Characterization, cloning and expression of NS3 protein gene of hepatitis C genotype 4a

Article *in* Journal of the Egyptian Society of Parasitology · December 2009

Source: PubMed

4 authors, including:

Salwa Sabet
Cairo University
40 PUBLICATIONS **183** CITATIONS
SEE PROFILE

Maged M Al-Sherbiny
Cairo University
28 PUBLICATIONS **622** CITATIONS
SEE PROFILE

Essam H Ibrahim
King Khalid University
59 PUBLICATIONS **231** CITATIONS
SEE PROFILE

J. Egypt. Soc. Parasitol., 39 (3), 2009: 865-880

CHARACTERIZATION, CLONING AND EXPRESSION OF NS3 PROTEIN GENE OF HEPATITIS C GENOTYPE 4A

By

SALWA F. SABET[1], MAGED M. AL-SHERBINY[2], ESSAM H. IBRAHIM[3] AND PAUL HAGAN[4]

Department of Zoology, Faculty of Science, Cairo University[1, 2] National Organization for Drug Control and Research (NODCAR) 51 Wezaret El Zeraa street, Cairo, Egypt[3], Division of Infection and Immunity, Joseph Black Building (B4-09d), Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK[4]

Abstract

Clone and express NS3 gene of the Egyptian strain ED43 of HCV genotype 4a in *E. coli* was studied. Gene and protein sequences of NS3 gene of the ED43 strain were first analyzed using PC/GENE program. DNA homology was 89% the homologies and that of the protein was 78.8% indicating that NS3 gene of the genotype 4a is different from those isolated from other strains.

DNA of NS3 region of genotype 4a was amplified from HCV_ED43/PUC19 plasmid. The PCR product was cloned and expressed in *E. coli* M15 using pQE-30 vector. Fusion protein containing the peptides coded by HCV NS3 (NS3_4a) was expressed by *Escherichia coli*. The specific HCV antigenicity of the NS3_4a fusion protein was identified by western blotting.

Keywords: HCV, Genotype 4a, Egyptian strain ED43, NS3_4a, Cloning, Expression

Introduction

Hepatitis C virus (HCV) infects more than 10% of the general population in Egypt when intravenous injection with an antimony compound for endemic schistosomiasis in the past was implicated (Tanaka et al, 2004). Egypt could be considered a candidate country for performing trials of prophylactic and therapeutic vaccines as it has higher rates of HCV than neighboring countries as well as other countries in the world with comparable socioeconomic conditions and hygienic standards for invasive medical, dental, or paramedical procedures (Lavanchy and McMahon, 2000). Little is known about genotype 4a which is the most widely distributed of type 4 sequences, being the principal genotype in the Middle East and Egypt, and accounts for a major

proportion of cirrhosis and hepatocellular carcinoma in these populations this genotype (Dusheiko *et al*., 1994; Ray *et al*., 2000).

The NS3 protein has approximately 70 kd and involved in critical events of viral replication, thus making it an attractive target for antiviral therapy (Hong *et al*., 2000, Sillanpää *et al*., 2009). Analysis showed motifs characteristic of RNA helicase enzymatic function (Tai *et al*., 1996). Genomic analysis of the NS3 protein could be characterized as a serine protease (De Francesco and Steinkühler 2000) as well as an ATP-driven motor activity (Penin *et al*., 2004).

In the present study, NS3 gene of the Egyptian strain ED43 of HCV genotype 4a was chosen to be cloned and expressed in *E. coli.* Gene and protein sequence of NS3 gene of was first analyzed using PC/GENE program, and then amplified from HCV_ED43/PUC19 plasmid. The PCR products were then cloned and expressed in M15 strain of *E. coli* using pQE-30 vector. The expression and antigenicity of the NS3 proteins in *E. coli* was evaluated by SDS-PAGE and Western blotting.

Materials and Methods

Analysis of NS3 gene of HCV genotype 4a genome: PUC19 vector con-taining HCV genotype 4a DNA genome of Isolate ED43 (HCV_ED43) was provided by Dr. Richard M. Elliott (Institute of Virology, University of Glasgow, Glasgow). The complete nucleotide sequence of HCV_ED43 which is 9355 nucleotide long is deposited in the EMBL database under accession number Y11604.

Analysis of the nucleic acid sequence and the deduction of amino acid sequences of the NS3 (NS3_ED43) gene of HCV_ED43 were performed by PC/GENE software programs (Intelligentics, Inc.). These programs have been used to determine the open reading frame, restriction enzymes cleavage sites as well as the primary, secondary structure analysis of the deduced amino acid sequences. The nucleotide sequence of the NS3_ED43 was subjected to similarity search using standard nucleotide-nucleotide BLASTN option of the internet (http://www.ncbi.nlm.gov/BLAST) (Altschul *et al*., 1997). Multiple DNA and protein alignments of NS3_ED43 were done with some of the isolates which showed a high degree of similarity using PC/GENE program.

Cloning Strategy of the NS3_ED43 gene

The original DNA concentration of the HCV_ED43 genome was 0.98 µg/µl and was diluted to a final concentration of 0.1 µg/µl by adding 10 µl of DNA to 90 µl sterile distilled water (dH_2O).

The cloning strategy of the NS3_ED43 and gene was done (Sambrook *et al.,* 1989) and the QIAexpressionist kit instruction manual as: amplification of DNA

region of interest using whole PUC19/HCV_ED43 plasmid as a template. A restriction enzyme target site was introduced into each of the PCR primers. The restriction sites of the enzymes that were not found to cut inside the target sequence and at the same time are present in the multiple clonal site (MCS) of the appropriate vector, were chosen to be added to the primers. The resulting PCR product and cloning vector were digested with the restriction enzymes to generate complementary ends at the PCR product and the vector which were then ligated and transformed into *E. coli.*. DNA sequences coding for the whole NS3 protein was amplified from PUC19/HCV_ED43. The sequence of primers used was: NS3_ED43F primer: 5'-CGCGGATCCGCCCCCATCACAGCATAC GC-3'.NS3_ED43R primer: 5'-GACGTCGAC TGTCA CTACC TC GAGATCAGC -3'. The amplified region was cloned into pQE30 cloning and expression vector between the *Bam*HI and *SalI* Sites.

Expression of NS3 (NS3_4a) fusion protein: Recombinant M15 cells were grown overnight at 37°C in 20 ml of Luria–Bertani (LB) broth (1% tryptone, 0.5% yeast extract, 1% NaCl) containing 100 µg/ml ampicillin, 25 µg/ml kanamycin ($LB_{kan,amp}$). The 20 ml culture was inoculated into 1 litre of $LB_{kan,amp}$ media and cultured at 37°C with vigorous shaking for 1 h. Expression of recombinant protein was induced by adding isopropyl *b*-D-thiogalactoside (IPTG) to a final concentration of 1 mM. After 4 h of induction, cells were harvested by centrifugation 4000 rpm for 20 min & frozen overnight at -20°C until needed for purification.

Preparation of pure lysates was done under denaturing conditions on Ni2+-nitrilotriacetate (NTA)–agarose (Qiagen) according to the manufacturer's instructions. The cell pellet was thawed for 15 min on ice and resuspended in buffer B (100 mM NaH_2PO_4, 10 mM Tris·Cl, 8 M urea, pH to 8.0) at 5 ml/g wet weight (25 ml buffer was added to 5 g cells). The cells were gently vortexed for 20 min at RT until the solution became translucent. The lysate was centrifuged at 10,000 rpm for 30 min at RT to pellet the cellular debris. The supernatant (cleared lysate) was saved and the cell debris was discarded. Five ml of the 50% Ni-NTA slurry was added to 25 ml lysate and mixed gently by shaking (200 rpm on a rotary shaker) for 60 min at RT. The lysate–resin mixture was carefully loaded into an empty column with the bottom cap still attached. The bottom cap was removed and the flow-through was collected. The column was washed twice with 4 ml buffer C (100 mM NaH_2PO_4, 10 mM Tris·Cl, 8 M urea, pH to 6.3). The recombinant protein was eluted by adding 2.5 ml buffer D (100 mM NaH_2PO_4, 10 mM Tris·Cl, 8 M urea, pH to 5.9) 4 times and followed with 2.5 ml buffer E (100 mM NaH_2PO_4, 10 mM Tris·Cl, 8 M urea, pH to 4.5) 4 times. A 10

µl 5X SDS-PAGE sample buffer was added to 40 µl of each elution and stored at –20 °C for SDS-PAGE analysis. The samples were analysed in 12% SDS-PAGE. The protein content of purified samples was measured using BioRad protein assay reagent (BioRad)

For Western-blot analysis, fusion protein was electro-transferred on to nitrocellulose membrane (Wattman). The membrane was blocked with 1% bovine serum albumin in Tris-buffered saline (TBS, 20 mM Tris-Cl, pH 7.5, 500 mM NaCl) for 2 h at RT with shaking.. The membrane was then incubated with anti-HCV pooled human sera (from Egyptian blood donors infected with hepatitis C), washed, then incubated with second antibody [horseradish peroxidase (HRP)-labelled Protein A (Sigma)]for 1 h at 37 °C. After washing, the membrane was developed using the enhanced chemiluminescent (ECL2) kit from Amersham. The same steps were repeated using sera from healthy Egyptian donors to ensure the proteins recognition specify.

Results

HCV Sequence, NS3_ED43 gene of the HCV_ED43 is 1893 bp and lies between nucleotide bases 3358 and 5250 of the genome. Amino acid translation of the NS3_ED43 gene using PC/GENE program revealed a protein that consists of 631 aa and have a molecular weight of 67.5 kDa.

Amino acid analysis of NS3_ED43 protein showed threonine (11%), glycine (9.3%), valine (8.5%), leucine (8.2%) and alanine (8.2%) (Tab. 1).

Table 1: Absolute and % of amino acid residues of NS3_ED43 deduced protein.

Code	Nb	%	Code	Nb	%
Thr	70	11	Glu	22	3.4
Gly	59	9.3	Lys	20	3.1
Val	54	8.5	Gln	19	3
Leu	52	8.2	Phe	19	3
Ala	52	8.2	Cys	18	2.8
Ser	43	6.8	His	14	2.2
Pro	43	6.8	Asn	13	2
Asp	31	4.9	Met	10	1.5
Arg	30	4.7	Trp	7	1.1
Ile	30	4.7	Xaa	1	0.1
Tyr	24	3.8			

The estimated half-life is considered to be 4.4 hour in mammalian reticulocytes, in vitro, more than 20 hour in yeast, in vivo, and more than 10 hour in *E. coli*, in vivo.

The isoelectric point (pI) of the NS3_ED43 deduced protein was found to be 6.54

The three highest points of hydrophilicity of the NS3_ED43 protein were determined. The first was located from amino acid 371 to 376 (Lys-Lys-Lys-Cys-Asp-Glu) with an average hydrophilicity (Ah) of 2.33, second was located from amino acid 117 to 123(Arg-Arg-Arg-Gly-Asp-Thr-Arg) with an Ah of 1.93 and third was from amino acid 457 to 462 (Ser-Arg-Ser-Gln-Arg-Arg) with an Ah 1.63.

Sites annotated analysis of the deduced NS3_ED43 protein showed one N-glycosylation site, one tyrosine sulfatation site, 11 protein kinase C phos-phorylation sites, 10 casein kinase II phosphorylation sites, 16 N-myristoylation sites, one gram-positive cocci surface proteins `anchoring' hexapeptide, one cell attachment sequence and one ATP/GTP-binding site motif A (P-loop).

The BlASTN (for DNA) program was used to search the similarity between NS3_ED43 DNA sequence and other DNA sequences, using non redundant GenBank+EMBL+DDBJ+PDB sequences database. Three isolates were chosen for comparison with NS3_ED43. BLASTN showed similarities of 89% with isolate 25 of genotype 4a (Is. 25), 82% (homology search covered 90% only of NS3_ED43 sequence) with isolate 24 of genotype 4d (Is. 24), and 87% (covered 45% of NS3_ED43 sequence) with isolate MD4-2 of genotype 4 (Is. MD4-2). DNA of the isolates was multiple aligned with NS3_ED43 DNA sequence using PC/GENE program.

Amino acid translation of the nucleotide sequences of the three isolates that were aligned with NS3_ED43 gene was done and multiple protein alignment of NS3_ED43 protein with the deduced proteins were done using PC/GENE program (Fig. 1). Alignment revealed an identity of 78.8% between the NS3_ED43 and the three proteins.

DNA sequence coding for NS3_ED43 was amplified from PUC19/ HCV _ED43 and the expected size (1.9 kb) of PCR product was obtained (Fig.2).

PCR product of the amplified NS3_ED43 DNA was purified from gel by High Pure PCR Purification Kit (Roche), double digested with the restriction enzymes *BamHI* and *SalI*., and purified from solution. The purified double digested DNA concentration was 20ng/µl for each band (Fig. 3). PCR product cloned to pQE30, miniprep took place and electrophoresis of the double digested DNA with *BamHI* and *SalI* showed a band of around 1.9 kb which is the size of the NS3_4a insert (Fig. 4). 5´ DNA sequence analysis of NS3_4a region of recombinant pQE-30/NS3_4a plasmid showed that NS3_4a insert was ligated in frame with the 6xHis-tagged DNA of pQE-30 vector (Fig. 5) and alignment search with origi-

nal HCV_ED43 by BLAST program showed that the sequence of the insert is similar to that of the original NS3_ED43 (Fig. 6).

M15 bacterial cells containing recombinant plasmids pQE 30/NS3_4a was induced with 1mM IPTG. Four hours after induction, the cells were harvested, samples of whole-cell lysates were prepared and analysed by SDS-PAGE. A band of approximately 70 kDa was in good agreement with predicted molecular mass (Fig. 7). Nearly all of the expressed fusion protein was found in the insoluble fraction after sonication (Fig. 8). NS3_4a protein was solubilized with 8 M urea and was purified on Ni2+-NTA–agarose under denaturing conditions. After washing, bound proteins were eluted at low pH, and fusion proteins of high purity were obtained (Fig. 9). The yield was about 8.44 mg/2litre culture.

Western blot analysis checked the NS3_4a protein antigenicity. Expressed NS3_4a protein was recognized by the anti-HCV antibodies present in the pooled human sera of Egyptian patients infected with HCV and a band was approximately at 70 kDa corresponding to NS3_4a (Fig. 10). Specificity by using sera from healthy donors showed no specific recognition of the NS3_4a protein.

Discussion

Egypt could be considered a candidate country for performing trials of preventive and therapeutic vaccines as it has higher rates of HCV than neighboring countries as well as other countries in the world with comparable socioeconomic conditions and hygienic standards for invasive medical, dental, or paramedical procedures (Lavanchy and McMahon, 2000). Although Egypt has a very high prevalence of HCV and a high morbidity and mortality from chronic liver disease, cirrhosis, and hepatocellular carcinoma, not much is known about genotype 4a which is the most widely distributed of type 4 sequences, being the principal genotype in the Middle East and Egypt (Dusheiko *et al*., 1994; Ray *et al*., 2000).

NS3 protein was chosen in the present study for its importance as a multifunctional virus-specific protein that possesses multiple enzymatic activities (Misialek *et al*., 2009). NS3 contains serine protease activity in its N-terminal region and accounts for processing of the viral polyprotein at four cleavage sites, NS3/ 4A, NS4A/4B, NS4B/5A, and NS5A/5B. The serine proteinase activity of NS3 is an attractive target for new drugs that could block viral replication efficiently (Lordini *et al*., 2003; Penin *et al*., 2004). Also, helicase and nucleic acid-stimulated nucleoside triphosphatase activities are found in its C-terminal region (Bartenschlager, 1997; Kwong et al., 1998). The NS3 helicase– NTPase domain probably has multiple functions, including RNA-stimulated NTPase activity,

RNA binding, and unwinding of RNA regions with extensive secondary structure by coupling unwinding and NTP hydrolysis. This enzyme acts as an ATP-driven motor and is thought to switch between alternative conformations during active unwinding of double-stranded RNA (Penin *et al*., 2004).

In the present study, prior to cloning, the DNA sequence of the NS3 region (NS3_ED43) of the HCV-ED43 was analyzed using PC/GENE program. NS3 (NS3_ED43) gene of the HCV-ED43 is 1893 bp and lies between nucleotide bases 3358 and 5250 of the genome. The DNA sequence analysis revealed that it has a relatively higher proportion of GC bases (56.7%) than AT bases (43.2%) with a base percentage of 27.3 % G, 29.4 % C, 22.8% A and 20.3% T. Amino acid translation of the NS3_ED43 gene using PC/GENE program revealed a protein that consists of 631 aa and has a molecular weight of 67.5 kDa. Amino acid composition analysis of the NS3_ED43 protein revealed the presence of a high percentage of threonine (11%), glycine (9.3%), valine (8.5%), leucine (8.2%) and alanine (8.2%). The N-terminal of the NS3_ED43 protein sequence is A (Ala).The estimated half-life is considered to be 4.4 hour in mammalian reticulocytes, in vitro, more than 20 hour in yeast, in vivo, and more than 10 hour in *E. coli*, in vivo.

The three highest points of hydrophilicity of NS3_ED43 protein were determined. The first point was located from amino acid 371 to 376 with an average hydrophilicity (Ah) of 2.33, second point was located from amino acid 117 to 123 with an Ah of 1.93 and third was from amino acid 457 to 462 with an Ah 1.63. No previous data of the antigenic determinants of the NS3 proteins of the other HCV genotypes was available. Sites annotated analysis of the predicted protein encoded by NS3_ED43 revealed the presence of one N-glycosylation site. It is known that potential N-glycosylation sites are specific to the consensus sequence Asn-Xaa-Ser/thr. The presence of the consensus tripeptide is not sufficient to conclude that an asparagine residue is glycosylated, due to the fact that the folding of the protein plays an important role in the regulation of N-glycosylation (Pless and Lennarz, 1977). No glycosylation was mentioned for NS3 in the previous studies which suggest that NS3 folding might have depressed this site activity. Also, one Tyrosine sulfatation site, 11 protein kinase C phosphorylation sites, 10 Casein kinase II phosphorylation sites that might be responsible for interactions with protein kinases A and C. Numerous NS3 interactions with cellular components have been reported, including protein kinases A and C, p53, and histones H2B and H4, but their significance is unclear (Tellinghuisen and Rice, 2002). One Gram-

positive cocci surface protein `anchoring' hexapeptide has been reported as well as one Cell attachment sequence and one ATP/GTP-binding site motif A (P-loop) which might be responsible for the NS3 helicase– NTPase domain interaction with ATP. 16 N-myristoylation sites were also found along the sequence, Myristoylation is essential for the biological function of most proteins. As attachment of the myristoyl residue to glycine residues provides hydrophobicity and promotes protein-protein interactions (Johnson *et al*., 1994), it might have other vital functions in the NS3 protein other than providing hydrophobocity as NS3 is not known to have hydrophobic regions.

The BlASTN (for DNA) program gave the similarity between NS3_ED43 DNA sequence and other DNA sequences, Three NS3 isolates from those which showed significant similarity were chosen for multiple alignments with NS3_ED43. BLASTN Homology search revealed similarities of 89% with isolate 25 of genotype 4a (Is. 25), 82% (homology search covered 90% only of NS3_ED43 sequence) with isolate 24 of genotype 4d (Is. 24), and 87% (covered 45% of NS3_ED43 sequence) with isolate MD4_2 of genotype 4 (Is. MD4_2). These results support the findings that suggested that the Egyptian genotype 4a is different from those isolated from other parts of the world (Stuyver *et al*., 1994; Bukh *et al*., 1995; Alfonso *et al*., 2001).

Amino acid translation of the nucleotide sequences of the three isolates that were aligned with NS3_ED43 gene and multiple protein alignment of NS3_ED43 protein with the deduced proteins were done using PC/GENE program. Alignment revealed an identity of 78.8% between the NS3_ED43 and the three proteins. Amino acid composition comparison between NS3_ED43 and the proteins of the three isolates was done and revealed that the amino acid composition is almost conserved in the four isolates.

Computer analysis of NS3_ED43 was followed by cloning and expression of NS3_ED43 DNA region in pQE-30/M15 *E.coli*. In previous studies, Fragments Vishnuvardhan *et al*., 1997; Jiao *et al*., 2004, Frick, 2007), whole (Poliakov *et al*., 2002, Cheng *et al*., 2002), as well as fusion proteins from different ligated gene fragments such as NS3-NS4a (Du *et al*., 2002; Thibeault *et al*., 2004).

Amplification of NS3_ED43 DNA began by cloning of NS3_ED43 DNA via PUC19/HCV_ED43 plasmid as a template. Core insert into pQE30 vector, t *BamHI* and *SalI* restriction sites were introduced to the forward and reverse primers respectively. pQE-30 vector with PCR product of the core region (designated as NS3_4a) was used to transform M15(pREP4) bacteria supplied with the kit. Successful transformation was confirmed by plasmid miniprep

and gel electrophoresis showing the DNA bands of the expected size (1.9kb).

Optimized conditions for expression of NS3_4a protein was done according to the manufacturer's manual of QIAexpressionist kit with the only exception that before induction, the bacterial cultures of NS3_4a was grown at 37ºC, SDS-PAGE of the whole induced cultures before purification showed the bands of NS3_4a fusion protein of 69 kDa (Jiao *et al*., 2004). The solubility of expressed NS3_4a protein was checked before preparing the proteins on large scales, and it was found that the NS3_4a protein was expressed in the insoluble form. Thus, large scale preparations of the NS3_4a protein was done under denaturing conditions and Core_4a protein content was estimated to be around 4 mg/liter. Reactivity of NS3_4a against human sera in Western blot was analysed. The expressed NS3_4a protein was recognized by the anti-HCV present in the pooled human sera of Egyptian patients infected with hepatitis C and a band appeared approximately at 69 corresponding to the NS3_4a. Specificity of this recognition was confirmed by using sera from healthy donors which showed no specific recognition of the two proteins. So, this *E. coli*-derived NS3_4a protein displayed specific antigenicity.

Further studies are recommended in experimental animals to specify these proteins antigenicity. Moreover, characterization of more HCV isolates, extracted from the blood of Egyptian patients, will be important for the improvement of diagnostic, epidemiological and clinical treatment regimens as well as the development of a candidate vaccine against HCV genotype 4a.

Conclusion

Analysis of NS3 region of the isolate ED43 revealed that the Egyptian genotype 4a is different from other genotypes. Construction of pQE-30/NS3_4a recombinant plasmid and expression of NS3_4a gene in *E. coli* (M15) was successful and have immunogenicity in Western blot useful in vaccine development studies against the Egyptian genotype 4a. Further studies of NS3 region are recommended in experimental animals and cell cultures to explore its role in diagnostic, epidemiological, treatment and development of a candidate vaccine against HCV genotype 4a.

Acknowledgment

The authors would like to thank the staff, the Institute of Virology, Glasgow University for kind cooperation, especially Prof. Dr. Richard Elliot for his valuable assistance and advices.

References

Alfonso, V.; Flichman, D.; Sookoian, S.; Mbayed, V. A.; Campos, R. H.; 2001: Phylogenetic characterization of genotype 4 hepatitis

C virus isolates from Argentina. J. Clin. Microbiol. 39, 5:1989-92.

Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J.; 1997: Gapped BLAST nad PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res., 25:3389-402.

Bartenschlager, R.; 1997: Molecular targets in inhibition of hepatitis C virus replication. Antivir. Chem. Chemother. 8:281-301.

Bukh, J.; Miller, R.H.; Purcell, R.H.; 1995: Defining the genotypes of hepatitis C virus and their epidemiology. INTER-action 3:10-15.

Cheng, J.; Zhong, Y.; Liu, Y.; Dong, J., Yang, J.; Yang, S.; 2002: Inducible expression of non-structural protein 3 of hepatitis C virus in *E. coli*. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi, 16, 1:85-7.

De Francesco, R.; Steinkühler, C.; 2000: Structure and function of the hepatitis C virus NS3-NS4A serine proteinase. Curr. Top. Microbiol. Immunol. 242:149-69.

Du, G.X.; Hou, L.H.; Guan, R.B.; Tong, Y.G.; Wang, H.T.; 2002: Establishment of a simple assay in vitro for hepatitis C virus NS3 serine protease based on recombinant substrate and single-chain protease. Wld. J Gastroenterol. 8, 6:1088-93.

Dusheiko, G.; Weiss, H.; Brown, D.; McOmish, F.; Yap, P.L.; *et al*.; 1994: Hepatitis C virus genotypes: an investigation of type-specific differences in geographic origin and disease. Hepatol., 19:13-8.

Frick, D.N.; 2007: The hepatitis C virus NS3 protein: a model RNA helicase and potential drug target: Review. Mol. Biol., 9, 1:1-20.

Hong, Z.; Standring, D.N.; Baroudy, B.; Lau, J.Y.; 2000: Development of novel anti-HCV therapies: HCV protease, helicase, and polymerase as thera-peutic targets. Acta Gastroenterol. Belgica, 63:210-2.

Jiao, X.; Wang, R.Y.; Feng, Z.; Hu, G.; Alter, H.J.; Shih, J.; 2004: DNA immunization encoding the secreted nonstructural protein 3 (NS3) of hepatitis C virus and enhancing the Th1 type immune response. J. Viral. Hepat. 11, 1:18-26.

Johnson, D.R.; Bhatnagar, R.S; Knoll, L.J.; Gordon, J.I.; 1994: Genetic and biochemical studies of protein N-myristoylation. Ann. Rev. Biochem., 63: 869-914.

Kwong, A.D.; Kim, J.L.; Rao, G.; Lipovsek, D.; Raybuck, S.; 1998: He-patitis C virus NS3/4A protease. Antivir. Res. 40:1-18.

Lavanchy, D.; McMahon, B.; 2000: Worldwide Prevalence and Prevention of HC. In: Hepatitis C. By Liang TJ, Hoofnagle JH, San Diego, Academic Press.

Lodrini, S.; Bagaglio, S.; Canducci, F.; De Mitri, M.; Andreone, P.; *et al*.; 2003: Sequence analysis of NS3 protease gene in clinical strains of hepatitis C virus. J. Biol. Regul. Homeost. Agents 17, 2:198-204.

Misialek, S.; Rajagopalan, R.; Stevens, S.K.; Beigelman, L.; Seiwert, S.D.; Kossen, K.; 2009: Optimization of the multiple enzymatic activities of the hepatitis C virus NS3 protein. Anal. Biochem., 394,1:138-40.

Penin, F.; Dubuisson, J.; Rey, F. A.; Moradpour, D.; Pawlotsky, J. M.; 2004: Structural biology of hepatitis C virus. Hepatol., 39, 1:5-19.

Pless, D.D.; Lennarz, W. J.; 1977: Enzymatic conversion of proteins to glycoproteins. Proc. Natl. Acad. Sci. USA., 74, 1:134-8.

Poliakov A.; Hubatsch I.; Shuman C.F.; Stenberg G.; Danielson U. H.; 2002: Expression and purification of recombinant full-length NS3 protease-helicase from a new variant of Hepatitis C virus. Protein Exp. Purif., 25, 3: 363-71.

Ray, S.C.; Arthur, R.R.; Carella, A.; Bukh, J.; Thomas, D.L.; 2000: Genetic epidemiology of hepatitis C virus throughout Egypt. J. Infect. Dis., 182: 698-707.

Sambrook, J.; Fritsch, E.F.; Maniatis, T.; 1989: Molecular Cloning: Cold Spring Harbor Laboratory Press.

Sillanpää, M. ; Melén, K.; Porkka, P.; Fagerlund, R.; Nevalainen, K.; *et al*.; 2009 : HCV core, NS3, NS4B and NS5A are the major immunegenic proteins in humoral immunity in chronic HCV infection. Virol. J., 24, 6:86.

Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Hernandez, F.; Delaporte, E.; Maertens, G.; 1994: Classification of hepatitis C viruses based on phylo-genetic analysis of the envelope 1 and nonstructural 5B regions and identi-fication of five additional subtypes. Proc. Natl. Acad. Sci. USA, 91, 10:104-13.

Tai, C.L.; Chen, D.S.; Hwang, L.H.; 1996: The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). J. Virol., 70:8477-84.

Tanaka, T.; Tsukiyama, K.; Yamaguchi, K.; Yagi, S.; Tanaka, S.; *et al*.; 1994: Significance of specific antibody assay for genotyping of hepatitis C virus. Hepatol., 19, 6:1347-53.

Tellinghuisen, T.L.; Rice, C.M.; 2002: Interaction between hepatitis C virus proteins and host cell factors. Curr Opin Microbiol., 5:419-27.

Thibeault, D.; Bousquet, C.; Gingras, R.; Lagace, L.; Maurice, R.; *et al*.; 2004: Sensitivity of NS3 serine proteases from hepatitis C virus genotypes 2 and 3 to the inhibitor BILN 2061. J. Virol., 78, 14:7352-9.

Vishnuvardhan, D.; Kakiuchi, N.; Urvil, P.T.; Shimotohno, K.; Kumar, P.; Nishikawa, S.; 1997: Expression of highly active recombinant NS3 protease domain of hepatitis C virus in *E. coli*. FEBS Letters, 402, 2-3:209-12.

```
NS3_ ED43       APITAYXQQTRGLFSTIVTSLTGRDTNENCGEVQVLSTATQSFLGTAVNG    50
Is. 25          APITAYAQQTRGLFSTIITSLTGRDTNENCGEVQVLSTATQSFLGTAVNG    50
Is. 24          APITAYAQQTRGMLGTIITSLTGRDTNENCGEVQVLSTATQSFLGSAING    50
St. MD4-2       APITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCVNG    50
                ****** *****... *.*******.*.  *****.********.. .**

NS3_ ED43       VMWTVYHGAGAKTISGPKGPVNQMYTNVDQDLVGWPAPPGVRSLAPCTCG   100
Is. 25          VMWTVYHGAGSKTISGPKGPVNQMYTNVDQDLVGWPAPPGVKSLTPCTCG   100
Is. 24          VMWTVYHGAGSKTISGPKGPVNQMYTNVDQDLVGWPAPPGVKSLAPCTCG   100
St. MD4-2       ACWTVFHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCG   100
                . ***.****.**..*****..*************.****..**.*****

NS3_ ED43       SADLYLVTRHADVIPVRRRGDTRGALLSPRPISILKGSSGGPLLCPMGHR   150
Is. 25          ASDLYLVTRHADVVPVRRRGDTRGALLSPRPISTLKGSSGGPLLCPMGHA   150
Is. 24          SSDLFLVTRHADVVPVRRRGDTRGALISPRPISTLKGSSGGPLLCPLGHA   150
St. MD4-2       SSDLYLVTRHADVIPVRRRGDTRGSLLSPRPVSYLKGSSGGPLLCPSGHA   150
                ..**.********.**********.*.****.* ************ **

NS3_ ED43       AGIFRAAVCTRGVAKAVDFVPVESLETTMRSPVFTDNSTPPAVPQTYQVA   200
Is. 25          AGLFRAAVCTRGVAKAVDFVPVESLETTMRSPVFTDNSTPPAVPQAYQVA   200
Is. 24          AGIFRAAVCTRGVAKTVDFVPVESLETTMRSPVFSDNSTPPAVPQTYQVA   200
St. MD4-2       VGIFRAAVCTRGVAKAVDFIPVEAMETTMRSPVFTDNSSPPAVPQTFQVA   200
                .*.*************.***.***..*********.***.******..***

NS3_ ED43       HLHAPTGSGKSTKVPAAHAAQGYKVLVLNPSVAATLGFGVYMSKAYGIDP   250
Is. 25          HLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAYGIDP   250
Is. 24          HLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDP   250
St. MD4-2       HLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGTEP   250
                *****************.**********************.*****.*..*

NS3_ ED43       NIRSGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECYSTDSTTI   300
Is. 25          NIRSGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDSTTI   300
Is. 24          NIRSGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATTI   300
St. MD4-2       NIRTGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDSTTI   300
                ***.****************************************.***.***

NS3_ ED43       LGIGTVLDQAETAGVRLTVLATATPPGSVTTPHSNIEEVALPTTGEIPFY   350
Is. 25          LGIGTVLDQAETAGVRLVVLATATPPGSVTTPHSNIEEVALPTTGEVPFY   350
Is. 24          MGIGTVLDKPKTPEPRLAGPPTPTQPGPGKTPHHKKKHAARPTTTXIHLY   350
St. MD4-2       LGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSNTGEIPFY   350
                .*******...*.. **.  .*.*.**. ..** . ...* ..*. ...*

NS3_ ED43       GKAIPLELIKGGRHLIFCHSKKKCDELARQLTSLGLNAVAYYRGLDVSVI   400
Is. 25          GKAIPLELIKGGRHLIFCHSKKKCDELAKQLTSLGLNAVAYYRGLDVSVI   400
Is. 24          GRAIPLSLVKGGRHLIFCHSKKKCDELAKQLSSLGLNAVAYYRGLDVSVI   400
St. MD4-2       GKAIPIEVIKGGRHLIFCHSKKKCDELAAKLSALGLNAVAYYRGLDVSVI   400
                *.***....****************** .*..*****************

NS3_ ED43       PTSGDVVVCATDALMTGFTGDFDSVIDCNTSVIQTVDFSLDPTFSIEITT   450
Is. 25          PTSGDVVVCATDALMTGFTGDFDSVIDCNTSVIQTVDFSLDPTFSIETTT   450
Is. 24          PLSGDVVVCATDALMTGFTGDFDSVIDCNTSVIQTVDFSLDPTFSIETTT   450
St. MD4-2       PTSGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTT   450
                * ****** **********************.*.***********.**.**
```

To be continue

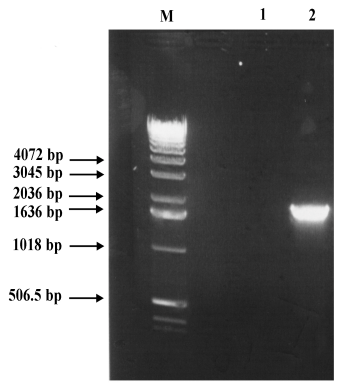
```
NS3_ ED43       VPQDAVSRSQRRGRTGRGRLGTYRYVTPGERPSGMFDTAELCECYDAGCA   500
Is. 25          VPQDAVSRSQRRGRTGRGRLGIYRYVTPGERPSGIFDTSVICECYDAGCA   500
Is. 24          VPQDAVSRSQRRGRTGRGRLGIYRYVTPGERPSGIFDSSVLCECYDAGCA   500
St. MD4-2       VPQDAVSRSQRRGRTGRGRRGIYRFVTPGERPSGMFDSSVLCECYDAGCA   500
                ****************** *.**.*********.**.. .*********

NS3_ ED43       WYELTPAETTTRLKAYFDTPGLPVCQDHLEFWESVFTGLTHIDGHFLSQT   550
Is. 25          WYELTPAETTTRLRAYFNTPGLPVCQDHLEFWESVFTGLTQIDGHFLPQT   550
Is. 24          WYELTPAETTVRLRAYFNTPGLPVCQDHLEFWEGVFTGLTHIDGHFLSQT   550
St. MD4-2       WYELTPAETSVRLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQT   550
                *********..**.**..***************.******.**.***.**

NS3_ ED43       KQSGENFPYLVAYQATVSAKVWLAPPSWDTMWKCLIRLKPTLHGPTPLLY   600
Is. 25          KQSGENFPYLVAYQATVCARALAPPPSWDTMWKCLIRLKPTLHGPTPLLY   600
Is. 24          KQAGDNYPYLVAYQATVCAKALAPPPSWDTMWKCLLRLKPTLRGPTPLLY   600
St. MD4-2       KQAGDNFPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLY   600
                **.*.*.**********.*..  .***** *****.******.*******

NS3_ ED43       RLGSVQNEVVLTHPITKYIMACMSADLEVVT   631
Is. 25          RLGSVQNEVTLTHPITKYIMACMSADLEVVT   631
Is. 24          RLGPVQNEVVLTHPITKYIAACMSADLEVVT   631
St. MD4-2       RLGAVQNEVTLTHPITKFIMACMSADLEVVT   631
                ***.*****.*******.* ***********
```

Fig. 1: Multiple alignments between NS3_ED43 protein and proteins of different HCV isolates. Consensus length: 631, Position in alignment perfectly conserved: '*' Position well conserved: '.', Identity: 497 (78.8%), Similarity: 08 (17.1%)



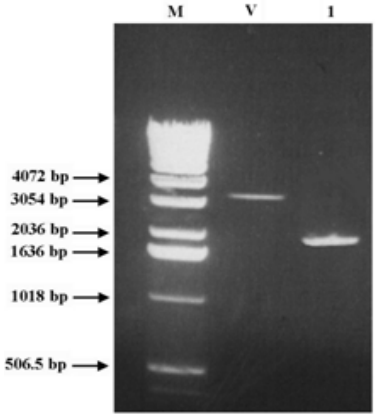


Fig. 2 Fig. 3

Fig. 2: Agarose gel electrophoresis of PCR product after amplification of NS3_ED43 DNA from HCV_ED43 genome. Lane M: 1 kb ladder marker. Lane 2: control sample. Lane 3: PCR product of NS3_ED43 gene.

Fig. 3: Agarose gel electorophoresis of PCR product of NS3_ED43 DNA digested with *BamHI* and *SalI* and purified from solution. Lane M:1 kb ladder marker. Lane V: pQE-30 DNA digested with *BamHI*, Lane 1: pure double digested (*BamHI/SalI*) NS3_ED43 PCR product.

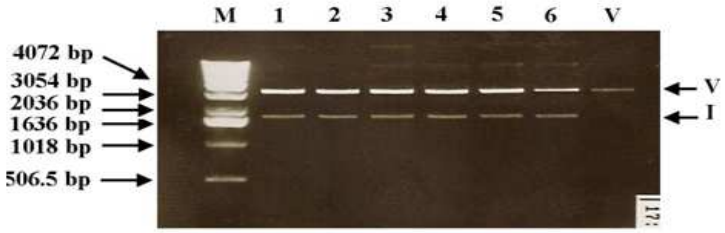


Fig. 4: Double digestion of miniprep samples of recombinant pQE-30/NS3_4a plasmid with *BamHI/XhoI.* Lane M: 1 kb Ladder marker. Lane V: pQE-30 digested with *BamHI/SalI.* Lane 1-5: DNA resulted from minipreps digested with *BamHI/XhoI.* V: vector. I: insert.

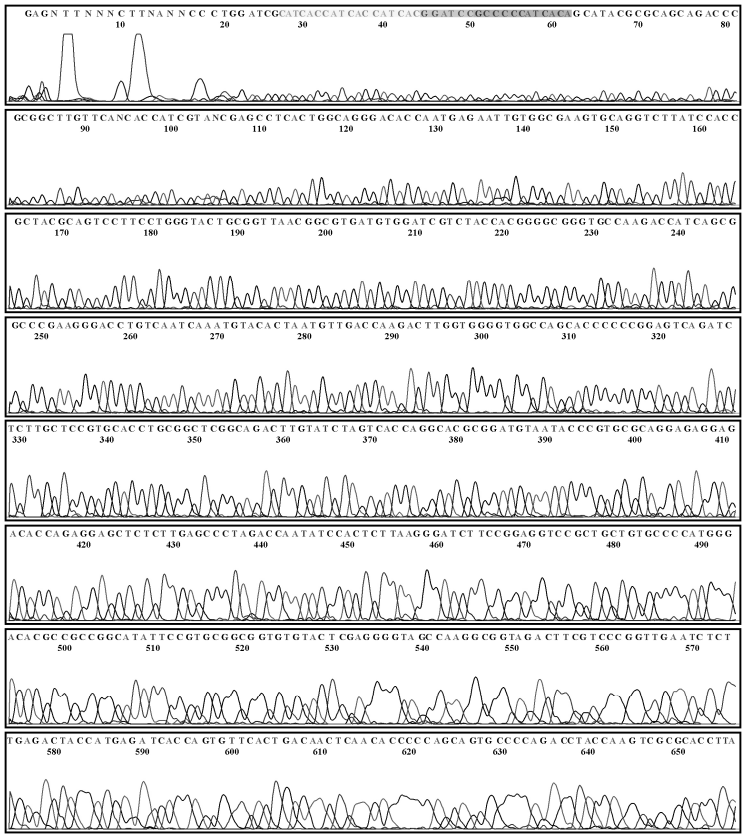


Fig. 5: DNA sequence analysis of NS3_4a insert (NS3-1) inside pQE-30 by ABI PRISM model 310 DNA automated sequencer. Bases 1-50 are pQE-30 vector sequence, bases starting from 51 ~ end partial NS3_4a sequence, Sequence of 6xHis, *BamHI* site, Beginning of NS3_4a sequence.

```
Query  51    GCCCCCATCACAGCATACGCGCAGCAGACCCGCGGCTTGTTCANCACCATCGTANCGAGC  110
             ||||||||||||||||||| |||||||||||||||||||| |||||||||| |||||
Sbjct  3358  GCCCCCATCACAGCATACGCNCAGCAGACCCGCGGCTTGTTCAGCACCATCGTAACGAGC  3417

Query  111   CTCACTGGCAGGGACACCAATGAGAATTGTGGCGAAGTGCAGGTCTTATCCACCGCTACG  170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3418  CTCACTGGCAGGGACACCAATGAGAATTGTGGCGAAGTGCAGGTCTTATCCACCGCTACG  3477

Query  171   CAGTCCTTCCTGGGTACTGCGGTTAACGGCGTGATGTGGATCGTCTACCACGGGGCGGGT  230
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  3478  CAGTCCTTCCTGGGTACTGCGGTTAACGGCGTGATGTGGACCGTCTACCACGGGGCGGGT  3537

Query  231   GCCAAGACCATCAGCGGCCCGAAGGGACCTGTCAATCAAATGTACACTAATGTTGACCAA  290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3538  GCCAAGACCATCAGCGGCCCGAAGGGACCTGTCAATCAAATGTACACTAATGTTGACCAA  3597

Query  291   GACTTGGTGGGGTGGCCAGCACCCCCCGGAGTCAGATCTCTTGCTCCGTGCACCTGCGGC  350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3598  GACTTGGTGGGGTGGCCAGCACCCCCCGGAGTCAGATCTCTTGCTCCGTGCACCTGCGGC  3657

Query  351   TCGGCAGACTTGTATCTAGTCACCAGGCACGCGGATGTAATACCCGTGCGCAGGAGAGGA  410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3658  TCGGCAGACTTGTATCTAGTCACCAGGCACGCGGATGTAATACCCGTGCGCAGGAGAGGA  3717

Query  411   GACACCAGAGGAGCTCTCTTGAGCCCTAGACCAATATCCACTCTTAAGGGATCTTCCGGA  470
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  3718  GACACCAGAGGAGCTCTCTTGAGCCCTAGACCAATATCCATTCTTAAGGGATCTTCCGGA  3777

Query  471   GGTCCGCTGCTGTGCCCCATGGGACACGCCGCCGGCATATTCCGTGCGGCGGTGTGTACT  530
             |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Sbjct  3778  GGTCCGCTGCTGTGCCCCATGGGACACCGCGCCGGCATATTCCGTGCGGCGGTGTGTACT  3837

Query  531   CGAGGGGTAGCCAAGGCGGTAGACTTCGTCCCGGTTGAATCTCTTGAGACTACCATGAGA  590
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3838  CGGGGGGTAGCCAAGGCGGTAGACTTCGTCCCGGTTGAATCTCTTGAGACTACCATGAGA  3897

Query  591   TCACCAGTGTTCACTGACAACTCAACACCCCCAGCAGTGCCCCAGACCTACCAAGTCGCG  650
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  3898  TCACCAGTGTTCACTGACAACTCAACACCCCCAGCAGTGCCCCAGACCTACCAGGTCGCG  3957

Query  651   CACCTTACACGCACCAACA  669
             |||| ||||||||||||||
Sbjct  3958  CACC-TACACGCACCAACA  3975
```

Fig. 6: Similarity search result between sequences of NS3_4a and NS3_ED43 genes. Query: DNA Sequence of NS3 _4a gene. Subject: sequence of original NS3_ED43 DNA. Identities: 609/619 (98%). Gaps: 1/619 (0%). Strand: Plus/Plus.

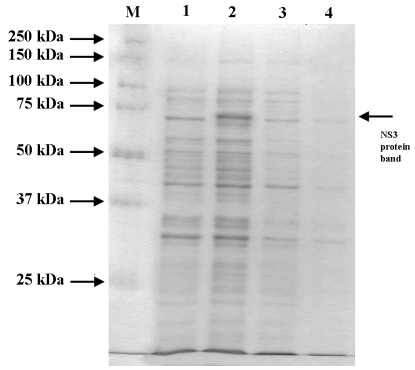


Fig. 7: SDS-PAGE (12% gel) of bacterial cultures resulted from small expression of NS3_4a protein in M15 bacteria. Lane M: prestained wide range MW marker. Lane 1: uninduced culture. Lane 2: induced culture. Lane 3: flow-through. Lane 4: wash.

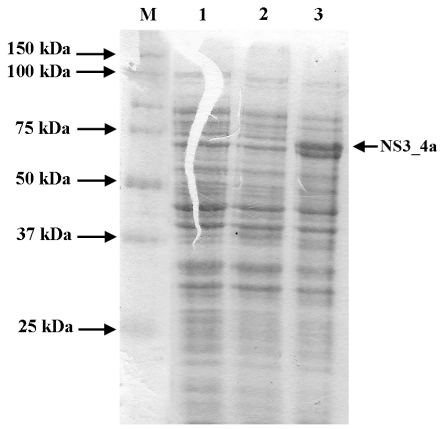


Fig. 8: SDS- PAGE (12% gel) of bacterial cultures resulted from expression of Core_4a and NS3_4a proteins in M15 bacteria to determine solubility. Lane M: prestained wide range MW marker. Lane 1: uninduced culture of NS3_4a protein. Lane 2: soluble fraction of induced culture of NS3_4a protein. Lane 3: insoluble fraction of induced culture of NS3_4a protein.

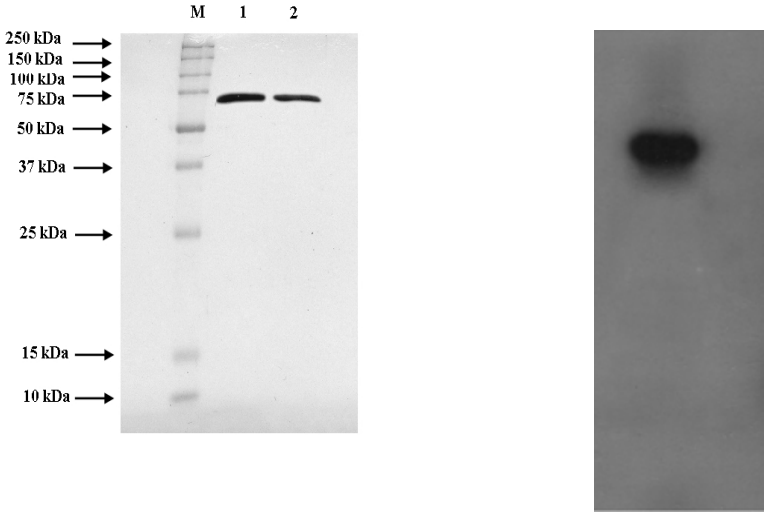


Fig. 9: SDS- PAGE (12% gel) of the NS3_4a protein after purification from large-scale culture. Lane M: prestained wide range MW marker. Lane 1: elution 1 of NS3_4a protein. Lane 2: elution 2 of NS3_4a protein.

Fig. 10: Western blot of purified NS3_4a protein recognized by pooled sera of Egyptian patients with HCV.