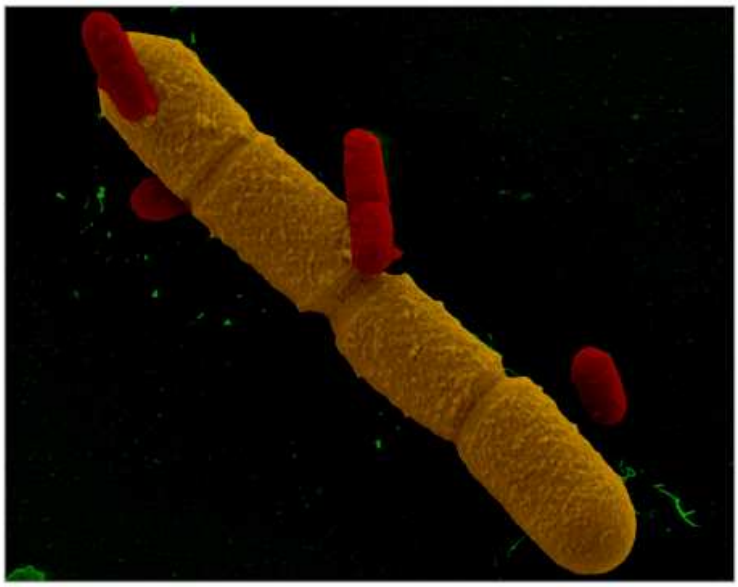
SINGLE CHAIN ANTIBODY FRAGMENT PRODUCTION AND METABOLIC ENGINEERING OF HYALURONIC ACID IN *BACILLUS* AND *ESCHERICHIA COLI*



Nadia Cavallarin

UNIVERSITÀ DI PADOVA

Dipartimento di Scienze Sperimentali Veterinarie

Scuola di dottorato di ricerca in Scienze Veterinarie

Indirizzo Scienze Biomediche Veterinarie e Comparate

Ciclo XXIII

SINGLE CHAIN ANTIBODY FRAGMENT PRODUCTION AND METABOLIC ENGINEERING OF HYALURONIC ACID IN *BACILLUS* AND *ESCHERICHIA COLI*

Nadia Cavallarin

Università degli Studi di Padova

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Sperimentali Veterinarie

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE VETERINARIE

INDIRIZZO SCIENZE BIOMEDICHE VETERINARIE E COMPARATE

CICLO XXIII

SINGLE CHAIN ANTIBODY FRAGMENT PRODUCTION AND METABOLIC ENGINEERING OF HYALURONIC ACID IN *BACILLUS* AND *ESCHERICHIA COLI*

Direttore della Scuola: Ch. mo Prof. Massimo Morgante

Supervisore: Ch.mo Prof. Alessandro Negro

Dottoranda : Nadia Cavallarin

Contents

Summary

Escherichia coli is the "workhorse" for the production of recombinant proteins due to numerous advantages. It is well known for its ease of handling with short time generation and for its ability to accumulate high amount of foreign protein. Nevertheless *E. coli* systems present also some limits; serious problems can occur during the process of heterologous gene expression and purification, as low expression rates, formation of inclusion bodies, improper protein-folding and inability to produce complex disulfide bonds. This is due to lacks fundamental prerequisites for efficient secretion, due to the membrane structure, the low chaperone and foldase level and the high periplasmatic protease concentration. Alternative expression hosts are the Gram-positive *Bacillus* strains, particularly *B.subtilis* and *B.megaterium*, designated as GRAS (generally recognized as safe) organisms, free of any endotoxin. Furthermore, compared with *E.coli*, *B.subtilis* and *B.megaterium* offer high biosynthetic capacity and an efficient secretion apparatus that guides the expressed proteins directly into the culture supernatant.

The aim of this project was to evaluate these microorganisms as expression hosts 1) either for production of heterologous eukaryotic proteins, as the anti-prion 8H4 antibody fragment (cap.I-III), 2) either for engineering of a natural metabolic pathway, as biosynthesis of hyaluronic acid (cap.IV).

8H4 scFv (Single chain Fragment variable) is an eukaryotic complex molecule with three disulphide bonds, that has been recently used as therapeutic approach in prion diseases, inhibiting prion replication and delaying the development of prion disease. In order to optimize this therapeutic perspective and address the 8H4 scFv to its targets in SCN, 8H4 scFv has been cloned in fusion with peptide-transduction domain of HIV-1 TAT protein, that penetrates efficiently into cells translocating across the plasma membrane and also the blood-brain barrier. TAT 8H4svFv protein was been produced previously in our laboratory in *E.coli*, but with very low yields.

In this study, WB800N *B.subtilis* (deleted of 8 extracellular proteases) and MS941 *B.megaterium* were engineered to produce and secrete anti-prion TAT-8H4scF protein .

Several expression-secretion vectors have been constructed, cloning the wild-type or a synthetic *E. coli* optimized TAT 8H4 scFv sequence, under the control of different promoters (IPTG-inducible Pgrac or strong constitutive P43 promoters in *B.subtilis*, or xylose-inducible T7 promoter in *B.megaterium*), in fusion with secretion signal peptide AmyQ. Although the chosen *Bacillus* strains were deleted of several extracellular proteases, no clearly TAT 8H4scFv secreted was revealed in our hand, (neither after IMAC purification of culture medium, neither after ammonium

sulphate precipitation). Antibody was present overall in insoluble intracellular fraction, and a small fraction was revealed in periplasm. (cap.II)

Because the high secretory capacity of *Bacillus* strains is not appeared, we have tried to take advantage from the use of fusion protein technology to increase the yields and solubility of TAT8H4scFv in *E.coli* cytoplasm. At the aim bacterial chaperone DnaK and like-chaperone α-synuclein protein were been chosen as fusion tags, for their activity favouring refolding together physic-chemical characteristics. Although fusions of TAT 8H4scFv to α-synuclein and DnaK increase expression resulting in accumulation of significant levels of antibody, these fusion proteins show to be largely insoluble. Thus, we have tried to co-express DnaKJE and GroELS bacterial chaperones together α-syn or DnaK fusion-8H4scFvs, and purify α-synTAT8H4 from more oxidizing periplasmic environment. Since the fusion-scFvs yields in soluble form remained not significant, we have purified Dnak TAT8H4 and α-syn TAT8H4scFvs in denaturing conditions by immobilized metal affinity chromatography (IMAC) and in vitro refolding. In this way has been possible to obtain correctly refolded α-syn 8H4, that is able to recognize human prion protein by immunoblotting. When Dnak TAT8H4 and α-syn TAT8H4 were added to CHO cells culture medium, they were rapidly delivered inside the cells, and displayed mainly a nuclear localization. α-syn TAT8H4 antibody fragment is able to deplete the superficial membrane bound prion protein in HeLa transfected cells with plasmids GFP-PrP but not the analog GFP-Doppel. Time course of these protein revealed that α-syn TAT8H4 shows a half-life major than Dnak TAT8H4 (cap.III).

In conclusion fusion protein technology is revealed effective to increase the yields of TAT8H4scFv in *E.coli* cytoplasm. α–synuclein is demonstrated preferable to DnaK in fusion with TAT8H4 scFv, either for lower molecular weight in fusion protein, than for higher yields and ability to refold. Due to its specificity to deplete PrP, α–syn TAT8H4 scFv could be effective in spongiform transmissible diseases treatment. Moreover, the use of cell-permeable antibodies, due to TAT transduction domain, would avoid the safety and ethical concerns surrounding the direct application of recombinant DNA technology in human clinical therapy and could be extended to treatment of other pathology.

The second part of this study concerns the engineering of a natural metabolic pathway, as the biosynthesis of hyaluronic acid HA. In our project the HA biosynthesis pathway present in natural producers as streptococci has been has been adapt to *Bacillus*.

Genes involved in the pathway of synthesis for the precursor sugars, that are disperse in the genoma of *Bacillus subtilis* , were here linked in a unique polycistronic mRNA togheter with hyaluronan synthase from *Streptococcus equi*, as is present in natural pathway of streptococci. In

our case we have developed several episomal vector where several genes are under the control of strong and inducible promoters. This expression system, using of a plasmid with relatively higher copies, it has advantage to express more higher levels of mRNA than integrative system on the chromosome.

By PCR amplification of *hasA* and *tuaD* genes from *S.zooepidemicus* and *B.subtilis* respectively, and cloning in pHT *B.subtilis/E.coli* shuttle vector under Pgrac, we have selected stable metabolic engineered 1012 and WB800N *Bacillus subtilis* strains, secreting HA with molecular weights higher of 800 kDa of *Streptococcus* with a yield more than 5g/L, that are actually produced by *Strepctococcus* in industrial HA and largely exceeding that is published. To further increase yields and HA molecular weights, we have construct and cloned the cassette-operons *hasA-tuaD* and *hasA-tuaD-gtaB-pgi* under inducible promoter T7 in *B. megaterium*. Althought T7 expression system is present also in *E.coli* , we have demonstrated that recombinant *E.coli* cells produce only low amounts of HA. Indeed these engineered *B.megaterium* strains in the optimal expression conditions identified produce about 2g/L in shake flask, that are very promising results in view of batch fermentation cultures. Moreover *hasA-tuaD-gtaB-pgi* overexpressing *B.megaterium* seem produce higher HA MW, about 1800 kDa, comparable also for polydispersity to commercially available *Streptococcus*, suggesting that *gtaB-gpi* overexpression result in a molecular weights enhancement (cap.IV).

In conclusion, although high potential secretory capacity of *Bacillus* was not appear in secretion of a heterologous protein, as anti-prion 8H4 scFv, *B. subtilis* and *B.megaterium* have proven to be superior expression hosts for engineering of a natural metabolic pathway, as biosynthesis of HA, based on several criteria: 1) good quality of HA, comparable to commercial *streptococcus* standards regards to molecular mass and polydispersity, and superior regards yields. 2) In addition, unlike *Streptococcus*, the *B. subtilis* and *B.megaterium*-derived HA products are exotoxin free and secreted directly into the surrounding medium and are not cells associate, simplifying the recovery process. 3) Finally, while *Streptococcus* A and C require more expensive complex media for growth, *Bacillus* strains grow on minimal media, assuring a final products more pure and toxin-free.

Sommario

Escherichia coli è " il cavallo di battaglia" nella produzione di proteine ricombinanti per i numerosi vantaggi. E'nota la sua facilità di manipolazione con tempi brevi di duplicazione e la sua capacità di produrre grandi quantità di proteine eterologhe. Tuttavia *E.coli* presenta anche alcuni limiti; difficoltà possono insorgere durante l'espressione e purificazione di proteine eterologhe, quali bassi livelli di espressione, formazione di corpi di inclusione, ripiegamento proteico improprio e incapacità di formare complessi legami disolfuro. Questo è sostanzialmente dovuto alla mancanza dei prerequisiti necessari per una efficiente secrezione, a causa della struttura della membrana, del basso livello di "chaperoni" e dell'elevata concentrazione di proteasi periplasmatiche.

Ospiti di espressione alternativi sono i ceppi batterici gram-positivi di *Bacillus*, in particolare *B.subtilis* e *B.megaterium*, classificati come organismi GRAS (generally recognized as safe) che non producono endotossine. Inoltre, rispetto a *E.coli*, *B.subtilis* e *B.megaterium*, possiedono un'alta capacità biosintetica e secretoria naturale con secrezione delle proteine espresse direttamente nel mezzo di coltura.

Scopo del progetto era la valutazione di questi microrganismi come ospiti di espressione 1) sia per la produzione di proteine eterologhe eucariotiche, come il frammento anticorpale anti-proteina prionica 8H4 (cap I-III), 2) sia per l'ingegnerizzazione di una via metabolica naturale, quale la biosintesi dell'acido ialuronico (cap-IV).

Il frammento anticorpale a catena singola (scFv) 8H4 è una complessa molecola eucariotica con tre ponti disolfuro, recentemente utilizzata nell'approccio terapeutico verso le malattie prioniche, in quanto in grado di inibire la replicazione dei prioni ritardando lo sviluppo della patologia. Nel tentativo di ottimizzare questa prospettiva terapeutica si è pensato di indirizzare il frammento anticorpale 8H4 al suo target nel sistema nervoso centrale, clonandolo in fusione con il dominio di trasduzione peptidica della proteina HIV-1 TAT, che penetra efficientemente nelle cellule traslocando attraverso la membrana plasmatica e anche la barriera ematoencefalica.

Il frammento 8H4 era già stato prodotto nel nostro laboratorio in *E.coli*, ma con rese molto basse.

In questo studio, i ceppi WB800N *B.subtilis* (deleto di otto proteasi extracellulari) e MS941 *B.megaterium* sono stati ingegnerizzati a produrre e secernere la proteina TAT-8H4.

Sono stati costruiti vari vettori di espressione-secrezione, clonando la sequenza naturale TAT 8H4 o l'analoga ottimizzata per *E.coli*, sotto il controllo di promotori differenti (IPTG inducibile Pgrac o costitutivo forte P43 in *B.subtilis*, oppure xilosio-inducibile T7 in *B.megaterium*) in fusione con il segnale di secrezione AmyQ. Benché i ceppi di *Bacillus* utilizzati fossero privi di parecchie proteasi extracellulari, non siamo riusciti ad ottenere chiaramente frammento anticorpale secreto

(né dopo purificazione IMAC del mezzo di coltura, né dopo precipitazione con ammonio solfato). L'anticopro era presente principalmente nella frazione intracellulare insolubile e in piccola parte nel periplasma (cap.II).

Poiché l'alto potenziale secretorio di *Bacillus* non appariva, abbiamo provato ad incrementare le rese e la solubilità del frammento anticorpale TAT 8H4 nel citoplasma di *E.coli* sfruttando la tecnologia delle proteine di fusione. Allo scopo sono stati scelti il chaperone batterico DnaK e la α-sinucleina (che sembra avere attività di chaperone), per la loro azione di facilitare il ripiegamento e per le loro caratteristiche chimico-fisiche.

Benché l'α-synucleina e il DnaK in fusione con TAT 8H4scFv ne incrementassero l'espressione a livelli significativi, queste proteine rimanevano largamente insolubili. Abbiamo quindi provato ad esprimerle in ceppi di *E.coli* ingegnerizzati per sovraesprimere i chaperoni batterici DnaKJE e GroELS, e a purificare α-synTAT8H4 dall'ambiente periplasmatico più ossidante. Poiché le rese di anticorpi di fusione in forma solubile rimanevano non significative, abbiamo purificato Dnak TAT8H4 e α-syn TAT8H4scFvs in condizioni denaturanti mediante cromatografia di affinità per ioni metallici (IMAC). Mediante processi di ripiegamento in vitro, è stato possibile ottenere in forma correttamente ripiegata solo α-syn TAT 8H4, così in grado di riconoscere la proteina prionica umana mediante immunoblotting. In cellule CHO trattate con Dnak TAT8H4 e α-syn TAT8H4, le proteine internalizzano in breve tempo, mostrando una localizzazione principalmente nucleare. Il frammento α-syn TAT8H4 è in grado di internalizzare gran parte della proteina prionica legata alla superficie cellulare, in cellule HELA transfettate con il plasmide GFP-PRP ma non con l'analogo GFP doppel. L'emivita della proteina α-syn TAT8H4 si è dimostrata superiore a quella del Dnak TAT8H4 (cap.III).

In conclusione la tecnologia delle proteine di fusione si è dimostrata efficace nell'incrementare le rese di produzione di TAT8H4scFv nel citoplasma di *E.coli*. In fusione con il frammento anticorpale, α–sinucleina si è dimostrata preferibile a DnaK per il minor peso molecolare, le maggiori rese di produzione e la sua azione di facilitare il ripiegamento.

Quindi per la sua specificità nell'eliminare in parte la proteina prionica, α–syn TAT8H4 scFv potrebbe essere efficace nel trattamento delle malattie spongiformi trasmissibili. Inoltre l'utilizzo di anticorpi permeabili alle cellule grazie al dominio di traslocazione TAT, eviterebbe le questioni etiche e di salute riguardanti l'applicazione della tecnologia del DNA ricombinante alla terapia clinica umana, e potrebbe essere esteso al trattamento di altre patologie.

La seconda parte di questo studio riguarda l'ingegnerizzazione di una via metabolica naturale, quale la biosintesi di acido ialuronico. Nel nostro progetto la via metabolica di sintesi dell'acido ialuronico, presente in produttori naturali come gli streptococchi, è stata adattata a *Bacillus*. I geni

coinvolti nella via metabolica biosintetica degli zuccheri precursori che in *Bacillus subtilis* sono dispersi nel genoma, sono stati associati in un unico mRNA policistronico assieme alla ialuronato sintasi da *Streptococcus equi*, come nella via naturale di produzione negli streptococchi. Abbiamo sviluppato parecchi vettori epitomali con i geni sotto il controllo di promotori forti e inducibili. Questo sistema di espressione, che utilizza plasmidi con un numero di copie relativamente alto, ha il vantaggio di esprimere livelli di mRNA maggiori rispetto ad un sistema integrativo sul cromosoma.

Mediante amplificazione PCR dei geni *hasA* e *tuaD* rispettivamente da *S.zooepidemicus* e *B.subtilis*, e loro clonaggio nel plasmide shuttle pHT *B.subtilis/E.coli* sotto il controllo del promotore Pgrac, abbiamo selezionato ceppi stabili di *Bacillus subtilis* 1012 e WB800N metabolicamente ingegnerizzati e secernenti, acido ialuronico con pesi molecolari maggiori di 800 kDa, e con rese molto superiori ai 5g/L, attualmente prodotti industrialmente da *Streptococcus*, e ai dati finora pubblicati.

Successivamente, per incrementare le rese e i pesi molecolari dell'acido ialuronico, abbiamo costruito e clonato le cassette-operone *hasA-tuaD* e *hasA-tuaD-gtaB-pgi* sotto il controllo del promotore inducibile T7 in *B. megaterium*.

Poiché il sistema di espressione T7 è presente anche in *E.coli*, abbiamo dimostrato che cellule ricombinanti di *E.coli* producono acido ialuronico in quantità molto basse. Invece i ceppi di *B. megaterium* ingegnerizzati, nelle migliori condizioni di espressione individuate, hanno rese di circa 2g/L in fiasca, che sono molto promettenti in vista di processi di fermentazione in bioreattori. Inoltre i ceppi di *B.megaterium* che sovraesprimono i geni *hasA-tuaD-gtaB-pgi* sembrano produrre acido ialuronico con peso molecolare maggiore, circa 1800 kDa, comparabile anche per polidispersione agli standard commerciali di *Streptococcus*, suggerendo che la sovraespressione di *gtaB-gpi* influisce sui pesi molecolari incrementandoli, con produzione di acido ialuronico a catena più lunga (cap.IV).

In conclusione se da una parte l'alto potenziale secretorio di *Bacillus* non si è manifestato nella secrezione di una proteina eterologa, quale il frammento anticorpale anti-PrP TAT8H4 scFv, *B. subtilis* and *B.megaterium* si sono rivelati ospiti di espressione superiori per l'ingenerizzazione di una via metabolica naturale, quale la biosintesi di acido ialuronico, sulla base di diversi criteri:1) buona qualità dell'acido ialuronico, comparabile per massa molecolare e polidispersione agli standard commerciali di Streptococcus e superiore per resa. 2) Inoltre, a differenza di *Streptococcus*, l'acido ialuronico prodotto da *B. subtilis* e *B.megaterium* non contiene tossine e non è associato alle cellule, ma secreto direttamente nel mezzo, semplificando i processi di purificazione. 3) Infine, mentre *Streptococcus* richiede mezzi complessi per la crescita, i ceppi di *Bacillus* crescono su terreni minimi, garantendo prodotti finali più puri ad esenti da tossine.

Chapter I

General introduction

1.1.Prion protein and prion disease

Prion protein is a normal product of a specific gene *Prnp*. The protein product of this gene (PrP^C) appears physiologically as an approximately 250 amino acid sialoglycoprotein. It is a highly conserved, ubiquitous protein that is glycosylphosphatidylinositol (GPI-) anchored on all cell membranes, but appears under normal circumstances to be principally located in neuronal and glial cells of the central nervous system (CNS). It has two N-glycosylation sites, and its C-terminal is folded largely into α-helices stabilised by a single disulphide bond and it is monomeric. The N-terminus contains an octapeptide repeat region with tight bonding sites for Cu^{2+} (Prusiner 1998; Harris 2003).

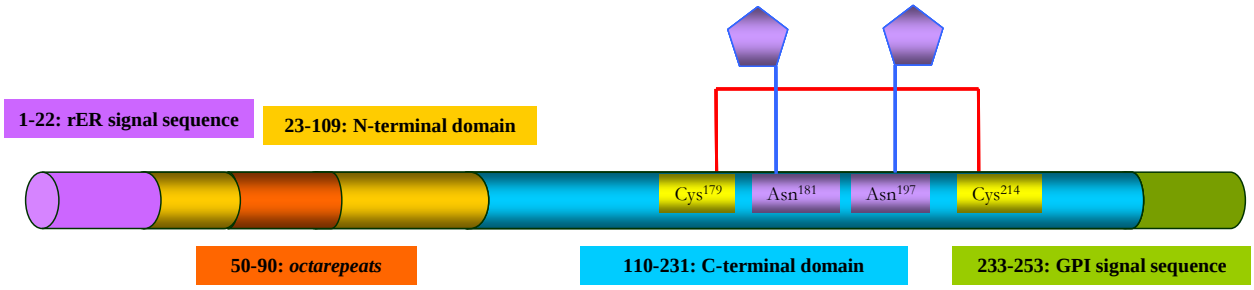


Fig. 1.1 Domain organization of the human PrP protein.

PrP^C is glycosilated in the endoplasmic reticulum after the removal of an N-terminal signal peptide. The GPI anchor is attached after cleavage of a second signal peptide at the C-terminus (Stahl et al, 1987). The protein is then trafficked via the Golgi to the cell surface (Shyng S.L., 1993), where it is internalized in clathrin-coated pits and endocytosed for recycling (Shyng S.L., 1994). Studies in cultured cells show that it is degraded with a half-life of approximately 6 hours (Taraboulos et al., 1992).

The functional role of PrP^C remains poorly understood but is associated with central nervous system development, synaptic transmission, copper homeostasis, neuroprotection, regulation of circadian rhythms, and memory and cognitive functions (Hu W., 2008). The conversion of endogenous cellular form of the prion protein (PrP^C) to so-called scrapie isoform of prion protein (PrP^{SC}) is a common feature of all prion diseases (Prusiner S.B., 1998). Prion diseases or transmissible spongiform encephalopathies (TSE), are progressive, fatal neurodegenerative disorders that affect humans and other animals.

In humans, TSE are classified into: (i) sporadic forms such as Creutzfeldt-Jakob disease (CJD) and fatal insomnia; (ii) inheritable forms such as Gerstmann-Sträussler-Scheinker disease, fatal familial insomnia and a small percentage of CJD of familial origin; and (iii) iatrogenic forms such as kuru and the new variant of CJD (vCJD) (Collinge J., 2001). In animals, TSE was

traditionally recognised in sheep and goats as scrapie, in mink as transmissible mink encephalopathy, in deer as chronic wasting disease, and in cattle as bovine spongiform encephalopathy (BSE). This latter disease, which is also known as 'mad cow' disease, has had dramatic consequences for public health in the UK and other European countries (Caramelli et al., 2006).

TSE are characterised by spongiform changes in the CNS (because of the vacuolation and cavitation give the "sponge-like" histologic appearance of the infected brain tissue), neuronal loss, astrocyte proliferation and microglial activation, due to accumulation of scrapie aggregates in the brain.

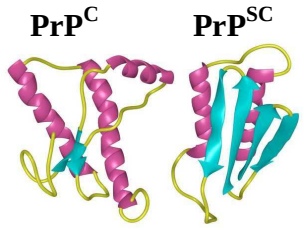


Fig.1.2 Conversion of PrP^{C} to PrP^{SC}

The prion disease related isoform (PrP^{SC}) is derived from by a post-translational process that involves conformational change and aggregation. PrP^{C} is rich in α-helical structure, whereas PrP^{SC}, which is found as insoluble aggregates, seem to be predominantly composed of β-sheet structure. In contrast to PrP^{C} ,PrP^{SC} has other important features including resistance to proteinases, radiation, and detergents that have particular significance for transmission and expression of disease.

The mechanisms involved in their pathogenesis are not completely understood, but they are associated with the accumulation in the brain of misfolded form, PrP^{SC} (Prusiner S.B., 1998).

Direct interaction between the pathogenic PrP^{SC} template and the endogenous PrP^{C} substrate and the following conformational conversion of PrP^{C} into PrP^{SC} are proposed to exert a neurotoxic effect via aberrant signalling cascades (Harris D.A, et al. 2006; Aguzzi A., 2005).

1.2. Therapeutic strategies

Various lines of evidence indicate that the process of conversion of PrP^{C} into PrP^{SC} is the key element in prion pathogenesis. Knockout mice (*Prnp* 0/0) are resistant to prion diseases and do not propagate infectivity (Bueler et al., 1993), and brain tissue homogenates from *Prnp* 0/0 cows are resistant to prion propagation in vitro (Richt et al., 2007). Moreover, *Prnp* 0/0 brain tissue surrounding prion-infected *Prnp*+/+ neurografts does not develop neuropathological changes typical of prion diseases (Brandner et al., 1996).

Despite the uncertainties about the possible mechanisms of neurotoxicity, prevention of this conversion in neurons can prevent disease progression and reverse early degenerative changes. Depleting neuronal PrP in prion infections prevents disease and reverse spongiosis (Mallucci et al., 2003).

Reagents specifically binding either prion-protein conformer may interrupt prion production by inhibiting this interaction. Over the past several years, many anti-prion compounds have been identified in vitro models of prion replication, such as polysulphated polyanionic compounds, polyamine, tetrapyrroles, polyene antibiotics, peptides, tetracyclic and tricyclic compounds.

These prion antagonists can be targeted towards the selective binding of PrP^{C} and/or PrP^{SC} to the process of conversion (Weissmann and Aguzzi, 2005; Ludewigs et al., 2007; Trevitt et al., 2006). However, most of these molecules were found to be toxic or ineffective in animal models of prion diseases.

Alternatively, anti-PrP antibodies have been shown to have an anti-prion effect in cellular and animal models (Peretz et al., 2001; Heppner et al., 2004; Enari et al., 2001; Sigurdsson et al., 2003). The use of antibodies that bind to and stabilize or sequester PrP^{C} (which make it less available for conversion) is a possible strategy for preventing conversion. Early indications of the potential of antibody therapy for prion disease came from in vitro studies showing a reduction in infectivity of prions after incubation with an anti-PrP antibody (Gabizon et al., 1988).

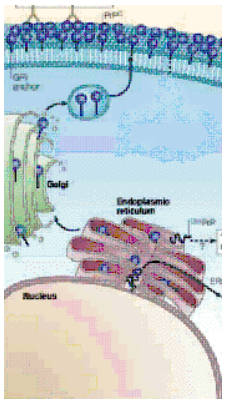


Fig.1.3 PrP^{C} cell trafficking (adapted from Mallucci &Collinge, 2005).

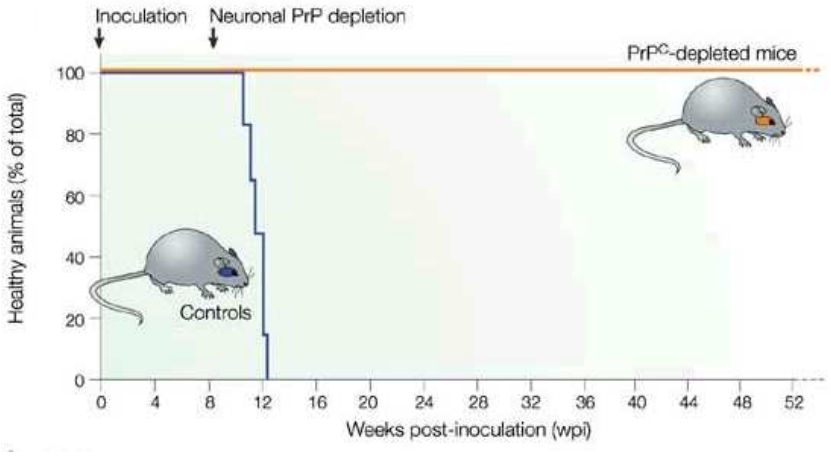


Fig.1.4 Survival of prion-infected mice after PrP^{C} depletion (adapted from Mallucci &Collinge, 2005).

It is showed a significant delay in the onset of clinical symptoms in scrapie-infected mice using monoclonal 8H4 anti-PrP antibodies directed against 175-185 residues of PrP inoculated intraperitoneally following exposure. Immunization with PrP peptides reduced PrP^{SC} in a peripheral model of mouse scrapie (Souan, 2001). Different publications report that antibodies directed against the middle portion of PrP (residues 91–110 and 132–156) can cure scrapie-infected cells in culture and reduce PK (proteinase K)-resistant PrP in spleens of infected mice (Peretz et al., 2001; White et al., 2003).

In cell systems, anti-prion antibodies were able to purge the cells entirely of PrP^{SC} (Peretz et al., 2001), and their potency correlated with their ability to recognize the total population of PrP^{C}

molecules on the cell surface. Despite these encouraging results, development of effective immunotherapy presents several problems in both active and passive approaches. An important obstacle in the development of efficacious regimens for active immunization is host tolerance to endogenous PrP^C, which limited the therapeutic efficacy of this immunization approach (Sigurdsson et al., 2003; Polymenidou et al.,2004). In comparison, passive immunization suffers from the intrinsic problem of poor antibody diffusion from vessels into tissues, especially in the nervous tissue: administration of monoclonal antibodies has been shown to prevent the pathogenesis only when applied simultaneously, or shortly after, peripheral prion infection (White et al., 2003). As antibody do not readily cross the blood-brain-barrier (BBB), there was no protective effect in the intracerebrally infected mice. The major problem in using anti-prion molecules in vivo is that peptide and protein therapeutics are generally excluded from transport to the brain, owing to the negligible permeability of these drugs to the brain capillary endothelial wall, which makes up the BBB (blood-brain barrier) in vivo. Moreover, production of large amounts of monoclonal antibodies for therapy is technically challenging and expensive.

1.3. Anti-PrP single-chain variable fragments

One promising solution to this problem is offered by the use of single-chain variable fragments (scFv). ScFv are recombinant antibody (rAb) fragments becoming popular therapeutic alternatives to full length monoclonal antibodies (Abs) since they are smaller, possess different properties that are advantageous in certain medical applications, can be produced more economically and are easily amendable to genetic manipulation. ScFv Abs are one of the most popular rAb format and are easily expressed by several expression systems.

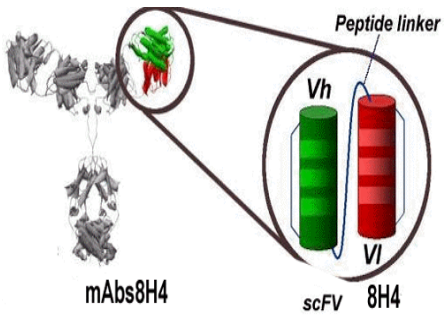


Fig. 1.5 Representation of scFv engineering.

scFvs Abs (26–28 kDa) are monovalent mini-antibodies containing the complete antigen binding site. They constituted by a single fusion polypeptide comprising the variable region of the light and heavy chain (V_H) (V_L) of an Ab, linked by an introduced flexible polypeptide linker (Maynard and Georgiou, 2000). The first scFv molecules were developed independently by Huston et al. (1988) and Bird et al.(1988) and represent the smallest functional V_L V_H domains of an Ab necessary for high-affinity binding of antigen.

scFv fragments retain the binding properties of their parent Abs, they maintain antigen specificity and can be engineered for intracellular expression or secretion. So far, several groups have

investigated the use of scFv in prion-infected cell culture systems. One group showed the paracrine inhibition of prion replication by RD-4 cells expressing and secreting scFv(6H4) on co-cultured ScN2a cells (Donofrio G. et al., 2005).

Another interesting study showed the retention of PrP^C in the ER (endoplasmic reticulum) of HEK (human embryonic kidney)-293 and PC12 cells after the expression of a scFv from antibody 8H4 and 8F9, containing the ER retention sequence KDEL. 8H4 intracellular expression causes a marked impairment of prion maturation and translocation towards the membrane compartment, with a strong reduction of the PrP^C membrane fraction. As a consequence, the formation and accumulation of the pathogenic scrapie isoform are blocked in infected cells (Cardinale A. et al., 2005).

Starting from these results, a subsequent set of in vivo analyses was carried out showing that mice intracerebrally injected with KDEL-8H4-NGF-differentiated PC12 cells infected with scrapie did not develop scrapie clinical signs or show any brain damage (Vetrugno V. et al,2005). Moreover, scFvs directed against the LRP/LR (37/67 kDa laminin receptor) have been recently used as therapeutic approach in prion diseases (Zuber, C. et al., 2008 a).

Delivery of these scFv both by passive immunotransfer refand by AAV (adeno-associated virus)-mediated gene transfer (Zuber, C. et al., 2008 b) resulted in significant reduction of the peripheral PrP Sc propagation. Therefore, these results support the use of scFv as a therapeutic approach.

1.4. Recombinant antibody expression systems

Diverse prokaryotic and eukaryotic expression systems have been developed for rAb expression. These have included bacterial (Martin C. et al., 2006), yeast and filamentous fungus ref, eukaryotic alga, insect cell (Bruenke J. et al., 2004), plant (Makvandi-Nejad S. et al., 2005), mammalian cell and transgenic animal systems (Natsume A. et al., 2006).

For therapeutic purposes, large doses of Ab are required, and in some cases exceed a gram per patient per year. Thus, there is a need to develop production systems to make these molecules efficiently and cost effectively (Andersen and Reilly, 2004).

Using a single production platform for most, if not all, proteins would be ideal for the Ab production industry; however, protein expression is variable and yields vary for each rAb fragment and/or fragment type (Weisser N. and Hall J.C., 2009). Expressed rAb fragments differ in their yield and activity because of several factors, such as protein size, solubility, stability and amino acid sequence. Thus, more than one production platform is usually required and expression of each protein must be optimized. The optimal expression system

depends on the type of rAb fragments being expressed as well as the required purity and quantity of the final product.

For example, a requirement for certain therapeutic IgGs to have appropriate glycosylation necessitates their expression in mammalian cells. Also, due to their complex structure, whole antibodies have been preferably expressed in eukaryotic systems which have the appropriate cellular machinery for efficient folding and assembly. Due to much simpler structure of Fab and, scFvs and no requirement of glycosylation, bacterial expression, and almost exclusively *E.coli* expression, has been the method of choice for expression of these molecules.

1.5. Bacterial expression *E.coli* expression system

Bacterial expression systems for heterologous protein production remain most attractive due to low cost, high productivity, and rapid use.

The Gram-negative bacterium *E.coli* is still the most common workhorse for recombinant proteins production, and its importance is also recognized for biopharmaceutical production, which tipically has several options in selecting the expression system (Schmidt F.R., 2004). This is because of it is able to grow rapidly and at high density in large-scale on inexpensive substrates, thus simple conditions and low capital costs for fermentation. Moreover it is the best-characterized host with ease of genetic manipulation and many available expression systems. A large pool of cloning and expression plasmids, strains and purification systems is presently available for *E.coli*. Finally no concerns about viruses that are harmful to humans and relatively for fermentation (Lee et al.1996; Makrides et al. 1996).

Therefore, most protein expression strategies in microbiologic research focus on this organism. However, aside from the obvious advantages of *E.coli* systems there are limits, because of *E. coli* cannot produce mammalian proteins that require post-translational modification for activity as glycosylation or some multi-domain molecules containing complex disulfide bonds such as IgGs . Moreover, serious problems can occur during the process of heterologous, particularly eukaryotic, gene expression and purification: (1) low expression rates, because of yields are dependent on the individual Ab fragment or protein (2) formation of inclusion bodies from which soluble and biologically active proteins can only be recovered by complicated and costly denaturation and refolding processes, with final yields usually very low (3) improper protein-folding, (4) toxicity problems depending on protein sequence and (5) the possibility of bacterial endotoxin contamination of purified products (Baneyx, et al, 1999).

Although eukaryotic expression hosts are sometimes able to overcome these problems, they are not without their own difficulties in terms of ease of use, time, cost and experimental flexibility.

So in order to enhancing the production of soluble proteins, in particulary scFvs, in more economical and easy of use and scale bacterial systems, significant researches into field has been done and a variety of techniques have been developed.

E.coli is the bacterial production system of choice for small non-glycosylated rAb fragments, including scFvs (Wang et al., 2008). Compared to mammalian cell lines, the generation of Ab-producing *E.coli* cell lines is faster and easier and thus allows for the easy progression from Ab selection to large-scale manufacturing (Andersen and Reilly, 2004). Moreover, large rAb concentrations can be attained in *E.coli*; expression levels of up to 2 g/L have been reported (Chen C. et al., 2004).

Because of a scFv molecule requires generally two disulfide bridges for correct folding and structural stability, indispensable for its biological activity two basic strategies have been applied to express various formats of antibody fragments, including Fabs, scFvs and sdAbs, in *E. coli*. The two approaches involve directing the antibody product to either the reducing environment of the cytoplasm or the oxidizing environment of the periplasmic space between the cytoplasmic and outer membranes or the culture medium.

- Expression in the reducing environment of the cytoplasm can be achieved at high concentrations but often results in the formation of inclusion, bodies due to their foreign nature, high expression rate and lack of disulphide bonds (i.e. reduced and unfolded Abs), that require solubilisation with denaturing agents (e.g. 8 M urea) and subsequent *in vitro* refolding, to convert to active species. However, correct in vitro refolding and purification of functional product is a complex and time-consuming process, requiring expertise and involving many steps. Moreover, refolding efficiency is highly variable depending on the specific antibody fragment with yields varying from 10–40% for Fab and Fv fragments.

- As an alternative to Ab recovery from inclusion bodies, cytoplamic expression can be done in *E. coli* strains that promote proper folding and oxidation in vivo, as *E. coli* cells with mutations in both the thioredoxin reductase (trxB) and glutathione reductase (gor) genes (Jurado P. et al., 2002)

The cytoplasm of the *E.coli* trxB/gor double mutant strains Origami and FA113 provides an oxidizing environment, while the *E.coli* cytoplasm is maintained in a reduced state under physiological conditions. Venturi et al.(2002) and Jurado et al. (2002) showed that functional antibody Fab and scFv fragments were produced efficiently in the cytoplasm of the mutant *E.coli* cells, respectively. Using such strains, scFv have been produced at levels equal to or greater than those achieved with periplasmic expression (Levy et al., 2001).

-Moreover, cytoplasmic expression using fusion partners such as glutathione S-transferase (GST), thioredoxin (Trx), and maltose-binding protein (MBP) generally results in high

productivity and enhanced solubility (Bach H. et al., 2001). It was previously reported that MBP-fused scFvs are expressed at high levels in the cytoplasm of *E.coli* as soluble and active proteins (Bach H. et al., 2001). Zheng et al.(2003) showed that a fusion method with *E.coli* protein NusA, which had been identified to have a highest solubility when over-expressed in the cytoplasm of *E.coli* cells, has potential as a valuable tool for overproduction of soluble and functional scFvs (Jurado P. et al., 2006).

- Another approach used for cytoplasmic expression of soluble heterologous proteins is co-expression of molecular chaperones in the cytoplasm. Although there are several reports that co-expression of molecular chaperones, such as GroES/GroEL (GroELS), DnaKJ/GrpE (DnaKJE), trigger factor (TF), DsbC, and Skp, with Ab fragments improves their production (Heo M.A. et al., 2006; Levy R. et al., 2001; Hu X. et al., 2007). Unfortunately, determination of a successful match between the target protein and the chaperone is a trial-and-error process (Georgiou G. et al., 1996; Hannig G. et al.,1998).

- Alternatively, scFv expression can be directed to either the oxidizing environment of the periplasmic space, where disulfide bonds in proteins are stable refunder regulation of the pelB leader sequence (Padiolleau-Lefèvre et al., 2006), or to the culture supernatant using the α-hemolysin (HylA) system of *E.coli* (Fernandez et al., 2000; Fraile et al., 2004). The HylA system is a type I secretory apparatus that forms a protein channel between the inner and outer membranes of *E.coli* through which the hemolysin toxin is secreted; this system has proved competent in the secretion of heterologous protein hybrids including scFv (Fernandez et al., 2000).

In some cases, periplasmic expression results in the rAbs leaking through the outer membrane and into the culture supernatant. When this does not occur, the rAb can be easily obtained from a periplamic extract via osmotic lysis.

Both periplamic and culture supernatant expressions are advantageous in that a high concentration of reasonably pure Ab can be easily obtained. In secretion, proteins are released directly into the growth medium, which greatly simplifies and reduces the costs of downstream purification steps. In addition, in case of efficient secretion, the formation of inclusion bodies in the cytoplasm is decreased, leading to higher amounts of properly folded and active enzymes.

However, production rate and efficiency of secretion depends a great deal on the type of target protein and individual sequence of the antibody, and secretion of a target protein into the oxidizing periplasm generally results in low yields. Moreover, Gram-negative bacteria like *E. coli* secrete proteins mostly into the periplasm (Ward et al., 1993). Only in rare cases can the scFv be isolated in higher amounts from the supernatant (Lauer et al., 2005).

The use of a Gram-positive bacterium could facilitate the scFv production due to the lack of an outer membrane allowing direct secretion of proteins into the growth medium. Among the

Gram-positive bacteria *Bacillus brevis* (Inoue et al., 1997; Shiroza et al., 2003) *Bacillus subtilis* (Wu et al., 1998; Wu et al. 2002) and *Bacillus megaterium* (Jordan et al., 2007). Thus an alternative expression host also for secretion of foreign proteins could be the Gram-positive *Bacillus* strains.

1.6. *Bacillus* strains

Bacillus species are Gram-positive, ubiquitous rod-shaped bacteria; they can be either obligate or facultative aerobes and show positive reaction in the catalase test. Members of the genus *Bacillus* are known to form spores under stressful conditions. These endospores are highly resistant to heat and radiation and are viable for extremely long periods.

1.6.1. *Bacillus subtilis*

Bacillus subtilis is an ubiquitous bacterium that exhibits a number of lifestyles, each characterized by a distinguishing pattern of gene expression. This ‘‘simple’’ organism resides in the region of soil surrounding plant roots, called the rhizophere, where it forms multicellular, organized biofilms, and secretes numerous enzymes to degrade a variety of substrates, enabling the bacterium to survive in a continuosly chancing environment. It exhibits a number of additional lifestyle adaptations as sporulation, competence, motile-sessile switch behaviour (Earl et al. 2008).

B. subtilis is one of the best known Gram-positive bacteria at both the genetic and physiological level, thus it is often used in molecular biology and as a general model organism. Its harmless nature, brilliant genetic amenability and relatively large gene size make the organism a highly valuable tool for science and demonstration purposes. It is been used to help demonstrate biochemical differentiation, gene/protein regulation and cell cycle events in bacteria.

The entire sequence of its chromosome is known (1997) and efficient tools for the genetic modification of this bacterium are available.

B.subtilis and related *Bacillus* species are well known for their contributions to agricultural, medical and food biotechnology and for large-scale production of recombinant proteins, in particular for the production of secreted enzymes (Schallmey M. et al., 2004).

Besides the well-known reputation in the industrial production of enzymes, such as proteases and amylases, some *Bacillus* species, with particular emphasis on *B.subtilis*, have been explored as a

Fig. 1.6 Phylogenetic tree of some members of the genus Bacillus based on 165 rRNA sequence analysis (Ash et al., 1991). Modified from Priest (1993). Kingdom: Bacteria; Division: Firmicutes; Class: Bacilli; Order: Bacillales; Family: Bacillaceae; Genus: *Bacillus.*

host for the expression of foreign proteins with pharmacological or immunological activities (Harwood 1992, Wong 1995). Such interest stems from a plethora of very solid reasons:

(1) compared with *E.coli*, the Gram-positive bacterium *B.subtilis* offers an efficient secretion apparatus that guides the expressed protein directly into the culture supernatant. *Bacillus* strains do not have an outer membrane and, thus, all secreted proteins are released directly into the growth

medium, which greatly simplifies and reduces the costs of downstream purification steps bypassing the time-consuming cell disruption (that makes subsequent protein purification much easier); (at present, about 60% of the commercially available enzymes are produced by *Bacillus* species). In addition, in case of efficient secretion, the formation of inclusion bodies in the cytoplasm is decreased, leading to higher amounts of properly folded and active enzymes (van Wely et al., 2001).

(2) the available knowledge on genetics and physiology of *B.subtilis* finds parallel only with *Escherichia coli* K12, making easier the development of controllable gene expression systems and adaptation to large-scale stream-line fermentation processes, Actually a large body of information concerning transcription, translation, protein folding and secretion mechanisms, genetic manipulation and large-scale fermentation has been acquired.

(3) *B.subtilis* strains are non-pathogenic and have a well-established safety record and have deserved the GRAS (generally regarded as safe) status, being organism free of any endotoxin (Westers L. et al., 2004)

(4) production of spores, the most resistant life form found on earth, warrants easy preservation of strains even under harsh environmental conditions;

(5) the ability to grow in simple and non-expensive media at fast growth rates and no significant bias in codon usage confer to this bacterial species a top candidate position for the expression of heterologous proteins, including those with potential use in vaccine development (Henner et al. 1990, Wong et al., 1995).

But there are also two obstacles reducing the use of *B subtilis*: (i) production of a number of extracellular proteases which recognize and degrade heterologous proteins, and (ii) stable vector plasmids. The first obstacle has been largely solved by the construction of protease-deficient strains. And the second has been completely overcome by introducing plasmids using the theta-mode of replication such as those derived from the natural plasmids pAME1 and pBS72 (Jannière et al., 1990; Titok et al., 2003).

1.6.2. *Bacillus megaterium*

1.6.2.1. General features

First described over 100 years ago, *B.megaterium* has recently been gaining more and more importance in scientific as well as industrial applications.

The source of the significant name "megaterium" was the large size of the vegetative cells (4 x 1.5 μm) and the spores. Compared to the model organism for Gram-negative bacteria, *Escherichia coli*, it has an up to 100-times higher volume (Fig. 1.7).

This large size has been exploited in several morphological studies. It has been used effectively to study cell wall synthesis as well as membrane and spore structure (Archibald et al., 1993). Moreover, the capability of sporulation has made *B. megaterium* an important tool for examining spore-mediated disease and cell development.

In the taxonomy of *Bacillus*, as can be seen in Fig. 1.6, it is within the *B.subtilis* group, but much more distantly related to *B.subtilis* than *B.licheniformis*, *B.cereus*, *B.anthracis* or *B. pumilus* by 16s rRNA sequence analysis (Ash et al., 1991 ; Priest, 1993).

It is interesting especially because of its physiology, unusual and useful enzymes and products, and wide range of ecological habitats. Although it is generally considered a soil organism, *B.megaterium* is able to grow on a wide variety of carbon sources and thus has been

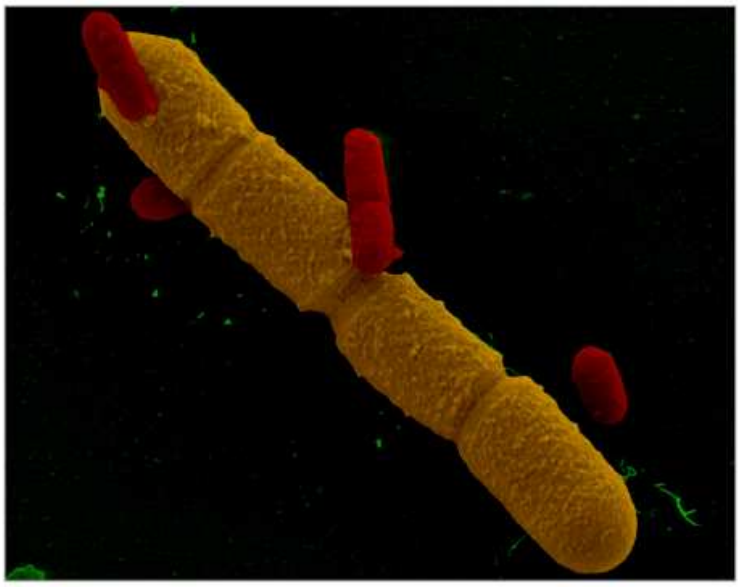


Figure 1.7 Electron microscope image of *Bacillus megaterium* (yellow) and *Escherichia coli* (red) vegetative cells. *B. megaterium* and *E. coli* cells were aerobically cultivated separately in LB medium at 37°C. They were grown until reaching the stationary phase. Compared to the *E. coli* volume of 0.5 μm^3 (0.5 x 0.5 x 2), *B. megaterium* has at least up to 100-times higher volume.(From Bunk et al., 2010).

found in diverse environments from rice paddies to dried food, seawater, sediments, fish, normal flora, and even in bee honey. Strains are often isolated on unusual substrates, such as herbicides, waste from meat industry or petrochemical effluents, in the company of pseudomonas and actinomicetes. Also documented has been the ability to degrade persistent insecticides and utilize them as carbon sources by *B. megaterium* (Sexana et al., 1987) offering potential applications as detoxifying agent.

Several *B.megaterium* proteins are of importance. A family of P450 cytochrome monooxygenases e.g. is similar to eukaryotic P450 playing a role in many diseases. Industrial applications of enzymes excreted by *B.megaterium* are diverse, starting from amylases used in bread industry to penicillin amidase, which is used for the generation of new synthetic antibiotics to steroid hydrolases, glucose L-valine and L-alanine dehydrogenases to synthesize valine and

alanine. It is the major aerobic producer of vitamin B12, and is one of the organisms involved in fish spoilage (Vary P.S.,1994).

Product/use	Comments	References
Alphostatin production	Inhibitor of alkaline phosphatase	Aoyagi *et al.* (1989)
α-Amylases	Can replace pullulanases	Takasaki (1989); Vinhinen & Matsala (1989)
β-Amylases	Bread industry	Metz *et al.* (1988); Hebeda *et al.* (1988)
Chitosanases	Yeast cell wall analysis	Pelletier & Sygusch (1990)
Fungicidal toxins	For *Rhizoctonia*, one isolated	Liu & Sinclair (1992); Bhattacharyya & Pukayastha (1989)
Glucose dehydrogenase	Generator of NADH, immobilization, biosensors	Nagao *et al.* (1992); Kittsteiner-Eberle *et al.* (1989)
Glutamate production		1987 Patent JP62048393
Modification of steroids	4-α-Glucanotransferase A	1989 Patent DD266592
Oxetanocin production	Inhibits HIV, hepatitis B, cytomegalovirus, herpes	Kohlbrenner *et al.* (1990); Tseng *et al.* (1992)
Phosphate solubilization	Phosphate fertilizer	Vary (1992)
Penicillin amidase	Construction of synthetic penicillins	Suga *et al.* (1990)
Sensitivity testing	Heat, sterilization, antimicrobials	Vary (1992)
Toxic waste cleanup	Herbicides, C–P bonds	Quinn *et al.* (1989); Saxena *et al.* (1987); Selvanayagam & Vijaya (1989)
Vitamin B_{12} production	Only aerobic producer	Wolf & Brey (1986); Robin *et al.* (1990)
Expression host	Secretes, processes, produces intact proteins	Ahn *et al.* (1993); Rygus & Hillen (1991); Von Tersch & Robbins (1990); Shivikumar *et al.* (1989)

Table 1.1 Examples of recent industrial uses of *B.megaterium*. (From Vary P., 1994).

During the 1980s, genetic techniques of transduction, plasmid transformation, protoplast fusion and transposition became developed enough in *B.megaterium* to apply them to the study of many of its metabolic and developmental functions. Moreover, it is increasingly used as a host to produce foreign genes since it has been found to express, secrete and process foreign proteins without degradation. In fact, in contrast to *B. subtilis*, *B.megaterium* does not produce alkaline proteases.

B.megaterium, like most species of *Bacillus* (including most industrial strains) cannot be naturally transformed, but a poly-ethylene-glycol-mediated protoplast transformation has been developed (Von Tersch and Carlton, 1983; Vorobjeva et al., l980), and many recombinant plasmids have been introduced (Von Tersch and Robbins, 1990). All the Gram-positive vectors function well in *B.megaterium*, and there is excellent segregational and structural stability of both natural and recombinant plasmids in some strains QM B1551,216 and DSM 319 (Kieselburg et al., 1984; Meinhardt et al., 1989; Von Tersch and Robbins, 1990). Such protoplast transformation has facilitated the use of transposons to develop the genetics of *B.megaterium* and for its emergence as an effective cloning host for foreign DNA. Transposons and plasmids lacking Gram-positive origins of replication have been successfully integrated in *B. megaterium.* Major research strains of B. megaterium include Q M B1551 , KM, 216, DSM

319, ATTC 10778 and ATTC 19213. Strains Q M B1551, 216 and IWG3, as well as the plasmidless strains PV361, DSM 319 and VT1660 are well-known industrially.

Moreover, surveys involving a total of 18 strains of *B.megaterium* showed that almost all strains had at least four and up to seven plasmids, which stably carry a percentage of its cellular DNA as plasmid DNA, coding also for several megacins, or bacteriocins (Rostas et al., 1980; Von Tersch and Carlton, 1984).

1.6.2.2. *Bacillus megaterium* as expression host

In molecular biology, *B.megaterium* has proven to be an excellent host for the expression of non-homologous DNA. In 1991, Rygus and Hillen pushed *B.megaterium* research with the introduction of a strong and xylose-inducible promoter found in the genome of *B.megaterium*. Under the strict control of the repressor protein XylR, the system was used for the controllable production of intracellular recombinant proteins like β-galactosidase, glucose dehydrogenase, formate dehydrogenase and toxin A (Bäumchen C et al, 2007; Burger S. et al., 2003; Rygus T. et al. 1991a; Rygus T. et al. 1991b). This promoter provided the basis for the effective and commercialized xylose-inducible vector system for *B.megaterium*: "*…one of the most efficient expression systems described in any organism so far*" (MoBiTec, Göttingen, Germany).

In contrast to other bacilli strains *B.megaterium* has the advantage, that none of the alkaline proteases are present. This fact enables an excellent cloning and expression of foreign proteins without degradation (Meinhardt et al., 1989; Rygus and Hillen, 1991). In addition, there are no endotoxins found in the cell wall. Furthermore *B.megaterium* is able to stably maintain several extra-chromosomal DNA elements in parallel. This high stability of plasmids during growth (Vary P.S., 1994) , is very advantageous allowing a stable gene expression in long term cultivations and bioreactors and methods for large scale high cell density bioreactor cultivation were established (Hollmann et al., 2004).

1.7. *Bacillus* secretion system

The transport of proteins from their site of synthesis in the cytoplasm to their functional location is an essential characteristic of all living cells.

In Gram-positive bacteria, whose *B.subtilis* has become the model system, proteins can at least be delivered to, or retained at, five (sub)cellular locations: the cytoplasm, the cytoplasmic membrane, the membrane/cell wall interface, the cell wall, and the growth medium. The final destination of a protein is governed by the presence or absence of signal peptides and/or retention signals.

At least five distinct pathways for protein transport have been identified in *B.subtilis*. The majority of ~300 potentially secretory proteins appear to be translocated by the ‘‘Sec’’ pathway (Fig. 1) for protein *sec*retion (H. Antelmann et al., 2001; J.D.H. Jongbloed et al.,2000; J.D.H. Jongbloed et al., 2002). Typical proteins of this type include degradative enzymes (e.g. carbohydrases, DNAses, lipases, phosphatases, proteases and RNAses), proteins involved in cell wall biogenesis, substrate binding proteins, and even pheromones involved in sensing the cell population density for onset of developmental processes such as natural competence and sporulation. (However, the precise function of a large proportion of proteins that were shown or predicted to follow the Sec pathway remains to be determined.)

Other pathways for protein transport, such as the twin-arginine translocation ‘‘Tat’’ pathway, a pseudopilin export (Com) pathway involved in natural competence development, phage-like holins, and certain ATP-binding cassette (ABC) transporters, are ‘‘special-purpose’’ transporters, limited to the export of a small number of specific proteins (Sarvas M. et al., 2004).

1.7.1. SEC-dependent protein transport

Most bacterial proteins destined to leave the cytoplasm are exported via the highly conserved SecA-YEG (Sec) pathway. Protein secretion via the Sec pathway in *B.subtilis* can be divided into three functional stages: targeting, translocation, and folding and release.

The exported proteins are synthesized as precursors with an N-terminal signal peptide (von Heijne, G. 1990; von Heijne, G. 1998). These preproteins are first recognized by soluble targeting factors (as cytoplasmic chaperones, such as SRP/FtsY and CsaA) (Tjalsma H. et al., 2000; Tjalsma, H. Et al., 1998) that keep the precursors in a translocation competent state (an essentially unfolded form an unfolded or loosely folded staterefand facilitate their targeting to the translocation machinery SecA-YEG in the cell membrane.

Next, the polypeptide chain is transported through a proteinacious channel-pore SecYEG in the membrane, a process driven by the translocation motor Sec that binds and hydrolyzes nucleotide triphosphates. During or shortly after translocation, a signal peptidase at the surface of the cell removes the signal peptide by cleavage of preprotein at its recognitions sequence, releasing the mature protein into the medium.

The mature protein in the membrane–cell wall interface must then be folded into their native configuration in an environment that is dominated by a high density of immobilised negative charge—in essence an ion exchange resin. The folding depends on the activities of PrsA (Kontinen, V. P and Sarvas M., 1993), BdbBCD and/or SpoIIIJ/YqjG (Stephenson K. 1998). It is essential to the viability of the cell that these proteins do not block the translocation machinery in the membrane, form illegitimate interactions with the cell wall or, through intermolecular

interactions, form insoluble aggregates particularly if the cell is subjected to certain types of stress. Consequently Gram- positive bacteria such as *Bacillus subtilis* encode membrane- and cell wall-associated proteases that act as a quality control machine, as HtrA and HtrB , as well as WprA (Tjalsma H. et al.,2000) involved in the quality control of secretory proteins clearing misfolded or otherwise aberrant proteins from the translocase and the cell wall. Importantly, HtrA and HtrB have the potential to assist in the folding or, if folding is impossible, degradation of malfolded secretory proteins (Noone, D. et al., 2001). A model for the function of these main components of the Sec machinery of *B subtilis* is depicted in fig.1.8.

Additionally, post-translocational protein folding is of major biotechnological importance in view of the widespread use of *Bacillus* species for the biomanufacture of proteins and enzymes.

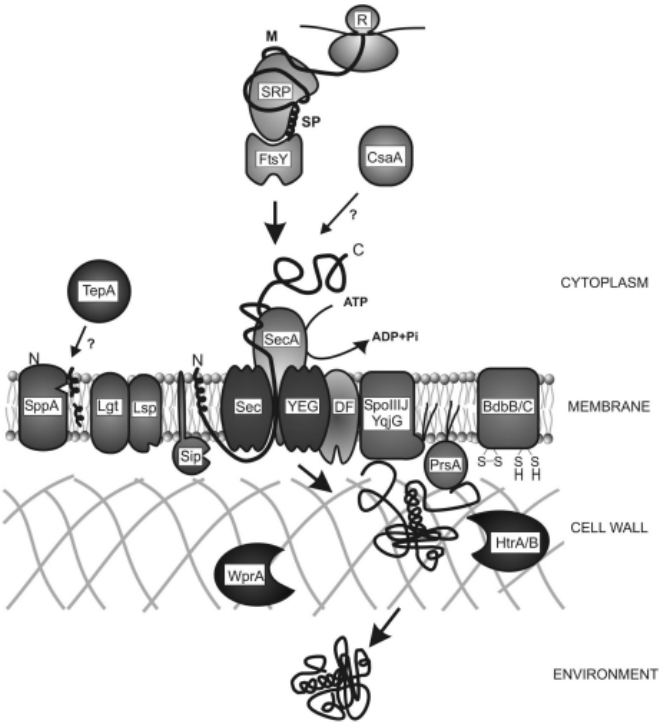


Fig. 1.8 Components involved in Sec-dependent protein export in *B.subtilis*. Secretory proteins are ribosomally synthesized as precursor proteins with an N-terminal signal peptide (SP). Cytoplasmic chaperones, such as SRP/FtsY and CsaA , keep the precursors in a translocation-competent state and facilitate their targeting to the translocase in the membrane, consisting of SecA, SecY, SecE, SecG, and SecDF. During or shortly after translocation, the preprotein is cleaved by one of the type I signal peptidases (SipS-W) or lipid modified by the diacylglyceryl-transferase (Lgt) and cleaved by the lipoprotein-specific signal peptidase (Lsp). SppA and TepA may be involved in the degradation of cleaved signal peptides (16), whereas the folding of several secreted proteins depends on the activities of PrsA , BdbBC, and/or SpoIIIJ/YqjG. HtrA, HtrB, and WprA are involved in the quality control of secretory proteins. It should be noted that for reasons of simplicity, HtrAB are depicted in the cell wall, although HtrA is detected in both the membrane and the medium . On passage through the cell wall, the mature protein is released into the environment. (Adapted from Tjalsma H. et al, 2004).

Chapter II

TAT 8H4 scFv secretion into culture medium in *Bacillus subtilis* and *Bacillus megaterium*

2.1. TAT 8H4 scFv

B.subtilis and *B.megaterium* are here evaluated for their capability to produce and secrete recombinant antibody fragments, in particular 8H4 anti- prion protein scFv, an eukaryotic complex molecule with three disulphide bonds indispensable for its stable structure and activity. The ScFv 8H4 protein was just produced previously in our laboratory in *E.coli*, but with very low yields. The scFv 8H4 is a small polypeptide (28.3 KDa) constituted by a single fusion polypeptide comprising the variable region of the light and heavy chain (V_H) (V_L) of anti-PrP 8H4 mAbs that recognizes residues 175-185 of helix B of PrP. This antibody recognizes all described PrP^C species (full-length and truncated, unglycosilated and glycosilated forms) and is able to react with both PrP^C and PrP^{SC} isoforms (Zanusso et al.,1998). The 8H4-scFv maintains the same immunogenic properties of the original mAbs 8H4 Moreover, both, mAbs 8H4 and 8H4 scfv, have shown therapeutic effects in vitro and in vivo, inhibiting prion replication and delaying the development of prion disease (Cardinale et al., 2005).

In order to optimize this therapeutic perspective and address the 8H4 scFv to its targets in SCN, it is thought to clone 8H4 scFv in fusion with peptide-transduction domain of HIV-1 Tat protein, that penetrates directly and efficiently into cells translocating across the plasma membrane and also the BBB (blood-brain barrier) (Frankel A. D. and Pabo C. O. (1988); Dietz et al.,2004; S.R. Schwarze et al.,1999).

Moreover this 'cell-permeable' antibody or 'transbody' possesses other advantages over conventional antibodies or intrabodies. For a start, 'correct' conformational folding and disulfide bond formation can take place prior to introduction into the target cell. More importantly, the use of cell-permeable antibodies would avoid the overwhelming safety and ethical concerns surrounding the direct application of recombinant DNA technology in human clinical therapy, which is required for intrabody expression within the cell by use of viral-based vectors.

Transbodies introduced into the cell would possess only a limited active half-life, without resulting in any permanent genetic alteration. This would allay any safety concerns with regards to their application in human clinical therapy (Boon Chin Heng and Tong Cao, 2005).

2.1.2. Cell penetrating peptides

Intracellular delivery of many therapeutic and diagnostic agents can be challenging as the plasma membrane forms a formidable barrier to the introduction of macromolecules into cells, being impermeable to most molecules greater than 500 Da. Entry into cells requires agents to transgress the lipid bilayer while being sufficiently hydrophilic to be easily formulated and distributed in aqueous media. The discovery of cell permeable peptides (CPPs) has made it possible to transduce

a broad range of physiologically and therapeutically active agents into living cells (Deshayes S.et al, 2005). Cell penetrating peptides (CPPs) or protein transduction domains (PTDs), also called permeation peptides are short (fewer than 30 residues), basic peptides, acting in a receptor- and energy-independent manner that penetrate into living cells. Efficiency of CPPs depends greatly on the cargo being transduced. The most frequently studied CPPs include HIV-1 Tat peptide (derived from transactivating protein Tat of human immunodeficiency virus-type 1, HIV-1) , HSV VP-22 peptide, penetratin (the third helix of the homeodomain of *antennapedia* transcription factorrefand simple synthetic polyarginine peptides. Fusion of CPPs to other proteins confers transducibility to the fusion protein and has been used to deliver a variety of proteins to living cells both in vitro and in vivo (Wadia, J.S. et al., 2005; Schwarze S.R. et al., 2005; Kanovsky M. et al., 2005; Phelan A., et al.1998). The Translocation across the cell membrane by PTDs is a rapid process, that occurs at 37°C as at 4°C, in dose-dependent manner practically in all cells type doing not have strong cellular specificity (Derossi D. et al.,1996;Vives E et al., 1997).

2.1.3. TAT sequence

HIV-tat peptide and its derivatives have recently received much attention, primarily because of their efficiency and short sequence. Tat peptide was originally derived from the HIV-1 tat protein, an 86-amino acid transactivation of transcription (TAT) protein involved in the replication of HIV-1 virus after it was shown that full length Tat protein could be taken up by cells and activate of the viral genome (Fawell, S. et al.,1994; Mann D.A et al., 1991; Vives, E. et al.,1997).

The transduction domain or region conveying the cell penetrating properties appears to be confined to a small (9 amino acids) stretch of basic amino acids derived from amino acids 47–57 of the HIV TAT protein, with the sequence RKKRRQRRR (Schwarze S.R. et al., 2000).

This short most basic (pI = 12.8) Tat peptide can deliver a large variety of cargoes, ranging from small particles of some hundreds of daltons to peptides, heterologous proteins of 110 kDa and nucleic acids and massive structures with a diameter up to 200 nm such as liposomes across most biomembranes without losing bioactivity, and it accumulates in the cell nucleus.

Several studies using tat peptide have demonstrated the breadth of strategies where TAT has been used successfully for efficient cargo delivery into cell culture and in vivo.

TAT PTD has been shown to deliver cargoes as large as iron nanobeads and fluorescent quantum dots into cells in culture TAT has also been used to deliver large, active proteins into the cells of live mice and TAT fusion proteins and peptides have been used to treat mouse models of cancer, inflammation and other diseases TAT has also been used to deliver phage encapsulated

DNA to cells, and liposome encapsulated DNA for gene expression in mice (Fawell S. et al, 1994; Schwarze and Dowdy 2000; Caron et al. 2001; Wills et al. 2001; Torchilin 2002; Ryu et al. 2003; Ryu et al. 2004; Snyder and Dowdy 2004; Wu et al.2006).

Despite the high number of biological applications using these peptides, and principally the Tat peptide, the precise mechanism of entry still appears controversial and certainly requires further investigations. The current model for TAT mediated protein transduction is uptake by way of macropinocytosis, a multistep process that involves binding of TAT to the cell surface, stimulation of macropinocytotic uptake of TAT and cargo into macropinosomes and endosomal escape into the cytoplasm (Gump and Dowdy, 2007). Macropinocytosis, a form of fluid phase endocytosis, is thought to occur in all cell types, also neuronal, because of the size of macropinosomes is such to take up large extracellular particles in the cells. TAT stimulation of macropinocytosis is important because it indicates that these molecules do not just passively enter cells but stimulate their own endocytic uptake.

In contrast with other amphipathic peptides with similar cell penetrating properties, Tat peptide has very low toxicity in cell culture. Being highly hydrophilic, tat peptide causes little perturbation to the plasma membrane, even during 1-hr incubations at a concentration of 100 μM (Hallbrink M. et al., 2001). So the TAT sequence using offers great advantages to delivery of antiPrP-scFv, because of allow a quick and easy administration in controlled amounts with nuclear and citoplasmic localization. Moreover TAT strategy allows to uptake the proteins also in denatured form because inside cells there are mechanisms and chaperones to refold the proteins. The use of TAT as a carrier for therapeutic proteins and peptides has many advantages over traditional gene therapy. The therapeutic is relatively short lived and lasts only as long as the lifetime of the protein or peptide. Most importantly, non-viral therapy does not require the integration of the nucleic acids into the genome of the patient, with potentially catastrophic consequences. Additionally, the use of TAT mediated delivery of macromolecules does not elicit the innate immune response of the body to viral repetitive structures, which can both limit the effectiveness of a gene based therapy and lead to toxic immunity (Boon Chin Heng and Tong Cao, 2005).

2.2. Materials and methods

2.2.1. Bacterial strains

The bacterial strain used in the present study are shown in Table 2.1.

Table 2.1 Bacterial strains used in this study

Strain

E. coli host

- Topl0F (F-*mcr*A Δ(*mrr-hds*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1*ara*D139 Δ(*ara-leu*) 7697*gal*U *gal*K *rps*L(strR)*end*A1 *nup*G). Novagen
- invaF' F´ *end*A1 *rec*A1 *hsd*R17 (rk-, mk+) *sup*E44 *thi*-1 *gyr*A96 *rel*A1 φ80*lac*ZΔM15Δ(*lac*ZYA-*arg*F)U169 λ- Novagen
- BL21(DE3) (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) Invitrogen
- BL21(DE3) Lys S (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) pLysS (CamR) Invitrogen

Topl0F, invaF' : used as recipient for cloning, transformation, propagation, and maintenance.
BL21[DE3], BL21[DE3] Lys S : specifically designed for expression of genes regulated by the T7 promoter. These strains have not been used for propagation or maintenance of interest plasmid.

Bacillus host

- *Bacillus subtilis* 1012 *leuA8 metB5 trpC2 hsdRM1* Schumann
- *Bacillus subtilis* WB800N *nprE aprE epr bpr mpr :: ble nprB :: bsr .vpr wprA :: hyg cm :: neo; NeoR* Schumann
- *Bacillus megaterium* MS941 MoBiTec

2.2.2. Growth condition

E.coli strains were grown aerobically in Luria Broth (LB) at 37°C supplemented, when necessary with specific antibiotic. *Bacillus subtilis* strains were grown aerobically in LB, MXR, 2xTY, Nutrient (Difco) or MSR at 25°C or indicated temperature supplemented, when indicated, with sugars (glucose and/or sucrose). For selection of plasmid, the medium was supplemented with chloramphenicol (10μg ml^{-1}) and/or neomycin (10μg ml^{-1}).

Bacillus megaterium strain was grown aerobically in MSR supplemented with sugars (glucose and sucrose) at 25°C or indicated temperature. For selection of plasmids, the medium was supplemented with chloramphenicol (4,5 μg ml^{-1}) and tetracyclin (10μg ml^{-1}).

Table 2.2. Culture media composition.

LB	MXR	2xTY	MSR
1%Bacto tryptone 0,5% Bacto yeast extract 0,5% NaCl	2.4% Bacto yeast extract 1.2% casein hydrolysate 0.4%glycerol 0.17 M KH_2PO_4 0.72 M K_2HPO_4	1.6%Bacto tryptone 1% Bacto yeast extract 1% NaCl 20 mM potassium phosphate buffer pH 7.0	2.5% Bacto yeast extract 1.5%Bacto tryptone 0.3% K_2HPO_4 1% glucose trace elements: 1ml/L 1000x stock solution (2M $MgCl_2$, 0,7M $CaCl_2$, 50mM $MnCl_2$, 5mM $FeCl_3$, 1 mM $ZnCl_2$)

2.2.3. DNA manipulation and transformation procedures

Procedures for DNA purification, restriction, ligation, polymerase chain reaction (PCR), agarose gel electrophoresis, preparation and transformation of competent *E.coli* cells were done using standard techniques (Sambrook et al. 1989).

2.2.4. Construcion of plasmids

All plasmids have been checked for correctness by cleavage with some restriction enzymes in agarose gels and by nucleotide sequencing.

2.2.4.1. pHT01/pHT43 *E.coli/B.subtilis* plasmids shuttle construction

pHT01-GFP or pHT43-GFP

The green fluorence protein (GFP) coding sequence was amplified using pHIS-GFP (Negro et al., 1997) as DNA template and the oligonucleotide primers

T37 5' GCAGATCTATGGCTAGCAAAGGAGAAGAAC 3' and

T36 5' GCTCTAGATCATTATTTGTAGAGCTCATCCAT 3' .

The PCR product was cleavaged with *BglII* and *XbaI* and cloned into *BamHI* and *XbaI* of pHT01 plasmid or pH43 (MoBiTec) to give pHT01-GFP or pHT43-GFP respectively.

pHT01-P43

To increase the production of GFP and scFv , the strong constitutive *Bacillus subtilis* P43 promoter sequence was in vitro synthesized (Genescript) and cloned in pUC57-PgracP43. The P43 was amplified from pUC57-P43 using the oligonucleotide primers BS11 5'GCGGATCCATGCGGGGTTCTCATCATCATCAT3' and BS13 5'GCTCTAGAGCGGTGGCAGCAGCCAACTCAG3' and cloned into *BamHI* and *XbaI* digested pHT01 to yield pHT01-P43.

pHT01- P43-GFP

The GFP coding sequence was amplified from pHT01-GFP using the oligonucleotide primers

T37 5' GCAGATCTATGGCTAGCAAAGGAGAAGAAC 3'and

T36 5' GCTCTAGATCATTATTTGTAGAGCTCATCCAT 3' .

The PCR product was released with *BglII* and *XbaI* and cloned into *BamHI* and *XbaI* digested pHT01+P43 to construct pHT01-P43-GFP.

pHT43-8H4

The eukaryotic TAT 8H4scFv sequence was amplified from pRSETB-8H4 using the oligonucleotide primers

T35 5' GCGGATCCATGCGGGGTTCTCATCATCATCAT 3' and

T38 5' GCTCTAGAGCGGTGGCAGCAGCCAACTCAG 3' .

The PCR product was cleavage with *BamHI* and *XbaI* and cloned in the same restriction sites of pHT43plasmid to give the secretion plasmid pHT43-8H4.

pHT43-8H4opt

The optimized 8H4scFv sequence was amplified from pUC57-8H4opt using the oligonucleotide primers DANBAC2 5' ATGGTAGATCTCATATGCGCGGTAG 3' and

DANBAC1 5' TCCTCTAGATTGGGAATTTCAGCCGCTAGCCG 3'

The PCR product was cut with *BglII* and *XbaI* and cloned in the same restriction sites of pHT43 plasmid to give pHT43-8H4opt.

pHT01- P43-8H4opt

The 8H4scFv optimized sequence was amplified from pHT43-8H4opt using the oligonucleotide primers BS14 5' GAAGGATCCATGAATCAAAAACGAAAGCGCAC 3' and

BS15 5' ACTCTAGATTGGGAATTCAGCCGCTAG 3' primers .

The PCR product was cleavage with *BamHI* and *XbaI* and cloned in the same restriction sites of pHT01+P43 plasmid to give the plasmid pHT01-P43-8H4opt.

pHT43 α–synuclein

The human α-synuclein coding sequence was amplified using pHIS-syn (Negro et al., 2004) as DNA template and the oligonucleotide primers BglII Syn and Xba Syn

5' GATCTAGATCAGGCTTCAGGTTCGTAGTCTTG 3'

5' GCAGATCTATGCGGGGTTCTCATCATCATCAT 3'

The PCR product was cleavage with *BglII* and *XbaI* and cloned into *BamHI* and *XbaI* restriction sites of pHT01 plasmid pHT43 (MoBiTec) to give pHT43 Syn.

2.2.4.2. pT7 *E.coli/B.megaterium* plasmids shuttle construction

pT7-8H4opt sec

The 8H4scFv optimized sequence was amplified from pHT01-P43-8H4opt using the oligonucleotide primers Dan end 5' ACTCTAGATTGGGAATTCAGCCGCTAGCG 3' and
Dan start T75' GCTGTACAATGATTCAAAAACGAAAGCGGACAGTT 3'
The PCR product was cleavage with *BsrGI* and *XbaI* and cloned into *BsrGI* and *SpeI* restriction sites of pT7plasmid to give pT7 -8H4opt.

2.2.5. Preparation of competent *Bacillus subtilis* cells

An overnight culture of *Bacillus subtilis* was prepared in 5 ml HS medium (table 2.4) at 37°C. 50 ml HS medium were inoculated with 0.5 ml of the overnight culture and incubated under vigorous shaking at 37°C. The growth curve was recorded, and when the cells reached the stationary phase, samples of 10 ml each were taken at 15 mm intervals, added 1 ml of sterile glycerol, mixed and left for 15 mm on ice. The cells were fractionated into 1 ml aliquots, frozen in liquid nitrogen and stored at -80°C. To identify the time point(s) yielding high level competent cells, one aliquot from each time point was checked with a reference plasmid DNA (pHT01).

2.2.6. Transformation of competent *Bacillus subtilis* cells

One competent cells aliquot was thaw at 37°C and used to inoculated 20 ml LS medium (table 2.4). Cells were shaked slowly in a 30°C water-bath to obtain maximal competence (about 2 h). Then 1-ml aliquots were taken into a class-tube, added 10 µ1of 0,1 M EGTA and incubated for 5 mm at room temperature. So plasmid was added and cell suspension was incubated for 2 h at 37°C while shaking. Cell suspension was transferred into an Eppendorf tube, centrifuged, supernatant discarded carefully and the cells were suspended into the final supernatant remaining on the pellet and plated on selective medium.

Table 2.4 Culture media composition.

HS medium	*LS medium*
66,5 ml dest. water	60 ml dest. water
10 ml 10x S-base	10 ml 10x S-base
12,5 ml 4% (w/v) glucose	12,5 ml 4% (w/v) glucose
5 ml 0.1% (w/v) L-tryptophan	0,5 ml 0.1% (w/v) L-tryptophan
2 ml 1% (w/v) casein	1 ml 1% (w/v) casein
25 ml 2% (w/v) yeast extract	5 ml 2% (w/v) yeast extract
10 ml 8% (w/v) arginine, 0,4% histidine	0,5 ml 0,5M $MgCl_2$
	0,5 ml 0,1M $CaCl_2$
	10 ml 1% (w/v) sodium citrate
	0,01 ml 1M $MgSO_4$

2.2.7. Preparation of competent *Bacillus megaterium* cells

50 ml LB medium were inoculated with 1 ml of a *Bacillus megaterium* overnight culture and incubated at 37°C in aerobic conditions. Cells were harvested by centrifugation (4500g, 15 min),

when the cell density has reached an optical density OD_{578} of 1. Cells were suspended in 5ml of fresh SMMP buffer (table 2.5). 50 µl of 1 mg/ml lysozyme were added in SMMP buffer for cell wall removal and incubated at 37°C for 60 min. Cells were harvested by centrifugation (1300g, 10 min), and gentle suspended in 5ml of fresh SMMP buffer. These two steps were repeated one more time. After suspension of the cell pellet, the protoplast were either directly used for transformation or frozen at -80°C in SMMP buffer containing 15%(v/v) glycerol for later use.

2.2.8. Transformation of competent *Bacillus megaterium* cells

500 µl of protoplast suspension are combined with 5 µg of DNA in SMMP, obtained mixing equal volumes of 2x SMM and 2x AB3 (Antibiotic Medium No. 3, DIFCO) in one 12 ml tube for each transformation. 1.5 ml of PEG-P were added and cells were incubated 2 minutes at room temperature (RT). 5 ml SMMP were added and mixed by rolling the tube carefully. Cells were harvested by gentle centrifugation (e.g. at 3,000 rpm for 10 minutes at RT), poured off supernatant immediately after centrifugation. 500 µl SMMP were added, and cells were incubate at 37 °C for 90 minutes with gentle shaking or rolling of tubes (max. 100 rpm). 2.5 ml aliquots of CR5-top agar were prepared in sterile tubes in a water-bath (max. 43 °C). After outgrowth, 50 to 200 µl of cells were added to 2.5 ml top agar, mix gently by rolling the tube between both hands and poured on a pre-warmed plate of LB containing the desired antibiotics. Cells were incubated overnight at 37 °C and streaked on fresh plates within two days.

Table 2.5 Culture media composition.

2x SMM	2x AB3	PEG-P	LB plates	SMMP
1M sucrose 40 mM maleic acid, disodium salt 40 mM $MgCl_2$ pH 6.5 autoclave for 12 minutes	7 g in 200 ml H_2O autoclave for 15 minutes	40 % (w/v) PEG6000 in 1x SMM autoclave for 12 minutes	Bacto-tryptone 10 g Bacto-yeast extract 5 g NaCl 10 g agar 15 g add 1 L adjust pH to 7.5 with sodium hydroxide	50% 2x SMM 50% 2x AB3

2.2.9. Analysis of GFP fluorescence

Fluorescent colonies expressing GFP gene (Negro et al., 1997) on solid agar were visualized at a wavelength of 360 nm on a UV illuminator (Reprostar II; CAMAG).

2.2.10. GFP fluorescence measurements

Recombinant GFP was quantified via fluorescence spectroscopy measurements (Luminescence Spectrometer LS50B, PerkinElmer). The fluorescence spectrum analysis of all GFP derivatives was carried out with a Photon Technology fluorimeter.

20 ml of LB or MSR added of glucose, sucrose or glucose-sucrose mixture in different percentage, as indicated, were inoculated with pHT01-GFP WB800N or pHT01-GFP 1012 *Bacillus subtilis* overnight culture up to OD_{600} =0.5, induced with 1 mM IPTG and incubated under vigorous shaking at 37°C. Cells samples of 1 ml were took at different time (2, 4, 7, 23 h), centrifuged 3 min at 10000 rpm, washed in phosphate-buffered saline (PBS) twice, resuspended in PBS and disrupted by sonication. Thus samples were diluited 1:10 in PBS prior to analysis.

For quantification of fluorescence, samples were excited at a wavelength of 400 nm, while fluorescence emission maxima were recorded at 512 nm.

2.2.11. Protein production studies in *Bacillus* strains

2.2.11.1. 8H4 scFv purification by medium

B. subtilis or *megaterium* strains carrying expression plasmid were cultivated in shake flasks at indicated temperature in MSR medium supplemented with glucose and sucrose or in other medium indicated. The media were supplemented with appropriate antibiotics to sustain the selective pressure on the stable replication of the corresponding plasmids. Tetracycline and chloramphenicol were used for *B.megaterium* at a final concentration of 10 and of 4.5µg ml^{-1} , respectively. Neomycin and chloramphenicol were used for WB800N *B.subtilis* at final concentration of 10 µg ml^{-1} each one. Recombinant expression of genes under transcriptional control of the xylose-inducible promoter was induced in *B.megaterium* by the addition of 0.5% (w/v) xylose at an optical density of 0.4 (at 578 nm). While recombinant expression of genes under transcriptional control of the IPTG-inducible promoter was induced in *B.subtilis* by the addition of 1mM IPTG at an optical density(OD) of 0.6 or indicated (at 600 nm). Samples were taken at indicated time points after induction. Cells were separated from the growth medium by centrifugation.

The scFv was purified by culture supernatant in native conditions by Ni-NTA affinity chromatography and eluited with imidazole. As control culture supernatant was concentrated by 44% of ammonium sulphate precipitation to yield the secreted protein fraction.

Recovery of antibody fragments via ammonium sulphate precipitation or affinity-purified, was assays by sodium dodecyl sulphate (SDS)-polyacrylamide gel and Western blot analysis using mouse anti-his primary antibody (1:3000; Sigma) and anti-mouse IgG-alkaline phosphatase secondary antibody (1:3000; Sigma).

2.2.11.2. 8H4 scFv purification by soluble cellular fraction

Cells were harvested by centrifugation at 7,000 x g. The cell pellets were resuspended in lysis buffer (500 mM NaCl, 10% glycerol, 0,1% Tween 20 mM Tris-HCl, pH 8.0) and disrupted by sonication. After centrifugation for 30 min at 10000 rpm at 4°C, the soluble fraction (supernatant) was collected and the scFv was carried out in native conditions using a NTA-Ni^{2+}column in the presence of 500 mM NaCl, 10% glycerol, 0,1% Tween, 250 mM imidazole, 20 mM Tris-HCl pH 6.5.

2.2.11.3. Protoplast preparation in *Bacillus* strains

To analyze the 8H4- scFv expression in cellular fractions, protoplast pellet and periplasmic fraction were prepared. Protoplasts were prepared as described by Merchante et al.(1995) with minor modifications.

Cells (10 ml) were harvested by centrifugation at 7,000 x g at 20°C. The culture supernatant was collected as the secreted fraction. The cell pellets were washed twice with SET buffer (20% sucrose, 50 mM EDTA50 mM Tris HCl, pH 7.6) and resuspended in 2 ml of protoplast buffer (66% sucrose, 16 mM $MgCl_{2,}$ 50 mM Tris HCl pH 8.0).The suspension was incubated at 37°C for 45 min in the presence of lysozyme (400 μg). The mixture was then centrifuged at 21,000 x g (15 min, 20°C) to yield the periplasmic fraction (supernatant) and protoplast pellet. Protoplasts were suspended in lysis buffer (5 mM $MgSO_{4,}$ 50 mM Tris HCl pH 8.0) and disrupted by sonication. The insoluble fraction was collected by centrifugation at 20,000 x g. As a control, the cell pellet from a similarly prepared 10-ml cell culture was directly suspended in lysis buffer without lyzozyme treatment and disrupted by sonication to yield the cellular soluble and insoluble fractions.

Protein samples were normalized against the cell density and were analyzed on an SDS 12% polyacrylamide gel and Western blot using mouse anti-his primary antibody (1:3000; Sigma) and anti-mouse IgG-alkaline phosphatase secondary antibody (1:3000; Sigma).

2.3. Results

2.3.1. Construction of the expression-secretion vector pHT43-8H4

A great advantage of *Bacillus subtilis* as host in recombinant protein production, is the its capacity of secreting functional extracellular proteins directly into the culture medium with high yields. The secretory production of recombinant protein has several advantages, such as simplicity of purification, authentic N-terminal amino acid sequence after cleavage by the signal peptidase, and a better chance of correct protein folding.The goal was the secretion of TAT-8H4 scFv directly into the culture mediumAt the aim, it has developed an expression-secretion vector to produce TAT 8H4scFv by replicative expression systems in *B.subtilis*. Using of an episomal vector with relatively high copies, it has advantage to express an higher level of interested gene (increasing the mRNA amount) than single copy in the chromosome. It was used vectors pHT, that are plasmid-based expression vectors for highly efficient intra- and extracellular production of recombinant proteins in *B.subtilis*. The plasmid pHT43 (MoBiTec) allows high-level expression of recombinant proteins directed into the medium. In fact, the sequence encoding target gene is fused with the coding region for the signal peptide of the amyQ gene, under the control of expression cassette Pgrac consisting of the *groE* promoter, the *lac* operator and the efficient *gsiB* Shine-Dalgarno sequence (Jürgen et al., 1998). Vector is based on the strong, costitutive σA - dependent promoter *groE* preceding the groE operon of *B.subtilis*, which has been converted into an efficiently controllable promoter by addition of the lac operator *lacO* of *E.coli*.

While the background level of expression of these expression cassettes is very low in the absence of the inducer IPTG an induction factor of about 1,300 was measured using the reporter gene (Phan et al., 2005). The amount of recombinant protein produced after addition of IPTG may represent 10-13% of the total cellular protein (Phan et al., 2005).

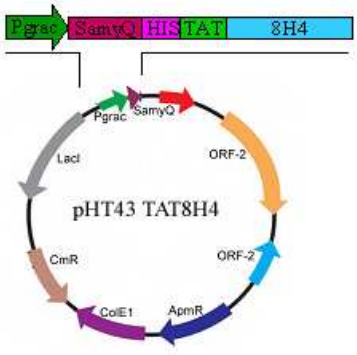


Fig. 2.1 pHT43 TAT-8H4 expression plasmid for secretion of scFv via AmyQ signal peptide.

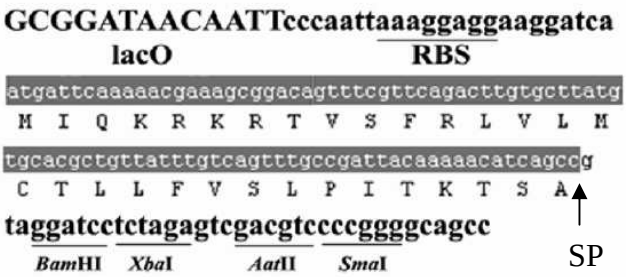


Fig. 2.2 MCS of plasmid pHT43 and complete signal sequence amyQ (in grey). The arrow indicates the site for scFv processing by signal peptidase (SP).

To obtain secretion of recombinant proteins, the coding region for the signal peptide of the amyQ gene encoding an α-amylase was fused to the SD sequence. High level secretion of amyQ α-

amylase of *B.amyloliquefaciens* and cellulase A and B of *Clostridium thermocellum* by sec A pathway was demonstrated in *B.subtilis* (Nguyen et al., 2006).

The expression-secretion vector pHT43 TAT8H4 (Fig.2.1) was constructed cloning the eukaryotic 8H4 scFv sequence, present in our laboratory, in pHT43 plasmid, under Pgrac promoter control and in fusion with secretion segnale peptide AmyQ (Fig. 2.2). At N-terminal of 8H4 scFv sequence there is an exa-his tag that allows to simplify IMAC purification and a TAT sequence for cell transduction.

The plasmid constructions correctness was controlled by recombinant plasmid sequencing and by restriction reactions.

```
10         20         30         40         50         60
MIQKRKRTVS FRLVLMCTLL FVSLPITKTS AVGSHMRGSH HHHHHGMARG YGRKKRRQRR

        70         80         90        100        110        120
RGARLAASGA HADILITQSP TTMAASPGEK ITIICSATSS ISPNYLHWYQ QKPGFSPRIL

       130        140        150        160        170        180
IFGTSDLASG VPVRFSGRGS GTSYSLTIVS MEAEDVATYY CQQGSSTPLT FGTGTKLEIK

       190        200        210        220        230        240
RSGGSTITSY NVYYTKLSSS GTEVMLVESG GALVQPGGSR KLSCAASGFT FSSFGMHWVR

       250        260        270        280        290        300
QAPEKELEWV AYISRGSSTG YCADTLKGRC TISRDNPKNT LFLEMTSLRS EDTAMYYCAR

       310        320
SGSSYGGWYF DVWGAGTTVT VSSASG
```

Fig. 2.3 Sequence of TAT 8H4scFv protein with amyQ signal secretion. In pink is showed His tag, in red amyQ signal secretion, in green TAT sequence, in blue 8H4 scFv sequence.

TAT8H4 scFv with signal sequence	properties
number of amino acids	326
molecular weight	35355.1
pI	9.83
ε	1.41

TAT8H4 scFv processed by SP	properties
number of amino acids	295
molecular weight	31790.6
pI	9.58
ε	1.57

Table 2.1 Some properties of TAT 8H4scFv protein, in the left the physic-chemical properties of TAT 8H4scFv with signal secretion and in the right physico-chemical properties of the mature TAT 8H4scFv protein

Because of *B.subtilis* low transformation efficiencies, in cloning it was used a shuttle vector pHT, that replicates in *E.coli and B.subtilis*. Therefore all cloning steps are carried out in *E.coli*, and the final recombinant plasmids pHT43-TAT-8H4 scFv is then used to transform the appropriate

B.subtilis cells. We have choose WB800N *B.subtilis* strain, because of it is a seven-extracellular protease and the wall-bound protease WprA deficient, that could degrade the scFv that should be produce in secreted form. In fact extracellular proteins unable to fold correctly are prone to degradation by protease. But even correctly folded heterologous proteins can be degraded by extracellular protease.

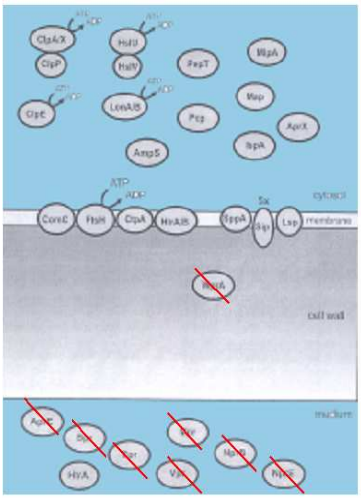


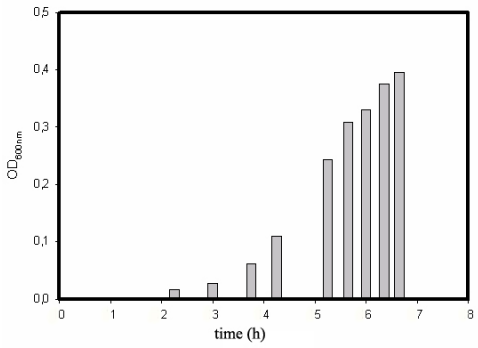
Fig.2.4 Proteases and peptidases of WB800N protease-deficient *B.subtilis* strain. The localization of known *B.subtilis* proteases and peptidases in the cellular compartments and the growth medium is indicated.
In WB800N strain structural genes encoding a total of seven soluble protease characterized [the genes AprE and NprE coding for alkaline(subtilisin) and neutral protease, the genes coding for extracellular protease (epr), the metalloprotease (mpr), the bacillopeptidase F (brp), the minor soluble protease vpr), and the cell-bound WprA protease] have been deleted. In red are barred the inactivated protease (*nprE; nprB; aprA; epr; mpr; bpf; vpr; wpr.)*
(Adapted from Quax W.J. et al., 2004.)

To the aim, a protocol of *B.subtilis* competent cells preparation and transformation was set because of initially in our laboratory the knowledge about *Bacillus* host-strain and expression system were absent.

2.3.2. Preparation and transformation of competent *Bacillus subtilis* cells

Different methods published allowing the introduction of genes into *B.subtilis* cells were been tested, with poor results. After several trials, *B.subtilis* competent cells preparation and transformation protocol based on natural competence was prepared. Many strains of *B.subtilis* are naturally capable of taking up DNA fragment under certain physiological conditions referred to as "competence" (Chen and Dubnau, 2004). This is a physiological state in which a bacterial cell stops dividing and expresses a set of genes for the internalization of environmental DNA, eventually incorporating this DNA into its own genome in a process known as transformation. A hallmark of competence is its bimodal expression. Only a minority of cells in a competent population expresses the DNA transport genes, utmost 20% of the cells in a culture shortly before the cessation of exponential growth. Individual cells therefore exist in two distinct states with respect to competence gene expression. This protocol has been applied to *Bacillus subtilis* WB800N and 1012 strains. To identify the time point(s) yielding high level competent cells, the growth curve was recorded and when the cells were reaching the stationary phase, samples were taken at 15-20 min intervals, and one aliquot from each time point was checked with a reference

plasmid DNA as pHT01 that confers resistance to cloramphenicol. As it can see in an example for WB800N, the aliquot with higher level competent cells is the number five, fig.2.5.



Aliquot tested (collection time)	Selection C=cloramphenicol N=neomicin	DNA transformant pHT01= control plasmid	N° colonies after transformation
1	C+N	pHT01	0
2	C+N	pHT01	21
3	C+N	pHT01	2
4	C+N	pHT01	83
5	C+N	pHT01	107
5	C+N	-	0
5	N	-	plate

Fig.2.5. *Bacillus subtilis* growth curve. In table an example of results after transformation of WB800N (resistant to neomycin) with control plasmid pHT01, that confers resistance to chloramphenicol, to identify the time point(s) yielding higher level competent cells.

2.3.3. Medium optimisation for the expression of recombinant protein in *B.subtilis*

At this point, it is looked for the optimal medium for *Bacillus subtilis* recombinant protein expression.

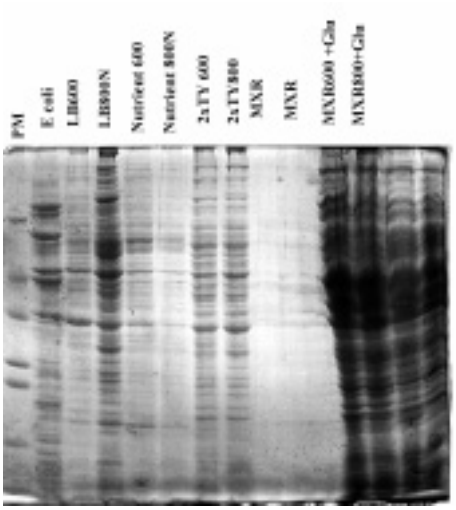


Fig.2.6 SDS-PAGE analysis of total intracellular protein production by WB600 and WB800N *Bacillus subtilis* cells grown in 2xTY, 1xMXR, LB, Nutrient or 1xMSR, after 5h from normalized against the cell density inoculum.

As a first approach to optimize recombinant protein production, total intracellular protein expression was tested by two *B.subtilis* strains (WB600 and WB800N), grown in 2xTY, 1xMXR, LB, Nutrient or 1xMSR. The highest cells optical densities and intracellular protein yields were obtained upon growth in LB and 1xMSR.

Then, in order to evaluate the abilities (strength and inducible expression regulation) of the plasmids used in our cloning (pHT01 and pHT43), to drive expression of heterologous genes in *B.subtilis*, have been constructed expression vector pHT01-GFP (Fig.2.7), carring the Pgrac promoter trascriptionally fused to the (reporter gene) fluorescent biotracer green fluorescent protein (GFP).

The two vectors pHT01 and pHT43 allow high-level expression of recombinant proteins within the cytoplasm, where pHT43 directs the recombinant proteins into the medium. They differ only for the presence or no of the signal secretion sequence amyQ.

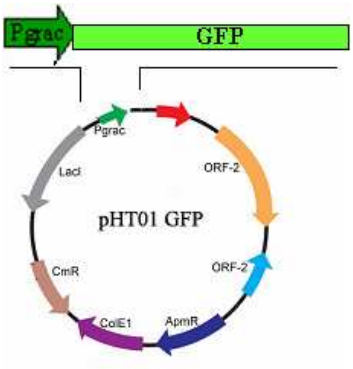


Fig. 2.7 pHT01 GFP expression plasmid.

So it is easier follow the protein expression, because of green fluorescence was visualized when GFP was excited at a wavelength of about 400 nm on a UV illuminator.

The pHT01-GFP was transformed before in BL21Lys S cells, that has been seen green after IPTG induction under UV excitation.

Then the pHT01-GFP was transformed simultaneously into 1012 and WB800N *B.subtilis* strains (Fig.2.8). GFP expression was evaluated in the time (2,4,7,23 h after IPTG induction) in the more productive media (MSR and LB) added of different carbonium sources (glucose and/or sucrose in different percentage) by scanning the fluorescence emitted by resulting cultures as described in materials and methods (Fig. 2.9).

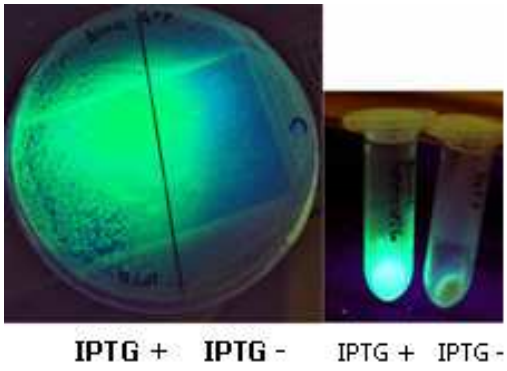


Fig. 2.8 1012 and WB800N *Bacillus subtilis* fluorescent colonies expressing *gfp* gene on solid agar or liquid medium were visualized at a wavelength of 360 nm on a UV illuminator. IPTG+ : induced colonies, IPTG- : no induced colonies.

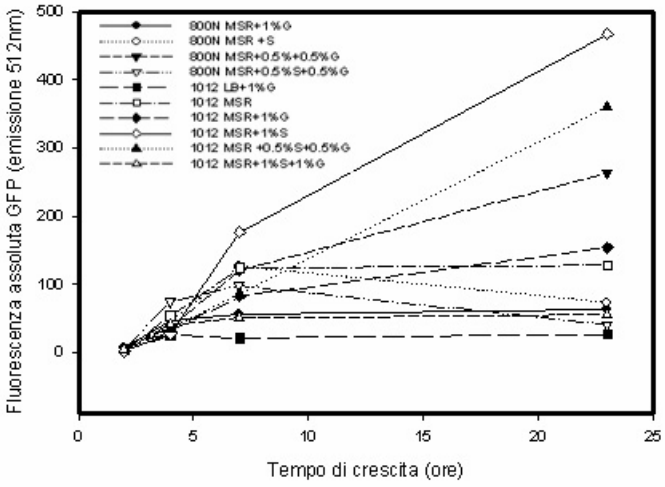


Fig. 2.9 GFP expression by pHT01-GFP 1012 and WB800N *Bacillus subtilis* strains was evaluated in the time and in MSR and LB media added of different carbonium sources (glucose and/or sucrose in different percentage) by scanning the fluorescence emitted at 512 nm by resulting cultures.

In conclusion the expression changes on base of growth conditions. In classical medium as LB with glucose, the protein expression level obtained is about 15 time minus than in more rich media. The major protein expression has been shown in MSR+1% sucrose for 1012 *B.subtilis*, or MSR added of mixture 0,5% sucrose and 0,5% glucose for 1012 and 800N *B.subtilis*.

2.3.4. Production of TAT-8H4 scFv in WB800N *B. subtilis* by secretion

2.3.4.1. TAT-8H4 scFv production in pHT43 TAT8H4 WB800N *B. subtilis*

In these optimised media, to individuate more productive conditions, WB800N *B. subtilis* growth curve trials were done in high rmp shaked flasks at different temperature 25°C or 37°C and measuring OD_{600} and pH medium (fig. 2.10).

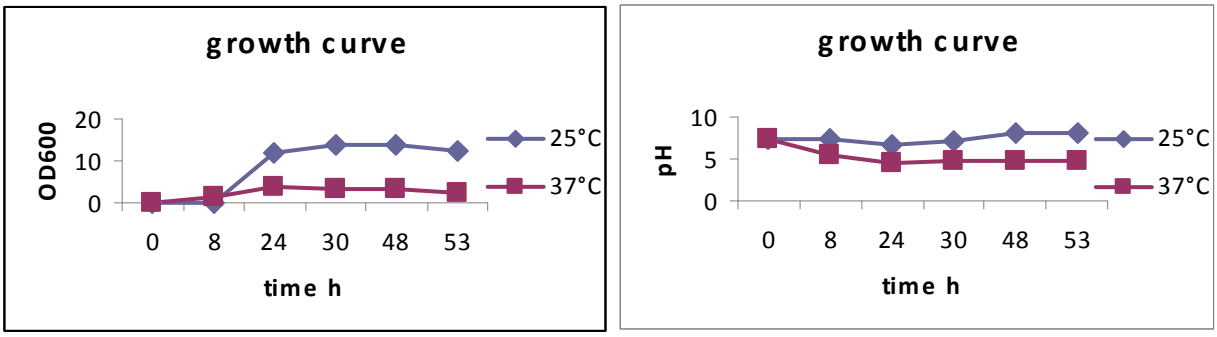


Fig. 2.10 WB800N *B. subtilis* growth curve trials in high rmp shaked flasks at different temperature 25°C or 37°C and measuring OD_{600} and pH medium to evaluate the best growth conditions.

Since, the best growth condition for *B.subtilis* were reached after 24-30 h at 25°C with pH values about neutrality, we started producing, secreting and purifying TAT 8H4scFv protein in these condition. In fact, it is noted that *Bacillus* species can secrete large amounts of industrial enzymes (mainly native secreted proteins) into the growth medium, in particular during the post-exponential growth phase (Fig.2.11) (S.Bron et al., 1998).

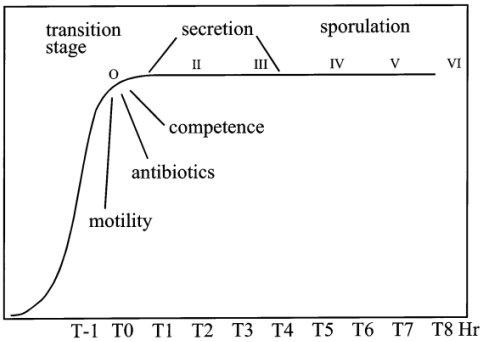


Fig. 2.11 Post-exponential gene expression in *B.subtilis*. A typical growth curve in minimal medium is shown. As part of an elaborate network of signal transduction, *B.subtilis* can express several regulons in a temporal way. Most of these are expressed during or after the transition from the exponential to the stationary phase of growth. The various post-exponential phase processes are indicated in relation to the growth phase. (Adapted from S.Bron et al., 1998).

WB800N *B.subtilis* was transformed with pHT43 TAT8H4 scFv and pHT43, as negative control. The expression of TAT 8H4scFv secrete protein in WB800N was tested, by SDS-PAGE and immunoblotting developed with anti-his antibody. The immunoblot analysis of culture medium after IMAC purification showed no TAT-8H4 scFv secreted at 24-30h . Nevertheless, the immunoblot analysis of cell pellet showed a main band of expected molecular weight of about 35-32 kDa (theoretical molecular weight for preprosequence 35300 Da and for mature 31790 Da). So scFv is expressed in WB800N cells even if with no very high yields, because it is visible only in

Western. By immunoblot analysis the intracellular produced scFv is shown prevalently as insoluble, even if about 35-40% is soluble (Fig.2.12). So in conclusion: 1) scFv is expressed in cells even if no with very high yields, 2) no TAT-8H4 scFv is detected in purified culture medium, 3) the intracellular produced scFv is shown as soluble and insoluble.

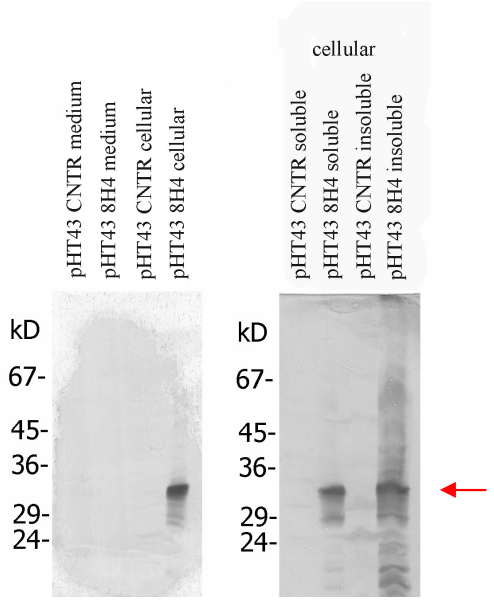


Fig. 2.12 Immunoblot detection of TAT-8H4scFv in the culture medium and in cell pellet homogenate of pHT43 TAT 8H4 WB800N *B.subtilis* at the best growth condition (24-30 h).

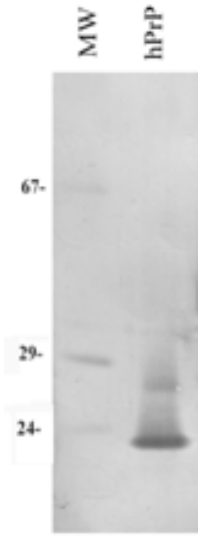


Fig. 2.13 Interaction between intracellular soluble purified TAT-8H4 from *B.subtilis*, and human PrP^C confirms the correctness and identity of TAT-8H4scFv as showed by immunoblotting. 50 ng of PrP^C was transferred to membrane and incubated with about 100 ng of TAT-8H4scFv purified in native conditions, and following with mouse anti-His and anti-mouse IgG.

2.3.4.2. Antigen binding specificity of intracellular soluble TAT-8H4scFv

To confirm the expression and correctness of TAT-8H4scFv produced by pHT43 TAT8H4 scFv WB800N *B.subtilis*, TAT 8H4 was purified from intracellular soluble fraction in native conditions (with low yields, about 0,3-0,4 μg/ml).

The specificity of soluble purified 8H4 scFv to recognize and interact with $hPrP^C$ was determined by immunoblotting. As shown in Fig. 2.13, the soluble correctly folded TAT 8H4scFv was active and able to recognize human prion protein, demonstrating the functionality of the construction.

In order to obtain the extracellular TAT- 8H4scFv production, several other experiments of expression and medium purification from pHT43-8H4 WB800N *B.subtilis* has been made, changing carbonium sources (sucrose, glucose) and amounts in media, changing culture incubation temperature before and after induction (37°C, 30°C ,25°C), testing different secretion times (2, 3, 4, 6, 7, 16, 24, 30, 48 h) and different OD induction (OD_{600} 0.6, 0.7, 0.8), and also changing purification conditions as elution (0.5M imidazol pH6.5, 250mM NaCl, 2mM TrisHCl pH8 or 0.5M imidazol 50 mM sodium acetato pH 5.9 or urea 8M, NaH_2PO_4, TrisHCl pH4.5), or

using a wild-type (eukaryotic) or a synthetic optimized scFV gene sequence for prokariotes but in any case it have been possible to visualize TAT 8H4scFv purified from culture medium.

2.3.4.3. TAT-8H4 scFv production in pHT01-P43 TAT8H4 WB800N *B. subtilis*

Promoters are important regulatory elements, that control influencing temporal expression of genes and their expression strength. At this purpose, it has been tested the expression of 8H4scFv sequence also under constitutive strong promoter P43 from *B subtilis*. P43 is a highly efficient promoter that comprises two overlapping promoters recognized by the *B.subtilis* σA and σB containing RNA polymerase holoenzymes. pHT01-P43 8H4scFv and pHT01-P43 GFP plasmids were constructed and transformed in WB800N *B.subtilis*. GFP expression under P43 promoter allowed to verify promoter activity either in *E.coli* either in *B.subtilis*. In the best conditions previously identified, immunoblot analysis showed that the 8H4 scFv expression in WB800N *B.subtilis* under P43 promoter control was not improved, with no scFv detected in purified culture medium and an intracellular scFv production, almost totally insoluble.

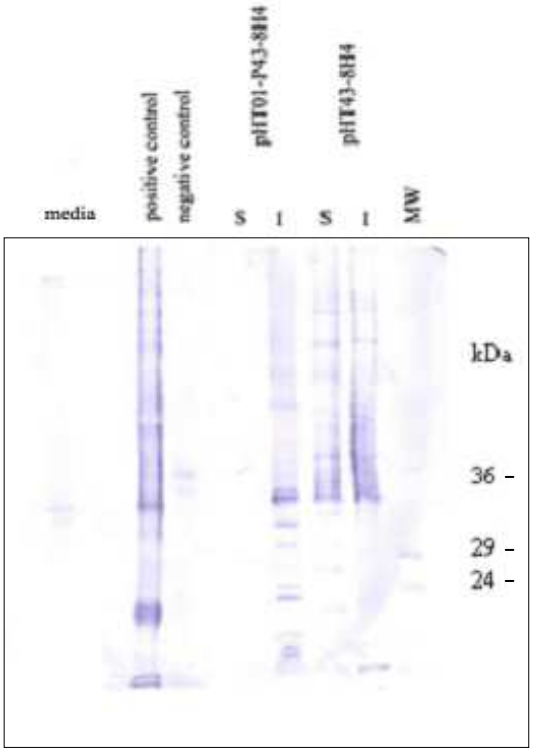


Fig. 2.14 Immunoblot detection of TAT-8H4scFv in the culture medium and in cell pellet homogenate of pHT43 TAT8H4 WB800N *B.subtilis*, in post-exponential growth phase (24-30 h).

To further verify the secretion, we have tried to produce other two proteins in secrete form (GFP and α-synuclein, a short and stable protein without disulphide bonds). Secretion vectors pHT43-GFP and pHT43-αsyn were constructed, and transformed inWB800N *B.subtilis*, but there was not visible fluorescence or protein secreted in medium (data not shown).

2.3.5. Production of TAT-8H4 scFv in *B. megaterium* by secretion

2.3.5.1. Construction of the expression-secretion vector pT7-TAT8H4

One important feature of efficient protein production is the use of strong, tightly controllable promoters, at this purpose, scFv secretion was evaluate in *B.megaterium* engineered with T7 RNA polymerase expression system.

This expression system, just developed and widely used for *E.coli* (Tabour and Richardson, 1985), offers the advantages of stringent selectivity and high transcriptional activity so that it is possible to lead to a saturation of the protein-synthesizing machinery in host organism. Consequently 50% or more of the total cellular protein can consist of the desired protein (Studier and Moffatt, 1986). The T7 RNAP expression system for *Bacillus megaterium* based on two parallel-replicating plasmids: pT7-RNAP and pP T7 (Gamer et al. 2009) distribuited by MoBiTec and combines the features of this system in *E.coli* with the regulation by the xylose operon.

Plasmid pP T7 contains all structural elements necessary for a T7 RNAP-dependent expression of target genes under the control of T7 promoter.

Plasmid pT7-RNAP contains the gene encoding RNA polymerase of the bacteriophage T7 (T7 RNAP) under the transcriptional control of xylose-inducible promoter PxylA . This promoter is regulated by repressor coded by XylR gene, that blocks xyl operon transcription in absence of xylose. In the presence of xylose, repressor takes off and starts transcription of T7 RNAP that recognizes its specific promoter beginning target genes transcription.

Using this promoter system localized on free replicating plasmids for the overexpression of recombinant genes in *B.megaterium* yielded an induction of up to 350-fold while adding xylose to the growth medium (Gamer et al 2009). T7 RNA polymerase is highly processive and recognise completely different promoters from *Bacillus* cells. So it can completely be used selectively to transcript target gene, without interfere with the growth and synthesis of bacterial protein.

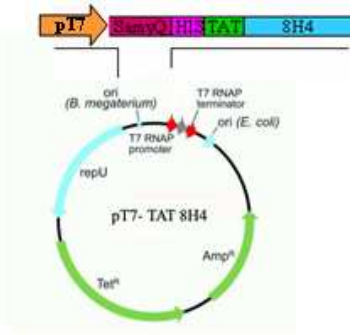


Fig. 2.15 pT7 TAT-8H4 expression plasmid.

The expression-secretion vector pT7-TAT8H4 was constructed, cloning the TAT8H4scFv sequence in pT7 plasmid, under T7 promoter control and in fusion with secretion signal peptide AmyQ of *B.subtilis* amylase Q gene, as described in materials and methods. At N-terminal of 8H4 scFv sequence there are an exa-His tag allowing to simplify IMAC purification and a TAT sequence for cells transduction (fig.2.15).

Because of the low transformation efficiency of *Bacillus* strains, all plasmid constructions were

done in *E.coli* TOP10 with the pT7shuttle vector, able to replicate in both *E.coli* and *B.megaterium*. The plasmid constructions correctness was controlled by recombinant plasmid sequencing and by restriction reactions.

2.3.5.2. TAT-8H4 scFv expression in pT7 TAT8H4 *E.coli*

The expression of recombinant protein produced in *E.coli* (DE3) LysS cells (provided of T7 RNA polymerase expression system) transformed with pT7 TAT8H4 , has been verified by SDS-PAGE and immunoblotting (fig.2.16). A main band of expected molecular weight of about 32-35 kDa (theoretical molecular weight preprosequence 35300 Da, mature 31790 Da), is visible in transformed and IPTG induced cells. It is comparable with TAT 8H4scFv without secretion sequence produced by plasmid pRsetB just present in our laboratory, and both are produced in very low amounts, being visible in western, thus indicating that the protein although express at low amount is correctly processed in mature form.

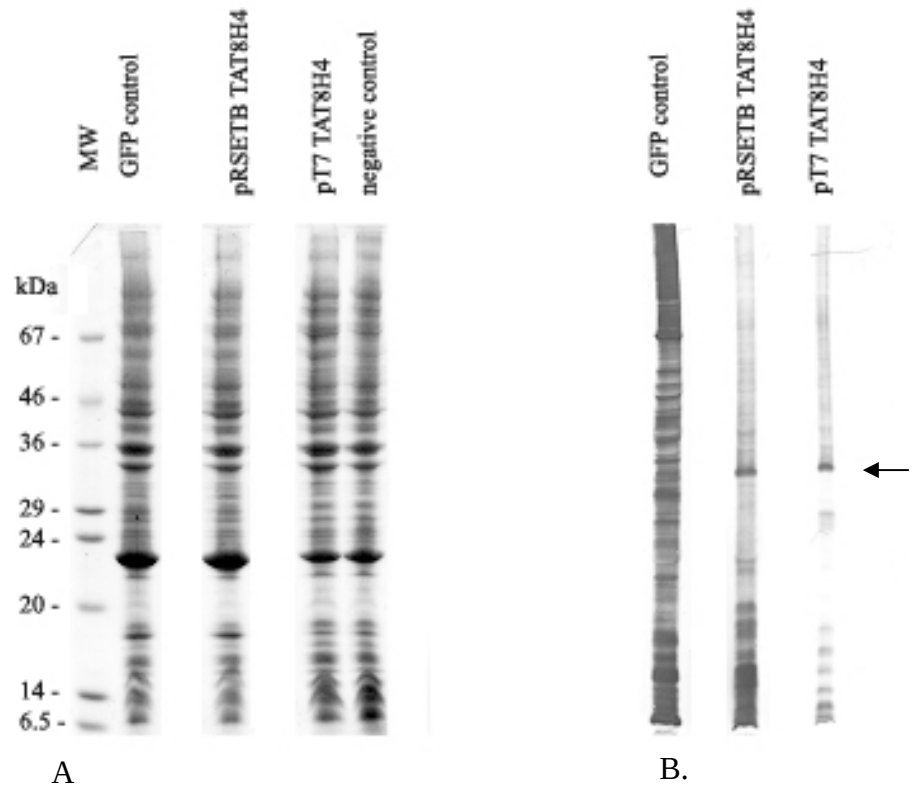


Fig. 2.16 SDS-PAGE (A) and western blotting analysis (B) of TAT-8H4scFv in cell pellet homogenate of *E.coli*. BL21[DE]LysS cells were transformed with the indicated plasmids and protein expression was induced with IPTG. Cell lysates corresponding to 40 µl of culture were loaded in Coomassie blue-stained gel (A), while corresponding to 10 µl of culture were loaded in western.

2.3.5.3. TAT-8H4 scFv expression in pT7 TAT8H4 *B. megaterium*

pT7-8H4 plasmid was transformed in MS941 *B.megaterium* competent cell. This strain, deleted of extracellulare protease NprM, provides a suitable host for the production and secretion of

heterologous proteins (Wittchen and Meinhardt, 1995). At the aim has been applied a poly-ethylene-glycol-mediated protoplast transformation protocol (Barg et al., 2005) to remove the solid cell wall of this bacterium and isolate the protoplasts. Two of about 20 transforming clones, were controlled by PCR to verify the presence of recombinant plasmid pT7-TAT8H4. These two pT7-8H4 *B.megaterium* transformed clones, were grown in sugar added MSR medium at 25°C and by time course trials, with samples taken at different culture times (5, 24, 30 h), the TAT8H4scFv expression in culture medium and cell pellet were analyzed. The immunoblot analysis of culture medium after ammonium sulphate precipitation (and IMAC purification) showed no clearly TAT-8H4 scFv secreted (Fig. 2.17). Nevertheless, the immunoblot analysis of soluble and insoluble intracellular fractions show that 1) scFv is expressed in cells even if no very high yields, because it is visible only in western 2) the intracellular produced scFv is shown so far as insoluble.

To elucidate the rate-limiting step(s) in the secretory production of 8H4 scfv with the T7 RNAP-dependent expression system, subcellular protein fractions were analysed. Immunological detection show that 1) 8H4 scfv is revealed as far as in insoluble fraction of protoplasts after 24-30 h from induction 2) a small fraction is revealed in soluble fraction of protoplasts and also in periplasmic fraction. This last result should prove that the scFv is secreted in periplasmic space between plasma membrane and cell wall, after signal peptidase-mediated processing.

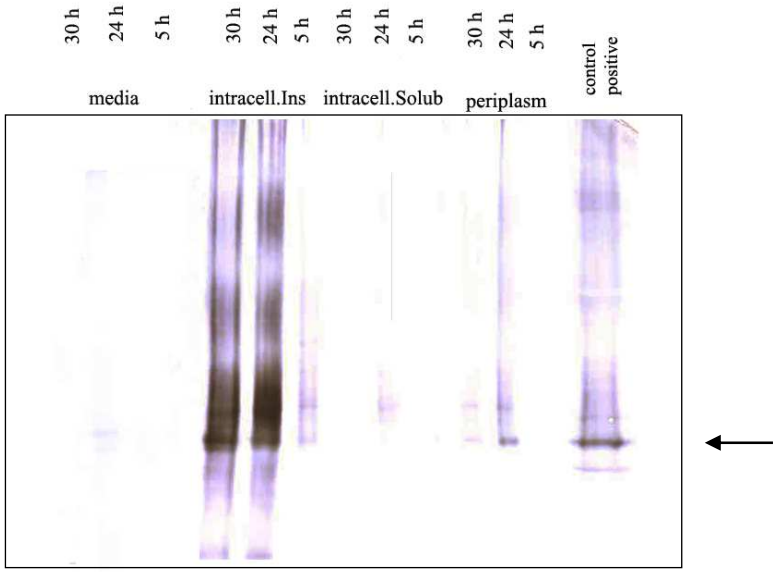


Fig. 2.17 Immunoblot detection of TAT-8H4scFv in the culture medium, intracellular soluble and insoluble protoplasmic fractions, and in periplasmic fraction of pT7 TAT8H4 *B.megaterium* on the time (5, 24, 30 h).

2.3.6. Discussion

Scfv are recombinant antibody fragments becoming popular therapeutic alternatives to full length monoclonal antibodies (Abs) since they are smaller, can be produced more economically, retain the binding properties and antigen specificity of their parent Abs and are easily amendable to genetic manipulation with possibility of engineering for intracellular expression or secretion. Several groups have investigated the use of anti-prion scFvs in prion-infected cell culture systems or in vivo analyses (with mice) and have been recently used as therapeutic approach in prion diseases resulting in significant reduction of the peripheral PrP^{SC} propagation (Gabizon et al., 1988; Peretz et al., 2001; White et al., 2003). In contrast to full length antibodies, these antigen binding fragments can conveniently be produced in *E.coli*. Although this production system is widely used, corresponding yields are somehow limited and in particular hampered by the inefficient secretion, leaving more than 90% of the produced scFv inside of the cell. Moreover, the majority of protein is found denatured in form of insoluble inclusion bodies.

Novel production systems could enhance the yield of functional antibody fragments or allow cheaper and easier production and purification. Various microorganism beside *E.coli* have been used for the production of recombinant antibody fragments. Among these, Gram-positive bacteria, like *Bacillus subtilis* (Wu et al., 1998 and 2002) *B.brevis* (Inoue et al., 1997; Shiroza et al., 2003), *B.megaterium* (Jordan et al., 2007) in particular, are advantageous due their capacity to secrete scFvs directly into the medium.

In this study, 800N *B.subtilis* and MS941 *B.megaterium* are evaluated for their capability to produce and secrete anti-prion TAT-8H4scFv protein, an eukaryotic complex molecule with three disulphide bonds indispensable for its stable structure and activity.

8H4 ScFv protein was just produced previously in our laboratory in *E.coli*, but with very low yields. Several expression vectors were constructed, cloning the wild-type or a synthetic *E. coli* optimized 8H4 scFv sequence, with TAT sequence for cell transduction, in pHT plasmid, under control of IPTG-inducible Pgrac or strong constitutive P43 promoters, and in fusion with signal peptide AmyQ. The immunoblot analysis showed that scFv is expressed in 800N *Bacillus subtilis* cells intracellulary even if with no very high yields and primary as insoluble, while no TAT-8H4 scFv is detected in IMAC purified culture medium although high level secretion by *Bacillus* secA pathway was widely demonstrated in these system pHT with high-level secretion of exoproteins (Schumann W. et al., 2006).

Because of the promoters influence the spatial and temporal expression of genes, it has been tested the expression-secretion of TAT 8H4scFv also under the strong, tightly controllable, xylose-inducible T7 promoter control in MS941 *B.megaterium* no secreting alkaline protease. The immunoblot analysis at different times of culture medium after IMAC purification and ammonium

sulphate precipitation showed no clearly TAT-8H4 scFv secreted, while a small fraction is revealed in periplasmic fraction and primary in insoluble intracellular fraction. This result should show that the scFv is secreted in periplasmic space between plasma membrane and cell wall, after signal peptidase-mediated processing. Neither scFv extracellular secretion was improved by utilization of media with increased Mg-ion concentration an lower Co and Ca concentration, optimized to directly favour higher scFvs formation and secretion in *Bacillus megaterium* as showed in study of David (David F. et al., 2010) about the role of metal ion on scFv secretion in *B.megaterium*.

Notwithstanding stronger promoters to further increase protein production, and highly secreting *Bacillus* strains, deleted of extracellular proteases (that could degrade secreted proteins) were been chosen, no clearly TAT-8H4 scFv secreted was revealed in our hand. Although examples of scFv protein secreted into the growth medium are described with more than 400 μg/L of recombinant anti -His6-tagged lisozyme scFv in *B.megaterium* (Jordan et al., 2007), and 12 mg/L of antidigoxin scFv in *B. subtilis* WB600B(six-extracellular-proteases deficient) and WB600BHM (costitutively producer of major intracellular molecular chaperones) *B.subtilis* strains, respectively (Wu et al., 1998 and 2002).

Surely the yield is strongly dependent on the individual sequence of the produced antibody fragment (Jordan et al., 2007), but it is noted that, despite high capacity for secretion of homologous proteins (industrial enzymes, such as proteases and α-amylases with yields of ~20 g/L) the secretion of various heterologous proteins by bacilli, in particular proteins of eukaryotic origin, is sometime inefficient, which limits the application potential of these organisms.

The causes for these failures are a combination of the properties of the secretion pathway, the *Bacillus* cell envelope, quality control proteases and the secreted target proteins themselves (Harwood and Cranenburgh, 2008).

Five potential bottleneckes have been found and documented in the secretion pathway of *B.subtilis*. First, heterologous proteins may form insoluble aggregates in the cytoplasm due to limited activity of chaperones (Wu et al., 1998). Second, the Spase can be a limiting factor in preprotein processing (van Dijl J., 1992). Third, it has been shown that the folding catalyst PrsA, which is attached to the extracytoplamic side of the membrane by lipid modification, sets a limit to the high-level secretion of certain secretory protein (Kontinen and Sarvas, 1993). Fourth, it has been suggested that the cell wall forms a barrier for at least one secreted heterologous protein, human serum albumin (Saunders C., 1987). Fifth, it has been known for a long time that *B.subtilis* secretes large amounts of proteases into the medium, which can degradate secreted heterologous protein (Simonem M. and Palva I., 1993; Nagarajan V.,1993). Recent studies suggest that not only the secreted proteases but also cell-associated proteases,(as Htr) are responsible for the

degradation of secreted heterologous proteins, particularly in stress conditions (Stephenson K.and Harwood C.R., 1998).

Thus TAT8H4scFv could be degraded by protease (notwithstanding bacilli strains used were deleted of most extracellular protease), or could occur interactions between secretory protein and the wall likely affected by the physico-chemical properties of both (the secretory proteins and the wall, composed of a matrix of peptidoglycan (Rogers, H. J.) and of covalently attached anionic polymers teichoic or teichuronic acid (Ward J. B.,1981), which confer a high degree of negative charge and contribute to the specific surface-charge properties of the cell (Archibald A. R., 1993).

It is shown that proteins that are naturally secreted into the culture medium by *B.subtilis* (strain W23) are neutral in terms of charge , pI values in the range 6.5-7.0 (Coxon, R. D.,1990; Kunst F., 1997); while proteins which function within the wall , have pI values of ~10 (as autolysin LytC or serine protease CWBP52 with pI=10.6 and 9.9 respectively (Lazarevic V., 1992; Margot P.,1996) which ensure that these proteins remain firmly attached to the wall by electrostatic interactions This may represent an important evolutionary adaptation to prevent proteins which are destined to exert their function in the extracellular environment from attaching inappropriately to the cell wall. Moreover a greater proportion of an engineering AmyL variant with modified net charge (pI about 10) binds to the wall following translocation across the cytoplasmic membrane, as compared with a normally well-secreted *Bacillus* α-amylase, AmyL, which is neutral in terms of net charge. The data demonstrates that the wall, as a consequence of electrostatic interactions, has the potential to retard the passage of positively charged secretory proteins to a greater extent than neutral or negatively charged proteins.

TAT 8H4scFv has a calculated pI of 9.83 due to high positive charge of TAT sequence, so the cell wall of Bacilli strains could act as a `barrier ' to the secretion, as it has been suggested for other certain heterologous proteins, but other investigations should be done about this.

Moreover, the fate of proteins emerging from the secretory translocase at the trans side of the cytoplasmic membrane, is largely determined by their ability to fold rapidly in the micro-environment created at the membrane–wall interface, where there are membrane- and wall-associated proteases that function to maintain the quality of the secreted proteins, and possibly to keep the translocase and cell wall free of misfolded or aggregated proteins.These proteases are induced in response to stresses that are likely to lead to protein misfolding.

Proteins emerging from the translocase are therefore subjected to opposing activities that, ultimately, decide their fate. To some extent, this fate is different for native and heterologous proteins. Native proteins have evolved to coexist with the quality control machine that is designed to ensure that secretory proteins neither block the translocase nor growth sites within the cell wall. They fold very rapidly into their native structures that are resistant to the quality

control proteases and which are compatible with the cell wall. Their rapid folding is aided by folding factors such as propeptides, PrsA, disulfide bond forming enzymes and/or metal cations. These proteins are now free to interact with, or traverse, the cell wall. Reducing the folding kinetics of normally well secreted proteins leads to an increase in their susceptibility to proteolysis and a reduction in their yield.

The fate of heterologous proteins is largely determined by their intrinsic folding kinetics in the environment at the membrane–wall interface, the ability of host folding factors to assist their folding and their compatibility with the cell wall (Sarvas M. et al., 2004).

However, the increased understanding of the factors involved now holds out the prospect of engineering production strains for the high-level production of heterologous proteins.

Interesting could be using of WB800 *B.subtilis* strain (Wu S.C. et al, 1998 and 2002), engineered for enhanced coproduction of both intracellular (GroES/EL and DnaK/DnaJ/GrpE) and extracytoplasmic (PrSA) molecular chaperones that mediate the proper folding, assembly, and higher secretory production of scFvs (e.g.MH-1 SCA at a level of 10 to 15 mg/liter), (Wu S.C. et al, 1998 and 2002).

In conclusion, because of the high potential secretory of *Bacillus* strains was not appeared in TAT 8H4 scFv production and secretion, a different strategy was experimented.

Chapter III

TAT-8H4 scFv production in *E.coli*

3.1. Introduction

3.1.1. Recombinant protein production in *E.coli*

The high secretory capacity of *Bacillus* strains is not appeared in secretion and purification of TAT8H4 scFv (our data cap. II), therefore other strategies have been proved to increase the scFv yields in *E.coli* .Gram-negative bacterium *E.coli* has been the "workhorse" for the production of recombinant proteins for obvious advantages offered, and the highest yields are in cytoplasm. Previous experiments in our laboratory of 8H4 scFv production in BL21 *E.coli* cytoplasm, were resulted in low yields of insoluble antibody. The soluble expression of heterologous proteins in *E.coli* remains a serious bottleneck in protein production. Although the percentages of soluble heterologous proteins expressed in *E.coli* ranging from 13% to 23% (Schlieker C. et al., 2002; Chambers S.P., 2004), many of the most biochemically interesting families of proteins, including kinases, phosphatases, membrane-associated proteins and many other enzymes or scFvs, are extremely difficult to produce as soluble forms in *E. coli*.

There are several reasons why it is difficult to express soluble mammalian proteins, but two are the main factors that might contribute in this field, the rate of translation and the rate of protein folding, which are almost an order of magnitude faster in *E. coli* as compared with eukaryotic systems (Widmann M., 2000). Although eukaryotic expression hosts are sometimes able to overcome these problems, they are not without their own difficulties in terms of ease of use, time, cost and experimental flexibility.

These problems in protein production have led to significant research into ways of enhancing the production of soluble proteins using currently available expression hosts. In the end, *E.coli* has significant benefits of cost, ease of use and scale, all of which make it essential to find ways to overcome the difficulty of generating soluble heterologous proteins in *E.coli*.

Improving the solubility of recombinant proteins in *E.coli* commonly involves some changing of the expression conditions. Factors such as reduced temperature (Hammarstrom M.,2002), changes in the *E.coli* expression strain (Miroux B., 1996), different promoters or induction conditions (Qing G., 2004), and co-expression of molecular chaperones and folding modulators (De Marco A. et al., 2004) have all been examined and in some specific cases lead to enhancements of soluble protein production. However, the improvements can be minimal and in many cases, none of these factors will solve the problem and proteins will be expressed in insoluble inclusion bodies when overproduced in *E.coli* (Chayen N.E.,2004). Overall, the production rate is strongly dependent on the individual sequence of the antibody. Although some efforts have been directed towards finding alternate expression conditions that can assist in making proteins soluble, the majority of the work in the field has focused on the discovery, development and refinement of solubility fusion tags.

As the name implies, these tags are proteins or peptides that are fused to the protein of interest and, in the best case, help to properly fold their partners leading to enhanced solubility in the protein of interest. This strategy of fusing a protein partner is effective in promoting the solubility of many target proteins (Esposito and Chatterjee, 2006).

3.1.2. Enhanced proteins productivity and solubility by fusion partners

Cytoplasmic expression using fusion partners such as glutathione S-transferase (GST), thioredoxin (Trx), and maltose-binding protein (MBP) generally results in high productivity and enhanced solubility (LaVallie E.R. and McCoy J.M., 1995). It was previously reported that MBP-fused scFvs are expressed at high levels in the cytoplasm of *E.coli* as soluble and active proteins (Bach H. et al., 2001). Zheng et al. (2003) showed that a fusion method with *E.coli* protein NusA, which had been identified to have a highest solubility when over-expressed in the cytoplasm of *E.coli* cells, has potential as a valuable tool for overproduction of soluble and functional scFvs (Jurado P. et al., 2006). Both improved production and correct folding of scFvs have been observed when scFvs are expressed as Trx fusion proteins in the cytoplasm of the *E.coli* FA113 cells. In the contest of a fusion protein, these fusions partner seem to function as molecular chaperones that promote the solubility and stability of scFv, that is fused to them.

It was further proposed that stability is probably the limiting factor for folding of an scFv in the cytoplasm. Since the intradomain disulphides contribute about 4-6 kcal/mol to the stability of antibody domains, (Frisch et al., 1994; Frisch et al.,1996; Woern A., 2001) antibodies fragments expressed in a reducing environment as cytoplasm are strongly destabilized, and a smaller fraction of these fragments is likely to fold into the correct native structure, possibly accounting for their tendency to aggregate.

Thus fusion partners with chaperone-like activity stabilize scFvs for efficient functional expression in the cell cytoplasm in a soluble, active form and since they do not release the protein that is fused to them, their stabilizing and solubilizing effects persist (Biocca S. et al,1995).

Moreover, recently C.Kyratsous et al.(2009) demonstrated that also bacterial chaperones as DnaK or GroEL placed in frame of a target polypeptide, facilitate production of large amounts of soluble recombinant proteins, that are normally highly insoluble when expressed in bacteria. So the chaperone fusions are soluble and when these fusion proteins are co-expressed with their cognate co-chaperone, solubility is further increased.

3.1.3. Bacterial Chaperones

Chaperones are protein machines essential for the correct folding of proteins in the cell under physiological and stress conditions. As Anfisen (1973) demonstrated in protein folding in vitro, the formation of the native protein from the unfolded state is a spontaneous process determined by the global free energy minimum and by amino acid sequence of the protein.

But in vitro experiments using small single-domain proteins, even with ideal conditions as at low concentrations, long incubation times, in mild conditions of temperature, pH, ionic strength and solvent composition, it is difficult to obtain 100% folding efficiency because of unfavourable side-reactions such as misfolding and aggregation.

In living cells, the conditions for folding are far from ideal. First, for nascent chains and for polypeptides during their translocation through membranes, folding is topologically restricted because folding conditions apply to some, but not all parts of the chain (Rothman, 1989). Secondly, most cells operate at ambient or homeothermically set temperatures (e.g. 37°C) where the hydrophobic effect is stronger and the time available for successful folding is short. Moreover the cytoplasm of a living cell is so incredibly crowded with macromolecules, e.g. the total protein and RNA concentration in an *Escherichia coli* cell is ≈ 350 mg ml^{-1}(Zimmerman and Trach, 1991), that protein aggregation is a serious problem.

Nature has endowed cells with complex and elegant protein molecular machines, referred to as heat shock proteins or chaperones, many of which are ATP dependent, that function to assist protein folding and to reverse or inhibit misfolding and aggregation.

E.coli cells have evolved an elaborate protection system, consisting of a set of molecular chaperones, which prevent aggregation and assist refolding of misfolded proteins (table 3.1).

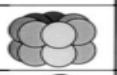
chaperone family	structure	ATP	*E. coli* member	number of active species per cell	null mutant phenotype	function
HSP100	6-mer 	+	ClpA		No phenotype	protein degradation together with the ClpP Protease
			ClpB	500	Impaired thermotolerance	disaggregation of protein aggregates together with the DnaK system
HSP90	dimer 	+	HtpG	1050	Reduced growth rate at 44°C	unknown
HSP70	monomer 	+	DnaK co-chaperone DnaJ/GrpE	9900	Temperature-sensitive growth (39°C)	de novo protein folding prevention of protein aggregation at high temperatures regulation of the heat shock response disaggregation of protein aggregates together with ClpB
HSP60	14-mer 	+	GroEL co-chaperone GroES	1230	lethal	de novo protein folding prevention of protein aggregation at high temperatures
sHSP	8-24-mer 		IbpA IbpB	<600	No phenotype	prevention of protein aggregation at high temperatures
Trigger-factor	monomer 		Trigger-factor	20000	No phenotype	ribosome-associated chaperone de novo protein folding

Table 3.1 Major cytosolic chaperone families of *E.coli*: structure and function.
(Adapted from Schlieker C., 2002)

H olded native protein. In proteins partially unfolded or de novo synthesized these surfaces are exposed and can interact

with other hydrophobic constituents of the crowded intracellular environment (Baneyx and Mujacic, 2004) accumulating as inclusion bodies aggregates.

Bacterial chaperones, binding transiently to exposed hydrophobic surfaces of a significant fraction of de novo synthesized or misfolded polypeptides, stabilize their substrate and facilitate conformational processing 1) by preventing non productive hydrophobic interactions resulting in aggregation of unfolded polypeptides and 2) by helping them acquire their correct tertiary conformation assisting in the correct refolding of chaperone-bound denatured polypeptides (Fink,1999; Hartl and Hayer-Hartl, 2002). The release of the substrate is often coupled to an ATP-directed conformational change in the chaperone. Following release, the substrate can proceed to fold correctly, or rebind to the chaperone again until a native conformation is reached. 3) Moreover an other activity of molecular chaperones is catalytic solubilization and refolding of stable protein aggregates. 4) They may maintain newly synthesized proteins in an unfolded conformation suitable for translocation across membranes and bind to nonnative proteins during cellular stress.

Chaperones are catalysts in the sense that they transiently interact with their substrate proteins but are not present in final folded product, and also in that they increase the yield of folded protein.

E.coli employs two major chaperone systems, the DnaK (Hsp70) system, composed of DnaK and its co-chaperones DnaJ and GrpE, and the GroEL (Hsp60)/GroES system (fig.3.1) (Fink, 1999; Hartl and Hayer-Hartl, 2002). Expression of these proteins is positively regulated under elevated temperatures or other forms of cellular stress that affect protein conformation (Hoffmann and Rinas,2004).

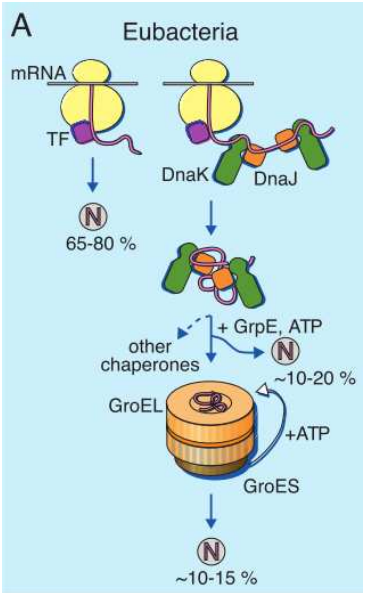


Fig. 3.1 Model for the chaperone-assisted folding of newly synthesized polypeptides in the cytosol of Eubacteria. TF, trigger factor; N, native protein. Nascent chains probably interact generally with TF, and most small proteins (65 to 80% of total) fold rapidly upon synthesis without further assistance. Longer chains (10 to 20% of total) interact subsequently with DnaK and DnaJ and fold upon one or several cycles of ATP-dependent binding and release. About 10 to 15% of chains transit the chaperonin system—GroEL and GroES—for folding. GroEL does not bind to nascent chains and is thus likely to receive an appreciable fraction of its substrates after their interaction with DnaK. (Adapted from Hartl and Hayer-Hartl, 2002).

Since protein aggregation is frequently observed upon synthesis of heterologous proteins in *E.coli*, molecular chaperones commonly have been applied in biotechnology by their co-overproduction

with the desired protein to increase the yield of properly folded, recombinant proteins in bacterial cell factories. In fact insufficient availability of molecular chaperones is observed as a major bottleneck for proper protein folding in recombinant proteins production.

Over-expression of recombinant proteins in the crowded milieu of the *E.coli* cytoplasm, where transcription and translation are tightly coupled, makes folding an extraordinary challenge. Expression of complex and heterologous proteins frequently results in kinetically trapped, slow-folding, non-productive intermediates that are prone to aggregate (Baneyx and Mujacic, 2004) and are usually deposited as dense refractile particles called inclusion bodies or degraded by bacterial proteolytic systems (Baneyx and Mujacic, 2004). Proteins deposited into inclusions bodies require complex solubilization and refolding procedures to gain biological activity.

Finally, small chaperones in fusion with scFvs or insoluble proteins have stabilizing and solubilizing effects persistent, since they do not release the protein that is fused to them. During expression in *E.coli*, fusion chaperones emerges from the ribosome and may fold before the fusion partner linked at its C terminus has been translated in full. It is at that stage where chaperone can presumably bind to its fusion partner and shield it from interacting with other cell constituents that would promote its aggregation mainly by interacting with the solvent-exposed hydrophobic residues. That would account for the increased production yield of chaperone fusion protein and further protects it from denaturation, staying attached to its fusion partner also after correct folding (Kyratsous et al. 2009).

3.1.4. DnaK chaperone

DnaK, homolog of the eukaryotic Hsp70, is the major cytosolic chaperone in *E.coli*, and plays an important role in the control of conformational quality. In fact, DnaK is involved in different activities such as prevention of aggregation, folding and refolding of misfolded species and protein disaggregation (Gragerov A. et al., 1992; Langer T. et al, 1992; Mogk A. et al., 2003; Schlieker C. 2004; Thomas JG., 1996). For this reason, DnaK has often been used in co-production approaches, either together with its co-chaperone DnaJ or both with DnaJ and their nucleotide exchange factor GrpE to minimize aggregation and to enhance solubility of the recombinant protein.

Structurally, DnaK consists of a ~44-kD NH 2 -terminal ATPase domain and a ~27-kD COOH-terminal peptide-binding domain (Bukau B., 1998) . The peptide-binding domain is divided into a β-sandwhich subdomain with a peptide-binding cleft and an α-helical latchlike segment (Zhu X., 1996). The target peptide interacts with a deep hydrophobic pocket of the β-sandwich domain. Target peptides are ~seven-eight residues long (NRLLLLTG) and have an interior

hydrophobic core (with leucine and isoleucine residues being preferred by DnaK) flanked by basic residues (Gragerov et al., 1994). These binding sites occur statistically every ~40 residues in proteins and are recognized with affinities of 5 nM to 5 μM .

Rapid peptide binding occurs in the ATP-bound state of DnaK in which the α-helical latch over the peptide-binding cleft is in an open conformation (Fig. 3.2A). Stable holding of peptide involves closing of the latch, a conformational change that is achieved by hydrolysis of bound ATP to adenosine-diphosphate (ADP). The cycling of DnaK between these states is regulated by co-chaperone DnaJ (41kD) and by nucleotide exchange factor GrpE, a homodimer of 20-kD subunits (Bukau B., 1998). The NH 2 -terminal J domain of DnaJ binds to DnaK and accelerates hydrolysis of ATP by DnaK, thus facilitating peptide capture (ref). The COOH-terminal domain of DnaJ (and of other Hsp40s) functions as a chaperone in recognizing hydrophobic peptides and can thus recruit DnaK to nascent chains (Harrison C.J.,1997). GrpE induces the release of ADP from DnaK , thus promotes the high- to low-affinity transition (Packschies et al., 1997), and upon rebinding of ATP the DnaK-peptide complex dissociates, completing the reaction cycle (Fig. 3.2B).

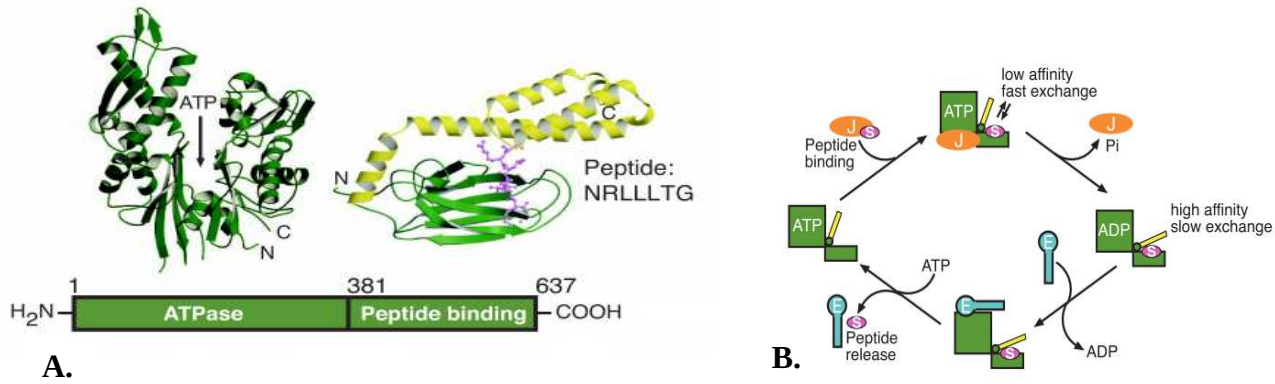


Fig. 3.2 Structure and function of chaperones with the ability to bind nascent chains. **A**. Structures of the ATPase domain and the peptide-binding domain of Hsp70 shown representatively for *E.coli* DnaK. The α-helical latch of the peptide binding domain is shown in yellow and a ball-and-stick model of the extended peptide substrate in pink. ATP indicates the position of the nucleotide binding site. The amino acid sequence of the peptide is indicated in single-letter code (D, Asp; E, Glu; G, Gly; L, Leu; N, Asn; R, Arg; T, Thr; and V, Val). **B**. Simplified reaction cycle of the DnaK system with DnaK colored as in A. J, DnaJ; E, GrpE; S, substrate peptide. GrpE is drawn to reflect the extended shape of the protein. Not all substrates are presented to DnaK by DnaJ. The intermediate DnaK-DnaJ-substrate-ATP is probably very transient, as this is the fastest step of the cycle. (Adapted from Hartl and Hayer-Hartl, 2002).

Repeated cycles of substrate binding and release promote the folding, assembly, translocation and proteolysis of proteins in *E.coli* cells (Bukau and Horwich, 1998; Ellis and Hartl, 1999; Agashe and Hartl, 2000).

Whereas some chains transit DnaK with half-lives of less than 1 min, consistent with rapid folding upon completion of synthesis, other newly synthesized proteins are released from DnaK

slowly, with half-lives of 10 min or more, facilitating the posttranslational folding of multidomain proteins (14-90 kDa) through cycles of binding and release.

Thus, DnaK recognizing and transiently as monomer binding to exposed hydrophobic regions common to most nascent chains or misfolding intermedies, assistes protein folding co- and post-translationally both as a delaying device to prevent premature folding before a complete folding unit has been synthesized, and as an anti-aggregation device to prevent interaction with hydrophobic residues on adjacent nascent chains. Moreover, (as activation of the cellular stress response,) DnaK mediates disaggregation and refolding of protein aggregates by local folding/unfolding and is also involved in the degradation of aggregation-prone polypeptides by targeting them to proteases such as Lon and ClpP (Fig.3.3).

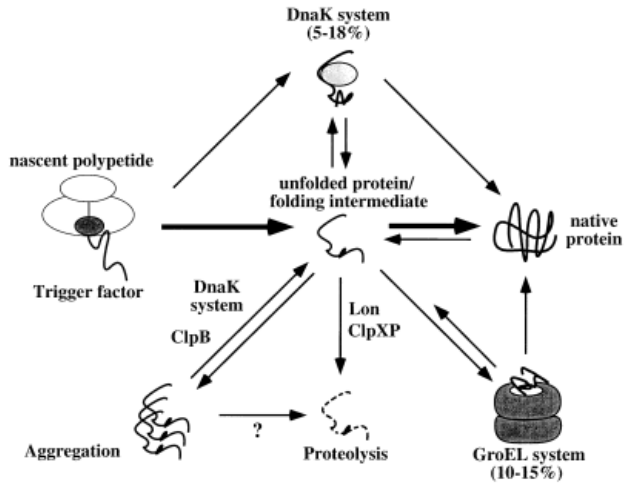


Fig. 3.3. The life cycle of a protein is guided by molecular chaperones. A polypeptide is synthesized at the ribosome and folds to its native tertiary structure. Folding intermediates, which are prone to aggregation, can be protected by association with molecular chaperones, which outcompete off-pathways of the folding process. TF is associated with the large subunit of the ribosome and interacts with virtually all nascent chains as they emerge from the ribosome. 5–18% and 10–15% of newly synthesized proteins interact with the DnaK or GroEL systems, respectively, as calculated from co-immunoprecipitation experiments with [35S]methionine-labeled spheroblasts. Aggregated species that escaped the cellular protection machinery can be rescued by the ClpB/DnaK bi-chaperone system. Nonnative species which fail to fold correctly can also be degraded, mainly by the ClpXP and Lon protease.(Adapted from Schlieker C., 2002)

Moreover, recently C.Kyratsous et al (2009) demonstrated that bacterial chaperones as DnaK or GroEL placed in frame of a target polypeptide facilitate production of large amounts of soluble recombinant proteins that are normally highly insoluble when expressed in bacteria.

3.1.5. A-synuclein

A-Synuclein is a highly conserved protein predominately expressed in the presynaptic terminals of central nervous system in brain, particularly in the neocortex, hippocampus, striatum, thalamus and cerebellum (Nakajo, S., 1990). It is also expressed in hematopoietic cells (Hashimoto et al. 1997; Shin et al. 2000) and in other tissues, such as the heart, skeletal muscle, pancreas, and placenta, but it is less abundant than in the brain. It is a synuclein family member with β-Synuclein, γ-Synuclein and synoretin, and they are not found outside vertebrates, although they have some conserved structural similarity with plant ‘late-embryo-abundant’

proteins. A-synuclein is the major filamentous component of intracytoplasmic Lewy bodies (Clayton & George, 1998), neuronal inclusion bodies in the substantia nigra, that are a pathological hallmark of Parkinson's disease (PD), a neurodegenerative disorder associated with dopaminergic nerve cell loss, but Lewy bodies are also accumulated in brains with dementia and multiple system atrophy and in some cases of Alzheimer's disease (Spillantini M.G. et al., 1997 and 1998).

Although the normal physiological functions of α-synuclein has not yet been established, several lines of evidence suggest a role in membrane-associated processes at the presynaptic terminal. A-synuclein has been implicated to play a critical role in synaptic events, such as neuronal plasticity during development, learning, and degeneration under pathological conditions (George J. M. et al., 1995; Stefanis L.,2001).

Moreover, α-synuclein has gested to function as a chaperone protein in vivo because it appears to bind many cellular proteins. In particular, α-synuclein shares regions of homology with 14-3-3 proteins (Ostrerova, N.,1999), which are a family of ubiquitous cytoplasmic chaperones, and binds to 14-3-3 proteins as well as to the ligands of 14-3-3 including PKC, BAD, and ERK (Skoulakis E.M.,1998). Recently, the chaperone activity of α-synuclein in vitro has been demonstrated by two research groups.

More importantly, α-synuclein is overexpressed under stress conditions and it is able to prevent the thermally and chemically induced aggregation of substrate proteins.

A-Synuclein is extremely heat-resistant and globally it possesses poorly ordered structure under physiological conditions in vitro, and although namely as natively unfolded protein under several condition the conformational behaviour is most sensitive to is environment, perhaps explaining its ability to interact with many other proteins or ligands (Cole N.B., 2002). Presence of several metals, lipids, α–synuclein is able to forms several oligomers amorphous aggregates and also organized amyloid fibrils (Davidson et al., 1998).

Fig. 3.4 Structure of α-synuclein. A. Subdivision in three regions: amphipathic N-terminal, the central hydrophobic NAC and the acidic C-terminal. B. A-synuclein, with very little secondary structure in aqueous solution, associates with small acidic phospholipid vesicles acquires increased level of secondary structure (Davidson et al., 1998).

A.

B.

Structurally, α-synuclein is an acidic protein of 140 amino acids with a molecular weight of 14460 Da and consists of three distinct regions, the amphipathic N-terminal, the central hydrophobic NAC region and the C-terminal region (Fig.3A).

The N-terminal domain of α-synuclein (residues 1-65) contain six repeats of an unusual 11-amino-acid imperfect repeat that display variation of a KTKEGV consensus motif, and is unordered in solution, but can schift to an α-helical conformation with amphipathic α-helixes that are reminiscent of the lipid-binding domain of apolipoproteins (Nielsen M.S., 2001). Consistent with these structural elements, α-synuclein binds strongly to negatively charged phospholipids and become α-helical, suggesting that the protein may normally be membrane associated (Fig. 3.B).

The central region of α-synuclein (NAC residues 66-95), is the second major component of brain amyloid plaques in Alzheimer's disease (Ueda K.,1993). These region contain three imperfect repeats and comprises the highly amyloidogenic part of the molecule that is responsible for α-synuclein ability to undergo a conformational change form random coil to β-sheet structure and to form Aβ-like protofibrils and fibrils.

The acidic C-terminal region (residues 96-140) of α-synuclein has no recognized structural elements but has a strong negative charge composed primarily of acidic amino acids. It is organized in random structure in most conditions and contains an acid domain also rich in proline residues (residues 125-140) that appears critical for the chaperone-like activity of α-synuclein (Park et al, 2002). Structural features and the conformational change of α-synuclein led us to the hypothesis that α-synuclein may bind to hydrophobic regions of partly unfolded proteins by stresses and act as a molecular chaperone. A study showed that the removal of the C-terminal acid tail of α-synuclein abolished its chaperone activity. Other data indicate that the C-terminal acid tail

was indeed necessary but not sufficient for the chaperone function of α-synuclein (Park et al, 2002). The acidic tail itself does not have chaperone activity, and does not appear to interact with the substrate protein. It is highly likely that the role of introduced acid tail is to increase protein solubility by electrostatic repulsions. It is well demonstrated that the solubility of a protein is approximately proportional to the square of the net charge on the protein. It showed that the N-terminal region of α-synuclein plays a crucial role in the efficiency of the chaperone function, binding the substrate protein and forming a soluble high molecular weight (HMW) complex. So N-terminal-binding domain governs the efficiency and the substrate specificity of chaperone proteins forming HMW complex, whereas the C-terminal acid tail solubilizes the HMW during the chaperone action (Park et al, 2002).

A-synuclein can act as a molecular chaperone in its native state in vitro, and conformational change by organic solvent or metal ions (zinc ions) induces the aggregation of α-synuclein, while heat-incubated α-synuclein enhances the chaperone like function.

These results suggest that structural conformational change of α-synuclein might explain the aggregation property (aggregation kinetics of α-synuclein) and loss of chaperonic-like activity of α-synuclein, which might be related to the formation of Lewy body fibrils in Parkinson's disease.

These data together with its physic-chemical characteristics, as a small protein dimension, that contributes at low molecular weights in fusion protein, high expression levels in *E.coli* (30% of total *E.coli* proteins), its facilitated purification due its high resistance to protease and to high temperature, make α-synuclein a good candidate as fusion partner.

3.2. Materials and methods

3.2.1. Bacterial strains

The bacterial strains used in the present study are shown in Table 3.2.

Table 3.2 Bacterial strains used in this study

Strains
E. coli host
Genotype of :
Topl0F (F-*mcr*A Δ(*mrr-hds*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1*ara*D139 Δ(*ara-leu*) 7697*gal*U *gal*K *rps*L(strR)*end*A1 *nup*G).
invaF' F´ *end*A1 *rec*A1 *hsd*R17 (rk-, mk+) *sup*E44 *thi*-1 *gyr*A96 *rel*A1 φ80*lac*ZΔM15Δ(*lac*ZYA-*arg*F)U169 λ-
BL21(DE3) (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3)
BL21(DE3) Lys S (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) pLysS (CamR)
Topl0, invaF' : suitable for cloning, transformation, propagation, and maintenance.
BL21[DE3], BL21[DE3] Lys S : specifically designed for expression of genes regulated by the T7 promoter.
The chaperone plasmid pG-KJE8 (for GroELS and DnaKJE expression), was obtained from Takara (Shiga, Japan) (Nishihara K., 1998).

3.2.2. *E.coli* plasmids construction

All plasmids have been checked for correctness by cleavage with some restriction enzymes in agarose gels and by nucleotide sequencing.

3.2.2.1. α-Syn plasmids

pMut4 syn

The entere coding sequence of human α-synuclein was amplified with following primers 5’ TGGCTAGCATGGATGTATTCATGAAAGGACTT3’

5’TCGGATCCACGCGGGTCAAGTGAGGCTTCAGGTTC3’

and cloned in plasmid pRSETB (Invitrogen) between *NheI* and *HindIII* restriction sites. By site directed mutagenesis the two sites *EcoRI* and *BamHI* were deleted using the following primers

5’ATACTCGAGGATATGCCTGTGGACCCTGACAA 3’

5’TATCCTCGAGTATTCCTTCCTGTGGGGCTCCTTC 3’

A synthetic coding domain was linked at C-terminal of α-synuclein in order to create a site for thrombin cleavage.

pSyn-TAT 8H4

The 8H4 scFv coding sequence was amplified using the oligonucleotide primers

5’CTGAATTCAGCCGCTAGCTGAGGAGACGGTG 3’

5’ CTGGATCCGCAAGCGGCGCGCATGCCGACAT 3’

The PCR product was cloned in plasmid pMut4 syn between the *BamHI* and *EcoRI* restriction sites, to give pSyn-TAT 8H4. In order to increase the production of scFv protein, a prokariotic *E.coli* optimized 8H4scFv nucleotide sequence was in vitro synthetized by Genescript. The plasmid carring this sequence was pUC57 8H4. The 8H4scFv optimized nucleotide sequence was amplified from pUC57-8H4 using the oligonucleotide primers Dannew1-FW 5'AGGGATCCGGCTATGGTCGTAAAAAACGTC 3' and Dannew2-RV 5' GGGAATTCAGCCGCTAGCGCTAGAAACG 3'. The PCR product was cloned between *BamHI* and *EcoRI* restriction sites of plasmid pMut4 syn in order to give plasmid pSyn-TAT 8H4opt.

3.2.2.2. DnaK plasmids

pDnaK

The *dnaK* coding sequence was amplified using plasmid pKJE7 as template and the oligonucleotide primers Dnk2- RV 5' CGTGGGACTAATTCAAATTCAGCGTCGACAAC 3' and Dnk3-FW 5' AGTCTAGAATGGGTAAAATAATTGGTATCG 3' in a first amplification, and primers Dnk3-FW 5' AGTCTAGAATGGGTAAAATAATTGGTATCG 3' dnk3-FW and Dnk1-RV 5' CTCGGATCCGCGTCCGACTAATTCAAATTCAG 3' in a following second PCR. The final PCR product cloned between the restriction sites *NheI* and *BamHI* of plasmid pRSETB in order to give pDnaK. The resulting chaperone protein has a thrombin cleavage site at its C-termini that should permit release of the protein of interest from the fusion.

pDnaK- TAT 8H4

The 8H4scFv optimized nucleotide sequence was cleavage with *BamHI* and *HindIII* restriction sites and cloned in the same restriction sites of plasmid pDnaK in order to give plasmid pDnaK-TAT 8H4.

3.2.3. Growth conditions

E.coli strains were grown aerobically at 37°C or at the indicated temperature in Luria Broth (LB) supplemented, when necessary with ampicillin (100μg ml^{-1}).

3.2.4. Solubility assays

Overnight bacterial cultures were used to inoculate 50 ml of fresh media. When the cultures reached an optical density (OD) of 0.5 at 600 nm the protein expression was induced by adding Isopropyl- β–thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM for 3 h and then

bacteria were harvested and washed twice with phosphate-buffered saline (PBS: 1mM KH_2PO_4, 10 mM Na_2HPO_4, 137 mM NaCl, 2.7 mM KCl, pH7.4). The bacterial pellets were suspended in PBS supplemented with 10 mM 2-β–mercaptoethanol and 0.2 mg/ml lysozyme. After stirring at 4°C for 30 min, Triton X-100 was added to a final concentration of 1% and stirring was continued for an additional 5 min. The lysates were sonicated 3 times for 20 s at 4°C and centrifuged at 12,000 g for 30 min at 4°C. 20 μl of supernatants (soluble lysates) was add to SDS sample buffer (100 mM2-β–mercaptoethanol, 2% SDS, 0.1% bromophenol blue, 10% glycerol, 50 mM Tris HC1 pH 6.8,). The pellets (inclusion bodies) were suspended in an equivalent volume of SDS sample buffer.

3.2.5. α-syn and α-syn TAT8H4 scFv periplasmic purification

Osmotic shock was carried out basically as described by Shevchik et al. (1994). Briefly, after growth and 3h of induction the cell pellet centrifugated (20min 6000 rpm 4°C) from 1 l of culture was resuspended in 100 ml ice-cold osmotic shock buffer (40% sucrose, 2 mM Na_2 -EDTA, 30 mM Tris-HCl, pH 7.2) and incubated for 10 min at room temperature. The pellet collected by centrifugation at 12000 rpm for 20 min was resuspended rapidly with 90 ml ice-cold water, followed by adding 37 μl of saturated $MgCl_2$ solution, and kept on ice for 3 min. The supernatant collected after centrifugation at 12,000 rpm for 20 min contained periplasmic proteins. Periplasmic supernatant added of 100 ml solution (300 mM NaCl, 20 mM Tris HCl pH8) was incubated with equilibrated NTA-Ni^{2+}resin, and purified protein eluted with 250mM imidazole, 300 mM NaCl, , 20 mM Tris HCl pH7.5. The purified proteins amounts were measured with spectroscopy at 280 nm.

3.2.6. Expression and purification of a-synTAT8H4and dnaKTAT8H4 fusion proteins

E. coli strain BL21 (DE3) LysS was transformed with expression plasmids pSyn-TAT8H4 or pDnaK-TAT8H4opt. Bacterial cultures were incubated in Luria Broth (LB) growth medium supplemented with ampicillin (100μg ml^{-1}) at 37°C. When the cultures reached an OD of 0.7 at 600 nm the protein expression was induced by adding IPTG to a final concentration of 0.5 mM for 3 h. The bacteria were harvested by centrifugation at 6000 rpm for 20 min at 4°C.

The Syn-TAT8H4 pellet was resuspended in 4M urea, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8 and sonicated (20 s burst/ 20 cooling x 3 cycles). The soluble and insoluble fraction were separated by centrifugation at 12000 rpm for 30min at 4°C. The inclusion bodies in insoluble fraction were further solubilized in 6M guanidine, 20 mM β-mercaptoethanol, 10 mM Tris/HCl,

100 mM NaH_2PO_4 pH 8 and the insoluble pellets were removed by centrifugation at 12000 rpm for 30min at 4°C. The solubilized inclusion bodies (supernatant) were added to a equivalent volume of 6M guanidine, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8 to decrease β-mercaptoethanol concentration at less of 10 mM. Purification of the protein was carried out in NTA-Ni^{2+}column
After loading, the column was first washed with 10 column volumes of 6M guanidine 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8 and then with 8M urea, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8. The Syn-TAT8H4 fusion protein was eluted in 8M urea, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 4.5.
The DnaK-TAT8H4opt pellet was resuspended in 6M guanidine-HCl, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8 and sonicated (3x 20 sec). The soluble and insoluble fraction were separated by centrifugation at 12000 rpm for 30min at 4°C. The solubilized protein was batch incubated with NTA- Ni^{2+} resin, and loaded in column. After washing with 10 column volumes of 6M guanidine, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8 and then with 8M urea, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 8, the DnaK-TAT8H4 fusion protein was eluted using 8M urea, 10 mM Tris/HCl, 100 mM NaH_2PO_4 pH 4.5. The column fractions were collected and analyzed by SDS-PAGE. The fusion protein titres concentration were measured with spectroscopic analysis at 280 nm.

3.2.7. Refolding of scFv fusion proteins

Refolding was performed according to the method on-column chemical refolding described by Oganesyan et al. 2004 with minor modifications. Briefly, the denaturate purified proteins collected in column fractions, were diluited in buffer A (8M urea, 20 mM Tris, 10 mM β-mercaptoethanol pH8) and incubated with equilibrated NTA-Ni^{2+} resin. After binding, resin was loaded into column and washed with 5 column volumes of buffer A, and then slowly with 10 column volumes of 0.1% Triton X100, 20 mM Tris pH 7.5, 0.5M NaCl. The column was further washed with 10 column volumes of 5 mM β-cyclodextrin, 0.1.M NaCl, 20 mM Tris, pH7.5, and after with 10 column volumes of 0.1 M NaCl, 20 mM Tris, pH 7.5. Refolded proteins were eluted in 500 mM imidazole, 20 mM Tris, pH 7.5, 0.1M NaCl.
The column fractions were collected and analyzed by SDS-PAGE.
The fusion protein titres concentration were measured with spectroscopic analysis at 280 nm.
The "refolding yield" was calculated by equation A_{280R}/A_{280D}. A_{280R} represent absorbance of the refolded solution, at 280 nm after refolding, and A_{280D} represent the absorbance of the denaturated solution at 280 before refolding.

3.2.8. Buffer-exchange by gel filtration

Dnak 8H4 and α-syn 8H4 scFv fusion proteins purified in denaturanting conditions from inclusion bodies were further purified by gel filtration using a PD10 SephadexTM G-25M column equilibrated with 274 mm NaCl, 20% glycerol, 0.1% pluronic acid, 0.02% Tween-80, 10 mM Tris-HCl pH 9.

3.2.9.Immunocytochemistry and Westernblot

Syn TAT-8H4scFv and dnaK TAT-8H4scFv translocations inside cells were evaluated in CHO cells grown on glass coverslips with Ham's F12 medium containing 10% fetal calf serum, penicillin (100U/ml) and streptomycin (100 pg/ml) at 37°C, in a humidified incubator under 5% CO_2 .These cells were treated with 10μg/ml Syn TAT-8H4scFv or with 10μg/ml dnaK TAT-8H4scFv for 2-3 h. After being rinsed twice with PBS, cells were fixed with 2% paraformaldehyde in PBS for 30 min at 4°C, rinsed twice with PBS, and finally were permeabilized with 0.1% Triton X-100 in PBS for 10 min at 4°C. After being rinsed twice with PBS, cells were incubated with antibody-mouse anti-His H1029 1:3000, Sigma, diluted in PBS with 1% BSA overnight at 4°C. After washing in PBS, a rhodamine isothiocyanate (TRITC)-conjugated secondary antibody (Dako, Milano) (1:5000 diluted in PBS with BSA 1%), was added for 1 h at 37°C followed by washing and incubation with nuclear staining Hoechst (1:5000 diluted in PBS) 10 min and washing in PBS. Finally, coverslips were mounted in glycerol for observation of cells using the fluorescence microscope coupled to a digital camera (Leica Imaging Systems). In order to evaluate the half-life of Syn TAT-8H4scFv and dnaK TATproteins, HeLa cells were treated with 10μg/ml of Syn TAT-8H4scFv or 10μg/ml dnaK TAT-8H4scFv and at different times cells were removed from the flask with trypsin, lysed in SDS-PAGE buffer.

3.2.10. SDS–PAGE and Western blot analysis

SDS–PAGE was carried out according to the method of Laemmli, using 10-12% polyacrylamide gels under reducing conditions. A constant current of 25 mA was applied. Proteins were reduced by treatment with 5% 2-β-mercaptoethanol SDS at 100°C for 5 min. The gels were then stained with 0.25% Coomassie brilliant blue R250 and destained with 5% acetic acid and 7% methanol solution or transferred to a polyvinylidene PVDF membrane (Millipore). The membranes were pretreated with 3% bovine serum albumin, the transferred membrane was incubated with a mouse anti-His tag monoclonal antibody (Sigma) as primary antibody and an

alkaline phosphatase-conjugated anti-mouse IgG antibody (Sigma) as secondary antibody; or a HRP-coniugated goat anti-mouse IgG (Perkin-Helmer) was used as secondary antibody, and luminol with 4-p-jodophenol (Sigma-Aldrich) as a staining substrate .

3.2.11. Translocation of prion protein inside the cells

HeLa cells were cultured in DMEM containing 10% fetal calf serum, penicillin (100U/ml) and streptomycin (100 pg/ml) on glass coverslips at 37^0C, in a humidified incubator under 5% CO_2. At approximately 80% confluence, cells were transfected with the expression plasmids $pPrP^C$-GFP or its analog pDoppel-GFP used as control using lipofectamine Plus reagent (Gibco-BRL), following the manufacturer instructions. After 36 h, the transfected Hela cells were treated with 10μg/ml Syn TAT-8H4scFv (diluted in Opti-MEM (GIBCO) for 3 h or 24 h at 37°C . After being rinsed twice with PBS, cells were fixed (30 min 4°C) with 2% paraformaldehyde in PBS rinsed again, and finally were permeabilized (10 min 4°C) using 0.1% Triton X-100 in PBS. After being rinsed twice with PBS, cells were incubated with the desired antibody (mouse anti-his 1:3000, Sigma, diluted in PBS with BSA 1%) overnight at 4°C. After washing in PBS, a rhodamine isothiocyanate (TRITC)-conjugated secondary antibody (Dako, Milano) (1:5000 diluted in PBS with BSA 1%), was added for 1 h at 37° followed by washing in PBS. Finally, coverslips were mounted in glycerol for observation of cells using the fluorescence microscope coupled to a digital camera (Leica Imaging Systems).

3.2.12. Fluorescence Microscopy and Image Analysis

Fluorescence microscopy (excitation at 488 nm; emission at 509 nm) was carried out using an Olympus IMT-2 set-up, equipped with a 12-bit digital CCD videocamera (Micromax, Princeton Instruments). Data was acquired and analyzed using the Metamorph software (Universal Imaging) and elaborated with Photoshop 8.

3.3. Results

3.3.1. Sequence of the 8H4 scFv gene

One of the most important factors that affects the expression of eukaryotic genes in prokaryotic systems is the codon usage for translational proteins (Gustafsson et al., 2004).

Under this aspect, in the codon usage for *E.coli* reported in Table 3.4, we can see that some codon are rare in use; for example AGA (Arg), AGG (Arg), or AUA (Ile) CUA (Leu), CCC (Pro) GGA (Gly) same of these are supply with six tRNA in case of Rosetta Gami *E. coli* (Novagen) strain, prokaryotic system.

Escherichia coli K12 **[gbbct]: 14 CDS's (5122 codons)**

fields: [triplet] [frequency: **per thousand**] ([number])

UUU	19.7(101)	UCU	5.7(29)	UAU	16.8(86)	UGU	5.9(30)
UUC	15.0(77)	UCC	5.5(28)	UAC	14.6(75)	UGC	8.0(41)
UUA	15.2(78)	UCA	7.8(40)	UAA	1.8(9)	UGA	1.0(5)
UUG	11.9(61)	UCG	8.0(41)	UAG	0.0(0)	UGG	10.7(55)
CUU	11.9(61)	CCU	8.4(43)	CAU	15.8(81)	CGU	21.1(108)
CUC	10.5(54)	CCC	6.4(33)	CAC	13.1(67)	CGC	26.0(133)
CUA	5.3(27)	CCA	6.6(34)	CAA	12.1(62)	CGA	4.3(22)
CUG	46.9(240)	CCG	26.7(137)	CAG	27.7(142)	CGG	4.1(21)
AUU	30.5(156)	ACU	8.0(41)	AAU	21.9(112)	AGU	7.2(37)
AUC	18.2(93)	ACC	22.8(117)	AAC	24.4(125)	AGC	16.6(85)
AUA	3.7(19)	ACA	6.4(33)	AAA	33.2(170)	AGA	1.4(7)
AUG	24.8(127)	ACG	11.5(59)	AAG	12.1(62)	AGG	1.6(8)
GUU	16.8(86)	GCU	10.7(55)	GAU	37.9(194)	GGU	21.3(109)
GUC	11.7(60)	GCC	31.6(162)	GAC	20.5(105)	GGC	33.4(171)
GUA	11.5(59)	GCA	21.1(108)	GAA	43.7(224)	GGA	9.2(47)
GUG	26.4(135)	GCG	38.5(197)	GAG	18.4(94)	GGG	8.6(44)

Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%

Table 3.*4* *Escherichia coli* codon usage. (Adapted from NCBI).

To avoid problems in translational efficiency it has therefore been made a synthetic version of TAT 8H4 scFv sequence in which the rare codons have been changed with the most frequent for *E.coli*. The sequences of natural and synthetic TAT 8H4 scFv are reported:

```
Synthetic8H4scFv   1 ATGCGCGGTAGCCATCACCATCACCATCACGGCATGGCACGCGGCTATGGTCGTAAAAAA
Wild-type8H4scFv   1 ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGAGGTTACGGTCGTAAAAAG

                       M  R  G  S  H  H  H  H  H  H  G  M  A  R  G  Y  G  R  K  K

Synthetic8H4scFv  60 CGTCGCCAGCGTCGCCGTGGTGCGCGTCTGGCGGCCTCTGGCGCGCATGCCGATATTCTG
Wild-type8H4scFv  60 GTCGCCAGCGTCGCCGTGGCGCTAGACTCGCAGCAAGCGGCGCGCATGCCGACATTTTG
```

```
                    R  R  Q  R  R  R  G  A  R  L  A  A  S  G  A  H  A  D  I  L

Synthetic8H4scFv120 ATCACCCAGAGCCCGACCACCATGGCAGCGTCTCCGGGTGAAAAAATTACCATTATCTGC
Wild-type8H4scFv120 ATAACTCAGTCTCCAACCACCATGGCTGCATCTCCCGGGGAGAAGATCACTATCATTTGC

                    I  T  Q  S  P  T  T  M  A  A  S  P  G  E  K  I  T  I  I  C

Synthetic8H4scFv180 TCTGCCACGAGCTCTATCAGTCCGAACTATCTGCACTGGTACCAGCAGAAACCGGGCTTT
Wild-type8H4scFv180 ATGGCCACGTCAAGCATAAGTCCCAATTACTTGCATTGGTATCAGCAGAAGCCAGGATTC

                    S  A  T  S  S  I  S  P  N  Y  L  H  W  Y  Q  Q  K  P  G  F

Synthetic8H4scFv240 AGCCCGCGTATTCTGATCTTCGGTACCAGCGATCTGGCATCTGGCGTGCCGGTTCGCTTT
Wild-type8H4scFv240 TCCCCTAGAATCTTGATTTTTGGGACATCCGATCTGGCTTCTGGTGTCCCAGTTCGCTTC

                    S  P  R  I  L  I  F  G  T  S  D  L  A  S  G  V  P  V  R  F

Synthetic8H4scFv300 AGTGGCCGTGGTAGCGGCACCAGCTATTCTCTGACGATTGTGAGCATGGAAGCAGAAGAT
Wildtype8H4scFv300 AGTGGCCGTGGGTCTGGGACCTCTTACTCTCTCACAATTGTTTCCATGGAGGCTGAAGAT

                    S  G  R  G  S  G  T  S  Y  S  L  T  I  V  S  M  E  A  E  D

Synthetic8H4scFv360GTTGCGACCTATTACTGCCAGCAGGGTAGTAGCACCCCGCTGACGTTCGGTACCGGCACG
Wildtype8H4scFv360 GTTGCCACTTACTACTGCCAGCAGGGTAGTAGTACACCACTCACGTTCGGTACTGGGACC

                    V  A  T  Y  Y  C  Q  Q  G  S  S  T  P  L  T  F  G  T  G  T

Synthetic8H4scFv420 AAACTGGAAATTAAACGCAGTGGCGGTAGCACCATCACGTCTTACAATGTGTATTACACG
Wildtype8H4scFv420  AAGCTGGAAATAAAACGTTCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACG

                    K  L  E  I  K  R  S  G  G  S  T  I  T  S  Y  N  V  Y  Y  T

Synthetic8H4scFv480 AAACTGTCTAGTAGCGGCACCGAAGTGATGCTGGTTGAAAGTGGCGGTGCGCTGGTTCAG
Wildtype8H4scFv480  AAGTTATCCTCGAGCGGTACCGAAGTGATGCTGGTGGAGTCTGGGGGAGCCTTAGTGCAG

                     K  L  S  S  S  G  T  E  V  M  L  V  E  S  G  G  A  L  V  Q

Synthetic8H4scFv540 CCGGGCGGTTCTCGTAAACTGAGTTGTGCCGCAAGCGGTTTTACCTTCTCTAGTTTTGGC
Wildtype8H4scFv540  CCTGGAGGGTCCCGGAAACTCTCCTGTGCAGCCTCTGGATTCACTTTCAGTAGCTTTGGA

                    P  G  G  S  R  K  L  S  C  A  A  S  G  F  T  F  S  S  F  G

Wildtype8H4scFv600 ATGCATTGGGTGCGCCAGGCGCCGGAAAAAGAACTGGAATGGGTTGCCTATATTAGTCGT
Wildtype8H4scFv600 ATGCACTGGGTTCGTCAGGCTCCAGAGAAGGAACTGGAGTGGGTCGCATATATTAGTCGT

                    M  H  W  V  R  Q  A  P  E  K  E  L  E  W  V  A  Y  I  S  R

Wildtype8H4scFv660 GGTAGCTCTACCGGCTACTGCGCCGATACGCTGAAAGGTCGCTGTACCATCAGCCGTGAT
Wildtype8H4scFv660 GGTAGTAGTACCGGCTACTGTGCAGACACATTGAAGGGCCGATGCACCATCTCCAGAGAC

                    G  S  S  T  G  Y  C  A  D  T  L  K  G  R  C  T  I  S  R  D

Wildtype8H4scFv720 AACCCGAAAAATACGCTGTTTCTGGAAATGACCTCTCTGCGCAGTGAAGATACGGCCATG
Wildtype8H4scFv720 AATCCCAAGAACACCCTGTTCCTGGAAATGACCAGTCTAAGGTCTGAAGACACGGCCATG
                    N  P  K  N  T  L  F  L  E  M  T  S  L  R  S  E  D  T  A  M

Wildtype8H4scFv780 TATTACTGTGCACGTAGCGGCAGTAGCTATGGCGGTTGGTACTTCGATGTG
Wildtype8H4scFv780 TACTACTGTGCAAGATCGGGTAGTAGCTACGGGGGGTGGTATTTCGATGTC

                    Y  Y  C  A  R  S  G  S  S  Y  G  G  W  Y  F  D  V

Wildtype8H4scFv840 TGGGGTGCGGGCACCACGGTGACCGTTTCTAGCGCTAGCGGC
```

```
Wildtype8H4scFv840 TGGGGCGCAGGGACCACGGTCACCGTCTCCTCAGCTAGCGGC
                   W  G  A  G  T  T  V  T  V  S  S  A  S  G
```

3.3.2. Construction of fusion vector pDnaK and pSYNmut4 and expression in *E.coli*

Preliminary experiments, in our laboratory, of 8H4 scFv protein expression under the control of strong promoter T7 in *E.coli* BL21[DE3] Lys S, showed low expression with protein present only in insoluble fraction. Therefore we have tried to take advantage from the use of fusion protein technology to increase the efficiency of translation together with the chaperones activities contained in α-synuclein and DnaK. Therefore we have made two expression vectors where the scFv 8H4 was fused in frame to α-synuclein (pSYNmut4) or DnaK (pDnaK). At this purpose, in order to over express the proteins, *α-syn* or the *dnak* genes were cloned into pRSETB under control of inducible T7 promoter.

SDS-PAGE analyses revealed that a protein of about 70 kDa corresponding to DnaK (theoretical molecular weight of 68996.8 Da) was strongly expressed in soluble cell fraction after 3h incubation with IPTG, so as an abundant band with an apparent molecular mass of 17 kDa corresponding to α-synuclein (theoretical molecular weight of 14460 Da) was detected in soluble cell fraction after IMAC purification (Fig.3.5). A-syn and DnaK were expressed from identical plasmid vectors to minimize any effects from expression in different backgrounds

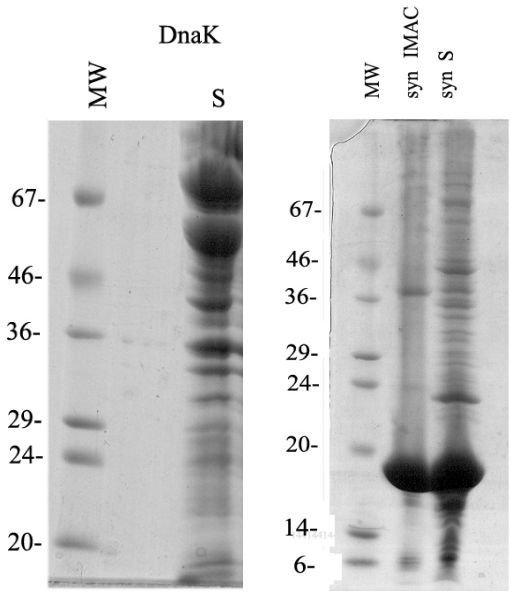


Fig. 3.5 BL21[DE]LysS cells were transformed with the indicated plasmids and protein expression was induced with IPTG. Cell proteins after SDS-lysis buffer were analyzed by SDS-PAGE and stained with Comassie Brilliant Blue.

3.3.3. Construction of fusion vector αsyn-TAT-8H4 scFv and dnaK TAT8H4 scFv and expression in *E.coli*

After that the fusion partners showed an abundant soluble expression, the TAT 8H4scFv sequence was cloned in frame at C-terminal of α-synuclein or DnaK respectively (fig. 3.6).

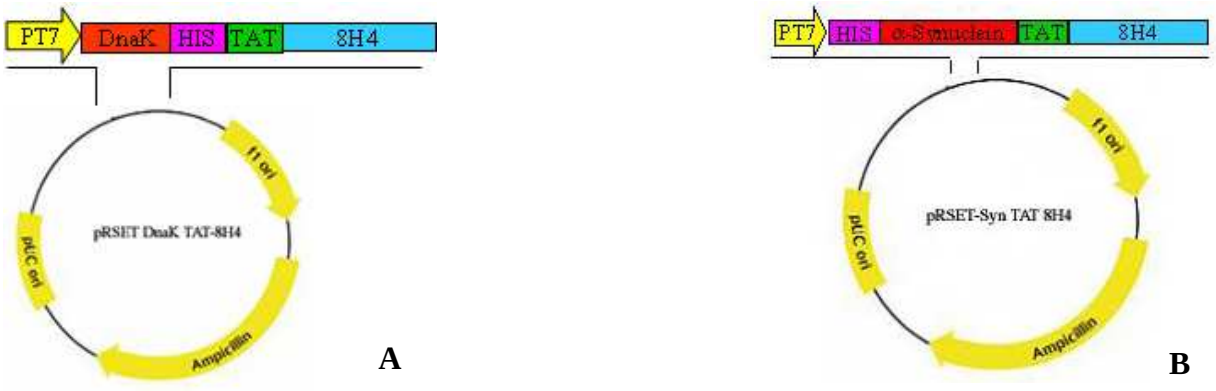


Fig.3.6 pDnaK TAT8H4 (A) and pSyn TAT8H4 (B) vectors, resulting by cloning of TAT8H4 sequence in frame at C-terminal of DnaK in pDnaK, and of a-syn in pSynMut4, respectively. His tag was cloned to facilitate the purification.

The sequences below correspond to recombinant fusion proteins, respectively:

A. DnaK TAT8H4

```
        10         20         30         40         50         60
MGKIIGIDLG TTNSCVAIMD GTTPRVLENA EGDRTTPSII AYTQDGETLV GQPAKRQAVT

        70         80         90        100        110        120
NPQNTLFAIK RLIGRRFQDE EVQRDVSIMP FKIIAADNGD AWVEVKGQKM APPQISAEVL

       130        140        150        160        170        180
KKMKKTAEDY LGEPVTEAVI TVPAYFNDAQ RQATKDAGRI AGLEVKRIIN EPTAAALAYG

       190        200        210        220        230        240
LDKGTGNRTI AVYDLGGGTF DISIIEIDEV DGEKTFEVLA TNGDTHLGGE DFDSRLINYL

       250        260        270        280        290        300
VEEFKKDQGI DLRNDPLAMQ RLKEAAEKAK IELSSAQQTD VNLPYITADA TGPKHMNIKV

       310        320        330        340        350        360
TRAKLESLVE DLVNRSIEPL KVALQDAGLS VSDIDDVILV GGQTRMPMVQ KKVAEFFGKE

       370        380        390        400        410        420
PRKDVNPDEA VAIGAAVQGG VLTGDVKDVL LLDVTPLSLG IETMGGVMTT LIAKNTTIPT

       430        440        450        460        470        480
KHSQVFSTAE DNQSAVTIHV LQGERKRAAD NKSLGQFNLD GINPAPRGMP QIEVTFDIDA
```

```
        490         500         510         520         530         540
DGILHVSAKD KNSGKEQKIT IKASSGLNED EIQKMVRDAE ANAEADRKFE ELVQTRNQGD

        550         560         570         580         590         600
HLLHSTRKQV EEAGDKLPAD DKTAIESALT ALETALKGED KAAIEAKMQE LAQVSQKLME

        610         620         630         640         650         660
IAQQQHAQQQ TAGADASANN AKDDDVVDAE FELVPRGSHM RGSHHHHHHG MARGYGRKKR

        670         680         690         700         710         720
RQRRRGARLA ASGAHADILI TQSPTTMAAS PGEKITIICS ATSSISPNYL HWYQQKPGFS

        730         740         750         760         770         780
PRILIFGTSD LASGVPVRFS GRGSGTSYSL TIVSMEAEDV ATYYCQQGSS TPLTFGTGTK

        790         800         810         820         830         840
LEIKRSGGST ITSYNVYYTK LSSSGTEVML VESGGALVQP GGSRKLSCAA SGFTFSSFGM

        850         860         870         880         890         900
HWVRQAPEKE LEWVAYISRG SSTGYCADTL KGRCTISRDN PKNTLFLEMT SLRSEDTAMY

        910         920         930
YCARSGSSYG GWYFDVWGAG TTVTVSSASG
```

Fig. 3.7 Sequence of fusion protein DnaK TAT8H4 scFv. In pink is showed His tag, in red α-syn sequence, in green TAT sequence, in black 8H4 scFv sequence.

DnaKTAT8H4 scFv	properties
number of amino acids	390
molecular weight	100526.2
pI	5.66
ε	0.65

Table 3.5 Some properties of fusion protein DnaK TAT8H4 scFv

B. Syn TAT8H4

```
        10          20          30          40          50          60
MRGSHHHHHH GMASMDVFMK GLSKAKEGVV AAAEKTKQGV AEAAGKTKEG VLYVGSKTKE

        70          80          90         100         110         120
GVVHGVATVA EKTKEQVTNV GGAVVTGVTA VAQKTVEGAG SIAAATGFVK KDQLGKNEEG

       130         140         150         160         170         180
APQEGILEDM PVDPDNEAYE MPSEEGYQDY EPEASLVPRG SGYGRKKRRQ RRRGARLAAS

       190         200         210         220         230         240
GAHADILITQ SPTTMAASPG EKITIICSAT SSISPNYLHW YQQKPGFSPR ILIFGTSDLA
```

```
       250         260         270         280         290         300
SGVPVRFSGR GSGTSYSLTI VSMEAEDVAT YYCQQGSSTP LTFGTGTKLE IKRSGGSTIT

       310         320         330         340         350         360
SYNVYYTKLS SSGTEVMLVE SGGALVQPGG SRKLSCAASG FTFSSFGMHW VRQAPEKELE

       370         380         390         400         410         420
WVAYISRGSS TGYCADTLKG RCTISRDNPK NTLFLEMTSL RSEDTAMYYC ARSGSSYGGW

       430
YFDVWGAGTT VTVSSASG
```

Fig. 3.8 Sequence of fusion protein α-syn TAT8H4 scFv. In pink is showed His tag, in red α-syn sequence, in green TAT sequence, in black 8H4 scFv sequence.

A-synTAT8H4 scFv	properties
number of amino acids	438
molecular weight (Da)	46480.1
pI	8.54
ε (M^{-1} cm^{-1})	1.20

Table 3.6 Some properties of fusion protein α-syn TAT8H4 scFv

SDS-page analyses revealed a large band of expected molecular weight about 46 kDa for α-syn TAT8H4 (theoretical 46480 Da) and about 100 kDa for DnaK TAT8H4 (theoretical 100526 Da) (Fig. 3.9). Since we have not found differences in α-syn TAT8H4 expression using *E.coli* cells transformed with TAT8H4 wild-type eukaryotic sequence (pSyn TAT8H4) or TAT8H4 optimized sequence (pSyn 8H4opt), has been used TAT8H4 optimized sequence in all following experiments. In any case fusions of 8H4 scFv to α-synuclein and DnaK result in expression and accumulation of significant levels of antibody respect of low levels of TAT8H4 scFv (data not shown).

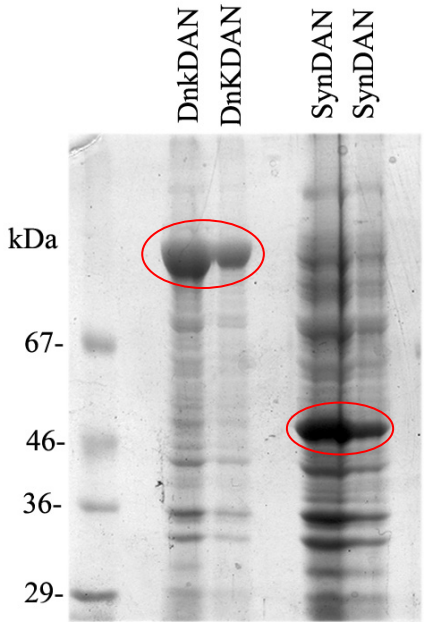


Fig. 3.9 BL21[DE]LysS cells were transformed with the indicated plasmids and protein expression was induced with IPTG. Cell proteins after SDS-lysis buffer were analyzed by SDS-PAGE and stained with Comassie Brilliant Blue.

Although the unfusion partner shows high solubility in several buffer, the fusion counterpart shows to be largely insoluble.

To increase the solubility of fusion proteins different conditions have been tested. Because of, in general, cultivating recombinant *E. coli* cells at low temperatures increases solubility and stability by preventing aggregation of recombinant proteins, Bl21(DE) LysS cells expressing DnaK or *a*-syn TAT8H4 scFv fusion proteins, were grown and induced at 25° or 37 °C in LB medium.

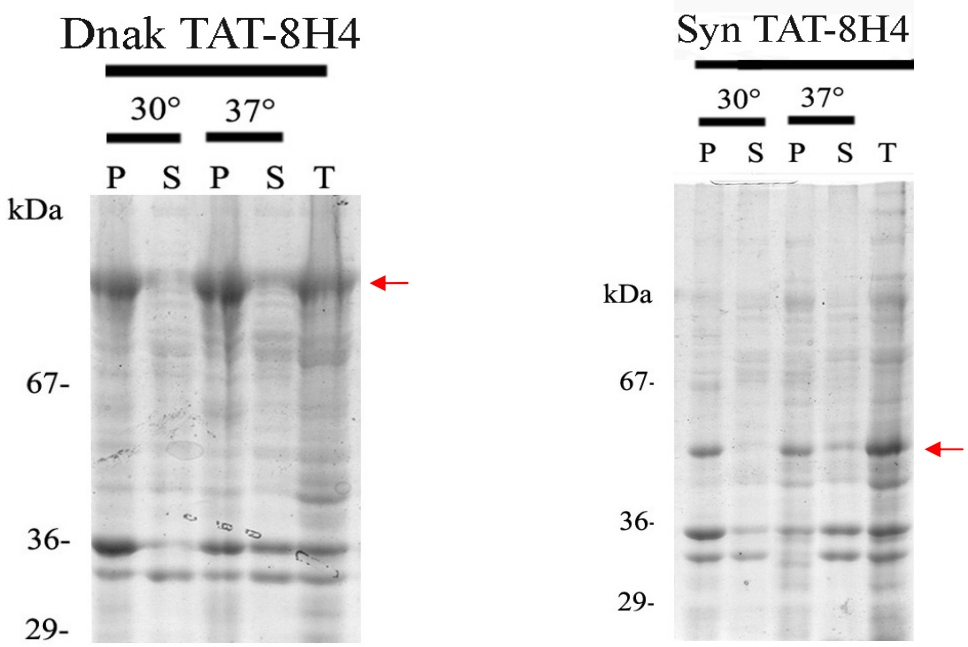


Fig. 3.10 TAT 8H4-scFv partition in the insoluble fraction when expressed as chaperone or chaperon-like fusion. Bl21(DE) LysS cells were transformed with the indicated plasmids and protein expression was induced with 0.5 mM IPTG at 25°C or 37°C. Cells lysates were prepared and proteins were fractionated, analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue. Soluble fraction (S), Insoluble fraction (I), Total cell extract (T) .

Different solubilisation buffers are been tested in solubility and purification assays, as 1)PBS, or 2) PBS supplemented with 10 mM β-mercaptoethanol, 0.2 mg/ml lysozyme, 1% Triton X-100, or 3), 274 mM NaCl, 20% glycerol, , 0,1% pluronic acid, 0.02% Tween-80, 10mM Tris pH 8. Nevertheless the fusion proteins solubility was not significantly improved, as it can see in an example (Fig. 3.10). To confirm, maxi purification of the fusion scFvs in native conditions were resulted in not significant amounts of proteins purified, suggesting the formation of inclusion bodies inside the bacterial cytoplasm.

3.3.4. αsyn TAT8H4 and dnaK TAT8H4 scFvs co-expression together DnaKJE and GroELS chaperones in *E.coli*

Since the behaviours of the fusion proteins indicate the formation of inclusion bodies inside the bacterial cytoplasm we have tried to co-express DnaKJE and GroELS chaperones together α-syn TAT8H4 scFv or DnaK TAT8H4 scFv. At this purpose in Bl21(DE) LysS cells were transformed with chaperone plasmid pG-KJE8 (obtained from Takara, Shiga, Japan), for over-expression of the two principal cytosolic *E.coli* chaperones systems, GroELS and DnaKJE.

Nevertheless, as it can see in a Western blot example (fig. 3.11), cytosolic chaperones over expression did not increase significantly the fusion-scFvs amounts in soluble fraction.

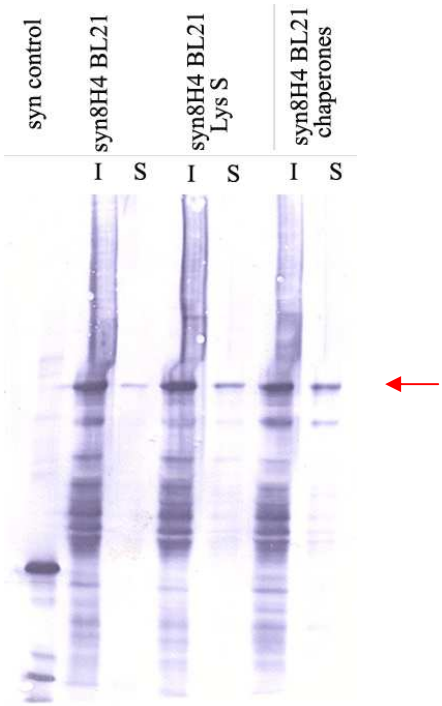


Fig. 3.11 TAT 8H4-scFv partition in the insoluble-soluble fractions when expressed as chaperone or chaperon-like fusion. Bl21(DE) cells, Bl21(DE) LysS cells, or Bl21(DE) pG-KJE8 cells, overexpressing chaperones GroELS and DnaKJE, were transformed with the indicated plasmid and protein expression was induced with 0.5 mM IPTG at 25°C. Cells lysates were prepared and proteins were fractionated, and analyzed by Western blot with anti-his antibodies. Soluble fraction (S), Insoluble fraction (I) .

3.3.5. Periplasm purification of α-syn 8H4 scFv

It is know that in the more oxidizing environment of the periplasm, correct folding and disulphide bonds are supported. Thus it was thought to take advantage of abundant periplasmic expression of recombinant α-synuclein in *E.coli*. A-synuclein translocation appears to be mediated by the SRP-dependent pathway (signal recognition particle-dependent pathway) and the C-terminal (99-to-140 portion) cooperating with the central 61-to-95 section, plays a signal-like role for the its translocation into the periplasm (Guoping Ren et al., 2007).

By taking advantage of this, and hypothesizing that α-syn in fusion with 8H4-scFv can mediate the translocation, expression of active α-syn TAT8H4scFv into *E.coli* periplasmic space and purification were assessed.

After osmotic lysis and IMAC purification, an abundant band with an apparent molecular mass of 17 kDa corresponding to α-syn was detected (fig. 3.12), confirming that the expressed α-syn was transported into the periplasm, with yields also of about 20mg/L human α-syn after IMAC purification. Nevertheless the expression of α-syn 8H4scFv fusion protein was not detected in periplasmic fraction before and after IMAC purification (fig.3.13), either in *E. coli* BL21(DE3) Lys S either in pG-KJE8 *E.coli* Bl21(DE) cells, overexpressing chaperones GroELS and DnaKJE. Probably 8H4 scFv fusion partner inhibits the translocation of α-syn into the periplasm.

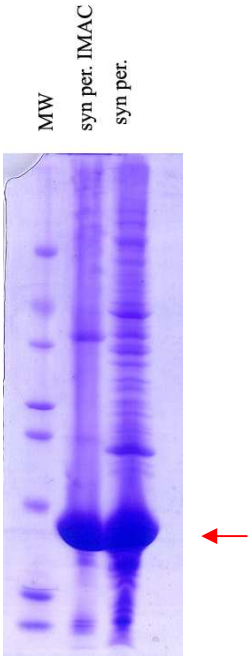


Fig 3.12 *E. coli* BL21(DE3) Lys S were transformed with pSYNmut4 and protein expression was induced with 0.5 mM IPTG at 25°C. After osmotic lysis and IMAC purification samples were subjected to SDS-poly-acrylamide gel electrophoresis and following staining with Coomassie Brilliant Blue.

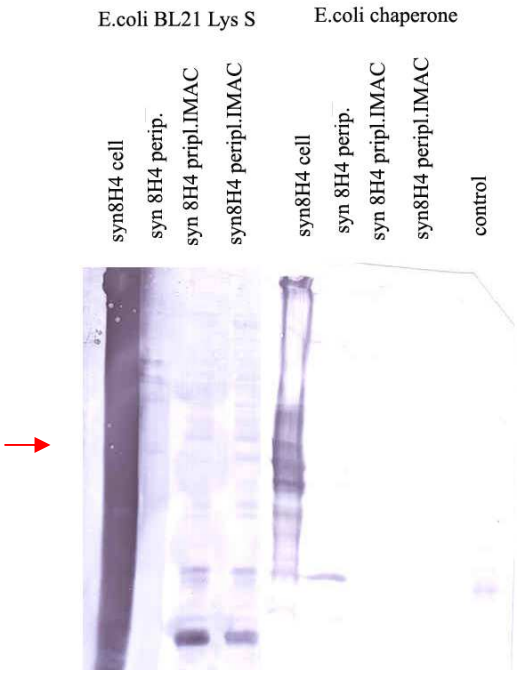


Fig.3.13 *E. coli* BL21(DE3) Lys S and pG-KJE8 Bl21(DE) cells, overexpressing chaperones GroELS and DnaKJE, were transformed with pSyn8H4opt and protein expression was induced with 0.5 mM IPTG at 25°C. After osmotic lysis and IMAC purification samples were subjected to Western blot with anti-his antibody.

3.3.6. Expression and purification in denaturing conditions of the fusion 8H4scFvs

Since production yields of soluble fusion-scFvs are not significant, while large quantities were observed in the insoluble fractions, revealing high tendency of TAT8H4 to form inclusion bodies inside *E.coli* cytoplasm upon expression, it was appeared more practical (as well as economical) purify Dnak TAT8H4 and α-syn TAT8H4 in denaturing conditions and subsequently regenerate to their active form using in vitro refolding procedures.

For expression, *E.coli* BL21(DE3) LysS cells were transformed with pSyn TAT8H4opt and pDnaK TAT8H4 were grown in LB at 37°C and harvested 3h post-induction. To improve the yield, DnaK TAT8H4scFv was purified in denaturanting condition solubilizing into guanidinium (fig.3.15). A-syn TAT8H4 was purified from inclusion bodies, that were collected by cells distruption in 4M urea followed by centrifugation. Purification of the non-native DnaK 8H4 scFv and α-syn TAT8H4 scFv solubilized in guanidinium under reducing condition, was carried out in using Ni-chelating His Trap columns (NTA-Ni^{+2}) in the presence of 8 M urea pH 4.5 (fig.3.14).

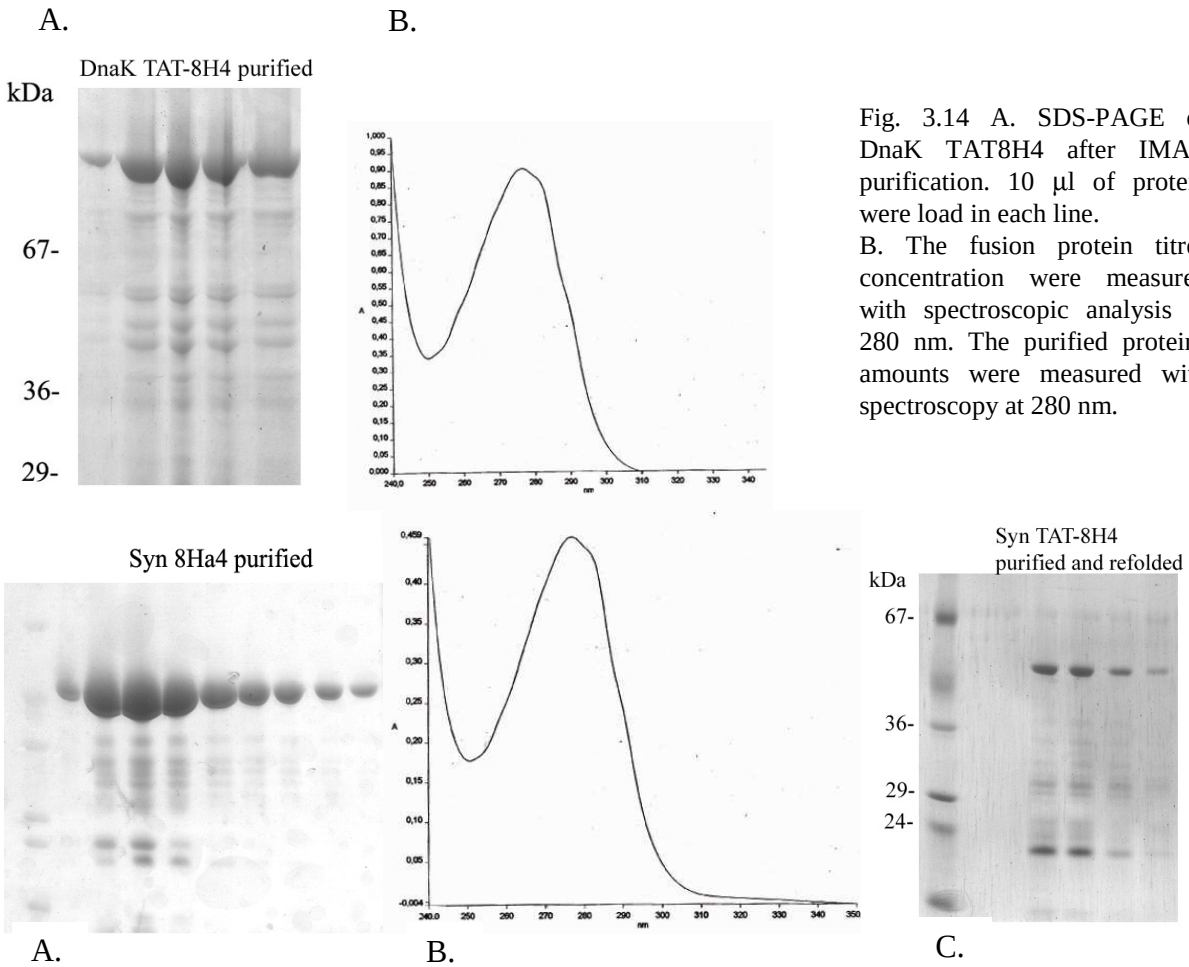


Fig. 3.14 A. SDS-PAGE of DnaK TAT8H4 after IMAC purification. 10 µl of protein were load in each line.
B. The fusion protein titres concentration were measured with spectroscopic analysis at 280 nm. The purified proteins amounts were measured with spectroscopy at 280 nm.

Fig. 3.15 A. SDS-PAGE of α-Syn TAT8H4 after IMAC purification. B. The fusion protein titres concentration were measured with spectroscopic analysis at 280 nm. The purified proteins amounts were measured with spectroscopy at 280 nm. C. SDS-PAGE of α-Syn TAT8H4 after on-column refolding. 10 μl of protein were load in each line in A and C.

The yields, on base of three purification processes, are 5.84 ± 0.89 mg/L for α-syn TAT8H4 (calculated $\varepsilon = 1{,}2\ M^{-1}\ cm^{-1}$) and 3.88 ± 1.28 mg/L for dnaK TAT8H4 (calculated $\varepsilon = 0.65\ M^{-1}\ cm^{-1}$).

3.3.7. Refolding of fusion-scFvs proteins

To recover the bioactivity of antibody fragment-fusion proteins, the purified scFvs were refolded with on-column refolding procedure. Two different protocols and buffers system are been tested.
In the first one, the purified fusion proteins were again incubated with Ni^{2+}-affinity resin in 8M urea pH8 with 7mM β-mercaptoethanol (to reduce the disulphide bonds) and then washed in three successive steps with decreasing guanidinium concentration from 2M to 0M (i.e. from 2 to 1 M, then from 1 to 0,5 M and finally from 0,5 to 0 M) keeping the concentrations of the oxidized glutathione and arginine constant (200mM arginine, 375μM GSSG). Analisys of samples, eluted in 200mM NaCl, 50 mM Tris/HCl pH6, 200mM arginine, 500mM imidazol by A_{280} and SDS-PAGE, showed no refolding efficiency either for DnaK 8H4scFv either α-syn 8H4 scFv.
The second protocol tested, based on-column chemical refolding method described by Oganesyan (Oganesyan et al., 2004) showed α-syn 8H4 refolded with ~ 10% -12% yield, but no DnaK 8H4. This refolding method for insoluble His-tagged proteins expressed in *E.coli*, has been called "artificial chaperone-assisted refolding" because has been claimed to mimic the GroEL-GroES chaperonin action in vivo.
In the original protocol, scFv purification and refolding is in only one steps of batch-absorption to Ni-NTA resin: solubilized inclusion bodies were bound to Ni-NTA resin, exposed to a detergent-containing solution to prevent misfolding and aggregation, followed by stripping of the detergent with cyclodextrin to promote correct folding, finally the purified proteins were eluted with imidazole. To achieve higher yield of refolded protein, we have modified the original protocol, adjusting scFv purification and refolding processes in two steps of batch-absorption to Ni-NTA resin, rather then one step as demonstrate also for TIMP2, another protein with six disulphide bridges (Negro et al., 1995). As described in more detail in material an methods, solubilized inclusion bodies were bound to Ni-NTA resin, washed and the purified protein eluted. In the next step, the denatured purified scFv was again bound to Ni-NTA and first was exposed to a detergent-containing solution to prevent misfolding and aggregation, followed by stripping of the detergent with cyclodextrin to promote correct folding. Refolded antibody was finally eluted with imidazole.

Following the original protocol with a single step on Ni-NTA affinity column, heavy precipitates of scFv were likely produced into column because refolded fusion scFv proteins were not showed in eluted samples by A_{280} and SDS-PAGE analysis.

Thus, a modified on-column chemical refolding method, based on protocol described by Oganesyan (Oganesyan et al., 2004), permitted to obtain α-syn 8H4scFv fusion protein refolded with ~ 10% -12% yield. A-syn 8H4scFv refolded recovered was about 0,6 mg/L (fig.3.15C).

3.3.8. Antigen binding specificity of refolded α-syn TAT8H4 scFv

The specifity of the refolded α-syn 8H4 scFv to recognize and interact with human PrP^C was determinated by immunoblotting. As shown in fig. 3.16 the soluble correctly folded scFv was active and able to recognize human prion protein.

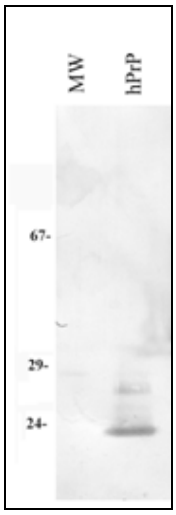


Fig. 3.16 Interaction between purified refolded TAT8H4 and human PrP^C as showed by immunoblotting. 50 ng of PrP^C was transferred to membrane and incubated with refolded α-syn 8H4scFv, and following with antibodies mouse anti-His and anti-mouse.

3.3.9. Buffer-exchange by gel filtration

Dnak TAT8H4 and α-syn TAT8H4 scFvs fusion proteins purified in denaturanting conditions from inclusion bodies were further purified by gel filtration column Use of desalting column allows to separate proteins from denaturant urea , toxic at high concentrations (8M titres) for cells treatments, replacing with above buffer that maintains the protein in solution without aggregation and is not toxic in cell treatments. Moreover use of protein-size column fractionates protein species, permitting purification from small protein degradation products and salts. Column matrix may also help proteins disperse, reducing aggregation.

3.3.10. 8H4 scFv fusion proteins trasduction inside CHO cells

Dnak TAT8H4 and α-syn TAT8H4 scFvs fusion proteins so obtained, were used to transduction of CHO cells, in order to verify their entry inside cells, localization and their time-life. The TAT-fusion proteins were added in culture medium at ~10μg/ml corresponding to final concentration of ~215 nM α-syn or ~100nM DnaK. After 2-3 hours, cells grown on glass cover-slips were fixed, permeabilized and incubated with anti-His primary antibody, that bound exa-His tag of proteins, followed by TRITC-conjugated secondary antibody for observation of cells using the fluorescence microscope (red staining). To visualize cells nuclei, Hoechst staining was performed (blue staining). As shown in fig. 3.17, Dnak TAT8H4 and α-syn TAT8H4 scFv fusion proteins localize inside the cell, in the nucleus and cytoplasm.

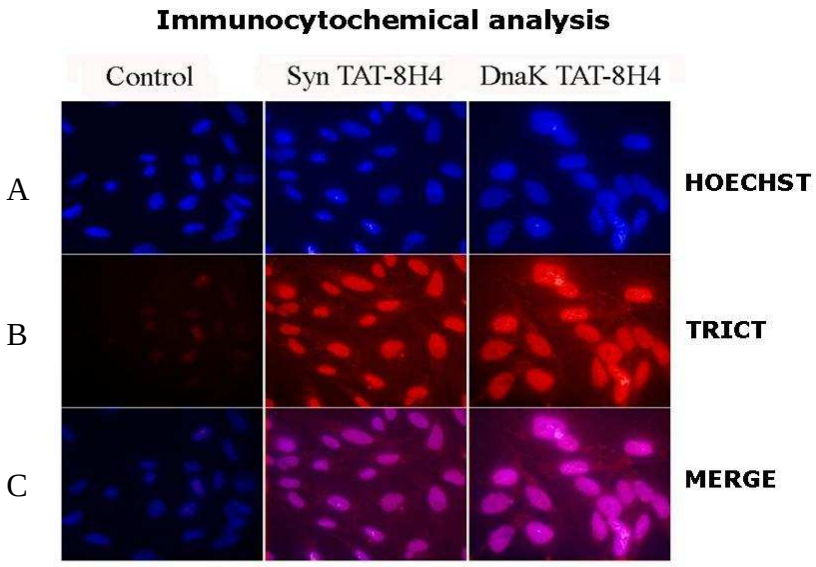


Fig. 3.17 Fluorescence microscope images show the α-syn TAT8H4 and DnaK TAT8H4 diffusion into CHO cells and their localization.
CHO cells incubated with ~10 μg/ml of α-syn TAT8H4 or DnaK TAT-8H4 for 2-3 hours. Hoechst staining of nuclei (line A) and anti-mouse II,TRICT of scFv marked (line B) are superimposed (line C).

In CHO cells untreated negative control , there is Hoechst staining detection of cellular nuclei (blu staining, images A) but not TRICT staining detection. A very weak signal revealed, is likely due to aspecific bound of antibody TRICT coniugated to cell.

In CHO cells treated, it is evident that α-Syn TAT-8H4 and DnaK-8H4 diffuse inside the cell (red staining revealed by TRICT-signalling, images B) in cytoplasma and nucleus, as showed by nuclear colocalization as revealed by superimposed images in pink staining (images C).

Thus, DnaK and α–syn fusion scFv, diffuse inside CHO cells due TAT sequence.

3.3.11. Time course and Western blot analysis

Every protein has a structural stability inside the cell, fundamental to its physiological role. Time-life test of a protein inside cells gives information about treatment system. At the aim, CHO cells

were incubated with ~10 μg/ml of α-syn (~215 nM) or DnaK (~100nM) fusionTAT-8H4. The cells harvested at different time 3, 20, 28, 51 hours. After SDS-PAGE and Western blot of cellular lysates, the proteins were visualized, using primary anti-His antibody and secondary anti-mouse-HRP conjugated with luminol detection.

As showed in fig. 3.18, α-syn TAT-8H4 is evident inside CHO cells at least until 28 hours from administration, while DnaK TAT-8H4 is degraded much more in time, because of it is not visible in western just after 3 hours. It gets inside the cells, as seen in previous experiments with CHO cells treatments, but likely it is early degraded.

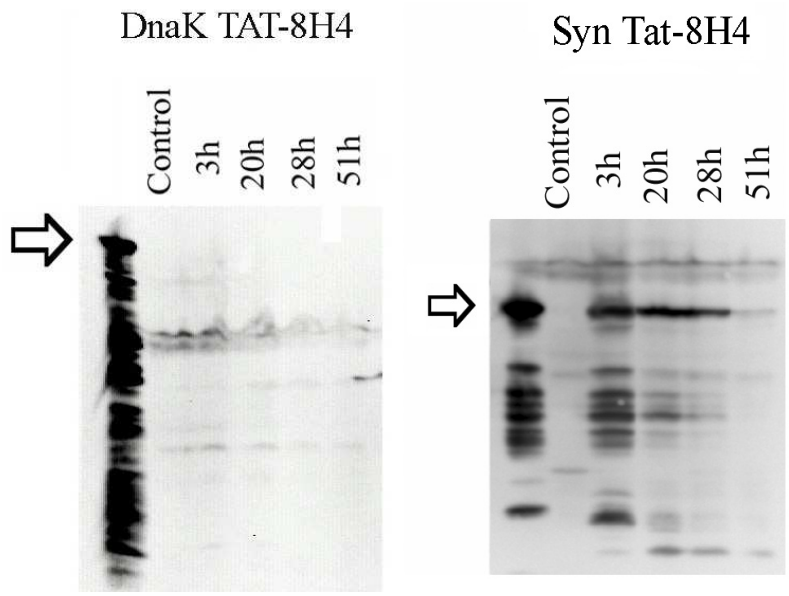


Fig. 3.18 CHO cells were incubated with 10 μg/ml of α-syn (~215 nM) or DnaK (~100nM) of α-syn or DnaK fusionTAT8H4. After SDS-PAGE and Western blot, the proteins were visualized, using anti-His and anti-mouse-HRP antibodies. A-syn TAT8H4 remains inside CHO cells, at least until 28 hours, much more than

Every protein have a variable time-life, depending of protein nature. For example, TAT-DJ1 protein no showed significant degradation after 48 hours (Batelli et al., 2008) and in vivo TAT α-synuclein was also present after 30 dies from sthereotassic injection in brain (Recchia et al.,2008). Protein time-life can be also influenced by more or less fast metabolism of cell. CHO cells are replicating about every 20 hours, while nervous cells have more long time.

These results are important to programme cellular treatments, because of proteins with short time-life will have to be added more frequently.

3.3.12. A-syn TAT8H4 scFv trasduction inside PrP transfected HeLa cells

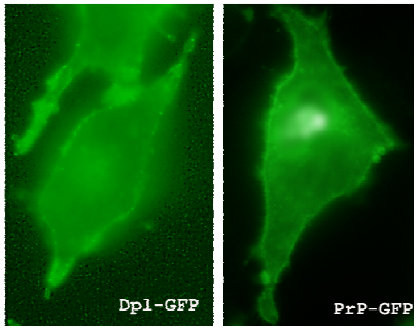
At the light of short time-life of DnaK-8H4scFv inside CHO, further functionality experiments in HeLa cells were conducted with α-syn TAT8H4 scFv. To test for the α-syn TAT8H4 scFv

capacity to deplete the superficial membrane prion protein, HeLa cells were transiently transfected with plasmid able to express GFP-Doppel and plasmid expressing GFP-PrP (Negro et al., 2001).

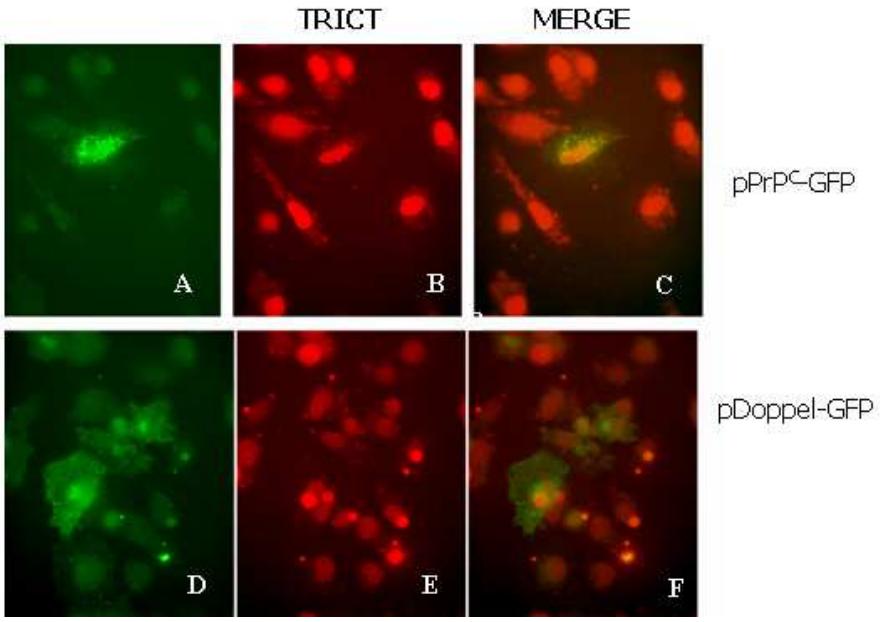
Doppel (Dpl) is identified as prion (PrP)-like protein due to the structural and biochemical similarities, both glycosylphosphatidylinositol(GPI)-anchored cell membrane proteins, with the same cellular localization pathway and topological location in membrane rafts. But Dpl is not recognized by 8H4scFv because of different amino acid sequence, so it represents an optimal control of 8H4 scfv antibody functionality and specificity. When Dpl and PrP are in fusion with GFP reporter gene, they show, by fluorescence in HeLa cells, their cellular localization in plasma membrane and partially in Golgi apparatus, one of sites of pathway of prion protein synthesis (Fig.3.19 panel A).

If these HeLa cells, after 36 hours post-transfection, were treated with 10 μg α-syn8H4 scFv (~215 nM) for 3h and fixed after 3 (Fig. 3.19 panel B), we can see scFv antibodies in any case inside HeLa cells, localizing prevalently in perinuclear region (photos B, E). Thus indicating overall that antibody fragments penetrate into all cells with high efficiency, these data confirming the results with CHO cells. But while GFP-Doppel fusion protein remains localized into plasma membrane and does not co-localize with 8H4 antibody (photos D, F), in the case of GFP-PrP transfected HeLa cells, we can see (photos A, C) that PrP became with difficulty at plasma membrane surface (poorly membrane fluorescence), showing strongly to co-localize with antibody fragment in perinuclear region. These data prove functionality and specificity of α-syn8H4scFv in cells, moreover in vitro by immunoblotting.

Nevertheless, administration of same amounts (10 μg/ml) of refolded or buffer-exchange gel-filtrated α-syn TAT8H4, demonstrated (data no shown) that antibody fragments penetrate both into all cells with high efficiency, but that α-syn TAT8H4 (particularly buffer-exchange gel-filtrated α-syn TAT8H4) is more efficient to deplete the superficial membrane prion protein.



Panel A



Panel B

Fig 3.19 Localization of GFP-PrP or its analog GFP-Doppel before and after treatment with α-syn TAT8H4 antibodies. HeLa were transfected with pPrPC-GFP or its analog pDoppel-GFP used as control (panel A) and then treated with α-syn TAT8H4 antibodies (panel B).
While α-synTAT8H4 scFv antibody was internalized and localized in perinuclear region (photos B, E), Doppel protein is anchored to plasma membrane via a GPI-anchor and its distribution remains invariable after TAT-8H4scFv antibody treatment (photos D,F). Prion protein is also bounds to plasma membrane via a GPI-anchor, but its distribution changes after TAT-8H4scFv treatment (photos A,C).

3.4. Discussion

Previous experiments in our laboratory of anti-prion TAT 8H4scFv protein production in BL21 *E.coli* cytoplasm, were resulted in low yields of insoluble antibody.In order to enhance the yields of functional antibody fragments and allow cheaper and easier production and purification, we have tried to produce anti-prion TAT-8H4scFv protein (eukaryotic complex molecule with three disulphide bonds indispensable for its stable structure and activity) in secreted form by gram-positive 800N *B.subtilis* and MS941 *B.megaterium* bacteria.

Nevertheless, in spite of numerous trials, we have shown that scFv is expressed in *Bacillus* cells intracellulary (even if with no very high yields and primary as insoluble), but no TAT-8H4 scFv is clearly detected in ammonium sulphate precipitation or IMAC purified culture medium, although high level of secretion by *Bacillus* secA pathway was widely demonstrated in other case (Schumann W. et al., 2006).

So other strategies were tried. In attempt to overcome the difficulty of generating soluble heterologous scFv protein in *E.coli*, one of more promising approach, on which the majority of the

work in the field is focusing, is development and refinement of solubility fusion tags. As the name implies, these tags are proteins or peptides that are fused to the protein of interest and, in the best case, help to properly fold their partners leading to enhanced solubility in the protein of interest. This strategy of fusing a protein partner is effective in promoting the solubility of many target proteins (Esposito and Chatterjee, 2006).

Cytoplasmic expression using fusion partners such as glutathione S-transferase (GST), thioredoxin (Trx), and maltose-binding protein (MBP) generally results in high productivity and enhanced solubility (LaVallie E.R. and McCoy J.M., 1995). In the contest of a fusion protein, these fusion partners seem to function as molecular chaperones that promote the solubility and stability of scFvs in active form in the cytoplasm, and since they do not release the protein that is fused to them, their stabilizing and solubilizing effects persist. Recently C.Kyratsous et al. demonstrated that also bacterial chaperones as DnaK or GroEL placed in frame of a target polypeptide, facilitate production of large amounts of soluble recombinant proteins, that are normally highly insoluble when expressed in bacteria.

Therefore we have tried to take advantage from the use of fusion protein technology to increase the yields and solubility of TAT8H4scFv. At the aim we have chosen as fusion tags, bacterial chaperone DnaK and like-chaperone α-synuclein protein, that for its activity together physic-chemical characteristics, as a small protein dimension, (that contributes at low molecular weights in fusion protein), high expression levels in *E.coli* (30% of total *E.coli* proteins), its facilitated purification (due its high resistance to protease and to high temperature), could be a good candidate as fusion partner.

Although fusions of TAT 8H4scFv to α-synuclein and DnaK result in expression and accumulation of significant levels of antibody respect of low levels of TAT 8H4scFv, these fusion proteins show to be largely insoluble in different buffer and temperature, differently from unfusion partners α-synuclein and DnaK.

To increase the soluble and active fusion scFvs yields, other trials were done because of there is no universal approach that can be applied to all cases but it depends by protein sequence in a trial and error process. Since the behaviours of the fusion proteins indicate the formation of inclusion bodies inside the bacterial cytoplasm we have tried to co-express DnaKJE and GroELS bacterial chaperones together α-syn or DnaK fusion-8H4scFvs. Nevertheless, cytosolic chaperones over expression did not increase significantly the fusion-scFvs amounts in soluble fraction, differently from satisfying yields obtained by Kyratsous (2009) for DnaK-PrP fusion protein. In attempt to explain the divergence of results obtained (also in literature) upon over expression of DnaK in fusions and co-expression of its co-chaperones, it was recently focused attention on dual role of

the DnaK chaperone, which acts both as a folding modulator and as proteolytic enhancer (Martìnez-Alonso et al., 2010).

So, being know that in the more oxidizing environment of the periplasm correct folding and disulphide bonds are supported, we have tried to take advantage of abundant periplasmic expression of recombinant α-synuclein in *E.coli* (Guoping Ren et al., 2007), hypothesizing that α-syn in fusion with 8H4-scFv can mediate the translocation. Nevertheless the expression of α-syn 8H4scFv fusion protein was not detected in periplasmic fraction before and after IMAC purification. Probably 8H4 scFv fusion partner inhibits the translocation of α-syn into the periplasm.

Since, nevertheless various trials, the yields of soluble fusion-scFvs are not significant, while large quantities were observed in the insoluble fractions (revealing high tendency of TAT8H4 to form inclusion bodies inside *E.coli* cytoplasm upon expression), we have tried to purify Dnak TAT8H4 and α-syn TAT8H4scFvs in denaturing conditions by IMAC, taking advantage by His-tag and subsequently regenerate ones to their active form using in vitro refolding procedures. Between the different protocols tested, an our modified two-steps refolding procedure, based on-column chemical refolding method described by Oganesyan (Oganesyan et al., 2004) showed α-syn 8H4 refolded with ~ 10% -12% yield, but no DnaK 8H4; while the original one-step Oganesyan protocol, was resulted inefficiently for both the fusion-scFvs. Although the yields are lower than expectations (yields of 30-100% of the proteins refolded in seven of 10 tested protein by Oganesyan method) the efficacy of modified refolding procedure is been confirmed by immunoblotting, showing the soluble correcly folded α-syn 8H4 scFv active and able to recognize specific human prion protein.

Thus, in our case the soluble production rate, that is strongly dependent on the individual sequence of the antibody, it would seem depend by correct disulphide bonds formation.

At this point, we have assay the ability of obtained TAT 8H4 fusion proteins to transduce inside CHO cells. Because of low or no yields of soluble refolded proteins, Dnak TAT8H4 and α-syn TAT8H4 scFvs fusion proteins, purified in denaturanting conditions from inclusion bodies, were further purified, without great losses (with high yields), by gel filtration against a buffer that maintains the proteins in solution without aggregation and is not toxic in cell treatments.

Treatments of CHO cells with Dnak TAT8H4 and α-syn TAT8H4 scFv show that the recombinant proteins localize inside the cell, in the nucleus and cytoplasm and that α-syn TAT8H4 has time-life major (at least 28h) than Dnak TAT8H4. This probably due to major dimension and minor concentration of DnaK fusion protein (more than 100 kDa respect of α-syn fusion protein about 46 kDa), although several TAT transduction domain are able to cross the plasma membranes (Dowdy

et al., 2001). Every protein have a variable time-life, depending of protein nature. For example, TAT-DJ1 protein no showed significant degradation after 48 hours (Batelli et al., 2008) and in vivo TAT α- synuclein was also present after 30 dies from sthereotassic injection in brain (Recchia et al.,2008). Protein time-life can be also influenced by more or less fast metabolism of cell. CHO cells are replicating about every 20 hours, while nervous cells have more long time. These results are important to programme cellular treatments, because of proteins with short half time- will have to be added more frequently.

So, at the light of short time-life of DnaK-8H4scFv inside CHO, further functionality experiments were conducted with α-syn TAT8H4 scFv in HeLa cells. HeLa cells were transiently transfected with plasmids able to express GFP-Doppel (control) and GFP-PrP (Negro et al., 2001) that localize both in cell membrane, glycosylphosphatidylinositol(GPI)-anchored. Administration of same amounts (10 μg/ml) of refolded or buffer-exchange gel-filtrated α-syn TAT8H4, demonstrate that antibody fragments penetrate into all cells with high efficiency, (these data confirming the results with CHO) and that α-syn TAT8H4 (particularly buffer-exchange gel-filtrated α-syn TAT8H4) is specific to deplete the superficial membrane prion protein, because of while GFP-Doppel fusion protein remains localized into plasma membrane and does not co-localize with 8H4 antibody, GFP-PrP arrives with difficulty at plasma membrane surface, showing strongly to co-localize with antibody fragment in perinuclear region.

We have observed a major efficiency of buffer-exchange gel-filtrated α-syn TAT8H4 respect of refolded one, nevertheless the first is denatured while the second is correctly refolded, as demonstrated by immunoblotting. This probably because denaturated proteins have a higher potential for biological effects than their counterparts prepared under soluble conditions (Nagahara et al., 1998). It was hypothesized that, due to reduced structural constraints, higher-energy (ΔG), denatured proteins may transduce more efficiently into cells than lower-energy, correctly folded ones. Once inside the cell, transduced, denatured proteins, can be correctly refolded by chaperones (Gottesman et al.,1997) and make their activity into cell . Indeed, an analysis of p27Kip1 protein revealed that when this protein was urea denaturated, the biological phenotypes were enhanced dramatically over protein prepared under soluble conditions (Nagahara et al., 1998). Thus preparing TAT-fusion proteins in this waycould be advantageous in most situations and result in dramatic yield and efficient activity increases (Dowdy et al., 2001). On the other hand, denatured proteins are more subject to degradation in circle.

In conclusion α–synuclein was resulted better than DnaK in fusion with TAT8H4 scFv, either for small dimension that contributes at low molecular weights in fusion protein, than for higher yields, ability to refold and specific activity to deplete PrP, so it could be effective in spongiform

transmissible diseases treatment. Moreover, the use of cell-permeable antibodies or “transbody”, due to TAT domain that penetrates efficiently into cells translocating across the plasma membrane and also the blood-brain barrier (Dowdy et al., 2001), would avoid the overwhelming safety and ethical concerns surrounding the direct application of recombinant DNA technology in human clinical therapy, which is required for intrabody expression within the cell by use of viral-based vectors. In fact, although the TAT transduction domain was originally identified from HIV, no infectious materials are contained within this domain; moreover it has very low toxicity in cell culture, because being highly hydrophilic, TAT peptide causes little perturbation to the plasma membrane (even during 1-hr incubations at a concentration of 100 μM).

Moreover transbodies introduced into the cell would possess only a limited active half-life, without resulting in any permanent genetic alteration. This would allay any safety concerns with regards to their application in human clinical therapy.

Chapter IV

Metabolic engeeniring of hyaluronic acid in *Bacillus subtilis*, *Bacillus megaterium* and *Escherichia coli*

4.1.1. Hyaluronic acid

Hyaluronic acid (HA), first described in 1934 (Meyer and Palmer,1934) is a linear, unbranched polysaccharide, glycosaminoglycan consisting of repeating disaccharide units (fig. 4.1) of alternated β-(1–4)-glucuronic acid (GlcA) and β-(1–3)-N-acetylglucosamine (GlcNAc) moieties (Weissman and Myer, 1954) and its molecular weight (MW) is high generally ranging from 10kDa to 10000 kDa in vertebrates and bacteria (De Angelis 1999).

In marked contrast with other gycosaminoglycans, HA is synthesized, not in the Golgi apparatus, but on the cytoplasmic surface of the plasma membrane (Prehm, 1984) and it is not sulphated or modified any other way troughout its length.

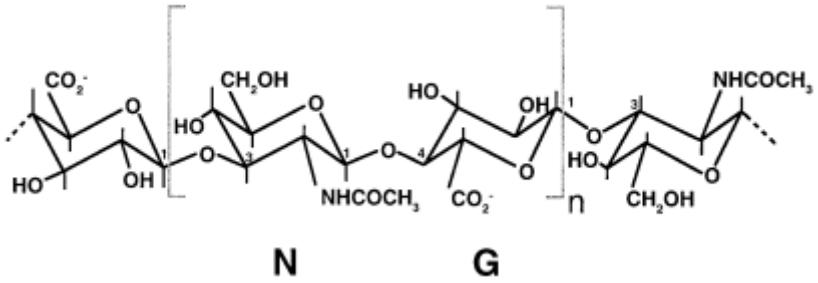


Fig. 4.1 Primary structure of hyaluronan. Hyaluronan is a regular, unbranched bipolymer consisting of a repeating disaccharide unit of 2-acetamido-2-deoxy-β-D-glucopyranose (N-acetylglucosamine) (N) and β-D glucopyranosyluronic acid (glucuronic acid) (G). The glycosidic linkages between N and G, and G and N are via the (1 → 4) and (1 → 3) ring positions respectively. (Adapted from Adam and Ghosh, 2001).

Existing in nature as a hydrated gel, HA is ubiquitous in the extracellular matrix (of connective tissues) in vertebrates, and it is particularly abundant in cartilage, synovial fluid, dermis, and the vitreous humor of the eye, where it serves specialized functions. This macromolecule influences cell behavior and has significant structural, rheological, physiological, and biological functions in the body (Knudson et al.,1993, Laurent et al., 1992, Toole, 1991), and also plays a critical role during fertilization and embryogenesis.

In many group A and C streptococci, HA forms a capsule that helps these microbes evade the host immune system (Wessels et al., 1991).

HA molecular weight is important for the physiochemical as well as biological properties of HA, thus its roles correlate with the length of the HA chain. High molecular weight is important for HA to exert its unique rheological properties (Fouissac et al., 1993), for mucoadherence (Saettone et al., 1991; Saso et al., 1999), and anti-inflammatory effects (Suzuki et al., 1993) whereas low molecular weight is a potent signaling molecule (Spurlock et al., 1999).

Its distinctive viscoelastic properties, coupled with its lack of immunogenicity or toxicity, have led to a wide range of applications in the cosmetic and pharmaceutical industries, including

skin moisturizers, osteoarthritis treatment, ophthalmic surgery, adhesion prevention after abdominal surgery, and wound healing (Goa and Benfield, 1994; Wohlrab and Reinhard, 2000).
In addition, many potential applications of HA are being developed, expecially in the area of cancer treatment, drug delivery and tissue engineering (Zeng et al. 1998; Toole et al. 2002; Galassi et al. 2000). At present the worldwide market for HA.

4.1.2. Commercial hyaluronic acid sources

At present the worldwide market for HA is estimated at over one billion dollars per annual, and is satisfied primarily by extraction from rooster combs and certain attenuated strains of group C *Streptococcus* which synthesize this compound naturally as part of their outer capsule (Balasz, 1979; Hascall and Laurent, 1997). However, these are less-than-ideal sources. All rooster comb-based HA products carry warnings directed to those who are allergic to avian products, and at least one has been reported to cause inflammatory reactions upon injection (Puttick et al., 1995). Moreover purification requires expensive purification processes because of HA is complexed with proteoglycans. Instead streptococci can be difficult or expensive to ferment, are challenging to genetically manipulate, and have the potential to produce exotoxins, moreover extraction requires organic solvents.
Thus, it would be advantageous to develop an alternative source of HA that avoids these serious pitfalls, since HA is primarily used in the biomedical field with many applications requiring injection into the human body. Moreover, to extend the applications of HA and make better HA-containing biomedical products, it is necessary to obtain specially designated MW or uniform-size definited HA.
Recently, recombinant synthesis in microorganisms through metabolic engineering emerged as an attractive alternative that could relieve safety concerns associated with the use of pathogenic microorganism and avian products. Both Gram-positive and Gram-negative bacteria were used as host, including *Bacillus subtilis* (Widner et al. 2005, Chien and Lee 2007a), *Lactococcos lactis* (Chien and Lee 2007b; Prasad et al., 2009; Sheng et al., 2009), *Agrobacterium* sp. (Mao and Chen 2007), and *Escherichia coli* (Yu and Stephanopoulos 2008; Mao et al. 2009, Chen et al. 2009). For a successful metabolic engineering effort, adequate expression of heterologous enzymes is a prerequisite.

4.1.3. HA biosynthesis in metabolically engineering *Bacillus* strains

The goal of this second part of the work, in collaboration with Fidia S.p.a., is a new HA production method in metabolically engineering by replicative system *Bacillus* strains, particularly *Bacillus subtilis* and *Bacillus megaterium.*

Bacillus species have long been established as industrial workhorses for the production of products ranging from hydrolytic enzymes, such as proteases and alpha-amylases, to specialty chemicals, such as amino acids and vitamins (Schallmey M. et al., 2004; Harwood et al., 1992, Wong et al.,1995; Vary P., 1994). These organisms are capable of secreting copious amounts of product, indicative of their highly developed biosynthetic capacity, and are very economical to grow in industrial fermentors. *Bacillus subtilis* is free of exotoxins and endotoxins and, consequently, many products produced in this organism have been awarded a GRAS (generally recognized as safe) designation. In addition, *B. subtilis* does not produce, nor does the genome sequence encode, a hyaluronidase which could degrade HA. Finally, *B. subtilis* is one of the most-well-characterized gram-positive microorganisms, its genome has been sequenced, and there are wide arrays of tools available for genetic manipulation. Moreover, in contrast to other bacilli strains *B.megaterium* has the advantage, that none of the alkaline proteases are present. This fact enables an excellent cloning and expression of foreign protein without degradation. Furthermore *B.megaterium* is able to stable maintain several extra-chromosomal DNA elements in parallel-replicating. Thus, these nonpathgenic bacteria offer several advantages as a possible expression hosts for HA production.

4.2. Material and methods

4.2.1. Bacterial strains

The bacterial strains used in the present study are shown in Table 4.1.

Table 4.1 Bacterial strains used in this study

Strains:
E. coli host

- Topl0F (F-*mcr*A Δ(*mrr-hds*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1*ara*D139 Δ(*ara-leu*) 7697*gal*U *gal*K *rps*L(strR)*end*A1 *nup*G). Novagen
- invαF' F´ *end*A1 *rec*A1 *hsd*R17 (rk-, mk+) *sup*E44 *thi*-1 *gyr*A96 *rel*A1 φ80*lac*ZΔM15Δ(*lac*ZYA-*arg*F)U169 λ- Novagen

- BL21(DE3) (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) Invitrogen
- BL21(DE3) Lys S (F^-, *omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm* (DE3) pLysS (CamR) Invitrogen
- JM 110 Stratagene

Topl0F, invaF' : used as recipient for cloning, transformation, propagation, and maintenance.
BL21[DE3], BL21[DE3] Lys S : specifically designed for expression of genes regulated by the T7 promoter.

Bacillus host

- *Bacillus subtilis* 1012 *leuA8 metB5 trpC2 hsdRM1* Schumann (Germany)
- *Bacillus subtilis* WB800N *nprE aprE epr bpr mpr :: ble nprB :: bsr .vpr wprA :: hyg cm :: neo; NeoR* Schumann
- *Bacillus megaterium* MS941 MoBiTec

4.2.2. Growth condition

E.coli strains were grown aerobically in Luria Broth (LB) at 37°C supplemented, when necessary with specific antibiotic. *Bacillus subtilis* strains were grown aerobically in LB, MXR, 2xTY, Nutrient (Difco) or MSR (Table 4.2) at 25°C or indicated temperature supplemented, when indicated, with sugars (glucose and/or sucrose). For selection of plasmid, the medium was supplemented with chloramphenicol (10μg ml^{-1}) and/or neomycin (10μg ml^{-1}). *Bacillus megaterium* strain was grown aerobically in MSR supplemented with sugars (glucose and sucrose) at 25°C or indicated temperature. For selection of plasmids, the medium was supplemented with chloramphenicol (4,5 μg ml^{-1}) and tetracyclin (10μg ml^{-1}).

Table 4.2 Culture media composition.

LB	MXR	2xTY	MSR
1%Bacto tryptone 0,5% Bacto yeast extract 0,5% NaCl	2.4% Bacto yeast extract 1.2% casein hydrolysate 0.4%glycerol 0.17 M KH_2PO_4 0.72 M K_2HPO_4	1.6%Bacto tryptone 1% Bacto yeast extract 1% NaCl 20 mM potassium phosphate buffer pH 7.0	2.5% Bacto yeast extract 1.5%Bacto tryptone 0.3% K_2HPO_4 1% glucose trace elements: 1ml/L 1000x stock solution (2M $MgCl_2$, 0,7M $CaCl_2$, 50mM $MnCl_2$, 5mM $FeCl_3$, 1 mM $ZnCl_2$)

4.2.3. DNA manipulation and transformation procedures

Procedures for DNA purification, restriction, ligation, polymerase chain reaction (PCR), agarose gel electrophoresis, preparation and transformation of competent *E.coli* cells were done using standard techniques (Sambrook et al. 1989).

4.2.4. Construction of plasmids

All plasmids have been checked for correctness by cleavage with some restriction enzymes in agarose gels and by nucleotide sequencing.

4.2.4.1. Construction of expression cassette-operon *hasA-tuaD*

UDP-glucose dehydrogenase gene (*tuaD*) cloning from *Bacillus subtilis* in pRSETB *tuaD*

The UDP-glucose dehydrogenase gene (*tuaD*), 1402 bases pairs bp of length, (Genebank access number AF015609), is one of 8 genes of teichuronic acid operon tuaABCDEFGH. *TuaD* gene was amplified by PCR using the *Bacillus subtilis* 168 strain chromosome DNA (obtained by extraction according to the Qiagen procedure) as template, and the following primers 5'ATGAAAAAATAGCTGTCATTGGAACAG 3' and 5'TTATAAATTGTCGTTCCCAAGTCT 3'. The amplified PCR product was of expected molecular weight and was further amplified with the primers 5'GCTGGATCCATGAAAAAATAGCTGTCATTGG3' and 5'CTCGCTAGCTTATAAATTGACGCTTCCCAAG 3'. This second amplified PCR product was cut by *BamHI* and *NheI,* and cloned in the same restriction sites of plasmid pRSETB (Invitrogen). In order to introduce the Shine-Dalgarno sequence at 5' of the gene for *tuaD*, the entire gene was amplified with the following primers 5'CGACATATGAAAAAATAGCTGTCATTGG3' and 5'CTCGCTAGCTTATAAATTGACGCTTCCCAAG3'.The amplified fragment was digested with restriction enzyme *NdeI* and *NheI* and cloned in the plasmid pRSETB between the same restriction sites. In this way the plasmid pRSETBtuaD contains the gene encoding for *tuaD* under the control of T7 promoter and an efficient ribosome-site binding (Shine-Dalgarno).

pRSETtuaD *XbaI-- NdeI*----------**TuaD**---------- *NheI –BamHI-- BglI –XhoI*

In order to maximize the expression of *tuaD* in *Bacillus subtilis* the initial start codon GTG for *tuaD* (able to encode for Valine), has been replaced with more frequent and efficient ATG that encodes for Methionine using the primer for mutagenesis 5'TACATATGAAAAAAATAGCTGTCATTGGAACAGG3'.

Hyaluronan synthase gene (*hasA*) cloning from *Streoptocooccus zooepidemicus* in pGEM4Z *hasA*

The gene sequence, 1254 bp of length, (*hasA*) coding the hyaluronic acid sintetase, was amplified by PCR using the chromosomal DNA of *Streptococcus zooepidemicus equi* (obtained by extraction kit Quiagen) as template , and the following primers 5'ATGAGAACATTAAAAAACCTCATAAC 3' and 5'TAATAATTTTTTACGTGTTCCCCAG 3'. The amplified PCR product was of expected molecular weight and was further amplified using the primers 5' GGAGGATCCATGAGAACATTAAAAAACCTCAT 3' and 5' CAGTCTAGATTATAATAATTTTTACGTGTCC 3' . The final PCR product *szHasA* was digested with *BamHI* and *XbaI* and cloned between the same sites in vector pGEM4Z.

pGEM4*hasA* *HindIII-BamHI* ----------**HasA**----------*XbaI-SalI*

Construction of cassette operon *hasA-tuaD*

To realize the cassette operon *hasA-tuaD*, with the two genes in tandem under the control of inducible Pgrac promoter in *B.subtilis* expression vector pHT01 (MoBiTec), pGEM4-*hasA-tuaD* has been created.

At the aim, the pRSETB*tuaD* has been digested *XbaI* and *XhoI* and the resulting *tuaD* gene sequence has been cloned in the digested pGEM4*hasA* vector between *XbaI* and *SalI* sites *(Xho I* and *SalI* are compatible restriction sites), resulting in plasmid pGEM4*hasA-tuaD*.

pRSETB*tuaD* *XbaI-- NdeI*----------**TuaD**---------- *NheI –BamHI-- BglI – XhoI*

pGEM4*hasA* *HindIII-BamHI* ----------**HasA**----------*XbaI-SalI*

pGEM4*hasA-tuaD* *HindIII-BamHI* ---------**HasA**---------*XbaI--NdeI*---------**TuaD**---------- *NheI –BamHI-- BglI –XhoI/SalI*

4.2.4.2. Construction of pHT01*hasA-tuaA E.coli/Bacillus subtilis* plasmid shuttle

To construct the *Bacillus subtilis* expression vector pHT01*hasA-tuaA*, pGEM4*hasA-tuaD* has been digested *BamHI* and *NheI*, and the outing fragment *hasA-tuaD* has been cloned in pHT01 digested *BamHI* and *XbaI* (*XbaI* and *NheI* are compatible restriction sites) originating pBS5 plasmid.

Here the complete nucleotide and amino acid sequences of plasmid pBS5:

```
   0  TTAAGTTATTGGTATGACTGGTTTTAAGCGCAAAAAAAGTTGCTTTTTCGTACCTATTAA
  60  TGTATCGTTTTAGAAAACCGACTGTAAAAAGTACAGTCGGCATTATCTCATATTATAAAA
 120  GCCAGTCATTAGGCCTATCTGACAATTCCTGAATAGAGTTCATAAACAATCCTGCATGAT
 180  AACCATCACAAACAGAATGATGTACCTGTAAAGATAGCGGTAAATATATTGAATTACCTT
 240  TATTAATGAATTTTCCTGCTGTAATAATGGGTAGAAGGTAATTACTATTATTATTGATAT
 300  TTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAATAGAAAGAGAAAAAGCATTTTCAG
 360  GTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATTATATTTCTCTACATCAGAAAGGT
 420  ATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGGAGTCCAAATACCAGAGAATGTTT
 480  TAGATACACCATCAAAAATTGTATAAAGTGGCTCTAACTTATCCCAATAACCTAACTCTC
 540  CGTCGCTATTGTAACCAGTTCTAAAAGCTGTATTTGAGTTTATCACCCTTGTCACTAAGA
 600  AAATAAATGCAGGGTAAAATTTATATCCTTCTTGTTTTATGTTTCGGTATAAAACACTAA
 660  TATCAATTTCTGTGGTTATACTAAAAGTCGTTTGTTGGTTCAAATAATGATTAAATATCT
 720  CTTTTCTCTTCCAATTGTCTAAATCAATTTTATTAAAGTTCATTTGATATGCCTCCTAAA
 780  TTTTTATCTAAAGTGAATTTAGGAGGCTTACTTGTCTGCTTTCTTCATTAGAATCAATCC
 840  TTTTTTAAAAGTCAATATTACTGTAACATAAATATATATTTTAAAAATATCCCACTTTAT
 900  CCAATTTTCGTTTGTTGAACTAATGGGTGCTTTAGTTGAAGAATAAAGACCACATTAAAA
 960  AATGTGGTCTTTTGTGTTTTTTTAAAGGATTTGAGCGTAGCGAAAAATCCTTTTCTTTCT
1020  TATCTTGATAATAAGGGTAACTATTGCCGATCGTCCATTCCGACAGCATCGCCAGTCACT
1080  ATGGCGTGCTGCTAGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCG
1140  GTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTA
                                                          EcoRI
1200  AGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTC
1260  GAGCTCAGGCCTTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGA
1320  AACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGT
1380  ATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTT
1440  CACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCG
1500  AAAATCCTGTTTGATGGTGGTTGACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTC
1560  GTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCAT
1620  TGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATT
1680  CAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGC
1740  TATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGC
1800  CGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAG
1860  ATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGT
1920  CTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAAT
1980  GGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAG
2040  ATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCAC
2100  GCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTG
2160  CAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTG
2220  TGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGT
2280  TTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACC
2340  GGCATACTCTGCGACATCGTATAACGTTACTGGTTTCATCAAAATCGTCTCCCTCCGTTT
2400  GAATATTTGATTGATCGTAACCAGATGAAGCACTCTTTCCACTATCCCTACAGTGTTATG
2460  GCTTGAACAATCACGAAACAATAATTGGTACGTACGATCTTTCAGCCGACTCAAACATCA
2520  AATCTTACAAATGTAGTCTTTGAAAGTATTACATATGTAAGATTTAAATGCAACCGTTTT
2580  TTCGGAAGGAAATGATGACCTCGTTTCCACCGGAATTAGCTTGGTACCAGCTATTGTAAC
2640  ATAATCGGTACGGGGGTGAAAAAGCTAACGGAAAAGGGAGCGGAAAAGAATGATGTAAGC
2700  GTGAAAAATTTTTTATCTTATCACTTGAAATTGGAAGGGAGATTCTTTATTATAAGAATT
                                                    BamHI
2760  GTGGAATTGTGAGCGGATAACAATTCCCAATTAAAGGAGGAAGGATCCATGAGAACATTA
   1                                                  M  R  T  L

2820  AAAAACCTCATAACTGTTGTGGCCTTTAGTATTTTTTGGGTACTGTTGATTTACGTCAAT
   1  K  N  L  I  T  V  V  A  F  S  I  F  W  V  L  L  I  Y  V  N
                           HindIII
2880  GTTTATCTCTTTGGTGCTAAAGGAAGCTTGTCAATTTATGGCTTTTTGCTGATAGCTTAC
   1  V  Y  L  F  G  A  K  G  S  L  S  I  Y  G  F  L  L  I  A  Y

2940  CTATTAGTCAAAATGTCCTTATCCTTTTTTTACAAGCCATTTAAGGGAAGGGCTGGGCAA
   1  L  L  V  K  M  S  L  S  F  F  Y  K  P  F  K  G  R  A  G  Q

3000  TATAAGGTTGCAGCCATTATTCCCTCTTATAACGAAGATGCTGAGTCATTGCTAGAGACC
   1  Y  K  V  A  A  I  I  P  S  Y  N  E  D  A  E  S  L  L  E  T

3060  TTAAAAAGTGTTCAGCAGCAAACCTATCCCCTAGCAGAAATTTATGTTGTTGACGATGGA
   1  L  K  S  V  Q  Q  Q  T  Y  P  L  A  E  I  Y  V  V  D  D  G

3120  AGTGCTGATGAGACAGGTATTAAGCGCATTGAAGACTATGTGCGTGACACTGGTGACCTA
   1  S  A  D  E  T  G  I  K  R  I  E  D  Y  V  R  D  T  G  D  L

3180  TCAAGCAATGTCATTGTTCACCGGTCAGAAAAAAAATCAAGGAAAGCGTCATGCACAGGCC
   1  S  S  N  V  I  V  H  R  S  E  K  N  Q  G  K  R  H  A  Q  A

3240  TGGGCCTTTGAAAGATCAGACGCTGATGTCTTTTTGACCGTTGACTCAGATACTTATATC
   1  W  A  F  E  R  S  D  A  D  V  F  L  T  V  D  S  D  T  Y  I
```

```
3300  TACCCTGATGCTTTAGAGGAGTTGTTAAAAACCTTTAATGACCCAACTGTTTTTGCTGCG
   1  Y  P  D  A  L  E  E  L  L  K  T  F  N  D  P  T  V  F  A  A

3360  ACGGGTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACGCTTGACAGATATT
   1  T  G  H  L  N  V  R  N  R  Q  T  N  L  L  T  R  L  T  D  I

3420  CGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACAGGTAATATTCTC
   1  R  Y  D  N  A  F  G  V  E  R  A  A  Q  S  V  T  G  N  I  L

3480  GTTTGCTCAGGCCCGCTTAGCGTTTACAGACGCGAGGTGGTTGTTCCTAACATAGATAGA
   1  V  C  S  G  P  L  S  V  Y  R  R  E  V  V  V  P  N  I  D  R

3540  TACATCAACCAGACCTTCCTGGGTATTCCTGTAAGTATCGGTGATGACAGGTGCTTGACC
   1  Y  I  N  Q  T  F  L  G  I  P  V  S  I  G  D  D  R  C  L  T

3600  AACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAATGTATTACAGAT
   1  N  Y  A  T  D  L  G  K  T  V  Y  Q  S  T  A  K  C  I  T  D

3660  GTTCCTGACAAGATGTCTACTTACTTGAAGCAGCAAAACCGCTGGAACAAGTCCTTCTTT
   1  V  P  D  K  M  S  T  Y  L  K  Q  Q  N  R  W  N  K  S  F  F

3720  AGAGAGTCCATTATTTCTGTTAAGAAAATCATGAACAATCCTTTTGTAGCCCTATGGACC
   1  R  E  S  I  I  S  V  K  K  I  M  N  N  P  F  V  A  L  W  T

3780  ATACTTGAGGTGTCTATGTTTATGATGCTTGTTTATTCTGTGGTGGATTTCTTTGTAGGC
   1  I  L  E  V  S  M  F  M  M  L  V  Y  S  V  V  D  F  F  V  G

3840  AATGTCAGAGAATTTGATTGGCTCAGGGTTTTGGCCTTTCTGGTGATTATCTTCATTGTT
   1  N  V  R  E  F  D  W  L  R  V  L  A  F  L  V  I  I  F  I  V

3900  GCTCTTTGTCGTAATATTCACTATATGCTTAAGCACCCGCTGTCCTTCTTGTTATCTCCG
   1  A  L  C  R  N  I  H  Y  M  L  K  H  P  L  S  F  L  L  S  P

3960  TTTTATGGGGTACTGCATTTGTTTGTCCTACAGCCCTTGAAATTGTATTCTCTTTTTACT
   1  F  Y  G  V  L  H  L  F  V  L  Q  P  L  K  L  Y  S  L  F  T
                                                XbaI
4020  ATTAGAAATGCTGACTGGGGAACACGTAAAAAATTATTATAATCTAGAAATAATTTTGTT
   1  I  R  N  A  D  W  G  T  R  K  K  L  L

4080  TAACTTTAAGAAGGAGATATACATATGAAAAAAATAGCTGTCATTGGAACAGGTTATGTA
   1                        M  K  K  I  A  V  I  G  T  G  Y  V

4140  GGACTCGTATCAGGCACTTGCTTTGCGGAGATCGGCAATAAAGTTGTTTGCTGTGATATC
   1  G  L  V  S  G  T  C  F  A  E  I  G  N  K  V  V  C  C  D  I

4200  GATGAATCAAAAATCAGAAGCCTGAAAAATGGGGTAATCCCAATCTATGAACCAGGGCTT
   1  D  E  S  K  I  R  S  L  K  N  G  V  I  P  I  Y  E  P  G  L

4260  GCAGACTTAGTTGAAAAAAATGTGCTGGATCAGCGCCTGACCTTTACGAACGATATCCCG
   1  A  D  L  V  E  K  N  V  L  D  Q  R  L  T  F  T  N  D  I  P

4320  TCTGCCATTCGGGCCTCAGATATTATTTATATTGCAGTCGGAACGCCTATGTCCAAAACA
   1  S  A  I  R  A  S  D  I  I  Y  I  A  V  G  T  P  M  S  K  T

4380  GGTGAAGCTGATTTAACGTACGTCAAAGCGGCGGCGAAAACAATCGGTGAGCATCTTAAC
   1  G  E  A  D  L  T  Y  V  K  A  A  A  K  T  I  G  E  H  L  N

4440  GGCTACAAAGTGATCGTAAATAAAAGCACAGTCCCGGTTGGAACAGGGAAACTGGTGCAA
   1  G  Y  K  V  I  V  N  K  S  T  V  P  V  G  T  G  K  L  V  Q

                                                                EcoRI
4500  TCTATCGTTCAAAAAGCCTCAAAGGGGAGATACTCATTTGATGTTGTATCTAACCCTGAA
   1  S  I  V  Q  K  A  S  K  G  R  Y  S  F  D  V  V  S  N  P  E

4560  TTCCTTCGGGAAGGGTCAGCGATTCATGACACGATGAATATGGAGCGTGCCGTGATTGGT
   1  F  L  R  E  G  S  A  I  H  D  T  M  N  M  E  R  A  V  I  G

4620  TCAACAAGTCATAAAGCCGCTGCCATCATTGAGGAACTTCATCAGCCATTCCATGCTCCT
   1  S  T  S  H  K  A  A  A  I  I  E  E  L  H  Q  P  F  H  A  P

4680  GTCATTAAAACAAACCTAGAAAGTGCAGAAATGATTAAATACGCCGCGAATGCATTTCTG
   1  V  I  K  T  N  L  E  S  A  E  M  I  K  Y  A  A  N  A  F  L

4740  GCGACAAAGATTTCCTTTATCAACGATATCGCAAACATTTGTGAGCGAGTCGGCGCAGAC
   1  A  T  K  I  S  F  I  N  D  I  A  N  I  C  E  R  V  G  A  D

4800  GTTTCAAAAGTTGCTGATGGTGTTGGTCTTGACAGCCGTATCGGCAGAAAGTTCCTTAAA
   1  V  S  K  V  A  D  G  V  G  L  D  S  R  I  G  R  K  F  L  K
```

```
 4860  GCTGGTATTGGATTCGGCGGTTCATGTTTTCCAAAGGATACAACCGCGCTGCTTCAAATC
    1  A  G  I  G  F  G  G  S  C  F  P  K  D  T  T  A  L  L  Q  I

 4920  GCAAAATCGGCAGGCTATCCATTCAAGCTCATCGAAGCTGTCATTGAAACGAACGAAAAG
    1  A  K  S  A  G  Y  P  F  K  L  I  E  A  V  I  E  T  N  E  K

 4980  CAGCGTGTTCATATTGTAGATAAACTTTTGACTGTTATGGGAAGCGTCAAAGGGAGAACC
    1  Q  R  V  H  I  V  D  K  L  L  T  V  M  G  S  V  K  G  R  T

 5040  ATTTCAGTCCTGGGATTAGCCTTCAAACCGAATACGAACGATGTGAGATCCGCTCCAGCG
    1  I  S  V  L  G  L  A  F  K  P  N  T  N  D  V  R  S  A  P  A

 5100  CTTGATATTATCCCAATGCTGCAGCAGCTGGGCGCCCATGTAAAAGCATACGATCCGATT
    1  L  D  I  I  P  M  L  Q  Q  L  G  A  H  V  K  A  Y  D  P  I
                 HindIII
 5160  GCTATTCCTGAAGCTTCAGCGATCCTTGGCGAACAGGTCGAGTATTACACAGATGTGTAT
    1  A  I  P  E  A  S  A  I  L  G  E  Q  V  E  Y  Y  T  D  V  Y

 5220  GCTGCGATGGAAGACACTGATGCATGCCTGATTTTAACGGATTGGCCGGAAGTGAAAGAA
    1  A  A  M  E  D  T  D  A  C  L  I  L  T  D  W  P  E  V  K  E

 5280  ATGGAGCTTGTAAAAGTGAAAACCCTCTTAAAACAGCCAGTCATCATTGACGGCAGAAAT
    1  M  E  L  V  K  V  K  T  L  L  K  Q  P  V  I  I  D  G  R  N

 5340  TTATTTTCACTTGAAGAGATGCAGGCAGCCGGATACATTTATCACTCTATCGGCCGTCCC
    1  L  F  S  L  E  E  M  Q  A  A  G  Y  I  Y  H  S  I  G  R  P

 5400  GCTGTTCGGGGAACGGAACCCTCTGACAAGTATTTTCCGGGCTTGCCGCTTGAAGAATTG
    1  A  V  R  G  T  E  P  S  D  K  Y  F  P  G  L  P  L  E  E  L

                                        Nhe/XbaI      SmaI
 5460  GCTAAAGACTTGGGAAGCGTCAATTTATAAGCTAGAGTCGACGTCCCCGGGGCAGCCCGC
    1  A  K  D  L  G  S  V  N  L

 5520  CTAATGAGCGGGCTTTTTTCACGTCACGCGTCCATGGAGATCTTTGTCTGCAACTGAAAA
 5580  GTTTATACCTTACCTGGAACAAATGGTTGAAACATACGAGGCTAATATCGGCTTATTAGG
 5640  AATAGTCCCTGTACTAATAAAATCAGGTGGATCAGTTGATCAGTATATTTTGGACGAAGC
 5700  TCGGAAAGAATTTGGAGATGACTTGCTTAATTCCACAATTAAATTAAGGGAAAGAATAAA
 5760  GCGATTTGATGTTCAAGGAATCACGGAAGAAGATACTCATGATAAAGAAGCTCTAAAACT
 5820  ATTCAATAACCTTACAATGGAATTGATCGAAAGGGTGGAAGGTTAATGGTACGAAAATTA
                                          HindIII
 5880  GGGGATCTACCTAGAAAGCCACAAGGCGATAGGTCAAGCTTAAAGAACCCTTACATGGAT
 5940  CTTACAGATTCTGAAAGTAAAGAAACAACAGAGGTTAAACAAACAGAACCAAAAAGAAAA
 6000  AAAGCATTGTTGAAAACAATGAAAGTTGATGTTTCAATCCATAATAAGATTAAATCGCTG
                                     EcoRI
 6060  CACGAAATTCTGGCAGCATCCGAAGGGAATTCATATTACTTAGAGGATACTATTGAGAGA
 6120  GCTATTGATAAGATGGTTGAGACATTACCTGAGAGCCAAAAAACTTTTTATGAATATGAA
 6180  TTAAAAAAAAGAACCAACAAAGGCTGAGACAGACTCCAAACGAGTCTGTTTTTTTAAAAA
 6240  AAATATTAGGAGCATTGAATATATATTAGAGAATTAAGAAAGACATGGGAATAAAAATAT
 6300  TTTAAATCCAGTAAAAATATGATAAGATTATTTCAGAATATGAAGAACTCTGTTTGTTTT
 6360  TGATGAAAAAACAAACAAAAAAAATCCACCTAACGGAATCTCAATTTAACTAACAGCGGC
 6420  CAAACTGAGAAGTTAAATTTGAGAAGGGGAAAAGGCGGATTTATACTTGTATTTAACTAT
 6480  CTCCATTTTAACATTTTATTAAACCCCATACAAGTGAAAATCCTCTTTTACACTGTTCCT
 6540  TTAGGTGATCGCGGAGGGACATTATGAGTGAAGTAAACCTAAAAGGAAATACAGATGAAT
 6600  TAGTGTATTATCGACAGCAAACCACTGGAAATAAAATCGCCAGGAAGAGAATCAAAAAAG
 6660  GGAAAGAAGAAGTTTATTATGTTGCTGAAACGGAAGAGAAGATATGGACAGAAGAGCAAA
 6720  TAAAAAACTTTTCTTTAGACAAATTTGGTACGCATATACCTTACATAGAAGGTCATTATA
 6780  CAATCTTAAATAATTACTTCTTTGATTTTTGGGGCTATTTTTTAGGTGCTGAAGGAATTG
 6840  CGCTCTATGCTCACCTAACTCGTTATGCATACGGCAGCAAAGACTTTTGCTTTCCTAGTC
 6900  TACAAACAATCGCTAAAAAAATGGACAAGACTCCTGTTACAGTTAGAGGCTACTTGAAAC
 6960  TGCTTGAAAGGTACGGTTTTATTTGGAAGGTAAACGTCCGTAATAAAACCAAGGATAACA
 7020  CAGAGGAATCCCCGATTTTTAAGATTAGACGTAAGGTTCCTTTGCTTTCAGAAGAACTTT
 7080  TAAATGGAAACCCTAATATTGAAATTCCAGATGACGAGGAAGCACATGTAAAGAAGGCTT
 7140  TAAAAAAGGAAAAAGAGGGTCTTCCAAAGGTTTTGAAAAAAGAGCACGATGAATTTGTTA
 7200  AAAAAATGATGGATGAGTCAGAAACAATTAATATTCCAGAGGCCTTACAATATGACACAA
 7260  TGTATGAAGATATACTCAGTAAAGGAGAAATTCGAAAAGAAATCAAAAAACAAATACCTA
 7320  ATCCTACAACATCTTTTGAGAGTATATCAATGACAACTGAAGAGGAAAAAGTCGACAGTA
 7380  CTTTAAAAAGCGAAATGCAAAATCGTGTCTCTAAGCCTTCTTTTGATACCTGGTTTAAAA
 7440  ACACTAAGATCAAAATTGAAAATAAAAATTGTTTATTACTTGTACCGAGTGAATTTGCAT
 7500  TTGAATGGATTAAGAAAAGATATTTAGAAACAATTAAAACAGTCCTTGAAGAAGCTGGAT
 7560  ATGTTTTCGAAAAAATCGAACTAAGAAAAGTGCAATAAACTGCTGAAGTATTTCAGCAGT
 7620  TTTTTTTATTTAGAAATAGTGAAAAAAATATAATCAGGGAGGTATCAATATTTAATGAGT
 7680  ACTGATTTAAATTTATTTAGACTGGAATTAATAATTAACACGTAGACTAATTAAAATTTA
 7740  ATGAGGGATAAAGAGGATACAAAAATATTAATTTCAATCCCTATTAAATTTTAACAAGGG
 7800  GGGGATTAAAATTTAATTAGAGGTTTATCCACAAGAAAAGACCCTAATAAAATTTTTACT
 7860  AGGGTTATAACACTGATTAATTTCTTAATGGGGGAGGGATTAAAATTTAATGACAAAGAA
                          HindIII
 7920  AACAATCTTTTAAGAAAAGCTTTTAAAAGATAATAATAAAAAGAGCTTTGCGATTAAGCA
 7980  AAACTCTTTACTTTTTCATTGACATTATCAAATTCATCGATTTCAAATTGTTGTTGTATC
```

```
 8040  ATAAAGTTAATTCTGTTTTGCACAACCTTTTCAGGAATATAAAACACATCTGAGGCTTGT
 8100  TTTATAAACTCAGGGTCGCTAAAGTCAATGTAACGTAGCATATGATATGGTATAGCTTCC
 8160  ACCCAAGTTAGCCTTTCTGCTTCTTCTGAATGTTTTTCATATACTTCCATGGGTATCTCT
 8220  AAATGATTTTCCTCATGTAGCAAGGTATGAGCAAAAAGTTTATGGAATTGATAGTTCCTC
 8280  TCTTTTTCTTCAACTTTTTTATCTAAAACAAACACTTTAACATCTGAGTCAATGTAAGCA
 8340  TAAGATGTTTTTCCAGTCATAATTTCAATCCCAAATCTTTTAGACAGAAATTCTGGACGT
 8400  AAATCTTTTGGTGAAAGAATTTTTTTATGTAGCAATATATCCGATACAGCACCTTCTAAA
 8460  AGCGTTGGTGAATAGGGCATTTTACCTATCTCCTCTCATTTTGTGGAATAAAAATAGTCA
 8520  TATTCGTCCATCTACCTATCCTATTATCGAACAGTTGAACTTTTTAATCAAGGATCAGTC
 8580  CTTTTTTTCATTATTCTTAAACTGTGCTCTTAACTTTAACAACTCGATTTGTTTTTCCAG
 8640  ATCTCGAGGGTAACTAGCCTCGCCGATCCCGCAAGAGGCCCGGCAGTCAGGTGGCACTTT
 8700  TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTA
 8760  TCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTAT
 8820  GAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGT
 8880  TTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACG
 8940  AGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGA
 9000  AGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCG
 9060  TATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGT
 9120  TGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG
 9180  CAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGG
 9240  AGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA
 9300  TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCC
 9360  TGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTC
 9420  CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
 9480  GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG
 9540  CGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
 9600  GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTC
 9660  ACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTT
 9720  AAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGAC
 9780  CAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA
 9840  AGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACC
 9900  ACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGT
 9960  AACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGG
10020  CCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC
10080  AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTT
10140  ACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGA
10200  GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCT
10260  TCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCG
10320  CACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCA
10380  CCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAA
10440  CGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTT
10500  CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
10560  TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
10620  GCGCCCAATACG
```

Fig. 4.2 Complete nucleotide and amino acid sequences of plasmid pBS5.

In this nucleotide sequence, the *hasA* gene is present between 2808 and 4062bp and before this, it is present an Shine Dalgarno sequence. GGAGGA sequence, after *hasA* gene, is able to transcript the *tuaD* gene between 4105bp and 5490bp. The initial codon for TuaD has been mutated from Valine to Methionine. The plasmid, checked by restriction analyses with *EcoRI* and *HindIII* results in bands with the expected following molecular weights: 3957bp, 1650bp, 1650bp, 1243bp and 610bp.

4.2.4.3. pT7 *E.coli/B.megaterium* plasmids shuttle construction

Construction of *Bacillus megaterium* pT7 *hasA-tuaA*

To realize pT7 *hasA-tuaA*, the cassette-operon *hasA-tuaA* previously created, coding for HA synthetase and UDP-glucose -dehydrogenase, was amplified by PCR using pGEM4*hasA-tuaD* as template , and the following primers 5'GCTTGTACATGAGAACATTAAAAAACCTCA 3' and 5'AGGGATCCTTATAAATTGACGCTTCCCAAG 3'. The amplified PCR product was of

expected molecular weight and was digested with *BsrGI* and *BamHI* and cloned between the same sites in vector pT7, under the control of T7 promoter and of a very efficient Shine-Dalgarno sequence, resulting in plasmid pT7 *hasA-tuaA*.

pT7 *hasA-tuaA* *BsrGI*---------**HasA**---------*XbaI--NdeI*---------**TuaD**----------*BamHI-- SphI*

Here the complete nucleotide and amino acid sequences of plasmid pT7 *hasA-tuaA*:

```
   0  CTTTTTAGGTTCTAAATCGTGTTTTTCTTGGAATTGTGCTGTTTTATCCTTTACCTTGTC

  60  TACAAACCCCTTAAAAACGTTTTTAAAGGCTTTTAAGCCGTCTGTACGTTCCTTAAGGCG

 120  AAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCGAATATTAATTAACCAAG
                Bsp1407I
 180  GAGGTGAAATGTACAATGAGAACATTAAAAAACCTCATAACTGTTGTGGCCTTTAGTATT
   1                  M  R  T  L  K  N  L  I  T  V  V  A  F  S  I
                                                      HindIII
 240  TTTTGGGTACTGTTGATTTACGTCAATGTTTATCTCTTTGGTGCTAAAGGAAGCTTGTCA
   1  F  W  V  L  L  I  Y  V  N  V  Y  L  F  G  A  K  G  S  L  S

 300  ATTTATGGCTTTTTGCTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTTTAC
   1  I  Y  G  F  L  L  I  A  Y  L  L  V  K  M  S  L  S  F  F  Y

 360  AAGCCATTTAAGGGAAGGGCTGGGCAATATAAGGTTGCAGCCATTATTCCCTCTTATAAC
   1  K  P  F  K  G  R  A  G  Q  Y  K  V  A  A  I  I  P  S  Y  N

 420  GAAGATGCTGAGTCATTGCTAGAGACCTTAAAAAGTGTTCAGCAGCAAACCTATCCCCTA
   1  E  D  A  E  S  L  L  E  T  L  K  S  V  Q  Q  Q  T  Y  P  L

 480  GCAGAAATTTATGTTGTTGACGATGGAAGTGCTGATGAGACAGGTATTAAGCGCATTGAA
   1  A  E  I  Y  V  V  D  D  G  S  A  D  E  T  G  I  K  R  I  E

 540  GACTATGTGCGTGACACTGGTGACCTATCAAGCAATGTCATTGTTCACCGGTCAGAAAAA
   1  D  Y  V  R  D  T  G  D  L  S  S  N  V  I  V  H  R  S  E  K

 600  AATCAAGGAAAGCGTCATGCACAGGCCTGGGCCTTTGAAAGATCAGACGCTGATGTCTTT
   1  N  Q  G  K  R  H  A  Q  A  W  A  F  E  R  S  D  A  D  V  F

 660  TTGACCGTTGACTCAGATACTTATATCTACCCTGATGCTTTAGAGGAGTTGTTAAAAACC
   1  L  T  V  D  S  D  T  Y  I  Y  P  D  A  L  E  E  L  L  K  T

 720  TTTAATGACCCAACTGTTTTTGCTGCGACGGGTCACCTTAATGTCAGAAATAGACAAACC
   1  F  N  D  P  T  V  F  A  A  T  G  H  L  N  V  R  N  R  Q  T

 780  AATCTCTTAACACGCTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCT
   1  N  L  L  T  R  L  T  D  I  R  Y  D  N  A  F  G  V  E  R  A

 840  GCCCAATCCGTTACAGGTAATATTCTCGTTTGCTCAGGCCCGCTTAGCGTTTACAGACGC
   1  A  Q  S  V  T  G  N  I  L  V  C  S  G  P  L  S  V  Y  R  R

 900  GAGGTGGTTGTTCCTAACATAGATAGATACATCAACCAGACCTTCCTGGGTATTCCTGTA
   1  E  V  V  V  P  N  I  D  R  Y  I  N  Q  T  F  L  G  I  P  V

 960  AGTATCGGTGATGACAGGTGCTTGACCAACTATGCAACTGATTTAGGAAAGACTGTTTAT
   1  S  I  G  D  D  R  C  L  T  N  Y  A  T  D  L  G  K  T  V  Y

1020  CAATCCACTGCTAAATGTATTACAGATGTTCCTGACAAGATGTCTACTTACTTGAAGCAG
   1  Q  S  T  A  K  C  I  T  D  V  P  D  K  M  S  T  Y  L  K  Q

1080  CAAAACCGCTGGAACAAGTCCTTCTTTAGAGAGTCCATTATTTCTGTTAAGAAAATCATG
```

```
   1  Q  N  R  W  N  K  S  F  F  R  E  S  I  I  S  V  K  K  I  M

1140  AACAATCCTTTTGTAGCCCTATGGACCATACTTGAGGTGTCTATGTTTATGATGCTTGTT
   1  N  N  P  F  V  A  L  W  T  I  L  E  V  S  M  F  M  M  L  V

1200  TATTCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATTGGCTCAGGGTTTTG
   1  Y  S  V  V  D  F  F  V  G  N  V  R  E  F  D  W  L  R  V  L

1260  GCCTTTCTGGTGATTATCTTCATTGTTGCTCTTTGTCGTAATATTCACTATATGCTTAAG
   1  A  F  L  V  I  I  F  I  V  A  L  C  R  N  I  H  Y  M  L  K

1320  CACCCGCTGTCCTTCTTGTTATCTCCGTTTTATGGGGTACTGCTTTGTTTGTCCTACAGC
   1  H  P  L  S  F  L  L  S  P  F  Y  G  V  L  L  C  L  S  Y  S

1380  CCTTGAAATTGTATTCTCTTTTTACTATTAGAAATGCTGACTGGGGAACACGTAAAAAAT
   1  P
            XbaI                                              NdeI
1440  TATTATAATCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAAAAAAA
   3                                                         M  K  K  I

1500  TAGCTGTCATTGGAACAGGTTATGTAGGACTCGTATCAGGCACTTGCTTTGCGGAGATCG
   3    A  V  I  G  T  G  Y  V  G  L  V  S  G  T  C  F  A  E  I  G
                           EcoRV  ClaI
1560  GCAATAAAGTTGTTTGCTGTGATATCGATGAATCAAAAATCAGAAGCCTGAAAAATGGGG
   3    N  K  V  V  C  C  D  I  D  E  S  K  I  R  S  L  K  N  G  V

1620  TAATCCCAATCTATGAACCAGGGCTTGCAGACTTAGTTGAAAAAAATGTGCTGGATCAGC
   3    I  P  I  Y  E  P  G  L  A  D  L  V  E  K  N  V  L  D  Q  R
                        EcoRV
1680  GCCTGACCTTTACGAACGATATCCCGTCTGCCATTCGGGCCTCAGATATTATTTATATTG
   3    L  T  F  T  N  D  I  P  S  A  I  R  A  S  D  I  I  Y  I  A

1740  CAGTCGGAACGCCTATGTCCAAAACAGGTGAAGCTGATTTAACGTACGTCAAAGCGGCGG
   3    V  G  T  P  M  S  K  T  G  E  A  D  L  T  Y  V  K  A  A  A

1800  CGAAAACAATCGGTGAGCATCTTAACGGCTACAAAGTGATCGTAAATAAAAGCACAGTCC
   3    K  T  I  G  E  H  L  N  G  Y  K  V  I  V  N  K  S  T  V  P

1860  CGGTTGGAACAGGGAAACTGGTGCAATCTATCGTTCAAAAAGCCTCAAAGGGGAGATACT
   3    V  G  T  G  K  L  V  Q  S  I  V  Q  K  A  S  K  G  R  Y  S
                              EcoRI
1920  CATTTGATGTTGTATCTAACCCTGAATTCCTTCGGGAAGGGTCAGCGATTCATGACACGA
   3    F  D  V  V  S  N  P  E  F  L  R  E  G  S  A  I  H  D  T  M

1980  TGAATATGGAGCGTGCCGTGATTGGTTCAACAAGTCATAAAGCCGCTGCCATCATTGAGG
   3    N  M  E  R  A  V  I  G  S  T  S  H  K  A  A  A  I  I  E  E

2040  AACTTCATCAGCCATTCCATGCTCCTGTCATTAAAACAAACCTAGAAAGTGCAGAAATGA
   3    L  H  Q  P  F  H  A  P  V  I  K  T  N  L  E  S  A  E  M  I
                                                          EcoRV
2100  TTAAATACGCCGCGAATGCATTTCTGGCGACAAAGATTTCCTTTATCAACGATATCGCAA
   3    K  Y  A  A  N  A  F  L  A  T  K  I  S  F  I  N  D  I  A  N

2160  ACATTTGTGAGCGAGTCGGCGCAGACGTTTCAAAAGTTGCTGATGGTGTTGGTCTTGACA
   3    I  C  E  R  V  G  A  D  V  S  K  V  A  D  G  V  G  L  D  S

2220  GCCGTATCGGCAGAAAGTTCCTTAAAGCTGGTATTGGATTCGGCGGTTCATGTTTTCCAA
   3    R  I  G  R  K  F  L  K  A  G  I  G  F  G  G  S  C  F  P  K

2280  AGGATACAACCGCGCTGCTTCAAATCGCAAAATCGGCAGGCTATCCATTCAAGCTCATCG
   3    D  T  T  A  L  L  Q  I  A  K  S  A  G  Y  P  F  K  L  I  E

2340  AAGCTGTCATTGAAACGAACGAAAAGCAGCGTGTTCATATTGTAGATAAACTTTTGACTG
```

```
   3    A  V  I  E  T  N  E  K  Q  R  V  H  I  V  D  K  L  L  T  V

2400  TTATGGGAAGCGTCAAAGGGAGAACCATTTCAGTCCTGGGATTAGCCTTCAAACCGAATA
   3    M  G  S  V  K  G  R  T  I  S  V  L  G  L  A  F  K  P  N  T
                                                      PstI
2460  CGAACGATGTGAGATCCGCTCCAGCGCTTGATATTATCCCAATGCTGCAGCAGCTGGGCG
   3    N  D  V  R  S  A  P  A  L  D  I  I  P  M  L  Q  Q  L  G  A
                                          HindIII
2520  CCCATGTAAAAGCATACGATCCGATTGCTATTCCTGAAGCTTCAGCGATCCTTGGCGAAC
   3    H  V  K  A  Y  D  P  I  A  I  P  E  A  S  A  I  L  G  E  Q

                                                            SphI
2580  AGGTCGAGTATTACACAGATGTGTATGCTGCGATGGAAGACACTGATGCATGCCTGATTT
   3    V  E  Y  Y  T  D  V  Y  A  A  M  E  D  T  D  A  C  L  I  L

2640  TAACGGATTGGCCGGAAGTGAAAGAAATGGAGCTTGTAAAAGTGAAAACCCTCTTAAAAC
   3    T  D  W  P  E  V  K  E  M  E  L  V  K  V  K  T  L  L  K  Q

2700  AGCCAGTCATCATTGACGGCAGAAATTTATTTTCACTTGAAGAGATGCAGGCAGCCGGAT
   3    P  V  I  I  D  G  R  N  L  F  S  L  E  E  M  Q  A  A  G  Y

2760  ACATTTATCACTCTATCGGCCGTCCCGCTGTTCGGGGAACGGAACCCTCTGACAAGTATT
   3    I  Y  H  S  I  G  R  P  A  V  R  G  T  E  P  S  D  K  Y  F

                                                          BamHI
2820  TTCCGGGCTTGCCGCTTGAAGAATTGGCTAAAGACTTGGGAAGCGTCAATTTATAAGGAT
   3    P  G  L  P  L  E  E  L  A  K  D  L  G  S  V  N  L

               SphI
2880  CCGGCCGCATGCCGGCTAATCGCGACCGGTTAACTAGCATAACCCCTTGGGGCCTCTAAA

2940  CGGGTCTTGAGGGGTTTTTTGCTAAAGGAGGAACTATATCCGGTCCAAGAATTGGAGCCA

3000  ATCAATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGAACATATCCATCG

3060  CGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCCGCGTTGCTGGCGTTTTTC

3120  CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA

3180  AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT

3240  CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG

3300  GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG

3360  CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT

3420  CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC

3480  AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC

3540  TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC

3600  GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT

3660  TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC

3720  TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG

3780  AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA

3840  ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
```

3900 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG

3960 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC

4020 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC

4080 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
PstI
4140 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATC

Hpy99I
4200 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG

4260 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC

4320 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT
ScaI
4380 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG

4440 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGAT

4500 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG

4560 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA

4620 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA

4680 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC

4740 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA

4800 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG

4860 CCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATC
EcoRI
4920 ACGAGGCCCTTTCGTCTTCAAGAATTCCTGTTATAAAAAAAGGATCAATTTTGAACTCTC

4980 TCCCAAAGTTGATCCCTTAACGATTTAGAAATCCCTTTGAGAATGTTTATATACATTCAA
-
5040 GGTAACCAGCCAACTAATGACAATGATTCCTGAAAAAAGTAATAACAAATTACTATACAG

5100 ATAAGTTGACTGATCAACTTCCATAGGTAACAACCTTTGATCAAGTAAGGGTATGGATAA

5160 TAAACCACCTACAATTGCAATACCTGTTCCCTCTGATAAAAAGCTGGTAAAGTTAAGCAA

5220 ACTCATTCCAGCACCAGCTTCCTGCTGTTTCAAGCTACTTGAAACAATTGTTGATATAAC

5280 TGTTTTGGTGAACGAAAGCCCACCTAAAACAAATACGATTATAATTGTCATGAACCATGA

5340 TGTTGTTTCTAAAAGAAAGGAAGCAGTTAAAAAGCTAACAGAAAGAAATGTAACTCCGAT

5400 GTTTAACACGTATAAAGGACCTCTTCTATCAACAAGTATCCCACCAATGTAGCCGAAAAT
ScaI
5460 AATGACACTCATTGTTCCAGGGAAAATAATTACACTTCCGATTTCGGCAGTACTTAGCTG

5520 GTGAACATCTTTCATCATATAAGGAACCATAGAGACAAACCCTGCTACTGTTCCAAATAT

5580 AATTCCCCCACAAAGAACTCCAATCATAAAAGGTATATTTTTCCCTAATCCGGGATCAAC

5640 AAAAGGATCTGTTACTTTCCTGATATGTTTTACAAATATCAGGAATGACAGCACGCTAAC

5700 GATAAGAAAAGAAATGCTATATGATGTTGTAAACAACATAAAAAATACAATGCCTACAGA
EcoRV

5760 CATTAGTATAATTCCTTTGATATCAAAATGACCTTTTATCCTTACTTCTTTCTTTAATAA

5820 TTTCATAAGAAACGGAACAGTGATAATTGTTATCATAGGAATGAGTAGAAGATAGGACCA

5880 ATGAATATAATGGGCTATCATTCCACCAATCGCTGGACCGACTCCTTCTCCCATGGCTAC

ClaI

5940 TATCGATCCAATAAGACCAAATGCTTTACCCCTATTTTCCTTTGGAATATAGCGCGCAAC

6000 TACAACCATTACGAGTGCTGGAAATGCAGCTGCACCAGCCCCTTGAATAAAACGAGCCAT

6060 AATAAGTAAGGAAAAGAAAGAATGGCCAACAAACCCAATTACCGACCCGAAACAATTTAT

6120 TATAATTCCAAATAGGAGTAACCTTTTGATGCCTAATTGATCAGATAGCTTTCCATATAC

6180 AGCTGTTCCAATGGAAAAGGTTAACATAAAGGCTGTGTTCACCCAGTTTGTACTCGCAGG

6240 TGGTTTATTAAAATCATTTGCAATATCAGGTAATGAGACGTTCAAAACCATTTCATTTAA

6300 TACGCTAAAAAAAGATAAAATGCAAAGCCAAATTAAAATTTGGTTGTGTCGTAAATTCGA

6360 TTGTGAATAGGATGTATTCACATTTCACCCTCCAATAATGAGGGCAGACGTAGTTTATAG

6420 GGTTAATGATACGCTTCCCTCTTTTAATTGAACCCTGTTACATTCATTACACTTCATAAT

6480 TAATTCCTCCTAAACTTGATTAAAACATTTTACCACATATAAACTAAGTTTTAAATTCAG

6540 TATTTCATCACTTATACAACAATATGGCCCGTTTGTTGAACTACTCTTTAATAAAATAAT

6600 TTTTCCGTTCCCAATTCCACATTGCAATAATAGAAAATCCATCTTCATCGGCTTTTTCGT

6660 CATCATCTGTATGAATCAAATCGCCTTCTTCTGTGTCATCAAGGTTTAATTTTTTATGTA

6720 TTTCTTTTAACAAACCACCATAGGAGATTAACCTTTTACGGTGTAAACCTTCCTCCAAAT

6780 CAGACAAACGTTTCAAATTCTTTTCTTCATCATCGGTCATAAAATCCGTATCCTTTACAG

6840 GATATTTTGCAGTTTCGTCAATTGCCGATTGTATATCCGATTTATATTTATTTTTCGGTC

6900 GAATCATTTGAACTTTTACATTTGGATCATAGTCTAATTTCATTGCCTTTTTCCAAAATT

6960 GAATCCATTGTTTTTGATTCACGTAGTTTTCTGTATTCTTAAAATAAGTTGGTTCCACAC

7020 ATACCAATACATGCATGTGCTGATTATAAGAATTATCTTTATTATTTATTGTCACTTCCG

7080 TTGCACGCATAAAACCAACAAGATTTTTATTAATTTTTTTATATTGCATCATTCGGCGAA

7140 ATCCTTGAGCCATATCTGACAAACTCTTATTTAATTCTTCGCCATCATAAACATTTTTAA

7200 CTGTTAATGTGAGAAACAACCAACGAACTGTTGGCTTTTGTTTAATAACTTCAGCAACAA

7260 CCTTTTGTGACTGAATGCCATGTTTCATTGCTCTCCTCCAGTTGCACATTGGACAAAGCC

7320 TGGATTTACAAAACCACACTCGATACAACTTTCTTTCGCCTGTTTCACGATTTTGTTTAT

7380 ACTCTAATATTTCAGCACAATCTTTTACTCTTTCAGCCTTTTTAAATTCAAGAATATGCA

7440 GAAGTTCAAAGTAATCAACATTAGCGATTTTCTTTTCTCTCCATGGTCTCACTTTTCCAC

7500 TTTTTGTCTTGTCCACTAAAACCCTTGATTTTTCATCTGAATAAATGCTACTATTAGGAC

7560 ACATAATATTAAAAGAAACCCCCATCTATTTAGTTATTTGTTTGGTCACTTATAACTTTA

7620 ACAGATGGGGTTTTTCTGTGCAACCAATTTTAAGGGTTTTCAATACTTTAAAACACATAC

7680 ATACCAACACTTCAACGCACCTTTCAGCAACTAAAATAAAAATGACGTTATTTCTATATG

7740 TATCAAGATAAGAAAGAACAAGTTCAAAACCATCAAAAAAAGACACCTTTTCAGGTGCTT

7800 TTTTTATTTTATAAACTCATTCCCTGATCTCGACTTCGTTCTTTTTTTACCTCTCGGTTA

7860 TGAGTTAGTTCAAATTCGTT

Fig. 4.3 Complete nucleotide and amino acid sequences of plasmid pT7*hasA-tuaD*.

the control of strong promoter pT7. The sequence for *hasA* from *Streptococcus equi* is present between 196bp and 1383. The sequence for *tuaD* from *Bacillus subtilis* is present between 1430 and 2873bp.

UDP-Glucose pyrophosphorylase gene(*gtaB*) cloning from *Bacillus subtilis* in pGEM4*hasA-gtaB*

GtaB gene was amplified by PCR using the *Bacillus subtilis* 168 strain chromosome DNA as template, and the following primers 5'ATGTCTAGAATAATAAGGAAGGTGCCTTTTAAATGAA 3' and 5'CTCTCGAGCTAGCTTAGATTTCTTCTTTGTTTAGTAAAG 3'. The amplified PCR product was of expected molecular weight (925bp) and was cut by *XbaI* and *XhoI* restriction enzymes, and cloned between *XbaI* and *SalI* of plasmid pGEM*hasA*, resulting in pGEM4*hasA-gtaB*.

pGEM4*hasA-gtaB* *HindIII-BamHI* **---------HasA**----------*XbaI*-----**GtaB**-----*NheI--XhoI/SalI-PstI*

Glucose pyrophosphorylase gene (*pgi*) cloning from *Bacillus subtilis* in pRSETB*gpi*

Pgi gene was amplified by PCR using the *Bacillus subtilis* 168 strain chromosome DNA as template, and the following primers 5'TACATATGACGCATGTACGCTTGACTACTCCAAAAG 3' and 5'ATGCTAGCTCATTTATAATCTTCCAGACGTTTTTCAAG 3'. The amplified PCR product was cut by *NdeI* and *NheI* restriction enzymes and cloned in the same restriction sites of plasmid pRSETB, resulting pRSETB*pgi*. So the *pgi* gene is under the control of T7 promoter.

pRSETB*gpi* *XbaI-- NdeI*----------**gpi** ---------- *NheI –BamHI- BglI –XhoI- PstI*

Construction of *B. megaterium* pT7 *hasA-tuaD-gtaB-pgi* or pT7hyal

pRSETB*gpi* was cleavage with *XbaI* and *PstI* and the resulting fragment (1340 bp) has been cloned in the plasmid pGEM4*hasA-gtaB* between *NheI* and *PstI* resulting pGEM4*hasA-gtaB-pgi*.

pGEM4*hasA-gtaB-pgi* *HindIII-BamHI* ----------**HasA**----------*XbaI*-----**GtaB**-----*Nhe/XbaI-- NdeI*----------**gpi** ---------- *NheI –BamHI-- BglI –XhoI-PstI*

pGEM4*hasA-gtaB-pgi* was cleavage with *XbaI* and *XhoI* and the resulting fragment containing *gtaB* and *pgi* sequences has been cloned in pRSETB*tuaD* between *NheI* and *XhoI* restriction sites resulting plasmid pRSETB*tuaD -gtaB-pgi*.

pRSETB*tuaD-gtaB-pgi* *XbaI-- NdeI*----------**TuaD**---------- *NheI/ XbaI*-----**GtaB**----*Nhe/XbaI-- NdeI*----------**gpi** ---------- *NheI –BamHI-- BglI –XhoI*

To construct the *Bacillus megaterium* expression vector, pRSETB*tuaD-gtaB-pgi* was cleavage with *XbaI* and *BamHI* and the outing fragment containing tuaD, *gtaB* and *pgi* sequences has been cloned in pT7*hasA-tuaD* between the same restriction sites, resulting pT7 *hasA-tuaD-gtaB-pgi* or pT7hyal.

pT7 *hasA-tuaD-gtaB-pgi* *BsrGI*---------**HasA**--------- *XbaI-- NdeI*----------**TuaD**---------- *NheI/ XbaI*-----**GtaB**-----*Nhe/XbaI-- NdeI*----------**gpi** ---------- *NheI –BamHI-- SphI*

Here the complete nucleotide and amino acid sequences of plasmid pT7 *hasA-tuaA-gtaB-pgi:*

```
   0  CTTTTTAGGTTCTAAATCGTGTTTTTCTTGGAATTGTGCTGTTTTATCCTTTACCTTGTC

  60  TACAAACCCCTTAAAAACGTTTTTAAAGGCTTTTAAGCCGTCTGTACGTTCCTTAAGGCG

 120  AAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCGAATATTAATTAACCAAG
                Bsp1407I
 180  GAGGTGAAATGTACAATGAGAACATTAAAAAACCTCATAACTGTTGTGGCCTTTAGTATT
   1                  M  R  T  L  K  N  L  I  T  V  V  A  F  S  I
                                                          HindIII
 240  TTTTGGGTACTGTTGATTTACGTCAATGTTTATCTCTTTGGTGCTAAAGGAAGCTTGTCA
   1  F  W  V  L  L  I  Y  V  N  V  Y  L  F  G  A  K  G  S  L  S

 300  ATTTATGGCTTTTTGCTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTTTAC
   1  I  Y  G  F  L  L  I  A  Y  L  L  V  K  M  S  L  S  F  F  Y

 360  AAGCCATTTAAGGGAAGGGCTGGGCAATATAAGGTTGCAGCCATTATTCCCTCTTATAAC
   1  K  P  F  K  G  R  A  G  Q  Y  K  V  A  A  I  I  P  S  Y  N

 420  GAAGATGCTGAGTCATTGCTAGAGACCTTAAAAAGTGTTCAGCAGCAAACCTATCCCCTA
   1  E  D  A  E  S  L  L  E  T  L  K  S  V  Q  Q  Q  T  Y  P  L

 480  GCAGAAATTTATGTTGTTGACGATGGAAGTGCTGATGAGACAGGTATTAAGCGCATTGAA
   1  A  E  I  Y  V  V  D  D  G  S  A  D  E  T  G  I  K  R  I  E

 540  GACTATGTGCGTGACACTGGTGACCTATCAAGCAATGTCATTGTTCACCGGTCAGAAAAA
   1  D  Y  V  R  D  T  G  D  L  S  S  N  V  I  V  H  R  S  E  K

 600  AATCAAGGAAAGCGTCATGCACAGGCCTGGGCCTTTGAAAGATCAGACGCTGATGTCTTT
```

```
   1  N  Q  G  K  R  H  A  Q  A  W  A  F  E  R  S  D  A  D  V  F

 660  TTGACCGTTGACTCAGATACTTATATCTACCCTGATGCTTTAGAGGAGTTGTTAAAAACC
   1  L  T  V  D  S  D  T  Y  I  Y  P  D  A  L  E  E  L  L  K  T

 720  TTTAATGACCCAACTGTTTTTGCTGCGACGGGTCACCTTAATGTCAGAAATAGACAAACC
   1  F  N  D  P  T  V  F  A  A  T  G  H  L  N  V  R  N  R  Q  T

 780  AATCTCTTAACACGCTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCT
   1  N  L  L  T  R  L  T  D  I  R  Y  D  N  A  F  G  V  E  R  A

 840  GCCCAATCCGTTACAGGTAATATTCTCGTTTGCTCAGGCCCGCTTAGCGTTTACAGACGC
   1  A  Q  S  V  T  G  N  I  L  V  C  S  G  P  L  S  V  Y  R  R

 900  GAGGTGGTTGTTCCTAACATAGATAGATACATCAACCAGACCTTCCTGGGTATTCCTGTA
   1  E  V  V  V  P  N  I  D  R  Y  I  N  Q  T  F  L  G  I  P  V

 960  AGTATCGGTGATGACAGGTGCTTGACCAACTATGCAACTGATTTAGGAAAGACTGTTTAT
   1  S  I  G  D  D  R  C  L  T  N  Y  A  T  D  L  G  K  T  V  Y

1020  CAATCCACTGCTAAATGTATTACAGATGTTCCTGACAAGATGTCTACTTACTTGAAGCAG
   1  Q  S  T  A  K  C  I  T  D  V  P  D  K  M  S  T  Y  L  K  Q

1080  CAAAACCGCTGGAACAAGTCCTTCTTTAGAGAGTCCATTATTTCTGTTAAGAAAATCATG
   1  Q  N  R  W  N  K  S  F  F  R  E  S  I  I  S  V  K  K  I  M

1140  AACAATCCTTTTGTAGCCCTATGGACCATACTTGAGGTGTCTATGTTTATGATGCTTGTT
   1  N  N  P  F  V  A  L  W  T  I  L  E  V  S  M  F  M  M  L  V

1200  TATTCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATTGGCTCAGGGTTTTG
   1  Y  S  V  V  D  F  F  V  G  N  V  R  E  F  D  W  L  R  V  L

1260  GCCTTTCTGGTGATTATCTTCATTGTTGCTCTTTGTCGTAATATTCACTATATGCTTAAG
   1  A  F  L  V  I  I  F  I  V  A  L  C  R  N  I  H  Y  M  L  K

1320  CACCCGCTGTCCTTCTTGTTATCTCCGTTTTATGGGGTACTGCTTTGTTTGTCCTACAGC
   1  H  P  L  S  F  L  L  S  P  F  Y  G  V  L  L  C  L  S  Y  S

1380  CCTTGAAATTGTATTCTCTTTTTACTATTAGAAATGCTGACTGGGGAACACGTAAAAAAT
   1  P
            XbaI                                          NdeI
1440  TATTATAATCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAAAAAAA
   3                                                     M  K  K  I

1500  TAGCTGTCATTGGAACAGGTTATGTAGGACTCGTATCAGGCACTTGCTTTGCGGAGATCG
   3    A  V  I  G  T  G  Y  V  G  L  V  S  G  T  C  F  A  E  I  G
                          EcoRV ClaI
1560  GCAATAAAGTTGTTTGCTGTGATATCGATGAATCAAAAATCAGAAGCCTGAAAAATGGGG
   3    N  K  V  V  C  C  D  I  D  E  S  K  I  R  S  L  K  N  G  V

1620  TAATCCCAATCTATGAACCAGGGCTTGCAGACTTAGTTGAAAAAAATGTGCTGGATCAGC
   3    I  P  I  Y  E  P  G  L  A  D  L  V  E  K  N  V  L  D  Q  R
                       EcoRV
1680  GCCTGACCTTTACGAACGATATCCCGTCTGCCATTCGGGCCTCAGATATTATTTATATTG
   3    L  T  F  T  N  D  I  P  S  A  I  R  A  S  D  I  I  Y  I  A

1740  CAGTCGGAACGCCTATGTCCAAAACAGGTGAAGCTGATTTAACGTACGTCAAAGCGGCGG
   3    V  G  T  P  M  S  K  T  G  E  A  D  L  T  Y  V  K  A  A  A

1800  CGAAAACAATCGGTGAGCATCTTAACGGCTACAAAGTGATCGTAAATAAAAGCACAGTCC
   3    K  T  I  G  E  H  L  N  G  Y  K  V  I  V  N  K  S  T  V  P

1860  CGGTTGGAACAGGGAAACTGGTGCAATCTATCGTTCAAAAAGCCTCAAAGGGGAGATACT
```

```
   3    V  G  T  G  K  L  V  Q  S  I  V  Q  K  A  S  K  G  R  Y  S
                              EcoRI
1920  CATTTGATGTTGTATCTAACCCTGAATTCCTTCGGGAAGGGTCAGCGATTCATGACACGA
   3    F  D  V  V  S  N  P  E  F  L  R  E  G  S  A  I  H  D  T  M

1980  TGAATATGGAGCGTGCCGTGATTGGTTCAACAAGTCATAAAGCCGCTGCCATCATTGAGG
   3    N  M  E  R  A  V  I  G  S  T  S  H  K  A  A  A  I  I  E  E

2040  AACTTCATCAGCCATTCCATGCTCCTGTCATTAAAACAAACCTAGAAAGTGCAGAAATGA
   3    L  H  Q  P  F  H  A  P  V  I  K  T  N  L  E  S  A  E  M  I
                                                       EcoRV
2100  TTAAATACGCCGCGAATGCATTTCTGGCGACAAAGATTTCCTTTATCAACGATATCGCAA
   3    K  Y  A  A  N  A  F  L  A  T  K  I  S  F  I  N  D  I  A  N

2160  ACATTTGTGAGCGAGTCGGCGCAGACGTTTCAAAAGTTGCTGATGGTGTTGGTCTTGACA
   3    I  C  E  R  V  G  A  D  V  S  K  V  A  D  G  V  G  L  D  S

2220  GCCGTATCGGCAGAAAGTTCCTTAAAGCTGGTATTGGATTCGGCGGTTCATGTTTTCCAA
   3    R  I  G  R  K  F  L  K  A  G  I  G  F  G  G  S  C  F  P  K

2280  AGGATACAACCGCGCTGCTTCAAATCGCAAAATCGGCAGGCTATCCATTCAAGCTCATCG
   3    D  T  T  A  L  L  Q  I  A  K  S  A  G  Y  P  F  K  L  I  E

2340  AAGCTGTCATTGAAACGAACGAAAAGCAGCGTGTTCATATTGTAGATAAACTTTTGACTG
   3    A  V  I  E  T  N  E  K  Q  R  V  H  I  V  D  K  L  L  T  V

2400  TTATGGGAAGCGTCAAAGGGAGAACCATTTCAGTCCTGGGATTAGCCTTCAAACCGAATA
   3    M  G  S  V  K  G  R  T  I  S  V  L  G  L  A  F  K  P  N  T
                                                       PstI
2460  CGAACGATGTGAGATCCGCTCCAGCGCTTGATATTATCCCAATGCTGCAGCAGCTGGGCG
   3    N  D  V  R  S  A  P  A  L  D  I  I  P  M  L  Q  Q  L  G  A
                                        HindIII
2520  CCCATGTAAAAGCATACGATCCGATTGCTATTCCTGAAGCTTCAGCGATCCTTGGCGAAC
   3    H  V  K  A  Y  D  P  I  A  I  P  E  A  S  A  I  L  G  E  Q

                                                       SphI
2580  AGGTCGAGTATTACACAGATGTGTATGCTGCGATGGAAGACACTGATGCATGCCTGATTT
   3    V  E  Y  Y  T  D  V  Y  A  A  M  E  D  T  D  A  C  L  I  L

2640  TAACGGATTGGCCGGAAGTGAAAGAAATGGAGCTTGTAAAAGTGAAAACCCTCTTAAAAC
   3    T  D  W  P  E  V  K  E  M  E  L  V  K  V  K  T  L  L  K  Q

2700  AGCCAGTCATCATTGACGGCAGAAATTTATTTTCACTTGAAGAGATGCAGGCAGCCGGAT
   3    P  V  I  I  D  G  R  N  L  F  S  L  E  E  M  Q  A  A  G  Y

2760  ACATTTATCACTCTATCGGCCGTCCCGCTGTTCGGGGAACGGAACCCTCTGACAAGTATT
   3    I  Y  H  S  I  G  R  P  A  V  R  G  T  E  P  S  D  K  Y  F

2820  TTCCGGGCTTGCCGCTTGAAGAATTGGCTAAAGACTTGGGAAGCGTCAATTTATAAGCTA
   3    P  G  L  P  L  E  E  L  A  K  D  L  G  S  V  N  L

2880  GAATAATAAGGAAGGTGCCTTTTAAATGAAAAAAGTACGTAAAGCCATAATTCCAGCAGC
   2                         M  K  K  V  R  K  A  I  I  P  A  A

2940  AGGCTTAGGAACACGTTTTCTTCCGGCTACGAAAGCAATGCCGAAAGAAATGCTTCCTAT
   2   G  L  G  T  R  F  L  P  A  T  K  A  M  P  K  E  M  L  P  I

3000  CGTTGATAAACCTACCATTCAATACATAATTGAAGAAGCTGTTGAAGCCGGTATTGAAGA
   2   V  D  K  P  T  I  Q  Y  I  I  E  E  A  V  E  A  G  I  E  D

3060  TATTATTATCGTAACAGGAAAAAGCAAGCGTGCGATTGAGGATCATTTTGATTACTCTCC
   2   I  I  I  V  T  G  K  S  K  R  A  I  E  D  H  F  D  Y  S  P
```

```
3120  TGAGCTTGAAAGAAACCTAGAAGAAAAAGGAAAAACTGAGCTGCTTGAAAAAGTGAAAAA
   2   E  L  E  R  N  L  E  E  K  G  K  T  E  L  L  E  K  V  K  K

3180  GGCTTCTAACCTGGCTGACATTCACTATATCCGCCAAAAAGAACCTAAAGGTCTCGGACA
   2   A  S  N  L  A  D  I  H  Y  I  R  Q  K  E  P  K  G  L  G  H

3240  TGCTGTCTGGTGCGCACGCAACTTTATCGGCGATGAGCCGTTTGCGGTACTGCTTGGTGA
   2   A  V  W  C  A  R  N  F  I  G  D  E  P  F  A  V  L  L  G  D

3300  CGATATTGTTCAGGCTGAAACTCCAGGGTTGCGCCAATTAATGGATGAATATGAAAAAAC
   2   D  I  V  Q  A  E  T  P  G  L  R  Q  L  M  D  E  Y  E  K  T

3360  ACTTTCTTCTATTATCGGTGTTCAGCAGGTGCCCGAAGAAGAAACACACCGCTACGGCAT
   2   L  S  S  I  I  G  V  Q  Q  V  P  E  E  E  T  H  R  Y  G  I

3420  TATTGACCCGCTGACAAGTGAAGGCCGCCGTTATCAGGTGAAAAACTTCGTTGAAAAACC
   2   I  D  P  L  T  S  E  G  R  R  Y  Q  V  K  N  F  V  E  K  P

3480  GCCTAAAGGCACAGCACCTTCTAATCTTGCCATCTTAGGCCGTTACGTATTCACGCCTGA
   2   P  K  G  T  A  P  S  N  L  A  I  L  G  R  Y  V  F  T  P  E
      BglII
3540  GATCTTCATGTATTTAGAAGAGCAGCAGGTTGGCGCCGGCGGAGAAATTCAGCTCACAGA
   2   I  F  M  Y  L  E  E  Q  Q  V  G  A  G  G  E  I  Q  L  T  D

3600  CGCCATTCAAAAGCTGAATGAAATTCAAAGAGTGTTTGCTTACGATTTTGAAGGCAAGCG
   2   A  I  Q  K  L  N  E  I  Q  R  V  F  A  Y  D  F  E  G  K  R

3660  TTATGATGTTGGTGAAAAGCTCGGCTTTATCACAACAACTCTTGAATTTGCGATGCAGGA
   2   Y  D  V  G  E  K  L  G  F  I  T  T  T  L  E  F  A  M  Q  D

3720  TAAAGAGCTTCGCGATCAGCTCGTTCCATTTATGGAAGGTTTACTAAACAAAGAAGAAAT
   2   K  E  L  R  D  Q  L  V  P  F  M  E  G  L  L  N  K  E  E  I

                                                        NdeI
3780  CTAAGCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACGCATGTACG
   2                                                     M  T  H  V  R

3840  CTTGACTACTCCAAAAGCGTTGACTTTCTTTCCAACGGAACATGAACTTACATACCTGCG
   2   L  T  T  P  K  A  L  T  F  F  P  T  E  H  E  L  T  Y  L  R

3900  GGACTTTGTAAAAACAGCACACCATAATATCCATGAGAAAACAGGCGCGGGCAGCGATTT
   2   D  F  V  K  T  A  H  H  N  I  H  E  K  T  G  A  G  S  D  F
                                                 EcoRI
3960  TCTAGGCTGGGTGGACCTCCCTGAACATTATGATAAAGAAGAATTCGCGCGCATCCAAAA
   2   L  G  W  V  D  L  P  E  H  Y  D  K  E  E  F  A  R  I  Q  K

4020  AAGCGCGGAAAAAATCCAATCTGACTCTGATGTCTTGCTTGTTGTCGGCATCGGCGGTTC
   2   S  A  E  K  I  Q  S  D  S  D  V  L  L  V  V  G  I  G  G  S

4080  TTATCTTGGAGCGCGGGCAGCGATTGAAGCGCTGAATCACGCGTTTTATAACACTTTGCC
   2   Y  L  G  A  R  A  A  I  E  A  L  N  H  A  F  Y  N  T  L  P

4140  AAAAGCCAAACGCGGCAATCCGCAAGTCATTTTTAACTTCTCTATTAATGTGATTTCTAA
   2   K  A  K  R  G  N  P  Q  V  I  F  N  F  S  I  N  V  I  S  K
                                                    HindIII
4200  ATCAGGTACGACAACTGAACCTGCAATCGCTTTCCGTATTTTCCGCAAGCTTCTTGAAGA
   2   S  G  T  T  T  E  P  A  I  A  F  R  I  F  R  K  L  L  E  E

4260  GAAATACGGTAAAGAAGAAGCGAAAGCGCGGATTTATGCAACAACTGATAAAGAGCGCGG
   2   K  Y  G  K  E  E  A  K  A  R  I  Y  A  T  T  D  K  E  R  G

4320  CGCATTAAAAACGCTTTCTAACGAAGAAGGCTTTGAATCATTCGTAATTCCTGACGATGT
   2   A  L  K  T  L  S  N  E  E  G  F  E  S  F  V  I  P  D  D  V
```

```
4380  CGGCGGCCGTTATTCAGTTTTAACAGCTGTAGGTCTCTTGCCGATTGCTGTCAGCGGCGT
   2   G  G  R  Y  S  V  L  T  A  V  G  L  L  P  I  A  V  S  G  V

4440  CAACATTGACGACATGATGAAAGGCGCCCTGGATGCGAGCAAAGATTTTGCAACATCTGA
   2   N  I  D  D  M  M  K  G  A  L  D  A  S  K  D  F  A  T  S  E

4500  ACTGGAAGATAACCCAGCATACCAATATGCGGTTGTTCGCAATGTCCTTTATAATAAGGG
   2   L  E  D  N  P  A  Y  Q  Y  A  V  V  R  N  V  L  Y  N  K  G

4560  CAAAACAATTGAAATGCTCATCAACTACGAACCGGCGCTTCAATACTTTGCGGAATGGTG
   2   K  T  I  E  M  L  I  N  Y  E  P  A  L  Q  Y  F  A  E  W  W

4620  GAAGCAGCTGTTCGGAGAAAGCGAAGGGAAAGATGAGAAGGGCATTTATCCTTCTTCAGC
   2   K  Q  L  F  G  E  S  E  G  K  D  E  K  G  I  Y  P  S  S  A

4680  GAACTATTCAACAGACCTTCATTCTTTAGGCCAGTATGTACAAGAAGGCCGCAGAGATTT
   2   N  Y  S  T  D  L  H  S  L  G  Q  Y  V  Q  E  G  R  R  D  L

4740  ATTCGAAACGGTCCTGAACGTAGAGAAGCCTAAACATGAACTGACAATTGAGGAAGCGGA
   2   F  E  T  V  L  N  V  E  K  P  K  H  E  L  T  I  E  E  A  D

4800  TAACGATCTTGACGGCTTGAACTATTTAGCCGGTAAAACTGTTGATTTCGTTAACAAAAA
   2   N  D  L  D  G  L  N  Y  L  A  G  K  T  V  D  F  V  N  K  K

4860  AGCATTCCAAGGTACAATGCTTGCCCATACAGACGGAAATGTTCCGAACTTAATCGTTAA
   2   A  F  Q  G  T  M  L  A  H  T  D  G  N  V  P  N  L  I  V  N

4920  CATTCCTGAGCTGAATGCATATACTTTTGGATACCTTGTATATTTCTTCGAAAAAGCCTG
   2   I  P  E  L  N  A  Y  T  F  G  Y  L  V  Y  F  F  E  K  A  C

4980  CGCGATGAGCGGTTACCTCCTTGGCGTCAATCCGTTTGACCAGCCTGGTGTAGAAGCGTA
   2   A  M  S  G  Y  L  L  G  V  N  P  F  D  Q  P  G  V  E  A  Y

5040  TAAAGTCAATATGTTTGCGTTACTCGGCAAACCTGGCTTTGAAGAGAAAAAAGCAGAGCT
   2   K  V  N  M  F  A  L  L  G  K  P  G  F  E  E  K  K  A  E  L

                                    NheI

5100  TGAAAAACGTCTGGAAGATTATAAATGAGCTAGCATGACTGGTGGACAGCAAATGGGTCG
   2   E  K  R  L  E  D  Y  K

                              BamHI KpnI     SphI          AgeI
5160  GGATCTGTACGACGATGACGATAAGGATCCGGTACCGGCCGCATGCCGGCTAATCGCGAC

5220  CGGTTAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTAAA

5280  GGAGGAACTATATCCGGTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAAT

5340  GCGCAAACCAACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACG

5400  CGGCGCATCTCGGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA

5460  TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA

5520  GGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG

5580  ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAG

5640  GTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT

5700  TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
```

5760 CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG

5820 CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT

5880 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC

5940 CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCG

6000 CAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG

6060 GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTA

6120 GATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTG

6180 GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG

6240 TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACC

6300 ATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATC

6360 AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC

6420 CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAG

PstI

6480 TTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTAT

6540 GGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTG

6600 CAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGT

6660 GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAG

ScaI

6720 ATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG

6780 ACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTT

6840 AAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCT

6900 GTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTAC

6960 TTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT

7020 AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCAT

7080 TTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACA

7140 AATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTAT

EcoRI

7200 TATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATT

7260 CCTGTTATAAAAAAAGGATCAATTTTGAACTCTCTCCCAAAGTTGATCCCTTAACGATTT

7320 AGAAATCCCTTTGAGAATGTTTATATACATTCAAGGTAACCAGCCAACTAATGACAATGA

7380 TTCCTGAAAAAAGTAATAACAAATTACTATACAGATAAGTTGACTGATCAACTTCCATAG

7440 GTAACAACCTTTGATCAAGTAAGGGTATGGATAATAAACCACCTACAATTGCAATACCTG

7500 TTCCCTCTGATAAAAAGCTGGTAAAGTTAAGCAAACTCATTCCAGCACCAGCTTCCTGCT

7560 GTTTCAAGCTACTTGAAACAATTGTTGATATAACTGTTTTGGTGAACGAAAGCCCACCTA

7620 AAACAAATACGATTATAATTGTCATGAACCATGATGTTGTTTCTAAAAGAAAGGAAGCAG

7680 TTAAAAAGCTAACAGAAAGAAATGTAACTCCGATGTTTAACACGTATAAAGGACCTCTTC

7740 TATCAACAAGTATCCCACCAATGTAGCCGAAAATAATGACACTCATTGTTCCAGGGAAAA

ScaI

7800 TAATTACACTTCCGATTTCGGCAGTACTTAGCTGGTGAACATCTTTCATCATATAAGGAA

7860 CCATAGAGACAAACCCTGCTACTGTTCCAAATATAATTCCCCCACAAAGAACTCCAATCA

7920 TAAAAGGTATATTTTTCCCTAATCCGGGATCAACAAAAGGATCTGTTACTTTCCTGATAT

7980 GTTTTACAAATATCAGGAATGACAGCACGCTAACGATAAGAAAAGAAATGCTATATGATG

EcoRV

8040 TTGTAAACAACATAAAAAAATACAATGCCTACAGACATTAGTATAATTCCTTTGATATCAA

8100 AATGACCTTTTATCCTTACTTCTTTCTTTAATAATTTCATAAGAAACGGAACAGTGATAA

8160 TTGTTATCATAGGAATGAGTAGAAGATAGGACCAATGAATATAATGGGCTATCATTCCAC

8220 CAATCGCTGGACCGACTCCTTCTCCCATGGCTACTATCGATCCAATAAGACCAAATGCTT

8280 TACCCCTATTTTCCTTTGGAATATAGCGCGCAACTACAACCATTACGAGTGCTGGAAATG

8340 CAGCTGCACCAGCCCCTTGAATAAAACGAGCCATAATAAGTAAGGAAAAGAAAGAATGGC

8400 CAACAAACCCAATTACCGACCCGAAACAATTTATTATAATTCCAAATAGGAGTAACCTTT

8460 TGATGCCTAATTGATCAGATAGCTTTCCATATACAGCTGTTCCAATGGAAAAGGTTAACA

8520 TAAAGGCTGTGTTCACCCAGTTTGTACTCGCAGGTGGTTTATTAAAATCATTTGCAATAT

8580 CAGGTAATGAGACGTTCAAAACCATTTCATTTAATACGCTAAAAAAAGATAAAATGCAAA

8640 GCCAAATTAAAATTTGGTTGTGTCGTAAATTCGATTGTGAATAGGATGTATTCACATTTC

8700 ACCCTCCAATAATGAGGGCAGACGTAGTTTATAGGGTTAATGATACGCTTCCCTCTTTTA

8760 ATTGAACCCTGTTACATTCATTACACTTCATAATTAATTCCTCCTAAACTTGATTAAAAC

8820 ATTTTACCACATATAAACTAAGTTTTAAATTCAGTATTTCATCACTTATACAACAATATG

8880 GCCCGTTTGTTGAACTACTCTTTAATAAAATAATTTTTCCGTTCCCAATTCCACATTGCA

8940 ATAATAGAAAATCCATCTTCATCGGCTTTTTCGTCATCATCTGTATGAATCAAATCGCCT

9000 TCTTCTGTGTCATCAAGGTTTAATTTTTTATGTATTTCTTTTAACAAACCACCATAGGAG

9060 ATTAACCTTTTACGGTGTAAACCTTCCTCCAAATCAGACAAACGTTTCAAATTCTTTTCT

9120 TCATCATCGGTCATAAAATCCGTATCCTTTACAGGATATTTTGCAGTTTCGTCAATTGCC

9180 GATTGTATATCCGATTTATATTTATTTTTCGGTCGAATCATTTGAACTTTTACATTTGGA

9240 TCATAGTCTAATTTCATTGCCTTTTTCCAAAATTGAATCCATTGTTTTTGATTCACGTAG

9300 TTTTCTGTATTCTTAAAATAAGTTGGTTCCACACATACCAATACATGCATGTGCTGATTA

9360 TAAGAATTATCTTTATTATTTATTGTCACTTCCGTTGCACGCATAAAACCAACAAGATTT

9420 TTATTAATTTTTTTATATTGCATCATTCGGCGAAATCCTTGAGCCATATCTGACAAACTC

9480 TTATTTAATTCTTCGCCATCATAAACATTTTTAACTGTTAATGTGAGAAACAACCAACGA

```
 9540  ACTGTTGGCTTTGTTTAATAACTTCAGCAACAACCTTTTGTGACTGAATGCCATGTTTC
 9600  ATTGCTCTCCTCCAGTTGCACATTGGACAAAGCCTGGATTTACAAAACCACACTCGATAC
 9660  AACTTTCTTTCGCCTGTTTCACGATTTTGTTTATACTCTAATATTTCAGCACAATCTTTT
 9720  ACTCTTTCAGCCTTTTTAAATTCAAGAATATGCAGAAGTTCAAAGTAATCAACATTAGCG
 9780  ATTTTCTTTTCTCTCCATGGTCTCACTTTTCCACTTTTTGTCTTGTCCACTAAAACCCTT
 9840  GATTTTTCATCTGAATAAATGCTACTATTAGGACACATAATATTAAAAGAAACCCCCATC
 9900  TATTTAGTTATTTGTTTGGTCACTTATAACTTTAACAGATGGGGTTTTTCTGTGCAACCA
 9960  ATTTTAAGGGTTTTCAATACTTTAAAACACATACATACCAACACTTCAACGCACCTTTCA
10020  GCAACTAAAATAAAAATGACGTTATTTCTATATGTATCAAGATAAGAAAGAACAAGTTCA
10080  AAACCATCAAAAAAAGACACCTTTTCAGGTGCTTTTTTTATTTTATAAACTCATTCCCTG
10140  ATCTCGACTTCGTTCTTTTTTTACCTCTCGGTTATGAGTTAGTTCAAATTCGTT.
```

Fig. 4.4 Complete nucleotide and amino acid sequences of plasmid pT7 *hasA-tuaD-gtaB-pgi* or pT7hyalu.

This plasmid has a molecular weight of 10194bp. All genes are under the control of strong promoter T7. Between 196 and 1383 bp is present the sequence *hasA* for HA synthase from *Streptococcus equi*, between 1430 and 2873bp it is present the sequence for *tuaD* gene, between 2905 and 3781bp it is present the sequence for *gtaB* and 3824 and 5125bp it is present the sequence for *pgi* gene, the last three genes derived from *Bacillus subtilis* genome.

The plasmid, checked by restriction analyses with *EcoRI* (column1), *EcoRI* and *HindIII* (column2), *HindIII* (column3), *XbaI* (column4) results in bands with the expected molecular weights (fig. 4.5).

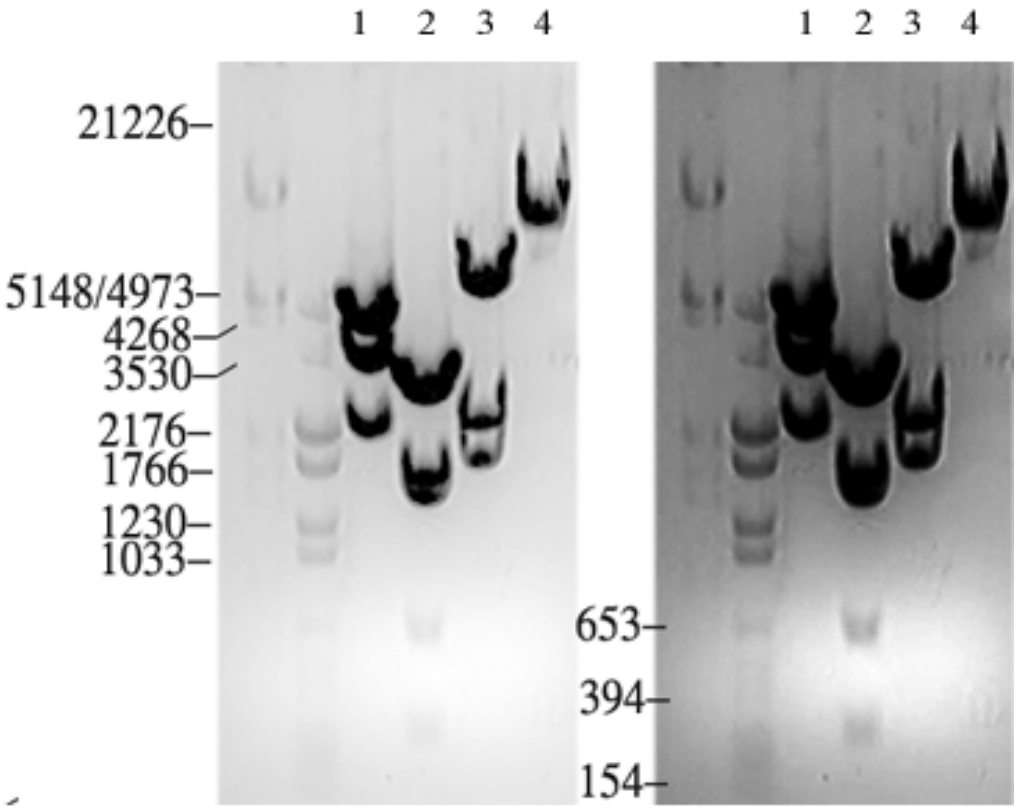


Fig. 4.5 Restriction analysis map of pT7hyalu.

4.2.5. Preparation and transformation of competent *B.subtilis* and *B.megaterium* cells

Preparation and transformation of competent *B.subtilis* and *B.megaterium* cells, was performed as described in cap. II par. 2.2.5-2.2.8.

4.2.6. Qualitative assays for HA production

Plate assay

To study the morphology of the HA-producing 1012 and WB800N *B.subtilis* transformants, clones were plated on the Difco Tryptose Blood Agar Base agar petri plates - chloramphenicol agar plates added containing 1 mM IPTG or a IPTG gradient, and 10 µg/ml chloramphenicol (1012*B.subtilis*) while 10 µg/ml chloramphenicol and 10 µg/ml neomycin (WB800N *B.subtilis*) (selection marker). The cells were streaked and incubated at 25°C for 24–48 h.

HA-producing *B.megaterium* transformants, clones were plated on LB agar petri plates - containing 0,5% (w/v) xylose or a xylose gradient, and chloramphenicol (4,5 µg ml^{-1}) and tetracyclin (10µg ml^{-1}) (selection marker). The cells were streaked and incubated at 25°C for 24–48 h.

HA precipitation

After collection of cells by centrifugation, culture medium was added of 2 volumes of absolute ethanol in order to precipitate HA from medium.

4.2.7. Hyaluronan analyses

HA titers were estimated by the carbazole assay after first precipitating HA from medium samples with 2 volumes of absolute ethanol, one washing in absolute ethanol, and then redissolving in H_2O. The assay detects the glucuronic acid released after the sample has been hydrolyzed with H_2SO_4 (Bitter and Muir, 1962). Carbazole assay shows absorbance at 530 nm typical of glucuronic acid, constituent of HA. This assay is not highly specific (compounds such as sucrose are cross-reactive),and so it was necessary to determine the background reading. The HA titer is assumed to be 2.05 times the glucuronic acid titer.

4.2.8.Colony stability assay

HA-producing bacteria from selective Difco Tryptose Blood Agar Base or LB agar plates were used to inoculate growth in selective liquid LB or MSR medium under 1 mM IPTG induction. At mid-log phase, the cells were diluted, plated on selective IPTG induced agar plates, and incubated at 25°C for the growth of well-separated colonies. Then, a mixture of colonies was diluted and plated again for more times. Replica plates experiments with and without antibiotic selection and /or induction show plasmid stability.

4.2.9. HA Molecolar Weight determination by agarose gel electrophoresis

The HA was isolated from supernatant culture medium. After centrifugation, HA in supernatant collected was precipitated and washed by adding two volumes of absolute ethanol at -20°C. The precipitate was redissolved in milliQ water. Samples contained approximately 10 μg HA were mixed with about 2 μl of a 0.02% solution of bromophenol blue in TAE buffer containing 2M sucrose. Agarose gel electrophoresis was performed essentially as described by Lee and Cowman using 0.5% (w/v) gel in 10% TAE buffer. Electrophoresis was at 20-35 V overnight and the gels were stained with 0.005% Stainsall in 50% ethanol. Gels were destained by washing and exposure to light.

4.3. Results

4.3.1. Expression of HA via artificial operons in *B.subtilis*

The pathway for HA biosynthesis in the group A and group C streptococci has been determined (Matsubara et al.,1991; O'Regan et al., 1994), and it has been optimized to produce HA in *B.subtilis* (Widner et al., 2005), as shown in fig.4.6; in fact the only lacking *B.subtilis* enzyme is the hyaluronan synthase, which is encoded by the *hasA* gene.

HA is synthesized by the polymerization, catalyzed by HA synthase or HAS, of the monosaccharides from the two nucleotide sugars UDP-glucuronic acid (UDP-GlcUA) and UDP-N-acetylglucosamine (UDP-GlcNAc), final products of two parallel metabolic branches, that, starting from glucose as a carbon source, also produce in this pathway multiple sugar intermediates precursors required for important cellular functions.

HA biosynthesis is an expensive process for the cell, both in terms of carbon and energy consumption. For every mole of HA disaccharide unit produced, 2 mol of glucose, 5 mol of nucleoside triphosphates (3 as ATP and 2 as UTP), and 1 mol of acetyl-coenzyme A are consumed. If large amounts of HA are produced, this could pose a substantial metabolic burden on the cell.

In addition, the nucleotide sugars UDP-Glc and UDP-GlcNAc are both required for cell wall biosynthesis, while the phosphosugars Glc-6-P and Fru-6-P funnel directly into the pentose phosphate and glycolytic pathways, respectively, both of which are essential for cell growth.

To ensure that adequate levels of various precursor sugars and metabolites are available to support both large HA biosynthesis and cell growth, it has been interesting note that different streptococcal species have evolved a dedicated expression system via operon, with transcription of the essential genes involved in the HA pathway in an unique polycistronic mRNA. Few bacterial species are able to produce HA, namely the group A and group C streptococci (gram positive) and *Pasteurella multocida* (gram negative).

The *hasA* gene has been incorporated as the first gene of *has* operon along with one or more genes which encode enzymes involved in the synthesis of some of the aforementioned precursor sugars, based on different species of *Streptococcus*. For example, the operon from *Streptococcus pyogenes* contains three genes: the first, *hasA*, encodes hyaluronan synthase; the second, *hasB*, encodes UDP-Glc dehydrogenase; and the third, *hasC*, encodes UDP-Glc pyrophosphorylase (Crater and van de Rijn, 1995). The operon from *Streptococcus uberis* is somewhat different in that the *hasC* gene is not present as part of the operon but is located elsewhere on the chromosome (Ward et al., 2001).

The operons from *S.equisimilis* and *S zooepidemicus* subsp. *equi* contain the *hasA*, *hasB*, and *hasC* genes as well as a fourth gene, *hasD*, which encodes UDP-GlcNAc pyrophosphorylase (Sloma et al., 2003). In addition, the latter strain contains a fifth gene, *hasE*, which encodes phosphoglucoisomerase.

Thus, these streptococcal species have evolved expression systems via operons with the genes coding enzymes involved in HA pathway transcripted in a unique polycistronic mRNA for ensuring that adequate levels of various precursor sugars are available to support both HA biosynthesis and cell growth (Fig. 4.6).

In order to produce HA in non-HA producer *B.subtilis*, it seemed logical to take a cue from nature and mimic this approach. Since it is also a gram-positive bacterium, it was thought to adapt to this strain the HA biosynthesis pathway, present in natural producers such as *S. zooepidemicus.* Enzymes involved in the synthesis pathway of the precursor sugars are also present in *Bacillus subtilis* in which the homologues for the genes *HasB, HasB, e HasD, HasE* of *Streptococcus* are *tuaD, gtaB* e *gcaD, pgi* respectively. But *B. subtilis* lacks an equivalent hyaluronan synthase.

Since endogenous genes are frequently expressed more efficiently in their native host, it was thought to construct artificial operons which are comprised of *B.subtilis* genes, rather than to introduce into *B.subtilis has* operon from a group A or C *Streptococcus*.

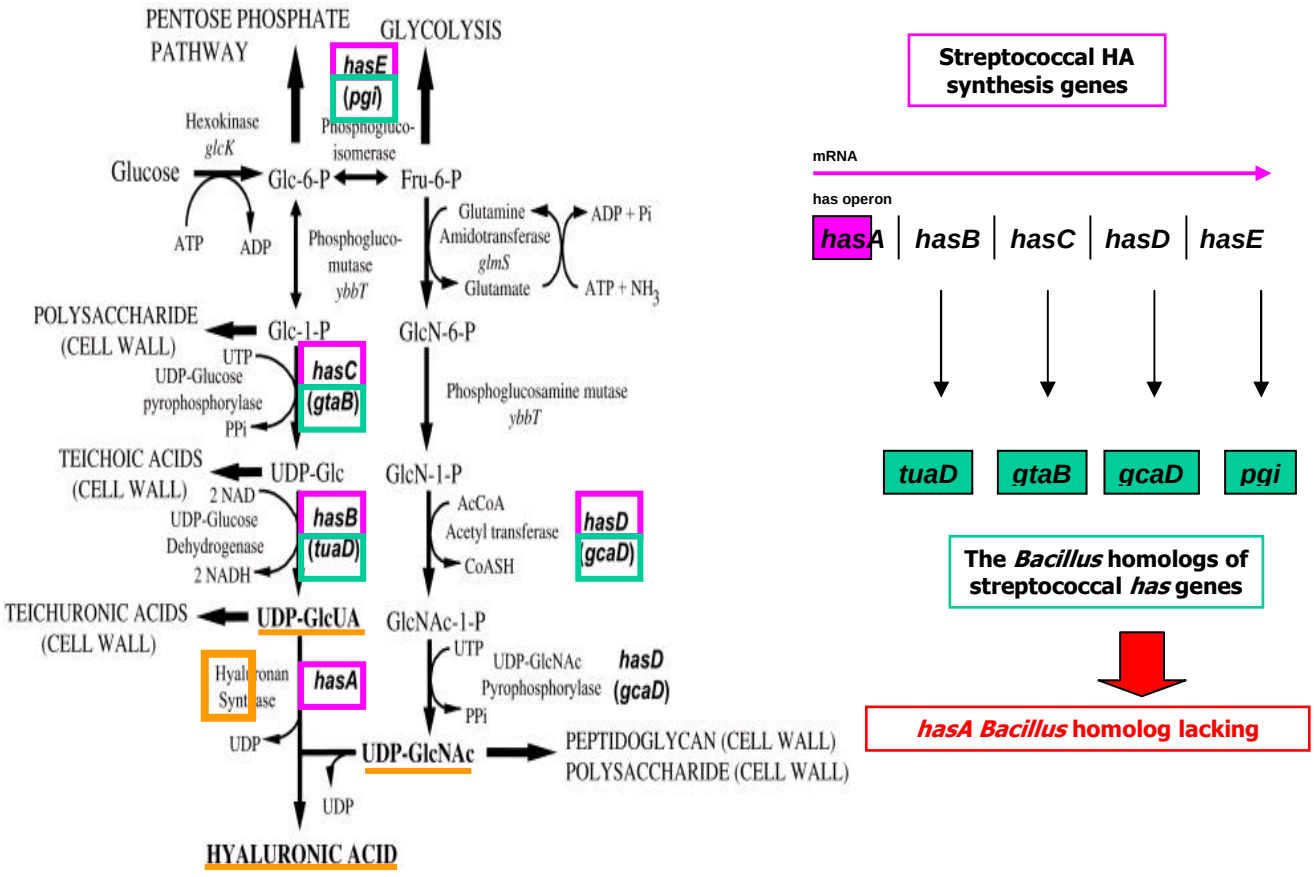


Fig. 4.6 Biosynthetic pathway for HA production in recombinant *B.subtilis* strains. This pathway is based on the reported HA biosynthetic pathway in group A and C streptococci (Matsubara et al.,1991; O'Regan et al., 1994). Gene designations are highlighted in boldface. Streptococcal has gene are in pink and yellow, while the *B.subtilis* homologs are in green. (Adapted from Widner et al., 2005).

Has been found that at least two genes (*hasA* and *hasB*) are crucial for production of HA in significant quantities comparable to native strains (Widner et al. 2005), thus we have been produced an artificial operon *hasA-tuaD*, that mimics which one present in the *S.uberis* locus.
While in Widner work (Widner et al., 2005), an artificial cassette-operon was created with an integrative system on the chromosome of *B.subtilis*, we have developed an expression system using of an episomal vector (i.e. a plasmid) with relatively high copies. It has advantage to express an more higher level of mRNA for pathway of HA than a single copy present on the chromosome. At the aim, therefore have been developed expression cassette-operons vectors, using a recent available vector (Nguyen D.H. et al., 2005) that is shuttle *E.coli/B.subtilis*, used for expression and production of recombinant proteins in higher efficiency in *B.subtilis*. These vectors (pHT series) are now available from MoBitec.

4.3.2. Construction of cassette operon *hasA-tuaD* and *Bacillus subtilis* pHT01*hasA-tuaA* or pBS5

The gene sequence (*hasA*) coding the hyaluronic acid synthase has been rescued by PCR from the chromosomal DNA of *Streoptocooccus zooepidemicus*. It has been choosen the *has A* gene from *Streptococcus equisimilis*, since the hyaluronan synthase from this organism is especially interesting due to its higher intrinsic polymerization rate (Kumari et al., 1997) and also because this strain is used from Fidia S.p.A. for production of HA. Since the HA production with high molecular weight has been obtained from Fidia s.p.a after several passages from the original strain of *Streptococcus equisimilis*, maybe that the *hasA* gene was modified in this strain. After cloning we have found that the *hasA* gene presented the same nucleotide sequence that published. The expression of recombinant HA synthase induced by 0.5 mM IPTG in *E.coli* BL21(DE3) LysS, showed a protein band of apparent expected MW about 42 kDa, according with MW reported in literature. In fact, the HAS is 417 amino acids long protein, with calculated MW of 47778Da, but enzyme migrates anomalously fast in SDS-PAGE polyacrilamide gel electrophoresis ~ 42000 Da, as reported in literature (Kumari et al., 1997) (Fig. 4.7).

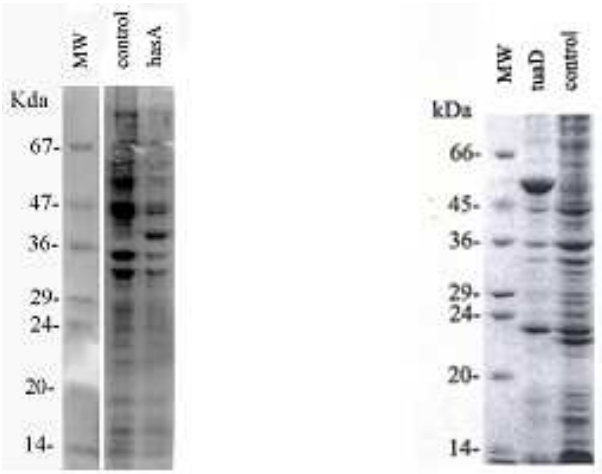


Fig.4.7 BL21[DE]LysS cells were transformed with the plasmids pRSETB tuaD and pGEM4HasA, and protein expression was induced with 0.5 mM IPTG. Cells were suspended in SDS-lysis buffer and proteins were analyzed by SDS-PAGE and stained with Comassie Brilliant Blue.

The UDP-glucose dehydrogenase gene (*tuaD*) was rescued by PCR from the chromosome DNA of *Bacillus subtilis* 168 strain and cloned also in vector pRSETB, resulting in plasmid pRSETB*tuaD*. It has been choosen to rescue the *tuaD* gene from *Bacillus subtilis*, rather than homologue *hasB* of *Streptococcus equisimilis*, since endogenous genes are frequently expressed more efficiently in their native host. The expression of recombinant UDP-glucose dehydrogenase induced by 0.5 mM IPTG in *E.coli* BL21(DE3)LysS, revealed a protein band of expected MW 54 kDa as showed in SDS-PAGE analyses (fig. 4.7). After construction of the cassette operon *hasA-TuaD* in pGEM4-*hasA-tuaD* (Fig.4.8), the two genes in tandem, were cloned under the control of IPTG inducible Pgrac promoter in *B.subtilis* expression vector pHT01resulting in plasmid pBS5 (Fig.4.9).

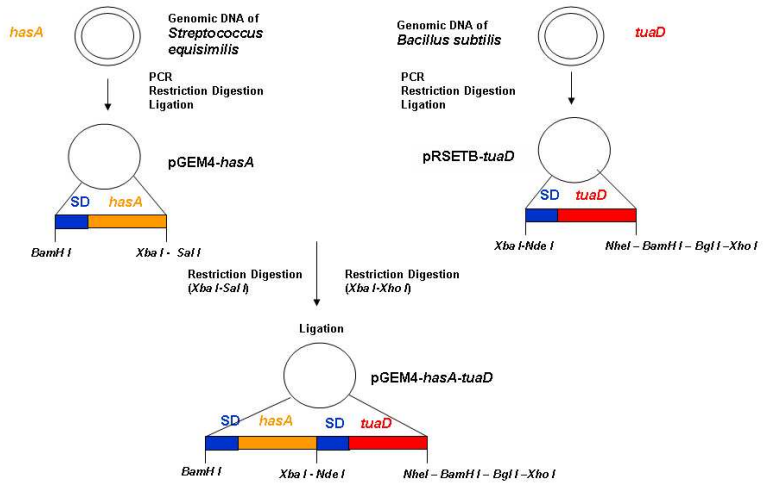


Fig. 4.8 Representation of cassette-operon *hasA-tuaD* construction.

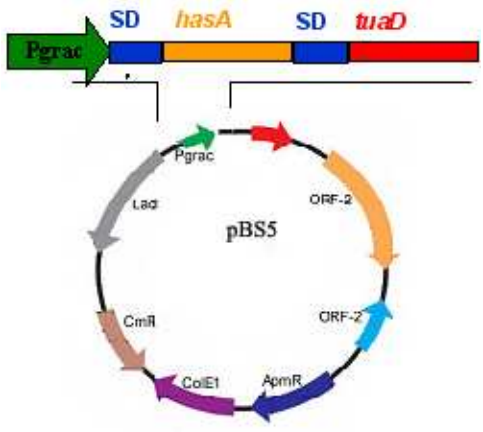


Fig. 4.9 pBS5 vector.

The plasmid pHT01 (MoBiTec) allows relatively high-level expression of intracellular recombinant proteins (Phan et al., 2005). The sequence encoding target genes is under the control of expression cassette Pgrac consisting of the *groE* promoter, the *lac* operator and the efficient

gsiB Shine-Dalgarno sequence (Jürgen et al., 1998). Vector is based on the strong, costitutive σA - dependent promoter *groE* preceding the groE operon of *B.subtilis*, which has been converted into an strong inducible promoter by addition of the lac operator *lacO* of *E.coli*. While the background level of expression of these expression cassettes is very low in the absence of inducer IPTG, an induction factor of about 1,300 was measured using reporter genes. The amount of recombinant proteins produced after addition of IPTG may represent 10-13% of the total cellular protein (Phan et al., 2005).

Because of the low transformation efficiency of *Bacillus subtilis* strains, plasmid construction were done in *E.coli* with the shuttle vector *E.coli/B.subtilis* pHT01 and then transformed in competent cells of *B.subtilis*.

This cloning has been particularly difficult since TOP 10 *E.coli* colonies were very small for size, showing growth problems (Fig.4.10) and when the expression of UDP-glucose dehydrogenase was induced in BL21 by adding IPTG, we have found that the cells do not survive, indicating that *tuaD* gene expression is toxic in *E.coli*. It's likely that high UDP-glucose dehydrogenase expression level drain UDP-glucuronic acid (precursor of HA) to produce teichuronic acid, impoverishing the cell of glucose.

During cloning, we have observed that amount of plasmid DNA from bacterial cells TOP10 transformed with pBS5, was about 1/10 concentration than amount obtained for parental plasmid pHT01 using a Qiagen protocol. Therefore pBS5 was transformed, growth and amplified in different *E.coli* strains in order to obtain a more high purification. Under this aspect we have found that INVaF' cells, that express costitutively *lacI* gene, codifying lactose repressor, were the best. We have found that TOP 10 *E.coli* transformed with pBS5 was able to produce HA although at low levels, indicating that the cassette and the Pgrac promoter is functional also in *E.coli*.

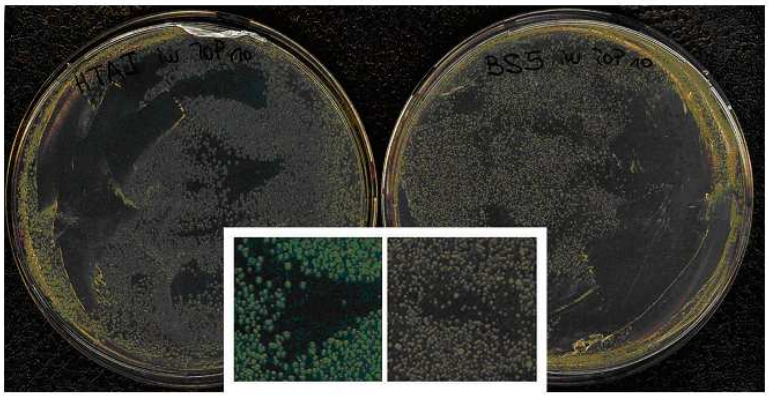


Fig.4.10 *HasA* and *tuaD* gene expression in pBS5 TOP10 *E.coli* cells. These cells are smaller than parental pHT01 TOP10 *E.coli* cells.

4.3.3. Metabolic engineering of WB800N and 1012*B.subtilis*

In order to increase the amount of HA production pBS5 plasmid has been transformed in *B.subtilis* cells. The transformation protocol used was the same described in cap II.

WB800N and 1012 *B.subtilis* strains have been chosen. The first has been engineered of seven-extracellular protease and wall-bound protease WprA deficient, that could degrade the secreted protein. Although hyaluronan synthase, is a transmembrane protein there is always the possibility that proteolisys occurs. 1012 strain was chosen due to its superior growth characteristics and high yield of secreted protein products in industrial fermentors and because is used often in pHT system. In any case pBS5 is an episomal non-integrative plasmid, so the artificial cassette operon *hasA-tuaD* is present in more copies than a single copy integrated in chromosome. The presence and integrity of expression cassette in PBS5 was confirmed by PCR analysis.

4.3.4. Static shaker flask culture studies

The HA production capability of the metabolically engineered pBS5 *B.subtilis* strains was investigated in static culture experiments in MSR medium added with sugars. Previous expression experiments in our laboratory of recombinant GFP-pHT01 cloned under Pgrac control in *B.subtilis,* with different media and carbonium sources, have allowed to select the more productive media for recombinant proteins expression (this work cap. II) .

The best medium used in these experiments was MSR media including 0,5-1% glucose and 0.5-1% sucrose, at 25°C. After 24 h from IPTG induction, the HA production was evaluate by ethanol precipitation and carbazole assay. It was observed that very important is shaker cultures at relatively high rpm (220-250 rpm) to favour oxygenation, this could to be important for oxidation of glucuronic acid.

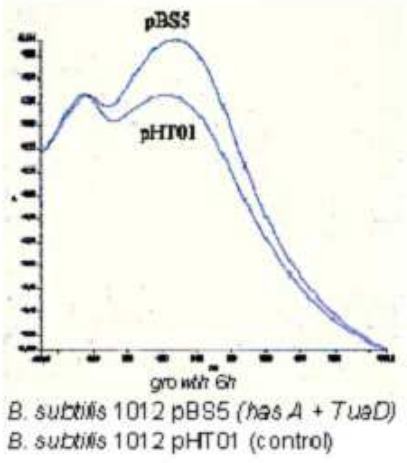


Fig.4.11 Glucuronic acid analysis by reaction with carbazole, measuring absorbance of the samples at 530 nm, against the blank. Example with pBS5 1012 *B.subtilis*; pHT01 *Bacillus* transformed represents the blank to value background.

Fig. 4.12 HA precipitation from *B.subilis* culture medium by absolute ethanol.

4.3.5. Production of HA in fermentors

In order to optimize and increase the HA yield, further production trials were made by Fidia S.p.a. in bioreactors, that assure a controlled environment, essential for high HA yield. By applying fed-batch technology in bioreactors (Kiss and Stephanopoulos, 1991), it is expected that the results will be significantly enhanced by appropriate glucose feed strategies and pH control, optimal feeding of sucrose, and suitable regulation of agitation and aeration corresponding to the enhanced viscosity of the fermentation broth by HA accumulation. Generally, these methods can enhance HA accumulation while maintaining robust culture viability through controlled cell growth.

The results showed significant HA production in minimal medium, comparable on long time with that obtained in rich media, as LB, with advantage that using of precise chemical composition medium, reduces accumulation of contaminants in production and purification of biological drugs, assuring a final products very pure and toxin-free.

In fermentation trials (Fig. 4.12), the HA yields varies from 0.48 to 0.8 g/L in the two recombinant WB800N and 1012 *B.subtilis* strains, comparable with that reported in literature for other recombinant HA-producing microorganisms (Widner et al. 2005; Chien and Lee 2007a; Chien and Lee 2007b). Nevertheless pBS5 recombinant plasmid stability at the fermentation end was very low, since presenting in max 8% of bacterial cells.

A

Data Fermentazione:	29/07/08
Clone *B.subtilis*:	1012-pB55 (colonia 1)
Volume Fermentazione:	7 L di minimal medium (in Fermentatore da 20L)
Agitazione:	1200 RPM
Areazione:	20-22 Litri air/ min
Temperatura:	37°
pH:	6,9-7,1
Fonte di Carbonio:	2% Saccarosio
Induzione O.D 0,6-0,8:	0,4-0,5 mM IPTG
Antibiotico (Cloranfenicolo):	20uM
Durata della fermentazione:	71h
Durata dopo Induzione:	67h

Tempo Induzione	tempo/h	conc HA (dg/L)-carbaz
Induzione	3,5	0,19
18h ind	21,5	1,06
28h ind	31,5	1,53
42h ind	45,5	2,77
52h ind	55,5	4,06
66h ind	68,5	4,62
73h ind	76,5	3,02

B

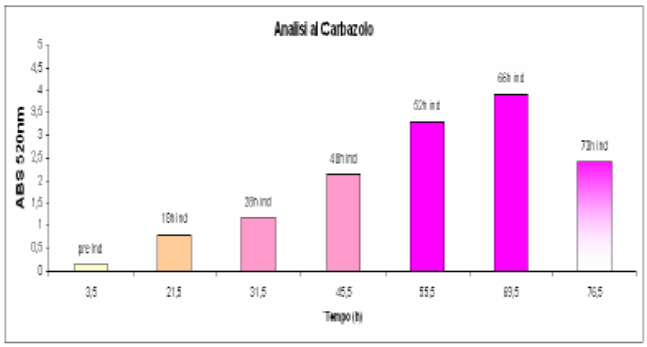


Fig. 4.12 A. Example of fermentation process for a pBS5 1012 *B.subtilis* colony . B. Carbazole assay shows a maxima production of 0.48 g/L at 66h from induction.

4.3.6. Selection of HA secerning *B.subtilis* strains with plasmid stability

In our laboratory, it has been tested stability of plasmid, by replica plating experiments with and without antibiotic selection and/or IPTG induction. While the number of bacterial colonies in plate is comparable either with either without antibiotic selection of cloramphenicol (Fig. 4.13), when bacteria were plate in presence of IPTG we have not observed develop of colonies (data not shown), thus indicating and confirming the toxicity observed in fermentator.

Therefore we have streaked our pBS5 transformed *B.subtilis*, in plate with IPTG concentration gradient under cloramphenicol selection.

In this case we have observed that pBS5 transformed cells died with high IPTG concentration, while survived without IPTG or in lower IPTG amounts, exhibiting a small and dry colony morphology or a large and mucoid colony morphology, respectively. This mucoid colony morphology is typical of *B.sutilis* HA-producing (Fig. 4.14).

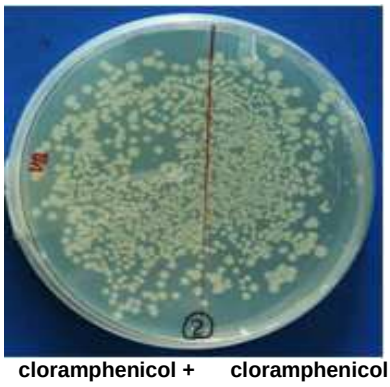


Fig.4.13 pBS5 plasmid stability with no IPTG induction.

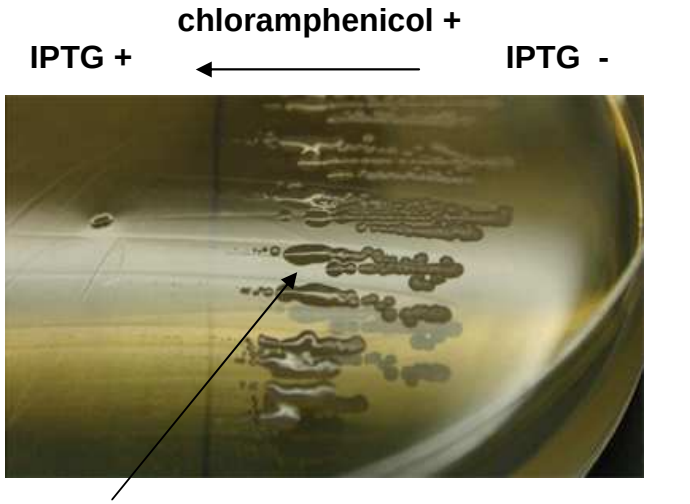


Fig.4.14 HA expression plate assays. pBS5 *B.subtilis* cells died with high IPTG concentration, while survived in a IPTG gradient, exhibiting a small and dry colony morphology or a large and mucoid colony morphology, this latter typical of *B.sutilis* HA-producing. The plates were incubated at 25°C for 24–48 h after induction.

Some of these single mucoid colonies HA-producing were selected and grown more times with IPTG and antibiotic selection, showing plasmid stability and the presence of the cassette for HA (Fig. 4.15).

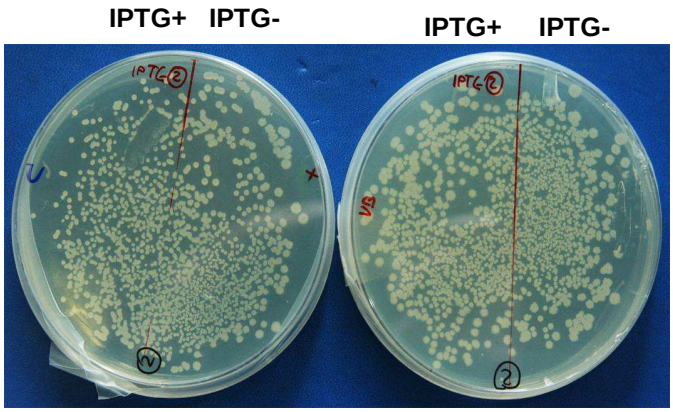


Fig.4.15 pBS5 plasmid stability with and without IPTG induction, after more selection cycles with IPTG and antibiotic selection.

Fermentazione 14/10/08 (10l2 colonia1 IPTG resisitente)		
tempo induzione	HA dg/L al Carbazole	Stabilità Plasmide
Induzione	0,87	100%
10h ind	4,9	100%
25h ind	8,89	100%
35h ind	14,7	100%
49h ind	33,8	100%

Fig. 4.16 pBS5 plasmid stability of HA producing cells selected after more selection cycles with IPTG and antibiotic selection, in fermentation process. After 49h from induction, plasmid stability was 100% and yields were increased at 3 up to 5 g/L.

These selected HA producing cells survived also in higher IPTG concentration and after fermentation processes they showed plasmid stability, with plasmid present in 100% of bacterial cells. This selection system makes possible obtain stable bacterial HA producing strains also after many growth cycles and cellular divisions. As consequence of this plasmid stability the yields and MW of HA resulting by fermentation processes, were incremented up to 5 g/L that is in large exceed that is published.

At the light of these results, it is possible hypothesize that HA synthase and UDP-Glucose-Dehydrogenase or its protein products are toxic . Perhaps these two genes are not synthesized in the same time and amounts. Although *hasA-tuaD* cassette-operon is transcripted in a unique polycistronic mRNA after IPTG induction, HA synthase, that polymerases, is a complex plasma membrane enzyme with several trans-membrane domains, that probably requires more long time for its synthesis, folding and functional localization, than UDP-Glucose-Dehydrogenase, a simplex cytoplasmic enzyme that converts glucose in glucuronic acid and for which it is probably required lower complexity for its folding (Fig.4.17). So maybe that after induction, UDP-Glucose-Dehydrogenase is traduced by polycistronic mRNA, becoming functional before than HA synthase. If it is so, dehydrogenase deprives the cells of glucose converting in glucuronic acid that, not early forming HA and excreting by HA-synthase, accumulates with toxic action and cell death as observed in *E.Coli*. Therefore there is necessity of a precise expression regulation of genes involved in HA biosynthetic pathway, with a more precise and equilibrated amount of enzymes. When we maintain the cell under IPTG induction these two genes become constitutive

and the amount of glucuronic acid produced by UDP-Glucose-Dehydrogenase is export out of cell under HA by HA synthase.

To obtain a more tightly controlled gene expression regulation, it was proved to clone *hasA* and *tuaD* under different promoters control. But during the HA expression plasmids construction, it has been discovered that *hasA* and *tuaD* each are toxic in *E.coli* when controlled by a constitutive promoter as P43.

These results maybe to support of hypotheses that, 1) a functional HA synthase overexpression deprive the cells of HA precursor N-acetylglucosamine and UDP-glucuronic acid necessaries also for cell wall construction, 2) while a functional UDP- glucose-dehydrogenase overexpression causes not controlled synhtesis of glucuronic acid that, not being incorporate in HA and excrete by synthase, acidific the cells impoverishing of glucose, inducing in both situations the death of cell. To further support of this toxicity hypothesis, it was observed during the pBS5 cloning in *E.coli*, that it can be grown and purified efficiently only in *E.coli* invαF'. The growth in TOP10, normally used in transformation and cloning, showed very small colonies. invαF' cells have lac repressor costitutively expressed, so they are not able to transcript the gene under Plac promoter. The plasmid pHT01, used to construct pBS5, contains hybrid Pgrac promoter with a lac operator, which binds lac repressor blocking *hasA* and *tuaD* gene expression. Top10 that do not contain the lac repressor are not able to maintain repress the Pgrac and therefore growth with more difficulty.

These results indicate that there is necessity of a precise concerted in time expression regulation of genes involved in HA biosynthetic pathway and organized in operons in *Bacillus* and other bacteria.

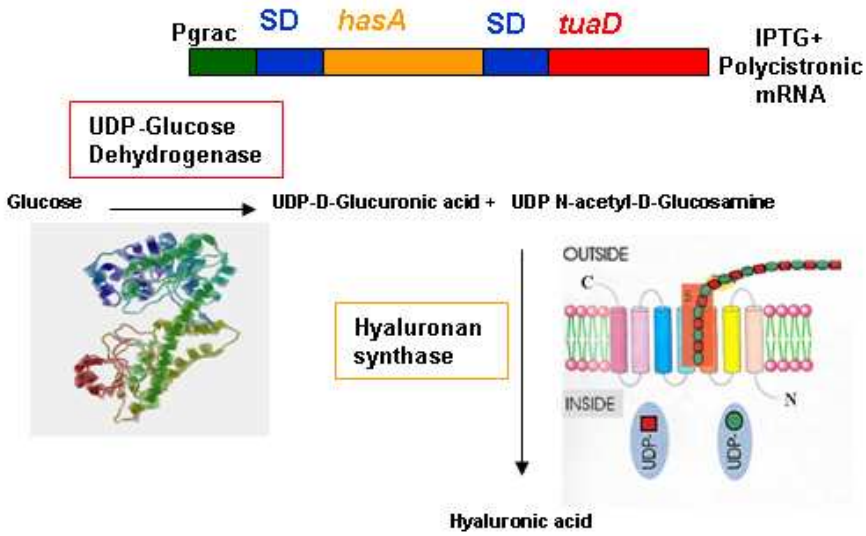


Fig. 4.17 Representation of *hasA-tuaD* cassette-operon, and concerted action of UDP-Glucose-Dehydrogenase and HA synthase in HA metabolic pathway.

4.3.7. Effect of metilate DNA in the HA expression

During the cloning of HA cassette we have found that the restriction sites *XbaI* was loss in pBS5 plasmid purified by invαF' *E.coli* . After checking byDNA sequences we have observed that this sites is always present, therefore we have hypothesized that *XbaI* sites are metilated in invαF'. To test this hypothesis we have transformed, amplified and purified the DNA plasmid pBS5 in JM110 *E.coli* strain that is deleted of dam and dcm metilase.

Restriction analysis with *XbaI* restriction enzyme confirm the presence of this site after purification by JM110 and therefore we have concluded that invαF' *E.coli* strain is able to metilate the pBS5 plasmid.

Therefore in order to ckeck if the metilation has an effect on HA expression, we have compared HA produced by *B.subtilis* transformed either with pBS5 DNA plasmid purified by invαF' *E.coli* or purified from JM110 strain. After several experiments we have found that there were not substantially differences between HA produced by two strains, and therefore seems that metilation no influence transcription of HA pathway (Fig.4.18).

Fig. 4.18 Agarose gel electrophoresis of HA produced by pBS5 800N and 1012 *B.subtilis* at 24-36 h from induction. pBS5 plasmid DNA was purified from invαF' or JM110 *E.coli*, to verify if metilation influence production of HA. Commercial standard HA of about 800 kDa produced by Fidia was used as references.

4.3.8. HA production via artificial operons in *Bacillus megaterium*

More recently has been develop an expression system using *B.megaterium* as host and the T7 RNA polymerase based promoter, and since this system shows a more tightly precise expression regulation in *E.coli*, we thought to use this system for HA expression in this host.

Advantages of this system are the stringent selectivity and the high transcriptional activity so that it is possible to lead to a saturation of the protein-synthesizing machinery in host organism. Consequently 50% or more of the total cellular protein can consist of the desired protein could be obtained in *E.Coli* (Studier and Moffatt, 1986). The T7 RNAP expression system for *B.megaterium* combines the features of this system in *E.coli* with the regulation by the xylose operon and is based on two parallel-replicating plasmids: pT7-RNAP and pP T7 (Gamer et al. 2009), available fromMoBiTec.

Plasmid pP T7 contains all structural elements necessary for a T7 RNAP-dependent expression of target genes under the control of T7 promoter. Plasmid pT7-RNAP contains the gene encoding RNA polymerase of the bacteriophage T7 (T7 RNAP) under the transcriptional control of xylose-inducible promoter PxylA . This promoter is regulated by repressor coded by XylR gene, that blocks xyl operon transcription in absence of xylose. In the presence of xylose, repressor takes off and starts transcription of T7 RNAP that recognizes its specific promoter beginning target genes transcription.

Plasmid pT7-RNAP contains the gene encoding RNA polymerase of the bacteriophage T7 (T7 RNAP) under the transcriptional control of xylose-inducible promoter PxylA.

Using this promoter system localized on free replicating plasmids for the overexpression of recombinant genes in *B.megaterium,* yielded an induction of up to 350-fold while adding xylose to the growth medium.

T7 RNA polymerase is highly processive and recognise only genes that are place under its control. So it could be used to selectively transcript the target gene, without interfere with the growth and synthesis of bacterial proteins.

B.megaterium as well as *B.subtilis* does not produce toxins associated with the outer membrane and grow on low cost substances, using of a wide variety of carbon sources (Vary,1994).

B.megaterium respect to other bacilli has the advantage that none of the alkaline proteases are present. This fact enables an excellent cloning and expression of foreign proteins without degradation (Meinhardt et al., 1989; Rygus & Hillen, 1991). Furthermore *B. megaterium* is known for the stable replication and maintenance of several extra-chromosomal DNA elements in parallel (Kim, 2003; Vary, 1992;von Tersch and Robbins, 1990) and this can be very important to cloning all the genes-operons of HA pathway without plasmid instability events. In fact, trials to construct cassette-operons in *Bacillus subtilis* with more genes than *hasA-tuaD (*example *hasA-tuaD-gtaB)* were failed probably due plasmid instability phenomena. Cloning cassette-operons containing more genes implicated in HA pathway could be interesting to favour higher HA yields and molecular weights.

4.3.9. HA molecular weight

HA molecular weight (MW) is important for the physiochemical as well as biological properties of HA. Although the HA biosynthetic mechanism is well established (see above par.4.3.1.), little is known about what controls HA molecular weight. The molecular mechanisms underlying chain termination and hence molecular weight control, remain poorly understood, not only for hyaluronan synthases but also for other β -polysaccharide synthases, (e.g. cellulose, chitin, and 1,3-betaglucan synthases) (Chen et al.,2009).

It is noted that MW is partly an intrinsic parameter of the HA synthase (Weigel P.H., 2002), for example, has been demonstrated that, mutation of conserved cysteine or polar residues in streptococcal HA synthases results in reduced molecular weight with limited effect on biosynthetic rate. Chen and others (2009) have demonstrated that in vivo MW is also affected by culture parameters, e.g. temperature and aeration (Park M. et al., 1997), under this aspect, a more likely explanation is that molecular weight is affected by the availability of activated sugar substrates (UDP-GlcUA and UDP-GlcNAc) as well as the concentration of possible effector molecules, such as free UDP.

More has been demonstrated that the MW hyaluronic acid could be increased by over-expressing genes involved in these pathways (e.g., phosphoglucoi-somerase (*pgi*) in *Streptococcus equi* subspecies *zooepidemicus* (Chen et al., 2009).

ChenW.Y. and others manipulated metabolite concentrations in the hyaluronan pathway by overexpressing the five genes of the hyaluronan synthesis operon in *S.zooepidemicus*, concluding that overexpression of genes involved in UDP-glucuronic acid biosynthesis decreased MW, whereas overexpression of genes involved in UDP-N-acetylglucosamine biosynthesis increased molecular weight. The data indicate that 1) high molecular weight is achieved when an appropriate balance of UDP-N-acetylglucosamine and UDP-glucuronic acid is achieved, 2) UDP-N-acetyl-glucosamine exerts the dominant effect on molecular weight, and 3) the wild-type strain has suboptimal levels of UDP-N-acetylglucosamine. Consistent herewith molecular weight correlated strongly with the concentration of UDP-N-acetylglucosamine (Chen et al., 2009).

Similarly, Yu and Stephanopoulos (2008) demonstrated that *pgi* plays an important role in the production of high MW HA in recombinant *Escherichia coli.*

At the light of these results and considering, as demonstrated by Widner (2005), that UDP-glucuronic acid is limiting for HA synthesis in *B. subtilis* (yields from *hasA* were negligible, whereas yields from *hasA/tuaD* were comparable to strain *hasA/tuaD/gtaB*, indicating that the

level of UDP-GlcUA is limiting in the native *B.subtilis* background due to insufficient levels of UDP-Glc dehydrogenase), it is thought in our laboratory to construct an artificial cassette-operon cloning *hasA-tuaD-gtaB* and *pgi* genes under T7 promoter control of pT7 plasmid for *B.megaterium*.

In fact, trials to construct cassette-operons in *Bacillus subtilis* with more genes as *hasA-tuaD-gtaB* were failed probably due plasmid instability phenomena. Because of *B.megaterium* is known for the stable replication and maintenance of several extra-chromosomal DNA elements in parallel (Kim et al., 2003), it could be advantageous to overexpression of HA pathway in order to increase yields and MW of produced HA.

4.3.10. Metabolic engineering of pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi*

As *B.subtilis*, also *B.megaterium* is missing of HA synthase enzyme and we have engineered Ms941 *B.megaterium* strain to overexpress the *hasA* derived from *Streptococcus equi* together with the other gene involved in HA pathway . Therefore as well as *B.subtilis* we have put the *hasA-tuaD* cassette under the control of T7 RNA polymerase dependent promoter of plasmid pT7 (Fig.4.19).

The expression vector pT7 plasmid for *B. megaterium* have a molecular weight significant more little that pHT plasmid series for *Bacillus subtilis* (5243bp vs 7955bp) therefore in pT7 plasmid additional gene are possible incorporate without instability. Therefore in order to increase HA yields and MW, in *B.megaterium*, two other gene involved in HA pathway are engineered in pT7*hasA-tuaD*. This two genes are *gtaB* and *pgi* genes, homologues of streptococci *hasC* and *hasE*. Gene sequences *gtaB*, coding for UDP-glucose pyrophoshorylase, and *pgi*, coding for phosphoglucoisomerase, have been rescued by PCR from *Bacillus subtilis* 168 strain. Also in this case the entire cassette *hasA-tuaD-gtaB-pgi* has been put under the control of T7 RNA polymerase dependent promoter (Fig.4.20).

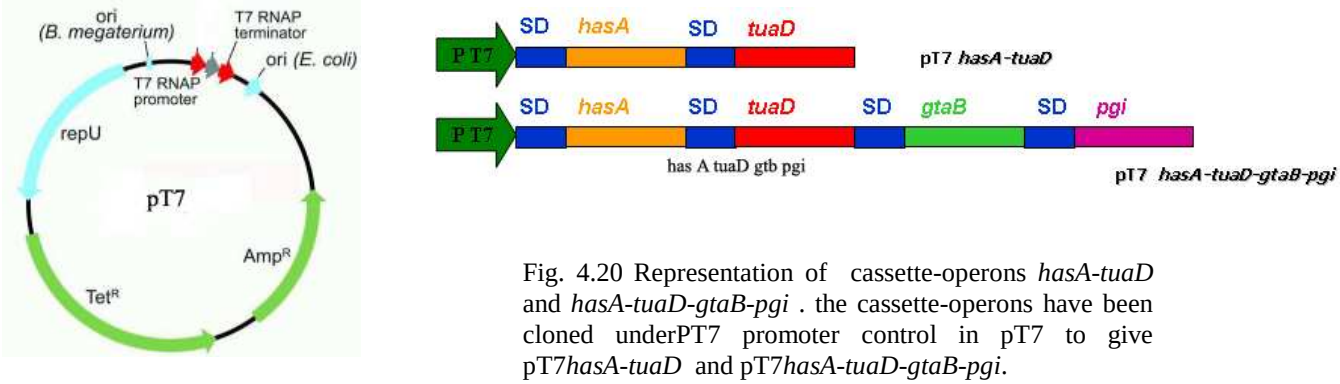


Fig. 4.20 Representation of cassette-operons *hasA-tuaD* and *hasA-tuaD-gtaB-pgi* . the cassette-operons have been cloned underPT7 promoter control in pT7 to give pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi*.

Fig. 4.19 pT7 expression plasmid.

For the low efficiency of transformation of *Bacillus* strains, all constructions were done in *E.coli* invαF' with the pT7shuttle vector, able to replicate in both *E.coli* and *B.megaterium*.

The advantage of pT7 plasmid is that is able to express the recombinant proteins also in *E.coli* when transformed in strains that support the synthesis of T7 RNA polymerase, and therefore it is possible found the correctness of constructions before the transformation in *B.megaterim*. At this purpose, pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi*, after construction in *E.coli* invαF', were transformed in *E.coli* (DE3) LysS cells and the expressed protein has been verified by SDS-PAGE. As shown in Fig. 4.21 a protein of about 42kDa according with HA synthase *hasA* MW reported in literature and a protein of 54 kDa corresponding to UDP-glucose dehydrogenase *tuaD*, are visible in pT7*hasA-tuaD* transformed and IPTG induced cells. In addition, two further protein bands are evident in pT7*hasA-tuaD-gtaB-pgi* or pT7hyal transformed and IPTG induced cells: a protein of 51 kDa corresponding to phosphoglucoisomerase *pgi* and a protein of 32 kDa corresponding to UDP-glucose pyrophoshorylase *gtaB*. In conclusion both plasmids produce enzymes necessary to HA production.

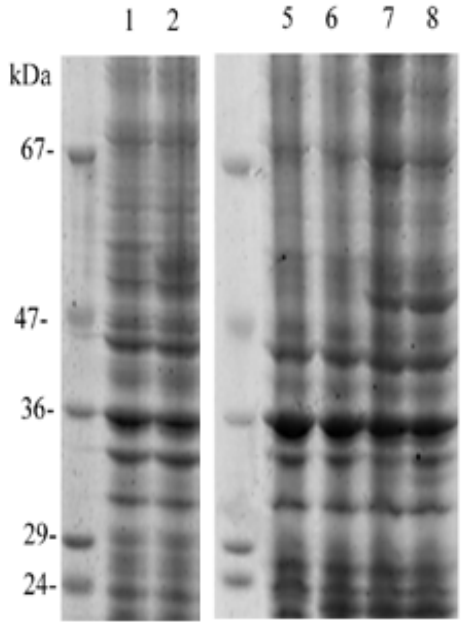


Fig.4.21 BL21[DE]LysS cells were transformed with the plasmids pT7*hasA-tuaD* (line2) and pT7*hasA-tuaD-gtaB-pgi* (line 7 and 8), and protein expression was induced with 0.5 mM IPTG. Cells were suspended in SDS-lysis buffer and proteins were analyzed by SDS-PAGE and stained with Comassie Brilliant Blue. Line 1,5and 6 represent the controls.

Fig. 4.22 Plate assays. pT7*hasA-tuaD* (colony 2) and pT7*hasA-tuaD-gtaB-pgi* (colony 6) *E.coli* cells were streaked on LB agar plates, pT7 control cells grew more easy and resulted larger and planer than pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* transformed cells.

4.3.11. Hyaluronic acid production by recombinant *E.coli*

Because of both plasmids produce enzymes necessary to HA production, it is thought to test the ability of engineered *E.coli* strains to produce HA in medium and in plate assays.

We have observed that when recombinant *E.coli* cells were transformed with pT7*hasA-tuaD* or pT7*hasA-tuaD-gtaB-pgi* , the bacteria cells grew more slowly than cells transformed with the parental plasmid and resulted with smaller colonies. Also in this case when pT7*hasA-tuaD* or pT7*hasA-tuaD-gtaB-pgi* transformed *E.coli* cells, were plated with IPTG concentration gradient and selection, these cells died with high IPTG concentrations, while some exhibited a shiner colony morphology, typical of HA-producing cells at lower IPTG concentration (fig. 4.22).

Some of these single mucoid colonies HA producing, streaked with IPTG survived; thus using a IPTG concentration gradient it is possible select stable colonies producing high HA levels from *E.coli* (fig.4.23).

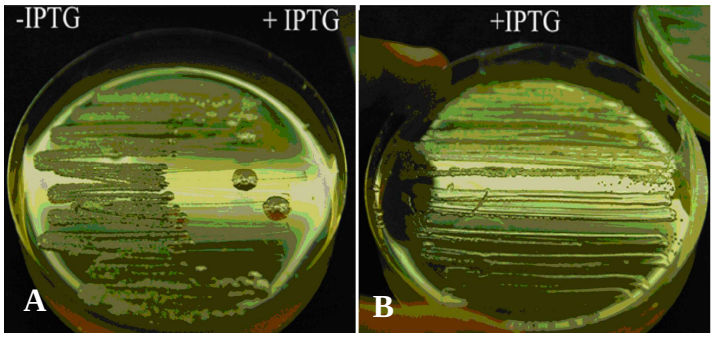


Fig. 4.23 A. HA expression plate assays. pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* *E.coli* cells died with high IPTG concentration, while survived in a gradient IPTG , some exhibiting a mucoid colony morphology, typical of HA-producing cells. B. Selected stable colonies HA-producing.

Thus, pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* HA-producing colonies were grown in LB solution added with 1% sucrose and IPTG for 48h. Cells, harvested by centrifugation from 1 ml of culture, were diluted with an equal volume of 0,1% w/v sodium-dodecyl-sulfate (SDS) and incubated at room temperature for 10 min to free the capsular HA (Chong and Nielsen, 2003). Subsequently the HA product was precipitated with 2 volumes of absolute ethanol. Only pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* colonies produced a HA precipitate (fig. 4.25). The HA precipitate, collected by centrifugation, was also verified by carbazole assay. However, as shown in gel electrophoresis (fig.4.24), recombinant *E.coli* cells are able to produce only low amounts of HA.

Fig. 4.24 Agarose gel electrophoresis of HA produced by pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* engineered colonies obtained in *E.coli*. HA produced by pBS5 WB800N and 1012 *Bacillus subtilis* was used as control.

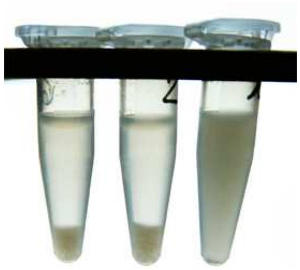


Fig. 4.25 HA precipitation from pT7, pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* transformed *E.coli* culture medium by absolute ethanol. pT7 *E.coli* (eppendorf 1) no produce HA (control).

4.3.12. Hyaluronic acid production by recombinant MS941 *B.megaterium*

pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* were transformed in MS941 *B.megaterium* competent cell provided of pT7-RNAP. At the aim has been set a poly-ethylene-glycol-mediated competent *B.megaterium* protoplasts preparation and transformation protocol (Barg et al., 2005) to remove the solid cell wall of this bacterium and isolate the protoplasts, as described with more details in materials and methods. We have obtained about twenty colonies transformed with pT7*hasA-tuaD* or pT7*hasA-tuaD-gtaB-pgi* in this *B.megaterium* strain. Most of these colonies were screened for ability to produce HA either in plate assay than in shaker flask culture.

Plate assay

When recombinant pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi B.megaterium* colonies were plated on the Difco Tryptose Blood -xylose- chloramphenicol tetracicline added agar plates, a weakly mucoid translucent morphlogy, typical of producing HA cells, (fig. 4.26) were observed with transformants, differently by dry morphology of un-induced recombinant colonies or transformed with parental plasmid pT7. Using a IPTG concentration gradient we have select stable colonies producing high HA levels.

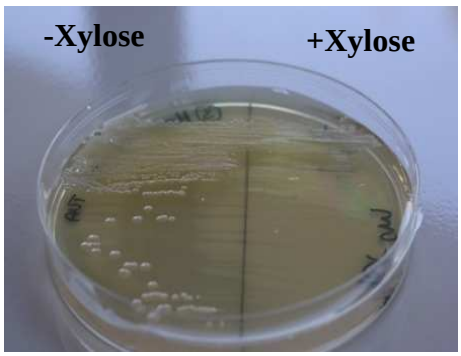


Fig. 4.26 HA expression plate assays. pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi B.subtilis* cells died with high IPTG concentration, while survived in a gradient IPTG, exhibiting a small and dry colony morphology or a mucoid colony morphology, this latter typical of HA-producing cells.

Shake flask culture studies

The ability to produce HA, was investigated in shaker culture experiments with MSR medium incorporating 1% glucose and 1%-5% sucrose and 10 mM $MgSO_4$ favouring HA synthase catalyst, being an important metal co-factor of HAS in the natural *Streptococcus*. As in *B.subtilis*, it is observed importance of high rpm shakering, to favour cell growth in aerobic environment and glucuronic acid oxidation, important to create the HA precursors. HA production was observed in all selected colonies, unless one, (colony 3 Fig. 4.27) after 24 h from 0,5% (w/v) D-xylose induction by ethanol precipitation and carbazole assay.

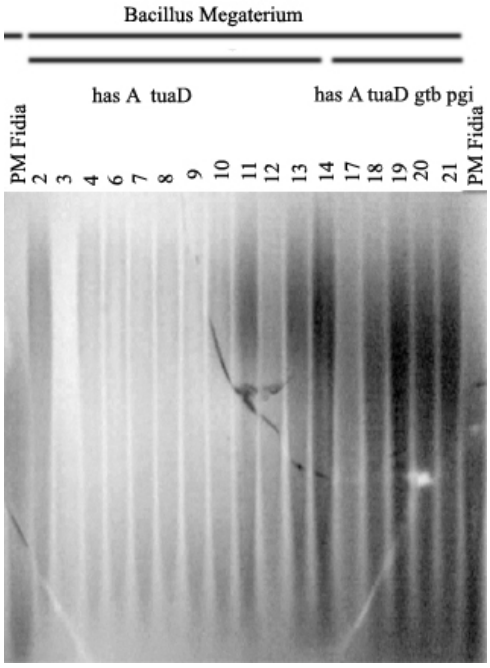


Fig. 4.27 Agarose gel electrophoresis of HA produced by some of pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* engineered colonies obtained in *B.megaterium*. Commercial standard HA of about 800 kDa produced by Fidia was used as references.

Temperature and time effects

Three different temperatures 25°C, 37°C and 41°C were evaluate for growth in MSR medium added of 1% glucose and 5% sucrose for 24 h.

We have found that a significant HA production was obtained only at 25°C either for pT7*hasA-tuaD* either for pT7*hasA-tuaD-gtaB-pgi B.megaterium* colonies. In this condition we have examined the level of synthesis of hyaluronic acid at different time (fig.4.28). While after 8h post-induction we have not found HA production, more long time show an increase of molecular weight.

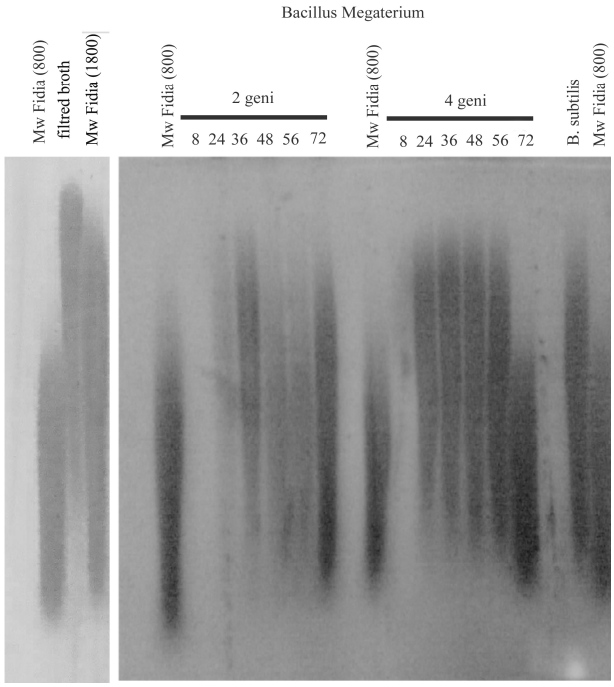


Fig. 4.28 Agarose gel electrophoresis of HA produced on time (8, 24, 36, 48, 56, 72 h from induction) by pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi* engineered *B.megaterium* strains. Commercial standard HA of 800 kDa [line 1] and 1800 kDa [line 3], and HA from filtred *Streptococcus* broth of 4000 kDa [line 2], produced by Fidia S.p.a., were used as references.

In these shake flask cultures the better HA production with highest WM and minor polydispersity is obtained at about 24-36 hours from induction. A minor HA polydispersity is commercially appreciable because means more homogeneity of product. A longer cultivation time resulted in a light reduction of HA MW, likely due perhaps to lower carbonium sources available to metabolic pathways in cells, because of in these shake flask experiments there were not further feeds.

Between these two constructs and in these condition we have observed higher molecular weights for engineered pT7*hasA-tuaD-gtaB-pgi B.megaterium*, demonstrating also in these case that the overexpression of *gtaB* and *pgi* seem influence MW.

A further comparison of HA yields produced between pBS5 *B.subtilis* strains (WB800N and 1012) and pT7*hasA-tuaD* and pT7*hasA-tuaD-gtaB-pgi B.megaterium* strains, shows that pBS5 1012 *B.subtilis* HA synthesis was resulted better than HA of pBS5 800N *B.subtilis* (strain deleted of 8 protease) as MW and polidispersity.

Moreover, the molecular weights of HA produced by engineered pBS51012 *B.subtilis* and pT7*hasA-tuaD B.megaterium* are comparable, while the best HA production is by pT7*hasA-tuaD-gtaB-pgi B.megaterium* with the highest weights, measured in range of 1.8 MDa, with a polydispersity comparable to that reported for commercially available streptococcus sources, as it can see by gel electophoresis (Fig. 4.29).

Because of HA samples examined are produced by shake flask cultures, these results are very promising in view of batch fermentation cultures.

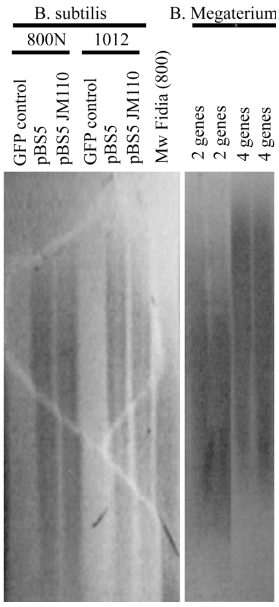


Fig. 4.29 Agarose gel electrophoresis of HA produced by pT7*hasA-tuaD*(2 genes) and pT7*hasA-tuaD-gtaB-pgi* (4 genes) engineered *B.megaterium* , and by pBS5 WB800N and 1012 *Bacillus subtilis*. GFP *Bacillus* strains were used as control.
Commercial standard HA of 800 kDa produced by Fidia S.p.a., were used as references.

4.4. Discussion

Bacillus species are well known for their contributions to agricultural, medical and food biotechnology and for large-scale production of recombinant proteins, in particular for the industrial production of secreted enzymes, and also for the expression of foreign proteins with pharmacological or immunological activities (Harwood et al. 1992, Wong et al. 1995). Such interest seems from several reasons: some *Bacillus* species, with particular emphasis on *B.subtilis*, offer high biosynthetic efficiency and a valid secretion apparatus with release of the target proteins directly into the culture supernatant which greatly simplifies and reduces the costs of downstream purification; at present, about 60% of the commercially available enzymes are produced by *Bacillus* species (van Wely et al., 2001). The large body of information concerning genetic and physiology, together easy of genetic manipulation and large-scale fermentation, make these non-pathogenic organisms, designed GRAS (generally regarded as safe), particularly interesting not only for recombinant proteins production but also for vaccine production and engineering of metabolic pathway, as hyaluronic acid pathway.

At present the worldwide market for HA is estimated at over one billion dollars per annual, and is satisfied primarily by extraction from rooster combs and certain attenuated strains of group C *Streptococcus* which synthesize this compound naturally as part of their outer capsule. However, these are less-than-ideal sources, because of rooster comb-based HA products have been reported to cause allergic and inflammatory reactions, instead streptococci have the potential to produce exotoxins. Thus, it would be advantageous to develop an alternative source of HA that avoids these serious pitfalls, since HA is primarily used in the biomedical field with many applications requiring injection into the human body. Actually, several papers report about recombinant HA synthesis in microorganisms through metabolic engineering (Widner et al. 2005, Chien and Lee 2007a, Chien and Lee 2007b, Prasad et al., 2009 Mao and Chen 2007, Yu and Stephanopoulos 2008, Mao et al. 2009, Sheng et al. 2009, Chen et al. 2009).

Previous works have demonstrated that non- natural producing *Bacillus subtilis* 168 can produce HA, modifying its HA biosynthetic pathway. It was determined that the production of UDP-glucuronic acid is limiting in *B. subtilis* 168 and that overexpressing the *Streptococcus hasA* gene along with the endogenous *tuaD* gene is sufficient for high-level production of HA. The system utilizes a chromosomal integrative strategy, resulting HA yields of about 1g/L in the 1MDA range (Widner et al. 2005). In an other work, to replenish the energy consumed for HA biosynthesis, *Vitreoscilla* hemoglobin was coexpressed with HA-expressing genes, resulting in about 1.8 g/L of HA (Chien and Lee, 2007).

In this work HA production has been obtained in WB800N (strain deleted of 8 protease) and 1012 *Bacillus subtilis* by replicative strategy. Using of an episomal vector (i.e. a plasmid) with

relatively high copies, it has advantage to express an higher level of interested gene (increasing the mRNA amount) than a single copy on the chromosome. Using of pHT plasmid-based expression vector for highly efficient intracellular overexpression of cassette-operon *hasA-tuaD* (codifying HA synthase from *Streptococcus equisimilis*, and UDP-glucose dehydrogenase form *B.subilis* 168), it was been possible select stable metabolic engineered *Bacillus subtilis* strains HA secreting, with yields major than 5g/L in bioreactor, with MW bigger than the 800-1000 kDa.

Moreover, in order to extend the applications of HA and make better HA-containing biomedical products, it is necessary to obtain high MW and the more possible uniform-size definited HA. Although little is known about what controls HA molecular weight, it is recently demonstrated that the MW of HA could be increased by over-expressing genes in the HA pathway (e.g., phosphoglucoisomerase *pgi*) in *Streptococcus equi* subspecies *zooepidemicus* (Chen et al., 2009). Similarly, Yu and Stephanopoulos (2008) demonstrated that *pgi* plays an important role in the production of high MW HA in recombinant *Escherichia coli*.

At the aim, in this work it was tried to construct artificial cassette-operons cloning *hasA-tuaD-gtaB* and *pgi* genes in pHT01 under Pgrac promoter in *B.subtilis*. Nevertheless numerous cloning trials, these plasmids have not been obtained probably due to plasmid instability phenomena.

So it is thought to construct cassette-operons in an other bacillus strain *B. megaterium*, known for the stable replication and maintenance of several extra-chromosomal DNA elements in parallel (Kim, 2003; Vary, 1992; von Tersch and Robbins, 1990) and this has been very important to cloning more genes-operons of HA pathway without plasmid instability events.

Moreover, to our knowledge no report exist on HA production in *B.megaterium*, and in this work for the first time was been engineered *B.megaterium* strains overexpressing *hasA-tuaD* or *hasA-tuaD-gtaB-pgi* artificial cassette-operons under T7 promoter control by replicative genetic. At the moment, in the optimal expression conditions (temperature, media, production time) here set, these recombinant *B.megaterium* strains produce about 2g/L in shake flask, that are good yields, being flask cultures, and very promising results in view of batch fermentation cultures.

In addition, the HA molecular weights in *hasA-tuaD* overexpressing *B.megaterium* are comparable to those of pBS5 *Bacillus subtilis* (superior at 800-1000kDa) while in *hasA-tuaD-gtaB-pgi* overexpressing *B.megaterium* seem higher, about 1800 kDa, comparable also for polydispersity to commercially available *Streptococcus* sources. We deduced that *gtaB-gpi* overexpression resulted in a molecular weights enhancement, and further trials overexpressing different HA pathway genes are doing to investigate about what influence and control HA molecular weight.

In conclusion although high potential secretory capacity of *Bacillus* was not appear in secretion of a foreign eukaryotic protein, as anti-prion 8H4 scFv here tested, *B. subtilis* and *B.megaterium* have proven to be superior expression hosts for engineering of a natural metabolic pathway, as

biosynthesis of HA, based on several criteria: 1) good quality of HA, comparable to commercial *streptococcus* standards regards to molecular mass and polydispersity, and optima HA yields. In fact, engineered pBS5 *Bacillus subtilis* strains produce HA with yields in bioreactor, much more than 4-5g/L that are actually produced by *Strepctococcus* in industrial commercial HA production by Fidia S.p.a. Moreover, innovative and promising is the HA production in engineered *B.megaterium* strains (about 2g/L in shake flask cultures).

In addition, unlike *Streptococcus*, the *B. subtilis* and *B.megaterium*-derived HA products are exotoxin free and secreted directly into the surrounding medium and are not cells associate, which should greatly simplify the recovery process and facilitate downstream processing. The production economics are further enhanced due to the ability of *Bacillus* strains to grow on minimal media, in contrast to the *Streptococcus* A and C, which are fastidious organisms and require more expensive complex media for growth. In addition to economic advantage using of minimal media with precise chemical composition, reduces accumulation of contaminants in production and purification of biological drugs, assuring a final products very pure and toxin-free.

References

References

Adam N., Ghosh P. (2001) *bHyaluronan molecular weight and polydispersity in some commercial intra-articular injectable preparations and in synovial fluid*. Inflamm. res. 50, 294–299

Agashe V.R., Hartl F.U. (2000) *Roles of molecular chaperones in cytoplasmic protein folding.* Semin Cell Dev Biol 11, 15–25.

Aguzzi A. (2005) *Prion toxicity: all sail and no anchor*. Science 308.

Andersen DC, Reilly DE. (2004) *Production technologies for monoclonal antibodies and their fragments.* Curr Opin Biotechnol.15, 456–62.

Anfisen C.B. (1973) *Principles that govern the folding of protein chains.* Science 181,223-230.

Antelmann H., Tjalsma H., Voigt B., Ohlmeier S., Bron S., van Dijl J.M., Hecker M. (2001) *A proteomic view on genome-based signal peptide predictions.* Genome Res. 11, 1484– 1502.

Archibald A. R., Hancock C., Harwood C.R. (1993) *Cell wall structure, synthesis, and turnover. In Bacillus subtilis and other Gram positive Bacteria.* Edited by Sonenshein A. L., Hoch J. A., Losick R.. Washington, DC: American Society for Microbiology pp. 381-410.

Archibald A. R., Hancock I. C., Harwood C. R. (1993) *Cell wall structure, synthesis and turnover. In Bacillus subtilis and other Gram-positive Bacteria: Biochemistry, Physiology and Molecular Genetics* (Sonenshein, A. L., Hoch, J. A., Losick, R., eds.), pp. 381-410, American Society for Microbiology, Washington, DC

Ash C., Farrow J.A., Walbanks E., Collins M. D. (1991) *Phylogenetic heterogeneity of the genus Bacillus revealed by comparative analysis of small-subunit-ribosomal RNA sequences.* Lett Appl Microbiol. 13, 202-206.

Bach H., Mazor Y., Shaky S., Shoham-Lev A., Berdichevsky Y., Gutnick D.L., Benhar I. (2001) *Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies.* J. Mol. Biol. 312, 79–93.

Balasz, E. A. (1979) *Ultrapure hyaluronic acid and use thereof.* U.S.patent 4,141,973.

Baneyx F, Mujacic M (2004) *Recombinant protein folding and misfolding in Escherichia coli.* Nat Biotechnol 22, 1399–1408.

Baneyx F. (1999) *Recombinant protein expression in Escherichia coli.* Curr Opin Biotechnol 10, 411-421

Barg H., Malten M., Jahn M. andJahn D. (2005) *Protein and vitamin production in Bacillus megaterium* Methods in Biotcnology, 18, 205-223

Batelli S.,AlbaniD.,Rametta R., Polito L., Prato F., Pesaresi M., Negro A., Forloni G., (2008) *DJ-1 modulates alpha-synuclein aggregation state in a cellular model of oxidative stress: relevance for Parkinson's disease and involvement of HSP70*.PLosOne 3, 4

Battistuta R., Negro A., Zanotti G. (2000) *Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the green fluorescent protein.* Proteins. 41, 429-437.

Bäumchen C, Roth AH, Biedendieck R, Malten M, Follmann M, Sahm H, et al. (2007) *D-mannitol production by resting state whole cell biotransformation of D-fructose by heterologous mannitol and formate dehydrogenase gene expression in Bacillus megaterium.* Biotechnol J. 2,1408-16.

Bird RE, Hardman KD, Jacobsen JW, Johnson S, Kaufman BM, Lee SM, et al. (1988) *Single-chain antigen-binding proteins.* Science;242: 426–32.

Brandner S., Isenmann S., Raeber A., Fischer M., Sailer A., Kobayashi Y., Marino S., Weissmann C., Aguzzi, A. (1996) *Normal host prion protein necessary for scrapie-induced neurotoxicity.* Nature 379, 339–343

Bruenke J, Fischer B, Barbin K, Schreiter K, Wachter Y, Mahr K, et al. (2004) *A recombinant bispecific single-chain Fv antibody against HLA class II and FcgammaRIII (CD16) triggers effective lysis of lymphoma cells.* Br J Haematol.125, 167–79.

Bueler H., Aguzzi A., Sailer A., Greiner R. A., Autenried P., Aguet M., Weissmann C. (1993) *Mice devoid of PrP are resistant to scrapie.* Cell 73, 1339–1347

Bukau B., Horwich A.L. (1998) *The Hsp70 and Hsp60 chaperone machines.* Cell 92, 351–366.

Bunk B., Schulz A., Stammen S., Münch R., Warren M.J., Rohde M., Jahn D., Rebekka (March/April 2010) *A short story about a big magic bug.* Biedendieck , Bioengineered Bugs 1, 2, 1-7;

Burger S, Tatge H, Hofmann F, Genth H, Just I, Gerhard R. (2003) *Expression of recombinant Clostridium difficile toxin A using the Bacillus megaterium system.* Biochem Biophys Res Commun. 307,584-8.

Caramelli M., Ru G., Acutis G., Forloni G. (2006) *Prion diseases.* CNS Drug 20, (1), 15-28.

Cardinale A., Filesi I., Vetrugno V., Pocchiari M., Sy M. S., Biocca S. (2005) *Trapping prion protein in the endoplasmic reticulum impairs PrPC maturation and prevents PrPSc accumulation.* J. Biol. Chem. 280, 685–694

Caron NJ, Torrente Y et al (2001) *Intracellular delivery of a TAT-EGFP fusion protein into muscle cells.* Mol Ther 3, 310–318

Chambers SP, Austen DA, Fulghum JR, Kim WM (2004) *High-throughput screening for soluble recombinant expressed kinases in Escherichia coli and insect cells.* Protein Expr Purif, 36:40-47.

Chayen NE (2004) *Turning protein crystallisation from an art into a science.* Curr Opin Struct Biol. 14,577-583.

Chen WY, Marcellin E, Hung J, Nielsen LK (2009) *Hyaluronan molecular weight is controlled by UDP-N-acetylglucosamine concentration in Streptococcus zooepidemicus.* J Biol Chem 284,18007–18014

Chien LJ, Lee CK (2007a) Enhanced *hyaluronic acid production in Bacillus subtilis by co-expressing bacterial hemoglobin.* Biotechnol Prog 23, 1017–1022

Chien LJ, Lee CK (2007b) Hyaluronic acid *production by recombinant Lactococcus lactis.* Appl Microbiol Biotechnol 77, 339–346

Christos A. Kyratsous, Saul J. Silverstein, Christine R. DeLong, Christos A. Panagiotidis (2009) *Chaperone-fusion expression plasmid vectors for improved solubility of recombinant proteins in Escherichia coli.* Gene 440, 9–15.

Clayton D. F., George J. M. (1998) *The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease.* Trends Neurosci. 21, 249–254.

Cole N. B., Murphy D. D., Grider T., Ruter S., Brasaemle D., Nussbaum R. I. (2002) *Lipid droplet binding and oligomerization properties of the Parkinson's disease protein a-synuclein.* J. Biol. Chem. 277, 6344–6352.

Collinge J. (2001) Prion *diseases of humans and animals : their causes and molecular basis.* Annu. Rev. Neurosci. 24, 519-550.

Coxon R. D. (1990) *Factors affecting protein export from Bacillus subtilis.* Ph.D. Thesis, University of Newcastle upon Tyne

David F., Westphal R., Bunk B., Jahn D., Franco-Lara E., (2010) *Optimization of antibody fragment production in Bacillus megaterium:The role of metal ions on protein secretion.* J Biotechnol. doi:10.1016/j.jbiotec.2010.07.023

Davidson W. S., Jonas A., Clayton D. F. and George J. M. (1998) *Stabilization of a-synuclein secondary structure upon binding to synthetic membranes.* J. Biol. Chem. 273, 9443-9449.

De Marco A, De Marco V (2004) *Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning.* J Biotechnol, 109,45-52.

DeAngelis PL (1999) *Hyaluronan synthases: fascinating glycosyl-transferases from vertebrates, bacterial pathogens, and algal viruses.* Cell Mol Life Sci 56, 670–682

Derossi D, Calvet S, Trembleau A, Brunissen A, Chassaing G, Prochiantz A. (1996) *Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent.* J Biol Chem. 271, 18188–93.

Deshayes S., Morris M. C., Divita G., Heitz F. (2005) *Cell-penetrating peptides: tools for intracellular delivery of therapeutics CMLS,* Cell. Mol. Life Sci. 62, 1839–1849

Dietz G.P., Bahr M. (2004) *Delivery of bioactive molecules into the cell: the Trojan horse approach,* Mol. Cell. Neurosci. 27 85–131.

Donofrio G., Heppner F. L., Polymenidou M., Musahl C., Aguzzi A. (2005) *Paracrine inhibition of prion propagation by anti-PrP single-chain Fv miniantibodies.*J. Virol. 79, 8330-8338

Earl AM, Losick R, Kolter R. (2008) *Ecology and genomics of Bacillus subtilis.* Trends Microbiol 16, 269–275.

Ellis R.J., Hartl F.U. (1999) *Principles of protein folding in the cellular environment.* Curr Opin Struct Biol 9, 102–110.

Enari M., Flechsig E., Weissmann C. (2001) *Scrapie prion protein accumulation by scrapie-infected neuroblastoma cells abrogated by exposure to a prion protein antibody.* Proc. Natl. Acad. Sci. U.S.A. 98, 9295–9299

Esposito D., Chatterjee D.K. (2006) *Enhancement of soluble protein expression through the use of fusion tags* Current Opinion in Biotechnology, 17:353–358

Fawell S, Seery J et al (1994) *Tat-mediated delivery of heterologous proteins into cells.* Proc Natl Acad Sci USA 91, 664–668

Fernandez LA, Sola I, Enjuanes L, de Lorenzo V. (2000) *Specific secretion of active single-chain Fv antibodies into the supernatants of Escherichia coli cultures by use of the hemolysin system.* Appl Environ Microbiol. 66, 5024–9.

Fink A.J. (1999) *Chaperone-mediated protein folding,* Physiological Reviews, 79, 425-449.

Fouissac E, Milas M, Rinaudo M (1993) *Shear rate, concentration, molecular weight and temperature viscosity dependences of hyaluronate, a wormlike polyelectrolytes.* Macromolecules 26, 6945–6951

Fraile S, Munoz A, de Lorenzo V, Fernandez LA. (2004) *Secretion of proteins with dimerization capacity by the haemolysin type I transport system of Escherichia coli.* Mol Microbiol.53, 1109–21.

Frankel A. D., Pabo C. O. (1988) *Cellular uptake of the tat protein from human immunodeficiency virus.* Cell 55, 1189–1193

Gabizon R., McKinley M. P., Groth D., Prusiner S. B. (1988) *Immunoaffinity purification and neutralization of scrapie prion infectivity.* Proc. Natl. Acad. Sci. U.S.A. 85, 6617–6621

Galassi G, Brun P, Radice M, Cortivo R, Zanon GF, Genovese P, Abatangelo G (2000) *In vitro reconstructed dermis implanted in human wounds: degradation studies of the HA-based supporting scaffold.* Biomaterials 21, 2183–2191

Gamer M., Fröde D., Biedendieck R., Stammen S.,Jahn D. (2009). *A T7 RNA polymerase-dependent gene expression system for Bacillus megaterium.* Appl. Microbiol. Biotechnol 82, 1195-1203.

George J. M., Jin H., Woods W. S. and Clayton D. F. (1995) *Characterization of a novel protein regulated during the critical period for song learning in the zebra finch.* Neuron 2, 361–372.

Georgiou G., Valax P. (1996) *Expression of correctly folded proteins in Escherichia coli*, Curr. Opin. Biotechnol. 7, 190–197.

Goa K. L., Benfield P. (1994) *Hyaluronic acid: a review of its pharmacology and use as a surgical aid in ophthalmology and its therapeutic potential in joint disease and wound healing.* Drugs 47,536–566.

Gottesman S., Wickmer S., Maurizi M.R. (1997) *Protein quality control: Triage by chaperones and proteases.* Genes Dev. 11, 815–823.

Gragerov A, Nudler E, Komissarova N, Gaitanaris GA, Gottesman ME, Nikiforov V (1992) *Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in Escherichia coli.* Proc Natl Acad Sci USA, 89, 10341-10344.

Gragerov A., Zeng L., Zhao X., Burkholder W., Gottesman M.E. (1994) Specificity *of DnaK-peptide binding.* J Mol Biol 235, 848–854.

Gump J.M., Dowdy S.F. (2007) *TAT transduction: the molecular mechanism and therapeutic prospects.* Trends Mol.Med. 13, 443-448.

Gustafsson C., Govindarajan S., Minshull J., (2004) *Codon bias and heterologous protein expression.* Trends biotechnol.22, 346-353.

Hallbrink M, Floren A, Elmquist A, Pooga M, Bartfai T, Langel U. (2001) *Cargo delivery kinetics of cell-penetrating peptides.* Biochim Biophys Acta 1515, 101–109.

Hammarstrom M, Hellgren N, van Den Berg S, Berglund H, Hard T (2002) *Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli.* Protein Sci, 11,313-321.

Hannig G., Makrides S.C. (2009) *Strategies for optimizing heterologous protein expression in Escherichia coli.* Trends Biotechnol.

Harris D.A, True H.L. (2006) *New insight into prion structure and toxicity*. Neuron 50, 353-357.

Harris D.A. (2003) *Trafficking, turnover and membrane topology of PrP*. Br. Med. Bull. 66, 71-85.

Harrison C.J., Hayer-Hartl M., Di Liberto M., Hartl F.-U., Kuriyan, J. (1997) *Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK.* Science 276, 431–435.

Hartl UF., Hayer-Hartl M. (2002) *Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein,* Science. 295, 1852-1858.

Harwood C. R. (1992) *Bacillus subtilis and its relatives: molecular biological and industrial workhorses.* Trends Biotechnol, 10 (7), 247–256.

Harwood C. R., Cranenburgh R. (2008) *Bacillus protein secretion: an unfolding story* Trends in Microbiology 16, 2

Hascall, V. C.,. Laurent T. C. (1997) *Hyaluronan: structure and physical properties.* In V. C. Hascall and M. Yanagishita (ed.), Science of hyaluronan today. Seikagaku Corp., Tokyo, Japan. [Online.] www.GlycoForum.gr.jp.

Hashimoto M., Yoshimoto M., Sisk A., Hsu L. J., Sundsmo M., Kittel A., Saitoh T., Miller A., Masliah E. (1997) *NACP, a synaptic protein involved in Alzheimer's disease, is differentially regulated during megakaryocyte differentiation.* Biochem. Biophys. Res. Commun. 237, 611–616.

Heng B.C., Cao T. (2005) *Making cell-permeable antibodies (Transbody) through fusion of protein transduction domains (PTD) with single chain variable fragment (scFv) antibodies: Potential advantages over antibodies expressed within the intracellular environment (Intrabody).* Medical Hypotheses 64, 1105–1108

Henner DJ. (1990) *Expression of heterologous genes in Bacillus subtilis.* Meth Enzymol. 185, 199–201.

Heo M.A., Kim S.H., Kim S.Y., Kim Y.J., Chung J., Oh M.K., Lee S.G (2006) *Functional expression of single-chain variable fragment antibody against c-Met in the cytoplasm of Escherichia coli,* Protein Expr. Purif. 47, 203–209.

Heppner F. L., Aguzzi A. (2004) *Recent developments in prion immunotherapy.* Curr. Opin. Immunol. 16, 594–598

Hoffmann F, Rinas U (2004) *Roles of heat-shock chaperones in the production of recombinant proteins in Escherichia coli.* Adv Biochem Eng Biotechnol 89, 143–161

Hollmann R, Deckwer WD (2004) *Pyruvate formation and suppression in recombinant Bacillus megaterium cultivation.* J Biotechnol 111, 89–96

Hu W., Kieseier B., Frohman E. et al. (2008) Prions *proteins: physiological functions and role in neurological disorders.* J Neuro Sci 264, 1-8.

Hu X., O'Hara L., White S., Magner E., Kane M., Wall J.G. (2007) *Optimisation of production of a domoic acid-binding scFv antibody fragment in Escherichia coli using molecular chaperones and functional immobilisation on a mesoporous silicate support,* Protein Expr. Purif. 52, 194–201.

Huston J.S., Levinson D., Mudgett-Hunter M., Tai M-S, Novotny J., Margolies MN., et al. (1988) *Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli.* Proc Natl Acad Sci U S A.85,5879-83.

Jannière L., Bruand C., Ehrlich S.D. (1990). *Structurally stable Bacillus subtilis cloning vectors,* Gene 87,53-61.

Jongbloed J.D.H., Antelmann H., Hecker M., Nijland R., Bron S., Airaksinen U. , Pries F., Quax W.J. , van Dijl J.M., Braun P.G. (2002) *Selective contribution of the twin-arginine translocation pathway to protein secretion in Bacillus subtilis,* J. Biol. Chem. 277, 44068– 44078.

Jongbloed J.D.H., Martin U., Antelmann H., Hecker M., Tjalsma H., Venema G., Bron S., van Dijl J.M., Muller J.P. (2000) *TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway.* J. Biol. Chem. 2275, 41350– 41357.

Jordan E., Hust M., Roth A., Biedendieck R., Schirrmann T., Jahn D., Dübel S. (2007) *Production of recombinant antibody fragments in Bacillus megaterium* Microbial Cell Factories, 6:2

Jurado P., de Lorenzo V., Fernández L.A. (2006) *Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of Escherichia coli: evidence that chaperone activity is the prime effect of thioredoxin,* J. Mol. Biol. 357, 49–61.

Jurado P., Ritz D., Beckwith J., de Lorenzo V., Fernandez L.A. (2002) *Production of functional single-chain Fv antibodies fragments in the cytoplasm of Escherichia coli.* J. Mol. Biol. 320, 1–10.

Jürgen B., Schweder T., Hecker M. (1998) *The stability of mRNA from the gsiB gene of bacillus subtilis is dipendent on the presence of a strong ribosome binding site.* Mol.Gen.Genet. 538-545

Kanovsky M., Raffo A., Drew L., Rosal R., Do T., Friedman F.K., Rubinstein P., Visser J., Robinson R., Brandt-Rauf P.W., Michl J., Fine R.L., Pincus M.R. (2001) *Peptides from the amino terminal mdm-2-binding domain of p53, designed from conformational analysis, are selectively cytotoxic to transformed cells*. Proc. Natl. Acad. Sci. U. S. A. 98(22), 12438–12443.

Kieselburg M. K., Weickert M., Vary P. (1984) *Analysis of resident and transformant plasmids in Bacillus megaterium.* Biotechnolou 2, 254-259.

Knudson, C. B., W. Knudson (1993) *Hyaluronan-binding proteins in development, tissue homeostasis, and disease.* FASEB J. 7,1233–1241.

Kontinen V. P, Sarvas M. (1993) *The PrsA lipoprotein is essential for protein secretion in Bacillus subtilis and sets a limit for high-level secretion.* Mol. Microbiol. 8:727–737.

Kunst F., Ogasawra N., Moszer I., Albertini A. M., Alloni G., Azevedo V., Bertero M. G., Bessieres P., Bolotin A., Borchert S. et al. (1997) *The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.* Nature (London) 390, 249-256

Langer T, Lu C, Echols H, Flanagan J, Hayer MK, Hartl FU (1992) *Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding.* Nature, 356, 683-689.

Laurent, T. C., Fraser J. R. E. (1992*) Hyaluronan.* FASEB J. 6:2397–2404.

LaVallie E.R., McCoy J.M. (1995) Gene *fusion expression systems in Escherichia coli*. Curr. Opin. Biotechnol. 6, 501–506.

Lazarevic V., Margot P., Soldo B., Karamata D. (1992) *Sequencing and analysis of the Bacillus subtilis lytRABC divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl- L -alanine amidase and its modifier*. J. Gen. Microbiol. 138, 1949-1961

Lee SY (1996) *High cell density cultivation of Escherichia coli*. Trends Biotechnol 14,98–105

Levy R., Weiss R., Chen G., Iverson B.L., Georgiou G. (2001) *Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones.* Protein Expr. Purif. 23, 338–347.

Ludewigs H., Zuber C., Vana K., Nikles D., Zerr I., Weiss S. (2007) *Therapeutic approaches for prion disorders.* Expert Rev. Anti Infect. Ther. 5, 613–630

Makrides SC (1996) *Strategies for achieving high-level expression of genes in Escherichia coli.* Microbiol Rev. 60, 512–538

Makvandi-Nejad S, McLean M, Hirama T, Almquist K, MacKenzie C, Hall J. (2005) *Transgenic tobacco plants expressing a dimeric single-chain variable fragment (scFv) antibody against Salmonella enterica serotype Paratyphi B.* Transgenic Res. 14, 785–92.

Mallucci G. et al. (2003). *Depleting neuronal PrP in prion infection prevents disease and reverses spongiosis*.Science 302, 871-874

Mallucci G., Collinge J. (2005) *Rational targeting for prion therapeutics.* Nature reviews. 6, 23-34

Mann D.A., Frankel A.D. (1991) *Endocytosis and targeting of exogenous HIV-1 Tat protein*. EMBO J. 10, 1733–1739

Mao Z, Chen RR (2007) *Recombinant synthesis of hyaluronan by Agrobacterium sp.* Biotech Prog 23, 1038–1042

Mao Z, Shin H-D, Chen R (2009) A *recombinant E. coli bioprocess for hyaluronan synthesis.* Appl Microbiol Biotechnol 84, 63–69

Margot P., Karamata D. (1996) *The wprA gene of Bacillus subtilis 168, expressed during exponential growth, encodes a cell-wall-associated protease.* Microbiology 142, 3437-3444

Martin C, Rojas G, Mitchell J, Vincent K, Wu J, McCafferty J, et al. (2006) *A simple vector system to improve performance and utilisation of recombinant antibodies.* BMC Biotechnol 6: 46.

Martinez-Alonso M, Vera A, Villaverde A (2007) *Role of the chaperone DnaK in protein solubility and conformational quality in inclusion body-forming Escherichia coli cells.* FEMS Microbiol Lett, 273,187-195.

Maynard J., Georgiou G. (2000) *Antibody Engineering.* Annu Rev Biomed Eng. 2, 339-76.

Meinhardt F., Stahl U., Ebeling W. (1989) *Highly efficient expression of homologous and heterologous genes in Bacillus megaterium.* Appl Microbiol Biotechnol. 30, 343-350.

Mercante R., Pooley H. M., Karamata D. (1995) *A periplasm in Bacillus subtilis.* J. Bacteriol. 177:6176-6183.

Meyer K, Palmer JW (1934) *The polysaccharide of the vitreous.* J Biol Chem 107:629–634

Miroux B, Walker JE (1996) *Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.*J Mol Biol, 260,289-298.

Mogk A, Deuerling E, Vorderwulbecke S, Vierling E, Bukau B (2003) *Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation.* Mol Microbiol, 50,585-595.

Nagahara H., Vocero-AkbaniA., Snyder E., Ho E.L.,Latham A., Lissy D., Becker-Hapac, Dowdy S:F: (1998) Nat.Med. 4, 1449-1452.

Nagarajan V.(1993) Protein secretion In A.L.Sonenshein, J.A. Hoch, R.Losick(ed.) *Bacillus subtilis and other gram-positive bacteria.* American Society for Microbiology, Washington D.C. 713-726.

Nakajo S., Omata K., Aiuchi T., Shibayama T., Okahashi I., Ochiai H., Nakai Y., Nakaya K., Nakamura Y. (1990) *Purification and characterization of a novel brain-specific 14 kDa protein.* J. Neurochem. 55, 2031-2038.

Nakase I. et al. (2004) *Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement.* Mol. Ther. 10, 1011–1022

Natsume A, Wakitani M, Yamane-Ohnuki N, Shoji-Hosaka E, Niwa R, Uchida K, et al. (2006) *Fucose removal from complex-type oligosaccharide enhances the antibody-dependent cellular cytotoxicity of single-gene-encoded bispecific antibody comprising of two single-chain antibodies linked to the antibody constant region.* J Biochem.140, 359–68.

Negro A, Ballarin C, Bertoli A, Massimino ML, Sorgato MC. (2001) *The metabolism and imaging in live cells of the bovine prion protein in its native form or carrying single amino acid substitutions.* Mol Cell Neurosci. Mar;17, 521-38.

Negro A, Onisto M, Masiero L, Garbisa S. (1995) *Synthesis and refolding of human TIMP-2 from E. coli, with specific activity for MMP-2.* FEBS Lett. Feb 20:360(1), 52-6.

Nguyen D.H., Nguyen Q.A., Ferreira R.C., Ferreira L.C.S., Tran L.T., Schumann W. (2005) *Construction of plasmid-based expression vectors for Bacillus subtilis.* Plasmid. 54, 241-8.

Nielsen, M. S., H. Vorum, E. Lindersson, P. H. Jensen. (2001) *Ca 2+binding to alpha-synuclein regulates ligand binding and oligomerization.* J. Biol. Chem. 276, 22680–22684.

Nina E., Weisser J., Hall C. (2009) *Applications of single-chain variable fragment antibodies in therapeutics and diagnostics* Biotechnology Advances. 27, 502–520

Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T (1998) *Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli* Appl Environ Microbiol,64:1694-1699.

Noone D., Howell A., Collery R., Devine K. M. (2001) *YkdA and YvtA, HtrA- like serine proteases in Bacillus subtilis, engage in negative autoregulation and reciprocal cross-regulation of ykdA and yvtA gene expression*. J. Bacteriol. 183,654–663.

Oganesyan N., Kim S-H, Kim R. (2004) On-*column chemical refolding of proteins.* PharmaGenomics, 22-24

Ostrerova N., Petrucelli L., Farrer M., Mehta N., Choi P., Hardy J., Wolozin B. (1999) *α-Synuclein shares physical and functional homology with 14-3-3 proteins.* J. Neurosci. 19, 5782-5791.

Padiolleau-Lefèvre S, Débat H, Phichith D, Thomas D, Friboulet A, Avalle B. (2006) *Expression of a functional scFv fragment of an anti-idiotypic antibody with a beta-lactam hydrolytic activity*. Immunol Lett.103, 39–44.

Park S. M., Jung H. Y., Kim H. O., Paik S. R., Chung K. C., Kim J.(2002) *Evidence that a-synuclein functions as a negative regulator of Ca 2+ -dependent a-granule release from human platelets.* Blood 100, 2506–2514.

Peretz D., Williamson R. A., Kaneko K., Vergara J., Leclerc E., Schmitt-Ulms G., Mehlhorn I. R., Legname G., Wormald M. R., Rudd P. M. et al. (2001) *Antibodies inhibit prion propagation and clear cell cultures of prion infectivity.* Nature 412, 739–743

Phan, T.T.P., Nguyen, H.D., Schumann, W. (2005). *Novel plasmid-based expression vectors for intra- and extracellular production of recombinant proteins in Bacillus subtilis.* Protein Expr. Purif. 46, 189-95

Phelan A., Elliott G., O'Hare P. (1998) *Intercellular delivery of functional p53 by the herpesvirus protein VP22.* Nat. Biotechnol. 16(5), 440–443.

Prasad S., Jayaraman G., Ramachandran K. (2010) *Hyaluronic acid production is enhanced by the additional co-expression of UDP-glucose pyrophosphorylase in Lactococcus lactis.* Appl Microbiol Biotechnol 86, 273–283

Prehm P (1984) *Hyaluronate is synthesised at plasma membrane.* Biochem J 220:597–600

Priest F. G., Coodfellow M., Todd C. (1988). *A numerical classification of the genus Bacillus*. J Gen Microbiol l34, 1847-1 882.

Prusiner S.B. (1998a) *Novel proteinaceous infectious particles cause scrapie*. Science 216, 136-144.

Prusiner S.B. (1998b) *Prions.* Prot. Natl Acad. Sci. USA 95, 13363-13383.

Puttick M. P., Wade J. P., Chalmers A., Connell D. G., Rangno K. K. (1995) *Acute local reactions after intraarticular hylan for osteoarthritis of the knee.* J. Rheumatol. 22,1311–1314.

Qing G, Ma LC, Khorchid A, Swapna GV, Mal TK, Takayama MM, Xia B, Phadtare S, Ke H, Acton T et al.(2004) *Cold-shock induced highyield protein production in Escherichia coli.* Nat Biotechnol, 22, 877-882.

Recchia A., Rota D., Debutto P., Peroni D., Guidolin D., Negro A., Skaper S.D., Giusti P. (2008) *Generation of a alpha-synuclein-based rat model of Parkinson's disease.* Neurobiol.Dis. 30, 8-18.

Ren G., Wang X, Hao S., Hu H., Wang C. (Apr. 2007) *Translocation of □-Synuclein Expressed in Escherichia coli.* Journal of bacteriology, 2777–2786

Richt J. A., Kasinathan P., Hamir A. N., Castilla J., Sathiyaseelan T., Vargas F.,Sathiyaseelan J., Wu H., Matsushita H., Koster J. et al. (2007) *Production of cattle lacking prion protein*. Nat. Biotechnol. 25, 132–138

Rogers H. J., Perkins H. R., Ward J. B. (1980) *Microbial Cell Walls and Membranes*, Chapman and Hall, London.

Rostas K., Dobritsa S. V., Dobritsa A. P., Konez C. S., Alfoldi L. (1980). *Megacinogenic plasmid from Bacilluls megaterium 216*. Mol Gen Genet 180, 323-329.

Rothman,J.E. (1989) *polypeptide chain binding ptoyrins: catalysts of protein folding and related processes in cells.* Cell 59, 591-601.

Rygus T, Scheler A, Allmansberger R, Hillen W. (1991b) *Molecular cloning, structure, promoters and regulatory elements for transcription of the Bacillus megaterium encoded regulon for xylose utilization.* Arch Microbiol. 155, 535-42.

Rygus T, Hillen W. (1991a) *Inducible high-level expression of heterologous genes in Bacillus megaterium using the regulatory elements of the xylose-utilization operon.* Appl Microbiol Biotechnol. 35, 594-599.

Ryu J, Han K et al (2003) Enhanced *uptake of a heterologous protein with an HIV-1 Tat protein transduction domains (PTD) at both termini.* Mol Cells 16, 385–391

Ryu J, Lee HJ et al (2004) *Intracellular delivery of p53 fused to the basic domain of HIV-1 Tat*. Mol Cells 17, 353–359

Saettone MF, Giannaccini B, Chetoni P, Torracca MT et al (1991) *Evaluation of high-molecular-weight and low-molecular-weight fractions of sodium hyaluronate and an ionic complex as adjuvants for topical ophthalmic vehicles containing pilocarpine.* Int J Pharm 72,131–139

Sambrook J., Fritschi E., Maniatis T. (1989) *Molecular cloning: a laboratory manual.* (New York: Cold Spring Harbor Laboratory Press).

Sarvas M., Harwood C.R., Bron S.,van Dijl J.M. (2004) *Post-translocational folding of secretory proteins in Gram-positive bacteria* Biochimica et Biophysica Acta. 1694, 311 –327

Saso L, Bonanni G, Grippa E, Gatto MT et al (1999) *Interaction of hyaluronic acid with mucin, evaluated by gel permeation chromatography.* Res Commun Mol Pathol Pharmacol 104,277–284

Saunders C.W., Schmit B.J., Mallonee R.L., Guyer M.S. (1987) *Secretion of human serum albumin from Bacillus subtilis.* J.Bacteriol. 169, 2917-2925

Saxena A., Zhang R. W., Bollag M. (1987) *Microorganism incapable of metabolizing the herbicide metolachlor.* Appl Envirorr Microbiol53, 390-396.

Schallmey M, Singh A, Ward OP (2004) *Developments in the use of Bacillus species for industrial production.* Can J Microbiol 50:1–17

Schlieker C, Bukau B, Mogk A (2002) *Prevention and reversion of protein aggregation by molecular chaperones in the E. coli cytosol: implications for their applicability in biotechnology.* J Biotechnol, 96,13-21.

Schlieker C, Bukau B., Mogk A. (2002) *Prevention and reversion of protein aggregation by molecular chaperones in the E. coli cytosol: implications for their applicability in biotechnology,* Journal of Biotechnology, 96,13–21

Schlieker C, Tews I, Bukau B, Mogk A (2004) *Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides.* FEBS Lett, 578, 351-356.

Schmidt FR (2004) *Recombinant expression systems in the pharmaceutical industry*. Appl Microbiol Biotechnol 65,363–372

Schwarze SR, Dowdy SF (2000) *In vivo protein transduction: intracellular delivery of biologically active proteins, compounds and DNA.* Trends Pharmacol Sci 21, 45–48

Schwarze S.R., Ho A., Vocero-Akbani A., Dowdy S.F. (1999) *In vivo protein transduction: delivery of a biologically active protein into the mouse.* Science 285(5433), 1569–1572.

Schwarze S.R., Hruska K.A., Dowdy S.F. (2000) *Protein transduction: unrestricted delivery into all cells?* Trends Cell Biol.10, 290-295.

Shevchik V.E., Condemine G., Baudouy J.R. (1994) *Characterization of DsbC, a periplasmic protein of Erwinia chrysanthemi and Escherichia coli with disulfide isomerase activity,* EMBO J. 13 2007–2013.

Shin E. C., Cho S. E., Lee D. K., Hur M. W., Paik S. R., Park J. H., Kim J. (2000) *Expression patterns of a-synuclein in human hematopoietic cells and in Drosophila at different developmental stages.* Mol. Cells 10, 65–70.

Shyng S. L., Heuser J.E., Harris D.A. (1994) *A glycolipid-anchoredprion protein is endocytosed via clathrin-coated pits.* J.Cell. Biol. 125,1239-1250.

Shyng S.L., Huber M.T., Harris D.A. (1993) *A prion protein cycles between the cell surface and the endocytic compartment in cultered neuroblastoma cells.* J.Biol. Chem. 268, 15922-15928.

Sigurdsson E. M., Sy M. S., Li R., Scholtzova H., Kascsak R. J., Kascsak R., Carp R., Meeker H. C., Frangione B., Wisniewski T. (2003) *Anti-prion antibodies for prophylaxis following prion exposure in mice.* Neurosci. Lett. 336, 185–187

Simonem M., Palva I, (1993) *Protein secretion in bacillus species.* Microbiol.Rev. 57, 109-137

Skoulakis E.M., Davis R.L. (1998) Mol. Neurobiol. 16, 269-284.

Snedecor B, Nishihara J, Joly J, McFarland N, Andersen D, et al. (2004*) High-level accumulation of a recombinant antibody fragment in the periplasm of Escherichia coli requires a triple-mutant (degP prc spr) host strain.* Biotechnol Bioeng. 85, 463–74.

Snyder EL, Dowdy SF (2004) *Cell penetrating peptides in drug delivery*. Pharm Res 21, 389–393

Spillantini M.G., Crowther RA., Jakes, R., Hasegawa M., Goedert M (1998) *α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies.* Proc. Natl. Acad. Sci. USA. 95, 6469-6473.

Spillantini M.G., Schmidt M.L., Lee V.M., Trojanowski J.Q., Jakes R., Goedert M. (1997) *α-Synuclein in Lewy bodies.* Nature, , 388, 839-840.

Spurlock SL, Spurlock GH, Bernstad S, Michanek P et al. (1999) *Treatment of acute superficial flexor tendon injuries in performance horses with high molecular weight sodium hyaluronate.* J Equine Vet Sci 19, 338–344

Stefanis L., Kholodilov N., Rideout H. J., Burke R. E., Greene L. A. (2001) *Synuclein-1 is selectively up-regulated in response to nerve growth factor treatment in PC12 cells.* J. Neurochem. 76, 1165–1176.

Stephenson K., Harwood C.R. (1998) *Influence of a cell-wall associated protease on production of alpha-amylase by Bacillus subtilis.* Appl.Environ.Microbiol 64, 2875-2881

Sthal N., Borchelt D.R., Hsiao K.,Prusiner S.B. (1987) *Scrapie prion proteins contain a phpsphatidylinositolglycolipid.* Cell 51, 229-240.

Studier F. W., Moffatt B. A. (1986). *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.* J. Mol. Biol 189, 113-130.

Suzuki Y, Yamaguchi T (1993) *Effect of hyaluronic acid on macrophage phagacytosis and active oxygen release.* Agents Actions 38:32–37

Tabor S., Richardson C. C. (1985) *A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes.* Proc. Natl. Acad. Sci. U.S.A 82, 1074-1078.

Taraboulos A., Raeber A., Borchelt D.R., Serban D., Prusiner S.B. (1992) *Synthesis and trafficking of prion proteins in cultered cells.* Mol. Biol. Cell 3, 851-863.

Thomas JG, Baneyx F (1996) *Protein folding in the cytoplasm of Escherichia coli: requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines.* Mol Microbiol, 21,1185-1196.

Titok M.A., Chapuis J., Selezneva Y.V., Lagodich A.V., Prokulevich V.A., Ehrlich S.D., Jannière, L. (2003). *Bacillus subtilis soil isolates: plasmid replicon analysis and construction of a new theta-replicating vector.* Plasmid 49, 53-62.

Tjalsma H., Antelmann H., Jan D.H. Jongbloed, Braun P. G., Darmon E., Dorenbos R.,Jean-Yves F. Dubois, Westers H., Zanen G., Quax W., Kuipers O., Bron S., Hecker M., van Dijl M. (2004) *Proteomics of Protein Secretion by Bacillus subtilis: Separating the "Secrets" of the Secretome* Microbiology and molecular biology reviews, 68, 2, 207–233

Tjalsma H., Bolhuis A., van Roosmalen M. L., Wiegert T., Schumann W., Broekhuizen C. P., Quax W., Venema G., Bron S., van Dijl J. M. (1998) *Functional analysis of the secretory precursor processing machinery of Bacillus subtilis: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases.* Genes Dev. 12:2318–2331.

Tjalsma H., Bolhuis A.,. Jongbloed J. D. H, Bron S., van Dijl. J. M (2000) *Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome.* Microbiol. Mol. Biol. Rev. 64:515–547.

Toole B. P. (1991) Proteoglycans and hyaluronan in morphogenesis and differentiation, p. 305–341. In E. D. Hay (ed.), *Cell biology of the extracellular matrix.* Plenum Press, New York, N.Y.

Toole BP, Wight TN, Tammi MI (2002) *Hyaluronan–cell interactions in cancer and vascular disease.* J Biol Chem 277,4593–4596

Torchilin VP (2002) *TAT peptide-modified liposomes for intracellular delivery of drugs and DNA.* Cell Mol Biol Lett 7, 265–267

Trevitt C. R., Collinge J. (2006) *A systematic review of prion therapeutics in experimental models.* Brain 129, 2241–2265

Ueda K., Fukushima H., Masliah E., Xia Y., Iwai A., Yoshimoto M., Otero D. A., Kondo J., Ihara Y., Saitoh T. (1993) *Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease.*Proc. Natl. Acad. Sci. USA 90,11282–11286.

van Wely KHM, Swaving J, Freudl R, Driessen AJM (2001) *Translocation of proteins across the cell envelope of Gram-positive bacteria.* FEMS Microbiol Rev 25, 437–454

van Dijl J.M., de Jong A., Vehmaanpera, Venema G., Bron S. (1992) *signal peptidase I of Bacillus subtilis: patterns of conserved aminoacids in prokaryotic and eukaryotic type I signale peptidase.* EMBO J.11, 2819-2828.

Vary P.S., 1994. *Prime time for Bacillus megaterium.* Microbiology 140, 1001–1013.

Venturi M., Seifert C., Hunte C. (2002) *High level production of functional antibody Fab fragments in an oxidizing bacterial cytoplasm.* J. Mol. Biol. 315, 1–8.

Vetrugno V., Cardinale A., Filesi I., Mattei S., Sy M. S., Pocchiari M., Biocca S.(2005) *KDEL-tagged anti-prion intrabodies impair PrP lysosomal degradation and inhibit scrapie infectivity.* Biochem. Biophys. Res. Commun. 338, 1791–1797

Vives E. et al. (1997) *A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus.* J. Biol. Chem. 272, 16010–16017

von Heijne, G. (1990b). *The signal peptide.* J. Membr. Biol. 115,195–201.

von Heijne, G. (1998) *Life and death of a signal peptide.* Nature 396,111-113.

Von Tersch M. A., Carlton B. C. (1984) *Molecular cloning of structural and immunity genes for megacins A-216 and A-19213 in Bacillzls megaterizlm.* J Bacteriol 160, 854-859.

Von Tersch M. A., Carlton B.C. (1983) *Megacinogenic plasmids of Bacillus megaterium.* J Bacteriol 155, 872-877.

Von Tersch M. A., Robbins H. L. (1990). *Efficient cloning in Bacillus megaterium : comparison to Bacillus subtilis and Escherichia coli cloning hosts.* FEMS Microbiol Lett 70, 305-310.

Vorobjeva P., Khmel A., Alfoldi L. (1980) *Transformation of Bacillus megacin protoplasts by plasmid DNA.* FEMS Microbiol Lett 7, 261-263.

Wadia J.S. et al. (2004) *Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis.* Nat. Med. 10, 310–315

Wadia J.S., Dowdy S.F. (2005) *Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer.* Adv. Drug Deliv. Rev. 57(4), 579–596.

Wang H, Dai J, Li B, Fan K, Peng L, Zhang D, et al. (2008) *Expression, purification, and characterization of an immunotoxin containing a humanized anti-CD25 single-chain fragment variable antibody fused to a modified truncated Pseudomonas exotoxin A.* Protein Expr Purif. 58, 140–7.

Ward J. B. (1981) Teichoic *and teichuronic acids: Biosynthesis, assembly and location*. Microbiol. Rev. 45, 211-243

Weissman, B., K. Myer. (1954) *The structure of hyalbiuronic acid and hyaluronic acid from umbilical cord*. J. Am. Chem. Soc. 76,1753–1757.

Weissmann C., Aguzzi A. (2005) *Approaches to therapy of prion diseases*. Annu. Rev. Med. 56, 321–344

Wessels M.R, Moses A, Goldberg JB, DiCesare TJ (1991) *Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci.* Proc Natl Acad Sci USA 88,8317–8321

Westers L., Westers H., Quax W.J. (2004) *Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism.* Biochimica et Biophysica Acta. 1694 (1-3), 299– 310

Widmann M, Christen P (2000) *Comparison of folding rates of homologous prokaryotic and eukaryotic proteins.* J Biol Chem, 275, 18619-18622.

Widner B., Behr R., Dollen S.V., Tang M., Heu T., Sloma A., Sternberg A, DeAngelis P., Weigel P., Brown S. (July 2005) *Hyaluronic Acid Production in Bacillus subtilis,* Applied and environmental microbiology , 3747–3752

Wills KN, Atencio IA et al (2001) *Intratumoral spread and increased efficacy of a p53-VP22 fusion protein expressed by a recombinant adenovirus.* J Virol 75, 8733–8741

Wittchen K., Meinhardt F. (1995) *Inactivation of the major extracellular protease from Bacillus megaterium DSM319 by gene replacement.* Applied Microbiology and Biotechnology 42, 871-877.

Wohlrab W., Reinhard N. (2000) *Skin protection agents containing a fragment mixture produced from hyaluronic acid by hydrolysis.* World patent application WO00/38647.

Wong S.-L. (1995) *Advances in the use of Bacillus subtilis for the expression and secretion of heterologous proteins*. Curr. Opin. Biotechnol. 6 (5), 517–522.

Wu SP, Fu AL et al (2006) *A novel therapeutic approach to 6-OHDA-induced Parkinson's disease in rats via supplementation of PTD-conjugated tyrosine hydroxylase.* Biochem Biophys Res Commun 346, 1–6

Yu H, Stephanopoulos G (2008) *Metabolic engineering of Escherichia coli for biosynthesis of hyaluronic acid.* Metab Eng 10, 24–32

Zanusso G., Liu D., Hegyi I., Yin X., Aguzzi A., Hornemann S., Liemann S., Glockshuber R., Manson J.C., Brown P., Petersen R.B., Gambetti P., Sy M.S. (1998) *Prion protein expression in different species: analysis with a panel of new mAbs,* proc. Natl. Acad. Sci.USA 95, 8812-8816

Zeng C, Toople BP, Kinney SD, Kuo JW, Stamenkovic I (1998) *Inhibition of tumor growth in vivo by hyaluronan oligomers.* Int J Cancer 77,396–410

Zheng L. , Baumann U. , . Reymond J.L (2003) *Production of a functional catalytic antibody ScFv–NusA fusion protein in bacterial cytoplasm,* J. Biochem. 133, 577–581.

Zhu X., Zhao X., Burkholder W.F., Gragerov, A., Ogata C.M., Gottesman M.E., Hendrickson W.A. (1996) *Structural analysis of substrate binding by the molecular chaperone DnaK.* Science 272, 1606–1614.

Zimmerman S.B., Trach, S.O. (1991) *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli.* J.Mol.Biol. 222, 599-620.

Zuber C., Knackmuss S., Rey C., Reusch U., Rottgen P., Frohlich T., Arnold G. J.,Pace C., Mitteregger G., Kretzschmar H. A. et al. (2008) *Single chain Fv antibodies directed against the 37 kDa/67 kDa laminin receptor as therapeutic tools in prion diseases.* Mol. Immunol. 45, 144–151

Zuber C., Mitteregger G., Schuhmann N., Rey C., Knackmuss S., Rupprecht W., Reusch U., Pace C., Little M., Kretzschmar H. A. et al. (2008) *Delivery of single-chain antibodies (scFvs) directed against the 37/67 kDa laminin receptor into mice via recombinant adeno-associated viral vectors for prion disease gene therapy.* J. Gen. Virol.89, 2055–2061