Journal of Biochemistry and Molecular Biology, Vol. 32, No. 5, September 1999, pp. 497–501

JOURNAL OF
Biochemistry
and
Molecular Biology

Short communication

Novel Vectors for the Convenient Cloning and Expression of *In Vivo* Biotinylated Proteins in *Escherichia coli*

Eun-Wie Cho, Jung-Hyun Park, Shin-Young Na, and Kil Lyong Kim*
Peptide Engineering Research Unit, Korea Research Institute of Bioscience and Biotechnology, Yusong, P.O. Box 115, Taejon 305-600, South Korea

Received 24 April 1999, Accepted 15 May 1999

Biotinylation of recombinant proteins is a powerful tool for the detection and analysis of proteins of interest in a large variety of assay systems. The recent development of *in vivo* biotinylation techniques in *E. coli* has opened new possibilities for the production of site-specifically biotinylated proteins without the need for further manipulation after the isolation of the recombinantly expressed proteins. In the present study, a novel vector set was generated which allows the convenient cloning and expression of proteins of interest fused with an *N*-terminal *in vivo* biotinylated thioredoxin (TRX) protein. These vectors were derived from the previously reported pBIOTRX vector into which was incorporated part of the pBluescript II+ phagemid multiple cloning site (MCS), amplified by PCR using a pair of sophisticated oligonucleotide primers. The functionality of these novel vectors was examined in this system by recombinant expression of rat transforming growth factor-β. Western-blot analysis using TRX-specific antibodies or peroxidase-conjugated streptavidin confirmed the successful induction of the fusion protein and the *in vivo* conjugation of biotin molecules, respectively. The convenience of molecular subcloning provided by the MCS and the effective *in vivo* biotinylation of proteins of interest makes this novel vector set an interesting alternative for the production of biotinylated proteins.

Keywords: *In vivo* biotinylation, Fusion protein, TGF-β, Thioredoxin.

Introduction

The biotinylation of recombinantly expressed proteins is a powerful method for the detection analysis of proteins in a variety of experimental systems (Updyke *et al.*, 1984; Gretch *et al.*, 1987). Because of its small size (244 Da) and extremely high affinity to avidin and streptavidin ($K_a = 10^{15}\ M^{-1}$), biotin has been used as the label of choice for the specific labeling of proteins. Until recently, the conjugation of biotin molecules to proteins of interest has been mainly performed *in vitro* using chemical modifying agents such as biotinyl-*N*-hydroxysuccinimide (NHS) ester and biotinyl-*N*-hydroxysulfosuccinimide (sulfo-NHS) esters (Bayer and Wilchek, 1990). However, limitations of this method, like the uncontrollable biotinylation of every lysine residue within the target protein and the harsh conditions required for the effective conjugation of biotin, have been a great hindrance to the full analysis of the functional and structural aspects of the biotinylated protein.

The identification of a short stretch of amino acids (13 residues) by peptide library screening (Schatz, 1993), which is used as substrate by the *E. coli* biotin holoenzyme synthetase, BirA (Barker and Campbell, 1981), allows for the controlled biotinylation of recombinant proteins *in vivo*. In fact, the fusion of the 13-mer BirA substrate sequence on either end of a given recombinant protein was found to result in the successful biotinylation of the peptide sequence without any modification of the protein itself (Schatz, 1993). Based on this observation, there has been a drastic increase of reports describing the powerful application of site-specifically biotinylated proteins in a variety of systems as diverse as the labeling of recombinant protein kinases (Tsao *et al.*, 1996) and the generation of MHC/peptide tetramers (Crawford *et al.*, 1998). Furthermore, a novel vector system termed pBIOTRX-BirA (Smith *et al.*, 1998) was recently described which allows for the recombinant expression and

* To whom correspondence should be addressed.
Tel: 82-42-860-4163; Fax: 82-42-860-4593
E-mail: kimkl@kribb4680.kribb.re.kr

in vivo biotinylation of proteins of interest as a fusion protein with thioredoxin (LaVallie *et al.*, 1993; Lu *et al.*, 1996), renowned for its ability to confer high solubility to recombinantly expressed fusion proteins (Sachdev and Chirgwin, 1998). Another special feature of this system is the inclusion of an extra *birA* gene in the expression vector, ensuring the quantitative biotinylation of the fusion protein. Despite the elegant design of this expression system, the actual molecular subcloning of cDNAs for the expression of proteins of interest has been hampered. Neither multiple-cloning sites (MCS) nor convenient single-cut restriction enzyme sites are available in this vector, which makes the generation of recombinant expression vectors very complicated. In the present study, the previously reported pBIOTRX-BirA system (GenBank accession no. AF044308) was redesigned and improved by destroying the *trxA* stop codon and replacing it with a multiple-cloning site inserted in-frame. The successful cloning of the novel vectors was confirmed by DNA sequencing of the multiple-cloning site. Finally, the functionality of this novel vector was demonstrated by the actual cloning and expression of rat transforming growth factor (TGF)-β cDNA using this system.

Materials and Methods

Materials The original vector for the expression of *in vivo* biotinylated thioredoxin proteins, pBIOTRX-BirA, was a generous gift from the Genetics Institute, Cambridge, USA. Plasmid DNA was isolated using a commercial kit (5 prime → 3 prime, Inc., Boulder, USA) and restriction enzymes were purchased from Posco Chem., Kyunggi-Do, Korea. PCR was performed in a GeneAmp™ PCR system 2400 thermal cycler (Perkin-Elmer, Foster City, USA) using conventional PCR reagents. Oligonucleotide primers were obtained from GenoTech Inc. (Taejon, Korea), and sequencing of the plasmids was performed with a dye terminator cycle sequencing kit from Perkin Elmer. The DNA sequence was analyzed with an ABI Prism Model 377 sequence analyzer at the Genome Center, Korea Research Institute of Bioscience and Biotechnology, Taejon, Korea.

Molecular cloning of the pBIOTRX-II expression vectors The pBIOTRX-BirA vector, which served as a template for the construction of the pBIOTRX-II vectors, encodes the expression of a thioredoxin gene fused to a short stretch of *in vivo* biotinylation sequences. The expression of the biotinylated fusion protein is controlled by the P_L promoter of bacteriophage λ, which in turn lies under the control of the cI repressor whose expression is inhibited by the addition of exogenous tryptophan. In the case of pBIOTRX-BirA, to subclone cDNA of interest into this vector, either PCR or the use of the single restriction enzyme site, *SfiI*, which is located just above the stop codon of the thioredoxin gene, is required. In the present study, the stop codon of thioredoxin was replaced by a multiple cloning site, which contains unique restriction sites for the convenient cloning of foreign genes. The insert fragment containing this multiple cloning site was generated by PCR using a pair of oligonucleotide primers, each containing a putative stop codon that is activated when placed as a 3′-downstream primer. In this way, the blunt end PCR fragment containing the multiple cloning site could be inserted in two different directions into the blunted *SfiI* site of the pBIOTRX-BirA vector. This resulted in either pBIOTRX-II (S/K) or (K/S) depending on the order of appearance of the restriction enzyme sites. For blunting the single-strand overhang after *SfiI* digestion, the reaction was incubated with T4 DNA polymerase for 15 min at room temperature, and then extracted with phenol/chloroform and the DNA was finally recovered by ethanol precipitation. Ligation, transformation, and identification of successfully cloned plasmids were performed using standard molecular biological techniques.

Moleclular subcloning of rat TGF-β cDNA into the pBIOTRX-II (S/K) vector cDNA encoding rat TGF-β was amplified by reverse transcription-PCR from the total RNA of concanavalin A-activated (Sigma, St. Louis, USA) whole splenocytes of adult Lewis rats. The sequences of the oligonucleotide primers used for amplification were; rat-TGF-up, 5′-tatagatctgccctggataccaactac-3′; rat-TGF-down, 5′-aatagtcgactcagctgcacttgcagga-3′. PCR was performed for 37 cycles with each cycle consisting of denaturation for 30 s at 96°C, annealing for 90 s at 56°C, and elongation for 60 s at 72°C. The PCR product was then purified using a 1.5% TAE-agarose gel and eluted using the QIAEX II gel extraction kit (Qiagen, Hilden, Germany). After digestion with *BglII* and *SalI*, the PCR fragment was inserted into the *BamHI*/*XhoI* site of the pBIOTRX-II (S/K) vector. The resulting vector was named pBIOTRX-II/TGF. Correct insertion of the cDNA was confirmed by restriction enzyme digestion.

SDS-PAGE and Western-blot analysis of recombinantly expressed biotinylated TRX-TGF-β fusion proteins For expression of the BIOTRX-TGF-β fusion proteins, *E. coli* strain GI826 (Invitrogen, Carlsbad, USA) transformed with the pBIOTRX-II/TGF vector was cultured overnight in RM base media (M9 salt supplemented with 2% casamino acids, 1% glycerol, 1 mM $MgCl_2$, and 100 μg/ml ampicillin) with continuous shaking at 30°C. The next day, an aliquot of this dense overnight culture was used to inoculate fresh IMC media (M9 salt supplemented with 0.2% casamino acids, 0.5% glucose, 1 mM $MgCl_2$, and 100 μg/ml ampicillin) containing biotin to a concentration of 100 μg/ml. The culture was then incubated at 37°C with continuous shaking to an OD_{600} of 0.5. Cells were then induced for expression of the recombinant proteins by the addition of tryptophan to a final concentration of 100 μg/ml, and then incubated for a further 4 h at 37°C. After harvesting, the cell pellet was resuspended in TE buffer and used for further analysis. SDS-polyacrylamide gel electrophoresis (PAGE) of the cell lysate of transformed and non-transformed GI826 cells was performed as described previously (Park *et al.*, 1998a), mainly following the method as described by Laemmli (1970).

Western-blot analysis was performed as had been described in a previous study (Park *et al.*, 1998b). In brief, whole cell lysate of transformed and non-transformed host cells were separated by 12.5% SDS-PAGE and then transferred onto a nitrocellulose membrane (Schleicher & Schuell, Keene, USA) by electroblotting. After blocking the membrane with 3% casein, the BIOTRX-TGF-β fusion proteins were then detected either with

peroxidase-conjugated streptavidin (Sigma), or with anti-TRX antiserum followed by peroxidase-conjugated anti-mouse IgG antibodies. The anti-TRX antiserum was developed in-house using BALB/c mice by immunization with recombinant thioredoxin. Specific binding was visualized by developing the blot with 4-chloro-1-naphtol (Sigma) and H_2O_2.

Results and Discussion

The object of this study was the generation of a versatile vector system for the expression of *in vivo* biotinylated recombinant proteins in *E. coli*. The previously described pBIOTRX-BirA system (Smith *et al.*, 1998) possesses all the features for the successful production of such proteins, but it is incomplete in that no polylinker nor unique restriction enzyme sites are available for the convenient subcloning of cDNA sequences into this vector. By analysis of the pBIOTRX-BirA sequence and the design of sophisticated primer sequences, a strategy was developed by which a multiple-cloning site could be introduced in-frame to the biotinylated thioredoxin without deletion of any residues from the original *trxA* gene, but encoding a new sequence-predictable fusion protein. The overall cloning strategy for the generation of the novel vector is shown in Fig. 1. In practice, the pBIOTRX-BirA plasmid was first cut with *Sfi*I, which recognizes the sequence 5′-GGCCNNN^NGGCC-3′ and which cleaves only once in the vector, just after the stop codon of the *trxA* gene (Fig. 1). The overhanging ends were then blunted by a T4-DNA polymerase reaction, which resulted in the destruction of the stop codon by deleting the last two nucleotides of the TAG stop. The opened plasmid was then dephosphorylated with shrimp alkaline phosphatase to prevent vector self-ligation. The insert, consisting of a partial polylinker region of the pBluescript II+ (K/S) vector, was prepared by amplification of this vector by PCR using the primers 5′-cttattgggtaccgggcccc-3′ and 5′-cttatggccgctctagaact-3′, and **Pfu**-DNA polymerase. The primers were designed in such a way that the insertion of the PCR fragment in either direction results in the in-frame transcription of the whole multiple-cloning site with an artificial stop codon at their ends (Fig. 2). Therefore, the induction of the novel vectors would result in the expression of a biotinylated TRX fusion protein with the addition of 33 amino acids to the C-terminus of the

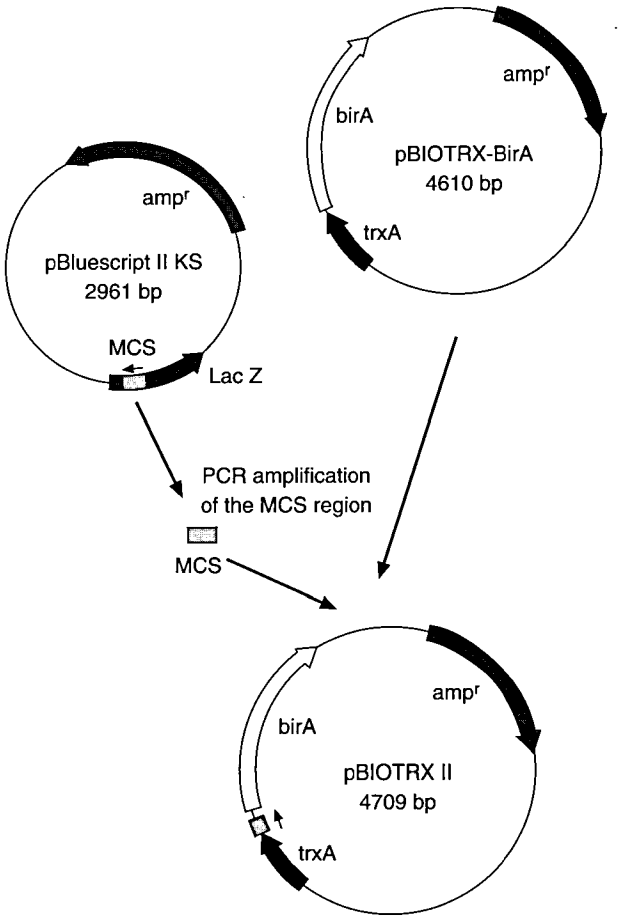

Fig. 1. Schematic presentation of the cloning strategy for the generation of the pBIOTRX-II fusion protein expression vectors.

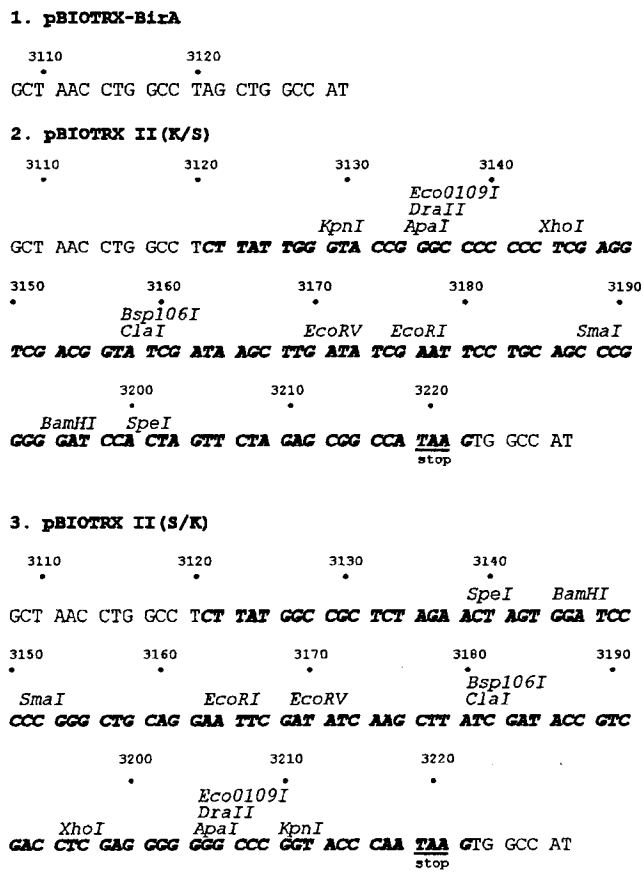

Fig. 2. Nucleotide sequence of the polyclonal site of the novel expression vectors in comparison to the original vector. Nucleotide sequences in italics represent the newly-inserted PCR fragment.

BIOTRX protein. Depending on the order of appearance of the restriction sites (*Spe*I or *Kpn*I) within the multiple-cloning site, the vectors were termed either pBIOTRX-II (K/S) or pBIOTRX-II (S/K).

To confirm the functionality of this expression system, we attempted to clone and express a foreign cDNA sequence in this system, which would result in the successful expression and *in vivo* biotinylation of the resulting protein. For this, partial cDNA encoding the mature rat TGF-β polypeptide was subcloned into the MCS of pBIOTRX-II (K/S) and transformed into GI826 *E. coli* cells for expression. TGF-β is a cytokine which consists of a disulfide-linked dimer of 112 amino acids each (Sporn and Roberts, 1992), and that has been implicated in playing a critical role in the mediation of a large number of biological responses (Roberts and Sporn, 1993). In particular, it is an important modulator of growth and differentiation of cells of the immune system, and shows both stimulatory and suppressive activities depending on the cell type and conditions. In fact, while TGF-β strongly suppresses the proliferation of T- and B-lymphocytes as well as inhibits NK cell activity (Hooper, 1991), it also induces the secretion of several growth factors in macrophages and up-regulates the secretion of IgA in B-cells under certain conditions (Lebman *et al.*, 1990). The precise cellular mechanism for these rather contradictory results is still not clear. However, since TGF-β is regarded to be heavily involved in the suppression of autoimmune diseases, as well as in the inhibition of host versus graft reaction and so on, further studies concerning the method of action of this cytokine are expected to have consequences in the treatment and prevention of such immunological diseases. In this regard, the generation of a site-specifically biotinylated TGF-β is of great interest in determining the distribution and regulation of cellular receptors specific for TGF-β, and thereby the mode of function in the activation and/or suppression of cellular activities.

After introduction of the pBIOTRX-II/TGF vector into GI826 cells, expression of the recombinant proteins was induced by the addition of tryptophan. Protein expression in pBIOTRX vectors is under the control of the P_L promoter of bacteriophage λ, which itself is regulated by the bacteriophage λ cI repressor that binds to the operator region just before the promoter. In the present system, the expression of the cI repressor is under the control of the *trp* promoter so that by addition of tryptophan, a tryptophan-*trp* repressor complex is formed that inhibits the synthesis of cI repressor proteins. This in turn releases the P_L promoter for expression of the transcription of the recombinant protein. This multi-step controlled regulation of transcription ensures tight control over the expression of recombinant proteins.

In Fig. 3 is shown the SDS-PAGE analysis of whole cell lysate of pBIOTRX-II/TGF transformed GI826 cells. A specific but faint band of about 25 kDa, which corresponds to the expected molecular weight of biotinylated TRX-TGF-β fusion proteins was observed in transformed cells. To confirm the identity of this protein and prove the successful *in vivo* biotinylation, Western-blot analysis was performed. The specific binding of anti-TRX antibodies as well as the biotin-specific streptavidin to the same band of induced proteins (Figs 4A, 4B), proves the validity and functionality of this novel expression vector system.

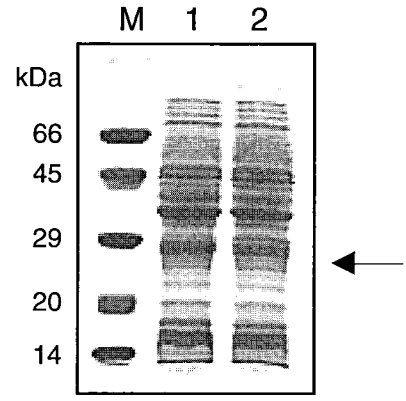

Fig. 3. SDS-PAGE analysis of whole cell lysate pBIOTRX-II/TGF transformed *E. coli*. M, molecular weight marker; 1, GI826 cells; 2, pBIOTRX-II/TGF transformed GI826 cells induced with tryptophan.

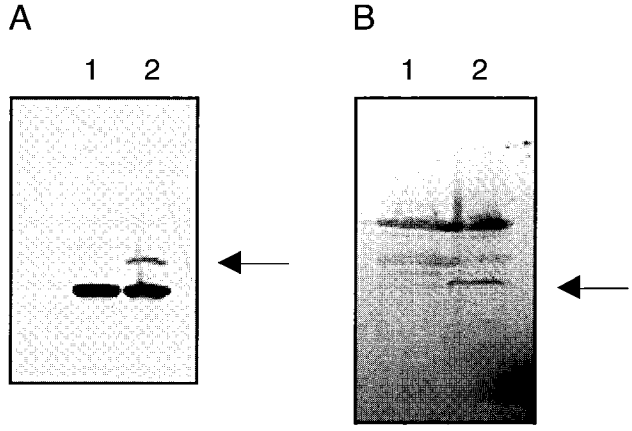

Fig. 4 Western-blot analysis of whole cell lysate of pBIOTRX-II/TGF transformed *E. coli*. **A.** Western-blot analysis of the recombinant BIOTRX-TGF-β fusion protein using peroxidase-conjugated streptavidin. 1, GI826 cells; 2, GI826 cells expressing BIOTRX-TGF-β. The other band detected by peroxidase-conjugated strepavidin is the only protein that is naturally biotinylated in *E. coli* by the BirA biotin holoenzyme synthetase (Schatz, 1993). **B.** Western-blot analysis of the recombinant BIOTRX-TGF-β fusion protein with anti-TRX antiserum. 1, GI826 cells; 2, GI826 cells expressing BIOTRX-TGF-β. The arrow indicates the position of the recombinant protein. Additional bands observed in the blot are nonspecifically detected throughout the various cell lysates, and are generated because of the polyclonal nature of the anti-TRX serum.

Judging from these results, the novel vectors as described in the present study are indeed functional, and the successful cloning of foreign cDNA into the newly generated MCS region also leads to the expression and *in vivo* biotinylation of these proteins. Some points still to be dealt with are the relative low expression level and the poor inducibility of these recombinant proteins. It is our belief that these difficulties are inter-related, and that the transcriptional leakage of non-induced culture is responsible for both the low expression level and bad induction. In fact, in most preparations of casamino acids, essential constituents of RM and IMC media, trace amounts of tryptophan contamination have been observed, which definitively leads to leakage. Such an effect was also observed by other groups (Lu, Z., Genetics Institute, personal communication), but this problem was overcome by using the highest grade of casamino acid for media preparation.

In summary, this novel vector set enables the convenient design and molecular cloning of foreign proteins in the form of *in vivo* biotinylated recombinant proteins in *E. coli*. The advantages of this novel set of expression vectors are obvious. The convenient cloning and a larger choice of cloning strategies will bring this vector system closer to a readily applicable system for recombinant expression of *in vivo* biotinylated proteins.

Acknowledgment This work was supported in part by a grant (BG420M) from the Korean Ministry of Health and Welfare.

References

Barker, D. F. and Campbell, A. M. (1981) The *birA* gene of *Escherichia coli* encodes a biotin holoenzyme synthetase. *J. Mol. Biol.* **146**, 451–467.

Bayer, E. A. and Wilchek, M. (1990) Protein biotinylation. *Methods Enzymol.* **184**, 138–160.

Crawford, F., Kozono, M., White, J., Marrack, P. and Kappler, J. (1998) Detection of antigen-specific T cells with multivalented soluble class II MHC covalented peptide complexes. *Immunity* **8**, 675–682.

Gretch, D. R., Suter, M. and Stinski, M. F. (1987) The use of biotinylated monoclonal antibodies and streptavidin-affinity chromatography to isolate herpes virus hydrophobic proteins or glycoproteins. *Anal. Biochem.* **163**, 270–277.

Hooper, W. C. (1991) The role of transforming growth factor-beta in hematopoiesis. A review. *Leuk. Res.* **15**, 179–184.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. and McCoy, J. M. (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology* **11**, 187–193.

Lebman, D. A., Lee, F. D. and Coffman, R. L. (1990) Mechanism for transforming growth factor-beta and IL-2 enhancement of IgA expression in lipopolysaccharide-stimulated B cell cultures. *J. Immunol.* **144**, 952–959.

Lu, Z., DiBlasio-Smith, E. A., Grant, K. L., Warne, N. W., LaVallie, E. R., Collins-Racie, L. A., Follettie, M. T., Williamson, M. J. and McCoy, J. M. (1996) Histidine patch thioredoxins. *J. Biol. Chem.* **271**, 5059–5065.

Park, J. H., Choi, E. A., Cho, E. W., Hahm, K. S. and Kim, K. L. (1998a) Maltose binding protein (MBP) fusion proteins with low or no affinity to amylose resins can be single-step purified using a novel anti-MBP monoclonal antibody. *Mol. Cells* **8**, 709–716.

Park, J. H., Cho, E. W., Lee, Y. J., Chung, J., Hahm, K. S. and Kim, K. L. (1998b) Generation of anti-HLA-DR4 specific antibodies by immunization of the recombinantly expressed allelic subtype-specific region of the HLA-DRB1*0405 molecules. *J. Biochem. Mol. Biol.* **31**, 111–116.

Roberts, A. B. and Sporn, M. B. (1993) Physiological actions and clinical applications of transforming growth factor-β. *Growth Factors* **8**, 1–9.

Sachdev, D. and Chirgwin, J. M. (1998) Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin *Prot. Exp. Purif.* **12**, 122–132.

Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *E. coli*. *Bio/Technology* **11**, 1138–1143.

Smith, P. A., Tripp, B. C., DiBlasio-Smith, E. A., Lu, Z., LaVallie, E. R. and McCoy, J. M. (1998) A plasmid expression system for quantitative *in vivo* biotinylation of thioredoxin fusion proteins in *Escherichia coli*. *Nucleic Acids Res.* **26**, 1414–1420.

Sporn, M. B. and Roberts, A. B. (1992) Transforming growth factor-beta: recent progress and new challenges. *J. Cell. Biol.* **119**, 1017–1021.

Tsao, K. L., DeBarbieri, B., Michel, H. and Waugh, D. S. (1996) A versatile expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in *Escherichia coli*. *Gene* **169**, 59–64.

Updyke, T. V. and Nicholson, G. L. (1984) Immunoaffinity isolation of membrane antigens with biotinylated monoclonal antibodies and immobilized streptavidin matrices. *J. Immunol. Meth.* **73**, 83–95.