Overexpression and Purification of PreS Region of Hepatitis B Virus Antigenic Surface Protein adr Subtype in *Escherichia coli*

Naaz Abbas, Aftab Ahmad and Abdul Rauf Shakoori*
School of Biological Sciences, University of the Punjab, New Campus, Lahore 54590, Pakistan

Received 12 March 2007, Accepted 27 July 2007

PreS domain of Hepatitis B virus (HBV) surface antigen is a good candidate for an effective vaccine as it activates both B and T cells besides binding to hepatocytes. This report deals with overexpression and purification of adr subtype of surface antigen that is more prevalent in Pakistan. PreS region, comprising 119 aa preS1 region plus a 55 aa preS2 region plus 11 aa from the N-terminal S region, was inserted in pET21a+ vector, cloned in *E. coli* DH5α cells and expressed in *E. coli* BL21 codon+ cells. The conditions for overexpression were optimized using different concentrations of IPTG (0.01-5 mM), and incubating the cells at different temperatures (23-41°C) for different durations (0-6 h). The cells were grown under the given optimized conditions (0.5 mM IPTG concentration at 37°C for 4 h), lysed by sonication and the protein was purified by ion exchange chromatography. On the average, 24.5 mg of recombinant protein was purified per liter of culture. The purified protein was later lyophilized and stored at −80°C.

Keywords: Adr subtype of surface antigen, Hepatitis B surface antigen, Hepatitis B virus, Hepatitis in Pakistan, PreS

Introduction

Hepatitis B virus (HBV) is a hepadnavirus that is noncytopathic and causes significant morbidity and mortality worldwide. It has a 3.2 kb partially double-stranded, open circular DNA with four open reading frames (ORF) encoding envelope, core, and X proteins as well as the DNA polymerase (Ganem and Varmus, 1987). Due to serological differences in hepatitis B surface antigen, HBV is characterized into four major subtypes: adw; adr; ayw, and ayr (Tiollais *et al*., 1985), although additional serological specificities have allowed the identification of five more subtypes (Norder *et al*., 1994). At present, based on the complete HBV genome eight genotypes, designated A to H, have been defined. It is estimated that 370 million people worldwide (6% of the world’s population) have chronic hepatitis B (CHB) (Perz, 2003). Virally encoded small (HBs), middle (MHBs), and large (LHBs) surface proteins together with cellular phospholipids form the envelope of hepatitis B virus. These proteins are translated from separate initiation codons, but share a common reading frame and stop codon. The HBs protein contains 226 amino acids and is the major component of the viral envelope. The MHBs protein has 55 extra amino acids (PreS2) located at the N-terminus of HBs and the LHBs protein carries an additional 119 amino acids (or 109 amino acids depending on the viral subtype, PreS1) at N-terminal extension with respect to the MHBs protein (Neurath and Kent, 1988). All envelope proteins are co-translationally inserted into the endoplasmic reticulum membrane directed by the topogenic elements in HBs (Ostapchuk *et al*., 1994; Prange and Strecek, 1995).

LHBs proteins play pivotal roles in infection and budding processes during the HBV life cycle. The N-terminal PreS region of the LHBs protein can adopt one of the two topological conformations depending on whether it undergoes a posttranslational translocation. Thus, LHBs proteins in virions exhibit a mixed population with their PreS region (PreS1 and PreS2) located either inside or outside of viral envelopes (Ostapchuk *et al*., 1994; Prange and Strecek, 1995). Encapsulation of viral nucleocapsids and secretion of mature viral particles require LHBs proteins with cytoplasmic PreS region (Bruss and Ganem, 1991; Bruss and Vieluf, 1995; Bruss, 1997; Petitson *et al*., 1997; Ponsel and Bruss, 2003). A stretch of amino acids across the PreS1 and the PreS2 is thought to be involved in the interaction with the cytosolic nucleocapsid before the budding event (Bruss, 1997; Ponsel and Bruss, 2003). During the infection process, externally exposed PreS region may mediate the binding of virion to a putative cellular receptor (De Meyer *et al*., 1997; Cooper *et al*., 2003). Amino acids 21-47 of the PreS1 domain likely bear

*To whom correspondence should be addressed.
Tel: 92-42-9231248; Fax: 92-42-9230980
E-mail: arshak@brain.net.pk, arshak.sbs@yahoo.com
the major epitope for cell attachment (Neurath et al., 1986; De Meyer et al., 1997; Cooper et al., 2003). Other important functions or roles have also been assigned to PreS, including regulation of viral replication and transactivation of a variety of promoter elements (Natoli et al., 1992; Lenhoff and Summers, 1994). Thus, multiple virological functions of the PreS region provide a useful target for anti-HBV drug intervention. Inhibition of viral infection by E. coli expressed PreS occurs through a direct interference with the binding of HBV to the putative cell surface receptor (Urban and Gripon, 2002). Recombinant PreS product could also be developed as a vaccine that elicits B and T cell immune responses to a humoral response, both the preS1 and preS2 domains (Caselmann, 1995).

The purpose of this study was to over express and purify the PreS region of Hepatitis B surface antigen from locally isolated Hepatitis B virus, for its later use as recombinant hepatitis B vaccine after completing the animal trials. The objects was also to develop a method for large scale production of PreS protein for various other studies, such as diagnostic use. The preS region in this study comprises 119 aa preS1 region plus a 55 aa preS2 region plus 11 aa from the N-terminal S region.

**Materials and Methods**

**Primer designing.** Oligonucleotide primers NHBFP2 (5’-GCTACCATATGGGAGGCATTGCCTC-3’) and NHBRI (5’-CAAAGCTTGAATTCCGTGAGTG-3’) were designed by aligning sequences of different HBV genotypes. To facilitate cloning in expression vector pET21a+, Ndel restriction site was inserted in the forward primer and HindIII site in the reverse primer sequence. A stop codon TAA was also added to the reverse primer sequence.

**PCR amplification.** Hepatitis B viral DNA was extracted from blood plasma according to Abbas and Shakoori (2005). The isolated HBV DNA (10 µl) was amplified using 10 mM Tris-Cl pH 8.8, 50 mM KCl, 1.5 mM MgCl2, 200 mM dNTPs, 1 µM each primer and 2.5 U Taq DNA polymerase in 25 µl reaction. PCR conditions were an initial denaturation step at 95°C for 5 min, followed by 35 cycles at 95°C for 30 s, annealing at 55°C for 30 s, and 1 min extension at 72°C. Finally an additional dengation step was carried out at 72°C for 10 min. The PCR products were resolved in 1% agarose gel stained with ethidium bromide.

**Cloning.** PCR product (580 bp) was purified from agarose gel in 1X TAE buffer using GeneClean® II kit BIO101. The purified PCR product was ligated in pET21a+ (Novagen, Madison, WI, USA) in 1:3 vector: insert ratio. Competent cells of E. coli DH5α were transformed with pET21a+PreS (hereafter designated as NHB1). Transformants were screened using X-Gal/ IPTG and ampicillin selection. Plasmids were isolated from white colonies according to Sambrook et al. (1989). For confirmation of the size of insert, NHB1 was restricted with Ndel and HindIII.

**Sequence analysis.** Plasmids for sequencing were prepared using QIAprep® Spin Miniprep Kit. Fluorescent sequencing was performed using one forward and one reverse primer on Beckman Coulter CEQ 8000 Automated Capillary Genetic analyzer in the facilities of School of Biological Sciences, University of the Punjab, Lahore. Newly sequenced PreS region from pakistani HBV isolate was analyzed with BLAST and ClustalW (http://www.ncbi.nlm.nih.gov/BLAST/). Sequence was deposited in GenBank with accession No. AY873842.

**Protein expression.** For protein expression studies, E. coli BL2 todon+ was transformed with NHB1. The expression of NHB1 clone (pET21a+preS in BL21 codon+ cells) was optimized after induction with different concentrations of IPTG incubated at different temperatures for different time periods. For optimizing IPTG concentration, two sets, control and experimental, each of nineteen 100 ml conical flasks with 15 ml sterile LB broth containing 100 µg/ml ampicillin were inoculated with 200 µl of overnight culture. The overnight culture for the experimental set was prepared by inoculating 5 ml LB broth containing 100 µg/ml sterile ampicillin, with single colony of NHB1 clone. The inoculum for control set was prepared by inoculating 5 ml LB broth containing 100 µg/ml sterile ampicillin with 200 µl
of pET2. All flasks were placed in shaking incubator at 37°C until 0.3-0.4 O.D. at 600 nm. This was designated as 0 h O.D. After that, 1.4, 4.2, 7.0, 9.8, 14.0, 28.0, 56.0, 70.0, 98.0, 140.0, 168.0, 210.0, 280.0, 350.0, 420.0, 490.0, 560.0, 630.0, 700.0 μl of 100 mM IPTG stock solution were added in the respective flasks to get 0.01, 0.03, 0.05, 0.07, 0.1, 0.2, 0.4, 0.5, 0.7, 1.0, 1.2, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5 and 5.0 mM IPTG concentrations. An aliquot of 1ml was taken from each flask 1, 3, and 4h after incubation, centrifuged at 12,000 rpm (5,100 x g) for 1 min to get cell pellet, to which 50-100 μg gel loading buffer (4v stacking gel buffer, 4% SDS, 20% glycerol, 0.2% bromophenol blue; 0.2 M DTT) was added. Pellet was vortexed, lysed in boiling water bath for 3-5 min, and then put immediately on ice for 1 min. Centrifugation was done at 12,000 rpm (5,100 x g) for 1 min. Samples were stored at ~20°C.

For determining the effect of temperature on protein expression, three sets of IPTG concentrations of 0.5, 1.0 and 3.0 mM, each of eight 100 ml conical flasks with 15 ml sterile LB broth containing 100 μg/ml sterile ampicillin were inoculated with 200 μl of overnight culture and incubated at 23, 25, 28, 31, 35, 37, 39 and 41°C.

For analysis of expression of protein at all these parameters, 15% discontinuous SDS-PAGE was run, stained with coomassie brilliant blue R-250, destained with destaining solution (methanol: glacial acetic acid, 3:1), and preserved in the destainer.

In order to determine plasmid stability, 100 μl cell culture (OD600 0.6) was plated on 1mM IPTG-containing LB plates at a dilution of 103, with and without ampicillin (100 μg/ml) and incubated overnight at 37°C. The number of colonies on each plate was counted manually.

**PreS recombinant protein purification.** LB broth (one litre) containing 100 μg/ml ampicillin was inoculated with 15 ml overnight culture of NHBI clone and incubated at 37°C until the O.D600 reached 0.7-0.8. Inoculum was prepared by inoculating 15 ml LB broth containing 100 μg/ml ampicillin with single colony of NHBI clone and incubating at 37°C for 18 h. Induction was done with 0.5 mM IPTG, and cells were harvested by centrifugation at 8000 rpm (2,500 x g) for 10 min at 4°C using JA-10 rotor (Beckman Coulter). Cells were washed with PBS (8 g NaCl, 0.2 g KCl, 1.44 g Na2HPO4, 0.24 g KH2PO4, dissolved in 1 litre distilled water, pH 7.4), and then mixed with 25 ml of 0.5 M citric acid on ice. Cells were lysed by sonication, 8 cycles, each of 30 s followed by 2 min of rest on ice. The sonicate was centrifuged at 20,000 rpm (2,500 x g) for 30 min at 4°C. The supernatant was dialyzed against large volume of distilled water for 24 h with constant stirring at 4°C. CM-Sepharose column was prepared and equilibrated with 10 mM phosphate buffer. The pH of dialyzing solution was raised to 6.8 with 0.1 M NaOH and applied to column. The bound protein was eluted with 50, 100, and 200 mM phosphate buffer and fractions were collected manually. Absorbance at 260 nm and 280 nm of each fraction was also measured. The protein was quantified according to Lowry et al. (1951).

**Results**

**HBV PreS sequence.** The PreS insert (0.6 Kb) in pET21a+ was sequenced and blasted against NCBI data, which revealed that Pakistani HBV PreS sequence had 100% homology with adr4C subtype, 96.6% with ayrC, 90.2% with adw2A, 82.4% with aywD, 79.4% with ayw4E and 76.5% with adw4F (Table 1).

<table>
<thead>
<tr>
<th>Subtypes/Genotypes</th>
<th>References</th>
<th>Homology</th>
<th>Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>adr4C</td>
<td>(Fujiyama et al., 1983)</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>ayrC</td>
<td>(Okamoto et al., 1986)</td>
<td>96.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td>adnqC</td>
<td>(Nordler et al., 1994)</td>
<td>95.5%</td>
<td>4.5%</td>
</tr>
<tr>
<td>adw2A</td>
<td>(Valenzuela et al., 1979)</td>
<td>90.2%</td>
<td>9.8%</td>
</tr>
<tr>
<td>adwB</td>
<td>(Okamoto et al., 1988)</td>
<td>86.8%</td>
<td>13.2%</td>
</tr>
<tr>
<td>aywD</td>
<td>(Tong et al., 1990)</td>
<td>82.4%</td>
<td>17.6%</td>
</tr>
<tr>
<td>aywE</td>
<td>(Nordler et al., 1994)</td>
<td>79.4%</td>
<td>20.6%</td>
</tr>
<tr>
<td>adw4F</td>
<td>(Nordler et al., 1994)</td>
<td>76.5%</td>
<td>23.5%</td>
</tr>
</tbody>
</table>

**Plasmid stability.** In the presence of IPTG, the transformed cells carrying PreS gene do not grow because the cells have dedicated all their resources to the production of the recombinant protein instead of cell maintenance. Consequently, the cells containing plasmid will not be able to grow in LB plates containing IPTG though they will survive in the medium with antibiotic. With unstable target plasmids, the fraction of cells that have lost the plasmid will be reflected by an increase in colonies on the LB plate plus IPTG and a decrease on the LB plate plus antibiotic.

Table 2 showed that almost all cells formed colonies both on the LB plate and on the LB plate plus antibiotic; about 2% of the cells formed colonies on the LB plate plus IPTG; and of the total only ~1% formed colonies on the LB plate plus antibiotic plus IPTG.

**Effect of different IPTG concentrations on gene expression.** The expression of NHB1 (pET21a+preS+ in BL21codon+) clone was optimized after induction with concentration of IPTG incubated at different temperatures for different period.

**Table 1. Comparison of NHB1 sequence with other HBV subtypes/genotypes**

<table>
<thead>
<tr>
<th>Plates</th>
<th>No. of colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB plate</td>
<td>237</td>
</tr>
<tr>
<td>LB plate + antibiotic</td>
<td>225</td>
</tr>
<tr>
<td>LB plate + IPTG (1 mM)</td>
<td>6</td>
</tr>
<tr>
<td>LB plate + antibiotic + IPTG (1 mM)</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 2. Results of plasmid stability test**
Overexpression and Purification of HBV Subtype adr Surface Protein

Effect of different temperatures on gene expression. The *E. coli* strain that had a plasmid with insert showed maximum growth at 39°C, as compared with other temperatures (Fig. 3). Although the cells grew well at 39°C, the protein expression was not maximum at this temperature. The maximum expression was recorded at 35°C and 37°C at IPTG concentration of 0.05, 0.1, 0.5 and 1.0 mM after 3 h of induction (Fig. 4).

Protein purification. Fig. 5 shows that after purification of the recombinant PreS protein by using cation exchange chromatography, there is only a single band of desired protein and there is no other band of the bacterial protein. Phosphate buffer at a concentration of 200 mM worked very well in eluting the bound protein. Moreover, the protein was eluted in small buffer volume.

PAGE pattern in Fig. 6 represents a flow diagram of purification of antigenic protein, which was retained in the pellet (ASP) after the NHB1 cells were sonicated, purified after dialysis (AD), and was eluted in fractions 3, 4, 5 (F3, F4, F5) as single band.

A total amount of 24.5 mg of recombinant PreS protein was prepared from 1 liter of culture that was grown in flasks without the use of fermenter. The protein was desalted by dialysis, for 24 h at 4°C, with 3 liter of distilled water. Finally the PreS protein was concentrated by lyophilization and the volume of protein solution was reduced from 35 ml to 3 ml. The 20 mg of recombinant PreS protein was stored at −80°C in double distilled water.

Fig. 1. The overall growth pattern of NHB1 (pET21a+preS−) cells in different IPTG concentrations used for optimization of expression (0.01-5.0 mM). The strain (having vector with insert of preS) showed maximum growth at 0.01 mM IPTG concentration due to less stress at this concentration on the growth. The strain showed optimum growth at 0.4 to 1.0 mM IPTG concentrations, and these are also the IPTG concentration at which there is also maximum expression of protein. The O.D.600nm is along the Y axis and the IPTG concentration is along the X axis.

Fig. 2. SDS-PAGE pattern of total protein of BL21codon+ *E. coli* clones with (preS+) and without (preS−) preS regions. This pattern of expression of preS region of HBV is after 4 h of administration of IPTG concentration ranging between 0.2-1.0 mM. The figure shows most prominent and distinct bands of expressed preS protein. The arrow head shows the position of desired size protein.

Fig. 3. The growth of *E. coli* strain (pET21a+preS+) at different temperatures and IPTG concentrations. As indicated in the figure, the increase in IPTG concentration affects the growth of strain. The pET21a+preS− strain showed best overall growth at 37°C and 39°C.

Fig. 4. SDS-PAGE pattern of total protein of BL21codon+ *E. coli* clones, preS+ and preS−, after administration of IPTG concentration ranging between 0.01-3.0 mM at 35°C. The figure shows distinct preS bands in 0.1-1.0 mM IPTG range, while no protein band is in control that have plasmids without preS inserts. Arrow head indicates the position of desired protein. Lysozyme (14 kDa) was used as marker.
The complete PreS region of Hepatitis B surface antigen was also expressed in many researchers used the pET system for expression (Delos et al., 1991; Wei et al., 2002) but other expression vectors were also used like pT7-7 (Nunez et al., 2001), pQE40, pMalC2x (NEB), pGEX2T, pThioHisA, and pTXB1 (NEB) (Deng et al., 2005), pS300 (Chen et al., 2003) for the expression of the same gene. In other laboratories different purification procedures have been used.

For purification of preS protein devoid of any tag, the protein was purified from culture by first suspending the cells in cold water containing 50 mM-EDTA, 10% imidodiacetate and 1 mM-phenylmethanesulphonyl fluoride. All subsequent procedures were performed at 4°C, and all buffers contained 0.02% NaN3. The cells were disrupted by a single passage through a French press at 138,000 kPa (20,000lb/in2). Protein was diluted and purified by using CM-Sepharose column (pH 6.8). The amount of expressed protein was 30 mg/liter of culture (Delos et al., 1991). In another way of purification of PreS protein, cells were lysed by sonication and proteins were precipitated with 20% ammonium sulfate and finally purified by gel filtration chromatography. It is also reported however that PreS protein precipitated with ammonium sulfate is hard to get in pellet after centrifugation (Fujiyama et al., 1983). PreS domain that was fused with Cellulose Binding Domain (CBD) was purified by using chitin matrix in column, that bound to CBD and then eluted out, but it is very difficult to elute the protein from the chitin matrix. Thioredoxin tagged preS protein was purified by using the nickel chelating sepharose, and the recombinant protein was purified under non-denaturing conditions, and protein was produced in active form. But even with stringent washing impurities were still present in the protein. In addition many other tags were also used with preS such as 6-Histidine, Glutathione S-Transferase (GST), Dihydrofolate reductase (DHFR), Maltoose binding protein (MBP), and preS protein was purified with the respective affinity column (Deng et al., 2005).

In another method the cells were sonicated in 6 M urea solution, and 10 mM Tris, and PreS protein was finally purified using Sepharose CL-6B Nitrilotriacetic acid (NTA) column. This preS protein has histidine tag, but after purification there was very less activity of the protein due to its denaturation. We followed almost the same procedure (Fujiyama et al., 1983; Chen et al., 2003) and got 24.5 mg of purified PreS protein. The purity of the protein was confirmed by running 15% SDS-PAGE. We did not use any tag or fusion protein in our experiment, because after purification process there is always need to first cleave and remove the tag or fusion protein and then again there is need to purify the protein so there is need of extra steps in purification that make the commercial production of that particular protein unfeasible. The procedure that we have developed in this experiment for PreS purification is rapid (2 days), efficient and cost effective as compared to other procedures being used for purification of PreS.

Discussion

The complete PreS on the surface of HBV virions has been shown as an independent domain with a native structure (Cooper et al., 2003). Thus, recombinant PreS proteins expressed in the E.coli system may have the same structure as it has in its normal form.

The preS region of Hepatitis B surface antigen may have the same structure as it has in its normal form. The preS region of Hepatitis B surface antigen in E.coli, or bacterial system, due to abundance of production and ease of purification. For expression of preS protein in E.coli, many researchers used the pET system for expression (Delos et al., 1991; Wei et al., 2002) but other expression vectors were also used like pT7-7 (Nunez et al., 2001), pQE40, pMalC2x (NEB), pGEX2T, pThioHisA, and pTXB1 (NEB) (Deng et al., 2005), pS300 (Chen et al., 2003) for the expression of the same gene. In other laboratories different purification procedures have been used.

For purification of preS protein devoid of any tag, the protein was purified from culture by first suspending the cells in cold water containing 50 mM-EDTA, 10% imidodiacetate and 1 mM-phenylmethanesulphonyl fluoride. All subsequent procedures were performed at 4°C, and all buffers contained 0.02% NaN3. The cells were disrupted by a single passage through a French press at 138,000 kPa (20,000lb/in2). Protein was diluted and purified by using CM-Sepharose column (pH 6.8). The amount of expressed protein was 30 mg/liter of culture (Delos et al., 1991). In another way of purification of PreS protein, cells were lysed by sonication and proteins were precipitated with 20% ammonium sulfate and finally purified by gel filtration chromatography. It is also reported however that PreS protein precipitated with ammonium sulfate is hard to get in pellet after centrifugation (Fujiyama et al., 1983). PreS domain that was fused with Cellulose Binding Domain (CBD) was purified by using chitin matrix in column, that bound to CBD and then eluted out, but it is very difficult to elute the protein from the chitin matrix. Thioredoxin tagged preS protein was purified by using the nickel chelating sepharose, and the recombinant protein was purified under non-denaturing conditions, and protein was produced in active form. But even with stringent washing impurities were still present in the protein. In addition many other tags were also used with preS such as 6-Histidine, Glutathione S-Transferase (GST), Dihydrofolate reductase (DHFR), Maltoose binding protein (MBP), and preS protein was purified with the respective affinity column (Deng et al., 2005).

In another method the cells were sonicated in 6 M urea solution, and 10 mM Tris, and PreS protein was finally purified using Sepharose CL-6B Nitrilotriacetic acid (NTA) column. This preS protein has histidine tag, but after purification there was very less activity of the protein due to its denaturation. We followed almost the same procedure (Fujiyama et al., 1983; Chen et al., 2003) and got 24.5 mg of purified PreS protein. The purity of the protein was confirmed by running 15% SDS-PAGE. We did not use any tag or fusion protein in our experiment, because after purification process there is always need to first cleave and remove the tag or fusion protein and then again there is need to purify the protein so there is need of extra steps in purification that make the commercial production of that particular protein unfeasible. The procedure that we have developed in this experiment for PreS purification is rapid (2 days), efficient and cost effective as compared to other procedures being used for purification of PreS.

To conclude, the complete PreS region of Hepatitis B surface antigen without fusion form or tag has been successfully
expressed. A simple yet efficient method has been established for purification of the PreS protein. The PreS protein purified is stable and soluble. The purified PreS protein may be a valuable candidate for studying the structure of the PreS region as well as for screening antivirals and use as recombinant Hepatitis B vaccine.

References


Nordh, H., Courouce, A. M. and Magnius, L. O. (1994) Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new


