

SIB EX4836R

G0164-SM-EPT

Traduzione in lingua italiana del Brevetto Europeo
domanda n° 13723086.8, pubblicazione n° 2850100

a nome di UCB Biopharma SPRL

di Allée de la Recherche 60, 1070 Brussels, Belgio

“CELLULE BATTERICHE RICOMBINANTI PER L’ESPRESSIONE DI PROTEINE”

Jacopo de Benedetti
USBM-043R B

Descrizione

L'invenzione riguarda un ceppo batterico ospite, in particolare *E. coli*. L'invenzione riguarda anche un metodo per produrre una proteina di interesse in tale cellula.

Retroscena dell'invenzione

Le cellule batteriche, come *E. coli*, sono comunemente utilizzate per produrre proteine ricombinanti. Vi sono numerosi vantaggi nell'utilizzare cellule batteriche, quali *E. coli*, per produrre proteine ricombinanti in particolare data la natura versatile delle cellule batteriche come cellule ospiti che permettono l'inserzione di geni mediante plasmidi. *E. coli* è stato utilizzato per produrre molte proteine ricombinanti inclusa l'insulina.

Nonostante i numerosi vantaggi nell'uso delle cellule batteriche per produrre proteine ricombinanti, vi sono ancora importanti limitazioni che includono la difficoltà nel produrre proteine sensibili alle proteasi. Le proteasi giocano un ruolo importante nel degradare proteine vecchie, danneggiate o mal ripiegate nel periplasma e nel citoplasma di

E. coli. Le proteasi batteriche agiscono in modo da degradare le proteine ricombinanti di interesse, riducendo in modo significativo, quindi, la resa di proteina attiva.

Sono state identificate numerose proteasi. In *E. coli* sono state identificate proteasi che includono la proteasi III (ptr), DegP, OmpT, Tsp, prlC, ptrA, ptrB, pepA-T, tsh, espC, eatA, clpP e Ion.

Tsp (anche nota come Prc) è una proteasi periplasmica da 60kDa. Il primo substrato noto della Tsp è stata la proteina legante la penicillina 3 (PBP3) (Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3 of *Escherichia coli*; Nagasawa H, Sakagami Y, Suzuki A, Suzuki H, Hara H, Hirota Y. *J Bacteriol.* 1989 Nov;171(11):5890-3 e Cloning, mapping and characterization of the *Escherichia coli* Tsp gene which is involved in C-terminal processing of penicillin-binding protein 3; Hara H, Yamamoto Y, Higashitani A, Suzuki H, Nishimura Y. *J Bacteriol.* 1991 Aug;173 (15):4799-813) ma è stato successivamente scoperto che la Tsc è anche in grado di tagliare proteine della coda fagica, ed è stata quindi rinominata come proteasi coda specifica "Tail Specific Protease" (Tsp) (Silber et al., *Proc. Natl. Acad. Sci. USA*, 89: 295-299 (1992)). Silber et al. (Deletion of the *prc(tsp)* gene provides evidence for additional tail-specific proteolytic

activity in *Escherichia coli* K-12; Silber, K.R., Sauer, R.T.; *Mol Gen Genet* 1994 242:237-240) descrive una un ceppo di delezione *prc* (KS1000) in cui la mutazione è stata creata rimpiazzando un segmento del gene *prc* con un frammento comprendente un marcatore Kan^r.

La riduzione dell'attività della Tsp (*prc*) è desiderabile al fine di ridurre la proteolisi di proteine di interesse. Tuttavia, è stato trovato che cellule prive della proteasi *prc* mostrano una crescita termoselettiva a bassa osmolarità. Hara et al hanno isolato revertanti termoresistenti contenenti mutazioni soppressore extrageniche (*spr*) (Hara et al., *Microbial Drug Resistance*, 2: 63-72 (1996)). *Spr* è una proteasi periplastica legata alla membrana da 18kDa e i substrati di *spr* sono Tsp peptidoglicani nella membrana esterna coinvolti nell'idrolisi della parete cellulare durante la divisione cellulare. Il gene *spr* è designato come UniProtKB/Swiss-Prot P0AFV4 (SPR_ECOLI).

Ceppi con deficienza proteasica migliorati comprendenti un gene mutante *spr* sono stati descritti. Chen et al descrive la costruzione di ceppi di *E. coli* che portano diverse combinazioni di mutazioni in *prc* (Tsp) e in un'altra proteasi, DegP, create amplificando le regioni a monte e a valle del gene e legandole insieme in un vettore comprendente marcatori di selezione e una mutazione *sprW174R* (High-level accumulation of a recombinant antibody fragment in the periplasm of *Escherichia coli* requires a triple-mutant (Δ DegP Δ *prc sprW174R*) host strain (Chen C, Snedecor B, Nishihara JC, Joly JC, McFarland N, Andersen DC, Battersby JE, Champion KM. *Biotechnol Bioeng.* 2004 Mar 5;85(5):463-74). La combinazione delle mutazioni Δ DegP, Δ *prc* e *sprW174R* è stata trovata fornire i maggiori livelli di catena leggera di anticorpo, di catena pesante di anticorpo e F(ab')₂-LZ. EP1341899 rivela un ceppo di *E. coli* deficiente nei DegP e *prc* cromosomici codificanti le proteasi DegP e Prc, rispettivamente, e porta un gene *spr* mutante che codifica una proteina che sopprime i fenotipi di crescita mostrati dai ceppi con mutanti *prc*.

Sono stati descritti altri ceppi proteasi-deficienti migliorati contenenti mutazioni sia in Tsp che in *spr* in WO2011/086136, e in WO 2012/013930.

I ceppi rivelati in WO02/48376 sono lac⁻ e non possono crescere in colture in cui sono impiegati timidina, fucosio o maltosio come fonte di carbonio. Questo può essere un grave svantaggio per ceppi da utilizzare su

scala commerciale. Vi possono essere altri svantaggi associati ai ceppi, ad esempio la mancanza di produzione di fosfatasi alcalina. Quest'ultima è una proteina periplasmica coinvolta nell'utilizzazione di fosfato dal mezzo di coltura.

Certe proteine mostrano attività peptidil-prolil isomerasica e/o isomerasica e/o di chaperone ed è stato trovato che forniscono proprietà vantaggiose quando impiegate in linee cellulari utilizzate per l'espressione di proteine.

La presente invenzione fornisce nuovi ceppi batterici che hanno mutazioni sia Tsp che spr e almeno un gene codificante una proteina o proteine in grado di facilitare il ripiegamento proteico che forniscono mezzi vantaggiosi per produrre proteine ricombinanti.

Sommario dell'invenzione

È qui descritta una cellula batterica gram-negativa ricombinante comprendente:

a. un gene spr mutato codificante una proteina avente una mutazione in uno o più degli amminoacidi selezionati tra D133, H145, H157, N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, V135, L136, G140, R144 and G147 and

b. un gene o geni in grado di esprimere o sovra esprimere una o più proteine in grado di facilitare il ripiegamento proteico, come FkpA, Skp, SurA, PPIA and PPIID

in cui la cellula ha una attività ridotta della proteina Tsp paragonata ad una cellula selvatica.

E' inoltre rivelata una cellula gram-negativa ricombinante codificante:

a. un gene spr mutato codificante una proteina avente una mutazione in uno o più degli amminoacidi selezionati tra D133, H145, H157, N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, V135, L136, G140, R144 and G147,

b. un gene o geni in grado di esprimere o sovra esprimere una o più proteine in grado di facilitare il ripiegamento proteico, come FkpA, Skp, SurA, PPIA and PPIID,

c. un gene in grado di esprimere una proteina di interesse, ad esempio un anticorpo o un suo frammento legante in cui la cellula ha una attività ridotta della proteina Tsp paragonata ad una cellula selvatica e il resto del DNA genomico è isogenico rispetto alla cellula di tipo selvatico da cui è stata derivata la cellula ricombinante.

Secondo un aspetto, il genoma cellulare è isogenico rispetto ad una cellula batterica di tipo selvatico eccetto per il gene spr mutato, la modifica essendo richiesta per ridurre l'attività della proteina Tsp rispetto ad una cellula di tipo selvatico ed il gene o i geni esprimendo una proteina in grado di facilitare il ripiegamento proteico.

E' inoltre qui rivelata una cellula batterica ricombinante gram-negativa avente un'attività ridotta della proteina Tsp rispetto ad una cellula selvatica e comprendente un gene spr mutante codificante una proteina spr, in cui il genoma cellulare è isogenico rispetto ad una cellula batterica selvatica esclusa la modifica necessaria a ridurre l'attività della proteina Tsp rispetto alla cellula selvatica, il gene spr mutato e il gene o i geni introdotti per esprimere una proteina in grado di facilitare il ripiegamento proteico.

Le cellule qui rivelate mostrano fenotipi vantaggiosi di crescita e di produzione proteica.

È ulteriormente qui rivelato un metodo per produrre una proteina di interesse comprendente l'esprimere la proteina di interesse in una cellula batterica gram-negativa come sopra definite.

Breve descrizione delle figure

Figura 1

mostra i risultati di fermentazioni su scala 5L effettuate con varie combinazioni di cellule ospite e "chaperone". W3110 è un ceppo selvatico di *E coli*. le varie combinazioni erano: selvatico senza chaperone; selvatico con FkpA e Skp; MXE016 mutante spr e Δ Tsp come pubblicato in WO2011/086136; MXE016 e FkpA; MXE016 e Skp; MXE016 e FkpA e Skp; MXE017 rivelato in WO2011/086136; MXE017 e FkpA e Skp

Figura 2

mostra i risultati di esperimenti di velocità di alimentazione a velocità di alimentazione post induzione 5.4, 6.0 e 7.0g/h, per MXE016 la maggior parte del Fab' aggiuntivo fatto a velocità di alimentazione maggiori era perso nel sovrantante.

Figure 3A,B

Mostra la vitalità cellulare (3A) e i titoli di Fab' (3B) per MXE016 +/- FkpA.

Figura 4

Mostra i dati primari di recupero per una produzione pilota su scala 20L.

Figura 5

Mostra il recupero primario su SDS-PAGE colorato in condizioni non riducenti di una produzione pilota su scala 20L. Oltre alle bande correlate a FkpA il profilo proteico appare molto simile tra i due ceppi.

Figura 6

Mostra un His-tag western in condizioni non riducenti per un procedimento su scala pilota da 20L. La FkpA a lunghezza intera rivelata corrisponde alla banda da ~30kDa e non è stato rivelato segnale nelle sole MXE016 come atteso.

Figura 7A-C

Mostra varie mutazioni in vari geni.

Figura 7D

Mostra una rappresentazione diagrammatica della creazione di un vettore comprendente una sequenza polinucleotidica codificante una catena leggera di un anticorpo (LC), una catena pesante di un anticorpo (HC), una sequenza polinucleotidica FkpA e/o un polinucleotide Skp

Figura 8

Mostra varie sequenze polinucleotidiche e amminoacidiche.

Breve descrizione delle sequenze

SEQ ID NO:1 è la sequenza di DNA del gene Tsp selvatico inclusi i sei nucleotidi ATGAAC a monte del codone d'inizio.

SEQ ID NO:2 è la sequenza amminoacidica della proteina Tsp selvatica.

SEQ ID NO:3 è la sequenza di DNA di un gene knockout Tsp mutato inclusi i sei nucleotidi ATGAAT a monte del codone d'inizio.

SEQ ID NO:4 è la sequenza di DNA del gene selvatico della Proteasi III.

SEQ ID NO:5 è la sequenza amminoacidica della proteina Proteasi III selvatica.

SEQ ID NO:6 è la sequenza di DNA di un gene knockout mutato della Proteasi III.

SEQ ID NO:7 è la sequenza di DNA del gene DegP selvatico.

SEQ ID NO:8 è la sequenza amminoacidica della proteina selvatica DegP.

SEQ ID NO:9 è la sequenza di DNA di un gene DegP mutato.

SEQ ID NO:10 è la sequenza amminoacidica di una proteina mutata DegP.

SEQ ID NO: 11 è la sequenza del primer oligonucleotidico 5' per la regione del gene mutato DegP comprendente il sito di restrizione Ase I.

SEQ ID NO: 12 è la sequenza del primer oligonucleotidico 3' per la regione del gene mutato DegP comprendente il sito di restrizione Ase I.

SEQ ID NO: 13 è la sequenza del primer oligonucleotidico 5' per la regione del gene mutato DegP comprendente il sito di restrizione Ase I.

SEQ ID NO: 14 è la sequenza del primer oligonucleotidico 3' per la regione del gene mutato Proteasi III comprendente il sito di restrizione Ase I.

SEQ ID NO: 15 è la sequenza del primer oligonucleotidico 5' per la regione del gene mutato Proteasi III comprendente il sito di restrizione Ase I.

SEQ ID NO: 16 è la sequenza del primer oligonucleotidico 3' per la regione del gene mutato DegP comprendente il sito di restrizione Ase I.

SEQ ID NO: 17 è la sequenza di DNA del gene selvatico spr.

SEQ ID NO: 18 è la sequenza del gene selvatico spr inclusa la sequenza segnale che sono i primi 26 residui amminoacidici.

SEQ ID NO: 19 è la sequenza del gene spr non mutato senza sequenza segnale.

SEQ ID NO: 20 è la sequenza nucleotidica di una sequenza mutata OmpT comprendente le mutazioni D210A e H212A.

SEQ ID NO: 21 è la sequenza amminoacidica di una sequenza OmpT mutata comprendente le mutazioni D210A e H212A.

SEQ ID NO: 22 è la sequenza nucleotidica di una sequenza knockout mutata OmpT.

SEQ ID NO: 23 mostra la sequenza di un adattatore oligonucleotidico OmpA.

SEQ ID NO: 24 mostra la cassetta oligonucleotidica codificante la sequenza intergenica 1 (IGS1) per l'espressione in *E. coli* di Fab.

SEQ ID NO: 25 mostra la cassetta oligonucleotidica codificante la sequenza intergenica 2 (IGS2) per l'espressione in *E. coli* di Fab.

SEQ ID NO: 26 mostra la cassetta oligonucleotidica codificante la sequenza intergenica 3 (IGS3) per l'espressione in *E. coli* di Fab.

SEQ ID NO: 27 mostra la cassetta oligonucleotidica codificante la sequenza intergenica 4 (IGS4) per l'espressione in *E. coli* di Fab.

SEQ ID NO: 28 è la sequenza di DNA del gene selvatico FkpA.

SEQ ID NO: 29 è la sequenza proteica del gene selvatico FkpA.

SEQ ID NO: 30 è la sequenza di DNA del gene FkpA con his tag.

SEQ ID NO: 31 è la sequenza proteica del gene FkpA con his tag.

SEQ ID NO: 32 è la sequenza di DNA del gene selvatico skp.

SEQ ID NO: 33 è la sequenza proteica del gene selvatico skp.

SEQ ID NO: 34 è la sequenza di DNA del gene skp con his tag.

SEQ ID NO: 35 è la sequenza proteica del gene skp con his tag.

SEQ ID NO: 36 - 74 mostrano varie sequenze amminoacidiche e di DNA per anticorpi FcRn o loro frammenti, che sono idonee all'espressione della linea cellulare della presente invenzione. In particolare, SEQ ID NO: 50 è la sequenza amminoacidica della regione variabile della catena leggera di una catena leggera di anticorpo anti-FcRn 1519gH20 e SEQ ID NO:58 è la sequenza amminoacidica della regione variabile della catena pesante di una catena pesante di anticorpo anti-FcRn 1519gH20.

Descrizione dettagliata delle forme di realizzazione preferite dell'invenzione

I presenti inventori hanno fornito cellule batteriche ricombinanti gram-negative migliorate, idonee all'espressione di una proteina ricombinante di interesse.

In una forma di realizzazione la proteina è un anticorpo o un suo frammento legante, in particolare un anticorpo terapeutico.

In particolare, gli inventori hanno fornito cellule batteriche ricombinanti gram-negative migliorate idonee all'espressione di una proteina ricombinante di interesse incorporando un gene o più geni, codificanti una proteina che facilita il ripiegamento proteico, in cellule batteriche gram-negative aventi un gene Tsp mutato e un gene spr mutato

In una forma di realizzazione il gene o i geni, codificanti la proteina per facilitare il ripiegamento proteico, sono integrati nel genoma cellulare, ad esempio a fornire una linea cellulare stabile. In una forma di realizzazione una proteina ricombinante da esprimere (come una proteina terapeutica) è trasfettata in una linea cellulare stabile a dare l'espressione della proteina ricombinante desiderata.

In una forma di realizzazione il gene o i geni, codificanti la proteina per facilitare il ripiegamento proteico, sono forniti in uno o più plasmidi, ad esempio i plasmidi sono trasfettati in modo transiente in una cellula a dare una linea cellulare della presente descrizione.

In una forma di realizzazione il gene o i geni, codificanti la proteina per facilitare il ripiegamento proteico, sono forniti in un plasmide contenente anche la sequenza codificante per una proteina ricombinante di interesse.

In una forma di realizzazione il gene o i geni, codificanti la proteina per facilitare il ripiegamento proteico, sono forniti in un plasmide che non contiene la sequenza codificante per una proteina ricombinante di interesse.

In una forma di realizzazione l'invenzione fornisce nuovi ceppi con un fenotipo di crescita cellulare migliorato rispetto alle cellule batteriche selvatiche e cellule che hanno solamente un gene Tsp mutato o un gene Tsp e un gene spr mutati.

Le cellule della presente invenzione hanno numerosi vantaggi. Gli inventori hanno sorprendentemente trovato che cellule secondo la presente descrizione possono mostrare una vitalità cellulare aumentata rispetto alla cellula selvatica o ad una cellula che comprende un gene Tsp e un gene spr mutati.

La vitalità cellulare è di particolare importanza in termini pratici perché cellule che non sono vitali tendono a lissarsi e creare frammenti di DNA nella coltura. Questi frammenti di DNA aumentano la difficoltà, i costi e le

spese per purificare la proteina desiderata. Quindi, minimizzare i frammenti di DNA da cellule non-vitali lisate è una questione importante nel produrre in modo efficiente proteine ricombinanti, vedere ad esempio US 6,258,560.

La vitalità cellulare si può misurare in qualsiasi tra varie tecniche di routine, ad esempio impiegando un colorante fluorescente ed analisi FACS o simili.

Specificamente, le cellule secondo l'espressione, mostrano generalmente un fenotipo di lisi cellulare ridotta rispetto a cellule con un gene Tsp e un gene spr mutati.

Inoltre, i nuovi ceppi possono ridurre la perdita di proteine dalle cellule e permettere un accumulo periplasmico prolungato rispetto a cellule con un gene Tsp e un gene spr mutati. Questo è particolarmente importante perché, ad esempio, quando i livelli totali di espressione della proteina di interesse sono simili, ma è accumulata più proteina nel periplasma o ne è accumulata meno nel sovrantante dato che il ceppo perde meno proteina, allora le cellule con queste proprietà di perdere meno proteina saranno generalmente più idonee alla produzione in larga scala perché facilitano il recupero della proteina.

Inoltre, le cellule secondo la presente descrizione possono mostrare una maggiore resa di proteina di interesse rispetto a cellule batteriche selvatiche o a cellule che comprendono un gene Tsp e uno spr mutati in assenza di un gene o geni che codificano una proteina come FkpA. La resa proteica migliorata può essere la resa di proteina periplasmica e/o la resa di proteina nel sovrantante. In una forma di realizzazione le cellule della presente invenzione mostrano una resa di proteina periplasmica paragonate ad una cellula che ha un gene Tsp e un gene spr mutati a causa della riduzione di perdita dalla cellula.

Le cellule batteriche ricombinanti possono essere in grado di avere una maggiore velocità di produzione di una produzione di interesse e, quindi, la stessa quantità di proteina di interesse può essere prodotta in un tempo minore rispetto ad una cellula batterica selvatica o a una cellula con un gene Tsp e un gene spr mutati. La velocità di produzione maggiore di una proteina di interesse può essere particolarmente significativa nel periodo di crescita cellulare iniziale, ad esempio nelle prime 5, 10, 20 o 30 ore successive all'induzione dell'espressione proteica.

Le cellule secondo la presente invenzione preferibilmente mostrano una resa massima nel periplasma e/o nel terreno di circa 1.0g/L, 1.5g/L, 1.8g/L, 2.0g/L, 2.4g/L, 2.5g/L, 3.0g/L, 3.5g/L o 4.0g/L di una proteina di interesse.

Inoltre, l'espressione di una proteina o di proteine che facilitano il ripiegamento, ottimizza ulteriormente l'espressione aumentando la proteina fornita con un ripiegamento idoneo. L'esperto del settore sa bene che un ripiegamento idoneo è necessario alla funzione biologica e quindi isolare proteine con il ripiegamento desiderato è di vitale importanza. Questo è particolarmente importante quando la proteina è espressa in cellule gram-negative perché la proteina espressa non sarà una proteina naturale per la cellula e quindi la cellula potrebbe non essere in grado di esprimere automaticamente la proteina con il ripiegamento giusto. Un ripiegamento errato si può manifestare come una aggregazione o altre impurità. L'isolamento della proteina desiderata può richiedere ampie purificazioni che hanno implicazione di costo e possono anche risultare in una bassa resa di proteina desiderata. Massimizzare la quantità di proteina correttamente ripiegata espressa minimizza la quantità di purificazione necessaria e può ottimizzare la resa utilizzabile ed è quindi vantaggioso.

Vantaggi associati con l'espressione di una proteina che facilita il ripiegamento includono uno o più dei seguenti: titolo più elevato (ad esempio aumentato a circa 1.05g/L rispetto al selvatico 0.5g/L); maggiore raccolta di vitalità (ad esempio >95%); titolo aumentato con maggiore velocità di alimentazione (migliori prospettive per lo sviluppo del procedimento); titolo maggiore a scale di produzione commerciali come scale da 20L e maggiore vitalità alla raccolta; e chiarificazione dell'estratto più facile, in particolare ad una scala da 20L.

Le cellule fornite dalla presente invenzione hanno una ridotta attività proteasica di Tsp rispetto a una cellula selvatica, che può ridurre la proteolisi di una proteina di interesse, in particolare di proteine di interesse che sono proteoliticamente sensibili a Tsp. Pertanto, le cellule fornite dalla presente invenzione possono fornire una maggiore resa di proteine intatte, preferibilmente della proteina di interesse e una resa inferiore, o preferibilmente nessun frammento proteolitico di proteine, preferibilmente della proteina di interesse, rispetto a una cellula batterica selvatica.

Sono qui descritte cellule che portano solo le minime mutazioni al genoma necessarie per introdurre le modifiche

secondo la presente descrizione. La cellula batterica può differire da una cellula batterica selvatica solo per una o più mutazioni nel gene spr e per la modifica necessaria per ridurre l'attività della proteina Tsp rispetto a una cellula selvatica perché per esempio il gene o geni codificanti la proteina per facilitare il ripiegamento proteico può essere introdotto in modo transiente nella cellula, ad esempio in un plasmide. Le cellule qui descritte non hanno altre mutazioni che possono avere effetti deleteri sulla capacità della cellula di crescere e/o di esprimere una proteina di interesse.

Di conseguenza, una o più delle cellule ospiti ricombinanti qui divulgate può mostrare una migliore espressione proteica e/o caratteristiche di crescita migliorate rispetto a cellule comprendenti ulteriori mutazioni geneticamente ingegnerizzate nella sequenza genomica. Le cellule qui divulgate sono anche più adatte per l'uso per produrre proteine terapeutiche rispetto a cellule comprendenti ulteriori interruzioni del genoma cellulare.

La persona esperta sarebbe facilmente in grado di testare un clone di cellule candidate per vedere se ha la resa desiderata di una proteina di interesse usando metodi ben noti nell'arte, incluso un metodo di fermentazione, ELISA e proteina G HPLC. I metodi di fermentazione adatti sono descritti in Humphreys D P, et al. (1997). Formation of dimeric Fabs in E. coli: effect of hinge size and isotype, presence of interchain disulphide bond, Fab' expression levels, tail piece sequences and growth conditions. J. IMMUNOL. METH. 209: 193-202; Backlund E. Reeks D. Markland K. Weir N. Bowering L. Larsson G. Fedbatch design for periplasmic product retention in Escherichia coli, Journal Article. Research Support, Non-U.S. Gov't Journal of Biotechnology. 135(4):358-65, 2008 Jul 31; Champion KM. Nishihara JC. Joly JC. Arnott D. Similarity of the Escherichia coli proteome upon completion of different biopharmaceutical fermentation processes. [Journal Article] Proteomics. 1(9):1133-48, 2001 Sep; and Horn U. Strittmatter W. Krebber A. Knupfer U. Kujau M. Wenderoth R. Muller K. Matzku S. Pluckthun A. Riesenberg D. High volumetric yields of functional dimeric miniantibodies in Escherichia coli, using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions, Journal Article. Research Support, Non-U.S. Gov't Applied Microbiology & Biotechnology. 46(5-6):524-32, 1996 Dec. The skilled person would also easily be able to test secreted protein to see if the protein is correctly folded using methods well known in the art, such as protein G HPLC, circular

dichroism, NMR, X-Ray crystallography e epitope affinity measurement methods.

La presente invenzione sarà ora descritta in maggior dettaglio.

I termini "proteina" e "polipeptide" sono usati in modo intercambiabile nel presente contesto, a meno che il contesto non indichi diversamente. "Peptide" è inteso riferirsi a 10 o meno amminoacidi.

I termini "polinucleotide" includono un gene, DNA, cDNA, RNA, mRNA ecc. a meno che il contesto non indichi diversamente.

"Attività ridotta" come qui impiegata si riferisce a "livelli più bassi di attività enzimatica, come l'attività enzimatica di Tsp in confronto alla corrispondente attività enzimatica in un ceppo di tipo selvatico quando misurati in condizioni comparabili in un saggio adatto. In una forma di realizzazione l'attività ridotta è del 50% o meno, 40% o meno, 30% o meno, 20% o meno, 10% o meno o 5% o meno dell'attività enzimatica di un comparatore selvatico. In una forma di realizzazione l'analisi per determinare i livelli di attività enzimatica è eseguita concomitantemente quando i risultati devono essere impiegati in un confronto diretto.

Il confronto diretto come qui impiegato si riferisce al punto in cui il valore numerico di due o più risultati è confrontato allo scopo di valutare se vi è un'attività ridotta come qui definita o classificare i risultati di attività ottenuti da un saggio.

Come qui usato, il termine "comprendente" nel contesto della presente specifica dovrebbe essere interpretato come "includente".

Cella selvatica come qui impiegato può essere utilizzato in modo intercambiabile con cellula di controllo o cellula non mutata.

La cellula di controllo o cellula non mutata nel contesto della presente invenzione indica una cellula dello stesso tipo della cellula gram-negativa ricombinante dell'invenzione in cui la cellula non è stata modificata per realizzare la suddetta riduzione di attività della proteina Tsp e per portare il gene mutante spr. Ad esempio, una cellula non mutata può essere una cellula selvatica e può essere derivata dalla stessa popolazione di cellule ospiti come le cellule dell'invenzione prima della modifica per introdurre qualsiasi mutazione.

Le espressioni "cellula", "linea cellulare", "coltura cellulare" e "ceppo" sono usate in modo intercambiabile.

L'espressione "fenotipo di una cellula comprendente un gene Tsp mutato" nel contesto della presente invenzione significa il fenotipo esibito da una cellula avente un gene mutante di Tsp. Tipicamente le cellule che comprendono un gene mutante di Tsp possono lisarsi, specialmente ad alte densità cellulari. La lisi di queste cellule causa la fuoriuscita di eventuali proteine ricombinanti nel sovrinatante. Le cellule portatrici del gene Tsp mutato possono anche mostrare una crescita termosensibile a bassa osmolarità. Ad esempio, le cellule non mostrano alcun tasso di crescita ridotto o le cellule muoiono in mezzi ipotonici ad alta temperatura, ad esempio a 40 ° C o più.

Il termine "isogenico" nel contesto della presente invenzione significa che il genoma della cellula della presente invenzione ha sostanzialmente la stessa o la stessa sequenza genomica rispetto alla cellula selvatica da cui è derivata la cellula eccetto per un gene spr mutato e la modifica richiesta per ridurre l'attività della proteina Tsp rispetto a una cellula selvatica. In questa forma di realizzazione il genoma della cellula non comprende ulteriori mutazioni geneticamente modificate non presenti in natura. In una forma di realizzazione la cellula secondo la presente invenzione può avere sostanzialmente la stessa sequenza genomica rispetto alla cellula selvatica eccetto per il gene spr mutato e la modifica richiesta per ridurre l'attività della proteina Tsp rispetto a una cellula selvatica tenendo in considerazione qualsiasi mutazione naturalmente presente in natura che possa verificarsi. In una forma di realizzazione, la cellula secondo la presente invenzione può avere esattamente la stessa sequenza genomica rispetto alla cellula selvatica eccetto per il gene spr mutato e la modifica richiesta per ridurre l'attività della proteina Tsp rispetto a una cellula selvatica.

Il termine "selvatico" nel contesto della presente invenzione indica un ceppo di una cellula batterica gram-negativa così come può essere presente in natura o può essere isolato dall'ambiente, che non porta mutazioni geneticamente ingegnerizzate. Un esempio di ceppo selvatico di *E. coli* è W3110, come il ceppo W3110 K-12.

Qualsiasi batterio gram-negativo adatto può essere usato come cellula parentale per produrre la cellula ricombinante della presente invenzione. Batteri gram-negativo adatti includono *Salmonella typhimurium*, *Pseudomonas fluorescens*, *Erwinia carotovora*, *Shigella*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Acinetobacter baumannii* e *E. coli*. Preferibilmente la cellula parentale è *E. coli*.

Qualsiasi ceppo adatto di *E. coli* può essere usato nella presente invenzione, ma preferibilmente è usato un ceppo selvatico W3110, come K-12 W3110.

Un inconveniente associato ai ceppi batterici proteasi deficienti precedentemente creati e utilizzati per esprimere proteine ricombinanti è che essi comprendono ulteriori mutazioni di geni coinvolti nel metabolismo cellulare e nella replicazione del DNA come, per esempio *phoA*, *fhuA*, *lac*, *rec*, *gal*, *ara*, *arg*, *thi* e *pro* in ceppi di *E. coli*. Queste mutazioni possono avere molti effetti deleteri sulla cellula ospite, compresi gli effetti sulla crescita cellulare, la stabilità, la resa di espressione della proteina ricombinante e la tossicità. I ceppi che presentano una o più di queste mutazioni genomiche, in particolare ceppi con un numero elevato di queste mutazioni, possono presentare una perdita di fitness che riduce il tasso di crescita batterica a un livello che non è adatto per la produzione industriale di proteine. Inoltre, una qualsiasi delle suddette mutazioni genomiche può influenzare altri geni in *cis* e/o in *trans* in modi dannosi imprevedibili alterando così il fenotipo del ceppo, la fitness e il profilo proteico. Inoltre, l'uso di cellule fortemente mutate non è generalmente adatto per produrre proteine ricombinanti per uso commerciale, in particolare terapeutiche, poiché tali ceppi hanno generalmente pathway metabolici difettosi e quindi possono crescere male o per nulla in un terreno minimo o chimicamente definito.

In una forma di realizzazione una cellula secondo la presente invenzione è isogenica ad una cellula batterica selvatica ad eccezione del gene *spr* mutato e della modifica richiesta per ridurre l'attività della proteina Tsp rispetto a una cellula selvatica. Sono fatte al genoma della cellula solo le mutazioni minimali per introdurre le mutazioni. Le cellule non portano altre mutazioni che possono avere effetti deleteri sulla capacità della cellula di crescere e/o di esprimere una proteina di interesse. Di conseguenza, una o più delle cellule ospiti ricombinanti della presente invenzione può mostrare una migliore espressione proteica e/o caratteristiche di crescita migliorate rispetto a cellule comprendenti ulteriori mutazioni geneticamente ingegnerizzate nella sequenza genomica. Le cellule fornite dalla presente invenzione sono anche più adatte per l'uso nella produzione di proteine terapeutiche rispetto a cellule comprendenti ulteriori interruzioni del genoma cellulare.

In una forma di realizzazione preferita, la cellula è isogenica a una cellula *E. coli* selvatica, come il ceppo W3110, eccetto per il gene *spr* mutato e la modifica richiesta per ridurre l'attività della proteina Tsp rispetto a

una cellula selvatica.

La cellula della presente invenzione può inoltre differire da una cellula selvatica nel fatto di comprendere un polinucleotide codificante per la proteina di interesse. La sequenza polinucleotidica codificante per la proteina di interesse può essere esogena o endogena. Il polinucleotide codificante per la proteina di interesse può essere contenuto in un vettore di espressione adatto trasformato nella cellula e/o integrato nel genoma della cellula ospite. Nella forma di realizzazione in cui il polinucleotide codificante per la proteina di interesse è inserito nel genoma dell'ospite, la cellula della presente invenzione differirà da una cellula selvatica anche a causa della sequenza polinucleotidica inserita codificante la proteina di interesse. Preferibilmente il polinucleotide si trova nella cellula all'interno di un vettore di espressione, causando in tal modo un'interruzione minima del genoma della cellula ospite.

La proteina spr è una proteasi periplasmatica legata alla membrana *E. coli*.

La sequenza di amminoacidi selvatico della proteina spr è mostrata in SEQ ID NO: 21 con la sequenza del segnale all'N-terminale e in SEQ ID NO: 22 senza la sequenza segnale di 26 amminoacidi (secondo il numero di accesso UniProt P0AFV4). La numerazione degli amminoacidi della sequenza della proteina spr nella presente invenzione include la sequenza segnale. Di conseguenza, l'amminoacido 1 della proteina spr è il primo amminoacido (Met) mostrato in SEQ ID NO: 21.

Il gene spr mutato è il gene spr cromosomico della cellula

Il gene spr mutato codifica per una proteina spr capace di sopprimere il fenotipo di una cellula comprendente inoltre un gene Tsp mutato. Cellule con un gene Tsp mutato possono avere un buon tasso di crescita cellulare, ma una limitazione di queste cellule è la loro tendenza a lisare, specialmente ad alte densità cellulari. Di conseguenza il fenotipo di una cellula comprendente un gene Tsp mutato è una tendenza a lisare, specialmente ad alte densità cellulari. Le cellule con un gene Tsp mutato mostrano anche una crescita termosensibile a bassa osmolarità. Tuttavia, le mutazioni di spr nelle cellule della presente invenzione, quando introdotte in una cellula che ha un'attività ridotta di Tsp, sopprimono il fenotipo Tsp ridotto e, pertanto, la cellula presenta una ridotta lisi, in particolare ad un'alta densità cellulare. Il fenotipo di crescita di una cellula può essere facilmente misurato da

una persona esperta nell'arte durante tecniche di agitazione della beuta o di fermentazione ad alta densità cellulare. La soppressione del fenotipo di lisi cellulare si può vedere dal tasso di crescita migliorato e/o dalla produzione di proteine ricombinanti, in particolare nel periplasma, mostrato da una cellula con un spr mutante e con attività Tsp ridotta rispetto a una cellula con del Tsp mutante e spr selvatico.

Le cellule secondo la presente invenzione comprendono un gene spr mutato codificante una proteina avente una mutazione in uno o più degli amminoacidi selezionati tra N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144, H145, G147 e H157, preferibilmente una mutazione di uno o più degli amminoacidi selezionati tra C94, S95, V98, Y115, D133, V135, H145, G147 e H157. In questa forma di realizzazione, la proteina spr preferibilmente non ha ulteriori mutazioni.

La mutazione di uno o più degli amminoacidi precedenti può essere una qualsiasi mutazione missenso adatta a uno, due o tre dei nucleotidi che codificano l'amminoacido. La mutazione modifica il residuo amminoacidico di un qualsiasi amminoacido adatto che risulta in una proteina mutata in grado di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato. La mutazione missenso può cambiare l'amminoacido con uno che ha una dimensione diversa e/o ha proprietà chimiche diverse rispetto all'amminoacido selvatico.

In una forma di realizzazione il gene mutante spr codifica una proteina spr avente una o più mutazioni selezionate tra C94A, S95F, V98E, Y115F, D133A, V135D o G, H145A, G147C e H157A.

In una forma di realizzazione la mutazione riguarda uno, due o tre della triade catalitica di residui amminoacidici di C94, H145, e H157 (Solution NMR Structure of the NlpC/P60 Domain of Lipoprotein Spr from Escherichia coli Structural Evidence for a Novel Cysteine Peptidase Catalytic Triad, Biochemistry, 2008, 47, 9715-9717).

Di conseguenza, il gene spr mutato può comprendere: una mutazione in C94; o una mutazione in H145; o una mutazione in H157; o una mutazione a C94 e H145; o una mutazione a C94 e H157; o una mutazione in H145 e H157; o una mutazione in C94, H145 e H157.

In questa forma di realizzazione, la proteina spr preferibilmente non ha ulteriori mutazioni.

Uno, due o tre tra C94, H145 e H157 possono essere mutati in qualsiasi amminoacido adatto che porti ad una proteina spr capace di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato. Ad esempio, uno,

due o tre tra C94, H145, e H157 possono essere mutati in un amminoacido piccolo come Gly o Ala. Di conseguenza, la proteina spr può avere una, due o tre delle mutazioni C94A, H145A e H157A. In una forma di realizzazione, il gene spr comprende la mutazione missenso C94A, che è stata trovata produrre una proteina spr capace di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato. In un'altra forma di realizzazione, il gene spr comprende la mutazione missenso H145A, che è stata trovata produrre una proteina spr capace di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato.

La designazione di una sostituzione mutante qui consiste in una lettera seguita da un numero seguito da una lettera. La prima lettera indica l'amminoacido nella proteina selvatica. Il numero si riferisce alla posizione dell'amminoacido in cui è effettuata la sostituzione dell'amminoacido, e la seconda lettera designa l'amminoacido usato per sostituire l'amminoacido selvatico.

In una forma di realizzazione la proteina mutante spr comprende una mutazione in uno o più amminoacidi scelti tra N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 e G147, preferibilmente una mutazione in uno o più amminoacidi scelti tra S95, V98, Y115, D133, V135 e G147. In questa forma di realizzazione, la proteina spr preferibilmente non ha ulteriori mutazioni. Di conseguenza, il gene mutato può comprendere: una mutazione in N31; o una mutazione in R62; o una mutazione in I70; o una mutazione in Q73; o una mutazione in S95; o una mutazione in V98; o una mutazione in Q99; o una mutazione in R100; o una mutazione in L108; o una mutazione in Y115; o una mutazione in D133; o una mutazione in V135; o una mutazione in L136; o una mutazione in G140; o una mutazione in R144; o una mutazione in G147.

In una forma di realizzazione la proteina mutante spr comprende mutazioni multiple negli amminoacidi: S95 e Y115; o N31, Q73, R100 e G140; o Q73, R100 e G140; o R100 e G140; o Q73 e G140; o Q73 e R100; o R62, Q99 e R144; o Q99 e R144.

Uno o più degli amminoacidi N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 e G147 possono essere mutati in qualsiasi amminoacido adatto che risulti in una proteina spr capace di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato. Ad esempio, uno o più di N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140 e

R144 possono essere mutati in un amminoacido piccolo come Gly o Ala. In una forma di realizzazione la proteina spr comprende una o più delle seguenti mutazioni: N31Y, R62C, I70T, Q73R, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D o V135G, L136P, G140C, R144C e G147C. In una forma di realizzazione la proteina spr comprende una o più delle seguenti mutazioni: S95F, V98E, Y115F, D133A, V135D o V135G e G147C. In questa forma di realizzazione, la proteina spr preferibilmente non ha ulteriori mutazioni. In una forma di realizzazione la proteina spr ha una mutazione selezionata tra N31Y, R62C, I70T, Q73R, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D o V135G, L136P, G140C, R144C e G147C. In questa forma di realizzazione, la proteina spr preferibilmente non ha ulteriori mutazioni.

In un'ulteriore forma di realizzazione la proteina spr ha più mutazioni scelte tra: S95F e Y115F; N31Y, Q73R, R100G e G140C; Q73R, R100G e G140C; R100G e G140C; Q73R e G140C; Q73R e R100G; R62C, Q99P e R144C; o Q99P e R144C.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione C94A.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione V103E.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione D133A.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione V135D.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione V135A.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione H145A.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione G147C.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione H157A.

In una forma di realizzazione il gene spr mutato codifica per una proteina spr con una mutazione scelta tra H145A, H157A e D133A.

In una forma di realizzazione della presente invenzione, può essere apportata al gene spr qualsiasi mutazione o mutazioni adatta che risulti in una proteina spr capace di sopprimere il fenotipo di una cellula comprendente un gene Tsp mutato. Preferibilmente, la proteina spr può avere una o più delle seguenti mutazioni: N31Y, R62C,

I70T, Q73R, C94A, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D, V135G, L136P, G140C, R144C, H145A, G147C, H157A e W174R. In una forma di realizzazione la proteina spr non comprende la mutazione W174R. Preferibilmente, il gene spr comprende una o più mutazioni discusse sopra.

Le cellule secondo la presente invenzione hanno una attività della proteina Tsp ridotta rispetto a una cellula selvatica. L'espressione "attività della proteina Tsp ridotta rispetto a una cellula selvatica" significa che l'attività Tsp della cellula è ridotta rispetto all'attività Tsp di una cellula selvatica. La cellula può essere modificata con qualsiasi mezzo adatto per ridurre l'attività di Tsp.

In una forma di realizzazione l'attività ridotta di Tsp deriva dalla modifica del polinucleotide endogeno codificante Tsp e/o delle sequenze associate alla regolazione dell'espressione. La modifica può ridurre o interrompere la trascrizione e la trascrizione del gene Tsp o può fornire una proteina Tsp espressa con attività proteasica ridotta rispetto alla proteina TsPC selvatica.

In una forma di realizzazione è modificata una sequenza associata alla regolazione dell'espressione per ridurre l'espressione di Tsp. Ad esempio, il promotore del gene Tsp può essere mutato per prevenire l'espressione del gene.

In una forma di realizzazione preferita le cellule secondo la presente invenzione portano un gene Tsp mutato che codifica per una proteina Tsp avente attività proteasica ridotta o un gene Tsp mutato knockout.

Preferibilmente è mutato il gene Tsp cromosomico.

Come usato qui, "gene Tsp" significa un gene codificante una proteasi Tsp (nota anche come Prc) che è una proteasi periplasmatica in grado di agire sulle proteine della coda del fago e sulla proteina-3 legante la penicillina (PBP3). La sequenza del gene Tsp selvatico è mostrata in SEQ ID NO: 1 e la sequenza della proteina Tsp selvatica è mostrata in SEQ ID NO: 2.

Il riferimento al gene Tsp mutato o a un gene Tsp mutato che codifica per Tsp, si riferisce a un gene Tsp mutato che codifica per una proteina Tsp con attività proteasica ridotta o a un gene Tsp mutato knockout, a meno che non sia indicato diversamente. L'espressione "gene Tsp mutato che codifica per una proteina Tsp con attività proteasica ridotta"

nel contesto della presente invenzione significa che il gene Tsp mutato non ha l'attività proteasica completa rispetto al gene Tsp selvatico non mutato. Preferibilmente, il gene Tsp mutato codifica per una proteina Tsp avente il 50% o meno, il 40% o meno, il 30% o meno, il 20% o meno, il 10% o meno o il 5% o meno dell'attività proteasica di una proteina Tsp selvatica non mutata. Più preferibilmente, il gene Tsp mutato codifica per una proteina Tsp senza attività proteasica. In questa forma di realizzazione la cellula non ha una deficienza di Tsp cromosomico, cioè la sequenza del gene Tsp non è stata cancellata o mutata per impedire l'espressione di alcuna forma di proteina Tsp. Qualsiasi mutazione adatta può essere introdotta nel gene Tsp al fine di produrre una proteina avente attività proteasica ridotta. L'attività proteasica di una proteina Tsp espressa da un batterio gram-negativo può essere facilmente testata da una persona esperta nella tecnica da qualsiasi metodo adatto nella tecnica, come il metodo descritto in Keiler et al (Identification of Active Site Residues of the Tsp Protease* THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 270, No. 48, Issue of December 1, pp. 28864-28868, 1995 Kenneth C. Keiler and Robert T. Sauer)

in cui è stata testata l'attività proteasica di Tsp.

E' stato riportato in Keiler et al (*supra*) che Tsp ha un sito attivo comprendente residui S430, D441 e K455 e residui G375, G376, E433 e T452 sono importanti per il mantenimento della struttura di Tsp. Keiler et al (*supra*) hanno riportato la scoperta che i geni Tsp mutati S430A, D441A, K455A, K455H, K455R, G375A, G376A, E433A e T452A non presentavano alcuna attività proteasica rilevabile. È stato inoltre riportato che il gene Tsp mutato S430C mostrava circa 5-10% di attività selvatica. Di conseguenza, la mutazione Tsp per produrre una proteina con attività proteasica ridotta può comprendere una mutazione, come una mutazione missenso a uno o più dei residui S430, D441, K455, G375, G376, E433 e T452. Preferibilmente, la mutazione Tsp per produrre una proteina avente attività proteasica ridotta può comprendere una mutazione, come una mutazione missenso in uno, due o tutti e tre i residui del sito attivo S430, D441 e K455.

Di conseguenza, il gene Tsp mutato può comprendere: una mutazione a S430; o una mutazione in D441; o una mutazione in K455; o una mutazione in S430 e D441; o una mutazione in S430 e K455; o una mutazione in

D441 e K455; o una mutazione in S430, D441 e K455.

Uno o più di S430, D441, K455, G375, G376, E433 e T452 possono essere mutati in qualsiasi amminoacido adatto che si traduca in una proteina con attività proteasica ridotta. Esempi di mutazioni idonee sono S430A, S430C, D441A, K455A, K455H, K455R, G375A, G376A, E433A e T452A. Il gene Tsp mutato può comprendere una, due o tre mutazioni ai residui del sito attivo, ad esempio il gene può comprendere: S430A o S430C; e/o D441A; e/o K455A o K455H o K455R.

Preferibilmente, il gene Tsp ha la mutazione puntiforme S430A o S430C.

L'espressione "gene Tsp mutato knockout" nel contesto della presente invenzione significa che il gene comprende una o più mutazioni che impediscono l'espressione della proteina Tsp codificata dal gene selvatico per fornire una cellula deficiente nella proteina Tsp. Il gene knockout può essere parzialmente o completamente trascritto ma non tradotto nella proteina codificata. Il gene Tsp knockout mutato può essere mutato in qualsiasi modo adatto, per esempio da una o più tra delezioni, inserzioni, mutazioni puntiformi, missenso, nonsense e di spostamento del modulo di lettura, a causare assenza di espressione della proteina. Ad esempio, il gene può essere knocked out mediante l'inserimento di una sequenza di DNA estraneo, come un marcatore di resistenza agli antibiotici, nella sequenza di codifica del gene.

In una forma di realizzazione preferita, il gene Tsp non viene mutato mediante l'inserimento di una sequenza di DNA estraneo, come un marcatore di resistenza agli antibiotici, nella sequenza di codifica del gene. In questa forma di realizzazione il gene Tsp può comprendere una mutazione al codone di inizio del gene e/o uno o più codoni di stop collocati a valle del codone di inizio del gene e a monte del codone di arresto del gene impedendo così l'espressione della proteina Tsp. La mutazione al codone di inizio può essere una mutazione missenso di uno, due o tutti e tre i nucleotidi del codone di inizio. In alternativa o in aggiunta, il codone di inizio può essere mutato mediante una mutazione di spostamento del modulo di lettura da inserzione o delezione. Il gene Tsp comprende due codoni ATG all'estremità 5 'della sequenza codificante, uno o entrambi i codoni ATG possono essere mutati da una mutazione missenso. Il gene Tsp può essere mutato al secondo codone ATG (codone 3) in TCG.

Il gene Tsp può alternativamente o ulteriormente comprendere uno o più codoni di stop collocati a valle del codone di inizio del gene e a monte del codone di arresto del gene. Preferibilmente il gene Tsp mutato knockout comprende sia una mutazione missenso al codone di inizio sia uno o più codoni di arresto inseriti. In una forma di realizzazione preferita il gene Tsp è mutato per eliminare "T" dal quinto codone provocando così uno spostamento del modulo di lettura che risulta in codoni di arresto nei codoni 11 e 16. In una forma di realizzazione preferita il gene Tsp è mutato ad inserire un sito di restrizione Ase I per creare un terzo codone di stop nel modulo di lettura, al codone 21.

In una forma di realizzazione preferita il gene Tsp mutato knockout ha la sequenza di DNA di SEQ ID NO: 3, che comprende i 6 nucleotidi ATGAAT a monte del codone d'inizio. In una forma di realizzazione il gene Tsp mutato ha la sequenza di DNA dei nucleotidi da 7 a 2048 di SEQ ID NO: 3.

Nella presente invenzione le cellule portano anche uno o più geni capaci di esprimere o sovraesprimere una o più proteine in grado di facilitare il ripiegamento proteico. Esempi includono proteine come FkpA, Skp, SurA, PPIA e PPIID.

In una forma di realizzazione la proteina per facilitare il ripiegamento delle proteine è FkpA, Skp o una loro combinazione.

In una forma di realizzazione la proteina per facilitare il ripiegamento delle proteine è selezionata tra FkpA o una combinazione di FkpA e Skp. FkpA è un peptidil-prolil cis-trans isomerasi con il numero Swiss-Prot P45523.

Skp è una proteina chaperone con il numero Swiss-Prot P0AEU7.

La proteina per facilitare il ripiegamento delle proteine può essere codificata da un gene nel genoma delle cellule o trasfettato transitoriamente in esso, per esempio in un plasmide o una combinazione degli stessi, a seconda dei casi.

La cellula batterica gram-negativa ricombinante dell'invenzione non comprende un polinucleotide ricombinante DsbC.

In una forma di realizzazione della presente invenzione la cellula batterica gram-negativa ricombinante

comprende inoltre un gene DegP mutato che codifica per una proteina DegP avente attività di chaperone e attività proteasica ridotta e/o un gene ptr mutato, in cui il gene mutato ptr codifica per una proteina Proteasi III avente ridotta attività proteasica o è un gene ptr mutante knockout e/o un gene OmpT mutato, in cui il gene OmpT mutato codifica una proteina OmpT con attività proteasica ridotta o un gene OmpT mutato knockout.

In questa forma di realizzazione il genoma della cellula è isogenico a una cellula batterica selvatica ad eccezione delle mutazioni di cui sopra.

Come qui usato, "DegP" significa un gene che codifica per la proteina DegP (anche noto come HtrA), che ha una doppia funzione come chaperone e proteasi (amilies of serine peptidases; Rawlings ND, Barrett AJ. *Methods Enzymol.* 1994; 244: 19- 61). La sequenza del gene DegP non mutato è mostrata in SEQ ID NO: 7 e la sequenza della proteina DegP non mutata è mostrata in SEQ ID NO: 8.

A basse temperature DegP funziona come una chaperone e ad alte temperature DegP preferisce funzionare come proteasi (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M) e The proteolytic activity of the HtrA (DegP) protein from Escherichia coli at low temperatures, Skorko-Glonek J et al *Microbiology* 2008, 154, 3649-3658).

Nelle forme di realizzazione in cui la cellula comprende la mutazione DegP la mutazione DegP nella cellula fornisce un gene DegP mutato che codifica per una proteina DegP avente attività di chaperone ma non attività proteasica completa.

L'espressione "avere attività di chaperone" nel contesto della presente invenzione significa che la proteina DegP mutata ha la stessa o sostanzialmente la stessa attività di chaperone rispetto alla proteina DegP selvatica non mutata. Preferibilmente, il gene DegP mutato codifica per una proteina DegP che ha il 50% o più, il 60% o più, il 70% o più, l'80% o più, il 90% o più o il 95% o più dell'attività chaperone di una proteina selvatica non mutata DegP. Più preferibilmente, il gene DegP mutato codifica per una proteina DegP avente la stessa attività di chaperone rispetto a DegP selvatico.

L'espressione "avendo ridotta attività proteasica" nel contesto della presente invenzione significa che la proteina

DegP mutata non ha l'attività proteasica completa rispetto alla proteina DegP selvatica non mutata. Preferibilmente, il gene DegP mutato codifica per una proteina DegP con il 50% o meno, il 40% o meno, il 30% o meno, il 20% o meno, il 10% o meno o il 5% o meno dell'attività proteasica di una proteina DegP selvatica non mutata. Più preferibilmente, il gene DegP mutato codifica per una proteina DegP senza attività proteasica. La cellula non è deficiente nel DegP cromosomico, cioè le sequenze del gene DegP non sono state eliminate o mutate per impedire l'espressione di qualsiasi forma di proteina Deg

Qualsiasi mutazione adatta può essere introdotta nel gene DegP al fine di produrre una proteina avente attività di chaperone e attività proteasica ridotta. L'attività di proteasi e chaperone di una proteina DegP espressa da un batterio gram-negativo può essere facilmente testata da una persona esperta nella tecnica da qualsiasi metodo adatto come il metodo descritto in Spiess et al. In cui sono state testate le attività di proteasi e chaperone di DegP su MalS, un substrato naturale di DegP (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M) anche il metodo descritto in The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures, Skorko-Glonek J et al *Microbiology* 2008, 154, 3649-3658.

DegP è una serina proteasi e ha un centro attivo costituito da una triade catalitica di residui amminoacidici di His105, Asp135 e Ser210 (Families of serine peptidases, *Methods Enzymol.*, 1994, 244: 19-61 Rawlings N e Barrett A). La mutazione DegP per produrre una proteina con attività chaperone e attività proteasica ridotta può comprendere una mutazione, come una mutazione missenso in uno, due o tre di His105, Asp135 e Ser210.

Di conseguenza, il gene DegP mutato può comprendere: una mutazione in His105; o una mutazione in Asp135; o una mutazione in Ser210; o una mutazione a His105 e Asp135; o una mutazione in His105 e Ser210; o una mutazione in Asp135 e Ser210; o una mutazione in His105, Asp135 e Ser210.

Uno, due o tre di His105, Asp135 e Ser210 possono essere mutati in qualsiasi amminoacido adatto che risulti in una proteina avente attività di chaperone e attività proteasica ridotta. Ad esempio, uno, due o tre tra His105, Asp135 e Ser210 possono essere mutati in un piccolo amminoacido come Gly o Ala. Un'ulteriore mutazione adatta è quella di cambiare uno, due o tre tra His105, Asp135 e Ser210 in un amminoacido avendo proprietà

opposte come mutare Asp135 in Lys o Arg, His105 polare in un amminoacido non polare come Gly, Ala, Val o Leu e mutare il piccolo Ser210 idrofilo in un residuo grande o idrofobo come Val, Leu, Phe o Tyr. Preferibilmente, il gene DegP comprende la mutazione puntiforme S210A, come mostrato in Figura 11c, che è stato trovato produrre una proteina avente attività di chaperone ma non attività proteasica (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M).

DegP ha due domini PDZ, PDZ1 (residui 260-358) e PDZ2 (residui 359-448), che mediano l'interazione proteina-proteina (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M). In una forma di realizzazione della presente invenzione il gene degP è mutato per eliminare il dominio PDZ1 e/o il dominio PDZ2. La delezione di PDZ1 e PDZ2 determina la perdita completa dell'attività proteasica della proteina DegP e diminuisce l'attività di chaperone rispetto alla proteina selvatica DegP mentre la delezione di PDZ1 o PDZ2 risulta in una attività proteasica al 5% e attività chaperone simile rispetto alla proteina DegP selvatica (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M).

Il gene DegP mutato può anche comprendere un sito di restrizione silente non presente in natura, come Ase I, per aiutare nei metodi di identificazione e screening, ad esempio come mostrato in Figura 7C.

La sequenza preferita del gene DegP mutato comprendente la mutazione puntiforme S210A e un sito marcatore di restrizione Ase I è fornita in SEQ ID NO: 9 e la sequenza proteica codificata è mostrata in SEQ ID NO: 10.

Nelle forme di realizzazione della presente invenzione in cui la cellula comprende un gene DegP mutato che codifica per una proteina DegP avente attività di chaperone e attività di proteasi ridotta, una o più delle cellule fornite dalla presente invenzione possono fornire una resa migliorata di proteine correttamente piegate dalla relativa cella a cellule mutate in cui il gene DegP è stato mutato in DegP knockout che impedisce l'espressione di DegP, come una deficienza cromosomica di DegP. In una cellula comprendente un gene DegP knockout mutato che impedisce l'espressione DegP, l'attività chaperone di DegP si perde completamente mentre nella cellula

secondo la presente invenzione l'attività di chaperone di DegP è mantenuta mentre l'attività proteasica completa è persa. In queste forme di realizzazione, una o più cellule secondo la presente invenzione hanno un'attività proteasica inferiore per impedire la proteolisi della proteina pur mantenendo l'attività di chaperone per consentire il corretto trasporto e ripiegamento della proteina nella cellula ospite.

In queste forme di realizzazione, una o più cellule secondo la presente invenzione possono avere una crescita cellulare migliorata rispetto alle cellule portatrici di un gene DegP knockout mutato che impedisce l'espressione di DegP. Senza voler essere vincolati dalla teoria, potrebbe essere mostrata una migliore crescita cellulare a causa della proteasi DegP che mantiene l'attività di un chaperone che può aumentare la capacità della cellula di processare tutte le proteine che richiedono l'attività di uno chaperone. Di conseguenza, la produzione di proteine correttamente piegate necessarie per la crescita e la riproduzione della cellula può essere aumentata in una o più delle cellule della presente invenzione rispetto alle cellule portatrici di una mutazione knockout DegP migliorando così le vie cellulari che regolano la crescita.

Inoltre, i ceppi noti deficienti in proteasi DegP sono generalmente sensibili alla temperatura e in genere non crescono a temperature superiori a circa 28 °C. Tuttavia, le cellule secondo la presente invenzione non sono sensibili alla temperatura e possono essere coltivate a temperature di 28° C o superiori, comprese temperature da circa 30° C a circa 37° C, che sono tipicamente utilizzate per la produzione su scala industriale di proteine da batteri.

In una forma di realizzazione della presente invenzione la cellula porta un gene ptr mutato. Come qui usato, "gene ptr" significa un gene che codifica per Proteasi III, una proteasi che degrada le proteine ad alto peso molecolare. La sequenza del gene ptr non mutato è mostrata in SEQ ID NO: 4 e la sequenza della proteina Proteasi III non mutata è mostrata in SEQ ID NO: 5.

Il riferimento al gene ptr mutato o del gene ptr mutato codificante Proteasi III, si riferisce a un gene ptr mutato che codifica per una proteina Proteasi III avente attività proteasica ridotta o un gene ptr knockout mutato, se non diversamente indicato.

L'espressione "gene mutato ptr che codifica per una proteina di Proteasi III con attività proteasica ridotta" nel

contesto della presente invenzione significa che il gene mutato di ptr non ha l'attività proteasica completa rispetto al gene ptr non mutato selvatico.

Preferibilmente, il gene ptr mutato codifica per una Proteasi III avente il 50% o meno, il 40% o meno, il 30% o meno, il 20% o meno, il 10% o meno o il 5% o meno dell'attività proteasica di una Proteina Proteasi III selvatica non mutata. Più preferibilmente, il gene ptr mutato codifica per una proteina Proteasi III senza attività proteasica. In questa forma di realizzazione la cellula non è deficiente in ptr cromosomico, cioè la sequenza del gene ptr non è stata cancellata o mutata per impedire l'espressione di alcuna forma di proteina Proteasi III.

Qualsiasi mutazione adatta può essere introdotta nel gene ptr al fine di produrre una proteina Proteasi III con attività proteasica ridotta. L'attività proteasica di una proteina Proteasi III espressa da un batterio gram-negativo può essere facilmente testata da una persona esperta del settore mediante qualsiasi metodo idoneo nella tecnica.

L'espressione "gene ptr knockout mutato" nel contesto della presente invenzione significa che il gene comprende una o più mutazioni, causando così assenza di espressione della proteina codificata dal gene per fornire una cellula carente nella proteina codificata dal gene mutato knockout. Il gene knockout può essere parzialmente o completamente trascritto ma non tradotto nella proteina codificata. Il gene ptr knockout mutato può essere mutato in qualsiasi modo adatto, per esempio mediante una o più delezioni, inserzioni, mutazioni puntiformi, missenso, nonsense e di slittamento del modulo di lettura, per ottenere l'assenza di espressione della proteina. Ad esempio, il gene può essere knocked out mediante l'inserimento di una sequenza di DNA estraneo, come un marcatore di resistenza agli antibiotici, nella sequenza codificante il gene. In una forma di realizzazione preferita, il gene non è mutato mediante l'inserimento di una sequenza di DNA estraneo, come un marcatore di resistenza agli antibiotici, nella sequenza codificante del gene. Preferibilmente il gene della Proteasi III comprende una mutazione nel codone di inizio del gene e/o uno o più codoni di stop collocati a valle del codone di inizio del gene e a monte del codone di arresto del gene impedendo in tal modo l'espressione della proteina Proteasi III. Una mutazione al codone di inizio del gene knockout bersaglio causa la perdita di funzione del codone di inizio e garantisce quindi che il gene bersaglio non comprenda un codone di inizio adatto all'inizio della sequenza

codificante. La mutazione al codone di inizio può essere una mutazione missenso di uno, due o tutti e tre i nucleotidi del codone di inizio. In alternativa o in aggiunta, il codone di inizio può essere mutato mediante una mutazione di spostamento del modulo di lettura da inserzione o delezione.

In una forma di realizzazione preferita il gene ptr è mutato per cambiare il codone di inizio ATG in ATT. Il gene ptr knockout mutato può comprendere alternativamente o in aggiunta uno o più codoni di stop collocati a valle del codone di inizio del gene e a monte del codone di stop del gene. Preferibilmente il gene ptr mutante knockout comprende sia una mutazione missenso nel codone di inizio sia uno o più codoni di stop inseriti. Gli uno o più codoni di stop inseriti sono preferibilmente codoni di stop nel modulo di lettura. Tuttavia, gli uno o più codoni di stop inseriti possono essere alternativamente o in aggiunta codoni di stop fuori dal modulo di lettura. Potrebbero essere necessari uno o più codoni di stop fuori dal modulo di lettura per interrompere la traduzione in cui un codone di inizio fuori modulo di lettura viene modificato in un codone di inizio in modulo di lettura mediante una mutazione di spostamento del modulo di lettura da inserzione o delezione. L'uno o più codoni di stop possono essere introdotti da qualsiasi mutazione idonea tra cui una mutazione puntiforme e una mutazione di spostamento del modulo di lettura. Gli uno o più codoni di stop sono preferibilmente introdotti mediante una mutazione di spostamento del modulo di lettura e/o una mutazione di inserzione, preferibilmente mediante sostituzione di un segmento della sequenza genica con una sequenza comprendente un codone di arresto. Ad esempio può essere inserito un sito di restrizione Ase I, che comprende il codone di stop TAA.

In una forma di realizzazione preferita, il gene ptr è mutato per inserire un codone di arresto in-modulo di lettura mediante l'inserimento di un sito di restrizione Ase I, come mostrato in Figura 7A. In una forma di realizzazione preferita, il gene ptr mutante knockout ha la sequenza di DNA di SEQ ID NO: 6.

Le suddette mutazioni knockout sono vantaggiose perché causano una minima o nessuna interruzione del DNA cromosomico a monte o a valle del sito del gene knockout target e non richiedono l'inserimento e il mantenimento di DNA estraneo, come marcatori di resistenza agli antibiotici, che possono influenzare l'idoneità della cellula ad esprimere una proteina di interesse, in particolare proteine terapeutiche. Di conseguenza, una o più delle cellule secondo la presente invenzione possono mostrare caratteristiche di crescita e/o espressione

della proteina migliorate rispetto alle cellule in cui il gene della proteasi è stato knocked out mediante l'inserimento di DNA estraneo nella sequenza di codifica del gene. In una forma di realizzazione le cellule secondo la presente invenzione portano un gene mutato OmpT. Come qui usato, "gene OmpT" significa un gene che codifica per proteasi OmpT (proteasi di membrana esterna T) che è una proteasi di membrana esterna. La sequenza del gene OMPT selvatico non mutato è SWISS-PROT P09169. Il riferimento a un gene OmpT mutato o gene mutato OmpT che codifica OmpT, si riferisce a un gene OmpT mutato che codifica per una proteina OmpT con attività proteasica ridotta o un gene OmpT mutato knockout, a meno che non sia diversamente indicato.

L'espressione "mutato OmpT gene codifica per una proteina OmpT avente ridotta attività proteasica" nel contesto della presente invenzione significa che il gene mutato OmpT non ha l'attività proteasica completa rispetto al gene non mutato selvatico OmpT. Il gene mutato OmpT può codificare una proteina OmpT con il 50% o meno, il 40% o meno, il 30% o meno, il 20% o meno, il 10% o meno o il 5% o meno dell'attività proteasica di una proteina OmpT selvatica non mutata. Il gene mutato OmpT può codificare una proteina OmpT senza attività proteasica. In questa forma di realizzazione la cellula non è deficiente in OmpT cromosomico, cioè la sequenza del gene OmpT non è stata cancellata o mutata per impedire l'espressione di alcuna forma di proteina OmpT. Qualsiasi mutazione adatta può essere introdotta nel gene OmpT al fine di produrre una proteina avente attività proteasica ridotta. L'attività proteasica di una proteina OmpT espressa da un batterio gram-negativo può essere facilmente testata da una persona esperta nella tecnica da qualsiasi metodo adatto nella tecnica, come il metodo descritto in Kramer et al. (Identification of essential acidic residues of outer membrane protease OmpT supports a novel active site, FEBS Letters 505 (2001) 426-430) e Dekker et al (Substrate Specificity of the Integral Membrane Protease OmpT Determined by Spatially Addressed Peptide Libraries, Biochemistry 2001, 40, 1694-1701).

E' stato riportato in Kramer et al (Identification of active site serine e histidine residues in Escherichia coli outer membrane protease OmpT FEBS Letters 2000 468, 220-224) che OmpT mostra che la sostituzione di serine, istidine e residui acidi con alanine risulta in attività ridotta di circa 10 volte per Glu27, Asp97, Asp208 o His101,

un'attività ridotta di circa 500 volte per Ser99 e ~ 10000 volte per Asp83, Asp85, Asp210 o His212. Vandeputte-Rutten et al (Crystal Structure of the Outer Membrane Protease OmpT from Escherichia coli suggests a novel catalytic site, The EMBO Journal 2001, Vol 20 No 18 5033-5039) come avente un sito attivo che comprende una coppia Asp83-Asp85 e una coppia His212-Asp210. Further Kramer et al (Lipopolysaccharide regions involved in the activation of Escherichia coli outer membrane protease OmpT, Eur. J. Biochem. FEBS 2002, 269, 1746-1752) descrive che le mutazioni D208A, D210A, H212A, H212N, H212Q, G216K / K217G, K217T e R218L nell'ansa L4 hanno dato tutte luogo a una perdita parziale o quasi totale dell'attività enzimatica.

Di conseguenza, la mutazione OmpT per produrre una proteina con attività proteasica ridotta può comprendere una mutazione, ad esempio una mutazione missenso in uno o più residui E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 & E250.

Uno o più tra E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 ed E250 possono essere mutati in qualsiasi amminoacido adatto che si traduca in una proteina con ridotta attività proteasica. Ad esempio, uno o più tra E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 ed E250 possono essere mutati in alanina. Esempi di mutazioni idonee sono E27A, D43A, D83A, D85A, D97A, S99A, H101A E111A, E136A, E193A, D206A, D208A, D210A, H212A, H212N, H212Q, G216K, K217G, K217T, R218L e E250A. In una forma di realizzazione il gene mutato OmpT comprende le mutazioni di D210A e H212A. Una sequenza OmpT mutata adatta comprendente le mutazioni di D210A e H212A è mostrata in SEQ ID NO: 23.

L'espressione "gene OmpT mutato knockout" nel contesto della presente invenzione significa che il gene comprende una o più mutazioni, che causano l'assenza di espressione della proteina codificata dal gene per fornire una cellula deficiente nella proteina codificata dal gene mutato knockout. Il gene knockout può essere parzialmente o completamente trascritto ma non tradotto nella proteina codificata. Il gene OMPT mutato può essere mutato in qualsiasi modo adatto, per esempio da una o più delezioni, inserzioni, mutazioni puntiformi, missenso, nonsense e di spostamento del modulo di scrittura, causando l'assenza di espressione della proteina. Ad esempio, il gene può essere knocked out mediante l'inserimento di una sequenza di DNA estraneo, come un

marcatore di resistenza agli antibiotici, nella sequenza di codifica del gene.

In una forma di realizzazione il gene OmpT comprende una mutazione al codone di inizio del gene e/o uno o più codoni di stop collocati a valle del codone di inizio del gene e a monte del codone di stop del gene impedendo in tal modo l'espressione della proteina OmpT. La mutazione al codone di inizio può essere una mutazione missenso di uno, due o tutti e tre i nucleotidi del codone di inizio. Una sequenza OmpT knockout mutata adatta è mostrata in SEQ ID NO: 24. In alternativa o in aggiunta, il codone di inizio può essere mutato mediante una mutazione di spostamento del modulo di lettura da inserzione o delezione.

In una forma di realizzazione, la cellula batterica gram-negativa secondo la presente invenzione non ha un gene ompT knockout mutato, come ad esempio essendo deficiente nell'ompT cromosomico.

In una forma di realizzazione la cellula batterica gram-negativa secondo la presente invenzione non ha un degP knockout mutato, ad esempio essendo deficiente nel degP cromosomico. In una forma di realizzazione la cellula batterica gram-negativa secondo la presente invenzione non ha un gene degP mutato.

In una forma di realizzazione la cellula batterica gram-negativa secondo la presente invenzione non ha un gene ptr knockout mutato, come ad esempio essendo deficiente nel ptr cromosomico.

Molte mutazioni geneticamente ingegnerizzate incluse le mutazioni knockout implicano l'uso di marcatori di resistenza antibiotica che consentono la selezione e l'identificazione di cellule mutate con successo. Tuttavia, come discusso sopra, ci sono una serie di svantaggi nell'uso di marcatori di resistenza agli antibiotici. Un'ulteriore forma di realizzazione della presente invenzione la cellula non comprende un marcatore di resistenza agli antibiotici e supera gli svantaggi sopra menzionati dell'uso di marcatori di resistenza antibiotica in cui il gene Tsp mutato, il gene spr mutato e facoltativamente il gene DegP mutato e/o un gene mutato ptr e/o un gene OmpT mutato, sono mutati in modo da comprendere uno o più siti di marcatori di restrizione.

I siti di restrizione sono geneticamente ingegnerizzati nel gene e non sono presenti in natura. I siti dei marcatori di restrizione sono vantaggiosi perché consentono lo screening e l'identificazione di cellule correttamente modificate che comprendono le mutazioni cromosomiche richieste.

Le cellule che sono state modificate per portare uno o più geni mutanti della proteasi possono essere analizzate

mediante PCR di DNA genomico da lisati di cellule usando coppie di oligonucleotidi progettate per amplificare una regione del DNA genomico comprendente un sito marcatore di restrizione non presente in natura.

Il DNA amplificato può quindi essere analizzato mediante elettroforesi su gel di agarosio prima e dopo l'incubazione con un enzima di restrizione adatto in grado di digerire il DNA nel sito del marcatore di restrizione non presente in natura. La presenza di frammenti di DNA dopo l'incubazione con l'enzima di restrizione conferma che le cellule sono state modificate con successo per portare l'uno o più geni mutati.

Nella forma di realizzazione in cui la cella porta un gene ptr mutante knockout avente la sequenza di DNA di SEQ ID NO: 6, le sequenze di primer oligonucleotidici mostrate in SEQ ID NO: 17 e SEQ ID NO: 18 possono essere usate per amplificare la regione del DNA comprendente il sito di restrizione Ase I non presente in natura dal DNA genomico delle cellule trasformate. Il DNA genomico amplificato può quindi essere incubato con l'enzima di restrizione Ase I e analizzato mediante elettroforesi su gel per confermare la presenza del gene ptr mutato nel DNA genomico.

Nella forma di realizzazione in cui la cella comprende un gene Tsp mutato knockout avente la sequenza di DNA di SEQ ID NO: 3 o nucleotidi da 7 a 2048 di SEQ ID NO: 3, le sequenze di primer oligonucleotidici mostrate in SEQ ID NO: 15 e SEQ ID NO: 16 possono essere usate per amplificare la regione del DNA comprendente il sito di restrizione Ase I non presente in natura dal DNA genomico delle cellule trasformate. Il DNA genomico amplificato può quindi essere incubato con l'enzima di restrizione Ase I e analizzato mediante elettroforesi su gel per confermare la presenza del gene Tsp mutato nel DNA genomico.

Nella forma di realizzazione in cui la cella comprende un gene DegP mutato avente la sequenza di DNA di SEQ ID NO: 9, le sequenze di primer oligonucleotidici mostrate in SEQ ID NO: 19 e SEQ ID NO: 20 possono essere usate per amplificare la regione del DNA che comprende il sito di restrizione Ase I non presente in natura dal DNA genomico delle cellule trasformate. Il DNA genomico amplificato può quindi essere incubato con l'enzima di restrizione Ase I e analizzato mediante elettroforesi su gel per confermare la presenza del gene DegP mutato nel DNA genomico.

Gli uno o più siti di restrizione possono essere introdotti da qualsiasi mutazione adatta includendo una o più

mutazioni di delezione, inserzioni, mutazioni puntiformi, missenso, nonsense e di spostamento del modulo di lettura. Un sito di restrizione può essere introdotto dalla mutazione del codone di inizio e/o per introdurre uno o più codoni di stop, come descritto sopra. Questa forma di realizzazione è vantaggiosa perché il sito del marcatore di restrizione è un marcatore diretto e unico delle mutazioni knockout introdotte.

Può essere inserito un sito di restrizione marcatore che comprende un codone di arresto nel modulo di lettura, come un sito di restrizione Ase I. Ciò è particolarmente vantaggioso poiché il sito di restrizione inserito agisce sia da sito di restrizione marcatore che da codone di stop per impedire la completa trascrizione della sequenza di codifica del gene. Ad esempio, nella forma di realizzazione in cui si introduce un codone di stop nel gene ptr mediante introduzione di un sito Ase I, questo crea anche un sito di restrizione, come mostrato in Figura 7A. Ad esempio, nella forma di realizzazione in cui è introdotto un codone di arresto nel gene Tsp nel codone 21 mediante l'introduzione di un sito Ase I, questo crea anche un sito di restrizione, come mostrato in Figura 7B.

Un sito di restrizione marcatore può essere inserito dalla mutazione del codone di inizio e opzionalmente una o più ulteriori mutazioni puntiformi. In questa forma di realizzazione il sito di restrizione marcatore è preferibilmente un sito di restrizione EcoRI. Questo è particolarmente vantaggioso perché la mutazione del codone di inizio crea anche un sito di restrizione marcatore. Ad esempio, nella forma di realizzazione in cui il codone di inizio del gene ptr è cambiato in ATT, questo crea un sito marcatore Eco RI, come mostrato in Figura 11a. Ad esempio, nella forma di realizzazione in cui il codone di inizio (codone 3) del gene Tsp è cambiato da ATG a TCG, come mostrato in Figura 1b, un'ulteriore mutazione puntiforme del codone 2 da AAC a AAT e una mutazione del codone 3 da ATG a TCG crea un sito marcatore di restrizione EcoRI, come mostrato in Figura 7B.

Nella forma di realizzazione della presente invenzione in cui la cellula ha un gene OmpT mutato, gli uno o più siti di restrizione possono essere introdotti da qualsiasi mutazione adatta incluse una o più delezioni, inserzioni, mutazioni puntiformi, missenso, nonsense e di spostamento del modulo di lettura. Ad esempio, nella forma di realizzazione in cui il gene OmpT comprende le mutazioni D210A e H212A, queste mutazioni introducono un sito di restrizione silente HindIII che può essere utilizzato come marcatore di selezione.

Nel gene DegP del gene spr, un sito di restrizione marcatore può essere introdotto usando cambiamenti silenti

nei codoni. Per esempio, un sito di Ase I può essere usato come sito di restrizione marcatore silente, in cui il codone di stop TAA è fuori modulo di lettura, come mostrato in Figura 7C per il gene DegP mutato. Nelle forme di realizzazione della presente invenzione, in cui il gene ptr e/o il gene Tsp sono mutati per codificare un Proteasi III o Tsp con attività proteasica ridotta, possono essere introdotti uno o più siti di restrizione marcatori usando cambiamenti silenti nei codoni.

La cellula batterica gram-negativa ricombinante secondo la presente invenzione può essere prodotta con qualsiasi mezzo adatto. La persona esperta conosce tecniche idonee che possono essere utilizzate per sostituire una sequenza genica cromosomica con una sequenza genica mutata. Possono essere impiegati vettori adatti che consentono l'integrazione nel cromosoma ospite mediante ricombinazione omologa.

Ad esempio, sono descritti metodi di sostituzione genica adatti in Hamilton *et al* (New Method for Generating Deletions e Gene Replacements in Escherichia coli, Hamilton C. M. et al., Journal of Bacteriology Sept. 1989, Vol. 171, No. 9 p 4617-4622), Skorupski *et al* (Positive selection vectors for allelic exchange, Skorupski K e Taylor R. K., Gene, 1996, 169, 47-52), Kiel *et al* (A general method for the construction of Escherichia coli mutants by homologous recombination e plasmid segregation, Kiel J.A.K.W. et al, Mol Gen Genet 1987, 207:294-301), Blomfield *et al* (Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene e a temperature sensitive pSC101 replicon, Blomfield I. C. et al., Molecular Microbiology 1991, 5(6), 1447-1457) e Ried *et al*. (An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in Gram-negative bacteria by marker exchange-eviction mutagenesis, Ried J. L. e Collmer A., Gene 57 (1987) 239-246). Un plasmide adatto che consente la ricombinazione/sostituzione omologa è il plasmide pKO3 (Link et al., 1997, Journal of Bacteriology, 179, 6228-6237).

I ceppi mutati con successo possono essere identificati usando metodi ben noti nello stato della tecnica compreso il sequenziamento del DNA della colonia per PCR e mappatura di restrizione per PCR della colonia.

Nella forma di realizzazione in cui la cellula comprende due o più geni cromosomici mutati, i geni mutati possono essere introdotti nel batterio gram-negativo sugli stessi o su diversi vettori.

In una forma di realizzazione, la cellula batterica gram-negativa secondo la presente invenzione non ha un gene

ompT knockout mutato, come ad esempio essendo deficiente nell'omT cromosomico.

In una forma di realizzazione le cellule secondo la presente descrizione contengono solo le tre mutazioni caratterizzanti dello spr mutato, Tsp e FkpA ridotto, Skp o una combinazione di FkpA e Skp.

Una forma di realizzazione la linea cellulare secondo la presente descrizione è costituita da un gene Tsp mutato e un gene spr mutato, e un gene FkpA.

In una forma di realizzazione la linea cellulare secondo la presente descrizione è costituita da un gene Tsp mutato e un gene spr mutato, e un gene Skp.

In una forma di realizzazione la linea cellulare secondo la presente descrizione è costituita da un gene Tsp mutato e un gene spr mutato, un gene FkpA e un gene Skp.

In una forma di realizzazione il gene dei geni in grado di esprimere o sovraesprimere una o più proteine in grado di facilitare il ripiegamento delle proteine, come FkpA, e Skp, viene trasformato in modo transiente nella cellula, per esempio in un vettore di espressione che comprende facoltativamente una sequenza di polinucleotidi codificante un anticorpo o un suo frammento legante.

In una forma di realizzazione la sequenza polinucleotidica codificante l'anticorpo e il polinucleotide codificante FkpA e/o Skp sono inserite in vettori di espressione separati.

Per la produzione di prodotti comprendenti catene sia pesanti che leggere, la linea cellulare può essere trasformata con due vettori, un primo vettore che codifica un polipeptide a catena leggera e un secondo vettore che codifica un polipeptide a catena pesante. In alternativa, può essere usato un vettore singolo, il vettore includendo sequenze che codificano polipeptidi a catena leggera e a catena pesante.

In alternativa, la sequenza polinucleotidica codificante per l'anticorpo e il polinucleotide codificante FkpA e/o Skp è inserita in un vettore. Preferibilmente il vettore comprende le sequenze che codificano i polipeptidi a catena pesante e leggera dell'anticorpo.

La presente invenzione fornisce anche un vettore di espressione comprendente un polinucleotide ricombinante codificante FkpA e/o Skp e un anticorpo o un suo frammento legante l'antigene specifico per FcRn umano. Il vettore di espressione è un vettore multi-cistronico che comprende la sequenza polinucleotidica codificante

FkpA e/o Skp e la sequenza polinucleotidica codificante l'anticorpo.

Il vettore multicistronico può essere prodotto mediante un metodo di clonazione vantaggioso che consente la clonazione sequenziale ripetuta di sequenze di polinucleotidi in un vettore. Il metodo utilizza estremità coesive compatibili di una coppia di siti di restrizione, come le estremità "AT" dei siti di restrizione Ase I e Nde I. Una sequenza polinucleotidica comprendente una sequenza codificante e avente estremità coesive compatibili, come un frammento AseI-NdeI, può essere clonata in un sito di restrizione nel vettore, come Nde I. L'inserimento della sequenza polinucleotidica distrugge il sito di restrizione 5' ma crea un nuovo sito di restrizione 3', come NdeI, che può quindi essere usato per inserire un'ulteriore sequenza polinucleotidica comprendente estremità coesive compatibili. Il processo può quindi essere ripetuto per inserire ulteriori sequenze. Ciascuna sequenza polinucleotidica inserita nel vettore comprende una sequenza non codificante 3' al codone di arresto che può comprendere un sito Ssp I per lo screening, una sequenza di legame del ribosoma Shine Dalgarno, un distanziatore ricco in A e un sito NdeI che codifica un codone di inizio.

Una rappresentazione schematica della creazione di un vettore comprendente una sequenza polinucleotidica che codifica una catena leggera di un anticorpo (LC), una catena pesante di un anticorpo (HC), una sequenza di polinucleotidi FkpA e un'ulteriore sequenza polinucleotidica è mostrata in Figura 7d.

La cella secondo la presente invenzione preferibilmente comprende un vettore di espressione come sopra definito. È qui ulteriormente divulgato un esempio in cui la cellula esprime anche una o più ulteriori proteine come segue: una o più proteine in grado di facilitare il ripiegamento di proteine, come, skp, SurA, PPIA e PPIB; e, opzionalmente una o più proteine in grado di facilitare la secrezione proteica o la traslocazione, come SecY, SecE, SecG, SecYEG, SecA, SecB, FtsY e Lep; l'una o più ulteriori proteine possono essere espresse da uno o più polinucleotidi inseriti nello stesso vettore della sequenza polinucleotidica codificante FkpA e/o della sequenza polinucleotidica codificante l'anticorpo. In alternativa gli uno o più polinucleotidi possono essere inseriti in vettori separati.

Il vettore di espressione può essere prodotto inserendo una o più cassette di espressione come definite sopra in un vettore adatto. Alternativamente, le sequenze di regolazione dell'espressione per dirigere l'espressione della

sequenza polinucleotidica possono essere contenute nel vettore di espressione e quindi può essere richiesta solo la regione codificante del polinucleotide per completare il vettore di espressione.

Il polinucleotide codificante FkpA e/o Skp e/o il polinucleotide codificante l'anticorpo è opportunamente inserito in un vettore replicabile, tipicamente un vettore di espressione a replicazione autonoma, per l'espressione nella cellula sotto il controllo di un opportuno promotore per la cellula. Molti vettori sono noti nella tecnica per questo scopo e la scelta del vettore appropriato può dipendere dalla dimensione dell'acido nucleico e dal tipo di cellula in particolare.

Esempi di vettori di espressione che possono essere impiegati per trasformare la cellula ospite con un polinucleotide secondo l'invenzione includono:

- un plasmide, come pBR322 o pACYC184, e/o
- un vettore virale come il fago batterico
- un elemento genetico trasponibile come un trasposone

Tali vettori di espressione di solito comprendono un'origine di replicazione del DNA plasmidica, un marcatore antibiotico di selezionabile, un promotore e terminatore trascrizionale separati da un sito di multi-clonazione (cassetta di espressione) e una sequenza di DNA codificante per un sito di legame ribosomiale.

I promotori impiegati nella presente invenzione possono essere collegati al relativo polinucleotide direttamente o in alternativa possono essere collocati in una posizione appropriata, ad esempio in un vettore tale che, quando il relativo polipeptide è inserito, il relativo promotore può agire sullo stesso. In una forma di realizzazione il promotore si trova prima della porzione codificante del polinucleotide su cui agisce, ad esempio un promotore attinente prima di ciascuna porzione codificante del polinucleotide. "Prima" come qui usato è inteso implicare che il promotore si trova al termine 5 primo rispetto alla porzione codificante del polinucleotide.

I promotori possono essere endogeni o esogeni alle cellule ospiti. Promotori adatti includono lac, tac, trp, phoA, Ipp, arabo, tet e T7.

Uno o più promotori impiegati possono essere promotori inducibili. Nella forma di realizzazione in cui il polinucleotide codificante FkpA e/o Skp e il polinucleotide codificante l'anticorpo sono inseriti in un vettore, le

sequenze nucleotidiche codificanti FkpA e l'anticorpo possono essere sotto il controllo di un singolo promotore o promotori separati. Nella forma di realizzazione in cui le sequenze nucleotidiche codificanti FkpA e/o Skp e l'anticorpo sono sotto il controllo di promotori separati, il promotore può essere promotori inducibili indipendentemente.

I promotori per l'uso in sistemi batterici contengono anche generalmente una sequenza Shine-Dalgarno (S.D.) operativamente legata al DNA che codifica il polipeptide di interesse. Il promotore può essere rimosso dal DNA di origine batterica mediante digestione con enzimi di restrizione e inserito nel vettore contenente il DNA desiderato.

Il vettore di espressione preferibilmente comprende anche un messaggio dicistronico per produrre l'anticorpo o il suo frammento di legame all'antigene come descritto in WO03/048208 o WO2007/039714 (il cui contenuto è qui incorporato per riferimento). Preferibilmente il cistrone a monte contiene il DNA codificante la catena leggera dell'anticorpo e il cistrone a valle contiene il DNA codificante la corrispondente catena pesante, e la sequenza intergenica dicistronica (IGS) preferibilmente comprende una sequenza selezionata tra IGS1 (SEQ ID NO: 23), IGS2 (SEQ ID NO: 24), IGS3 (SEQ ID NO: 25) e IGS4 (SEQ ID NO: 26).

I terminatori possono essere endogeni o esogeni alle cellule ospiti. Un terminatore adatto è rrnB. Ulteriori regolatori trascrizionali adatti inclusi promotori e terminatori e metodi di targeting proteico si possono trovare in "Strategies for Achieving High-Level Expression of Genes in Escherichia coli" Savvas C. Makrides, Microbiological Reviews, Sept 1996, p 512-538.

Il polinucleotide FkpA inserito nel vettore di espressione comprende preferibilmente l'acido nucleico che codifica le sequenze segnale di FkpA e la sequenza di codifica FkpA. Il vettore contiene preferibilmente una sequenza di acido nucleico che consente al vettore di replicarsi in una o più cellule ospite selezionate, preferibilmente di replicarsi indipendentemente dal cromosoma ospite. Tali sequenze sono ben note per una varietà di batteri.

In una forma di realizzazione la FkpA e/o Skp e/o la proteina di interesse comprende un tag di istidina al N-terminale e/o al C-terminale.

La molecola di anticorpo può essere secreta dalla cellula o indirizzata al periplasma mediante sequenze segnale idonee. Alternativamente, le molecole di anticorpo possono accumularsi nel citoplasma della cellula. Preferibilmente la molecola di anticorpo è indirizzata al periplasma.

Il polinucleotide codificante l'anticorpo può essere espresso come una fusione con un altro polipeptide, preferibilmente una sequenza segnale o un altro polipeptide avente un sito di taglio specifico all'N-terminale del polipeptide maturo. La sequenza segnale eterologa selezionata dovrebbe essere una che è riconosciuta ed elaborata dalla cellula ospite. Per le cellule ospite procariotiche che non riconoscono e processano la sequenza segnale del polipeptide eucariotico o nativo, la sequenza del segnale è sostituita con una sequenza segnale procariotica. Sequenze segnale adatte includono OmpA, PhoA, LamB, PelB, DsbA e DsbC. In una forma di realizzazione in cui la cellula comprende una sequenza polinucleotidica codificante una catena pesante dell'anticorpo e una sequenza polinucleotidica codificante una catena leggera dell'anticorpo, ciascun polinucleotide può comprendere una sequenza segnale, come OmpA.

La costruzione di vettori adatti contenenti uno o più dei componenti sopra elencati impiega tecniche di ligazione standard. Plasmidi isolati o frammenti di DNA sono tagliati, adattati, e ri-ligati nella forma desiderata per generare i plasmidi richiesti. Metodi generali mediante i quali si possono costruire i vettori, metodi di trasfezione e metodi di coltura sono ben noti agli esperti del settore. A tale riguardo, si fa riferimento a "Current Protocols in Molecular Biology", 1999, F. M. Ausubel (ed), Wiley Interscience, New York e the Maniatis Manual prodotto da Cold Spring Harbor Publishing.

Possono essere utilizzate per preparare sequenze di DNA che codificano per l'anticorpo tecniche standard di biologia molecolare. Le sequenze di DNA desiderate possono essere sintetizzate completamente o in parte usando tecniche di sintesi di oligonucleotidi. Se opportuno si possono utilizzare tecniche di mutagenesi sito-diretta e di reazione di polimerasi a catena (PCR).

Forme di attuazione dell'invenzione qui descritte con riferimento al polinucleotide si applicano ugualmente a forme di realizzazione alternative dell'invenzione, ad esempio possono essere applicati ugualmente vettori, cassette di espressione e/o cellule ospite comprendenti i componenti ivi impiegati, per quanto riguarda l'aspetto

rilevante.

La cella secondo la presente invenzione può inoltre comprendere una sequenza polinucleotidica codificante una proteina di interesse. La sequenza polinucleotidica codificante la proteina di interesse può essere esogena o endogena. La sequenza polinucleotidica codificante la proteina di interesse può essere integrata nel cromosoma dell'ospite o può essere non-integrata in un vettore, tipicamente un plasmide. In una forma di realizzazione la cellula secondo la presente invenzione esprime una proteina di interesse. "Proteina di interesse" nel contesto della presente descrizione intende riferirsi al polipeptide per l'espressione, solitamente un polipeptide ricombinante. Tuttavia, la proteina di interesse può essere una proteina endogena espressa da un gene endogeno nella cellula ospite.

Come qui usato, un "polipeptide ricombinante" si riferisce a una proteina che è costruita o prodotta usando la tecnologia del DNA ricombinante. La proteina di interesse può essere una sequenza esogena identica a una proteina endogena o una sua versione mutata, ad esempio con attività biologica attenuata, o suo frammento, espressa da un vettore esogeno. In alternativa, la proteina di interesse può essere una proteina eterologa, normalmente non espressa dalla cellula ospite. La proteina di interesse può essere qualsiasi proteina adatta compresa una proteina terapeutica, profilattica o diagnostica.

La persona esperta sarebbe facilmente in grado di testare la proteina secreta per vedere se la proteina è correttamente piegata usando metodi ben noti nella tecnica, come G HPLC proteico, dicromismo circolare, NMR, i metodi di cristallografia a raggi X e di misurazione dell'affinità all'epitopo. In una forma di realizzazione la proteina di interesse è utile nel trattamento di malattie o disturbi tra cui disturbi immunologici e/o malattie autoimmuni.

In una forma di realizzazione la malattia autoimmune è selezionata da encefalomielite acuta disseminata (ADEM), leucoencefalite emorragica necrotizzante acuta, morbo di Addison, agammaglobulinemia, alopecia areata, amiloidosi, vasculite ANCA-associata, spondilite anchilosante, nefrite anti-GBM/anti-TBM, sindrome antifosfolipide (APS), angioedema autoimmune, anemia aplastica autoimmune, disautonomia autoimmune,

epatite autoimmune, iperlipidemia autoimmune, immunodeficienza autoimmune, malattia autoimmune dell'orecchio interno (AIED), miocardite autoimmune, pancreatite autoimmune, retinopatia autoimmune, porpora trombocitopenica autoimmune (ATP), malattia tiroidea autoimmune, orticaria autoimmune, neuropatie assonali e neuronali, malattia di Balo, malattia di Behçet, pemfigoide bolloso, cardiomiopatia, malattia di Castleman, celiachia, malattia di Chagas, sindrome da stanchezza cronica, polineuropatia demielinizzante infiammatoria cronica (CIDP), ostomelite multifocale ricorrente cronica (CRMO), sindrome Churg-Strauss, pemfigo cicatriziale/pemfigoide mucoso benigno, morbo di Crohn, sindrome di Cogans, malattia da agglutinina fredda, blocco cardiaco congenito, miocardite di coxsackie, malattia CREST, crioglobulinemia mista essenziale, neuropatie demielinizzanti, dermatite erpetiforme, dermatomiosite, morbo di Devic (neuromielite ottica), cardiomiopatia dilatativa, lupus discoidale, sindrome di Dressler, endometriosi, fibrosi angio-centrica eosinofila, fascite eosinofila, eritema nodoso, encefalomyelite allergica sperimentale, sindrome di Evans, fibromialgia, alveolite fibrosante, arterite a cellule giganti (arterite temporale), glomerulonefrite, sindrome di Goodpasture, granulomatosi con poliangioite (GPA) vedere sindrome di Wegener, malattia di Graves, sindrome di Guillain-Barre, encefalite di Hashimoto, tiroidite di Hashimoto, anemia emolitica, porpora di Henoch-Schonlein, herpes gestationis, ipogammaglobulinemia, nefropatia idiopatica ipocomplementemica tubulointestinale, porpora trombocitopenica idiopatica (ITP), nefropatia a IgA, malattia correlata a IgG4, malattia sclerosante correlata a IgG4, lipoproteine immunoregolatrici, aneurisma infiammatorio aortico, pseudotumore infiammatorio, miosite con corpi inclusi, diabete insulino-dipendente (tipo 1), cistite interstiziale, artrite giovanile, diabete giovanile, sindrome di Kawasaki, tumore di Kuttner, Sindrome di Lambert-Eaton, vasculite leucocitoclastica, lichen planus, lichen sclerosus, congiuntivite lignea, malattia lineare da IgA (LAD), Lupus (LES), malattia di Lyme, fibrosi cronica mediastinica, malattia di Meniere, poliangite microscopica, sindrome di Mikulicz, malattia del tessuto connettivo misto (MCTD), ulcera di Mooren, malattia di Mucha-Habermann, fibrosclerosi multifocale, sclerosi multipla, miastenia grave, miosite, narcolessia, neuromielite ottica (Devic's), neutropenia, pemfigoide cicatriziale oculare, neurite ottica, malattia di Ormond (fibrosi retroperitoneale), reumatismo palindromico, PANDAS (Disturbi neuropsichiatrici autoimmuni pediatrici associati a Streptococcus), degenerazione cerebellare

paraneoplastica, polineuropatie paraproteinemiche, emoglobinuria parossistica notturna (PNH), sindrome di Parry Romberg, sindrome di Parsonnage-Turner, planite di pars (uveite periferica), pemfigo volgare, periaortite, periarterite, neuropatia periferica, encefalomielite perivenosa, anemia perniciosa, sindrome di POEMS, poliartrite nodosa, sindromi polighiandolari autoimmuni di tipo I, II e III, polimialgia reumatica, polimiosite, sindrome da infarto post-miocardico, sindrome postpericardiotomica, dermatite da progesterone, cirrosi biliare primitiva, colangite sclerosante primitiva, psoriasi, artrite psoriasica, fibrosi polmonare idiopatica, pioderma gangrenoso, aplasia pura delle cellule rosse, fenomeno di Raynauds, distrofia simpatica riflessa, sindrome di Reiter, policondrite recidivante, sindrome delle gambe senza riposo, fibrosi retroperitoneale (malattia di Ormond), febbre reumatica, artrite reumatoide, tiroidite di Riedel, sarcoidosi, sindrome di Schmidt, sclerite, scleroderma, sindromedi Sjogrens, autoimmunità spermatica e testicolare, sindrome della persona rigida, endocardite batterica subacuta (SBE), sindrome di Susac, oftalmia simpatica, arterite di Takayasu, arterite temporale/arterite a cellule giganti, trombotica, porpora trombocitopenica (TTP), sindrome di Tolosa-Hunt, mielite trasversa, colite ulcerativa, malattia del tessuto connettivo indifferenziato (UCTD), uveite, vasculite, dermatosi vescicolobulosa, vitiligine, macroglobulinemia di Waldenstrom, anemia emolitica tiepida idiopatica e granulomatosi di Wegener (ora chiamata Granulomatosi con poliangite (GPA).

In una forma di realizzazione il disturbo neurologico è selezionato da polineuropatia demielinizzante infiammatoria cronica (CIDP), sindrome di Guillain-Barre, polineuropatie paraproteinemica, neuromielite ottica (NMO, disturbi dello spettro NMO o malattie dello spettro NMO), e miastenia grave.

In una forma di realizzazione il disturbo ematologico immunologico è selezionato tra porpora trombocitopenica idiopatica (ITP), porpora trombotica trombocitopenica (TTP), anemia emolitica tiepida idiopatica, sindrome di Goodpasture, e mancata corrispondenza tra donatore e trapianto a causa di anticorpi anti-HLA.

In una forma di elaborazione la malattia è selezionata tra miastenia grave, neuro-mielite ottica, CIDP, Sindrome di Guillaume-Barre, Poli Neuropatia paraproteinemica, Epilessia refrattaria, ITP/TTP, anemia emolitica, Sindrome di Goodpasture, disfunzione ABO, nefrite da Lupus, vasculite renale, scleroderma, alveolite fibrosante, cardiomiopatia dilatativa, malattia di Grave, diabete di tipo 1, diabete autoimmune, pemfigo,

scleroderma, lupus, vasculite ANCA, dermatomiosite, malattia di Sjogren e artrite reumatoide.

In una forma di realizzazione il disturbo dermatologico è selezionato tra pemfigoide bolloso, pemfigo volgare, vasculite associata a ANCA e cardiomiopatia dilatativa.

In una forma di realizzazione gli anticorpi o frammenti secondo la presente descrizione possono essere impiegati nella profilassi o nel trattamento di malattie alloimmuni associate a trapianto di organi o di tessuti allogenici o in certe condizioni neonatali.

La proteina può essere un polipeptide proteoliticamente sensibile, cioè proteine che sono inclini ad essere tagliate, suscettibili di taglio, o tagliate da uno o più batteri gram-negativi, come *E. coli*, proteasi, sia nello stato nativo che durante la secrezione. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile a una proteasi selezionata tra DegP, Proteasi III e Tsp. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile alla proteasi Tsp. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile alle proteasi DegP e Proteasi III. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile alle proteasi DegP e Tsp. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile alle proteasi Tsp e Proteasi III. In una forma di realizzazione la proteina di interesse è proteoliticamente sensibile alle proteasi DegP, Proteasi III e Tsp. Preferibilmente la proteina è un polipeptide eucariotico. a proteina di interesse espressa dalle cellule secondo l'invenzione può, per esempio essere un immunogeno, una proteina di fusione comprendente due proteine eterologhe o un anticorpo. Gli anticorpi da utilizzare come proteine di interesse includono anticorpi monoclonali, multivalenti, multi-specifici, umanizzati, completamente umani o chimerici. L'anticorpo può essere di qualsiasi specie, ma è preferibilmente derivato da un anticorpo monoclonale, un anticorpo umano o un frammento umanizzato. L'anticorpo può essere derivato da qualsiasi classe (ad esempio IgG, IgE, IgM, IgD o IgA) o sottoclasse di molecole di immunoglobuline e può essere ottenuto da qualsiasi specie tra cui topo, ratto, squalo, coniglio, maiale, criceto, cammello, lama, capra o uomo. Parti del frammento di anticorpo possono essere ottenute da più di una specie, ad esempio i frammenti di anticorpi possono essere chimerici. In un esempio le regioni costanti sono di una specie e le regioni variabili di

un'altra.

L'anticorpo può essere una molecola anticorpale completa avente catene a lunghezza totale pesanti e leggere o un loro frammento, ad esempio VH, VL, VHH, Fab, Fab modificato, Fab' , F (ab')₂, Fv, frammento scFv, o un anticorpo a doppia specificità, come Fab-dAb o Fab-Fv, come descritto in WO2009/040562 e WO2010/035012.

In una forma di realizzazione la proteina è un Fab'.

L'anticorpo può essere specifico per qualsiasi antigene bersaglio. L'antigene può essere una proteina associata alla cellula, ad esempio una proteina di superficie cellulare su cellule come cellule batteriche, cellule di lievito, cellule T, cellule endoteliali o cellule tumorali, o può essere una proteina solubile. Antigeni di interesse possono essere anche qualsiasi proteina clinicamente rilevante come quelle proteine sovregolate durante una malattia o infezione, per esempio recettori e/o i loro ligandi corrispondenti. Esempi particolari di proteine di superficie cellulare includono molecole di adesione, ad esempio integrine come integrine β 1 ad esempio VLA-4, E-selectina, P selectina o L-selectina, CD2, CD3, CD4, CD5, CD7, CD8, CD11a, CD11b, CD18, CD19, CD20, CD23, CD25, CD33, CD38, CD40, CD40L, CD45, CDW52, CD69, CD134 (OX40), ICOS, BCMP7, CD137, CD27L, CDCP1, CSF1 o Recettore CSF1, DPCR1, DPCR1, dudulina2 , FLJ20584, FLJ40787, HEK2, KIAA0634, KIAA0659, KIAA1246, KIAA1455, LTBP2, LTK, MAL2, MRP2, tipo-nettina2, NKCC1, PTK7, RAIG1, TCAM1, SC6, BCMP101, BCMP84, BCMP11, DTD, antigene carcinoembrionale (CEA) , globulina di grasso di latte umano (HMFG1 e 2), antigeni MHC di classe I e MHC di classe II, KDR e VEGF, e, quando opportuno, loro recettori.

Gli antigeni solubili includono interleuchine come IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-8, IL-12, IL-13, IL-14, IL-16 o IL-17, come IL17A e/o IL17F, antigeni virali per esempio antigeni del virus respiratorio sinciziale o antigeni del citomegalovirus, immunoglobuline, come IgE, interferoni come l'interferone α , l'interferone β o l'interferone γ , il fattore di necrosi tumorale TNF (precedentemente noto come fattore di necrosi tumorale- α), fattore di necrosi tumorale- β , fattori stimolanti la colonia come G-CSF o GM-CSF, e fattori di crescita derivati da piastrine come PDGF- α , e PDGF- β e, ove appropriato, i loro recettori. Altri antigeni includono antigeni di superficie delle cellule batteriche, tossine batteriche, virus come l'influenza, EBV, HepA, B e C, agenti di

bioterrorismo, radionuclidi e metalli pesanti, veleni di serpente e ragno e tossine.

In una forma di realizzazione l'antigene è FcRn.

Gli anticorpi per l'uso nella presente divulgazione si possono ottenere usando qualsiasi metodo adatto noto nella tecnica. Il polipeptide/proteina FcRn comprendente proteine di fusione e suoi mutanti, comprese le cellule (ricombinanti o naturalmente) che esprimono il polipeptide (come le cellule T-attivate) può essere usato per produrre anticorpi che riconoscono specificamente FcRn. Il polipeptide può essere il polipeptide "maturo" o un suo frammento o derivato biologicamente attivo. La proteina umana è registrata in Swiss-Prot con il numero P55899. In una forma di realizzazione l'immunogeno è la catena alfa FcRn o un suo frammento.

I polipeptidi, per l'uso per immunizzare un ospite, possono essere preparati mediante processi ben noti nella tecnica da cellule ospite geneticamente ingegnerizzate comprendenti sistemi di espressione o possono essere recuperati da fonti biologiche naturali. Nella presente domanda, il termine "polipeptidi" include peptidi, polipeptidi e proteine. Questi sono usati in modo intercambiabile se non diversamente specificato. Il polipeptide FcRn può in alcuni casi essere parte di una proteina più grande come una proteina di fusione, ad esempio fuso con un tag di affinità.

Si possono ottenere anticorpi generati contro il polipeptide FcRn, quando è necessaria l'immunizzazione di un animale, somministrando i polipeptidi ad un animale, preferibilmente un animale non umano, usando protocolli di routine e ben noti, si veda ad esempio Handbook of Experimental Immunology, D. M. Weir (ed.), Vol 4, Blackwell Scientific Publishers, Oxford, England, 1986). Possono essere immunizzati molti animali a sangue caldo, come conigli, topi, ratti, pecore, mucche, cammelli o maiali. Tuttavia, topi, conigli, maiali e ratti sono generalmente più adatti.

In una forma di realizzazione, l'anticorpo può essere usato per alterare funzionalmente l'attività dell'antigene di interesse. Ad esempio, l'anticorpo può neutralizzare, antagonizzare o agonizzare l'attività di detto antigene, direttamente o indirettamente o semplicemente bloccare il legame del ligando normale ad esso. La presente invenzione fornisce anche una cellula batterica gram-negativa ricombinante comprendente un gene Tsp mutato, in cui il gene Tsp mutato codifica per una proteina Tsp avente attività proteasica ridotta o è un gene

Tsp mutato knockout, un gene mutante spr che codifica per un spr mutante, un gene in grado di esprimere o sovraesprimere una o proteine in grado di facilitare il ripiegamento proteico e una sequenza polinucleotidica codificante un anticorpo o un suo frammento legante l'antigene specifico per FcRn.

Vari frammenti e anticorpi anti-FcRn sono mostrati nelle sequenze da 36 a 74.

In una forma di realizzazione la catena pesante comprende 1, 2 o 3 CDR selezionati indipendentemente tra le SEQ ID NO: 36, 37 e 38.

In una forma di realizzazione la catena leggera comprende 1, 2 o 3 CDR selezionati indipendentemente tra le SEQ ID NO: 39, 40 e 41.

In una forma di realizzazione la catena pesante di anticorpo comprende la sequenza data in SEQ ID NO: 36 per CDR-H1, la sequenza data in SEQ ID NO: 37 per CDR-H2 e la sequenza data in SEQ ID NO: 38 per CDRH3.

In una forma di realizzazione la catena leggera di anticorpo comprende la sequenza data in SEQ ID NO: 39 per CDR-L1, la sequenza data in SEQ ID NO: 40 per CDR-L2 e la sequenza data in SEQ ID NO: 41 per CDRL3.

In una forma di realizzazione la catena pesante di anticorpo comprende la sequenza indicata in SEQ ID NO: 36 per CDR-H1, la sequenza indicata in SEQ ID NO: 37 per CDR-H2, la sequenza indicata in SEQ ID NO: 38 per CDRH3, la sequenza indicata in SEQ ID NO: 39 per CDR-L1, la sequenza indicata in SEQ ID NO: 40 per CDR-L2 e la sequenza indicata in SEQ ID NO: 41 per CDRL3.

In una forma di realizzazione la cellula esprime una molecola anticorpale con una sequenza qui descritta. Le molecole anticorpali includono anticorpi e loro frammenti leganti.

Opportunamente, l'anticorpo umanizzato secondo la presente invenzione ha un dominio variabile comprendente regioni di struttura di accettori umani nonché uno o più dei CDR qui forniti specificamente. Pertanto, in una forma di realizzazione è fornito un anticorpo umanizzato che lega FcRn umano in cui i domini variabili comprendono regioni di struttura accettrici umane e CDR di donatori non umani. In un esempio il dominio variabile della catena leggera comprende la sequenza data in SEQ ID NO: 50 e il dominio variabile catena pesante comprende la sequenza data in SEQ ID NO: 58. In un esempio la catena leggera comprende la sequenza data in SEQ ID NO: 54 e la catena pesante comprende la sequenza data in SEQ ID NO: 62.

Dopo l'espressione, i frammenti di anticorpi possono essere ulteriormente elaborati, ad esempio mediante coniugazione con un'altra entità come una molecola effettrice. Il termine molecola effettrice come qui usato include, ad esempio, agenti antineoplastici, farmaci, tossine (come tossine enzimaticamente attive di origine batterica o vegetale e loro frammenti ad esempio ricina e relativi frammenti) proteine biologicamente attive, ad esempio enzimi, altro anticorpo o frammenti di anticorpo, polimeri sintetici o presenti in natura, acidi nucleici e loro frammenti ad esempio DNA, RNA e loro frammenti, radionuclidi, in particolare radioioduro, radioisotopi, metalli chelati, nanoparticelle e gruppi reporter come composti fluorescenti o composti che possono essere rilevati da NMR o spettroscopia ESR. La molecola effettrice può essere attaccata all'anticorpo o suo frammento con qualsiasi metodo adatto, per esempio un frammento di anticorpo può essere modificato per attaccare almeno una molecola effettrice come descritto in WO05/003171 o WO05/003170 (il cui contenuto è qui incorporato da riferimento). WO05/003171 o WO05/003170 descrivono anche opportune molecole effettrici.

In una forma di realizzazione l'anticorpo o il suo frammento, come un Fab, è PEGilato per generare un prodotto con le proprietà richieste, ad esempio simili a tutti gli anticorpi, se necessario. Ad esempio, l'anticorpo può essere un anti-TNF- α Fab' PEGilato, come descritto in WO 01/094585, preferibilmente collegato a uno dei residui di cisteina all'estremità C-terminale della catena pesante un gruppo derivato da lisil-maleimmide in cui ciascuno dei due gruppi amminici del residuo lisilico ha legato covalentemente ad esso un residuo metossipoli (etilenglicole) avente un peso molecolare di circa 20.000 Da, tale che il peso molecolare medio totale dei residui di metossipoli (etilenglicole) è circa 40.000Da, più preferibilmente il gruppo derivato da lisil-maleimmide è [1-[[[2-[[3-(2,5-diosso-1-pirrolidinil)-1-ossopropil] ammino] etil] ammino]-carbonil]-1, 5-pentandiil] bis (imminocarbonile).

In una forma di realizzazione un Fab o Fab' secondo la presente descrizione è coniugato con una molecola di albumina di siero umano o una molecola di amido. La cellula può anche comprendere ulteriori sequenze di polinucleotidi che codificano una o più proteine di interesse.

Inoltre sono qui descritte cellule in cui una o più proteine ospiti di *E. coli* che nel tipo selvatico sono note per co-purificare con la proteina ricombinante di interesse durante la purificazione sono selezionate per la modificazione genetica, come descritto in Humphreys et al. "Engineering of Escherichia coli to improve the purification of periplasmic Fab' fragments: changing the pI of the chromosomally encoded PhoS/PstS protein", *Protein Expression e Purification* 37 (2004) 109-118 e WO04/035792 (il cui contenuto è incorporato qui per riferimento). L'uso di tali proteine ospiti modificate migliora il processo di purificazione per le proteine di interesse, in particolare di anticorpi, prodotti in *E. coli* alterando le proprietà fisiche di proteine di *E. coli* selezionate in modo che non siano più co-purificante con l'anticorpo ricombinante. Preferibilmente, la proteina *E.coli* che viene alterata è scelta tra una o più delle proteine leganti il fosfato (PhoS / PstS), la proteina legante del dipeptide (DppA), la proteina legante il maltosio (MBP) e la tioredoxina. In un esempio, una proprietà fisica di una proteina ospite contaminante è alterata mediante l'aggiunta di un tag amminoacidico al C-terminale o al N-terminale. In una forma di realizzazione preferita, la proprietà fisica che è alterata è il punto isoelettrico e il tag amminoacidico è un tag di acido poli-aspartico attaccato al C-terminale. In una forma di realizzazione le proteine di *E.coli* alterate dall'aggiunta di tale tag sono la proteina legante dipeptidi (DppA), la proteina legante il maltosio (MBP), la proteina legante tioredoxina e fosfato (PhoS / PstS). In una forma di realizzazione specifica, il pI della proteina legante fosfato di *E. coli* (PhoS / PstS) è ridotto da 7,2 a 5,1 mediante l'aggiunta di un tag di acido poliaspartico (polyD), contenente 6 residui di acido aspartico al terminale C.

È anche descritta la modifica di specifici residui della proteina *E. coli* contaminante per alterarne le proprietà fisiche, da sola o in combinazione con l'aggiunta di Tag N o C terminali. Tali cambiamenti possono includere inserzioni o delezioni per alterare la dimensione della proteina o sostituzioni di amminoacidi per alterare il pI o l'idrofobicità. Questi residui possono essere localizzati sulla superficie della proteina. In una forma di realizzazione, particolari residui di superficie della proteina PhoS sono alterati al fine di ridurre il pI della proteina. Residui che sono ritenuti essere importanti nel legame al fosfato (Bass, US 5,404,472) sono evitati al fine di mantenere una proteina PhoS funzionale.

Preferibilmente, sono bersagliati i residui di lisina che proiettano lontano dalla superficie della proteina o sono in prossimità di grandi gruppi di residui basici. In un esempio, la proteina PhoS ha un tag di esa-acido poliaspartico attaccato al terminale C, mentre i residui di superficie all'estremità opposta della molecola sono destinati alla sostituzione. Preferibilmente, residui di lisina selezionati sono sostituiti con acido glutammico o acido aspartico per conferire un maggiore potenziale di variazione del pI rispetto a quando si cambiano i residui neutri in acidi. La designazione di una sostituzione mutante qui consiste in una lettera seguita da un numero seguito da una lettera.

La prima lettera indica l'aminoacido nella proteina selvatica. Il numero si riferisce alla posizione dell'amminoacido in cui è effettuata la sostituzione dell'amminoacido, e la seconda lettera designa l'amminoacido usato per sostituire l'amminoacido selvatico. Nelle mutazioni preferite di PhoS nella presente invenzione i residui di lisina (K) 275, 107, 109, 110, 262, 265, 266, 309, 313 sono sostituiti con acido glutammico (E) o glutammina (Q), come mutazioni singole o combinate, inoltre, la lisina (K) 318 può essere sostituita con l'acido aspartico (D) come mutazione singola o combinata. Preferibilmente le mutazioni singole sono K262E, K265E e K266E. Preferibilmente le mutazioni combinate sono K265 / 266E e K110/265/266E. Più preferibilmente, tutte le mutazioni sono combinate con il tag di acido poliaspartico (polyD) attaccato al C-terminale e opzionalmente anche con la sostituzione K318D. In un ulteriore esempio, le mutazioni determinano una riduzione del pI di almeno 2 unità. Preferibilmente le mutazioni riducono il pI di PhoS da 7,2 a tra circa 4 e circa 5,5. In un esempio il pI della proteina PhoS di *E. coli* è ridotto da 7,2 a circa 4,9, circa 4,8 e circa 4,5 utilizzando le mutazioni poliD K318D, poliD K265/266E e poliD K110/265/266E rispettivamente.

Il polinucleotide codificante per la proteina di interesse può essere espresso come fusione con un altro polipeptide, preferibilmente una sequenza segnale o un altro polipeptide avente un sito di taglio specifico all'N-terminale del polipeptide maturo. La sequenza segnale eterologa selezionata dovrebbe essere una sequenza che è riconosciuta ed elaborata dalla cellula ospite. Per le cellule ospite procariotiche che non riconoscono e processano la sequenza segnale del polipeptide eucariotico o nativo, la sequenza segnale è sostituita con una sequenza segnale procariotica. Sequenze segnale idonee includono OmpA, PhoA, LamB, PelB, DsbA e DsbC.

La costruzione di vettori adatti contenenti uno o più dei componenti sopra elencati impiega tecniche standard di ligazione. Plasmidi isolati o frammenti di DNA sono tagliati, adattati, e riligati nella forma desiderata per generare i plasmidi richiesti. In una forma di realizzazione, è impiegata una cassetta di espressione nella presente invenzione per trasportare il polinucleotide codificante per la proteina di interesse che tipicamente comprende una o più sequenze codificanti proteine, che codificano una o più proteine di interesse e una o più sequenze di regolazione dell'espressione. Le una o più sequenze di regolazione dell'espressione possono includere un promotore. Le una o più sequenze di regolazione dell'espressione possono anche includere una regione 3' non tradotta come una sequenza di terminazione. Promotori adatti sono discussi in maggior dettaglio di seguito. In una forma di realizzazione, la cellula secondo la presente invenzione comprende un vettore, come un plasmide. Il vettore preferibilmente comprende una o più delle cassette di espressione come sopra definite. Nella forma di realizzazione in cui la proteina di interesse è un anticorpo comprendente entrambe le catene pesanti e leggere, la linea cellulare può essere trasfettata con due vettori, un primo vettore che codifica un polipeptide a catena leggera e un secondo vettore che codifica un polipeptide a catena pesante. In alternativa, può essere usato un vettore singolo, il vettore includendo sequenze che codificano polipeptidi a catena leggera e catena pesante. Il vettore per l'uso nella presente invenzione può essere prodotto inserendo una cassetta di espressione come definita sopra in un vettore adatto. In alternativa, le sequenze regolazione dell'espressione per dirigere l'espressione della sequenza polinucleotidica codificante una proteina di interesse, possono essere contenute nel vettore e quindi può essere richiesta solo la regione codificante del polinucleotide per completare il vettore. Esempi di vettori che possono essere impiegati per trasformare la cellula ospite con un polinucleotide secondo l'invenzione includono: un plasmide, come pBR322 o pACYC184, e/o un vettore virale come fago batterico, un elemento genetico trasponibile come un trasposone. Sono disponibili molte forme di vettore di espressione. Tali vettori di solito comprendono un'origine plasmidica della replicazione del DNA, un marcatore antibiotico selezionabile un

promotore e un terminatore trascrizionale separati da un sito di multi-clonazione (cassetta di espressione) e una sequenza di DNA codificante per un sito di legame ribosomiale. I promotori impiegati nella presente invenzione possono essere direttamente legati al relativo polinucleotide o, in alternativa, possono essere collocati in una posizione appropriata, ad esempio in un vettore tale che quando il relativo polipeptide è inserito, il relativo promotore può agire sullo stesso. In una forma di realizzazione il promotore si trova prima della porzione codificante del polinucleotide su cui agisce, ad esempio un promotore adatto prima di ciascuna porzione polinucleotidica codificante. "Prima" come qui usato è inteso implicare che il promotore si trova al 5' primo in relazione alla porzione di polinucleotide codificante.

I promotori possono essere endogeni o esogeni alle cellule ospite. Promotori adatti includono lac, tac, trp, phoA, Ipp, Arab, tet e T7. Uno o più promotori impiegati possono essere promotori inducibili.

Le unità di espressione per l'uso in sistemi batterici, generalmente contengono anche una sequenza ribosomiale Shine-Dalgarno (S. D.) operativamente legata al DNA che codifica il polipeptide di interesse. Il vettore di espressione preferibilmente comprende anche un messaggio dicistronico per produrre l'anticorpo o un suo frammento che lega l'antigene come descritto in WO03/048208 o WO2007/039714 (il cui contenuto è qui incorporato per riferimento). Preferibilmente il cistrona a monte contiene il DNA codificante la catena leggera dell'anticorpo e il cistrona a valle contiene il DNA codificante la corrispondente catena pesante, e la sequenza intergenica dicistronica (IGS) preferibilmente comprende una sequenza selezionata tra IGS1 (SEQ ID NO: 38), IGS2 (SEQ ID NO: 39), IGS3 (SEQ ID NO: 40) e IGS4 (SEQ ID NO: 41).

I terminatori possono essere endogeni o esogeni alle cellule ospiti. Un terminatore adatto è rrnB.

Ulteriori regolatori trascrizionali adatti inclusi promotori e terminatori e metodi di targeting proteico si possono trovare in "Strategies for Achieving High-Level Expression of Genes in Escherichia coli" Savvas C. Makrides, Microbiological Reviews, Sept 1996, p 512-538.

La molecola anticorpale può essere secreta dalla cellula o indirizzata al periplasma mediante opportune sequenze segnale. In alternativa, le molecole anticorpali possono accumularsi all'interno del citoplasma della cellula.

Preferibilmente la molecola anticorpale è indirizzata al periplasma.

Forme di realizzazione dell'invenzione qui descritte con riferimento al polinucleotide si applicano ugualmente a forme di realizzazione alternative dell'invenzione, ad esempio vettori, cassette di espressione e/o cellule ospite comprendenti i componenti ivi impiegati, fintanto che l'aspetto rilevante possa essere ad essi applicato.

La presente invenzione fornisce anche un metodo per produrre una proteina ricombinante di interesse comprendente esprimere la proteina ricombinante di interesse in una cellula batterica gram-negativa ricombinante come descritto sopra per la presente invenzione.

La proteina batterica gram-negativa di interesse preferibilmente impiegata nel metodo della presente invenzione è descritta in dettaglio sopra.

Quando il polinucleotide codificante la proteina di interesse è esogeno, il polinucleotide può essere incorporato nella cellula ospite usando qualsiasi mezzo adatto noto nella tecnica. Tipicamente, il polinucleotide è incorporato come parte di un vettore di espressione che viene trasformato nella cella. Di conseguenza, in un aspetto la cellula secondo la presente invenzione comprende una cassetta di espressione comprendente il polinucleotide codificante la proteina di interesse.

La sequenza polinucleotidica può essere trasformata in una cellula usando tecniche standard, ad esempio impiegando cloruro di rubidio, PEG o elettroporazione.

Come ulteriormente descritto qui, il metodo può anche impiegare un sistema di selezione per facilitare la selezione di cellule stabili che sono state trasformate con successo con il polinucleotide codificante la proteina di interesse. Il sistema di selezione impiega tipicamente la co-trasformazione di una sequenza polinucleotidica che codifica un marcatore di selezione. In un esempio, ciascun polinucleotide trasformato nella cellula comprende inoltre una sequenza polinucleotidica che codifica uno o più marcatori di selezione. Di conseguenza, la trasformazione del polinucleotide codificante per la proteina di interesse e gli uno o più polinucleotidi codificanti per il marcatore avviene insieme e il sistema di selezione può essere impiegato per selezionare quelle cellule che producono le proteine desiderate.

Le cellule in grado di esprimere gli uno o più marcatori sono in grado di sopravvivere/crescere/moltiplicarsi in

determinate condizioni artificialmente imposte, ad esempio l'aggiunta di una tossina o di un antibiotico, grazie alle proprietà fornite dal componente polipeptide/gene o polipeptide del sistema di selezione incorporato (ad esempio resistenza agli antibiotici). Quelle cellule che non possono esprimere uno o più marcatori non sono in grado di sopravvivere/crescere/moltiplicarsi nelle condizioni imposte artificialmente. Le condizioni imposte artificialmente possono essere scelte per essere più o meno vigorose, come richiesto. Qualsiasi sistema di selezione adatto può essere impiegato nella presente invenzione. In genere il sistema di selezione può essere basato sull'inclusione nel vettore di uno o più geni che forniscono resistenza ad un antibiotico noto, ad esempio un gene della resistenza alla tetraciclina, al cloramfenicolo, alla kanamicina o all'ampicillina. Le cellule che crescono in presenza di un antibiotico adatto possono essere selezionate in quanto esprimono sia il gene che conferisce resistenza all'antibiotico sia la proteina desiderata. In una forma di realizzazione, il metodo secondo la presente invenzione comprende inoltre la fase di coltivare la cellula trasformata in un terreno per esprimere così la proteina di interesse. Può essere usato un sistema di espressione inducibile o un promotore costitutivo noto per esprimere la proteina di interesse. Sistemi di espressione inducibili adatti e promotori costitutivi sono ben noti nella tecnica. Può essere usato qualsiasi terreno adatto per coltivare la cellula trasformata. Il terreno può essere adattato per uno specifico sistema di selezione, ad esempio il terreno può comprendere un antibiotico, per consentire solo a quelle cellule che sono state trasformate con successo per crescere nel terreno. Le cellule ottenute dal mezzo possono essere sottoposte a ulteriore screening e/o purificazione secondo necessità. Il metodo può inoltre comprendere uno o più passaggi per estrarre e purificare la proteina di interesse come richiesto.

Il polipeptide può essere recuperato dal ceppo, incluso dal citoplasma, dal periplasma, dal sovratanante.

In una forma di realizzazione l'anticorpo è isolato dal periplasma.

In una forma di realizzazione la velocità di alimentazione post induzione è compresa tra 5 e 7,5 g/ora, ad esempio circa 7 g/ora.

Il/i metodo/i specifico/i usato/i per purificare una proteina dipende/ono dal tipo di proteina. Metodi idonei

includono il frazionamento su colonne di immuno-affinità o di scambio ionico; precipitazione con etanolo; HPLC a fase inversa; cromatografia di interazione idrofobica; cromatografia su silice; cromatografia su una resina a scambio ionico come S-SEPHAROSE e DEAE; chromatofocusing; precipitazione con solfato di ammonio; e filtrazione di gel.

Gli anticorpi possono essere opportunamente separati dal terreno di coltura e/o estratto di citoplasma e/o estratto di periplasma mediante procedure di purificazione di anticorpi convenzionali quali, ad esempio, protein A-Sepharose, cromatografia di proteina G, cromatografia di proteina L, tiofilica, resine a modalità mista, His -tag, FLAGTag, cromatografia su idrossiapatite, elettroforesi su gel, dialisi, cromatografia per affinità, solfato di ammonio, etanolo o frazionamento/precipitazione con PEG, membrane a scambio ionico, cromatografia ad adsorbimento a letto espanso (EBA) o cromatografia a letto mobile simulato.

Il metodo può anche includere un'ulteriore fase di misurazione della quantità di espressione della proteina di interesse e selezione di cellule con livelli di espressione elevati della proteina di interesse.

Il metodo può anche includere una o più fasi successive di trattamento a valle come la PEGilazione della proteina di interesse, quale un anticorpo o frammento di anticorpo.

Una o più fasi del metodo qui descritto possono essere eseguite in combinazione in un contenitore adatto quale un bioreattore.

Gli anticorpi e i frammenti secondo la presente descrizione possono essere impiegati nel trattamento o nella profilassi.

Quando tecnicamente opportuno, si possono combinare forme di realizzazione dell'invenzione. Le forme di realizzazione sono qui descritte come comprendenti certe caratteristiche/elementi. La descrizione si estende anche a forme di realizzazione separate che consistono o consistono essenzialmente in dette caratteristiche/elementi. I riferimenti tecnici qui, come brevetti e applicazioni sono qui incorporati per riferimento.

La presente invenzione è ulteriormente descritta solo a titolo illustrativo nei seguenti esempi, che si riferiscono alle Figure allegate.

ESEMPI

Esempio 1: Generazione di Linee Cellulari

La generazione di MXE016 e MXE017 è fornita in WO2011/086136.

Generazione di plasmidi per l'espressione anti-FcRn Fab' e anti-FcRn Fab' con FkpA, anti-FcRn Fab' con skp e anti-FcRn Fab con espressione di entrambi FkpA e skp

È stato costruito un plasmide contenente entrambe le sequenze di catena leggera e pesante di un anti-FcRn Fab (SEQ ID NO: 63 e 55, rispettivamente). Il plasmide pTTOD 1519.g57 Fab', è stato costruito usando metodologie convenzionali di clonazione a restrizione che possono essere trovate in Sambrook et al 1989, *molecular cloning: a laboratory manual*. CSHL press, N.Y. Il plasmide pTTOD 1519.g57 Fab' conteneva le seguenti caratteristiche: una sequenza di promotore forte *tac* e operatore *lac*. Il plasmide conteneva un sito di restrizione EcoRI unico dopo la regione codificante la catena pesante di Fab', seguita da una sequenza non codificante contenente un sito di legame al ribosoma forte e quindi un sito di restrizione NdeI unico.

I geni per la catena leggera e la catena pesante di Fab sono stati trascritti come un unico messaggio policistronico. Il DNA codificante per il peptide segnale dalla proteina OmpA di *E. coli*, che ha diretto la traslocazione dei polipeptidi nel periplasma di *E. coli*, è stato fuso all'estremità 5' di entrambe le sequenze dei geni della catena leggera e della catena pesante. La trascrizione è stata terminata utilizzando un terminatore di trascrizione doppio *rrnB t1t2*. Il gene *lacIq* codificava la proteina repressore Lac I costitutivamente espressa. Questo ha represso la trascrizione dal promotore *tac* fino a quando la de-repressione è stata indotta mediante presenza di allolattosio o IPTG. L'origine della replicazione utilizzata era p15A, che manteneva un numero basso di copie. Il plasmide conteneva un gene di resistenza alla tetraciclina per la selezione antibiotica. Il plasmide pTTOD 1519.g57 Fab' FkpA, un vettore di espressione per l'anti-FcRn Fab e FkpA (un polipeptide periplasmico), è stato costruito legando FkpA al termine 3' della sequenza Fab' di pTTOD 1519.g57 Fab' usando i siti EcoRI e NdeI (vedere Figura 7d). L'FkpA (SEQ ID NO: 30) è stato costruito sinteticamente per rimuovere 2 siti Puv II, un sito SfuI, un sito BamHI e un EcoRI, in modo tale che il nuovo costrutto codificasse per un sito EcoRI al 5' seguito da un forte sito di legame ribosomiale, seguito dal codone di inizio nativo, sequenza segnale

e sequenza matura di FkpA, terminando con un His-tag C-terminale e infine un forte sito di legame al ribosoma seguito da un sito NdeI non codificante. Il frammento di restrizione EcoRI-NdeI è stato digerito e ligato nel vettore di espressione in modo tale che tutti e tre i polipeptidi: catena leggera Fab, catena pesante Fab e FkpA fossero stati codificati su un singolo mRNA policistronico.

Il plasmide pTTOD 1519.g57 Fab' Skp, un vettore di espressione per l'anti-FcRn Fab e Skp (un polipeptide periplasmatico), è stato costruito legando skp al termine 3' della sequenza Fab del plasmide pTTOD 1519.g57 Fab' usando i siti EcoRI e NdeI. Lo skp (SEQ ID NO: 34) è stato costruito sinteticamente per rimuovere 4 siti Pst I e un sito EcoRV, in modo tale che il nuovo costrutto codificasse per un sito EcoRI 5' seguito da un sito di legame ribosomiale forte, seguito da un codone di inizio nativo, una sequenza del segnale e una sequenza matura di Skp, terminando in un His-tag C-terminale e infine un sito di legame al ribosoma forte seguito da un sito NdeI non codificante. Il frammento di restrizione EcoRI-NdeI è stato digerito e ligato nel vettore di espressione in modo tale che tutti e tre i polipeptidi: catena leggera Fab, catena pesante Fab' e skp fossero codificati su un singolo mRNA policistronico.

Il plasmide pTTOD 1519.g57 Fab' FkpA Skp, un vettore di espressione per l'anti-FcRn Fab, FkpA e Skp (entrambi polipeptidi periplasmatici), è stato costruito legando Skp nel plasmide pTTOD 1519.g57 Fab' FkpA, al termine 3' della Sequenza FkpA usando il sito NdeI. Lo Skp (SEQ ID NO: 34) è stato costruito sinteticamente per rimuovere 4 siti Pst I e un sito EcoRV, in modo tale che il costrutto codificasse per un sito Ase I includendo il codone di inizio nativo, la sequenza segnale e la sequenza matura di skp, terminando in un His-tag C-terminale e infine un sito NdeI non codificante. Il frammento di restrizione AseI-NdeI è stato digerito e ligato nel vettore di espressione in modo tale che tutti e quattro i polipeptidi: catena leggera Fab, catena pesante Fab, FkpA e Skp fossero codificati su un singolo mRNA policistronico.

Fermentazione di pTTOD 1519.g57 Fab', pTTOD 1519.g57 Fab' FkpA, pTTOD 1519.g57 Fab' Skp e pTTOD 1519.g57 Fab' FkpA skp in *E.coli* W3110, MXE012 (*E. coli* W3110 spr H145A), MXE016 (*E. coli* W3110 Δ Tsp, spr C94A) e MXE017 (*E. coli* W3110 Δ Tsp, spr H145A). I ceppi di *E. coli* W3110, MXE012, MXE016 e MXE017 sono stati trasformati con i plasmidi pTTOD 1519.g57 Fab', pTTOD 1519.g57 Fab' FkpA, pTTOD

1519.g57 Fab' Skp e pTTOD 1519.g57 Fab' FkpA Skp generati nell'esempio 1. La trasformazione dei ceppi è stata effettuata utilizzando il metodo trovato in Chung CT et al Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989). Questi ceppi trasformati sono stati saggiati per l'espressione mediante fermentazione.

Esempio 2: Effetto dei chaperones sull'espressione di Fab: I ceppi mostrati in **Figura 1** sono stati saggiati in esperimenti di fermentazione di 1L e 2.5L confrontando l'espressione di Fab' : **Terreno di crescita, fasi di inoculo e fermentazione.** Il processo di fermentazione è stato avviato preparando un inoculo da una fiala della banca cellulare e amplificando attraverso diversi stadi pre-coltura (fiasche e reattori) prima della semina nel fermentatore di produzione. Nel fermentatore di produzione, le cellule sono state coltivate in terreno definito ad alta densità in modalità batch e fed-batch. Quando è stata raggiunta la densità cellulare desiderata, l'espressione di Fab' è stata indotta mediante l'aggiunta di IPTG. L'espressione di Fab' è indirizzata allo spazio periplasmatico di *E. coli*, dove Fab' si accumula durante il corso della fase di induzione. È stata applicata un'alimentazione di una fonte di carbonio durante la fase di induzione per controllare l'espressione la crescita delle cellule. Temperatura, ossigeno disciolto (pO₂) e pH sono stati controllati per mantenere la coltura in condizioni di coltura ottimali.

Misura della concentrazione della biomassa e del tasso di crescita. La concentrazione della biomassa è stata determinata misurando la densità ottica delle colture a 600 nm.

Estrazione periplasmatica. Le cellule sono state raccolte da campioni di colture mediante centrifugazione. La frazione sovranatante è stata mantenuta (a -20 ° C) per ulteriori analisi. La frazione di pellet cellulare è stata risospesa al volume di coltura originale in tampone di estrazione (100 mM Tris-HCl, 10 mM EDTA, pH 7,4). Dopo incubazione a 60°C per circa 10-12 ore, l'estratto è stato chiarificato mediante centrifugazione e la frazione di sovranatante è stata usata fresca o conservata (a -20°C) per analisi.

Quantificazione di Fab'. Sono state determinate le concentrazioni di Fab' negli estratti periplasmatici e nei sovranatanti della coltura usando la HPLC proteina G. Una colonna HiTrap Protein-G HP da 1 ml (GE-Healthcare o equivalente) è stata caricata con l'analita (pH circa neutro, 30°C, 0,2 µm filtrata) a 2 ml/min, la

colonna è stata lavata con fosfato 20 mM, 50 mM NaCl pH 7,4 e quindi Fab' è stato eluito utilizzando un'iniezione di 50 mM di glicina / HCl a pH 2,7. Il Fab' è stato misurato con A280 su un sistema HPLC Agilent 1100 o 1200 e quantificato facendo riferimento a una curva standard di una proteina Fab' purificata di concentrazione nota.

La Figura 1 mostra i risultati di 46 fermentazioni in scala 5L effettuate con varie combinazioni di cellule ospite e "chaperone". W3110 è un ceppo selvatico di *E. coli*. le varie combinazioni erano: selvatico senza chaperone, selvatico con FkpA e Skp, MXE016 mutante spr e Δ Tsp pubblicati in WO2011/086136, MXE016 e FkpA, MXE016 e Skp, MXE016 e FkpA e Skp, MXE017 rivelati in WO2011/086136, MXE017 e FkpA e Skp.

I risultati mostrano che MXE016, MXE016 e FkpA, MXE016 e FkpA e Skp, e MXE017 e FkpA e Skp presentavano i migliori livelli di espressione.

Esempio 3: Effetto della velocità di alimentazione post induzione

L'effetto di tre diverse velocità di alimentazione post induzione 5,4, 6,0 e 7,0g/ora è stato saggiato su MXE016 e MXE016 + FkpA.

Terreno di crescita, fasi di inoculo e fermentazione. Il processo di fermentazione è stato avviato preparando un inoculo da una fiala della banca cellulare e amplificando attraverso diversi stadi pre-coltura (fiasche e reattori) prima della semina del fermentatore di produzione. Nel fermentatore di produzione, le cellule sono state coltivate in terreni definiti ad alta densità in modalità batch e fed-batch. Quando è stata raggiunta la densità cellulare desiderata, l'espressione di Fab' è stata indotta dall'aggiunta di IPTG. L'espressione di Fab' è indirizzata allo spazio periplasmatico di *E. coli*, dove Fab' si accumula durante il corso della fase di induzione. È stata applicata un'alimentazione di fonte di carbonio durante la fase di induzione per controllare l'espressione e la crescita delle cellule, in questo esperimento la velocità di alimentazione era impostata su tre setpoint separati (mostrato in figura). Temperatura, ossigeno disciolto (pO_2) e pH sono stati controllati per mantenere la coltura in condizioni di coltura ottimali.

La Figura 2 mostra il ceppo MXE016 trasformato con plasmidi che non esprimono chaperone o FkpA. La figura mostra l'effetto di aumento della velocità di alimentazione post induzione sulla produzione e ritenzione di

Fab' nel periplasma sia con che senza FkpA. I dati mostrano che quando la velocità di alimentazione è aumentata senza l'espressione FkpA, il Fab' prodotto in eccesso è perso nel sovranatante. Questo avviene con l'espressione di FkpA in cui il Fab' supplementare prodotto a velocità di alimentazione più elevate è trattenuto all'interno del periplasma risultando in un titolo più alto.

Esempio 4: Analisi della resa e della vitalità cellulare

Le figure 3A e B mostrano l'effetto sulla vitalità cellulare e il titolo di Fab' in MXE016 trasformato con plasmidi che non esprimono chaperone o FkpA in scala 20L. La Figura 3A mostra che la vitalità cellulare è inferiore dove non è presente FkpA. 3b mostra che i titoli di Fab' erano più alti e meno variabili per MXE016 + FkpA.

La Figura 4 mostra che il titolo più alto con MXE016 + FkpA risulta nel 30% in più di Fab' dopo il recupero del prodotto Fab' primario.

Le figure 5 e 6 mostrano western SDS-PAGE e anti-his-tag di campioni dal recupero primario di MXE016 senza chaperone e MXE016 con espressione FkpA.

I campioni sono stati caricati in base alla concentrazione di Fab' con 1 µg di Fab' caricato per corsia. I gel erano Tris-Glycine Novex al 4-20% e venivano fatti funzionare in condizioni non riducenti a 125 V per 2 ore. I gel sono stati colorati con Sypro Ruby stain e fotografati. I Western Blot sono stati trasferiti su una membrana PVDF utilizzando un sistema Invitrogen iBlot e quindi bloccati con una soluzione di caseina al 1%. Le membrane sono state poi esaminate con un anticorpo coniugato anti-his HRP (Novagen). I blot sono stati quindi sviluppati con la soluzione ECL e con l'utilizzo di un HyperProcessor di Amersham Life Sciences.

LISTA DI SEQUENZA

<110> UCB Pharma S.A

<120> Cellula ricombinante batterica per espressione proteica

<130> G0164-WO

<150> GB1208367.1

<151> 2012-05-14

<160> 82

<170> versione PatentIn 3.5

<210> 1

<211> 2049

<212> DNA

<213> Escherichia coli

<400> 1

atgaacatgt ttttaggct taccgctta gctggcctgc ttgcaatagc aggccagacc	60
ttcgctgtag aagatatac gcgtgctgat caaattccgg tattaagga agagacgcag	120
catgacgacg taagtgagcg cgtaacgtcg cgctcaccg gttctatta tcgccagttc	180
gacctcgatc aggcattttc ggccaaaatc tttgaccgct acctgaatct gctcgattac	240
agccacaacg tgctgctggc aagcgatggt gaacagttcg cgaaaagaa aaccgagtta	300
ggcgatgaac tgcgttcagg caaactcgac gttttctacg atctctacaa tctggcgcaa	360
aagcgccggt ttgagcggtta ccagtacgct ttgtcggtag tggaaaagcc gatggatttc	420
accggcaacg acaactataa ccttgaccgc agcaaagcgc cctggccgaa aaacgaggct	480
gagttgaacg cgctgtggga cagtaaagtc aaattcgacg agttaagcct gaagctgaca	540
ggaaaaacgg ataaagaaat tcgtgaaacc ctgactogcc gctacaaat tgccattcgt	600
cgtctggcgc aaaccaacag cgaagatggt ttctcgtcgg caatgacggc gtttgcgcgt	660
gaaatcgacc cgcataccaa ctatctttcc ccgcgtaata ccgaacagtt caaactgaa	720
atgagtttgt cgtcgggaag tattggcgca gtgctgcaaa tggatgatga ctacaccggt	780
atcaattcga tggcggcagc tggcggcga gcgaagagta aagctatcag cgttggtgac	840
aaaattgctg gtgttggtca aacaggcaag ccgatgggtg acgtgattgg ctggcgtctt	900
gatgatgtgg ttgccttaat taaagggccg aagggcagta aagttcgtct ggaaatttta	960
cctgctggta aagggaccaa gaccgtact gtaacgttga cccgtgaacg tattcgtctc	1020
gaagaccgcg cggttaaaat gtcgggtaag accgtcggta aagagaaagt cggcgtgctg	1080
gatattccgg gcttctatgt gggtttgaca gacgatgtca aagtgcaact gcagaaactg	1140
gaaaaacaga atgtcagcag cgtcatcatc gacctcgtga gcaatggcgg tggggcgtaa	1200
actgaagccg taccgctctc cggctctgtt attcctgcgg gtcccattgt tcaggtccgc	1260

gataacaacg gcaaggttcg tgaagatagc gataaccgacg gacaggtttt ctataaaggc 1320
ccgctggtgg tgctggttga ccgcttcagt gcttcggctt cagaaatctt tgcccgggca 1380
atgcaggatt acggtcgtgc gctggttggtg ggtgaaccga cgtttggtaa aggcaccggt 1440
cagcaatacc gttcattgaa ccgtatttac gatcagatgt tacgtcctga atggccagcg 1500
ctgggttctg tgcaagtacac gatccagaaa ttctatcgcg ttaacggcgg cagtacgcaa 1560
cgtaaaggcg taacgccaga catcatcatg ccgacgggta atgaagaaac ggaaacgggt 1620
gagaaattcg aagataacgc gctgccgtgg gatagcattg atgccgacgac ttatgtgaaa 1680
tcaggagatt taacggcctt tgaaccggag ctgctgaagg aacataatgc gcgtatcgcg 1740
aaagatcctg agttccagaa catcatgaag gatatcgcgc gttcaacgc tatgaaggac 1800
aagcgcaata tcgtttctct gaattacgct gtgctgaga aagagaataa tgaagatgat 1860
gcgacgcgctc tggcgcggtt gaacgaacgc tttaaaccgcg aaggtaaacc ggagttgaag 1920
aaactggatg atctaccgaa agattaccag gagccggatc cttatctgga tgagacggtg 1980
aatatcgcac tcgatctggc gaagcttgaa aaagccagac ccgcggaaca acccgctccc 2040
gtcaagtaa 2049

<210> 2

<211> 680

<212> PRT

<213> Escherichia coli

<400> 2

Met Phe Phe Arg Leu Thr Ala Leu Ala Gly Leu Leu Ala Ile Ala Gly
1 5 10 15

Gln Thr Phe Ala Val Glu Asp Ile Thr Arg Ala Asp Gln Ile Pro Val
20 25 30

Leu Lys Glu Glu Thr Gln His Ala Thr Val Ser Glu Arg Val Thr Ser
35 40 45

Arg Phe Thr Arg Ser His Tyr Arg Gln Phe Asp Leu Asp Gln Ala Phe
50 55 60

Ser Ala Lys Ile Phe Asp Arg Tyr Leu Asn Leu Leu Asp Tyr Ser His
65 70 75 80

Asn Val Leu Leu Ala Ser Asp Val Glu Gln Phe Ala Lys Lys Lys Thr
85 90 95

Glu Leu Gly Asp Glu Leu Arg Ser Gly Lys Leu Asp Val Phe Tyr Asp
100 105 110

Leu Tyr Asn Leu Ala Gln Lys Arg Arg Phe Glu Arg Tyr Gln Tyr Ala
115 120 125

Leu Ser Val Leu Glu Lys Pro Met Asp Phe Thr Gly Asn Asp Thr Tyr
130 135 140

Asn Leu Asp Arg Ser Lys Ala Pro Trp Pro Lys Asn Glu Ala Glu Leu
145 150 155 160

Asn Ala Leu Trp Asp Ser Lys Val Lys Phe Asp Glu Leu Ser Leu Lys
165 170 175

Leu Thr Gly Lys Thr Asp Lys Glu Ile Arg Glu Thr Leu Thr Arg Arg
180 185 190

Tyr Lys Phe Ala Ile Arg Arg Leu Ala Gln Thr Asn Ser Glu Asp Val
195 200 205

Phe Ser Leu Ala Met Thr Ala Phe Ala Arg Glu Ile Asp Pro His Thr
210 215 220

Asn Tyr Leu Ser Pro Arg Asn Thr Glu Gln Phe Asn Thr Glu Met Ser
225 230 235 240

Leu Ser Leu Glu Gly Ile Gly Ala Val Leu Gln Met Asp Asp Asp Tyr
245 250 255

Thr Val Ile Asn Ser Met Val Ala Gly Gly Pro Ala Ala Lys Ser Lys
260 265 270

Ala Ile Ser Val Gly Asp Lys Ile Val Gly Val Gly Gln Thr Gly Lys
275 280 285

Pro Met Val Asp Val Ile Gly Trp Arg Leu Asp Asp Val Val Ala Leu
290 295 300

Ile Lys Gly Pro Lys Gly Ser Lys Val Arg Leu Glu Ile Leu Pro Ala
305 310 315 320

Gly Lys Gly Thr Lys Thr Arg Thr Val Thr Leu Thr Arg Glu Arg Ile
325 330 335

Arg Leu Glu Asp Arg Ala Val Lys Met Ser Val Lys Thr Val Gly Lys
340 345 350

Glu Lys Val Gly Val Leu Asp Ile Pro Gly Phe Tyr Val Gly Leu Thr
355 360 365

Asp Asp Val Lys Val Gln Leu Gln Lys Leu Glu Lys Gln Asn Val Ser
370 375 380

Ser Val Ile Ile Asp Leu Arg Ser Asn Gly Gly Gly Ala Leu Thr Glu
385 390 395 400

Ala Val Ser Leu Ser Gly Leu Phe Ile Pro Ala Gly Pro Ile Val Gln
405 410 415

Val Arg Asp Asn Asn Gly Lys Val Arg Glu Asp Ser Asp Thr Asp Gly
420 425 430

Gln Val Phe Tyr Lys Gly Pro Leu Val Val Leu Val Asp Arg Phe Ser
435 440 445

Ala Ser Ala Ser Glu Ile Phe Ala Ala Ala Met Gln Asp Tyr Gly Arg
450 455 460

Ala Leu Val Val Gly Glu Pro Thr Phe Gly Lys Gly Thr Val Gln Gln
465 470 475 480

Tyr Arg Ser Leu Asn Arg Ile Tyr Asp Gln Met Leu Arg Pro Glu Trp
485 490 495

Pro Ala Leu Gly Ser Val Gln Tyr Thr Ile Gln Lys Phe Tyr Arg Val
500 505 510

Asn Gly Gly Ser Thr Gln Arg Lys Gly Val Thr Pro Asp Ile Ile Met
515 520 525

Pro Thr Gly Asn Glu Glu Thr Glu Thr Gly Glu Lys Phe Glu Asp Asn
530 535 540

Ala Leu Pro Trp Asp Ser Ile Asp Ala Ala Thr Tyr Val Lys Ser Gly
545 550 555 560

Asp Leu Thr Ala Phe Glu Pro Glu Leu Leu Lys Glu His Asn Ala Arg
565 570 575

Ile Ala Lys Asp Pro Glu Phe Gln Asn Ile Met Lys Asp Ile Ala Arg
580 585 590

Phe Asn Ala Met Lys Asp Lys Arg Asn Ile Val Ser Leu Asn Tyr Ala
595 600 605

Val Arg Glu Lys Glu Asn Asn Glu Asp Asp Ala Thr Arg Leu Ala Arg

atattccggg cttctatgtg ggtttgacag acgatgtcaa agtgcaactg cagaaactgg 1140
aaaaacagaa tgtcagcagc gtcacatcgc acctgcgtag caatggcggg ggggcgtaa 1200
ctgaagccgt atcgcctctcc ggtctgttta ttcctgcggg tcccattgtt caggcgcg 1260
ataacaacgg caaggctcgt gaagatagcg ataccgacgg acaggttttc tataaagcc 1320
cgctggtggt gctggttgac cgcttcagtg cttcggttc agaaatcttt gccgcggcaa 1380
tgcaggatta cggcgtgagc ctggttggtg gtgaaccgac gtttggtaaa ggcaccgttc 1440
agcaataccg ttcattgaac cgtatttacg atcagatggt acgtcctgaa tggccagcgc 1500
tgggttctgt gcagtacacg atccagaaat tctatcgcgt taacggcggc agtacgcaac 1560
gtaaaggcgt aacgccagac atcatcatgc cgacgggtaa tgaagaaacg gaaacgggtg 1620
agaaatcga agataacgcg ctgccgtggg atagcattga tgccgcgact tatgtgaaat 1680
caggagattt aacggccttt gaaccggagc tgctgaagga acataatgcg cgtatcgcga 1740
aagatcctga gttccagaac atcatgaagg atatcgcgcg cttcaacgct atgaaggaca 1800
agcgcfaatat cgtttctctg aattacgctg tgcgtgagaa agagaataat gaagatgatg 1860
cgacgcgtct gccgcgtttg aacgaacgct ttaaaccgca aggtaaaccg gagttgaaga 1920
aactggatga tctaccgaaa gattaccagg agccggatcc ttatctggat gagacggtga 1980
atatcgcact cgatctggcg aagcttgaaa aagccagacc cgcggaacaa cccgctcccg 2040
tcaagtaa 2048
<210> 4
<211> 2889
<212> DNA
<213> Escherichia coli
<400> 4
atgccccgca gcacctgggt caaagcatta ttgttgtag ttgcccttg gccaccctta 60
agtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120
aaccgccagt atcaggctat acgtctggat aacggtatgg tggctctgct ggtttctgat 180
ccgcaggcag ttaaatcgct ctccggcctg gtggtgcccg ttgggtcgct ggaagatccc 240
gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag 300
taccgcgagg ctgacagtct ggccgaatat ctcaaatgc acggcggtag tcacaatgcc 360
agcactgagc cgtatgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt 420
gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa 480
cgtgagcgtg atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgagc 540
atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600
ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660

ttccacgaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaacccg 720
ctgccggagt tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780
aaaaaacccg aatcacccgt gccggtagtc accgacgcgc aaaagggcat tatcattcat 840
tacgtccctg cgctgccgcg taaagtgttg cgcgttgagt ttcgcatcga taacaactca 900
gcgaagtcc gtagtaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960
ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020
gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080
ggcctggcta atcgcgatca ggttgtggcg gcaatthtta gctatctcaa tctgttacgt 1140
gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200
cgttatccgt cgatcacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260
cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320
gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380
ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgcctatca ggtcgataaa 1440
atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500
ccagagctta acccttatat tcctgatgat ttctcgctga ttaagtcaga gaagaaatac 1560
gaccatccag agctgattgt tgatgagtcg aatctgcgcg tgggtgatgc gccaaagccgt 1620
tattttgccg gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680
gacagcgcgc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740
gatcagttaa gcaaccaggc gtcggttggg ggcataagtt tttccaccaa cgctaacaac 1800
ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860
ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtccctgg 1920
tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgccc 1980
gcgcagatgc tctcgcaagt gccgtacttc tcgcgagatg aacggcgtaa aatthtggcc 2040
tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100
gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160
caaaaacagt tgggcgctga tggttcagag tgggtgcgaa acaaagatgt agtggctgat 2220
aaaaaacaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280
gtatttgtag cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340
cagatcgtac agccgtggtt ctacaatcag ttgcgtaccg aagaacaatt gggctatgcc 2400
gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttcct tttgcaaagc 2460
aatgataaac agccttcatt cttgtgggag cgttacaagg cgtthttccc aaccgcagag 2520

gcaaaattgc gagcgatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580
cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640
gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700
acgccgcaaa aacttgctga tttcttccat cagcgggtgg tcgagccgca aggcattggc 2760
attctgtcgc agatttccgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820
tggaaagtgt gggagaacgt cagcgcgttg cagcaaacaa tgcccctgat gagtgaaaag 2880
aatgagtga 2889
<210> 5
<211> 962
<212> PRT
<213> Escherichia coli
<400> 5
Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
1 5 10 15

Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu
20 25 30

Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg
35 40 45

Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val
50 55 60

Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro
65 70 75 80

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met
85 90 95

Gly Ser Lys Lys Tyr Pro Gln Ala Asp Ser Leu Ala Glu Tyr Leu Lys
100 105 110

Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala
115 120 125

Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg
130 135 140

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu
145 150 155 160

Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg

165 170 175

Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala
180 185 190

His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp
195 200 205

Lys Pro Gly Asn Pro Val Gln Gln Ala Leu Lys Asp Phe His Glu Lys
210 215 220

Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val Ile Tyr Ser Asn Lys Pro
225 230 235 240

Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro
245 250 255

Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp
260 265 270

Ala Gln Lys Gly Ile Ile Ile His Tyr Val Pro Ala Leu Pro Arg Lys
275 280 285

Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg
290 295 300

Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro
305 310 315 320

Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile
325 330 335

Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala
340 345 350

Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val
355 360 365

Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile
370 375 380

Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe
385 390 395 400

Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala
405 410 415

Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn
420 425 430

Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met
435 440 445

Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro
450 455 460

His Asn Lys Thr Ala Tyr Phe Val Asp Ala Pro Tyr Gln Val Asp Lys
465 470 475 480

Ile Ser Ala Gln Thr Phe Ala Asp Trp Gln Lys Lys Ala Ala Asp Ile
485 490 495

Ala Leu Ser Leu Pro Glu Leu Asn Pro Tyr Ile Pro Asp Asp Phe Ser
500 505 510

Leu Ile Lys Ser Glu Lys Lys Tyr Asp His Pro Glu Leu Ile Val Asp
515 520 525

Glu Ser Asn Leu Arg Val Val Tyr Ala Pro Ser Arg Tyr Phe Ala Ser
530 535 540

Glu Pro Lys Ala Asp Val Ser Leu Ile Leu Arg Asn Pro Lys Ala Met
545 550 555 560

Asp Ser Ala Arg Asn Gln Val Met Phe Ala Leu Asn Asp Tyr Leu Ala
565 570 575

Gly Leu Ala Leu Asp Gln Leu Ser Asn Gln Ala Ser Val Gly Gly Ile
580 585 590

Ser Phe Ser Thr Asn Ala Asn Asn Gly Leu Met Val Asn Ala Asn Gly
595 600 605

Tyr Thr Gln Arg Leu Pro Gln Leu Phe Gln Ala Leu Leu Glu Gly Tyr
610 615 620

Phe Ser Tyr Thr Ala Thr Glu Asp Gln Leu Glu Gln Ala Lys Ser Trp
625 630 635 640

Tyr Asn Gln Met Met Asp Ser Ala Glu Lys Gly Lys Ala Phe Glu Gln
645 650 655

Ala Ile Met Pro Ala Gln Met Leu Ser Gln Val Pro Tyr Phe Ser Arg
660 665 670

Asp Glu Arg Arg Lys Ile Leu Pro Ser Ile Thr Leu Lys Glu Val Leu
675 680 685

Ala Tyr Arg Asp Ala Leu Lys Ser Gly Ala Arg Pro Glu Phe Met Val
690 695 700

Ile Gly Asn Met Thr Glu Ala Gln Ala Thr Thr Leu Ala Arg Asp Val
705 710 715 720

Gln Lys Gln Leu Gly Ala Asp Gly Ser Glu Trp Cys Arg Asn Lys Asp
725 730 735

Val Val Val Asp Lys Lys Gln Ser Val Ile Phe Glu Lys Ala Gly Asn
740 745 750

Ser Thr Asp Ser Ala Leu Ala Ala Val Phe Val Pro Thr Gly Tyr Asp
755 760 765

Glu Tyr Thr Ser Ser Ala Tyr Ser Ser Leu Leu Gly Gln Ile Val Gln
770 775 780

Pro Trp Phe Tyr Asn Gln Leu Arg Thr Glu Glu Gln Leu Gly Tyr Ala
785 790 795 800

Val Phe Ala Phe Pro Met Ser Val Gly Arg Gln Trp Gly Met Gly Phe
805 810 815

Leu Leu Gln Ser Asn Asp Lys Gln Pro Ser Phe Leu Trp Glu Arg Tyr
820 825 830

Lys Ala Phe Phe Pro Thr Ala Glu Ala Lys Leu Arg Ala Met Lys Pro
835 840 845

Asp Glu Phe Ala Gln Ile Gln Gln Ala Val Ile Thr Gln Met Leu Gln
850 855 860

Ala Pro Gln Thr Leu Gly Glu Glu Ala Ser Lys Leu Ser Lys Asp Phe
865 870 875 880

Asp Arg Gly Asn Met Arg Phe Asp Ser Arg Asp Lys Ile Val Ala Gln
885 890 895

Ile Lys Leu Leu Thr Pro Gln Lys Leu Ala Asp Phe Phe His Gln Ala
900 905 910

Val Val Glu Pro Gln Gly Met Ala Ile Leu Ser Gln Ile Ser Gly Ser
915 920 925

Gln Asn Gly Lys Ala Glu Tyr Val His Pro Glu Gly Trp Lys Val Trp
930 935 940

Glu Asn Val Ser Ala Leu Gln Gln Thr Met Pro Leu Met Ser Glu Lys
945 950 955 960

Asn Glu

<210> 6

<211> 2889

<212> DNA

<213> Sequenza Artificiale

<220>

<223> gene mutato knockout Proteasi III

<400> 6

```
attccccgca gcacctggtt caaagcatta ttgttgtag ttgcccttg ggcacattaa      60
tgtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat     120
aacgcccagt atcaggctat acgtctggat aacggtatgg tggctctgct ggtttctgat     180
ccgcaggcag ttaaatcgct ctcggcgctg gtggtgcccg ttgggtcgct ggaagatccc     240
gaggcgtagc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag     300
taccgcaggc ctgacagtct ggccgaatat ctcaaatgc acggcggtag tcacaatgcc     360
agcactgccc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt     420
gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa     480
cgtgagcgtg atgcgggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgccc     540
atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt     600
ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat     660
ttccacgaga agtactattc cgccaatttg atgaaaggcg ttatttacag taataaaccg     720
ctgccggagt tggcaaaaat ggccggcgac acctttggtc gcgtgccgaa caaagagagc     780
aaaaaacccg aatcaccgtg gccggtagtc accgacgcgc aaaagggcat tatcattcat     840
tacgtccctg cgctgccgcg taaagtgttg cgcgttgagt ttcgcatcga taacaactca     900
gcgaagttcc gtagtaaaac cgatgaattg attacctatc tgattggcaa tcgcagccca     960
ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc    1020
gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa    1080
ggcctggcta atcgcgatca ggttggtggc gcaattttta gctatctcaa tctgttacgt    1140
gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc    1200
```

cgttatccgt cgatcacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260
cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320
gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380
ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgcggtatca ggtcgataaa 1440
atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500
ccagagctta acccttatac tcctgatgat ttctcgctga ttaagtcaga gaagaaatac 1560
gaccatccag agctgattgt tgatgagtcg aatctgcgcg tgggtgatgc gccaaagccgt 1620
tattttgcca cgcagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680
gacagcgcgc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagc gctggcgctt 1740
gatcagttaa gcaaccaggc gtcggttggt ggcataagtt tttccaccaa cgctaacaac 1800
ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860
ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtcctgg 1920
tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgccc 1980
gcgcagatgc tctcgcgaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040
tccattacgt tgaagagggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100
gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcatgtg 2160
caaaaacagt tgggcgctga tggttcagag tgggtgcgaa acaaagatgt agtggtcgat 2220
aaaaaacaaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280
gtatttgtag cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340
cagatcgtac agccgtgggt ctacaatcag ttgctaccg aagaacaatt gggctatgcc 2400
gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttccct tttgcaaagc 2460
aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520
gcaaaattgc gagcagatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580
cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taagatttc 2640
gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700
acgccgcaaa aacttgctga tttcttccat caggcgggtg tcgagccgca aggcattggct 2760
attctgtcgc agatttccgg cagccagaac gggaaaagccg aatatgtaca ccctgaaggc 2820
tggaagtggt gggagaacgt cagcgcgttg cagcaacaa tgcccctgat gagtgaag 2880
aatgagtga 2889
<210> 7
<211> 1425

<212> DNA

<213> Escherichia coli

<400> 7

```
atgaaaaaaaa ccacattagc actgagtgc ctggctctga gtttaggtt ggcgttatct      60
ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaaagc     120
cttgaccgga tgctcgaaaa ggtgatgcct tcagtgggtca gcattaacgt agaaggtagc     180
acaaccgtta atacgccgcg tatgccgcgt aatttccagc agttcttcgg tgatgattct     240
ccgttctgcc aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc     300
ggtaatggtg gcggccagca acagaaattc atggcgctgg gttccggcgt catcatgat      360
gccgataaag gctatgtcgt caccaacaac cacgttggtg ataacgcgac ggtcattaaa     420
gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct     480
gatatcgcgc tgatccaaat ccagaaccgg aaaaacctga ccgcaattaa gatggcggat     540
tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaaccctgt tggctctgggc     600
gagacggtaa cttccgggat tgtctctcgc ctggggcgta gcggcctgaa tgccgaaaac     660
tacgaaaact tcattccagc cgatgcagcg atcaaccgtg gtaactccgg tggtgcgctg     720
gttaacctga acggcgaact gatcggatc aacaccgca tcctcgacc ggacggcggc     780
aacatcggtg tcggttttgc tatcccagat aacatggtga aaaacctgac ctcgcagatg     840
gtggaatacg gccaggtgaa acgcggtgag ctgggtatta tggggactga gctgaaactcc     900
gaaactggca aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg     960
cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac    1020
ggtaagccga tcagcagctt tgccgcactg cgtgctcagg tgggtactat gccggtaggc    1080
agcaaaactg ccctgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg    1140
cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaagcgct    1200
gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcaact    1260
ccggctgcgc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggca    1320
gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc    1380
aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa      1425
```

<210> 8

<211> 474

<212> PRT

<213> Escherichia coli

<400> 8

```
Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly
1           5           10           15
```

```
Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala
```

20 25 30

Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val
35 40 45

Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn
50 55 60

Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser
65 70 75 80

Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln
85 90 95

Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala
100 105 110

Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr
115 120 125

Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser
130 135 140

Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser
145 150 155 160

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile
165 170 175

Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala
180 185 190

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

Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe
210 215 220

Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ser Gly Gly Ala Leu
225 230 235 240

Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala
245 250 255

Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met
260 265 270

Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg
275 280 285

Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys
290 295 300

Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu
305 310 315 320

Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile
325 330 335

Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala
340 345 350

Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu
355 360 365

Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser
370 375 380

Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala
385 390 395 400

Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val
405 410 415

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val
420 425 430

Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg
435 440 445

Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg
450 455 460

Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln
465 470

<210> 9

<211> 1425

<212> DNA

<213> Sequenza Artificiale

<220>

<223> DegP mutato

<400> 9

atgaaaaaaa ccacattagc actgagtgc ctggctctga gtttaggttt ggcggtatct

60

ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaagc 120
cttgaccgga tgctcgaaaa ggtgatgcct tcagtgggtca gcattaacgt agaaggtagc 180
acaaccgtta atacgccgcg tatgccgctg aatttccagc agttcttcgg tgatgattct 240
ccgttctgcc aggaagggtc tccgttccag agctctccgt tctgccaggg tggccagggc 300
ggtaatggtg gcggccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360
gccgataaag gctatgtcgt caccaacaac cacgttggtg ataacgcgac ggtcattaaa 420
gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480
gatatcgcgc tgatccaaat ccagaaccgg aaaaacctga ccgcaattaa gatggcggat 540
tctgatgcac tgcgctggg tgattacacc gtagcgttg gtaaccctgt tggctctggc 600
gagacggtaa cttccgggat tgtctctgcg ctggggcgta gcggcctgaa tgccgaaaac 660
tacgaaaact tcattccagac cgatgcagcg attaactcgtg gtaacgccgg tgggtgcgctg 720
gttaacctga acggcgaact gatcggatc aacaccgca tcctcgcacc ggacggcggc 780
aacatcggtg tcggttttgc tatcccgagt aacatggtga aaaaacctgac ctgcgagatg 840
gtggaatacg gccagggtgaa acgcggtgag ctgggtatta tggggactga gctgaactcc 900
gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960
cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac 1020
ggtaagccga tcagcagctt tgccgcactg cgtgctcagg tgggtactat gccggtaggc 1080
agcaaaactga ccttgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg 1140
cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaagcgct 1200
gagatgagca acaaggcaaa agatcagggc gtggtagtga acaacgtgaa aacgggcact 1260
ccggctgcgc agatcggcct gaagaaaggat gatgtgatta ttggcgcgaa ccagcaggca 1320
gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc 1380
aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

<210> 10

<211> 474

<212> PRT

<213> Sequenza Artificiale

<220>

<223> DegP mutato

<400> 10

Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly
1 5 10 15

Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala
20 25 30

Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val
35 40 45

Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn
50 55 60

Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser
65 70 75 80

Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln
85 90 95

Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala
100 105 110

Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr
115 120 125

Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser
130 135 140

Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser
145 150 155 160

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile
165 170 175

Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala
180 185 190

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

Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe
210 215 220

Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ala Gly Gly Ala Leu
225 230 235 240

Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala
245 250 255

Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met
260 265 270

Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg

275 280 285

Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys
290 295 300

Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu
305 310 315 320

Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile
325 330 335

Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala
340 345 350

Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu
355 360 365

Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser
370 375 380

Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala
385 390 395 400

Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val
405 410 415

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val
420 425 430

Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg
435 440 445

Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg
450 455 460

Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln
465 470

<210> 11
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> primer oligonucleotidico 5' per la regione del gene DegP mutato comprendente il sito di
restrizione Ase I
<400> 11
ctgcctgcga ttttcgccgg aacg 24

<210> 12
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> 3' primer oligonucleotidico 3' per la regione del gene DegP mutato comprendente il sito di restrizione Ase I
<400> 12
cgcatggtac gtgccacgat atcc 24
<210> 13
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> Primer oligonucleotidico 5' primer per la regione del gene mutato DegP comprendente il sito di restrizione Ase I
<400> 13
gggaaatgaa cctgagcaaa acgc 24
<210> 14
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> primer oligonucleotidico 3' per la regione del gene mutato Proteasi III comprendente il sito di restrizione Ase I
<400> 14
gggaaaggcg gcggaaccgc ctacg 24
<210> 15
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> primer oligonucleotidico 5' per la regione del gene mutato Proteasi III comprendente il sito di restrizione Ase I
<400> 15
ctactgtgcc agcggtgta atgg 24
<210> 16
<211> 24
<212> DNA
<213> Sequenza Artificiale
<220>
<223> primer oligonucleotidico 3' per la regione del gene mutato DegP comprendente il sito di restrizione Ase I
<400> 16
gcatcataat ttcttttta cctc 24
<210> 17
<211> 564
<212> DNA
<213> Sequenza Artificiale
<220>
<223> gene spr selvatico

<400> 17
atggtcaaat ctcaaccgat tttgagatat atcttgcgcg ggattcccgc gattgcagta 60
gcggttctgc tttctgcatg tagtgcaaat aacaccgcaa agaatatgca tcctgagaca 120
cgtgcagtgg gtagtgaaac atcatcactg caagcttctc aggatgaatt tgaaaacctg 180
gttcgtaatg tcgacgtaaa atcgcgaatt atggatcagt atgctgactg gaaagcgta 240
cgttatcgtc tgggcggcag cactaaaaaa ggtatcgatt gttctggttt cgtacagcgt 300
acattccgtg agcaatttgg cttagaactt ccgcgttcga cttacgaaca gcaggaaatg 360
ggtaaatctg tttcccgtag taatttgcgt acgggtgatt tagttctggt ccgtgccggt 420
tcaacgggac gccatgctcg tatttataatc ggcaacaatc agtttgtcca tgcttcacc 480
agcagtgggtg ttattatttc cagcatgaat gaaccgtact ggaagaagcg ttacaacgaa 540
gcacgccggg ttctcagccg cagc 564

<210> 18

<211> 188

<212> PRT

<213> Escherichia coli

<400> 18

Met Val Lys Ser Gln Pro Ile Leu Arg Tyr Ile Leu Arg Gly Ile Pro
1 5 10 15

Ala Ile Ala Val Ala Val Leu Leu Ser Ala Cys Ser Ala Asn Asn Thr
20 25 30

Ala Lys Asn Met His Pro Glu Thr Arg Ala Val Gly Ser Glu Thr Ser
35 40 45

Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu Asn Leu Val Arg Asn Val
50 55 60

Asp Val Lys Ser Arg Ile Met Asp Gln Tyr Ala Asp Trp Lys Gly Val
65 70 75 80

Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys Gly Ile Asp Cys Ser Gly
85 90 95

Phe Val Gln Arg Thr Phe Arg Glu Gln Phe Gly Leu Glu Leu Pro Arg
100 105 110

Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys Ser Val Ser Arg Ser Asn
115 120 125

Leu Arg Thr Gly Asp Leu Val Leu Phe Arg Ala Gly Ser Thr Gly Arg
130 135 140

His Val Gly Ile Tyr Ile Gly Asn Asn Gln Phe Val His Ala Ser Thr
145 150 155 160

Ser Ser Gly Val Ile Ile Ser Ser Met Asn Glu Pro Tyr Trp Lys Lys
165 170 175

Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser Arg Ser
180 185

<210> 19
<211> 162
<212> PRT
<213> Escherichia coli
<400> 19
Cys Ser Ala Asn Asn Thr Ala Lys Asn Met His Pro Glu Thr Arg Ala
1 5 10 15

Val Gly Ser Glu Thr Ser Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu
20 25 30

Asn Leu Val Arg Asn Val Asp Val Lys Ser Arg Ile Met Asp Gln Tyr
35 40 45

Ala Asp Trp Lys Gly Val Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys
50 55 60

Gly Ile Asp Cys Ser Gly Phe Val Gln Arg Thr Phe Arg Glu Gln Phe
65 70 75 80

Gly Leu Glu Leu Pro Arg Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys
85 90 95

Ser Val Ser Arg Ser Asn Leu Arg Thr Gly Asp Leu Val Leu Phe Arg
100 105 110

Ala Gly Ser Thr Gly Arg His Val Gly Ile Tyr Ile Gly Asn Asn Gln
115 120 125

Phe Val His Ala Ser Thr Ser Ser Gly Val Ile Ile Ser Ser Met Asn
130 135 140

Glu Pro Tyr Trp Lys Lys Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser
145 150 155 160

Arg Ser

<210> 20

<211> 954

<212> DNA

<213> Sequenza Artificiale

<220>

<223> sequenza nucleotidica di una sequenza mutata OmpT comprendente le mutazioni D210A e H212A

<400> 20

```
atgcgggcga aacttctggg aatagtcctg acaaccoccta ttgcgatcag ctcttttgct      60
tctaccgaga ctttatcgtt tactcctgac aacataaatg cggacattag tcttggaaact      120
ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt      180
caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg      240
atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat      300
atggctcgatc aggactggat ggattccagt aaccccgaa cctggacgga tgaaagtaga      360
caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc      420
ctcaacgaac ccaattaccg cctgggactc atggccgat atcaggaaag cgttatagc      480
tttacagcca gaggtggttc ctatatctac agttctgagg agggattcag agatgatatc      540
ggctccttc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac      600
attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac      660
agcggctggg tggaatcatc tgataacgct gaagcttatg acccgggaaa aagaatcact      720
tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac      780
gtcacaccta acgcaaaagt ttatggtgaa ggcgcatgga atcgggttac gaataaaaaa      840
ggtaataactt cactttatga tcacaataat aacacttcag actacagcaa aatggagca      900
ggtatagaaa actataactt catcactact gctggtotta agtacacatt ttaa      954
```

<210> 21

<211> 317

<212> PRT

<213> Sequenza Artificiale

<220>

<223> sequenza OmpT mutata comprendente le mutazioni D210A e H212A

<400> 21

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

Ser Ser Phe Ala Ser Thr Glu Thr Leu Ser Phe Thr Pro Asp Asn Ile
20 25 30

Asn Ala Asp Ile Ser Leu Gly Thr Leu Ser Gly Lys Thr Lys Glu Arg
35 40 45

Val Tyr Leu Ala Glu Glu Gly Gly Arg Lys Val Ser Gln Leu Asp Trp
50 55 60

Lys Phe Asn Asn Ala Ala Ile Ile Lys Gly Ala Ile Asn Trp Asp Leu
65 70 75 80

Met Pro Gln Ile Ser Ile Gly Ala Ala Gly Trp Thr Thr Leu Gly Ser
85 90 95

Arg Gly Gly Asn Met Val Asp Gln Asp Trp Met Asp Ser Ser Asn Pro
100 105 110

Gly Thr Trp Thr Asp Glu Ser Arg His Pro Asp Thr Gln Leu Asn Tyr
115 120 125

Ala Asn Glu Phe Asp Leu Asn Ile Lys Gly Trp Leu Leu Asn Glu Pro
130 135 140

Asn Tyr Arg Leu Gly Leu Met Ala Gly Tyr Gln Glu Ser Arg Tyr Ser
145 150 155 160

Phe Thr Ala Arg Gly Gly Ser Tyr Ile Tyr Ser Ser Glu Glu Gly Phe
165 170 175

Arg Asp Asp Ile Gly Ser Phe Pro Asn Gly Glu Arg Ala Ile Gly Tyr
180 185 190

Lys Gln Arg Phe Lys Met Pro Tyr Ile Gly Leu Thr Gly Ser Tyr Arg
195 200 205

Tyr Glu Asp Phe Glu Leu Gly Gly Thr Phe Lys Tyr Ser Gly Trp Val
210 215 220

Glu Ser Ser Asp Asn Ala Glu Ala Tyr Asp Pro Gly Lys Arg Ile Thr
225 230 235 240

Tyr Arg Ser Lys Val Lys Asp Gln Asn Tyr Tyr Ser Val Ala Val Asn
245 250 255

Ala Gly Tyr Tyr Val Thr Pro Asn Ala Lys Val Tyr Val Glu Gly Ala
260 265 270

Trp Asn Arg Val Thr Asn Lys Lys Gly Asn Thr Ser Leu Tyr Asp His
275 280 285

Asn Asn Asn Thr Ser Asp Tyr Ser Lys Asn Gly Ala Gly Ile Glu Asn
290 295 300

Tyr Asn Phe Ile Thr Thr Ala Gly Leu Lys Tyr Thr Phe
305 310 315

<210> 22

<211> 954

<212> DNA

<213> Sequenza Artificiale

<220>

<223> sequenza mutata knockout OmpT

<400> 22

attcggggcga aacttctggg aatagtcctg acaaccoccta ttgcgatcag ctcttttgc 60

tctaccgaga ctttatcggt tactcctgac aacataaatg cggacattag tcttggaaact 120

ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180

caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240

atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300

atggctgatc aggactggat ggattccagt aaccccgga cctggacgga tgaagtaga 360

cacctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420

ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccggtatagc 480

tttacagcca gaggtgggtc ctatatctac agttctgagg agggattcag agatgatatc 540

ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 600

attggcttga ctggaagtta tcggtatgaa gattttgaac tcggtggcac atttaaatac 660

agcggctggg tggaatcatc tgataacgat gaacactatg acccgggaaa aagaatcact 720

tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780

gtcacaccta acgcaaaaagt ttatggtgaa ggcgcatgga atcgggttac gaataaaaaa 840

ggtaataactt cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900

ggtatagaaa actataactt catcactact gctggtotta agtacacatt ttaa 954

<210> 23

<211> 45

<212> DNA
<213> Sequenza Artificiale
<220>
<223> adattatore oligonucleotidico OmpA
<400> 23
cgattgaatg gagaactga atcgggcga aactctggg aatag 45
<210> 24
<211> 33
<212> DNA
<213> Sequenza Artificiale
<220>
<223> cassetta oligonucleotidica codificante la sequenza intergenica 1 (IGS1) per l'espressione di Fab
in E. coli
<400> 24
aagttttaat agaggagagt gttaatgaag aag 33
<210> 25
<211> 35
<212> DNA
<213> Sequenza Artificiale
<220>
<223> cassetta oligonucleotidica codificante la sequenza intergenica 2 (IGS2) per l'espressione di Fab
in E. coli
,
<400> 25
aagttttaat agaggggagt gttaaatga agaag 35
<210> 26
<211> 47
<212> DNA
<213> Sequenza Artificiale
<220>
<223> cassetta oligonucleotidica codificante la sequenza intergenica 3 (IGS3) per l'espressione di Fab
in E. coli

<400> 26
aagctttaat agaggagagt gttgaggagg aaaaaaaaaat gaagaaa 47
<210> 27
<211> 47
<212> DNA
<213> Sequenza Artificiale
<220>
<223> cassetta oligonucleotidica codificante la sequenza intergenica 4 (IGS4) per l'espressione di Fab
in E. coli

<400> 27
aagctttaat agaggagagt gttgacgagg attatataat gaagaaa 47
<210> 28
<211> 810
<212> DNA
<213> Escherichia coli

<400> 28
atgaaatcac tgtttaaagt aacgctgctg gcgaccacaa tggccgttgc cctgcatgca 60
ccaatcactt ttgctgctga agctgcaaaa cctgctacag ctgctgacag caaagcagcg 120
ttcaaaaatg acgatcagaa atcagcttat gcaactgggtg cctcgctggg tcgttacatg 180
gaaaactctc taaaagaaca agaaaaactg ggcatcaaac tggataaaga tcagctgacg 240
gctggtgttc aggatgcatt tgctgataag agcaaaactct ccgaccaaga gatcgaacag 300
actctacaag cattcgaagc tcgctggaag tcttctgctc aggcgaagat ggaaaaagac 360
gctgctgata acgaagcaaa aggtaaagag taccgcgaga aatttgccaa agagaaaagg 420
gtgaaaaact cttcaactgg tctggtttat caggtagtag aagccggtaa aggcgaagca 480
ccgaaaagaca gcgatactgt tgtagtgaac taaaaaggta cgctgatcga cggtaaagag 540
ttcgacaact cttacaccgg tggggaaccg ctttctttcc gtctggacgg tggtatcccg 600
ggttggacag aaggtctgaa gaacatcaag aaaggcggta agatcaaact ggttattcca 660
ccagaactgg cttacggcaa agcgggtggt ccggggatcc caccgaattc taccctgggtg 720
tttgacgtag agctgctgga tgtgaaacca ggcgccgaagg ctgatgcaaa gccggaagct 780
gatgcgaaag ccgcagattc tgctaaaaaa 810

<210> 29

<211> 270

<212> PRT

<213> Escherichia coli

<400> 29

Met Lys Ser Leu Phe Lys Val Thr Leu Leu Ala Thr Thr Met Ala Val
1 5 10 15

Ala Leu His Ala Pro Ile Thr Phe Ala Ala Glu Ala Ala Lys Pro Ala
20 25 30

Thr Ala Ala Asp Ser Lys Ala Ala Phe Lys Asn Asp Asp Gln Lys Ser
35 40 45

Ala Tyr Ala Leu Gly Ala Ser Leu Gly Arg Tyr Met Glu Asn Ser Leu
50 55 60

Lys Glu Gln Glu Lys Leu Gly Ile Lys Leu Asp Lys Asp Gln Leu Ile
65 70 75 80

Ala Gly Val Gln Asp Ala Phe Ala Asp Lys Ser Lys Leu Ser Asp Gln

85 90 95

Glu Ile Glu Gln Thr Leu Gln Ala Phe Glu Ala Arg Val Lys Ser Ser
100 105 110

Ala Gln Ala Lys Met Glu Lys Asp Ala Ala Asp Asn Glu Ala Lys Gly
115 120 125

Lys Glu Tyr Arg Glu Lys Phe Ala Lys Glu Lys Gly Val Lys Thr Ser
130 135 140

Ser Thr Gly Leu Val Tyr Gln Val Val Glu Ala Gly Lys Gly Glu Ala
145 150 155 160

Pro Lys Asp Ser Asp Thr Val Val Val Asn Tyr Lys Gly Thr Leu Ile
165 170 175

Asp Gly Lys Glu Phe Asp Asn Ser Tyr Thr Arg Gly Glu Pro Leu Ser
180 185 190

Phe Arg Leu Asp Gly Val Ile Pro Gly Trp Thr Glu Gly Leu Lys Asn
195 200 205

Ile Lys Lys Gly Gly Lys Ile Lys Leu Val Ile Pro Pro Glu Leu Ala
210 215 220

Tyr Gly Lys Ala Gly Val Pro Gly Ile Pro Pro Asn Ser Thr Leu Val
225 230 235 240

Phe Asp Val Glu Leu Leu Asp Val Lys Pro Ala Pro Lys Ala Asp Ala
245 250 255

Lys Pro Glu Ala Asp Ala Lys Ala Ala Asp Ser Ala Lys Lys
260 265 270

<210> 30

<211> 828

<212> DNA

<213> Sequenza Artificiale

<220>

<223> gene FkpA con tag

<400> 30

atgaaatcac tgtttaaagt aacgctgctg gcgaccacaa tggccgttgc cctgcacgca 60

ccaatcactt ttgctgctga agctgcaaaa cctgctactg ctgctgacag caaagcagcg 120

ttcaaaaatg acgatcagaa atcagcttat gcaactgggtg cctogctggg tcggttacatg 180

gaaaactctc taaaagaaca agaaaaactg ggcataaac tgataaaga tcaactgatc 240

gctggtgttc aggatgcatt tgctgataag agcaaactct ccgaccaaga gatcgaacag 300
actctacaag ctttgaagc tcgcgtaag tcttctgctc aggcgaagat ggaaaaagac 360
gcggtctgata acgaagcaaa aggtaaagag taccgcgaga aatttgccaa agagaaaggt 420
gtgaaaacct cttcaactgg tctggtttat caggtagtag aagccggtaa aggcgaagca 480
ccgaaagaca gcgatactgt tgtagtgaac tacaaaggta cgctgatcga cggtaaagag 540
ttcgacaact cttacaccgg tgggtaaccg ctttctttcc gtctggacgg tgttatcccg 600
ggttggacag aaggtctgaa gaacatcaag aaagcggta agataaaact ggttattcca 660
ccagaactgg cttacggcaa agcgggtggt ccggggattc caccaaattc taccctgggtg 720
tttgacgtag agctgctgga tgtgaaacca gcgccgaagg ctgatgcaaa gccggaagct 780
gatgcgaaag ccgcagattc tgctaaaaaa caccatcacc atcaccac 828

<210> 31

<211> 276

<212> PRT

<213> Sequenza Artificiale

<220>

<223> FkpA con his tag

<400> 31

Met Lys Ser Leu Phe Lys Val Thr Leu Leu Ala Thr Thr Met Ala Val
1 5 10 15

Ala Leu His Ala Pro Ile Thr Phe Ala Ala Glu Ala Ala Lys Pro Ala
20 25 30

Thr Ala Ala Asp Ser Lys Ala Ala Phe Lys Asn Asp Asp Gln Lys Ser
35 40 45

Ala Tyr Ala Leu Gly Ala Ser Leu Gly Arg Tyr Met Glu Asn Ser Leu
50 55 60

Lys Glu Gln Glu Lys Leu Gly Ile Lys Leu Asp Lys Asp Gln Leu Ile
65 70 75 80

Ala Gly Val Gln Asp Ala Phe Ala Asp Lys Ser Lys Leu Ser Asp Gln
85 90 95

Glu Ile Glu Gln Thr Leu Gln Ala Phe Glu Ala Arg Val Lys Ser Ser
100 105 110

Ala Gln Ala Lys Met Glu Lys Asp Ala Ala Asp Asn Glu Ala Lys Gly
115 120 125

Lys Glu Tyr Arg Glu Lys Phe Ala Lys Glu Lys Gly Val Lys Thr Ser
130 135 140

Ser Thr Gly Leu Val Tyr Gln Val Val Glu Ala Gly Lys Gly Glu Ala
145 150 155 160

Pro Lys Asp Ser Asp Thr Val Val Val Asn Tyr Lys Gly Thr Leu Ile
165 170 175

Asp Gly Lys Glu Phe Asp Asn Ser Tyr Thr Arg Gly Glu Pro Leu Ser
180 185 190

Phe Arg Leu Asp Gly Val Ile Pro Gly Trp Thr Glu Gly Leu Lys Asn
195 200 205

Ile Lys Lys Gly Gly Lys Ile Lys Leu Val Ile Pro Pro Glu Leu Ala
210 215 220

Tyr Gly Lys Ala Gly Val Pro Gly Ile Pro Pro Asn Ser Thr Leu Val
225 230 235 240

Phe Asp Val Glu Leu Leu Asp Val Lys Pro Ala Pro Lys Ala Asp Ala
245 250 255

Lys Pro Glu Ala Asp Ala Lys Ala Ala Asp Ser Ala Lys Lys His His
260 265 270

His His His His
275

<210> 32

<211> 486

<212> DNA

<213> Escherichia coli

<400> 32

gtgaaaaagt ggttattagc tgcaggtctc ggttagcac tggcaacttc tgctcaggcg 60

gctgacaaaa ttgcaatcgt caacatgggc agcctgttcc agcaggtagc gcagaaaacc 120

ggtgtttcta acacgctgga aatgagttc aaaggccgtg ccagcgaact gcagcgtatg 180

gaaaccgatc tgcaggctaa aatgaaaaag ctgcagtcca tgaaagcggg cagcgatcgc 240

actaagctgg aaaaagacgt gatggctcag cgccagactt ttgctcagaa agcgcaggct 300

tttgagcagg atcgcgcacg tcgttccaac gaagaacgcg gcaaactggt tactcgtatc 360

cagactgctg tgaatatccgt tgccaacagc caggatatcg atctggttgt tgatgcaaac 420

gccgttgctt acaacagcag cgatgtaaaa gacatcactg ccgacgtact gaaacaggtt 480

aaataa
<210> 33
<211> 161
<212> PRT
<213> Escherichia coli
<400> 33
Met Lys Lys Trp Leu Leu Ala Ala Gly Leu Gly Leu Ala Leu Ala Thr
1 5 10 15

Ser Ala Gln Ala Ala Asp Lys Ile Ala Ile Val Asn Met Gly Ser Leu
20 25 30

Phe Gln Gln Val Ala Gln Lys Thr Gly Val Ser Asn Thr Leu Glu Asn
35 40 45

Glu Phe Lys Gly Arg Ala Ser Glu Leu Gln Arg Met Glu Thr Asp Leu
50 55 60

Gln Ala Lys Met Lys Lys Leu Gln Ser Met Lys Ala Gly Ser Asp Arg
65 70 75 80

Thr Lys Leu Glu Lys Asp Val Met Ala Gln Arg Gln Thr Phe Ala Gln
85 90 95

Lys Ala Gln Ala Phe Glu Gln Asp Arg Ala Arg Arg Ser Asn Glu Glu
100 105 110

Arg Gly Lys Leu Val Thr Arg Ile Gln Thr Ala Val Lys Ser Val Ala
115 120 125

Asn Ser Gln Asp Ile Asp Leu Val Val Asp Ala Asn Ala Val Ala Tyr
130 135 140

Asn Ser Ser Asp Val Lys Asp Ile Thr Ala Asp Val Leu Lys Gln Val
145 150 155 160

Lys
<210> 34
<211> 501
<212> DNA
<213> Sequenza Artificiale
<220>
<223> skp con his tag

<400> 34
atgaaaaagt ggttattagc cgcaggtctc ggtttagcac tggcaacttc tgctcaggcg 60
gctgacaaaa ttgcaatcgt caacatgggc agcctgttcc agcaggtagc gcagaaaacc 120
ggtgttttcta acacgctgga aatgagttc aaaggccgtg ccagcgaact ccagcgtatg 180
gaaaccgatc tccaggctaa aatgaaaaag ctgcaatcca tgaagcggg cagcgatcgc 240
actaagctgg aaaaagacgt gatggctcag cgccagactt ttgctcagaa agcgcaggct 300
tttgagcagg atcgcgcacg tcgttccaac gaagaacgcg gcaaactggt tactcgtatc 360
cagactgctg tgaatccgt tgccaacagc caggaaatcg atctggttgt tgatgcaaac 420
gccgttgctt acaacagcag cgatgtaaaa gacatcactg ccgacgtact gaaacaggtt 480
aaacaccatc accatcacca c 501

<210> 35

<211> 167

<212> PRT

<213> Sequenza Artificiale

<220>

<223> skip con his tag

<400> 35

Met Lys Lys Trp Leu Leu Ala Ala Gly Leu Gly Leu Ala Leu Ala Thr
1 5 10 15

Ser Ala Gln Ala Ala Asp Lys Ile Ala Ile Val Asn Met Gly Ser Leu
20 25 30

Phe Gln Gln Val Ala Gln Lys Thr Gly Val Ser Asn Thr Leu Glu Asn
35 40 45

Glu Phe Lys Gly Arg Ala Ser Glu Leu Gln Arg Met Glu Thr Asp Leu
50 55 60

Gln Ala Lys Met Lys Lys Leu Gln Ser Met Lys Ala Gly Ser Asp Arg
65 70 75 80

Thr Lys Leu Glu Lys Asp Val Met Ala Gln Arg Gln Thr Phe Ala Gln
85 90 95

Lys Ala Gln Ala Phe Glu Gln Asp Arg Ala Arg Arg Ser Asn Glu Glu
100 105 110

Arg Gly Lys Leu Val Thr Arg Ile Gln Thr Ala Val Lys Ser Val Ala
115 120 125

Asn Ser Gln Glu Ile Asp Leu Val Val Asp Ala Asn Ala Val Ala Tyr
130 135 140

Asn Ser Ser Asp Val Lys Asp Ile Thr Ala Asp Val Leu Lys Gln Val
145 150 155 160

Lys His His His His His His
165

<210> 36

<211> 10

<212> PRT

<213> Sequenza Artificiale

<220>

<223> CA170 1519 Ab CDRH1

<400> 36

Gly Phe Thr Phe Ser Asn Tyr Gly Met Val
1 5 10

<210> 37

<211> 17

<212> PRT

<213> Sequenza Artificiale

<220>

<223> CA170_1519 Ab CDRH2

<400> 37

Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val Lys
1 5 10 15

Gly

<210> 38

<211> 8

<212> PRT

<213> Sequenza Artificiale

<220>

<223> CA170_1519 Ab CDRH3

<400> 38

Gly Ile Val Arg Pro Phe Leu Tyr
1 5

<210> 39

<211> 16

<212> PRT

<213> Sequenza Artificiale

<220>

<223> CA170_1519 Ab CDRL1

<400> 39

Lys Ser Ser Gln Ser Leu Val Gly Ala Ser Gly Lys Thr Tyr Leu Tyr
1 5 10 15

<210> 40

<211> 7

<212> PRT

<213> Sequenza Artificiale

<220>

<223> CA170_1519 Ab CDRL2

<400> 40
Leu Val Ser Thr Leu Asp Ser
1 5
<210> 41
<211> 9
<212> PRT
<213> Sequenza Artificiale
<220>
<223> CA170_1519 Ab CDRL3
<400> 41
Leu Gln Gly Thr His Phe Pro His Thr
1 5
<210> 42
<211> 112
<212> PRT
<213> Sequenza Artificiale
<220>
<223> regione Ab 1519 VL di ratto
<400> 42
Asp Val Val Met Thr Gln Thr Pro Leu Ser Leu Ser Val Ala Leu Gly
1 5 10 15

Gln Pro Ala Ser Ile Ser Cys Lys Ser Ser Gln Ser Leu Val Gly Ala
20 25 30

Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Arg Ser Gly Gln Ser
35 40 45

Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp Ser Gly Ile Pro
50 55 60

Asp Arg Phe Ser Gly Ser Gly Ala Glu Thr Asp Phe Thr Leu Lys Ile
65 70 75 80

Arg Arg Val Glu Ala Asp Asp Leu Gly Val Tyr Tyr Cys Leu Gln Gly
85 90 95

Thr His Phe Pro His Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
100 105 110
<210> 43
<211> 336
<212> DNA
<213> Sequenza Artificiale
<220>
<223> regione Ab 1519 VL di ratto

<400> 43
gatgttgga tgaccagac tccactgtct ttgtcggtg cccttgaca accagcctcc 60
atctcttga agtcaagtca gaggctcgtg ggtgctagt gaaagacata tttgtattgg 120
ttatttcaga ggtccggcca gtctccaaag cgactaatct atctggtgtc cacactggac 180
tctggaattc ctgataggtt cagtggcagt ggagcagaga cagattttac tcttaaaatc 240
cgcagagtgg aagccgatga tttgggagtt tattactgct tgcaaggtac acattttcct 300

cacacgtttg gagctgggac caagctggaa ttgaaa 336

<210> 44

<211> 132

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione Ab 1519 VL di ratto

<400> 44

Met Met Ser Pro Ala Gln Phe Leu Phe Leu Leu Met Leu Trp Ile Gln
1 5 10 15

Gly Thr Ser Gly Asp Val Val Met Thr Gln Thr Pro Leu Ser Leu Ser
20 25 30

Val Ala Leu Gly Gln Pro Ala Ser Ile Ser Cys Lys Ser Ser Gln Ser
35 40 45

Leu Val Gly Ala Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Arg
50 55 60

Ser Gly Gln Ser Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp
65 70 75 80

Ser Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ala Glu Thr Asp Phe
85 90 95

Thr Leu Lys Ile Arg Arg Val Glu Ala Asp Asp Leu Gly Val Tyr Tyr
100 105 110

Cys Leu Gln Gly Thr His Phe Pro His Thr Phe Gly Ala Gly Thr Lys
115 120 125

Leu Glu Leu Lys
130

<210> 45

<211> 396

<212> DNA

<213> Sequenza Artificiale

<220>
<223> regione Ab 1519 VL di ratto
<400> 45
atgatgagtc ctgcccagtt cctgtttctg ctgatgctct ggattcaggg aaccagtgg 60
gatgttgga tgaccagac tccactgtct ttgtcggtg cccttgaca accagcctcc 120
atctcttgca agtcaagtca ggcctcgtg ggtgctagt gaaagacata tttgtattgg 180
ttatttcaga ggtccggcca gtctccaaag cgactaatct atctgggtgc cacactggac 240
tctggaattc ctgataggtt cagtggcagt ggagcagaga cagattttac tcttaaaatc 300
cgcagagtgg aagccgatga tttgggagtt tattactgct tgcaaggtac acattttcct 360
cacacgtttg gagctgggac caagctggaa ttgaaa 396

<210> 46

<211> 116

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione Ab 1519 VH di ratto

<400> 46

Glu Val Pro Leu Val Glu Ser Gly Gly Gly Ser Val Gln Pro Gly Arg
1 5 10 15

Ser Met Lys Leu Ser Cys Val Val Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Gly Met Val Trp Val Arg Gln Ala Pro Lys Lys Gly Leu Glu Trp Val
35 40 45

Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asn Asn Ala Lys Ser Thr Leu Tyr
65 70 75 80

Leu Gln Met Asp Ser Leu Arg Ser Glu Asp Thr Ala Thr Tyr Tyr Cys
85 90 95

Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln Gly Thr Thr
100 105 110

Val Thr Val Ser
115

<210> 47

<211> 348

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione Ab 1519 VH di ratto

<400> 47
gagggtgccgc tgggtggagtc tggggggcggc tcagtgacgc ctggggaggtc catgaaactc 60
tcctgtgtag tctcaggatt cactttcagt aattatggca tggctctgggt ccgccaggct 120
ccaaagaagg gtctggagtg ggtcgcatat attgattctg atggtgataa tacttactac 180
cgagattccg tgaagggccg attcactatc tccagaaata atgc metaag caccctatat 240
ttgcaaatgg acagtctgag gtctgaggac acggccactt attactgtac aacagggatt 300
gtccggccct ttctctattg gggccaagga accacgggtca ccgtctcg 348

<210> 48

<211> 135

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione Ab 1519 VH di ratto

<400> 48

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

Val Arg Cys Glu Val Pro Leu Val Glu Ser Gly Gly Gly Ser Val Gln
20 25 30

Pro Gly Arg Ser Met Lys Leu Ser Cys Val Val Ser Gly Phe Thr Phe
35 40 45

Ser Asn Tyr Gly Met Val Trp Val Arg Gln Ala Pro Lys Lys Gly Leu
50 55 60

Glu Trp Val Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg
65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asn Asn Ala Lys Ser
85 90 95

Thr Leu Tyr Leu Gln Met Asp Ser Leu Arg Ser Glu Asp Thr Ala Thr
100 105 110

Tyr Tyr Cys Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln
115 120 125

Gly Thr Thr Val Thr Val Ser
130 135

<210> 49

<211> 405

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione Ab 1519 VH di ratto

<400> 49

```
atggacatca gtctcagctt ggcttttcctt gtccttttca taaaagggtgt cgggtgtgag      60
gtgccgctgg tggagtctgg gggcggtca gtgcagcctg ggaggtccat gaaactctcc      120
tgtgtagtct caggattcac tttcagtaat tatggcatgg tctgggtccg ccaggctcca      180
aagaagggtc tggagtgggt cgcataatatt gattctgatg gtgataatac ttactaccga      240
gattccgtga agggccgatt cactatctcc agaaataatg caaaaagcac cctatatttg      300
caaatggaca gtctgaggtc tgaggacacg gccacttatt actgtacaac agggattgtc      360
cgccctttc tctattgggg ccaaggaacc acggtcaccg tctcg                          405
```

<210> 50

<211> 112

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione 1519 gL20 V

<400> 50

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

Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser Leu Val Gly Ala
          20          25          30

Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Lys Pro Gly Lys Ala
          35          40          45

Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp Ser Gly Ile Pro
          50          55          60

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile
65          70          75          80

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Gly
          85          90          95

Thr His Phe Pro His Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys
          100         105         110
```

<210> 51

<211> 336

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione 1519 gL20 V

<400> 51
gatatccaga tgaccagag tccaagcagt ctctccgcca gcgtaggcga tcgtgtgact 60
attacctgta aaagctccca gtccctgggtg ggtgcaagcg gcaaaccta cctgtactgg 120
ctcttccaga aaccgggcaa agctccgaaa cgcctgatct atctgggtgtc taccctggat 180
agcggatttc cgtctcgttt ctccggtagc ggtagcggta ccgaattcac gctgaccatt 240
agctccctcc agccggagga ctttgctacc tattactgcc tccagggcac tcattttccg 300
cacactttcg gccagggtag caaactggaa atcaaa 336

<210> 52

<211> 133

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione 1519 gL20 V

<400> 52

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

Thr Val Ala Gln Ala Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
20 25 30

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln
35 40 45

Ser Leu Val Gly Ala Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln
50 55 60

Lys Pro Gly Lys Ala Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu
65 70 75 80

Asp Ser Gly Ile Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu
85 90 95

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
100 105 110

Tyr Cys Leu Gln Gly Thr His Phe Pro His Thr Phe Gly Gln Gly Thr
115 120 125

Lys Leu Glu Ile Lys
130

<210> 53

<211> 399

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione 1519 gL20 V

<400> 53

```
atgaaaaaga cagctatcgc aattgcagtg gccttggctg gtttcgctac cgtagcgcaa      60
gctgatatcc agatgaccca gagtccaagc agtctctccg ccagcgtagg cgatcgtgtg      120
actattacct gtaaaagctc ccagtcctcg gtgggtgcaa gcggcaaac ctacctgtac      180
tggctcttcc agaaaccggg caaagctccg aaacgcctga tctatctggt gtctaccctg      240
gatagcggta ttccgtctcg tttctccggt agcggtagcg gtaccgaatt cacgctgacc      300
attagctccc tccagccgga ggactttgct acctattact gcctccaggg cactcatttt      360
ccgcacactt tcggccaggg taccaaactg gaaatcaaa      399
```

<210> 54

<211> 219

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519 gL20 (V + costante)

<400> 54

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

Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser Leu Val Gly Ala
20 25 30

Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Lys Pro Gly Lys Ala
35 40 45

Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp Ser Gly Ile Pro
50 55 60

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile
65 70 75 80

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Gly
85 90 95

Thr His Phe Pro His Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys
100 105 110

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
115 120 125

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
130 135 140

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
145 150 155 160

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
165 170 175

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
180 185 190

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
195 200 205

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
210 215

<210> 55

<211> 657

<212> DNA

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519 gL20 (V + costante)

<400> 55
gatatccaga tgaccagag tccaagcagt ctctccgcca gcgtaggcga tcgtgtgact 60
attacctgta aaagctccca gtccctggtg ggtgcaagcg gcaaaccta cctgtactgg 120
ctcttccaga aaccgggcaa agctccgaaa cgcctgatct atctggtgtc taccctggat 180
agcggatttc cgtctcgttt ctccggtagc ggtagcggta ccgaattcac gctgaccatt 240
agctccctcc agccggagga ctttgctacc tattactgcc tccagggcac tcattttcog 300
cacactttcg gccagggtac caaactggaa atcaaacgta cggtagcggc cccatctgtc 360
ttcatcttcc cgccatctga tgagcagttg aaatctggaa ctgcctctgt tgtgtgcctg 420
ctgaataact tctatcccag agaggccaaa gtacagtgga aggtggataa cgccctccaa 480
tcgggtaact cccaggagag tgtcacagag caggacagca aggacagcac ctacagcctc 540
agcagcacc tgacgctgag caaagcagac tacgagaaac acaaagtcta cgcctgcgaa 600
gtcaccatc agggcctgag ctcaccagta acaaaaagtt ttaatagagg ggagtgt 657

<210> 56

<211> 240

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519 gL20

<400> 56

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

Thr Val Ala Gln Ala Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
20 25 30

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln
35 40 45

Ser Leu Val Gly Ala Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln
50 55 60

Lys Pro Gly Lys Ala Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu
65 70 75 80

Asp Ser Gly Ile Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu
85 90 95

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
100 105 110

Tyr Cys Leu Gln Gly Thr His Phe Pro His Thr Phe Gly Gln Gly Thr
115 120 125

Lys Leu Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe
130 135 140

Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys
145 150 155 160

Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val
165 170 175

Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln
180 185 190

Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser
195 200 205

Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His
210 215 220

Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225 230 235 240

<210> 57

<211> 720

<212> DNA

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519 gL20

<400> 57

atgaaaaaga cagctatcgc aattgcagtg gccttggtcg gtttcgctac cgtagcgcgca 60

gctgatatcc agatgaccca gagtccaagc agtctctccg ccagcgtagg cgatcgtgtg 120

actattacct gtaaaagctc ccagtcctcg gtgggtgcaa gcggcaaac ctacctgtac 180

tggctcttcc agaaaaccggg caaagctccg aaacgcctga tctatctggt gtctaccctg 240

gatagcggta ttccgtctcg tttctccggt agcggtagcg gtaccgaatt cacgctgacc 300

attagctccc tccagccgga ggactttgct acctattact gcctccaggg cactcatttt 360

ccgcacactt tcggccaggg taccaaactg gaaatcaaac gtacggtagc ggccccatct 420

gtcttcatct tcccgccatc tgatgagcag ttgaaatctg gaactgcctc tgttgtgtgc 480

ctgctgaata acttctatcc cagagaggcc aaagtacagt ggaaggtgga taacgccctc 540

caatcgggta actcccagga gagtgtcaca gagcaggaca gcaaggacag cacctacagc 600

ctcagcagca cctgacgct gagcaaagca gactacgaga aacacaaagt ctacgcctgc 660

gaagtcaccc atcagggcct gagctcacca gtaacaaaaa gttttaatag aggggagtg 720

<210> 58

<211> 116

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione 1519 gH20 V

<400> 58

Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Ser Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln Gly Thr Leu
100 105 110

Val Thr Val Ser
115

<210> 59

<211> 348

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione 1519 gH20 V

<400> 59

gagggtccgc tggtcgagtc tggaggcggg cttgtccagc ctggaggag cctgcgtctc 60

tcttgtgcag tatctggctt cacgttctcc aactacggta tgggtgtgggt tcgtcaggct 120

ccaggtaaag gtctggaatg ggtggcgtat attgactccg acggcgacaa cacctactat 180

cgcgactctg tgaaggctcg cttcaccatt tcccgcgata acgccaatc cagcctgtac 240

ctgcagatga acagcctgcg tgctgaagat actgcggtgt actattgcac cactggcatc 300

gtgcgtccgt ttctgtattg gggtcagggt accctcgta ctgtctcg 348

<210> 60

<211> 137

<212> PRT

<213> Sequenza Artificiale

<220>

<223> regione 1519 gH20 V

<400> 60

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

Thr Val Ala Gln Ala Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu
20 25 30

Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe
35 40 45

Thr Phe Ser Asn Tyr Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys
50 55 60

Gly Leu Glu Trp Val Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr
65 70 75 80

Tyr Arg Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala
85 90 95

Lys Ser Ser Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110

Ala Val Tyr Tyr Cys Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp
115 120 125

Gly Gln Gly Thr Leu Val Thr Val Ser
130 135

<210> 61

<211> 411

<212> DNA

<213> Sequenza Artificiale

<220>

<223> regione 1519 gH20 V

<400> 61

atgaagaaga ctgctatagc aattgcagtg gcgctagctg gtttcgccac cgtggcgcaa 60
gctgaggttc cgctggtcga gtctggaggc gggcttgctc agcctggagg gagcctgcgt 120
ctctcttgctg cagtatctgg cttcacgttc tccaactacg gtatggtgtg ggttcgtcag 180
gctccaggta aaggctctgga atgggtggcg tatattgact ccgacggcga caacacctac 240
tategcgact ctgtgaaagg tcgcttcacc atttcccgcg ataacgcaa atccagcctg 300
tacctgcaga tgaacagcct gcgtgctgaa gatactgcgg tgtactattg caccactggc 360
atcgtgcgctc cgtttctgta ttggggctcag ggtaccctcg ttactgtctc g 411

<210> 62

<211> 228

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 Fab' (V + gamma-1 CH1 umano + hinge)

<400> 62

Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Ser Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln Gly Thr Leu
100 105 110

Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
115 120 125

Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys
130 135 140

Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser
145 150 155 160

Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser
165 170 175

Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser
180 185 190

Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn
195 200 205

Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
210 215 220

Thr Cys Ala Ala
225

<210> 63

<211> 684

<212> DNA

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 Fab' (V + gamma umano-1 CH1 + hinge)

<400> 63

gagggtccgc tggtcgagtc tggaggcggg cttgtccagc ctggaggag cctgcgtctc 60

tcttgtgcag tatctggctt cacgttctcc aactacggta tgggtgtgggt tcgtcaggct 120

ccaggtaaag gtctggaatg ggtggcgtat attgactccg acggcgacaa cacctactat 180

cgcgactctg tгааaggctg cttcaccatt tcccgcgata acgccaatc cagcctgtac 240

ctgcagatga acagcctgcg tgctgaagat actgcgggtg actattgcac cactggcatc 300

gtgcgtccgt ttctgtattg gggtcagggt accctcgta ctgtctcgag cgcttctaca 360

aagggcccat cggtcttccc cctggcacc cctccaaga gcacctctgg gggcacagcg 420

gccctgggct gcctgggtcaa ggactacttc cccgaaccgg tgacgggtgc gtggaactca 480

ggcgccctga ccagcggcgt gcacaccttc ccggctgtcc tacagtcctc aggactctac 540

tccctcagca gcgtgggtgac cgtgccctcc agcagcttgg gcaccagac ctacatctgc 600

aacgtgaatc acaagcccag caacaccaag gtcgacaaga aagttgagcc caaatcttgt 660

gacaaaactc acacatgccc cgcg 684

<210> 64

<211> 249

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519 gH20 Fab'

<400> 64

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

Thr Val Ala Gln Ala Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu
20 25 30

Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe
35 40 45

Thr Phe Ser Asn Tyr Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys
50 55 60

Gly Leu Glu Trp Val Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr
65 70 75 80

Tyr Arg Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala
85 90 95

Lys Ser Ser Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110

Ala Val Tyr Tyr Cys Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp
115 120 125

Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro
130 135 140

Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr
145 150 155 160

Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr
165 170 175

Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro
180 185 190

Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr
195 200 205

Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn
210 215 220

His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser
225 230 235 240

Cys Asp Lys Thr His Thr Cys Ala Ala
245

<210> 65
<211> 747

<212> DNA
<213> Sequenza Artificiale
<220>
<223> catena pesante 1519 gH20 Fab'
<400> 65
atgaagaaga ctgctatagc aattgcagtg gcgctagctg gtttcgccac cgtggcgcaa 60
gctgaggttc cgctggtcga gtctggaggc gggcttgtcc agcctggagg gagcctgcgt 120
ctctcttggt cagtatctgg cttcacgttc tccaactacg gtatggtgtg ggttcgtcag 180
gctccaggta aaggctctgga atgggtggcg tatattgact ccgacggcga caacacctac 240
tatcgcgact ctgtgaaagg tcgcttcacc atttcccgcg ataacgcaa atccagcctg 300
tacctgcaga tgaacagcct gcgtgctgaa gatactgcgg tgtactattg caccactggc 360
atcgtgcgtc cgtttctgta ttggggtcag ggtaccctcg ttactgtctc gagcgttct 420
acaaagggcc catcggctctt ccccctggca ccctctcca agagcacctc tgggggcaca 480
gcggccctgg gctgcctggt caaggactac ttccccgaac cggtgacggt gtcgtggaac 540
tcaggcgccc tgaccagcgg cgtgcacacc ttcccggctg tcctacagtc ctcaggactc 600
tactccctca gcagcgtggt gaccgtgccc tccagcagct tgggcacca gacctacatc 660
tgcaacgtga atcacaagcc cagcaacacc aaggtcgaca agaaagttga gcccaaatct 720
tgtgacaaaa ctcacacatg cgccgcg 747

<210> 66

<211> 444

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 IgG4 (V + gamma umano-4P costante)

<400> 66

Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Ser Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln Gly Thr Leu
100 105 110

Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
115 120 125

Ala Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys
130 135 140

Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser
145 150 155 160

Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser
165 170 175

Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser
180 185 190

Leu Gly Thr Lys Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn
195 200 205

Thr Lys Val Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro
210 215 220

Pro Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe
225 230 235 240

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
245 250 255

Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe
260 265 270

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro
275 280 285

Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr
290 295 300

Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val
305 310 315 320

Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala
325 330 335

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln
340 345 350

Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly
355 360 365

Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
370 375 380

Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
385 390 395 400

Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu
405 410 415

Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
420 425 430

Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys
435 440

<210> 67

<211> 1939

<212> DNA

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 IgG4 (V + gamma umano-4P costante, esoni)

<400> 67

```
gaggtagccac ttgtggaag cggaggaggt cttgtgcagc ctggaggaag tttacgtctc      60
tcttgtgctg tgtctggctt caccttctcc aattacggaa tggctctgggt cagacaagca      120
cctgaaagg gtcttgaatg ggtggcctat attgactctg acggggacaa cacctactat      180
cgggattccg tgaaggagc cttcacaatc tcccgagata acgccaagag ctcactgtac      240
ctgcagatga atagcctgag agccgaggat actgccgtgt actattgcac aacgggaatc      300
gttaggcctt ttctgtactg gggacagggc accttgggta ctgtctcgag cgcttctaca      360
aagggcccat ccgtcttccc cctggcgccc tgctccagga gcacctccga gagcacagcc      420
gccctgggct gcctgggcaa ggactacttc cccgaaccgg tgacgggtgc gtggaactca      480
ggcgccctga ccagcggcgt gcacacctc ccggctgtcc tacagtctc aggactctac      540
tccctcagca gcgtgggtgac cgtgccctcc agcagcttgg gcacgaagac ctacacctgc      600
aacgtagatc acaagcccag caacaccaag gtggacaaga gagttggtga gaggccagca      660
cagggagggg ggggtgtctgc tggagccag gctcagccct cctgcctgga cgcaccccg      720
```

ctgtgcagcc ccagcccagg gcagcaaggc atgccccatc tgtctcctca cccggaggcc 780
tctgaccacc ccaactcatgc ccagggagag ggtcttctgg atttttccac caggctccgg 840
gcagccacag gctggatgcc cctaccccag gccctgcgca tacaggggca ggtgctgcgc 900
tcagacctgc caagagccat atccgggagg accctgcccc tgacctaagc ccaccccaaa 960
ggcctaaactc tccactccct cagctcagac accttctctc ctcccagatc tgagtaactc 1020
ccaatcttct ctctgcagag tccaaatatg gtccccatg cccaccatgc ccaggtgaagc 1080
caaccaggc ctcgcctcc agctcaaggc gggacagggtg ccctagagta gcctgcatcc 1140
agggacaggc cccagccggg tgctgacgca tccacctcca tctcttcctc agcacctgag 1200
ttcctggggg gaccatcagt ctctctgttc cccccaaaac ccaaggacac tctcatgatc 1260
tcccggacct ctgaggtcac gtgcgtggtg gtggacgtga gccaggaaga ccccagggtc 1320
cagttcaact ggtacgtgga tggcgtggag gtgcataatg ccaagacaaa gccgcgggag 1380
gagcagttca acagcacgta ccgtgtggtc agcgtcctca ccgtcctgca ccaggactgg 1440
ctgaacggca aggagtacaa gtgcaaggtc tccaacaaag gcctcccgtc ctccatcgag 1500
aaaaccatct ccaaagccaa aggtgggacc cacgggggtgc gagggccaca tggacagagg 1560
tcagctcggc ccaccctctg ccctgggagt gaccgctgtg ccaacctctg tccctacagg 1620
gcagccccga gagccacagg tgtacaccct gccccatcc caggaggaga tgaccaagaa 1680
ccaggtcagc ctgacctgcc tggtaaaagg cttctacccc agcgacatcg ccgtggagtg 1740
ggagagcaat gggcagccgg agaacaacta caagaccacg cctcccgtgc tggactccga 1800
cggctccttc ttctctaca gcaggtaac cgtggacaag agcaggtggc aggaggggaa 1860
tgtcttctca tgctccgtga tgcatgaggc tctgcacaac cactacacac agaagagcct 1920
ctccctgtct ctgggtaaa 1939

<210> 68

<211> 347

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519gL20 FabFv

<400> 68

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

Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser Leu Val Gly Ala
20 25 30

Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Lys Pro Gly Lys Ala
35 40 45

Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp Ser Gly Ile Pro
50 55 60

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Ile
65 70 75 80

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Gly
85 90 95

Thr His Phe Pro His Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys
100 105 110

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
115 120 125

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
130 135 140

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
145 150 155 160

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
165 170 175

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
180 185 190

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
195 200 205

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Ser Gly Gly Gly Gly
210 215 220

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr
225 230 235 240

Gln Ser Pro Ser Ser Val Ser Ala Ser Val Gly Asp Arg Val Thr Ile
245 250 255

Thr Cys Gln Ser Ser Pro Ser Val Trp Ser Asn Phe Leu Ser Trp Tyr
260 265 270

Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Glu Ala Ser
275 280 285

Lys Leu Thr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly

290 295 300

Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala
305 310 315 320

Thr Tyr Tyr Cys Gly Gly Gly Tyr Ser Ser Ile Ser Asp Thr Thr Phe
325 330 335

Gly Cys Gly Thr Lys Val Glu Ile Lys Arg Thr
340 345

<210> 69
<211> 1041
<212> DNA
<213> Sequenza Artificiale
<220>
<223> catena leggera 1519gL20 FabFv
<400> 69

gatatccaga	tgaccagag	cccattctagc	ttatccgctt	ccgttggtga	tcgctgaca	60
attacgtgta	agagctccca	atctctcgtg	ggtgcaagtg	gcaagaccta	tctgtactgg	120
ctctttcaga	agcctggcaa	ggcaccaaaa	cggtgatct	atctggtgtc	tacccttgac	180
tctgggatac	cgtcacgatt	ttccggatct	gggagcggaa	ctgagttcac	actcacgatt	240
tcatcgctgc	aaccogagga	ctttgctacc	tactactgcc	tgcaaggcac	tcatttccct	300
cacactttcg	gccaggggac	aaaactcgaa	atcaaacgta	cggtagcggc	cccattctgtc	360
ttcatcttcc	cgccatctga	tgagcagttg	aaatctggaa	ctgcctctgt	tgtgtgcctg	420
ctgaataact	tctatccag	agaggccaaa	gtacagtgga	aggtggataa	cgccctccaa	480
tcgggtaact	cccaggagag	tgtcacagag	caggacagca	aggacagcac	ctacagcctg	540
agcagcacc	tgacgctgtc	taaagcagac	tacgagaaac	acaaagtgt	cgctgcgaa	600
gtcaccatc	agggcctgag	ctcaccagta	acaaaaagtt	ttaatagagg	ggagtgtagc	660
ggtggcggg	gcagtgggtg	gggaggctcc	ggaggtggcg	gttcagacat	acaaatgacc	720
cagagtcctt	catcggatc	cgcgctcgtt	ggcgataggg	tgactattac	atgtcaaagc	780
tctcctagcg	tctggagcaa	ttttctatcc	tggtatcaac	agaaaccggg	gaaggctcca	840
aaacttctga	tttatgaagc	ctcgaaactc	accagtggag	ttccgtcaag	attcagtggc	900
tctggatcag	ggacagactt	cacgttgaca	atcagttcgc	tgcaaccaga	ggactttgcg	960
acctactatt	gtggtggagg	ttacagtagc	ataagtgata	cgacatttgg	gtgcggtact	1020
aaggtgaaa	tcaaactgac	c				1041

<210> 70
<211> 367

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519gL20 FabFv

<400> 70

Met Ser Val Pro Thr Gln Val Leu Gly Leu Leu Leu Leu Trp Leu Thr
1 5 10 15

Asp Ala Arg Cys Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser
20 25 30

Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser
35 40 45

Leu Val Gly Ala Ser Gly Lys Thr Tyr Leu Tyr Trp Leu Phe Gln Lys
50 55 60

Pro Gly Lys Ala Pro Lys Arg Leu Ile Tyr Leu Val Ser Thr Leu Asp
65 70 75 80

Ser Gly Ile Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe
85 90 95

Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr
100 105 110

Cys Leu Gln Gly Thr His Phe Pro His Thr Phe Gly Gln Gly Thr Lys
115 120 125

Leu Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro
130 135 140

Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
145 150 155 160

Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp
165 170 175

Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp
180 185 190

Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys
195 200 205

Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln
210 215 220

Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Ser
225 230 235 240

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp
245 250 255

Ile Gln Met Thr Gln Ser Pro Ser Ser Val Ser Ala Ser Val Gly Asp
260 265 270

Arg Val Thr Ile Thr Cys Gln Ser Ser Pro Ser Val Trp Ser Asn Phe
275 280 285

Leu Ser Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
290 295 300

Tyr Glu Ala Ser Lys Leu Thr Ser Gly Val Pro Ser Arg Phe Ser Gly
305 310 315 320

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
325 330 335

Glu Asp Phe Ala Thr Tyr Tyr Cys Gly Gly Gly Tyr Ser Ser Ile Ser
340 345 350

Asp Thr Thr Phe Gly Cys Gly Thr Lys Val Glu Ile Lys Arg Thr
355 360 365

<210> 71

<211> 1101

<212> DNA

<213> Sequenza Artificiale

<220>

<223> catena leggera 1519gL20 FabFv

<400> 71

```
atgtctgtcc ccaccaagt cctcggactc ctgctactct ggcttacaga tgccagatgc      60
gatatccaga tgaccagag cccatctagc ttatccgctt ccgttggtga tcgctgaca      120
attacgtgta agagctccca atctctcgtg ggtgcaagtg gcaagaccta tctgtactgg      180
ctctttcaga agcctggcaa ggcacaaaaa cggctgatct atctggtgtc tacccttgac      240
tctgggatac cgtcacgatt ttccggatct gggagcggaa ctgagttcac actcacgatt      300
tcatcgtctg aaccogagga ctttgctacc tactactgcc tgcaaggcac tcatttccct      360
cacactttcg gccaggggac aaaactcgaa atcaaacgta cggtagcggc cccatctgtc      420
ttcatcttcc cgccatctga tgagcagttg aaatctggaa ctgcctctgt tgtgtgcctg      480
ctgaataact tctatcccag agaggccaaa gtacagtggg aggtggataa cgccctccaa      540
```

tctgggtaact cccaggagag tgtcacagag caggacagca aggacagcac ctacagcctg 600
agcagcaccc tgacgctgtc taaagcagac tacgagaaac acaaagtgtg cgcctgcgaa 660
gtcaccatc agggcctgag ctccaccagta acaaaaagtt ttaatagagg ggagtgtagc 720
ggtggcgggtg gcagtgggtg gggaggctcc ggaggtggcg gttcagacat acaaatgacc 780
cagagtcctt catcgggtatc cgcgtccggtt ggcgataggg tgactattac atgtcaaagc 840
tctcctagcg tctggagcaa ttttctatcc tggatcaac agaaaccggg gaaggctcca 900
aaacttctga tttatgaagc ctcgaaactc accagtggag ttccgtcaag attcagtggc 960
tctggatcag ggacagactt cacgttgaca atcagttcgc tgcaaccaga ggactttgcg 1020
acctactatt gtggtggagg ttacagtagc ataagtata cgacatttgg gtgctgtact 1080
aaggtggaaa tcaaactgac c 1101

<210> 72

<211> 357

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 FabFv

<400> 72

Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Ser Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln Gly Thr Leu
100 105 110

Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
115 120 125

Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys
130 135 140

Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser
145 150 155 160

Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser
165 170 175

Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser
180 185 190

Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn
195 200 205

Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Ser Gly Gly Gly
210 215 220

Gly Ser Gly Gly Gly Gly Thr Gly Gly Gly Gly Ser Glu Val Gln Leu
225 230 235 240

Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu
245 250 255

Ser Cys Ala Val Ser Gly Ile Asp Leu Ser Asn Tyr Ala Ile Asn Trp
260 265 270

Val Arg Gln Ala Pro Gly Lys Cys Leu Glu Trp Ile Gly Ile Ile Trp
275 280 285

Ala Ser Gly Thr Thr Phe Tyr Ala Thr Trp Ala Lys Gly Arg Phe Thr
290 295 300

Ile Ser Arg Asp Asn Ser Lys Asn Thr Val Tyr Leu Gln Met Asn Ser
305 310 315 320

Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Thr Val Pro
325 330 335

Gly Tyr Ser Thr Ala Pro Tyr Phe Asp Leu Trp Gly Gln Gly Thr Leu
340 345 350

Val Thr Val Ser Ser
355

<210> 73

<211> 1071

<212> DNA
<213> Sequenza Artificiale
<220>
<223> catena pesante 1519gH20 FabFv
<400> 73
gagggtaccac ttgtggaaag cggaggaggt cttgtgcagc ctggaggaag tttacgtctc 60
tcttgtgctg tgtctggctt caccttctcc aattacggaa tggctctgggt cagacaagca 120
cctggaaagg gtcttgaatg ggtggcctat attgactctg acggggacaa cacctactat 180
cgggattccg tgaaggagc cttcacaatc tcccagata acgccaagag ctactgtac 240
ctgcagatga atagcctgag agccgaggat actgccgtgt actattgcac aacgggaatc 300
gttaggcctt ttctgtactg gggacagggc accttgggta ctgtctcgag cgcgtccaca 360
aagggcccat cgtcttccc cctggcacc cctccaaga gcacctctgg gggcacagcg 420
gccctgggct gcctgggtcaa ggactacttc cccgaaccag tgacgggtgc gtggaactca 480
ggtgccctga ccagcggcgt tcacacctc ccggctgtcc tacagtcttc aggactctac 540
tccttgagca gcgtgggtgac cgtgccctcc agcagcttgg gcaccagac ctacatctgc 600
aacgtgaatc acaagcccag caacaccaag gtcgataaga aagttgagcc caaatcttgt 660
agtggagggtg ggggctcagg tggaggcggg accggtggag gtggcagcga ggttcaactg 720
cttgagtctg gaggaggcct agtccagcct ggaggagacc tgcgtctctc ttgtgcagta 780
agcggcatcg acctgagcaa ttacgccatc aactgggtga gacaagctcc ggggaagtgt 840
ttagaatgga tcggtataat atgggccagt gggacgacct tttatgctac atgggcgaaa 900
ggaaggttta caattagccg ggacaatagc aaaaacaccg tgtatctcca aatgaactcc 960
ttgcgagcag aggacacggc ggtgtactat tgtgctcgca ctgtcccaggt ttatagcaact 1020
gcaccctact tcgatctgtg gggacaaggg accctgggtga ctgtttcaag t 1071

<210> 74

<211> 376

<212> PRT

<213> Sequenza Artificiale

<220>

<223> catena pesante 1519gH20 FabFv

<400> 74

Met Glu Trp Ser Trp Val Phe Leu Phe Phe Leu Ser Val Thr Thr Gly
1 5 10 15

Val His Ser Glu Val Pro Leu Val Glu Ser Gly Gly Gly Leu Val Gln
20 25 30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Val Ser Gly Phe Thr Phe

35 40 45

Ser Asn Tyr Gly Met Val Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
50 55 60

Glu Trp Val Ala Tyr Ile Asp Ser Asp Gly Asp Asn Thr Tyr Tyr Arg
65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Ser
85 90 95

Ser Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
100 105 110

Tyr Tyr Cys Thr Thr Gly Ile Val Arg Pro Phe Leu Tyr Trp Gly Gln
115 120 125

Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val
130 135 140

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
145 150 155 160

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
165 170 175

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
180 185 190

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
195 200 205

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
210 215 220

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Ser
225 230 235 240

Gly Gly Gly Gly Ser Gly Gly Gly Gly Thr Gly Gly Gly Gly Ser Glu
245 250 255

Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser
260 265 270

Leu Arg Leu Ser Cys Ala Val Ser Gly Ile Asp Leu Ser Asn Tyr Ala
275 280 285

Ile Asn Trp Val Arg Gln Ala Pro Gly Lys Cys Leu Glu Trp Ile Gly
290 295 300

Ile Ile Trp Ala Ser Gly Thr Thr Phe Tyr Ala Thr Trp Ala Lys Gly
305 310 315 320

Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Val Tyr Leu Gln
325 330 335

Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
340 345 350

Thr Val Pro Gly Tyr Ser Thr Ala Pro Tyr Phe Asp Leu Trp Gly Gln
355 360 365

Gly Thr Leu Val Thr Val Ser Ser
370 375

<210> 75

<211> 22

<212> PRT

<213> Sequenza Artificiale

<220>

<223> Tsp 5' selvatico

<400> 75

Met Asn Met Phe Phe Arg Leu Thr Ala Leu Ala Gly Leu Leu Ala Ile
1 5 10 15

Ala Gly Gln Thr Phe Ala
20

<210> 76

<211> 66

<212> DNA

<213> Sequenza Artificiale

<220>

<223> Tsp 5' selvatico

<400> 76

atgaacatgt ttttaggct taccgcgta gctggcctgc ttgcaatagc aggccagacc 60

ttcgct 66

<210> 77

<211> 19

<212> PRT

<213> Sequenza Artificiale

<220>

<223> frammento mutato Tsp 5'

<400> 77

Met Asn Ser Phe Leu Gly Leu Pro Arg Leu Ala Cys Leu Gln Gln Ala
1 5 10 15

Arg His Leu

<210> 78
<211> 66
<212> DNA
<213> Sequenza Artificiale
<220>
<223> frammento mutato Tsp 5'
<400> 78
atgaattcgt ttttaggctt accgcgtag ctggcctgct tgcaatagca ggccagacat 60

taattg 66
<210> 79
<211> 20
<212> PRT
<213> Artificiale
<220>
<223> delta ptr (proteasi III) 5' mutata
<400> 79
Ile Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
1 5 10 15

Trp Ala His Cys
20
<210> 80
<211> 66
<212> DNA
<213> Artificiale
<220>
<223> delta ptr (proteasi III) 5' mutata
<400> 80
tgaattcccc gcagcacctg gttcaaagca ttattgttgt tagttgccct ttgggcacat 60

taatgt 66
<210> 81
<211> 21
<212> PRT
<213> Artificiale
<220>
<223> ptr (proteasi III) 5' selvatica
<400> 81
Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
1 5 10 15

Trp Ala Pro Leu Ser
20
<210> 82
<211> 66
<212> DNA
<213> Artificiale
<220>
<223> ptr (proteasi III) 5' selvatica

<400> 82
tgaatgccc gcagcacctg gttcaaagca ttattgttgt tagttgccct ttgggcaccc 60
ttaagt 66

Rivendicazioni

1. Cellula batterica ricombinante gram-negativa comprendente:

a. un gene spr mutato codificante una proteina avente una mutazione in uno o più degli amminoacidi selezionati tra D133, H145, H157, N31, R62, 170, Q73, C94, S95, V98, Q99, R100, L108, Y115, V135, L136, G140, R144 e G147 e

b. un vettore di espressione comprendente un gene che esprime o sovraesprime una o più proteine in grado di facilitare il ripiegamento proteico scelte tra FkpA, Skp o una combinazione di FkpA e Skp, in cui la cellula ha una attività ridotta della proteina Tsp paragonata ad una cellula selvatica e in cui la cellula non comprende un polinucleotide ricombinante codificante DsbC e in cui la cellula batterica è isogenica rispetto a una cellula *E. coli* selvatica ad eccezione del gene mutato spr e della modifica necessaria a ridurre l'attività della proteina Tsp rispetto ad una cellula selvatica.

2. Cellula secondo la rivendicazione 1, in cui il gene spr mutante codifica una proteina spr avente una o più mutazioni selezionate tra D133A, H145A, H157A, N31Y, R62C, I70T, Q73R, C94A, S95F, V98E, Q99P, R100G, L108S, Y115F, V135D, V135G, L136P, G140C, R144C e G147C.

3. Cellula secondo la rivendicazione 1, in cui il gene spr mutante codifica una proteina spr avente le mutazioni C94 e H145A.

4. Cellula secondo la rivendicazione 1, in cui il gene spr mutante codifica una proteina spr avente una mutazione selezionata tra D133A, H145A e H157A.

5. Cellula secondo la rivendicazione 1, in cui il gene spr mutante codifica una proteina spr avente una mutazione selezionata tra C94A.

6. Cellula secondo una qualsiasi rivendicazione precedente, in cui il gene spr mutante è integrato nel genoma della cellula.

7. Cellula secondo una qualsiasi rivendicazione precedente in cui un gene codificante una proteina in grado di facilitare il ripiegamento proteico è trasfettato in modo transiente nella cellula.

8. Cellula secondo una qualsiasi rivendicazione precedente in cui la proteina in

grado di facilitare il ripiegamento proteico è FkpA.

9. Cellula secondo una qualsiasi rivendicazione precedente, in cui la cellula comprende ulteriormente uno o più dei seguenti geni mutati:

- a. un gene DegP codificante una proteina DegP con attività di chaperone e attività proteasica ridotta;
- b. un gene ptr mutato, in cui il gene ptr mutato codifica una proteina Proteasi III con attività proteasica ridotta o è un gene ptr mutato knockout; e
- c. un gene OmpT mutato, in cui il gene OmpT mutato codifica una proteina OmpT con attività proteasica ridotta o è un gene OmpT mutato knockout.

10. Cellula secondo una qualsiasi rivendicazione precedente, in cui la cellula comprende un gene Tsp mutato codificante una proteina Tsp con attività proteasica ridotta o è un gene Trp mutato.

11. Cellula secondo una qualsiasi rivendicazione precedente, in cui la cellula ha un gene Tsp mutato knockout che comprende una mutazione al codone di inizio del gene e/o uno o più codoni di terminazione collocati a valle del codone d'inizio del gene e a monte del codone di terminazione del gene.

12. Cellula secondo la rivendicazione 11, in cui il gene Tsp mutato knockout comprende un sito di restrizione marcatore creato da una mutazione missenso al codone di inizio del gene e opzionalmente una o più ulteriori mutazioni puntiformi.

13. Cellula secondo la rivendicazione 12, in cui il gene Tsp mutato knockout comprende SEQ ID NO:3.

14. Cellula secondo una qualsiasi rivendicazione precedente, in cui la cellula è *E. coli*.

15. Cellula secondo la rivendicazione 14 in cui il ceppo di *E. coli* è K12 o W3110.

16. Cellula secondo una qualsiasi rivendicazione precedente, in cui la cellula comprende una sequenza polinucleotidica codificante una proteina di interesse selezionata tra un anticorpo o un suo frammento legante un antigene.

17. Cellula secondo la rivendicazione 16, in cui l'anticorpo o il suo frammento legante un antigene è specifico per FcRn.

18. Metodo per produrre una proteina di interesse comprendente coltivare una cellula batterica ricombinante

gram-negativa come definita in una qualsiasi delle rivendicazioni da 1 a 17 in un terreno di coltura in condizioni efficaci per esprimere la proteina ricombinante di interesse e recuperare la proteina ricombinante di interesse dal periplasma della cellula batterica ricombinante gram-negativa e/o dal terreno di coltura.

19. Metodo secondo la rivendicazione 18 in cui il metodo comprende inoltre il recuperare la proteina di interesse dalla cellula.

20. Metodo secondo la rivendicazione 19 in cui la proteina di interesse è recuperata dal periplasma e/o dal sovrantante.

Disegno

Figura 1

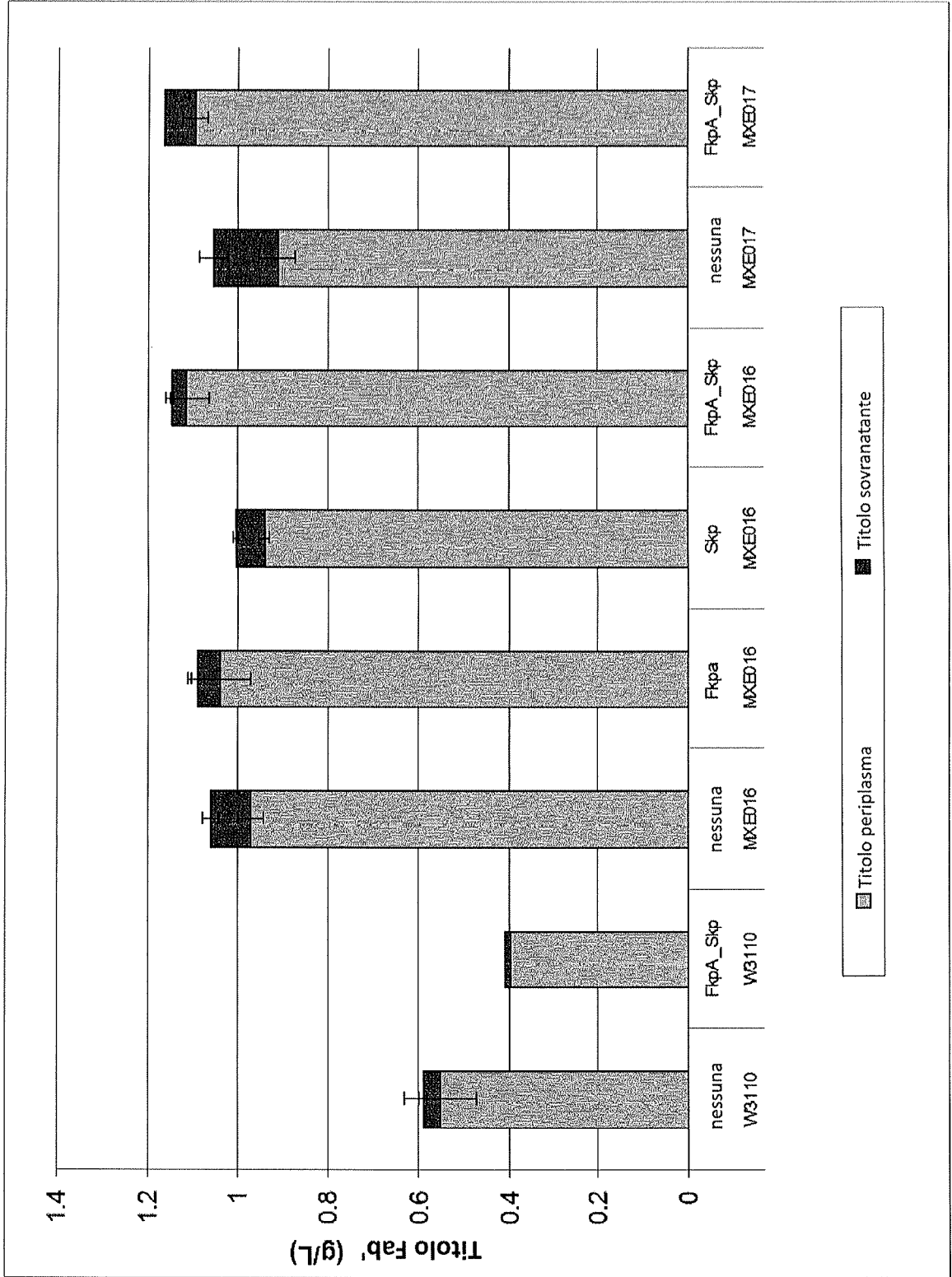


Figura 2 esperimenti di variazione della velocità di alimentazione

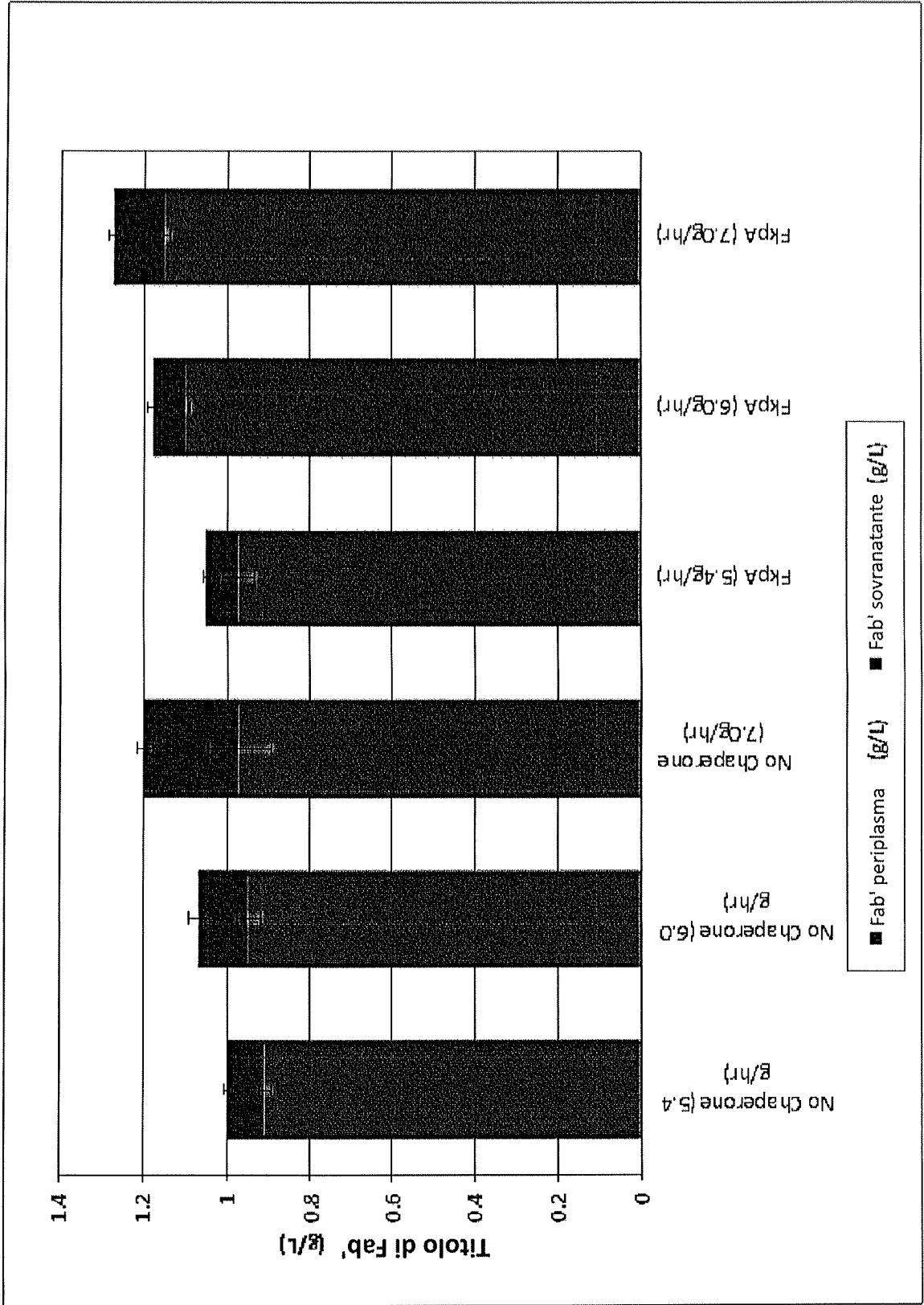


Figura 3A Vitalità cellulare

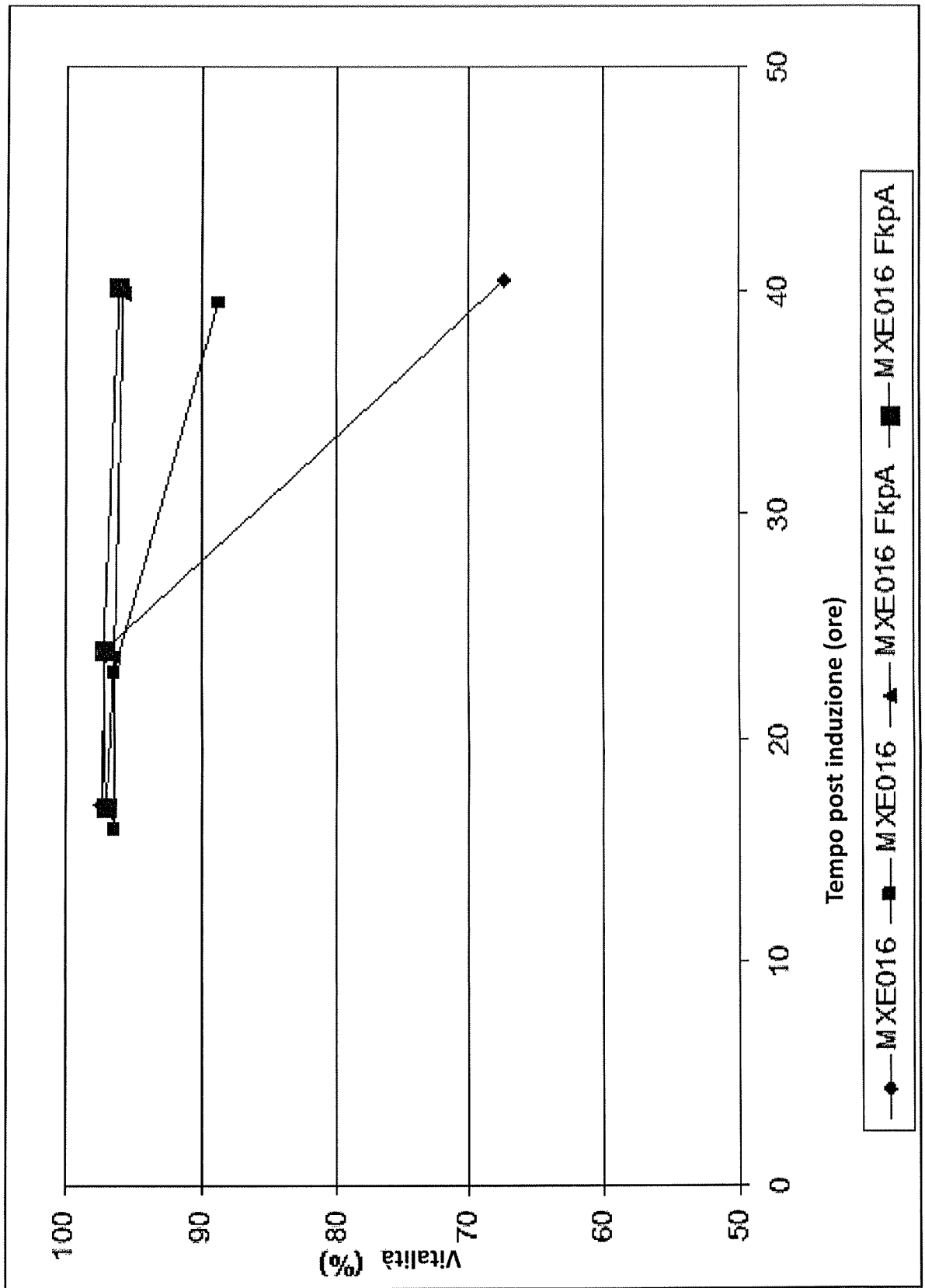
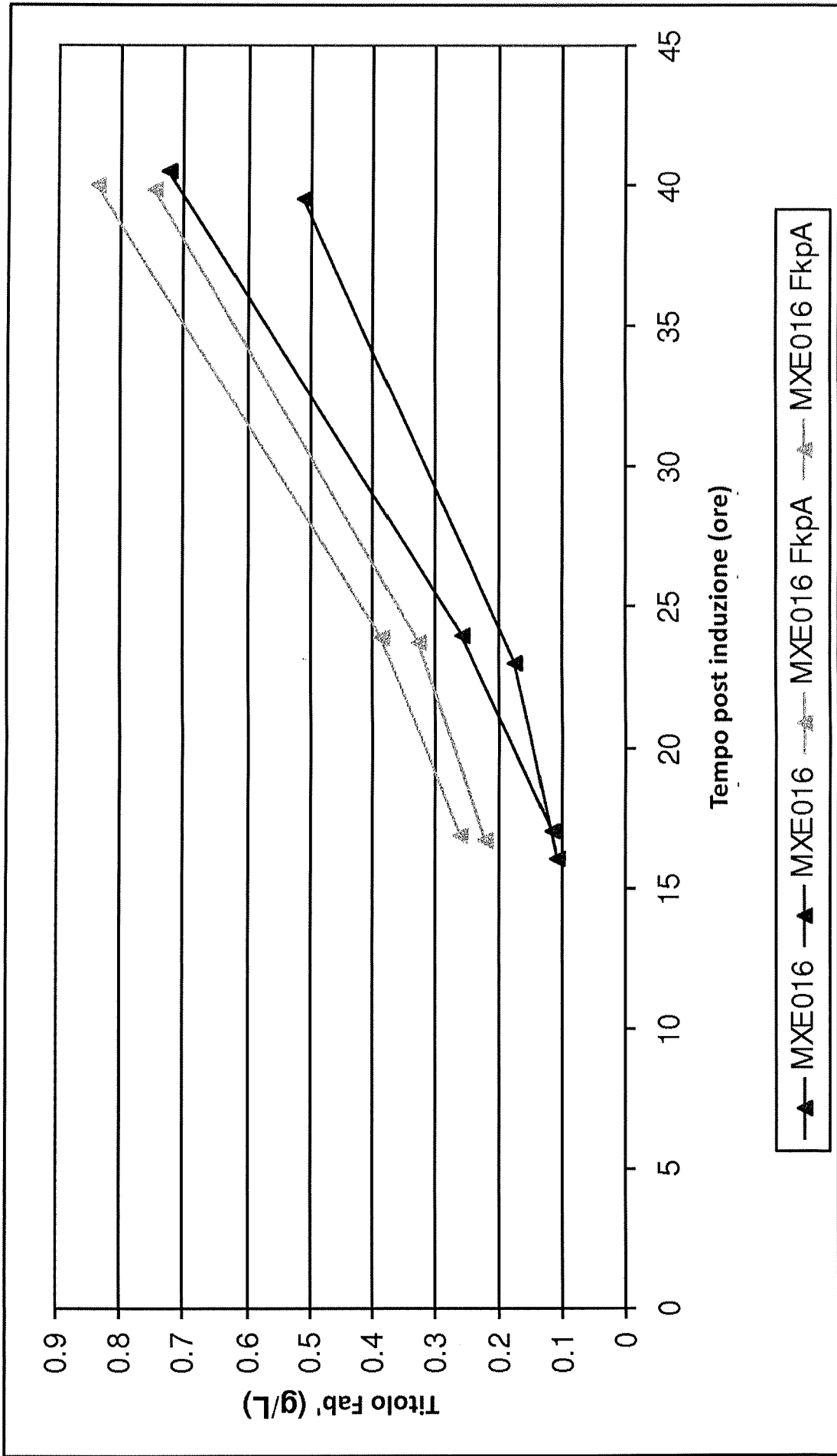


Figura 3B



4 Raccolta da Produzione in scala 20L pilota e Recupero primario per linee cellulari "MXE016"

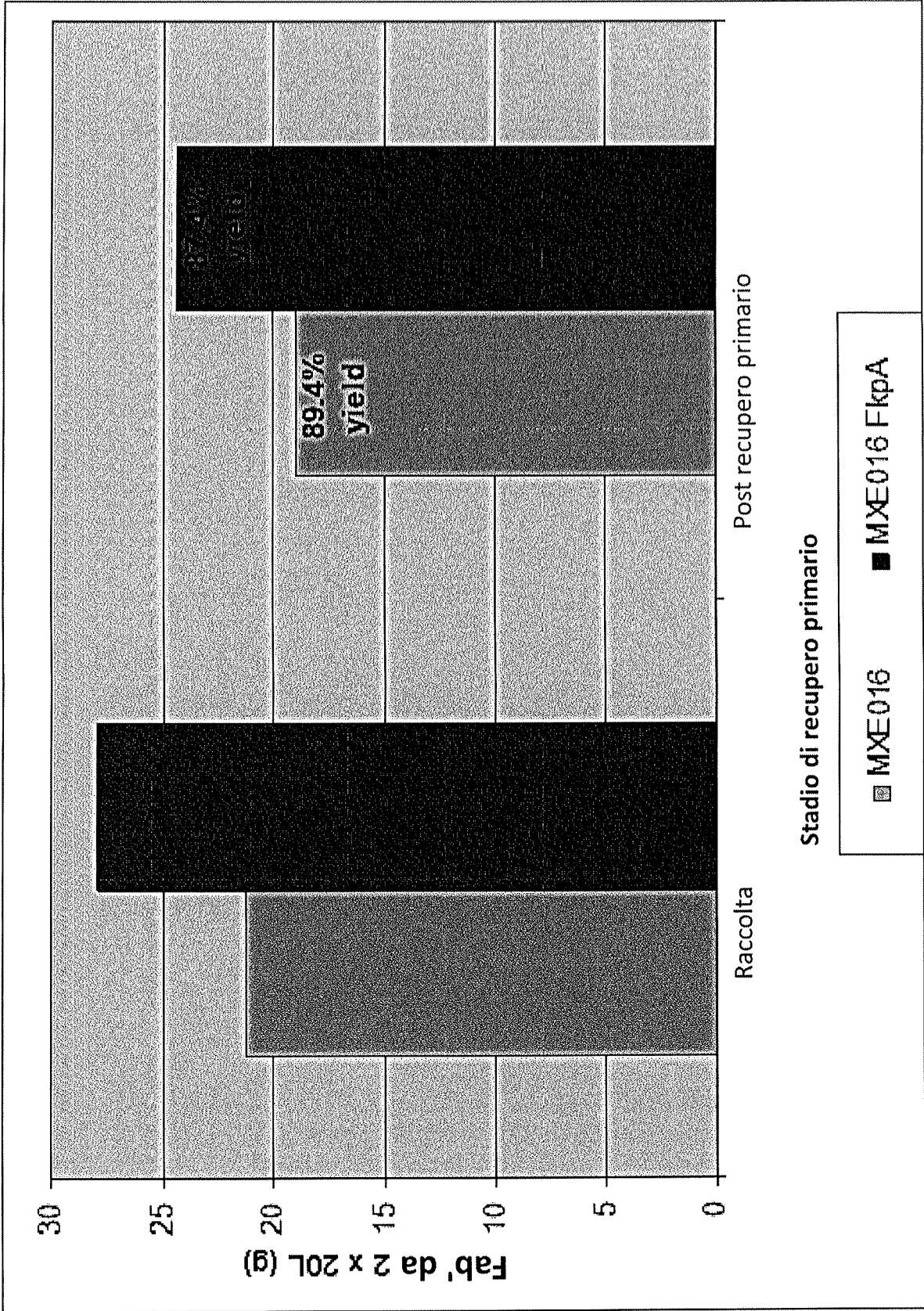
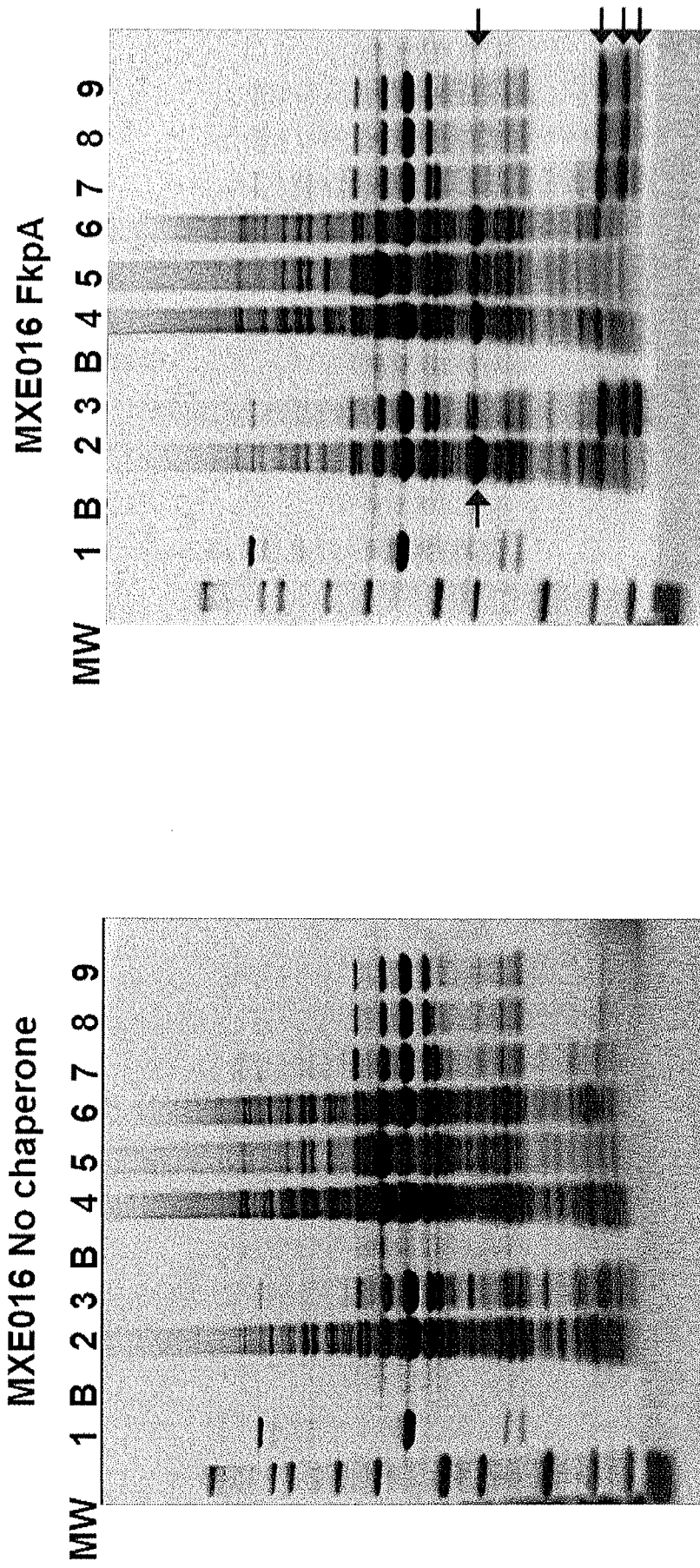
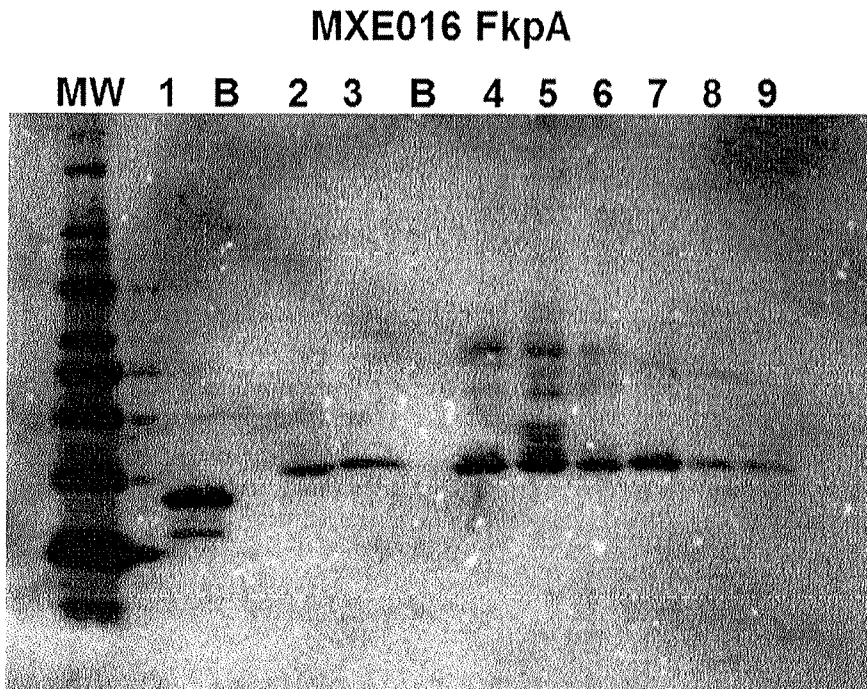


Figura 5 Raccolta primaria da Batch Run 20L su gel SDS-Page colorato in condizioni non riducenti



- | | |
|----------------------------------|------------------------------------|
| 1 -XFab' Standard; | 2- 30°C estratto su piccola scala; |
| 3-60°C estratto su piccola scala | 4- Sovranatante fase pesante; |
| 5-Fase Leggera | 6- Aggiunta Post buffer; |
| 7-Post estrazione; | 8- Centrifugazione post estratto |
| 9-Post filtrazione | |

Figura 6 Western Blot in condizioni non riducenti per un Batch scala pilota da 20L



- 1-Controllo positivo His-tag (DbsC)
- 2-30°C estratto su piccola scala
- 3-60°C estratto su piccola scala
- 4-Sovranatante fase pesante
- 5-Fase leggera
- 6-Aggiunta post buffer
- 7-Post estrazione
- 8-Centrifugazione post estratto
- 9-Post filtrazione

Figura 7A Mutazioni in vari geni

(proteasi III) 5' Selvatica (SEQ ID NO 81 e 82)

* M P R S T W F K A L L L L V
 TGA **ATG** CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

A L W A P L S
 GCC CTT TGG GCA CCC TTA AGT

Δ ptr (proteasi III) 5' Mutata (SEQ ID NO 79 e 80)

EcoR I
 ~~~~~  
 \* I P R S T W F K A L L L L V  
 TGA **ATT** CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

*Ase I*  
 ~~~~~  
 A L W A H * C
 GCC CTT TGG GCA **CAT TAA** TGT

Figura 7b

Tsp 5' Selvatica (SEQ ID NO 75 e 76)

M N M F F R L T A L A G L L A
ATG AAC **ATG** TTT TTT AGG CTT ACC GCG TTA GCT GGC CTG CTT GCA

I A G Q T F A
 ATA GCA GGC CAG ACC TTC GCT

Δ Tsp 5' Mutata (SEQ ID NO 77 e 78)

EcoR I
 ~~~~~  
 M N S F L G L P R \* L A C L Q  
**ATG** AAT **TCG** TTT TTA GGC TTA CCG CGT TAG CTG GCC TGC TTG CAA

*Ase I*  
 ~~~~~  
 * Q A R H * L
 TAG CAG GCC AGA **CAT TAA** TTG

Figura 7C

: DegP Selvatica

202 D A A I N R G N S G G
 949 GAT GCA GCG ATC AAC CGT GGT AAC **TCC** GGT GGT

DegP S210A Mutata

Ase I

202 D A A I N R G N A G G
 949 GAT GCA GCG **ATT AAT** CGT GGT AAC **GCC** GGT GGT

fFigura 7D inserzione genica

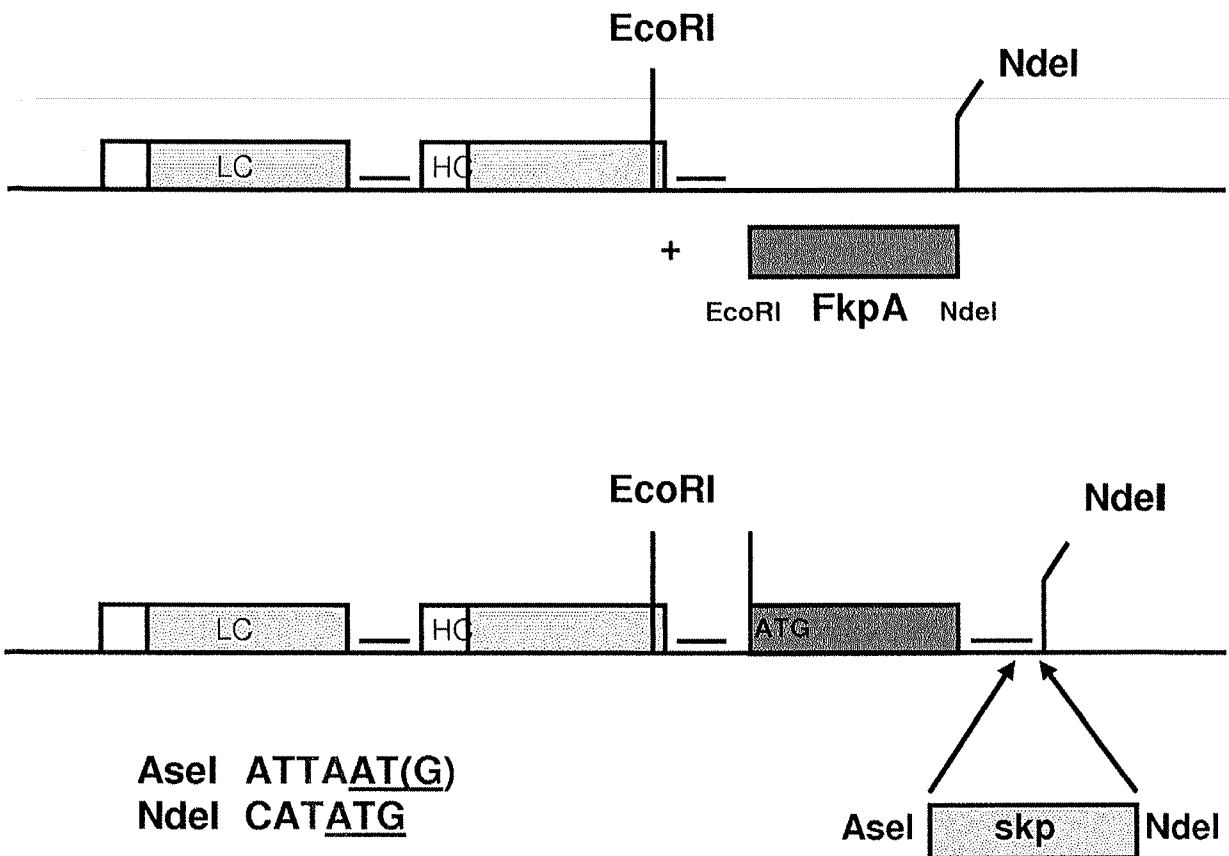


Figura 8a SEQ ID NO 1: è la sequenza di DNA del gene Tsp selvatico inclusi i 6 nucleotidi ATGAAC a monte del codone d'inizio

ATGAACATGTTTTTTAGGCTTACCGCGTTAGCTGGCCTGCTTGCAATAGCAGGCCAGACCTTCGCTGTAGAA
 GATATCACGCGTGCTGATCAAATTCGGTATTAAGGAAGAGACGCAGCATGCGACGGTAAGTGAGCGCGTA
 ACGTCGCGCTTACCCGTTCTCATTATCGCCAGTTCGACCTCGATCAGGCATTTTCGGCCAAAATCTTTGAC
 CGCTACCTGAATCTGCTCGATTACAGCCACAACGTGCTGCTGGCAAGCGATGTTGAACAGTTCGCGAAAAAG
 AAAACCGAGTTAGGCGATGAACGCGTTACAGGCAAACGACGTTTTCTACGATCTCTACAATCTGGCGCAA
 AAGCGCCGTTTTGAGCGTTACCAGTACGCTTTGTGGTACTGGAAAAGCCGATGGATTTACCCGGCAACGAC
 ACTTATAACCTTGACCGCAGCAAAGCGCCCTGGCCGAAAAACGAGGCTGAGTTGAACGCGCTGTGGGACAGT
 AAAGTCAAATTCGACGAGTTAAGCCTGAAGCTGACAGGAAAAACGGATAAAAGAAATTCGTGAAACCCTGACT
 CGCCGCTACAAATTTGCCATTCGTCTGCTGGCGCAAACCAACAGCGAAGATGTTTTCTCGCTGGCAATGACG
 GCGTTTTGCGCGTGAAATCGACCCGCATACCAACTATCTTTCCCCGCGTAATACCGAACAGTTCAACACTGAA
 ATGAGTTTGTGCTGGAAGGTATTGGCGCAGTGTGCAAATGGATGATGACTACCCGTTATCAATTCGATG
 GTGGCAGGTGGTCCGGCAGCGAAGAGTAAAGCTATCAGCGTTGGTGACAAAAATTGTCGGTGTGGTCAAACA
 GGCAAGCCGATGGTTGACGTGATTGGCTGGCGTCTTGATGATGTGGTTGCCCTAATTAAGGGCCGAAGGGC
 AGTAAAGTTCGTCTGGAATTTTACCTGCTGGTAAAGGGACCAAGACCCGTAAGTAAAGGGCCGAAGGGC
 CGTATTCGTCTCGAAGACCGCGCGTTAAAATGTGGTGAAGACCGTGGTAAAGAGAAAGTCCGCGTGTG
 GATATTCGGGCTTCTATGTGGGTTTACAGACGATGTCAAAGTGAACGCAACTGCAGAACTGGAAAAACAGAAT
 GTCAGCAGCGTCATCATCGACCTGCGTAGCAATGGCGGTGGGGCGTTAACTGAAGCCGATCGCTCTCCGGT
 CTGTTTATTCCTGCGGGTCCCATTGTTTCAGGTCCGCGATAACAACGGCAAGGTTTCGTGAAGATAGCGATACC
 GACGGACAGTTTTTCTATAAAGGCCCGCTGGTGGTGCTGGTTGACCGCTTCAGTGCTTCGGCTTCAGAAATC
 TTTGCCGCGCAATGCAGGATTACGGTCTGCGCTGGTTGTGGGTGAACCGACGTTTGGTAAAGGCACCGTT
 CAGCAATACCGTTCATTGAACCGTATTTACGATCAGATGTTACGTCCTGAATGGCCAGCGCTGGGTTCGTG
 CAGTACACGATCCAGAAATTCATCGCGTTAACGGCGCAGTACGCAACGTAAGGCGTAACGCCAGACATC
 ATCATGCCGACGGTAATGAAGAAACGGAACGGGTGAGAAATTCGAAGATAACGCGCTGCCGTGGGATAGC
 ATTGATGCCGCGACTTATGTGAAATCAGGAGATTTAACGGCCTTTGAACCGGAGCTGCTGAAGGAACATAAT
 GCGCGTATCGCGAAAGATCCTGAGTTCAGAACATCATGAAGGATATCGCGCGCTTCAACGCTATGAAGGAC
 AAGCGCAATATCGTTTTCTCTGAATTACGCTGTGCGTGAGAAAGAGAATAATGAAGATGATGCGACGCGTCTG
 GCGCGTTTGAACGAACGCTTTAAACGCGAAGGTAAACCGGAGTTGAAGAACTGGATGATCTACCGAAAGAT
 TACCAGGAGCCGATCCTTATCTGGATGAGACGGTGAATATCGCACTCGATCTGGCGAAGCTTGAAAAAGCC
 AGACCCGCGGAACAACCCGCTCCCGTCAAGTAA

Figura 8b SEQ ID NO 2 è la sequenza amminoacidica della proteina Tsp selvatica

MFFRLTALAGLLAIAGQTFAVEDITRADQIPVLKEETQHATVSEVTSRFRTRSHYRQFDLDQAFSAKIFDRY
 LNLLDYSHNVLLASDVEQFAKKKTELGDLELRSGKLDVFDLYNLAQKRRFERYQYALSVLEKPMDFGTNDY
 NLDRSKAPWPKNEAELNALWDSKVKFDELSLKLTKGTDKEIRETLTRYKFAIRRLAQTNSEDVFLAMTAF
 AREIDPHTNYLSPRNTEQFNTEMSLSLEGIGAVLQMDDDYTVINSMVAGGPAAKSKAISVGDKIVGVGQTGK
 PMVDVIGWRLDDVVALIKGPKGSKVRLEILPAGKGTKTRTVTLTRERIRLEDRAVKMSVKTVGKEKVGVLDI
 PGFYVGLTDDVKVQLQKLEKQNVSSVIIDLRSNGGGALTEAVSLSGLFIPAGPIVQVRDNNKQVREDSDTDG
 QVFKYKGPLVVLVDRFSASASEIFAAMQDYGRALVVGEPTFGKGTQQYRSLNRIYDQMLRPEWPALGVSQY
 TIQKFYRVNGGSTQRKGVTPDIIMPTGNEETETGEKFDNALPWDSIDAATYVKSGLDTAFEPPELLKEHNAR
 IAKDPEFQNI MKDIARFNAMKDRNIVSLNYAVREKENNEDATRLARLNERFKREGKPELKKLDDL PKDYQ
 EPDPYLDETVNIALDLAKLEKARPAEQPAPVK*

Figura 8C

SEQ ID NO 3 è la sequenza di DNA di un gene Tsp mutato knockout inclusi i 6 nucleotidi ATGAAT a monte del codone d'inizio

ATGAATTCGTTTTTAGGCTTACCGCGTTAGCTGGCCTGCTTGCAATAGCAGGCCAGACATTAATTGTAGAAG
 ATATCACGCGTGCTGATCAAAATCCGGTATTAAGGAGAGACGCAGCATGCGACGGTAAGTGAGCGCGTAA
 CGTCGCGCTTACCCGTTCTCATTATCGCCAGTTCGACCTCGATCAGGCATTTTCGGCCAAAATCTTTGACC
 GCTACCTGAATCTGCTCGATTACAGCCACAACGTGCTGCTGGCAAGCGATGTTGAACAGTTCGCGAAAAAGA
 AAACCGAGTTAGGCGATGAACTGCGTTCAGGCAAACTCGACGTTTTCTACGATCTCTACAATCTGGCGCAAA
 AGCGCCGTTTTGAGCGTTACCAGTACGCTTTGTGCGTACTGGAAAAGCCGATGGATTTACCCGGCAACGACA
 CTTATAACCTTGACCGCAGCAAAGCGCCCTGGCCGAAAAACGAGGCTGAGTTGAACGCGCTGTGGGACAGTA
 AAGTCAAATTCGACGAGTTAAGCCTGAAGCTGACAGGAAAAACGGATAAAGAAAATTCGTGAAACCCTGACTC
 GCCGCTACAAATTTGCCATTTCGTGCTCTGGCGCAAAACACAGCGAAGATGTTTTCTCGCTGGCAATGACGG
 CGTTTGGCGGTGAAATCGACCCGCATACCAACTATCTTTCCCGCGTAATACCGAACAGTTCAACACTGAAA
 TGAGTTTGTGCTGGAAGGTATTGGCGCAGTGCTGCAAAATGGATGATGACTACACCGTTATCAATTCGATGG
 TGGCAGGTGGTCCGGCAGCGAAGAGTAAAGCTATCAGCGTTGGTGACAAAATTTGTGCGGTGTTGGTCAAACAG
 GCAAGCCGATGGTTGACGTGATTGGCTGGCGTCTTGATGATGTGGTTGCCTTAATTAAGGGCCGAAGGGCA
 GTAAGTTTCGTCTGGAATTTTACCTGCTGGTAAAGGGACCAAGACCCGTAACGTTGACCCGTGAAC
 GTATTCGTCTCGAAGACCGCGCGTTAAATGTGCGTGAAGACCGTTCGGTAAAGAGAAAAGTCCGGCGTGGTGG
 ATATTCGGGGCTTCTATGTGGGTTTACAGACGATGTCAAAGTGCAACTGCAGAACTGGAAAAACAGAATG
 TCAGCAGCGTCATCATCGACCTGCGTAGCAATGGCGGTGGGGCGTTAACTGAAGCCGTATCGCTCTCCGGTC
 TGTTTATTCCTGCGGGTCCCATTTGTTTCCAGGTCCGCGATAACAACGGCAAGGTTTCGTGAAGATAGCGATAACG
 ACGGACAGGTTTTCTATAAAGGCCCGCTGGTGGTGTGTTGACCGCTTCAGTGTCTCGGCTTCAGAAATCT
 TTGCCGCGGCAATGCAGGATTACGGTCTGCGCTGGTGTGGTGAACCGACGTTTGGTAAAGGCACCGTTC
 AGCAATACCGTTCATTGAACCGTATTTACGATCAGATGTTACGTCTGAATGGCCAGCGCTGGGTTCTGTGC
 AGTACACGATCCAGAAATTCATCGCGTTAACGGCGGCAGTACGCAACGTAAAGGCGTAACGCCAGACATCA
 TCATGCCGACGGGTAATGAAGAAACGGAAACGGGTGAGAAATTCGAAGATAACGCGCTGCCGTGGGATAGCA
 TTGATGCCGCGACTTATGTGAAATCAGGAGATTTAACGGCCTTTGAACCGGAGCTGCTGAAGGAACATAATG
 CGCGTATCGCGAAAGATCCTGAGTTCAGAACATCATGAAGGATATCGCGCGCTTCAACGCTATGAAGGACA
 AGCGCAATATCGTTTTCTCTGAATTACGCTGTGCGTGAGAAAGAGAATAATGAAGATGATGCGACGCGTCTGG
 CGGTTTTGAACGAACGCTTTAAACGCGAAGGTAACCGGAGTTGAAGAACTGGATGATCTACCGAAAGATT
 ACCAGGAGCCGGATCCTTATCTGGATGAGACGGTGAATATCGCACTCGATCTGGCGAAGCTTGA AAAAGCCA
 GACCCGCGGAACAACCCGCTCCCGTCAAGTAA

SEQ ID NO 4 è la sequenza di DNA del gene selvatico della Protasi III

ATGCCCCGAGCACCTGGTTCAAAGCATTATTGTTGTTAGTTGCCCTTTGGGCACCCTTAAGTCAGGCAGAA
 ACGGGATGGCAGCCGATTACAGGAAACCATCCGTAAAAGTGATAAAGATAACCGCCAGTATCAGGCTATACGT
 CTGGATAACGGTATGGTGGTCTTGCTGGTCTTGATCCGCAGGCAGTTAAATCGCTCTCGGCGCTGGTGGTG
 CCCGTTGGGTGCTGGAAGATCCCAGGCGTACCAGGGCTGGCACATTACCTTGAACATATGAGTCTGATG
 GGGTCGAAAAAGTACCCGACAGGCTGACAGTCTGGCCGAATATCTCAAAATGCACGGCGGTAGTCACAATGCC
 AGCACTGCGCCGATCGCACGGCTTTCTATCTGGAAGTTGAGAACGACGCCTTGCTGGTGGGTGACCCG
 CTGGCCGATGCTATTGCTGAACCTTTGCTCGACAAGAAATATGCCGAACGTGAGCGTAATGCGGTGAACGCT
 GAATTAACCATGGCGCGTACGCGTGACGGGATGCGCATGGCACAGGTCAGCGCAGAAACCATTAACCCGGCA
 CACCCCGTTCAAAGTTTTCTGGTGGTAACCTCGAAACTTTAAGCGACAAACCTGGTAATCCGGTGCAGCAG
 GCGCTGAAAGATTTCCACGAGAAGTACTATTCCGCCAATTTGATGAAGGCGGTTATTTACAGTAATAAACCG
 CTGCCGGAGTTGGCAAAAATGGCGGCGGACACCTTTGGTCCGCTGCCGAACAAAGAGAGCAAAAAACCGGAA
 ATCACCGTGCCGGTAGTCACCGACGCGCAAAAGGGCATTATCATTACCTTACGTCCTGCGCTGCCGCGTAAA
 GTGTTGCGGTTGAGTTTCGCATCGATAACAACCTCAGCGAAGTTCCGTAGTAAAACCGATGAATTGATTACC
 TATCTGATTGGCAATCGCAGCCAGGTACACTTTCTGACTGGCTGCAAAAGCAGGGATTAGTTGAGGGCATT
 AGCGCAACTCCGATCCTATCGTCAACGGCAACAGCGGCGTATTAGCGATCTCTGCGCTTTAAACCGATAAA
 GGCCTGGCTAATCGCGATCAGGTTGTGGCGGCAATTTTGTAGCTATCTCAATCTGTTACGTGAAAAAGGCATT

Figura 8C continuata

GATAAACAATACTTCGATGAACTGGCGAATGTGCTGGATATCGACTTCCGTTATCCGTCGATCACCCGTGAT
 ATGGATTACGTCGAATGGCTGGCAGATACCATGATTCCGCGTTCCTGTTGAGCATAACGCTGGATGCAGTCAAT
 ATTGCCGATCGGTACGATGCTAAAGCAGTAAAGGAACGCTCTGGCGATGATGACGCCGCAGAAATGCCGCTATC
 TGGTATATCAGCCCGAAAGAGCCGCACAACAAAACGGCTTACTTTGTTCGATGCGCCGTATCAGGTCGATAAA
 ATCAGCGCACAAACTTTCGCCGACTGGCAGAAAAAAGCCGCCGACATTGCGCTCTCTTTGCCAGAGCTTAAC
 CCTTATATTCCTGATGATTTCTCGCTGATTAAGTCAGAGAAGAAAATACGACCATCCAGAGCTGATTGTTGAT
 GAGTCGAATCTGCGCGTGGTGTATGCGCCAAGCCGTTATTTTCCAGCGAGCCCAAAGCTGATGTCAGCCTG
 ATTTTGCCTAATCCGAAAGCCATGGACAGCGCCCGCAATCAGGTGATGTTTGCCTCAATGATTATCTCGCA
 GGGCTGGCGCTTGATCAGTTAAGCAACCAGGCGTGGTGGTGGCATAAGTTTTTCCACCAACGCTAACAAC
 GGCCTTATGGTTAATGCTAATGGTTACACCCAGCGTCTGCCGAGCTGTTCCAGGCATTGCTCGAGGGGTAC
 TTTAGCTATAACCGCTACGGAAGATCAGCTTGAGCAGGCGAAGTCCCTGGTATAACCAGATGATGGATTCGCA
 GAAAAGGGTAAAGCGTTGAGCAGGCGATTATGCCCGCGCAGATGCTCTCGCAAGTGCCGTACTTCTCGCGA
 GATGAACGGCGTAAAATTTTGCCTCCATTACGTTGAAAGAGGTGCTGGCCTATCGCGACGCCTTAAAATCA
 GGGGCTCGACCAGAGTTTATGGTTATCGGCAACATGACCGAGGCCAGGCAACAACGCTGGCACGCGATGTG
 CAAAACAGTTGGGCGCTGATGGTTTACAGAGTGGTGTGCAAAACAAAGATGTAGTGGTTCGATAAAAAACAATCC
 GTCATCTTTGAAAAAGCCGGTAACAGCACCGACTCCGCACTGGCAGCGGTATTTGTACCGACTGGCTACGAT
 GAATACACCAGCTCAGCCTATAGCTCTCTGTTGGGGCAGATCGTACAGCCGTGGTCTACAATCAGTTGCGT
 ACCGAAGAACAATTGGGCTATGCCGTGTTTGCCTTCCAATGAGCGTGGGGCGTCAGTGGGGCATGGGCTTC
 CTTTTCGAAAGCAATGATAAACAGCCTTCAATCTTGTGGGAGCGTTACAAGGCGTTTTTCCCAACCGCAGAG
 GCAAAATTGCGAGCGATGAAGCCAGATGAGTTTGCGCAAATCCAGCAGGCGGTAATTACCCAGATGCTGCAG
 GCACCGCAAACGCTCGGCGAAGAAGCATCGAAGTTAAGTAAAGATTCGATCGCGGCAATATGCGCTTCGAT
 TCGCGTGATAAAATCGTGGCCCAGATAAAAAGCTGCTGACGCCGCAAAAAGCTTGCTGATTCTTCCATCAGGCG
 GTGGTCGAGCCGCAAGGCATGGCTATTCTGTGCGAGATTTCCGGCAGCCAGAACGGGAAAAGCCGAATATGTA
 CACCCTGAAGGCTGGAAAGTGTGGGAGAACGTCAGCGGTTGCAGCAAACAATGCCCTGATGAGTGAAAAG
 AATGAGTGA

SEQ ID NO 5 è la sequenza amminoacidica della proteina Proteasi III selvatica

MPRSTWFKALLLLVALWAPLSQAETGWQPIQETIRKSDKDNRYQAIRLDNGMVVLLVSDPQAVKSLSALVV
 PVGSLEDPEAYQGLAHYLEHMSLMGSKKYPQADSLAEYLKMHGGSHNASTAPYRTAFYLEVENDALPGAVDR
 LADAI AEP LLDK KYA ER ERNAVNAELTMARTRDGM RMAQVSAETINPAHPGSKFSGGNLETLSDKPGNPVQQ
 ALKDFHEKYY SANLMKAVIY SNKPLPELAKMAADTFGRVPNKESKKEITVPVVTDAQGII IHYVPALPRK
 VLRVEFRIDNNSAKFRSKTDELITYLIGNRSPGTLSDWLQKQGLVEGISANS DPIVNGNSGVL AISASLTDK
 GLANRDQVVAAI FSYLNLLREK GIDKQYFDELANVLDIDFRYPSITRDM DYVEWLADTMI RVPVEHTLDAVN
 IADRYDAKAVKERLAMMTPQNARIWYISPKEPHNK TAYFVDAPYQVDKISAQTFADWQKKAADIALSLPELN
 PYPIDDFSLIKSEKKYDHP ELIVDES NLRVVYAPSRYFASEPKADVSLILRNPKAMDSARNQVMFALNDYLA
 GLALDQLSNQASVGGISFSTNANNGLMVNANGYTQRLPQLFQALLEGYFSYTATEDQLEQAKSWYNQMMDSA
 EKGKAFEQAIMPAQMLSQVPYFSRDERRKILPSITLKEVLAYRDALKSGARPEFMVIGNMTEAQATTLARDV
 QKQLGADGSEWCRNKDVVVDKQSVIFEKAGNSTDSALAAVFPVPTGYDEYTS SAYSSLLGQIVQPW FYNQLR
 TEEQLGYAVFAFPMSVGRQWGMGFL LQSNKQPSFLWERYKAFFPTAEAKLRAMKPDEFAQIQQAVITQMLQ
 APQTLGEEASKLSKDFDRGNMRFDSRDKIVAQIKLLTPQKLADFFHQAVVEPQGMALISQISGSQNGKAEYV
 HPEGWKVWENV SALQQTMPLMSEKNE*

Figura 8D

SEQ ID NO 6 è la sequenza di DNA di un gene mutato knockout della Proteasi III

ATCCCCGCAGCACCTGGTTCAAAGCATTATTGTTGTTAGTTGCCCTTTGGGCACATTAATGTCAGGCAGAA
ACGGGATGGCAGCCGATTCAGGAAACCATCCGTA AAAAGTGATAAAGATAACCGCCAGTATCAGGCTATACGT
CTGGATAACGGTATGGTGGTCTTGCTGGTTTCTGATCCGCAGGCAGTTAAATCGCTCTCGGCGCTGGTGGTG
CCCGTTGGGTGCTGGAAGATCCCCGAGGCGTACCAGGGGCTGGCACATTACCTTGAACATATGAGTCTGATG
GGGTCGAAAAAGTACCCGCAGGCTGACAGTCTGGCCGAATATCTCAAAATGCACGGCGGTAGTCACAA TGCC
AGCACTGCGCCGATTCGCACGGCTTTCTATCTGGAAGTTGAGAACGACGCCCTTGCC TGGTGC GGTTAGACCGC
CTGGCCGATGCTATTGCTGAACCTTTGCTCGACAAGAAATATGCCGAACGTGAGCGTAATGCGGTGAACGCT
GAATTAACCATGGCGCGTACGCGTGACGGGATGCGCATGGCACAGGTCAGCGCAGAAACCATTAACCCGGCA
CACCCCGGTTCAAAGTTTCTGTTGGTAACCTCGAAACTTTAAGCGACAAACCTGGTAATCCGGTGCAGCAG
GCGCTGAAAGATTTCCACGAGAAGTACTATTCCGCCAATTTGATGAAGGCGGTTATTTACAGTAATAAACCG
CTGCCGGAGTTGGCAAAAATGGCGGCGGACACCTTTGGTGC GCGTGCCGAACAAAGAGAGCAAAAACCCGGAA
ATCACCGTGC GCGTAGTCACCGACGCGCAAAAGGGCATTATCATTCATTACGTCCCTGCGCTGCCGCGTAAA
GTGTTGCGCGTTGAGTTTCGCATCGATAACAACCTCAGCGAAGTTCCGTAGTAAAACCGATGAATTGATTACC
TATCTGATTGGCAATCGCAGCCCAGGTACACTTTCTGACTGGCTGCAAAAGCAGGGATTAGTTGAGGGCATT
AGCGCCAACCTCCGATCCTATCGTCAACGGCAACAGCGGCGTATTAGCGATCTCTGCGCTTTTAACCGATAAA
GGCCTGGCTAATCGCGATCAGGTTGTGGCGGCAATTTT TAGCTATCTCAATCTGTTACGTGAAAAAGGCATT
GATAACAATACTTCGATGAACTGGCGAATGTGCTGGATATCGACTTCCGTTATCCGTCGATCACCCGTGAT
ATGGATTACGTCGAATGGCTGGCAGATAACCATGATTCGCGTTCCGTGTTGAGCATAACGCTGGATGCAGTCAAT
ATTGCCGATCGGTACGATGCTAAAGCAGTAAAGGAACGCTCTGGCGATGATGACGCCGAGAATGCGCGTATC
TGGTATATCAGCCGAAAGAGCCGCACAACAAAACGGCTTACTTTGTGATGCGCCGATCAGGTCGATAAA
ATCAGCGCACAAACTTTCCCGACTGGCAGAAAAAAGCCGCCGACATTGCGCTCTCTTTGCCAGAGCTTAAC
CCTTATATTCCTGATGATTTCTCGCTGATTAAGTCAGAGAAGAAATACGACCATCCAGAGCTGATTGTGAT
GAGTCGAATCTGCGCGTGGTGTATGCGCCAAGCCGTTATTTTGGCAGCGAGCCCAAAGCTGATGTCAGCCTG
ATTTTGGTAAATCCGAAAGCCATGGACAGCGCCCCGCAATCAGGTGATGTTTGGCTCAATGATTATCTCGCA
GGGCTGGCGCTTGATCAGTTAAGCAACCAGGCGTCGGTTGGTGGCATAAGTTTTCACCAACGCTAACCAAC
GGCCTTATGGTTAATGCTAATGGTTACACCCAGCGCTCTGCCGAGCTGTTCCAGGCATTGCTCGAGGGGTAC
TTTAGCTATACCGCTACGGAAGATCAGCTTGAGCAGGCGAAGTCCTGGTATAACCAGATGATGGATTCCGCA
GAAAAGGGTAAAGCGTTT GAGCAGGCGATTATGCCCGCGCAGATGCTCTCGCAAGTGCCGTA CTCTCGCGA
GATGAACGGCGTAAAATTTGCCCTCCATTACGTTGAAAGAGGTGCTGGCCTATCGCGACGCCTTAAAATCA
GGGGCTCGACCAGAGTTTATGGTTATCGGCAACATGACCGAGGCCCAGGCAACAACGCTGGCACGCGATGTG
CAAAAACAGTTGGGCGCTGATGGTTCAGAGTGGTGTGAAACAAAGATGTAGTGGTCGATAAAAAACAATCC
GTCATCTTTGAAAAAGCCGTAACAGCACCGACTCCGCACTGGCAGCGGTATTTGTACCGACTGGCTACGAT
GAATACACCAGCTCAGCCTATAGCTCTCTGTTGGGGCAGATCGTACAGCCGTGGTCTACAATCAGTTGCGT
ACCGAAGAACAATTTGGGCTATGCCGTGTTTGGCTTTCCAATGAGCGTGGGGCGTCAGTGGGGCATGGGCTTC
CTTTTGC AAAGCAATGATAAACAGCCTTCATTCTTGTGGGAGCGTTACAAGGCGTTTTTCCCAACCGCAGAG
GCAAAATTCGAGCGATGAAGCCAGATGAGTTTGGCGAAATCCAGCAGGCGGTAATTACCCAGATGCTGCAG
GCACCGCAAACGCTCGGCGAAGAAGCATCGAAGTTAAGTAAAGATTTTCGATCGCGGCAATATGCGCTTCGAT
TCGCGTGATAAAATCGTGGCCAGATAAACTGCTGACGCCGCAAAAACCTGCTGATTTCTTCCATCAGGCG
GTGGTCGAGCCGCAAGGCATGGCTATTCTGTGCGAGATTTCCGGCAGCCAGAACGGGAAAGCCGAATATGTA
CACCTGAAGGCTGGAAGTGTGGGAGAACGTGACGCGTTGCAGCAAACAATGCCCTGATGAGTGAAAAG
AATGAGTGA

Figura 8E

SEQ ID NO 7 è la sequenza di DNA del gene DegP selvatico

ATGAAAAAAAAACCACATTAGCACTGAGTGCCTGGCTCTGAGTTTAGGTTTGGCGTTATCTCCGCTCTCTGCA
 ACGGCGGCTGAGACTTCTTCAGCAACGACAGCCCAGCAGATGCCAAGCCTTGACCCGATGCTCGAAAAGGTG
 ATGCCTTCAGTGGTCAGCATTAACTAGAAAGGTAGCACAACCGTTAATACGCCGCGTATGCCGCGTAATTTT
 CAGCAGTTCTTCGGTATGATTCTCCGTTCTGCCAGGAAGGTTCTCCGTTCCAGAGCTCTCCGTTCTGCCAG
 GGTGGCCAGGGCGGTAATGGTGGCGGCCAGCAACAGAAATTCATGGCGCTGGGTTCCGGCGTCATCATTGAT
 GCCGATAAAGGCTATGTCGTCACCAACAACCACGTTGTTGATAACCGGACGGTCATTAAAGTTCAACTGAGC
 GATGGCCGTAAGTTTCGACGCGAAGATGGTTGGCAAAGATCCGCGCTCTGATATCGCGCTGATCCAAATCCAG
 AACCCGAAAAACCTGACCGCAATTAAGATGGCGGATTCTGATGCACTGCGCGTGGGTGATTACACCGTAGCG
 ATTGGTAACCCGTTTGGTCTGGGCGAGACGGTAACCTCCGGGATTGTCTCTGCGCTGGGGCGTAGCGGCCTG
 AATGCCGAAAACCTACGAAAACCTCATCCAGACCGATGCAGCGATCAACCGTGGTAACCTCCGGTGGTGCCTG
 GTTAACCTGAACGGCGAACTGATCGGTATCAACACCGCGATCCTCGCACCGGACGGCGGCAACATCGGTATC
 GGTTTTGCATCCCGAGTAACATGGTGA AAAACCTGACCTCGCAGATGGTGAATACGGCCAGGTGAAACGC
 GGTGAGCTGGGTATTATGGGGACTGAGCTGAACTCCGAACTGGCGAAAAGCGATGAAAGTTGACGCCAGCGC
 GGTGCTTTCGTAAGCCAGGTTCTGCCTAATTCCTCCGCTGCAAAAAGCGGGCATTAAAGCGGGTGATGTGATC
 ACCTCACTGAACGGTAAGCCGATCAGCAGCTTTGCCGCACTGCGTGCTCAGGTGGTACTATGCCGGTAGGC
 AGCAAACTGACCCTGGGCTTACTGCGCGACGGTAAGCAGGTTAACGTGAACCTGGAACCTGCAGCAGAGCAGC
 CAGAATCAGGTTGATTCCAGCTCCATCTTCAACGGCATTGAAGGCGTGAGATGAGCAACAAAGGCAAAGAT
 CAGGGCGTGGTAGTGAACAACGTGAAAACGGGCACTCCGGCTGCGCAGATCGGCCTGAAGAAAGGTGATGTG
 ATTATTGGCGCGAACCAGCAGGCAGTGA AAAACATCGCTGAACTGCGTAAAGTTCTCGACAGCAAACCGTCT
 GTGCTGGCACTCAACATTAGCGCGGGCGACAGCACCATCTACCTGTTAATGCAGTAA

SEQ ID NO 8 è la sequenza amminoacidica della proteina DegP selvatica

MKKTTLALSALALSLGLALSPLSATAAETSSATTAQQMPSLAPMLEKVMPSVVSINVEGSTTVNTPRMPRNF
 QOFFGDDSPFCQEGSPFQSSPFCQGGQGGNGGGQQQKFMALGSGVIIDADKGYVVTNNHVVDNATVIKVQLS
 DGRKFDKMGKDPKRSALIQIQNPKNLTAIKMADSDALRVGDYTVAIIGNPFGLGETVTSIVSALGRSGL
 NAENYENFIQTDAAINRGNSSGALVNLNGLIGINTAILAPDGGNIGIGFAIPSNMVKNLTSQMVEYGVKVR
 GELGIMGTELNSELAKAMKVDARQGAFFVSVLPNSSAAKAGIKAGDVITSLNGKPISSFAALRAQVGTMPVG
 SKLTLGLLRDQVNVNLELQQSSQNQVDSSTIFNGIEGAEMSNGKDKQGVVVNVNKTGTFAAQIGLKKGDV
 IIGANQQAVKNIAELRKVLDSKPSVLALNIQRGDSTIYLLMQ

SEQ ID NO 9 è la sequenza di DNA di un gene DegP mutato

ATGAAAAAAAAACCACATTAGCACTGAGTGCCTGGCTCTGAGTTTAGGTTTGGCGTTATCTCCGCTCTCTGCA
 ACGGCGGCTGAGACTTCTTCAGCAACGACAGCCCAGCAGATGCCAAGCCTTGACCCGATGCTCGAAAAGGTG
 ATGCCTTCAGTGGTCAGCATTAACTAGAAAGGTAGCACAACCGTTAATACGCCGCGTATGCCGCGTAATTTT
 CAGCAGTTCTTCGGTATGATTCTCCGTTCTGCCAGGAAGGTTCTCCGTTCCAGAGCTCTCCGTTCTGCCAG
 GGTGGCCAGGGCGGTAATGGTGGCGGCCAGCAACAGAAATTCATGGCGCTGGGTTCCGGCGTCATCATTGAT
 GCCGATAAAGGCTATGTCGTCACCAACAACCACGTTGTTGATAACCGGACGGTCATTAAAGTTCAACTGAGC
 GATGGCCGTAAGTTTCGACGCGAAGATGGTTGGCAAAGATCCGCGCTCTGATATCGCGCTGATCCAAATCCAG
 AACCCGAAAAACCTGACCGCAATTAAGATGGCGGATTCTGATGCACTGCGCGTGGGTGATTACACCGTAGCG
 ATTGGTAACCCGTTTGGTCTGGGCGAGACGGTAACCTCCGGGATTGTCTCTGCGCTGGGGCGTAGCGGCCTG
 AATGCCGAAAACCTACGAAAACCTCATCCAGACCGATGCAGCGATTAATCGTGGTAACGCCGGTGGTGCCTG
 GTTAACCTGAACGGCGAACTGATCGGTATCAACACCGCGATCCTCGCACCGGACGGCGGCAACATCGGTATC
 GGTTTTGCATCCCGAGTAACATGGTGA AAAACCTGACCTCGCAGATGGTGAATACGGCCAGGTGAAACGC
 GGTGAGCTGGGTATTATGGGGACTGAGCTGAACTCCGAACTGGCGAAAAGCGATGAAAGTTGACGCCAGCGC
 GGTGCTTTCGTAAGCCAGGTTCTGCCTAATTCCTCCGCTGCAAAAAGCGGGCATTAAAGCGGGTGATGTGATC
 ACCTCACTGAACGGTAAGCCGATCAGCAGCTTTGCCGCACTGCGTGCTCAGGTGGTACTATGCCGGTAGGC
 AGCAAACTGACCCTGGGCTTACTGCGCGACGGTAAGCAGGTTAACGTGAACCTGGAACCTGCAGCAGAGCAGC
 CAGAATCAGGTTGATTCCAGCTCCATCTTCAACGGCATTGAAGGCGTGAGATGAGCAACAAAGGCAAAGAT
 CAGGGCGTGGTAGTGAACAACGTGAAAACGGGCACTCCGGCTGCGCAGATCGGCCTGAAGAAAGGTGATGTG
 ATTATTGGCGCGAACCAGCAGGCAGTGA AAAACATCGCTGAACTGCGTAAAGTTCTCGACAGCAAACCGTCT
 GTGCTGGCACTCAACATTAGCGCGGGCGACAGCACCATCTACCTGTTAATGCAGTAA

Figura 8F

SEQ ID NO 10 è la sequenza amminoacidica di una proteina DegP mutata

MKKTTLALSALALSGLALSPLSATAAETSSATTAQQMPSLAPMLEKVMPSVVSINVEGSTTVNTPRMPRNF
 QQFFGDDSPFCQEGSPFQSSPFCQGGQGGNGGGQQQKFMALGSGVIIDADKGYVVTNNHVVDNATVIKQVLS
 DGRKFDKAMVKGKDPKSDIALIQIQNPKNLTAIKMADSDALRVGDYTVAIIGNPFGLGETVTSGIVSALGRSGL
 NAENYENFIQTDAAINRGNAGGALVNLNDELIGINTAILAPDGGNIGIGFAIPSNMVKNLTSQMVEYGVKVR
 GELGIMGTELNSELAKAMKVDAQRGAFVVSQVLPNSSAAKAGIKAGDVITSLNGKPISSFAALRAQVGTMPVG
 SKLTLGLLRDQKQVNVNLELQQSSQNQVDSSEIFNGIEGAEMSNKGDQGVVVNNVKTGTTPAAQIGLKKGDV
 IIGANQQAVKNIAELRKVLDSKPSVLAALNIQRGDSTIYLLMQ

SEQ ID NO 11 è la sequenza del primer oligonucleotidico 5' per la regione del gene DegP mutato
 comprendente il sito di restrizione AseI

CTGCCTGCGATTTTCGCCGGAACG

SEQ ID NO 12 è la sequenza del primer oligonucleotidico 3' per la regione del gene DegP mutato
 comprendente il sito di restrizione AseI

CGCATGGTACGTGCCACGATATCC

SEQ ID NO 13 è la sequenza del primer oligonucleotidico 5' per la regione del gene Tsp mutato
 comprendente il sito di restrizione AseI

GGGAAATGAACCTGAGCAAACGC

SEQ ID NO 14 è la sequenza del primer oligonucleotidico 3' per la regione del gene Tsp mutato
 comprendente il sito di restrizione AseI

GGGAAAGGCGCGGAACCGCCTAG

SEQ ID NO 15 è la sequenza del primer oligonucleotidico 5' per la regione del gene Proteasi III
 mutato comprendente il sito di restrizione AseI

CTACTGTGCCAGCGGTGGTAATGG

SEQ ID NO 16 è la sequenza del primer oligonucleotidico 3' per la regione del gene Proteasi III
 mutato comprendente il sito di restrizione AseI

GCATCATAATTTCTTTTTACCTC

SEQ ID NO 17 è la sequenza di DNA del gene spr selvatico

ATGGTCAAATCTCAACCGATTTTGAGATATATCTTGCGCGGGATCCCGCGATTGCAGTAGCGGTTCTGCTT
 TCTGCATGTAGTGCAAATAACACCGCAAAGAATATGCATCCTGAGACACGTGCAGTGGGTAGTGAAACATCA
 TCACTGCAAGCTTCTCAGGATGAATTTGAAAACCTGGTTCGTAATGTCGACGTAAAAATCGCGAATTATGGAT
 CAGTATGCTGACTGGAAGGCGTACGTTATCGTCTGGGCGGCAGCACTAAAAAAGGTATCGATTGTTCTGGT
 TTCGTACAGCGTACATTCGGTGAGCAATTTGGCTTAGAACTTCCGCGTTCGACTTACGAACAGCAGGAAATG
 GGTAAATCTGTTTCCCGCAGTAATTTGCGTACGGGTGATTTAGTTCTGTTCCGTGCCGTTCAACGGGACGC
 CATGTCCGTAATTTATATCGGCAACAATCAGTTTGTCCATGCTTCCACCAGCAGTGGTGTATTATTTCCAGC
 ATGAATGAACCGTACTGGAAGAAGCGTTACAACGAAGCACGCCGGGTTCTCAGCCGCGAGC

Figura 8G

SEQ ID NO 18 è la sequenza amminoacidica del gene spr selvatico inclusa la sequenza segnale che sono i primi 26 residui amminoacidici

MVKSQPILRYILRGIPAI AVAVLLSACSANNTAKNMHPETRAVGSETSSLQASQDEFENLVRNVDVKSRI MDQYADWKGVRYRLGGSTKKGIDCSGFVQRTFREQFGLLELPSTYEQQEMGKSVSRNLRTGDLVLFragSTGRHVGIYIGNNQFVHASTSSGVI I SSMNEPYWKKRYNEARRVLSRS

SEQ ID NO 19 gene spr non mutato senza sequenza segnale

CSANNTAKNMHPETRAVGSETSSLQASQDEFENLVRNVDVKSRI MDQYADWKGVRYRLGGSTKKGIDCSGFVQRTFREQFGLLELPSTYEQQEMGKSVSRNLRTGDLVLFragSTGRHVGIYIGNNQFVHASTSSGVI I SSMNEPYWKKRYNEARRVLSRS

SEQ ID NO 20 sequenza OmpT mutata comprendente le mutazioni D210A e H212A

ATGCGGGCGAACTTCTGGGAATAGTCCTGACAACCCCTATTGCGATCAGCTCTTTTGCTTCTACCGAGACTTTATCGTTTACTCCTGACAACATAAATGCGGACATTAGTCTTGGAACCTCTGAGCGGAAAAACAAAAGAGCGTGTTTATCTAGCCGAAGAAGGAGGCCGAAAAGTCAGTCAACTCGACTGGAAATTC AATAACGCTGCAATTATTAAAGGTGCAATTAATTGGGATTTGATGCCCCAGATATCTATCGGGGCTGCTGGCTGGACAACCTCTCGGCAGCCGAGGTGGCAATATGGTCGATCAGGACTGGATGGATTCCAGTAACCCCGGAACCTGGACGGATGAAAGTAGACACCCTGATACACAACCTCAATTATGCCAACGAATTTGATCTGAATATCAAAGGCTGGCTCCTCAACGAACCC AATTACCGCTGGGACTCATGGCCGGATATCAGGAAAGCCGTTATAGCTTTACAGCCAGAGGTGGTTCCCTATATCTACAGTTCTGAGGAGGGATTCAGAGATGATATCGGCTCCTCCCGAATGGAGAAAGAGCAATCGGCTACAAACAACGTTTTTAAAATGCCCTACATTGGCTTGACTGGAAGTTATCGTTATGAAGATTTTGAACCTCGGTGGCACATTTAAATACAGCGGCTGGGTGGAATCATCTGATAACGCTGAAGCTTATGACCCGGGAAAAAGAATCACTTATCGCAGTAAGGTCAAAGACCAAAATTA CTATTCTGTGTCAGTCAATGCAGGTTATTACGTCACACCTAACGCAAAAGTTTTATGTTGAAGGCGCATGGAATCGGGTTACGAATAAAAAAGGTAATACTTCACTTTTATGATCAC AATAATAACACTTCAGACTACAGCAAAAATGGAGCAGGTATAGAAAACCTATAACTTCATCACTACTGCTGGTCTTAAGTACACATTTTAA

SEQ ID NO 21 sequenza OmpT mutata comprendente le mutazioni D210A e H212A

MRAKLLGIVLTPPIAIISSFASTETLSFTPDNINADISLGTLSGKTKERVYLAEEGGRKVSQLDWKFNNAAII KGAINWDLMPQISIGAGWTTLGSRRGNMVDQDWMDSNPGTWTDSESRHPDTQLNYANEFDLNIKGWLLNEP NYRLGLMAGYQESRYSFTARGGSYIYSSEEGFRDDIGSFPNGERAI GYKQRFKMPYIGLTGSYRYEDFELGGTFKYSGWVSSDNAEAYDPGKRIT YRSKVKDQNYYSVAVNAGYYVTPNAKVYVEGAWNRVTNKKGNTSLYDH NNNTSDYSKNGAGIENYNFITTAGLKYTF

SEQ ID NO 22 sequenza nucleotidica di una sequenza mutata knockout OmpT

ATTCGGGCGAACTTCTGGGAATAGTCCTGACAACCCCTATTGCGATCAGCTCTTTTGCTTCTACCGAGACTTTATCGTTTACTCCTGACAACATAAATGCGGACATTAGTCTTGGAACCTCTGAGCGGAAAAACAAAAGAGCGTGTTTATCTAGCCGAAGAAGGAGGCCGAAAAGTCAGTCAACTCGACTGGAAATTC AATAACGCTGCAATTATTAAAGGTGCAATTAATTGGGATTTGATGCCCCAGATATCTATCGGGGCTGCTGGCTGGACAACCTCTCGGCAGCCGAGGTGGCAATATGGTCGATCAGGACTGGATGGATTCCAGTAACCCCGGAACCTGGACGGATGAAAGTAGACACCCTGATACACAACCTCAATTATGCCAACGAATTTGATCTGAATATCAAAGGCTGGCTCCTCAACGAACCC AATTACCGCTGGGACTCATGGCCGGATATCAGGAAAGCCGTTATAGCTTTACAGCCAGAGGTGGTTCCCTATATCTACAGTTCTGAGGAGGGATTCAGAGATGATATCGGCTCCTCCCGAATGGAGAAAGAGCAATCGGCTACAAACAACGTTTTTAAAATGCCCTACATTGGCTTGACTGGAAGTTATCGTTATGAAGATTTTGAACCTCGGTGGCACATTTAAATACAGCGGCTGGGTGGAATCATCTGATAACGATGAACACTATGACCCGGGAAAAAGAATCACTTATCGCAGTAAGGTCAAAGACCAAAATTA CTATTCTGTGTCAGTCAATGCAGGTTATTACGTCACACCTAACGCAAAAGTTTTATGTTGAAGGCGCATGGAATCGGGTTACGAATAAAAAAGGTAATACTTCACTTTTATGATCAC AATAATAACACTTCAGACTACAGCAAAAATGGAGCAGGTATAGAAAACCTATAACTTCATCACTACTGCTGGTCTTAAGTACACATTTTAA

Figura 8H

SEQ ID NO 23 mostra la sequenza dell'oligonucleotide adattatore OmpA

CGATTGAATGGAGAACCTTGAATTCGGGCGAAACTTCTGGGAATAG

SEQ ID NO 24 mostra una cassetta codificante la sequenza intergenica 1 (IGS1) per l'espressione di Fab in E.coli

AAGTTTTAATAGAGGAGAGTGTTAATGAAGAAG

SEQ ID NO 25 mostra la cassetta oligonucleotidica codificante la sequenza intergenica 2 (IGS2) per l'espressione di Fab in E.coli

AAGTTTTAATAGAGGGGAGTGTAAAATGAAGAAG

SEQ ID NO 26 mostra una cassetta codificante la sequenza intergenica 3 (IGS3) per l'espressione di Fab in E.coli

AAGCTTTAATAGAGGAGAGTGTGAGGAGGAAAAAAAAATGAAGAAA

SEQ ID NO 27 mostra una cassetta codificante la sequenza intergenica 4 (IGS4) per l'espressione di Fab in E.coli

AAGCTTTAATAGAGGAGAGTGTGACGAGGATTATATAATGAAGAAA

SEQ ID NO 28 p la sequenza di DNA del gene FkpA selvatico

ATGAAATCACTGTTTTAAAGTAACGCTGCTGGCGACCACAATGGCCGTTGCCCTGCATGCACCAATCACTTTT
GCTGCTGAAGCTGCAAAACCTGCTACAGCTGCTGACAGCAAAGCAGCGTTCAAAAATGACGATCAGAAATCA
GCTTATGCACTGGGTGCCTCGCTGGGTGCTTACATGGAAAACCTCTCTAAAAGAACAAGAAAACTGGGCATC
AAACTGGATAAAGATCAGCTGATCGCTGGTGTTCAGGATGCATTTGCTGATAAGAGCAAACCTCTCCGACCAA
GAGATCGAACAGACTCTACAAGCATTCGAAGCTCGCGTGAAGTCTTCTGCTCAGGCCGAAGATGGAAAAAGAC
GCGGCTGATAACGAAGCAAAAAGGTAAAGAGTACCGCGAGAAAATTTGCCAAAGAGAAAAGGTGTGAAAACCTCT
TCAACTGGTCTGGTTTATCAGGTAGTAGAAGCCGGTAAAGGCGAAGCACCAGAAAGACAGCGATACTGTTGTA
GTGAACTACAAAGGTACGCTGATCGACGGTAAAGAGTTCGACAACCTTTACACCCGTGGTGAACCGCTTTCT
TTCCGCTCTGGACGGTGTATCCCGGGTTGGACAGAAGGTCTGAAGAACATCAAGAAAGGCGGTAAGATCAAA
CTGGTTATTCACCAGAACTGGCTTACGGCAAAGCGGGTGTTCGGGGATCCCACCGAATTCTACCCTGGTG
TTTGACGTAGAGCTGCTGGATGTGAAACCAGCGCCGAAGGCTGATGCAAAGCCGGAAGCTGATGCGAAAGCC
GCAGATTCCTGCTAAAAAA

SEQ ID NO 29 è la sequenza proteica del gene FkpA selvatico

MKSLFKVTLTLLATTMAVALHAPITFAAEAAKPATAADSKAAFKNDDQKSAYALGASLGRYMENSLKEQEKLGI
KLDKDLQLIAGVQDAFADKSKLSDQEIEQTLQAFEARKVSSAQAKMEKDAADNEAKGKEYREKFAKEKGVKTS
STGLVYQVVEAGKGEAPKSDTVVVVNYKGLIDGKEFDNSYTRGEPLSFRLDGVIPGWTEGLKNIKKGGKIK
LVI PPELAYGKAGVPGI PPNSTLVFDVELLDVVKPAPKADAKPEADAKAADSARK

SEQ ID NO 30 è la sequenza del gene FkpA con His-tag

ATGAAATCACTGTTTTAAAGTAACGCTGCTGGCGACCACAATGGCCGTTGCCCTGCACGCACCAATCACTTTT
GCTGCTGAAGCTGCAAAACCTGCTACTGCTGCTGACAGCAAAGCAGCGTTCAAAAATGACGATCAGAAATCA
GCTTATGCACTGGGTGCCTCGCTGGGTGCTTACATGGAAAACCTCTCTAAAAGAACAAGAAAACTGGGCATC
AAACTGGATAAAGATCAACTGATCGCTGGTGTTCAGGATGCATTTGCTGATAAGAGCAAACCTCTCCGACCAA
GAGATCGAACAGACTCTACAAGCATTTGAAGCTCGCGTGAAGTCTTCTGCTCAGGCCGAAGATGGAAAAAGAC
GCGGCTGATAACGAAGCAAAAAGGTAAAGAGTACCGCGAGAAAATTTGCCAAAGAGAAAAGGTGTGAAAACCTCT
TCAACTGGTCTGGTTTATCAGGTAGTAGAAGCCGGTAAAGGCGAAGCACCAGAAAGACAGCGATACTGTTGTA
GTGAACTACAAAGGTACGCTGATCGACGGTAAAGAGTTCGACAACCTTTACACCCGTGGTGAACCGCTTTCT
TTCCGCTCTGGACGGTGTATCCCGGGTTGGACAGAAGGTCTGAAGAACATCAAGAAAGGCGGTAAGATAAAA
CTGGTTATTCACCAGAACTGGCTTACGGCAAAGCGGGTGTTCGGGGATTCACCAAATTCTACCCTGGTG
TTTGACGTAGAGCTGCTGGATGTGAAACCAGCGCCGAAGGCTGATGCAAAGCCGGAAGCTGATGCGAAAGCC
GCAGATTCCTGCTAAAAAACACCATCACCATCACCAC

Figura 8 I

SEQ ID NO 31 è la sequenza proteica del gene FkpA con His-tag

MKSLFKVTL LATTMAVALHAPITFAAEAAKPAATAADSKAAAFKNDDQKSAYALGASLGRYMENSLKEQEKLGI
 KLDKQDLIAGVQDAFADKSKLSDQEIEQTLQAFEARVKSSAQAKMEKDAADNEAKGKEYREKFAKEKGVKTS
 STGLVYQVVEAGKGEAPKDSDTVVVNYKGTLDGKEFDNSYTRGEPLSFRLDGVIPGWTEGLKNIKKGGKIK
 IVIPELAYGKAGVPGIIPNSTLVFDVELLDVVKPAPKADAKPEADAKAADS AKKHHHHHH

SEQ ID NO 32 è la sequenza del gene *skp* selvatico

GTGAAAAAGTGGTTATTAGCTGCAGGTCTCGGTTTAGCACTGGCAACTTCTGCTCAGGCGGCTGACAAAATT
 GCAATCGTCAACATGGGCAGCCTGTTCCAGCAGGTAGCGCAGAAAACCGGTGTTTCTAACACGCTGGAAAAT
 GAGTTCAAAGGCCGTGCCAGCGAACTGCAGCGTATGGAAACCGATCTGCAGGCTAAAATGAAAAAGCTGCAG
 TCCATGAAAGCGGGCAGCGATCGCTACTAAGCTGGAAAAGACGTGATGGCTCAGCGCCAGACTTTTGTCTCAG
 AAAGCGCAGGCTTTTGTGAGCAGGATCGCGCACGTCGTTCCAACGAAGAACCGGCAAACCTGGTTACTCGTATC
 CAGACTGCTGTGAAATCCGTTGCCAACAGCCAGGATATCGATCTGGTTGTTGATGCAAACGCCGTTGCTTAC
 AACAGCAGCGATGTAAAAGACATCACTGCCGACGTACTGAAACAGGTTAAATAA

SEQ ID NO 33 è la sequenza proteica del gene *skp* selvatico

MKKWLLAAGLGLALATSAQAADKIÀIVNMGSLFQQVAQKTGVSN̄TLENEFKGRASELQRMETDLQAKMKKLQ
 SMKAGSDRTKLEKDVMAQRQTFAQKAQAFEQDRARRSNEERGKLVTRIQTAVKSVANSQDIDLVDANAVAY
 NSSDVKDI TADVLKQVK

SEQ ID NO 34 è la sequenza del gene *skp* con His-tag

ATGAAAAAGTGGTTATTAGCCGCAGGTCTCGGTTTAGCACTGGCAACTTCTGCTCAGGCGGCTGACAAAATT
 GCAATCGTCAACATGGGCAGCCTGTTCCAGCAGGTAGCGCAGAAAACCGGTGTTTCTAACACGCTGGAAAAT
 GAGTTCAAAGGCCGTGCCAGCGAACTCCAGCGTATGGAAACCGATCTCCAGGCTAAAATGAAAAAGCTGCAA
 TCCATGAAAGCGGGCAGCGATCGCTACTAAGCTGGAAAAGACGTGATGGCTCAGCGCCAGACTTTTGTCTCAG
 AAAGCGCAGGCTTTTGTGAGCAGGATCGCGCACGTCGTTCCAACGAAGAACCGGCAAACCTGGTTACTCGTATC
 CAGACTGCTGTGAAATCCGTTGCCAACAGCCAGGAAATCGATCTGGTTGTTGATGCAAACGCCGTTGCTTAC
 AACAGCAGCGATGTAAAAGACATCACTGCCGACGTACTGAAACAGGTTAAACACCATCACCATCACCAC

SEQ ID NO 35 è la sequenza proteica del gene *skp*

MKKWLLAAGLGLALATSAQAADKIÀIVNMGSLFQQVAQKTGVSN̄TLENEFKGRASELQRMETDLQAKMKKLQ
 SMKAGSDRTKLEKDVMAQRQTFAQKAQAFEQDRARRSNEERGKLVTRIQTAVKSVANSQEIDLVDANAVAY
 NSSDVKDI TADVLKQVKHHHHHH

Sequenze Ab CA170_1519**CDRH1**

GFTFSNYGMV SEQ ID NO: 36

CDRH2

YIDSDGDNTYYRDSVKG SEQ ID NO: 37

CDRH3

GIVRPFLY SEQ ID NO: 38

CDRL1

KSSQSLVGASGKTYLY SEQ ID NO: 39

CDRL2

LVSTLDS SEQ ID NO: 40

CDRL3

LQGTHFPHT SEQ ID NO: 41

Figura 8J

regione Ab 1519 VL di ratto SEQ ID NO: 42

DVVMTQTPLS LSV[~]ALGQPAS ISCKSSQSLV GASGKTYLYW LFQRSGQSPK RLIYLVSTLD
SGIPDRFSGS GAETDFTLKI RRVEADDLGV YYCLQGTHFP HTFGAGTKLE LK

regione Ab 1519 VL di ratto SEQ ID NO: 43

gatg[~]ttgtga tg[~]accagac tccactgtct ttgtcgggtg cccttggaca accagcctcc
atctcttgca agtcaagtca gagcctcgtg ggtgctagtg gaaagacata tttgtattgg
ttatttcaga ggtccggcca gtctccaaag cgactaatct atctggtgct cacactggac
tctggaattc ctgataggtt cagtggcagt ggagcagaga cagattttac tcttaaaatc
cgcagagtgg aagccgatga tttgggagtt tattactgct tgcaaggtac acattttcct
cacacgtttg gagctgggac caagctggaa ttgaaa

regione Ab 1519 VL di ratto con sequenza segnale in corsivo sottolineata SEQ ID NO: 44

MMSPAQFLFLMLWIQGTSGDVVMTQTPLSLSVALGQPASISCKSSQSLVGASGKTYLYWLFQRSGQSPK
RLIYLVSTLDSGIPDRFSGSGAETDFTLKI RRVEADDLGVYYCLQGTHFPHTFGAGTKLELK

regione Ab 1519 VL di ratto con sequenza segnale in corsivo sottolineata SEQ ID NO: 45

atgatgagtc ctgcccagtt cctg[~]tttctg ctgatgctct ggattcaggg aaccagtggt
gatg[~]ttgtga tg[~]accagac tccactgtct ttgtcgggtg cccttggaca accagcctcc
atctcttgca agtcaagtca gagcctcgtg ggtgctagtg gaaagacata tttgtattgg
ttatttcaga ggtccggcca gtctccaaag cgactaatct atctggtgct cacactggac
tctggaattc ctgataggtt cagtggcagt ggagcagaga cagattttac tcttaaaatc
cgcagagtgg aagccgatga tttgggagtt tattactgct tgcaaggtac acattttcct
cacacgtttg gagctgggac caagctggaa ttgaaa

regione Ab 1519 VH di ratto SEQ ID NO: 46

EVPLVESGGG SVQPGRSMKL SCVVS[~]GFTFS NYGMVWVRQA PKKGLEWVAY IDSDGDNTYY
RDSVKGRFTI SRNNAKSTLY LQMDSLRS[~]ED TATYYCTGI VRPFLYWQGG TTVTVS

regione Ab 1519 VH di ratto SEQ ID NO: 47

gaggtgccgc tgg[~]tggagtc tggggg[~]cggc tcagatgcagc ctgggaggtc catgaaactc
tcctgtgtag tctcaggatt cactttcag[~]t aattatggca tgg[~]tctgggt ccgccaggct
caaagaagg gtctggagtg ggtcgcata[~]t attgattctg atgggtgataa tacttactac
cgagattccg tgaagggccg attcactatc tccagaaata atgcaaaaag caccctatat
ttgcaaatgg acagtctgag gtctgaggac acggccactt attactgtac aacagggatt
gtccggccct ttctctattg gggccaagga accacggtca ccgtctcg

regione Ab 1519 VH di ratto con sequenza segnale sottolineata in corsivo SEQ ID NO: 48

MDISLSLAF[~]L VLFIKGVRCE VPLVESGGGS VQPGRSMKLS CVVSGFTFSN YGMVWVRQAP
KKGLEWVAYI DSDGDNTYYR DSVKGRFTIS RNNAKSTLYL QMDSLRS[~]EDT ATYYCTTGIV
RPFLYWQGGT TTVTVS

regione Ab 1519 VH di ratto con sequenza segnale sottolineata in corsivo SEQ ID NO: 49

atggacatca gtctcagctt ggctttcctt gtccttttca taaaagggtg ccggtgtgag
gtgccgctgg tggagtctgg gggcggctca gtgcagcctg ggaggtccat gaaactctcc
tgtgtagtct caggattcac tttcagta[~]at tatggcatgg tctgggtccg ccaggctcca
aagaagggtc tggagtgggt cgcata[~]tatt gattctgatg gtgataatac ttactaccga
gattccgtga agggccgatt cactatctcc agaaataatg caaaaagcac cctatatttg
caaatggaca gtctgaggtc tgaggacacg gccacttatt actgtacaac agggattgtc
cggccctttc tctattgggg ccaaggaacc acggtcaccg tctcg

Figura 8k

regioneV 1519 gL20 SEQ ID NO: 50

DIQMTQSPSS LSASVGDRV ITCKSSQSLV GASGKTYLYW LFQKPGKAPK RLIYLVSTLD
SGIPSRFSGS GSGTEFTLTI SSLQPEDFAT YYCLQGTHFP HTFGQGTKLE IK

regioneV 1519 gL20 SEQ ID NO: 51

gatatccaga tgaccagag tccaagcagt ctctccgcca gcgtaggcga tcgtgtgact
attacctgta aaagctccca gtcctgggtg ggtgcaagcg gcaaaccta cctgtactgg
ctctccaga aaccgggcaa agctccgaaa cgcctgatct atctggtgtc taccctggat
agcggatatt cgtctcgttt ctccggtagc ggtagcggta ccgaattcac gctgaccatt
agctccctcc agccggagga ctttgctacc tattactgcc tccagggcac tcattttccg
cacactttcg gccagggtac caaactggaa atcaaa

regioneV 1519 gL20 con sequenza segnale sottolineata in corsivo

SEQ ID NO: 52

MKKTAI[~]AI[~]AV[~]AL[~]AG[~]FATVAQ[~]ADIQMTQSPSSLSASVGDRVITITCKSSQSLVGASGKTYLYWLFQKPGKAPKR
LIYLVSTLDSGIPSRFSGSGSGTEFTLTISSLQPEDFATYYCLQGTHFPHTFGQGTKLEIK

regioneV 1519 gL20 con sequenza segnale sottolineata in corsivo

SEQ ID NO: 53

atgaaaaaga cagctatcgc aattgcagtg gccttgctg gtttcgctac cgtagcgcaa
gctgatatcc agatgaccca gagtccaagc agtctctccg ccagcgtagg cgatcgtgtg
actattacct gtaaaagctc ccagtcctgt gtgggtgcaa gcggcaaac ctacctgtac
tggctcttcc agaaaccggg caaagctccg aaacgcctga tctatctggt gtctaccctg
gatagcggta ttccgtctcg tttctccggt agcggtagcg gtaccgaatt cacgctgacc
attagctccc tccagccgga ggactttgct acctattact gcctccaggg cactcatttt
ccgcacactt tcggccaggg taccaaactg gaaatcaaa

catena leggera 1519 gL20 (V+costante)SEQ ID NO: 54

DIQMTQSPSS LSASVGDRV ITCKSSQSLV GASGKTYLYW LFQKPGKAPK RLIYLVSTLD
SGIPSRFSGS GSGTEFTLTI SSLQPEDFAT YYCLQGTHFP HTFGQGTKLE IKRTVAAPSV
FI[~]FPPSDEQL KSGTASV[~]VCL LNNFY[~]PREAK VQWKVDNALQ SGNSQESVTE QDSKDSTYSL
SSTLTLSKAD YEKHKVYACE VTHQGLSSPV TKSFN[~]RGEC

catena leggera 1519 gL20 (V+costante)SEQ ID NO: 55

gatatccaga tgaccagag tccaagcagt ctctccgcca gcgtaggcga tcgtgtgact
attacctgta aaagctccca gtcctgggtg ggtgcaagcg gcaaaccta cctgtactgg
ctctccaga aaccgggcaa agctccgaaa cgcctgatct atctggtgtc taccctggat
agcggatatt cgtctcgttt ctccggtagc ggtagcggta ccgaattcac gctgaccatt
agctccctcc agccggagga ctttgctacc tattactgcc tccagggcac tcattttccg
cacactttcg gccagggtac caaactggaa atcaaacgta cggtagcggc cccatctgtc
ttcatcttcc cgccatctga tgagcagttg aaatctggaa ctgcctctgt tgtgtgctg
ctgaataact tctatcccag agaggccaaa gtacagtgga aggtggataa cgccctcaa
tcgggtaact ccagggagag tgtcacagag caggacagca aggacagcac ctacagcctc
agcagcacc[~] tgacgctgag caaagcagac tacgagaaac acaaagtcta cgctgcgaa
gtcaccatc[~] agggcctgag ctaccagta acaaaaagtt ttaatagagg ggagtgt

catena leggera 1519 gL20 con sequenza segnale sottolineata e in corsivo SEQ ID NO: 56

MKKTAI[~]AI[~]AV[~] AL[~]AG[~]FATVAQ[~]ADIQMTQSPSSLSASVGDRV TITCKSSQSL VGASGKTYLY
WLFQKPGKAP KRLIYLVSTL DSGIPSRFSG SSGTEFTLT ISSLQPEDFA TYYCLQGTHF
PHTFGQTKL EIKRTVAAPS VFIFPPSDEQ LKSGTASVVC LLNNFYPREA KVQWKVDNAL
QSGNSQESVT EQDSKDSTYS LSSTLTLSKA DYKHKVYAC EVTHQGLSSP VTKSFNRGEC

Figura 8L

catena leggera 1519 gL20 con sequenza segnale sottolineata in corsivo SEQ ID NO: 57
atgaaaaaga cagctatcgc aattgcagtg gccttggctg gtttcgctac cgtagcgcaa
gctgatatcc agatgacca gagtccaagc agtctctccg ccagcgtagg cgatcgtgtg
actattacct gtaaaagctc ccagtcctcg gtgggtgcaa gcggcaaac ctacctgtac
tggetcttcc agaaaccggg caaagctccg aaacgcctga tctatctggt gtctaccctg
gatagcggta ttccgtctcg tttctccggg agcggtagcg gtaccgaatt cacgctgacc
attagctccc tccagccgga ggactttgct acctattact gcctccaggg cactcatttt
ccgcacactt tccggccaggg taccaaactg gaaatcaaac gtacggtagc ggccccatct
gtcttcatct tcccgccatc tgatgagcag ttgaaatctg gaactgcctc tgttgtgtgc
ctgctgaata acttctatcc cagagaggcc aaagtacagt ggaagggtga taacgccctc
caatcgggta actcccagga gagtgtcaca gagcaggaca gcaaggacag cacctacagc
ctcagcagca ccctgacgct gagcaaagca gactacgaga aacacaaagt ctacgcctgc
gaagtcaccc atcagggcct gagctcacca gtaacaaaaa gttttaatag aggggagtg

1519 gH20 regione V

SEQ ID NO: 58

EVPLVESGGG LVQPGGSLRL SCAVSGFTFS NYGMVWRQA PGKGLEWVAY IDSDGDNTYY
RDSVKGRFTI SRDNAKSSLY LQMNSLRAED TAVYYCTTGI VRPFLYWGGQ TLVTVS

1519 gH20 regione V

SEQ ID NO: 59

gaggttccgc tggctcgagtc tggaggcggg cttgtccage ctggagggag cctgcgtctc
tcttgtgcag tatctggctt caggttctcc aactacggta tgggtgtggg tgcgcaggct
ccaggtaaag gtctggaatg ggtggcgtat attgactccg acggcgacaa cacctactat
cgcgactctg tgaaggtcg cttcaccatt tcccgcgata acgcaaaatc cagcctgtac
ctgcagatga acagcctgcg tgctgaagat actgcggtgt actattgcac cactggcatc
gtgcgtccgt ttctgtattg gggtcagggt accctcgta ctgtctcg

1519 gH20 regione V con sequenza segnale sottolineata in corsivo

SEQ ID NO: 60

MKKTAIATAV ALAGFATVAQ AEVPLVESGG GLVQPGGSLR LSCAVSGFTF SNYGMVWRQA
APGKGLEWVA YIDSDGDNTY YRDSVKGRFT ISRDNAKSSL YLQMNSLRAE DTAVYYCTTG
IVRPFLYWGGQ GTLVTVS

1519 gH20 regione V con sequenza segnale sottolineata in corsivo

SEQ ID NO: 61

atgaagaaga ctgctatagc aattgcagtg gcgctagctg gtttcgccac cgtggcgcaa
gctgaggttc cgctggtcga gtctggaggc gggcttgtec agcctggagg gagcctgctg
ctctcttgtg cagtatctgg cttcacgttc tccaactacg gtatggtgtg ggttcgtcag
gctccaggta aaggctgga atgggtggcg tatattgact ccgacggcga caacacctac
tatcgcgact ctgtgaaagg tcgcttcacc atttcccgcg ataacgcaa atccagcctg
tacctgcaga tgaacagcct gcgtgctgaa gatactgcgg tgtactattg caccactggc
atcgtgctgc cgtttctgta ttggggtcag ggtaccctcg ttactgtctc g

1519 gH20 catena pesante Fab' (V+gamma umano CH1+ hinge)

SEQ ID NO: 62

EVPLVESGGG LVQPGGSLRL SCAVSGFTFS NYGMVWRQA PGKGLEWVAY IDSDGDNTYY
RDSVKGRFTI SRDNAKSSLY LQMNSLRAED TAVYYCTTGI VRPFLYWGGQ TLVTVSSAST
KGPSVFPLAP SSKSTSGGTA ALGCLVKDYF PEPVTVSWNS GALTSGVHTF PAVLQSSGLY
SLSSVVTVPS SSLGTQTYIC NVNHKPSNTK VDKKVEPKSC DKTHTCOA

Figura 8M

1519gH20 catena pesante Fab' (V+gamma umano I-CH1+ hinge) SEQ ID NO: 63
gaggttccgc tggtcgagtc tggaggcggg cttgtccagc ctggaggag cctgcgtctc
tcttgtgcag tatctggctt cacgttctcc aactacggta tgggtgtgggt tcgtcaggct
ccaggtaaag gtctggaatg ggtggcgtat attgactccg acggcgacaa cacctactat
cgcgactctg tgaaaggctg cttcaccatt tcccgcgata acgccaatc cagcctgtac
ctgcagatga acagcctgcg tgetgaagat actgcggtgt actattgcac cactggcatc
gtgctgcccgt ttctgtattg gggtcagggt accctcgta ctgtctcgag cgcttctaca
aagggcccat cggctctccc cctggcacc cctccaaga gcacctctgg gggcacagcg
gccctgggct gcctgggtcaa ggactacttc cccgaaccgg tgacgggtgtc gtggaactca
ggcgccctga ccagcggcgt gcacaccttc ccggctgtcc tacagtctc aggactctac
tcctcagca gcgtgggtgac cgtgccctcc agcagcttg gcaccagac ctacatctgc
aacgtgaatc acaagcccag caacaccaag gtcgacaaga aagttgagcc caaatcttgt
gacaaaactc acacatgctc cgcg

1519gH20 catena pesante Fab' con sequenza segnale sottolineata in corsivo SEQ ID NO: 64
MKKTAIATAV *ALAGFATVAQ* *AEVPLVESGG* *GLVQPGSLR* *LSCAVSGFTF* *SNYGMVWRQ* *APGKGLEWVA*
YIDSDGDNTY YRDSVKGRFT ISRDNKSSL YLQMNSLRAE DTAVYYCTTG IVRPFLYWQO GTLTVSSAS
TKGPSVFPLA PSSKSTSGGT AALGCLVKDY FPEPVTWSN SGALTSVHT FPAVLQSSGL YLSSVVTVP
SSSLGTQTYI CNVNHKPSNT KVDKKEPKS CDKTHCAA

1519gH20 catena pesante Fab' con sequenza segnale sottolineata in corsivo SEQ ID NO: 65
atgaagaaga ctgctatagc aattgcagtg gcgctagctg gtttcgccac cgtggcgcaa
gctgaggttc *cgctgggtcga* *gtctggagge* *gggctgtcc* *agcctggagg* *gagcctgctg*
ctctcttgtg cagtatctgg cttcacgttc tccaactacg gtatgggtgtg ggttcgtcag
gctccaggta aaggtctgga atgggtggcg tatattgact ccgacggcga caacacctac
tatcgcgact ctgtgaaagg tcgcttcacc atttcccgcg ataacgcaa atccagcctg
tacctgcaga tgaacagcct gcgtgctgaa gatactgctg tgtactattg caccactggc
atcgtgctgc cgtttctgta ttgggggtcag ggtaccctcg ttactgtctc gagcgttct
acaaagggcc catcgggtctt ccccctggca ccctcctcca agagcacctc tgggggcaca
gcggccctgg gctgcctggt caaggactac ttccccgaac cggtgacgggt gtcgtggaac
tcaggcgccc tgaccagcgg cgtgcacacc ttcccggctg tctacagtc ctcaggactc
tactccctca gcagcgtggt gaccgtgccc tccagcagct tgggcacca gacctacatc
tgcaacgtga atcacaagcc cagcaacacc aaggtcgaca agaaagttga gcccaaactc
tgtgacaaaa ctcacacatg cgcgcg

1519gH20 IgG4 catena pesante (V+gamma umano-4P costante) SEQ ID NO: 66
EVPLVESGGG LVQPGSLRL SCAVSGFTFS NYGMVWRQA PGKGLEWVAY IDSDGDNTYY
RDSVKGRFTI SRDNKSSLY LQMNSLRAED TAVYYCTTGI VRPFLYWQO GTLTVSSAST
KGPSVFPLAP CSRSTSESTA ALGCLVKDYF PEPVTWSNS GALTSVHTF PAVLQSSGLY
YLSSVVTVPS SSGTKTYTC NVDHKPSNTK VDKRVESKYG PPCPPCPAPE FLGGPSVFLF
PPKPKDTLMI SRTPEVTCVV VDVSQEDPEV QFNWYVDGVE VHNAKTKPRE EQFNSTYRVV
SVLTVLHQDW LNGKEYKCKV SNKGLPSSIE KTISKAKGQP REPQVYTLPP SQEEMTKNQV
SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSRLTVD KSRWQEGNVF
SCSVMHEALH NHYTQKLSL SLGK

Figura 8N

1519gH20 IgG4 catena pesante (V+gamma umano-4P costante, esoni sottolineati) SEQ ID NO: 67
gaggtaccacttgtggaaagcggaggaggtcttgtgcagcctggaggaagtttacgtctctcttgtgctgtgtctggctt
caccttctccaattacggaatgggtctgggtcagacaagcacctggaaaggtcttgaatgggtggcctatatgtactctg
acggggacaacacctaactatcgggatcccgtaagggagccttcacaatctcccagagataacgccaagagctcactgtac
ctgcagatgaatagcctgagagccgaggatactgccgtgtaactattgcacaacgggaatcgtttaggccttttctgtactg
gggacagggcaccttgggttactgtctcgagcgttctacaaagggccatccgctcttccccctggcgccttgcctccagga
gcacctccgagagcacagccgcccctgggtgcttcaaggactacttcccgaaccgggtgacgggtgctgtggaactca
ggcgcctgaccagcggcgtgcacaacttcccggctgtctacagctctcaggactctactccctcagcagcgtgggtgac
cgctgcccctccagcagcttgggacgaagacctacacctgcacaagcccagcaacaccaaggtggacaagagagttcgggga
gaggccagcacagggaggagggtgtctgtctggaagccaggtcagccctcctgctggagcaccocgctgtgcagcc
ccagcccagggcagcaaggcatgccccatctgtctcctcaccggaggcctctgaccaccccactcatgcccagggagag
ggtcttctggatttttccaccaggtccgggcagccacaggctggatgcccctaccccaggccctgcgcatacaggggca
ggtgtctgcgtcagacctgccaagagccataccggggaggacctgcccctgacctaaagcccacccaaagccaaactc
tccactccctcagctcagacaccttctctcctcccagatctgagtaactcccaatcttctctctgagagtcctcaaatg
gtcccccatgcccacatgcccaggtaagccaaccaggcctcgcctccagctcaaggcgggacaggtgcccctagagta
gcctgcatccagggacaggcccagccgggtgctgacgcatccacctccatctcttctcagcaccctgagttcctcggggg
gacctcagcttctcctgttcccccaaaaacccaaggacactctcatgactctcccggaccctgaggtcagctgctgggtg
gtggagctgagccaggaagaccccaggtccagttcaactggtagctggatggcgtggagggtgcataatgccaagacaaa
gcccggggaggagcagttcaacagcagcgtaccgtgtgggtcagcgtcctcaccgctcaccaggactggctgaacggca
aggagtacaagtgcaaggtctccaacaaaggtcctcccctccatcgagaaaacacctcccaagccaaaggtgggacc
cacggggtgagggccacatggacagaggtcagctcggcccacctctgcccctgggagtgaccgctgtgccaacctctg
tccctacagggcagcccggagagccacaggtgtacacctgccccatcccaggaggagatgaccaagaaccagggtcagc
ctgacctgcttgggtcaaaggcttctaccccagcgacatcgccgtggagtgaggagcaatgggcagccggagaaacta
caagaccacgctcccgtgctggactccgacggctccttcttctctacagcaggttaaccgtggacaagagcaggtggc
aggaggggaatgtcttctcatgtcctcgatgcatgaggtctgtgcacaaccactacacacagaagagcctctccctgtct
ctgggtaaa

1519gL20 FabFv catena leggera

SEQ ID NO: 68

DIQMTQSPSS LSASVGRDVT ITCKSSQSLV GASGKTYLYW LFQKPGKAPK RLIYLVSTLD
 SGIPSRFSGS GSGTEFTLTI SSLQPEDFAT YYCLQGFHP HTFGQGTKLE IKRTVAAPSV
 FIFPPSDEQL KSGTASVVCL LNNFYPRKAK VQWKVDNALQ SGNSQESVTE QDSKSTYSL
 SSSLTLLSKAD YEKHKVYACE VTHQGLSSPV TKSFNRRGECG GGGGSGGGGS GGGGSDIQMT
 QSPSSVSASV GDRVTITCQS SPSVWSNFLS WYQQKPGKAP KLLIYEASKL TSGVPSRFSG
 SSGSTDFTLT ISSLQPEDFA TYYCGGYS ISDTTFGCGT KVEIKRT

1519gL20 FabFv catena leggera

SEQ ID NO: 69

gatataccaga tgaccacagag cccatctagc ttatccgctt ccggttgggtga tgcgctgaca
 attacgtgta agagctccca atctctcgtg ggtgcaagtg gcaagacctt tctgtactgg
 ctctttcaga agcctggcaa ggcacaaaaa cggctgatct atctgggtgc tacccttgac
 tctgggatac cgtcacgatt ttccggatct gggagcggaa ctgagttcac actcacgatt
 tcatcgtgc aaccagagga ctttctacc tactactgcc tgcaaggcac tcatctccct
 cacactttcg gccaggggac aaaactcgaa atcaaacgta cggtagcggc cccatctgtc
 ttcactttcc cgccatctga tgagcagttg aaatctggaa ctgectctgt tgtgtgcctg
 ctgaataact tctatcccag agaggccaaa gtacagtga aggtggataa cgccctccaa
 tccgggtaact cccaggagag tgtcacagag caggacagca aggacagcac ctacagcctg
 agcagcacc tgacgctgtc taaagcagac tacgagaaac acaaagtga cgcctgcaaa
 gtcacccatc agggcctgag ctcaccagta acaaaaagt ttaatagagg ggaggtgagc
 ggtggcgggtg gcagtggtgg gggaggctcc ggaggtggcg gttcagacat acaaatgacc
 cagagtcctt catcggatc cgcgtccggt ggcgataggg tgactattac atgtcaaaagc
 tctcctagcg tctggagcaa ttttctatcc tggatcaac agaaaccggg gaaggctcca
 aaacttctga tttatgaagc ctcgaaactc accagtgag ttccgtcaag attcagtggc
 tctggatcag ggacagactt cacgttgaca atcagttcgc tgcaaccaga ggactttgag
 acctactatt gtgggtggag ttacagtagc ataagtgata cgacatttgg gtgocgggtact
 aagggtgaaa tcaaacgtac c

Figura 80**1519gL20 FabFv catena leggera con sequenza segnale sottolineata in corsivo SEQ ID NO 70**

*MSVPTQVLGLLLLLWLT*DARCDIQMTQSPSSLSASVGDVRTITCKSSQSLVGASGKTYLYWLFQKPGKAPKRLIYLV
 STLDSGIPSRFSGSGSGTEFTLTISLQPEDFATYYCLOGTHFPHTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKS
 GTASVVCLLNFFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSTLTLSKADYEKHKVYACEVTHQGLSS
 PVTKSFNRGECSSGGGSGGGGSDIQMTQSPSSVSASVGDVRTITCQSSPSVWSNFLSWYQQKPGKAPKLLI
 YEASKLTSGVPSRFSGSGSGTDFTLTISLQPEDFATYYCGGGYSSISDTTFGCGTKVEIKRT

1519gL20 FabFv catena leggera con sequenza segnale sottolineata in corsivo SEQ ID NO 71

*atgtctgtccccacccaagt**cctcggactcctgctactctggcttacagatgccagatg*cgataatccagatg
 acccagagcccattctagcttattccgcttccgttgggtgatcgcgtgacaattacgtgtaagagctccaatct
 ctcgtgggtgcaagtggcaagacctatctgactggctctttcagaagcctggcaaggcaccaaaacggctg
 atctatctggtgtctacccttgactctgggataccgtcacgattttccggatctgggagcggaaactgagttc
 aactcacgatttcatcgctgcaacccgaggactttgctacctactgctgcaaggcactcatttccct
 cacactttcggccaggggacaaaactcgaaatcaaacgtacggtagcggccccatctgtcttcatcttcccg
 ccatctgatgagcagttgaaatctggaactgcctctgttgtgtgctgctgaataacttctatcccagagag
 gccaaagtacagtgggaaggtggataacgccctccaatcgggtaactccaggagagtgtcacagagcaggac
 agcaaggacagcacctacagcctgagcagcaccctgacgctgtctaaagcagactacgagaaacacaaagtg
 tacgectgcgaagtacccatcagggcctgagctcaccagtaacaaaaagtttaatagaggggagtgtagc
 ggtggcggtggcagtggtgggggaggctccggaggtggcggttcagacatacaaatgaccagagtccttca
 tcggtatccgctccgttggcgataggtgactattacatgtcaaagctctcctagcgtctggagcaat
 ctatcctggtatcaacagaaaccggggaaggctccaaaacttctgatttatgaagcctcgaaactcaccagt
 ggagttccgtcaagattcagtggtctggatcaggacagacttcagttgacaatcagttcgtgcaacca
 gaggactttgcgacctactattgtggtggaggttacagtagcataagtgatacgacatttgggtgcggtact
 aaggtggaatcaacgtacc

1519gH20 FabFv catena pesante SEQ ID NO 72

EVPLVESGGGLVQPGGSLRLSCAVSGFTFSNYGMVWRQAPGKGLEWVAYIDSDGDNTYYRDSVKGRFTISR
 DNAKSSLYLQMNLSRAEDTAVYYCTTGIVRPFYWGQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGC
 LVKDYFPEPVTVSWNSGALTSKVHTFPAVLQSSGLYSLSSVTVPSLGLTQTYICNVNHKPSNTKVDKKE
 PKSCSGGGGSGGGTGGGGSEVQLLESGLLVQPGGSLRLSCAVSGIDLSNYAINWVRQAPGKCLEWIGI IW
 ASGTTFYATWAKGRFTISRDNKNTVYLQMNLSRAEDTAVYYCARTVPGYSTAPYFDLWGQGLTVTVSS

1519gH20 FabFv catena pesante SEQ ID NO 73

gaggtaccacttgtggaagcggaggaggtcttgtgcagcctggaggaagtttacgtctctcttgtgctgtg
 tctggcttcccttctccaattacggaatggctctgggtcagacaagcacctggaaagggctctgaaatgggtg
 gcctatattgactctgacgggacacacctactatcgggattccgtgaaaggacgcttcacaatctccga
 gataacgccaaagagctcactgtacctgcagatgaatagcctgagagccgaggataactgccgtgtactattgc
 acaacgggaatcgttaggccttttctgtactggggacagggcaccttggttactgtctcgagcgcgtccaca
 aagggcccatcggctctccccctggcacctcctccaagagcacctctgggggcacagcggccctgggctgc
 ctggtcaaggactacttccccgaaccagtgcaggtgtcgtggaaactcaggtgccctgaccagcggcgttca
 acctccccggtgtcctacagctctcaggactctactccctgagcagcgtggtgaccgtgccctccagcagc
 ttgggcacccagacctacatctgcaacgtgaatcacaagccagcaacaccaaggtcgataagaaagttag
 ccaaatcttgtagtgagggtgggggctcaggtggaggcgggaccggtggaggtggcagcagaggttcaactg
 ctgagctctggaggaggcctagtcagcctggaggagcctgctctcttctgtgcagtaagcggcatcgac
 ctgagcaattacgccatcaactgggtgagacaagctccgggggaaggtttagaatggatcggtataatattgg
 gccagtgggagcagcttttatgctacatgggcgaaaggaaggtttacaattagccgggacaatagcaaaac
 accgtgtatctccaaatgaactccttgcgagcagaggacagggcgggtgactattgtgctcgactgtccca
 ggttatagcactgcacctacttcgatctgtggggacaagggaccctggtgactgtttcaagt

1519gH20 FabFv catena pesante con sequenza segnale sottolineata in corsivo SE1 ID NO 74

*MEWSWVFLFLLSVTTGVHSE*EVPLVESGGGLVQPGGSLRLSCAVSGFTFSNYGMVWRQAPGKGLEWVAYIDS
 DGDNTYYRDSVKGRFTISRDNKSSLYLQMNLSRAEDTAVYYCTTGIVRPFYWGQGLTVTVSSASTKGPSV
 FPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSKVHTFPAVLQSSGLYSLSSVTVPSLGLTQ
 TYICNVNHKPSNTKVDKKEPKSCSGGGGSGGGTGGGGSEVQLLESGLLVQPGGSLRLSCAVSGIDLSNYA
 INWVRQAPGKCLEWIGI IWASGTTFYATWAKGRFTISRDNKNTVYLQMNLSRAEDTAVYYCARTVPGYSTA
 PYFDLWGQGLTVTVSS