UNIVERSITA' DEGLI STUDI
DI MILANO – BICOCCA

SCUOLA DI DOTTORATO DI SCIENZE

Facoltà di Scienze Matematiche, Fisiche e Naturali

Dipartimento di Biotecnologie e Bioscienze

Corso di Dottorato di Ricerca in Biologia, XXIII ciclo

An evaluation of an innovative technological platform for straightforward purification of recombinant proteins

Dott. Daniele Cappellini

Anno Accademico 2009-2010

Dottorato in Biologia, XXIII ciclo

Dott. Daniele Cappellini

Matricola: 040925

Tutor: Prof. Paolo Tortora

Il lavoro presentato in questa tesi è stato realizzato presso i laboratori di Biologia Molecolare e Biochimica di DiaSorin S.p.A., sotto la supervisione del dott. Pier Natale Brusasca

Università degli Studi di Milano-Bicocca
Piazza dell'Ateneo Nuovo 1, 20126, Milano

Dipartimento di Biotecnologie e Bioscienze
Piazza della Scienza 2, 20126, Milano

TABLE OF CONTENT

INTRODUCTION

In process biotechnology, purification of proteins from complex biological mixtures involves a series of complicated recovery steps, each of which can compromise the purity and yield of the desired product.

Although the expression of proteins has been greatly simplified by the advent of recombinant DNA techniques, their purification remains a critical problem. In the past decades, a number of protein expression systems have been designed to solve this problem by the expression of recombinant proteins incorporating a purification tag: a carrier protein or peptide that displays highly specific and reversible binding to a ligand (Nilsson B. 1992). In these systems the fusion protein is typically purified from contaminants by affinity chromatography using an immobilized, moderate-affinity ligand, specific to the purification tag: these systems typically allow one-step purification of a recombinant protein from cell extract by affinity chromatography (Nilsson J. 1997). A large number of purification tags and immobilized affinity ligands are now commercially available; some examples are: maltose binding protein (Maina C.V. 1988), glutathione S-transferase (Smith D.B. 1988), biotin carrier protein (Tsao K.L. 1996), thioredoxin (LaVallie E.R. 1993) and cellulose binding domain (Ong E. 1989).Similarly, vectors for fusion to short peptide tags, such as oligohisitidine(Smith M.C. 1988), S-peptide (Kim J.S. 1993), and the FLAG™ peptide (Su X. 1992) are also available.

A typical protocol for expression and purification of a maltose binding protein-fused protein is shown in figure 1: the target gene is cloned into the expression vector, then after protein expression, the cell lysate is loaded on an amylose resin: only the species containing the MBP tag, which is affine to amylase, are retained, while all the others are washed away. Elution is achieved by addition of maltose which competes with amylase ligands for MBP. An additional step is the cleavage of the target protein from the tag by a protease specific to a site previously inserted between the target protein and the tag. Further purification is needed to separate the target protein from MBP.

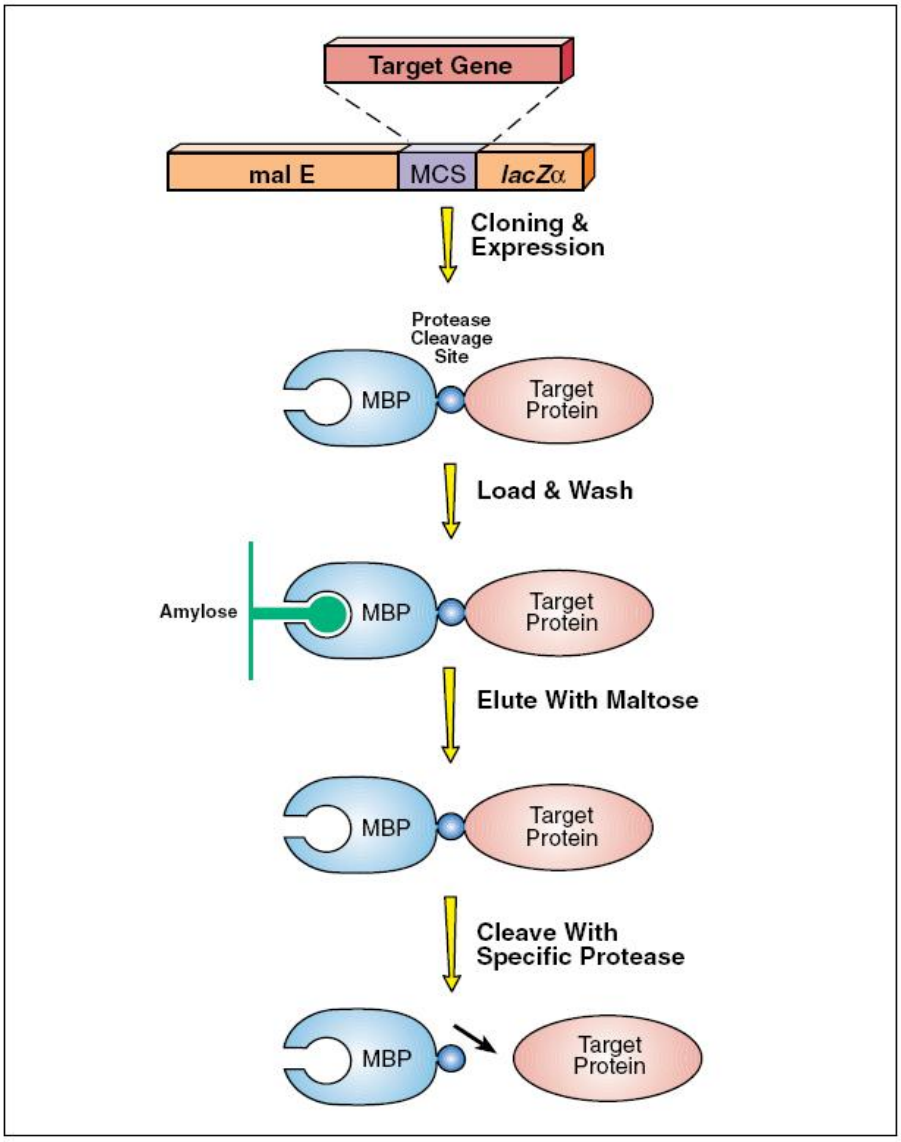


Figure 1. Typical expression and purification scheme of an MBP-fused protein.

Despite this abundance of purification tags and complementary affinity ligands that are now available, affinity chromatography is not always an efficient and economical method for purification of recombinant proteins. Affinity chromatography is not particularly well suited to the purification of recombinant proteins at the preparative scale because of the substantial costs associated with scale-up: limitations arise from the equipment and consumable resin costs associated with these procedures (Wood 2005); other

drawbacks are the possibility of ligand leakage from the resin (and this is particuarly undesirable when the protein has to be used as therapeutics); the need of proteases for the cleavage of the affinity tag together with the risk of aspecific cleavage.

These limitations of current bioseparation techniques, therefore, provide a compelling rationale for the development of new, non chromatographic methods for the purification of recombinant proteins.

In this thesis I worked on the optimization of one of these non-chromatographic methods, which is focused on the combination of two technologies: Elastin-like polypeptides (ELPs) and self-cleaving inteins.

Dramatic improvements are possible for large-scale antigen purification processes through the combination of these two systems. These methods are immediately attractive for large-scale industrial proteins, where small levels of impurities are tolerable. In some other applications where high purity is required, these methods can act as a first-capture step, delivering substantially purified material for downstream polishing. We believe that this technology offers an innovative method for protein/polypeptide purification and avoids costly chromatographic steps.

PROTEIN SPLICING (INTEIN)

Protein splicing is a process analogous to RNA splicing, in which a gene encoding intervening amino acids (intron) that are absent from the mature protein is removed from RNA prior to translation of the mature mRNA; the two flanking sequences, called exons, are joined. In the same way protein splicing is defined as a posttranslational event that involves the removal of an internal polypeptide sequence (the **INTEIN** (Perler F.B. 1994)) from a protein precursor and the concomitant ligation of the two flanking sequences (the **EXTEINS**) to produce a mature protein and the free intein. This process is schematically represented in figure 2.

The first intein was found in 1990 in *Saccharomyces cerevisiae* vacuolar ATPase (*Sce* VMA) independently by two groups (Hirata R. 1990; Kane P.M. 1990) and in less than 20 years more than 100 protein splicing elements have been discovered in organisms from all three domains of life as well as in viruses (Pietrokovski S. 1998; Perler F.B. 2000).

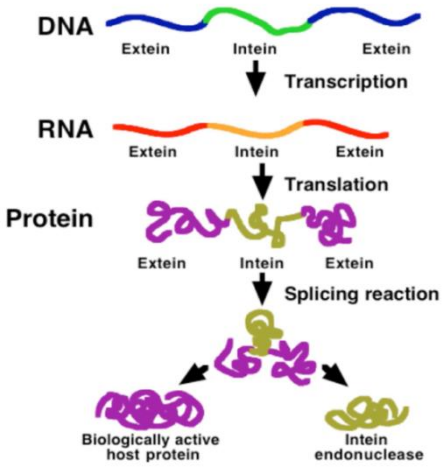


Figure 2. Natural intein function.

Inteins are remarkable single-turnover enzymes that utilize the same mechanisms as enzymes to maximize catalytic efficiency. The intein-mediated protein splicing pathway consists of a coordinated set of four nucleophilic displacement reactions resulting in peptide bond cleavage at both intein–extein splice junctions, and ligation of the flanking sequences with formation of a new peptide bond to yield a mature extein protein and an excised intein (Xu M.Q. 1996). No exogenous cofactors or energy sources are necessary (Xu M. 1993; Kawasaki M. 1997). Proximal extein sequences can be thought of as the "substrate" of the intein by a self catalyzed protein-splicing reaction. All this is accomplished by an intein splicing domain that works in concert with the first C-extein residue.

Structure

The organization of a splicing precursor is well shown in Figure 3. Inteins (red boxes) are bifunctional proteins that sometimes have homing endonuclease activity (blue box), which is essential for mobility of intein genes (Gimble F.S. 1992). The homing endonuclease domain is not essential in protein splicing activity, and it can be often excided from the amino acid sequence without loss of protein splicing activity (Hodges R.A. 1992).

The letters above the precursor represents the intein motifs, while the letters below the precursor indicates catalytic residues. Capital letters mean that the corresponding aa occurs frequently in intein sequence, while small letters indicates that the corresponding amino acid is rarely found in that position. Boxes surround nucleophilic residues.

The numbering of amino acids is by convention as follows: intein amino acids are numbered from N-terminus starting with 1. N-exteins are denoted from C-terminus with negative numbers starting with -1 at the splice junction and coming back towards the N-terminus of the precursor. C-extein is numbered from its N-terminus (the C-terminal splice junction) beginning with +1 and including a plus sign.

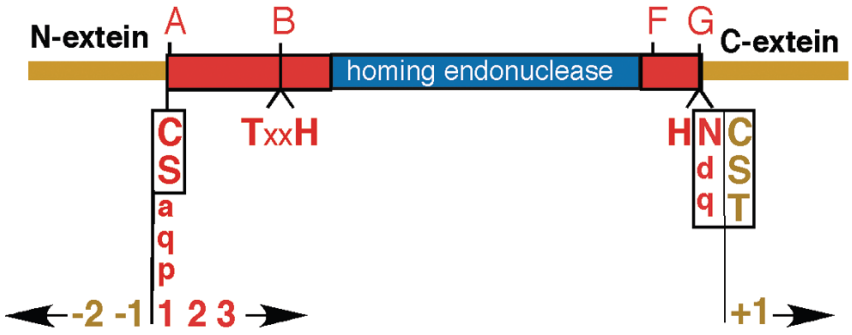


Figure 3. Organization of a splicing precursor.

Inteins have a low degree of sequence similarity but few positions are highly conserved (Pietrokovski S. 1998; Pietrokowski S. 1994; Perler F. B. 1997):

- **Position 1** of intein is often Cys or Ser. Intein with these residues splices according the mechanism shown in figure 4. Few inteins with Ala, Gln and Pro in position 1 have been discovered; these inteins splice with a slight different mechanism (Southworth M.W. 2000).
- **C-terminal** often ends in His-Asn.
- **Position +1** of C-extein is often Cys, Ser or Thr.
- The sequence **TxxH** in motif B3 of intein is highly conserved and structural data indicate that it is involved in the stabilization of the intermediate in step 1 of the protein splicing (Lee J.J. 1994).

Mechanism

Figure 4 shows the mechanism for protein splicing of most inteins; it consists of four steps (Chong S. 1996):

1 **STEP 1: N-S Acyl shift.** Attack by the side chain of the first intein residue (Cys or Ser) on the preceding carbonyl group, resulting in an acyl shift of the N-extein to the side chain of the first intein residue. Structural data have underlined that the thioester bond formation passes through a tetrahedral intermediate (an oxythiazolidine anion) with negatively charged carbonyl oxygen. This polarized intermediate is stabilized by an oxyanion hole formed mainly by conserved residues TxxH in motif B (Poland B.W., 2000).

2 **STEP 2: transesterification.** A transesterification reaction in which the -SH (or -OH) of the first C-extein residue (Cys, Ser or Thr) attacks the (thio)ester linkage, resulting in transfer of the N-extein to the C-terminal splice junction. The resulting branched intermediate, which was found to have two N-termini, was detected in 1993 (Xu M. 1993) and migrates abnormally slowly in SDS-PAGE.

3 **STEP 3: Asn cyclization.** Cleavage of the amide linkage at the intein C terminus by side chain of Asn (or Gln) cyclization to release the free intein.

4 **STEP 4: S-N Acyl shift.** Uncatalyzed and spontaneous rearrangement of the (thio)ester linkage between the ligated exteins to the more stable peptide bond.

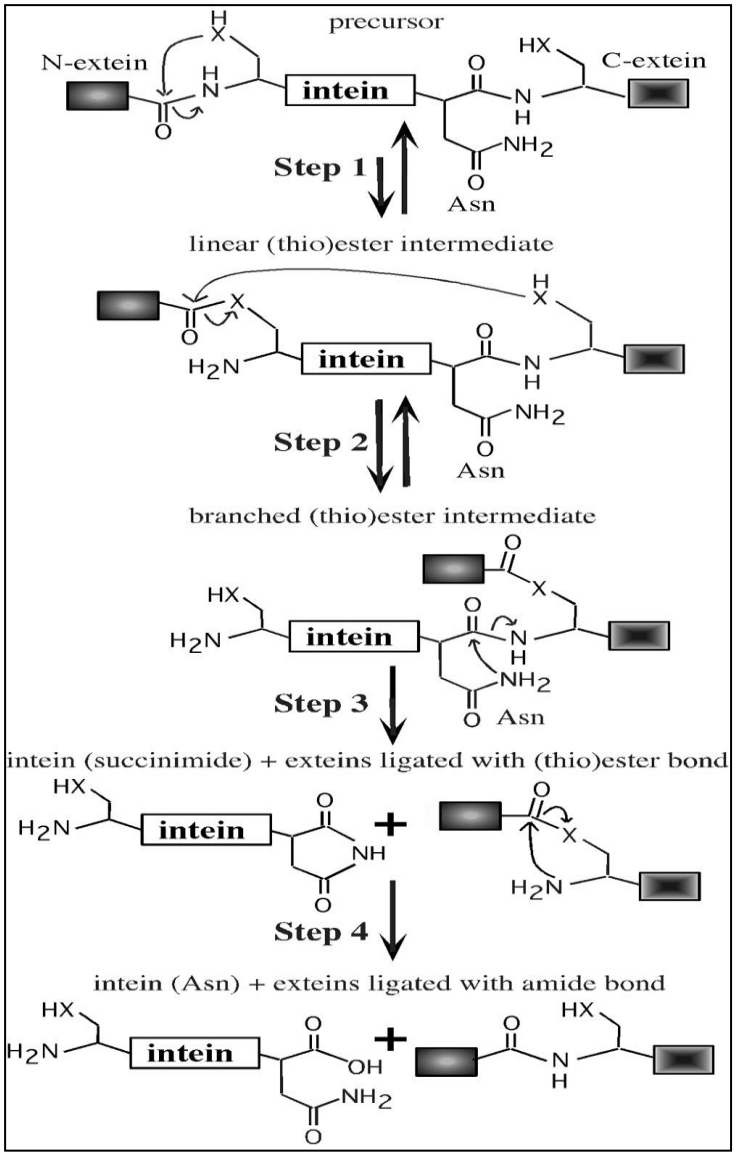


Figure 4.Mechanism of protein splicing

The identification of the residues directly participating in the breakage and formation of peptide bonds has paved the way for modulating the protein splicing reaction for various application in protein chemistry, like expressed protein ligation (Dawson P.E. 1994), *trans*-splicing (Mills K.V. 1998) and protein purification.

Specifically to the field of recombinant protein purification the following intein's features have been exploited:

Inteins share little sequence homology in their extein sequences, and the native extein can often be replaced by a foreign protein sequences without a dramatically adverse effect on the splicing or cleavage activity (Noren C.J., 2000).

1. The protein splicing activity can be inhibited at one of the splice junctions by mutagenesis of the amino acid sequence: intein maintains its ability to self cleave at only the non-mutated splice junctions shown in figure 5 (Xu M.Q. 1996).

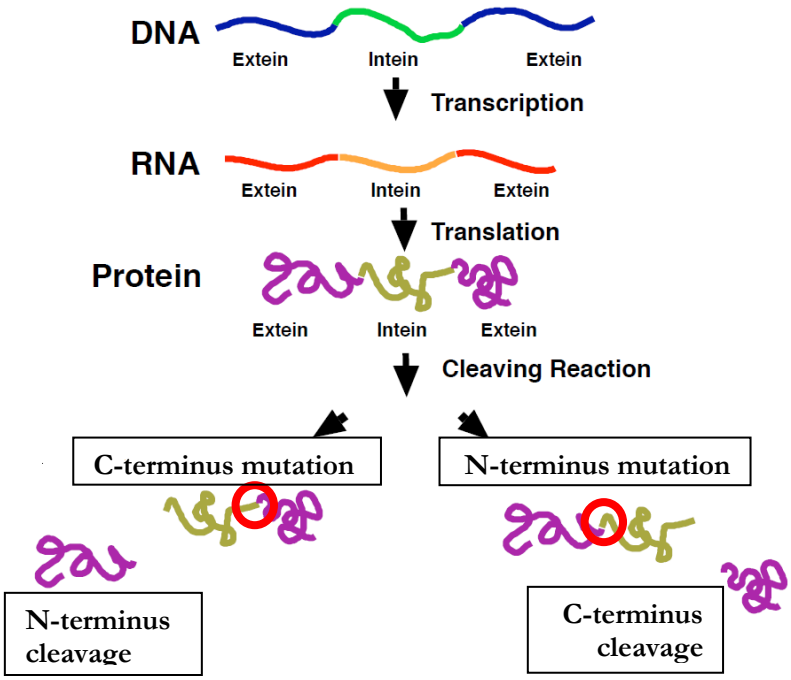


Figure 5. Different cleavage patterns for the N- or C-terminal cleavage upon modification of the splice junctions.

N-terminus cleavage occurs when C-terminal amino acid of the intein is mutated, i.e. Asn→Ala, see figure 6; substitution of the intein C-terminal residue (Asn) prevents cyclization and subsequent cleavage of the peptide bond at the C-terminal junction. N-terminus cleavage requires the addition/presence of a nucleophilic group (typically thiol group or hydroxylamine) able to perform a transesterification reaction on the -1 carbonyl residue.

This property was first employed in a commercial expression vector that used a modified *S.cerevisiae* VMA intein, in which the C-terminal asparagine residue was replaced by alanine, fused to a chitin-binding domain (CBD) as the C-extein and the target protein as the N-extein (Chong S 1997). The target protein is immobilized through the CBD on a chitin binding resin, and cleaved by addition of thiols whenever desired. Similar types of vectors were subsequently designed using the *M.xenopi*GyrA mini-intein (Evans TC 1998), the *M.thermoautotrophicum* RIR1 mini-intein (Evans T.C. 1999) or a mini-intein derived from the DnaE intein of *Synechocystis* sp. PCC6803 (Mathys S 1999). These last inteins brought some advantages, such as higher expression efficiencies on account of smaller intein size, wider tolerance to C-terminal aminoacids and in one case (*Ssp* DnaB) the cleavage can be induced just increasing the pH to 8.0 or above, avoiding then the use of thiols.

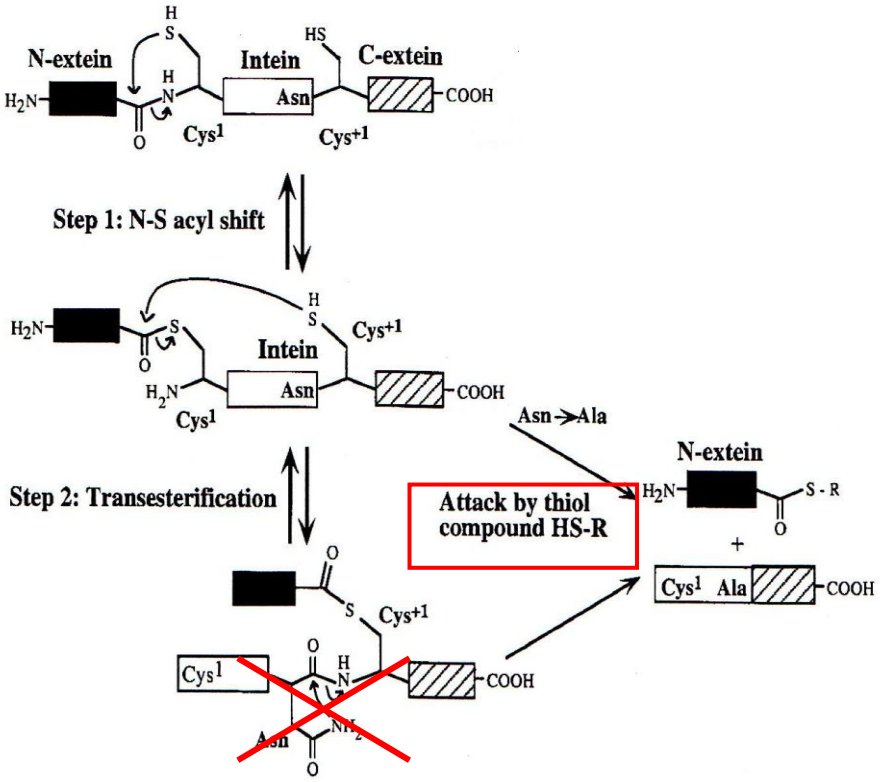


Figure 6.N-terminal cleavage.

C-terminus cleavage is obtained by mutating the amino acid in position 1 of the intein, i.e. $Cys^1 \rightarrow Ala$, see figure 7; N-S acyl shift in step 1 can no longer happen and intein can only splice at the C-terminus. C–terminal cleavage is usually induced by change in temperature and/or pH conditions.

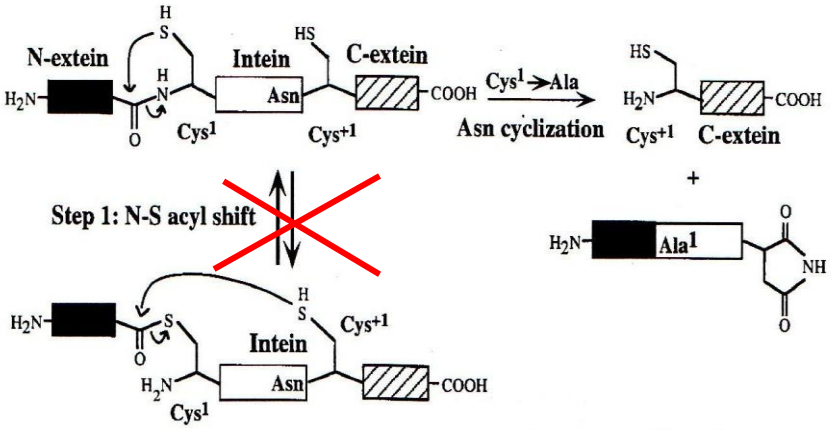


Figure 7. C-terminal cleavage.

As an example, in figure 8, a commercial available expression system from New England Biolabs is shown (IMPACT™ Kit). This kit employs an engineered intein fused to CBD (chitin binding domain). A vector comprising the gene for intein-CBD tag, together with a multi-cloning site, where the protein of interest can be inserted, is provided. The target protein can be cloned at the N-terminus of the intein (pTYB21 vector) or at the C-terminus of the intein (pTXB1 vector) as preferred by the operator. The fusion construct is expressed in *E.coli* and is made of the target protein, the intein and the CBD. When the bacterial lysate is loaded on chitin beads, only the CBD-fused protein remains bound, while the *E.coli* endogenous proteins, lacking a chitin binding domain, are washed away. When all the contaminants are washed and the fusion construct is still on the beads, the target protein is cleaved thanks to the intein, and eluted without contamination of the tag, still bound to the beads.

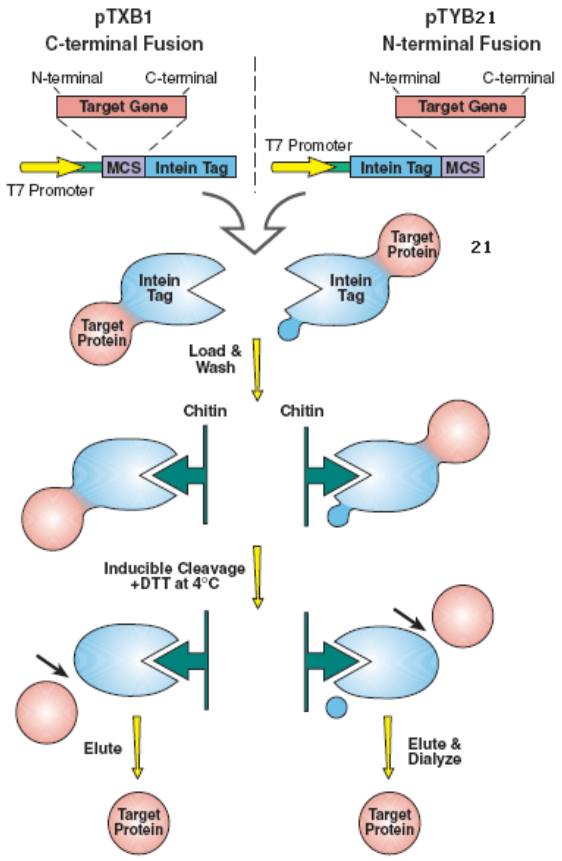


Figure 8. Schematic drawing of the IMPACT system.

After the discovery of intein mechanism, inteins found application in many protein-engineering processes, such as protein purification, polypeptide ligation (trans splicing and intein-mediated expressed protein ligation), isotopic labeling of proteins for NMR analysis, peptide binding to solid supports.

Protein purification. Inteins cleave the peptide bonds between the intein and the exteins. To be useful for purification vectors, these cleavage reactions must be isolated and controlled, and an affinity tag introduced. The control over the cleavage site is obtained by mutagenesis of some of the interested aminoacids, as seen above.

- Polypeptide ligation. Two intein-based technologies have been used to ligate polypeptides: trans-splicing and intein-mediated expressed protein ligation (EPL or IPL) (David R., 2004; Giriat, I., 2001). In trans-splicing, intein fragments of split precursors spontaneously

reassemble to yield a functional intein: some naturally split inteins have been found, and are easy to work with; alternatively artificially split inteins can be created, but they are usually less efficient (Figure 9). EPL is an in vitro technique that ligates polypeptides having an intein generated C-terminal a-thioester with polypeptides beginning with Cys. A standard peptide bond is formed after ligation.

- Isotopic labelling for NMR analysis. Protein NMR is limited by the loss of resolution with increasing sequence length. In segmental isotopic labelling, a segment (or a domain) of a larger protein is labelled with ^{13}C or ^{15}N in bacteria, and ligated by mean of EPL to the unlabeled rest of the protein. By a series of experiment different protein segments can be labelled, allowing the study of the labelled region in the context of the entire folded protein (Xu R., 1999).
- Peptide binding to solid support. Inteins have been used to synthesize protein microarrays (protein chips) by linking the protein with a solid surface binding domain. In one example, target proteins with α-thioesters were immobilized on a glass surface by reacting with Cys linked to the glass by polyethylene glycol (Camarero, J.A., 2004). In other examples, biotin was added to proteins using EPL and the proteins bound to avidin-functionalized solid supports.

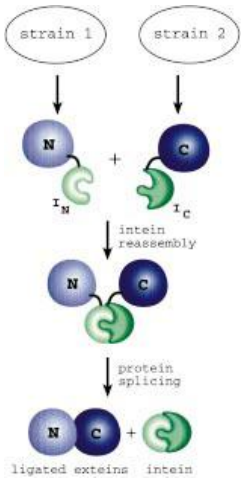


Figure 9. Trans-splicing: control of splicing by precursor fragmentation. The precursor gene can be split within the intein and each fragment expressed in separate cultures. Reassembly occurs in vitro. The intein fragments (IN.IC) reconstitute an active splicing element that then mediates ligation of the extein fragments (N, C).

One of the most popular intein among the ones employed for the purification of proteins is a 198 aminoacids in-frame insertion in the *gyrA* gene of *Mycobacterium xenopi* (called then *Mxe*GyrA. Telenti A., 1997). This intein, compared to other mycobacterial gyrA inteins, was found to have lost more than 200 aminoacids of the central intein domain, which is involved in the homing endonuclease activity, and gained a linker of unrelated residues. This linker is necessary to maintain the splicing activity, as the lacking of the endonuclease domain without replacement results in a failure to splice. Such inteins, which lost the endonuclease activity, but not the splicing capability, are named mini-inteins. The mini-intein *Mxe*GyrA has then a primary structure made of 198 amino acids and a molecular weight of 21.3 kDa (Telenti A, 1997). Its sequence is shown in figure 10, while its crystal structure in figure 11 (Klabunde T., 1998).

YCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFH
SGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVI
QRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTD
GRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHNT

Figure 10. *Mxe*GyrA mini-intein aminoacid sequence. In red the -1 and +1 exteins residues, in yellow the linker substituting the endonuclease motif.

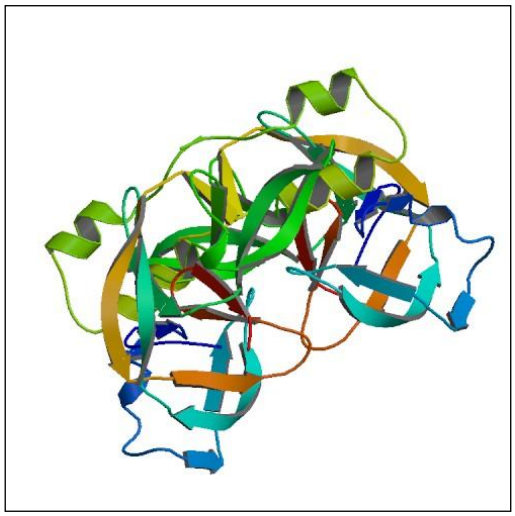


Figure 11. Crystal structure of MxeGyrA intein. (Klabunde T., 1998)

ELASTIN-LIKE POLYPEPTIDE (ELP) SYSTEM

Elastin is an extracellular matrix protein that is found predominantly in the connective tissue of arteries, skin, lung and ligament (Vrhovski B 1998). Tropoelastin, the soluble precursor of elastin, is composed of alternating hydrophobic and hydrophilic crosslinking domains. Once secreted into the extracellular space, tropoelastin is enzymatically crosslinked via its lysine residues, creating fibrils of insoluble elastin (J. S. Vrhovski B 1997). Elastin-like polypeptides are repetitive artificial polypeptides derived from recurring amino acid sequences found in the hydrophobic domain of tropoelastin: the most common motif is the sequence $(VPGXG)_m$, where X (guest residue) can be any amino acid other than proline, as its presence at the fourth position would prevent helix formation, and *m* is the number of pentapeptide repeats.

NOMENCLATURE: the current nomenclature for the $(VPGXG)_m$ type of ELP system uses the following notation: ELP[$X_iY_jZ_k$-*m*] where the bracketed capital letters are single letter amino acid codes of a guest residue, their corresponding subscripts denote the ratios of that guest residue in the monomer, and *m* is the number of pentapeptides in the ELP. For example, the most common ELP used in literature is ELP[$V_5A_2G_3$-90]; it is an ELP containing 90 repeats of the pentapeptide VPGXG, in which five pentapeptides contain V at the fourth position, two have A and three pentapeptides have G (Meyer DE 1999).

The main feature of ELPs composed of the $(VPGXG)_m$ repeated motifs consists in a reversible, inverse phase transition which takes place in response to thermal changes. ELPs are highly soluble in aqueous solutions under a certain temperature called **Transition temperature (T_t)**, but undergo a sharp phase transition (2-3°C range) when the temperature is raised above its T_t, leading to aggregation of the polypeptide and formation of an insoluble, polymer-rich "coacervate" (Urry DW 1997). The ELP aggregation is a reversible process, so the polypeptide is completely re-solubilised in buffer when the solution temperature is reduced below its transition temperature. The T_t of ELPs can be tuned at the design level by the composition of the guest residue as well as ELP chain length, as will been explained below.

The inverse phase transition can be monitored by assaying the turbidity of the ELP solution as a function of temperature by mean of a turbidimeter or even a spectrophotometer at 600 nanometer of wavelength.

Increase the temperature beyond the critical point T_t results in a sharp turbidity increase to a maximum. T_t is defined as the temperature at which the 50% of the maximum turbidity is reached. T_t is a convenient parameter to describe the inverse phase transition.

ELPs are members of a broad class of stimulus responsive polymers that exhibit inverse temperature phase transition in response to their solution environment (Rodríguez-Cabello JC 2006;Hoffman AS 2007; Ista LK 1998; Shimoboji T 2003; Ding 2001; Urry DW. 1992). Differently from other classes of polymers which transition phase is only affected by temperature, ELPs' behaviour can be isothermally triggered by cosolutes, especially kosmotropic salts (Cho Y 2008), and variants of these polymers can be designed to exhibit the same behaviour in response to changes in pH (Urry DW. 1993), redox triggers (Urry DW. 1992) and light (Shimoboji T. 2002).

T_t is affected by several factors (Urry D.W. 1985; Urry D W 1991; Meyer DE 1999):

Identity and mole fractions of the guest residues Xs in the fourth position.

ELP chain length

ELP concentration

Medium ionic strength

The guest residues X have a strong influence on the T_t given by their hydrophobicity, as quantified by Urry and co-workers (Urry D.W., 1991): the more hydrophobic the X side chain, the lower the transition temperature; vice versa the higher the polarity, the higher the T_t (see table 1; Urry D.W., 1997) This is consistent with the observation that aggregation and precipitation above the T_t is due to a hydrophobic collapse. The most common amino acid combinations reported in literature (Xin G., 2005; Lim D.W., 2007) are Val, Ala and Gly in the ratio 5:2:3, Lys, Val, Phe in ratios of 1:7:1 and 1:2:1, respectively, and Val only.

Residue	R Group	Abbreviation	Letter	T_t[a]	ΔH_t kcal/mol[d] ± 0.05	ΔS_t kcal/mol[d] ± 0.05
Tryptophan		Trp	W	−90°C	2.10	7.37
Tyrosine		Tyr	Y	-55°C	1.87	6.32
Phenylalanine		Phe	F	−30°C	1.93	6.61
Histidine		His	H	−10°C		
Proline(calc.)[b]	$-CH_2CH_2CH_2-$	Pro	P	(−8°C)	1.51	5.03
Leucine	$-CH_2CH(CH_3)_2$	Leu	L	5°C	1.43	4.60
Isoleusine	$-CH(CH_3)CH_2CH_3$	Ile	I	10°C	1.00	3.29
Methionine	$-CH_2CH_2SCH_3$	Met	M	20°C	1.20	3.90
Valine	$-CH(CH_3)_2$	Val	V	24°C		
Histidine		His^+	H^+	30°C		
Glutamic Acid	$-CH_2CH_2COOH$	Glu	E	30°C	0.96	3.14
Cysteine	$-CH_2SH$	Cys	C	30°C		
Lysine	$-CH_2CH_2CH_2CH_2NH_2$	Lys^+	K^+	35°C	0.71	2.26
Proline(exptl)[c]	$-CH_2CH_2CH_2-$	Pro	P	40°C	0.92	2.98
Alanine	$-CH_3$	Ala	A	45°C	0.85	2.64
Aspartic Acid	$-CH_2COOH$	Asp	D	45°C	0.78	2.57
Threonine	$-CH(OH)CH_3$	Thr	T	50°C	0.82	2.60
Asparagine	$-CH_2CONH_2$	Asn	N	50°C	0.71	2.29
Serine	$-CH_2OH$	Ser	S	50°C	0.59	1.86
Glycine	$-H$	Gly	G	55°C	0.70	2.25
Arginine	$-CH_2CH_2CH_2NHC(NH)NH_2$	Arg	R	60°C		
Glutamine	$-CH_2CH_2CONH_2$	Gln	Q	60°C	0.55	1.76
Lysine	$-CH_2CH_2CH_2CH_2NH_3^+$	Lys	K	120°C		
Tyrosinate		Tyr^-	Y^-	120°C	0.31	0.94
Asporate	$-CH_2COO^-$	Asp^-	D^+	120°C		
Glutamate	$-CH_2CH_2COO^-$	Glu^-	E^-	250°C		

Table 1. The effects of guest residue composition on the Tt of ELP homopolymers.

b. ELP length has an inverse proportional influence on the T_t (Figure 12): as ELP becomes longer its transition temperature decreases. To date the smaller ELP employed in a protein purification process is composed of 40 aa (MW: 4.3 kDa) and its T_t (comprising a fused Trx) overtakes 90 °C (Lim DW 2007). To be useful for protein purification application, fusion of elastin-like polypeptides with a target protein should have a transition temperature slightly higher than 37 °C, in order to remain soluble in *E.coli* cultures, but aggregate in response to small increase in temperature; since higher temperatures might affect protein stability.

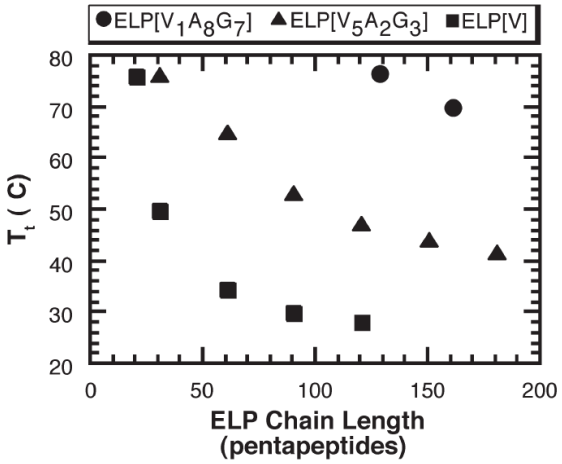


Figure 12.T_t as a function of ELP length of three different polypeptides (from Urry D.W., 1985).

c. ELP concentration affects T_t by a logarithmical relation (Figure 13), then it is of crucial importance to keep the polypeptide concentration higher than ≈100 μM, otherwise the transition temperature becomes too high to be useful in a protein purification process.

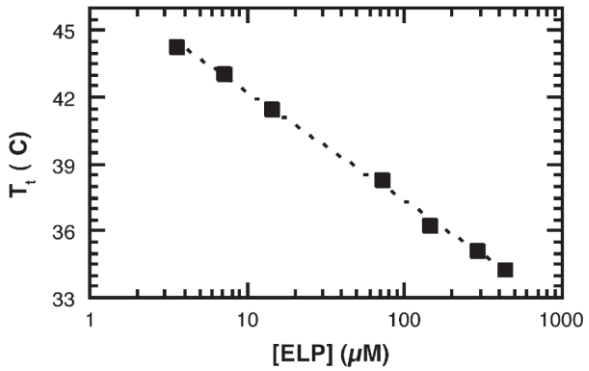


Figure 13. T_t as a function of ELP[$V_5A_2G_3$-180] concentration (from Urry D.W., 1985).

d. The buffer ionic strength is crucial in the tuning of T_t. The control of this parameter allows T_t to be tuned over a linear 50 °C range, and then provides a means to isothermally trigger the phase transition (Figure 14). This is due to the mechanism of hydrophobic collapse of the ELPs: enhancing the medium dielectric constant, the hydrophobic collapse will be favoured and less energy will occur to make it happen.

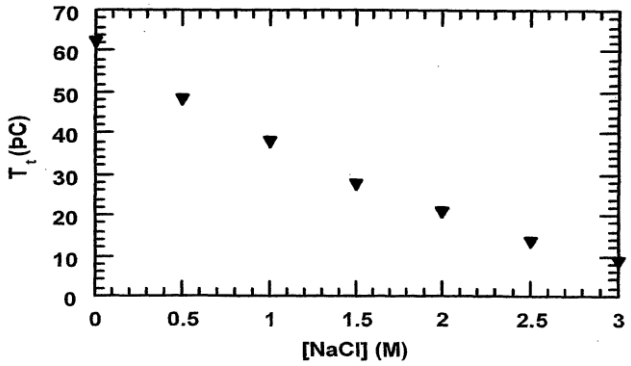


Figure 14.T_t plotted as a function of NaCl concentration for 25 uM Trx-ELP[$V_5A_2G_3$-60] (from Meyer DE, 1999).

The inverse transition behaviour is maintained even upon fusion of an ELP sequence in protein (Meyer DE 1999). Figure 15 shows the turbidity profiles of some ELP[$V_5A_2G_3$] $_{90}$ and fusion constructs of different proteins with the same ELP; the overall behaviour does not change, while the differences in transition temperature are noticeable. These differences are due to differences in the solvent-accessible surface area of the fusion protein. As hydrophobicity is crucial in the collapse which causes the phase transition,

the more hydrophobic the fusion protein is the lower T_t will be. A study of Trabbic-Carlson *et al* demonstrated that the parameter to be considered is the fraction of hydrophobic surface area, rather than the total protein hydrophobicity. They considered the hydrophobic fraction of the surface area of 500 proteins, and plotted it versus the transition temperature of their ELP-fusion (figure 16): the T_t varies in a range of almost 20 °C, showing a crucial influence of the fusion protein on the phase behaviour of the complex.

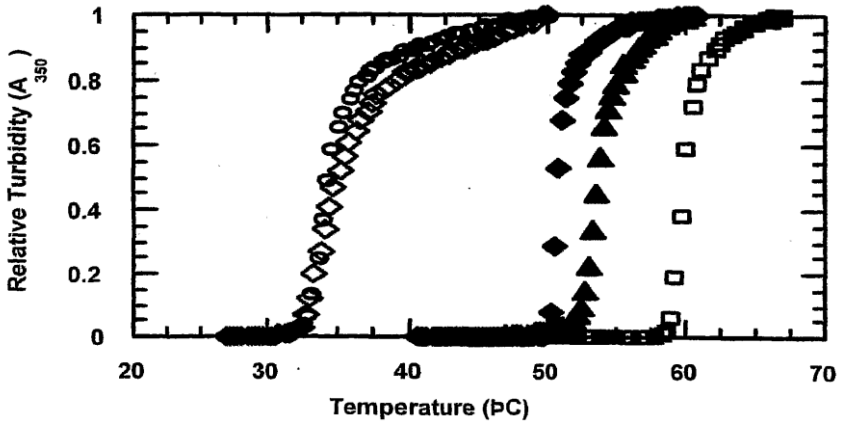


Figure 15. Turbidity profiles of some fusion protein-ELP[$V_5A_2G_3$-90] complexes and free ELP[$V_5A_2G_3$-90] profiles. Solid rumble: free ELP; solid triangle: thioredoxin-ELP; hollow circle: thioredoxin-ELP-tendamistat; hollow rumble: ELP-tendamistat; hollow square: thioredoxin-ELP (cleaved from thioredoxin-ELP-tendamistat) (from Meyer DE, 1999).

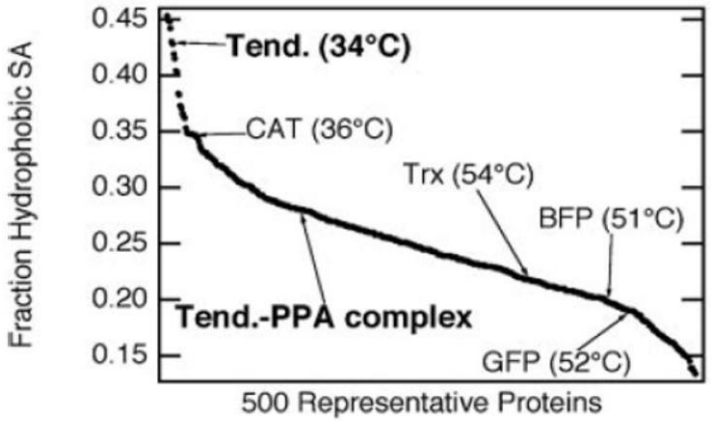


Figure 16. T_t trend vs. hydrophobic fraction of surface area of some example proteins. (From Trabbic-Carlson K., 2004)

Inverse transition Cycling (ITC)

Inverse Transition Cycling (ITC) is a method developed at the end of the last century in the Chilkoti's laboratory at the Duke University (North Carolina) for the purification of recombinant proteins (Meyer DE 1999). This method has the main advantage of eliminating the need for column chromatography. Specifically, it exploits the fact that the ELP phase transition behaviour is retained when the polypeptides are fused to soluble protein.

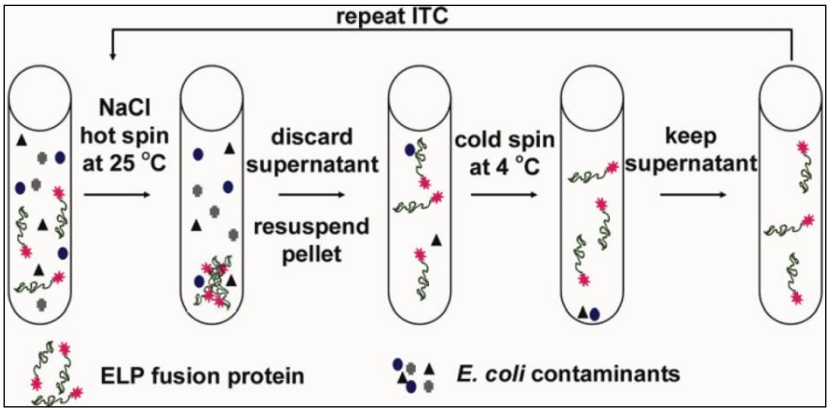


Figure 17.Scheme of a typical ITC protein purification.

A schematic ITC purification is shown in Figure 17. The phase transition of an ELP fusion protein in its soluble cell lysate is isothermally triggered by salt (i.e. NaCl) addition to the lysate at a concentration that depresses the T_t of the ELP fusion protein below the solution temperature (typically ambient). When the expression level of an ELP fusion protein is high enough, the aggregates of the ELP fusion protein that are formed cause the cell lysate to visibly turn turbid, thereby providing a convenient visual confirmation that the phase transition has occurred in the cell lysate. The turbid suspension is then centrifuged at room temperature (termed ‘‘hot spin’’, HS): the pellet containing the fusion protein is retained while the supernatant is discarded. The pellet is then re-dissolved in low ionic strength buffer at a temperature below the T_t of the fusion protein. At this point, a centrifugation step at 4 °C (termed ‘‘cold spin’’, CS) is carried out to remove insoluble contaminants that may have been physically trapped in the ELP fusion protein pellet during centrifugation.

In the cold spin step, whereas insoluble contaminants precipitate and are separated in the pellet fraction, the ELP-fusion protein is soluble and remains in solution. The supernatant from the cold spin is retained, and the pellet is discarded. The whole process, salt-induced precipitation of ELP followed by a hot spin, re-suspension in a low ionic buffer followed by a cold spin, constitutes one round of ITC. Several rounds of ITC can be performed, but only few (generally two or three) are sufficient to obtain the ELP fusion protein with a high yield and purity.

Protein purification using an ELP tag has several advantages over conventional chromatography:

The ELP tag that is co-expressed as a fusion partner with the target protein acts as a capture mechanism, so no chromatography beads or media are needed. This reduces the expenses associated with protein purification.

The ELP based purification method does not require a concentration step to recover the final product. Because the recombinant protein is precipitated and concentrated during each step of the purification process, loss of the final product is lower than in chromatography.

Purification of proteins with ELP tags by ITC appears to be universal for soluble recombinant proteins. Notwithstanding, the ELP should be chosen taking into account the characteristics of the fusion partner and the desired transition temperature.

One main drawback of this method is shared with the affinity tags already described, and is that ELP tag must be cleaved from the fusion protein if it is not necessary or even worse if it interferes with the biological activity of the protein of interest. Before the application of inteins, a solution to this problem was the insertion between the fusion protein and the ELP genes of a DNA sequence encoding for an amino acid sequence recognized by a protease. After the ITC process, protease addition let the cleavage of the target protein from the ELP. Three problems arise with the use of a protease:

Protease adds additional cost to the purification

An additional step to the purification process is needed, as the protease has to be separated from the target protein.

Due to their limited specificity proteases could act on the target protein.

AIM OF THE PROJECT

The purification of recombinant proteins is a crucial process in the biopharmaceutical industry. In the last decades, the introduction of affinity tags allowed recovery of the target protein with good purity and yield. Some crucial disadvantages are though brought by these tools: limitations for high-throughput protocols; ligand leakage from the chromatographic resin; use of proteases to cleave the affinity tag increases a lot the unitary cost and the purification time of the target recombinant protein, with aspecific cleavage risk; scalability from bench to industrial scale is expensive and not simple; affinity resins has usually short life cycles.

A new method which could represent a viable alternative is then needed. One method, that in principle would not require any chromatographic step, was developed in 1999 based on the combination of two technological tools, a phisico-chemical (ELP aggregation ad solvation) and a biochemical one (intein auto-cleavage activity). ELP act as an affinity tag which has no need to be immobilized, then it can circumvent all the problems (and costs) given by an affinity resin; insertion of an intein at the right position in the fusion sequence gives the opportunity to avoid the use proteases with all its drawbacks (i.e. cost, aspecificity). In literature some example of this kind of purification are present, but no optimization work has been found up to our grade of knowledge on a specific protein.

The work described in this thesis developed the purification of three model proteins employing two different kinds of ELPs and one intein. The three model proteins were: thioredoxin (Trx), p24M and a chimerical protein derived from the major outer membrane protein (MOMP) of *Chlamydia trachomatis* (Ct_mono). The ELPs we tested were ELP[$V_5A_2G_3$]90 (ELP90)and ELP[KV_7F]36 (ELP36), while the only intein tested was the previously described DNA gyrase subunit A from *Mycobacterium xenopi* (*Mxe*GyrA).

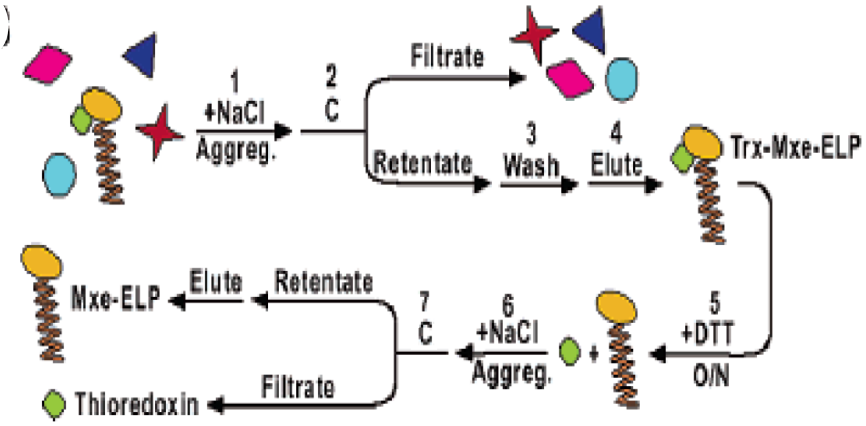


Figure 18.Purification path of the fusion protein.

MATERIALS AND METHODS

REAGENTS

Tryptone, peptone were purchased from BD, IPTG was purchased from Inalco, NaCl, Urea were purchased by Carlo Erba Reagenti, Sodium phosphate, potassium phosphate, Trizma base, DTT, TFA, Ampicillin, Kanamicin, and CH_3CN (HPLC grade) were purchased by Sigma-Aldrich;.

STRAINS

E. coli XL1-Blue strain (Stratagene, Agilent Technologies, Santa Clara, USA) was used for mutagenesis and cloning operations and plasmid maintenance.
E. coli BL21(DE3) strain (Novagen, Merck Chemicals Ltd) was used for protein expression.

CLONING OF EXPRESSION VECTORS

A) pTME-MCS expression vector

pTME vector is a modification of pET32 expression vector (Novagen, Merck Chemicals Ltd) and includes the gene encoding for the Trx-*Mxe*GyrA-ELP90 fusion protein; it was a gentle gift from Prof. Chilkoti (Duke University, Durham, NC, USA).

The expressed protein had the following amino acids sequence, figure 19:

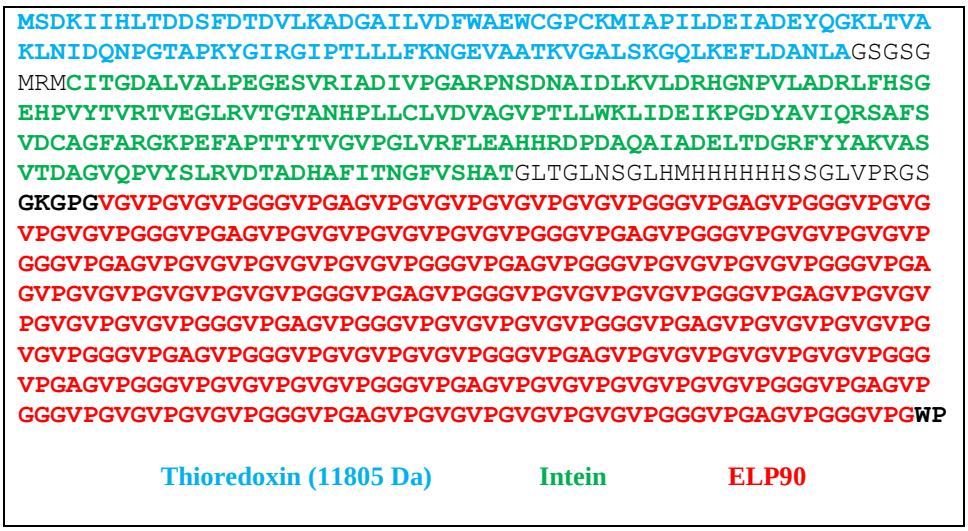


Figure 19. Trx-*Mxe*GyrA-ELP90 fusion protein sequence (799 amino acids; 72.6 kDa) encoded by the pTME vector

The main drawback of pTME vector was that the Trx gene could not be replaced by other DNA sequences because no endonuclease restriction sites were anymore available at both the 5' and 3' terminus of the Trx sequence. For this reason it was modified by replacing the Trx gene with a multiple cloning site according to the following protocols, figure 20:

1. pTME plasmid was digested with *Nde*I, in order to excise the Trx-Intein DNA sequence.

2. A multiple cloning site–Intein fragment was generated by PCR amplification using as template the Intein sequence of pTME vector. This fragment is characterized by sticky ends to *Nde*I.

   The following primers were used:

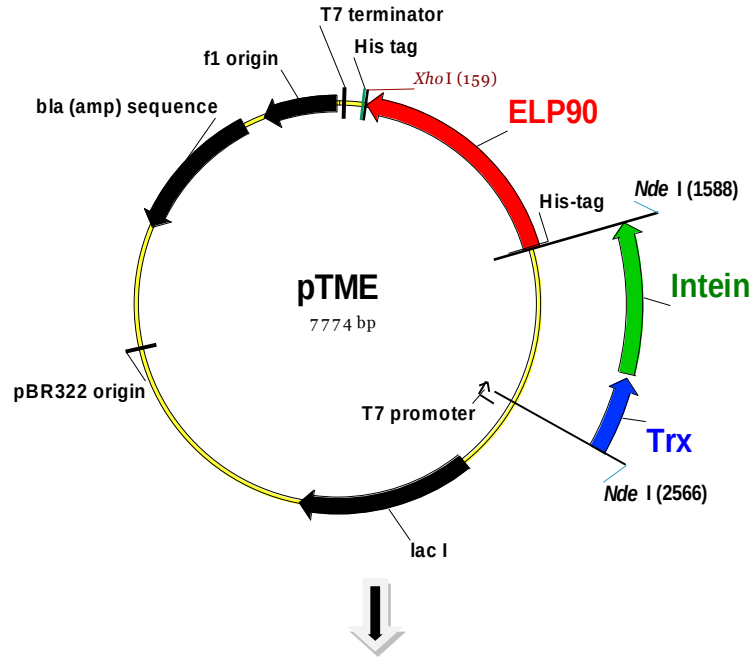
   Fw:

   NdeI *EcoRI* **BamHI** *HindIII*
   5'_AG**CATATG***GAATTC***GGATCC***AAGCTT*AGATCTGG
   TTCTGGTTCTGGCATGCGTATG_3'

Rev:

NdeI
5’_GC**CATATG**AGATCTGAGGCCTGAGTTCAGACCG
GTGAG_3’

3. Ligation of the open vector from step1 with the PCR product of step 2 digested with *Nde*I gave the new expression vector pTME-MCS.



Trx-Intein sequence was excised by *Nde*I and replaced by a MCS-Intein sequence obtained by PCR

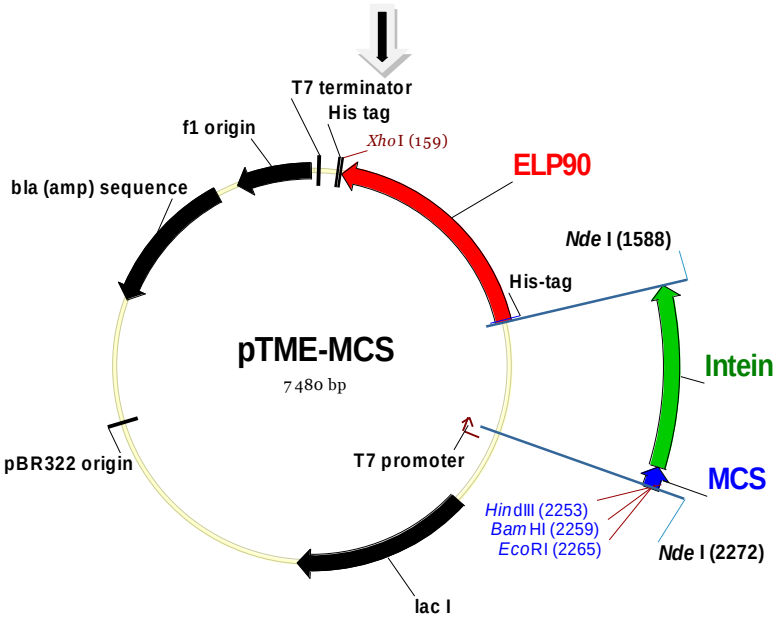


Figure 20. A schematic representation of the cloning strategy of pTME-MCS vector

pTME-MCS vector is characterized by the following multiple cloning site, figure 21:

5’	atg	gaa	ttc	gga	tcc	aag	ctt	aga	tct	3’
	M	E	F	G	S	K	L	R	S	
	START	EcoRI*		BamHI*		HindIII*		BglII**		

Figure 21: Nucleotide sequence of the multiple cloning site of pTME-MCS vector. (*) Unique restriction sites usable for cloning. (**) The BglII restriction site is present twice in the pTME-MCS vector, before and after the Intein sequence and cannot be used for cloning but for eventual Intein deletion.

The amino acid sequence of the MCS-*Mxe* GyrA-ELP90 protein encoded by pTME-MCS vector resulted then as shown in Figure 22

MEFGSKLRSGSGSGMRMCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRH
GNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKP
GDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELT
DGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLRSHMHHH
HHHSSGLVPRGSGKGPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGA
GVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVP
GVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVG
VPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPG
AGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGV
PGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGV
GVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVP
GAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGG
VPGWP

MCS Intein ELP90

Figure 22. Protein sequence encoded by the pTME-MCS vector

B) pTME-T3C-MCS vector

Site-directed mutagenesis was carried out using the QuickChange™ mutagenesis kit (Stratagene, CA) to prepare the Thr3 → Cys mutant in the intein protein.

pTME-MCS was used as template to obtain pTME-T3C-MCS, whereas pTME was used to obtain pTME-T3C vector (for the expression of Trx-*Mxe*GyrA(T3C)-ELP90).

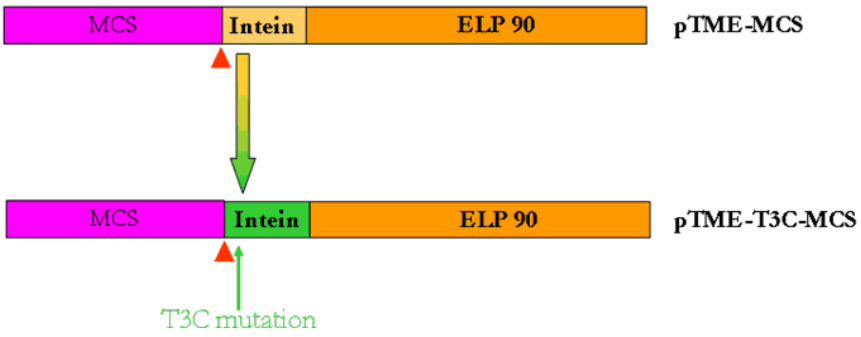


Figure 23.Flowchart of the pTME vector mutagenesis

The following primers were used:

pTME_T3C_FW
5'_TGGCATGCGTATGTGCATCTGCGGAGATGCACTAGTTGCCCTAC-_3'

pTME_T3C_REV
5'_GTAGGGCAACTAGTGCATCTCCGCAGATGCACATACGCATGCCA-_3'

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTV
AKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSG
SGMRMCICGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLF
HSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQR
SAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRFYY
AKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLHMHHHHHHSSG
LVPRGSGKGPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVP
GGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVP
GVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVP
GVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVP
GGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVP
GAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVP
GVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVP
GVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVP
GVGVPGGGVPGAGVPGGGVPGWP

Thioredoxin (11805 Da) Intein (underlined: mutated T) ELP90

Figure 24. Trx-*Mxe*GyrA(T3C)-ELP90 sequence

C) pET24-MCS-Intein-ELP(KV7F-36) vector

A synthetic gene encoding for a new ELP linked to the *Mxe*GyrA gene and a multiple cloning site was purchased by GeneArt (Regensburg, Germany) and cloned via *Nde*I-*Xho*I in pET24a vector (Novagen, Merck Chemicals Ltd).

The amino acid sequence of the MCS-*Mxe*GyrA- ELP[$KV_7F]_{36}$ protein encoded by pET24-MCS-Intein-ELP[$KV_7F]_{36}$ vector resulted as shown in Figure 25:

MEFTSKLASGSGMRMCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRH
GNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEI
KPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIA
DELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGGGSGGGSGG
GSGGGSGGGSGGGSGGGSGGGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPG
VGVPGFGVPGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPGV
GVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPGVGVPGKGVPGVG
VPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPG

MCS Intein ELP36

Figure 25. protein encoded by the pET24-MCS-Intein-ELP]$KV_7F]_{36}$

D) Cloning of p24M

A synthetic gene encoding for HIV-1 capsid protein p24 (group M, from here on "p24M") was purchased by GeneArt and cloned via *EcoR*I-*Hind*III in:

- pTME-MCS to obtain pTME-p24M vector
- pTME-T3C-MCS to obtain pTME-T3C-p24M vector
- pET24-MCS-Intein-ELP(KV7F-36) to obtain pET24-p24M-Intein-ELP(KV7F-36) vector

p24 sequence differs at the N-terminus from the authentic sequence in that the residues Pro-Ile- are replaced by Met-Asn-Ser-Ala-Met- as in Hausdorf et *al.* 1994

pTME-p24M vector encodes for p24M-*Mxe*GyrA-ELP90 fusion protein (933 amino acids; 87.8 kDa).

MEFMNSAMVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQ
DLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAG
TTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFR
DYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC
QGVGGPGHKARVLKLRSGSGSGMRMCITGDALVALPEGESVRIADIVPGARPNSDN
AIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVP
TLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAH
HRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHAT
GLTGLNSGLRSHMHHHHHHSSGLVPRGSGKGPGVGVPGVGVPGGGVPGAGVPGVGV
PGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVP
GVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPG
GGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGA
GVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGG
VPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGV
PGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVGVPGVP
GGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPG
AGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGWP

p24M (25902 Da) Intein ELP90

Figure 26. p24M-*Mxe*GyrA-ELP90 sequence

pTME-T3C-p24M vector encodes for p24M-*Mxe*GyrA(T3C)-ELP90 fusion protein (933 amino acids; 87.8 kDa).

MEFMNSAMVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQ
DLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAG
TTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFR
DYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC
QGVGGPGHKARVLKLRSGSGSGMRMCICGDALVALPEGESVRIADIVPGARPNSDN
AIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVP
TLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAH
HRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHAT
GLTGLNSGLRSHMHHHHHHSSGLVPRGSGK**GPG**VGVPGVGVPGGGVPGAGVPGVGV
PGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVP
GVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPG
GGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGA
GVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGG
VPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGV
PGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVP
GGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPG
AGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPG**WP**

p24M (25902 Da) **Intein** **ELP90**

Figure 27. p24M-*Mxe*GyrA(T3C)-ELP90 sequence

pET24-p24M-Intein-ELP[KV$_7$F-36] vector encodes for p24M-*Mxe*GyrA-ELP36 fusion protein (657 amino acids; 65.7 kDa).

MEFMNSAMVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQ
DLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAG
TTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFR
DYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC
QGVGGPGHKARVLKLASGSGMRMCITGDALVALPEGESVRIADIVPGARPNSDNAI
DLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTL
LWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHR
DPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGG
GSGGGSGGGSGGGSGGGSGGGSGGGSGG**G**VGVPGKGVPGVGVPGVGVPGVGVPGVG
VPGVGVPGVGVPGFGVPGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV
PGFGVPGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPGVGVP
GKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPG

p24M (25902 Da) **Intein** **ELP36**

Figure 28. p24M-*Mxe*GyrA-ELP36 sequence

E) Cloning of *Chlamydia trachomatis* multiple antigen

A synthetic gene encoding for a *Chlamydia trachomatis* major outer-membrane protein (MOMP) polyantigen was purchased by GeneArt and cloned via *EcoR*I-HindIII in:

- pTME-MCS to obtain pTME-chlamydia_T_mono vector
- pTME-T3C-MCS to obtain pTME-T3C-chlamydia_T_mono vector
- pET24-MCS-Intein-ELP(KV7F-36) to obtain pET24-Chlamydia_T_mono-Intein-ELP(KV7F-36) vector

The polyantigen is constituted by four immunogenic peptides derived from *C. trachomatis* MOMP linked by a short flexible spacer (Gly-Gly-Asp/Glu-Lys-Gly-Gly).

pTME-chlamydia_T_mono vector encodes for Ct_mono-*Mxe*GyrA-ELP90 fusion protein (806 amino acids and 71.9 kDa of molecular weight).

MEF**GGDKGGLAEAILDVTTLNPTITGKAVVSKGGEKGGCVFDVTTLNPTIAGAGDV**
KGGDKGGIFDTTLNPTIAGAGDVKGGDKGGVDITTLNPTIAGCGSVAKGGDKGGKL
RSGSGSGMRM**CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVL**
ADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDY
AVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTD
GRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLRSHMHH
HHHHSSGLVPRGSGK**GPG**VGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGV
PGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVP
GGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPG
VGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGV
GVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGG
VPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGV
PGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVP
GVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPG
VGVPGGGVPGAGVPGGGVPG**WP**

Ct_mono (10085 Da) **Intein** **ELP90**

Figure 29. Ct_mono-*Mxe*GyrA-ELP90 sequence.

pTME-T3C-chlamydia_T_mono vector encodes for Ct_mono-*Mxe*GyrA(T3C)-ELP90 fusion protein (806 amino acids; 71.9 kDa).

MEFGGDKGGLAEAILDVTTLNPTITGKAVVSKGGEKGGCVFDVTTLNPTIAGAGDV
KGGDKGGIFDTTLNPTIAGAGDVKGGDKGGVDITTLNPTIAGCGSVAKGGDKGGKL
RSGSGSGMRMCICGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVL
ADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDY
AVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTD
GRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLRSHMHH
HHHHSSGLVPRGSGKGPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGGGV
PGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVP
GGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPG
VGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGV
GVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGG
VPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGV
PGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVP
GVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPG
VGVPGGGVPGAGVPGGGVPGWP

Ct_mono (10085 Da) Intein ELP90

Figure 30. Ct_mono-*Mxe*GyrA(T3C)-ELP90 sequence.

pET24-Chlamydia_T_mono-Intein-ELP(KV7F-36) vector encodes for Ct_mono-*Mxe*GyrA-ELP36 fusion protein (530 amino acids; 49.9 kDa).

MEFGGDKGGLAEAILDVTTLNPTITGKAVVSKGGEKGGCVFDVTTLNPTIAGAGDV
KGGDKGGIFDTTLNPTIAGAGDVKGGDKGGVDITTLNPTIAGCGSVAKGGDKGGKL
ASGSGMRMCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLAD
RLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAV
IQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGR
FYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGGGSGGGSGGGSGGGSG
GGSGGGSGGGSGGGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFG
VPGVGVPGKGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPGVGVPGKGV
PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGFGVPGVGVPGKGVPGVGVPGVGVP
GVGVPGVGVPGVGVPGVGVPGFGVPG

Ct_mono (10085 Da) Intein ELP36

Figure 31. Ct_mono-*Mxe*GyrA-ELP90 sequence.

FERMENTATIONS

For the expression a pre-inocule was performed for all the samples: a few µL of glycerinated BL21(DE3) cells were poured in 50 mL of broth (LB or TB, depending on the protocol) supplemented with antibiotics (ampicillin 50µg/ml or kanamycin 30µg/ml, depending on the vector), and left on oscillating device (180 rpm) at 37°C overnight. The following day this sample was diluted (1:50 or 1:35 depending on the conditions) for fermentation.
Bacterial cells, grown in LB or TB with antibiotics (ampicillin 50µg/ml or kanamycin 30µg/ml, depending on the vector), were induced with 0.1 mM or 1mM IPTG when the absorbance at 600 nm reached a value of 0.6.

PROTEIN INDUCTION TEST PROTOCOL

Aliquots of a fermentation culture corresponding to 0,6 optical density at 600 nm of wavelength per ml of cellular colture were withdrawn at the moment of protein induction (IPTG induction), after fixed times from protein induction, i.e. 1 hour, 2 hours, .and at the end of the fermentation. Samples were spinned-down at 10,000xg in a bench top centrifuge and the related bacteria pellets treated with 100 µL of reducing loading buffer at +99ºC for five minutes and the analyzed by SDS-Page.

SOLUBILITY TEST PROTOCOL

At the end of fermentation, 15 milliliters of the cell colture were centrifuge and the corresponding pellet suspended in one milliliter of PBS. Then, the following working scheme is carried on, table 2:

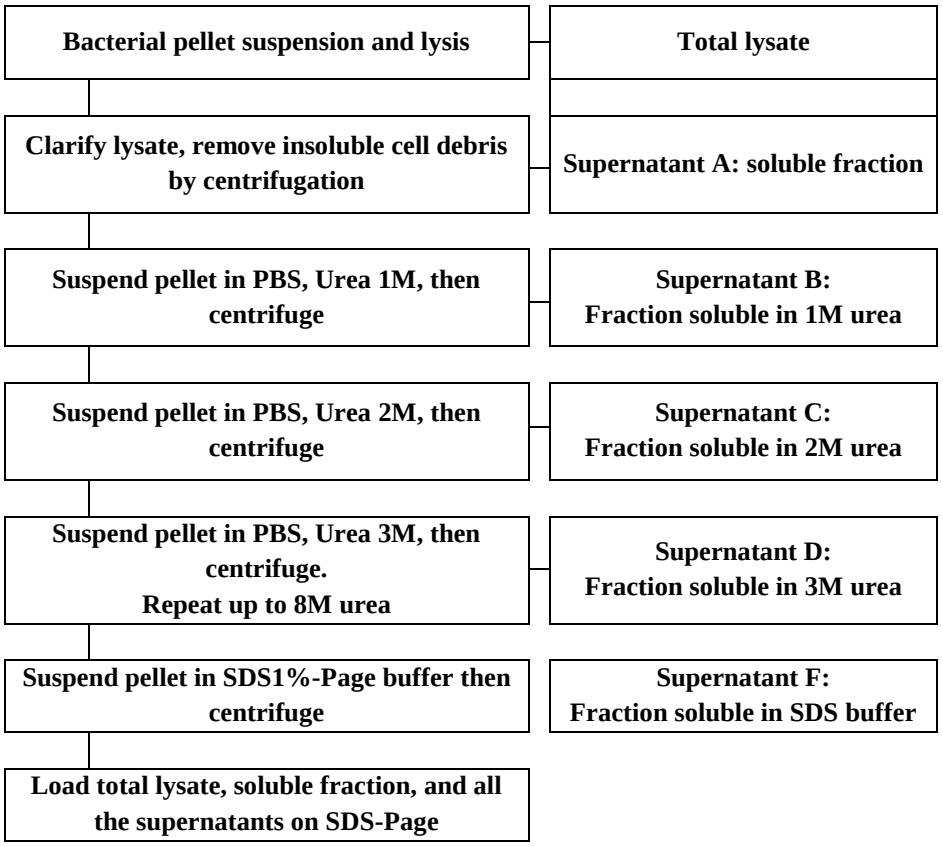


Table 2. Solubility test protocol

SDS-PAGE PROTOCOL

Electrophoresis on polyacrylamide gels was performed using the Mini-PROTEAN® II electrophoresis cell apparatus from BIO-RAD. Separation gels were home casted following the recipe:

Compound	10%	12.5%	Stacking gel
Acrylamide/Bis 19:1	2.0 mL	2.5 mL	0.75 mL
MilliQ water	2.6 mL	2.1mL	3.71 mL
Tris 1M pH 8.7	3 mL	3 mL	-
Tris pH 6.8	-	-	0.625 mL
SDS 20%	40 µL	400 µL	25 µL
APS 10%	70 µL	70 µL	50 µL
TEMED	8 µL	8 µL	6 µL

Running buffer composition:

Compound	Concentration
Tris base	25 mM
Glycine	192 mm
SDS	0.1 (w/v)

Reducing loading buffer:

Compound	Concentration
Tris base	125mM (pH 6.8)
Glycerol	22 % (v/v)
β-OH	2.5 % (v/v)
SDS	2.5 % (w/v)
Bromophenol Blue	0.0025 (w/v)
Pyronin	0.0025 (w/v)

Electrophoresis condition: 80-120 Volts for 1.5-2 hours using the power supplier Power-Pac 3000 from BIO-RAD.

Staining solution: Coomassie blue G250 3 g/Liter, Methanol 50 % (v/v), Acetic acid 10 (v/v).

Destaining solution: Methanol 20 % (v/v), Acetic acid 10 (v/v).

WESTERN BLOT PROTOCOL

After electrophoresis separation on a polyacrylamide gel, protein bands were transfer to a PVDF membrane (Immobilon-P, MILLIPORE) using the Mini Trans-Blot Electrophoretic Transfer Cell from BIO-RAD for 1 hour at 100 Volts of voltage.
Non saturated membrane sites were blocked with the *Diluent buffer*: PBS, Proclin 0.03%, Casein 0.5%; pH 7.4 over night at 4 °C.
Primary antibody incubation: Monoclonal antibody anti p24 was properly diluted in the *Diluent buffer*. Incubation for 1.5 h at room temperature.
Secondary antibody (HRP-conjugated): Anti-mouse HRP conjugated (GE Healthcare) was diluted 1 to 300 in the *Diluent buffer*. Incubation for 1 h at room temperature.
Development: PBS, 4-chloronaphtol, H_2O_2.

ANALYTICAL HPLC

Chromatographic analyses were performed with Beckman-Coulter System GOLD® 126 Solvent module system coupled to System GOLD® 168 Detector and 32Karat™ software. The column used for analysis was Jupiter 5u C5 300A 4.6x150mm (Phenomenex).

UPLC-MS

UPLC-MS analyses were performed with Waters Acquity® system and MassLynx 4.1 version (Waters); column: Acquity UPLC® BEH300 C4 1.7um 2.1x100 mm.

MALDI-TOF

MALDI-TOF and peptide mass fingerprinting characterizations were carried out by mean of an Ultraflex TOF/TOF Brucker Maldi

INVERSE TRANSITION CURVE DETERMINATION

The optical absorbance at 600 nm of *Mxe*GyrA-ELP fusion, p24M-*Mxe*GyrA-ELP fusion solutions (final protein concentrations: 1 µM, 10 µM,

50 µM and 100 µM; in presence of 0 M, 0.5 M, 1.0 M, 1.5 M, 2.0 M and 2.5 M NaCl) were monitored in the 25–45 °C range (1 °C interval) on an ELISA reader (Tecan) equipped with a multicell thermoelectric temperature controller.

PROTEIN PURIFICATION

Chromatographic steps were carried out using an Akta™ Explorer 100 HPLC system and the Unicorn™ software from GE; the following columns were used:

- Gel filtration chromatography: HiLoad xk 16/60 Superdex 75 prep grade (GE) for the purification of p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90.
- Anion exchange chromatography: Q Sepharose HP xk 16/75 (GE) for the purification of p24M-*Mxe*GyrA-ELP36.
- Reversed phase chromatography: DeltaPak C4 15 um 300A 7.8x300mm (Waters) for the purification of Ct_mono-*Mxe*GyrA-ELP90, Ct_mono-*Mxe*GyrA(T3C)-ELP90 and Ct_mono-*Mxe*GyrA-ELP36.

Thioredoxin purification protocol

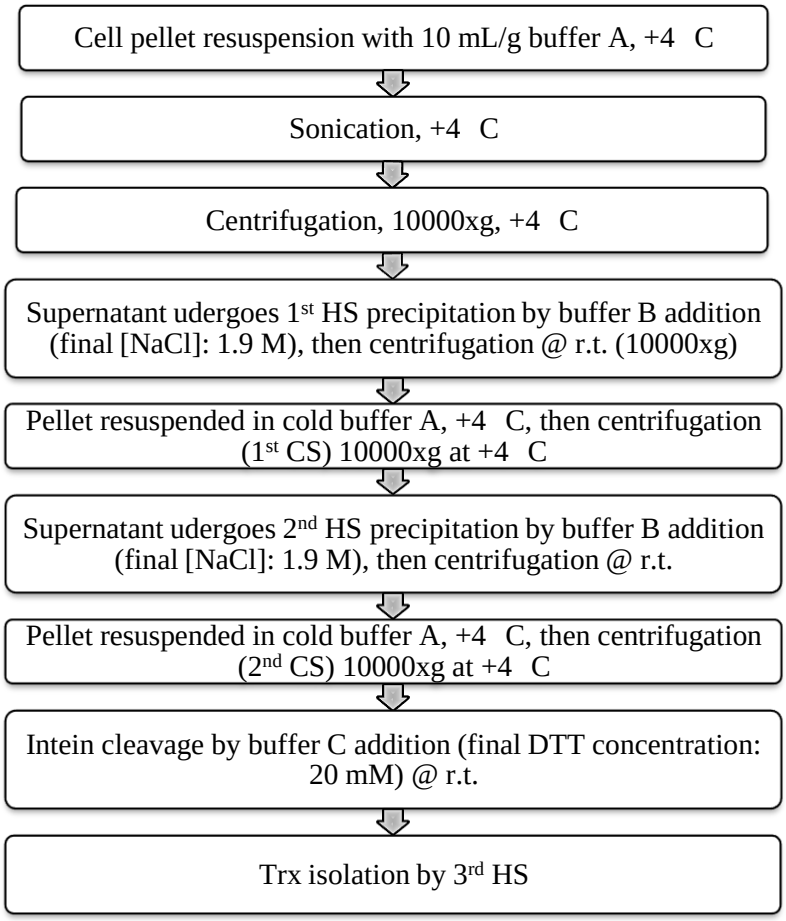


Figure 32. Purification scheme of Trx. HS: hot spin, CS: cold spin.

Buffer	Composition
Buffer A	10 mM TrisHCl, 2 mM EDTA, pH 8.5
Buffer B	5 M NaCl
Buffer C	2 M DTT

Table 3. Buffers used for Trx purification.

p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90 purification protocol

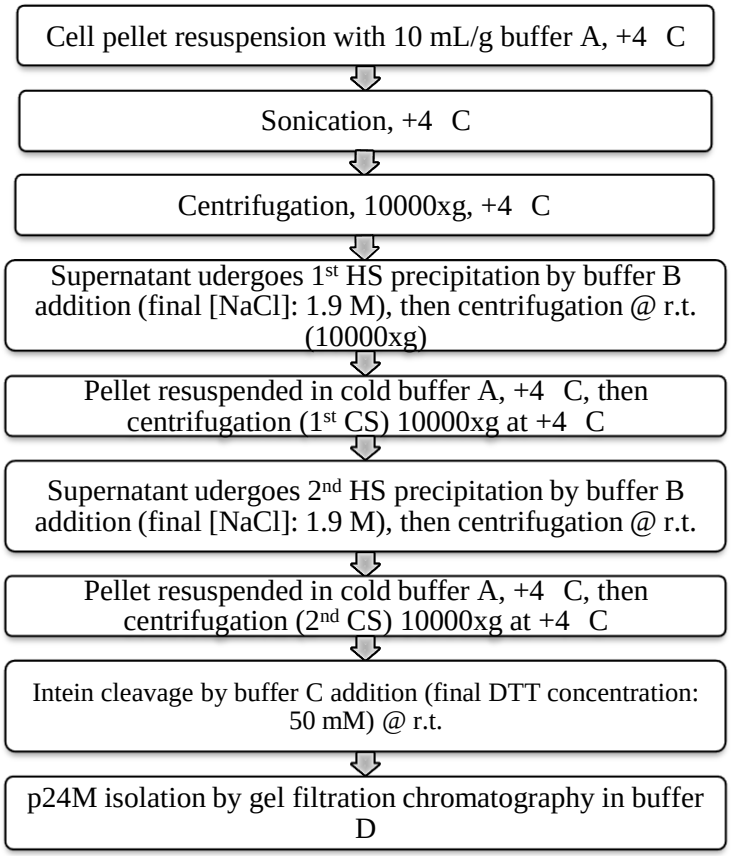


Figure 33. Purification scheme of p24M from p24M-*Mxe*GyrA-ELP90 and from p24M-*Mxe*GyrA(T3C)-ELP90

Buffer	Composition
Buffer A	10 mM TrisHCl, 2 mM EDTA, pH 8.5
Buffer B	5 M NaCl
Buffer C	2 M DTT
Buffer D	PBS 1x

Table 4. Buffers used in the purification of p24M from p24M-*Mxe*GyrA-ELP90 and from p24M-*Mxe*GyrA(T3C)-ELP90

Chromatographic conditions:

- column HiLoad xk 16/60 Superdex 75 prep grade
- flow rate: 1.2 mL/min
- isocratic elution in buffer D

p24M-MxeGyrA-ELP36 purification protocol

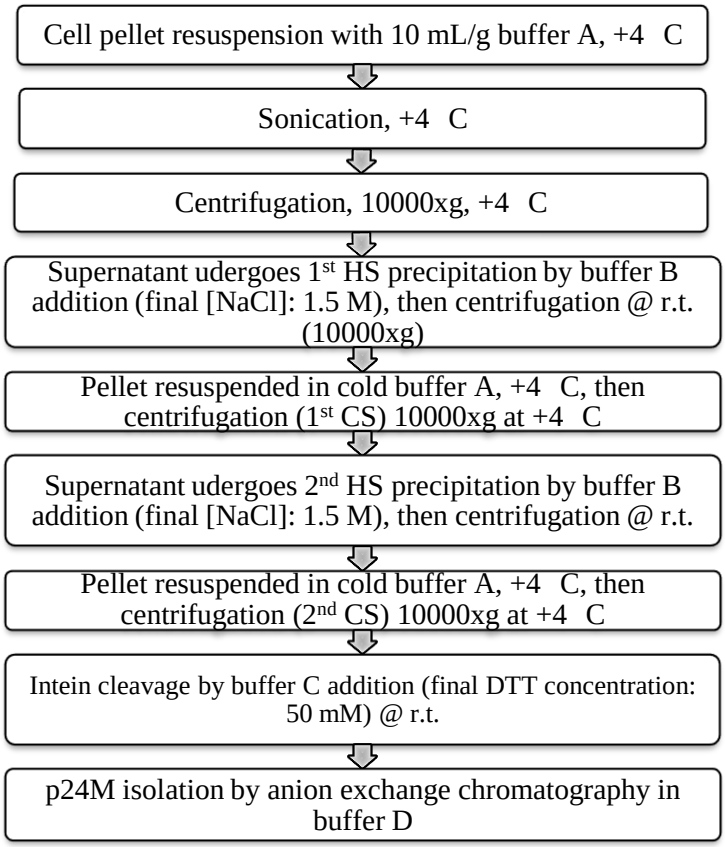


Figure 34. Purification scheme of p24M from p24M-*Mxe*GyrA-ELP36

Buffer	Composition
Buffer A	10 mM TrisHCl, 2 mM EDTA, pH 8.5
Buffer B	5 M NaCl
Buffer C	2 M DTT
Buffer D	PBS 1x
Buffer E	20 mM TrisHCl, pH 7.4
Buffer F	Buffer E, 1 M NaCl

Table 5. Buffers used in the purification of p24M from p24M-*Mxe*GyrA-ELP36

Chromatographic conditions:

- column: Q Sepharose HP xk 16/75
- flow rate: 2.0 mL/min
- Loading and washing buffer: Buffer E
- Protein elution: gradient: 0-100% buffer F in 3 CVs

Ct_mono-Mxe GyrA-ELP90, Ct_mono-Mxe GyrA-ELP36

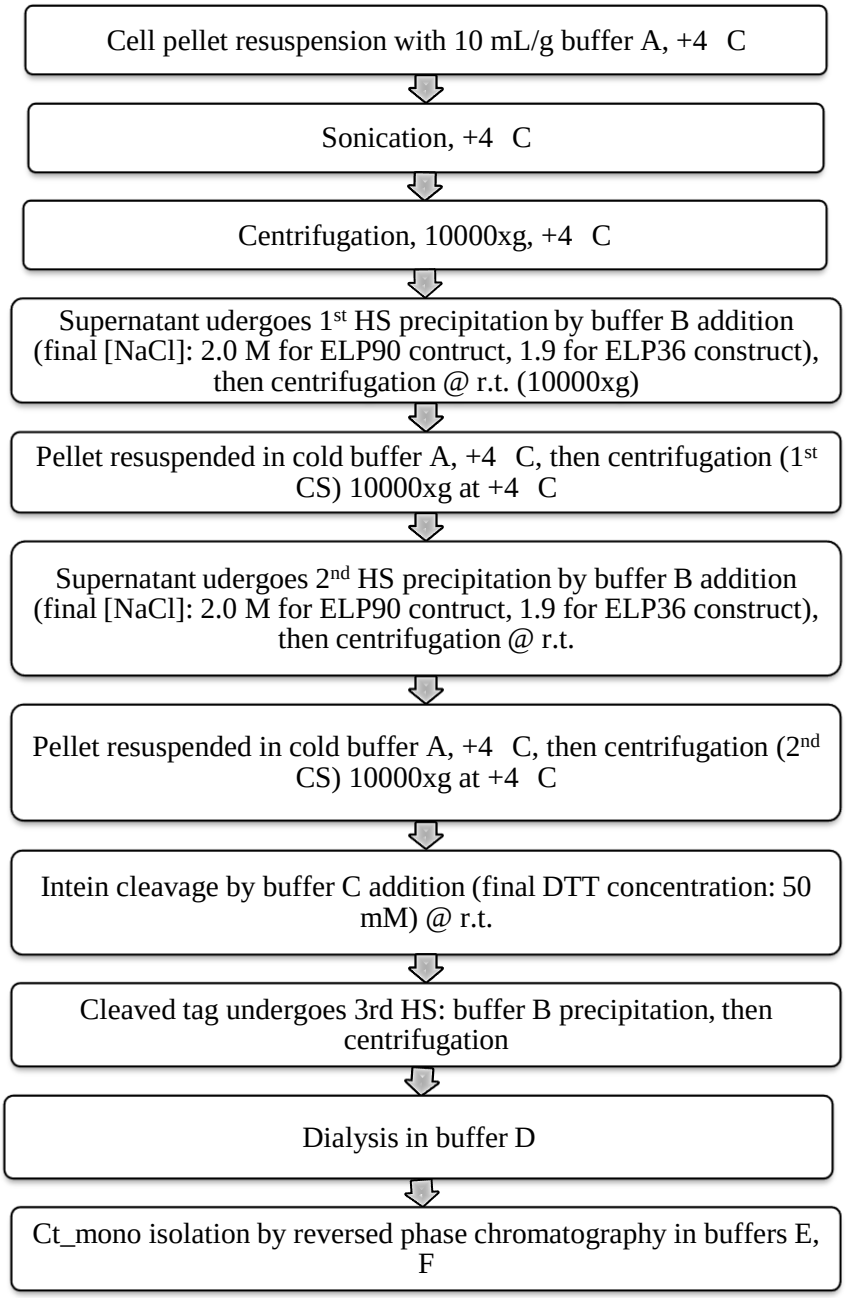


Figure 35. Ct_mono purification scheme from Ct_mono-*Mxe*GyrA-ELP90

Buffer	Composition
Buffer A	10 mM TrisHCl, 2 mM EDTA, pH 8.5
Buffer B	5 M NaCl
Buffer C	2 M DTT
Buffer D	PBS 1x
Buffer E	H_2O, Trifluoroacetic acid 0.1% (v/v)
Buffer F	CH_3CN, Trifluoroacetic acid 0.1% (v/v)

Table 6. Ct_mono purification scheme from Ct_mono-*Mxe*GyrA-ELP

Chromatographic conditions:

- column DeltaPak C4 15um 300A 7.8x300mm (Waters)
- flow rate: 1.5 mL/min
- Starting conditions: 90% buffer E, 10% buffer F
- Sample loading and washing: 90% buffer E, 10% buffer F
- Gradient: 10-60% buffer B in 50 mL
  60-80% buffer B in 7.5 mL
  80% for 10 mL
  80%-10% buffer B in 10 mL

LIAISON® ASSAY

LIAISON® SYSTEM

The Liaison® system is an instrument designed to perform immunometric analyses of biological fluid samples (such as serum or plasma) in a completely automated way. Up to 15 different tests can be performed at once on up to 144 samples in a sequential or random access mode. The output of the analysis is generated through the formation of an immune complex, followed by a chemiluminescent reaction that produces an emission of light.

The instrument is composed of two modules, designed to be both allocated on a single workbench (figure 36); the first module is a personal computer with touch screen, that hosts the user interface software and all the system data (assay protocols, reagent cartridges database, output of analyses, calibration history, network controls etc.). The second module is the actual analyzer, which performs the analysis from sample loading all the way to the final output for the user. Key components of the analyzer include:

- ✓ Cuvette loader and stacker: two conveyor belts allow continuous loading of the reaction modules, which are stored on a multilevel rack (7 levels).
- ✓ Sample rack slots: in the left-hand part of the instrument, a storage area can hold up to 12 sample racks, each carrying up to 12 samples. A barcode reader allows error-free cataloging of samples.
- ✓ Reagent slots: in the right-hand part of the instrument, another storage area can hold up to 15 different reagent cartridges simultaneously. This area is kept at a constant temperature of 15°C for optimal conservation of the reagents, while a stirring device keeps the microbeads always in homogenous suspension.
- ✓ Barcode reader for cartridge identification.
- ✓ Robot dispense arms: two robotic arms each carrying a dispensing needle. One arm is usually dedicated to dispensing samples, another to dispensing reagents. Each one has a separate washing well to clean the needle after each pipetting.
- ✓ Incubator: this area hosts the reaction cuvettes during incubation times, at a constant temperature of 37°C.
- ✓ Washing station: through the application of a magnetic field, this part allows retention of the paramagnetic microbeads and removal of the reaction liquids. Any number of washing steps with the desired washing buffer can be set.
- ✓ Read area: contains the injection devices of trigger reagents and the photomultiplier tube

Figure 36. Liaison® system.

The Liaison® system is based on two key features: the use of paramagnetic microbeads as the solid phase (figure 37) and the generation of signal by means of chemiluminescence.

The adoption of microbeads as the solid phase instead of the classic immunoassay supports, as the ELISA microwells gives a clear edge in terms of available reaction surface, which in turn increases the kinetic rates of the antigen-antibody complex formation. Moreover, diffusion of both the analyte and the solid phase in the reaction volume is allowed, while in ELISA system only the analyte can diffuse, decreasing the possibility of the immune complex formation. The microbeads adopted in the Liaison® system are colloidal particles composed of a ferric oxide core covered by a polystyrene layer formed by spontaneous coalescence of polystyrene linear chains. This structure is in turn coated with another layer composed of polyurethane activated with tosyl- groups. The tosyl- group (4-toluenesulfonyl chloride) can undergo nucleophilic attack, allowing the beads to covalently bind proteins through their available amino groups (ε-amino groups of lysines, N-terminal end). The paramagnetic properties of these microbeads allow easy manipulation through the application of a magnetic field. The particles respond to a magnet but are not magnetic themselves and retain no residual magnetism after removal of the magnet.

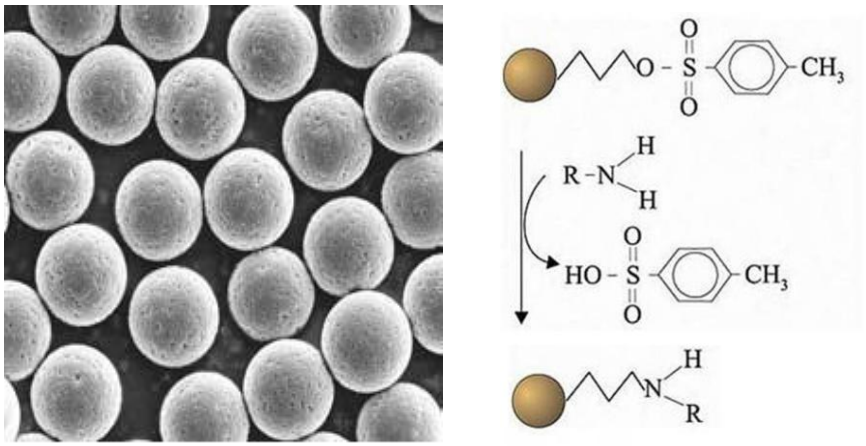


Figure 37.paramagneticmicrobeads are functionalized with a tosyl-group, to allow binding on amino-groups of proteins.

The tracer molecule is an antigen or antibody conjugated to a signal generating compound. Chemiluminescent tracers are formed by conjugating the antibody (or antigen) to a molecule that can generate a photon emission upon addition of certain reagents. The entity of this photon emission is measured with a luminometer equipped with a photomultiplier tube.

The chemiluminescent molecule used in the Liaison® system is the luminol derivate ABEI (N-(4-Amino-Butyl)-N-Ethyl-Isoluminol), which is converted to its activated ester to allow conjugation with the antibody or antigen (figure 38). In presence of H_2O_2 and a microperoxidase (deuteroferriheme), ABEI achieves an excited state in consequence of a chemical reaction. This excited level decays to the ground level generating energy in form of light. The emission of light is recorded by the photomultiplier tube for an interval of 3 seconds ("flash" chemiluminescence) and the signal is integrated over this interval. The final result is expressed in RLUs (Relative Light Units).

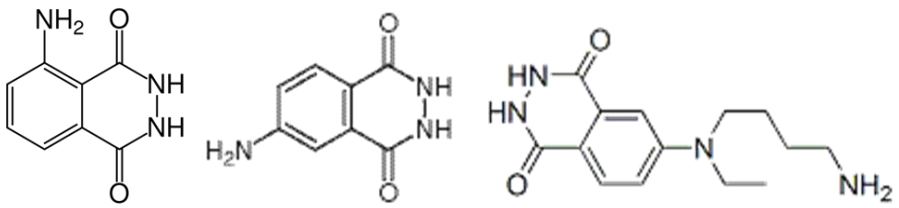


Figure 38.Luminol, isolumiol and ABEI

Employment of chemiluminescence is a great improvement over enzymatic signal generation of classic ELISA format assays. Sensitivity is highly increased and a greater dynamic range can be achieved. Lower molecular weight and steric hindrance of ABEI compared to horseradish peroxidase allow conjugation of more signal generating molecules per tracer molecule. Moreover, generation and recording of signal is completed in a very short time (3 seconds), with a sensible throughput increase.

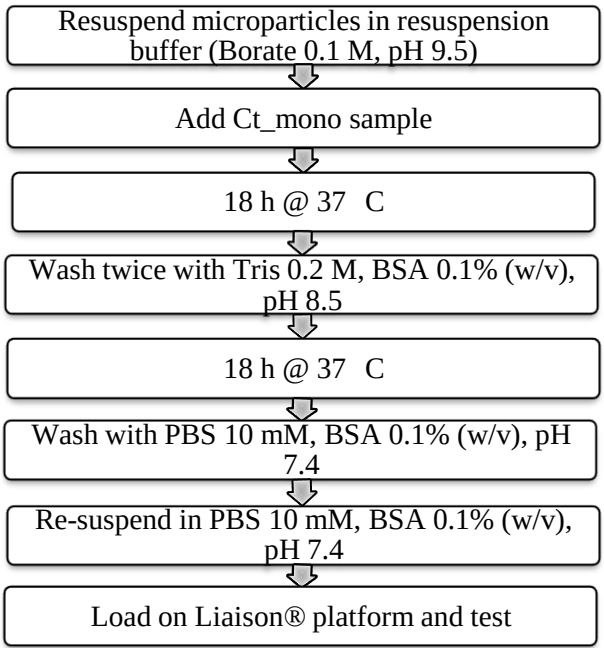
p24M

Microparticles for p24 antigen detection are streptavidinated. Monoclonal anti-p24 antibodies are added together with the sample. One of the monoclonal antibodies is biotinilated and the other is conjugated to ABEI:

the immunocomplex is then captured on the microparticles. In presence of H_2O_2 and a microperoxidase (deuteroferriheme), ABEI achieves an excited state in consequence of a chemical reaction. This excited level decays to the ground level generating energy in form of light. The emission of light is recorded by the photomultiplier tube for an interval of just 3 seconds and the signal is integrated over this interval. The final result is expressed in RLUs (Relative Light Units).

Ct_mono

Microparticles coating for anti-Chlamydia trachomatis antibody detection:



The Liaison® test proceeds as follows:
Microparticles coated with Ct_mono antigen are added together with the diluted serum (standard dilution 1:16), washed and incubated @37 °C for 10'. The tracer, made of a secondary antibody conjugated to ABEI, is then added and, in case the sample is positive, is bound by the serum antibody. After a washing step, the magnetic microparticles go to the reading chamber, where, in presence of H_2O_2 and a microperoxidase (deuteroferriheme), ABEI achieves an excited state, which decays to the ground level generating energy in form of light. The emission of light is recorded by the

photomultiplier tube for an interval of just 3 seconds and the signal is integrated over this interval. The final result is expressed in RLUs (Relative Light Units).

RESULTS

THIOREDOXIN

Thioredoxin (Trx) was the first model protein used for our experiments. Trx is part of a small disulphide-containing redox proteins system that has been found in all the kingdoms of living organisms. Thioredoxin serves as a general protein disulphide oxidoreductase. It interacts with a broad range of proteins by a redox mechanism based on reversible oxidation of two cysteine thiol groups to a disulphide, accompanied by the transfer of two electrons and two protons. The net result is the covalent interconversion of a disulphide and a dithiol (see figure 39).

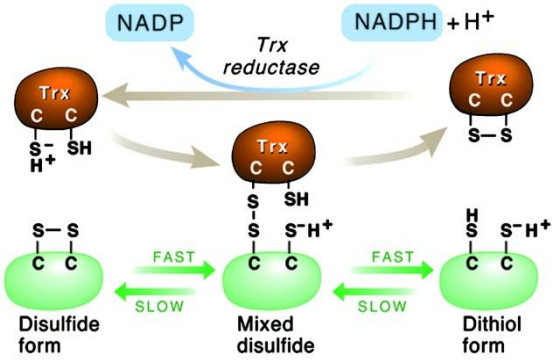


Figure 39. Regulation of protein activity by thioredoxin. (Kumar J.K., 2004)

Initially, a limited set of conditions for the expression of the Trx-*Mxe*GyrA-ELP protein were explored in order to find the best one. As many different expression conditions of ELP-fused proteins are reported in literature, without any univocal consensus on which is the one giving the highest expression yield, we explored some of them to find the one giving the best results in our case. In particular, the transformed BL21(DE3) *E.coli* strain with the pTME-Trx vector was grown in Luria-Bertani broth using four different expression conditions (see Table 7).

Fermentation	Temperature (°C)	Induction Time (h)	[IPTG] (mM)
A	37	3	1
B	37	24	0
C	20	24	0.3
D	20	48	0

Table 7. Pre-inoculus conditions were: 50 mL LB broth, 37 °C, o.n.

Analysis by SDS-Page (figure 40) revealed that the best expression conditions were those used in fermentation C (20 °C, 24 hours of induction time, 0.3 mM IPTG), followed by conditions of fermentation A. The two samples without induction (B and D, no addition of IPTG) did not show any protein induction at any time. All subsequent fermentations were performed using the parameters of fermentation C (20 °C, 24 hours, 0.3 mM IPTG induction). Fermentation volumes ranged from 0.5 to 1 litre.

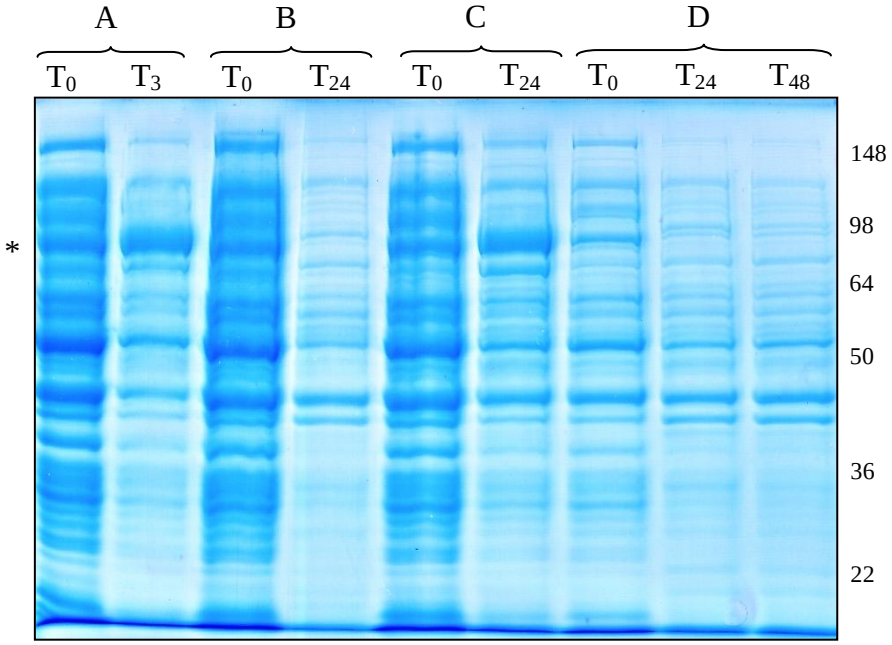


Figure 40. SDS-Page of the expression conditions tests. * indicates the band of Trx-*Mxe* GyrA-ELP90

The following step was dedicated to the setting up of the Trx purification process. As already mentioned before, the phase transition can occur either by gently heating/cooling the protein solution at temperatures around 30-40 °C depending on ELP properties, or in isothermal conditions by varying the salt concentration in the buffer. Regarding to the ELP precipitation steps, we decided to trigger the phase transition by adding NaCl, in order to achieve precipitation of our target protein without overcoming the room temperature: this method is more controllable and avoids eventual protein denaturation. In order to maximize the ELP dissolution and to speed up the purification process, the re-suspension steps

were conducted by acting on both temperature and salt concentration by bringing the sample at +4 °C and at low ionic strength.

First, in order to find the lowest NaCl concentration able to completely precipitate the Trx-*Mxe*GyrA-ELP90 fusion protein, a large range of salt concentrations was explored: 0 M, 1.0 M, 1.5 M and 2.0 M of NaCl (figure 41).

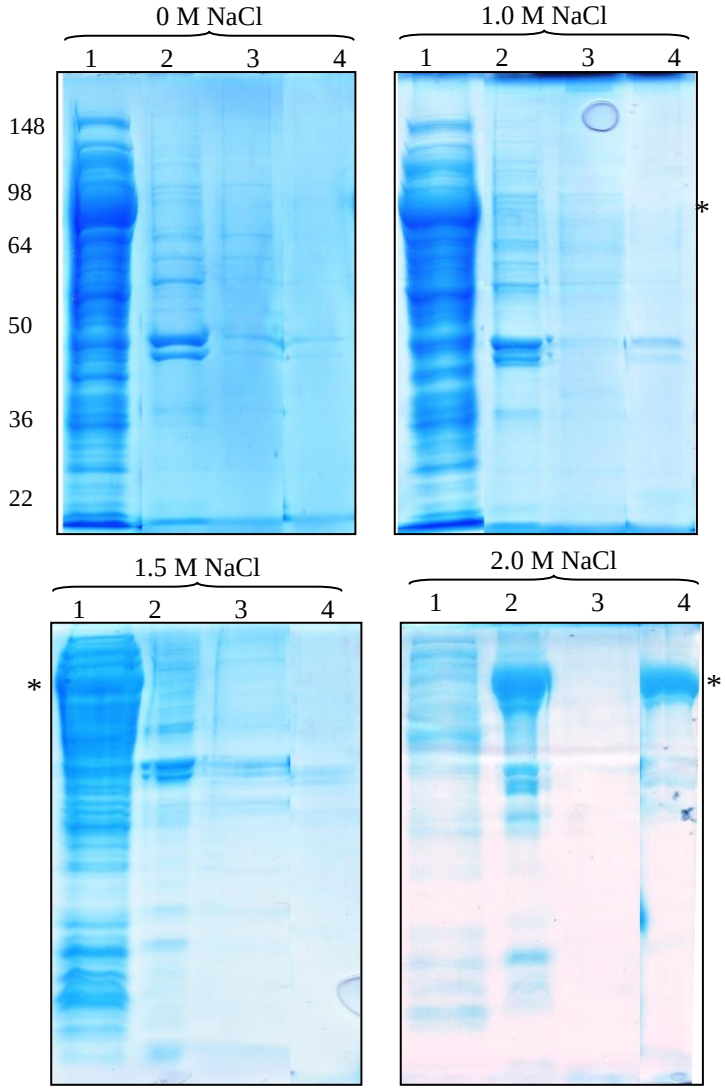


Figure 41. Precipitation tests at different NaCl concentrations. Lane 1: supernatant after 1st hot spin; lane 2: supernatant after 1st cold spin; lane 3: supernatant after 2nd hot spin; lane 4: supernatant after 2nd cold spin. Asterisk * indicated the band of Trx-*Mxe*GyrA-ELP90.

Only the highest NaCl concentration (2 M) permitted the complete recovery of the Trx-*Mxe* GyrA-ELP90 protein while a protein loss was still visible at 1.5 M NaCl concentration. The exploration range was further narrowed using the following points: 1.3 M, 1.5 M, 1.7 M, 1.9 M and 2.0 M NaCl, respectively (see figure 42).

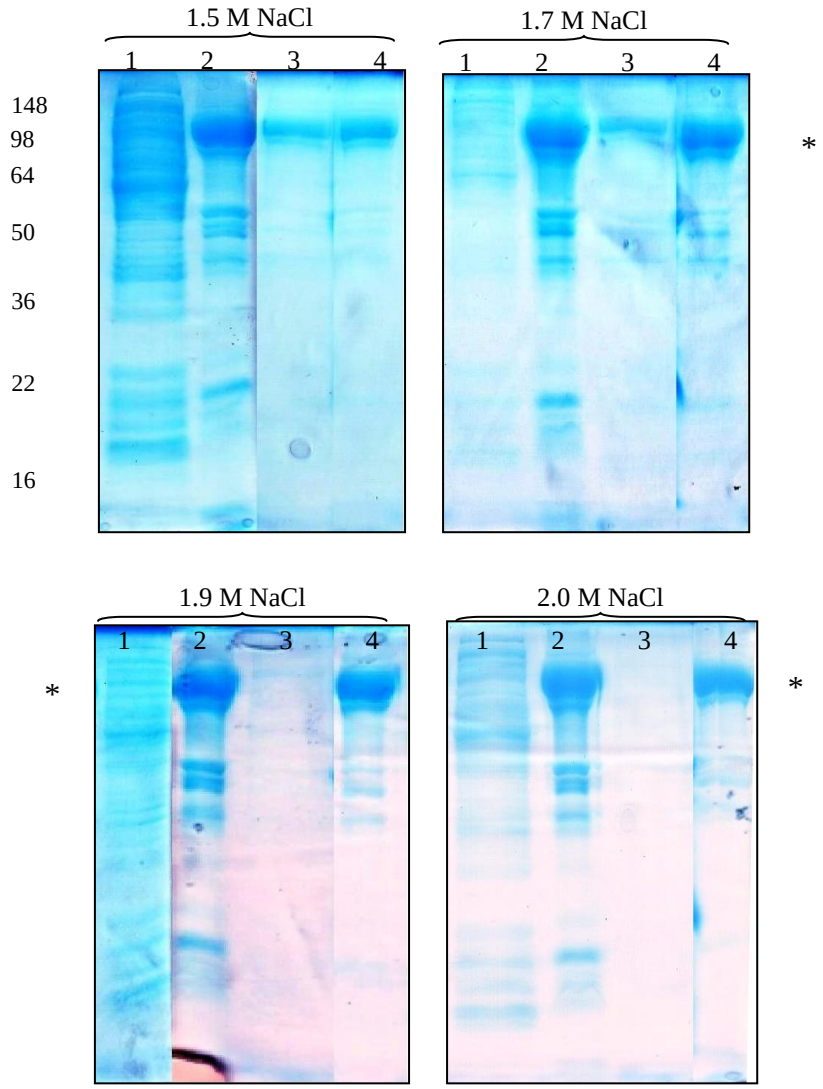
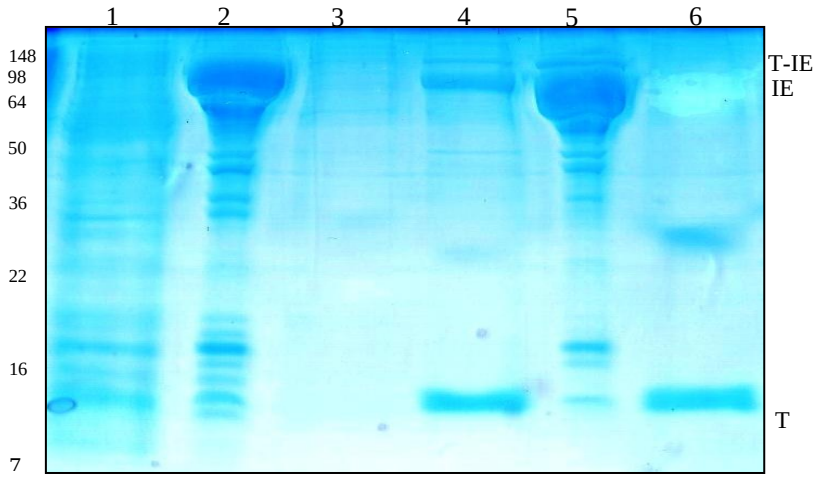


Figure 42. Precipitation tests at different NaCl concentrations. Lane 1: supernatant after 1st hot spin; lane 2: supernatant after 1st cold spin; lane 3: supernatant after 2nd hot spin; lane 4: supernatant after 2nd cold spin. Asterisk * indicated the band of Trx-*Mxe*GyrA-ELP90.

Concentration of 1.3, 1.5 and 1.7 M of NaCl used for precipitation were not enough to completely desolvate the fusion protein, in facts its band is still detectable in the supernatants after the second spin. The lowest NaCl concentration bringing about complete precipitation of the fusion protein without any protein loss in the supernatant was 1.9 M. Therefore all subsequent precipitation steps were performed at this salt concentration.

Two ITCs were performed for every purification batch; figure 43, lanes 1 to 3 show the ITC purification of the fusion protein. The next step was, according to literature, the intein cleavage reaction; it was performed at room temperature, in absence of salt, at pH 6.5 (figure 43, lane 4) and in presence of 20 mM DTT overnight. It is clearly seen that the cleavage reaction is almost quantitative, and only a small band at the precursor molecular weight remains. To separate the cleaved tag from the soluble protein a third transition cycle was performed by adding NaCl 1.9 M: the intein-ELP tag precipitates, while Trx remained in solution. As final step, the protein sample was dialysed against PBS to bring it back to physiological salt concentration.



Lane	Sample
1	1st Hot spin supernatant
2	1st Cold spin supernatant
3	2nd Hot spin supernatant
4	DTT cleavage
5	3rd Hot spin: *Mxe*GyrA-ELP90 tag
6	3rd Hot spin: Trx

Figure 43.SDS-Page analysis of Trx purification. T-IE: Trx-*Mxe*GyrA-ELP90, IE: *Mxe* GyrA-ELP90, T: Trx

This SDS-Page gel shows the protein content in some of the purification steps. After the first precipitation-solvation cycle, the fusion construct already reached a good purity grade (lanes 1 and 2) which was further increased by a second precipitation-resuspension cycle (lane 3: supernatant after precipitation). Lane 4 refers to the sample after overnight cleavage with DTT. The two cleavage products are visible: *Mxe* GyrA-ELP90 and Trx. Complete separation between the tag and the cleaved Trx was obtained with an additional ITC round, as showed in lanes 5 and 6. Lane 5 shows the ELP-intein tag with some minor impurities (probably degradation products) while lane 6 shows a band corresponding to quite pure Trx. No other band is visible in this lane, meaning a high purity level achieved by the method. Protein concentration was determined spectrophotometrically.

Thioredoxin activity assay

The thioredoxin activity, after purification with the intein-ELP method was determined by the insulin reduction assay. Native Trx (Trx-S_2) from *E.coli* or mammalian cells contains an active cysteine disulphide bridge that is reduced to the thiol form (Trx-$(SH)_2$) by NADPH and thioredoxin-reductase. The active form of thioredoxin is therefore Trx-$(SH)_2$ which acts as an effective protein disulphide reductase enzyme. The overall reaction carried on *in vivo* is:

$$\text{Trx-}S_2 + \text{NADPH} + H^+ \underset{}{\overset{\text{Trxreductase}}{\rightleftharpoons}} \text{Trx-}(SH)_2 + NADP^+$$

$$\text{Trx-}(SH)_2 + \text{protein-}S_2 \longrightarrow \text{Trx-}S_2 + \text{protein-}SH_2$$

The oxidoreductase activity of Trx is maintained *in vitro* in presence of reducing dithiols, such as dithiothreitol (DTT).
To verify whether our Trx, after being ITC-purified and intein-cleaved, is still active as biocatalyst, we assessed its ability to reduce insulin in presence of DTT, as reported by Holmgren (Holmgren A. 1979). Insulin is a protein consisting of two polypeptide chains (A and B) linked by two disulphide bridges (figure 44). When these two S-S bonds are cleaved by a reduction reaction, the free chains rapidly precipitate clouding the reaction solution. Turbidity is measured by recording absorbance at 630 nm.

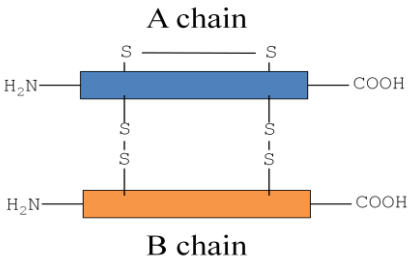


Figure 44. Scheme of insulin structure.

In our experiment, insulin was used at a final concentration of 0.75 mg/mL (≈ 0.13 mM), DTT at 1.0 mM (not included in the blank), while Trx concentrations were 0, 1.0, 3.0 and 6.0 µM, respectively. The buffer was PBS at pH 6.6, as this was reported in literature (Holmgren 1979) to be the ideal pH for this reaction; the final reaction volume was 300 µL. Along with Trx we also tested the cleaved ELP-intein tag recovered from the pellet in the 3rd ITC, to see whether traces of Trx remained uncleaved.
The absorbance at 630nm was measured up to two hours post-addition of the reagents.Table 8 shows the sample compositions (in µL), while figure 44 shows the Abs_{630nm} plot.

	Blank	**noTrx**	**Trx 1 µM**	**Trx 3 µM**	**Trx 6 µM**	**ELP-Intein 6 µM**
Insulin (stock 0.173 mM)	225	225	225	225	225	225
DTT (stock 10 mM)	0	30	30	30	30	30
Thioredoxin (stock 75 µM)	0	0	4	12	24	24
PBS	75	45	41	33	21	21

Table 8. All values expressed in µL

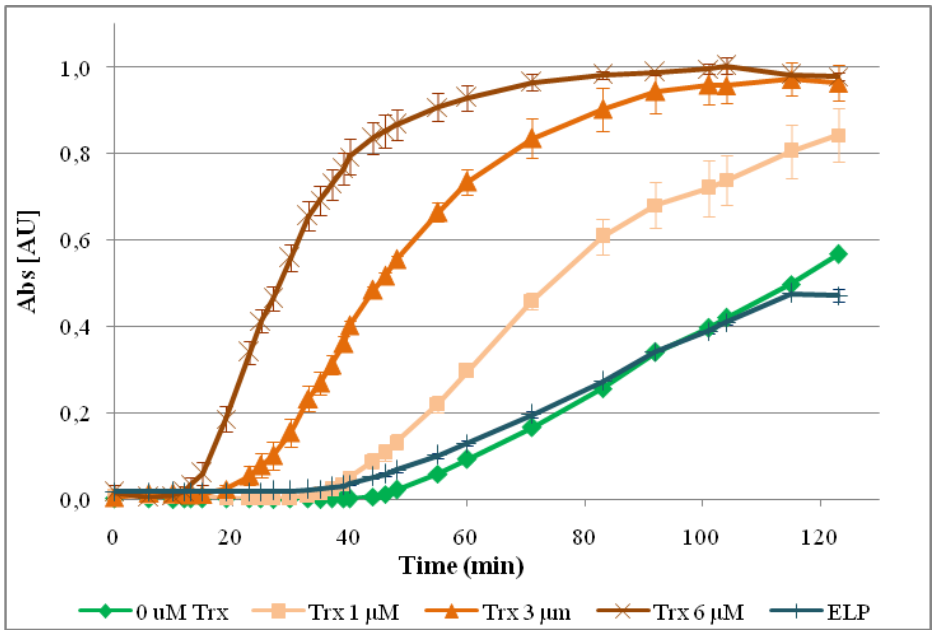


Figure 45.

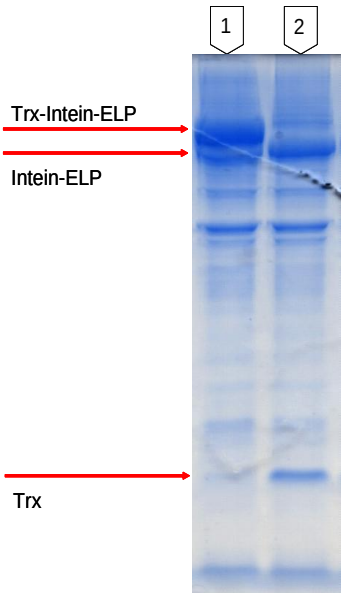
These data show that after ITC-based purification followed by cleavage from *Mxe*GyrA-ELP90 tag, thioredoxin retains its catalytic activity, as solution turbidity increases with the increase of Trx concentration. The curve obtained with cleaved *Mxe*GyrA-ELP90 tag resulted super imposable to that of the negative control, meaning the absence of residual Trx in the cleaved-tag sample.

*Mxe*GyrA point mutation

Both literature and experimental data show that *Mxe*GyrA(N198A) intein is affected by an *in vivo* auto-cleavage phenomenon which has the deleterious effect to decrease the final amount of purified target protein. This phenomenon is clearly visible in figure 46, where below the predominant band of Trx-intein-ELP fusion the one of fragment intein-ELP that is the product of the *in vivo* auto-cleavage reaction is visible.

In the next step, according to Cui *et al*'swork (Cui C. 2006), a single mutation was introduced in the intein sequence to try to eliminate or at least reduce the precocious auto-cleavage occurring *in vivo*: threonine in position 3 was mutated to cysteine.

Cui's group demonstrated that very low yield or no protein production of two target proteins (MESD and apoE), which were cloned at the N-terminus of the fusion protein *Mxe*GyrA(N198A)-chitin binding domain (CBD) in the commercially available pTWIN1 vector (NEB) already described, was caused by the *in vivo* auto-cleavage side reaction.



Cui's group hypothesizes that *in vivo* auto-cleavage follows a similar mechanism of the thiol induced-cleavage (figure 47), in which the free sulfhydryl group of Cys1 in the N-terminus of intein allows the *in vivo* N–S acyl rearrangement to form a thioester. The redox environment inside bacteria may mimic thiol reagents that cleave the thioester bond to separate the target protein from the intein–CBD fusion during bacterial expression. In this case, the free cysteine in the N-terminus of intein triggers the N–S acyl rearrangement and eventually results in this *in vivo* auto-cleavage.

Figure 46. *Mxe* GyrA(N198A) intein in vivo auto-cleavage. Lane 1: Trx-*Mxe* GyrA-ELP90 after two rounds of ITC. Lane 2: Trx-*Mxe* GyrA-ELP90 from line after overnight treatment with DTT (thiol induced-cleavage). In lane 1 the *Mxe* GyrA-ELP90 is already visible as results of the *in vivo* auto-cleavage.

In order to block the N–S shift, Cui proposed to replace this free sulfhydryl group of Cys1 of intein with a potential disulfide bond. A careful inspection of the X-ray crystal structure of *MxeGyrA*intein (figure 48) indicates that there are two residues closely located within a 5Å radius to the first cysteine residue and not involved on secondary structure elements: serine 53 and threonine 3. In Cui's laboratory two single mutants were prepared: S53C and T3C, respectively. While mutation S53C was able to partially block the *in vivo* auto-cleavage, the T3C mutant indicated that the *in vivo* auto-cleavage was completely blocked in all the test cases.

Based on these results, we decided to introduce the T3C mutation generating the pTME-T3C-MCS vector.

In order to obtain T3C mutant, site-directed mutagenesis was performed using the QuickChange™ mutagenesis kit from Stratagene. Finally, the mutation was confirmed by DNA sequence analysis. The new vector was called pTME-T3C-MCS.

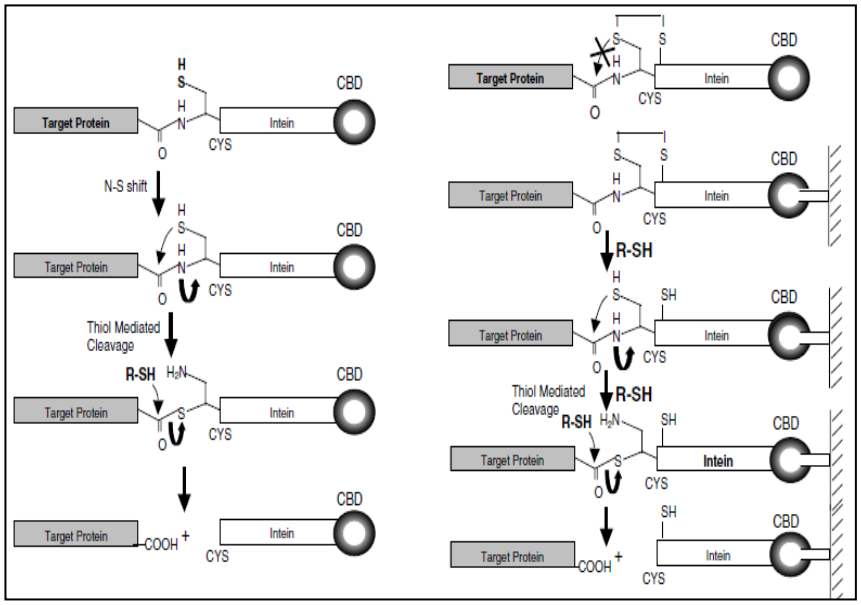


Figure 47. (A). Proposed mechanism of in vivo auto-cleavage according to Cui. (B). The possible mechanism of the single cysteine mutants T3C to block the un-intended in vivo auto-cleavage by replacing the free sulfhydryl group of cys1 of intein with a disulfide bond between the two cysteine in position 1 and 3, respectively (from Cui C., 2006).

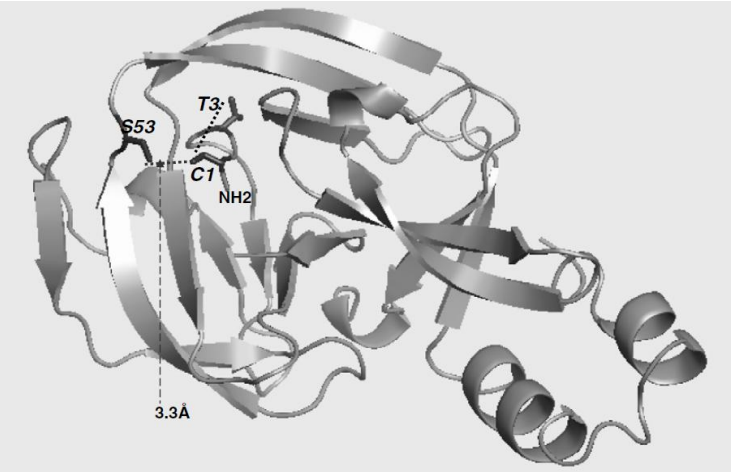


Figure 48. A ribbon representation of the X-ray crystal structure of an engineered *Mxe*GyrAintein (pdb code: 1AM2). The side chain atoms of residues S1, T3 and S53 are shown as sticks in dark colour to highlight the close distances among these residues (from Cui C., 2006).

Thioredoxin gene was inserted into the pTME-T3C-MCS vector and expressed and purified following the same protocol developed for the non mutated intein to assess the real effectiveness of the T3C mutation in limiting the *in vivo* auto-cleavage reaction. Experiments were performed in duplicate; volume of each fermentation was 500 mL. The results are shown in table 9.

Protein	**Trx-*Mxe* GyrA-ELP90**		**Trx-*Mxe*GyrA(T3C)-ELP90**	
Bacterial pellet weight (g)	4.01	3.08	3.06	3.03
Final amount of purified protein (mg)	18.25	15.10	19.03	19.53
Mean	16.47 (±0.10)		19.28 (±0.02)	
Δ% final yield (mg) T3C vs. wild type	+13.7%			

Table 9. Comparison of the expression and purification levels of the Trx-*Mxe* GyrA-ELP90 and Trx-*Mxe*GyrA(T3C)-ELP90 proteins.

Data comparison underlines that while the amount of pellets harvested from fermentations expressing the Trx-*Mxe*GyrA(T3C)-ELP90 protein is slightly lower than that expressing the wild type protein, analysis of the final purification yield, expressed as amount of purified protein, indicates that this is higher in the mutated protein compared to the its wild type: 13.7%. Although the limited amount of experimentally data available, in our case it can be concluded that the insertion of the 3TC mutation in the intein has not generate a sensible increase in the purification yield of the protein as claimed in literature. One possible explanation may be that in the reducing environment of the bacteria, Thioredoxin can act as a disulphide reductase and reduce the disulphide bridge between the two cysteine in position 1 and 3, nullifying the effectiveness of mutation in position 3. This consideration is indirectly supported by the fact that Cui did not used Trx as model protein in his experiments.

Densitometric analysis

The purified protein (two batches from wild type intein, two lots from T3C mutated *Mxe*GyrA) were run by SDS-Page and then analyzed by densitometry using ImageQuant TM (GE) software. The results are shown in figure 49 and summarized in table 10.
The overall purity is good: only minor contaminant are visible in the samples.

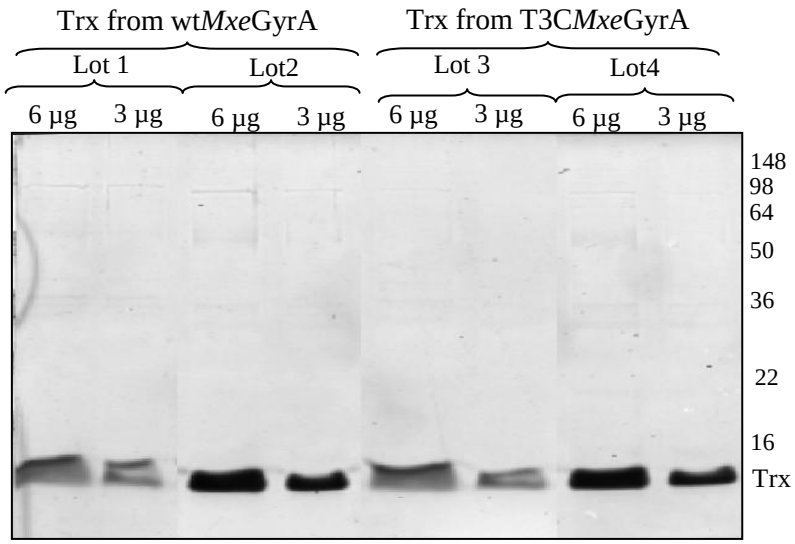


Figure 49. SDS-Page for densitometric analysis of the four batches of Trx

Protein	**Trx from wt*Mxe*GyrA**		**Trx from T3C *Mxe*GyrA**	
	Lot 1	**Lot 2**	**Lot 3**	**Lot 4**
Purity	95.5	99	99	99.5

Table 10.

p24 (HIV-1 VAR M)

p24M-Mxe**GyrA-ELP90**

As a second test protein for the application of the intein-ELP purification system, the structural p24 protein of HIV-1 M (from now on called p24M) was chosen. The capsid protein p24 is one of the main structural proteins of human immunodeficiency virus. There are about 1500 p24 molecules composing virus capsid in a mature virion, therefore p24 is the most abundant protein produced during virus replication and can be detected in the very early stage of asymptomatic phase after HIV infection. In the late stage of AIDS, p24 can be detected again because of rapid replication of viruses and decline of HIV specific antibody production. Hence p24 detection provides a means to aid the early diagnosis of HIV infection, track the progression of disease and assess the efficacy of antiretroviral therapy. Additionally, p24 is one of the most conserved viral proteins and shows very little antigenic variations among the isolated strains of HIV-1.

The gene of p24M was cloned into the pTME-MCS vector and expressed in *E.coli*. Several different condition of growth and expression were tested:

A LB broth, 37 °C, induction @0.6 OD_{600nm}, with 1 mM IPTG, 3 h induction;
B Terrific broth, 37 °C, induction @0.6 OD_{600nm} with 1 mM IPTG, 3 h induction;
C Terrific broth, 37 °C, no induction, 24 h induction;
D LB broth, 20 °C, induction @0.6 OD_{600nm}, with 0.3 mM IPTG, 24 h induction;
E Terrific broth, 20 °C, induction @0.6 OD_{600nm}, with 0.3 mM IPTG, 24 h induction;
F Terrific broth, 20 °C, induction @0.6 OD_{600nm}, with 0.3 mM IPTG (other 0.7 mM of IPTG were added 4 h after induction), 24 h induction.

p24M-*Mxe*GyrA-ELP90 protein induction levels were analyzed both by SDS-Page and western blot figures 50 and 51, respectively. SDS-Page gels were scanned and the p24M-*Mxe*GyrA-ELP90 optical band areas quantified

by densitometric analysis using the ImagQuant TL software from GE. The western blot analysis revealed that a large amount of protein is present in all samples after induction (and in some cases even at time zero); in all cases a lot of bands at lower mass than that of the fusion construct, which are responsive to the anti-p24M antibody are clearly visible.

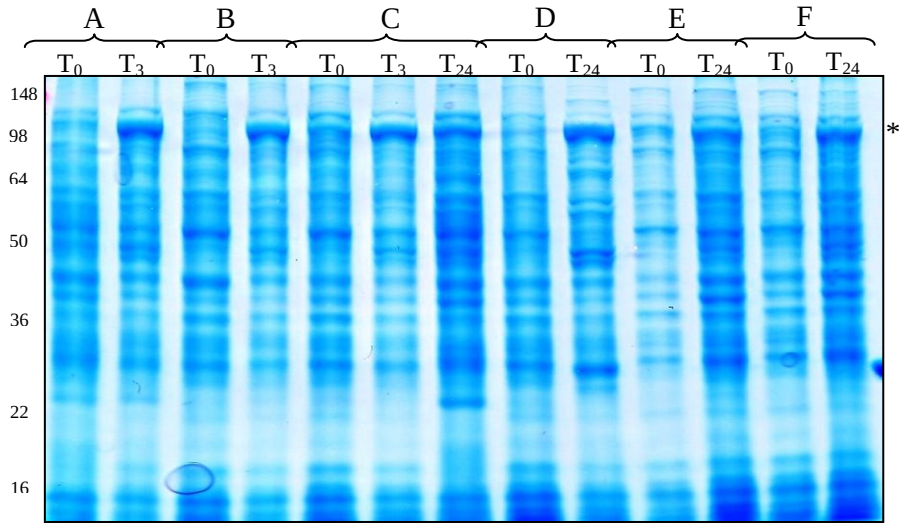


Figure 50. SDS-Page analysis of different expression conditions for p24M-*Mxe* GyrA-ELP90. * indicates the target protein band.

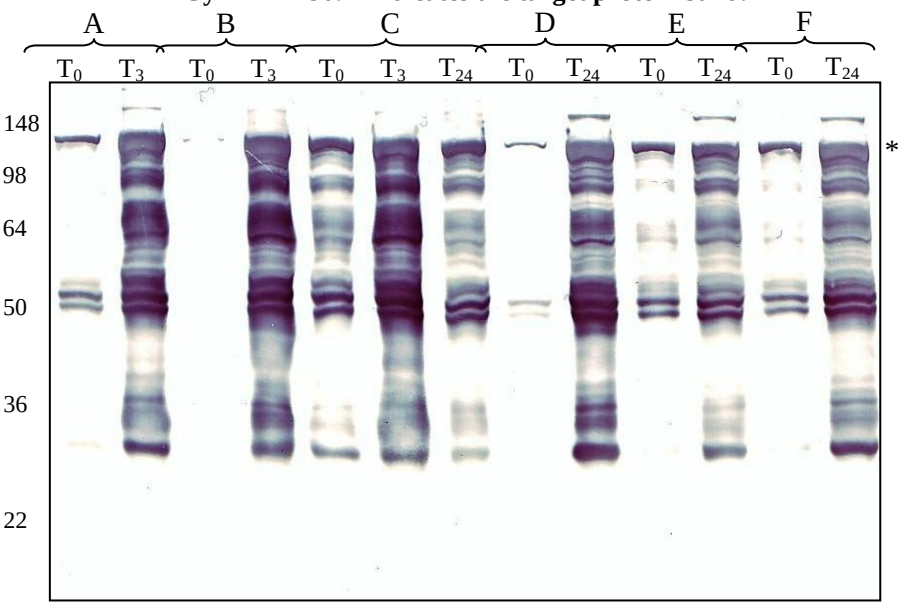


Figure 51. WB analysis of different expression conditions for p24M-*Mxe* GyrA-ELP90. * indicates the target protein band.

The densitometric analysis gave the following results, see figure 52. The optical band area is measured in arbitrary units (AU). The best conditions, looking only at this parameter, are those used in protocol D: LB broth, 20 °C, induction @0.6 OD_{600nm}, with 0.3 mM IPTG, 24 h induction;

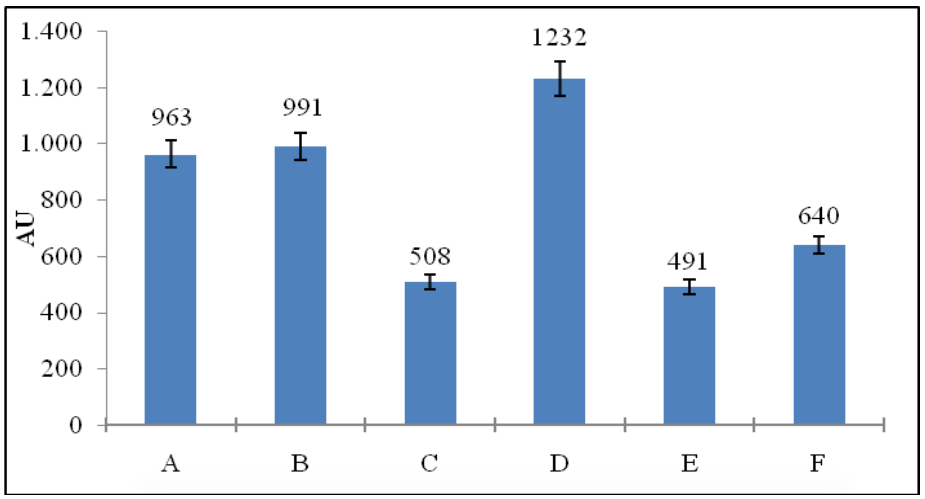


Figure 52. Plot of the p24M-Mxe GyrA-ELP90 optical band area.

The final cell density reached at the end of the fermentation, which was measured by light absorbance at 600 nm, was also determined in all six test conditions. These values are shown in figure 53. According ot this parameter the best fermentation conditions are those used in protocol C.

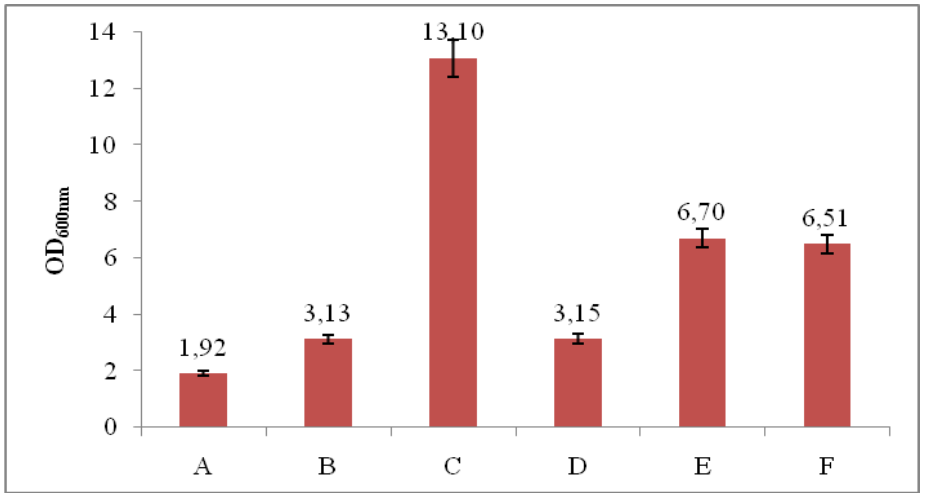


Figure 53. OD_{600nm} values at the end of each fermentation.

The plot of figure 54 shows the magnitude obtained multiplying the optical band area value (figure 52) times the final cell density value of each fermentation. This magnitude gives an estimation of the total amount of protein produced in the fermentation

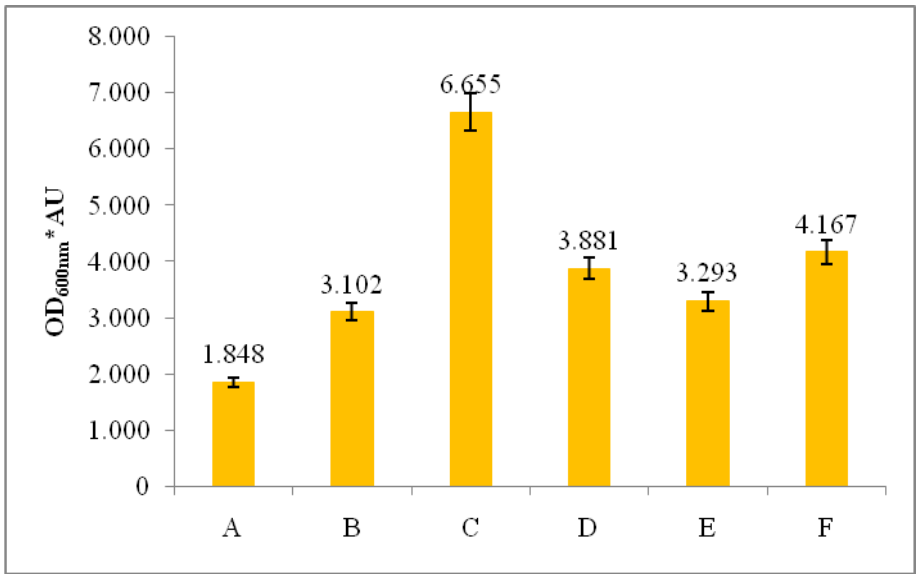
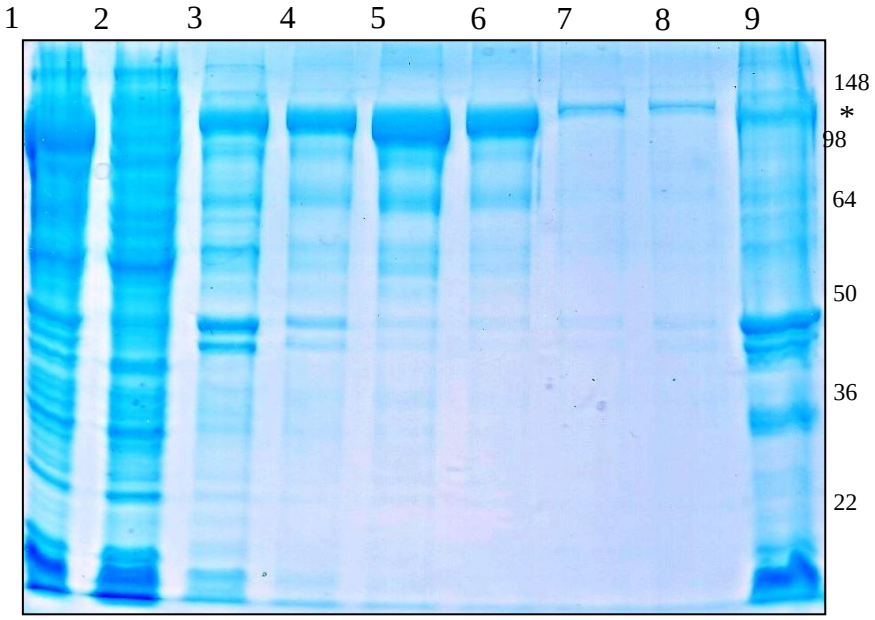


Figure 54. Plot of the OD_{600nm}*optical band area for each fermentation condition

The best fermentation condition seems to be C (Terrific broth, 37 °C, no induction, 24 h fermentation) having though a very low amount of produced protein per cell, but an incredibly high cell density. As this condition is very time-consuming, we chose to employ the protocol B (Terrific broth, 37 °C, induction @0.6 OD_{600nm} with 1 mM IPTG, 3 h fermentation) which appeared to be a good compromise between the total amount of produced protein and the total fermentation time.

After we choose the fermentation conditions, we explored the solubility level of the fusion construct. Protein solubility test is performed according to the protocol described in the "materials and methods" chapter. Samples from the solubility test were loaded on SDS-Page, see figure 55.

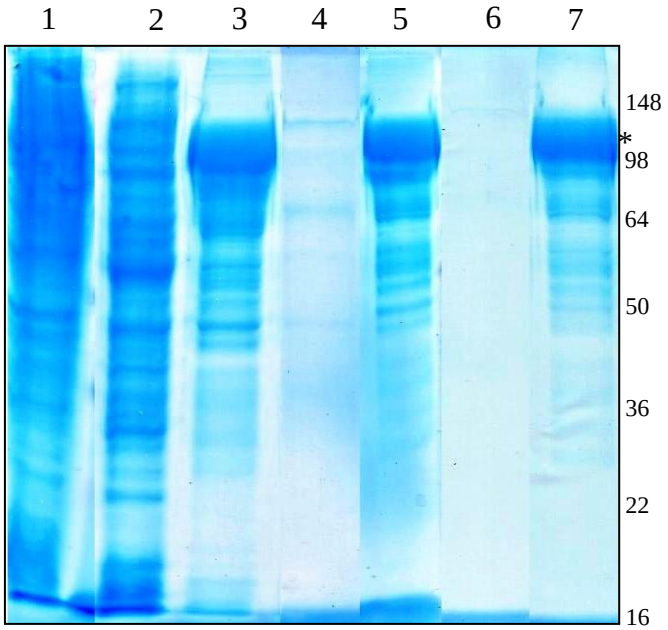


Lane	Sample
1	Total lysate
2	Soluble fraction
3	Sample soluble in urea 1 M
4	Sample soluble in urea 2 M
5	Sample soluble in urea 3 M
6	Sample soluble in urea 4 M
7	Sample soluble in urea 5 M
8	Sample soluble in urea 6 M
9	Sample soluble in SDS 1%

Figure 55.Solubility test of p24M-*Mxe* GyrA-ELP90. * indicates the band of the fusion construct.

The band corresponding to p24M-*Mxe* GyrA-ELP0 is visible in the total lysate (lane 1) and in the lanes 3-6 (buffer containing urea from 1 to 4 M). A

completely soluble protein would be the best purification condition, and up to now no data have been reported in literature about ITC purification and subsequent cleavage in urea-containing buffers.
The purification was carried out with two ITCs in presence of urea 4 M, and the precipitation was achieved with a final NaCl concentration of 2.0 M. The ITCs worked well, no protein was lost during the process, and after two cycles we had a fusion construct with good purity. See figure 56.



Lane	Sample
1	Total lysate
2	Soluble fraction
3	Insoluble fraction
4	1st HS
5	1st CS
6	2nd HS
7	2nd HS

Figure 56. SDS-Page of ITC purification in buffer containing urea 4.0 M. * indicates the band of p24M-*Mxe* GyrA-ELP90

The intein cleavage in urea 4 M did not work, as can be seen in figure 57. We also tried the intein cleavage in urea 2M: after the last precipitation, the sample was suspended in urea 2.0 M, and with the same urea concentration

the cleavage was carried on. In this case partial cleavage was seen: the band of the precursor (p24M-*Mxe*GyrA-ELP90) was still present, but also the products bands (p24M and *Mxe*GyrA-ELP90) were visible. This should be due to the chaotropic property of urea, which unfolding the *Mxe*GyrA intein, does not allow a correct conformation of the catalytic site.

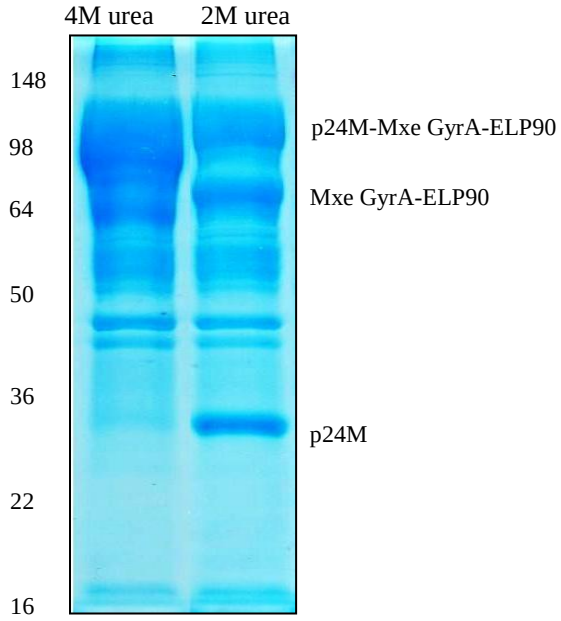


Figure 57. SDS-Page analysis of the cleavage reaction in 4 M urea and 2 M urea.

To overcome the problem of insolubility and consequent impossibility of proceed with an effective intein cleavage, we tried to increase the solubility of protein p24M-*Mxe*GyrA-ELP90 by lowering the fermentation temperature during protein induction from 37 °C to 15 °C.
The new expression conditions were:

- Pre-inoculus @ 37 °C, in LB broth
- Dilution 1:35 of pre-inoculus in LB broth, growth @ 25 °C up to OD_{600nm} of 0.6.
- @ OD_{600nm} = 0.6 induction with 0.1 mM IPTG; 15 °C, 24 h

In these conditions the protein was expressed in a good amount as shown in figure 58.

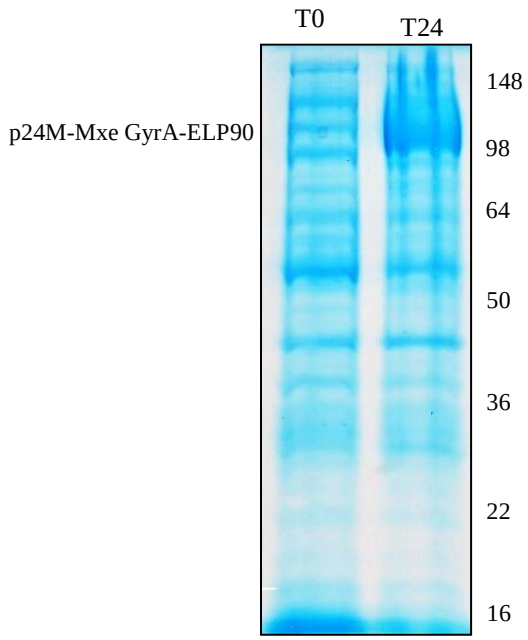
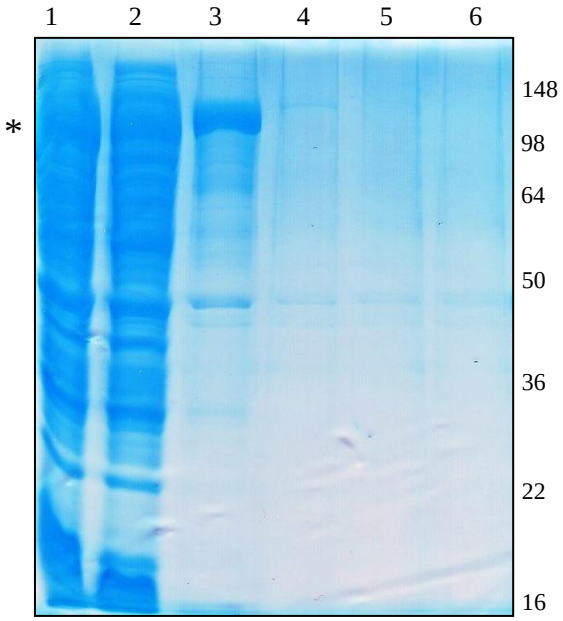


Figure 58. SDS-Page analysis at T0 and 24h after IPTG induction @ 15 °C

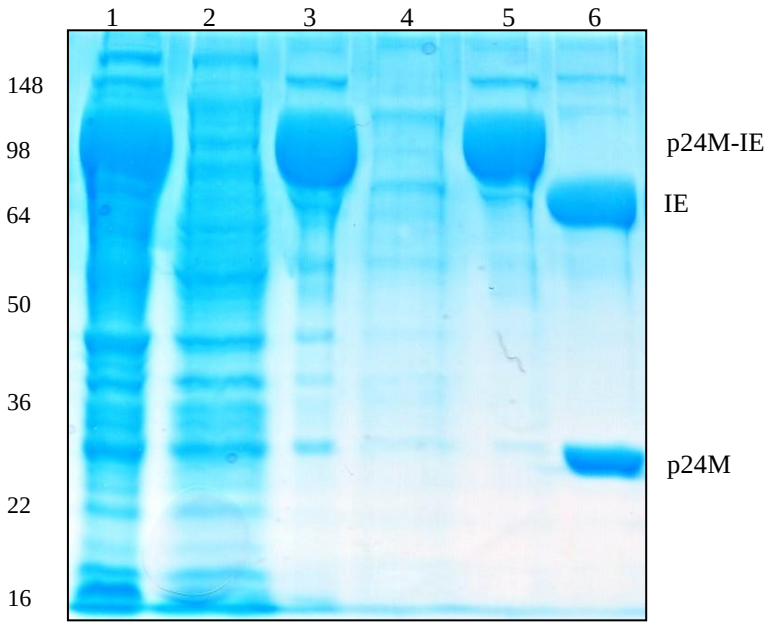
The solubility test gel (figure 59) shows the band of the fusion construct both in the soluble fraction and in the sample solubilised in buffer containing urea 1M, but no bands in the lanes of sample solubilised with higher urea concentrations.



Lane	Sample
1	Total lysate
2	Soluble fraction
3	Sample soluble in urea 1 M
4	Sample soluble in urea 2 M
5	Sample soluble in urea 3 M
6	Sample soluble in urea 4 M

Figure 59. Solubility test of p24M-Mxe GyrA-ELP90 expressed @ 15 °C. * indicates the position of the band corresponding to the target fusion construct.

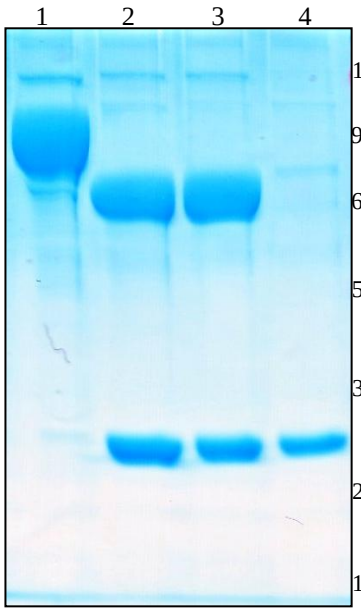
As the whole p24M-*Mxe*GyrA-ELP90 protein is soluble in buffer phosphate 50 mM, urea 1 M, pH 8.0, the ITC purification process and cleavage were performed in this buffer; p24M-*Mxe*GyrA-ELP90 fusion protein is precipitated in the spins with 1.9 M NaCl as final concentration and the intein cleavage reaction is performed by the addition of DTT 50 mM under magnetic stirring at room temperature overnight. Fusion construct is purified with a good purity in 1 M urea and the cleavage is quantitative (figure 60).



Lane	Sample
1	Soluble fraction
2	1st HS
3	1st CS
4	2ndHS
5	2nd CS
6	DTT cleavage

Figure 60.SDS-Page analysis of ITC purification and cleavage of p24M-*Mxe* GyrA-ELP90. P24M-IE stands for p24M-*Mxe* GyrA-ELP90, IE for *Mxe* GyrA-ELP90.

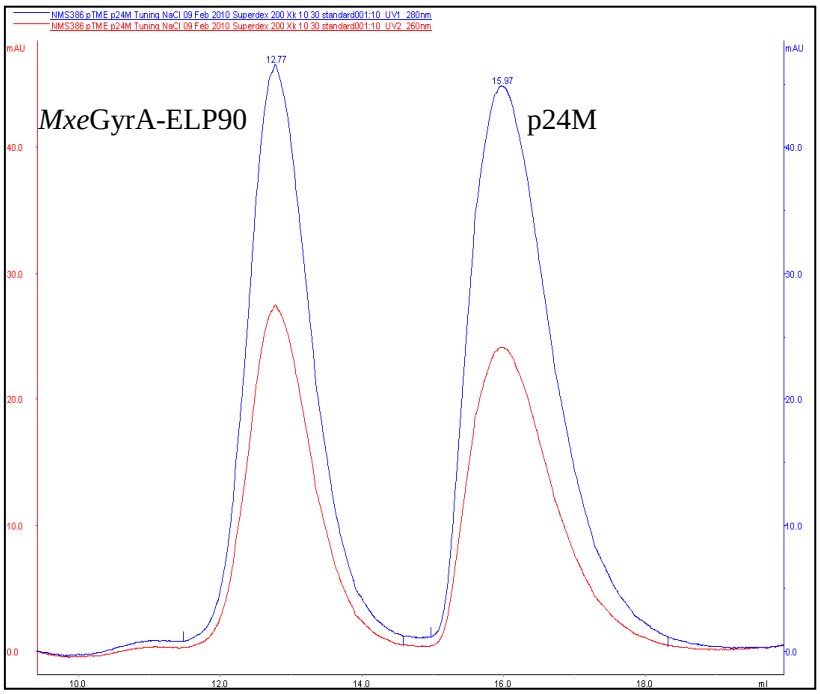
When the additional ITC was performed (adding 1.9 M NaCl) to separate the cleaved *Mxe*GyrA-ELP90 tag from the p24M, the separation was not perfect at all: a large part of p24M was found in the precipitated part together with the tag (figure 61). The attempt of precipitating the tag with NaCl 2.5 M was also unsuccessful (data not shown).

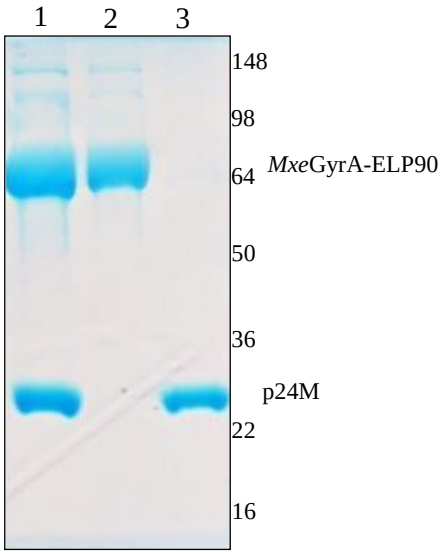


Lane	Sample
1	2nd CS
2	DTT cleavage
3	3rd HS pellet
4	3rd HS supernatant

Figure 61.

As the final ITC was not effective to separate the target protein (p24M) from the cleaved *Mxe*GyrA-ELP90 tag, as it was for thioredoxin, the difference in size of the two species allowed a perfect separation via size exclusion chromatography (figure 62). The first peak corresponds to the cleaved tag and the second one to p24M, respectively.





Lane	Sample
1	DTT cleavage
2	1st GFC peak
3	2nd GFC peak

Figure 62.Chromatogram of Size exclusion purification of the cleaved p24M and Mxe GyrA-ELP90 tag (above) and corresponding SDS-Page (below).

Three lots of p24M from pTME-p24M and three using the same vector, but with the mutated intein *Mxe*GyrA(T3C) were expressed and purified with this method. The yields were as reported in table 11.

p24M from p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90			
	Lot	**Yield (mg/L)**	**Mean (mg/L)**
T3C	1	11,4	16,4±6,2
	2	14,4	
	3	23,3	
WT	4	22,3	16,6±5,0
	5	13,0	
	6	14,6	

Table 11. p24M yields.

The yield values are lower than what expected and than what is usually purified by affinity chromatography (40-50 mg/L of fermentation): they are even lower than what we obtained in the purification of Trx, which is about half the mass. Also in this case, after Trx, we can see that in the case of T3C point mutatied intein the yield is not higher than for the wild type intein.

PURITY AND CHARACTERIZATION

All batch preparations obtained from construct p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90 were characterized first by SDS-Page. A total of 6 preparations of p24M, three for each construct, underwent an electrophoretic separation, according to the protocol described in "Matherials and methods". Figures of the gels and the determined p24M purity values are shown in figure 63 and in table 12, respectively.

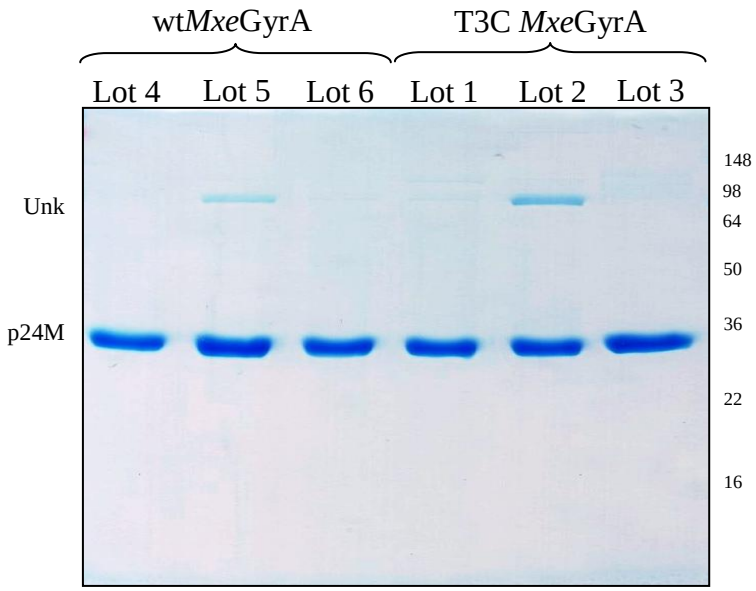


Figure 63. SDS-Page assessing p24M purity.

p24M from p24M-*Mxe*GyrA-ELP90, p24M-*Mxe*GyrA(T3C)-ELP90				
	Lot	**Purity SDS-PAGE(%)**	**Mean**	**Std dev**
T3C	1	100	98	2.3
	2	96		
	3	99		
WT	4	98	94	6.8
	5	86		
	6	98		

Table 12. Purity and yield summary of p24M from p24M-*Mxe*GyrA-ELP90

In all cases a good purity is achieved with this purification method: only in one case the purity goes under 90, while all other samples achieved an SDS-Page purity higher than 95%. In two cases (lots 2 and 5of p24M from p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90, respectively) independently from the intein mutation a small band, with an apparent double molecular weight of p24M is present in the sample. This band was hypothesized to be a dimeric form of p24M retained even in SDS-Page or a

degradation product of the whole fusion construct. We investigated the nature of this band by western blot and by PMF (peptide mass fingerprinting).

For the western blot analysis a high affinity monoclonal antibody was employed. The small band was recognized by the antibody, meaning that the band contains the p24M sequence: it could be either a dimeric form which retains its structure even in SDS-Page (see figure 64) or a degradation product from the p24M-intein-ELP90 construct.

For the peptide mass fingerprinting the proteins of the two bands corresponding to p24M and unknown compound (lot 5) were treated with Asp-N (which cleaves the N-term peptide bond of aspartic acid and, less frequently of glutamic acid) and with trypsin (cleaving the N-term of lysine and arginine) and the peptides formed from these treatments were analyzed by MALDI-TOF/TOF. The results confirm the western blot result: both bands contain peptides only from p24M. No peptides matching intein or ELP sequence were detected (figure 65), letting us conclude that the small upper band should made of a dimeric form of p24M.

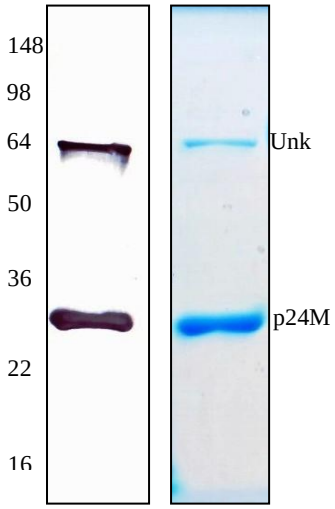


Figure 64.Western blot (left) and corresponding SDS-Page (right) of lot 2 of p24M.

MEFMNSAMVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMF
SALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPV
HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKR
WIILGLNKIVRMYSPTSIL**DIRQGPKEPFRDYVDRFYKTLRAEQAS**
QEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGP
GHKARVLKLRSGSGSGMRM

Legend:

In yellow the aminoacids found in both bands after trypsin digestion.

In cyan the aminoacids found only in band 2 after trypsin digestion.

In bold the aminoacids found in band 1 after Asp-N digestion.

Underlined the aminoacids found in band 2 after Asp-N digestion.

Figure 65.p24M sequence and peptide coverage after enzymatic treatments.

The purity was also assessed by UPLC-MS (column Acquity UPLC BEH300 C4 1.7µm, 2.1x100 mm, Waters; buffer A: H_2O, TFA 0.1% v/v, buffer B: CH_3CN, TFA 0.1% v/v; gradient 30-60% buffer B in 4'). All the batches show the same chromatographic profile of figure 66: a high and steep peak, followed by a smaller peak.

As reported in table 13, the main peak is in good agreement with the theoretical molecular mass, while the second peak has a molecular mass of 117 Da higher than that of the main peak. Considering that p24M contains a couple of cysteines, it can be hypothesized that a post-translational modification occurred on a fraction of p24M.

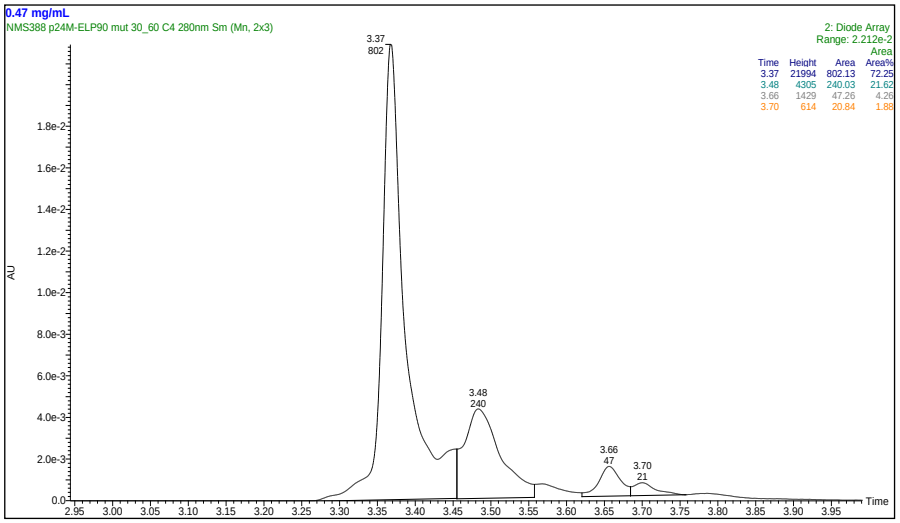


Figure 66.UPLC profile of p24M cleaved and purified from *Mxe*GyrA-ELP90 tag.

	T3C *Mxe*GyrA			**wt*Mxe*GyrA**		
Lot	**1**	**2**	**3**	**4**	**5**	**6**
Purity 214nm (%)	74.30	95.19	84.03	78.88	95.13	80.55
Purity 280nm (%)	72.25	94.64	88.15	83.52	94.45	81.20
Main peak (uma)	27565±1					
2ndary peak (uma)	27682±2 (+117)					
Theorethical	27558					

Table 13. UPLC-MS purity data.

To assess the nature of the second peak, the sample was loaded on a semi-preparative HPLC column (DeltaPak C4 300A 7.8x300 mm, Waters). The chromatographic profile (figure 67), recorded at 214 nm, was similar to the one obtained by UPLC: some fractions of the secondary peaks were then analyzed by SDS-Page and western blot (figure 68). The only band detected in SDS-Page (the putative only protein present in the eluted fraction) were also detected by the anti-p24 monoclonal antibody employed in the western blot. This confirms again that the only band present in the purified sample is p24M, and the overall purity should be given by the sum of the two UPLC peaks.

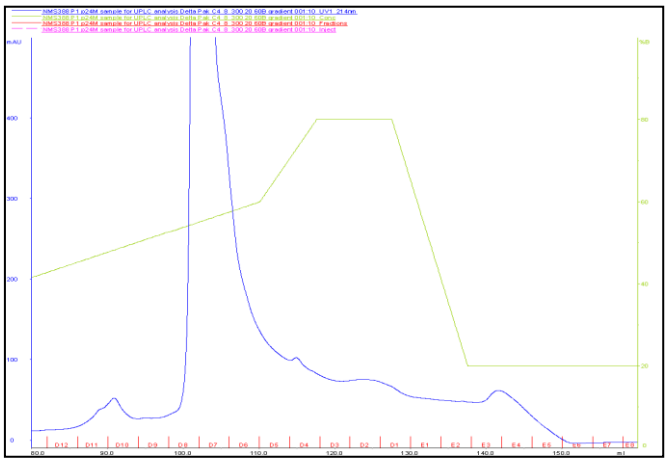
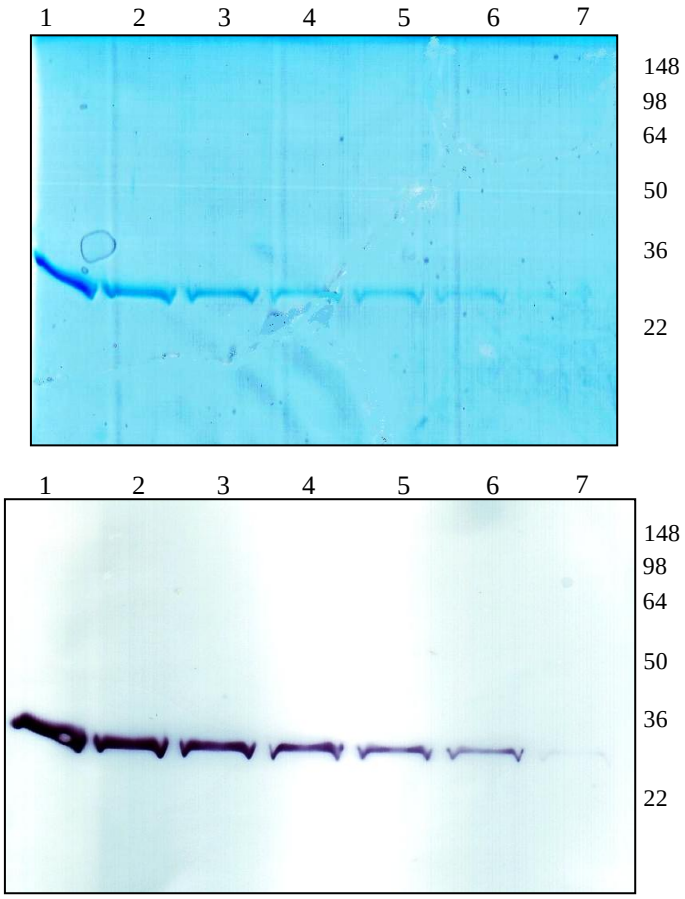


Figure 67.Semi-preparative HPLC analysis of one p24M lots.



Lane	Sample
1	Fraction D4
2	Fraction D3
3	Fraction D2
4	Fraction D1
5	Fraction E1
6	Fraction E2
7	Fraction E3

Figure 68. SDS-Page (above) and western blot (below) analysis of some eluted fraction from semi-preparative HPLC.

The protein was further characterized by mass spectrometry (MALDI-TOF); the molecular weight found in this analysis is in accordance with the molecular weight (27558 Da). The MALDI spectra of lot 2 and lot 5 are shown in figure 69.

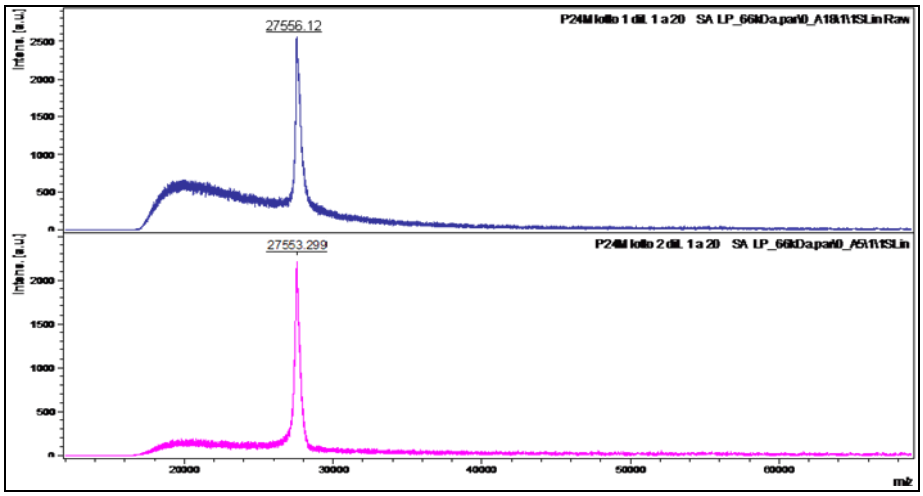


Figure 69. MALDI spectra of lot 2 (above) and lot 5 (below) of protein p24M.

p24M-*Mxe*GyrA-ELP36

The purification yields of p24M expressed by the pTME-*Mxe*GyrA-ELP90 constructs were about 16 milligrams per litre of fermentation, a value which is quite lower than that we would expect (around 40-50 mg/L).
Fusion protein p24M-*Mxe*GyrA-ELP90 is 933 residues long for a theoretical mass of 87788. 98 Da. Protein p24M, which is composed of 234 amino acids, covers only 25% of the all fusion protein and 32% if the molecular weight is considered. This means that three fourth of the expressed protein is lost because it does not constitute our target protein.
In order to increase the final purification yield, in literature there are few examples that demonstrate how varying appropriately the guest residues nature, it is possible to decrease the total number of repeated motifs in the ELP sequence, and thus shortening the fusion protein: Meyer D.E., 2001, Lim D.W., 2007
Among the different ELP molecules described in those articles, the one that appeared to us the most promising one was the [KV_7F-36] polypetide, the crucial difference with ELP[$V_5A_2G_3$] is the composition of the guest residues. The polar Lys residue is found to enhance the salt sensitivity of the whole ELP, but on the other hand its polarity has a negative effect on the hydrophobic collapse, increasing the transition temperature. To counterbalance the lysine effect on T_t, Phe, a very hydrophobic aminoacid has to be inserted. The ELP[KV_7F] of a determined mass need much less salt for precipitation than the corresponding ELP[$V_5A_2G_3$] of the same lenght, then a much shorter sequence can be employed: 36 repetitions of the model pentapeptide, instead of 90 were enough to have a transition temperature close to the desired one (Lim D.W., 2007).

p24M-*Mxe*GyrA-ELP[KV_7F-36] was the new fusion protein we desired to express. Cloning of the relate construct is described in material and methods.

p24M-*Mxe*GyrA-ELP[KV_7F-36] fusion protein is 657 residues long for a theoretical mass of 65708.66 Da. The p24M protein weight in this fusion protein is 36% considering the amino acid sequence and 42% considering the molecular weight. Other differences with the fusion p24M-*Mxe*GyrA-ELP90 are in the linker between the intein and the ELP, here made of seven repletion of the motif Gly-Gly-Gly-Ser: the reason for this substitution is that the previous linker, containing also an histidine tag, was called to be a

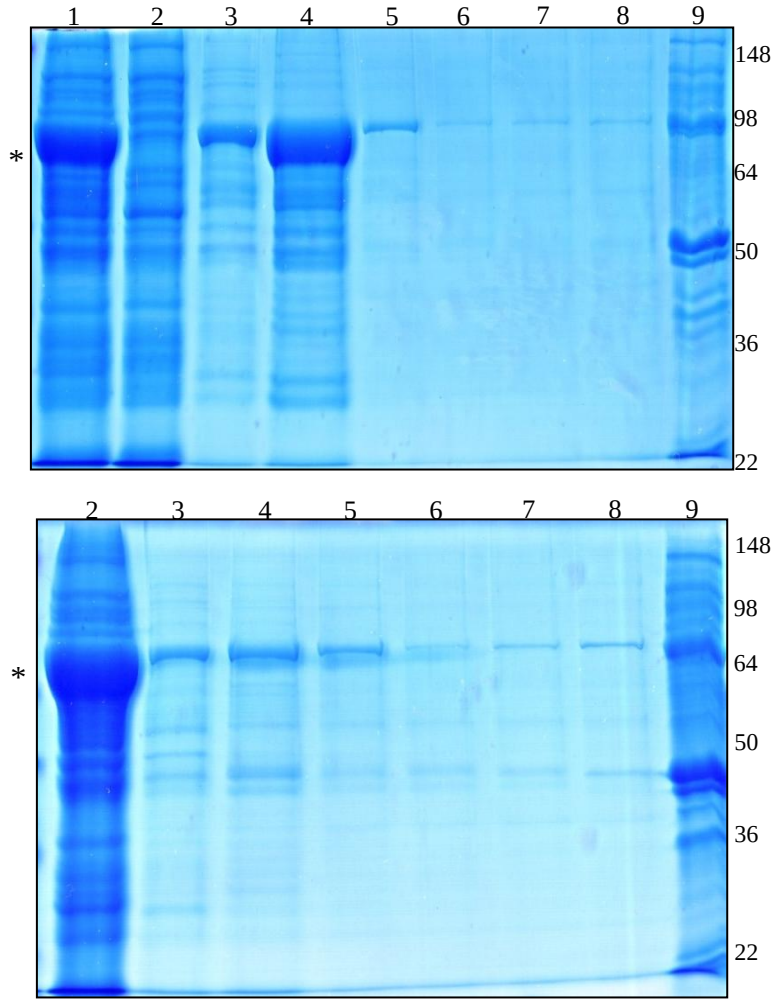
probable protease restriction site (Lim D.W., 2007). This new linker eliminates this site, keeping the same distance between intein and ELP than in the previous construct.

Preliminar experiments showed up that the new fusion protein was induced in a good yield. Since we aimed to express a soluble fusion protein, we explored two different conditions for protein expression:

A LB broth, 37 °C, induction with 1 mM, IPTG at 0.6 OD_{600nm}, 3h

B LBbroth, growth @ 25 °C, induction with 0.1 mM, IPTG at 0.6 OD_{600nm}, then 24 h @ 15 °C.

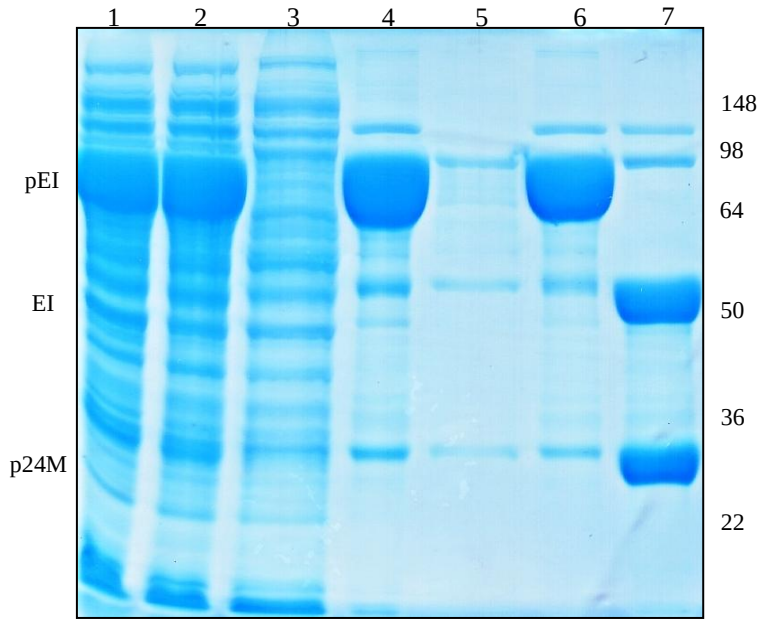
The protein was well expressed in both the conditions; while the protein expressed under condition B was completely soluble without using any urea containing buffer, see figure 70, the one expressed under condition A was completely solved only in urea 2 M. See figure 70.
We then decided to proceed with the B conditions for the expression of the target protein.



Lane	Sample
1	Total lysate
2	Soluble fraction
3	Urea 1M
4	Urea 2M
5	Urea 3M
6	Urea 4M
7	Urea 5M
8	Urea 6M
9	SDS 1%

Figure 70.Solubity test of p24M-Mxe GyrA -ELP36 expressed both in conditions A and B. * indicates the band corresponding to the fusion construct.

In order to find the best NaCl concentration to be used in the purification process, we explored a range of NaCl concentrations between 1.5 and 2.0 M. As all these concentration were sufficient to completely precipitate the protein, lower concentrations were tested: 1.1, 1.3, 1.5 and 1.7 M. Only with the last two concentrations no protein was lost: the chosen NaCl concentration for the precipitation process was then 1.5 M.
An SDS-Page analyzing a typical purification is shown in figure 71.

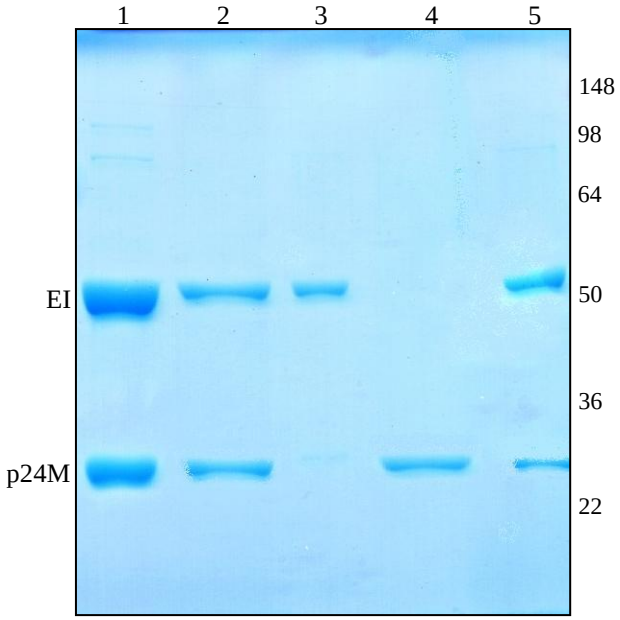


Lane	Sample
1	Total lysate
2	Soluble fraction
3	1st HS
4	1stCS
5	2nd HS
6	2nd CS
7	DTT cleavage

Figure 71. ITC purification and cleavage of p24M-*Mxe* GyrA-ELP36.pEI indicates the fusion p24M-Mxe GyrA-ELP36, EI indicates the tag.

The overnight cleavage was quantitative: the precursor species was nearly eliminated; after that, the sample containing the two separated fragments underwent a third ITC, with 1.5 or 2.0 M NaCl. The SDS-Page of figure 72

shows that this third ITC was not effective in separating the two species: in one case (1.5 M NaCl), not all the free tag precipitated, probably because its transition temperature was higher for intein-ELP tag than for the whole fusion construct. This is confirmed by the second case (2.0 M NaCl): the cleaved tag is all precipitated, but also a part of p24M is found in the precipitated fraction.



Lane	Sample	
1	DTT cleavage	
2	1.5 M NaCl	3rd ITC: supernatant
3		3rd ITC: pellet
4	2.0 M NaCl	3rd ITC: supernatant
5		3rd ITC: pellet

Figure 72.SDS-Page analysis of the unsuccesful separation of *Mxe* GyrA-ELP36 tag and p24M by a third ITC. EI indicates the position of Mxe GyrA-ELP36.

At the beginning, a limited amount of sample was processed on a gel filtration column (HiLoad 26/60 Superdex 75 prep grade, GE). The intein-ELP36 peak and the cleaved p24M peak were well resolved, even though their heights were quite low (figure 73).

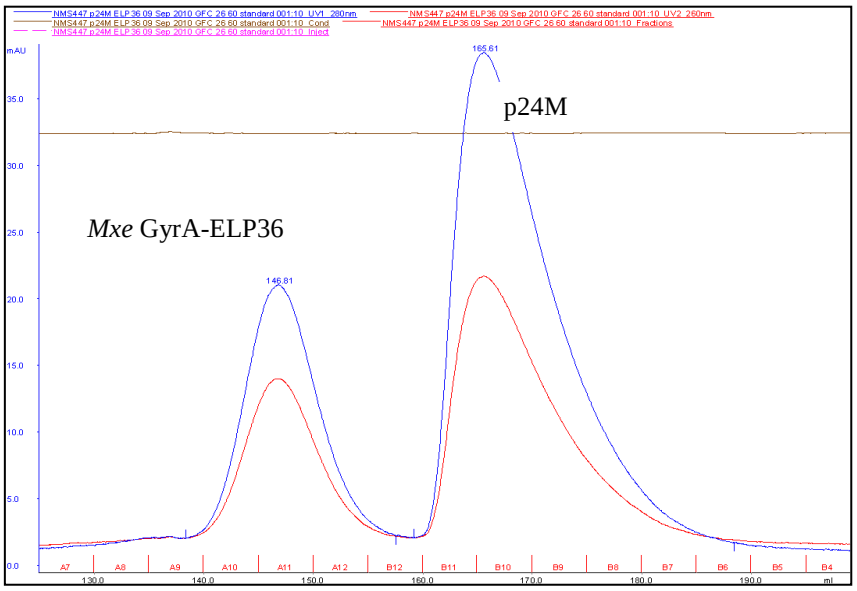


Figure 73.Chromatogram of SEC separation of the cleaved tag and p24M.

This purification strategy revealed its limits when it was applied on a higher scale (preparative): the two peaks merged and the two fragments were eluted together, as seen in figure 74.

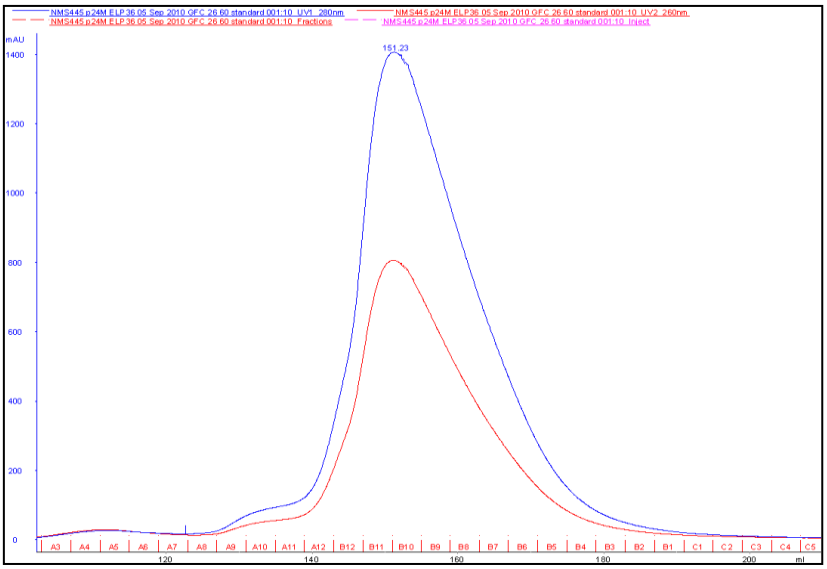


Figure 74.Cromatogram of unresolved SEC separation of the cleaved tag and p24M.

We then tried to employ other chromatographic techniques to separate the two fragments.
As first attempt, the sample was loaded on a reversed phase column (DeltaPak C4, see materials and methods) previously equilibrated with water/TFA and CH_3CN/TFA. Elution was performed with a linear gradient of 10-60% CH_3CN/TFA in 50 mL, but no separation of p24M from the protein from the tag was achieved (figure 75).

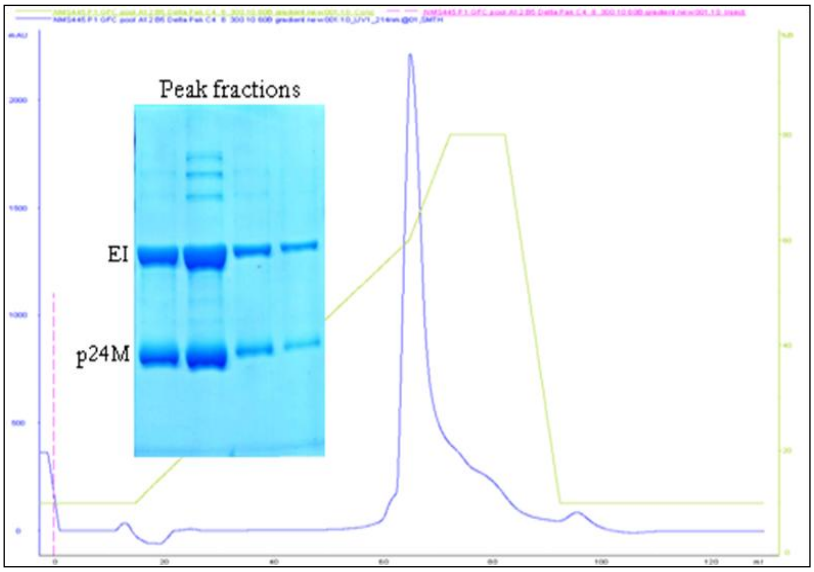


Figure 75.RPC chromatogram and SDS-Page analysis of the main peak fractions.

Another tentative to purify p24M was done with anionic exchange chromatography. The intein-cleaved sample was first dialysed against PBS, then diluted with 2 sample volumes of loading buffer and then loaded onto a Q Sepharose HP column. Elution was performed with a linear gradient of NaCl (0-1.0 M) in two column volumes. Surprisingly p24M protein did not bind to the resin and was collected in the flow through with a high purity grade. *Mxe*GyrA-ELP36 tag is instead eluted in the saline gradient. See figure 76.

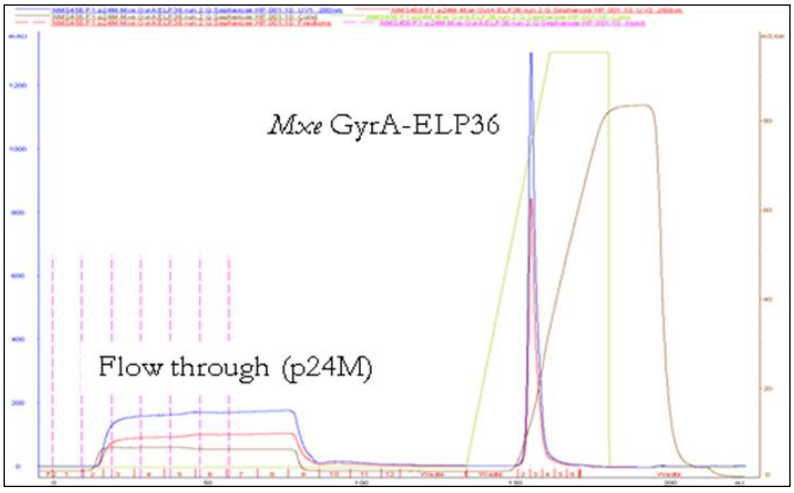


Figure 76.Chromatographic profile of AEC for separation of cleaved tag and p24M.

An SDS-Page analysis (figure 77) confirmed that the two species were separated during anion exchange chromatography. In the final pool only the first four flow through fractions were joined, as the last ones were contaminated by the intein-ELP36 tag and by some other impurities of high molecular weight (and probably a small fraction of uncleaved precursor). This phenomenon could be due to the large amount of protein species present in the sample, that overcame the column binding capacity or to the weak interactions between the resin and the tag: this could be seen from the fact that the intein-ELP peak was eluted as the conductivity started to grow.

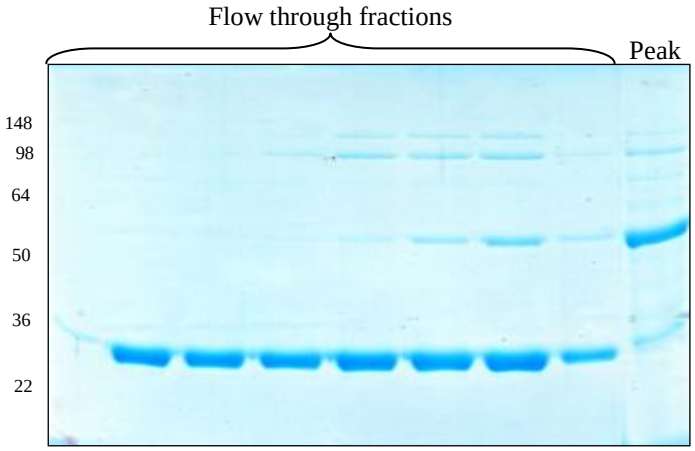
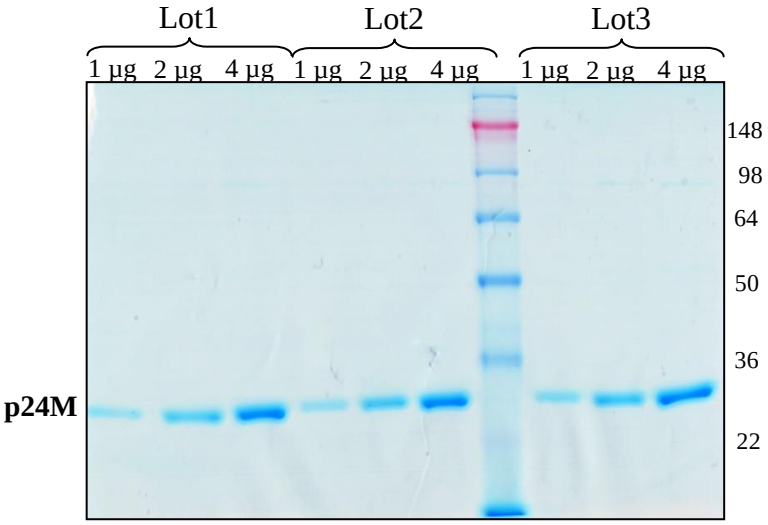


Figure 77.

Three batches from 1 liter of fermentation each were purified according to the following protocol:
Cell lysis
ITC (2 runs)
Dyalisis against PBS
Anion exchange chromatography

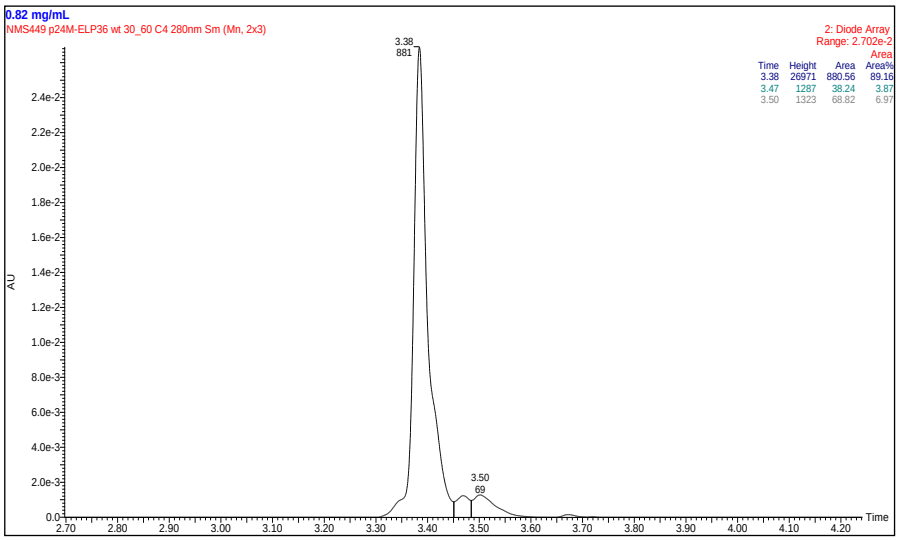
PURITY AND CHARACTERIZATION

The purity of the three lots was assessed by densitometry of SDS-Page bands (load 1, 2 and 4 µg of protein of each lot; software: ImageQuant TL, figure 78) and by UPLC-MS (column Acquity UPLC BEH300 C4 1.7µm, 2.1x100 mm, Waters; buffer A: H_2O, TFA 0.1% v/v, buffer B: CH_3CN, TFA 0.1% v/v; gradient 30-60% buffer B in 4', figure 79). The SDS-Page purities were very high (figure 78), while the one assessed by UPLC-MS was lower (figure 79); the chromatograms of all the three lots resembled the ones obtained with ELP90: a high and sharp peak, followed by a much smaller one. The main peak (molecular mass: 27337 Da) was in good agreement with the expected molecular mass of p24 (27330 Da), while the second peak, as in the case of ELP90, was found to have a mass higher of 117 Da than the expected p24M (27454 Da). The same hypoteses done for p24M from ELP90 can be done for p24M obatined with ELP36, but these hypoteses were not furhter investigated. Anyway we can confidently affirm that the second peak is due a specific form of p24M.



p24M from p24M-*Mxe* GyrA-ELP36			
Lot	**Purity SDS-PAGE(%)**	**Mean (%)**	**Std dev**
1	100	100	0.6
2	100		
3	99		

Figure 78. SDS-Page employed for the densitometric analysis (above) and table summary of the results (below)



	Lot 1	Lot 2	Lot 3
Purity 214nm (%)	88.93	88.66	84.45
Purity 280nm (%)	90.25	89.16	84.97
SDS-PAGE	100	100	99
Main peak (uma)	27337±1		
2ndary peak (uma)	27454 (+117)		
Theorethical	27330		

Figure 79. UPLC profile of p24M purified from p24M-*Mxe*GyrA-ELP36 (above) and table summary of the results (below)

The yields of the three lots of p24M obtained from the fusion construct p24M-*Mxe*GyrA-ELP36 are reported in table 14. As a comparison also the yield obtained from ELP90 constructs are reported. The dramatic four-fold increase than what obtained from p24M-*Mxe*GyrA-ELP90 is even larger than what was reported in literature (Lim D.W., 2007), and the final yields are now comparable with the typical yields of recombinant proteins purified with affinity chromatography. Also the decrease in NaCl concentration from 1.9 to 1.5 M in the precipitation step is noticeable.

p24M from p24M-*Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA(T3C)-ELP90			
	Lot	**Yield (mg/L)**	**Mean (mg/L)**
T3C	1	11,4	16,4±6,2
	2	14,4	
	3	23,3	
WT	4	22,3	16,6±5,0
	5	13,0	
	6	14,6	
p24M from p24M-*Mxe*GyrA-ELP36			
	1	78.1	69.0±6.4
	2	65.4	
	3	63.5	

Table 14. Purity and yield summary of p24M from p24M-*Mxe*GyrA-ELP36

p24M activity test

Protein p24M from our expression and purification method was tested in an immunodiagnostic kit manufactured by DiaSorin. This kit is the Liaison® HIV-Ag that is used for the detection of the HIV viral capsid on human sera. This kit uses a commercial p24 as a calibrator and positive control. In this experiment the immunoreactivity of the p24 used in the Liaison® HIV-Ag kit was compare with one lot of our p24M produced with *Mxe* GyrA-ELP90 tag and one lot of p24M from p24M-*Mxe* GyrA-ELP36 fusion constructs.
The test was conducted in triplicate at eight different concentrations and the response, in RLU, was plotted versus the concentration, figure 80.
Results shows that both batches of "our" p24M are immunoresponsive: the RLUs follows the concentration increase, but the response is lower, about half in both cases, than p24M already used in the kit. This could be due to some differences in the protein sequence, that lower the affinity factor for the antibody, or to an incorrect protein folding, which shield some epitops, or to aggregate forms. All these hypotheses were not investigated.

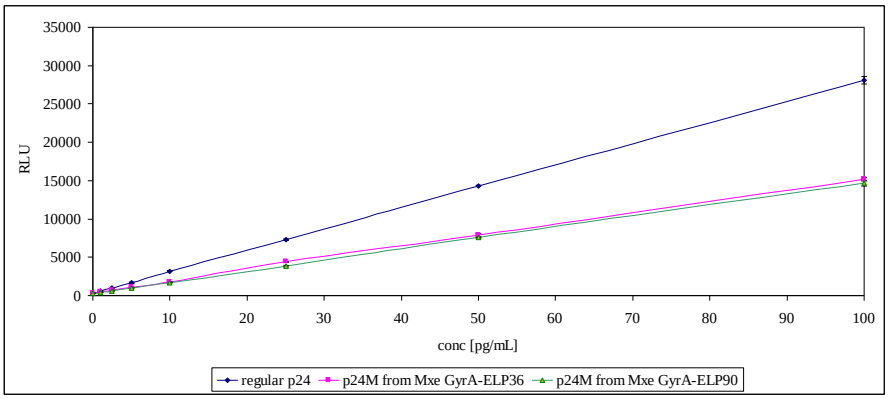
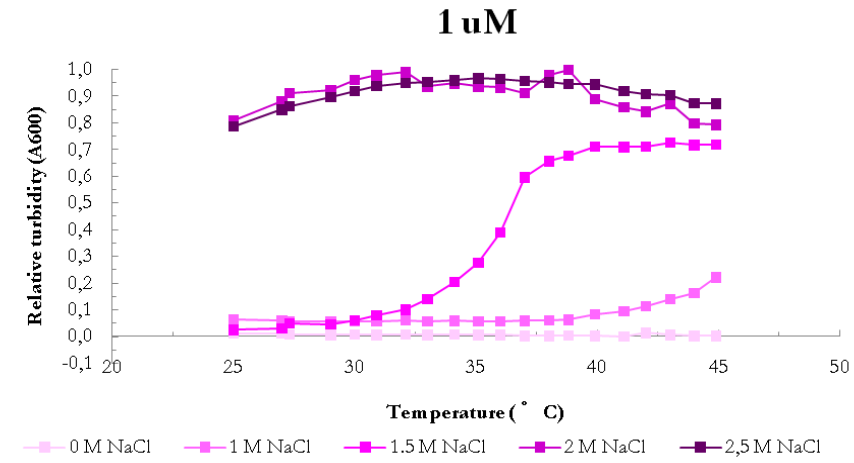


Figure 80. Plot of RLUs vs concentration of p24M (blue: standard p24M, purple: p24M from ELP36, green: p24M from ELP90)

INVERSE TRANSITION CURVES

At this point we studied the *Mxe*GyrA-ELP90 tag phase behaviour as reported in "materials and methods".

Fusion proteins *Mxe*GyrA-ELP90 and p24M-*Mxe*GyrA-ELP90 were incubated at fixed protein concentrations (1, 10, 50 and 100 μM) and fixed NaCl concentrations (0.0, 0.5, 1.0, 1.5, 2.0 and 2.5 M). Protein-salt solutions were heated from 25 to 45 °C.
Protein precipitation was assessed by measuring the UV absorbance at 600 nm (turbidity), then plotted it vs the temperature (figure 81). The results coincide with what was expected: keeping the tag concentration constant, the NaCl has a strong effect on the transition temperature: the more the NaCl, the lower the T_t. Moreover, increasing protein concentrations lower the transition temperature. As an example we can take the curves of 1.5 M NaCl at 1, 10 and 100 μM protein concentration: in the first case (1 μM protein concentration) the curve has a transition point around 37 °C, in the second (10 μM protein concentration) the protein already precipitates around 27 °C, while in the last case (100 μM protein concentration) we cannot assign a transition point, as at 25°C the turbidity is already approximating the maximum.
Two other things are worth to say: in presence of 2.0 and 2.5 M NaCl the protein is already precipitated at 25 °C, no matter its concentration; the other thing is that after reaching a maximum the precipitation curves tend to descend: what was hypotesized is that it could be due to sedimentation of the precipitated proteins (though the solution was stirred before every measurement) or to the size of the aggregates (as the aggregates become larger, they were not seen as turbidity, but as flakes).
In figure 82 we made a comparison between some curves obtained with *Mxe*GyrA-ELP90 and others obtained in the same boundary conditions, but using p24M-*Mxe*GyrA-ELP90. As expected, we see that the presence of p24M in the fusion has a negative effect on the trasitions temperature, enhancing it of several degrees (estimated in 8-9 °C). This increase is seen independently of the protein and NaCl concentrations.



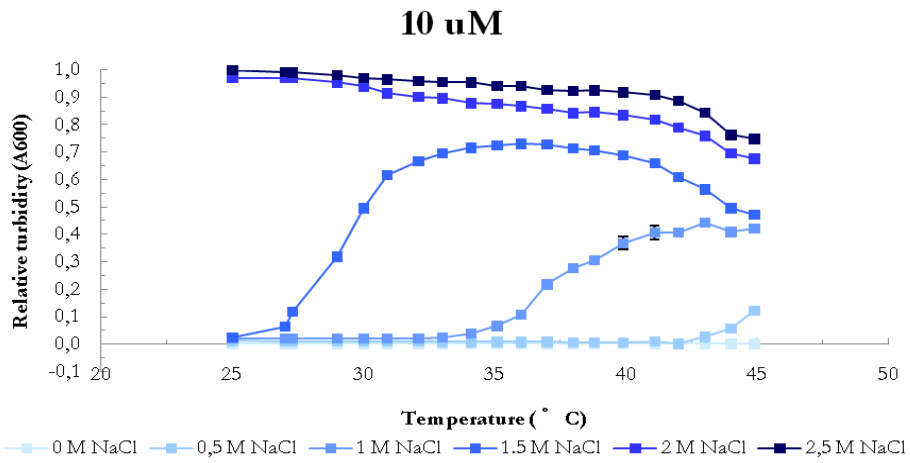


Figure 81. (Continues on the next page)

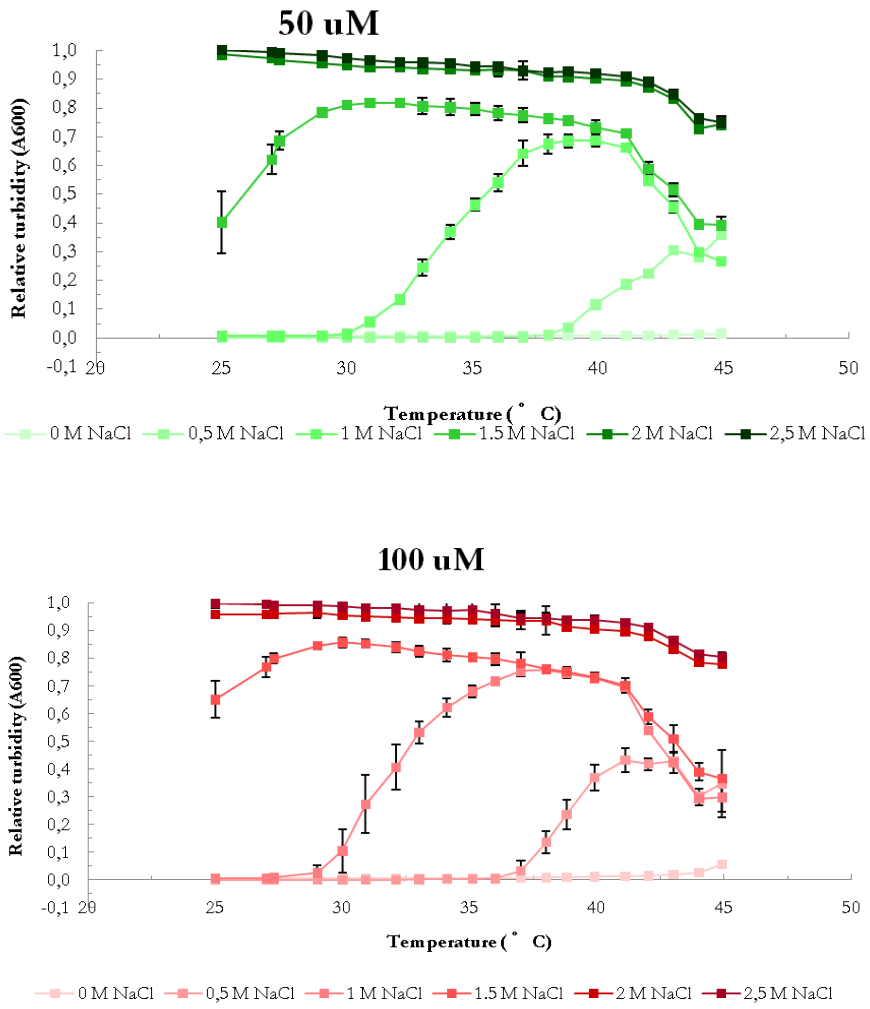


Figure 81. (from prev. page) Plots of the turbidity vs temperature at different protein and NaCl concentration.

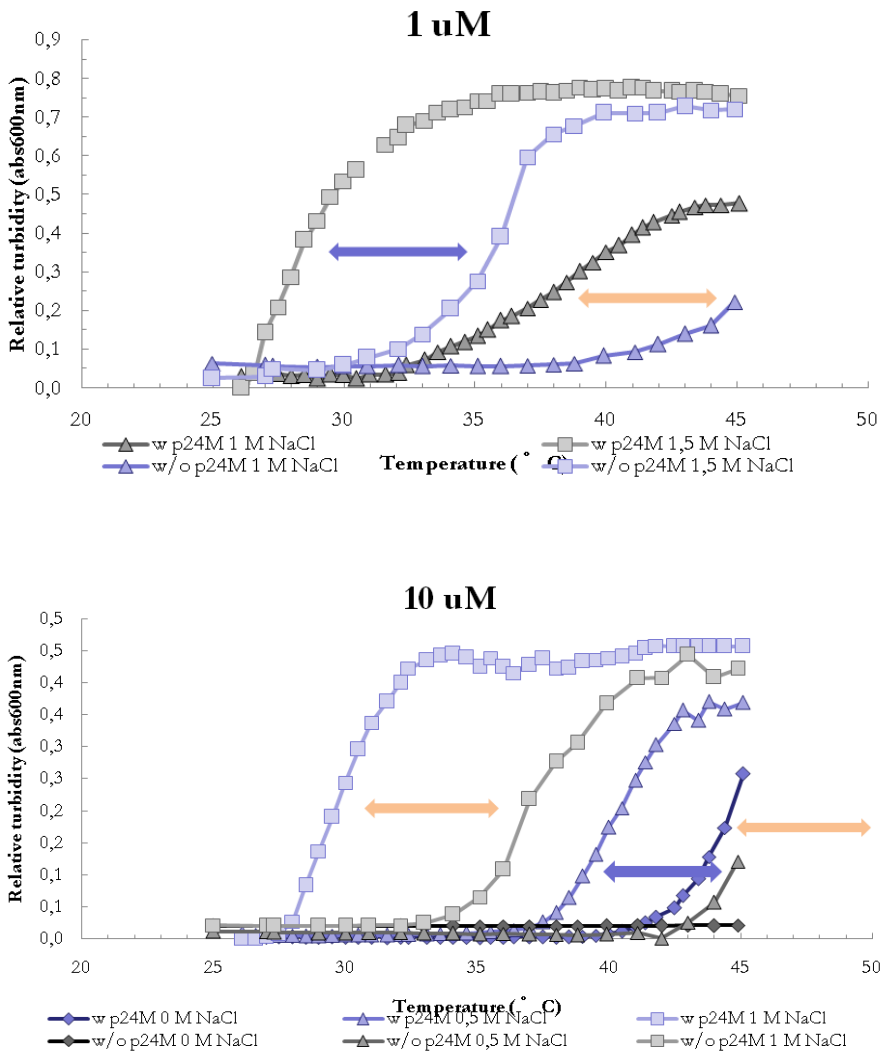


Figure 82. Comparisons of some transition curves of p24M-*Mxe* GyrA-ELP90 and *Mxe* GyrA-ELP90.

INTEIN KINETICS

The next study was conducted on the intein cleavage rate. The fusion construct p24M-*Mxe* GyrA-ELP36 was purified as already described and before cleavage it was divided into two aliquots.
The cleavage reaction was followed for 20 hours covering a range of DTT concentration of one order of magnitude (5 and 50 mM, respectivley). The results obtained by densitometric analysis of the Coomassie-stained SDS-Page are summarized in figure 83: the total volume of the three bands (fusion construct, cleaved p24M and cleaved *Mxe* GyrA-ELP36 tag, respectively) was considered 100% (the minor contaminant bands were not taken into account), and was built up by the relative contribution of each band. It can be seen that from the same starting point, the decrease of the fusion construct band contribution is much faster for the sample with 50 mM DTT than that where only 5 mM was addded. After 20 hours the cleavage of the 50 mM DTT sample went nearly to completion, while the starting fusion protein contribution for the sample with 5 mM DTT was still around 30%.
In figure 84 we also plotted the reaction efficiency vs time. The reaction efficience was calculated as follows: we considered that 42% of the calculated molecular weight of the whole fusion construct comes from p24M and we stated a 100% reaction efficiency when the p24M band constitued the 42% of the sum of the three bands. It can be noticed that at the end of 20 hours, the sample treated with 50 mM DTT overcame an efficiency of 90%, while 5 mM DTT was enough to reach about 65%. The different cleavage rate can be seen from the beginnig of the reaction.

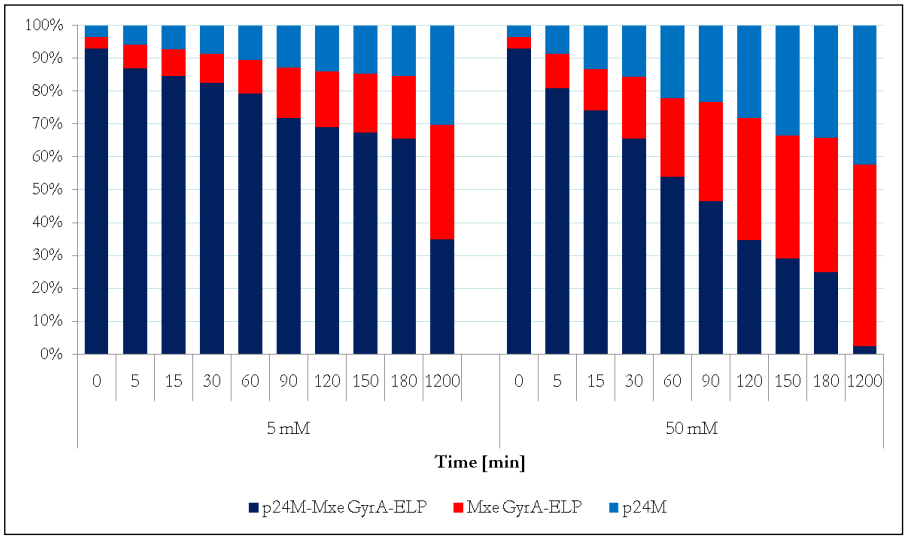


Figure 83. Graphical representation of the band distribution in the cleavage reactions.

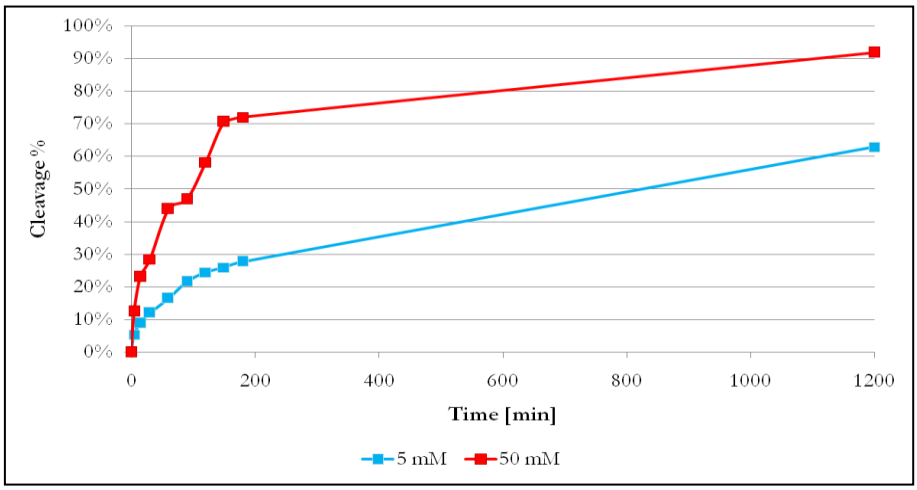


Figure 84. Cleavage efficiency vs time of p24M-Mxe GyrA-ELP36 with two different DTT concentration (5 and 50 mM, respectively)

CT_MONO

Ct_mono-*Mxe* GyrA-ELP90

The target species used in these experiments is a quadruple antigen deriving from the fusion of four different antigenic peptides used as a mix in an ELISA test by Savyon (Israel) for the diagnosis of *Chlamydia Trachomatis* currently on the market.
The sequences of these peptides are shown in figure 85:

Ct4A	**IFDTTLNPTIAGAGDVK**	**MW 1733**
Ct4B	**VDITTLNPTIAGCGSVAK**	**MW 1760**
Ct4C	**CVFDVTTLNPTIAGAGDVK**	**MW 1921**
Ct4D	**LAEAILDVTTLNPTITGKAVVSK**	**MW 2355**

Described in patent US6699678

Figure 85. Primary equence of the four peptides in the Savyon Clamydia trachomatis ELISA test.

The design of the quadruple antigen was based on:
Introduction of flexible and hydrophilic linkers between the four peptides. In particular the linkers were formed by 6 amino acids consisting in the pair Asp (or Glu)-Lys flanked by two pairs of Gly residues.
Proper order of the four peptides within the chimeric construct, in order to avoid formation of foreign aminoacid sequence, potential source of aspecific cross-reactivity.

The four peptides were then assembled in this order: Ct4D-Ct4C-Ct4A-Ct4B. This final antigen, named Ct_mono, consists of 107 total residues and its sequence is reported in figure 86.

```
GGDKGGLAEAILDVTTLNPTITGKAVVSKGGEKGGCVFDVTTLNPTIAGAGDVK
GGDKGGIFDTTLNPTIAGAGDVKGGDKGGVDITTLNPTIAGCGSVAKGGDKGG
```

Figure 86. Primary structure of Ct_mono peptide.

Ct_mono was fist cloned in the pTME vector to express the Ct_mono-*Mxe*GyrA-ELP90 fusion protein. Two different fermentation conditions were tried: (A) at 37 °C for 3 hours with 1 mM IPTG and (B) at 15 °C for 24

hours with 0.1 mM IPTG. The induction protein test, figure 87, showed that condition B produced a fusion protein band much larger than condition A. Moreover, the fusion protein expressed in condition B is totally soluble, figure 88.

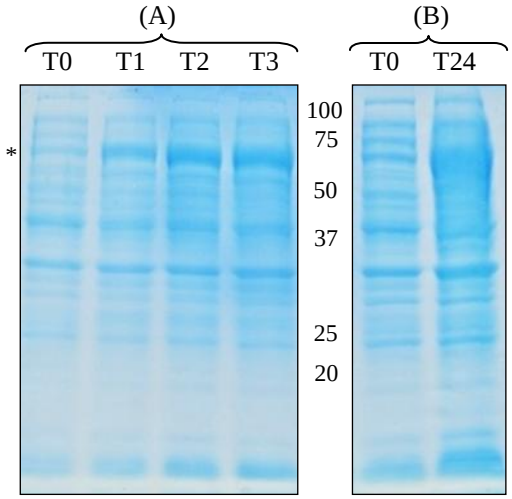
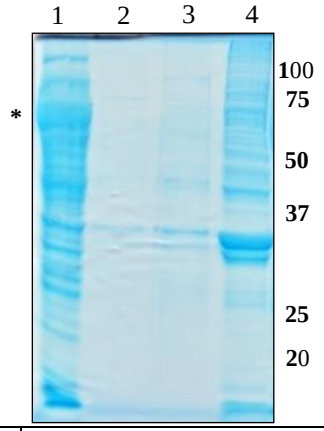


Figure 87. Fermentation control of Ct_mono-Mxe GyrA-ELP90. * indicates the corresponding band of the fusion construct.

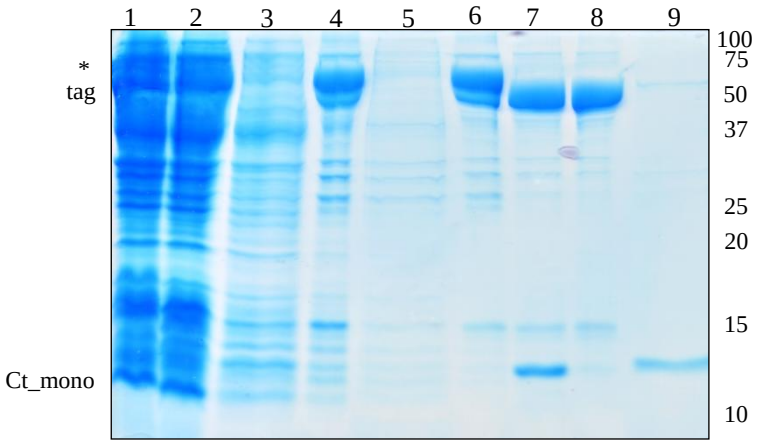


Lane	Sample
1	Soluble fraction
2	Sample soluble in urea 1 M
3	Sample soluble in urea 8 M
4	Sample soluble in SDS 1%

Figure 88. Solubility test of Ct_mono-Mxe GyrA-ELP90. * indicates the position of the fusion construct band.

After exploring a large range of NaCl concentrations (0, 0.5, 1.0, 1.5, 2.0, 2,5 M) we found out that the lowest concentration to get the fusion construct precipitated at ambient temperature was 2.0 M. This salt concentration is higher than the one needed in the p24M case, even though p24M is a larger protein then Ct_mono. This behaviour could be due to the higher hydrophobicity of the Ct_mono construct.

An SDS-Page monitoring a typical ITC purification and cleavage processes is shown in figure 89. After two cycles of ITC, fusion protein is purified to a quite high degree of purity, lane 6 of figure 89. The following DTT cleavage, which is performed with 50 mM DTT and keeping the sample under magnetic stirring at room temperature overnight, went nearly to completion, line 7 of figure 89. The third ITC process was able to separate the *Mxe*GyrA-ELP90 tag from the target protein Ct_mono, line 8 and 6, figure 89.



Lane	Sample
1	Total lysate
2	Soluble fraction
3	1st HS supernatant
4	1st CS supernatant
5	2nd HS supernatant
6	2nd CS supernatant
7	DTT cleavage
8	3rd HS pellet
9	3rd HS supernatant

Figure 89. SDS-Page monitoring a typical ITC purification and cleavage. * indicates the band corresponding to Ct_mono-*Mxe* GyrA-ELP90 fusion construct, tag **indicates the position of the band of *Mxe* GyrA-ELP90 tag.**

The purification did not show any critical point: the whole fusion construct is precipitated during the hot spins, and re-suspended during the cold spins (some soft sonication may be needed to completely solvate the re-suspended pellet during the hot spins); the DTT cleavage did not give any problem and also the third hot spin is able to separate the two cleaved species. The only note that could be done is that in the cold spins a smaller band can be seen under the main band of the entire fusion construct, at the corresponding molecular weight of the cleaved tag. This should comes from an undesired *in vivo* and/or *in vitro* cleavage phenomenon as already reported in literature. This phenomenon was not investigated more deeply.

After the third hot spin, the protein was found in a buffer condition unfavourable for storage (the sample is in buffer NaCl 2.0 M) and was still not pure enough for our scopes; for these reasons, it underwent a supplementary step, which consisted in a dialysis in PBS 10 mM followed by a chromatographic polishing step by reversed phase chromatography (see the chromatogram in figure 90). In the RPC a very high and sharp peak is eluted, with only minor impurities visible earlier in the gradient. This step is effective in eliminating the minor contaminants and brings the sample in a buffer suitable for lyophilization. The eluted fractions corresponding to the main peak were pooled together; and the pool, after freezing at -80°C, was lyophilized for 24 hours.

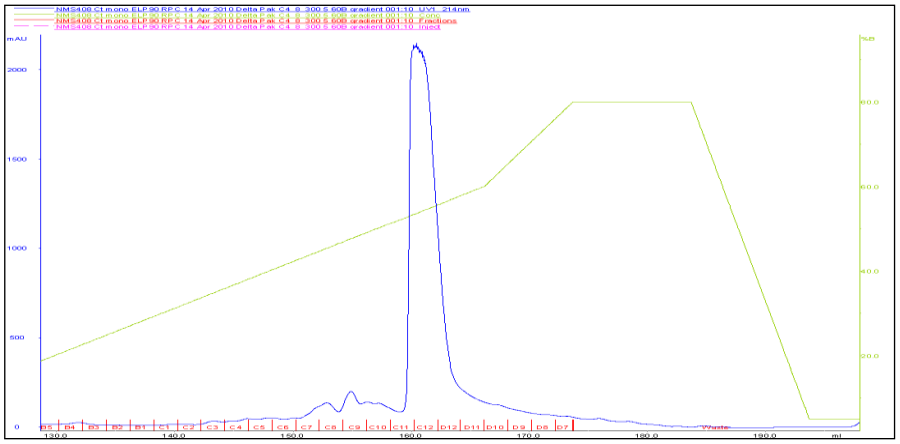


Figure 90. Chromatographic profile of RPC.

Three lots of this protein were expressed and purified according to protocol already reported in “Materials and methods”. In table 15 the final yields are

reported; although not very high, they can be considered acceptable, taking in account the small size of the target peptide (less than 12 kDa). In the same table, the purity values of the three lots on SDS-PAGE are also reported: a 100% purity for all the preparation batches was found. See figure 91.

Lot	Yield (mg/L)	Mean (mg/L)	Std dev	Purity (%)	Mean (%)
NMS408	4,24	5,4	2,1	100	100
NMS409	4,13			100	
NMS418	7,80			100	

Table 15.

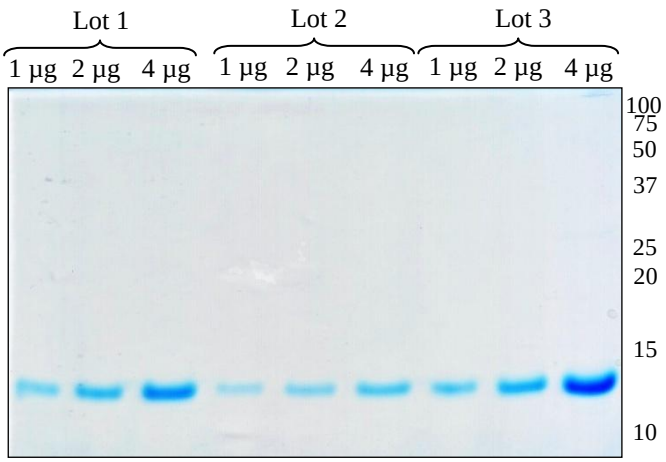
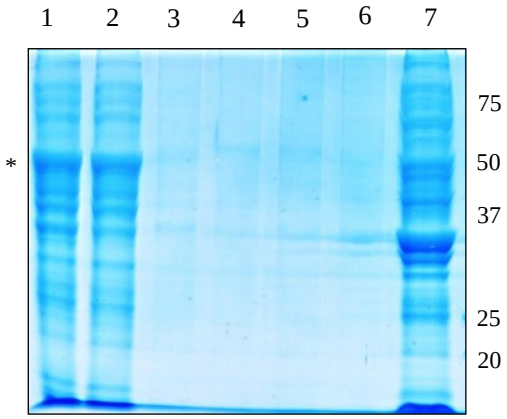


Figure 91. SDS-Page analyzed by densitometry.

Ct_mono-*Mxe* GyrA-ELP36

Given the good results obtained in the case of p24M protein by substituting ELP90 with a different (cationic and shorter) ELP36, we decided to apply the same strategy to Ct_mono antigen. The gene coding for Ct_mono was cloned into the pET24-MCS-Intein-ELP(KV7F-36) vector, giving a fusion construct of a calculated mass of 49892.09 Da (of which Ct_mono antigen contributes for 11512.54 Da, equal to 23%; in the case of ELP90 Ct_mono contribution was only 16.3%). The fermentation protocol of Ct_mono-*Mxe*GyrA-ELP90 was employed, giving a perfectly soluble fusion protein, figure 92.



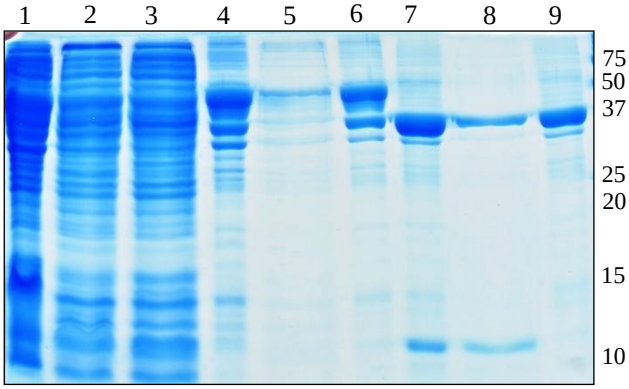
Lane	Sample
1	Total lysate
2	Soluble fraction
3	Sample soluble in urea 1 M
4	Sample soluble in urea 2 M
5	Sample soluble in urea 3 M
6	Sample soluble in urea 4 M
7	Sample soluble in SDS 1%

Figure 92. Solubility test of Ct_mono-*Mxe* GyrA-ELP36

NaCl concentrations ranging from 0 M to 2.5M were explored for the precipitation of the fusion protein in the hot spin: while in the case of p24M the NaCl concentration necessary for the precipitation of the construct

dropped from 1.9 M with ELP90 to 1.5 M with ELP36, in the case of Ct_mono, the salt concentration only decreased from 2.0 M with ELP90 to 1.9 M with ELP36.

The purification was then carried out using 1.9 M NaCl in the hot spins. Figure 93 shows a typical purification scheme for Ct_mono-*Mxe* GyrA-ELP36: the first two ITCs and the DTT cleavage outcomes were similar to those obtained with ELP90. The third ITC, differently from ELP90, did not separate completely the two cleaved species: *Mxe*GyrA-ELP36 tag did not completely precipitate and remained in the soluble fraction, lanes 8 and 9, figure 93. For this reason a final polishing step was again necessary for the purification of Ct_mono from *Mxe*GyrA-ELP36. We proceeded as we did for the ELP90 case: dialysis in PBS, to lower the NaCl concentration, followed by a reversed phase chromatography to isolate the Ct_mono antigen. The chromatogram was not as clean as it was in theELP90 case, but showed two main peaks and some other minor peaks due to impurities, figure 94.



Lane	Sample
1	Total lysate
2	Soluble fraction
3	1st HS supernatant
4	1st CS supernatant
5	2nd HS supernatant
6	2nd CS supernatant
7	DTT cleavage
8	3rd HS supernatant
9	3rd HS pellet

Figure 93. SDS-Page analysis monitoring a typical purification and DTT cleavage of Ct_mono-*Mxe* GyrA-ELP36.

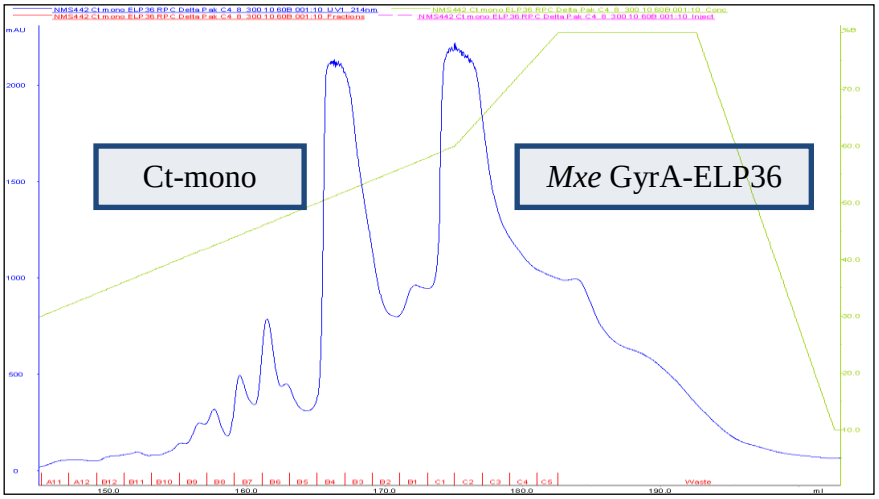
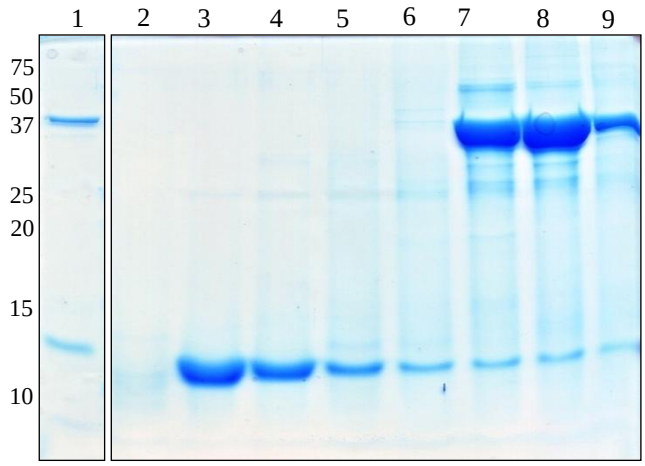


Figure 94. RPC chromatographic profile of 3rd spin supernatant after DTT cleavage.

The collected fractions corresponding to the main peaks were analysed by SDS-Page and HPLC, see figures 95 and 96. Electrophoretics analysis allowed an assignment of the peak identities: first peak is the target protein Ct_mono, the second is the *Mxe*GyrA-ELP36 tag. After RPC step, the Ct_mono pool was lyophilized and stored at -30 °C. Also for this construct three fermentation and purification batches were done.



Lane	Sample
1	Load
2	Fraction B5
3	Fraction B4
4	Fraction B3
5	Fraction B2
6	Fraction B1
7	Fraction C1
8	Fraction C2
9	Fraction C3

Figure 95. SDS-Page analysis of the most inteeresting fractions of RPC polishing step.

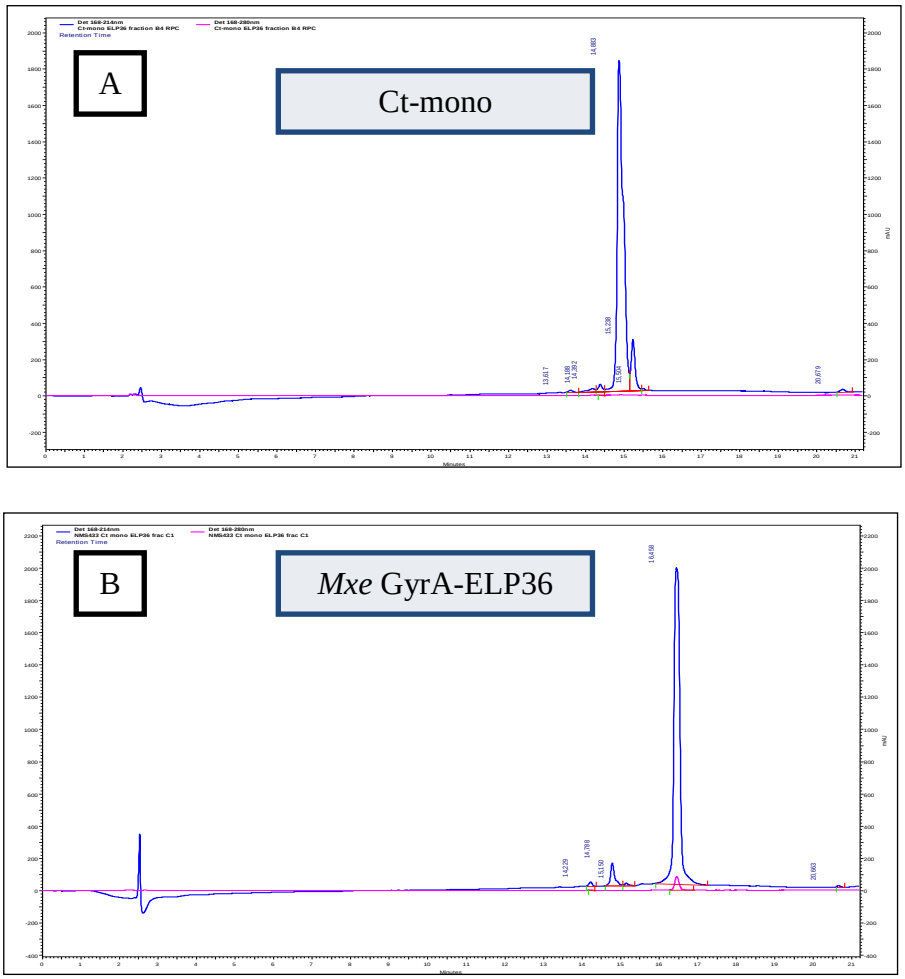


Figure 96. A: HPLC profile of Ct_mono peak. B: HPLC profile of the cleaved tag peak.

Chromatographic conditions:

Column: Jupiter 5u C5 300A 4.6*150 mm (Phenomenex)
Buffer A: H_2O/TFA 0,1%
Buffer B: CH_3CN/TFA 0,1%
Starting Buffer B%: 10%
Gradient: 10% B in 0,1'
100% B in 15'
100% B in 7'
10% B in 4'

Table 16 shows the purity on SDS-Page (figure 97) and the final yields of the three purified batches together with values obtained from ELP90 construct. It is worth noticing that despite a not perfect resolution was obtained in the chromatographic run, the SDS-Page analysis indicated that a good separation was achieved, figure 95. This consideration is also supported by the HPLC data, figure 96: only a negligible contamination was detected in the purified Ct_mono pool.

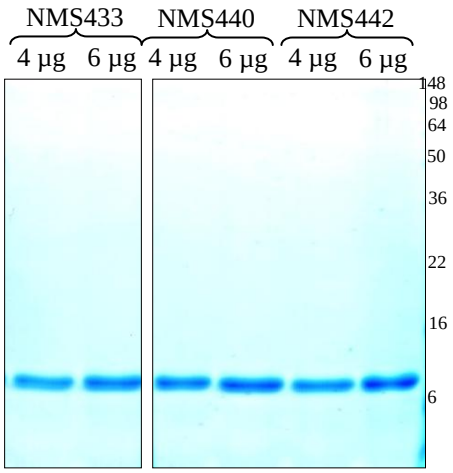


Figure 97. SDS-Paga analysis of the purified Ct_mono from Ct_mono-*Mxe*GyrA-ELP36

	Lot	Yield (mg/L)	Mean (mg/L)	Std dev	Purity (%)	Mean (%)
ELP36	NMS433	6,2	7,5	2,2	100	101
	NMS442	6,3			100	
	NMS458	10,0			100	
ELP90	NMS408	4,24	5,4	2,1	100	100
	NMS409	4,13			100	
	NMS418	7,80			100	

Table 16.

The yields obtained for Ct_mono from Ct_mono-*Mxe*GyrA-ELP36 are showed an increase of 39% compared to those obtained with Ct_mono-*Mxe*GyrA-ELP90. This percentage increase is quite good, but the absolute values are much lower than those found for p24M, and even of thioredoxin, which has a comparable mass.

IMMUNOCHEMICAL EVALUATION OF Ct_mono ANTIGEN.

The Ct_mono immunoresponse was evaluated by an immunochemical assay as described in the "materials and methods" and compared with a mix of the four peptides of the Savyon kit. Figures 98 and 99 shows the results: both compounds obtained from Ct_mono-*Mxe*GyrA-ELP90 and Ct_mono-*Mxe*GyrA-ELP36 are positively responsive to positive sera, while all the negative sera have low immunochemical response. Comparison with the "free" peptides shows that the response of the positive samples is very ofter higher for Ct_mono; some samples recognized as negative by the free peptides give now a positive response; the negative samples give slightly higher response with Ct_mono, but no false positive are found.

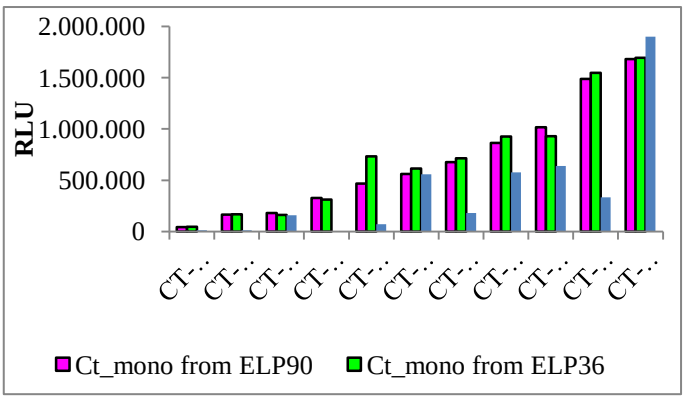


Figure 98. Immunochemical evaluation of Ct_mono by a panel of positive sera.

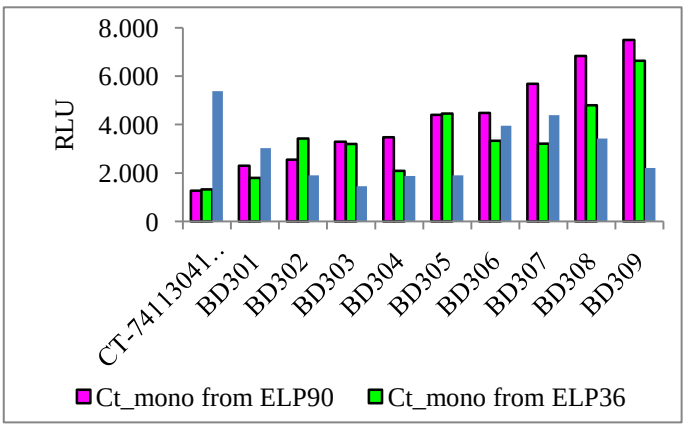


Figure 99. Immunochemical evaluation of Ct_mono by a panel of negative sera.

DISCUSSION

The combination of elastin-like polypeptide and intein tested in this work has demonstrated to be a valid alternative to standard chromatographic techniques in the recombinant protein purification field. Three different model proteins were chosen for expression and purification via inverse transition cycling: thioredoxin, a small soluble protein currently used as fusion partner for insoluble proteins; p24, a structural protein of HIV-1, and a chimeric polypeptide antigen from major outer membrane protein (MOMP) of *Chlamydia trachomatis*. These three proteins were purified with the contribution of this method and biochemically and immunometrically characterized.

Two different elastin-like polypeptide were tested: ELP[$V_5A_2G_3$-90] and ELP[KV_7F-36] in combination with *Mxe*GyrA intein. A point mutation (T3C) in the intein sequence, claimed to decrease the amount of unwanted cleavage, was also evaluated.

Thioredoxin was easily purified from two fusion constructs (Trx-*Mxe*GyrA-ELP90 and Trx-*Mxe*GyrA(T3C)-ELP90) without the need of any chromatographic technique. Inverse transition cycling and intein cleavage combination was able to achieve the target protein from the whole construct in an active form, as demonstrated by the activity assay. Also p24 and Ct_mono fusion proteins were easily purified with only two inverse transition cycles, but after intein cleavage the isolation of the two target species was not straightforward. In these two cases the use of the transition phase property of elastin-like polypeptide fused tag was not enough to obtain the target species at acceptable levels of purity and yields.

Different expression conditions were explored, and the ones which gave a soluble construct were preferred: in case the fusion construct was insoluble, the purification in presence of a chaotropic agent, such as urea 4 M was carried out. In these conditions, while no detrimental effect was observed in the ITC purification (precipitation occurred immediately after salt addition), cleavage of the intein to release the target proteins did not occur efficiently, most likely because of unfolding and inactivation of the intein Buffer containing urea 1M was sufficiently mild to keep the intein active for cleavage. This is a fundamental requirement to obtain the target protein: the whole construct should be soluble at most in 1 M urea.

The precipitation conditions were optimized not varying temperature with heating cycles, but keeping the temperature at ambient level, to avoid the risk of thermal stress. Phase transition was triggered by addition of NaCl. The concentration range of salt for precipitation was between 1.5 and 2.0 M, which is quite far from physiological conditions, but did not negatively

affect at all the structure and activity of the proteins taken as models. Following the Hofmeister ion series other salts could be even more effective in causing hydrophobic collapse of elastin-like polypeptides (i.e. ammonium sulphate), but their use was not explored in this work.
The resolvation transition phase was achieved by placing the fusion construct at low temperature (+4 °C) in low ionic buffer and, as this thermal change should not stress the proteins.
The intein insertion in the fusion construct allowed the avoidance of any protease for target protein release from the tag, differently from what would have been with affinity tags. Protease use would have increased process cost and time, introducing also the risk of unspecific cleavage. In our work, the release of the target protein was efficiently achieved in nearly physiological conditions independently from the intein flanking sequences (exteins).
Mutated T3C intein was reported in literature (Cui C., 2006) to increase up to 10-fold the expression and purification yield by elimination of the unwanted auto-cleavage. In our cases the differences in expression and purification yields were absolutely negligible, and putative small auto-cleavage bands were detected even when T3C intein was employed. The attempt of eliminating it by acting on the aminoacid directly involved in the cleavage mechanism was unsuccessful.
One of the main drawbacks of this method is that after the cleavage the main part of the expressed construct is constituted by the tag, while the target protein is a minor portion of the whole fusion. This problem has been partially overcome by employing a smaller elastin-like polypeptide: instead of having 90 repetitions of the pentapeptide VPGXG (which means 450 aminoacids only for the ELP), a 36-repetitions one was used. As reducing the size means increasing the transition temperature (or the salt concentration to be reached for precipitation), such small ELP should have a different guest residues composition than the ELP90: ELP[$V_5A_2G_3$-90] was substituted ELP[KV_7F-36]. The advantageous effects were:

- higher solubility of the whole construct;
- lower salt concentration to be reached for the ELP precipitation (lysine presence in the guest residue position increased the salt sensitivity, even though this effect was much more remarkable in the case of p24M than of Ct_mono);

- higher yields. This is particularly true for the p24M case, for which the increase was 4-fould, and the purification yield reached valued usually found with affinity chromatography.

In light of these results, we believe that this technology might represent a viable alternative protocol for the purification of soluble proteins, able to deliver products with comparable yield and purity of those obtained with the currently used purification procedures. The typical purification process involves a first capture step, followed by one or more further chromatographic steps; the overall process may require several days to obtain a purified protein. In contrast, the herein presented method, based on inverse transition cycling (ITC) combined with intein cleavage, shows noticeable advantages, like the low cost of the reagents and equipments and the easy scalability from bench scale to large scale production only by enlarging the volumes.

In other cases, like Ct_mono, it was the only way to obtain the polypeptide: its primary sequence is too large to be chemically synthesized with standard techniques, and founds some difficulties to be expressed by a classical expression system. In this borderline situation the fusion of Ct_mono with such a tag was the only viable way to have the antigen purified.

Further studies need to be conducted, especially on

- Design of shorter and more salt-sensitive ELPs, in order to: (a) further improve the yield, as the fusion protein will be less demanding for cell expression, (b) facilitate purification and make it cheaper.
- Use of different salts to optimize the inverse transition cycling step.
- Use of different inteins, particularly those able to release the target protein in native form, without any additional amino acid.

REFERENCES

Camarero, J.A., Kwon, Y., and Coleman, M.A. "Chemoselective attachment of biologically active proteins to surfaces by expressed protein ligation and its application for 'protein chip' fabrication". *J. Am. Chem. Soc*. 126 (2004), 14730 – 14731.

Cho Y, Zhang Y, Christensen T, Sagle LB, Chilkoti A, Cremer PS. "Effects of Hofmeister anions on the phase transition temperature of elastin-like polypeptides." *J Phys Chem B* 112 (2008): 13765–13771.

Chong S, Mersha FB, Comb DG, Scott ME, Landry D, Vence LM, Perler FB, Benner J, Kucera RB, Hirvonen CA, Pelletier JJ, Paulus H, Xu MQ. "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element." *Gene* 192 (1997): 271-281.

Chong S., Shao Y., Paulus H., Benner J., Perler F.B., Xu M.Q. "Protein splicing involving the Saccharomyces cerevisiae VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system." *J Biol Chem* 271 (1996): 22159-22168.

David, R., Richter, M. P., and Beck-Sickinger, A. G. Expressed protein ligation. Method and applications. *Eur. J. Biochem*. 271 (2004), 663-677

Dawson P.E., Muir T.W., Clark-Lewis I., Kent S.B. "Synthesis of proteins by native chemical ligation." *Science* 266 (1994): 776-779.

Ding, Z, Fong RB, Long CJ, Stayton PS, Hoffman AS. "Size-dependent control of the binding of biotinylated proteins to streptavidin using a polymer shield." *Nature* 411 (2001): 59-62.

Evans T.C., Benner J., Xu M.Q. "The in vitro ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum." *J Biol Chem* 274 (1999): 3923-3926.

Evans TC, Benner J, Xu MQ. "Semisynthesis of cytotoxic proteins using a modified protein splicing element." *Protein Sci* 7 (1998): 2256-2264.

Fong BA, Wood DW. "Optimization of ELP-intein mediated protein purification by salt substitution." *Protein Expr Purif* 66 (2009): 198-202.

Gimble F.S., Thorner J. "Homing of a DNA endonuclease gene by meiotic gene conversion in Saccharomyces cerevisiae." *Nature* 357 (1992): 301-306.

Giriat, I., Muir, T. W., and Perler, F. B. "Protein splicing and its applications." *Genet. Eng*. 23 (2001), 171 – 199)

Hausdorf G., Gewiess A., Wray V., Porstmann T. "A recombinant human immunodeficiency virus type-1 capsid protein (rp24): its expression,

purification and physico-chemical characterization" *J Virol Methods* (1994) 50: 1-9

Hirata R., Ohsumk Y., Nakano A., Kawasaki H., Suzuki K., Anraku Y. "Molecular structure of a gene, VMA1, encoding the catalytic subunit of H(+)-translocating adenosine triphosphatase from vacuolar membranes of Saccharomyces cerevisiae." *J Biol Chem* 265, no. 12 (1990): 6726-33.

Hodges R.A., Perler F.B., Noren C.J., Jack W.E. "Protein splicing Removes intervening sequences in an Archaea DNA polymerase." *Nucleic Acid Res.* 20 (1992): 6153-6157.

Hoffman AS, Stayton PS. "Conjugates of stimuli-responsive polymers and proteins." *Prog Polym Sci* 32 (2007): 922-932.

Ista LK, Lopez GP. "Lower critical solubility temperature materials as biofouling release agents." *J Ind Microbiol Biotechnol* 20 (1998): 121–125.

Kane P.M., Yamashiro C.T., Wolczyk D.F., Neff N., Goebl M. and Stevens T.H. "Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase." *Science* 250 (1990): 651-7.

Kawasaki M., Satow Y., Anraku Y. "Protein splicing in the yeast Vma1 protozyme: evidence for an intramolecular reaction." *FEBS lett.* 412 (1997): 518-520.

Kim J.S., Raines R.T. "Ribonuclease S-peptide as a carrier in fusion proteins." *Protein Sci.* 2, no. 3 (1993): 348-56.

Klabunde T., Sharma, S., Telenti A., Jacobs W.R., Jr., Sacchettini J.C. "Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing". *Nat Struct Biol* 5 (1998): 31-6

Kumar J.K., Tabor S., Richardson C.C.. "Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli". *Proc Natl Acad Sci* 101 (2004):3759-3764

LaVallie E.R., DiBlasio E.A., Kovacic S., Grant K.L., Schendel P.F., McCoy J.M. "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm." *Biotechnology*, (1993): 187-193.

Lee J.J., Ekker S.C., von Kessler D.P., Porter J.A., Sun B.I., Beachy P.A. "Autoproteolysis in hedgehog protein biogenesis." *Science* 266 (1994): 1528-37.

Lim D.W., Trabbic-Carlson K, MacKay JA, Chilkoti A. "Improved non-chromatographic purification of a recombinant protein by cationic elastin-like polypeptides." *Biomacromolecules* 8 (2007): 1417-1424.

Lue, R.Y., Chen, G.Y., Zhu, Q., Lesaicherre, M.L., and Yao, S.Q. “Site-specific immobilization of biotinylated proteins for protein microarray analysis”. *Methods Mol. Biol.* 264 (2004), 85 – 100.

Maina C.V., Riggs P.D., Grandea A.G. 3rd, Slatko B.E., Moran L.S., Tagliamonte J.A., McReynolds L.A., Guan C.D. “An Escherichia coli vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein.” *Gene* 74, no. 2 (1988): 365-373.

Mathys S, Evans TC, Chute IC, Wu H, Chong S, Benner J, Liu XQ, Xu MQ. “Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation.” *Gene* 231 (1999): 1-13.

Meyer DE, Chilkoti A. “Purification of recombinant proteins by fusion with thermally-responsive polypeptides.” *Nature Biotechnology* 17 (1999): 1112 - 1115.

Meyer D.E., Trabbic-Carlson K., Chilkoti A. “Protein purification by fusion with an environmentally responsive Elastine-like polypeptide: effect of polypeptide lenght on the purification of thioredoxin” *Biotechnol prog* 17 (2001): 720 - 728.

Mills K.V., Lew B.M., Jiang S., Paulus H. “Protein splicing in trans by purified N- and C-terminal fragments of the Mycobacterium tuberculosis RecA intein.” *PNAS* 95 (1998): 3453-3458.

Nilsson B., Forsberg G., Moks T., Hartmanis M., Uhlén M. “Fusion proteins in biotechnology and structural biology.” *Curr. Opin. Struct. Biol.* 2 (1992): 569-575.

Nilsson J., Stahl S., Lundenberg J., Uhlén M. and Nygren P.A. “Affinity fusion strategies for detection, purification and immobilization of recombinant proteins.” *Prot. Expr. Purif.* 11 (1997): 1-16.

Ong E., Greenwood J.M., Gilkes N. R., Kilburn D.G., Miller R.C. and Warren R.A.J. “The cellulose-binding domains of cellulases: tools for biotechnology.” *Trends Biotechnol.* 7 (1989): 239-243.

Perler F. B., Olsen, G. J., Adam, E. “Compilation and analysis of intein sequences.” *Nucleic Acids Res.* 25 (1997): 1087-1093.

Perler F.B. “InBase, the intein database.” *Nucleic acids Res* 28 (2000): 344-5.

Perler F.B., Davis E.O., Dean G.E., Gimble F.S., Jack W.E., Neff N., Noren C.J., Thorner J., Belfort M. “Protein splicing elements: inteins and exteins--a definition of terms and recommended nomenclature.” *Nucleic Acids Res* 22 (1994): 1125-7.

Pietrokovski S. “Modular organization of inteins and C-terminal autocatalytic domains.” *Protein Sci* 7, no. 1 (1998): 64-71.
Pietrokowski S. “Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins.” *Protein Sci* 3 (1994): 2340-2350.
Poland, B.W., Xu, M.Q., and Quiocho, F.A. “Structural insights into the protein splicing mechanism of PI-SceI.” *J. Biol. Chem.* 275, 16408 – 16413.
Rodríguez-Cabello JC, Reguera J, Girotti A, Arias FJ, Alonso M. “Genetic Engineering of Protein-Based Polymers: The Example of Elastinlike Polymers.” *Advances in Polymer Science* 200 (2006): 119-167.
Shimoboji T, Larenas E, Fowler T, Hoffman AS, Stayton PS. “Temperature-Induced Switching of Enzyme Activity with Smart Polymer−Enzyme Conjugates.” *Bioconjug Chem* 14 (2003): 517–525.
Shimoboji T., Ding ZL, Stayton PS, Hoffman AS. “Photoswitching of ligand association with a photoresponsive polymer-protein conjugate.” *Bioconjug Chem* 13 (2002): 915-919.
Smith D.B., Johnson K.S. “Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.” *Gene* 67, no. 1 (1988): 31-40.
Smith M.C., Furman T.C., Ingolia T.D., Pidgeon C. “Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography for recombinant proteins.” *J Biol Chem* 263, no. 15 (1988): 7211-5.
Southworth M.W., Benner J., Perler F.B. “An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile.” *EMBO J* 19 (2000): 5019-5026.
Su X., Prestwood A.K. and McGraw R.A. “Production of recombinant porcine tumor necrosis factor alpha in a novel E.coli expression system.” *Biotechniques* 13 (1992): 756-762.
Telenti A, Southworth M, Alcaide F, Daugelat S, Jacobs WR Jr, Perler FB. “The Mycobacterium xenopi GyrA protein splicing element: characterization of a minimal intein.” *J Bacteriol* 179, no. 20 (1997): 6378-6382.
Trabbic-Carlson K., Meyer D.E., Liu L., Piervincenzi R., Nath N., LaBean T., Chilkoti A. “Effect of protein fusion on the transition temperature of an environmentally responsive elastin-like polypeptide: a role for surface hydrophobicity?” Protein Engineering, Design & Selection 17 (2004): 57-66
Tsao K.L., DeBarbieri B., Michel H., Waugh D.S. “A versatile plasmid expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in Escherichia coli.” *Gene* 169, no. 1 (1996): 31-40.

Urry D W, Luan CH, Parker T.M., Gowda D.C., Prasad K.U., Reid M.C., Safavy A. “Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity.” *J. Am. Chem. Soc.* 113 (1991): 4346-4348.

Urry D.W., Trapane TL, Prasad KU. “Phase-structure transitions of the elastin polypentapeptide-water system within the framework of composition-temperature studies.” *Biopolymers* 24 (1985): 2345-2356.

Urry DW. “Physical Chemistry of Biological Free Energy Transduction As Demonstrated by Elastic Protein-Based Polymers.” *J Phis Chem B* 101, no. 51 (1997): 11007–11028.

Urry DW. “Molecular Machines: How Motion and Other Functions of Living Organisms Can Result from Reversible Chemical Changes.” *Angew Chem Int Ed Eng* 32 (1993): 819–841.

Urry DW., Hayes LC, Gowda DC, Harris CM, Harris RD. “Reduction-driven polypeptide folding by the delta Tt mechanism.” *Biochem Biophys Res Commun* 188 (1992): 611-617.

Vrhovski B, Jensen S, Weiss AS. “Coacervation Characteristics of Recombinant Human Tropoelastin.” *Eur J Biochem* 250 (1997): 92-98.

Vrhovski B, Weiss AS. “Biochemistry of tropoelastin.” *Eur J Biochem* 258 (1998): 1-18.

Wood D.W., Banki M.R. “Inteins and affinity resin substitutes for protein purification and scale up.” *Microb Cell Fact* 4, no. 32 (2005).

Xin G., Yang D.S.C., Trabbic-Carlson K., Kim B., Chilkoti A., Filipe C.D.M., “Self-cleavable stimulus responsive tags for protein purification without chromatography.” *J. Am. Chem. Soc. Comm* 127 (2005): 11228-11229.

Xu M., Southworth M.W., Mersha F.B., Hornstra L.J., and Perler F.B. “In vitro protein splicing of purified precursor and the identification of a branched intermediate.” *Cell* 75 (1993): 1371-7.

Xu M.Q., Perler F.B.,. “The mechanism of protein splicing and its modulation by mutation.” *EMBO J*, 1996: 5146-53.

Xu R., Ayers B., Cowburn D., Muir T.W. “Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies.” *PNAS*, 1999: 388-393.

RIASSUNTO

In questo lavoro di tesi si valuta l'applicazione di un metodo di purificazione di proteine ricombinanti basato sulla combinazione di due diverse specie biochimiche: inteine e polipeptidi elastino-simili. Le inteine sono una classe di proteine che operano una auto-eliminazione dalla sequenza in cui sono contenute unendo con legame peptidico le due sequenze precedente e successiva all'inteina stessa. Tale attività (detta di *splicing*) può essere modulata con delle mutazioni puntuali agli amminoacidi direttamente coinvolti nel meccanismo. I polipeptidi elastino-simili (ELP) ricalcano l'estesa ripetizione del motivo pentapeptidico tipico dell'elastina bovina. I polipeptidi ELP sono una famiglia di biopolimeri formata dal ripetersi dell'unità pentapeptidica "Val-Pro-Gly-Xaa-Gly", dove Xaa può essere qualsiasi amminoacido. Tali strutture artificiali presentano interessanti qualità tra le quali la capacità di auto-assemblaggio (aggregazione reversibile) in seguito a variazione di parametri chimico-fisici quali temperatura, pH e forza ionica.

I peptidi elastino-simili sono solubili in acqua se si trovano al di sotto di una caratteristica temperatura di transizione inversa (Tt) tipica della loro struttura, ma subiscono una transizione di fase causata dalla formazione di aggregati insolubili, quando la temperatura dell'ambiente supera la Tt. Questo peculiare comportamento dei polipeptidi ELP viene preservato anche quando sono fusi ad un'altra struttura proteica, pertanto queste strutture generate dalla fusione tra ELP e una proteina ricombinante rappresentano una nuova generazione di proteine "intelligenti", le cui proprietà chimico-fisiche e funzionali possono essere modulate in funzione delle caratteristiche ambientali. Oltre alla temperatura e alla forza ionica, altre variabili ambientali che possono influenzare la transizione di fase dei polipeptidi ELP e delle corrispondenti fusioni ELP-proteina, sono il pH, l'aggiunta di soluti organici e solventi, la concentrazione proteica.

Nella tecnica descritta in questo lavoro, la transizione di fase degli ELP viene quindi sfruttata in un ciclo di transizione inversa (ITC): il tampone in cui si trova il costrutto di fusione viene modificato (per aggiunta di cloruro di sodio) per far precipitare e risolubilizzare la proteina. In tal modo si permette l'eliminazione dal campione dei contaminanti solubili (quando la proteina è precipitata) e insolubili (quando è in soluzione). Dopo due ITC il costrutto è sufficientemente puro da poter operare un'auto-digestione attraverso l'attività dell'inteina. Per operare la digestione è sufficiente l'agguinta di un composto tiolico (es. DTT) in basse concentrazioni a temperatura ambiente. La reazione di digestione necessita di 20 ore per arrivare a completezza.

Il metodo è stato testato su tre proteine modello: tioredoxina, p24M (una proteina strutturale di HIV-1 var M) e Ct_mono (un polipeptide costituito dalla fusione di quattro peptidi antigenici di MOMP di *Chlamydia trachomatis*) con due differenti ELP (ELP[$V_5A_2G_3$-90] e ELP[KV_7F-36]) e due forme di una inteina (*Mxe*GyrA, in forma nativa e mutata T3C in modo da eliminare l'auto-digestione *in vivo*).

Tioredoxina è stata purificata in forma attiva senza l'ausilio di tecniche cromatografiche, mentre p24M e Ct_mono hanno avuto la necessità di uno step di *polishing* cromatografico dovuto alle difficoltà di separazione tra proteina e coda inteina-ELP dopo la digestione.

Le rese sono state inferiori rispetto a quelle che normalmente si ottengono con processi di purificazione standard cromatografici, ma l'impiego di ELP più corti e sensibili alla forza ionica del tampone di lavoro si può aumentare notevolmente la resa di espressione e di purificazione, come nel caso di ELP[KV_7F-36].

L'impiego di *Mxe*GyrA mutata T3C non ha portato ad alcun innalzamento significativo delle rese.

Il metodo si è dimostrato una valida alternativa all'utilizzo di tecniche cromatografiche nella purificazione di proteine ricombinanti, per la sua facilità di utilizzo, di applicazione e di scalabilità su scala industriale.