STUDIA UBB CHEMIA, LXII, 2,Tom II, 2017 (p. 333-345)
(RECOMMENDED CITATION)
DOI:10.24193/subbchem.2017.2.26

HETEROLOGOUS EXPRESSION AND PURIFICATION OF RECOMBINANT PROAPOPTOTIC HUMAN PROTEIN SMAC/DIABLO WITH EGFP AS FUSION PARTNER

PÁL SALAMON[a,b], ILDIKÓ MIKLÓSSY[a], BEÁTA ALBERT[a], MÓNIKA KORODI[a], KATALIN NAGY[c], ILDIKÓ BAKOS[c], SZABOLCS LÁNYI[a,c], CSONGOR ORBÁN[a,*]

ABSTRACT. New proteins as molecular targets in development of therapies are discovered every day. However, study of their interactions with other proteins or binding partners in complex cellular environments has its limits. Therefore, high-yield production of these proteins in heterologous systems is a valid necessity, while obtaining these proteins linked to suitable fluorescent markers represents a step ahead in protein-protein interaction studies and cellular or subcellular localization.

In this study, we present production of human SMAC/Diablo recombinant protein with EGFP as a fusion partner. High-yield expression of the fusion protein was carried out in E. coli Rosetta™(DE3)pLysS strain, and an acceptable purity of the protein was obtained after affinity chromatography purification and gel filtration. The obtained protein can be further used in protein-protein interaction studies, whereas our method represents a cost-effective and efficient production method for EGFP-fused proteins, applicable for a number of therapeutically important polypeptides.

Keywords: *SMAC/Diablo, apoptosis, heterologous expression, EGFP fusion proteins*

[a] *SAPIENTIA Hungarian University of Transylvania, Faculty of Economics, Socio-Human Sciences and Engineering, Department of Bioengineering, 1 Libertatii Square, RO-530104, Miercurea Ciuc, Romania*

[b] *University of Medicine and Pharmacy of Targu Mures, Faculty of Pharmacy, 38 Gheorghe Marinescu str., RO-540139, Targu Mures, Romania*

[c] *University POLITEHNICA of Bucharest, Faculty of Applied Chemistry and Materials Science, Str. Gh. Polizu, Nr. 1-7, Sector 1, 011061 Bucuresti, Romania*

** Corresponding author: orbancsongor@uni.sapientia.ro*

INTRODUCTION

A complex regulatory network of protein activation and inhibition maintains optimal cellular function. In case of malfunctions of this system, various diseases can develop. Apoptosis, or programmed cell death represents a vital mechanism in maintenance of tissue homeostasis and development of the immune system. Tumor cells are responsive to both internal and external stress, while being resistant to apoptosis [1]. Changes in the rate of apoptosis may have pathological consequences: an increase in apoptosis rate can cause neurodegenerative diseases, while a decrease can cause cancer.

SMAC/Diablo (second mitochondria - derived activator of caspase direct inhibitor of apoptosis-binding protein with low pI) is a newly discovered pro-apoptotic mitochondrial caspase activator protein [1]. Changes in permeability of the mitochondrial membrane SMAC/Diablo is released in concert with cytochrome-C, and it binds to the anti-apoptosis IAPs (inhibitor of apoptosis proteins) in the cytosol. By inhibition of these proteins, apoptosis is promoted [2].

SMAC/Diablo protein was identified as a cytochrome-C/APAF-1/caspase-9 pathway caspase activation mitochondrial protein [3]. It is known that SMAC/Diablo neutralizes XIAP in the cytosol, while generating additional initiator caspase activity [1, 2]. The first 55 amino acids of the protein sequence represent a mitochondrial signal peptide, while the mature protein presents an Amino-terminal AVPI-end, which is essential for the function of SMAC/Diablo protein as a natural antagonist of anti-apoptosis proteins (IAPs) [4]. The molecular weight of the mature protein is 20.8 kDa, which occurs in the cytosol in a tetrameric structure [5]. In the BioGrid (Biological General Repository for Interaction Datasets) database there are listed a number of 45 known protein-protein interactions involving SMAC/Diablo.

The discovery of GFP protein [6, 7] and the two-photon microscope [8] allowed the development of fluorescence microscopy applications in cell biology [9]. The wide-range use of GFP is a consequence of its unique properties: reduced sensitivity to higher temperatures, detergents, alkaline pH, photobleaching, organic salts, chaotropic salts and many proteases [10]. EGFP (enhanced green fluorescent protein, a mutant version of GFP) has emerged as a powerful fluorescent label for quantitative fluorescence microscopy applications [11]. EGFP contains two amino acid substitutions (Ser65Thr and Phe64Leu), that lead to a 35-fold enhancement of fluorescence over wild-type green fluorescent protein (wtGFP) [12]. Using an EGFP-fused protein we can track and quantify the fusion protein, examine protein-protein interactions, describe biological events and signals in cells, or even apply these fusion constructs in drug discovery processes.

In the light of the above, the aim of the study was fusion of human SMAC/Diablo recombinant protein with EGFP, expression of the fusion construct in prokaryotic host cells, namely in *E. coli* Rosetta™(DE3)pLysS strain and the subsequent purification by affinity chromatography followed by gel filtration (size exclusion chromatography). The EGFP-fusion construct was intended to assure easy and specific detection of SMAC/Diablo in further experiments.

Our data supports that using a prokaryotic expression system in bioreactor cultures we can cost-effectively produce the recombinant protein in large quantities. With the affinity and gel filtration chromatography we can isolate an adequately pure protein which can be further used in oncological or cell biology research.

RESULTS AND DISCUSSION

Modelling of SMAC_EFGP fusion protein structure

In the first phase of our study, in order to assess structural stability of the created SMAC_EGFP fusion protein, and more importantly, to predict structural changes which could affect the functionally essential Amino-terminal AVPI segment of Smac/Diablo, we performed analysis of the 3D structural model of the fusion protein.

The tridimensional structure model was obtained by homology modelling, whereas homologous sequences were compiled using the Phyre2 online search engine [13].

The Amino-terminus of the recombinant fusion protein sequence starts with a methionine (Figure 1, marked in yellow), which is removed during translation by methionyl aminopeptidase. Removal of the methionine results in the mature Amino-terminus beginning with the AVPI amino acid sequence (marked in red), which is considered crucial to its interaction with the IAP proteins. EGFP is marked with green, its Carboxyl-terminus being continued with a 6xHis tag (marked in purple), required for affinity chromatography.

The tridimensional model of the fusion protein reveals presence of the significant secondary structural elements of SMAC/Diablo reported in crystallographic studies [14], as well as the free Amino-terminal AVPI sequence, whereas the EGFP presents its characteristic beta-barrel [15]. Based on the above mentioned results, we concluded that the SMAC_EGFP_His recombinant fusion protein would retain the structural elements required for SMAC/Diablo function.

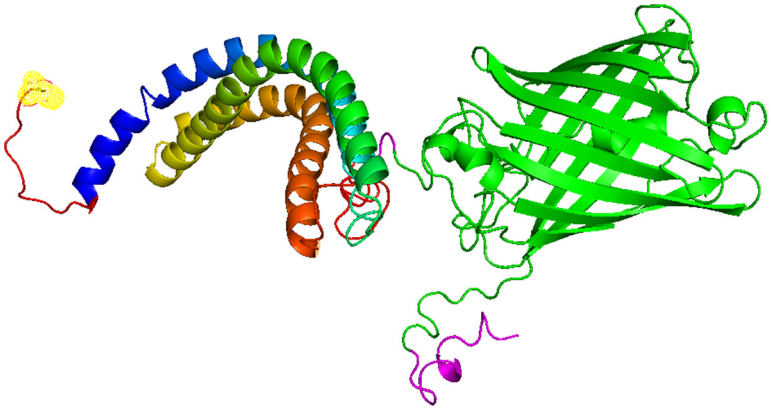

Figure 1. SMAC_EGFP 3D ribbon model.
(yellow – N terminal methionine, red – N-term AVPI, green – EGFP, purple – His tag. Structural model compiled by Phyre2 online search engine, visualized and edited in PyMol).

Design and assembly of the SMAC_EGFP_His construct

The SMAC_EGFP_His construct was designed taking into account the chosen heterologous *E. coli* pET-based expression system (pET20b vector with *E. coli Rosetta™(DE3)pLysS* as a host strain), as well as the downstream purification steps. Restriction map of the designed vector is presented in Figure 2, outlining the significant elements of the recombinant construct. Vector selection was based on our previous experience in high-yield protein production, according to which these expression plasmids proved to be extremely reliable for the production of a number of proteins.

Coding sequences for SMAC/Diablo and EGFP, respectively, were obtained by PCR, specific primers being designed (as described in the Experimental section) to assure the restriction sites selected for cloning (XbaI and BamHI for the SMAC sequence, and BamHI and NotI for EGFP).

The expression vector was obtained by directional cloning carried out with the above mentioned restriction endonucleases (details in Experimental section), whereas assembly of the two coding sequences was performed by one-step ligation. Verification of the correct plasmid assembly and conformation, respectively, and incorporation of the gene construct into the vector structure was carried out by subsequent restriction digestions. Correct integration of the SMAC_EGFP_His construct into the plasmid pET20b was verified by digestion with NotI, EcoRI and XmnI.

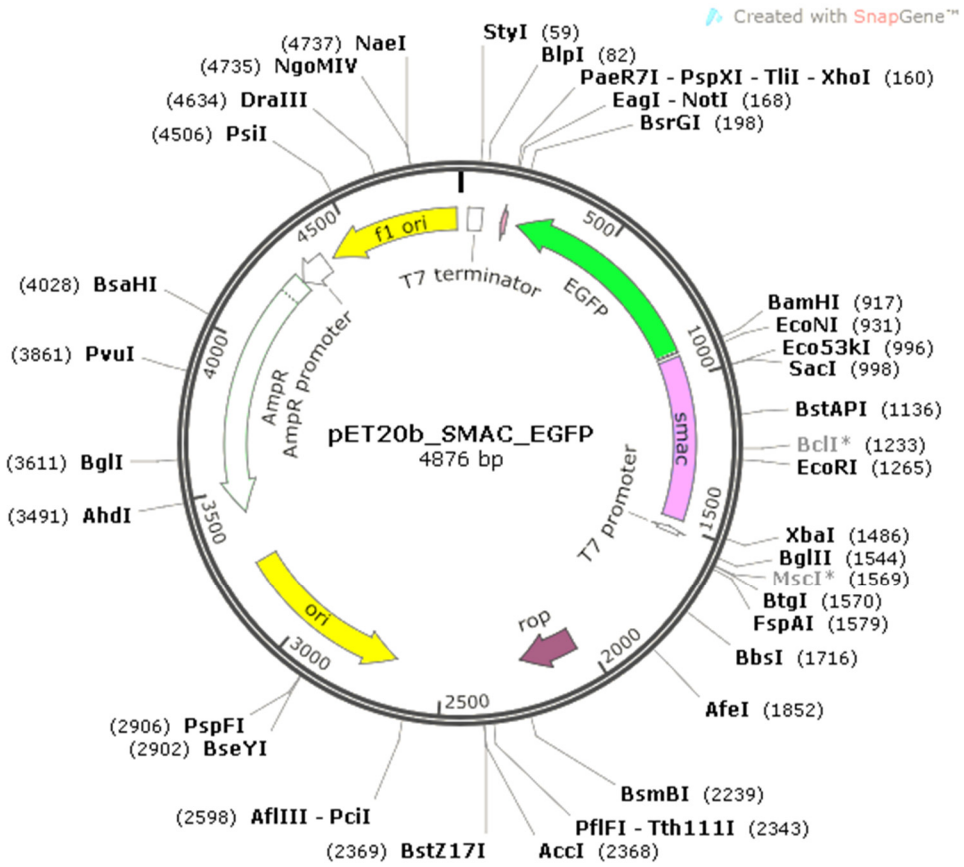

Figure 2. Restriction map of pET20b_SMAC_EGFP expression plasmid, outlining significant features of the vector (Map created in SnapGene).

The expected fragments were 3779 bp and 1097 bp for the NotI, EcoRI double digestion, respectively 2942 bp and1934 bp for the digestion performed by the double-cutter XmnI. The results of the restriction digestions are visualized in Figure 3, where presence of restriction fragments of the expected length supports formation of the correct structure of the recombinant vector and successful ligation of the SMAC_EGFP construct.

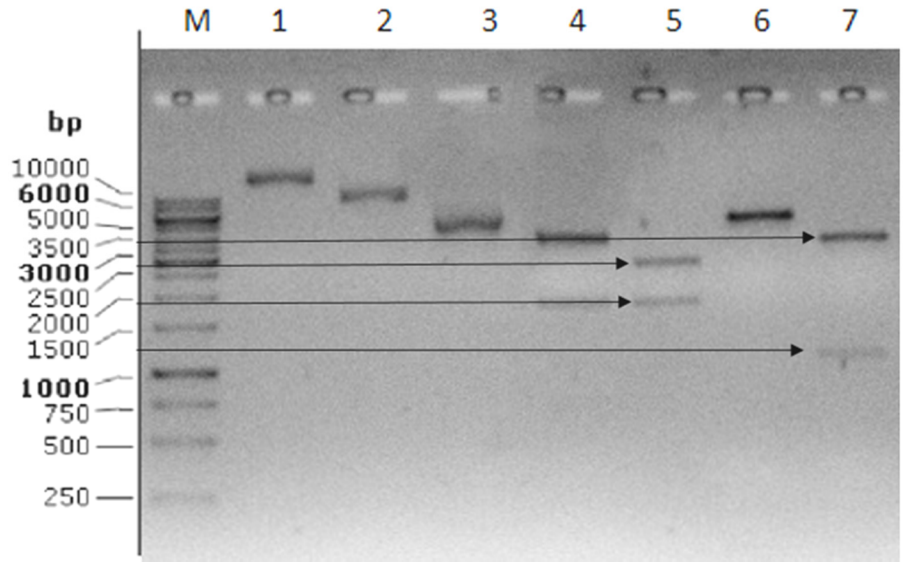

Figure 3. Result of pET20b_SMAC_EGFP plasmid digestion: (M) - 1 kb DNA ladder GeneRuler; (1) -pET20b_1; (2) - pET20b_2; (3) - pET20b; (4) - pET20b_1 XmnI; (5) - pET20b_2 XmnI; (6) - pET20b_1 Not + Eco RI; (7) - Not + Eco RI pET20b_2. (Reaction products were separated by electrophoresis on a 1% agarose gel and visualized by RedSafe staining).

High-yield expression of the SMAC_EGFP fusion protein

To ensure high-yield production of the SMAC_EGFP fusion protein, a bioreactor system was used, as cultivation of transformed expression strains under controlled process parameters offers the possibility to obtain high biomass levels, and consequently, high protein production rates.

The SMAC_EGFP recombinant protein was expressed in a Biostat A plus bioreactor, in cultures of *E. coli Rosetta™(DE3)pLysS* strain transformed with the pET20b_SMAC_EGFP expression vector. Fermentation conditions, determined previously in small-scale expression experiments (data not shown) were 37 °C, pH 7.0, using an M9 mineral medium with 2 g/L glucose as the sole carbon source. Induction of target protein expression by the strong T7 promoter was performed by addition of 1 mM IPTG to the culture media at OD~20. During the derepression period, cells were further cultivated at 18 °C to facilitate correct folding of the protein. After 16 hours of expression, we obtained a biomass yield of 40 g WCW/1 L culture. Our procedure for high-yield protein expression obtained in high-density bacterial cultures has

resulted in similar protein production rates reported in other studies concerning expression of alcohol-dehydrogenase [16], even labeled eukaryotic proteins [17] or for expression of genes from heterologous biosynthetic pathways [18].

Two-step purification of recombinant SMAC_EGFP

As downstream experiments often require a highly purified protein solution, purification of the obtained recombinant SMAC_EGFP was carried out in two consequent steps, by Ni-affinity binding followed by size exclusion chromatography. Results of the fusion protein purification by affinity chromatography were verified by SDS-PAGE gel analysis (Figure 4, lanes 6 and 7). As our results illustrate, the Ni-NTA affinity resin non-specifically bound other proteins from the production culture, along with the target protein. However, we found our SMAC_EGFP being the predominant protein fraction with a relative molecular weight of 52 kDa in the elution fractions 6 and 7 (Figure 4). Due to non-specific binding, further purification of the protein solution was implemented by size exclusion chromatography.

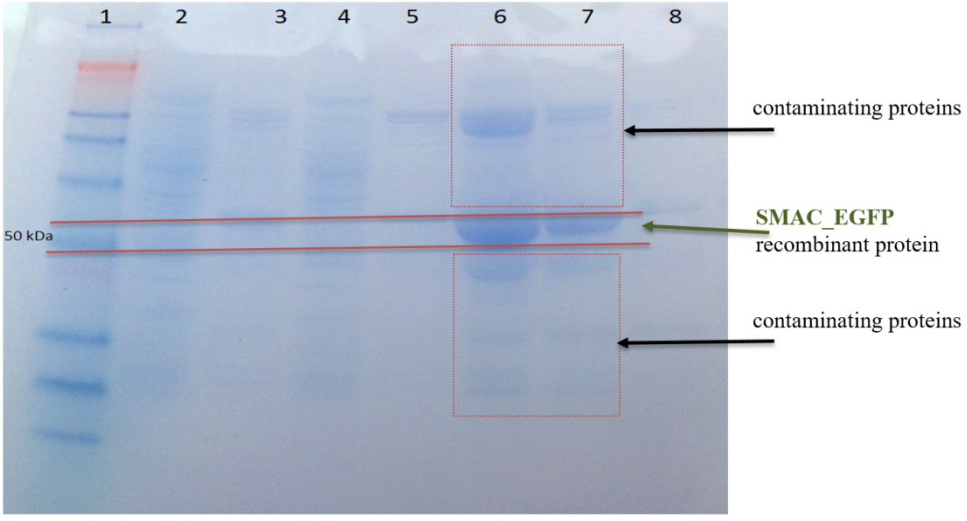

Figure 4. SDS-PAGE gel analysis of affinity chromatography purification of the obtained fusion protein. (1) - PageRuler™ Prestained NIR Protein Ladder, (molecular weight marker from Thermo Scientific™); (2) – total cell protein, (3) – total soluble protein, (4) – flow through, (5) – fraction 6, 250 mM imidazole elution, (6) – fraction 7, 250 mM imidazole elution, (7) – fraction 8, 250 mM imidazole elution, (8) – fraction 9, 250 mM imidazole elution

Based on our SDS-PAGE results (Figure 5) illustrating the downstream purification steps, the non-specifically bound proteins remaining in the eluted protein fractions after affinity purification were eliminated by size exclusion chromatography. Concentration of the recombinant protein after gel filtration was determined 2.5 (±0.03) mg/mL, obtained in a total volume of 8 mL, with a total quantity of 20 mg pure recombinant protein SMAC_EGFP.

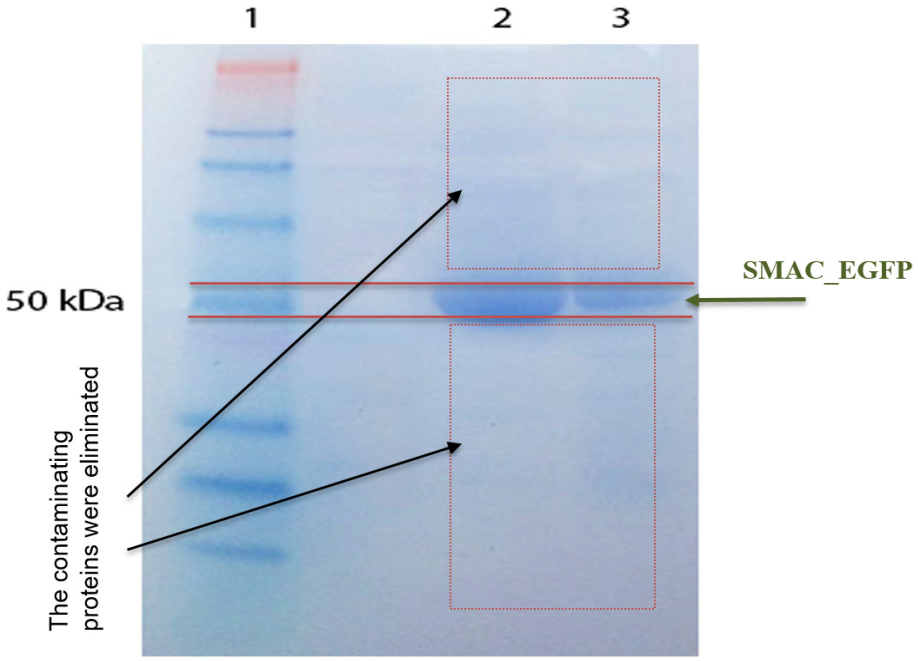

Figure 5. SDS-PAGE gel analysis of the second purification step of SMAC_EGFP by size exclusion chromatography. (1) - PageRuler™ Prestained NIR Protein Ladder (molecular weight marker from Thermo Scientific™); (2) - SMAC_EGFP eluted fraction, 20 µL; (3) - SMAC_EGFP eluted fraction, 10 µL

CONCLUSIONS

Based on the results of restriction endonuclease digestion confirmed by sequencing of the plasmid we have successfully assembled the SMAC_EGFP_His recombinant construct. For expression of the SMAC_EGFP recombinant protein, the *E. coli* expression strain Rosetta™(DE3) pLysS transformed with the recombinant vector was used successfully under the following conditions: 16 hours of cultivation after induction with 1 mM concentration of IPTG at 18 ºC.

A significant quantity of 20 mg/L of pure recombinant SMAC_EGFP protein was obtained by applying the above conditions in a bioreactor culture. Regarding the downstream purification procedure, our two-step affinity and size exclusion chromatography protocol proved to be successful, as we obtained a protein solution in adequately pure form for further investigations, e.g. protein-protein interactions in cancer and cell and biology research.

EXPERIMENTAL SECTION

Oligonucleotide design for PCR and DNA manipulations

The properties of primers, designed for cloning are shown in table 1. During the PCR reaction the FW_BamHI_EGFP and Rev_NotI_EGFP, FW_XbaI_SMAC and Rev_BamHI_SMAC primers, respectively, were used together. In the cloning PCR reactions, the pSmac-GFP vector (nr. 40881 from ADDGENE) was used to obtain the SMAC coding sequence.

Table 1. Properties of oligonucleotides used for cloning

Name	Sequence	Nr. of nucleotides	Melting temp.	Restriction site
FW_BamHI_EGFP	5'GCGTAGGATCCCCATATGGTGAGCAAGG3'	28 bp.	60 ºC	BamHI (5'…GGATCC…3')
Rev_NotI_EGFP	5'GTATTAGCGGCCGCTCTGAGTCCGGACTTGTACAG3'	35 bp.	60 ºC	NotI (5'…GCGGCCGC…3')
FW_XbaI_SMAC	5'GGTCTAGATAAGGGAAGCTTATGGCG3'	26 bp.	59 ºC	XbaI (5'…TCTAGA…3')
Rev_BamHI_SMAC	5'GTAGTAGTAGGATCCGCATCCTCACGCAG3'	29 bp.	60 ºC	BamHI (5'…GGATCC…3')

Thermal cycles and reaction components used to assemble the XbaI_SMAC_BamHI construct:

Cycles	Temp.	Time
1x	95 ºC	2 min
35x	95 ºC	0.5 min
	53 ºC	0.5 min
	72 ºC	1 min
1x	72 ºC	7 min

Reaction mix	
10x PFU buffer	5 µL
10 mM dNTP	1 µL
100 mM FW_XbaI_SMAC primer	0.5 µL
100 mM Rev_BamHI_SMAC primer	0.5 µL
Template DNA	1 µL
PFU enzyme	1 µL
Sterile water	41 µL
Total	50 µL

Thermal cycles and reaction components used to create the BamHI_EGFP_NotI construct:

Cycles	Temp.	Time
1x	95 ºC	2 min
35x	95 ºC	0.5 min
	47 ºC	0.5 min
	72 ºC	1 min
1x	72 ºC	7 min

Reaction mix	
10x PFU buffer	5 µL
10 mM dNTP	1 µL
100 mM FW_BamHI_EGFP primer	0.5 µL
100 mM Rev_NotI_EGFP primer	0.5 µL
Template DNA	1 µL
PFU enzyme	1 µL
Sterile water	41 µL
Total	50 µL

The PCR products were purified by Thermo Scientific GeneJET PCR Purification Kit, according to the manufacturer's recommendations, then in order to create cohesive ends, the samples were digested with restriction endonucleases (double digests), according to the reaction set-ups presented in Table 2. Both digests were incubated at 37 °C (Thermo-Shaker TS-100C) for 1 hour and inactivated at 80 ºC for 20 minutes. The pET20b plasmid was also digested at 37 °C for 1 hour, then incubated at 80 ºC for 20 minutes to inactivate the restriction enzyme.

Table 2. Restriction reaction set-up

XbaI + BamHI digestion		**BamHI + NotI digestion**		**XbaI + NotI digestion**	
XbaI_SMAC_BamHI PCR product	25 µL	BamHI_EGFP_NotI PCR product	24 µL	pET20b plasmid	10 µL
10x Tango buffer	3 µL	10x BamHI buffer	3 µL	10x Orange buffer	3 µL
XbaI enzyme (5 U/µL)	1 µL	BamHI enzyme (5 U/µL)	1 µL	XbaI enzyme (5 U/µL)	4 µL
BamHI enzyme (5 U/µL)	1 µL	NotI enzyme (5 U/µL)	2 µL	NotI enzyme (5 U/µL)	1 µL
Total	30 µL	Total	30 µL	Sterile water	12 µL
				Total	30 µL

To increase the efficiency of ligation of the digested plasmid pET20b and the digested PCR products, both were separated on a 1% agarose and the appropriate fragments were isolated with the GeneJet Gel Extraction Kit (Thermo).

The PCR products and the plasmid pET20b were ligated O.N., 16 ºC using T4 DNA ligase enzyme (Thermo Scientific). Ligation mixtures were transformed into chemically competent *E. coli* Top10F and selected on ampicillin containing LB plates.

Expression of human recombinant SMAC_EGFP protein in bioreactor cultures

Recombinant protein production was carried out in a 1 L capacity Sartorius Biostat®A Plus Bioreactor, using BioPAT®MFCS/DA Supervisory Control and Data Acquisition (SCADA) Software. The reactor was firstly loaded with 1 L basic M9 broth. The system was autoclaved for 20 minutes at 120 °C, in order to ensure sterility. After sterilization, the reactor was connected to aeration, acid and base solutions, temperature control system and control unit. After the reactor has cooled down (<40 ºC), thermally unstable compounds were added (Table 3. marked with *) through a sterile filter (0.25 µm). In order to homogenize the system, mixing, temperature, and pH control were launched: 400 RPM, 37 ºC, pH 6.9. After the stabilization of the reactor, the reactor was inoculated under sterile conditions with 10 mL inoculum. The culture at this stage was grown at 37 °C, with a dissolved oxygen level above 40%, and pH 6.9. When the cell density reached OD_{600}=20, the temperature was set to 18 ºC and the protein expression was induced with isopropyl-thiogalactopyranoside (IPTG) (1 mM final concentration). Protein expression was carried out at 18 ºC, O.N. (16 hours).

In order to harvest the cells, 1 L culture was centrifuged (12,000xg for 10 min at 4 ºC), and the cell pellets were stored at -80 °C until further processing (Thermo Scientific FORM 88000 series).

Cell lysis was performed as follows: 1 gram of cells were resuspended in 5 mL of lysis buffer (20 mM Tris-HCl (pH 8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF, protease inhibitor cocktail). Cell disruption was performed with a Microfluidizer LM10, in order to increase the efficiency, the micro-flight compression was performed twice. The resulting cell extract was centrifuged at 4 ºC, 60,000xg, for 60 minutes in order to separate the solubilized proteins from cellular debris.

Table 3. Composition of 1x M9 minimal broth

Component	Concentration	1000X TRACE elements	
Na_2HPO_4	3.54 g/L	$FeCl_3 \times 6H_2O$	50 mM
NaCl	0.50 g/L	$CaCl_2$	20 mM
KH_2PO_4	3.40 g/L	$MnCl_2 \times 4H_2O$	10 mM
NH_4Cl*	2.00 g/L	$ZnSO_4 \times 7H_2O$	10 mM
$C_6H_{12}O_6$*	2.00 g/L	$CoCl_2$	2 mM
$MgSO_4$*	2.00 mM	$CuCl_2 \times 2H_2O$	2 mM
$CaCl_2$*	0.02 mM	$NiCl_2 \times 6H_2O$	2 mM
Ampicillin*	100 µg/mL	$Na_2MoO_4 \times 2H_2O$	2 mM
TRACE*	1X	$Na_2SeO_3 \times 5H_2O$	2 mM
*- Added after sterilization		$H_3BO_3 \times 5H_2O$	2 mM
		HCl	60 mM

Purification of recombinant SMAC_EGFP protein

Affinity chromatographic purification was carried out using a 2x5 ml HisTrap (GE Healthcare) column with an ÄKTA FPLC system. Data acquisition and system control were carried out with a UNICORN 5.11 software package. The column, the protein solution and the buffers were kept at 4 ºC. The wash buffer contained 20 mM Tris-HCl (pH 8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF, 20 mM imidazole. The elution buffer contained 20 mM Tris-HCl (pH 8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF, 500 mM imidazole. Purification parameters were set at: 10 mL column volume, 1 mL/min flow rate, 2 mL fraction volume. A cell lysate of 200 mL volume was injected, while the elution step was performed in one step, at 250 mM imidazole concentration. Following affinity chromatography, the purified protein was dialyzed for 24 hours, at 4 ºC, under stirring, in the following buffer: 20 mM Tris-HCl (pH 8), 250 mM NaCl, 2 mM DTT, 1 mM PMSF.

Size exclusion chromatography purification was carried out using a HiLoad16/600, Superdex 75 column with a 20 mM Tris-HCl (pH 8.0, 250 mM NaCl, 2 mM DTT, 1 mM PMSF solution as the mobile phase.

ACKNOWLEDGMENTS

Our work was carried out with the support of the Romania Ministry of Education in the framework of PN-II-PT-PCCA-2013-4 project.

REFERENCES

1. G. Martinez-Ruiz, V. Maldonado, G. Ceballos-Cancino, J. Grajeda, J. Melendez-Zajgla, *Journal of experimental & clinical cancer research,* **2008**, *27*, 48.
2. A.M. Verhagen, P.G. Ekert, M. Pakusch, J. Silke, L. Connolly, G. Reid, D. Vaux, *Cell,* **2000**, *102*, 43.
3. C. Du, M. Fang, Y. Li, L. Li, X. Wang, *Cell,* **2000**, *102*, 33.
4. L. Ma, Y. Huang, Z. Song, S. Feng, X. Tian, W. Du, M. Wu, *Nature,* **2006**, *13*, 2079.
5. E. Mastrangelo, P. Vachette, F. Cossu, F. Malvezzi, M. Bolognesi, M. Milani, *Biophysical Journal*, **2015**, 714.
6. R.Y. Tsien, *Annu. Rev. Biochem,* **1998**, *67*, 509.
7. M. Chalfie, Y. Tu, G. Euskirchen, W.W. Ward, D.C. Prasher, *Science*, **1994**, 263, 802-805.
8. W. Denk, J.H. Strickler, W. W. Webb, *Science*, **1990**, *248*, 73.
9. R. Yuste, *Nature*, **2005**, 902-904.
10. M.A. Ehrmann, C.H. Scheyhing, R.F. Vogel, *Letters in Applied Microbiology,* **2001**, *32*, 230.
11. R.M. Hoffman, *Laboratory Investigation*, **2015**, *95*, 432.
12. J. Sun, G.H. Kelemen, J.M. Fernandez-Abalos, M.J. Bibb, *Microbiology*, **1999**, *145*, 2221.
13. L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, *Nature Protocols,* **2015**, *10*, 845.
14. G. Wu, J. Chai, T.L. Suber, J.W. Wu, C. Du, X. Wang, Y. Shi, *Nature*, **2000**, *408*, 1008.
15. A. Royant, M. Noirclerc-Savoye, *Journal of Structural Biology,* **2011***, 174*, 385.
16. J. Glazyrina, E. Materne, T. Dreger, D. Storm, S. Junne, T. Adams, G. Greller, P. Neubauer, *Microbial Cell Factories,* **2010**, *9*, 42.
17. A. Sivashanmugam, V. Murray, C. Cui, Y. Zhang, J. Wang, Q. Li, *Protein Science*, **2009**, *18(5),* 936.
18. Z. Lai, R. Rahim, A. Ariff, R. Mohamad, *Journal of Microbiology, Biotechnology and Food Sciences,* **2016**, *6(3),* 905-910.