STUDIA UNIVERSITATIS BABEŞ-BOLYAI, CHEMIA, LIV, Sp.Iss. 2, 2009

OBTENTION OF THE ANCESTRAL PROTEINASE STEMZYME-IDP-β BY HETEROLOGOUS EXPRESSION

ILDIKÓ MIKLÓSSY[a], LÁSZLÓ SZILÁGYI[a,b], BEÁTA ÁBRAHÁM[a], SZABOLCS SZILVESZTER[c], SZABOLCS LÁNYI[a]

ABSTRACT. Considering the ancestral proteinase stemzyme-IDP-β as a proper starter-molecule to redesign functional properties among serin-proteinases, we constructed several expression vectors and studied the expression of this hypotethical enzyme in different cellular systems. Expression studies were performed in *E. coli* cultures, as well as in cultures of the methylotrophic yeast *Pichia pastoris.*

Keywords: IDP serine-proteinases, heterologous expression, expression vector, Pichia pastoris

INTRODUCTION

Proteinases can be classified into five different catalytic types in which serine, threonine, cysteine, aspartic or metallo groups play primary role in the cleavage of the peptide bond. Enzymes whose catalytic process involves a nucleophilic attack to the peptide bound by the hydroxyl group of a serine residue are termed serine proteinases. There are more than 40 families of serine proteinases, which can be distinguished on the basis of amino acid sequences. The largest family is family S1 regarding both the number of know protein sequences and the variety in the chemical character of the cleavage sequence. This family is often called as trypsin superfamily according to its best-characterized member the digestive protease trypsin. In this superfamily there are enzymes that cut the polypeptide chain after positively charged residues (trypsin, thrombin, plasmin), after negatively charged residues (granzyme B), after large and small hydrophobic residues (chymotrypsin and pancreatic elastase, respectively) [1].

[a] *Sapientia University, Faculty of Sciences, Libertății Square, Nr. 1, RO-530100, Miercurea Ciuc, Romania, Facultatea de Babeş-Bolyai, Facultatea de Chimie şi Inginerie Chimică, Str. Kogălniceanu, Nr. 1, RO-400084, Cluj-Napoca, Romania, mtt@sapientia.siculorum.ro*

[b] *"Eotvos Lorand"University, Department of Biochemistry, Pazmany Street, Nr. 17/c, 1117, Budapest, Hungary, szilagyl@elte.hu*

[c] *"Politehnica" University of Bucharest, Faculty of Applied Chemistry and Materials Science, Splaiul Independenței 313, RO-060042, Bucharest, Romania, iovu@tsocm.pub.ro*

Wouters and colleagues reported the construction of a synthetic protein predicted theoretically to be the ancestor of the present-day immune defense proteases (IDP) [2]. By parsimony analysis of a multiple alignment of 56 IDP sequences a synthetic gene was constructed to express the recombinant protein that they called stemzyme-IDP-β.

According to this study, this ancestral enzyme displayed an unexceptional tolerance to mutations at the binding site, with different mutations resulting in activities similar to some of the synthetic enzyme's descendents; the ancestral enzyme also showing a remarkably plastic S1 substrate binding site, which can be explained by the fact that the common ancestor of serine proteinases having widely different substrate specificity must tolerate various changes in the substrate binding region during evolution.

This interesting property of the ancestral enzyme makes it a favourable target for rational design to create novel enzyme specificities. Although homology modelling is feasible since several high resolution X-ray structures are available for closely related serine proteinases it is of outmost significance to crystallize and determine the 3D structure stemzyme IDP-. For this purpose we synthesized its gene and worked out efficient heterologous expression systems using *E. coli* (this work) and *Pichia pastoris* (manuscript in preparation).

RESULTS AND DISCUSSION

Design of the coding sequence for stemzyme-IDP-β

Since stemzyme-IDP-β is a hypothetical protein, which cannot be cloned from natural sources, its entire coding region has to be designed and optimized in terms of codon usage and control regions. The coding sequence was thus obtained based on the published amino-acid sequence and optimized for codon usage considering the expression hosts to be expressed by.

Serine proteinases are synthesized in their native form as inactive zymogens, as they contain an N-terminal propetide sequence; activation of the zymogen occurs through a specific proteolytic cleavage, which liberates the alpha amino group of Ile16. The signal sequence and propeptide region of the serine proteinases is located generally in a separate first exon in their genes, which is variable among the members of this family; the reconstituted sequence of the ancestral proteinase does not contain this region, its sequence begins with the mentioned Ile [2].

Consequently, expression of a proteolytic enzyme, of whose propetide sequence is not known, raises some questions, which can be solved by the fact that the propetide regions of some proteases can be interchanged without loss of function [3]. Addition of the propetide sequence of well-studied serine-proteases (like trypsin) is one possibility which we considered feasible when designing the coding sequence for the ancestral proteinase.

Construction of the pET17-stem vectors

Fragments amplified by the pET-stemF-pET-stemR1 and pET-stemF-pET-stemR2 oligo pair were purified by agarose gel electrophoresis and cleaved with Hind III and Xho I. The modifications at the terminals of stemzyme IDP-β coding sequence made it possible to insert it into pET17c vector cleaved with the same enzymes. This vector was originally developed to express human trypsinogens [4]. By this way two vectors were obtained, pET17c-stemR1 and pET17c-R2. Both express the protease in form of inactive zymogen, requiring activation by enteropeptidase. In addition, pET17c-stemR2 results in an enzyme with C-terminal myc-tag and his-tag fusion. Positive clones were identified by restriction digestion (Figure 1, Part A) and sequencing. Maps of both plasmids are shown in Figure 1, Part B.

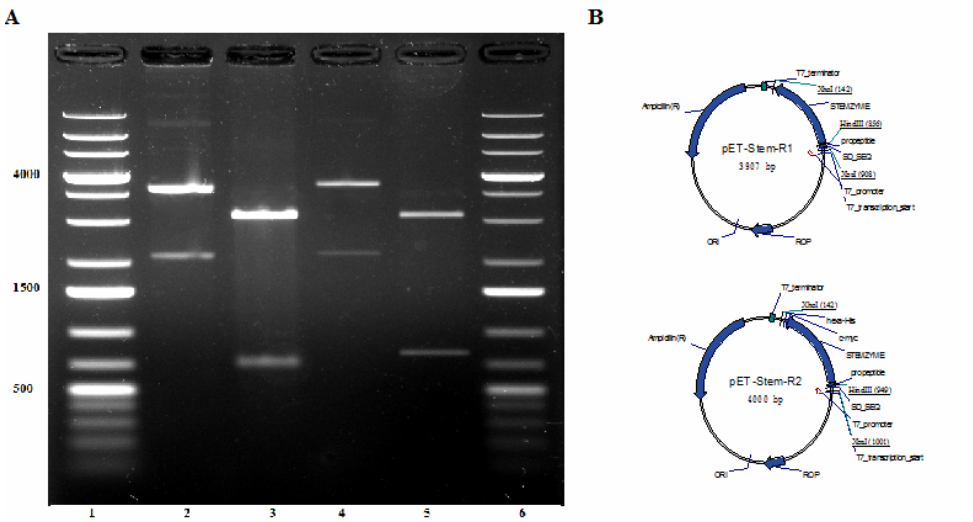

Figure 1. Illustration of the pET17c-stemR1 and pET17c-R2 expression vectors. *Part A*. Gel electrophoresis image of the recombinant constructs and verification by restriction digestion. Lane 1: 1kb DNA ladder, lane 2: pET17c-stemR1 plasmid; lane 3: pET17c-stemR1 plasmid digested with *XhoI* and *HindIII*; lane 4: pET17c-stemR2 plasmid; lane 5: pET17c-stemR2 plasmid digested with *XhoI* and *HindIII*; lane 6: 1 kb DNA ladder; *Part B*. Map of pET17c-stemzyme expression vectors. The drawing summarizing the features of the vector was composed by VNTI program.

Analysis of the expression of recombinant proteins

Expression vectors were introduced into *E. coli* BL21 (DE3) pLysS by chemical transformation or by electroporation. Transformed cells were plated on LB agar plates containing 50 ug/ml ampicillin supplemented with 1% glucose. Starter cultures from single colonies were grown overnight at 37ºC, at

250 rpm shaking in LB-amp+1% glucose. Expression cultures in 200-500 ml 2YT supplemented with 50 μg/ml ampicillin were inoculated with 10 % (v/v) starter cultures. Expression was induced by adding IPTG to the cultures in late-logarithmic growth phase (OD_{600} of 0.8-1.2) to a final concentration of 0.5 mM. Induced cultures were further grown in the same conditions for 3-4 hours. Before induction and at different times during induction 2 ml samples were withdrawn and analyzed for protein expression as recommended by Novagen (pET System Manual, 11.ed. p.34.). Figure 2. shows the time course of the expression.

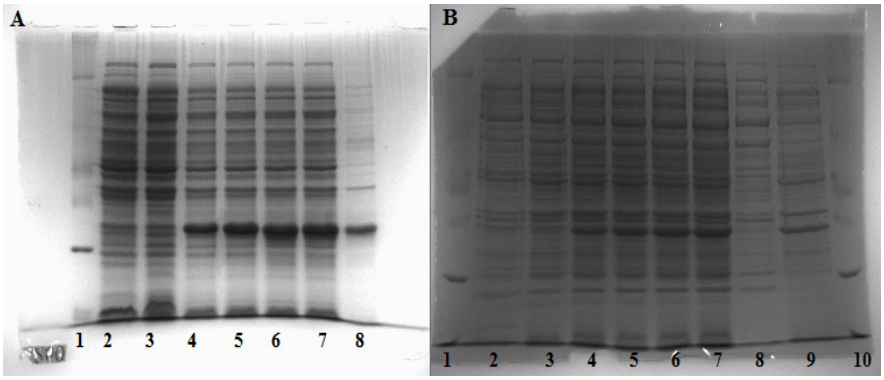

Figure 2. Time course of the expression.

Part A. Expression of stemzyme R1 - Comassie-stained proteins from the production culture

lane 1: protein molecular weight marker, Fermentas; lanes 2, 3: proteins from the expression culture before induction; lane 4: cellular proteins after 1 hour of expression; lanes 5, 6, 7: cellular proteins after 2, 3 and 4 hours of expression; lane 8: protein molecular weight marker, Fermentas

Part B. Expression of stemzyme R2

lane 1: protein molecular weight marker, Fermentas; lanes 2, 3: proteins from the expression culture before induction; lane 4: cellular proteins after 1 hour of expression; lanes 5, 6, 7: cellular proteins after 2, 3 and 4 hours of expression; lane 8: soluble proteins; lane 9: insoluble proteins, lane 10: protein molecular weight marker, Fermentas

On the basis of these gels the size of the protein produced by pET-stem R1 was 25 kDa and by pET-stemR2 was 29 kDa. These values agree well with the calculated molecular weights of 26,300 and 29,350, respectively. According to the apparent Mr data the two forms of the protein were designated as stemzyme-26 and stemzyme-29.

The recombinant proteins are insoluble

After the derepression period cells were harvested by centrifugation for 15' at 6000xg, 4°C and the cell pellet was stored at -80°C until purification. Harvested cell pellets were resuspended in 1xPBS in a ratio of 4 ml of PBS/100 ml of production culture, supplemented with protease inhibitor PMSF 1 mM (Sigma) and 1% TritonX-100 (Sigma). For disruption of cell walls and obtention of raw cell extract, ultrasound treatment was used (sonicator Dr. Hierschler) 5 cycles of 10s at 70% amplitude, with 10s pause between the cycles. Separation of cell debris and insoluble proteins was carried out by centrifugation for 25' at 4°C and 10000xg.

Purification of the His-tagged soluble proteins was performed by affinity binding, as described; however, a significant amount of the target protein was found to aggregate in inclusion bodies. The insoluble protein was solubilized by the reducing agent guanidine hydrochloride and refolded by cisteine treatment; protein bands corresponding to each step were separated by SDS-PAGE and bound to anti-His6x antibody as visualized in Figure nr. 3.

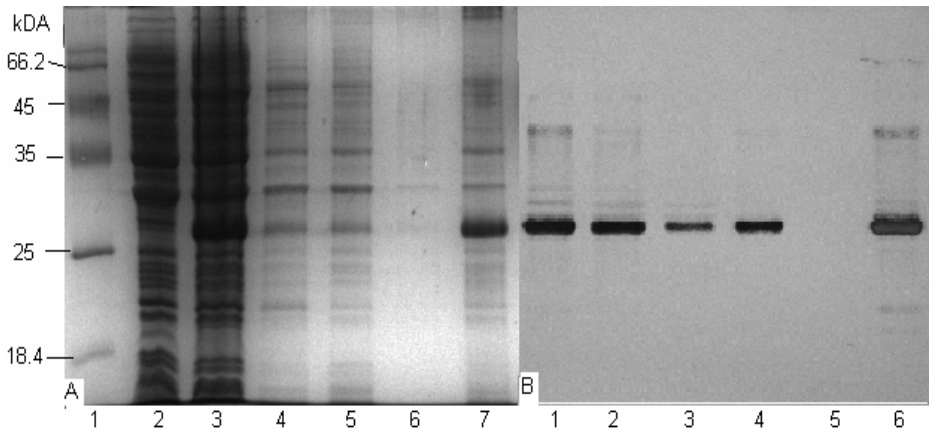

Figure 3. Illustration of stemyzme-R2 expressed in BL21(DE3)pLysS.

Part A. Comassie-stained proteins from the production culture.

Lane 1: protein molecular weight marker, Fermentas; lane 2: proteins from the expression culture before induction; lane 3: cellular proteins after 3 hours of expression; lanes 4, 5, 6: soluble proteins in the raw cell extract obtained in 3 steps of sonication and washing; lane 7: insoluble proteins

Part B. Western Blot of the obtained stemzyme presenting the purification procedure.

Lane 1: total insoluble proteins in GuHCl; lane 2: unbound proteins from the first purification step performed on 1 ml of Porfinity Ni-charged resin for 500 ml of production culture; lane 3: protein fraction collected after washing; lane 4: elution with 250 mM imidazole; lane 6: elution fraction obtained after a second step of purification.

Kinetic measurements

A preliminary determination of the activity of the obtained enzyme was carried out in order to compare the kinetic parameters to the ones presented by Wouters and collab. in 2003. Absorption spectrum of the purified stemzyme preparate was determined in order to appreciate the obtained enzyme quantity, the molar extinction coefficient for calculations being determined by the ProtParam Tool based on the amino acid sequence.

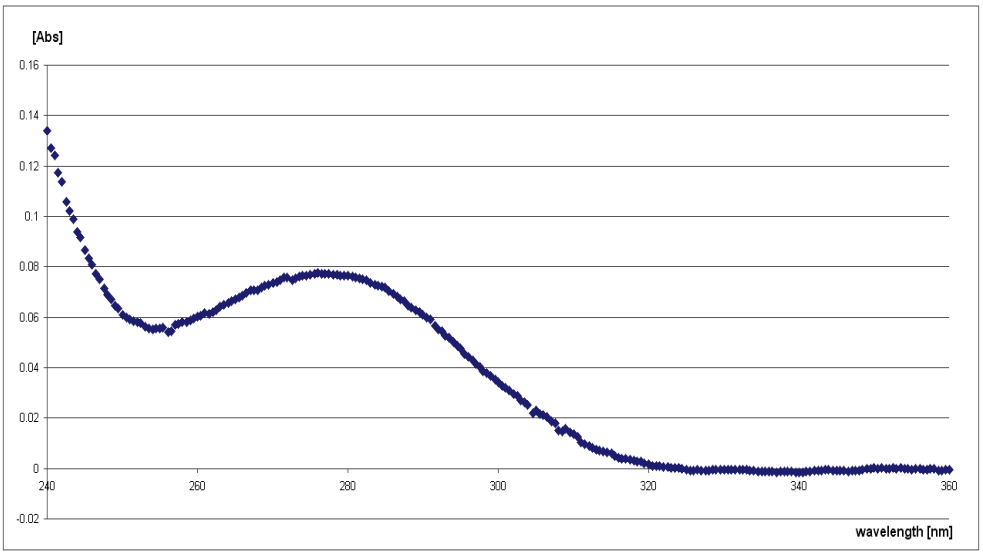

Figure 4. Graphical representation of the absorption spectrum of the obtained stemzyme.

For kinetic measurements, a 0.1 mg/ml stemzyme concentration was used, after activation with enteropeptidase (0.1 mg/ml). Series of measurements from the activation reaction were conducted after 0, 20, 60 and 240 minutes of activation. Using the fluorescent substrate Suc Ala Ala Pro Tyr AMC (Bachem), added in a concentration of 1 mM, absorption was read for 4 minutes at 340 nm wavelength (CaryWinUV, Varian). From the above-presented data calculating with a reaction rate of 1 mAbs min^{-1}, a k_{cat} value of 73 mol s^{-1} was determined, considered as an acceptable kinetic parameter.

Data series of 4 measurements are presented, taken in the time points of 0, 20, 60 and 240 minutes of activation with enterokinase. Equations of the trendlines and R^2 values are represented on the chart.

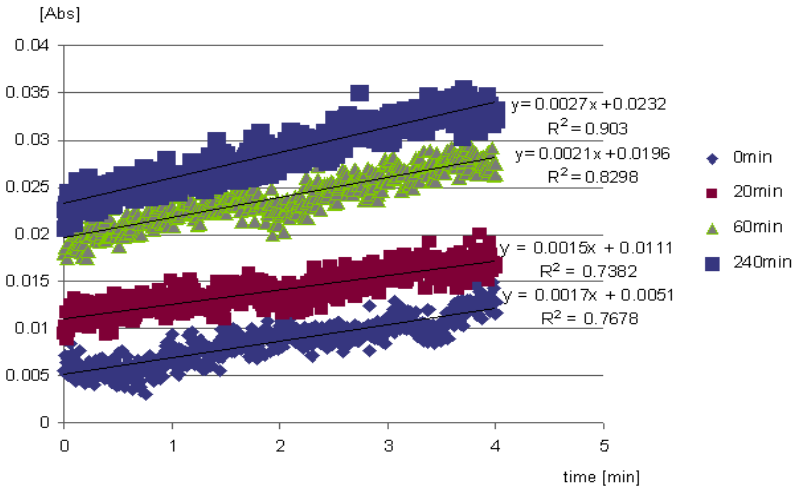

Figure 5. Graphical representation of the activity assay of the obtained stemzyme.

CONCLUSIONS

Expression in a prokaryotic system using *E.coli* BL21(DE3)pLysS as a host strain for the recombinant plasmids pET17-stemR1 and pET17-stemR2 proved to be successful. Based on the results presented, expression of the active stemyzme was achieved using this procedure.

Transformation efficiencies proved to be higher when using electroporation as method for introduction of the plasmid into the host cell. In order to prevent plasmid loss, production cultures of the transformed colonies were prepared shortly after transformation. Optimization of the expression procedures has yet to be performed, taking into consideration higher concentrations of the inducer agent IPTG (up to 1-5 mM concentrations), as well as growth temperature. Using lower temperatures for growth of the production cultures (25, 30°C) could result in the same expression levels, but formation of insoluble proteins aggregated in inclusion bodies may be decreased to some extent.

As showed in the results chapter, purification of the obtained recombinant protein was completed successfully, although several purification steps had to be performed. This fact may be caused by the insufficient binding capacity of the Ni-chelating matrix used. Extensive work has yet to be done considering solubilization and renaturation of the obtained insoluble protein. Considering determination of the kinetic parameters of the obtained enzyme, the preliminary results presented in this study are a good starting point in determination of substrate specificity of the stemzyme. For future experiments, considering determination of 3D structure of the ancestral proteinase by X-ray diffraction, a highly purified form of the enzyme should be obtained, towards which the work presented in this study offers useful information.

The presented results underline that, after a well elaborated design procedure of the coding sequence and careful choosing of the expression system, hypothetical proteins, with deduced sequences can be obtained even in prokaryotic systems. Moreover, using *E. coli* as an expression host assures the ease of manipulation and reduced costs with respect to other expression systems, providing also high efficiencies in protein production.

EXPERIMENTAL SECTION

Bacterial strains and plasmids

BL21(DE3) pLysS (genotype: F^-, ompT, hsdSB (rB–, mB–), dcm, gal, λ(DE3), pLysS, Cm^r) strain was purchased from Stratagene. Strain Top10F F'(genotype: {*lac*Iq Tn*10* (TetR)} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1*ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G') was from Invitrogen; pET17c plasmid vector was purchased from Novagen.

Vector construction

The deduced coding sequence for stemzyme-IDP-β was chemically synthesized by Eurofins MWG Shyntesis GmbH (Germany), and inserted by Topo-cloning into a pCR2.1 plasmid, its sequence being verified by the manufacturer. The synthesized sequence was aimed to be introduced into pPICZα vector for expression in the methylotroph yeast *Pichia pastoris*. The synthesized coding sequence included a carboxy terminal myc tag followed by a hexahistidine tag to facilitate the identification and purification of the recombinant protein

To create a bacterial expression vector, oligonucleotides were designed to generate a Hind III restriction site at the 5` end and an Xho I site at the 3` end of the stemzyme-IDP-β coding sequence. With the 5` oligo (pET-stemF) we also attached the trypsinogen propeptide containing an enteropeptidase cleavage sequence to the amino terminus of stemzyme IDP-β. For the modification of the 3` end two oligos were synthesized. pET-stemR1 added a stop codon to the end of the coding sequence of the protease thereby eliminating both the myc and the his tags, while pET-stemR2 retained both tags.

The sequences of the oligonucleotides were as follows (recognition sequences underlined):

pET-stemF:

5′CCCAAGCTTTTCCCGTGGACGATGATGACAAGATCATCGGTG GAACGGAGGCC-3′; pET-stemR1: 5′-

CGGCTCGAGTCATAGATTGGCCTTCAAGATCTTC-3′

pET-stemR2: GGTCTCGAGGAACAGTCATGTCTA (Eurofins MWG Shyntesis GmbH, Germany). The oligos were designed by Oligo Explorer 1.2 (www.genelink.com) The calculated melting point for the matching sequences in pET-stemF, pET-stemR1 and pET-stemR2 is 65.2 °C, 61.8°C and 65.2°C, respectively.

PCR amplification

The PCR protocol for 30 cycles was as follows: 95°C for 5 min, 94°C for 45 s, 58°C for 45 s, 72°C for 1 min, and a final 5 min at 72°C.

DNA manipulation, restriction digestion and ligation reactions were performed by standard techniques [5]. *HindIII* and *XhoI* were purchased from Fermentas, T4 DNA ligase was from Invitrogen.

Expression

For transformation of the pET17-stem vector in the chosen host strains, both $CaCl_2$ transformation and electroporation (Xcell GenePulser, BioRad) were used. Both chemically transformed and electroporated cell suspensions (including control probes without plasmid DNA) were plated on LB agar containing 50 μg/ml ampicilin and incubated overnight at 37°C in order to select transformed colonies.

Propagation of the transformants was carried out by inoculation of LB broth supplemented with 50 μg/ml ampicillin with an isolated colony. Induction of transcription of the gene of interest was realized by IPTG (isopropyl-β-thio-galactopyranoside, Sigma) in a concentration of 0.5 mM and the culture was further incubated for 3 h in the same conditions.

Purification

Denatured proteins forming inclusion bodies were solubilized by 6 M guanidine hydrochloride (Fluka) treatment, refolding was achieved using 1.5 mM cisteine solution in 0.1 M Tris pH 8.8 (Fluka).

Affinity purification of the His-tagged enzyme was carried out by adsorption on a Ni-charged polymer matrix (Porfinity IMAC Ni-charged Resin, BioRad). (Soluble proteins resulted from 1000 ml of production culture were adsorbed on 1 ml of Ni-NTA Sepharose resin, in 20 mM PBS (20 mM imidazole and 0.5 M NaCl), elution was carried out by 250 mM imidazole.)

PAGE

Specific binding of the proteins transferred to the nitrocellulose membrane was carried out based on the 6xHis fusion tag, by a mouse Anti-His primary monoclonal antibody (Invitrogen) and the WesternBreeze Chromogenic Kit (Invitrogen).

ACKNOWLEDGMENTS

The present work was supported by the National Authority for Management of Programmes (CNMP), project number 61027.

REFERENCES

1. A. Caputo, J. C. Harish, M. N. James, J. C. Powers, R. C. Bleackley, *Proteins,* **1999**, *35,* 415.
2. M. A. Wouters, K. Liu, P. Riek, A. Husain, *Mol. Cell,* **2003**, *12*, 343.
3. I. Venekei, L. Szilágyi, L. Gráf, W. J. Rutter, *FEBS Lett.,* **1996**, *383*, 143.
4. L. Szilágyi, E. Kénesi, G. Katona, G. Kaslik, G. Juhász, L. Gráf, *J. Biol. Chem.,* **2001**, *276,* 24574.
5. J. Sambrook, “Molecular Cloning: A Laboratory Manual (2nd ed.)”, Cold Spring Harbour Laboratory Press, New York, **1989**.