Molecular cloning and characterization of an antigenic protein with a repeating region from *Clonorchis sinensis*

Tae Yun KIM¹, Shin-Yong KANG¹, Il-Young Ahn¹, Seung-Yull CHO² and Sung-Jong HONG¹*  

Department of Parasitology¹, Chung-Ang University Faculty of Medicine, Seoul 156-756,  
Section of Molecular Parasitology², Department of Molecular Cell Biology,  
Sungkyunkwan University School of Medicine, Suwon 440-746, Korea

Abstract: In the course of immunoscreening of *Clonorchis sinensis* cDNA library, a cDNA CsRP12 containing a tandem repeat was isolated. The cDNA CsRP12 encodes two putative peptides of open reading frames (ORFs) 1 and 2 (CsRP12-1 and -2). The repetitive region is composed of 15 repeats of 10 amino acids. Of the two putative peptides, CsRP12-1 was proline-rich and found to have homologues in several organisms. Recombinant proteins of the putative peptides were bacterially produced and purified by an affinity chromatography. Recombinant CsRP12-1 protein was recognized by sera of clonorchiasis patients and experimental rabbits, but recombinant CsRP12-2 was not. One of the putative peptide, CsRP12-1, is designated CsPRA, proline-rich antigen of *C. sinensis*. Both the C-termini of CsRP12-1 and -2 were bacterially produced and analysed to show no antigenicity. Recombinant CsPRA protein showed high sensitivity and specificity. In experimental rabbits, IgG antibodies to CsPRA was produced between 4 and 8 weeks after the infection and decreased thereafter over one year. These results indicate that CsPRA is equivalent to a natural protein and a useful antigenic protein for serodiagnosis of human clonorchiasis.

Key words: *Clonorchis sinensis*, repetitive peptide, antigenic, proline-rich, CsPRA

INTRODUCTION

About 30-40% of eukaryotic genome is composed of repetitive sequences. The repetitive sequences have been understood as functional components of eukaryotic genome representing genome plasticity and adaptability and contributing to genomic evolution (Brosius and Tiedge, 1996; Kidwell and Lisch, 1997; Levin, 1997). Although the specific function of the repeats remains to be elucidated, they have been applied to determine phylogenetic relationships (Hamilton et al., 1990; Minnick et al., 1992), and to provide information about the organization of genomes (Porter, 1994; Connolly et al., 1995).

However, the genomic structures of parasites have not been intensively studied compared to those of other eukaryotes and prokaryotes. The genome of *Schistosoma mansoni* contains short repetitive DNA elements designated Smα elements (Spotila et al., 1989). The elements homologous to RNA polymerase III promoter are thought to disperse throughout the genome. Tandemly repeated DNA elements were cloned and
sequenced from *Echinococcus granulosus*, which is clustered in the genome (Rosenzvit et al., 1997). The repetitive elements showing sex-specific polymorphism were identified from *S. mansoni* (Quack et al., 1998).

In diagnosis of parasitic human infections, subtelomeric tandem repeat sequence specific to *Leishmania braziliensis* showed usefulness in diagnosis for the parasitic infection by PCR and endonuclease digestion (Fu, 1998). From *Clonorchis sinensis*, the gene containing a repetitive sequence, Cs44, was identified and its gene product was evaluated as a potent antigenic protein for serodiagnosis of human clonorchiasis (Yong et al., 1998).

The diagnosis of clonorchiasis mainly depends on the detection of eggs in the stool. However, the difficulty of collecting stools from patients in the field have urged the researchers to recruit another diagnostic tests such as serodiagnosis. Among them, the enzyme-linked immunosorbent assay (ELISA) has been employed as an effective serodiagnostic technique due to its specificity and sensitivity (Hahm et al., 1984; Yong et. al., 1996). Studies on appropriate immunodiagnostic techniques should be paralleled to search specific and sensitive antigenic proteins. The specific and sensitive antigenic proteins of *C. sinensis* have been studied for several decades but antigenic molecules superior to the crude worm extract have not been identified yet.

In this study, to identify genes encoding antigenic proteins, an expression cDNA library of adult *C. sinensis* was screened with a clonorchiasis patient serum. A clone with repetitive sequences was purified and its gene product produced bacterially was isolated and characterized. It was sensitive and specific for serodiagnosis of human clonorchiasis.

**MATERIALS AND METHODS**

**Cloning a repetitive cDNA**

A cDNA expression library of adult *C. sinensis* constructed previously (Hong et al., 2000) was used. The cDNA library, 6 x 10^6 pfu, was mixed with *Escherichia coli* XL1-Blue and cultured on a LB-agar plate. The plate was overlaid with a nitrocellulose membrane (BioTrace™ NT; Gelman Science, Ann Arbor, MI, USA) soaked previously in 10 mM isopropyl-D-thiogalactoside (IPTG), and incubated further at 37°C for 4 hr. The membrane alone was then incubated for 3 hr within a *C. sinensis*-infected human serum diluted by 1:100 at room temperature, followed by incubation with secondary antibody, alkaline phosphatase-conjugated anti-human IgG diluted by 1:2,000. A color signal was developed with the addition of 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium (BCIP/NBT; Sigma Chemical Co., St. Louis, MO, USA) solution. Positive clones were purified through secondary and tertiary screenings using the same serum. The purified clones were in vivo excised to single-stranded phagemids by employing a helper phage (EXAssist™, Stratagene, La Jolla, CA, USA) and converted to double-stranded plasmids by following manufacturer’s instruction.

**cDNA sequencing and data base search**

The cDNA sequences were determined by using BigDye™ Terminator Cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) at Korea Basic Science Institute (Taejeon, Korea). Briefly, plasmid DNA and primers were added to the BigDye mixture, and thermal cycling was carried out by 25 cycles at 96°C for 30 sec, at 50°C for 15 sec and at 60°C for 4 min. After the removal of unincorporated nucleotides, the products were denatured by heating at 90°C for 2 min and electrophoresed on an automatic DNA sequencer (Model 377, Applied Biosystems). To get overlapping and both strand sequences, specific internal oligonucleotides (Fig. 1) were synthesized at Bioneer Co. (Chungwon-gun, Korea) and used in the cycle DNA sequencing. Putative peptide sequence of the cDNA was analysed using DNASIS program (Hitachi Software Engineering Co., Yokohama, Japan). Sequence homology of the cDNA clone was searched using GenBank database by BLAST program (Altschul et al., 1997). Multiple alignments of peptide sequences were generated by CLUSTAL W program (Thompson et al., 1994) and were optimized manually in part.
Production of recombinant CsRP12-1 and CsRP12-2 proteins

To produce recombinant proteins corresponding to the open reading frames (ORFs) 1 and 2 of the CsRP12 cDNA, coding region of CsRP12 was enzymatically cut and ligated into an expression vector pRSET™ A or B (Invitrogen, Carlsbed, CA, USA) with respect to the frame set of each respective ORF. CsRP12 cDNA was digested with EcoRI endonuclease and run on 1% agarose gel. The cDNA fragment corresponding to the coding region was excised and recovered from the gel slice by gene cleaning method. pRSET A and B were also digested and recovered as aforementioned. The recovered EcoRI-fragment of CsRP12 was ligated into pRSET I-fragment of CsRP12 was ligated into pRSET A or B by using T4 DNA ligase. Bacterial host cells, E. coli XL-1 Blue, were transformed and in frame connection of CsRP12 ORF 1 and 2 to the tag peptide of the respective vector was confirmed by DNA sequencing. The plasmids, pRSET A and B constructs, were then transformed into E. coli BLR(DE3)pLysS (Novagen, Madison, WI, USA) by heat shock method. Positive colonies were grown to A_{600} = 0.6 and expression of the recombinant fusion proteins were induced by adding IPTG at 1 mM final concentration and further incubation at 37°C for 4 hr. The bacterial pellet harvested by spinning the culture medium was resuspended in mouse toxicity phosphate-buffered saline (15 mM NaCl, 160 mM NaH_2PO_4, 40 mM Na_2HPO_4, 0.5 mM phenylmethylsulfonyl fluoride), and electrophoresed in reducing condition on 13.5% SDS-polyacrylamide gel. Localization of the expressed fusion proteins were checked by Coomassie Brilliant blue staining and by immunoblotting using anti-Xpress™ mouse IgG (Invitrogen) as a primary antibody that specifically recognize the tag peptide of the pRSET vectors.

Since the expressed protein is localized in cytosolic fractions as an insoluble form, the recombinant fusion proteins were purified by an affinity chromatography employing nickel-nitrotriacetic acid (Ni-NTA) resin (Qiagen GmbH, Hilden, Germany) under denaturation condition according to the manufacturer’s instruction. The bacterial pellets were resuspended in buffer A (6 M Guanidine hydrochloride, 0.1 M NaH_2PO_4, 10 mM β-mercaptoethanol, 0.01 M Tris, pH 8.0). The suspension was spun, and supernatant was removed and loaded onto Ni-NTA column. The column was sequentially washed out with buffers B and C (8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris, pH 8.0 and 6.3, respectively). The bound fusion protein was eluted with buffers D and E (8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris, pH 5.9 and 4.5, respectively) containing 250 mM imidazole. The fractions collected were dialyzed exhaustively 3 times for 3 hr each against distilled water and deployed by SDS-PAGE. The fusion proteins were detected by immunoblotting with anti-Xpress IgG (Invitrogen).

Flukes and sera

Clonorchis sinensis metacercariae were collected from Pseudorasbora parva caught in Chinju, Kyeongsangnam-do and administered orally, 500 metacercariae each, to four New Zealand White rabbits. Adult C. sinensis were recovered from bile ducts of the experimental rabbits 6 months after the metacercarial infection. The recovered flukes were used immediately, or stored at -70°C. Sera were collected from the experimental rabbits at 1-6 week interval over one year, and stored at -20°C until used. After one year, the rabbits were sacrificed and adult worms were recovered. Sera was collected from 35 clonorchiasis and 5 paragonimiasis patients proven parasitologically infected with the respective flukes. Cysticercosis sera were obtained from 5 patients diagnosed by CT/MRI imaging. Sparganosis sera were collected from 5 patients proven by surgical removing of the worm(s).

Antigenicity of recombinant CsRP12-1 and CsRP12-2 proteins

The purified recombinant proteins (CsRP12-1 and -2) were separated on 13.5% SDS-PAGE and electrotransferred onto nitrocellulose membranes (Hybond™-C, Amersham Pharmacia Biotech, Uppsala, Sweden). The membranes were immunoblotted with sera of helminth-infected humans and experimental
rabbits at 1:200 dilution. Alkaline phosphatase conjugated anti-human or anti-rabbit IgG antibodies were used as a secondary antibody at 1:2,000 dilution. Color was developed within BCIP/NBT substrate (Sigma Chemical Co.).

**Production and analysis of recombinant C-terminus peptides.**

The DNA fragment corresponding to ORFs-1 and -2 after the repetitive region of CsRP12 cDNA were prepared by PCR employing synthetic oligonucleotides containing *Bgl* II and *Sal* I endonuclease recognition sequences. The PCR product was introduced into Topo™ TA vector (Invitrogen) and amplified bacterially. The inserts excised by *Bgl* II and *Sal* I double-digestion were ligated into an expression vector pQE™ 40 (Qiagen GmbH). The ligation constructs, designated CsRP12ctm1 and CsRP12ctm2, were screened and amplified bacterially. *E. coli* M15™ (Qiagen GmbH) was transformed with the ligation constructs, and the recombinant CsRP12ctm1 and CsRP12ctm2 proteins were overexpressed by IPTG induction and purified by Ni-NTA affinity chromatography under denaturing condition.

The purified recombinant CsRP12ctm1 and CsRP12ctm2 proteins were electrophoresed and transferred onto nitrocellulose membrane. The membrane was stripped and immunoblotted with an anti-RGS His antibody (Qiagen GmbH) recognizing the tag peptide of the recombinant proteins, or with *C. sinensis*-infected patient sera as aforementioned.

**RESULTS**

**cDNA sequence of CsRP12**

Among the 8 purified positive clones, one clone, designated CsRP12, was selected for further analysis. The CsRP12 contained a cDNA insert of 826 bp with poly (A) tail. The CsRP12 encoded two putative proteins corresponding to ORF-1 and ORF-2 of 198 and 233 amino acids, respectively. Both putative proteins comprised an N-terminus peptide of 22 amino acids, a repetitive region of 15 repeats of 10 amino acids and a C-terminus peptide (Fig. 1). Molecular mass of CsRP12-1 and -2 was estimated to be 19.3 and 27.2 kDa, respectively. The sequence data reported herein has been submitted to GenBank and assigned to accession number, AF343876.

A homology search with CsRP12 cDNA reported several polypeptides homologous with CsRP12-1 peptide with a moderate sequence identities (Fig. 2), while CsRP12-2 peptide did not. The CsRP12-1 peptide showed the highest sequence identity and secondary structural similarity to a protein, *C. sinensis* antigen 44 (Cs44; Yong et al., 1998). The sequence homology between CsRP12-1 and Cs44 in N-terminus and C-terminus peptides were strikingly high 63% and 71%, respectively.

In northern blot analysis using *EcoR* I-digested CsRP12 cDNA fragment as a probe, a signal was produced at approximately 1.0 kb (data not shown).

**Recombinant CsRP12-1 and -2 proteins**

As both recombinant CsRP12-1 and -2 proteins were localized in cytosolic insoluble fraction in bacterial host BLR(DE3)pLysS, both of the recombinant fusion proteins were purified by Ni-NTA affinity chromatography under denaturation condition. The recombinant fusion proteins CsRP12-1 and -2 were efficiently eluted as 37 and 30 kDa molecules.
respectively (Fig. 3) with an elution buffer, pH 5.9 containing 250 mM imidazole. Immunoblots employing anti-Xpress antibody revealed signals specific to the fusion proteins.

**Fig. 2.** A multiple alignment of Clonorchis sinensis repetitive peptide, CsRP12-1, with surface proteins of animals. Residues identical to CsRP12-1 were shaded. A gap (–) was introduced for maximal alignment. Cs44, C. sinensis antigen 44 (Yong et al., 1998); XP2pre, Xenopus laevis skin secretory protein XP2 precursor (Hauser et al., 1992); CePro, Caenorhabditis elegans proline and glycine rich protein (Wilson et al., 1994); TcSA, Trypanosoma cruzi surface antigen (Buschiazzo et al., 1992); LdA2pre, Leishmania donovani amastigote-specific protein A2 precursor (Charest and Matlashewski, 1994); PfSAg, Plasmodium falciparum surface antigen protein precursor (Brown et al., 1987).

**Antigenic property of recombinant CsRP12-1 and -2 proteins**

Recombinant CsRP12-1 and -2 proteins produced bacterially was evaluated for its antigenicity to the helminth-infected patient sera, and the result was summarized in Table
1. All of the clonorchiasis patient sera reacted to recombinant CsRP12-1 protein (Fig. 4). Cross-reactivity of the recombinant CsRP12-1 appeared to be 12-21% against the tissue invading helminth-infected patient sera. However, the signals in immunoblottings were marginal for positive. The recombinant CsRP12-2 protein did not react to clonorchiasis patient sera nor to any helminth-infected patient sera (Table 1).

Table 1. Seroreactivity of recombinant CsRP12 proteins to various helminth-infected patient sera.

<table>
<thead>
<tr>
<th></th>
<th>CsRP12-1</th>
<th>CsRP12-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of sera</td>
<td>Positive sera (%)</td>
<td>Positive sera (%)</td>
</tr>
<tr>
<td>Clonorchiasis</td>
<td>35</td>
<td>35 (100%)</td>
</tr>
<tr>
<td>Paragonimiasis</td>
<td>19</td>
<td>4 (21%)</td>
</tr>
<tr>
<td>Sparganosis</td>
<td>17</td>
<td>2 (12%)</td>
</tr>
<tr>
<td>Cysticercosis</td>
<td>17</td>
<td>3 (18%)</td>
</tr>
<tr>
<td>Normal control</td>
<td>2</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

Antibodies reactive to CsRP12-1 in C. sinensis-infected rabbits

As shown in Fig. 5, IgG antibodies reacting to the recombinant CsRP12-1 were produced between week 4 and 8 after the metacercarial infection. The IgG antibody level decreased steadily from week 8 to week 52 after the infection.

Antigenic property of recombinant CsRP12 C-terminal peptides.

Both of the recombinant C-terminus peptides, CsRP12ctm1 and CsRP12ctm2, were found in insoluble forms in E. coli lysate and thereafter purified by Ni-NTA affinity chromatography under denaturation condition (Fig. 6A). The immunoreactivities were analyzed by immunoblotting with C. sinensis-infected patient sera and anti-RGS His antibody. Of seven C. sinensis-infected sera, only one reacted with the recombinant CsRP12ctm1 protein. Recombinant CsRP12ctm2 protein did not produce any positive signal against the sera (Fig. 6B).

DISCUSSION

A cDNA, CsRP12, was cloned by immunoscreening the adult C. sinensis cDNA library. Of the two putative peptides deduced from the CsRP12, CsRP12-1 corresponding to ORF-1 showed the highest homology to a glycine-rich C. sinensis protein (GRCSP; Yang et al., 2000) and a wide range of sequential identities with various proteins of invertebrate animals (Fig. 2). The homologous proteins
contain the repetitive peptides rendering sequential identities to CsRP12-1. Of the two bacterially produced recombinant proteins of CsRP12, CsRP12-1 was recognized by the sera of C. sinensis-infected patients and experimental rabbits, but CsRP12-2 was not. These results indicate that the peptide corresponding to CsRP12-1 is the natural protein occurring in C. sinensis. With superior antigenicity and proline-richness in the repeating unit, the CsRP12-1 protein designates CsPRA (proline-rich antigen of C. sinensis) in this report.

The peptide structure of CsPRA is different from GRCSP in amino acid composition of the repeating units, proline-rich versus glycine-rich and number of repetition, 15-times versus 23 times (Yang et al., 2000). However, CsPRA and GRCSP share high sequential identities in their N- and C-terminus regions. The homologous proteins to CsPRA are identified to be surface-associated and antigenic. The proteins homologous to CsPRA were identified to be the surface proteins of the respective organisms. In this respect, it is assumed that CsPRA may be associated with surface membrane of C. sinensis.

The recombinant CsPRA protein, although its N-terminus was truncated, was a good antigen showing high sensitivity and specificity toward clonorchiasis patient sera. A counterpart, GRCSP, which shows the highest homology to CsPRA, showed a high antigenic specificity and moderate sensitivity for serodiagnosis of clonorchiasis (Yong et al., 1998). In protozoan and helminthic parasites, proteins containing repetitive peptides have been identified to be highly antigenic with acceptable sensitivities and specificities for serological diagnosis of the respective infections (Imboden et al., 1995; Charest and Matlashewski, 1994; Burns et al., 1993; Marin et al., 1992). Both of the recombinant C-terminal peptides of CsPRA, CsRP12ctm1 and CsRP12ctm2, were not antigenic toward C.
indicating that antigenic epitope of CsPRA protein is most probably in the repetitive region. In many antigenic proteins, the repetitive regions are found to contain the major antigenic determinant(s) (Kulane et al., 1999; Banic et al., 1998; Burns et al., 1993) when compared to less antigenic N- and C-terminus regions (Lopez et al., 1996; Sharma et al., 1996). The repetitive amino acid sequences are variable in length and the number of repeats and show low sequence similarities with each other. The variability of the repetitive peptides has been recognized to render antigenic specificities of the protein (Kulane et al., 1999; Imboden et al., 1995).

Interestingly, when kept at room temperature, the purified CsPRA was readily degraded showing a ladder-like feature in SDS-polyacrylamide gel (data not shown). The ladder pattern was highly ordered by 15 fragments equal to the number of repeats. The PEST sequences (proline, glutamate, serine and threonine-rich sequence) involved in protein degradation (Rogers et al., 1986)

\[
\begin{align*}
&\text{CsRP12ctm1} \\
&\text{CsRP12ctm2}
\end{align*}
\]

\[
\begin{align*}
&\text{kDa} \\
&43 - 30 - 20 -
\end{align*}
\]
creating structural domains recognized by proteolytic enzymes were not found in the CsRP12 protein sequence. Moreover, since proteinase inhibitor cocktail was added immediately after the purification, the degradation of recombinant CsPRA is assumed to be arisen from auto-proteolysis. This may support the smearing under the CsPRA major band in immunoblotting reacted with clonorchiasis patients sera (Fig. 3).

ACKNOWLEDGEMENT

This work was supported by the grant from the Health Technology Planing and Evaluation (HMP-96-M-2-1059), the Ministry of Health and Social Affairs, the Republic of Korea.

REFERENCES


Liu DW, Kato H, Kazuo S (1996) Isolation of a cDNA encoding an antigenic polypeptide containing repeating units of Spirometra...


