Malaysian Cocoa Journal
Volume 13(2)/2021

OPTIMIZATION OF PROTEIN PURIFICATION PROTOCOL FOR ARTIFICIAL TRANSCRIPTION FACTORS LEC2

Rahman M. Z. A., Jainuddin H. and Kasran R.

Biotechnology Division, Centre for Cocoa Biotechnology Research, Malaysian Cocoa Board, Commercial Zone 1, Norowot Road, South KKIP, 88460, Kota Kinabalu, Sabah, Malaysia.
Corresponding author: hilmi@koko.gov.my

Malaysian Cocoa J. (2021) 13(2): 91-95

ABSTRACT – *Transcription factors (TFs) are essential for the expression of all proteins, including those involved in the plant regeneration and totipotency. LEAFY COTYLEDON 2 (LEC2) exerts significant impacts on determining embryogenic potential and various metabolic processes through a complicated genetic regulatory network and is sufficient to induce somatic embryo development in vegetative cells. Previously, five new TcLEC2 artificial transcription factors (ATFs) were successfully constructed and expressed using a prokaryotic system (pET100/D-TOPO). The recombinant proteins were separately purified by his-tagged protein purification using the ÄKTA pure protein purification system. However, the expressed ATFs were co-purified with the host's chromosomal DNA causing problems for subsequent application. Although this is the common issue for the case of DNA binding protein, an optimization of purification protocol should be conducted in order to obtain DNA-free purified protein samples. Optimization of binding and elution buffer's composition for his-tagged affinity chromatography had been conducted with additional washing steps to remove the bound DNA from the recombinant protein samples. Introduction of high salt washing step (20mM phosphate, 2M NaCl, 50 mM Imidazole, pH 7.4) before protein elution step successfully dissociated and removed the bound DNA from the targeted recombinant proteins.*

Keywords: LEAFY COTYLEDON 2, artificial transcription factor, protein purification, gene expression, affinity chromatography.

INTRODUCTION

High degree of genetic heterogeneity is undesirable in crop production as it is difficult to preserve the desired traits and loss of these traits can reduce productivity significantly. To overcome this issue, different techniques have been developed to bypass the need to propagate plants through seeds and instead produce plants asexually. One of the preferable methods for *cacao* propagation is via somatic embryogenesis (SE) (Maximova *et al.*, 2002). However, cacao SE is very time-consuming, low efficiencies, genotype dependence and low conversion rate, thus, still presents a significant limitation for its propagation at commercial scales. The processes that induce a cell to change its fate and enter an embryonic program of development are not known, nor is it known whether a common pathway is used to initiate embryonic development in all of these diverse cell types. LEAFY COTYLEDON 2 (LEC2) regulates many distinct aspects of embryogenesis (Meinke *et al.*, 1994)(Stone *et al.*, 2001) and encodes a transcription factor with a B3 domain, a DNA binding region found thus far only in plant proteins (D. Gaj, 2001; Santos Mendoza *et al.*, 2005; Braybrook *et al.*, 2006) which binds specifically to the RY motifs in the 5′ flanking regions of LEC2-induced genes (Giraudat *et al.*, 1992). LEC2 is exclusively expressed in developing zygotic embryos during both the early development and maturation phases. It is required for development and maintenance of suspensors and cotyledons and for the acquisition of desiccation tolerance and inhibition of premature germination (D. Gaj, 2001). The LEC genes are also involved in regulation of fatty acid biosynthesis and storage lipid deposition during embryo development. The seed specific overexpression of ZmLEC1 and BnLEC1 led to 35% and 20% increase in seed oil contents in maize and canola, respectively (Shen *et al.*, 2010; Tan *et al.*, 2011) Ectopic expression of AtLEC2 in Arabidopsis leaves resulted in the accumulation of seed specific fatty acids (C20:0 and C20:1) and increased the mRNA level of

oleosin (Stone *et al.*, 2001). Furthermore, a direct downstream target of AtLEC2, AtWRI1 is known to control fatty acid metabolism through interactions with key genes upstream in the pathway (Maeo *et al.*, 2009).

Attempts have been made to enhance somatic embryogenesis in *cacao* by the overexpression of the BABYBOOM (BBM) gene and LEC2 gene in transgenic *cacao* tissue by promoting the transition of somatic *cacao* cells from the vegetative to embryonic state (Zhang *et al.*, 2014). The expression of TcLEC2 was higher in dedifferentiated cells competent for somatic embryogenesis (embryogenic calli), compared to non-embryogenic calli. Transient overexpression of TcLEC2 in immature zygotic embryos resulted in changes in gene expression profiles and fatty acid composition. Ectopic expression of TcLEC2 in cacao leaves changed the expression levels of several seed related genes. The overexpression of TcLEC2 in cacao explants greatly increased the frequency of regeneration of stably transformed somatic embryos. TcLEC2 overexpressing cotyledon explants exhibited a very high level of embryogenic competency and when cultured on hormone free medium, exhibited an iterative embryogenic chain-reaction (Shires *et al.*, 2017). The first demonstration of a titratable control over somatic embryo formation in a commercially relevant plant, *T. ca*cao, was achieved using a dexamethasone activatable chimeric transcription factor. This four-fold enhancement in embryo production rate utilized a glucocorticoid receptor fused to an embryogenic transcription factor LEAFY COTYLEDON 2 (Shires *et al.*, 2017). Where previous T. cacao somatic embryogenesis has been restricted to dissected flower parts, this construct confers an unprecedented embryogenic potential to leaves.

The works on these transcription factors have shown very promising results and provide confirmation that transcription factors can be used to enhance SE without compromising plant development. To gain insight into the mechanisms by which cells change their fate and become embryogenic, we analyzed protein-DNA structure interaction of LEC2 cotyledon from cacao. In this study, we determined the protein structure of TcLEC2 by using a comparative modeling approach followed by molecular docking to uncover the interaction between LEC2 and targeted recognition sites on the promoter region of AGL15.

MATERIALS AND METHODS

Protein expression of TcLEC2 ATFs using prokaryotic expression system

The TcLEC2 ATFs (L1, L2, L6, L7 and L8) proteins were expressed by 0.1mM IPTG and 0.1% rhamnose induction in 2000ml *E. coli* KRX batch culture under optimal conditions at 37°C for 16 hours. The bacterial cultures were pelleted down at 10,000 x *g* for 15 min at 4°C.

Crude protein samples preparation of Synthetic TcLEC2 ATFs

The recombinant KRX cells were then subjected to cell lysis using Q500 Sonicator (QSonica, USA) with sonication pulse rate at 15 seconds ON, 45 seconds OFF for the total cycle duration of 10 minutes. The soluble fractions of the expressed fusion proteins were isolated by centrifuging the lysed samples at 10,000 x *g* for 20 min at 4°C.

Protein Purification of L1, L2, L6, L7 and L8 TcLEC2 ATFs

The soluble fractions of the expressed fusion proteins were subjected to affinity chromatography using 1ml HisTrap HP columns and were run using ÄKTA pure protein purification system. Optimization of binding and elution buffer's composition for his-tagged affinity chromatography were conducted using varying NaCl concentration from 0.5M to 2M. The eluted purified fusion proteins were analyzed by using 12% SDS-PAGE.

RESULTS AND DISCUSSION

Protein expression of TcLEC2 ATFs using prokaryotic expression system

The *E. coli* strain KRX was chosen as the expression host for producing the TcLEC2 ATF proteins because it has the attributes that make it a good protein expression strain. The ompT– and ompP– mutations eliminate one source of

proteolysis of overexpressed protein in *E. coli*. KRX incorporates a chromosomal copy of the T7 RNA polymerase gene driven by the rhamnose promoter (rhaPBAD) BAD most widely used protein expression systems by virtue of the well-defined promoter, which is completely independent of *E. coli* RNA polymerase promoters, and the rapid elongation rate exhibited by T7 polymerase, which is about five times that of *E. coli* RNA polymerase (Golomb and Chamberlin, 1974)). This promoter is subject to catabolite repression by glucose and activated by adding rhamnose BAD recombinant protein production. rhaPBAD-driven expression is positively controlled through a regulatory cascade of BAD rhamose, activating transcription from rhaPBAD. Since the isomerase (RhaA), kinase (RhaB) and aldolase BAD and is not consumed during growth.

In our protein expression experiment, when the cultures reach an O.D.600 of 0.5–0.6 the cultures were induced by 0.1mM IPTG and 0.1% rhamnose for 16 hours which assist the production of soluble expressed protein. When comparing the protein expression between induction with and without the rhamnose, the yield of the expressed protein increased up to 80% when induced with 0.1% rhamnose compared to the culture without the addition of rhamnose. Hence, rhamnose is crucial to drive the rhamnose promoter (rhaPBAD) for optimal protein expression, even though IPTG alone can drive the expression of the ATF proteins via T7 promoter.

Protein Purification of L1, L2, L6, L7 and L8 TcLEC2 ATFs

Previously, few attempts have been made to purify DNA-free TcLEC2 ATFs using conventional His-tagged affinity chromatography. However, the resulting purified proteins were still tightly bound with the host's chromosomal DNA and making the downstream application such as ChIP-seq difficult. An efficient method of removing the contaminated bound DNA from the prepared protein samples was explored with the aim of single step DNA removal technique.

Based on the previous finding reported by Vingadassalon *et al.* (2016), the contaminated DNA can be removed either using heparin column, ion exchange column or DNase such as benzonase. All the described techniques required additional steps in the pre or post treatment and resulted lost in protein yield. Therefore, in this study we try to develop a simple step yet efficient technique in removal of bound DNA and recovery of apo-DNA binding protein (Figure 1). In our optimized method for purifying DNA-binding protein, three washing steps were introduced in our affinity chromatography protocol. The first washing step is conducted as soon as the targeted protein bound to the nickel sepharose column that composed low salt concentration washing buffer (20mM phosphate, 0.5M NaCl, 50 mM Imidazole, pH 7.4) for 5 column volumes. The first washing step was intended to remove the unspecific bound protein to the affinity column and the contaminated host chromosomal DNA still tightly bound to the DNA binding protein. In the second washing step, high concentration of salt (2M NaCl) was incorporated in the second wash buffer (20mM phosphate, 2M NaCl, 50 mM Imidazole, pH 7.4) and the column was washed for 5 column volumes. High salt decreases non-specific binding between the protein of interest and DNA impurities by interrupting the electrostatic interactions between them. The final washing step before protein elution was using the same buffer composition as the first washing step which is a low salt wash buffer (20mM phosphate, 0.5M NaCl, 50 mM Imidazole, pH 7.4) and the column was washed for another 5 column volumes. The final washing step is important to equilibrate the column with appropriate salt concentration before the proteins were eluted from the column. The purified proteins for L1, L2, L6, L7 and L8 TcLEC2 ATFs were analyzed on SDS PAGE as shown in Figure 2.

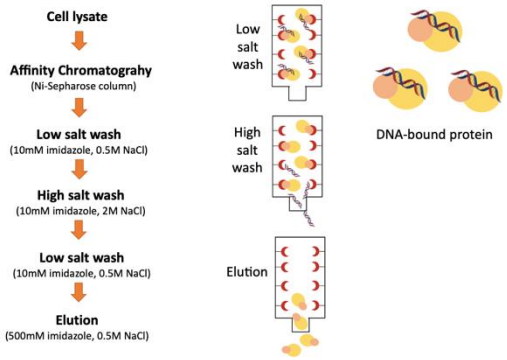

Figure 1. Overview and schematic diagram of optimized single step affinity chromatography for DNA-binding protein purification protocol.

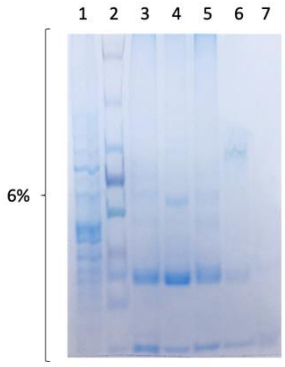

Figure 2. SDS PAGE of purified L1, L2, L6, L7 and L8 TcLEC2 ATFs. Lane 1 : Low range protein marker. Lane 2 : Broad range protein marker. Lane 3 : DNA-free purified L1 protein. Lane 4 : DNA-free purified L2 protein. Lane 5 : DNA-free purified L6 protein. Lane 6: DNA-free purified L7 protein. Lane 7: DNA-free purified L8 protein.

CONCLUSION

The synthetic TcLEC2 ATF constructs (L1, L2, L6, L7 and L8) were successfully expressed in *E. coli* KRX prokaryotic system under 0.1mM IPTG and and 0.1% rhamnose induction and were purified using his-tagged affinity protein chromatography. The protein purification protocol was optimized by introducing additional three washing steps with low salt and high salt concentration buffers to remove the contaminated host's chromosomal DNA from the purified protein. The DNA-free TcLEC2 ATF proteins will be further characterized with EMSA, ChIP-Seq and DNase footprinting assay in order to determine the interaction between ATFs and their corresponding targeted DNA.

ACKNOWLEDGEMENT

This work is supported by Developmental Fund (P20001011116013-Collection and Sequencing of Cocoa Genome) from the Ministry of Finance under the Eleventh Malaysia Plan.

REFERENCES

Braybrook, S.A., Stone, S.L., Park, S., Bui, A.Q., Le, B.H., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2006). 'Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis', *Proceedings of the National Academy of Sciences of the United States of America*. **103:** 3468–3473.

D. Gaj, M. (2001). 'Direct somatic embryogenesis as a rapid and efficient system for in vitro regeneration of *Arabidopsis thaliana*', *Plant Cell, Tissue and Organ Culture*. **64**: 39–46.

Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Parcy, F. and Goodman, H.M. (1992). 'Isolation of the Arabidopsis ABI3 Gene by Positional Cloning', *The Plant Cell Online*. **4:** 1251–1261.

Golomb, M. and Chamberlin, M. (1974). Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major in vitro transcripts by gel electrophoresis. *Journal of Biological Chemistry*. **249**: 2858–2863.

Maeo, K., Tokuda, T., Ayame, A., Mitsui, N., Kawai, T., Tsukagoshi, H., Ishiguro, S. and Nakamura, K. (2009). 'An AP2-type transcription factor, WRINKLED1, of *Arabidopsis thaliana* binds to the AW-box sequence conserved among proximal upstream regions of genes involved in fatty acid synthesis', *The Plant Journal*. **60:** 476–487.

Maximova, S. N., Alemanno, L., Young, A., Ferriere, N., Traore, A., & Guiltinan, M. J. (2002). Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L., *In Vitro Cellular & Developmental Biology - Plant*, 38: 252–259.

Meinke, D.W., Franzmann, L.H., Nickle, T.C. and Yeung, E.C. (1994). 'Leafy Cotyledon Mutants of *Arabidopsis*.', *The Plant cell*. **6:** 1049–1064.

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M. and Lepiniec, L. (2005). 'LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves', *FEBS Letters*. **579:** 4666–4670.

Shen, B., Allen, W.B., Zheng, P., Li, C., Glassman, K., Ranch, J., Nubel, D. and Tarczynski, M.C. (2010). 'Expression of ZmLEC1 and ZmWRI1 Increases Seed Oil Production in Maize', *Plant physiology*. **153:** 980–987.

Shires, M. E., Florez, S.L., Lai, T.S. and Curtis, W.R. (2017). 'Inducible somatic embryogenesis in *Theobroma cacao* achieved using the DEX-activatable transcription factor-glucocorticoid receptor fusion', *Biotechnology Letters*. **39:** 1747–1755.

Stone, S. L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001). 'LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development', *Proceedings of the National Academy of Sciences*. **98:** 11806–11811.

Tan, H., Yang, X., Zhang, F., Zheng, X., Qu, C., Mu, J., Fu, F., Li, J., Guan, R., Zhang, H., Wang, G. and Zuo, J. (2011). 'Enhanced Seed Oil Production in Canola by Conditional Expression of *Brassica napus* LEAFY COTYLEDON1 and LEC1-LIKE in Developing Seeds', *Plant Physiology*. **156:** 1577–1588.

Vingadassalon, A., Bouloc, P. and Rimsky, S. (2016). Removing nucleic acids from nucleoid-associated proteins purified by affinity column. *Journal of Biological Methods*. **3**:e35.

Zhang, Y., Clemens, A., Maximova, S.N., Guiltinan, M.J. (2014). 'The *Theobroma cacao* B3 domain transcription factor TcLEC2 plays a duel role in control of embryo development and maturation'. *BMC Plant Biology*. **14:** 106.