

# Molecular Cloning and Expression of Hepatitis C Virus Core Protein and Production of Monoclonal Antibodies to the Recombinant Protein

Sirirung Songsivilai, Tararaj Dharakul, Rojana Kunkitti and Charin Thepthai

Hepatitis C virus (HCV) is a major cause of post-transfusion non-A, non-B hepatitis, and is prevalent worldwide.<sup>1</sup> More than half of HCV-infected individuals develop chronic hepatitis which may progress to cirrhosis and hepatocellular carcinoma. The virus is transmitted parenterally via blood transfusion. However, most individuals with chronic hepatitis C who can transmit the virus are asymptomatic. Therefore, screening of blood and blood products for antibodies to HCV is currently the most effective method for the prevention of HCV transmission.<sup>2</sup>

HCV, a member of the family *Flaviviridae*, has a single stranded RNA genome 9.5 kilobases in length.<sup>3</sup> The genome contains a single open reading frame encoding viral structural and non-structural (NS2-NS5) proteins. Nucleotide sequence analysis of HCV isolates from different parts of the world indicates that there is a high degree of nucleotide sequence heterogeneity of the virus. HCV can be classified into various genotypes on the basis of nucleotide sequence similarity of the viral genome.<sup>4</sup> The degree of sequence variability is not uniformly distributed

**SUMMARY** The gene encoding nucleocapsid (core) protein of hepatitis C virus (HCV) was isolated from a Thai blood donor infected with HCV genotype 1b. The nucleotide sequence of this clone showed a high degree of homology to that of 4 HCV strains isolated from other Thai blood donors as well as that of the HCV prototypes of genotypes 1a, 1b, 2a and 3a. The entire region of the core gene was cloned into an expression plasmid pGEX-3X to be expressed as a fusion protein with glutathione-S-transferase (GST). *E. coli* transformants containing this plasmid did not express the fusion protein. However, GST-HCV core fusion protein could be produced when the core gene was truncated at the 3' end resulting in a gene encoding only the first 123 amino acid residues of the core protein. This fusion protein was insoluble in standard buffers, but could be solubilized by sarkosyl and thus subsequently purified using glutathione-Sepharose 4B. The purified fusion protein was immunogenic and could react with antibodies from blood donors infected with all genotypes of HCV found in Thailand. In addition, two murine hybridoma clones secreting monoclonal antibodies specific to the recombinant HCV core protein were produced. The purified HCV core protein and the monoclonal antibodies to the recombinant protein will be useful for developing assay systems for detecting anti-HCV antibodies and HCV antigen, respectively.

throughout the genome. The core gene (encoding 191 amino acids nucleocapsid protein) shows approximately 90% sequence similarity between different HCV types, whereas high degrees of sequence divergence are found in the genes encoding non-structural proteins. HCV cannot be cultured *in vitro* and thus molecular biological methods are usually used for obtaining large amounts of virus proteins for studying the virus and for developing diagnostic assays.

The first generation assay for

detecting antibodies to HCV, using a recombinant nonstructural HCV protein in the NS4 region (c100-3), has been implemented extensively for blood screening. However, the assay has both high false-positive and high false-negative rates.<sup>5,6</sup> HCV-infected patients may be tested false-negative for anti-HCV antibodies, while HCV RNA could be

From the Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

Correspondence : Sirirung Songsivilai,

amplified from their sera.<sup>7</sup> Anti-c100-3 negative donors may transmit HCV.<sup>8</sup> The antigen used in the assay is derived from a relatively non-conserved region of the virus genotype 1a, which is prevalent in the United States. Antibodies to the viral NS4 protein are later found to be type-specific and thus patients infected with HCV other than genotype 1 may have no antibodies to the NS4 protein and would therefore test negative.<sup>9</sup> Recent studies have shown that assays which include the well-conserved nucleocapsid or core protein, in addition to NS4 protein, are more sensitive and more effective than the original anti-c100-3 based assays for prevention of transfusion-transmitted HCV infection. Antibodies to HCV core protein can be found early in infection and the detection of anti-core antibodies may be useful for the early diagnosis of hepatitis C.<sup>10,11</sup>

Thailand is endemic for HCV with a prevalence of anti-HCV antibody in blood donor populations of about 1.5% (Songsivilai, unpublished data). HCV genomes isolated in Thailand can be classified into at least 4 groups, including genotypes 3a, 1b, 1a and 7 (Songsivilai & Kanistanon, unpublished data, and ref. 12). Genotypes 3 and 1b are the most common.<sup>12</sup> Genotype 1a, which is used in commercial diagnostic tests, is uncommon in Thailand. The differences among HCV genotype found in different geographical areas may have implications for the diagnosis, clinical course, pathogenesis, response to treatment, immunity and vaccine development of HCV. The effectiveness of diagnostic assays for HCV is depended on the genotype prevalence in a geographical area. This is particularly for Thailand, which has a different genotypic distribution of HCV than other areas. The ideal assays for Thailand should be based on common Thai genotypes of HCV. In addition, the assays should incor-

porate a combination of antigens from various parts of the viral genome, containing both structural and non-structural proteins, in order to avoid false-negative results arising from the heterogeneity of HCV genomes.

The objectives of the present study are to clone the gene encoding the core protein of HCV isolated from a Thai blood donor and to express the recombinant core protein using a prokaryotic expression system. The purified recombinant protein will be used as one of the antigens in diagnostic tests of HCV infection. Specific monoclonal antibodies to the recombinant HCV core proteins are also produced.

## MATERIALS AND METHODS

### Specimens

An anti-HCV antibody-positive plasma sample (designated BB51) was obtained from the Blood Transfusion Unit, Siriraj Hospital, Thailand. Antibodies to HCV were detected by enzyme immunoassay and immunoblot assay (HCV ELISA and HCV Blot 3.0, Diagnostic Biotechnology, Singapore). HCV RNA can be detected by the polymerase chain reaction using primers in the conserved 5' non-coding region. The virus genotype was identified as type 1b by a reverse hybridization assay. This donor had no history of

jaundice, intravenous drug use, or blood transfusion. The plasma was thawed once, aliquoted into 1.5 ml vials and immediately stored at -70°C. In addition, plasma samples from 12 other blood donors who had antibodies to HCV and whose viral genotypes were known were also used in this study. Three cases were infected with HCV genotype 1a, 3 with genotype 1b, 5 with genotype 3a, and 1 with genotype 7. Plasma from 4 cases of patients infected with Dengue virus and plasma from 4 healthy individuals who had no antibodies to HCV were used as negative controls.

### DNA amplification and nucleotide sequencing

Amplification of the HCV genome was performed using a nested reverse transcription polymerase chain reaction (RT-PCR) from 100 µl of plasma as previously described with some modifications.<sup>13</sup> A new set of oligonucleotide primers (A11L-A8R and A5L-A6R) was used for amplifying the entire core region of HCV genome (Fig. 1) in a nested PCR system. The nucleotide sequences of the primers used are shown in Table 1. Primers were synthesized using an automated DNA/RNA synthesizer (Model 392, Applied Biosystems, Inc., U.S.A.). The PCR product after 2 rounds of amplification was 614 base-pairs

**Table 1.** Nucleotide sequence of primers for amplifying the entire HCV core gene. The sequence of A5L has a *Bam* HI site whereas the sequence of A6R has an *Eco* RI site, to facilitate gene cloning.

Primer	Nucleotide sequence(5'- 3')	Position	Polarity
A11L	ACTACTGTCTTCACGCAGAAAGC	-288 to -266	sense
A8R	TCAGTCGCCTCRTACACAAT	634 to 653	antisense
A5L (nested)	TCTCGGATCCCGTGCAHCAT	-18 to 2	sense
A6R (nested)	CCGAATTCGTTGCGCACYTSRTA	574 to 596	antisense

(Note: R=A or G; H=A, T or C; Y=C or T; S=C or G)

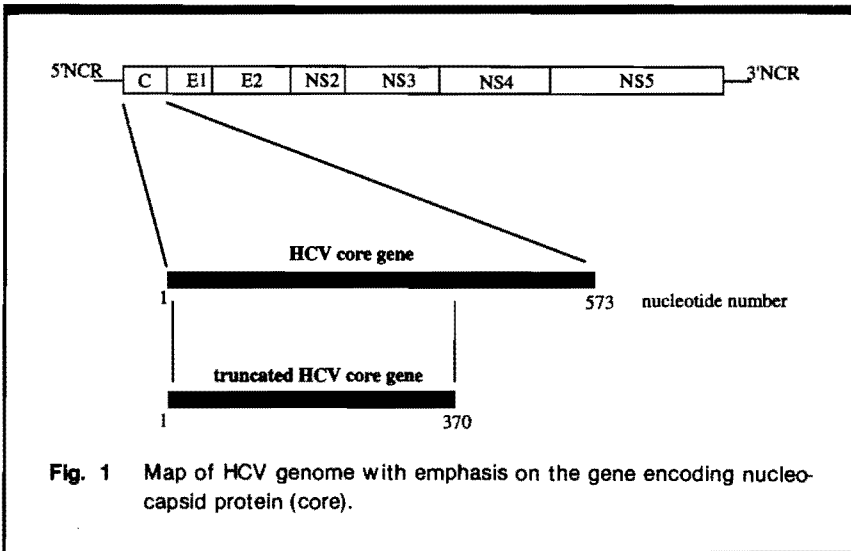


Fig. 1 Map of HCV genome with emphasis on the gene encoding nucleocapsid protein (core).

in length. The nucleotide sequence of the PCR product was obtained using the PRISM dye termination sequencing system (Perkin Elmer, U.S.A.) running on an automated DNA sequencer (Model 373, Applied Biosystems, Inc., U.S.A.). Nucleotide sequencing was performed in both directions. The sequence obtained was analyzed using SeqEd and Mac Vector programs (Applied Biosystems and Kodak Scientific Imaging System, respectively) and compared with nucleotide sequences of other HCV isolates cloned by our group (isolates BB10, BB16, BB43 and BB81) and with those deposited in the GenBank database release 85.0 (National Center for Biotechnology Information, U.S.A.).

**Construction of expression plasmids containing HCV core gene**

The PCR product containing the entire HCV core region was end-repaired using T4 DNA polymerase and Klenow DNA polymerase enzymes, then digested with restriction enzymes *Bam*HI and *Eco*RI. The product was cloned into the *Bam*HI-*Eco*RI site of plasmid pGEX-3X (Pharmacia, U.S.A.), and the recombinant clones were identified by restriction enzyme analysis. The selected clone, designated pGCBB51, contained the entire HCV core region (Fig. 2). The inserted HCV core gene of this clone was sequenced and the nucleotide sequence was identical to that derived from nucleotide sequencing directly from the PCR product. To truncate the 3' end of the inserted core gene, pGCBB51 was digested with restriction enzymes *Cla*I and *Eco*RI, end-filled and religated. The resulting clone, renamed pGCBB51-g (Figure 2), contained the first 370 nucleotides of the core gene, equivalent to the first 123 amino acids. The two plasmids were transformed into *Escherichia coli* strain JM101 for protein expression.

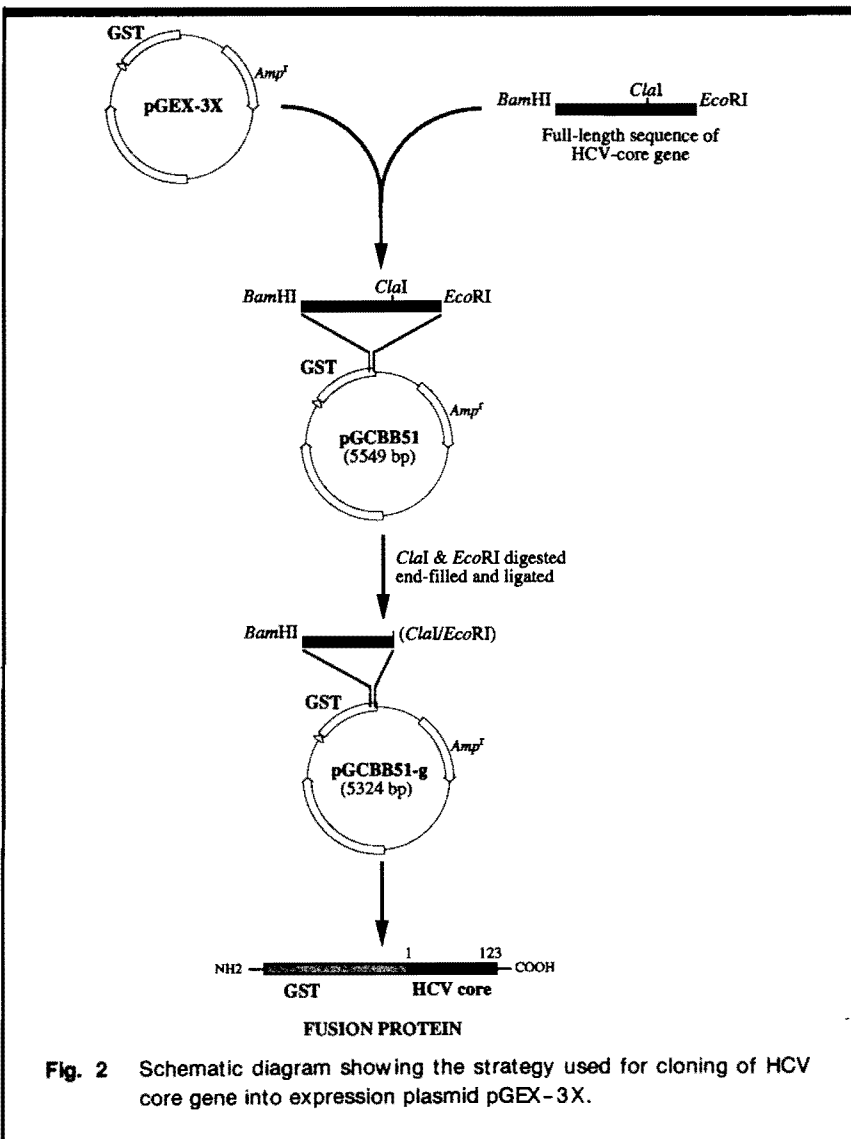


Fig. 2 Schematic diagram showing the strategy used for cloning of HCV core gene into expression plasmid pGEX-3X.

### Production and purification of HCV core-containing fusion proteins

Protein expression was carried out as described by Smith and Johnson with some modifications.<sup>14</sup> An overnight culture of bacteria containing recombinant pGCBB51, pGCBB51-g or pGEX-3X, was diluted 1/10 in culture medium containing 100 µg/ml ampicillin and incubated at 37°C for 1 hour with vigorous shaking. Isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.5 mM to induce the production of fusion protein and the culture was incubated for a further 3 hours at 37°C. The cells were pelleted by brief centrifugation, washed 3 times with phosphate-buffered saline (PBS) and resuspended in 1/40 volume of PBS containing 1% Triton-X-100. Alternatively, Triton-X-100 can be substituted with either 1% Tween-20, 10 mM dithiothreitol, 0.03% sodium dodecyl sulphate or 0.025% *N*-laurylsarcosine (sarkosyl). The suspension was sonicated to disrupt the bacterial cell membrane, centrifuged, and the supernatant was collected. For protein purification, the fusion protein was solubilized in PBS containing 0.025% sarkosyl, 25 mM triethanolamine, and 1 mM EDTA. The mixture was sonicated, centrifuged and the supernatant was collected. Triton-X-100 and CaCl<sub>2</sub> were added to final concentrations of 2% and 1 mM, respectively. The supernatant was applied to glutathione-Sepharose 4B beads (Pharmacia Biotech, U.S.A.), and incubated at room temperature on a rotating platform for 10 minutes. The beads were collected by brief centrifugation and washed three times with PBS containing 0.025% sarkosyl. Fusion protein was eluted using 2×5 bed volume of 50 mM TrisCl (pH 8.0) containing 10 mM reduced glutathione and 0.5% sarkosyl on a rotating platform for 5 minutes. The purified fusion protein was dialysed in 2×1,000 volume

of PBS at 4°C, overnight. Total cell lysate, pellet and supernatant after sonication and solubilization, and the purified protein were analyzed by SDS-polyacrylamide gel electrophoresis and immunoblotting. The concentration of protein was determined by colorimetric methods, using Bio-Rad protein assay kit (Bio-Rad, U.S.A.).

### Western blot analysis of the fusion proteins

The purified proteins were blotted onto nitrocellulose membrane and detected using an immunoenzyme method.<sup>15</sup> Plasma samples from 12 blood donors who were infected with HCV genotype 1a, 1b, 3a or 7 were used for immunostaining (1:100 dilution). Plasma from 4 healthy individuals who had no antibodies to HCV and from 4 patients infected with dengue virus were used as negative controls.

### Production of monoclonal antibodies to HCV core protein

Hybridoma production was carried out essentially as previously described.<sup>15</sup> Purified fusion protein expressed from plasmid pGCBB51-g was used for immunizing a BALB/c mouse. The mouse was injected intraperitoneally with 19 µg of the fusion protein in complete Freund's adjuvant twice at 3 month intervals, followed by a booster injection with the same amounts of protein without adjuvant six weeks after the second immunization. The animal was sacrificed 3 days after the final booster and the spleen was removed. Spleen cells were fused with myeloma cells X63-Ag8.653. The hybridoma cells were identified and tested for antibodies against the fusion protein and purified GST carrier protein. The selected clones were recloned three times. Monoclonal antibodies were collected from cell culture supernatants.

## RESULTS

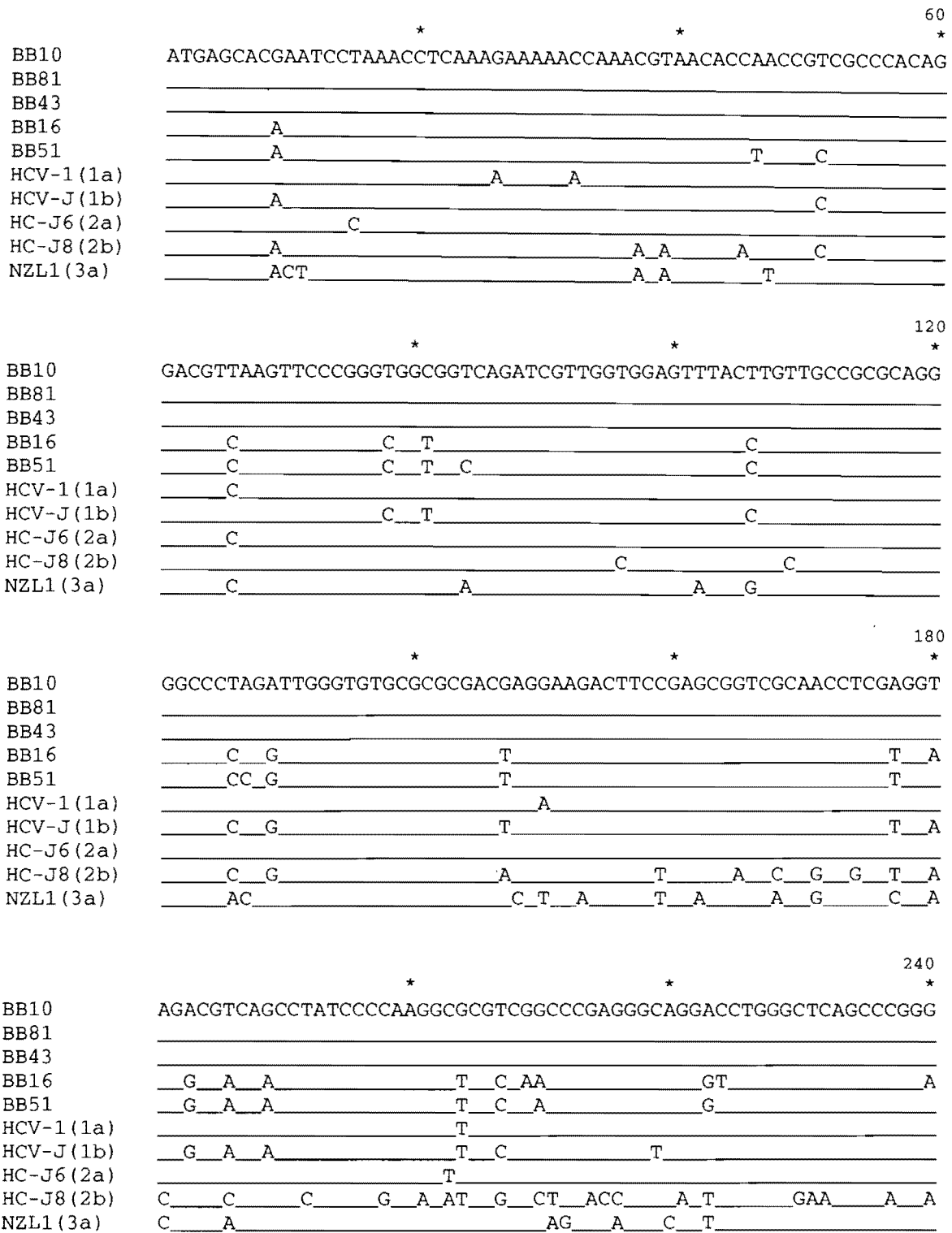
### Amplification and nucleotide se-

### quencing of the gene encoding the core protein

The gene encoding the HCV core protein was amplified from the plasma of donor BB51 as well as from other blood donors (BB10, BB16, BB43 and BB81). The entire core genes were sequenced (Fig. 3 and Fig. 4). Comparison of the nucleotide sequences demonstrated that the sequences of BB51 and BB16 were similar to that of prototype strain HCV-J (genotype 1b), whereas the nucleotide sequences of BB10, BB43, and BB81 were similar to prototype strain HCV-1 (genotype 1a). Deduced amino acid sequences in the entire core region of the Thai isolates were very similar to each other, with 99.3% homology among the genotype 1a isolates and 96.8% homology among genotype 1b isolates. The entire core gene of the Thai isolate BB51 used for cloning shared only 90.1% homology with of prototype strain genotype 3a (NZL1). The HCV core gene shared no significant amino acid sequence homology to that of other pathogens deposited in the GenBank. The degree of amino acid sequence homology was slightly higher in the first 123 amino acid residues of the HCV core protein, with 100% homology between genotype 1a isolates, 96.8% between genotype 1b isolates and 96.5% between genotype 1a and 1b isolates. Interestingly, the first 123 amino acid residues of isolate BB51 shared 94.3% homology with that of genotype 3a prototype.

### Expression of GST-HCV core fusion protein

pGCBB51, pGCBB51-g and pGEX-3X were used for protein expression. JM101 *E. coli* transformed with pGCBB51 containing the entire core gene did not express the recombinant fusion protein (data not shown), while the glutathione-S-transferase (GST) was successfully expressed from pGEX-3X. Failure to express the fusion protein was



**Fig. 3** Nucleotide sequence alignment of the core region of 5 strains of HCV isolated in Thailand, including genotype 1a (BB10, BB81 and BB43), genotype 1b (BB16 and BB51 - used for cloning and expression), in comparison with the sequences of the prototype strains of HCV genotypes 1a, 1b, 2a, 2b and 3a, deposited in GenBank database. Only the first 480 nucleotides of the core gene are shown. The *Cla*I restriction site is underlined.

300  
\*

BB10 TACCCTTGGCCCCTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCCCCC

BB81 T

BB43

BB16 C ATG A A

BB51 TG\_C ATG A

HCV-1 (1a) T

HCV-J (1b) C TATG A A

HC-J6 (2a) T

HC-J8 (2b) T G\_C A C T C T

NZL1 (3a) T\_C A G A

360  
\*

BB10 CGTGGCTCTCGGCCTAGCTGGGGCCCTACAGACCCCCGGCGTAGGTTCGCGCAATTTGGGT

BB81 M

BB43 T

BB16 C C T C T

BB51 T C AT T

HCV-1 (1a) C

HCV-J (1b) T C T

HC-J6 (2a) T C G

HC-J8 (2b) C G T CT C C A A A C

NZL1 (3a) C C T ATC A AT G C

420  
\*

BB10 AAGGTCATCGATACCCTCACGTGCGGCTTCGCCGACCTCATGGGGTACATCCCCTCGTC

BB81

BB43

BB16 A T T

BB51 A T T

HCV-1 (1a) T A A

HCV-J (1b) T A T T

HC-J6 (2a) A

HC-J8 (2b) GA A T T T T TG T

NZL1 (3a) A A A

480  
\*

BB10 GGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGAC

BB81 C

BB43 G C

BB16 CT\_A\_G T C A G

BB51 C\_A\_G T A G

HCV-1 (1a)

HCV-J (1b) C\_A\_G A T G

HC-J6 (2a)

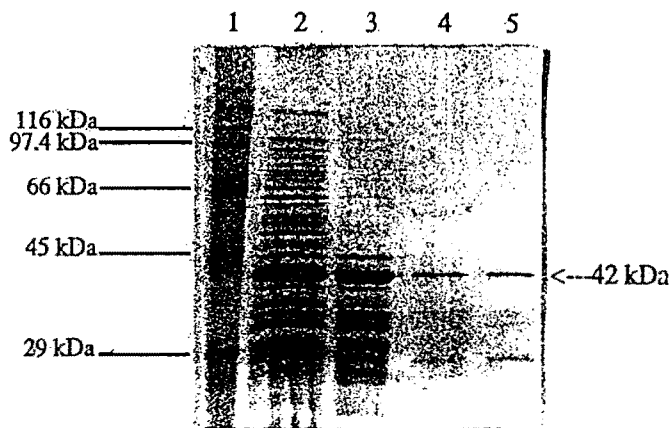
HC-J8 (2b) GG TC A T A C T TA C G

NZL1 (3a) T\_G\_A TC A A C GA CC T

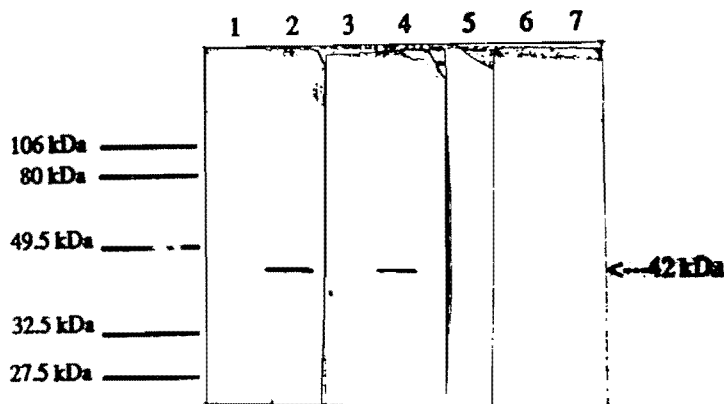
Fig. 3. (Continued).

		*		*	60
					*
BB10	MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSEERSQPRG				
BB81	_____				
BB43	_____				
BB16	_____				
BB51	_____				
HCV-1 (1a)	_____ K N _____				
HCV-J (1b)	_____				
HC-J6 (2a)	_____ I _____				
HC-J8 (2b)	_____				
NZL1 (3a)	_____ L _____ I _____ V _____				
					120
		*		*	*
BB10	RRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG				
BB81	_____				
BB43	_____				
BB16	_____ Q V _____ M _____				
BB51	_____ A _____ D M _____ N _____				
HCV-1 (1a)	_____				
HCV-J (1b)	_____ M _____				
HC-J6 (2a)	_____ V _____				
HC-J8 (2b)	_____ D ST KS GK _____ T _____ H _____				
NZL1 (3a)	_____ S S _____ N _____				
					180
		*		*	*
BB10	KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDSVNYATGNLPGCSFSIFLLA				
BB81	_____ G _____				
BB43	_____ G _____				
BB16	_____ V _____ G _____				
BB51	_____ G _____				
HCV-1 (1a)	_____ G _____				
HCV-J (1b)	_____ G _____				
HC-J6 (2a)	_____ G _____				
HC-J8 (2b)	R _____ I _____ V _____ V V _____ GI _____				
NZL1 (3a)	_____ V V _____ A _____ GI F _____				
					191
BB10	LLSCLTVPASA				
BB81	_____				
BB43	_____ P _____				
BB16	_____ I _____				
BB51	_____ I _____				
HCV-1 (1a)	_____				
HCV-J (1b)	_____ I _____				
HC-J6 (2a)	_____				
HC-J8 (2b)	_____ V V _____				
NZL1 (3a)	_____ F _____ IH _____ AS _____				

Fig. 4 Comparison of deduced amino acid sequences of the core region of 5 strains of HCV isolated in Thailand and the prototype strains of HCV genotypes 1a, 1b, 2a, 2b and 3a, deposited in the GenBank database.



**Fig. 5** Purification of GST-HCV core fusion protein. Fusion protein expressed by plasmid pGCBB51-g was purified using glutathione-Sepharose 4B and the proteins were analysed on 10% SDS-polyacrylamide gel. Lane 1: molecular weight markers, lane 2: total bacterial lysate, lane 3: bacterial lysate solubilized in sarkosyl, lane 4: protein fraction bound to glutathione-Sepharose 4B beads, lane 5: protein eluted from the beads.



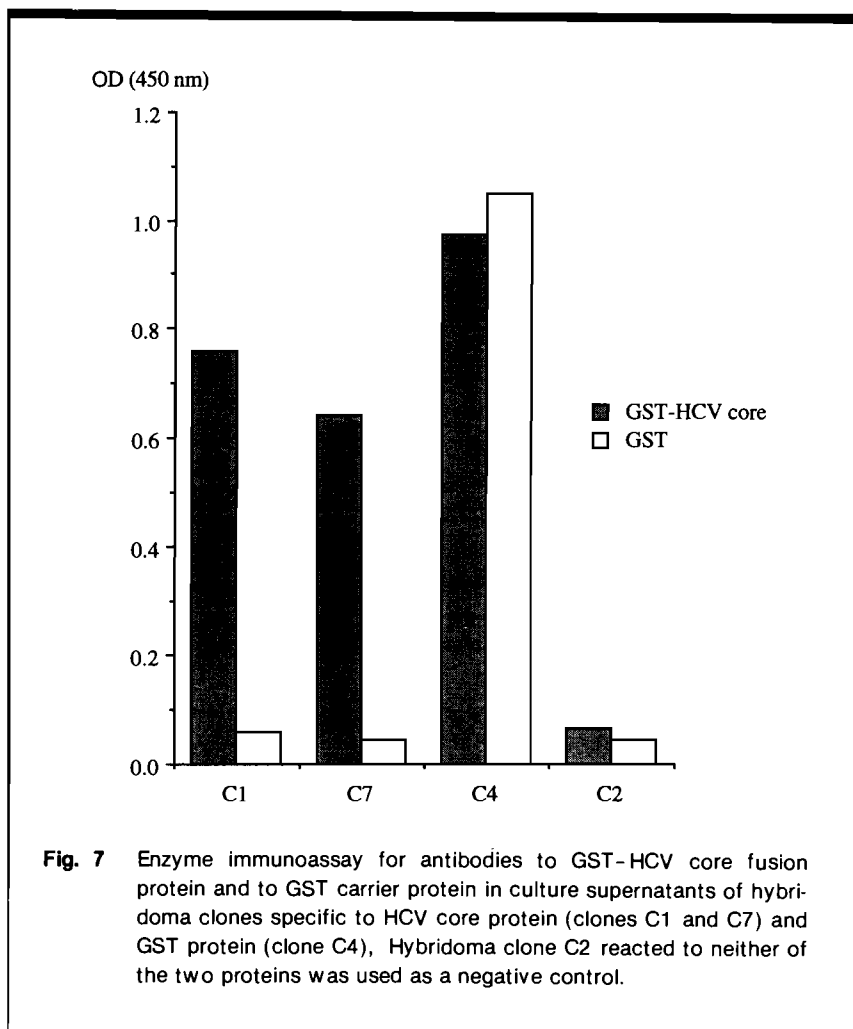
**Fig. 6** Immunoblot analysis of GST-HCV core fusion protein. Purified protein separated in SDS-PAGE was blotted onto nitrocellulose membrane and immunostained with plasma from patients who had antibodies to HCV. Lanes 1, 3, 6: GST. Lanes 2, 4, 7: GST-HCV core fusion protein. Lane 5: molecular weight markers. Lanes 1, 2: reacted with plasma from donor BB51 (autologous plasma). Lanes 3, 4: reacted with plasma from donor BB8 (infected with HCV genotype 1a). Lanes 6 and 7: reacted with plasma from a patient infected with dengue virus.

also demonstrated when pGCBB51 was retransformed into other *E. coli* strains, TG1 and DH5 $\alpha$ . Nucleotide sequence analysis of the plasmid in the junction between the GST and HCV sequences showed that pGCBB51 contained HCV core gene in the correct orientation and reading frame.

In contrast, bacteria transformed with pGCBB51-g were able to express the fusion protein of an expected molecular weight (42 kDa). This protein contained the HCV core protein (amino acid residues 1-123) fused to the carboxyl terminal of GST. Unfortunately, the protein was insoluble in buffer containing 1% Triton-X-100, the most commonly used buffer for affinity purification of GST-based fusion proteins. The protein was also insoluble in other mild detergents recommended for the GST fusion protein system, including 1% Tween-20, 10 mM dithiothreitol, and 0.03% SDS (data not shown). However, the fusion protein could be solubilized in a buffer containing sarkosyl. At a sarkosyl concentration of 0.025%, approximately 50% of the fusion protein was solubilized and about 20% of the protein bound to glutathione-Sepharose 4B. Solubilization increased with higher concentrations of sarkosyl and approximately 100% of the protein was soluble in 0.5% sarkosyl. However, high concentrations of sarkosyl interfered with the binding of the fusion protein to glutathione-Sepharose 4B. The procedure was optimized and a sarkosyl concentration of 0.025% was subsequently used for solubilization and purification of the GST-HCV core fusion protein.

Fusion protein could be purified from the bacterial cell lysate using glutathione-Sepharose 4B affinity chromatography (Fig. 5). A major band of 42 kDa containing purified GST-HCV core fusion protein was seen as expected. How-





ever, there were also several faint bands of lower molecular weights. Large scale production and purification of the fusion protein can be carried out in a 500 ml bacterial culture system, yielding 13.09 mg of purified protein. The yield of the purified protein was approximately 26.2  $\mu\text{g}/\text{ml}$  of bacterial culture.

#### Immunogenicity of the purified fusion protein

Immunoblot analysis of the fusion protein using plasma from blood donors who had antibodies to HCV showed that the antibodies could bind to the 42 kDa purified protein, but not to the GST carrier protein (Fig. 6). The antibodies could also bind to the small bands.

All plasma samples from donors infected with HCV genotypes 1a (n=3), 1b (n=3), 3a (n=5) and 7 (n=1) could react with the purified GST-HCV core fusion protein cloned from HCV genotype 1b (BB51), but not to the GST. Plasma from 4 patients infected with dengue virus, a member of family *Flaviviridae*, as well as that from 4 patients who had no antibodies to HCV did not react either to the fusion protein or to GST.

#### Production of monoclonal antibodies to the fusion protein

A total of 266 hybridoma clones were identified from 960 wells plated. Cell culture supernatants from 5 wells contained antibodies

to the GST-HCV core fusion protein, as detected by enzyme immunoassay. Four clones reacted to GST-HCV core fusion protein only, but not to GST, and were thus specific for the HCV core portion of the fusion protein. The other clone reacted with both the GST and GST-HCV core proteins suggesting that it was specific to the GST portion of the protein. Three clones were selected and recloned three times by limiting dilution. These HCV-specific (C1 and C7) and GST-specific (C4) clones were found to produce monoclonal antibodies to the purified HCV core and GST proteins, respectively, as detected by immunoblot analysis and enzyme immunoassay (Fig. 7).

#### DISCUSSION

The genes encoding core protein of HCV from 5 Thai blood donors were amplified and sequenced. Sequence analysis showed that the sequences of the entire core region of Thai strains were similar to the prototype HCV-1 (genotype 1a) and HCV-J (genotype 1b) strains. They were also similar to that of strain NZL1 of genotype 3a, especially at the 5' end of the sequence. These 3 genotypes are the most common HCV genotypes found in Thailand, comprising about 90% of HCV.<sup>12,16</sup> The nucleotide sequence of genotypes 7-9, which may also be found in Thailand, are not currently available and thus no comparison to known sequences could be made. However, these data suggest that the overall homology of the amino acid sequences in the core among various isolates and genotypes is high, approximately 95%. Therefore, in contrast to the NS4 protein, recombinant HCV core protein from one genotype should be recognized by antibodies from patients infected with HCV of other genotypes.

A prokaryotic expression plasmid pGEX-3X was used for

expressing the HCV core protein in this study. This cloning system allows the expression of cloned HCV protein as a fusion protein at the carboxyl terminal of GST which can be easily purified using glutathione-containing affinity matrices.<sup>14</sup> GST are not normally present in bacterial cells, thus the purification of GST fusion protein can be accomplished by expressing the appropriate plasmid in a bacterial host, preparing a crude detergent lysate, and doing affinity purification with glutathione-Sepharose 4B. Once purified, the protein of interest can be cleaved away from GST since the adapter between GST and the protein of interest contains a proteolytic cleavage site for either thrombin (in pGEX-2T vector) or Factor X<sub>a</sub> (in pGEX-3X vector).

The plasmid containing the entire HCV core gene did not express the fusion protein. Several *E. coli* strains, including JM101, TG1 and DH5 $\alpha$ , were transformed but none produced such a protein. This is most likely due to the toxicity of the protein for bacterial cells since some GST-fusion proteins are toxic to *E. coli*.<sup>17</sup> In addition, several GST-fusion proteins are unstable and subjected to proteolytic degradation.<sup>17</sup> It has also been shown that the carboxyl-terminal region of the E1 glycoprotein of HCV interfered with the expression of GST-based fusion proteins.<sup>18</sup> Truncation of the putative membrane association region circumvented the difficulty of expressing the E1 protein in *E. coli*.<sup>18</sup> Similar problems may occur with GST-HCV core protein expression.

To overcome the problem; the 3' end of the core gene was truncated resulting in a clone containing an equivalent of the first 123 amino acids of the core protein. This protein should still be immunogenic since the antigenic epitopes of this protein have been shown to reside in the first 40 amino acid residues.<sup>19</sup>

In the present study, GST-HCV core fusion protein could be expressed after induction with IPTG. However, this protein is insoluble in standard buffers recommended for purification of GST-based fusion proteins, including Triton-X-100, Tween-20, dithiothreitol and SDS. The insolubility of the protein is due to the hydrophobicity of the core protein since this protein functions as a RNA-binding protein in the virus.<sup>3</sup> However, the fusion protein could be solubilized in Sarkosyl, a stronger anionic detergent. Sarkosyl was chosen because other proteins solubilized by this agent were able to bind to glutathione and subsequently purified.<sup>20,21</sup> Sarkosyl acts by disrupting the interactions that maintain the expressed protein in coaggregate with bacterial outer membrane components.<sup>21</sup> Subsequent purification of GST-HCV core fusion proteins was successful. After purification, the protein was dialysed in PBS and the soluble fraction collected. Although binding to glutathione is reduced in Sarkosyl buffer, the overall yield of purified fusion protein remains satisfactory. The faint bands of lower molecular weights after purification are probably degradation products of the GST-HCV core fusion protein since they also reacted to HCV antibodies.

Immunoblotting confirmed that the purified GST-HCV core fusion protein is immunogenic. Antibodies from plasma of HCV-infected individuals bound to the fusion proteins but not to GST. Furthermore, this protein could be detected by antibodies from patients infected with all four Thai HCV genotypes. This protein is therefore useful as an antigen for detection of antibodies to HCV and could be incorporated into a diagnostic assay for HCV in Thailand. It may be used in fusion protein form or as a purified HCV core protein after cleavage by Factor X<sub>a</sub>.

In addition, murine monoclonal antibodies to the HCV core protein and GST have been produced. These antibodies will be valuable for further studies of the pathogenesis of HCV infection and for developing a system for detecting hepatitis C virus antigen in clinical specimens and in tissue sections.

#### ACKNOWLEDGEMENTS

This work was supported by the National Science and Technology Development Agency of Thailand, Mahidol University and the Anandamahidol Foundation. We thank Dr S Bejarachandra and Prof D Chandanayingyong for providing the clinical samples.

#### REFERENCES

1. Kuo G, Choo QL, Alter HJ, *et al.* An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* 1989; 244 : 362-4.
2. Anonymous. Public Health Service Interagency guidelines for screening donors of blood, plasma, organs, tissues, and semen for evidence of hepatitis B and hepatitis C. *MMWR* 1991; 40 (RR-4) : 1-17.
3. Choo QL, Richman KH, Han JH, *et al.* Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci USA* 1991; 88 : 2451-5.
4. Bukh J, Purcell RH, Miller RH. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proc Natl Acad Sci USA* 1993; 90 : 8234-8.
5. Barbara JAJ, Contreras M. Non-A, Non-B hepatitis and the anti-HCV assay. *Vox Sang* 1991; 60 : 1-7.
6. Van der Poel CL, Reesink HW, Schaasberg W, *et al.* Infectivity of blood seropositive for hepatitis C virus antibodies. *Lancet* 1990; 335 : 558-60.
7. Bukh J, Miller RH, Kew MC, Purcell RH. Hepatitis virus RNA in southern African Blacks with hepatocellular carcinoma. *Proc Natl Acad Sci USA* 1993; 90 : 1848-51.

8. Inaba S, Fukuda M, Okochi Y, *et al.* HCV transmission after receiving anti-c 100-negative blood units. *Lancet* 1991; 337 : 1354.
9. Simmonds P, Rose KA, Graham S, *et al.* Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV) : use of type specific peptides to serologically differentiate infection with HCV types 1, 2, and 3. *J Clin Microbiol* 1993; 31 : 1493-503.
10. Bresters D, Cuypers HTM, Reesink HW. *et al.* Enhanced sensitivity of a second generation ELISA for antibody to hepatitis C virus. *Vox Sang* 1992; 62 : 213-7.
11. Uyttendaele S, Claeys H, Mertens W, Verhaert H, Vermeylen C. Evaluation of third-generation screening and confirmatory assays for HCV antibodies. *Vox Sang* 1994; 66 : 122-9.
12. Luengrojanakul P, Vareesangthip K, Chainuvati T, *et al.* Hepatitis C virus infection in patients with chronic liver disease or chronic renal failure and blood donors in Thailand. *J Med Virol* 1994; 44 : 287-92.
13. Songsivilai S, Dharakul T, Senawong S. Hepatitis B- and hepatitis C-associated hepatocellular carcinoma : evaluation of alpha-fetoprotein as a diagnostic marker. *Asian Pac J Allergy Immunol* 1995; 13 : (in press).
14. Smith DB, Johnson KS. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 1988; 67 : 31-40.
15. Harlow E, Lane D. *Antibodies : A laboratory manual.* Cold Spring Harbor : Cold Spring Harbor Laboratory Press, 1988 : 53-328.
16. Apichartpitayakul C, Chittivudikarn C, Miyajima H, Homma M, Hotta H. Analysis of hepatitis C virus isolates among healthy blood donors and drug addicts in Chiang Mai, Thailand. *J Clin Microbiol* 1994; 32 : 2276-9.
17. Sharrocks AD. A T7 expression vector for producing N- and C-terminal fusion proteins with glutathione S-transferase. *Gene* 1994; 138 : 105-8.
18. Yan B-S, Liao L-Y, Leou K, Chang Y-C, Syu W-J. Truncating the putative membrane association region circumvents the difficulty of expressing hepatitis C virus protein E1 in *Escherichia coli*. *J Virol Methods* 1994; 49 : 343-52.
19. Nasoff MS, Zebedee SL, Inchauspe G, Prince AM. Identification of an immunodominant epitope within the capsid protein of hepatitis C virus. *Proc Natl Acad Sci USA* 1991; 88 : 5462-6.
20. Frangioni JV, Neel BG. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem* 1993; 210 : 179-87.
21. Frankel S, Sohn R, Leinwand L. The use of sakosyl in generating soluble protein after bacterial expression. *Proc Natl Acad Sci USA* 1991; 88 : 1192-6.