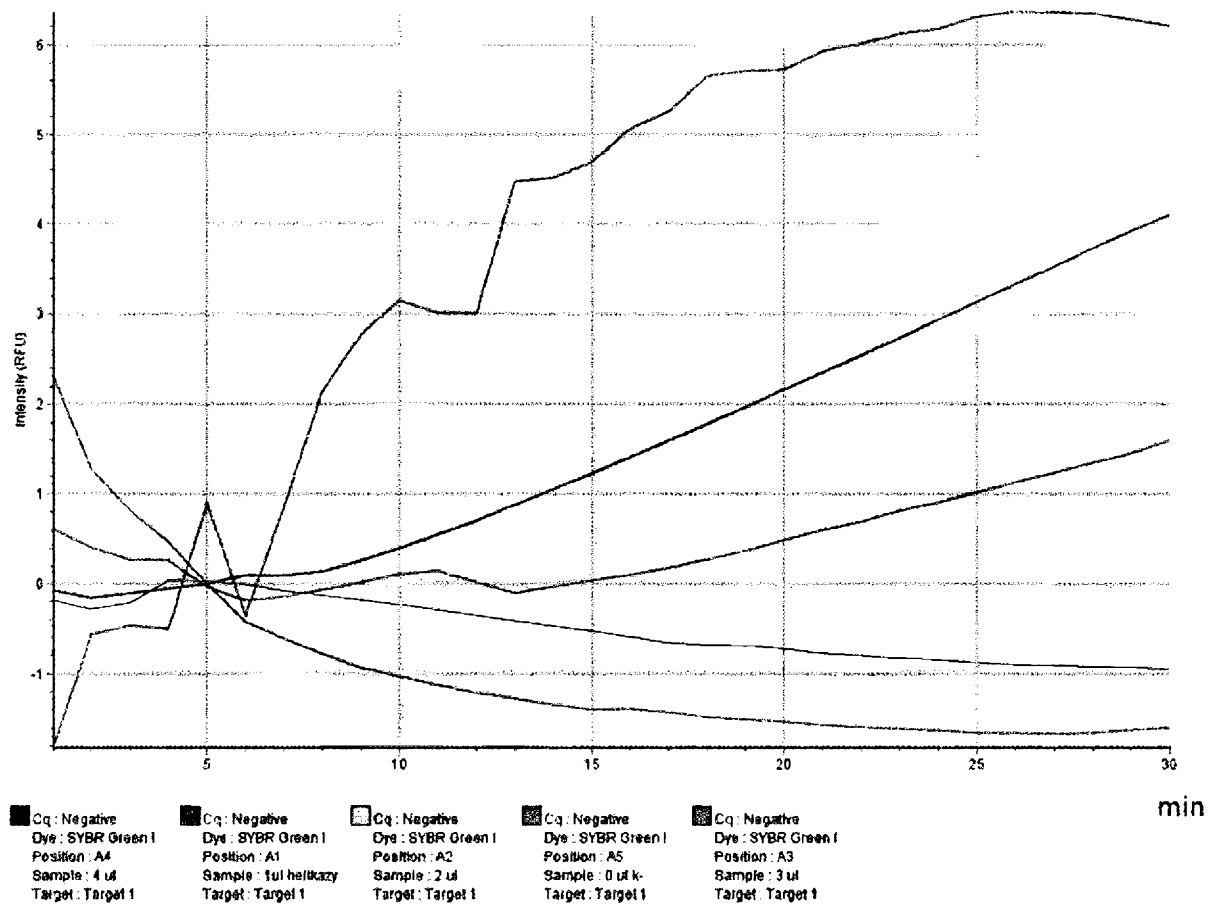


Fig. 13

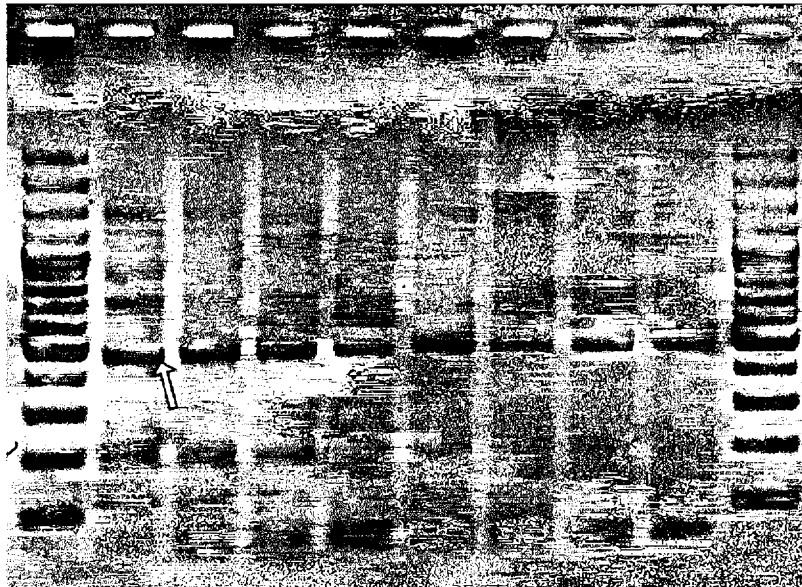


Fig. 14