International Journal of Biomedicine 2(1) (2012)

INTERNATIONAL JOURNAL OF BIOMEDICINE

Basic Research

Genetic Structures and Conditions of Their Expression, Which Allow Receiving Native Recombinant Proteins with High Output

Kirill V. Solovyov, PhD[1], Anna M. Kern[2], Natalya A. Grudinina, PhD[1], Tatyana D. Aleynikova, PhD[1], Dmitry S. Polyakov, PhD[1], Irina V. Morozova[1], Michael M. Shavlovsky, PhD, ScD[1]*

[1]*Federal State Budgetary Institution «Research Institute of Experimental Medicine», North Western Branch of the Russian Academy of Medical Sciences, Saint Petersburg, Russian Federation,*
[2]*Hannover Medical School, Hannover, Germany*

Abstract

We investigated the possibility of obtaining native recombinant amyloidogenic proteins by creating genetic constructs encoding fusion proteins of target proteins with Super Folder Green Fluorescent Protein (sfGFP). In this study, we show that the structures, containing the sfGFP gene, provide a synthesis, within a bacterial system, of fusion proteins with minimal formation of inclusion bodies. Constructs containing genes of the target proteins in the 3'-terminal region of the sfGFP gene followed by a polynucleotide sequence, which allows for affinity purification fusion proteins, are optimal. Heating bacterial cultures before the induction of the expression of recombinant genes in 42°C for 30 min (heat shock) was found to increase the output of the desired products, thus practically avoiding the formation of insoluble aggregates. IJBM 2012; 2(1):45-49.

Key words: *recombinant proteins, β2-microglobulin, transthyretin, heat shock, superfolder green fluorescent protein, fusion proteins, inclusion bodies.*

Introduction

Recently, much attention is being paid to a group of relatively rare diseases, characterized by post-translational modifications in the spatial organization of protein molecules. We are referring to amyloidosis [1]. Leading cause of amyloidosis is the occurrence of abnormal protein conformations. About 30 human proteins [2, 3], and their abnormal associates form the basis of amyloidosis. These proteins in vivo and in vitro are capable of forming filamentous associates - fibrils, which are the basis of the amyloid deposits that define the occurrence of amyloidosis. Widely known amyloidoses include Alzheimer's, Huntington's disease, prion diseases, and a few not-so-common conditions, not often mentioned. The study of amyloidosis, particularly the mechanisms of protein fibrillogenesis, can solve both the practical and fundamental problems. Practical problems, of course, involve medical tasks. Any decisions taken will help toward developing effective therapies for devastating and often fatal disease. The fundamental problems are primarily associated with the general theory of protein folding and the functions of the native molecules. Therefore, the study of amyloidosis becomes a very serious problem from all points of view. Our studies [4-8]

**Corresponding author: Prof Michael M. Shavlovsky, PhD, ScD, Head of Human Molecular Genetics Laboratory, Department of Molecular Genetics, Federal State Budgetary Institution «Research Institute of Experimental Medicine», North Western Branch of the Russian Academy of Medical Sciences, 12, Acad. Pavlov str., 197376, Saint Petersburg, Russian Federation.*
Tel: 7-812-2343356
Fax: 7-812-3205849
E-Mail: mmsch@rambler.ru

have been devoted to the investigation of two models, transthyretin protein (TTR) and β2-microglobulin (β2M), and it is their fibrillogenesis that is the cause for the development of the amyloidosis. The study of amyloidogenesis involving these proteins necessitates the production of recombinant analogs with various modifications. The most rational approach for the preparation of such part-modified proteins is by the synthesis of their recombinant analogs in bacterial systems. Therefore, we obtained genetic constructs, encoding the various derivatives of amyloidogenic proteins, and investigated the influence of some factors on the output of the recombinant products. During the study we were able to show that the introduction of an additional protein factor such as the green fluorescent protein superfoldera (sfGFP) and a change in the conditions of the cultivation of the bacteria (heat shock) can significantly improve the output of the native target protein and eliminate the formation of insoluble inclusion bodies.

Our preliminary experiments showed that the recombinant β2M and TTR, as well as their derivatives in the synthesis of the bacterial system, formed large numbers of inclusion bodies, in addition to soluble forms. Obtaining suitable proteins for further analysis required the dissolution of the inclusion bodies using a strong denaturant, followed by renaturation. During the renaturation stage, in most cases, there is no full recovery of the native conformation, and therefore, the proteins received cannot be applied for further conformational studies. For such studies it becomes necessary to obtain baseline-native proteins, i.e. without renaturation. From the literature [9] "improperly" folded proteins are known to have started to fold correctly when they had been sewn with sfGFP. This is the reason why, a construct, containing a gene fusion protein β2M with sfGFP for expression in a bacterial system, was obtained by us [7].

Material and Methods

Expression constructions:

The method of obtaining the expression genetic structures pTRC99a-TTR (L55P)-sfGFP and pTRC99a-β2M-sfGFP was previously described by Solovyov K., et al. [7].

Obtaining the expression genetic structures pTRC99a-sfGFP-β2M-His-tag:

The plasmid created by us earlier [7] and containing the cDNA of β2M and primers: forward 5'-caggtcgacatccagcgtactccaaag-3' with a restriction site for SalI (underlined) and reverse 5'-gtcaagcttatcagtgatggtgatggtgCATGTCTCGATCCCACTT-3' with restriction site for HindIII (underlined) was used as the matrix for amplification. The PCR product, after processing with the appropriate restriction enzymes followed by purification (using Promega, WizardSV Gel and PCR Clean-Up System) was ligated into the pTRC99a-P7 plasmid containing cDNA of sfGFP and processed by the same restriction endonucleases (SalI and HindIII). These restriction sites were chosen so that the β2M-sfGFP-His-tag vector obtained by us could have after ATG (start-codon) of the original plasmid in the same reading frame, the sequence encoding sfGFP, and then the sequence of β2M with the His-tag of six histidines and the stop codon.

Synthesis, production and purification of recombinant proteins: sfGFP-His-tag, TTR (L55P)-sfGFP-His-tag, β2M-sfGFP-His-tag and sfGFP-β2M-His-tag:

The competent cells of strain *E. coli* DH5 (alpha) were used for cloning the plasmid. Selection of the transformants containing the gene of the desired protein (TTR (L55P) or β2M) was performed by PCR analysis. The expressive strain of *E. coli* BL21 (DE3) was used to obtain the proteins. Selection of colonies, the best-producing target protein, was performed using laser scanning confocal microscopy, i.e. bacterial cells of one colony were placed on a glass slide under a cover slip, and the fluorescence was detected using sfGFP optical filter 505-550 nm by exciting the fluorescence using a laser of wavelength 488 nm.

E. coli BL21(DE3) cells, transformed with the appropriate expression constructions, were grown in LB medium at 37°C under aeration until the optical density of OD_{600}=0.8-1. Protein synthesis was induced by adding IPTG to a final concentration of 250 mM, and cultivation was continued overnight at 26°C in air. Bacterial cells were separated by centrifugation for 15 min at 10,000×g and washed once in PBS. PBS (at the rate of 30 ml per pellet obtained from 1 liter of medium), containing 1 mM PMSF, 10mM imidazole and 5 mM 2-mercaptoethanol, was added to the extract washed cell pellet. The cells were disrupted using ultrasound (40 kHz, 3 intervals for 30 seconds) and cooling with ice-water. Glass beads (Glass Beads, 500 microns, Sigma) in a 1:1 ratio to the volume of cell pellet were added to the suspension for a more complete destruction. After the destruction of the bacteria, the supernatant was separated by centrifugation for 15 min at 10,000×g. The supernanant was then filtered through the filter paper and processed to chromatographic purification on metal-chelate nickel sorbent (Invitrogen) (column volume 1.5 ml). The presence of polyhistidine sequences in the recombinant proteins made it possible to complete the purification in one step. After the adsorption of proteins, the column was washed in PBS, containing 20 mM imidazole. Desorption of proteins was performed with PBS with 200 mM final concentration of imidazole. We did not use solutions with high sodium chloride content, because most of our model proteins were capable of fibrillogenesis under these conditions.

The concentration and the yield of proteins were determined using the spectrophotometer NanoDrop2000C (ThermoFisher, USA) by absorbance at 280 nm and 490 nm. As our recombinant products contained sfGFP as a label, the number of fusion proteins was expressed in mg, based on the absorbance at 490 nm (absorption maximum sfGFP) and the sample volume. The molar extinction coefficient of sfGFP at 490 nm is 39.2×10^3 M^{-1} cm^{-1} [9]. We assumed that this ratio does not change, at least substantially, in the fusion protein. Absorbance performed at 490 nm precluded the inclusion of errors due to incomplete removal of the impurity of bacterial protein during affinity chromatography. In terms of recalculation of the absorption at 490 nm on the protein, we used the extinction coefficient of sfGFP and molecular weights of the expression products of the genes (molecular weights: sfGFP - 27.7 kDa, β2M - 11.8 kD, sfGFP-β2M - 39.5 kDa,

TTR - 14 kDa (tetramer - 55 kDa), TTR-sfGFP - 41.7 kDa). Herewith, 0.392 units of the absorption matched the concentration 0.277 mg/mL of sfGFP, 0.395 mg/mL of β2M-sfGFP and 0.417 mg/mL of TTR-sfGFP.

Heat shock

After attaining the stationary growth phase, the cultures were heated at 42°C for 30 minutes. Heating was performed in a water bath with constant stirring. Simultaneously, the control portion of the culture continued to grow at 37°C. Induction of the heated and control culture was performed by IPTG immediately after the heat shock.

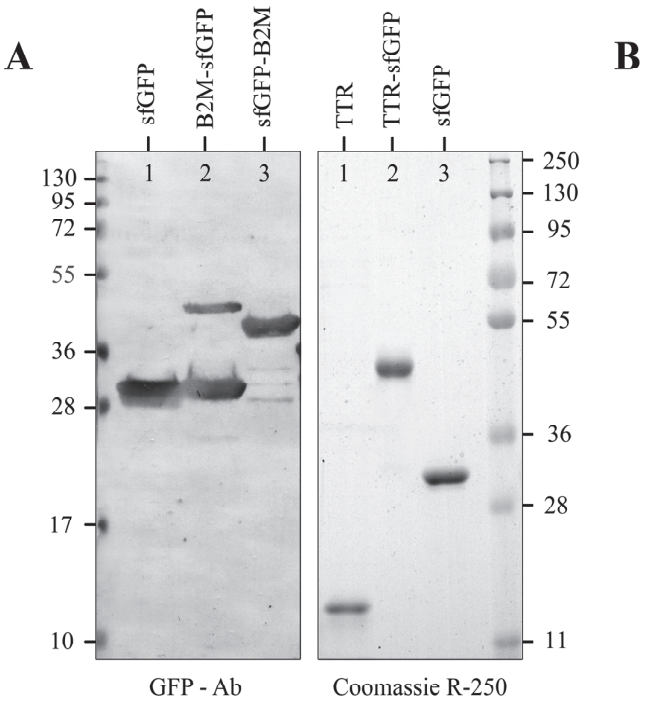
Polyacrylamide gel electrophoresis and Western blot

Electrophoretic analysis (EA) of proteins in polyacrylamide gel electrophoresis (PAGE) was performed according to standard procedure [10]. In our research, the EA method was applied under non-denaturing conditions, in the presence of SDS. To investigate the TTR, 8% PAGE was used, while for β2M – the 12% PAGE was employed. Staining of PAG was performed using Coomassie R-250. In the course of studying the fluorescence of sfGFP and fusion proteins, we photographed (in transmitted UV light) the unstained PAG after EA, under non-denaturing conditions. Western blot was performed using the PVDF-membranes (Sigma). We used HRP-conjugated goat anti-rabbit IgG (Sigma) and the standard staining technique, where 1-chloro-4 naphthol (Sigma) acted as the substrate. The polyclonal rabbit antibodies that we obtained were used for the specific detection of TTR and the proteins based on it. Also, the rabbit polyclonal antibodies to β2M and EGFP obtained by us were used. Coomassie R-250 staining was used to identify all the proteins (total protein) transferred from the PAG to the PVDF-membrane.

Results

In our research, we have showed the possibility of obtaining the fibrillogenic fusion protein β2M-sfGFP [7]. This protein is synthesized intracellularly in the bacterial system and can be obtained in soluble form; it possesses fluorescent properties (as well as native sfGFP) and can form fibrils. However, insoluble fluorescent material is presented in great amounts in the sediment after the extraction of the soluble products, which, according to fluorescence microscopy, represents the aggregates or inclusion bodies. Protein refolding after the dissolution of inclusion bodies by the denaturing agents is not possible. Fig. 1A shows the results of Western blot with staining by using antibodies to GFP. The fusion protein β2M-sfGFP was obtained from a non-denaturing fraction of bacterial lysate (Fig. 1A, lane 2) under denaturing conditions (in the presence of SDS) with the addition of mercaptoethanol and heat processing; the product thus obtained had a molecular weight of 40 kDa, although the preparation showed one an additional "band" (molecular weight of 27.7 kDa), that matched the molecular weight of sfGFP. The quantity of the material matching the fusion protein did not exceed the number of sfGFP. These data are consistent with the results of EA under non-denaturing conditions, where the presence of additional fluorescent bands was also detected

Figure 1.
(A) Western blot analysis after SDS-electrophoresis of sfGFP and fusion proteins β2M-sfGFP, sfGFP-β2M. Staining was done by using polyclonal antibodies directed against GFP.
(B) SDS-electrophoresis of TTR, sfGFP proteins and fusion protein TTR-sfGFP. Stained with Coomassie R-250.



in the samples of the fusion proteins. Apparently, this is due to the partial proteolytic degradation of the fusion protein. Thus, the C-terminal location of sfGFP with the His-tag does not allow the full-length fusion protein to be obtained as a single product with a good output. An attempt was made to change the design of the genetic structure to increase the output of the full-length fusion protein β2M-sfGFP. The B2M gene had been incorporated after the gene sfGFP, and His-tag, in this case, was found at the 3'-end of the gene β2M. The structure obtained was effectively expressed in the bacterial system. According to Western blot (Fig. 1, lane 3) the fusion protein sfGFP-β2M -His-tag migrates as a single band with an apparent molecular weight (corresponding to estimated) 40 kD, and was stained by using antibodies to GFP. These data led us to conclude that the genetic construction created by us, facilitates obtaining the target fusion protein sfGFP-β2M-His-tag with high degree of efficiency and without the formation of inclusion bodies and their degradation. To prevent the proteolytic degradation of our first fusion protein β2M-sfGFP-His-tag, we employed thermal effect on the bacteria transformed, with the relevant genetic construct. This effect contributed to an increase in the yield of the fluorescent product, as well as the full-length protein, although the overall electrophoretic pattern remained the same. The increase in the protein yield following heat exposure spurred us to further study the effect of heat shock on bacteria, with other structures. Table 1 shows the results of several independent

Table 1
Effect of heat shock on the yield of recombinant proteins

Sample	#	Culture conditions	E280	E490	Sample volume (мL)	Yield (mg)	Total yield (mg)	Ratio of yields
I sfGFP	1	no heating	0.142	0.313	1.795	3.95	3.95	
	2	42°C 30 min	0.223	0.467	1.875	6.19	6.19	1.57
II β2M-sfGFP	1	no heating	0.229	0.241	1.064	2.56	2.56	
	2	42°C 30 min	0.191	0.215	1.356	2.9	2.9	1.13
	1	no heating	0.249	0.162	1.777	2.88	2.88	
	2	42°C 30 min	0.231	0.164	2.085	3.4	3.4	1.18
III sfGFP-β2M	1	no heating	0.200	0.083	0.603	0.5	0.5	
	2	42°C 30 min	0.302	0.119	0.743	0.88	0.88	1.76
	1	no heating	0.22	0.075	1.5	1.125	1.125	
	2	42°C 30 min	0.32	0.075	1.5	1.875	1.875	1.67
	1	no heating	0.21	0.074	1.5	1.12	1.12	
	2	42°C 30 min	0.64	0.26	1.5	3.9	3.9	3.5
	1	no heating	0.63	0.25	1.5	3.75	3.75	
	2	42°C 30 min	1.47	0.58	1.5	8.7	8.7	2.32
IV TTR-sfGFP	1	no heating	0.611	0.421	1.540	6.9	6.9	
	2	42°C 30 min	0.82 0.165	0.63 0.01	1.624 0.842	10.88 0.09	10.97	1.59
	1	no heating	0.68	0.53	1.558	8.8	8.8	
	2	42°C 30 min	0.68 0.356	0.612 0.308	1.549 0.598	10.08 1.95	12.03	1.37

Note: *Protein yield was found to be mainly dependent on the volume of the culture medium. The results of the parallel experiments are shown: experience (1) and control (2). Using the Wilcoxon Matched Pairs Test it was determined that the group of experience "1" was significantly different from the control group "2". ($p<0,01$; $z=2,67$) (Valid # = 9, T=0.00, Z=2.665570, p-level=0.007686)*

experiments. Experiments were performed on the same culture, which was shared in equal parts immediately before heat exposure. As the Table shows, the amount of protein synthesized after heat shock (in all cases) was significantly higher than in bacterial cultures not exposed to heat. However, an increase in the duration of heat shock had almost no effect on the yield. Particularly, effective heat shock increased the yield of fusion protein sfGFP-β2M-His-tag. The increase in the yield of the product after heat shock, with respect to the genetic constructs encoding fusion proteins containing amyloidogenic proteins, could indicate that the effect is caused by a decrease in the proportion of the insoluble component. Therefore, control experiments were conducted on the effect of heat shock on the synthesis of unmodified sfGFP. As the experiments show, sfGFP did not form associates, and is entirely in a soluble form in the bacterial cells, and its yield after chromatography on Ni-agarose is quantitative. As the Table shows, thermal processing of the culture, in this case, also contributes to the yield increase. One more structure containing a gene fusion protein TTR-sfGFP-His-tag was investigated to confirm the effect of heat processing of a bacterial culture on the yield of recombinant proteins. Figure 1B shows the results of the electrophoretic analysis obtained after chromatography on Ni-agarose of the fusion protein. The experimental results of the heat shock effects on fusion protein TTR-sfGFP-His -tag gene expression are also shown in Table 1. Therefore, apparently, the heat shock increases the yield of the recombinant proteins, regardless of their structure.

Discussion

Recombinant proteins are increasingly being used as tools of molecular biology and as therapeutic alternatives to natural physiologically active proteins. Nowadays, the production of the recombinant genetic

constructs, encoding analogs of the natural protein, does not have principle difficulties. A technical base of genetic engineering allows the creation of all sorts of designs with predetermined sequences. However, serious problems related to an unpredictable influence of the cellular components on the synthesized proteins and the ability of the synthesized protein to form aggregates, in particular, the so-called inclusion bodies, often appear at the stage of expression of the recombinant genes. Therefore, to obtain native recombinant protein molecules, fully consistent with the natural, it often becomes necessary to use a non-bacterial system, in some cases, the eukaryotic cell. Naturally, *E. coli* is the most preferred producer of recombinant proteins because of the ease of a prior genetic engineering and the subsequent stages of extraction of the target proteins. Therefore, the development of common approaches in this bacterial system, for the synthesis of the recombinant native proteins is very useful. In our research on the study of the fibrillogenic (amyloidogenic) properties of the proteins, we were constantly confronted with the problem of obtaining these proteins in the form of the recombinant native analogs. Unfortunately, these proteins, as well as fusion proteins with the reporter groups often have the ability to produce the aggregated forms in the bacterial cells, which can be transferred into the solution only by denaturation. Following the renaturation stage, an incomplete return of the native conformation often occurs. As natural conformation is the main condition for the study of fibrillogenesis, we tried to apply several approaches to avoid the formation of inclusion bodies. First, this involved obtaining the target proteins in the form of fusion proteins with sfGFP, which would enable the natural folding of the affiliated polypeptides [9]. Second, this helped to increase the yield of the native conformers due to the induction of chaperones - heat shock proteins. According to our preliminary results, sfGFP really has the ability to ensure the correct folding of the affiliated target protein with the formation of soluble fusion proteins, which are readily extractable. Inclusion in the construction of the His-tag allows us to obtain almost pure preparations of the protein in a single chromatographic step. However, despite the effectiveness of this approach, part of the fusion protein cannot be extracted from the bacterial lysate due to its aggregation. Therefore, in order to increase the yield, it was decided to use the well-known ability of the *E. coli* cells to induce the synthesis of heat shock proteins, after short-term heat exposure (30 min, 42°C). Such a process is known to increase the protein synthesis [11]. Therefore, in light of our results, the structures, beginning with the gene sfGFP, followed by a gene (cDNA) of the target protein with a C-terminal His-tag (or another sequence, which is used for affinity cleaning), are most suitable for obtaining functionally active proteins in their natural conformation. Expression of these constructs in a bacterial system effectively enables obtaining a good yield of purified fusion proteins in the native condition. The inclusion of the heat-shock phase can significantly increase the product yield. Thus, the use of sfGFP and heat shock could be an effective tool for increasing the yield of the native recombinant protein, especially if these proteins are prone to the formation of inclusion bodies. To obtain the target protein without any additional components (sfGFP and His-tag) specific plots can be incorporated (by using genetic engineering), which are exposed to specific proteolysis. Purified and appropriate to the natural proteins, the target proteins may be, for example, obtained by affinity chromatography on Ni-agarose (or any other sorbent) followed by filtration of the eluate through a column with immobilized specific protease.

References

1. Shavlovsky MM. Molecular and genetic basis of pathogenesis of amyloidosis. Academic Medical Journal 2010; 10(4):63-81.
2. Buxbaum JN. Animal models of human amyloidoses: Are transgenic mice worth the time and trouble? FEBS Letters 2009; 583:2663-2673.
3. Sipe JD, Benson MD, Buxbaum JN, Ikeda S-I, Merlini G, Saraiva MJM, Westermark P. Amyloid fibril protein nomenclature: 2010 recomendations commttee of the International Society of Amyloidosis. Amyloid 2010; 17(3-4):101-104.
4. Solovyov KV, Gasteva AA, Egorov VV, Aleinikova TD, Sirotkin AK, Schwarzman AL, Shavlovsky MM. The role of the C-terminal fragment of human transthyretin in abnormal fibrillogenesis. Biochemistry 2006; 71: 672-680.
5. Egorov VV, Solovyov KV, Grudinina NA, Lebedev DV, Isaev-Ivanov VV, Kiselev OI, Shavlovsky M.M. Atomic force study of peptides homologous to beta-domain of alpha-lactalbuins. Protein and Peptide Letters 2007; 14:471-474.
6. Polyakov, DS, Grudinina NA, Solovyov K., Egorov VV, AK Sirotkin, Aleinikova TD, Totolyan Areg A. Shavlovsky MM. β2-microglobulin amyloidosis: fibrillogenesis of natural and recombinant human β2-microglobulins. Academic Medical Journal 2010; 10 (2):40-49.
7. Solovyov KV, Polyakov DS, Grudinina NA, Egorov VV, Morozova IV, Aleynikova TD, Shavlovsky MM. Expression in E. coli and purification of the fibrillogenic fusion proteins ttr-sfgfp and β2m-sfgfp, Preparative biochemistry and biotechnology 2011; 41(4):337-349.
8. Polyakov DS, Grudinina NA, Solovyov K, Egorov VV, Sirotkin AK, Aleinikova TD, Totolyan Areg A, Shavlovsky MM. Production of recombinant human β2-microglobulin and its fibrillogenesis at low and neutral pH. Molecular Medicine 2011; 2:36 -39.
9. Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 2006; 24:79-88.
10. Sambrook J, Fritsch EF, Maniatis T, Molecular cloning, USA, Cold Spring Harbor Laboratory Press 1989; 174-175.
11. Macrides SC. Strategies for achieving high-level expression of genes in Escherihia coli, Microbiological Reviews 1996; 60(3):512-538.