



UNIVERSITÀ DEGLI STUDI DI PARMA
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
CORSO DI LAUREA IN BIOTECNOLOGIE INDUSTRIALI

**ESPRESSIONE TERMOINDUCIBILE IN
ESCHERICHIA COLI DI INTERFERONE BETA IN
FORMA DI PROTEINA DI FUSIONE**

RELATORE

Prof. Simone Ottonello

TUTOR AZIENDALE

Dott. Carlo Maullu

STUDENTE

Federico Deidda

Anno Accademico 2004-2005

INDICE

<u>INTRODUZIONE</u>	pag. 5
<i>PRINCIPALI SISTEMI DI ESPRESSIONE</i>	pag. 6
<i>SISTEMI DI ESPRESSIONE UTILIZZATI NELLA PRODUZIONE DI PROTEINE TERAPEUTICHE</i>	pag. 6
<u>STRATEGIE DI ESPRESSIONE IN <i>Escherichia coli</i></u>	pag. 9
<i>COMPONENTI ESSENZIALI DI UN VETTORE DI ESPRESSIONE E STRATEGIE PER MIGLIORARE L'ESPRESSIONE</i>	pag. 9
Dosaggio genico.....	pag. 9
Marcatori di resistenza.....	pag. 9
Promotori.....	pag. 10
Terminatori della trascrizione.....	pag. 12
Stabilità dell'RNA messaggero.....	pag. 12
Inizio della traduzione.....	pag. 12
Tecnologia delle proteine di fusione.....	pag. 13
Sostituzione dei codoni rari.....	pag. 13
<i>SCELTA DEL COMPARTIMENTO CELLULARE IN CUI INDIRIZZARE LA PRODUZIONE</i>	pag. 13
Espressione citoplasmica.....	pag. 14
Espressione periplasmica.....	pag. 14
Secrezione extracellulare.....	pag. 14
<i>CARATTERISTICHE DEI CEPPI D'ESPRESSIONE</i>	pag. 14
<u>INTERFERONE BETA</u>	pag. 16
<i>INTERFERONE BETA</i>	pag. 16
<i>UNA IMPORTANTE APPLICAZIONE DELL'INTERFERONE BETA: GESTIONE DELLA SCLEROSI MULTIPLA (MS)</i>	pag. 16
Descrizione della malattia.....	pag. 16
Gestione della malattia grazie all'Interferone beta.....	pag. 17
<u>SCOPO DEL LAVORO</u>	pag. 20

<u>MATERIALI E METODI</u>	pag. 21
<i>CEPPI BATTERICI</i>	pag. 21
<i>TERRENI DI CRESCITA</i>	pag. 21
<i>PLASMIDI</i>	pag. 21
Plasmidi di partenza.....	pag. 21
Plasmidi costruiti in questa tesi.....	pag. 24
<i>OLIGONUCLEOTIDI</i>	pag. 25
<i>TRASFORMAZIONE DI Escherichia coli</i>	pag. 25
Preparazione cellule elettrocompetenti	pag. 25
Ligazioni.....	pag. 25
Trasformazione per elettroporazione di <i>Escherichia coli</i>	pag. 25
<i>ESTRAZIONE E PURIFICAZIONE DEL DNA</i>	pag. 26
Estrazione del DNA plasmidico (mini prep)	pag. 26
Estrazione da gel di frammenti di DNA.....	pag. 26
<i>ANALISI DEI TRASFORMANTI DI Escherichia coli</i>	pag. 26
Analisi dei trasformanti mediante screening per PCR	pag. 26
Analisi del DNA plasmidico per digestione enzimatica e PCR.....	pag. 27
<i>PROVE DI ESPRESSIONE IN Escherichia coli JM109</i>	pag. 27
Curve di crescita e raccolta dei campioni.....	pag. 27
Preparazione dei glicerinati.....	pag. 27
Curve di crescita a 28°C e a 42°C	pag. 27
Preparazione dei campioni per l' SDS-PAGE (Polyacrylamide Gel Electrophoresis)	pag. 27
SDS-PAGE	pag. 28
Soluzioni, reagenti e strumentazione.....	pag. 28
Preparazione gel	pag. 28
<u>RISULTATI</u>	pag. 30
<i>COSTRUZIONE DEI VETTORI PER L'ESPRESSIONE TERMOINDUCIBILE DI FP₃₄-CS-β IFN, FP₅₆-CS-β IFN e FP₇₅-CS-β IFN</i>	pag. 30
Clonaggio delle cassette d'espressione nel plasmide PL381.....	pag. 30
Controllo dei plasmidi PL381/FP ₃₄ -CS-β IFN, PL381/FP ₅₆ -CS-β IFN e PL381/FP ₇₅ -CS-β IFN mediante digestione e PCR.....	pag. 32

<i>COSTRUZIONE DEL VETTORE PER L'ESPRESSIONE TERMOINDUCIBILE DI β IFN</i>	pag. 34
<i>PROVE DI ESPRESSIONE IN Escherichia coli JM109</i>	pag. 35
Screening d'espressione su vari cloni derivanti da trasformazione con uno stesso plasmide.....	pag. 35
Curve di crescita e analisi d'espressione: JM109/PL382, JM109/PL385 (controlli JM109/PL381, JM109/PL378).....	pag. 36
<u>DISCUSSIONE</u>	pag. 40
<u>BIBLIOGRAFIA</u>	pag. 44

INTRODUZIONE (1)

Cinquanta anni dopo la scoperta della struttura a doppia elica del DNA, i nuovi sviluppi nel campo della scienza hanno rivoluzionato l'industria farmaceutica. Un gran numero di farmaci è stato prodotto grazie all'evoluzione delle conoscenze in campo biotecnologico.

Storicamente, gli approcci adottati dall'industria farmaceutica per identificare nuovi agenti terapeutici si basavano sulla scoperta e l'analisi di composti estratti da fonti naturali; tale approccio ha condotto alla scoperta di agenti medicinali di alto valore, come ad esempio i salicilati. Attualmente questo approccio gode di un nuovo impulso grazie alle campagne di esplorazione in aree caratterizzate da elevata biodiversità e che rappresentano un bacino potenziale per la scoperta di nuove molecole.

La prima importante svolta nella scoperta di farmaci innovativi fu provocata da studi sistematici effettuati su molecole organiche sintetizzate o modificate; per almeno un secolo tali molecole hanno determinato il destino dell'industria farmaceutica. Un'eccezione a ciò fu la scoperta che l'insulina ha effetti terapeutici, la quale dimostrò che le proteine possono essere usate come farmaci.

Del gran numero di medicine oggi usate, solo poche erano conosciute prima della seconda guerra mondiale; a quel tempo fu soprattutto l'industria farmaceutica tedesca a dare un notevole contributo alla scoperta di nuovi tipi di farmaci.

Dopo la guerra, grazie ad una più globale distribuzione delle industrie farmaceutiche, il primato tedesco venne scalzato dalle compagnie Anglo-Americane, che ancora oggi sono le regine nel settore.

La crescente necessità di insulina e l'aumento nella domanda di antibiotici, posero le basi per lo sviluppo di una nuova branca dell'industria farmaceutica che si basa sull'impiego delle biotecnologie piuttosto che sulla sintesi chimica.

Un campo nel quale i nuovi sviluppi hanno dato notevole impulso è quello della produzione di proteine terapeutiche. La difficoltà e il costo dell'estrazione delle proteine dalle loro fonti naturali erano, già nei primi anni '80, considerati un grosso problema: la resa di prodotto finito ottenuto mediante tale metodo di produzione era infatti insufficiente ad incontrare la domanda. Il problema è però stato superato elegantemente, come dimostra lo sviluppo del primo biofarmaco: nel 1982 venne immessa sul mercato una forma di insulina umana prodotta in *Escherichia coli* grazie alla tecnologia del DNA ricombinante, che ne consentiva la produzione in larga scala.

Oltre al limite della disponibilità, l'industria farmaceutica dovette, e deve ancora oggi, superare lo scoglio della sicurezza dei prodotti; infatti l'uso di proteine terapeutiche da fonte naturale può comportare dei rischi alla salute, come è dimostrato da patologie sviluppatesi in persone sottoposte a trattamento con prodotti di origine ematica (quali ormoni) contaminati da agenti patogeni. Un successo in tal senso provenne dalla produzione in *Escherichia coli* dell'ormone della crescita umano: il farmaco così ottenuto garantì, infatti, una sicurezza sanitaria che la molecola di derivazione naturale non era in grado di garantire, come dimostrano i casi di Creutzfeld-Jacob verificatisi in seguito alla sua somministrazione.

Un altro notevole successo fu la produzione del Fattore VIII, ossia una proteina di notevoli dimensioni e con un alto grado di glicosilazione che, in virtù di tali caratteristiche, richiese l'impiego di cellule di mammifero come piattaforma di espressione in grado di assicurare l'autenticità del prodotto.

Altre due innovazioni tecnologiche diedero nuovo impulso alle biotecnologie in campo farmaceutico: lo sviluppo nel 1975 del primo anticorpo monoclonale, che ha aperto la strada alla messa a punto di farmaci basati su anticorpi ricombinanti, e lo sviluppo di tecnologie per la produzione in larga scala di proteine ricombinanti, che ha reso possibile produrre citochine molto rare.

Dal 1982, anno di lancio dell'insulina, al 2000, sono stati approvati 84 biofarmaci totali per la somministrazione all'uomo. Un quarto di tutti i farmaci approvati nei tre anni successivi, cioè dal 2000 al 2003, è rappresentato da biofarmaci; essi sono costituiti da composti che si basano su

proteine, mentre nessuno si basa su acidi nucleici. Il dato esaltante è che rimarca l'importanza delle biotecnologie nel settore, è la constatazione che negli ultimi tre anni il numero di nuovi farmaci biotecnologici già eguaglia il numero di quelli sviluppati nei precedenti 18 anni.

È bene sottolineare che la maggior parte dei biofarmaci approvati di recente è rivolto ad indicazioni quali diabete, emofilia, patologie cardiovascolari e varie forme di tumore, ossia le maggiori cause di morte nei paesi industrialmente sviluppati; al contrario le malattie infettive, che provocano la morte di milioni di persone nel resto del mondo, non rappresentano un mercato altrettanto attraente per le case farmaceutiche.

È possibile prevedere che il futuro delle biotecnologie farmaceutiche sarà fortemente influenzato dalla determinazione della sequenza completa del genoma umano, la quale sta aprendo la strada ad un nuovo tipo di medicina che potremmo definire "personalizzata".

PRINCIPALI SISTEMI DI ESPRESSIONE

Gli aspetti più importanti nella produzione di farmaci e di altri composti destinati alla somministrazione umana sono la sicurezza e l'autenticità del prodotto. Ne consegue che la produzione di farmaci che hanno avuto approvazione è ristretta a sistemi di espressione ben conosciuti e validati, ossia *Escherichia coli*, alcuni lieviti e cellule di mammifero.

Con lo sviluppo della genomica e il conseguente incremento dei geni target, sta aumentando la richiesta di adatti sistemi di espressione che soddisfino le esigenze produttive. Attualmente la varietà di sistemi di espressione spazia dai procarioti (gram-positivi e gram-negativi), diversi lieviti e funghi filamentosi, fino alle cellule vegetali e di mammifero.

SISTEMI DI ESPRESSIONE UTILIZZATI NELLA PRODUZIONE DI PROTEINE TERAPEUTICHE(1)

Il mercato delle proteine terapeutiche può essere grossolanamente diviso in due segmenti:

1. Le proteine che non richiedono modificazioni post-traduzionali
2. Le proteine che richiedono modificazioni post-traduzionali (principalmente glicosilazione) per poter adempiere alla propria funzione biologica.

Le proteine del primo gruppo sono prodotte principalmente in *Escherichia coli* e nel lievito *Saccharomyces cerevisiae* e rappresentano il 40% del mercato delle proteine terapeutiche. Per quanto riguarda il secondo gruppo, la produzione di proteine che richiedono glicosilazione per essere attive è affidata, nella maggior parte dei casi, all'impiego di cellule di mammifero in grado di mimare la glicosilazione umana. In molte glicoproteine l'importanza della glicosilazione è notevole in quanto le controparti non glicosilate non si strutturano correttamente, sono biologicamente inattive e vengono eliminate rapidamente dalla circolazione sanguigna.

Alcune proteine che sono naturalmente glicosilate ma per la cui attività la glicosilazione non è indispensabile, possono essere prodotte in ospiti che non sono in grado di glicosilare; un esempio in tal senso è rappresentato dall'interferone β -1 prodotto sia in *Escherichia coli* (*Betaferon*) che in cellule di mammifero (*Avonex*).

In generale, la mancanza di glicosilazione nelle proteine può comprometterne la vita media e altre proprietà clinicamente rilevanti.

Un aspetto riguardante il sistema procariotico, è relativo alla produzione di proteine ricombinanti in forma di corpi di inclusione; sebbene essi rappresentino talvolta un buon materiale di partenza per la purificazione, spesso però sono costituiti da proteine in forma insolubile e biologicamente inattiva. Anche se in questi casi è necessario ricorrere a sofisticati e costosi processi di rinaturazione del prodotto inattivo (il cui buon esito non è sempre prevedibile né garantito) i corpi di inclusione forniscono una valida opzione per produrre diverse proteine, che sono entrate con successo nel mercato.

Diversamente dai sistemi di espressione procariotici, le cellule di lievito e di insetto sono in grado di glicosilare le proteine ma non nelle modalità attuate dai mammiferi. Nel lievito, ad esempio, la

natura peculiare dei glicani, caratterizzati da un elevato contenuto di mannosio, ha un impatto negativo sulla vita media della proteina a causa dell'affinità dei suddetti glicani per specifici recettori localizzati sulla superficie dei macrofagi e delle cellule epiteliali. Il risultato è una rapida rimozione dal circolo sanguigno della glicoproteina somministrata. Per tali ragioni le cellule di mammifero e, in particolare, le *Chinese hamster ovary (CHO) cells*, stanno emergendo come il sistema di espressione più largamente impiegato per la produzione di glicoproteine umane che rappresentano, da sole, il 60% del mercato delle proteine terapeutiche. I limiti dell'impiego delle cellule *CHO* sono l'elevato costo della procedura di fermentazione e la bassa produzione, che è molto inferiore a quella dei sistemi microbici.

Le cellule *CHO* tipicamente sono in grado di esprimere glicoproteine che hanno un pattern di glicosilazione quasi umano, anche se è riconosciuto che le strutture glicaniche introdotte dalle linee cellulari *CHO* differiscono da quelle riscontrate nelle glicoproteine umane e ciò può avere effetti sull'efficacia terapeutica. Va inoltre sottolineato che l'intrinseca variabilità nel processo di glicosilazione cellulare porta alla produzione di miscele eterogenee di glicoforme, la qual cosa impone un grosso limite alla standardizzazione della produzione: la constatazione del fatto che alcune glicoforme sono più attive di altre mette in luce l'importanza del controllo della glicosilazione, al fine di modulare e migliorare le funzioni terapeutiche delle glicoproteine.

Diversi approcci sono stati adottati per superare questo limite; tra questi, la modulazione del processo di glicosilazione nelle cellule di mammifero o l'uso di metodi enzimatici in vitro che permettano di modificare la glicosilazione, dopo che è stata purificata la proteina.

Un recentissimo filone di ricerca mira all'impiego di sistemi evolutivamente inferiori alle *CHO*, quali lieviti e funghi, modificati geneticamente allo scopo di effettuare una glicosilazione di tipo umano.

I vantaggi offerti dai sistemi fungini in generale sono notevoli: alte produzioni della proteina target, processi di produzione scalabili (fino a 100 m³) e disponibilità di un gran numero di strumenti genetici per modificare l'ospite. Più di 120 proteine ricombinanti sono state espresse in lievito, molte delle quali hanno un'origine umana o di mammifero. Sebbene tali sistemi offrano notevoli vantaggi nella produzione di proteine ricombinanti, il loro pattern di glicosilazione non-umano ne ha in passato escluso l'impiego come sistemi di espressione per la produzione di glicoproteine da impiegarsi in terapia sull'uomo. Con l'avvento dei ceppi di lievito ingegnerizzati per compiere una glicosilazione di tipo umano, sono emersi nettamente i vantaggi offerti rispetto alle cellule *CHO*: oltre alla riduzione del costo di produzione, il vantaggio più importante è senza dubbio la qualità del biofarmaco. Infatti, si dimostra che tali ceppi sono in grado di produrre una glicoproteina con un singolo tipo di N-glicosilazione anziché la miscela di glicoforme ottenuta con le cellule *CHO*.

La seguente tabella permette di riassumere le principali caratteristiche che contraddistinguono le piattaforme di espressione citate (2):

CARATTERISTICHE	<i>E. coli</i>	LIEVITO	CELLULE DI INSETTO	CELLULE DI MAMMIFERO
CRESCITA CELLULARE	Rapida (t _D 30 min.)	Rapida (t _D 30 min.)	Lenta (t _D 18-24 h)	Lenta (t _D 24 h)
COMPLESSITA' DEL MEDIUM DI CRESCITA	Bassa	Bassa	Elevata	Elevata
COSTO DEL MEDIUM DI CRESCITA	Basso	Basso	Alto	Alto
LIVELLO DI ESPRESSIONE	Alto	Basso-Alto	Basso-Alto	Basso-Moderato
ESPRESSIONE EXTRACELLULARE	Secrezione nel periplasma	Secrezione nel medium	Secrezione nel medium	Secrezione nel medium
MODIFICAZIONI POST-TRADUZIONALI				
Folding proteico	Refolding generalmente richiesto	Refolding generalmente richiesto	Folding adeguato	Folding adeguato
N-glicosilazione	Nessuna	Alto contenuto in	Semplice, no	Complessa

		Mannosio	acido sialico	
O-glicosilazione	No	Si	Si	Si
Fosforilazione	No	Si	Si	Si
Acetilazione	No	Si	Si	Si
Acilazione	No	Si	Si	Si
γ -carbossilazione	No	No	No	Si

STRATEGIE DI ESPRESSIONE IN *Escherichia coli*

Tra i diversi sistemi esistenti per la produzione di proteine eterologhe, il batterio gram-negativo *Escherichia coli*, rimane tra i più attraenti. A suo favore giocano: la capacità di crescere velocemente raggiungendo alte densità su substrati poco costosi, l'ottima conoscenza delle sue caratteristiche genetiche e la disponibilità di una grande varietà di vettori di clonaggio e di ceppi mutanti. Malgrado tali vantaggi, l'efficiente espressione di svariate proteine in *Escherichia coli* può non esser praticabile o può esser resa difficoltosa da alcune limitazioni del sistema.

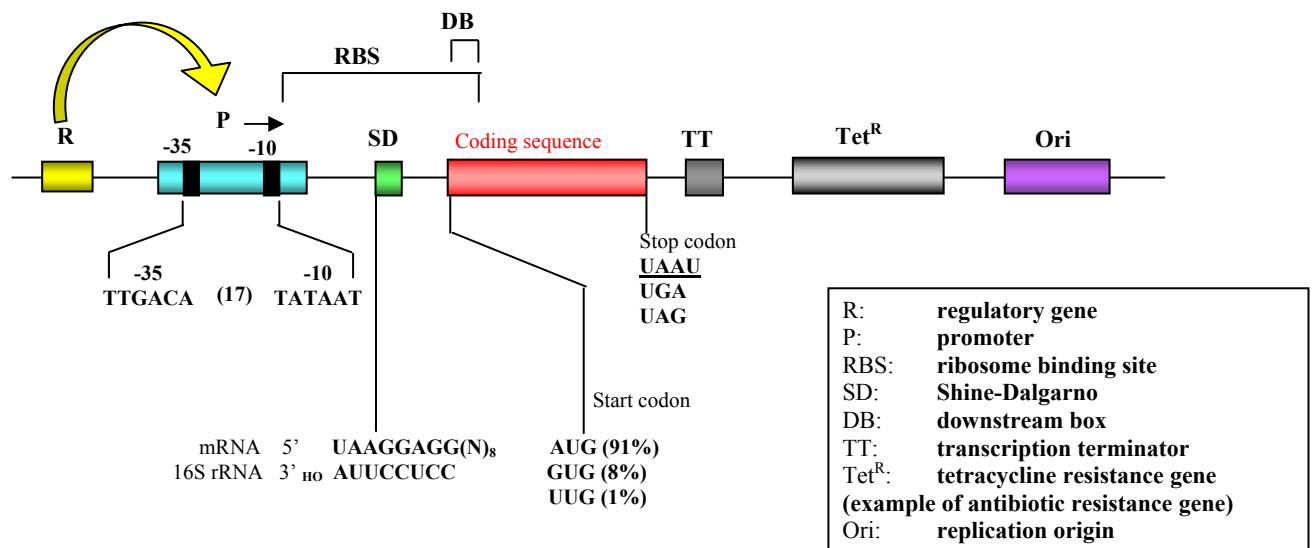
Partendo da quelle che sono le caratteristiche salienti del sistema di espressione *Escherichia coli*, sono state messe a punto diverse strategie per superare i limiti dello stesso.

COMPONENTI ESSENZIALI DI UN VETTORE DI ESPRESSIONE E STRATEGIE PER MIGLIORARE L'ESPRESSIONE

Un vettore d'espressione procariotico ben progettato, consiste di un set ottimale di elementi genetici che influenzano la produzione proteica sia a livello della trascrizione sia a livello della traduzione.

Altri elementi non secondari includono, la presenza di un gene che consenta la selezione fenotipica delle cellule che portano il vettore e la presenza di un'origine di replicazione (Ori) dalle cui caratteristiche dipende il numero di copie di plasmide per cellula.

La seguente figura riporta lo schema generico della struttura di un plasmide in forma lineare:



Dosaggio genico (3)

Più è rapido l'accumulo di prodotto intracellulare, più grande sarà la probabilità di aggregazione dello stesso. Questa tipica situazione è desiderabile nella produzione dei corpi di inclusione; infatti, come la produzione di insulina dimostra, essa può dar vita ad un processo produttivo economicamente valido.

Se l'obiettivo è la produzione di una proteina attiva e solubile, è preferita una forma di produzione più moderata e quindi un alto numero di copie di plasmidi per cellula può non esser desiderabile.

A tal fine l'impiego di vettori di espressione che si basano sui plasmidi F (a basso numero di copie) o la tecnica di inserzione cromosomale delle cassette d'espressione, permettono di avere un accumulo del prodotto che, pur rimanendo soddisfacente, consenta di evitare la sua aggregazione.

Marcatori di resistenza (4)

I marcatori di resistenza giocano un ruolo molto importante nel mantenimento e nella stabilità plasmidica, in particolare nel caso di plasmidi con un alto numero di copie.

Tale importanza è messa in luce dall'osservazione che in condizioni di laboratorio i plasmidi multicopia sono casualmente distribuiti durante la divisione cellulare e, in assenza di una pressione selettiva, sono persi a bassa frequenza (10^{-5} - 10^{-6}).

Inoltre, la perdita plasmidica può aumentare di molto nel caso di plasmidi con un numero molto alto di copie, in particolare nel caso in cui i plasmidi portino geni codificanti per proteine che siano tossiche per l'ospite o ne riducano la velocità di crescita.

Lo strumento più semplice per limitare tale inconveniente è l'impiego di plasmidi che portano un marcatore di resistenza ad un antibiotico e la somministrazione di quest'ultimo nel terreno di crescita al fine di uccidere le cellule che non portano il plasmide.

I più comuni marcatori per la resistenza agli antibiotici conferiscono resistenza al Cloramfenicolo, Tetraciclina, Ampicillina e Kanamicina.

Gli inconvenienti dell'impiego di antibiotici sono la perdita di pressione selettiva come risultato della degradazione dell'antibiotico durante la crescita, l'inattivazione degli enzimi periplasmici detossificanti (ad esempio le beta-lattamasi) e, nell'ottica della produzione di una proteina terapeutica, la possibile contaminazione del prodotto da parte dell'antibiotico. Poiché quest'ultimo aspetto è senza dubbio il meno accettabile, sono state sviluppate diverse strategie che evitano l'uso degli antibiotici.

Molte di queste strategie si basano su vettori di clonaggio che sono ingegnerizzati per portare geni o repressori che causano la morte cellulare in caso di perdita plasmidica. Una di esse consiste nell'impiego, come ospite d'espressione, di un ceppo che possiede, a livello cromosomale, un gene essenziale per la propria crescita posto sotto il controllo della regione promotore/operatore dell'operone Lac e, come vettore di espressione, di un plasmide multicopia che porta l'operatore Lac come marcatore di selezione. Le poche unità di repressore Lac, sintetizzate dalla cellula, sono sequestrate dalla sequenza operatore presente nei plasmidi, garantendo la sopravvivenza della cellula; al contrario, la perdita del plasmide ne compromette la sopravvivenza.

Promotori

I meccanismi di regolazione della trascrizione che si ritrovano in natura sono molteplici; la maggior parte di essi si basa sul legame di una proteina regolatrice alla sequenza promotore, ma ne esistono anche altri, tra cui: l'attenuazione trascrizionale, l'RNA antisense, i cambiamenti nel fattore sigma e i fattori anti-sigma.

Oltre alla natura dei meccanismi di regolazione, sono importanti anche le condizioni che conducono a cambiamenti nell'attività del promotore. Tali condizioni sono, ad esempio, la disponibilità della fonte di carbonio, la temperatura di crescita, il pH. Nella pratica, la regolazione dell'attività dei promotori avviene soprattutto grazie a proteine regolatrici in risposta a cambiamenti nella fonte di carbonio e nella temperatura di crescita. Tali promotori sono genericamente definiti "inducibili".

I promotori inducibili rientrano in due categorie: quelli a controllo negativo e quelli a controllo positivo; nei primi, una proteina repressore si lega al promotore o a valle di esso e inibisce la trascrizione, mentre nei secondi, una proteina attivatrice è necessaria per consentire il legame dell'RNA polimerasi.

Lo studio dei promotori inducibili ne ha permesso l'impiego per l'espressione di proteine eterologhe, in particolare di proteine che risultano tossiche per il ceppo ospite.

Un esempio dell'applicazione di un promotore inducibile a regolazione negativa riguarda l'operone Lac di *Escherichia coli*. Per rendere tale sistema più efficiente nel regolare l'espressione di proteine eterologhe è stato necessario superare diversi limiti dello stesso. Infatti, inserendo un gene target posto sotto il controllo del promotore-operatore Lac in un plasmide multicopia, la produzione cromosomale del repressore Lac non è quantitativamente in grado di garantire la repressione (si ha espressione costitutiva del gene target). Una prima strategia per migliorare il sistema, è consistita nell'isolare una variante mutante del gene che codifica per il repressore (LacI^q), che garantisce un incremento nella produzione del repressore di dieci volte. Nonostante i potenziali vantaggi della scoperta, la localizzazione cromosomale di LacI^q non garantiva ugualmente la completa repressione

del sistema; al contrario, inserendo LacI (o anche LacI^q) all'interno del plasmide multicopia insieme alla cassetta d'espressione, l'equilibrio tra i due elementi genetici garantisce la repressione del sistema. L'altro limite era di natura applicativa: la completa induzione del sistema era garantita dall'uso dell'isopropil-β-D-tiogalattoside (IPTG), analogo non idrolizzabile del lattosio, il cui impiego non è consigliato per la produzione di proteine terapeutiche a causa del suo costo e della sua tossicità. Al fine di evitare l'uso di tale induttore, si è sviluppato un repressore Lac termosensibile che è inducibile mediante uno shift di temperatura da 30°C a 42°C (4).

Un esempio dell'applicazione di un promotore inducibile a regolazione positiva riguarda il sistema L-arabinosio di *Escherichia coli*. In esso, il regolatore positivo araC modula l'espressione dell'operone in base alla presenza dell'L-arabinosio. Vettori di espressione che portano il gene araC e il promotore pBAD (al quale si lega la polimerasi), permettono di ottenere produzioni proteiche rilevanti e di avere un livello di espressione basale molto basso.

Dagli esempi indicati, emergono quelle che sono le caratteristiche principali che un promotore deve avere perché possa essere impiegato per l'espressione genica ad alti livelli (5):

1. Deve essere forte, al fine di garantire una produzione proteica di almeno il 10-30% delle proteine cellulari totali;
2. Deve essere un promotore regolato, se l'intento è quello di produrre proteine nocive per la crescita del ceppo ospite; i promotori regolati, a differenza dei costitutivi, permettono di regolare l'espressione genica garantendo un livello minimo di trascrizione basale in assenza dell'induttore, consentendo di superare il problema della tossicità del prodotto;
3. L'induzione di promotori regolati deve essere semplice e vantaggiosa dal punto di vista economico: attualmente, per la produzione in larga scala, sono utilizzate come tecniche di induzione quella termica e quella chimica.

Attualmente, per l'espressione eterologa sono disponibili numerosi sistemi che combinano l'azione di forti promotori e di proteine regolatrici molto efficaci. Alcuni esempi sono indicati nella seguente tabella (1):

Elementi di controllo usati per l'espressione di proteine ricombinanti in <i>E. coli</i>				
Promotore	Regolazione	Induttore	Livello di espressione	Commenti
T7 (Batteriofago)	lacI ^q	IPTG	Molto alto	Utilizza la RNA polimerasi T7. Induzione costosa
trc (ibrido) <i>E.coli</i>	lacI, lacI ^q	IPTG	Moderatamente alto	Ha un livello di espressione più basso rispetto al sistema T7, ma si può avere una regolazione dell'espressione più stretta. Induzione costosa
P _L (fago lambda)	cIts857	Shift di T° a 42°C	Moderatamente alto	Richiede l'impiego di un ceppo termosensibile. Minore probabilità di espressione non indotta.
araBAD	araC	L-Arabinosio	Variabile	Livelli di espressione dipendenti dalla dose dell'induttore. Induttore non costoso.

Terminatori della trascrizione (6)

L'impiego dei terminatori della trascrizione è una pratica raccomandata poiché inibisce l'azione impropria dell'RNA polimerasi in altri geni nel vettore o addirittura nell'origine di replicazione; infatti, la localizzazione di un terminatore a valle del gene di interesse aumenta la stabilità plasmidica.

Dei due meccanismi di terminazione della trascrizione esistenti nei batteri, uno Rho-dipendente e Rho-indipendente, quello che generalmente viene impiegato nei sistemi di espressione è il secondo. In esso, la formazione di una struttura stem-loop al 3' del messaggero garantisce, oltre al blocco della trascrizione, anche una maggiore stabilità del trascritto presumibilmente poiché tale struttura protegge il trascritto dall'azione delle esonucleasi 3'→5'.

Stabilità dell'RNA messaggero

L'espressione genica è strettamente dipendente dalla stabilità dell'mRNA.

Gli mRNA di *Escherichia coli* mostrano una vita media che si aggira da meno di un minuto a 10-20 minuti.

I principali enzimi coinvolti nella degradazione dei trascritti sono due esonucleasi 3'→5' (RNasi II e la polinucleotide fosforilasi) e l'endonucleasi RNasi E (7).

Studi sulla stabilità dei messaggeri di *Escherichia coli* hanno messo in luce che le stabili strutture secondarie che si formano alle estremità non tradotte (*UTR, untranslated regions*) del trascritto, possono incrementare la stabilità dello stesso; infatti si dimostra che l'inserzione di sequenze che codificano per strutture a forcina alle estremità dei trascritti, determina un incremento della vita media dei messaggeri (5).

Inizio della traduzione

Il processo di inizio della traduzione è influenzato da diversi fattori. Infatti, esiste una notevole variabilità nei valori di efficienza di tale processo dovuta, fra l'altro, alle strutture secondarie che si formano all'estremità 5' di ogni specie di messaggero. Non esiste, dunque, una sequenza universale in grado di garantire un'efficiente inizio di traduzione.

Le regioni di inizio della traduzione comprendono (8):

1. La sequenza di *Shine-Dalgarno*
2. Il codone di inizio
3. La regione compresa tra la sequenza *Shine-Dalgarno* e il codone di inizio
4. Enhancers traduzionali (downstream box)

A causa dello stretto accoppiamento, nei procarioti, tra trascrizione e traduzione, l'ingegnerizzazione delle regioni di inizio della traduzione è un ottimo strumento per modulare l'espressione genica.

Per migliorare l'espressione genica si può agire su più fronti:

1. La natura del codone di inizio della traduzione: l'uso preferenziale da parte di *Escherichia coli* del codone AUG (98%) coincide anche con una migliore efficienza di traduzione.
2. Stabili strutture secondarie, che coinvolgano la sequenza *Shine-Dalgarno* (cui corrisponde una sequenza anti-SD all'interno dell'RNA ribosomale 16S) e/o il codone di inizio, possono drammaticamente ridurre l'espressione genica interferendo con il legame dell'apparato traduzionale. Una mutagenesi volta a destabilizzare tali strutture può quindi avere un notevole effetto benefico sull'efficienza di traduzione.
3. La lunghezza della regione compresa tra *Shine-Dalgarno* e codone di inizio della traduzione, che varia mediamente da 5 a 13 nucleotidi, influenza l'efficienza della traduzione; estensivi studi a riguardo hanno stabilito che la spaziatura ottimale tra i due elementi è di cinque nucleotidi (9).

4. Un ruolo addizionale è svolto dalla *DB (downstream box)*, la quale è localizzata a valle del codone di inizio e mostra una più o meno spiccata complementarietà con una sequenza dell'rRNA 16S (anti-DB) diversa dalla anti-SD. Essa dunque coopera con la *Shine-Dalgarno* alla stabilizzazione del complesso ribosoma-messaggero. La mutagenizzazione di tale regione, al fine di aumentarne la complementarietà con la anti-DB senza alterare la sequenza codificante, può determinare un incremento dell'espressione. Proprio in virtù di tale caratteristica, tali sequenze sono state denominate "enhancers traduzionali" (10).

Tecnologia delle proteine di fusione (4)

Le proteine di fusione, o proteine chimeriche, generalmente includono un partner legato a una proteina target con interposizione di un sito di riconoscimento per una specifica proteasi.

I partner di fusione offrono diversi vantaggi, tra cui:

1. Alcuni partner di fusione aumentano la solubilità delle proteine target, che altrimenti si accumulerebbero nel citoplasma in forma di corpi di inclusione. La ragione di tale fenomeno è probabilmente insita nel fatto che il partner di fusione raggiunge rapidamente ed efficientemente la conformazione nativa nel momento in cui esso emerge dal ribosoma e ciò promuove l'acquisizione di una corretta struttura da parte della proteina ad esso fusa.
2. Vantaggi a livello di purificazione: l'uso di tag-poliistidinici, ad esempio, come partner di fusione, consente l'efficiente purificazione grazie alla cromatografia per affinità metallica.
3. Alti livelli di espressione possono essere ottenuti localizzando un opportuno partner di fusione, che deriva da un gene ben espresso in *Escherichia coli*, all'estremità amminoterminale di una proteina target che, in assenza del partner, risulterebbe debolmente espressa; tale risultato è probabilmente dovuto alla stabilizzazione dell'rRNA chimerico risultante e all'assenza, nel partner di fusione, di codoni rari che ostacolerebbero l'inizio della traduzione (7).

Possibili svantaggi derivanti dall'applicazione della tecnologia delle proteine di fusione riguardano l'elevato costo delle proteasi usate per la rimozione del partner, l'efficienza con cui avviene il taglio proteolitico non sempre soddisfacente, i passaggi addizionali richiesti per l'ottenimento del prodotto e, infine, il fatto che la solubilità del prodotto proteico non è, comunque, mai garantita (4).

Sostituzione dei codoni rari

Le differenze nell'uso dei codoni tra procarioti ed eucarioti può avere un impatto significativo sulla produzione di proteine eterologhe. I codoni per l'Arginina AGA e AGG si trovano raramente in *Escherichia coli*, mentre sono comuni in *Saccharomyces cerevisiae* e negli eucarioti in generale. La presenza di tali codoni in geni clonati influenza i livelli di accumulo proteico, la stabilità plasmidica e dell'mRNA e nei casi più gravi, inibisce la sintesi proteica e la crescita.

Per minimizzare gli effetti negativi dovuti al differente uso dei codoni riscontrato in *Escherichia coli* rispetto agli eucarioti, si è ricorsi ad una strategia che consiste nell'alterare geneticamente i codoni rari del gene target, senza modificare il prodotto proteico codificato; le modifiche apportate sono eseguite in modo da rispecchiare l'adeguato uso preferenziale dei codoni da parte di *Escherichia coli* (4).

Diversi studi, hanno posto l'accento sul fatto che l'effetto negativo dovuto alla presenza di codoni rari è maggiore quando questi sono localizzati nell'estremità 5' del trascritto (5).

SCelta DEL COMPARTIMENTO CELLULARE IN CUI INDIRIZZARE LA PRODUZIONE (5)

La decisione di indirizzare la proteina verso uno specifico compartimento cellulare, ossia citoplasma o periplasma, dipende dal bilancio vantaggi/svantaggi che è offerto da ogni compartimento cellulare.

Espressione citoplasmica

La formazione di intermedi di folding in forma aggregata ed insolubile, denominati corpi di inclusione, è la conseguenza frequente della produzione proteica ad alti livelli nel citoplasma.

La produzione di proteine ricombinanti come corpi di inclusione ha diversi vantaggi, sebbene il recupero della proteina biologicamente attiva da corpi di inclusione possa essere costoso per la produzione in larga scala.

Tra le strategie adottate per favorire il raggiungimento della conformazione nativa troviamo:

- (a) La coespressione di *chaperons proteins*, che favoriscono il folding e dunque la solubilità;
- (b) Uso di peptidi altamente solubili come partner di fusione.

Espressione periplasmica

Le proteine presenti nel citoplasma sono in forma ridotta e non contengono legami disolfuro. Il periplasma è invece l'ambiente ideale per il folding delle proteine contenenti ponti disolfuro, quali l'ormone della crescita e i frammenti anticorpali, perché esso possiede gli enzimi necessari a formare ed isomerizzare tali legami. Inoltre, poiché il pool di proteine batteriche nel periplasma è ben inferiore a quello del citoplasma, la purificazione del prodotto è più semplice.

Un aspetto fondamentale è la traslocazione attraverso la membrana interna: se da un lato essa può non essere efficiente e rappresentare uno step limitante nel complesso del processo produttivo, dall'altro comporta il taglio *in vivo* del peptide segnale, con conseguente liberazione della proteina target avente l'estremità ammino-terminale autentica.

Anche se l'ambiente periplasmico non si è evoluto per supportare il folding di grandi proteine complesse con tanti legami disolfuro, esso può essere modificato over-esprimendo opportune *chaperon proteins*.

Secrezione nel medium extracellulare

La secrezione del prodotto proteico nel terreno di crescita è la strategia più desiderabile ai fini dell'ottenimento del prodotto proteico.

I vantaggi della secrezione sono molteplici:

1. Basso livello di attività proteolitica nel medium
2. Purificazione più semplice a causa del ridotto numero di specie proteiche presenti
3. Autenticità dell'estremità ammino-terminale
4. Miglioramento del folding

Nonostante i vantaggi elencati, generalmente non si ha secrezione da parte di *Escherichia coli*. Inoltre, è bene sottolineare che, anche se la purificazione è più vantaggiosa in termini di ridotto numero di specie proteiche presenti nel terreno di crescita, essa risulta più complessa a causa della diluizione del prodotto stesso.

CARATTERISTICHE DEI CEPPI D'ESPRESSIONE (7)

Cepi dotati di caratteristiche particolari possono rivelarsi maggiormente apprezzabili ai fini della produzione di proteine terapeutiche. Alcuni esempi:

- L'utilizzo di ceppi deficienti nella produzione di *Tioredoxina* riduttasi fa sì che l'ambiente citoplasmatico sia meno riducente e favorisca così la formazione dei legami disolfuro, fondamentali per il folding di alcune proteine;
- Ceppi che hanno una velocità di sintesi proteica ridotta favoriscono il processo di folding;
- L'impiego di ceppi che crescono bene a temperature inferiori ai canonici 37°C favorisce il processo di folding;

- Ceppi deficienti in proteasi localizzate nella membrana citoplasmatica limitano la proteolisi delle proteine indirizzate nel periplasma;
- L'impiego di ceppi ingegnerizzati per codificare *tRNA* in grado di riconoscere codoni normalmente rari in *Escherichia coli*, può risultare vantaggioso per l'espressione di proteine eucariotiche.

INTERFERONE BETA

INTERFERONE BETA (IFN β) (11)

Gli interferoni sono una grande famiglia di proteine multifunzionali secrete, che appartengono ad una classe di sostanze denominate citochine.

Le funzioni principali attribuite agli interferoni sono:

- 1) Difesa antivirale
- 2) Regolazione della crescita cellulare
- 3) Attivazione del sistema immunitario

La famiglia degli interferoni è suddivisibile in due gruppi:

- Gli interferoni di tipo I (IFN α e IFN β)
- Gli interferoni di tipo II (IFN γ)

Nei primi dieci anni di ricerca, gli IFN di tipo I erano considerati fundamentalmente degli inibitori virali, mentre si riteneva che quelli di tipo II fossero coinvolti solamente nella risposta immune. Questa visione è cambiata considerevolmente negli ultimi anni, soprattutto per quanto riguarda gli Interferoni di tipo I; infatti, si è osservato che gli IFNs di tipo I hanno una notevole importanza nei processi infiammatori, nella immunoregolazione e nella risposta delle cellule T. Le nuove funzioni identificate hanno assunto notevole importanza nel trattamento di malattie come le epatiti virali croniche e la sclerosi multipla.

UNA IMPORTANTE APPLICAZIONE DELL'INTERFERONE BETA: GESTIONE DELLA SCLEROSI MULTIPLA (MS)

Descrizione della malattia

Un esempio di applicazione dell'Interferone beta in campo medico, riguarda il suo impiego nella terapia della sclerosi multipla.

Essa è una malattia infiammatoria che colpisce il Sistema Nervoso Centrale (cervello e midollo spinale). Poiché essa può colpire qualsiasi area del Sistema Nervoso Centrale, è caratterizzata dal punto di vista clinico, da una grande varietà di sintomi; essi possono riguardare differenti funzioni fisiologiche:

- Visione (offuscamento, visione doppia, perdita completa della vista)
- Coordinazione (perdita completa o parziale dell'equilibrio, dolori muscolari, tremori)
- La forza (debolezza, calo del tono muscolare, paralisi)
- La sensibilità (formicolii, perdita di sensibilità tattile)
- La sessualità (impotenza, disturbi dell'eccitazione sessuale)
- Controllo della vescica (disfunzioni vescicali)

Sebbene molti pazienti affetti da sclerosi multipla conducano una vita più o meno normale, essa può compromettere la mobilità e creare inabilità nei casi più gravi.

La Società della sclerosi multipla degli Stati Uniti ha classificato la malattia in sottotipi a seconda del pattern di progressione(12):

- 1) *Relapsing remitting*: è caratterizzata da attacchi non prevedibili seguiti da intervalli di tempo che vanno da mesi ad anni di relativa calma, con nessun segno di attività della malattia.
- 2) *Secondary progressive*: si manifesta con un progressivo declino neurologico tra una fase acuta e la successiva, senza alcun periodo di ripresa (remission). L'80% dei casi che inizialmente si manifestano come remitting-relapsing, evolve in secondary progressive.

- 3) *Primary progressive*: rappresenta il 10% dei casi; in essi non si ha mai un blocco dell'attività della malattia dopo la comparsa dei sintomi iniziali; si ha un costante declino senza l'insorgenza di manifestazioni acute.
- 4) *Progressive relapsing*: descrive quegli individui che, dall'inizio della MS, hanno un declino neurologico costante e soffrono anche di attacchi acuti.

Il decorso della malattia è difficile da predire, infatti un particolare sottotipo può, per ragioni poco chiare, evolvere, nel tempo, ad un altro sottotipo.

La sclerosi multipla colpisce le fibre nervose mielinizzate del cervello e del midollo spinale. Tali fibre sono circondate da una guaina di natura lipidica chiamata mielina, che è importante per l'adeguata conduzione degli impulsi nel sistema nervoso. Con la sclerosi multipla si ha la graduale distruzione della mielina a livello del cervello e del midollo spinale (demielinizzazione), e ciò causa differenti sintomi a seconda del segnale che viene interrotto.

Il processo di demielinizzazione è il risultato di una serie di eventi che comprende:

- a) L'attivazione del sistema immunitario
- b) Il danno alla barriera ematoencefalica
- c) Il passaggio nel Sistema Nervoso Centrale delle cellule T attivate
- d) L'aggressione alla guaina mielinica e la sua distruzione

Non è stato possibile identificare un unico agente quale causa determinante la sclerosi multipla. Nella sua genesi sembrano coinvolti più fattori, quali: una predisposizione genetica, fattori ambientali, l'età ed un alterato funzionamento del sistema immunitario. La sclerosi multipla, infatti, potrebbe essere la conseguenza di un'infezione virale (fattore esogeno) o di un fattore endogeno come l'alterazione del sistema immunitario che scatena una risposta nei confronti della mielina, che non è più riconosciuta come elemento "self".

Il processo si esplica attraverso l'attacco da parte del sistema immunitario al sistema nervoso ed è per questo che la sclerosi multipla è classificata come malattia autoimmune. In questo senso, giocano un ruolo importante cellule bianche del sangue denominate linfociti T.

In condizioni normali, la barriera ematoencefalica fornisce un'efficace separazione tra le cellule del sangue e il Sistema Nervoso Centrale, impedendo così alle cellule T di attaccare la mielina; le cellule T sono però in grado di produrre delle metallo-proteinasi della matrice (MMP) che, attivate dall'esposizione ad altri enzimi, indeboliscono la tenuta della barriera ematoencefalica. Un altro meccanismo che causa la destabilizzazione di tale barriera è probabilmente legato all'azione distruttiva delle cellule bianche del sangue nei confronti degli organismi invadenti; tale azione comporta la liberazione di superossidi e radicali liberi che sono estremamente attivi e possono distruggere molti tessuti con cui vengono a contatto.

La destabilizzazione della barriera ematoencefalica consente la diffusione delle cellule bianche del sangue nel Sistema Nervoso Centrale, determinando la distruzione della mielina e le conseguenze ad essa associate.

Gestione della malattia grazie all'Interferone beta (14).

La gestione della terapia della MS è complessa e richiede l'ausilio di più discipline. Poiché non esiste alcuno specifico trattamento per la MS, le terapie usate sono indirizzate a diversi aspetti della malattia; tra queste vi sono:

- Terapie che modificano il decorso della malattia
- Trattamenti impiegati nelle fasi di acutizzazione della malattia
- Farmaci usati per trattare le complicanze della malattia

Le terapie che modificano il decorso della malattia si basano fondamentalmente su farmaci che agiscono sul sistema immunitario sia mediante stimolazione sia mediante repressione; tra i farmaci che appartengono a questa categoria vi sono gli interferoni.

Nel caso della sclerosi multipla l'azione degli interferoni è fondamentale nel decorso della malattia. In essa l'IFN γ , prodotto dalle cellule T, è ritenuto il responsabile dell'attivazione dei macrofagi che

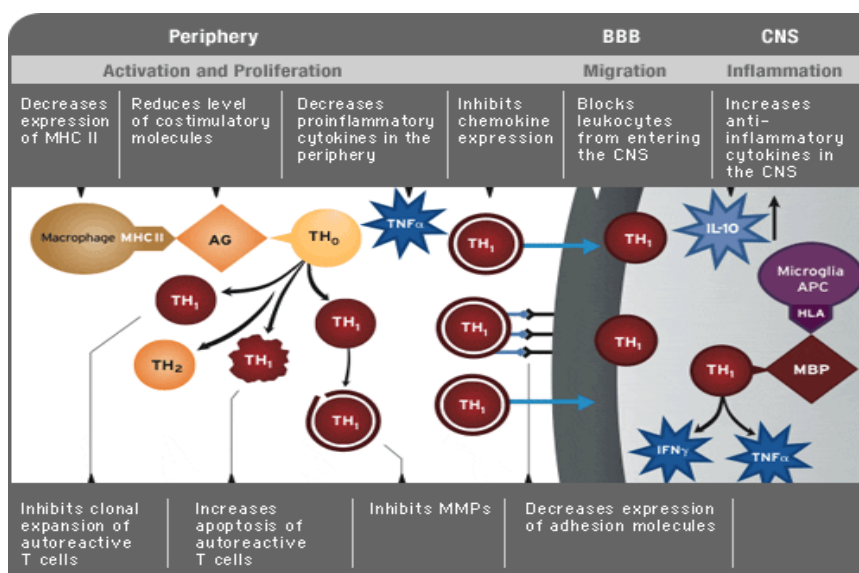
distrucono la mielina e sono in grado di acutizzare il decorso della malattia. Gli IFN α e IFN β sono prodotti invece da diversi tipi di cellule, tra cui i macrofagi, i fibroblasti, e le cellule endoteliali. Si sono rivelate di una certa importanza le attività immunomodulatorie esplicate dall'IFN β ai fini del trattamento della MS; esse consistono nella:

- Inibizione della proliferazione e della presentazione dell'antigene da parte delle cellule bianche del sangue
- Inibizione della migrazione delle cellule T attivate attraverso la barriera ematoencefalica
- Modulazione della produzione di citochine che inducono una condizione infiammatoria
- Aumento dell'attività delle cellule T suppressor deputate alla inibizione di altre cellule del sistema immune

Per quanto riguarda l'inibizione della migrazione delle cellule T attivate attraverso la barriera ematoencefalica, l'Interferone beta agisce a più livelli:

- Inibisce l'espressione delle molecole responsabili dell'adesione delle cellule T attivate sulla parete endoteliale
- Inibisce la produzione delle metallo-proteinasi di matrice da parte delle cellule T attivate, che sono responsabili della destabilizzazione della struttura della barriera ematoencefalica

Nella seguente figura sono indicati i livelli ai quali agisce l'Interferone beta:



La conoscenza di tali attività ha permesso lo sviluppo di tre formulazioni di IFN β , attualmente sul mercato, che sono attive contro la sclerosi multipla; esiste inoltre un farmaco sintetico che è il *Glatiramer acetato (GA)*.

Le caratteristiche dei farmaci indicati sono riportate nella tabella:

Farmaci immunomodulatori per il trattamento della Sclerosi Multipla			
Farmaco	Dosaggio	Effetti collaterali	Sistema di espressione
Interferone beta 1 a (Avonex)	Iniezione intramuscolare, 30 μ g una volta alla settimana	Sintomi simili all'influenza	<i>Chinese hamster ovary (CHO) cells</i>
Interferone beta 1 a (Rebif)	Iniezione sottocutanea, 22 μ g per tre volte alla settimana	Sintomi simili all'influenza e reazioni nel sito di iniezione	<i>Chinese hamster ovary (CHO) cells</i>

Interferone beta 1 b (Betaseron). Porta una mutazione <i>cys17→ser</i> .	Iniezione sottocutanea, 0,25 mg a giorni alterni	Sintomi simili all'influenza e reazioni nel sito di iniezione	<i>Escherichia coli</i>
Glatiramer acetato (GA) (Copaxone)	Iniezione subcutanea di 20 mg giornalmente	Reazioni nel sito di iniezione e, raramente, reazioni sistemiche benigne	

Il farmaco più interessante è l'IFN β 1b, che viene prodotto in un sistema procariotico, quale è appunto *Escherichia coli*. Tale sistema, nonostante non sia in grado di glicosilare la proteina, come avviene naturalmente, dà origine ad un prodotto proteico funzionale ai fini della terapia della sclerosi multipla.

In uno studio risalente al 1998 (15) è stato posto l'accento sulle differenze strutturali e funzionali della forma non glicosilata dell'IFN β 1b rispetto alla forma glicosilata dell'IFN β 1a. Si è evidenziato che il carboidrato della glicoproteina naturale gioca un ruolo diretto nello stabilizzare la struttura della proteina stessa. In particolare:

- Vi sono differenze nella temperatura di denaturazione fra le due molecole, come risultato della destabilizzazione della struttura in mancanza dello zucchero. Lo stesso effetto si manifestava asportando enzimaticamente lo zucchero dall'IFN β 1a.
- La presenza dello zucchero nella struttura minimizza l'esposizione al solvente di una superficie non carica che ha un gran numero di amminoacidi idrofobici; poiché tale esposizione è termodinamicamente sfavorita, l'aggregazione di IFN β non glicosilato in corpi di inclusione rappresenta un meccanismo di copertura della regione esposta al solvente. La condizione di aggregazione potrebbe indurre degli eventi di denaturazione, con effetti negativi sull'attività biologica.
- Lo studio conclude che, *in vitro*, la maggiore attività biologica dell'IFN β 1a rispetto all'IFN β 1b è dovuta fondamentalmente agli effetti stabilizzanti del carboidrato sulla struttura.

SCOPO DEL LAVORO

Alla luce di quanto precedentemente esposto a proposito:

- della validità di *Escherichia coli* come sistema d'espressione per proteine poco complesse e che non necessitano di glicosilazione per essere pienamente attive;
- delle potenzialità terapeutiche dell'Interferone beta nella gestione della Sclerosi Multipla;
- della non indispensabilità della glicosilazione di detta molecola ai fini della sua piena attività biologica *in vivo*

il lavoro di tesi si propone di testare un sistema d'espressione procariotico per la produzione di Interferone beta, in forma di proteina di fusione.

In particolare, il lavoro si propone di:

- 1) assemblare un vettore di espressione plasmidico con tutti gli elementi genetici necessari alla regolazione dell'espressione del gene per l'interferone beta; in dettaglio esso contiene:
 - a) gli elementi per l'espressione termoinducibile della proteina (sistema P_L/P_R del fago lambda regolato dal repressore termosensibile cI857), perché dati di letteratura e esperienze precedenti in Bioker hanno dimostrato che tale proteina è tossica per *Escherichia coli*;
 - b) al 5' del gene per l'Interferone beta, una sequenza codificante per un partner di fusione che ha dimostrato, in studi eseguiti precedentemente in Bioker, di incrementare notevolmente l'espressione di proteine altrimenti poco espresse in *Escherichia coli*.
 - c) tra il partner di fusione (FP) e il gene per l'Interferone beta, una sequenza codificante un sito di taglio riconosciuto specificamente da una proteasi che si è dimostrata, in esperienze precedenti, essere economicamente vantaggiosa e sufficientemente specifica.

Rappresentazione schematica della proteina di fusione



- 2) Testare l'effetto della lunghezza del partner di fusione sui livelli di espressione della proteina
- 3) Ottimizzare, almeno su piccola scala, terreni e modalità di induzione dell'espressione
- 4) Verificare se la proteina si presenta in forma solubile o in forma di corpi di inclusione

MATERIALI E METODI

CEPPI BATTERICI

Ceppi JM109: F' traD 36 proA⁺B⁺ lacI^q Δ(lac-proAB) glnV44 e14⁻ gyrA96 recA1 relA1 endA1 thi hsdR17

TERRENI DI CRESCITA

LS brodo: 1% peptone di farina di soia ottenuto per digestione papainica (MERCK), 0,5% di estratto di lievito (BD), 1% NaCl (FLUKA).

LS Agar: 1% peptone di farina di soia ottenuto per digestione papainica (MERCK), 0,5% di estratto di lievito (BD), 1% NaCl (FLUKA), 1,5% di Agar-Agar (MERCK).

LS brodo per cellule competenti: 1% peptone di farina di soia ottenuto per digestione papainica (MERCK), 0,5% di estratto di lievito (BD), 0,5 % NaCl (FLUKA).

FM₁ 0,5X : 18g/L Yeast Extract, 10g/L Soitone, 3,2g/L K₂HPO₄, 0,6g/L KHPO₄, 10g/L glicerolo. I terreni, a seconda dei casi, sono stati addizionati dell'opportuno antibiotico, ossia kanamicina ad una concentrazione finale di 30 µg/ml

PLASMIDI

Plasmidi di partenza

Alcuni dei plasmidi utilizzati in questa tesi derivano dalla ligazione fra prodotti di amplificazione e il plasmide commerciale pCRII-TOPO (*Invitrogen, No. 10351-021*).

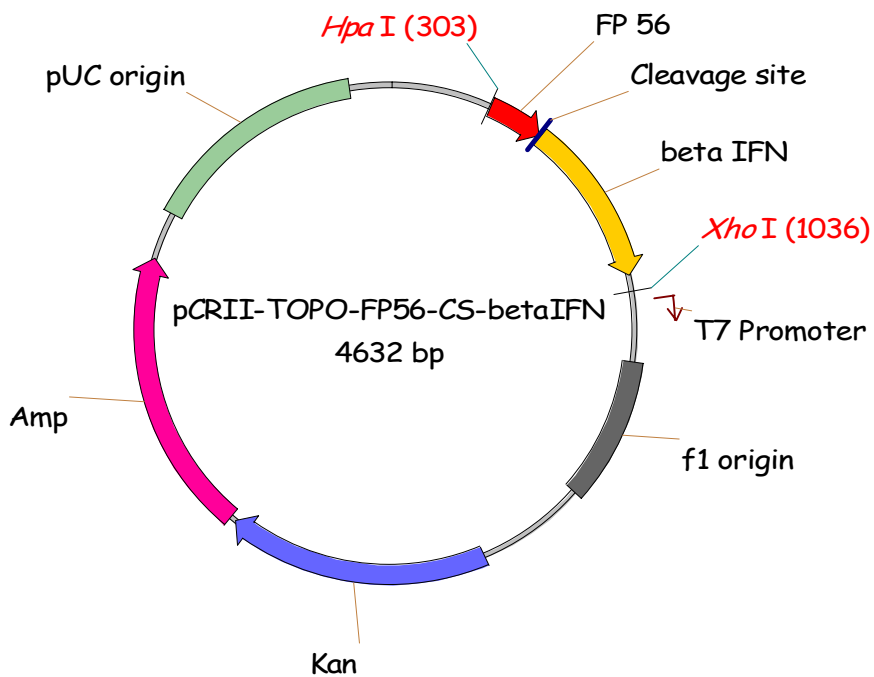
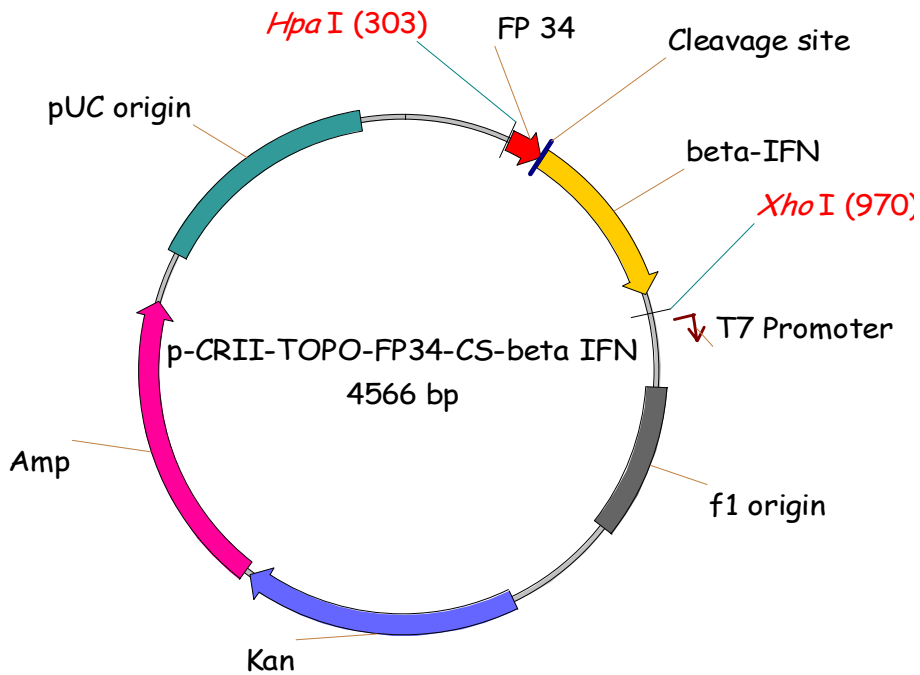
pCRII-TOPO è un vettore specifico per il clonaggio di amplificati per PCR aventi 3'A overhangs.

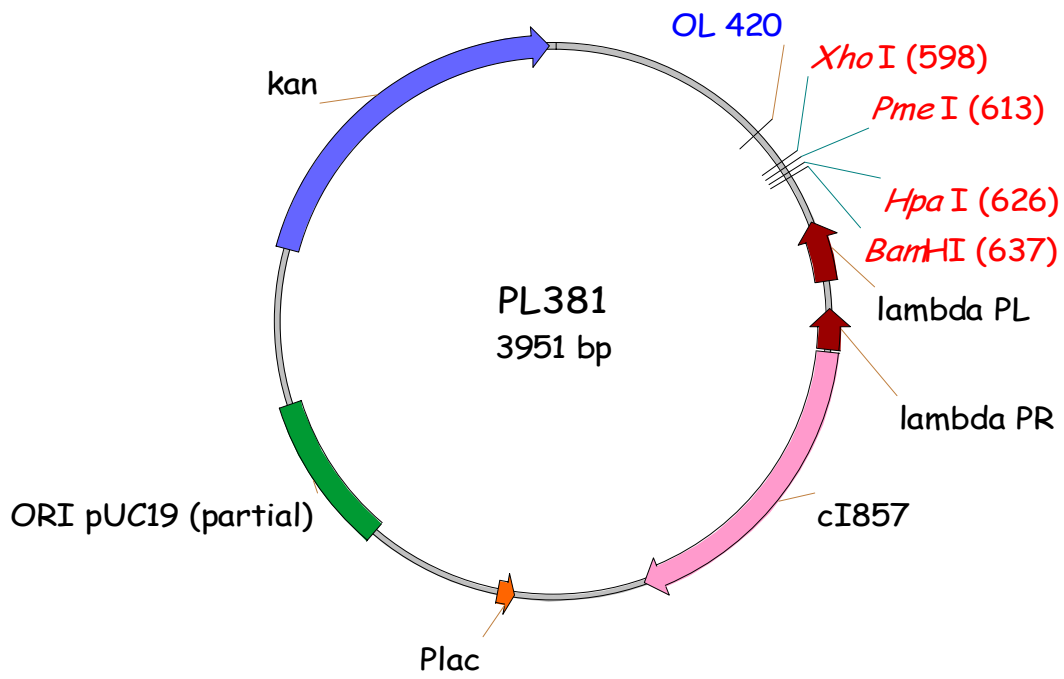
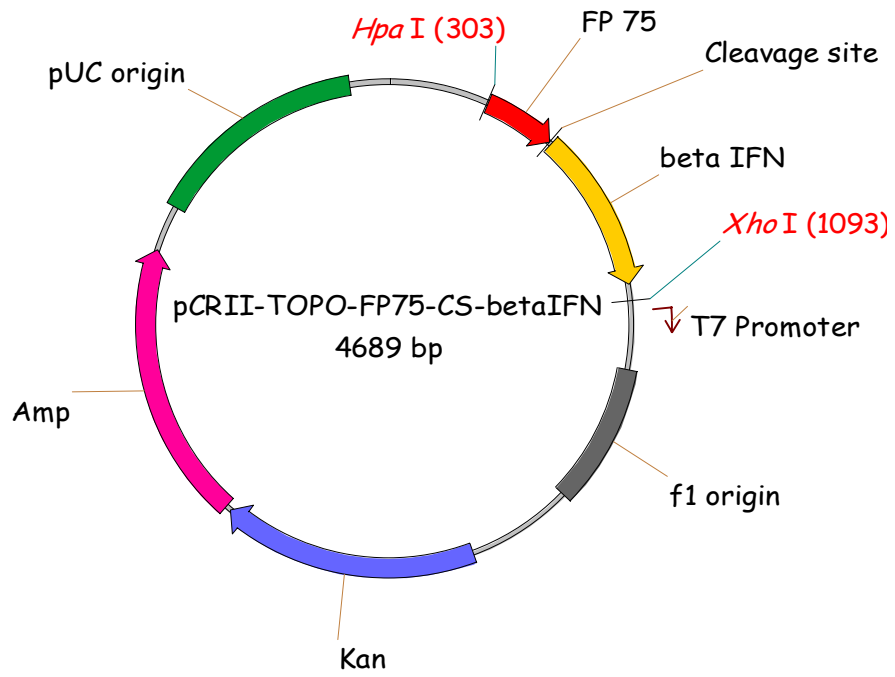
Abbreviazioni: **FP_n** **Fusion Partner** (n è il numero degli amminoacidi di cui è composto)
 CS **Cleavage Site** (sito di taglio per una specifica proteasi)

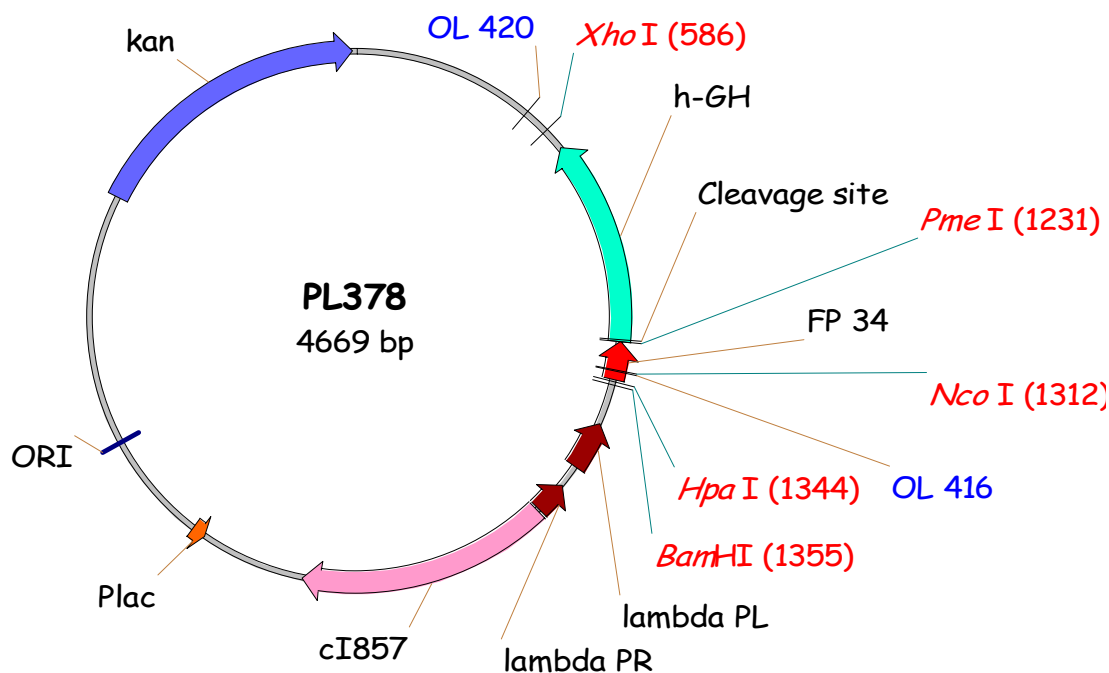
Nome	Dimensioni	Resistenza	Numero di copie	Proteina espressa	Caratteristiche
pCRII-TOPO/FP ₃₄ -CS-β-IFN	4566 bp	Ampicillina-Kanamicina	Alto		Plasmide intermedio, derivato dal pCRII-TOPO (Invitrogen).
pCRII-TOPO/FP ₅₆ -CS-β-IFN	4632 bp	Ampicillina-Kanamicina	Alto		Plasmide intermedio, derivato dal pCRII-TOPO (Invitrogen).
pCRII-TOPO/FP ₇₅ -CS-β-IFN	4689 bp	Ampicillina-Kanamicina	Alto		Plasmide intermedio, derivato dal pCRII-TOPO (Invitrogen).
PL381	3951 bp	Kanamicina	Alto		Plasmide per espressione inducibile. Gli elementi per l'espressione sono: -la coppia di promotori P _L /P _R del fago lambda seguita da un multicloning site. -il gene codificante per il repressore termosensibile cI857, sotto il controllo del proprio promotore naturale. L'induzione dell'espressione si ottiene mediante shift di temperatura da 28°C a 42°C.
PL378	4669 bp	Kanamicina	Alto	FP ₃₄ -CS-hGH	Plasmide per espressione inducibile. Il gene codificante per la fusione FP ₃₄ -CS-hGH è sotto il controllo dei promotori P _L /P _R del fago lambda. Il plasmide contiene anche il gene codificante il repressore termosensibile cI857, sotto il controllo del proprio promotore naturale. L'induzione dell'espressione si ottiene con shift di temperatura da 28° a 42°C.

Rappresentazione grafica dei plasmidi di partenza:

- 1) Gli elementi nominati in rosso sono i siti di riconoscimento per enzimi di restrizione;
- 2) Gli elementi nominati in blu sono i siti di annealing degli oligonucleotidi impiegati nello screening per PCR;
- 3) In nero sono riportati i geni e altri elementi genetici che caratterizzano i plasmidi.







Plasmidi costruiti in questa tesi

Nome	Dimensioni	Resistenza	Numero di copie	Proteina espressa	Caratteristiche
PL382	4594 bp	Kanamicina	Alto	FP ₃₄ -CS- β -IFN	Plasmide di espressione derivato da PL381. Il gene che esprime la fusione è sotto il controllo di P _L /P _R . L'induzione dell'espressione si ottiene mediante shift di temperatura da 28°C a 42°C.
PL383	4660 bp	Kanamicina	Alto	FP ₅₆ -CS- β -IFN	Plasmide di espressione derivato da PL381. Il gene che esprime la fusione è sotto il controllo di P _L /P _R . L'induzione dell'espressione si ottiene mediante shift di temperatura da 28°C a 42°C.
PL384	4717 bp	Kanamicina	Alto	FP ₇₅ -CS- β -IFN	Plasmide di espressione derivato da PL381. Il gene che esprime la fusione è sotto il controllo di P _L /P _R . L'induzione dell'espressione si ottiene mediante shift di temperatura da 28°C a 42°C.
PL385	4487 bp	Kanamicina	Alto	β -IFN	Plasmide di espressione derivato da PL384. Il gene che esprime β -IFN è sotto il controllo di P _L /P _R . L'induzione dell'espressione si ottiene mediante shift di temperatura da 28°C a 42°C.

OLIGONUCLEOTIDI

Nome oligonucleotidi	Sequenza	Impiego
OL 416	AACCTATGACCATGATTACGAATTCTTCC	Screening per PCR su colonia, PCR di controllo sui plasmidi d'espressione
OL 420	TGCAAGGCGATTAAGTTGGG	Screening per PCR su colonia, PCR di controllo sui plasmidi d'espressione

TRASFORMAZIONE DI *Escherichia coli* JM109

Preparazione cellule elettrocompetenti

Si effettua un preinocolo di alcune colonie, prelevate con ansa sterile da una piastra di LS precedentemente seminata, in 10ml di LS per cellule competenti (beuta da 100 ml). Si incuba overnight a 37°C in forte agitazione (220rpm).

Il giorno seguente 2,5ml di tale preinocolo vengono utilizzati per inoculare 250ml di brodo LS per cellule competenti. La coltura viene suddivisa in due aliquote da 125ml, ognuna in beuta da 1l, e queste vengono incubate a 37°C, 220rpm, fino al raggiungimento di un valore di densità ottica misurata a 600 nm (A_{600}) pari a 0,5.

La coltura è quindi aliquotata in Falcon da 50ml e le cellule vengono raccolte mediante centrifugazione a 4000xg per 15' a 4°C. Successivamente i pellets così ottenuti vengono risospesi in un totale di 240ml di glicerolo al 10% pre-raffreddato in ghiaccio e ri-centrifugati come visto. Seguono altri due lavaggi, eseguiti con le stesse modalità, effettuati rispettivamente con 120 e 60ml di glicerolo al 10%.

Dopo i tre lavaggi, le cellule sono risospese in un volume finale di circa 400µl di glicerolo al 10% freddo e suddivise in aliquote da 40µl ciascuna, in provette Eppendorf sterili.

Le aliquote sono quindi congelate in una miscela di etanolo/ghiaccio secco e conservate, per non più di sei mesi, a -80°C.

Ligazioni

Per le ligazioni dei plasmidi con i frammenti di DNA è stata utilizzata la *T4 DNA Ligasi* (BIOLABS catalogo n° M0202S), che è fornita dalla casa produttrice ad una concentrazione di 400U/µl.

Il protocollo utilizzato prevede, per ogni ligazione, l'allestimento di una miscela così composta:

- DNA plasmidico (60-100ng) e frammento in rapporto molare di 1:3
- Buffer di ligazione 2µl
- 400U o 10U di T4 DNA ligasi
- acqua milliQ q.b. ad un volume finale di 20µl

La miscela di reazione viene incubata over-night a 16°C.

Trasformazione per elettroporazione di *Escherichia coli*

Prima di poter effettuare l'elettroporazione è necessario dializzare la ligazione allo scopo di eliminare dal campione i sali che possono interferire negativamente con la procedura di elettroporazione. La dialisi è effettuata come segue:

- opportuni filtrini da dialisi Millipore (catalogo n° VSWP02500) vengono messi a galleggiare su acqua milliQ
- sul filtrino viene deposta la miscela di ligazione (20µl)

- si incuba per circa un'ora e mezza a temperatura ambiente
- si ritira la ligazione e si aggiunge acqua milliQ per riportarla ad un volume di 20µl

Per la trasformazione si usano 3µl della ligazione sottoposta a dialisi.

Per ogni trasformazione viene usata una aliquota di cellule competenti, prodotta con il protocollo sopra descritto. Le aliquote vengono scongelate in ghiaccio, viene aggiunto il DNA e, dopo un'incubazione in ghiaccio per un tempo non inferiore a un minuto, i campioni vengono trasferiti nelle opportune cuvette per elettroporazione, preraffreddate in ghiaccio. Per ogni singola trasformazione, la cuvetta, collocata nell'elettroporatore (BIORAD), subisce una scarica pari a 1,8V, come riportato nel manuale di istruzioni dello strumento. Nel caso in cui la trasformazione venga eseguita con DNA plasmidico circolare, ad ogni aliquota vengono aggiunti circa 300pg di plasmide, risospeso in acqua milliQ o in un buffer contenente 10 mM Tris-HCl a pH 8,5.

Le cellule, appena ricevuta la scarica, sono inoculate in 1ml di brodo LS sterile e mantenute in agitazione a 37°C a 220 rpm. Nel caso di trasformazione con un plasmide per espressione termoinducibile, si incuba a 28-30°C a 220rpm. Il tempo di incubazione varia a seconda della resistenza portata dal plasmide: nel caso della resistenza all'ampicillina e alla Kanamicina il tempo è un'ora. L'incubazione viene protratta per un tempo maggiore (1,5-2 ore) nel caso venga effettuata a 28-30°C.

Al termine del periodo di incubazione, le cellule vengono piastrate (1/10 e 9/10) in piastre Petri da 90mm, contenenti brodo LS + Agar e l'opportuno antibiotico.

Le piastre vengono esaminate dopo una notte a 37°C (28-30° C nel caso del sistema termoinducibile).

ESTRAZIONE E PURIFICAZIONE DEL DNA

Estrazione di DNA plasmidico (mini prep)

L'estrazione di DNA plasmidico da *Escherichia coli* è eseguita impiegando il Kit commerciale *QIAGEN Spin Miniprep* (numero catalogo: 27106). Il campione viene eluito in un tampone contenente 10 mM Tris-HCl a pH 8,5.

Estrazione da gel di frammenti di DNA

La purificazione da gel di agarosio è eseguita con il *QIAquick Spin Gel Extraction Kit* (QIAGEN catalogo n° 28704), che richiede una eluizione del campione in 30-50µl di tampone (10 mM Tris-HCl a pH 8,5), oppure acqua milliQ.

ANALISI DEI TRASFORMANTI DI Escherichia coli

Analisi dei trasformanti mediante screening per PCR

Consiste in una PCR effettuata contemporaneamente su più colonie ottenute dopo trasformazione ed ha lo scopo di evidenziare i cloni che hanno ricevuto i plasmidi portanti il frammento voluto.

L'analisi prevede la risospensione delle singole colonie in 30-40µl (a seconda della dimensione delle colonie stesse) di acqua milliQ in Eppendorf da 1,5ml. Successivamente si aliquotano, in provette per PCR da 200µl, 10 µl per ogni risospensione cellulare cui si aggiungono 10µl di mix per PCR. Le concentrazioni finali utilizzate sono le seguenti:

- 0,5µM ogni oligonucleotide
- 200µM ogni dNTP
- 1X tampone specifico (Fynnzyme)
- 0,3µl di Taq Dynazyme II (Fynnzyme, F 501L), corrispondenti a 0,15U/20µl.

Un esempio di ciclo termico in generale prevede:

N° cicli: 1	N° cicli: 25	N° cicli:1
94°C per 5'	94° C per 30"	72° C per 7'
	T° dipendente da quella di melting degli oligonucleotidi	
	72°C per 30"	

Al termine della PCR, ad ognuno dei campioni sono aggiunti 4µl di 6X loading dye (30%glicerolo in TE+blu di bromofenolo+xilencianolo) e si caricano, per ogni campione, circa 15µl su gel di agarosio allo 0,8%.

Alla quota di risospensione batterica che non è utilizzata come template per lo screening per PCR, si aggiungono 0,5ml di LS brodo addizionato di antibiotico. Tali minicolture vengono incubate a 37°C (o 28-30)°C ed utilizzate per l'inoculo dei cloni, risultati positivi allo screening, in un volume di LS che consenta l'estrazione dei plasmidi.

Analisi del DNA plasmidico per digestione enzimatica e PCR

Il DNA estratto dai cloni risultati positivi allo screening per PCR, viene analizzato mediante:

- elettroforesi su gel di agarosio;
- elettroforesi su gel di agarosio dopo digestione con opportuni enzimi di restrizione;
- PCR con i medesimi oligonucleotidi impiegati per lo screening su colonia.

PROVE DI ESPRESSIONE IN Escherichia coli JM109

Curve di crescita e raccolta dei campioni

Preparazione dei glicerinati

I glicerinati sono allestiti sterilmente in criovials da 2 ml. Essi sono costituiti da una aliquota di coltura batterica, cresciuta over-night a 28°C, portata ad una concentrazione finale di glicerolo pari al 16%. I campioni così preparati sono conservati a -80°C.

Curve di crescita a 28°C e a 42°C

Preinoculo: per ogni ceppo viene scongelato un glicerinato e si provvede ad inocularlo, in rapporto 1:100, in 10-20ml di terreno FM₁ (0,5X)+MgSO₄x7H₂O+Kanamicina (in beuta da 100ml). La coltura viene quindi incubata over-night a 28°C, 220 rpm di agitazione.

Inoculo: la mattina seguente le colture cresciute over-night vengono diluite, sulla base della lettura spettrofotometrica a 600nm, in un rapporto che oscilla tra 1:50 e 1:100 in 50ml dello stesso terreno, in beuta da 250ml; la diluizione è fatta in modo tale che i valori iniziali di assorbanza dei diversi campioni siano fra loro paragonabili.

Le colture sono poi poste a crescere alla temperatura permissiva di 28°C in shaker orbitale (220 rpm), fino al raggiungimento di un valore di assorbanza di circa 0,5. Raggiunto tale valore, ogni coltura viene suddivisa in due aliquote di circa 25ml ciascuna; una aliquota è posta a crescere alla temperatura di 28°C e l'altra a 42°C (temperatura di induzione), in beuta da 250ml e 220rpm di agitazione.

L'assorbanza a 600nm viene monitorata per circa 4 ore; al termine di tale periodo, si procede alla raccolta di pellets cellulari da 5ml per ogni coltura. I pellets, se non utilizzati subito, sono conservati a 4°C per non più di una notte.

Preparazione dei campioni per l'SDS-PAGE (Polyacrylamide Gel Electrophoresis)

I pellets ottenuti dopo 4 ore di crescita a 42°C e a 28°C, vengono risospesi in un opportuno volume di soluzione fisiologica (0,9% NaCl), in modo tale che, in termini di quantità di cellule raccolte (misurata in UA₆₀₀) per ogni µl di soluzione fisiologica, le concentrazioni cellulari nei vari campioni risospesi siano fra loro paragonabili. A 10µl di campione così risospeso vengono aggiunti 30µl di Sample Buffer; si procede quindi ad una incubazione per 10' a 99°C, in agitazione nel thermomixer. Al termine dell'incubazione, si centrifugano i campioni per 5' a 13000 rpm e si procede a caricarne 20-30µl sul gel di poliacrilammide. Come marker di peso molecolare si usa il Prestained SDS-PAGE Standard Low Range (catalogo 161-0305).

SDS-PAGE

La metodica è stata eseguita secondo il metodo di King e Laemmli (1971) utilizzando un apparato per minigel della Biorad

Soluzioni, reagenti e strumentazione

- 1) Tampone per running gel (Tris-HCl 1,5M, pH 8,8)
- 2) Tampone per stacking gel (Tris-HCl 0,5M, pH 6,8)
- 3) Acrilammide/bis-acrilammide 40% (soluzione Biorad cat.161-0148)
- 4) SDS al 10% (Biorad cat.161-0418)
- 5) APS (Ammonio PerSolfato) al 10% (Biorad cat.161-0700)
- 6) TEMED (cat.161-0800)
- 7) Tampone di corsa 10X(Tris-HCl 25mM+glicina 192mM+SDS 0,1%, pH 8,8; Biorad cat. 161-0732)
- 8) Sample Buffer 2X (Tris-HCl 0,5M pH 6,8+SDS 2%finale + glicerolo 10% finale + blu di bromofenolo 0,0025%; Biorad cat. 161-0737). Prima dell'uso, si aggiunge al Sample Buffer 2-mercapto-etanolo al 5% finale (corsa elettroforetica in condizioni riducenti).
- 9) Staining Solution (Blu di Coomassie): soluzione Biorad 1X (cat.161-0436)
- 10) Destaining Solution (acido acetico 10% + etanolo 15%)
- 11) Prestained SDS-PAGE Standard Low Range Cat. 161-0305 Biorad
- 12) Apparato elettroforetico BIORAD (mini-protean III-electrophoresis module-minitrans-blot module 165-3318)
- 13) Quantity One: è un software di acquisizione, ottimizzazione ed analisi di immagini di campioni chemiluminescenti, fluorescenti e colorimetrici
- 14) Spettrofotometro UV/visibile *Ultrospec 3100 pro* (Amersham Biosciences)

Preparazione gel

Il gel consiste di due parti, differenti fra loro per porosità e pH del tampone utilizzato per la preparazione, denominate "stacking gel" e "running gel". Lo stacking gel viene stratificato sul running gel dopo che quest'ultimo si è polimerizzato.

	STACKING GEL (ml)	RUNNING GEL (ml)			
	4 %	7,5 %	10 %	12 %	15 %
Acqua	6,32	10,93	9,69	8,69	7,19
Acrilamide	1,00	3,76	5,00	6,00	7,5
Tris-HCl pH 6,8	2,5	-	-	-	-
Tris-HCl pH 8,8	-	5,00	5,00	5,00	5,00
SDS 10%	0,100	0,200	0,200	0,200	0,200
TEMED	0,017	0,020	0,020	0,020	0,020
APS 10%	0,050	0,134	0,134	0,134	0,134

La corsa elettroforetica è condotta ad una corrente costante di 20mA per lastrina nello stacking gel e di 25mA per lastrina nel running gel, finchè il fronte di migrazione non fuoriesce dall'estremità inferiore del running gel. Si procede quindi alla colorazione con Coomassie blu per un'ora circa e alla decolorazione con Destain Solution per il tempo necessario.

RISULTATI

COSTRUZIONE DEI VETTORI PER L'ESPRESSIONE TERMOINDUCIBILE DI FP₃₄-CS-β IFN, FP₅₆-CS-β IFN e FP₇₅-CS-β IFN

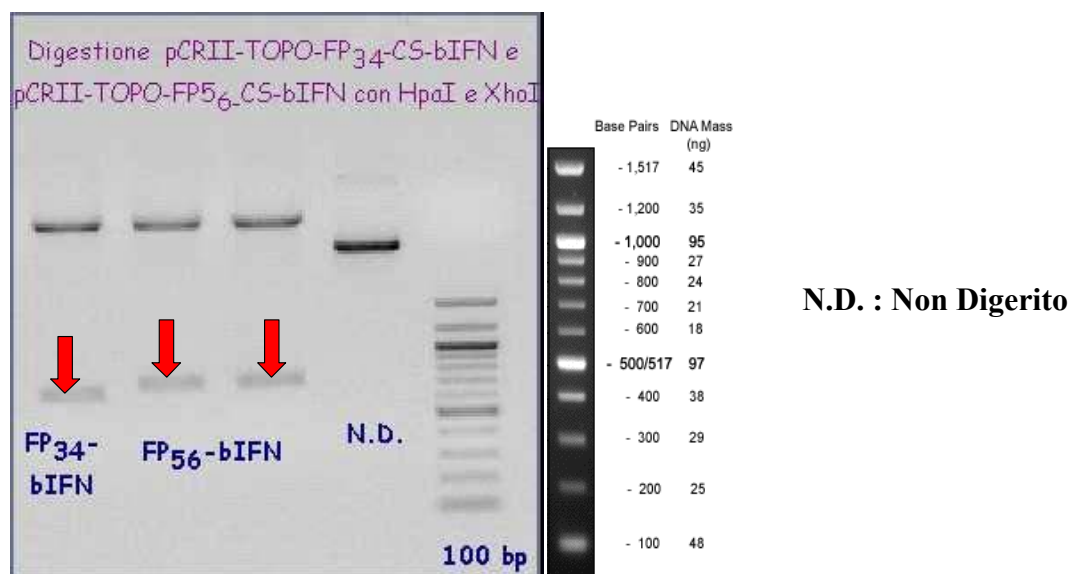
Clonaggio delle cassette d'espressione nel plasmide PL381

Per la costruzione dei plasmidi adatti all'espressione termoinducibile delle proteine di fusione FP₃₄-CS-β IFN, FP₅₆-CS-β IFN e FP₇₅-CS-β IFN, le sequenze codificanti le suddette proteine sono state excise dai vettori intermedi pCRII-TOPO-FP₃₄-CS-β IFN, pCRII-TOPO-FP₅₆-CS-β IFN e pCRII-TOPO-FP₇₅-CS-β IFN, descritti in "Materiali e Metodi", e clonate in PL381 a valle del promotore termoinducibile P_L/P_R.

In particolare, si sono allestite le seguenti digestioni, mantenute over-night a 37°C:

Plasmide	Enzimi	Frammenti ottenuti (bp)	Frammento utile
pCRII-TOPO-FP ₃₄ -CS-β IFN	XhoI+HpaI	3899+667	667 bp (codificante FP ₃₄ -CS-β IFN)
pCRII-TOPO-FP ₅₆ -CS-β IFN	XhoI+HpaI	3899+733	733 bp (codificante FP ₅₆ -CS-β IFN)
pCRII-TOPO-FP ₇₅ -CS-β IFN	Xho+HpaI	3899+790	790 bp (codificante FP ₇₅ -CS-β IFN)
PL381	XhoI+HpaI	3923+28	3923 bp (plasmide)

Esempio di excisione enzimatica dei frammenti codificanti le fusioni FP₃₄-CS-β IFN e FP₅₆-CS-β IFN dai plasmidi intermedi pCRII-TOPO:



I frammenti utili ai successivi passaggi di clonaggio sono stati estratti da gel mediante QIAquick Spin Gel Extraction Kit, si è verificata la resa dell'estrazione facendo correre su gel di agarosio un'aliquota di DNA estratto e sono state allestite le seguenti ligazioni:

Plasmide	Frammento
PL381 (estremità XhoI-HpaI)	FP ₃₄ -CS-β IFN (estremità HpaI-XhoI)
	FP ₅₆ -CS-β IFN (estremità HpaI-XhoI)
	FP ₇₅ -CS-β IFN (estremità HpaI-XhoI)
	Nessuno (ligazione di controllo)

Tali ligazioni sono state quindi utilizzate per trasformare *Escherichia coli* JM109. La selezione dei trasformanti è stata effettuata su LS/Kanamicina a 28°C e le colonie ottenute sono state sottoposte a screening per PCR utilizzando gli oligonucleotidi OL416 e OL420. Le dimensioni delle bande attese per i cloni positivi sono le seguenti:

756 bp per PL381/FP₃₄-CS-β IFN

822 bp per PL381/FP₅₆-CS-β IFN

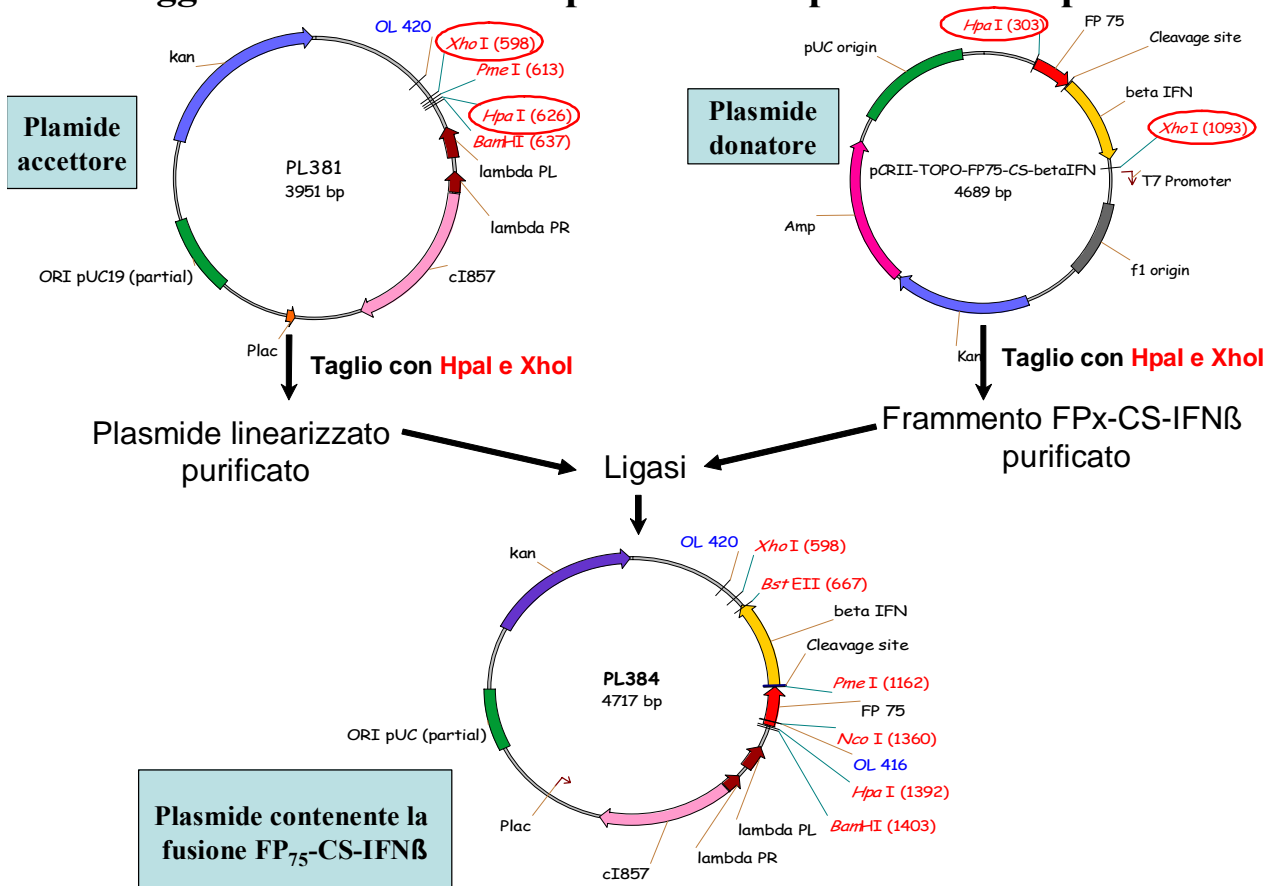
879 bp per PL381/FP₇₅-CS-β IFN

nessuna banda per cloni derivanti dalla trasformazione con ligazione di controllo.

Per ciascuna delle tre fusioni sono stati individuati diversi cloni positivi e questi sono stati inoculati in 10 ml di LS/Kanamicina e incubati over-night a 28°C a 220 rpm. Da tali inoculi sono stati estratti i plasmidi da sottoporre a successive verifiche mediante digestione e PCR.

Di seguito è riportato lo schema che descrive, come esempio, il trasferimento del frammento codificante FP₇₅-CS-β IFN dal plasmide intermedio pCRII-TOPO al plasmide di espressione PL381:

Clonaggio delle cassette di espressione nel plasmide di espressione



Controllo dei plasmidi PL381/FP₃₄-CS-β IFN, PL381/FP₅₆-CS-β IFN e PL381/FP₇₅-CS-β IFN mediante digestione e PCR

I plasmidi suddetti sono stati sottoposti ai seguenti controlli:

- Digestione con HpaI+XhoI, che excide il frammento codificante FP_x-CS-β-IFN. Tale frammento ha lunghezza differente a seconda del tipo di fusione. In particolare:

671bp nel caso di PL381/FP₃₄-CS-β IFN

737bp nel caso di PL381/FP₅₆-CS-β IFN

794bp nel caso di PL381/FP₇₅-CS-β IFN

- Digestione con BstEII, il quale taglia all'interno del gene β IFN;
- Digestione con PmeI, il quale taglia tra la sequenza codificata FP e quella codificante β IFN.
- PCR con OL416 e OL420

La analisi effettuate hanno dato i risultati attesi.

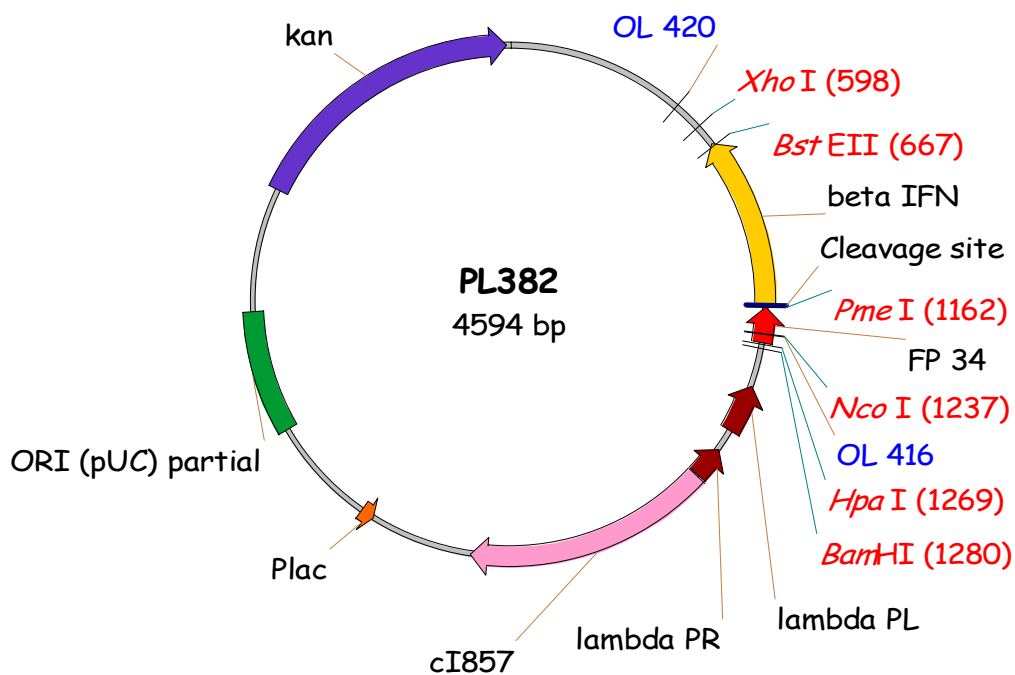
Dei cloni sottoposti a verifica per digestione e PCR, uno per ogni tipo è stato prescelto. I plasmidi sono stati nominati come segue:

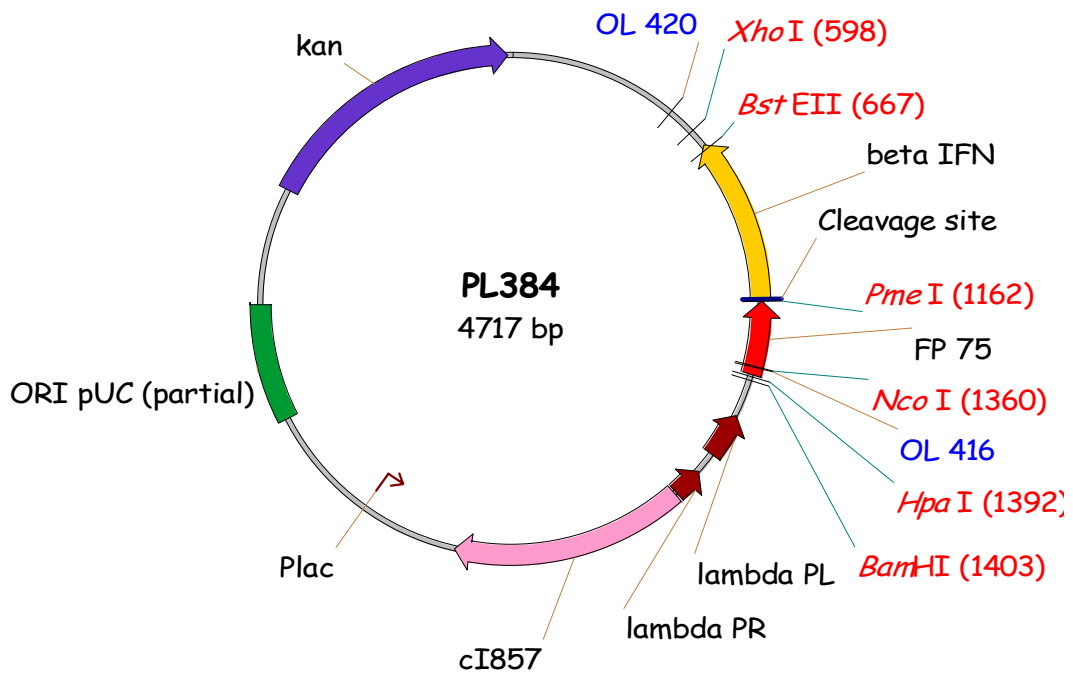
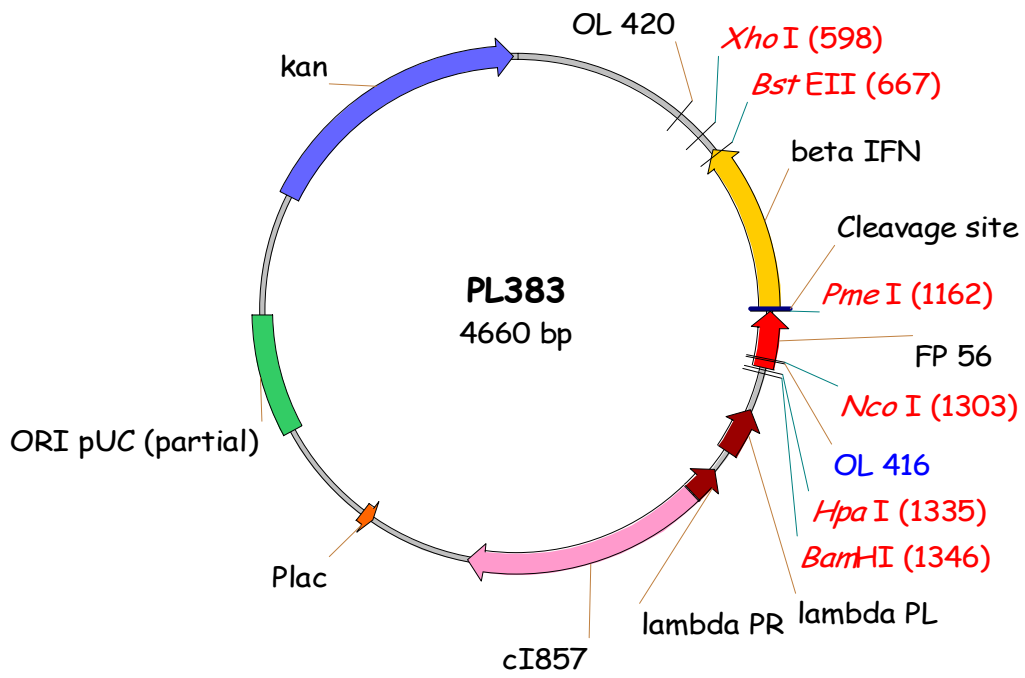
PL381/FP₃₄-CS-β IFN: PL382

PL381/FP₅₆-CS-β IFN: PL383

PL381/FP₇₅-CS-β IFN: PL384

Di seguito è riportata la rappresentazione grafica dei plasmidi costruiti (la simbologia è la stessa utilizzata per la rappresentazione grafica dei plasmidi di partenza):





COSTRUZIONE DEL VETTORE PER L'ESPRESSIONE TERMOINDUCIBILE DI β IFN

Allo scopo di confrontare i livelli di espressione ottenibili per β IFN espresso come fusione con FP e β IFN espresso in forma di proteina matura senza partner di fusione, è stato costruito un plasmide di espressione che manca della sequenza codificante il partner di fusione: β IFN è espresso cioè in forma di proteina matura naturale.

A tale scopo il plasmide PL384 è stato digerito con HpaI e PmeI a 37°C over-night; l'azione combinata dei due enzimi, permette il taglio dal plasmide della sequenza codificante il partner di fusione (FP₇₅, 230bp). La banda corrispondente al plasmide digerito è stata estratta da gel e il DNA così ottenuto è stato riligato su se stesso (sia HpaI che PmeI danno estremità piatte). La ligazione è stata quindi dializzata e utilizzata per trasformare *Escherichia coli* JM109. La selezione è stata effettuata su LS+Agar+Kanamicina a 28°C.

Le colonie ottenute sono state sottoposte a screening per PCR utilizzando gli oligonucleotidi OL416 e OL420. In questo caso sono da considerarsi positivi i cloni che non danno alcun segnale di amplificazione in quanto il plasmide atteso ha perso il sito di appaiamento per OL416. In parallelo, quale controllo positivo, è stato sottoposto a PCR il plasmide PL384, la cui amplificazione con i suddetti oligonucleotidi produce una banda di 879 bp.

Sei cloni, tra i positivi allo screening, sono stati inoculati in 10ml LS+Kanamicina e fatti crescere over-night a 28°C in agitazione a 220 rpm; dalle colture così ottenute è stato estratto il plasmide.

I plasmidi estratti dai sei cloni sono stati quindi sottoposti a controllo mediante digestioni enzimatiche e PCR. In particolare sono stati effettuate le seguenti verifiche:

- Digestione con NcoI+XhoI: il primo dei due enzimi taglia a livello della sequenza codificante il partner di fusione, quindi i plasmidi che mancano di essa vengono linearizzati (4487bp); al contrario, il plasmide PL384 (da cui deriva il PL385), utilizzato come controllo, per l'azione dei suddetti enzimi si divide in due bande distinte di 3955bp e 762bp;
- Digestione con BamHI+XhoI: le dimensioni attese per le bande risultanti da tale digestione doppia sono le seguenti

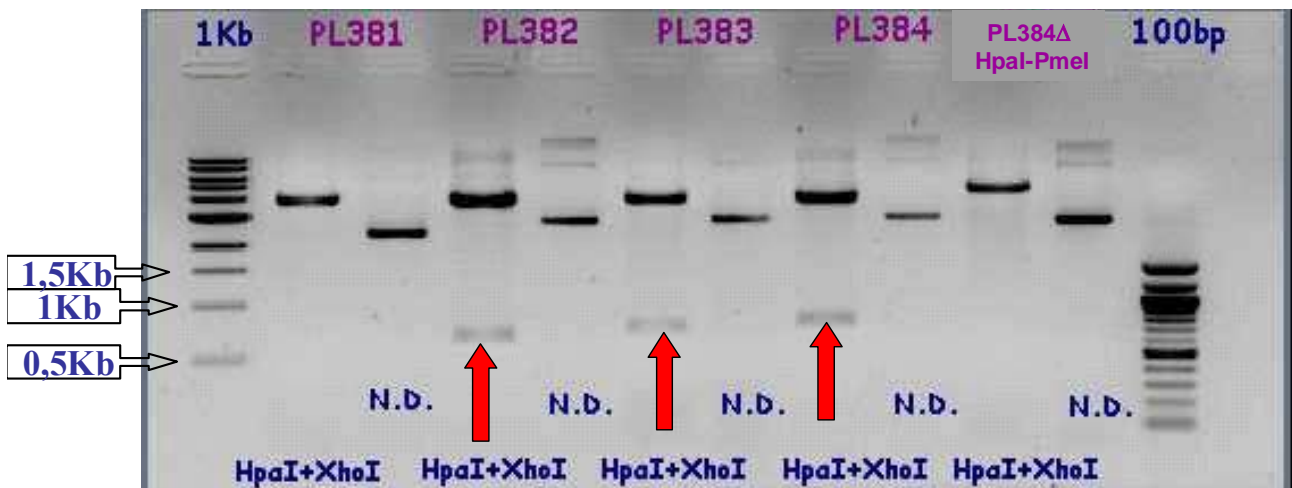
3912bp e 805bp per PL384

3912bp e 575bp per PL384 portante la delezione HpaI-PmeI (PL384 Δ HpaI-PmeI)

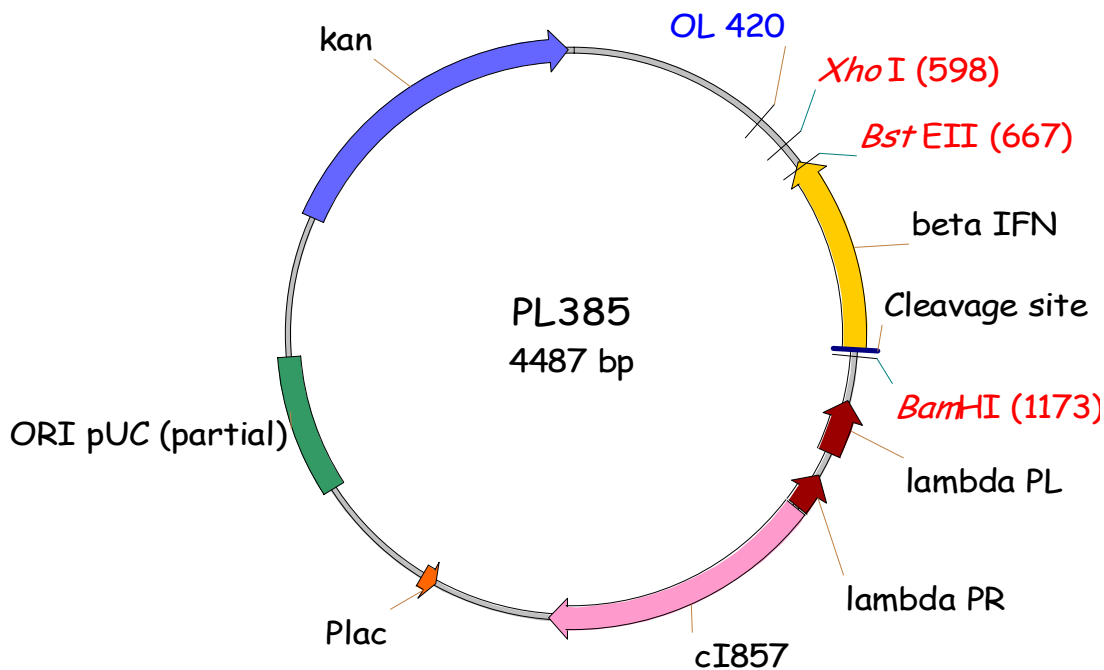
- Digestione con HpaI-XhoI: le dimensioni attese sono le seguenti

PL381	3923 bp+28bp
PL382	3923 bp+671bp
PL383	3923 bp+737bp
PL384	3923 bp+794 bp
PL384 Δ HpaI-PmeI	4487 bp

Di seguito è mostrato l'esito della digestione con HpaI e XhoI di PL381, PL382, PL383, PL384 e PL384 Δ HpaI-PmeI. Le bande contrassegnate dalle frecce rosse rappresentano i frammenti che codificano le fusioni FPx-CS- β IFN :



Sulla base delle verifiche eseguite sul plasmide mancante del partner di fusione (PL384ΔHpaI-PmeI), è stato prescelto un clone e denominato PL385.



PROVE DI ESPRESSIONE IN *Escherichia coli* JM109

Per verificare l'effettiva produzione proteica dell'Interferone beta sono state effettuate prove di espressione in terreno FM₁ 0,5X. La scelta di tale terreno è frutto di esperimenti effettuati su ceppi di *Escherichia coli* JM109 portanti il plasmide PL378 (espressione termoinducibile di FP₃₄-CS-hGH) che lo hanno indicato come un medium in grado di permettere livelli di espressione apprezzabili in SDS-PAGE.

Screening d'espressione su vari cloni derivanti da trasformazione con uno stesso plasmide

Inizialmente abbiamo voluto verificare se vi possono essere differenze nei livelli d'espressione fra diversi cloni derivanti dalla trasformazione con uno stesso plasmide.

I plasmidi:

PL378 codificante per FP₃₄-CS-hGH (controllo positivo)
 PL381 controllo negativo
 PL385 codificante per β IFN
 PL382 codificante per FP₃₄-CS- β IFN
 PL383 codificante per FP₅₆-CS- β IFN
 PL384 codificante per FP₇₅-CS- β IFN

sono stati quindi utilizzati per trasformare *Escherichia coli* JM109. Per ogni singola trasformazione sono stati utilizzati circa 300pg di plasmide circolare e la selezione è stata effettuata in LS/Kanamicina a 28°C.

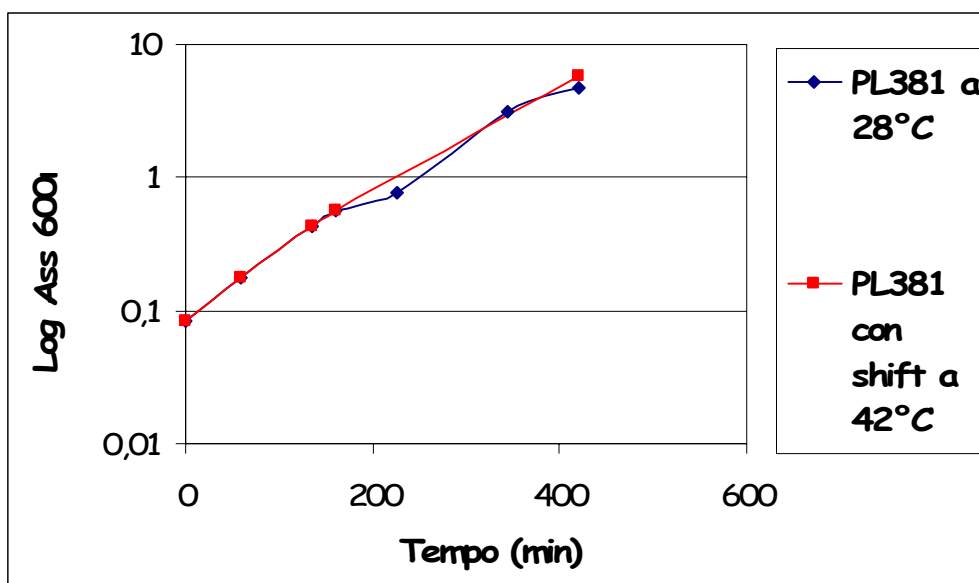
Per ogni vettore, sono state inoculate 4-5 colonie, derivanti dalla trasformazione, ognuna in 10ml di LS/Kanamicina. Dopo incubazione over-night a 28°C e 220rpm, tali colture sono state utilizzate per allestire alcuni glicerinati. Lo schema seguito è il seguente:

Trasformazione → clone 1 → glicerinati
 clone 2 → glicerinati
 clone 3 → glicerinati
 clone 4 → glicerinati

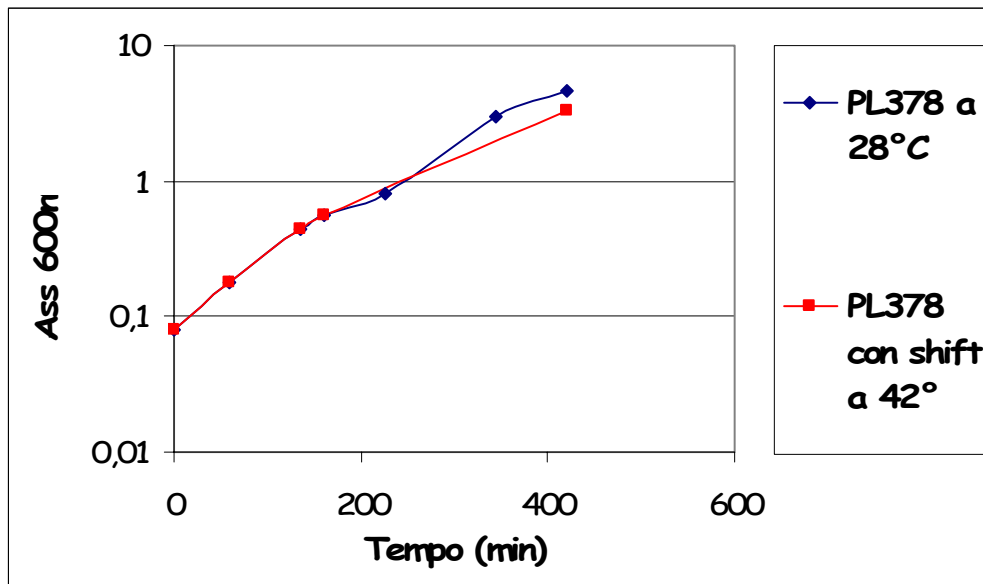
Per lo screening di espressione sono stati quindi analizzati più cloni per ogni plasmide. Le modalità seguite sono quelle descritte in “Materiali e Metodi”. In breve, è stato effettuato un preinoculo da glicerinato scongelato, un inoculo da preinoculo, è stata seguita l’assorbanza a 600nm delle colture e lo shift di temperatura da 28°C a 42°C è stato effettuato ad un valore di A₆₀₀ pari a circa 0,5. Le colture sono state quindi mantenute alla temperatura di induzione per 4 ore allo scadere delle quali sono stati raccolti i campioni per l’analisi SDS-PAGE. Il tempo di induzione è stato scelto sulla base di dati di letteratura e di risultati di esperimenti precedentemente effettuati con *Escherichia coli* JM109/PL378. Poiché non sono state rilevate significative differenze d’espressione per cloni diversi derivanti da trasformazione con uno stesso plasmide, per ogni plasmide è stato prescelto un clone singolo da sottoporre ad ulteriori prove d’espressione.

Curve di crescita e analisi d’espressione: JM109/PL382, JM109/PL385 (controlli: JM109/PL381, JM109/PL378)

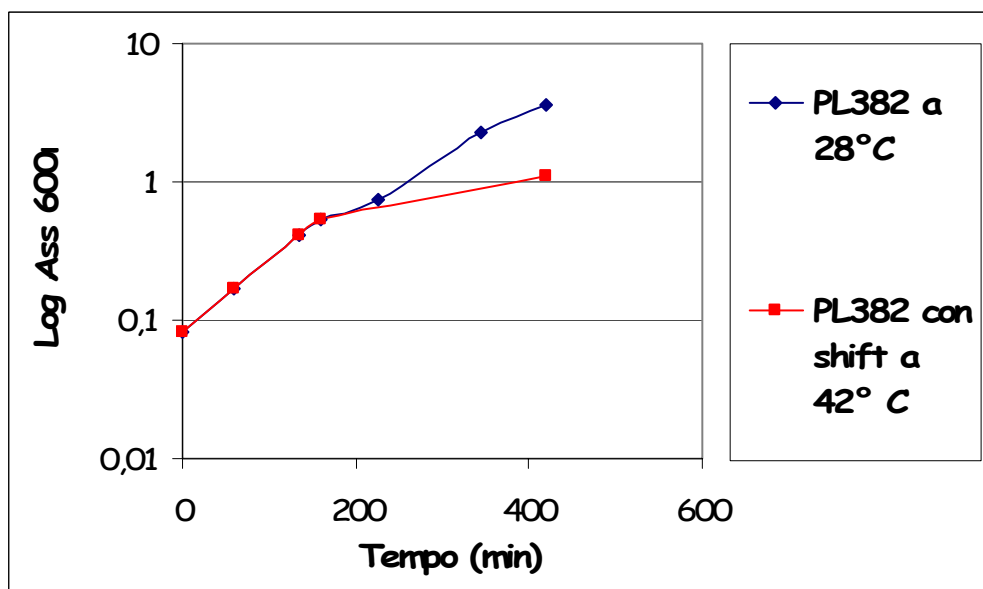
Di seguito sono mostrate le curve di crescita a 28°C e 42°C dei ceppi suddetti. Il metodo seguito è quello descritto in “Materiali e Metodi”, il terreno è FM₁ 0,5X:



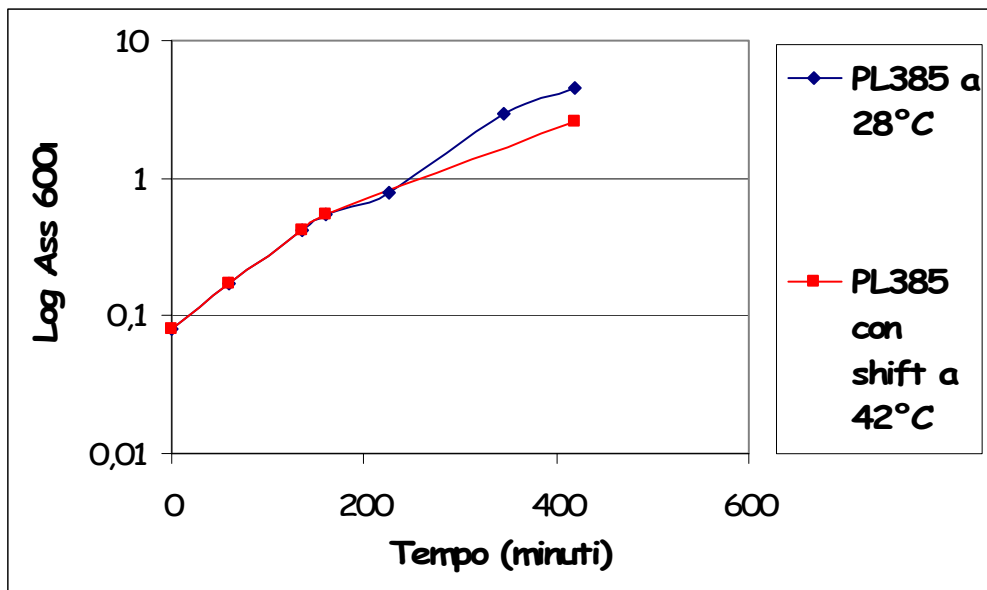
Il ceppo JM109/PL381 (plasmide vuoto, controllo negativo), non mostra alcuna differenza nella crescita alle due diverse temperature.



Il ceppo JM109/PL378 (plasmide codificante hGH, controllo positivo) non mostra differenze significative nella crescita alle due diverse temperature. Studi precedenti dimostrano che l'ormone della crescita umano in forma di proteina di fusione non è nocivo per la cellula ospite.



Il ceppo JM109/PL382 manifesta una marcata differenza nella crescita alle due temperature a seguito dell'induzione. Questa curva è rappresentativa anche delle modalità di crescita osservate nel caso dei ceppi JM109/PL383 e JM109/PL384.



Il ceppo JM109/PL385 (codificante Interferone beta senza partner di fusione) manifesta una differenza minima nella crescita alle due diverse temperature.

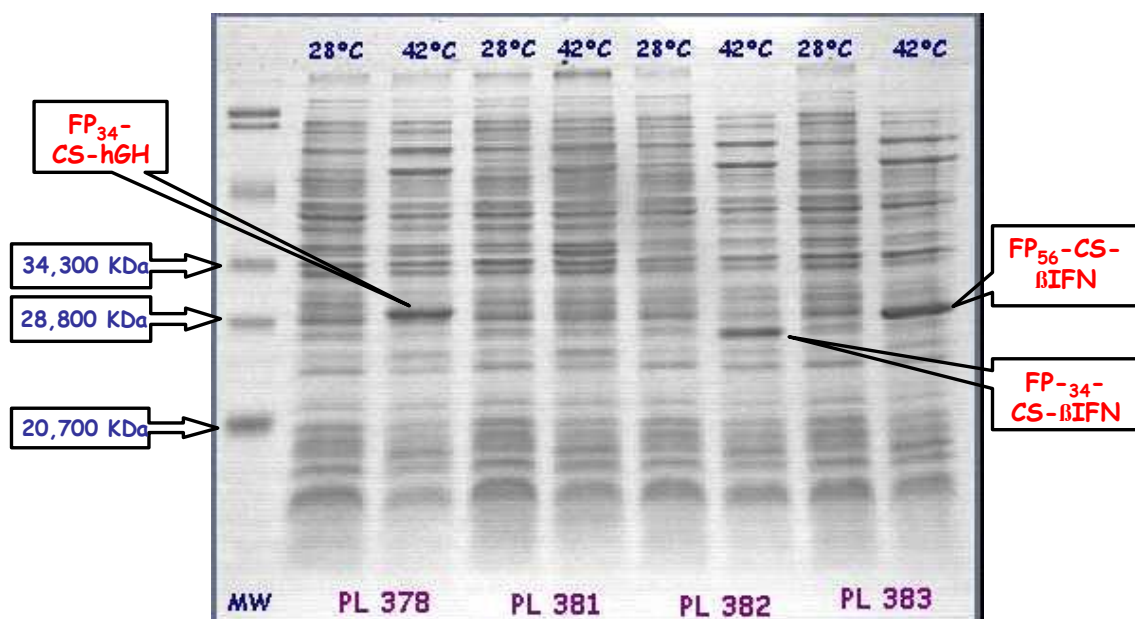
L'espressione è stata valutata mediante analisi SDS-PAGE secondo le modalità riportate in "Materiali e Metodi". Le concentrazioni di acrilammide-bisacrilammide utilizzate sono di 4% per lo stacking gel e 15% per il running gel.

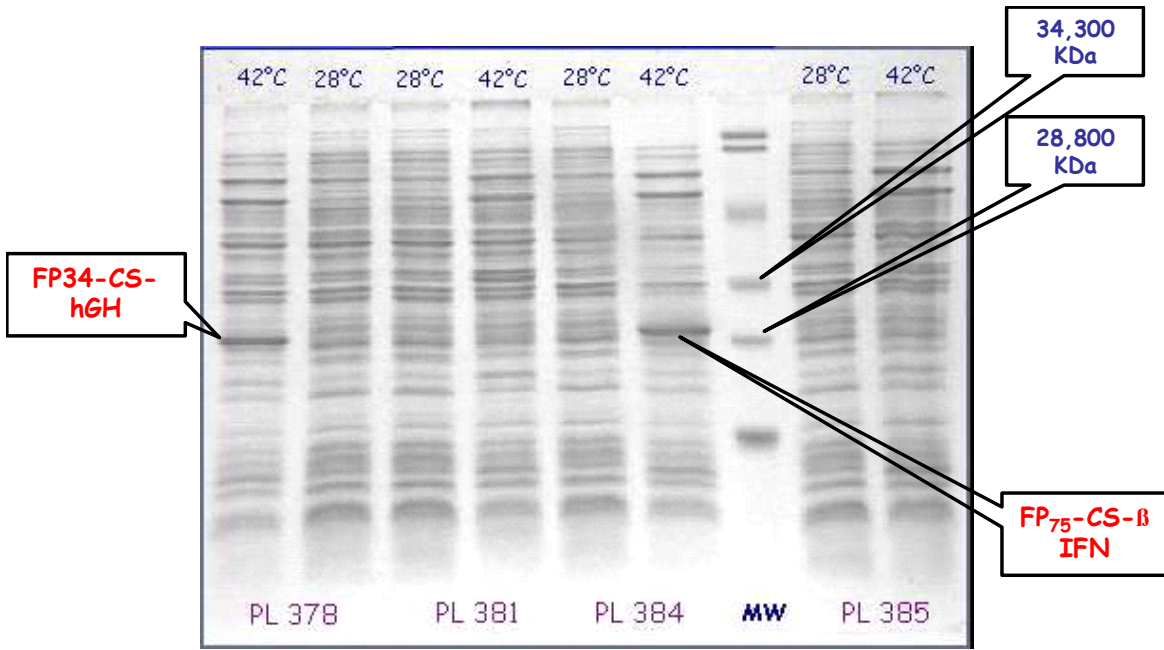
I pesi molecolari attesi per le proteine espresse sono i seguenti:

- 26,258 kDa per FP₃₄-CS-hGH (controllo positivo)
- 20,075 kDa per b IFN
- 23,937 kDa per FP₃₄-CS-β IFN
- 26,441 kDa per FP₅₆-CS-β IFN
- 28,524 kDa per FP₇₅-CS-β IFN

In blu: pesi molecolari del marker

In rosso: bande putativamente corrispondenti alle proteine di fusione





DISCUSSIONE

Il lavoro di tesi è consistito nel testare l'espressione termoinducibile in *Escherichia coli* di Interferone beta in forma di proteina di fusione.

La scelta dell'utilizzo di un sistema di espressione batterico per la produzione di Interferone beta, proteina naturalmente glicosilata, deriva dalla non indispensabilità della glicosilazione ai fini della sua attività biologica (15).

Il carico metabolico dovuto sia alla replicazione del plasmide sia all'espressione della proteina ricombinante può ridurre drasticamente la velocità di crescita di un batterio ingegnerizzato. In alcuni casi infatti, come nel caso dell'Interferone beta, una produzione costitutiva della proteina di interesse può risultare tossica per la cellula batterica. In queste condizioni è necessario utilizzare un sistema di espressione inducibile in cui la produzione della proteina ricombinante possa essere repressa durante la fase di crescita cellulare ed indotta successivamente, nella fase finale del processo fermentativo.

Il sistema adoperato nel lavoro di tesi è inducibile per shift di temperatura e consiste nell'impiego dei promotori p_L e p_R e del repressore cI del fago lambda. In esso l'attività dei due promotori può essere repressa dalla proteina repressore che è codificata dall'allele $cI857$ del gene cI ; quando il repressore è presente l'espressione della proteina posta sotto il controllo dei promotori è inibita. L'allele $cI857$ del gene cI codifica per un repressore termosensibile: esso reprime a temperature comprese fra i 30°C e i 37°C ed è inattivo a 42°C a causa della ridotta stabilità termodinamica. Questa caratteristica rende il sistema inducibile per shift di temperatura (16). L'induzione termica è fra i più semplici sistemi di induzione e l'impiego di plasmidi d'espressione che portano i promotori fagici P_L e P_R in tandem consente di ottenere alte densità cellulari; ciò rende il sistema potenzialmente utile per applicazioni industriali.

Il sistema termoinducibile presenta anche degli inconvenienti:

- la condizione di stress termico può determinare l'attivazione della risposta SOS e ciò determina un arresto della replicazione del DNA e, nel caso di ceppi lisogeni, anche lisi cellulare;
- l'incremento dell'espressione delle proteine da shock termico (*heat shock proteins*) che, a seconda della natura delle stesse, può risultare indesiderabile (nel caso, ad esempio, di proteasi).

Il ceppo da noi utilizzato (JM109) è deficiente nella risposta SOS, in quanto $recA^-$. Ciò dovrebbe ridurre l'attivazione di tale risposta.

I risultati del lavoro di tesi confermano la necessità di impiegare un sistema di espressione inducibile nella produzione di Interferone beta in *E. coli*. Infatti, come si evince dalle curve di crescita precedentemente viste, ma anche dalla letteratura e da precedenti esperienze in Bioker, l'espressione di tale proteina ha effetti negativi sulla crescita del ceppo batterico. Ciò sembra sia correlato alla natura idrofoba della proteina, che potrebbe determinarne l'adesione alla membrana cellulare con conseguenze negative sul metabolismo della cellula (17). E' bene altresì sottolineare che la temperatura di 42°C, necessaria all'induzione, potrebbe determinare altri effetti indesiderati, oltre a quelli già citati, quali ad esempio: alterazioni della parete cellulare, formazione di corpi di inclusione, folding non corretto (18).

Dato l'effetto negativo dell'espressione di Interferone beta sulla crescita batterica, si è scelto di far crescere i batteri alla temperatura di 28° finché non fosse raggiunta una adeguata biomassa cellulare. Alla temperatura di 28°C l'espressione della proteina è completamente repressa: ciò è stato verificato sia tramite SDS-PAGE che tramite Western Blot (dati non mostrati). In queste condizioni il tempo di duplicazione è di circa un'ora, indipendentemente dal tipo di proteina espressa. Tale bassa velocità di crescita non pare dunque imputabile a una minima produzione della proteina, ma più semplicemente alla temperatura subottimale. A seguito dell'induzione (42°C) la crescita subisce un forte rallentamento solo nei ceppi che producono la proteina di fusione, al

contrario di ciò che si osserva in ceppi che non esprimono la proteina di fusione (PL385, PL381, PL378). In particolare, confrontando fra loro i vari ceppi testati, si può dedurre che:

- il mancato rallentamento della crescita in JM109/PL381 (plasmide vuoto) può essere imputato al fatto che esso non produce Interferone beta

- il mancato rallentamento della crescita in JM109/PL385 (plasmide codificante Interferone beta senza partner di fusione) può essere correlato alla bassa o nulla espressione di Interferone beta (non rilevabile in SDS-PAGE, dato da confermare mediante Western Blot) alla temperatura di induzione. Per quanto riguarda il ceppo utilizzato come controllo positivo (JM109/PL378), in cui l'espressione dell'ormone della crescita umano in forma di proteina di fusione (FP₃₄-CS-hGH) è regolata dallo stesso promotore usato per l'Interferone beta, il mancato rallentamento della crescita a 42°C sembra dimostrare che l'effetto inibente sulla crescita sia dovuto principalmente al tipo di proteina prodotto. Il ceppo in questione, infatti, esprime una proteina che, come dimostrano precedenti esperienze in Bioker, non è nociva per *E. coli*.

I ceppi cresciuti nel terreno LS₁₀ hanno mostrato una ridotta crescita a 28°C e nessuna espressione proteica a seguito dell'induzione; al contrario, l'impiego di un terreno più nutriente come FM₁, ha permesso una crescita più sostenuta e una produzione della proteina di interesse rilevabile in SDS-PAGE a seguito dell'induzione.

Per garantire l'espressione di una proteina umana in un sistema batterico non sempre è sufficiente la presenza di un promotore forte. Altri fattori sono infatti da considerare, in particolare: la presenza di codoni rari nel gene di interesse e le caratteristiche delle regioni coinvolte nel processo di inizio della traduzione. Per garantire un'efficace traduzione del gene per l'Interferone beta si è ricorsi alla tecnologia delle proteine di fusione.

La tecnologia delle proteine di fusione è caratterizzata da due aspetti principali;

- A. Produzione di una proteina ricombinante in forma di polipeptide di fusione avente struttura generale: (FPx)-(CS)-(Proteina ricombinante)
- B. Identificazione ed uso di un reattivo chimico o di un enzima proteolitico in grado di riconoscere selettivamente il sito di taglio (CS) e di rimuovere in maniera specifica la porzione FPx liberando così la proteina di interesse (P) nella sua forma nativa.

Nel lavoro di tesi l'Interferone beta è espresso come proteina di fusione con un partner (FP) la cui presenza all'estremità ammino-terminale della proteina ha dimostrato, in studi precedenti in Bioker, incrementare notevolmente l'espressione di proteine altrimenti poco espresse (o per nulla espresse) in *Escherichia coli*. L'identità del partner di fusione fa parte del know-how aziendale, non protetto da brevetto, che non è possibile divulgare.

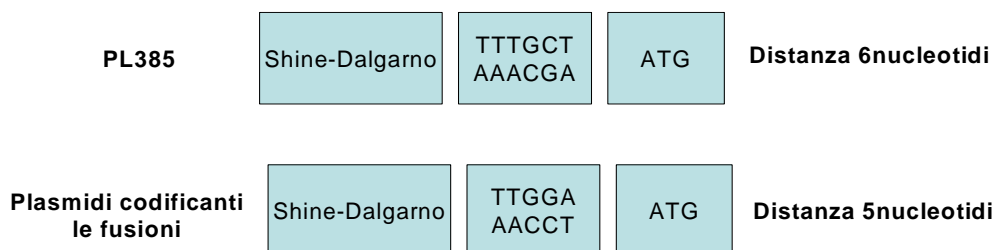
Il livello di espressione di una proteina ricombinante è spesso correlabile alla sequenza e/o alla struttura del gene codificante per la porzione N-terminale della proteina stessa, mentre la sequenza codificante per le porzioni distali della catena polipeptidica non pare direttamente implicata nel determinare l'efficienza dell'inizio della traduzione.

Avendo, in precedenza, empiricamente verificato alti livelli di espressione per una data combinazione ospite-vettore-gene codificante per un partner di fusione (FPx), abbiamo deciso di esprimere Interferone beta come fusione con il medesimo partner (FP). Si è altresì scelto di mantenere, nel costrutto, una distanza, tra Shine-Dalgarno e ATG d'inizio, pari a 5nucleotidi, in quanto dati di letteratura (8) e precedenti esperienze dimostrano che tale distanza favorisce un efficiente inizio di traduzione.

Tra il partner di fusione e il gene per l'Interferone beta umano è stata posta una sequenza che è riconosciuta in maniera specifica da una endoproteasi, che è prodotta in Bioker e di cui non è possibile, per motivi aziendali, citare l'identità.

Per valutare l'importanza del partner di fusione nell'espressione, si sono confrontati i livelli di espressione ottenibili per β IFN espresso come fusione con il partner di fusione (FPx) e β IFN espresso in forma di proteina matura senza FPx; a tal fine è stato costruito un plasmide controllo di espressione che manca della sequenza codificante il partner di fusione (PL385): β IFN è espresso cioè in forma di proteina matura naturale. I risultati dimostrano che l'assenza del partner di fusione

conduce ad un livello di espressione della proteina di interesse non rilevabile in SDS-PAGE. È bene sottolineare che il PL385 è stato ottenuto per digestione di uno dei plasmidi d'espressione contenenti il partner di fusione (nel dettaglio la digestione ha riguardato il plasmide PL384). Ciò ha come conseguenza lievi differenze, nella sequenza compresa tra Shine-Dalgarno e ATG, fra PL385 e i plasmidi codificanti le fusioni. In particolare si ha:



Tali differenze non paiono tali da giustificare il netto calo di espressione evidenziato con PL385. Ci proponiamo, comunque, di allestire un costrutto che sia un controllo perfetto in modo da verificare gli eventuali effetti di tali variazioni di sequenza.

Nel presente lavoro è stato testato l'effetto, sui livelli di espressione, di tre diverse lunghezze del partner di fusione; l'analisi in SDS-PAGE non ha messo in luce differenze significative nei livelli di espressione: la minima lunghezza impegnata (FP₃₄) determina livelli di espressione paragonabili a quelli ottenuti con FP_x di lunghezza maggiore. Questa condizione potrebbe avere notevole importanza in sede di purificazione della proteina e in particolare nel refolding della stessa, qualora venisse espressa in forma di corpi di inclusione. Dati preliminari attestano che Interferone beta, nelle condizioni da noi adottate, è prodotto in tale forma.

La produzione di proteine ricombinanti in batteri come *E. coli* frequentemente conduce alla produzione di una proteina inattiva, aggregata in forma dei cosiddetti "corpi di inclusione"; questa condizione è ulteriormente favorita anche dall'impiego di alte temperature, necessarie nel caso dell'applicazione di un sistema termoinducibile.

Sebbene la produzione di una proteina in forma di corpo di inclusione possa rappresentare uno svantaggio in quanto impone uno step di refolding, tale condizione presenta anche diversi vantaggi:

- Si ha un più alto grado di purezza della proteina target nella frazione aggregata
- La proteina in forma di corpo di inclusione è in una condizione di maggiore protezione nei confronti della degradazione proteolitica rispetto alla proteina nella frazione solubile.

In generale i passaggi attraverso i quali si ottiene la proteina biologicamente attiva a partire da corpi di inclusione sono i seguenti:

- 1) Rottura delle cellule per omogeneizzazione in presenza di detergenti.
- 2) Separazione dei corpi inclusi dal materiale cellulare grazie a diverse fasi di centrifugazione, sempre in presenza del detergente.
- 3) Si risospendono i corpi di inclusione in un buffer di lavaggio al fine di eliminare il detergente.
- 4) Si denaturano i corpi di inclusione (guanidina o altro); successivamente si riduce quasi a zero la concentrazione dell'agente denaturante: in queste condizioni avviene il refolding della proteina. A parità di altri fattori, non è da escludere che una maggior lunghezza del partner di fusione possa diminuire l'efficienza del refolding. Il fatto che la lunghezza minima di FP testata determini un livello di espressione paragonabile a quello ottenuto con lunghezze maggiori è dunque un vantaggio.

- 5) Si concentra la soluzione per limitare i volumi: volumi troppo grandi richiedono costi e tempi di purificazione notevoli.
- 6) Si dializza la soluzione concentrata per eliminare completamente il denaturante.
- 7) Si opera il taglio con la proteasi prodotta in Bioker: l'idrolisi proteolitica della porzione (FPx)-(CS)-(β IFN) è relativamente indipendente dalle caratteristiche strutturali della proteina fusa (β IFN), purché il sito di taglio sia in una porzione della catena polipeptidica relativamente flessibile ed esposta, così da permettere una facile accessibilità necessaria al riconoscimento del substrato da parte della proteasi.
La scelta del sistema di idrolisi della proteina di fusione costituisce un fattore critico dell'intero processo, in quanto deve rispondere a requisiti di specificità e selettività di taglio, deve avere una rapida cinetica di idrolisi e non dovrà modificare la struttura della proteina di interesse. La proteasi utilizzata in Bioker si è dimostrata essere economicamente vantaggiosa e sufficientemente specifica.
- 8) Si purifica la proteina per via cromatografica: a seconda della proteina si utilizzano metodi cromatografici differenti.

Per il futuro ci proponiamo dunque di:

- verificare che effettivamente Interferone beta sia prodotto in forma di corpi inclusi;
- allestire prove di refolding su piccola scala;
- ottimizzare le condizioni di taglio al fine di minimizzare, se presente, il taglio proteolitico a livello di un sito che è naturalmente posto entro la sequenza di Interferone beta;
- testare l'espressione utilizzando un gene codificante Interferone beta ottimizzato per il codon usage di *E. coli*.

BIBLIOGRAFIA

- (1): “ *Production of recombinant protein*” Gerd Gellissen (ed) WILEY-VCH.
- (2): From: *Gene Expression System. Using nature for the art of expression* (Fernandez, J.M. & Hoeffler, J.P. eds) Academic Press, San Diego, 1999.
- (3) “*Advances in Escherichia coli production of therapeutic proteins*” (James R Swartz) Current Opinion in Biotechnology 2001, 12:195-201.
- (4) “*Recombinant protein expression in Escherichia coli*” (Francois Baneyx) Current Opinion in Biotechnology 1999, 10:411-421.
- (5) “*Strategies for optimising heterologous protein expression in Escherichia coli*” (Gerhard Hanning and Savvas C. Makrides) Trends in Biotechnology 1998 Elsevier Trends Journals.
- (6): “*Strategies for Achieving High-Level Expression of Genes in Escherichia coli*” (Savvas C. Makrides) Microbiological Reviews 1996, p.512-538.
- (7): “*Advanced genetic strategies for recombinant protein expression in Escherichia coli*” (Hans Peter Sorensen, Kim Kusk Mortensen) Journal of Biotechnology 115 (2005) 113-128.
- (8): “*Production of recombinant proteins in Escherichia coli*” (Wolfgang Schumann and Luis Carlos S. Ferreira) Genetics and molecular Biology, 27,3, 442-453 (2004).
- (9): “*Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia E. coli mRNAs*” (Hongyun Chen, Matthew Bjerknes, Ravindra Kumar et. Al) Nucleic Acids Research 1994, Vol.22 , No.23 (4953-4957).
- (10): “*A vector with the downstream box of the initiation codon can highly enhance protein expression in Escherichia coli*” (Xiaolan Zhang, Peng Guo & Guozhong Jing) Biotechnology Letters 25: 755-760 (2003).
- (11): “*Interferons: cell signalling, immune modulation, antiviral responses and virus countermeasures*” (S. Goodbourn, L. Didcock and R. E. Randall) Journal of General Virology (2000), 81, 2341-2364.
- (12): “*Diagnosis and Management of Multiple Sclerosis*” (Peter A. Calabresi, M.D, Johns Hopkins) University School of Medicine Baltimore, Maryland.
- (13): “*Caleidoscopio Italiano*” (Ircano Cogato e Enrico Montanari) 1996.
- (14): “*Multiple Sclerosis: the disease and its treatment*” (A. Annapurna, V. Krishna Kumar, et. Al) Indian Journal of Pharmacology 2002; 34: 3-15.
- (15): “*Structural and functional differences between glycosylated and non-glycosylated form of human Interferon β* ” Pharmaceutical research vol.15 No.4 (1998).
- (16): “*Constitutive versus thermoinducible expression of heterologous proteins in E.coli based on strong $P_R P_L$ promoter from Phage Lambda*”; www.InterScience.wiley.com .

(17): "Inducible high level synthesis of mature human fibroblast interferon in *E. coli*" in Nuclear acids research volume 11 n°14 1983.

(18): "Fine regulation of cI857-controlled gene expression in continuous culture of recombinant *E. coli* by temperature" vol.59 n°10.