Cloning and expression of gene encoding P23 protein from Cryptosporidium parvum

**ABSTRACT**
We cloned the cp23 gene coding P23 (glyco)protein from Cryptosporidium parvum isolated from Thua Thien Hue province, Vietnam. The coding region of cp23 gene from C. parvum is 99% similar with cp23 gene deposited in NCBI (accession number: U34390). SDS-PAGE and Western blot analysis showed that the cp23 gene in E. coli BL21 Star™(DE3) produced polypeptides with molecular weights of approximately 37, 40 and 49 kDa. These molecules may be non-glycosylated or glycosylated P23 fusion polypeptides. Recombinant P23 protein purified by GST (glutathione S-transferase) affinity chromatography can be used as an antigen for C. parvum antibody production as well as to develop diagnostic kit for C. parvum.

**Key words:** cp23 gene, Cryptosporidium parvum, P23 protein

**Introduction**
Cryptosporidium parvum is a protozoan parasite that infects the intestinal epithelium of humans and animals causing the enteric disease. Cryptosporidiosis is a cause of life-threatening disease in immunodeficient people and it has been estimated that 10-15% of patients (or higher) died as a result of infection with Cryptosporidium in the world (Jenkins & Fayer, 1995; Wang et al., 2009).

To detect C. parvum infection, several methods were developed such as sedimentation-flotation technique (McNabb et al., 1985), oocysts staining (O'Donoghue, 1995), PCR amplification (Laxer et al., 1991). However, these approaches could not be applied widely because of its high cost (Wang et al., 2009). Rapid immuno-chromatographic assays for detecting infections with C. parvum in fecal samples were evaluated. The sensitivity (75%) and specificity (100%) of this assay were relatively high compare with modified Ziehl-Neelsen staining method (Klein et al., 2009; Luginbühl et al., 2005).

A 23 kDa (glyco)protein of C. parvum (P23 protein) is a sporozoite surface protein that is geographically conserved among isolated C. parvum strains. The P23 protein was identified as an antigen with neutralization-sensitive epitopes. Therefore, it is considered to be a potential candidate for the development of an effective vaccine against cryptosporidiosis, as well as for the diagnosis of cryptosporidiosis by ELISA since it is likely to detect C. parvum in various geographical regions (Perryman et al., 1999).

In this work, we report the cloning and expression of gene encoding P23 protein from C. parvum isolated in Thua Thien Hue province, Vietnam. Our recombinant P23 protein will be used as an antigen for production of anti-P23 antibody and to
develop diagnostic kit for *C. parvum* based on immunoassays in the future.

**Materials and Methods**

**Detection of *C. parvum* oocysts**

Bovine fecal samples were collected from barns in Thua Thien Hue province, Vietnam. Sample preparation for detection of *C. parvum* oocysts was performed as described by Wang et al. (2009).

The oocysts of *C. parvum* were detected by the modified Ziehl-Neelsen (MZN) method. The oocysts stained red on a pale green background and they appear as small discs with 4-6 μm in diameter, but the degree and proportion of staining varies with individual oocysts (World Organisation for Animal Health, 2008).

**DNA isolation and amplification of cp23 gene**

Genomic DNA of *C. parvum* was isolated by QIAamp DNA Stool Mini Kit (Qiagen, Max-Volmer-Straße, Hilden). Genomic DNA was then used as template in PCR amplification with the specific primers (Cp23F: 5'-ACGGATCCAAAAATGGGTTGTTT-3' and Cp23R: 5'-ACGGATCCTAATTTAGGCATCA-3'), which designed base on the 23 kDa protein synthetic region of the cp23 gene from NCBI (accession number: U34390). GGATCC sequence was designed for BamHI cloning site in pGEX 4T-3 expression vector.

The PCR component consisted of 2 µl DNA template, 12.5 µl 2× GoTag PCR master mix (Promega, Madison, WI), and 10 pmol each primer. Distilled water was added to a final volume of 25 µl. After 3 min of genomic DNA denaturation at 94°C, 40 cycles of 1 min for denaturation at 94°C, 45 seconds for annealing at 45°C and 1 min for polymerization at 72°C were carried out in the iCycler thermocycler (Bio-Rad, Hercules, CA). In the final cycle, the temperature of 72°C was held for an additional 7 min. The PCR products were purified by Quantum Prep™ Freeze N Squeeze DNA Gel Extraction Spin Columns Kit (Bio-Rad) after electrophoresis on 2% agarose gel.

**Cloning cp23 gene**

PCR products (cp23 gene) were ligated into pGEM®-T Easy vector (Promega), the ligation component consists of 50 ng vector, 5 µl buffer, 3 unit T4 DNA ligase, and 25 ng PCR product. Distilled water was added to a final volume of 10 µl. The ligation was incubated at 25°C for 1 h and at 4°C overnight. Ligation products were transformed into chemically competent *E. coli* DH5α cells by heat-shock method. The presence of the insert was determined by colony direct PCR followed by 2% agarose gel electrophoresis. Positive colonies were cultured on 5 ml of LB medium supplemented with 50 μg/ml ampicillin for biomass production. Recombinant vector (pGEM®-T Easy/cp23) was then isolated by PureLink™ Quick Plasmid Miniprep Kit (Invitrogen, Carlsbad, CA). The PCR product was sequenced by the dideoxyterminator method on the 3130 Genetic Analyzer (Applied Biosystems®).

**Expression of P23 protein in *E. coli***

The cp23 gene was cut from pGEM-T Easy/cp23 vector by BamHI, and purified by Wizard® SV Gel and PCR Clean-Up System Kit (Promega), they were then ligated into BamHI site of pGEX 4T-3 expression vector harboring glutathione S-transferase (GST) gene (GE Healthcare, Björkgatan, Uppsala). The ligation component consists of 0.5 µg vector, 5 µl 2× buffer, 4 µg DNA, 3 unit T4 DNA ligase, and distilled water was added to a final volume of 10 µl. The ligation was incubated at 4°C for overnight.

*E. coli* BL21(DE3) cells containing recombinant pGEX 4T-3/cp23 vector was cultured in 400 ml LB medium supplemented with 50 µg/ml ampicillin on a rotation shaker with speed of 200 rpm at 37°C to an OD₆₀₀ value (cell density) of 0.6. Isopropylthio-β-galactoside (IPTG) was added to a final concentration of 1 mM for induction and the culture was then incubated at 25°C for 4 h. The cell biomass was harvested by centrifugation at 10 000 rpm/4°C for 10 min, resuspended in 8 ml TNE buffer (50 mM Tris.HCl pH 7.5, 100 mM NaCl, 2 mM EDTA) and kept on ice bath for 1 hour. Total soluble protein was obtained after sonication for 5 min and centrifugation at 10 000 rpm/4°C for 5 min. Expression levels of P23 protein were determined by sodium dodecyl sulfate-12% polyacrylamide gel electrophoresis (SDS-PAGE, 12% w/v). The gel slice was then stained with Coomassie Blue R-250 and the image was analyzed by Quality One software (ver 4.1, Bio-Rad).

**Western blot**

Total soluble protein was separated on a SDS-12% PAGE and then electroblotted to a nitrocellulose membrane. Western blot was performed using WesternDotTM 625 Western Blot Kit (Invitrogen) according to the manufacturer’s instructions. Anti-P23 antibody was produced in mice by our laboratory (Immunology Laboratory, Institute of Biotechnology, Hue University). The image was collected under UV Transilluminator system (DyNA Light, Labnet,
Edison, NJ) and analyzed by Quality One software (ver. 4.1, Bio-Rad).

**Protein purification**

The P23 protein was purified by Glutathione Sepharose™ 4B Purification System (GE Healthcare) according to the manufacturer’s instructions. Eight milliliters of cell lysate was added to 2 ml prepared glutathione sepharose 4B and incubated for 30 min at room temperature, the mixture was then transferred to a 15-ml Falcon tube. The glutathione sepharose 4B resin was collected by centrifugation at 500× g for 5 min, washed by adding 5 ml 1× PBS and centrifuged at 500× g for 5 min again, repeat washing step 3 times. P23 protein was eluted by adding 1 ml elution buffer, incubated at room temperature for 30 min and centrifuged at 500× g for 5 min, repeat eluting step 2 times. Eluted protein from each step was mixed to run SDS-PAGE.

**Results**

**Cloning cp23 gene**

According to Perryman et al (1996), the cDNA of cp23 gene has a full-length of 602 bp with a coding region of 336 bp (from nucleotide 111 to 446). The coding frame of cp23 gene has 345 bp in length including two BamHI ends was amplified by specific primers.

Genomics DNAs from 10 oocysts of C. parvum were used as DNA templates for cp23 gene isolation by PCR amplification. Figure 1 showed bands were of expected size of approximately 350 bp, this size is similar with the coding frame of cp23 gene.

**Expression of cp23 gene**

The cp23 gene was cut from recombinant pGEM T-Easy/cp3 vector by BamHI and inserted into pGEX 4T-3 expression vector. The pGEX 4T-3/cp23 vector was transformed into E. coli BL21(DE3) cells and positive colonies from selection culture were checked by PCR. Our result showed that a DNA band occurred with the expected size of 345 bp (Figure 4), this suggested that the pGEX 4T-3/cp23 vector successfully transformed into E. coli cells.

Induction was then performed at low temperature of 25°C to avoid formation of inclusion bodies (Makrides, 1996). Expression of the cp23 gene was expected to produce protein of approximately 49 kDa (23 kDa of P23 protein and 26 kDa of glutathione S-transferase (GST) tag fusion fragment). The results of SDS-PAGE and Western blot showed that there are three protein bands with molecular weights approximate of 49, 40 and 37 kDa (Figure 5).
In *C. parvum*, the *cp23* gene encodes a polypeptide chain with 111 amino acids (molecular weight approximate of 11.3 kDa). After P23 protein was synthesized, a *N*-glycosylation was performed to produce a glycoprotein with molecular weight approximate of 23 kDa (or 27 kDa in some cases based on electrophoresis condition) (Priest et al., 1999; Wang et al., 2009). In this study, the P23 polypeptide chain may be glycosylated or not. If it was glycosylated, molecular weight (fusion protein) will be approximate of 49 kDa or 40 kDa (glycosylated in a different form); if not, molecular weight (fusion polypeptide) will be approximate of 37 kDa (11.3 kDa of P23 polypeptide and 26 kDa of GST tag).

In some previous reports, P23 glycoprotein from *C. parvum* were expressed successfully in *E. coli* based on different expression systems such as pET 30a(+) (Liu et al., 2010), pGEX 4T (Bonafonte et al., 2000), pGEX 4T-2 (Borad et al., 2012; Priest et al., 1999), or pGEX-5X-2 vector (Shahbazi et al., 2009; Shayan et al., 2008). Another host also was used for expression of P23 protein to be *Toxoplasma gondii* (Shirafuji et al., 2005). All of them are glycosylated forms with different molecular weight, i.e. 43 kDa (Shahbazi et al., 2009) or 46 kDa (Bonafonte et al., 2000; Wang et al., 2009) (fusion protein), 23 kDa (Inpankaew et al., 2009; Shahbazi et al., 2009; Shirafuji et al., 2005) or 27 kDa (Liu et al., 2010; Priest et al., 1999) (non-fusion protein). The difference in non- or fusion proteins indicated that P23 protein may be glycosylated in different forms. Other antigenic *C. parvum* protein also expressed in *E. coli* to be CP15 (Jenkins & Fayer, 1995).

The lysate of induced *E. coli* BL21(DE3) cells containing pGEX-4T-3/*cp23* vector were purified by GST chromatography using glutathione sepharose 4B resin. The results of SDS-PAGE showed two strong protein bands have molecular weights of approximately 37 kDa and 49 kDa, another one is very weak (appx. 40 kDa) (Figure 6).

In *C. parvum*, the *cp23* gene encodes a polypeptide chain with 111 amino acids (molecular weight approximate of 11.3 kDa). After P23 protein was synthesized, a *N*-glycosylation was performed to produce a glycoprotein with molecular weight approximate of 23 kDa (or 27 kDa in some cases based on electrophoresis condition) (Priest et al., 1999; Wang et al., 2009). In this study, the P23 polypeptide chain may be glycosylated or not. If it was glycosylated, molecular weight (fusion protein) will be approximate of 49 kDa or 40 kDa (glycosylated in a different form); if not, molecular weight (fusion polypeptide) will be approximate of 37 kDa (11.3 kDa of P23 polypeptide and 26 kDa of GST tag).

In some previous reports, P23 glycoprotein from *C. parvum* were expressed successfully in *E. coli* based on different expression systems such as pET 30a(+) (Liu et al., 2010), pGEX 4T (Bonafonte et al., 2000), pGEX 4T-2 (Borad et al., 2012; Priest et al., 1999), or pGEX-5X-2 vector (Shahbazi et al., 2009; Shayan et al., 2008). Another host also was used for expression of P23 protein to be *Toxoplasma gondii* (Shirafuji et al., 2005). All of them are glycosylated forms with different molecular weight, i.e. 43 kDa (Shahbazi et al., 2009) or 46 kDa (Bonafonte et al., 2000; Wang et al., 2009) (fusion protein), 23 kDa (Inpankaew et al., 2009; Shahbazi et al., 2009; Shirafuji et al., 2005) or 27 kDa (Liu et al., 2010; Priest et al., 1999) (non-fusion protein). The difference in non- or fusion proteins indicated that P23 protein may be glycosylated in different forms. Other antigenic *C. parvum* protein also expressed in *E. coli* to be CP15 (Jenkins & Fayer, 1995).

The lysate of induced *E. coli* BL21(DE3) cells containing pGEX-4T-3/*cp23* vector were purified by GST chromatography using glutathione sepharose 4B resin. The results of SDS-PAGE showed two strong protein bands have molecular weights of approximately 37 kDa and 49 kDa, another one is very weak (appx. 40 kDa) (Figure 6).
Conclusion

A cp23 gene coding P23 protein from *C. parvum* was cloned and its sequence has 99% similarity with the cp23 gene from NCBI. The cp23 gene in *E. coli* BL21(DE3) produced polypeptides with molecular weights of approximately 37, 40 and 49 kDa; they may be non-glycosylated and glycosylated fusion polypeptides. Western blot analysis showed that 3 fusion polypeptide chains have a GST tag and an epitope for immunological response. Protein purification using GST affinity chromatography was harvested two strong bands (37 and 49 kDa).

Acknowledgement

This study was funded by the Ministry of Science and Technology of Vietnam (No. 02/DTDL/2011-T/29/HĐ).

References


