Comparison of Three *Escherichia coli* Strains in Recombinant Production of Reteplase

Mehrnosn Fathi-Roudsari 1*, Asal Akhavian-Tehrani 1, and Nader Maghsoudi 2*

1. National Institute of Genetic Engineering and Biotechnology (NIGEB), Shahrah-e Pajoohesh, Tehran, Iran
2. Neurobiology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

**Abstract**

**Background:** *Escherichia coli* (*E. coli*) is the most extensively used host for the production of recombinant proteins. However, most of the eukaryotic proteins are typically obtained as insoluble, misfolded inclusion bodies that need solubilization and refolding. Reteplase as a highly disulfide-bonded recombinant protein is an example of difficult to express protein in *E. coli*.

**Methods:** In this study, a codon optimized reteplase gene was synthetically prepared and cloned under the control of an IPTG inducible T7 promoter. The vector was simultaneously transformed and expressed in three different *E. coli* strains. The ability of strains for expression of this recombinant pharmaceutical was compared. Also, an attempt was made to increase the soluble production of reteplase in SHuffle T7 *E. coli* with alterations of expression condition like temperature, inducer concentration and oxygen supply.

**Results:** High amounts of reteplase were expressed as inclusion bodies in all three strains. BL21 (DE3) showed the highest level of expression in inclusion bodies followed by Rosetta-gami (DE3) and Shuffle T7. Changes of expression conditions were insufficient for soluble expression of reteplase in SHuffle T7 as a genetically engineered host for production of disulfide bonded proteins.

**Conclusion:** The oxidizing cytoplasm of Rosetta-gami and Shuffle T7 in addition to alterations of cultivation parameters could not result in soluble production of reteplase, although the inclusion bodies produced in these two strains might increase the rate of refolding procedure likely due to formation of folding intermediates.

**Keywords:** *Escherichia coli*, Recombinant proteins, Reteplase, Tissue plasminogen activator

**Introduction**

Thrombolysis is the first choice of therapy for Acute Myocardial Infarction (AMI) 1. Tissue Plasminogen Activator (tPA) which derives from fibrinolytic system of blood vessel endothelial cells is a thrombolytic agent that activates plasminogen to form plasmin 2,3. Native tPA is a 70 kDa serine protease composed of 527 amino acid residues with five structural domains and 17 disulfide bonds 4.

Reteplase is a recombinant non-glycosylated form of human tPA produced in *Escherichia coli* (*E. coli*). It has been modified to contain 355 (1-3 and 176-527) of the 527 amino acids of the original protein with a total molecular mass of 39 kDa. This deleted variant of tPA has lost three structural domains of Kringle I, Finger, and EGF, while retaining the thrombolytic Kringle II and protease domains and contains 9 disulfide bonds. However, because of the deletion of three structural domains as well as the carbohydrate side chains, the plasma half-life of this mutated variant is increased to 13-16 min when compared to full length tPA 5,6.

*E. coli* is by far the most widely used host organism for biopharmaceutical production of heterologous recombinant proteins. *E. coli* expression platform is favored due to its ability to quickly reach high cell densities, simplicity, low cost and Food and Drug Administration (FDA) approved status for human applications 8,9. However, production of reteplase in *E. coli* remains challenging. Presence of several disulfide bonds, rare codon usage and cytotoxicity are the major drawbacks of reteplase expression in *E. coli* 9,10. Here, a synthetic reteplase gene optimized for expression in *E. coli* was prepared and its expression was induced under the control of strong T7 promoter in BL21 (DE3), Rosetta-gami (DE3) and SHuffle T7 strains. Also, an attempt was made to compare the ability of the three different strains in the T7-based production of this recombinant pharmaceutical.

Several research groups have also reported the ex-
pression of reteplase in *E. coli* strains, resulting in the formation of inactive inclusion bodies which need to be renatured to gain biological activity. Since there are 9 disulfide bonds, renaturation of reteplase is very difficult with low efficiency 11-14. Furthermore, the refolding procedures may not fully restore the native protein fold and may reduce recombinant protein function 15,16. At the same time, there are few reports which claim soluble production of reteplase in *E. coli*. For instance, display of reteplase on the surface of polyhydroxybutyrate granule using phasin as the affinity tag prevented the aggregation of this protein but the final protein needed a thrombin cleavage step for isolation of reteplase 2. Also, translocation of reteplase to periplasmic region of *E. coli* Top10 under the control of an arabinose promoter has produced negligible amounts of soluble protein 15. Since these methods are inefficient for high level production of soluble reteplase, still refolding of inclusion bodies is the method of choice.

Therefore, in this study, common approaches were used to increase the yield of soluble and active reteplase production in SHuffle T7 as a genetically engineered *E. coli* which is suitable for production of disulfide-bonded proteins 17. Moreover, the effect of isopropyl-beta-D-thiogalactopyranoside (IPTG) concentrations, temperature of expression and oxygen supply were investigated in order to find a way to produce remarkable amounts of biologically active and soluble form of reteplase in *E. coli*.

**Materials and Methods**

**Bacterial strains and plasmids**

*E. coli* DH5α was used as the cloning host for propagation of expression vector. Three different *E. coli* expression strains BL21(DE3), Rosetta-gami and SHuffle T7 were prepared and transformed using the standard protocols.

pET-21a was used as the expression vector. The gene coding for reteplase was synthesized by Bio-basic (Canada). Restriction enzymes and T4 DNA ligase were from Fermentas. Taq DNA polymerase master mix was obtained from Ampliqon. PfU DNA polymerase was purchased from iNtron. Complete protease inhibitor cocktail tablets were purchased from Roche. All chemicals were purchased from Sigma and Merck.

**Construction of the expression vector**

The gene coding for reteplase was synthetically prepared with codon preference of *E. coli* in pUC57 plasmid. The gene was amplified using the primers ret-F 5’ CATATGCTTTACCGGTAACACG 3’ and ret-R 5’ AAGCTTGGGGCGCATGTATCGC 3’. The forward and reverse primers contained *Nde*I and *Hind*III recognition sites, respectively. PCR was performed with the following conditions: denaturation at 94°C for 5 min, 30 cycles of 95°C for 45 s, 66°C for 45 s and 72°C for 1 min followed by a final extension of 10 min at 72°C. The PCR product was double digested with *Nde*I and *Hind*III restriction enzymes and subsequently cloned in pET-21a previously digested with the same enzymes. The ligation mixture was transformed into *E. coli* DH5α and selected on LB agar medium containing 100 μg/ml of ampicillin. The authenticity of the cloning procedure was confirmed by sequencing. The resulting plasmid was named pET-ret.

**Expression of reteplase in different *E. coli* hosts**

The pET-ret plasmid was transformed into chemically competent expression hosts. Single colonies from transformed cells were used for inoculation of 5 ml pre-culture medium containing ampicillin. The cultures were incubated for 12 hr at 37°C with the shake of 180 rpm. 200 µl of the pre-cultures were used for inoculation of 20 ml of LB medium at 37°C until reaching the OD$_{600}$ of 0.8. For protein expression, IPTG was added to the final concentration of 1 mM. Expression was continued at 37°C for 20 hr and cells were harvested within desired intervals. Quantification of protein expression was carried out by image analysis of the SDS-PAGE scans using ImageJ software.

**Determination of cultivation parameters**

For temperature optimization, cells were grown at 37°C until reaching the desired turbidity of 0.8, then the temperature was reduced to 25°C and 18°C and cells were incubated for 30 min. After reaching the desired temperature, IPTG was added and expression was continued for 4 hr.

For optimization of IPTG concentration, two different concentrations of inducer were used. IPTG was added to the final concentration of 1 mM and 0.1 mM after reaching the optimal density of bacterial growth. The expression was continued for four hours and samples were withdrawn with one hour intervals.

In order to find the effect of oxygen supply, two different shaking speeds were checked. After induction of expression with IPTG, the shaking speed of incubator was fixed to 180 or 120 rpm. All experiments were carried out in duplicates.

**SDS-PAGE analysis**

For analysis of reteplase production, cells were harvested by centrifugation at 5000 rpm for 5 min. The bacterial pellets were resuspended in lysis buffer containing 1X protease inhibitor cocktail and disrupted by 5 cycles of sonication (30 s) with 10 s of rest interval. The whole steps were carried out on ice. The supernatant was separated from cell debris by centrifugation at 13000 rpm for 30 min at 4°C. Both fractions were analyzed on 5% stacking, 10% resolving SDS-PAGE. The gels were stained with Coomassie Blue R-250 dye for 2 hr and destained overnight.

**Results**

**Construction of the expression vector**

The reteplase gene was cloned in *Nde*I and *Hind*III recognition sites of pET-21a, therefore located under

---

Avicenna Journal of Medical Biotechnology, Vol. 8, No. 1, January-March 2016

Fathi-Roudsari M, et al
Expression of Reteplase in *Escherichia coli*

Expression of reteplase in *E. coli* BL21 (DE3) strain

*E. coli* BL21 (DE3) is the most widely used prokaryotic expression host which is deficient for two main proteases OmpT and Lon. This strain has remained as the gold standard among expression hosts since commercialization of them started. Therefore, the pET-ret plasmid was first transformed and expressed in this strain. When the cell density reached to OD₆₀₀ of 0.8, IPTG was added to the final concentration of 1 mM and expression was continued for 20 hr at 37 °C. The shaking speed of incubator was 180 rpm during the growth and induction phases. As shown in figure 2, a clear band of overexpression was observed with the molecular mass of 39 kDa. Obviously, all of the expressed protein was deposited in the insoluble fraction of cells and no expression was found in the soluble fraction (Figure 2). Expression was also carried out at decreased temperatures (25 °C and 18 °C) but no accumulation of reteplase was observed in soluble fraction of cell lysates. Induction of expression at lower temperatures reduced the total amount of reteplase expression in inclusion bodies (Data not shown).

Expression of reteplase in *E. coli* Rosetta-gami (DE3) strain

Rosetta-gami B (DE3) is designed to enhance the expression of eukaryotic proteins in *E. coli* by overcoming the problem of codon bias. The strain supplies six tRNAs for rare codons. In addition, due to the mutation in glutaredoxin reductase and thioredoxin reductase (Δgor ΔtrxB), correct folding of disulfide-bonded recombinant protein in cytosolic fraction of *E. coli* is enhanced.

As described above, reteplase contains 9 disulfide bonds, therefore, Rosetta-gami as the expression host was used and the effect of oxidizing cytoplasm on expression of reteplase was found. Expression was induced with 1 mM IPTG at 37 °C and 180 rpm shaking. As shown in figure 3, recombinant protein was expressed and accumulated in the insoluble fraction. No overexpression bands were observed in soluble fraction of cell lysates. Reducing the temperature of expression to 25 and 18 °C did not also increase the expression of reteplase as a soluble protein (Data not shown).

Expression of reteplase in *E. coli* SHuffel T7 strain

SHuffel T7 is an *E. coli* engineered by New England Biolabs specifically for expression of disulfide-bonded recombinant proteins controlled by T7 promoter. In SHuffel strain, not only glutaredoxin reductase and thioredoxin reductase are deleted (Δgor ΔtrxB), but also a version of the periplasmic disulfide bond isomerase DsbC which lacks its signal sequence is expressed in cytoplasm. It has been claimed that the strain is able to express recombinant proteins with multiple disulfide bonds and also correct the mis-oxidized bonds and promote proper folding.
The pET-ret plasmid was introduced into *E. coli* SHuffle T7 and expression was induced with 1 mM IPTG at 37°C and 180 rpm. The expression was continued for 20 hr. As can be seen in figures 4A and B in all fractions that were analyzed on SDS-PAGE, the overexpression band of the recombinant protein was deposited in inclusion bodies and no expression band was observed in soluble fraction.

**Comparison of *E. coli* strains on the expression of reteplase in aggregated form**

In this experiment, *E. coli* strains BL21, Rosetta-gami and SHuffle were simultaneously induced with 1 mM IPTG at 37°C and 180 rpm and expression was continued for 2 and 4 hr. After sonication, pellets were isolated and equal amounts of total proteins were loaded on SDS-PAGE. Figure 5 shows the comparison of reteplase expression in insoluble fraction of the three strains. *E. coli* BL21 produced the highest amount of reteplase in inclusion bodies in comparison to Rosetta-gami and SHuffle T7. In addition, total amount of reteplase expression was higher in Rosetta-gami in comparison to SHuffle T7. Also, the amount of recombinant expression was quantified using ImageJ software version 1.49 (National Institutes of Health, Bethesda, MD). The percent of recombinant reteplase band to the total insoluble fraction of cells is shown in table 1. SHuffle T7 expressed reteplase in almost 50% of the insoluble fraction after 4 hr. This amount of expression is the lowest concentration of reteplase among strains that were used. BL21 (DE3) was able to express reteplase as 70% of the insoluble fraction after 4 hr which shows the high ability of this strain in expression of codon optimized reteplase gene. Rosetta-gami showed an average ability in expression of reteplase. Although the strain had the advantage of expressing rare tRNAs,
it was capable of expressing reteplase in lower concentrations in comparison to BL21 (DE3).

Effect of expression parameters on solubilization of reteplase

Among *E. coli* strains used here, SHuffle T7 is the most promising host for soluble expression of reteplase in cytoplasm. Since expression of reteplase at 37 °C and 180 rpm resulted in complete accumulation of reteplase in inclusion bodies, investigating the role of expression factors on solubilization of this protein in SHuffle T7 host was the main purpose. In order to find the effect of temperature, induction was carried out with 1 mM IPTG at 25 °C and 18 °C and 180 rpm. In spite of low temperature, most of the gene product still aggregated in inclusion bodies and no improvement was reached in soluble expression of reteplase (Figures 6A and 6B). Lowering the temperature also reduced the total amount of reteplase accumulated in insoluble fraction.

The concentration of IPTG was also reduced to 0.1 mM in order to reduce the amount of reteplase production which might be beneficial for correct folding of the protein. The temperature of expression was adjusted to 37 °C with 180 rpm shake. In this strategy, the amount of reteplase accumulation in insoluble fraction was decreased but still no significant amount of reteplase was observed in soluble fraction of cell lysate (Data not shown).

The effect of oxygen supply was also investigated on reteplase production in SHuffle T7 *E. coli*. The speed of shaking was reduced to 120 rpm when 1 mM IPTG was added to the expression flask. The expression was carried out at 37 °C. Figures 6C and 6D show that reduction of aeration cannot increase the expression of reteplase in soluble fraction. Interestingly, the accumulation of reteplase was increased in insoluble fraction.

Finally, an expression was carried out with all three modifications in induction protocol. 0.1 mM IPTG was added to culture medium at 18 °C and 120 rpm. Simultaneous modification of all three parameters was also unable to suppress the aggregation of reteplase in SHuffle T7 *E. coli* and no protein band was observed in supernatant fraction of cell lysate (Figures 6E and 6F).

**Discussion**

As a second-generation recombinant tissue type plasminogen activator, reteplase is developed to treat thrombotic diseases like myocardial infarction, ischemic stroke, peripheral arterial occlusion and venous thromboembolism 20-22. Reteplase shows decreased affinity to fibrin which allows free diffusion and better penetration in blood clots instead of binding only on the surface 23.

Reteplase is famous as an aggregation-prone protein in *E. coli* since contains multiple disulfide bonds and reductive environment of cytoplasm does not let correct disulfide bond formation resulting in generally misfolded and biologically-active inclusion bodies 24. Refolding of inclusion bodies elongates the time of drug production. The efficiency of refolding may be low and refolded enzymes might be unstable due to incomplete refolding or incomplete removal of detergent and denaturants used in this process. Therefore, it is often desirable to obtain high degrees of soluble protein. Till now there is no universal approach for efficient soluble production of aggregation-prone recombinant proteins 25,26 but several strategies like introduction of molecular chaperons to the host bacteria, modulation of expression level, optimization of culture medium, reduction of temperature, addition of chemical chaperons have helped in preventing protein aggregation 27.

In the current study, an attempt was made to express reteplase in three different strains of *E. coli* and compare their ability for soluble expression and production yield of this recombinant protein. *E. coli* BL21 (DE3) is the most common prokaryotic strain used for expression of recombinant proteins till now 19,28. Previous reports on tPA expression in BL21 (DE3) strain show that common strategies failed to produce soluble protein in this strain but aggregated inclusion bodies are extracted for refolding procedure 20.

Rosetta-gami B (DE3) is a hybrid strain engineered to enhance the expression of genes containing rare codons. At the same time, it can enhance disulfide bond formation in the cytosolic fraction due to mutations of trxB and gor genes. This strain can yield 10-fold more active protein than other hosts even when overall expression levels are similar 18,19,28. The protease domain of tPA has been expressed as a soluble protein in this strain but there is no report on expression of reteplase which contains kringle II and protease domains in this strain yet 10.

SHuffle T7 is a recently prepared *E. coli* capable of expressing disulfide bonded proteins in its cytoplasm. In addition to mutations in trxB and gor genes, SHuffle expresses a signal sequenceless DsbC gene in the cytoplasm. Using this strain, production of several disulfide-bond containing recombinant proteins have been dramatically increased in soluble fraction of cell lysates. Expression of vtPA as a truncated mutant of tPA which contains 6 disulfide bonds -formed between 12 cysteins- was also investigated in this strain and presence of vtPA in the soluble fraction of cell lysate was confirmed 17. There are no reports on expression of reteplase in this novel strain and there remains the answer whether SHuffle T7 may favor soluble expression of reteplase.

In the present study, the induction of reteplase expression with 1 mM IPTG at 37 °C and 180 rpm in all three strains resulted in high level expression of reteplase in aggregated fraction of cell lysates. Obviously, BL21 (DE3) showed the highest expression in comparison to Rosetta-gami and SHuffle T7 possibly due to its protease deficient genotype (lon, OmpT). Supplying
Although the total amount of reteplase was reduced.

Good or poor oxygen availability did obviously not affect soluble expression of recombinant reteplase.

Oxygen supply is known for variable effects on protein production in E. coli. Decreasing the temperature of expression is a well known technique for reduction of aggregation-prone recombinant proteins. Decreasing the temperature of expression, concentration of the inducer (IPTG) and oxygen supply was also investigated on reteplase expression in BL21 (DE3) cells. Limited oxygen supply results in lower acetate and higher ammonia accumulation and decreased the amount of reteplase in inclusion bodies. In renaturation procedure, the precipitation of aggregation-prone recombinant proteins will be increased. Here, two different concentrations of inducer (0.1 and 0.5 mM) were used but reduction of IPTG concentration to 0.1 mM could not result in solubilization of reteplase.

Another known strategy for reduction of aggregation during recombinant protein production is optimization of the expression of proteins. Therefore, the effect of oxygen supply on protein expression in BL21 (DE3) cells, limited oxygen supply results in lower acetate and higher ammonia accumulation and decreased the amount of reteplase in inclusion bodies. In renaturation procedure, the precipitation of aggregation-prone recombinant proteins will be increased. Here, two different concentrations of inducer were used but reduction of IPTG concentration to 0.1 mM could not result in solubilization of reteplase.

In an attempt to Wang et al. [14] we prepared several different E. coli strains, BL21 (DE3), Rosetta-gami and Origami 2 using an auto-induction procedure. The authors express their gratitude to the Iran National Science Foundation for the financial support during the course of this project.

Conclusion

It is concluded that for remarkable production of soluble reteplase in E. coli, strategies other than engineering host or environmental parameters should be tested. Manipulating the expression level of proteins or expression system with respect to the soluble and active production from inclusion bodies produced by each strain will be a key factor in improving the refolding procedure needs to be investigated in comparison to higher temperature at 77°C.

It is concluded that for remarkable production of soluble reteplase in E. coli, strategies other than engineering host or environmental parameters should be tested. Manipulating the expression level of proteins or expression system with respect to the soluble and active production from inclusion bodies produced by each strain will be a key factor in improving the refolding procedure needs to be investigated in comparison to higher temperature at 77°C.

Acknowledgement

The authors express their gratitude to the Iran National Science Foundation for the financial support during the course of this project.

References


Avicenna Journal of Medical Biotechnology, Vol. 8, No. 1, January-March 2016
Expression of Reteplase in *Escherichia coli*


