1997

Cloning and Expression of Capsid Proteins of Hepatitis a Virus in Escherichia Coli.

Lingaiah Chandrashekar
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/6537

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CLONING AND EXPRESSION OF CAPSID PROTEINS OF HEPATITIS A
VIRUS IN *Escherichia coli*

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Microbiology

by

Lingaiah Chandrashekar
M.B.B.S., Bangalore University, 1989
December 1997
DEDICATION

To

The Loving Memory of

My Grandparents,
Smt. & Sri K. Papanna, and
Smt. & Sri. Channamalle Gowda

and

My father, Sri. C. Lingaiah
ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my major professor, Dr. Ronald J. Siebeling, since without his benevolence, patience, timely advice, supervision, and unfailing moral and material support throughout the course of this research, the completion of this dissertation would have been impossible.

I wish to thank Dr. E.C. Achberger, Dr. R.C. Gayda, Dr. D.S. Shih, Dr. V.R. Srinivasan, and Dr. J. Caprio for serving on my graduate committee. Dr. Srinivasan was one of my guardian angels, throughout; played a devil’s advocate most of the time; and was my sparring partner during our Tuesday morning lab meetings. Dr. Achberger and Dr. Gayda, were the Gurus who taught me the fundamentals of molecular biology; Dr. Shih, along with Dr. Gus Kousoulas who I am very thankful to also, provided stimulating input of ideas and helpful advice whenever I found it necessary.

I am also grateful to Ms. Margaret C. Henk for her friendly advice and encouragement throughout; expert guidance during microscopic studies, preparation of slides and figures; and her generous hospitality time and again.

My special thanks to my very dear friends Ajay, Anand, Anil, Aravind, Arun, Giri, Janet, Jeff Corbett, Jeff Cruz, Kumar, Luis, and Raghu, for their love, friendship and encouragement. My thanks also to my collegues, Amy, Anne, Dan, Marilyn, Philip, Pierrette and Ron Holley, for their enthusiastic participation in the various aspects of my research. My deepest appreciation to the wonderful and always helpful Ms. Joyce Andrews and Ms. Priscilla Milligan, for their generosity throughout the time that I have known them.

iii
I am eternally indebted to Mrs & Mr. B.M. Karunesh, Mrs. & Mr. P. Dayananda, Mrs & Dr. C.K. Hiranya Gowda, Mrs. & Dr. Prabhakaran, Mrs. & Rev. George Corbett, Mrs. & Mr. Sai Pinnepalli, Mrs. & Dr. Yibing Wang for their love, trust and support whenever required.

Finally, I wish to thank my mother, Mrs. Vijayalakshmi; my wife, Kalpana; my grandmothers, Mrs. Narasamma and Mrs. Kalamma; and my brother, Purushotham, for being immeasurable sources of moral strength, inspiration, courage and strong motivation. Without their many sacrifices and their unsurpassed love, abiding faith, sympathy, and loyalty, I am certain that none of this would have been possible.

This research was supported in part by the Louisiana Sea Grant College Program, a part of the National Sea Grant College Program maintained by the National Oceanic and Atmospheric Administration, U.S. Department of Commerce (grant NA 89 AA-D-SG226).
TABLE OF CONTENTS

DEDICATION ................................................................. ii
ACKNOWLEDGEMENTS ................................................ iii
LIST OF TABLES .......................................................... vi
LIST OF FIGURES ......................................................... vii
ABSTRACT ................................................................. x
LITERATURE REVIEW .................................................. 1

CHAPTER I. CLONING AND EXPRESSION OF THE CAPSID
PROTEINS OF HEPATITIS A VIRUS IN *Escherichia coli*
AS β-GALACTOSIDASE FUSION PRODUCTS ................. 17
  INTRODUCTION ......................................................... 18
  MATERIALS AND METHODS .................................. 22
  RESULTS ............................................................... 40
  DISCUSSION .......................................................... 54

CHAPTER II. EFFECTS OF AN ANTI-TERMINATOR SEQUENCE
ON RECOMBINANT HAV PROTEIN EXPRESSION
IN *Escherichia coli* .................................................. 59
  INTRODUCTION ......................................................... 60
  MATERIALS AND METHODS .................................. 64
  RESULTS ............................................................... 73
  DISCUSSION .......................................................... 80

CHAPTER III. ALTERNATIVE APPROACHES FOR CLONING
OF HAV CAPSID GENES FOR EXPRESSION IN
*Escherichia coli* ..................................................... 87
  INTRODUCTION ......................................................... 88
  MATERIALS AND METHODS .................................. 95
  RESULTS ............................................................... 124
  DISCUSSION .......................................................... 144

CONCLUSIONS ............................................................ 156
REFERENCES ............................................................. 162
APPENDIX ................................................................. 177
VITA ............................................................................. 179
LIST OF TABLES

1. Bacterial strains and plasmid vectors used in cloning for expression of HAV capsid genes as β-galactosidase fusions ............................................. 23

2. ELISA with rabbit antiserum for detecting antigenicity of affinity purified recombinant HAV capsid proteins ...................................................... 53

3. Bacterial strains and plasmid vectors used to study the effects of *E. coli* rrn boxA antiterminator on recombinant HAV VP1 expression .................. 65

4. Analysis of N-terminal sequence of the 33-kD protein induced in *E. coli* BoxA-VP1 cells by sequence alignment against known sequences in protein databases ............................................................................................. 78

5. Results of sequence homology search for the N-terminal sequence of the 33-kD protein induced in *E. coli* BoxA-VP1 cells .............................................. 79

6. Bacterial strains, phage and plasmid vectors for surface expression of HAV capsid proteins in *E. coli* ............................................................................ 96

7. Bacterial strains and plasmid vectors for phage T7 RNA polymerase dependent expression of HAV capsid proteins ................................................. 97

8. A comparison of codon usage for selected amino acids in HAV VP genes... 142

9. A comparison of preferred codon usage frequency in *E. coli* for select amino acids encoded in the capsid protein genes of HAV ............................. 143
LIST OF FIGURES

1. The proposed genomic organization of hepatitis A virus based on the L434 nomenclature ................................................................. 11


3. A schematic diagram of the prokaryotic expression vector, pAX4b+, used for cloning and expression of proteins as β-galactosidase fusions............ 24

4. Oligonucleotide primer sequences for PCR amplification of insert HAV capsid DNA for forced directional cloning into the pAX4b+ vector........ 26

5. Agarose gel analysis of DNA segments amplified from pGEM-HAV by PCR and encoding HAV capsid proteins........................................ 41

6. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP1I1........................................................................... 43

7. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP2B2................................................................. 45

8. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP3T3...................................................................... 46

9. Western blot analysis of fusion proteins expressed in E.coli UT2300 cells from recombinant plasmids pAX4b-VP1I1, -VP2B2, and -VP3T3 ...... 47

10. Phase contrast microscopy of E.coli cells induced for expression of fusion proteins from recombinant plasmids encoding sequences for HAV capsid proteins ........................................................................ 49

11. Electron microscopy of E. coli cells induced for expression of recombinant fusion protein from recombinant plasmid pAX4b-VP1I1.............. 50

12. SDS-PAGE analysis of samples from various solubilization steps during the large scale extraction of the VP1I1 recombinant fusion protein French press-sarkosyl extraction procedure................................. 51
13. SDS-PAGE analysis of affinity-purified fusion proteins expressed from recombinant plasmids, pAX4b-VP1, -VP2, -VP3 in *E. coli* UT2300 cells... 52

14. Transcriptional antiterminator sequences from *phage* and *E. coli* genomes... 63

15. Oligonucleotide primer sequences used for PCR amplification of insert HAVBoxA VP1 capsid DNA ........................................................................ 66

16. Buffer matrix of the Opti-Prime™ PCR Optimization kit.......................... 69

17. Agarose gel analysis of HAV VP1 insert DNA amplified by PCR......... 74

18. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-BoxAVP1.......................................................... 75

19. Comparison of amino acid sequences in search for homology of the N-terminal sequence of the 33-kD protein expressed in recombinant *E. coli* BoxA-VP1 cells.............................................................................. 77

20. A schematic diagram of plasmid vector pT7-7 used in cloning HAV capsid genes for expression in a non-fusion format................................. 90

21. A schematic diagram of the SurfZAP™ vector and its derivative, pSurfscript™ employed in cloning for surface expression in *E. coli* and chimeric coliphages................................................................. 93

22. Oligonucleotide primer sequences for (A) PCR amplification of insert HAV VP1 gene for cloning into the SurfZAP™ vector, and (B) PCR based *in vitro* excision of phagemid SurfVP1.......................................................... 98

23. Oligonucleotide primer sequences for PCR amplification of insert HAV capsid genes for cloning into pT7-7 vector............................................. 99

24. A schematic diagram representing the design of insert DNA for cloning into the SurfZAP™ vector................................................................. 100

25. A schematic representation of the essential steps in the *in vitro*-Excision PCR procedure...................................................................................... 118

26. Agarose gel analysis of insert Surf VP1 DNA amplified by PCR for cloning into the SurfZAP™ vector................................................................. 125

viii
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Agarose gel analysis of restriction enzyme digests of recombinant pT7-derived plasmids carrying HAV genes</td>
<td>127</td>
</tr>
<tr>
<td>28</td>
<td>SDS-PAGE analysis for recombinant HAV VP3 expression from pT7-VP3 carrying <em>E. coli</em> BL-21 cells</td>
<td>129</td>
</tr>
<tr>
<td>29</td>
<td>Agarose gel analysis of PCR products to verify the presence of insert SurfVPl DNA during screening of recombinant λSurfVPl plaques</td>
<td>132</td>
</tr>
<tr>
<td>30</td>
<td>Agarose gel analysis of plasmid pSurfVPl DNA extracted from recombinant <em>E. coli</em> SOLR cells by the alkaline lysis method</td>
<td>134</td>
</tr>
<tr>
<td>31</td>
<td>Agarose gel analysis of plasmid pSurfVPl DNA extracted from recombinant <em>E. coli</em> SOLR cells by the PERFECTprep™ method</td>
<td>135</td>
</tr>
<tr>
<td>32</td>
<td>Agarose gel analysis of pSurfVPl DNA amplified by <em>in vitro</em>-Excision PCR</td>
<td>137</td>
</tr>
<tr>
<td>33</td>
<td>Agarose gel analysis of PCR products to verify the presence of insert SurfVPl DNA in pSurfVPl following the <em>in vitro</em>-Excision PCR</td>
<td>139</td>
</tr>
<tr>
<td>34</td>
<td>Agarose gel analysis of plasmid pSurfVPl DNA extracted from recombinant <em>E. coli</em> SurfVPl clones # 2 and # 26</td>
<td>141</td>
</tr>
</tbody>
</table>
ABSTRACT

Hepatitis A virus (HAV), the causative agent of food and water borne acute hepatitis worldwide, is very difficult and expensive to propagate in tissue culture. Large scale production of HAV proteins in E.coli by recombinant DNA techniques would be invaluable for antisera production, biochemical analyses and in developing sensitive diagnostic tests. Therefore, PCR amplified cDNA fragments encoding the HAV capsid proteins were cloned individually into prokaryotic expression vectors for production of either recombinant fusion or non-fusion proteins. Cloning and expression of HAV capsid genes, which induced toxic effects in certain host E.coli, was tolerated better by protease-deficient host cells. HAV capsid proteins expressed from the prokaryotic vector, pAX4b+, as β-galactosidase fusions reacted with antisera generated against intact HAV particles. However, fusion protein expression and subsequent affinity-based purification were complicated by co-expression of β-galactosidase and its cross-reactivity with anti-HAV antibodies. An attempt was made to improve transcription by incorporating the E.coli rnrB BoxA sequence in frame between the lacZ and HAV VP1 gene segments in the fusion construct. Induction of protein expression from recombinant pAX-BoxAVP1 failed to yield HAV related protein but resulted in an overexpression of a 33 kD protein with a strong homology to a hypothetical protein sequence derived from the origin of replication of F and other related E.coli plasmids. HAV capsid sequences were cloned successfully into the pT7-7 vector for expression of proteins in a non-fusion form, and into the λSurfZAP™ vector for expression of a membrane-directed, surface-displayed fusion protein in E.coli and for the generation of chimeric coliphages.
However, expression of a ~30kD protein following induction was seen in recombinants carrying pT7-VP3 plasmids. Expression of recombinant VP1 was not discernible in cells carrying either pT7-VP1 or pSurf-VP1. A study of the codon usage profiles revealed that the coding strategy in HAV might involve a higher frequency of codons recognized as rare in both *E.coli* and human systems, and may be responsible for the inefficient expression of HAV proteins in both *E.coli* and tissue culture.
LITERATURE REVIEW

Hepatitis is defined as inflammation of the liver caused by a variety of noxious stimuli such as chemicals and drugs as well as bacterial, viral and parasitic agents. The extent of injury suffered by the hepatic tissue can vary from minor cell damage that is detectable only by biochemical tests for liver function, to fulminant disease with massive necrosis followed by coma and death.

Viral hepatitis is a term used to refer to an episode of hepatitis caused by a hepatotrophic virus or a virus whose primary target is the liver. However, inflammatory conditions of the liver that are part of a more generalized infection with viruses such as yellow fever, varicella-zoster, herpes simplex, cytomegal-, coxsackie, Epstein-Barr, and rubella viruses, are not included under this definition. Viral hepatitis can occur either as an acute or as a chronic form of liver disease. At the present time, there are at least five different forms of acute viral hepatitis: hepatitis A, B, C, D and E. The viral agents that cause hepatitis A, B and D have been well characterized while those that cause hepatitis C and E are not well characterized yet (Ross, et al., 1991)

Hepatitis A is an acute, contagious, necroinflammatory disease of the liver caused by an infectious virus, the Hepatitis A virus (HAV) (Koff, et al 1992). It is one of the oldest diseases known to humankind and descriptions of this disease can be traced as far back as the fifth century BC (Coulepis, et al., 1987). It was noted by Hippocrates as a disease that was accompanied by fever and jaundice, occasionally occurring in epidemic form and probably due to alterations in the bile (Zuckerman, 1970). The earliest mention of the contagious nature of the disease is found in the correspondence between the
Archbishop of Mainz and Pope Zacharias in 751 AD. Since epidemic hepatitis is particularly common during military campaigns, it was earlier called "jaunesse des camps" by the French and "Kriegsickterus" by the Germans (Gust and Feinstone, 1988). The term "Infectious Hepatitis" was coined by Cockanye to emphasize the contagious nature of the jaundice. Outbreaks of Hepatitis A during wars around the world have been documented beginning in the seventh century, including the American Civil War, World War I & II, Israeli War of Independence, the Korean War, the war in Vietnam and more recently, during the Bosnian Civil War and the various civil wars in Africa and Asia. The term "Hepatitis A" was introduced by MacCallum in 1947 to distinguish infectious hepatitis caused by HAV from a second form of hepatitis that was first reported in 1885 to be transmitted through parenteral route following inoculation with unsterile syringes and transfusion of blood or blood products. The second form of hepatitis was shown subsequently to be different from infectious hepatitis in its incubation period and the etiologic agent and was referred to as Hepatitis B or "Serum Hepatitis" (Coulepis, et al., 1987).

Infection with HAV occurs throughout most parts of the world and is especially common in countries with a very high density of population such as China, India, Bangladesh and Indonesia. In these nations, it is most often associated with overcrowding, poor standards of sanitation and personal hygiene. The incidence of Hepatitis A is comparatively low in the developed countries such as the United States (8/100,000 population) and cases occur sporadically. Infection in the developed countries is generally confined to certain groups-at-risk, such as children and staff of day-care
centers, overseas travelers, inmates of institutions, homosexual men, and intravenous drug users. Still, the annual cost to the US economy due to Hepatitis A has been estimated at $300 million (Gust and Feinstone, 1988).

Hepatitis A occurs in epidemic waves with peaks during late autumn and early winter in many countries with temperate climates. However, in recent years, it has been observed to occur in an endemic form in the developed countries probably due to the decreased incidence of infection (Shapiro and Margolis, 1993). In developing countries, most of the population affected are children in whom the symptoms of the disease tend to be rather mild and provide the individual affected with a life-long immunity against the disease. Hence, this disease is not of major public health concern. In the developed countries, on the other hand, the risk of getting infected in early childhood is reduced due to improved living standards, resulting in an increased mean age of infection (Gust, et al., 1978). However, in the absence of childhood exposure, there appears to be an equal chance of acquiring the infection, regardless of age (Gust, 1984).

The severity of HAV related liver disease can vary widely between individuals and factors such as age, sex, physical condition and pre-existing illnesses can exert considerable affect on the clinical course of the disease. Age, however, appears to be the most significant factor influencing the outcome of HAV infection. A majority of adults, when infected with HAV, tend to develop overt clinical symptoms unlike children in whom the disease is most often asymptomatic. Symptomatic disease occurs among <10% of children younger than 6 years of age, 40-50% of older children, and 70-80% of adults (Alter and Mast, 1994). HAV infections can present as a subtle subclinical infection detectable only
by serological tests or they manifest in several different ways ranging from a relatively mild "flu-like" syndrome to the more serious fulminating hepatitis that can lead to coma and death. Case-fatality rates from Hepatitis A is highest (17.5 per 1000) amongst persons > 49 years of age, although the overall case-fatality rate is very low (4 per 1000). At least 100 deaths each year in the United States are attributed to fulminating hepatitis A (Alter and Mast, 1994).

Hepatitis A infection exhibits an average incubation period of 30 days with a range of 15 to 45 days. The onset of clinically overt disease is often preceded by a brief prodromal period or pre-icteric phase, lasting less than 1 week, during which patients may complain of loss of appetite, nausea and abdominal discomfort (Coulepis, et al., 1987). The prodromal phase is characterized by an abrupt onset of a "flu-like" state marked by the major symptoms like fever, chills, headache, fatigue, malaise, myalgia, anorexia, nausea and vomiting. However, Hepatitis A in children may present with atypical symptoms such as diarrhoea, cough and coryza. The acute or icteric phase which succeeds the prodromal phase begins with the appearance of dark urine (bilirubinuria) followed by a paler than usual feces (steatorrhoea) and an yellow discoloration of the skin and sclera (jaundice). An enlarged, tender liver may be palpable in the right upper quadrant of the abdomen during this stage and persist for two to three weeks. Liver function tests during this phase usually reveal an elevation in the levels of bilirubin (> 3 mg%) and the hepatic enzymes, alanine aminotransferase (ALT), aspartate aminotransferase (AST) and alkaline phosphatase, which suggests injury to liver tissue. The lesions in the hepatic tissue caused during HAV infection are, however, indistinguishable from those owing to other viral causes. Liver
biopsies usually reveal a centrilobular cholestasis accompanied by hepatocyte necrosis and portal inflammation with mononuclear cell infiltration (Gordon, et al., 1984).

Hepatitis A is normally a self-limiting disease which lasts for 3-4 weeks. The virus is shed in the feces of infected adults late in the incubation period and before onset of clinical illness. Children and infants may, however, excrete HAV for a longer period than in adults, possibly for several weeks following the onset of clinical illness. Viral RNA has been detected by the polymerase chain reaction (PCR) in fecal specimens for up to 6 months after infection in infants born prematurely. Chronic fecal shedding of HAV, however, has not been observed in any patient group. The clinical symptoms and signs of steatorrhoea, jaundice and hepatomegaly begin to subside during the second and third week of the acute phase. Liver function tests return to normal at about the same time, which signals recovery of hepatic function. The liver architecture will appear normal without scar formation in most patients and there are no known chronic sequelae. Occasionally, the disease might persist for up to 18 weeks as cholestatic hepatitis with high bilirubin levels along with the other major symptoms and signs (Gordon, et al., 1984). Although the disease is not known to exist in a chronic form, relapses do occur sometimes. Also, in 0.4% of all hospitalized patients, the disease may progress to the rare fulminant hepatitis with a case-fatality rate >50% despite aggressive therapy which includes liver transplantation. Fulminant hepatitis is accompanied by an increasingly deep jaundice, altered sensorium, fetor hepaticus, asterixis or flapping tremors, followed by encephalopathy and coma. A poor prognosis is indicated by prolongation of the prothrombin time, steeply rising serum bilirubin levels, and occasionally a decrease in the
levels of serum aminotransferases, all of which suggest an extensive liver damage. Death usually ensues within 14 days (Ross, et al., 1991).

Human volunteer studies done during World War II to determine etiology and the mode of transmission of hepatitis A showed that the disease could be acquired by healthy individuals following oral administration of a bacteria-free suspension of feces and serum from patients in their pre-icteric phase. These results implied that the disease was probably caused by a virus and that transmission from person to person was by the fecal-oral route. Experimentally infected individuals also failed to develop symptoms of the disease upon reinoculation, 6 to 9 months later, suggesting development of immunity to hepatitis A. Furthermore, convalescent serum from previously infected individuals when introduced into susceptible persons, afforded protection against HAV infection. Studies with animal models demonstrated that the disease is confined exclusively to humans and can be mimicked only in certain non-human primates such as marmosets and chimpanzees (Maynard, et al., 1975; Dienstag, et al., 1975). Further investigations in marmosets revealed that the virus enters the gut and via the bile duct, gains access to liver which is the major site of viral replication as shown by immunofluorescence (Mathieson, et al., 1980) and electron microscopy (Shimizu, et al., 1978). Studies performed to establish HAV pathogenesis failed to provide any direct evidence linking viral replication and hepatocellular damage. Instead, the fact that histopathological signs of liver damage occurred coincident with elevation in the levels of circulating antibody, suggested that the damage may be immunologically mediated. Later studies involving HAV-infected tissue culture cells (BS-C-1) have provided direct evidence that HAV-specific T-lymphocyte
mediated cytotoxicity and natural killer cell activity are more likely involved in hepatocellular injury (Kurane, et al., 1985).

HAV transmission occurs following ingestion of virus particles shed in the feces of an infected person. Feces of infected individuals during the incubation period can contain up to $10^8$ virions per ml and serves as the primary source of HAV. The virus is spread mainly by close personal contact, especially between young children in day-care centers. Among adults, HAV can also be acquired through sexual contact. Contaminated water and food, particularly shellfish, also serve as vehicles of HAV transmission. Oysters, which are usually eaten raw, from sewage-contaminated waters can concentrate enteroviruses up to 60-fold and are frequently implicated in outbreaks of hepatitis A (Rippey, 1994). Since HAV can survive for 12 weeks to 10 months in water and since gentle steaming alone is insufficient to inactivate HAV, outbreaks of shellfish-associated hepatitis A can potentially lead to large epidemics (Coulepis, et al., 1987). One such outbreak in Shanghai, China in 1988 caused more than 200,000 cases (Communicable Disease Intelligence, 1988). HAV transmission can also be water-borne and infections may be acquired by drinking fecally contaminated water or by swimming in contaminated pools and lakes. HAV transmission through parenteral route, although possible, is very rare.

Hepatitis A virus is a pathogenic, hepatotropic virus and belongs to the family Picornaviridae. It is one of nearly 200 human-adapted picornaviruses but the only one that has been documented as a cause of seafood-associated illness (Ahmed, 1991). Only one serotype of HAV has been observed among HAV isolates collected from various parts of
the world. HAV was first identified as virus-like particles of 27-nm diameter by immunoelectron microscopy (IEM) using fecal samples from infected experimentally human volunteers as source of antigen and convalescent sera from patients with either experimentally induced or naturally acquired hepatitis A infection, as a source of anti-HAV antibody (Feinstone, et al., 1973). Subsequent studies revealed that HAV is a non-enveloped virus consisting of a single-stranded RNA genome enclosed within an icosahedral protein shell or capsid. The majority of the virus particles have a buoyant density of 1.32-1.34 g/cm$^2$ when determined by density gradient equilibrium centrifugation in cesium chloride (Coulepis, et al., 1982). HAV, like other picornaviruses, exhibits remarkable stability against lipid solvents such as ether, acidic conditions such as pH 3.0 and prolonged exposure to heat (Rueckert, 1991). The virus, however, is readily inactivated by formaldehyde (Provost and Hilleman, 1978). Based on the biophysical and biochemical features, HAV was classified previously as a member of the genus Enterovirus within the Picornaviridae family (Gust, et al., 1983). However, HAV differs significantly from the other members of the genus Enterovirus with regard to stability at elevated temperatures and low pH (Flehmig, et al., 1985; Scholz, et al., 1989), mechanism of replication, genome organization, and behaviour in cell cultures (Ticehurst, et al., 1989). As a consequence, it has been proposed that HAV be classified under a separate genus Hepatovirus within the family Picornaviridae (Minor, 1991)

The capsid of HAV is made of at least three structural proteins VP1, VP2, and VP3 and a putative VP4 with inferred molecular weights 33,200 (VP1), 24,800 (VP2), 27,800 (VP3), and 2500 (VP4) based on nucleotide sequence data. However, analysis of the viral
structural proteins by sodium dodecyl sulfate-polyacrylamide gel electrophoresis have shown protein bands of apparent molecular weights 34,500 (VP1), 25,500 (VP2) and 23,000 (VP3) (Coulepis, et al., 1978). Previously published reports suggest that the migration of the HAV VP2 and VP3 capsid proteins upon gel electrophoresis varies depending on the individual gel systems and the presence or absence of urea (Linemeyer, et al., 1985; Gauss-Muller, et al., 1986; Wheeler, et al., 1986). The putative VP4 protein has not been characterized as well as the other HAV capsid proteins, and conflicting reports exist in the literature with respect to the molecular weight of VP4 and whether it is myristylated or not (Coulepis, et al., 1980; Palmenberg, 1989). However, like other picornaviruses, HAV particles also appear to contain a few copies of VP0 which is the precursor of VP2 and VP4 proteins (Gauss-Muller, et al., 1986).

The genome of HAV is a linear, positive-sense, single-stranded RNA molecule of approximately 7500 nucleotides (nt) or 7.5 kilobases (kb) (Coulepis, et al., 1981; Jameel, 1994). The HAV RNA, although transcribed in an eucaryotic system, does not possess the 5'-N-methylated Guanosine cap. Instead, the 5'-end of the RNA is protected by a covalent linkage to a tyrosine residue of one of the HAV-derived proteins, the Vpg terminal peptide. Like other eucaryotic RNAs, however, the HAV RNA does possess a poly(A) tract towards the 3'-end. The properties of the HAV genome were elucidated following the observations that the hepatitis A genomic material was located exclusively in the cytoplasm of infected hepatocytes, stained with Acridine Orange, sedimented at 33S with a buoyant density of 1.64 g/cm^2 and was sensitive to RNase but not DNase (Ross, et al., 1991). The naked RNA genome, like the positive-sense genomic RNA of other picornaviruses, has been shown to
be intrinsically infectious and readily translatable into HAV proteins once access to the cytoplasm of the host cell is gained (Anderson, et al., 1987; Cohen, et al., 1987b). Since there is no RNA-dependent RNA polymerase with the HAV genome in the intact virus particle, translation of the viral RNA within the cytoplasm of the host cell is the first step in the viral replicative cycle and hence the genome is considered intrinsically infectious.

According to the well-established L434 nomenclature for describing the genomic organization of picornaviruses (Rueckert and Wimmer, 1984), the HAV genome can be considered as functionally divided into a 5' untranslated region (UTR) of ~733 nt, followed by a single open reading frame (ORF) of ~6,681 nt and a 3' UTR of ~64 nt as shown in Figure 1. The single ORF, which encodes a polyprotein precursor of 2227 amino acids from which all HAV proteins are derived following post-translational processing, can be subdivided into three regions: P1, P2 and P3. Proteins derived from the P1 region (1A, 1B, 1C and 1D) constitute the structural capsid proteins VP1, VP2, VP3 and VP4. The P2 region accounts for a viral protease (2A) which cleaves the huge viral polyprotein precursor between the P1 and P2 regions, a second protein (2B) of unknown function, and a third protein that is believed to anchor the 5'-terminal VPg peptide to the viral genome. The P3 region encodes the precursor of VPg (3A), a second protein whose function remains unknown, a major viral protease (3C<sup>pro</sup>) involved in proteolytic processing of the polyprotein and viral RNA replication and the viral RNA-dependent RNA polymerase (3D<sup>pol</sup>). The non-structural P2 and P3 regions are actively involved in and are essential for the replication and production of infectious HAV particles.
Figure 1. The proposed genomic organization of hepatitis A virus based on the L434 nomenclature. The figure also schematically depicts the post-translational events contributing to the polyprotein-processing during poliovirus replication. Cleavage between P1 (1ABCD) and P2 (2ABC) is mediated by protease 2A and at all other sites except for 1A/1B, which is autocatalytic, by protease 3C.
Events in the replicative cycle of HAV follow the scheme postulated for the Picornaviridae virus family which is depicted in figures 2a, 2b and 2c. Following ingestion of contaminated food or water by a susceptible individual, HAV is believed to attach to an as-yet-uncharacterized cell-surface receptors in the gastrointestinal tract and gain access to the liver via the portal circulation. The HAV particles attach to target hepatocytes and penetrate the cell membrane by viropexis (a virus specific receptor-mediated-endocytosis) and lose their protein shell, an event termed uncoating. During uncoating, viral RNA is released into the cytoplasm over a period of approximately 12 hours mainly due to the high affinity of the viral RNA to the capsid proteins. Following the release, viral RNA serves as a messenger RNA for synthesis and translation of the large HAV ORF. The polyprotein is processed by viral proteases such as 2A and 3C which are themselves derived from the polyprotein. The polyprotein processing begins before translation of the ORF is fully completed and results in a group of structural capsid proteins (P1) and a second group of non-structural proteins (P2 and P3). The non-structural proteins participate in the transcription of the viral RNA into negative-strand templates from which more copies of positive-strand viral genomic RNA are generated. Since the viral genomic RNA possesses a high affinity for capsid proteins, encapsidation of the majority of the positive-strand copies occurs almost coincidental with the RNA synthesis, resulting in mature infectious virions. Thus, in contrast to other picornaviruses, particularly poliovirus, a much higher proportion of progeny RNA is seen to be packaged into mature virus particles throughout the replicative cycle of HAV (Anderson, et al., 1988).
Figure 2A. A schematic depiction of the events during the hepatitis A virus infection cycle. Virus particles gain access to the hepatocytes via the portal circulation following entry into the cells lining the gastrointestinal tract by a mechanism not completely elucidated. The virus particles then attach to the liver cells and enter the cytoplasm by viropexis. The delivery and uncoating processes also are not fully understood at the present time. Following uncoating, the viral RNA is released into the host cytoplasm whereupon the host protein synthesizing machinery is recruited to translate the ORF encoding the viral polyprotein that is subsequently processed into the precursor P1, P2 and P3 segments.
Figure 2B. Events during the hepatitis A virus infection cycle. Post-translational events in the processing of the polyprotein derived P1, P2 and P3 segments. The P1 segment is processed to yield the capsid proteins VP0, VP1 and VP3 which initially form an immature protomer (5S) consisting of one copy of each. The protomers then assemble into pentamers (12-14S). The P2 and P3 segments are processed to yield the non-structural proteins essential for the viral RNA synthesis via a complementary minus-stranded template.
Figure 2C. Events during hepatitis A virus infection cycle. A smooth endoplasmic reticulum-associated synthesis of (+) stranded genomic RNA copies through multi-stranded replicative intermediate. The replication of both (-) and (+) strands is carried out by the viral RNA-dependent-RNA polymerase (3Dpol) in steps that require the other non-structural proteins as well. The (+) RNA pool is used for translation, synthesis of more (-) RNA and for packaging into progeny virions. The packaged infectious virions (150-160S) then undergo a "maturation cleavage", an autoproteolytic event, in which the VP0 peptide is cleaved into VP2 and VP4. The progeny, mature virus particles are released following lysis of the host cell.
The specific diagnosis of naturally acquired acute HAV infection involves the detection of IgM antibodies to HAV during the acute or early convalescent phase of the infection. The IgM anti-HAV antibodies are present in the serum of patients by 5-10 days into the incubation period and at the onset of clinical illness. Antibody titer rises rapidly reaching a peak before onset of jaundice 1-2 weeks following onset of the initial symptoms and then declines gradually to non-detectable levels. IgG anti-HAV antibodies begin to appear following the peak IgM response and stabilize between 3-8 months. The IgM antibodies remain detectable for up to 6 months following the onset of symptoms while the IgG remain detectable for the lifetime of the individual thereby indicating a life-long protection against the disease (Kao, et al., 1984).
CHAPTER I

CLONING AND EXPRESSION OF THE CAPSID PROTEINS OF HEPATITIS A VIRUS IN *Escherichia coli* AS β-GALACTOSIDASE FUSION PRODUCTS
INTRODUCTION

Although clinical diagnosis of viral hepatitis is based on the recognition of the classical signs and symptoms compatible with acute inflammation of the liver, the identification of the precise etiological agent responsible mandates laboratory confirmation through specific tests. The diagnosis of acute hepatitis A is confirmed by serological detection of IgM anti-HAV antibodies in the serum of patients. IgM anti-HAV levels become detectable with onset of clinical symptoms, reach peak levels within one to two weeks, and remain elevated for up to four months (Liaw, et al., 1986). Later, as IgM antibody titers gradually decline, IgG anti-HAV antibodies appear in the serum and may persist for years following the acute infection. The IgG anti-HAV confer long-term immunity against HAV and is a serological indicator for previous exposure to this viral agent. Secretory immunity, mediated by IgA antibodies, which plays a significant role in protecting against a number of enteroviral infections at the site of primary entry of the viruses, however, does not appear to be effective against HAV (Stapleton, et al. 1991).

Serological detection of anti-HAV antibodies in the serum of victims can be accomplished by employing one of several approaches available such as complement fixation (Provost, et al., 1975), immunoelectron microscopy (Feinstone, et al., 1973; Feinstone, et al., 1974), radioimmunoassay (Blaine Hollinger, et al., 1975), and enzyme immunoassays (Eble, et al., 1991; Fayol and Ville, 1991; Locarnini, et al., 1978; Robbins, et al., 1991; Wang, et al., 1985). The commercial immunoassay kits currently in use facilitate detection of either total anti-HAV, a mixture of IgG, IgM and IgA, or
the acute phase IgM anti-HAV exclusively (Decker, et al., 1981; Safford, et al., 1980). Unfortunately, tests which detect HAV particles or antigens in clinical specimens directly are not commercially available. Those immunoassays which detect anti-HAV antibodies are predicated on the specific reactivity of HAV-specific antibodies with relevant antigens, that is, the HAV capsid proteins. Hence, most immunoassays require HAV particles, raised in tissue culture, as the antigen source. Propagation of HAV in tissue culture is expensive, requires containment facilities and is compounded by factors such as variation in the infectivity properties of individual virus isolates. Additional difficulties associated with cultivation of HAV are its poor and protracted growth in tissue culture (Cohen, et al., 1989; Wang, et al., 1985; Winokur, et al., 1991), frequently resulting in cell-line infections without discernible cytopathic effects (Bitton, et al., 1986; Fayol and Ville, 1991). These obstacles are further complicated by the necessity to rid the virus particles of host cell debris which accumulates during viral harvest and purification steps. Therefore, replacement of tissue culture generated virions required in immunoassays with an \textit{in vivo} or \textit{in vitro} generated, recombinant subunits of the viral capsid as antigens, holds promise as a simple, time-saving and inexpensive alternative. Recent advances in recombinant DNA technology coupled with availability of the complete nucleotide sequence of HAV genome for analysis, have provided a means by which the individual capsid proteins can be produced synthetically or be expressed in either prokaryotic or eukaryotic systems.

Several putative, dominant epitopes in HAV capsid proteins have been identified by various investigators for constructing recombinant DNA and synthetic peptides, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
their potential use in diagnostic assays and in vaccine development (Baroudy, et al., 1985; Emini, et al., 1985b; Gauss-Muller, et al., 1990; Hughes, et al., 1984; Hughes, e, 1985; Stapleton and Lemon, 1987). Mapping studies which identified these epitopes were done using monoclonal antibodies against HAV (Hughes, et al., 1984; McGregor, et al., 1983; Stapleton and Lemon, 1987) and the results suggests that capsid protein VP1 contains a major neutralization epitope. Reports from other studies (Ping, et al., 1988) suggested that capsid proteins VP3 and VP2 may also contribute amino acid residues to the neutralization epitope within VP1 or may contain neutralization epitope(s) of their own. Thus, while consensus has emerged that the capsid peptide VP1 possesses the dominant conformational epitope against which a neutralizing immune response is directed, the contribution of the VP2 and VP3 peptides to the conformation of the major neutralization epitope(s) on the capsid remains unresolved. Further immunological studies are required to clarify the immunochemistry of these capsid proteins before they are used as a efficient vaccine to confer a long-term protection against HAV infection or as an antigen in the development of inexpensive immunodiagnostic assays.

Molecular biological techniques have enabled investigators to characterize the single-stranded RNA genomes of various picornaviruses including HAV, and to identify those nucleotide sequences which correspond to the various HAV structural and non-structural proteins (Baroudy, et al., 1985a; Cohen, et al., 1989; Hughes, et al., 1984; Malcolm, et al., 1992; Ross, et al., 1991; Rueckert, 1991; Ticehurst, et al., 1983). Complementary DNA (cDNA) sequence corresponding to the HAV RNA genome has been cloned successfully (Baroudy, et al., 1985b), facilitating production of recombinant
HAV proteins in prokaryotic or eukaryotic systems, and has made studies directed towards a systematic and precise elucidation of the nucleotide sequences which encode immunodominant epitopes, more promising.

Several studies concerning the recombinant expression of HAV capsid proteins and their antigenicity have been reported (Emini, et al., 1985a; Ostermayr, et al., 1987; Gauss-Muller, et al., 1990; Stapleton, et al., 1991; Winokur, et al., 1991). Cloning and expression in *Escherichia coli* have resulted in the production of HAV fusion peptides containing limited portions of the different capsid proteins, especially VP1 and VP3 and *E.coli* proteins encoded by the *lacZ* (Ostermayr, et al., 1987) or *TrpE* gene (Johnston, et al., 1988). These recombinant fusion proteins, representing only portions of *E.coli* and HAV capsid proteins, were poorly soluble and failed, to induce any significant anti-HAV neutralizing antibody response in laboratory animals upon immunization (Emini, et al., 1985b; Johnston, et al., 1988) and to react with rabbit anti-HAV polyvalent sera but not with human convalescent sera (Gauss-Muller, et al., 1990; Ostermayr, et al., 1987). These reports suggest that the immunogenicity of recombinant HAV peptides can vary among species and attempts to improve the immunogenicity are a prerequisite towards successful development of a synthetic, recombinant vaccine. Also, synthesis of complete, not just portions of capsid proteins, or even "empty" capsids may be essential and crucial to achieve this goal. In this regard, eukaryotic expression in recombinant baculoviral (Stapleton, et al., 1991) and in Vaccinia viral (Winokur, et al., 1991) systems of whole or portions of HAV capsid has been achieved although the methods employed still involved time-consuming and expensive tissue culture procedures and
resulted in unpredictable yields. However, there are no reports thus far on any attempts to express any of the individual HAV capsid proteins in their entirety, either in the fusion or in the non-fusion form, in prokaryotic systems.

One of the main objectives of our investigations was to study the potential for cloning and expression of entire HAV capsid proteins in *E. coli*, either as a fusion or in the non-fusion format. The recombinant capsid proteins could be used subsequently in studies directed to assess their antigenicity and utility in immunoassays, in the production of diagnostic antibody reagents from laboratory animals, and possibly, as a potential recombinant vaccine. We report the individual cloning of the complete nucleotide sequences for VP1, VP2 and VP3 capsid protein genes of HAV, their subsequent expression as β-galactosidase fusion proteins in *Escherichia coli* (or *E. coli*), and the purification of the resulting recombinant products. The antigenicity of the β-galactosidase-HAV fusion proteins was assessed by ELISA and Western Blot techniques using polyclonal and monoclonal antibodies raised against tissue culture grown, intact HAV particles.

**MATERIALS AND METHODS**

**Bacterial Strains and Plasmids:** Bacterial strains and plasmids used in this study are listed in Table 1. Prokaryotic expression vector pAX4b is schematically represented in Figure 3.

**Anti-HAV Polyvalent and Monoclonal Antibodies:** Rabbit anti-HAV polyvalent serum and Mouse anti-HAV monoclonal antibodies (Mab) used in this study were generated previously in a separate study in Dr. Siebeling's lab by Dr. Baez. These
Table 1. Bacterial strains and plasmid vectors used in cloning for expression of HAV capsid genes as β-galactosidase fusions

<table>
<thead>
<tr>
<th>Strains/Plasmids</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em> GW1000</td>
<td>xyl-5, mtl-1, galK2, λ-rac, rpsL31, kdgk51, Δ(gpt-proA)62, lacY1, tsx-33, supE44, thi-1, leuB6, hisG4, mgl-51, arg-3, rfbD1, ara-14, thr-1, recA441, sulA, ivl, ΔlacU169, pro^+</td>
<td>G. Walker/ J.R. Battista (L.S.U.)</td>
</tr>
<tr>
<td><em>Escherichia coli</em> UT2300</td>
<td>azi-6, fhuA23, lacY1, leu-6, mlt-1, proC14, purE42, rpsL109, thi-1, trpE38, tsx-67, entA403, purE^+, fhuA^+, fepA, Δlon::Tn10</td>
<td>Elish, et al., 1988</td>
</tr>
<tr>
<td><em>Escherichia coli</em> Top10</td>
<td>mcrA, Δ(mrr-hsdRMS-mcrBC),Φ80ΔlacΔM15,ΔlacX74, deoR, recA1, araD139,Δ(ara,leu)7697 galU, galK, λ^- rpsL, endA1, nupG</td>
<td>Grant, et al., 1990</td>
</tr>
<tr>
<td>Plasmid pAX4b+</td>
<td></td>
<td>United States Biochemical Corporation, Cleveland, OH</td>
</tr>
<tr>
<td>Plasmid pGEM-HAV</td>
<td></td>
<td>R.H. Purcell (N.I.H)</td>
</tr>
</tbody>
</table>
Figure 3. A schematic diagram of the prokaryotic expression vector, pAX4b+, used for cloning and expression of proteins as β-galactosidase fusions. Plac = lacZ gene promoter; lacZ = Complete gene for β-galactosidase; CS = Collagen fragment which also contains endoproteinase Xa cleavage site; MCS = Multiple Cloning Site; t0 = Lambda t o terminator; f1 ori = Phage f1 origin of replication; and bla: Ampicillin resistance gene.
antibodies had been purified and characterized for specificity against HAV by various immunoassays using HAV HM-175 strain. MAb P2G9 and P1E11 were IgG isotypes while Mab P6B5 and P2G11 were IgM isotypes. P2G9 was also a neutralizing anti-HAV antibody as characterized in the earlier study (Baez, 1993).

**Amplification of Nucleotide Sequences Encoding HAV Capsid Proteins:** The plasmid pGEM-HAV (kindly provided by Dr. Robert H. Purcell, NIH) was used as a template to amplify individually the genes encoding the three major capsid proteins of the Hepatitis A virus. Oligonucleotide primers containing flanking restriction enzyme recognition sequences were designed to facilitate the directional cloning of the amplified viral genomic fragments and were synthesized by beta cyanoethyl phosphoramidite chemistry (Appligene Inc., Pleasanton, CA.). These oligonucleotide primer sequences are shown in Figure 4. The DNA regions encoding individual capsid protein genes were generated by PCR amplification of the relevant regions on pGEM-HAV. PCR reactions were set up in a final volume of 100μl containing 200μM concentrations of dATP, dCTP, cGTP and dTTP; 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 0.01% gelatin, 5 units AmpliTaq polymerase (Perkin-Elmer Cetus, Norwalk, CT), 1.0 μM of each primer, and 500 ng of template DNA. PCR amplification was performed using a Perkin-Elmer Cetus, Norwalk, CT). Thirty cycles of amplification were run under the following conditions: denaturation at 94°C for 90 sec, annealing at 50°C for 90 sec, extension at 72°C for 90 sec, with a cycle extension of 6 sec per cycle and then cooling to 4°C. The amplified products were detected by electrophoresing 10μl of the reaction mixture through a 2% agarose gel in Tris-Acetate-EDTA (TAE) buffer and
Ax-VP1-PS

\[5'-d(CCG\ AAT\ TCT\ GTT\ GGA\ GAT\ GAT\ TCT\ GGA)-3']\]

EcoRI

Ax-VP1-NS

\[5'-d(CCC\ TGC\ AGG\ CTC\ AAA\ TCT\ TTT\ ATC\ TTC)-3']\]

PstI

Ax-VP2-PS

\[5'-d(CCG\ AAT\ TCT\ ATG\ GAC\ ATT\ GAG\ GAA\ GAG)-3']\]

EcoRI

Ax-VP2-NS

\[5'-d(CCA\ GAT\ CTT\ GTG\ TAG\ AAA\ GAG\ GAG\ TTA)-3']\]

BglII

Ax-VP3-PS

\[5'-d(CCG\ AAT\ TCT\ ATG\ AGA\ AAT\ GAA\ TTT\ AGG)-3']\]

EcoRI

Ax-VP3-NS

\[5'-d(CCC\ TGC\ AGT\ TTG\ TGT\ AGT\ AAC\ ATC\ CAT)-3']\]

PstI

Figure 4. Oligonucleotide Primer Sequences for PCR Amplification of Insert HAV Capsid DNA for forced directional cloning into the pAX4B+ vector. (VP) Viral peptide, (PS) positive strand primer and (NS) negative strand primer. The restriction endonuclease recognition sequences (EcoRI, PstI, BglII) within each primer are underlined.
stained with ethidium bromide. The PCR-amplified DNA was extracted by ethanol precipitation and resuspended in sterile distilled water. The DNA concentration was calculated by measuring absorbance at 260 nm (A_{260}) of diluted samples.

Approximately 5 μg of the PCR amplified DNA was digested with 20 units restriction enzyme, EcoRI (GIBCO BRL, Gaithersburg, MD), in a 100 μl restriction enzyme reaction to generate the 5'-cohesive ends required for ligation into plasmid vector pAX4b+ (United States Biochemical Corporation, Cleveland, OH). The digestion reaction mix was incubated at 37°C for a minimum of 8 hours to ensure complete digestion of the DNA and then heated to 65°C for 10 minutes to inactivate EcoRI. The digested VP DNA was resolved on a 2% agarose II (Amresco, Salon, OH.) gel and recovered by using Geneclean II (Bio 101 Inc., La Jolla, CA). The recovered DNA was resuspended in 30 μl of sterile distilled water and then subjected to digestion with 20 units of the second restriction enzyme (PstI for DNA encoding VP1 and VP3, BglII for DNA encoding VP2; GIBCO BRL, Gaithersburg, MD) for at least 8 hours at 37°C to generate 3'-cohesive ends for ligation. The double-digested DNA was purified and recovered from a 2% agarose II gel by using Geneclean II and resuspended in 15 to 20 μl of sterile water. The DNA concentration was determined by measuring A_{260} and samples were stored at -20°C until needed.

Cloning of Digested PCR Products Encoding Viral Capsid Proteins: The procaryotic expression vector pAX4b+ was selected for the cloning the PCR amplified products as lacZ fusions (Markmeyer, et al., 1990). The plasmid was propagated in each Escherichia coli strain in the presence of ampicillin (100 μg/ml) following transformation
and purified by a modification of the alkaline lysis method (Birnboim and Doly., 1979). The plasmid DNA was subjected to appropriate restriction enzyme digestions (EcoRI and PstI or EcoRI and BglII) to generate compatible cohesive ends for ligation of insert VP DNA digested similarly as described earlier. The double -digested plasmid DNA representing mainly the larger vector fragments following the second restriction enzyme digestion was purified by the Geneclean II procedure to eliminate the smaller plasmid fragment from the multiple cloning site that could potentially allow reannealing of the vector in any subsequent ligation reaction. The purified vector DNA was resuspended in 20 μl of TE buffer (10mM Tris-HCl, pH 8.0, 1 mM EDTA) and its concentration estimated following A₂₆₀ measurement of diluted samples.

Prior to setting up ligation reactions, molar ratios for the plasmid vector and the individual VP DNA fragments were calculated using the formula as follows: μg dsDNA = fmoles of dsDNA ends X (1 μg/3000 fmol) X (size of dsDNA in bp/1000 bp) (Focus on applications, Technical Bulletin 5224-1, GIBCO BRL, Gaithersburg, MD). The relative amounts of insert and vector DNA necessary to yield insert:vector molar ratios of 1:1, 2:1, 4:1 and 10:1 in separate ligation reactions were thus determined. Ligation reactions containing 0.5 to 2 units of T4 DNA Ligase (GIBCO BRL, Gaithersburg, MD) in a 5 X DNA ligase reaction buffer (250 mM Tris-HCl, pH 7.6, 50 mM MgCl₂, 5 mM ATP, 5 mM DTT, and 25% (w/v) PEG 8000) in addition to appropriate amounts of vector and insert DNA were set up in final volumes of 15 or 20 microliters. The reactions were incubated overnight at 22°C (Room Temperature, RT) and used subsequently in the transformation experiments or stored at -20°C until needed.
*Escherichia coli* strains GW1000 and UT2300 were chosen as candidates for the transformation and amplification of recombinant plasmids. Transformation of these strains with recombinant DNA was performed by one of the following approaches: (i) The calcium chloride transformation protocol or (ii) The high efficiency electro-transformation of *E. coli* protocol.

To prepare competent *E. coli* cells for transformation by the calcium chloride method, a 500 μl sample from an overnight culture was diluted in 25 ml of fresh LB broth (1% Tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.0) and the cells were incubated at 37°C for 2-3 hours with vigorous shaking. Following the incubation, the culture was centrifuged at 3,000 X g to pellet cells. After carefully discarding the supernatant, the cell pellet was resuspended in 20 ml of ice cold 100 mM CaCl₂ and centrifuged again. After centrifugation, the cells were resuspended in 10 ml of the ice cold 100 mM CaCl₂ and incubated on ice for 45 min to 1 hour. Then, the cells were centrifuged at 3,000 X g for 15 min, resuspended in 2 ml of the ice cold 100 mM CaCl₂ and stored at 4°C overnight before use in a transformation reaction. Transformation reactions were performed by incubating 200 μl of the CaCl₂-treated competent cells with 15 μl of a 1:5 dilution of the ligation mix together on ice for 30 min and heat shocking at 42°C for 2 min. Following the transformation, one ml of SOC medium (2% tryptone, 0.5% yeast extract, 8.5 mM NaCl, 25 mM KCl, 10 mM MgCl₂, 20 mM glucose, pH 7.0) was added and the cells incubated at 37°C for 2-3 hours. Control transformation reactions with *E. coli* host cells and either uncut plasmid vector or single digested plasmid (*EcoRI* or *BglII* or *PstI*) or double digested plasmid (*EcoRI* and *PstI* or
EcoRI and BglII) or double digested and religated plasmid or plasmid pBR322 were set up to monitor the efficiency of transformation. In addition, a transformation reaction with *E. coli* host cells but without any plasmid DNA was set up as a negative control. The transformed cells were plated in triplicates at 200 µl/plate onto SOC agar containing 100 µg/ml ampicillin and incubated at 37°C overnight. The remaining ligation mix was stored at 22°C overnight and plated the next day, if necessary. The SOC plates were screened for colonies and representative colonies were selected and amplified in 10 ml SOC-Amp medium for plasmid isolation and further analysis.

To prepare competent cells for the high efficiency electro-transformation method, 1 liter of LB broth was inoculated with 1/100 volume of an overnight *E. coli* culture. The cells were then grown at 37°C with vigorous shaking to an $A_{600}$ of 0.5 to 1.0 and chilled on ice for 15 to 30 minutes. The culture was then centrifuged in a cold rotor at 4,000 × g for 15 minutes. After carefully removing the supernatant, the cell pellets were resuspended in a total of 1 liter of ice cold sterile distilled water. The resuspended cells were centrifuged again and resuspended in 0.5 liter of cold water. The centrifugation step was repeated again, and the resulting cell pellet was resuspended in ~20 ml of cold 10% glycerol. The glycerol suspension was centrifuged again and the cell pellet resuspended in a final volume of 2-3 ml in cold 10% glycerol. Aliquots of this suspension were used in the electroporation reactions or were frozen on dry ice and stored at -70°C or below until needed. To set up an electroporation reaction, 40 µl of the cell suspension (fresh or gently thawed if from frozen stock), was mixed with 1-2 µl of the ligation reaction mix and left on ice for 1 min. Then, the mixture of cells and
DNA was transferred into a cold, 0.2 cm Gene Pulser electroporation cuvette (Bio-Rad laboratories, Hercules, CA) and subjected to an electric pulse with a time constant of 4-5 ms using the Gene Pulser apparatus (Bio-Rad laboratories, Hercules, CA) set at 25 \( \mu \text{F} \) and 2.5 kV and with the Pulse Controller set at 200\( \Omega \). Following the pulse, 1 ml of SOC medium was added immediately to the mixture in the cuvette and the cells were quickly resuspended with a sterile Pasteur pipette. The resulting cell suspension was transferred into a sterile 17 X 100 polypropylene tube and incubated at 37°C for 1 hour with shaking. The cell suspension was then plated onto SOC-Amp medium as described earlier and incubated overnight at 37°C. The next day, plates were screened for recombinant colonies and representative colonies were chosen for further study.

Plasmid DNA from several selected recombinants was extracted by the alkaline-lysis method and DNA concentration of each sample was estimated by measuring \( A_{600} \). Plasmid DNA samples were linearized by restriction enzyme digestion with \textit{EcoRI} and the relative size of each linearized recombinant plasmid DNA was compared to that of the linearized control plasmid, pAX4b+ without any inserts, by 0.8% agarose gel electrophoresis. A relative shift in the position of the linear plasmid bands corresponding to 900 base pairs for VP1, 660 base pairs for VP2 and 738 base pairs for VP3 strongly suggested that cloning of the viral DNA segments was successful. Further, using the linearized recombinant plasmids as template, PCR reactions were performed with corresponding sets of VP1 or VP2 or VP3 primers. The reaction conditions for PCR were as already mentioned in the PCR amplification of viral DNA from pGEM-HAV. Successful amplification of specific VP regions, as revealed by agarose gel
electrophoretic analyses, confirmed the presence of insert DNA in the recombinant plasmids.

Expression of β-galactosidase-Viral Peptide Fusion Proteins: Three representative recombinant clones: *E.coli* cells carrying either AX-VP1II1 with VP1 DNA or AX-VP2B2 with VP2 DNA or AX-VP3T3 with VP3 DNA were chosen for studying the expression of fusion proteins as follows. Briefly, one ml of overnight cultures of each clone in LB broth containing 100 μg ampicillin/ml (Lbamp) and 0.5% glucose was subcultured into 10 ml of fresh LB-Amp-Glucose medium, incubated for 1 h at 37°C with vigorous shaking, and subcultured again into 10 ml of fresh LB-Amp-Glucose medium and incubated as before. After 1 h, a 5 ml volume of the culture was inoculated into 50 ml of fresh LB-Amp-Glucose medium and incubated at 37°C with vigorous shaking until the culture attained an A<sub>600</sub> of 0.5 units. Then, the culture was centrifuged at 3,000 X g for 5 min and resuspended in 50 ml pre-warmed, fresh LB-Amp medium in a 250-ml Erlenmeyer flask and the first 1 ml sample (T<sub>0</sub>) collected. To induce expression of the fusion product, iso-propyl-thio-β-D-galactoside (IPTG) was added to the cultures to yield a final concentration of 1.3 mM. Cultures were incubated at 37°C with vigorous shaking and 1 ml samples were removed every hour through four hours post-induction. The 1 ml samples were centrifuged at 12,000 rpm for 2 min at RT and the resulting cell pellets were resuspended in 100 μl of a SDS gel-loading buffer (50 mM Tris-HCl, pH 6.8, 100 mM Dithiothreitol (DTT), 2% SDS, 0.1% Bromophenol blue, 10% Glycerol). Then, the samples in SDS gel-loading buffer were heated to 100°C
for 3 min, centrifuged at 12,000 X g for 2 min at RT and stored at -20°C until further use.

**SDS-PAGE Analysis of Fusion Products:** Whole cell lysates of IPTG-induced recombinant clones were analyzed for expression of β-galactosidase fusion products by 7.5% and/or 12% discontinuous Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis (SDS-PAGE; Laemmli, 1970). Polyacrylamide gels of 1.5 mm thickness were used along with a Mighty Small R280 electrophoresis apparatus (Hoefer Scientific Instruments, San Francisco, CA). Following electrophoresis at a constant current of 18 milliamperes for 4 to 5 h, the gels were stained in a Coomassie blue staining solution (0.25% Coomassie Brilliant Blue R-250 (Sigma Chemical Co., St. Louis, MO.), 40% methanol, 7% acetic acid) for 3 to 4 h with gentle shaking and destained for 1 h in a destaining solution I (50% methanol, 10% acetic acid), followed by a destaining solution II (5% methanol, 7% acetic acid) for 4 to 6 hours.

**Western Blot Analysis of Fusion Products:** Bacterial cell lysates intended for Western Blot analysis were resolved first by SDS-PAGE under the conditions mentioned above. Following electrophoresis, the gels were equilibrated for 30 minutes in 500 ml of a Tris-glycine transfer buffer (25 mM Tris; pH 8.3, 192 mM glycine, 20% v/v methanol). While the gel was equilibrating, pieces of filter pads along with pieces of 3MM Whatman filter paper (Whatman LabSales, Hillsboro, OR) and pure nitrocellulose transfer and immobilization membrane (Schleicher & Schuell, Keene, NH), all of which were the same size as the polyacrylamide gel, were soaked to saturation in the Tris-glycine buffer in separate trays. A Western Blot transfer sandwich was assembled for
effecting transfer of proteins onto nitrocellulose using a Trans-Blot™ Cell (Bio-Rad Laboratories, Richmond, CA) according to the manufacturer's recommendations. First, a saturated filter pad was placed on the inner side of the gel holder. Next, a saturated piece of 3MM filter paper was placed on top of the pad followed by the pre-equilibrated gel. Any air bubbles trapped underneath the gel were removed by rolling a glass pipet on the surface of the gel and thereby ensuring that the gel remained in close contact with the filter paper support. Then, the nitrocellulose membrane was placed on top of the gel and any air bubbles trapped between the gel and the membrane were removed as before. This was followed by placement of a saturated filter paper with removal of air bubbles and then, a piece of buffer-saturated filter pad to complete the six-layered sandwich. The gel holder was closed with the sandwich inside and placed in the Trans-Blot™ Cell, with the nitrocellulose membrane side towards the anode. Blotting was performed overnight under a low field (30V/0.1A), with constant stirring and cooling conditions. A standard field 2 h transfer was performed at 60V/0.21A, following the overnight transfer. After completing the transfer procedure, the sandwich was disassembled and the nitrocellulose membrane utilized in the later steps for immunoscreening of fusion proteins with the ProtoblotTM immunoscreening system (Promega, Madison, WI) by the Protoblot™ Immunoscreening method.

The nitrocellulose membrane with the electro-blotted proteins was incubated in TBST-1% BSA (10 mM Tris-HCl; pH8.0, 150 mM NaCl, 0.05% Tween 20) buffer for 30 min at RT with gentle shaking, to saturate any nonspecific protein binding sites on the membrane. Next, the filters were incubated with the primary antibody (rabbit anti-
HAVpolyvalent; mouse anti-HAV monoclonal) in TBST buffer for 30 min at RT with gentle shaking. The primary antibody from the affinity purified rabbit polyvalent anti-HAV serum was diluted 1:1000 in TBST before use. Following the incubation, the membranes were washed thoroughly at least 3 times in 15-20 ml of TBST buffer with 5-10 min for each washing step. After the third wash, the membranes were transferred into TBST containing an appropriate dilution (1:5000 to 1:10,000) of an alkaline-phosphatase (AP)-secondary antibody conjugate (Cappel/Organon Teknika Corp., West Chester, PA) and incubated for 30 min at RT with gentle shaking. The membranes were then washed three times with TBST buffer as described earlier. Following the wash, damp membranes were dried briefly on clean, absorbent filter paper and transferred to a freshly prepared AP color development substrate solution (5 ml AP buffer containing 100 mM Tris-HCl; pH 9.5, 100 mM NaCl, 5 mM MgCl₂ and 66 μl of nitroblue tetrazolium (NBT; 50 mg/ml), 33 μl of 5-bromo-4-chloro-3-indoyl phosphate (BCIP; 50 mg/ml)). The color development was monitored carefully to avoid over exposure and resultant high non-specific background. The reaction was stopped by discarding the substrate development solution at the appropriate time, rinsing the membrane under running water for one min and treating it with a stop solution (20 mM Tris-HCl; pH 8.0, 5 mM EDTA) for 5-10 minutes at room temperature. Subsequently, the membranes were dried on a piece of clean filter paper and stored inside individual filter paper envelopes in the dark.

**Affinity Purification of Fusion Peptides:** Purification of the fusion proteins was done using a commercial APTG (para-aminophenyl-1-thio-β-D-galactopyranoside)
affinity matrix (United States Biochemical Corp., Cleveland, OH). Whole cell lysates were diluted in loading buffer (20 mM Tris-HCl, pH 7.4, 10 mM MgCl$_2$, 1.6 M NaCl, 10 mM 2-mercaptoethanol), to a final concentration of 4 mg/ml and were applied to the gel matrix to yield a flow rate of 1 ml per 10-15 minutes. After loading, the column was washed with 50 - 70 ml of the loading buffer. The bound protein was eluted with 15 - 20 ml of an elution buffer (0.1 M borate, pH 10, 10 mM 2-mercaptoethanol). The eluted fraction was immediately neutralized by collection into a dialysis bag immersed in borate buffered saline (0.67 M H$_3$BO$_3$, pH 8.4, 0.15 M NaCl). The column was washed thoroughly with additional 50 ml of the elution buffer followed by 50-70 ml of storage buffer (10 mM Tris-HCl, pH 7.4). The dialyzed peptides were concentrated with a carboxymethyl cellulose adsorbent (AQUACIDE, CALBIOCHEM Corp., La Jolla, CA). The total protein concentration was determined by the BCA method (Pierce, Rockford, IL), and the purified peptides were analyzed by SDS-PAGE.

**ELISA for antigenicity of Recombinant HAV Capsid Proteins:** Purified, formalin inactivated HAV HM-175 strain virus particles were kindly provided by Dr. Mark D. Sobsey, Univ. of North Carolina, Chapel Hill, NC. Ninety-six well microtiter plates (Immulon II; Dynatech Laboratories Inc., Alexandria, Va.) were coated in duplicate with each of the following components: 5 μg per well of each of the three partially purified recombinant HAV peptides (50 μg/ml), E. coli β-galactosidase (50 μg/ml; Sigma Chemical Co., St. Louis, MO.) and 10$^8$ virions per well of intact HAV (10$^9$ virions/ml) in carbonate coating buffer (600 mM NaHCO$_3$, 600 mM Na$_2$CO$_3$, pH 9.6). The proteins were allowed to adsorb overnight at 4°C. Following the incubation,
the coating buffer was removed and 200-μl of blocking buffer (50 mM Tris-HCl, pH 7.2, 1 mM EDTA, 150 mM NaCl, 2% BSA, 10% fetal calf serum) was added to each well. The plates were incubated for 45 min at room temperature (RT) and then washed twice with 200 μl of TEN buffer (50 mM Tris-HCl, pH 7.2, 1 mM EDTA, 150 mM NaCl, 0.05% Tween-20). Next, 100 μl aliquots of either rabbit pre-immune serum, affinity purified/absorbed polyvalent anti-HAV antiserum, or a purified rabbit anti-β-galactosidase antibody (Cappel/Organon Teknika Corp., West Chester, Pa.) in TEN-10% fetal calf serum (FCS) were added to appropriate wells in individual plates and incubated for 1 h at RT on a rotary platform. The plates were washed six times with TEN and 50 μl of goat-anti-mouse Ig-alkaline-phosphatase enzyme conjugate (Cappel/Organon Teknika Corp., West Chester, Pa.) in TEN-10% FCS were added to each well. After a 45 min incubation at RT, wells were washed as before and then 200 μl of the substrate para-nitrophenyl phosphate (1 mg/ml) (Sigma Chemical Co., St. Louis, Mo.) in diethanolamine buffer (9.7% (v/v) diethanolamine, 0.02% (w/v) NaN₃, 0.01% MgCl₂·6H₂O, pH 9.8) were added to each well. The plates were incubated at 37°C in the dark and the reaction was stopped by adding 50 μl of 3N NaOH per well after 30 min. The absorbance was read using a Bio-Tek EIA Reader (Bio-Tek Instruments) at a 406 nm.

**Phase Contrast and Electron Microscopy of E.coli Expressing Fusion Proteins:** Overnight cultures of recombinant E.coli cells carrying HAV genes were inoculated into 10 ml of fresh LB-amp-glucose media and incubated either at 30°C or 37°C with vigorous shaking for 1 h. One ml of these cultures were transferred into
appropriately labeled Erlenmeyer flasks with 10 ml of fresh LB-amp-glucose and incubated as before for 3-4 h. Each recombinant was inoculated into 2 separate flasks of LB-amp. Following the incubation, 1 ml samples of each culture were collected into individual sterile microfuge tubes (gr) and the remaining cells were centrifuged at 3,000 X g for 10 min. After discarding the supernatant carefully, one cell pellet of each recombinant was gently resuspended in fresh LB-amp without any glucose and the other with fresh LB-amp-glucose. Cells in media without glucose were subjected to IPTG induction as described earlier. One ml samples of each culture (Time Oh) were collected again immediately following the addition of IPTG to one set of cultures. The remaining cultures were incubated at 30°C or 37°C for 2h before collecting the next set of samples (Time 2h). *E.coli* UT2300 cells (i.e. without any recombinant plasmid) and cells carrying vector pAX4b+ without any insert DNA were used as negative and positive controls in this experiment.

The live cells from various time points mentioned above were transferred onto microscopic slides immediately after sample collection and the individual preparations were examined under a phase contrast microscope shortly thereafter. Selected microscopic fields were photographed with phase optics on a Nikon Microphot HFX-2A.

Bacteria from time 2h were also examined by Transmission Electron Microscopy after preparing cells as follows. Cells were fixed in 2% glutaraldehyde and 1-2% Osmium tetraoxide for 20 min and then *en bloc* stained in 0.5% Uranium acetate. The cells were then dehydrated in absolute ethanol and embedded in LR White resin; thin
sections were poststained with lead citrate. Grids were examined and photographed on a JEOL JEM 100CX transmission electron microscope.

**Large Scale Expression and Partial Purification of Recombinant Proteins by Sarkosyl Extraction of Inclusion Bodies:** Large scale production of the fusion peptides was accomplished by a modification of a method previously reported (Stark, 1987). Bacterial clones were inoculated onto LB-Amp-Glucose medium and incubated as described before. One liter cultures were grown at 37°C with continuous shaking to an optical density of 0.5 units at a wavelength of 600 nm. The cells were centrifuged at 5,900 x g for 15 min and the cell pellets were resuspended to the same density in LB media containing ampicillin (100 μg/ml), with no glucose. Expression of the recombinant plasmids was induced by addition of IPTG and the cultures were incubated at 37°C in a shaker incubator for 4 h. The fusion peptides were extracted by a modification of the method described previously (McNally, *et al.*, 1991). Following IPTG-induction, the cells were sedimented at 5,900 x g for 30 min and resuspended in 100 ml of washing buffer (100 mM Tris-HCl, pH 7.4, 10 mM EDTA, 10 mM 2-mercaptoethanol, 1 mM phenyl methyl sulfonyl fluoride (PMSF)). The cells were washed three times and the pellet was weighed. Cells from 1 liter of induced culture (6-8 g/L) were then resuspended in 15 ml of a low salt buffer (10 mM triethanolamine, 0.5 mM ATP, 0.5 mM DTT, 0.1 mM CaCl₂) containing protease inhibitors (Aprotinin 20 μg/ml, Leupeptin 10 μg/ml, Pepstatin 2.5 μg/ml, Chymostatin 2.5 μg/ml, O-Phenanthroline 0.5 mM, PMSF 1 mM) (27). The cells were disrupted by two passages through a French Press cell at 1000 lb/in² followed by centrifugation at 500 rpm for 10
min to remove unbroken cells. The supernatant fluid was then centrifuged at 116,500 x g for 8 min and the resulting pellet was resuspended in 10 ml of extraction buffer (1.5% Sarkosyl, 25 mM triethanolamine, 4 mM ATP, 0.8 mM DTT, 1 mM EDTA, 0.02% sodium azide) containing protease inhibitors using a Teflon-glass homogenizer and centrifuged again at 116,500 x g for 16 min. The supernatant was collected into a sterile beaker and 58.5 ml of octylglucoside (OG) buffer (to sequester Sarkosyl) was added. This resulted in a final concentration of 2% octylglucoside and 0.2% Sarkosyl in a solution containing 25 mM triethanolamine, 800 mM NaCl, 1.0 mM ATP, 0.8 mM DTT, 0.02% sodium azide, 0.13 mM EDTA, and protease inhibitors. After mixing thoroughly, CaCl₂ was added to a final concentration of 100 mM. The protein suspension was then transferred into Spectra/Por CE membranes (MWCO 100,000; Spectrum, Houston, TX) and dialyzed for 96 h at 4°C against a dialysis buffer (25 mM triethanolamine, 800 mM NaCl, 1 mM ATP, 0.2 mM DTT, 0.02% NaN₃, 0.13 mM EDTA and protease inhibitors). The dialysate was replaced with fresh buffer every 8 h. The total protein concentration of the dialyzed protein suspension was determined by the Bicinchoninic acid (BCA) method (Pierce, Rockford, IL.), followed by SDS-PAGE and Western Blot analysis.

RESULTS

The nucleotide sequences selected to design flanking oligonucleotide primers for each capsid sequence were reported previously (Cohen, 1987a). Restriction endonuclease sequences were incorporated to flank the viral sequences to facilitate directional cloning into the plasmid vector. Figure 5 shows the PCR-amplified products
Figure 5. Agarose gel analysis of DNA segments amplified from pGEM-HAV by PCR and encoding HAV capsid proteins. Lane 1: BioMarker low molecular weight DNA standard, Lane 2: VP1 (899 bp), Lane 3: VP2 (665 bp), Lane 4: VP3 (737 bp), Lane 5: BioMarker low molecular weight DNA standard.
encoding VP1, VP2, and VP3. The molecular size of each product corresponded to that of each specific capsid gene.

Following cloning and confirmation that the relevant insert DNA was present in the recombinant plasmids, Ax-VP1II, -VP2B2 and -VP3T3, expression of individual β-galactosidase fusion proteins was assessed in three different *E. coli* strains. Optimum protein expression occurred when overnight cultures were expanded through two consecutive passages in fresh medium supplemented with glucose and ampicillin prior to IPTG induction.

Expression from recombinant pAx-VP1II was assessed first in *E. coli* UT 2300. Analysis of protein expression under IPTG induction revealed two prominent overexpressed protein bands of 125 kD and 150 kD (Figure 6). The 125 kD band corresponding to β-galactosidase was also expressed from the vector pAX4b (Figure 6). The 150 kD corresponded to the fusion product and was detectable one hour following IPTG-induction. Western blot analysis revealed that both *E. coli* β-galactosidase and the recombinant fusion protein bands were stained by rabbit anti-HAV polyvalent antiserum (Figure 9). Surprisingly, the anti-HAV antiserum appeared to possess serological activity against *E. coli* β-galactosidase (MW = 116 kD) present in the commercial molecular weight markers (Figure 9; Biorad laboratories, Richmond, CA) and β-galactosidase expressed from the plasmid pAx4b in *E. coli* UT2300 (not shown). The anti-HAV serum was absorbed with commercial *E. coli* β-galactosidase (Sigma Chemical Co., St.Louis, MO) and the Western Blot experiment was repeated. Neither the recombinant protein nor the β-galactosidase protein bands were stained with the
Figure 6. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP111 in host *E. coli* UT2300 cells. *E. coli* lysates from various time points before and after IPTG induction were examined. Lane 1: Host cells without vector. Lane 2: Biorad molecular weight marker. Lane 3: Cells under glucose repression. Lane 4: Cells at 1 mM IPTG induction (Tₜ). Lane 5: 1 h after induction. Lane 6: 2 h. Lane 7: 3 h. Lane 8: 4 h. Lane 9: Rainbow™ molecular weight marker. Lane 10: Host cells carrying vector pAX4b+.
absorbed antiserum (Figure 9). When these experiments were performed with a neutralizing IgG mouse MAb (P2G9), and a non-neutralizing IgM mouse MAb (P6B5), both MAbs stained the fusion protein as well as β-galactosidase, similar to the polyvalent anti-serum before absorption with the β-galactosidase (Figure 9). Induction of protein expression from pAX-VP1I1 in *E. coli* GW1000 showed no discernible expression of the fusion product (not shown). Similarly, recombinant pAX-VP1I1 in *E. coli* Top 10 also exhibited poor expression of the fusion protein (not shown).

Induction of pAX-VP2B2 in *E. coli* UT2300 led to expression of the recombinant β-gal-VP2 protein as the predominantly overexpressed protein (Figure 7). A marked increase in the production of the recombinant product was observed by examination of the level of expression between time 0 and time 1h (Figure 7). Western blot studies revealed findings similar to those observed with pAX-VP1I1 (Figure 9). Expression of the recombinant β-gal-VP2 fusion protein in *E. coli* GW1000 revealed overexpression of the galactosidase protein only (Figure 9). Similarly, expression of both β-galactosidase and the β-gal-VP3 recombinant protein increased as the induction period progressed in the UT2300 host strain (Figure 8). Fusion protein expression from pAX-VP3T3 in the GW1000 host strain was low and β-galactosidase was the predominant overexpressed protein (not shown). Again, Western blot analyses showed that the anti-HAV antibodies reacted with both protein bands, β-galactosidase and the fusion proteins (Figure 8).

Large-scale production in one liter cultures and purification of the fusion proteins β-gal-VP1, -VP2 & -VP3 was done. The cells were fractured by sonication and the
Figure 7. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP2B2 in host *E. coli* UT2300 (A), and in GW1000 cells (B). *E. coli* lysates from various time points before and after IPTG induction were examined. Lane 1: host cells without vector. Lane 2: molecular weight marker. Lane 3: Cells under glucose repression. Lane 4: Cells at 1 mM IPTG induction (T₀). Lane 5: 1 h after induction. Lane 6: 2 h. Lane 7: 3 h. Lane 8: 4 h. Lane 9: molecular weight marker. Lane 10: Host cells carrying vector pAX4b+.
Figure 8. 7.5% SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-VP3T3 in host *E. coli* UT2300 cells. *E. coli* lysates from various time points before and after IPTG induction were examined. Lane 1: host cells without vector. Lane 2: Biorad molecular weight marker. Lane 3: Cells under glucose repression. Lane 4: Cells at 1 mM IPTG induction (T₀). Lane 5: 1 h after induction. Lane 6: 2 h. Lane 7: 3 h. Lane 8: 4 h. Lane 9: Rainbow™ molecular weight marker. Lane 10: Host cells carrying vector pAX4b+.
Figure 9. Western Blot analysis of fusion proteins expressed in *E. coli* UT2300 cells from recombinant plasmids pAX4b-VP1I1, -VP2B2, and -VP3T3. Whole cell lysates 4 h after IPTG induction were resolved by a 6.5% SDS-PAGE, transferred onto nitrocellulose and screened with rabbit anti-HAV polyvalent antiserum adsorbed with *E. coli* whole cell extract (A), rabbit anti-HAV polyvalent antiserum adsorbed with *E. coli* whole cell extract and β-galactosidase (B), mouse anti-HAV monoclonal IgM antibody P6B5 (C), neutralizing mouse anti-HAV monoclonal IgG antibody P2G9 (D). Lane 1: BioRad molecular weight marker. Lane 2: VP1I1/UT2300. Lane 3: VP2B2/UT2300. Lane 4: VP3T3/UT2300. Lane 5: Rainbow™ molecular weight marker.
fusion proteins precipitated at 30% ammonium sulfate concentration. However, upon attempts to dissolve the precipitate in buffer, a significant portion of this preparation formed insoluble aggregates. Further, both phase contrast and electron microscopic examination revealed aberrant changes in the morphology of the recombinant bacterial cells (Figure 10), which suggested the possibility that the fusion proteins were organized into inclusion bodies within the host cells (Figure 11). In an attempt to extract and then solubilize the fusion proteins in a more efficient manner, a combination of French Press lysis and Sarkosyl extraction was used and fusion protein β-gal-VP1II was recovered. A significant fraction of the fusion protein remained in solution following extraction and high speed centrifugation, and sequestration of the Sarkosyl by octylglucoside and dialysis employed to remove the detergents from the protein solution (Figure 12).

An APTG affinity column was used to immobilize the fusion proteins from the fractionated cell lysates obtained by the sonication-ammonium sulfate precipitation procedure, or by the French Press lysis-Sarkosyl extraction. Fusion proteins (~150 kD) obtained from the pAX4b+ recombinant plasmids, and containing the entire β-galactosidase molecule were copurified with β-galactosidase (~125 kD) (Figure 13).

The affinity purified recombinant HAV capsid proteins, VP1II, VP2B2 and VP3T3 were examined for serological activity (antigencity) by ELISA (Table 2). *Escherichia coli* β-galactosidase and intact HAV particles were included as controls in the immunoassay. The HAV particles and recombinant HAV proteins, VP1II and VP2B2 reacted with the absorbed rabbit polyvalent antibodies at high titer as indicated by OD<sub>405</sub> readings. The recombinant VP3T3 and *E. coli* β-galactosidase proteins reacted
Figure 10. Phase contrast microscopy of *E. coli* cells induced for expression of fusion proteins from recombinant plasmids encoding sequences for HAV capsid proteins. All fields are at same magnification (X 1200). (A) Host *E. coli* UT2300 cells. (B) VP1I1/UT2300. (C) VP2B2/UT2300. (D) VP3T3/UT2300. Some inclusion bodies (IB) are marked by arrows.
Figure 11. Electron microscopy of *E. coli* cells induced for expression of recombinant fusion proteins from recombinant plasmid pAX4b-VP1I1. The host cells display intracellular inclusion bodies (IB). Transmission electron micrograph of thin sections (< 100 nm) of *E. coli* AX-VP1I1/UT2300 cells. (A) Magnification X 13,000. (B) Magnification X 33,000. Some of the amorphous, homogeneous and predominantly polar inclusions are marked by arrows.
Figure 12. SDS-PAGE analysis of samples from various solubilization steps during the large scale extraction of VP111 fusion protein by the French Press-Sarkosyl extraction method. Lane 1: Whole cell lysate from host UT2300 strain. Lane 2: Biorad SDS-PAGE molecular weight standards-High. Lane 3: Whole cell lysate of VP111/UT2300 prepared from sample collected prior to the French Press lysis step. Lane 4: Whole cell lysate of VP111/UT2300 following French Press lysis at 1000 lb/in². Lanes 5 and 6: Samples prepared from pellet (Lane 5) and supernatant (Lane 6) obtained after French Press lysis and centrifugation at 500 rpm for 10 min. Lanes 7 and 8: Samples representing inclusion body pellet (Lane 7) and supernatant (Lane 8) following centrifugation at 116,500 x g for 8 min. The IB pellet was solubilized in extraction buffer. Lanes 9 and 10: Samples representing pellet (Lane 9) and supernatant (Lane 10) following centrifugation of sarkosyl-extracted IB sample at 116,500 x g for 16 min. The supernatant was mixed with buffer containing octylglucoside.
Figure 13. SDS-PAGE analysis of affinity-purified fusion proteins expressed from recombinant plasmid pAX4b-VP1, -VP2, and -VP3 in *E. coli* UT2300 cells. The individual recombinant proteins were affinity purified using an APTG column. Lane 1: Molecular weight marker. Lane 2: Fusion protein β-gal-VP3. Lane 3: Fusion protein β-gal-VP2. Lane 4: Fusion protein β-gal-VP1. Lanes 5 and 6: Molecular weight markers.
Table 2. ELISA with Rabbit Antiserum for Detecting Antigenicity of Affinity Purified Recombinant HAV Capsid Proteins.

<table>
<thead>
<tr>
<th></th>
<th>pre-immune serum</th>
<th>immune serum @ 1:500</th>
<th>anti-β-gal serum</th>
<th>no serum</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP1I1 (β-gal-VP1 fusion protein, 5 µg/well)</td>
<td>0.208&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.907</td>
<td>1.264</td>
<td>0.164</td>
</tr>
<tr>
<td>VP2B2 (β-gal-VP2 fusion protein, 5 µg/well)</td>
<td>0.168</td>
<td>0.992</td>
<td>1.227</td>
<td>0.197</td>
</tr>
<tr>
<td>VP3T3 (β-gal-VP3T3 fusion protein, 5 µg/well)</td>
<td>0.186</td>
<td>0.648</td>
<td>1.515</td>
<td>0.508</td>
</tr>
<tr>
<td>E. coli β-galactosidase, 5 µg/well</td>
<td>0.164</td>
<td>0.572</td>
<td>1.155</td>
<td>0.176</td>
</tr>
<tr>
<td>Formalinized intact HAV particles (10&lt;sup&gt;8&lt;/sup&gt; virions/well)</td>
<td>0.162</td>
<td>0.820</td>
<td>0.354</td>
<td>0.176</td>
</tr>
</tbody>
</table>

<sup>a</sup> absorbance @ 405 nm (in triplicate)
less strongly albeit still at higher titers when compared with the negative controls. The anti-β-galactosidase antibody, reacted very strongly with each of the recombinant proteins as well as with native β-galactosidase and only poorly with intact HAV particles. The rabbit pre-immune serum did not yield any significant titer against any of the antigens tested.

**DISCUSSION**

Cloning and expression of PCR amplified HAV capsid genes in prokaryotic or eukaryotic systems provide a convenient and safe alternative to isolation of viral proteins by disruption of infectious and potentially hazardous virus particles. Purification of HAV proteins from intact infectious virions is complicated by low and unreliable yields of the specific proteins extracted from either infected liver tissue or cell cultures (Hughes, *et al.*, 1984; Cromeans, *et al.*, 1989). A PCR amplified product which includes an entire viral gene cloned and then expressed in a prokaryotic vector may provide a superior and more convenient approach for immunogenic assessment of each protein. The antigenicity of each recombinant protein can then be compared with the intact capsid of native virions. It has been postulated that each of the three major viral capsid peptides contribute to the conformation of the immunologically dominant and relevant epitopes (Emini, *et al.*, 1985; Fayol, and Ville, 1991; Hughes, 1985). The goal of this study was to determine whether large-scale stable expression of HAV peptides in *E. coli* was possible and if it is, whether a substitution of tissue culture grown virus, currently used in immunodiagnostic and serology, by recombinant viral capsid peptides is practical and commercially feasible.
The prokaryotic expression vector, pAX4b+ (Figure 3), was selected since it would allow for fusion of the entire *E. coli* β-galactosidase protein with the N-terminus of the foreign protein of interest, thereby enabling easy identification and purification of the expressed fusion protein by APTG-affinity chromatography directed towards the β-galactosidase component. Furthermore, a hinge region of 62 amino acids derived from the structural protein, collagen, which is positioned between the β-galactosidase and the protein of interest following successful cloning, will permit independent folding of the two components of the protein fusion. An endoproteinase Xa specific sequence exists in the rigid hinge which permits cleavage of the protein of interest from β-galactosidase upon exposure to endoproteinase Xa. Application of digested protein mixture onto the same APTG-column will selectively bind and retain the *E. coli* β-galactosidase, while the protein of interest is eluted and recovered as a pure fraction.

The incorporation of unique restriction endonuclease sites as flanking sequences into each oligonucleotide primer facilitated directional cloning of the amplified PCR products. Recombinant clones, each of which exhibited variable expression efficiency the target gene products, were thus obtained. Significant morphological changes were observed in the recombinant strains during expression of the recombinant viral proteins. The appearance of elongated, filamentous cells which possessed polar inclusion bodies and exhibited slower growth rates suggested that the cloned viral genes or their products maybe toxic to and exert an impact on host bacterial cell metabolism.

The HAV capsid proteins expressed as fusions with the *E. coli* β-galactosidase exhibited antigenic properties when assessed by Western Blot analyses using both
polyvalent and monoclonal antibodies generated in animals immunized with intact Hepatitis A virus. The VP1 recombinant protein also reacted with MAb P2-G9 which had been previously determined to possess neutralizing activity against intact HAV particles (Baez, L.A., 1994). These findings suggest that the recombinant VP1 may contain epitopes that resemble those present on the intact Hepatitis A virion.

*E. coli* UT2300 cells exhibited a better level of fusion protein expression when compared to the other strains tested, probably due to the fact that they are protease deficient. A similar observation was made earlier (Lindler, *et al.*, 1991) regarding a marked difference in expression and stability of recombinant malarial peptides in *E. coli* which was shown to be associated to the expression of proteases linked to the bacterial heat shock response. The simultaneous expression of both the 125kD *E. coli* β-galactosidase and the ~150 kD recombinant β-gal-VP fusions may have been due to events such as premature termination of transcription by *E. coli* RNA polymerase due to fortuitous intragenic terminators within the viral sequences, or a rapid mRNA turnover due to formation of secondary structures recognized by host nucleases, or an inefficient translation resulting from a biased codon usage. In addition, the possibility that a specific autoproteolytic cleavage may have occurred *in vivo* during synthesis of the recombinant proteins cannot be ruled out. Such an event could be mediated by capsid sequences of the fusion proteins, although such an activity for HAV capsid proteins VP1 and VP3 has not been reported.

One persistent observation was serological cross reactivity which occurred between the two proteins, β-galactosidase and the fusion proteins, during Western blot
analysis with a panel of anti-HAV antibodies which included a neutralizing monoclonal antibody used to screen and detect the expressed proteins. Further, preliminary studies with human acute phase serum collected from patients who tested positive for HAV revealed a similar cross reactivity between β-galactosidase and the recombinant HAV proteins (data not shown) indicating that this phenomenon is not limited to a single source of anti-HAV antibodies. The possibility of non-specific binding to β-galactosidase by goat anti-rabbit or anti-mouse alkaline phosphatase conjugated antibodies was examined by omitting addition of anti-HAV antibodies during the blotting procedure. The recombinant proteins were not stained by the goat antibodies alone.

Affinity purification of recombinant fusion proteins was complicated by co-expression of β-galactosidase in vivo during IPTG-induction of recombinant E.coli and its interference in efficient purification of the individual fusion proteins. Early attempts to cleave the fusion proteins with endoproteinase Xa following immobilization onto the APTG-affinity column or attempts to purify capsid proteins by passage through the APTG column subsequent to digestion of the fusion proteins outside the column failed. Hence, we had to abandon this strategy to obtain individual HAV capsid proteins by subjecting the affinity-purified fusion proteins to endoproteinase Xa cleavage.

The ELISA was used to detect antigenicity of affinity purified recombinant HAV proteins. The ELISA results, while confirming specificity of absorbed rabbit antiserum against intact HAV particles, also suggested an interaction between the anti-HAV antibodies and E. coli β-galactosidase that resulted in a titer lower than any seen with HAV proteins but nevertheless significant. Moreover, rabbit anti-β-galactosidase
antibody, which reacted strongly with both native β-galactosidase and recombinant β-galactosidase-HAV capsid proteins, exhibited a weak reaction with intact HAV particles. It should be emphasized at this juncture that there have been no previously published reports regarding non-specific recognition of *E. coli* β-galactosidase by rabbit or mouse antibodies when used in either ELISA or Western blotting procedure. Therefore, the serological observations strongly suggest a possible coincidental existence of certain common epitopes between two ostensibly unrelated proteins, β-galactosidase and HAV capsid proteins, at least in the SDS-mediated denatured state.

HAV fusion proteins sequestered in inclusion bodies following IPTG induction of host cells could permit their isolation and manipulation as a distinct subcellular fraction. The French Press-Sarkosyl extraction procedure enables efficient solubilization of such inclusion bodies on a large scale.
CHAPTER II.

EFFECTS OF AN ANTI-TERMINATOR SEQUENCE ON RECOMBINANT HAV PROTEIN EXPRESSION IN *Escherichia coli*
INTRODUCTION

Expression of heterologous proteins in *Escherichia coli*, subsequent to successful cloning of the relevant gene into appropriate procaryotic plasmid vectors, requires efficient transcription and translation of the gene of interest. It is often possible by recombinant DNA techniques to engineer and produce successfully in *E.coli* any protein that is not too small, too large, too hydrophobic, and does not contain too many cysteines. However, expression of certain foreign cDNA or genes may present unique challenges that must be overcome to achieve high-level expression. Our earlier attempts at β-galactosidase-HAV VP fusion protein expression from expression vector pAX4b+ and their subsequent affinity-based purification were frustrated by the simultaneous expression of β-galactosidase without the viral protein components. Factors which may have contributed to the dichotomous expression of the fusion proteins in *E.coli* include premature termination of transcription, rapid decay of mRNA, poor translation, or extreme instability or toxicity of the proteins expressed (Balbas and Bolivar, 1990; Das, 1990; Powdrill and Johnston, 1991).

We speculated that premature termination of transcription of β-galactosidase-HAV VP gene fusions may have been due to fortuitous terminator-like sequences in the mRNA encoding the ~150 kD fusion proteins. Presence of an intragenic terminator within a foreign DNA insert would prevent a complete, high-level transcription of the cloned gene. Previously published reports suggest that an inclusion of various anti-termination sequences downstream of the promoter from which the recombinant genes are transcribed might improve and ensure complete transcription of such messages.

Such approaches for influencing gene expression have utilized, various components of the λ early or late antitermination systems in *E. coli* (Friedman, 1988; Podhajska, et al., 1985; Hasan and Szybalski, 1987; Shatzman and Rosenberg, 1987). In addition to the λ-derived antitermination systems, other mechanisms of transcriptional antitermination are operative in the *E. coli* (Greenblatt, et al., 1993). One such antitermination system, which shares several features in common with the λ-derived antitermination systems, utilizes elements of the *E. coli* ribosomal RNA (rrn) operon (Li, et al., 1984).

The bacteriophage λ antitermination systems in *E. coli* are mediated by nucleotide sequences and protein factors derived from both the infecting bacteriophage and the host bacterial cell (Friedman, 1988). Early transcriptional antitermination is mediated by the λ *N* gene product and a short nucleotide sequence, *nut* (the *N* utilization site) which is thought to be functional only after its transcription into the RNA form (Das, 1992). The *nut* site includes a seven-nucleotide boxA sequence followed by a stem-loop boxB motif. In addition to the λ *N* protein and the antiterminator nucleotide sequences, four *E. coli* Nus (N utilization substance) proteins namely NusA, NusB, NusE, and NusG, are required for successful N-mediated antitermination (Li, et al., 1992; Nodwell, et al., 1993). The late antitermination system is much simpler and depends on a λ *qut* sequence, the λ *Q* protein, and the *E. coli* NusA protein (Yarnell and Roberts, 1992).

The *E. coli* ribosomal RNA (rrn) operon leader sequences contain regions homologous to the λ *nut* site. The boxA and boxB elements are similar but seem to be
reversed in their corresponding locations (Figure 14). Recent studies have shown that in the presence of E. coli Nus factors, transcription of a short, 12-nucleotide boxA sequence, 5'-TGC TCT TTT AAC-3', is essential and maybe sufficient to render the E. coli RNA polymerase terminator-resistant both in vivo and in vitro and cause read-through of Rho-dependent terminators. However, the exact mechanism of antitermination in the E. coli rrn operons remain to be elucidated.

The λ and rrn antitermination systems along with the Nus factors induce and maintain a long-lived modification of the E. coli RNA polymerase in the elongation complex and enable transcription through multiple consecutive terminators (Albrechtsen, et al., 1990). An alteration in the RNA polymerase function can exert several consequences such as the protection of untranslated rRNA transcripts from premature Rho-dependent termination, regulation of protein synthesis, RNA folding patterns and maturation (Aksoy, et al., 1984; Morgan, 1986).

To improve the expression of β-galactosidase-HAV VP1 fusion-protein in E. coli, it was assumed that premature termination of transcription may have been responsible for the inefficient expression from the recombinant plasmids constructed earlier. A twelve nucleotide boxA sequence derived from the E. coli rrn antiterminator sequence was incorporated upstream of the HAV VP1 sequence during cloning into plasmid vector, pAX4b+. Subsequently, to study the effects of the boxA sequence on fusion-protein expression in vivo, the recombinant plasmid DNA was used to transform host E. coli cells as a prelude to IPTG induction of fusion-protein expression. The resultant
(i) Phage \( \lambda \) transcriptional antiterminator sequence of the \textit{nutR} region.

\[ 5'-\text{CGCTTTTACACATCCCAGGCTGAAGAGGCATCAAAATTAAACCACCCATGCTGTATG-3'} \]

\begin{align*}
\text{BoxA} & & \text{BoxB} & & \text{BoxC} \\
\end{align*}

(ii) \textit{E. coli} transcriptional antiterminator sequence of the \textit{rrnG} leader region.

\[ 5'-\text{ATCGATCG-CACTTTTTACACATTTTACTGAACTTGTTGCGGATCC-3'} \]

\begin{align*}
\text{BoxA} & & \text{BoxC} \\
\end{align*}

\textbf{Figure 14.} Transcriptional antiterminator sequences from phage \( \lambda \) and \textit{E. coli} genomes. Both sequences have been drawn as single-stranded DNA. The BoxA and BoxC sequences are underlined. The ability of BoxB sequence from \textit{E. coli} to form a stem-loop structure also is shown.
protein was subjected to analysis by SDS-PAGE, Western Blot and limited N-terminal amino-acid sequencing techniques.

**MATERIALS AND METHODS**

**Bacterial Strains and Plasmids**: Bacterial strains and plasmids used in this study are listed on Table 3.

**Oligonucleotide Primers**: The nucleotide sequences of each PCR primer used in this study are listed in Figure 15. The primers were synthesized by beta cyanoethyl phosphoramidite chemistry (Genelab, Dept. of Vet. Med. & Parasitol., School of Veterinary Medicine, LSU, Baton Rouge, LA). Primer AXBOXAVPIPS included the translational antiterminator BoxA sequence 5'-TGC TCT TTA ACA-3' of *E.coli* ribosomal RNA (*rrn*) operon.

**PCR Amplification and Cloning of HAV VP1 DNA with a 5' Extension Carrying the Antiterminator BoxA Sequence**: Primers AXBOXAVP1PS and AXVP1NS were used in conjunction with linearized template pGEM-HAV DNA to generate insert AXBoxAVP1 DNA by PCR. Individual reactions were set up in 100 μl volumes in accordance with the procedure followed earlier for PCR amplification of HAV VP1, VP2 and VP3 DNA. Thermal cycling was performed through 30 cycles as follows: denaturation at 94°C for 90 sec, annealing at 55°C for 60 sec and elongation at 72°C for 90 sec with an elongation autoextension step of 10 sec per cycle. Following a final elongation step of 10 min, the reaction mix was cooled to 4°C, separated from the mineral oil overlay and transferred to a fresh microfuge tube. A five μl volume of the products were screened by 2% agarose gel electrophoresis to determine if the specific
Table 3. Bacterial strains and plasmid vectors used to study the effects of *E.coli* *rrn* boxA antiterminator on recombinant HAVVPI expression

<table>
<thead>
<tr>
<th>Strains/Plasmids</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em> UT2300</td>
<td><em>azi-6, fhuA23, lacY1, leu-6, mlt-1, proc14, purE42, rpsL109, thi-1, trpE38, tsx-67, entA403, purE</em>, fhuA*, fepA, Δlon::Tn10*</td>
<td>Elish, <em>et al.</em>, 1988</td>
</tr>
<tr>
<td>Plasmid pAX4b +</td>
<td></td>
<td>United States Biochemical Corporation, Cleveland, OH</td>
</tr>
<tr>
<td>Plasmid pGEM-HAV</td>
<td></td>
<td>R.H. Purcell (N.I.H.)</td>
</tr>
</tbody>
</table>
Figure 15. Oligonucleotide primer sequences used for PCR amplification of insert HAV BoxA VP1 capsid DNA for forced directional cloning into the pAX4b+ vector and for verification of the presence of the insert DNA following successful cloning. (VP) Viral Peptide, (lacZ) E.coli gene encoding the β-galactosidase protein, (PS) positive strand primer and (NS) negative strand primer. The restriction endonuclease recognition sequences (EcoRI, PstI) within each primer are underlined.
-950 bp AXBoxAVP1 fragment had been amplified, before the remaining product mix was precipitated in ethanol. The resulting undigested insert AXBoxAVP1 DNA pellet was resuspended in 20 μl of TE buffer and the DNA concentration estimated by A<sub>260</sub> measurements. The concentrated DNA stock was stored at -20°C.

Restriction digestion with EcoRI and PstI was performed sequentially on 5 μg of insert DNA to generate cohesive-ended double-digested insert DNA for a forced directional cloning into a similarly digested pAX4b+ vector. The double digested insert and vector DNA samples were purified separately by the GENECLEAN II procedure. Ligation reactions were set up with Insert to Vector ratios of 1:1, 2:1 and 4:1 and performed under conditions as outlined earlier for cloning HAV genes into the pAX vector. Next, freshly prepared electrocompetent host E.coli UT2300 cells were transformed with DNA from the ligation reaction by electroporation. Recombinants on SOC-Amp plates were identified and selected for the subsequent plasmid DNA extraction, verification of the presence of insert AXBoxAVP1 insert by PCR and for IPTG induction studies by methods similar to those described in Chapter I. Conditions for PCR were optimized for the pBoxAVP1 template-AXBOAxAVPIPS and AXVP1NS primer set combination by using the Opti-Prime™ kit (Stratagene, La Jolla, CA) as follows.

A structured trial-and-error optimization of buffer conditions for amplifying insert AXBoxAVP1 DNA from pAXBoxAVP1 template was performed by constituting 12 individual 50 μl PCR reaction mixtures, each with identical template-primer-dNTP-AmpliTaq polymerase content but with a different formulation of the PCR reaction.
buffer, as outlined below. A master mix consisting of 12.5 μl of Master Mix buffer (20 mM Tris-HCl, pH8.0, 250nM EDTA), 12.5 μl each of dATP, dTTP, dCTP and dGTP (from individual dNTP stock solutions, each at 10 mM concentration), 2.5 μg of each oligonucleotide primer (AXBOXAVP1PS and AXVP1NS), 1.2 μg of pAXBoxAVP1 template DNA, 30 U of AmpliTaq DNA polymerase and χ μl of sterile, redistilled water to make up a final volume of 565 μl was prepared in a microfuge tube on ice. A set of 12 GeneAmp® thin-walled PCR reaction tubes were labeled 1-12 and 45 μl of the master mix dispensed into each. Next, from a set of twelve 10x PCR reaction buffers, each containing 100 mM Tris-HCl but different from the others in final pH, total MgCl2 and KCl concentrations (Figure 16), a 5 μl volume was added to the correspondingly numbered PCR reaction tube. The 10X PCR Reaction buffers were provided with the Opti-Prime™ kit. The contents of each reaction tube were vortexed and overlaid with 25 μl of mineral oil before thermal cycling reaction in a GeneAmp Thermal Cycler 480 (Perkin-Elmer Cetus, Norwalk, CT) as suggested in the protocol (Instruction Manual accompanying the Opti-Prime kit, Stratagene, La Jolla, CA). Thermal cycling parameters were as follows: Cycle 1, Denaturation at 94°C for 3 min followed by Annealing at 55°C for 2 min followed by Elongation at 72°C for 1.5 min; Cycle 2-30, Denaturation at 94°C for 1 min, Annealing at 55°C for 1 min and Elongation at 72°C for 1.5 min. After an additional 8 min elongation step at the end of cycle 30, each sample was cooled to 4°C, the aqueous layer from each reaction tube containing the PCR products was separated from the mineral oil overlay and transferred to a fresh labeled microfuge tube. Ten μl of each of the 12 PCR samples was subjected to 2% agarose gel
<table>
<thead>
<tr>
<th>pH</th>
<th>10 mM Tris-HCl</th>
<th>MgCl₂</th>
<th>25 mM KCl</th>
<th>75 mM KCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.3</td>
<td>1.5 mM</td>
<td>Buffer #1</td>
<td>Buffer #2</td>
<td></td>
</tr>
<tr>
<td>8.3</td>
<td>3.5 mM</td>
<td>Buffer #3</td>
<td>Buffer #4</td>
<td></td>
</tr>
<tr>
<td>8.8</td>
<td>1.5 mM</td>
<td>Buffer #5</td>
<td>Buffer #6</td>
<td></td>
</tr>
<tr>
<td>8.8</td>
<td>3.5 mM</td>
<td>Buffer #7</td>
<td>Buffer #8</td>
<td></td>
</tr>
<tr>
<td>9.2</td>
<td>1.5 mM</td>
<td>Buffer #9</td>
<td>Buffer #10</td>
<td></td>
</tr>
<tr>
<td>9.2</td>
<td>3.5 mM</td>
<td>Buffer #11</td>
<td>Buffer #12</td>
<td></td>
</tr>
</tbody>
</table>

Figure 16. Buffer matrix of the Opti-Prime™ PCR Optimization Kit (Stratagene, La Jolla, CA). The final reaction concentrations of the various buffer components are shown in the top row. The first column specifies the pH of Tris-HCl in the buffers appearing in the last two columns. The second column specifies the final concentration of MgCl₂ in the buffers appearing in the last two columns. The last two columns specify the final KCl concentration of each buffer. For example, buffer #5 is 10mM Tris-HCl (pH 8.8), 1.5mM MgCl₂ and 25mM KCl.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
electrophoretic analysis to determine the optimal pH, MgCl2 and KCl concentrations for each DNA template and primer set.

**IPTG-induction of Protein Expression in Recombinant *E.coli***

**carrying pAXBoxAVP1**: Recombinant *E.coli* UT2300 cells were grown to log-phase in LB-Amp-Glucose medium at 30°C with vigorous shaking following the procedure described (Chapter I; Materials and Methods). Prior to addition of IPTG, cells were removed from under glucose repression by centrifugation and resuspended in fresh LB-Amp without glucose. One ml samples of non-induced and IPTG-induced cells were collected periodically for up to 4 h post-induction. Whole cell lysates were prepared for SDS-PAGE analysis from cell pellets and stored at -20°C until required. Aliquots of whole cell lysates were resolved on 1.5 mm gels and either Coomassie Blue stained to screen for any IPTG-induced protein bands or transferred onto nitrocellulose for Western Blot analysis with anti-HAV antibodies as previously.

**Preparation of Recombinant Protein Samples for a Limited N-terminal Protein Sequencing Analysis**: Whole cell lysates of *E.coli* recombinant BoxAVP1 were resolved on 14 X 16 cm vertical slab gels by preparative denaturing gel electrophoresis using a Standard Cooled Dual Vertical Slab Unit-SE 600 Series (Hoefer Scientific Instruments, San Francisco, CA). Two SDS-polyacrylamide gels (12.5%) of 0.75 mm thickness were set up fully immersed in the SDS-PAGE running buffer (0.025 M Tris, 0.192 M Glycine, 0.1% SDS, pH 8.3) in the SE 600 electrophoresis unit following the manufacturer's instructions (Basic techniques and Exercises in Electrophoresis, Hoefer Scientific Instruments, San Francisco, CA) and aged for 24 h. The gels were then pre-
run in the presence of the SDS-PAGE running buffer containing 0.1 mM sodium thioglycollate (Sigma Chemical Company, St. Louis, MO) for 30 min at 30 mA before loading the protein samples. Whole cell lysate samples of IPTG-induced BoxAVP1 in SDS gel-loading buffer were heated in a boiling water bath for 90 sec, centrifuged at 12,000 X g for 2 min and loaded at 50 µg of total protein onto individual lanes. Electrophoresis was carried out at 10 mA overnight. One gel was Coomassie Blue-stained while the other was subjected to electroblotting onto a PVDF membrane in a procedure similar to the one described already for Western blotting.

For electroblotting, the SDS-PAGE gel was pre-equilibrated for 15 min in CAPS transfer buffer [10 mM CAPS (3-[Cyclohexylamino]-1-propanesulfonic acid); Sigma Chemical Company, St. Louis, MO), 10% methanol, pH 11.0] and placed on a 14 X 14 cm piece of WESTRAN™ PVDF protein transfer and sequencing membrane (Schleicher & Schuell, Keene, NH) that had been briefly immersed in 100% methanol (2-3 sec), rinsed with distilled water (2-3 sec) and also pre-equilibrated for 15 min in CAPS transfer buffer. Next, the gel and the membrane were sandwiched together with 3MM paper sheets and filter pads as done before for Western blotting with nitrocellulose membranes. The assembled sandwich was then placed into a hinged blotting cassette and electroblotted using a Transphor Tank Electro-Transfer Unit model TE-50 (Hoefer Scientific Instruments, San Francisco, CA) at a constant current of 0.5 A for 1 h.

Following transfer, the membrane was rinsed several times with sterile double-distilled deionized water. Protein bands were visualized either by transillumination by rewetting the blot in 20% methanol and was placed on a white-light illumination box or
by reversible staining with 0.5% Ponceau Red (Sigma Chemical Company, St. Louis, MO) in 1% acetic acid. The protein bands of interest were identified by comparing to molecular weight standards included on the gel, excised with a clean scalpel, cut into 2-3 mm² fragments, and placed into sterile 0.5 ml microfuge tubes. The samples were sent to Dr. Richard G. Cook, Director, Protein Chemistry Core Facility, Baylor College of Medicine, Houston, TX and a limited (20 cycles) N-terminal Amino acid sequencing was performed subsequently on 150 pmoles of the recombinant protein.

**BLAST and BLITZ Protein Sequence Homology Searches:** Protein sequence information obtained by the limited N-terminal aminoacid sequencing was compared with known protein sequences by performing online, non-redundant database searches using a Gateway 2000 P5-75 personal computer (Gateway 2000, Inc., N. Sioux City, SD). The various protein and nucleotide sequence databases for BLAST (Basic Local Alignment Search Tool) search are accessible on the internet at the worldwide website "http://www.ncbi.nlm.nih.gov/BLAST" by the National Center for Biotechnology Information, Bethesda, MD. Similarly, the BLITZ-Version 1.5 search tool for comparison of the sequence in question against the latest version of the Swiss-Prot Protein Sequence Database, is available and online at the web address "http://www.ebi.ac.uk/searches/blitz.html". BLITZ is an automatic electronic mail server for the Mpsrch program of Shane Sturrock and John Collins, Biocomputing Research Unit, University of Edinburgh, UK.
RESULTS

A positive strand PCR primer for amplification of HAV VP1 sequence which also contained the \textit{E.coli} \textit{rrn} operon anti-terminator BoxA sequence upstream of the translational start codon for VP1 was designed and used successfully to amplify insert BoxA-VP1 DNA from plasmid, pGEM-HAV. The PCR amplified insert DNA was then cloned into pAX and \textit{E.coli} cells carrying recombinant pAX-BoxAVP1 plasmid were subjected to PCR verification for the presence of BoxA-VP1 insert sequence prior to IPTG induction (Figure 17). Next, SDS-PAGE analysis of IPTG-induced BoxA-VP1 clones revealed the appearance of a new, 33-kD protein band following induction of BoxA-VP1 clones (Figure 18). However, in stark contrast to the earlier induction experiments with VP1II clones without the BoxA modification of insert VP1 sequence, the fusion protein band (expected size \textasciitilde 150 kD) was not seen. Western Blot analysis of whole cell lysates of IPTG-induced BoxA-VP1 showed the 33-kD protein reacted weakly or failed to react with any of the anti-HAV antibodies (rabbit, mouse or human sources) or with rabbit polyvalent anti-\(\beta\)-galactosidase antibody. The 125-kD protein band corresponding to non-fusion form of \(\beta\)-galactosidase continued to persist in the IPTG-induced BoxA-VP1 lysates.

\textit{N-terminal} amino acid sequence analysis performed at the Protein Sequencing Facility, Baylor College of Medicine, revealed that the 33-kD protein possessed the sequence \(\text{NH}_4^+\text{MRLASRGYNSIHREMPT-COO}^-\) which did not correspond to any stretch in either \textit{E.coli} \(\beta\)-galactosidase or HAV VP1 protein sequences. Therefore, the \textit{N-terminal} sequence of the 33-kD protein was analyzed further by comparison against
Figure 17. Agarose gel analysis of HAV VP1 insert DNA amplified by PCR (i) from parent plasmid pGEM-HAV prior to cloning and (ii) from recombinant plasmids pAXBoxAVP1 and pAXVP1I1 following successful cloning and transformation of E. coli UT2300 cells. Lane 1: BioMarker low molecular weight DNA standard, Lane 2: BoxAVP1 DNA amplified from pGEM-HAV template, Lane 3: Insert DNA amplified from pAXBoxAVP1 template, Lane 4: VP1 DNA (without BoxA sequence) amplified from pGEM-HAV template, and Lane 5: Insert DNA amplified from pAXVP1I1 template. The samples in lanes 4 and 5 were prepared from positive control reactions.
Figure 18. SDS-PAGE analysis of fusion proteins expressed from recombinant plasmid pAX4b-BoxAVP1. Whole cell lysates of recombinant host *E. coli* UT2300 cells from various time points before and after IPTG induction were examined. Lane 1: Host cells without any vector. Lane 2: Biorad molecular weight marker. Lane 3: Cells under glucose repression. Lane 4: Cells at 1 mM IPTG induction (T₀). Lane 5: 1 h after induction. Lane 6: 2 h. Lane 7: 3 h. Lane 8: 4 h. Lane 9: Rainbow™ molecular weight marker. Lane 10: Host cells carrying vector pAX4b+.
sequence data for a total of 208,180 sequences through a computerized search of protein sequence databases [Genebank, PDB, SwissProt, SPupdate and PIR using the basic BLAST algorithm (Altschul, et al., 1990) and the BLITZ/MPsrch program version 1.5 (Sturrock and Collins, 1993). Homologous protein sequences obtained following the database search are listed in Figure 19. The sequence of the 33 kD protein displayed a 97% identity with the N-terminus of a highly conserved open reading frame (O273) that lies within the leading region adjacent to the origin of transfer (OriT) of conjugative Plasmid F (Loh, et al., 1989), a 97% homology with a hypothetical 32 kD protein sequence (gene YFJQ) encoded by the E.coli ALPA-GABD intergenic region (Plunkett, 1995), a 89% homology with the N-terminus of a hypothetical 32 kD protein sequence from the leading region of the E.coli resistance plasmid R1(Graus, et al., 1990), and a 78% homology with protein derived from an open reading frame, ORF298, within the Rhs elements of E.coli K12 (Zhao, et al., 1993; Takemoto, et al., 1996). The Rhs elements are accessory genetic elements that share extensive homology with one another and are present in many but not all wild-type E.coli isolates. In addition, the N-terminal sequence of interest displayed varying degrees of homology (Table 4 & 5) with short motifs from a diverse set of previously documented protein sequences such as E.coli ADP-ribosyltransferase (Fitzmaurice, et al., 1989), Bordatella pertussis Outer membrane usher protein FIMC precursor (Willems, et al., 1992), Sindbis virus structural polyprotein (Strauss, et al., 1984), a hypothetical 31.4 KD protein encoded by E.coli gene f268b in the TSR-MDOB intergenic region (Burland, et al., 1995), an inclusion body matrix protein (viroplasmin) from Cauliflower Mosaic Virus (Stratford, R., et al.,
Figure 19. Comparison of sequences in search for homology of the N-terminal sequence of the 33-kD protein expressed in recombinant *E. coli* BoxA-VP1 cells, with other protein sequences retrieved from protein databases using the basic BLAST or BLITZ algorithms. The actual position occupied by each amino acid residue in a given protein sequence can be inferred by referring to the Indo-Arabic numerals preceding and following each sequence. The mismatched aminoacid residues are depicted in bold. A gap in a given sequence (necessitated only for the sake of comparison, not present in the actual protein) is shown by (-). Sequences listed and their original source are as follows: (i) the N-terminal sequence of the 33 kD protein/recombinant *E. coli* BoxA-VP1 clones, (ii) Hypothetical protein 273 (OriT 5' region)/*E. coli* Conjugative Plasmid F, (iii) Hypothetical Gene YFJQ Protein from the ALPA-GABD intergenic region/*E. coli*, (iv) Hypothetical 32K Protein/*E. coli* Plasmid R1, (v) Unknown Reading Frame/*E. coli*, (vi) ADP-Ribosyltransferase/Rhodospirillum rubrum, (vii) Outer Membrane Usher Protein FIMC Precursor/Bordetella pertussis, (viii) Sindbis Virus Structural Polyprotein/Alphaviruses; Togaviridae; ss-+ RNA enveloped viruses, (ix) Hypothetical 31.4 KD Protein in TSR-MDOB Intergenic Region (F268B)/*E. coli*, (x) Inclusion Body Matrix Protein (Viroplasmin)/Cauliflower Mosaic Virus; Caulimoviridae; dsDNA non-enveloped viruses, and (xi) Nuclear DNA-binding Protein/dead ringer gene product/Drosophila melanogaster.
Table 4. Analysis of N-terminal Sequence of the 33 kD Recombinant Protein from *E. coli* AXBoxAVP1, by Alignment Against Similar Sequences in Protein Databases

<table>
<thead>
<tr>
<th>Description/Source</th>
<th>Length</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hypothetical Protein 273 (OriT 5' region)/<em>E. coli</em> Plasmid F</td>
<td>273</td>
<td>Db 1 MRLAS RFGRY NSIHR ERPLT 20; Qy 1 MRLAS RFGRY NSIHR ERPLT 20</td>
</tr>
<tr>
<td>2 Hypothetical Gene YFJQ Protein from the ALPA-GABD intergenic region/<em>E. coli</em></td>
<td>273</td>
<td>Db 38 MRLAS RFGRY NSIHR ERPLT 57; Qy 1 MRLAS RFGRY NSIHR ERPLT 20</td>
</tr>
<tr>
<td>3 Hypothetical 32K Protein/<em>E. coli</em> Plasmid R1</td>
<td>273</td>
<td>Db 1 MRLAS RFGRY NSIHR ERPLT 20; Qy 1 MRLAS RFGRY NSIHR ERPLT 20</td>
</tr>
<tr>
<td>4 Open Reading Frame within <em>Rhs</em> elements/<em>E. coli</em></td>
<td>425</td>
<td>Db 155 RLAS RFGAA NLIRR DRPLT 173; Qy 2 RLAS RFGRY NSIHR ERPLT 20</td>
</tr>
<tr>
<td>5 ADP-Ribosyltransferase/<em>Rhodospirillum rubrum</em></td>
<td>276</td>
<td>Db 123 S RFGLF PTFHR E- PIT 137; Qy 4 S RFGRY NSIHR ERPLT 20</td>
</tr>
<tr>
<td>6 Outer Membrane Usher Protein FIMC Precursor/<em>Bordetella pertusis</em></td>
<td>873</td>
<td>Db 490 RFGRY NSLGR ERP 502; Qy 5 RFGRY NSIHR ERP 18</td>
</tr>
<tr>
<td>7 Sindbis Virus Structural Polyprotein/<em>Alphavirus</em>; <em>Togavirus</em>; ss+RNA enveloped viruses</td>
<td>1245</td>
<td>Db 485 RLKE TTAGY ITMHR PRP 501; Qy 2 RLAS RFGRY NSIHR ERP 18</td>
</tr>
<tr>
<td>8 Hypothetical 31.4 KD Protein in TSR-MDOB Intergenic Region (F268B)/<em>E. coli</em></td>
<td>268</td>
<td>Db 96 KFGRY NLIHS EK 107; Qy 6 RFGRY NSIHR ER 17</td>
</tr>
<tr>
<td>9 Inclusion Body Matrix Protein (Viroplasmin)/<em>Cauliflower Mosaic Virus</em>; <em>Caulimoviridae</em>; dsDNA non-enveloped viruses</td>
<td>522</td>
<td>Db 37 GG- NSLHR ETPV 47; Qy 8 GGY NSIHR ERPL 19</td>
</tr>
<tr>
<td>10 Nuclear DNA-binding Protein/dead ringer gene product/<em>Drosophila melanogaster</em></td>
<td>901</td>
<td>Db 395 S SYGQY EAMHN QMPMT 410; Qy 5 S RFGRY NSLHR E RPL T 20</td>
</tr>
</tbody>
</table>

**ABBREVIATIONS:** Db = Sequence from Protein Sequence Database; Qy = Query Sequence derived from the 33 kD recombinant protein; + = Perfect Match; O = Partial match or Conserved Substitution.
Table 5. Results of Sequence Homology Search for N-terminal Sequence of the 33 kD Protein derived from *E.coli* BoxAVP1 Clones with Protein Sequence Databases.

<table>
<thead>
<tr>
<th>Description</th>
<th>Length</th>
<th>Residue Matches</th>
<th>Partial Matches</th>
<th>Mismatches</th>
<th>% Match</th>
<th>Predicted Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothetical Protein 273 (OriT 5' region)- <em>E.coli</em> Plasmid F</td>
<td>273</td>
<td>19/20</td>
<td>0/20</td>
<td>1/20</td>
<td>97</td>
<td>4.07e-20</td>
</tr>
<tr>
<td>Hypothetical Gene YFJQ Protein/<em>E.coli</em></td>
<td>310</td>
<td>19/20</td>
<td>0/20</td>
<td>1/20</td>
<td>97</td>
<td>4.07e-20</td>
</tr>
<tr>
<td>Hypothetical 32K Protein-<em>E.coli</em> Plasmid R1</td>
<td>273</td>
<td>18/20</td>
<td>0/20</td>
<td>2/20</td>
<td>89</td>
<td>5.40e-17</td>
</tr>
<tr>
<td>Open Reading Frame within Rhs elements (<em>E.coli</em>)</td>
<td>425</td>
<td>14/19</td>
<td>1/19</td>
<td>4/19</td>
<td>78</td>
<td>3.54e-09</td>
</tr>
<tr>
<td>ADP-Ribosyltransferase</td>
<td>276</td>
<td>9/16</td>
<td>3/16</td>
<td>3/16</td>
<td>41</td>
<td>4.60e-01</td>
</tr>
<tr>
<td>Outer Membrane Usher Protein FIMC Precursor</td>
<td>873</td>
<td>8/13</td>
<td>2/13</td>
<td>3/13</td>
<td>39</td>
<td>1.66e+00</td>
</tr>
<tr>
<td>Sindbis Virus Structural Polyprotein</td>
<td>1245</td>
<td>8/17</td>
<td>2/17</td>
<td>7/17</td>
<td>39</td>
<td>1.66e+00</td>
</tr>
<tr>
<td>Hypothetical 31.4 KD Protein in TSR-MDOB Intergenic Region (F268B)-<em>E.coli</em></td>
<td>268</td>
<td>7/12</td>
<td>2/12</td>
<td>3/12</td>
<td>39</td>
<td>1.66e+00</td>
</tr>
<tr>
<td>Inclusion Body Matrix Protein (Viroplasmin)</td>
<td>522</td>
<td>8/11</td>
<td>2/11</td>
<td>1/11</td>
<td>39</td>
<td>1.66e+00</td>
</tr>
<tr>
<td><em>D.melanogaster</em> Nuclear DNA-binding Protein/dead ringer gene product</td>
<td>901</td>
<td>6/16</td>
<td>6/16</td>
<td>4/16</td>
<td>38</td>
<td>3.09e+00</td>
</tr>
</tbody>
</table>
1988), and a *Drosophila melanogaster* nuclear DNA-binding protein/dead ringer gene product (Gregory, *et al.*, 1996). A non-recombinant expression in *E. coli* of a protein corresponding to these sequences has not been documented previously.

**DISCUSSION**

Earlier attempts at β-galactosidase-HAV VP fusion protein expression and their affinity-based purification were frustrated by the simultaneous appearance of β-galactosidase, apparently without the viral protein components. Further, the anti-HAV antibodies displayed cross reactivity with native *E.coli* β-galactosidase when used to verify antigenicity of the recombinant HAV proteins by Western Blotting and ELISA. Subsequent attempts to express the HAV capsid proteins in the non-fusion form from various prokaryotic vectors were unsuccessful and found to induce toxic effects in host *E.coli* cells. Therefore, various strategies were considered to improve the efficiency of fusion protein expression from recombinant pAX vectors in *E.coli* and efforts were initiated simultaneously to identify suitable vector-host *E.coli* systems for expression of HAV capsid proteins in a non-fusion form.

Premature transcriptional termination of recombinant mRNA synthesis was one explanation considered which could lead to expression of the non-fusion form of β-galactosidase in addition to the fusion proteins. Hence, the hypothesis that incorporation of the *E.coli* *rrn* operon antiterminator sequence, BoxA, upstream of the HAV VP1 coding sequence, could lead to improved transcription of full length messages for fusion protein and render overall expression of fusion protein more efficient, was tested. If
successful, this approach would possibly minimize expression of the non-recombinant β-galactosidase protein encountered previously.

However, contrary to expectations, induction of fusion protein expression from recombinant *E. coli* carrying plasmid pAXBoxAVP1 resulted in overexpression of a new ~33 kD protein species and no expression of the expected ~150 kD β-galactosidase-HAVVP1 fusion protein. Host cells grown in the presence of glucose and prior to addition of IPTG did not display either ~125 kD β-galactosidase or the 33 kD proteins in detectable levels when analyzed by SDS-PAGE. The appearance of both protein bands on SDS-PAGE, following IPTG induction, strongly suggested the possibility of temporally synchronized fusion-protein expression and post-translational processing events *in vivo*. Host *E. coli* proteases could have cleaved the fusion protein into its constituent ~125 kD β-galactosidase and the 33 kD HAV VP1 components. However, subsequent Western Blot studies were inconclusive and a limited N-terminal aminoacid sequencing of the 33 kD protein yielded sequence information that did not corroborate either the cloned HAV sequence or any putative protein-encoding open reading frame (ORF) present on the plasmid pAX4b+ vector. Rather, amino acid sequence of the 33 kD protein, upon a computerized search of protein sequence databases, showed varying degrees of homology with a spectrum of protein sequences from diverse sources (Figure 19).

Nineteen of the possible twenty amino acid residues in the N-terminal region of the 33 kD protein displayed exact matches with the N-terminal protein sequence derived from a putative open reading frame, ORF273, within the origin of transfer
(OriT) or the leading region of the *E.coli* K12 conjugative F plasmid (Loh, *et al*., 1989). The leading region, which is that portion of plasmid DNA that is transferred first to the recipient cell during conjugation, including ORF273 shares extensive sequence homology with the leading regions of related plasmids such as the IncFII R group of plasmids R1, R6-5 and R100 (Sharp, *et al*., 1973) as well as a wide range of unrelated, non-F-like conjugative plasmids belonging to various Incompatibility (Inc) groups (Bergquist, 1988; Bukhari, *et al*., 1977). The well-conserved homology of the leading region may be relevant in conjugal DNA transfer and plasmid maintenance. Analysis of the putative ORF273 suggests that upon transcription and translation, this ORF could encode a mainly hydrophilic polypeptide of ~32 kD residing in the cytoplasm. In addition, the putative ORF273 appears to share limited homology, mostly within the N-terminal region, with several prokaryotic DNA-binding proteins involved in replication functions such as the *E.coli* single-stranded DNA-binding (SSB) protein and the F plasmid SSB (Chase, *et al*., 1983). SSB activity is essential in protecting DNA from nucleases during conjugal plasmid DNA transfer and DNA replication.

In addition to a shared homology with the F plasmid ORF273 derived protein sequence, the N-terminal sequence of the 33 kD protein of interest showed significant homology with proteins encoded by *E.coli* chromosomal DNA and *E.coli* plasmid R1. The sequence also exhibited limited homology with a diverse set of proteins, both procaryotic and eukaryotic in origin, and with potential for binding nucleotides or DNA (Table 4 & 5).
The reasons for apparent lack of expression of HAV VP1 component of the fusion protein and for the evident expression of a 33 kD protein with putative DNA-binding activity in *E.coli* AXBoxAVP1 clones remain unclear. In this context, we have to note that the antiterminator BoxA sequence is present within an untranslated leader region in the *E.coli* *rrn* operon. In contrast, in our construct of the recombinant plasmid, pAXBoxAVP1, the antiterminator sequence was placed within a open reading frame contiguous to the hinge region and following the β-galactosidase gene but upstream of the HAV VP1 gene. Also, caution was exercised during PCR amplification and preparation of insert VP1 DNA to ensure that the BoxA sequence would itself remain in the appropriate reading frame without causing any disruption in the ability to translate the downstream VP1 sequence. The boxA sequence would merely lengthen the recombinant protein in its hinge region by four amino acid residues which was assumed to be innocuous. The hinge region of the recombinant protein, which is designed to keep the β-galactosidase and the foreign protein components of the fusion protein discrete, is derived from collagen and contains recognition sites for collagenase, endoproteinase Xa and similar proteases (Markmeyer, *et al.*, 1990). It was possible that a full length fusion protein was expressed in the host cells subsequent to IPTG-induction, but the VP1 component was degraded soon after due to host proteolytic activity, leaving behind the ~125 kD β-galactosidase. The initial cleavage may have been aided by extension of the rigid hinge region of the expressed recombinant protein due to additional amino acid residues contributed by the boxA sequence. Also, another possibility is that expression of viral protein may have been toxic to host *E.coli* cells.
Another possibility is the misincorporation of nucleotides during the preparation of HAV VP1 insert DNA by PCR. The use of the AmpliTaq DNA Polymerase, which lacks the 3'→5' proofreading exonuclease activity, in PCR may be accompanied by incorporation of missense and nonsense errors within protein coding sequences of cloned DNA inserts at a rate of $10^{-4}$ per base per cycle (Lundberg, et al., 1991; Barnes, 1994). Such errors may lead to premature termination of translation or synthesis of proteins with altered amino acid sequences resulting in abnormal, truncated proteins or peptides. Also, a rapid, overexpression of recombinant proteins, whether normal or abnormal, may not allow for proper folding into their native, stable conformations. Instead, the improperly folded proteins may begin to aggregate, become sequestered into inclusion bodies or begin to precipitate in vivo. Such events can potentially trigger the SOS or heat shock responses in host E.coli cells and induce host proteases which degrade the recombinant proteins (Gottesman, 1984; Goff and Goldberg, 1985; Meyer, et al., 1985; Goldberg and Goff, 1986). One fallout of the SOS response could have been the induction of the 33 kD protein of interest although its purpose during such events remains to be defined.

The gene responsible for expression of the 33 kD protein in the recombinant AXBoxAVP1 cells remains to be identified. Plasmids pAX4b+ or pGEM-HAV were ruled out as possible candidates by a careful study of their nucleotide sequences. The genotype of host E.coli UT2300 cells, to the best of our knowledge, does not indicate
the presence or absence of any F or F-related plasmids which could have been potential sources (Elish, et al., 1987). Still, the presence of cryptic F or related plasmids cannot be ruled out. In addition, there is a much closer correspondence of the exact positions of individual amino acid residues (the first twenty residues from the N-terminus for each) between the sequence of the protein in question and those derived from plasmids F and R1, as versus the sequences derived from E. coli chromosomal DNA such as the gene YFJQ encoded hypothetical protein or the hypothetical protein encoded by ORF298 within the Rhs elements of E. coli K12. Hence, it is very likely that an unreported F or similar plasmid, present within the host cells as an episome, is the likely origin for the 33 kD protein seen in E. coli AXBoxAVP1 cells. However, the reasons for overexpression of such a protein upon IPTG induction of host cells to express fusion proteins still remain obscure. It is interesting to note that no such protein was evident upon SDS-PAGE analysis when we induced overexpression of HAV VP1, VP2 and VP3 as β-galactosidase fusions without additional residues owing to the antiterminator BoxA sequence (Chapter I, this dissertation) and that a non-recombinant overexpression of the ORF273 protein, even with F plasmid containing host cells, has not been hitherto reported.

In conclusion, our attempts to improve the expression of β-galactosidase-HAV VP1 fusion protein through the inclusion of an E. coli antiterminator sequence upstream of the viral DNA insert, although unsuccessful, revealed a rather novel phenomenon with an overexpression of a putative single-stranded DNA-binding protein, the reasons for which remain unclear. The expression of the 33 kD protein with putative DNA-binding
properties, although much needs to be learned about its exact nature and functions, may
have been triggered as part of a general stress response (such as akin to the heat shock
response) to overexpression of seemingly toxic proteins in the host cells. Such a protein
may be essential for replication or maintainence of the relevant plasmids that it is derived
from, while within host cells that are undergoing a stress-induced, active SOS response.
CHAPTER III.

ALTERNATIVE APPROACHES FOR CLONING OF HAV CAPSID GENES
FOR EXPRESSION IN Escherichia coli

87

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
INTRODUCTION

Recombinant DNA technology has permitted production of high-levels of various proteins in both prokaryotic and eukaryotic systems, thereby making available many proteins whose supply from natural sources were previously limited. *Escherichia coli* still remains the leading host organism for expression of heterologous proteins due to several reasons such as, a short generation time of cells, low-cost culture conditions essential for growth and ease of manipulation by simpler techniques as compared to the more time-consuming and expensive methods like tissue culture. Proteins over-expressed and purified from *E.coli* can be subsequently used, (i) to study their properties such as structure, biological activity, antigenic and immunological potential; (ii) to produce antibody reagents useful in research and diagnostics; (iii) as replacements in therapy of human diseases, for example, recombinant insulin for diabetes mellitus, and human somatostatin (hGH) for growth hormone deficiency; and (iv) in the development of safer, subunit vaccines (Itakura, *et al.*, 1977; Goeddel, *et al.*, 1979a; Goeddel, *et al.*, 1979b). However, for an overproduction of a heterologous protein in *E.coli* following a successful cloning of the gene encoding the protein of interest, several crucial events have to occur in an optimal fashion. These include an efficient transcription and translation of the cloned gene, and stabilization of the resulting mRNA and the protein product. In addition, the expressed protein should not be very toxic to the host cells. In this context, each gene can present its own unique expression problems. Despite standard cloning procedures having become routine and a large variety of host/vector systems being available, no single set of procedures is currently available that can
guarantee successful production of every protein in an useful format. Hence, in the absence of a universal solution when difficulties in expression of recombinant proteins are encountered, an investigator has to apply a host of strategies, some novel and some already well-documented, in seeking appropriate redressal of the individual and specific expression problems. (Goeddel, 1990; Das, 1990).

In our efforts to express HAV capsid proteins as *E. coli* β-galactosidase fusions, we discovered problems such as an apparent co-expression of recombinant fusion protein and the non-recombinant β-galactosidase, a cross-reactivity of anti-HAV antibodies with native β-galactosidase and potential toxic effects induced in host cells by the recombinant proteins or their overexpression (Chapter 1, this dissertation). Our subsequent attempt to improve the transcription of the recombinant genes through introduction of an *E. coli* antiterminator sequence also was not successful (Chapter 2, this dissertation). Hence, in continuing to pursue our primary objectives of cloning and expression of HAV capsid proteins in *E. coli*, we decided to explore alternatives such as new cloning vectors and host strains. In addition, we refocused our studies toward expression of HAV capsid proteins in their native forms and as chimeric proteins displayed on the surface of host *E. coli* cells and on the surface of filamentous coliphages that infect and replicate in *E. coli*.

Plasmid pT7-7 is a prokaryotic expression vector that allows cloning of genes for high-level expression and contains the bacteriophage T7 RNA polymerase promoter ϕ10 upstream of a multiple cloning site (Figure 20). The T7 promoter sequence is rarely encountered in DNA unrelated to that from phage T7 and the T7 RNA polymerase is
Figure 20. A schematic diagram of plasmid vector pT7-7 used in cloning HAV capsid genes for expression in a non-fusion format. Sequence of the cloning region containing the phage promoter, ribosome binding site, translational start codon (ATG) and the multiple cloning site are shown underneath the plasmid diagram. Pφ10=Phage T7 gene10 promoter which is recognized by T7 RNA Polymerase only; RBS=Ribosome Binding Site; MCS=Multiple Cloning Site which includes EcoRI and Clal/Bsp106I restriction enzyme sites; bla=beta-lactamase gene for antibiotic selection of recombinants; colEl Ori=Origin of plasmid DNA replication.
uniquely and extremely specific for this particular promoter (Dunn and Studier, 1983). Since foreign genes cloned under the transcriptional control of a T7 promoter will not be expressed until the T7 RNA polymerase is made available, the pT7-7 vector is suited well for a selective cloning and amplification of genes whose product may be potentially toxic to the host cells, such as those encoding HAV capsid proteins, and for their subsequent controlled exclusive expression (Tabor and Richardson, 1985). The T7 RNA polymerase exhibits at least a five-fold higher processivity (≈ 300 ribonucleotides per second) compared to that of *E. coli* RNA polymerase (≈ 50 bases per second). The phage transcribes DNA independent of translation, unlike its *E. coli* counterpart, thereby making the T7 expression system ideal for genes with a potential preponderance of low-usage codons. Transcription of the cloned genes from recombinant pT7-7 may occur relentlessly, if the T7 RNA polymerase is introduced into the host cells carrying the plasmid, and target RNA can accumulate to levels comparable to that of ribosomal RNA (Studier, et al., 1990).

*E. coli* B strain BL-21 are host cells deficient in certain proteases such as the *lon* protease and *ompT* outer membrane protease (Studier and Moffat, 1986). Both of these proteases can degrade foreign proteins expressed in *E. coli*, and potentially jeopardize expression and purification procedures. BL-21 (DE3)pLysS are BL-21 derivatives that carry a plasmid, pLysS, and are lysogenic for a lambda bacteriophage DE3. Bacteriophage DE3 is a lambda phage derivative that carries the phage 21 immunity region, the *lacI* gene and the *lac*UV5-driven T7 RNA polymerase expression cassette (Studier and Moffatt, 1986). Plasmid pLysS, derived from pACYC184, carries an
expression cassette for a low level expression of phage T7 lysozyme which binds and inhibits T7 RNA polymerase if a premature, low-level "leaky" production of the polymerase were to occur (Chang and Cohen, 1978). Thus, a tighter control of expression is exerted until expression of T7 RNA polymerase can be induced by addition of IPTG (Doherty, et al., 1994). Recombinant pT7 derived plasmids carrying certain toxic target genes reportedly are tolerated better and unstable plasmids become stable in host strains carrying the gene for T7 lysozyme (Studier, et al., 1990).

The SurfZAP™ vector is a bacteriophage vector derived from the Lambda ZAP® II vector which is suited for cloning and expression of foreign protein sequences on the surface of single-stranded filamentous (f1 or M13) phage as fusions with the aminoacids 198-406 (membrane-anchoring domain) of its minor coat protein, cpIII (Short, et al., 1988; Parmley, and Smith, 1988; Amberg, et al., 1993; Smith and Scott, 1993). Cloning into bacteriophage lambda derived vectors with the subsequent packaging and infection of host E.coli allows for a rapid and easy amplification of recombinant DNA, including those containing genes for potentially toxic proteins, at an efficiency that exceeds that of plasmid transformation by 100-fold. The cloning site for foreign DNA insertion is contained within a phagemid, pSurfscript™ SK (-), which itself is an integral part of the larger λSurfZAP™ vector (Figure 22). Since the pSurfscript is bound by the initiation (I) and termination (T) sequences of the phage f1 origin of replication, the entire phagemid (3.64 kb in size without any insert DNA) can be excised in vivo following infection of the host E.coli with helper f1 or M13 phage (Dotto, et al., 1984; Short and Sorge, 1992). Subsequently, the excised phagemid containing phage f1
\[ \text{\textcopyright{} SurfZAP™ vector (~45 kb) which includes an excisable phagemid, pSurfscript™.} \]

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure21.png}
\caption{A schematic diagram representing relevant regions of SurfZAP™ (Stratagene Corporation, La Jolla, CA) vector and its derivative, pSurfscript™ employed in cloning for surface expression in \textit{E. coli} and chimeric coliphages. Amp=Ampicillin Resistance marker; lacZ=Cloning region containing \textit{NotI} and \textit{SpeI} restriction enzyme sites for target foreign DNA insertion; \textit{colEl Ori} = Origin of plasmid DNA replication; and \textit{f1 Ori} = Origin of phage \textit{f1} replication.}
\end{figure}
or M13 origin of replication, bacterial colE1 origin, and ampicillin-resistance marker (Amp') can be propagated either as double-stranded plasmid DNA within appropriate host *E.coli* grown in the presence of ampicillin, or as single-stranded phagemid DNA that is packaged into M13 particles in the presence of an infecting M13 helper phage.

Recombinant protein expression from the λSurfZAP™ or its pSurfscript derivative is under the transcriptional control of the *lacZ* promoter. Following induction with IPTG, the expressed fusion protein, containing the pelB leader sequence and the membrane anchoring domain of cpIII, is directed for embedding into the inner membrane of the host *E.coli* with the foreign protein sequence oriented in the periplasmic space (Boeke and Model, 1982; Skerra and Pluckthun, 1988). If a helper phage is present within the host cells, the recombinant protein is utilized in packaging single-stranded phagemid DNA resulting in chimeric phagemid particles which display fusion protein of interest on their surface and carry the cloned insert DNA sequence within. Thus, this vector has been used previously to express on the surface of filamentous phage, a heterogenous group of peptide and protein molecules including Staphylococcal protein A, human growth hormone, bovine pancreatic trypsin inhibitor, alkaline phosphatase and ricin B, and to clone from large random libraries of peptides and antibodies (Smith, 1985; Huse, *et al.*, 1989; Bass, *et al.*, 1990; Cwirla, *et al.*, 1990; McCafferty, *et al.*, 1990; Pluckthun, 1990; Hogrefe, *et al.*, 1993; Kushwaha, *et al.*, 1994).

We report here our studies on cloning and expression strategies for HAV capsid protein production in *E.coli*, either as non-fusion proteins from a T7 expression system, or as chimeric proteins from a bacteriophage-derived surface-display system.
MATERIALS AND METHODS

Vectors, Phage and Bacterial Host Strains: The SurfZAP™ cloning kit which contains a phage λ derived vector, relevant bacterial host strains and f1 helper phage was purchased from Stratagene, La Jolla, CA (Table 6). The bacterial host strain which carries the pT7-7 plasmid was kindly provided by Dr. Stanley Tabor, Dept. of Biol. Chem., Harvard Medical School, Boston, MA (Table 7).

Insert Design for Cloning into SurfZAP™ and pT7-7 Vectors: Oligonucleotide primers were designed for PCR amplification of selected HAV capsid genes and to facilitate the cloning of the resultant DNA into the SurfZAP or pT7-7 vectors. These primer sequences are listed in Figures 22 and 23. The SurfZAP primers would result in DNA which contained a 5' NotI site and a 3' SpeI site. The need to accommodate a portion of the pelB DNA leader sequence downstream of the NotI site during PCR amplification of insert DNA for cloning into SurfZAP vector mandated designing two partially overlapping 5' (or positive strand) primers. The features of insert DNA that are essential for a successful cloning are shown schematically in Figure 24. Therefore, amplification of insert DNA for SurfZAP cloning by PCR required two separate PCR reactions with a specific 5' primer for each and in a defined sequence (i.e. SurfVP1PS1 primer in Reaction 1 and SurfVP1PS2 primer in Reaction 2). Also, different sources of template DNA (pGEM-HAV in Reaction 1 and SurfVP1 DNA from Reaction 1 as template in Reaction 2) were used. However, a single PCR reaction with pGEM-HAV as template and with primers for PT7-7 cloning would result in insert DNA with a 5' EcoRI site and a 3' ClaI/Bsp1061 site. The primers listed in Figures 22 and 23 were
Table 6. Bacterial strains, phage and plasmid vectors for surface expression of HAV capsid proteins

<table>
<thead>
<tr>
<th>Strains/helper phage/vectors</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> UT2300</td>
<td>azi-6 fhuA23 lacY1 leu-6 mlt-1 proC14 purE42 rpsL109 thi-1 trpE38 tss-67 entA403 purE* fhuA* fepA Λlon::Tn10</td>
<td>Elish, et al., 1988</td>
</tr>
<tr>
<td><em>Escherichia coli</em> XL-1 Blue</td>
<td>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacFZΔm15 Tn10 (Tet')]c</td>
<td>Bullock, et al., 1987</td>
</tr>
<tr>
<td><em>Escherichia coli</em> SOLR™</td>
<td>el4(McrA') Δ(mcrCB-hsdSMR-mrr)171 sbcC recB recJ uvrC umuC::Tn5 (Kan') lac gyrA96 relA1 thi-1 endA1λκ [F' proAB lacIqZΔM15]c Strw (nonsuppressing)</td>
<td>Stratagene, La Jolla, CA</td>
</tr>
<tr>
<td><em>Escherichia coli</em> BL21</td>
<td><em>E. coli</em> B F- dcm ompT hsdS(rB, mB) gal</td>
<td>Weiner, et al., 1994</td>
</tr>
<tr>
<td><strong>1 Helper Phage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExAssist™ interference-resistant helper phage</td>
<td>for in vivo excision of pSurfscript phagemid from λSurfZAP™ vector; single-strand size ~ 5 kb</td>
<td>Stratagene, La Jolla, CA</td>
</tr>
<tr>
<td><strong>Phage/Plasmid Vector and Template DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λSurfZAP™</td>
<td>a lambda phage derived vector for surface expression of cloned genes in <em>E. coli</em></td>
<td>Stratagene, La Jolla, CA</td>
</tr>
<tr>
<td>pGEM-HAV</td>
<td>plasmid carrying HAV DNA including the capsid genes; used as template in PCR to generate insert HAV DNA</td>
<td>R.H. Purcell (N.I.H.)</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 7. Bacterial strains and plasmid vectors for phage T7 RNA polymerase-dependent expression of HAV capsid proteins

<table>
<thead>
<tr>
<th>Strains/helper phage/vectors</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> BL21</td>
<td><em>E. coli</em> B F- <em>dcm ompT hsdS</em>(rB, mB) <em>gal</em></td>
<td>Weiner, <em>et al.</em>, 1994</td>
</tr>
<tr>
<td><em>Escherichia coli</em> BL21(DE3) pLysS</td>
<td><em>E. coli</em> B F- <em>dcm ompT hsdS</em>(rB, mB) <em>gal</em> λ(DE3) [pLysS Cam']</td>
<td>Studier, <em>et al.</em>, 1990</td>
</tr>
<tr>
<td><strong>Plasmid Vector and Template DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT7-7</td>
<td>plasmid vector for T7 RNA polymerase-dependent expression of cloned genes</td>
<td>S. Tabor, Harvard Medical School, Boston, MA</td>
</tr>
<tr>
<td>pGEM-HAV</td>
<td>plasmid carrying HAV DNA including the capsid genes; used as template in PCR to generate insert HAV DNA</td>
<td>R.H. Purcell (N.I.H.)</td>
</tr>
<tr>
<td>pSup2021</td>
<td>plasmid carrying the transposon Tn5 including the Kan' gene; used as template in PCR to generate insert Kan' DNA</td>
<td>A.J. Biel, L.S.U., Baton Rouge, LA</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
A. Primers for PCR amplification of insert HAV VP1 DNA for cloning into the λSurfZAP™ vector

1. SurfVP1PS2
5'-d(TTGCCCTACGGGCGCCGAGGCTCTCCTGCTCTTAGCACAACCA)-3'

2. SurfVP1PS1
5'-d(TTAGCAGCACAAACAAACAGCAATGGCGTTGGAGATGATGTTCTGGA)-3'

3. SurfVP1NS
5'-d(ACCCCTACCTCCTCTAGCTCTAAATCTTTATCTTTCTCCTCTGATCTAGG)-3'

B. Primers for in vitro excision of pSurfVP1 phagemid from recombinant λSurfVP1 DNA template

1. F1ORIPS
5'-d(CCCGATATCGCAACACTCAACCGCTCTATTC)-3'

2. F1ORINS
5'-d(CCCGATATCGCTCCAGTTTGGAACGAGG)-3'

Figure 22. Oligonucleotide primer sequences for (A) PCR Amplification of Insert HAV Capsid DNA for forced directional cloning into the λSurfZAP vector and (B) PCR based in vitro excision of phagemid pSurfVP1. (F1ORI) Phage f1 origin of replication, (VP) Viral peptide, (PS) positive strand primer and (NS) negative strand primer. The restriction endonuclease recognition sequences (EcoRV, NotI, SpeI) within each primer are underlined.
T7-VP1-PS2
5'-d(CCCGCTAGAATTCCAGTTGGAGATGATTCTGGAGGT)-3'
EcoRI

T7-VP1-NS2
5'-d(CCCGTCGACATCGATTCAAATCTTTTATCTTCCCTCT)-3'
Bsp106I/Clal

T7-VP3-PS2
5'-d(CCCGCTAGAATTCCAGATGATGAGAAATGAATTTAGG)-3'
EcoRI

T7-VP3-NS2
5'-d(CCCGTCGACATCGATTTGTGTAGTAACATCCATAGC)-3'
Bsp106I/Clal

T7-KAN-PS1
5'-d(ATGGCTAGAATTCCAGATGATTGAACAAGATGGATTG)-3'
EcoRI

T7-KAN-NS1
5'-d(CCCGTCGACATCGATTCAGAAGAACTCGTCAAGAAG)-3'
Bsp106I/Clal

Figure 23. Oligonucleotide primer sequences for PCR amplification of insert HAV capsid DNA for cloning into the pT7-7 vector. (KAN) Tn5 gene encoding kanamycin resistance, (VP) Viral peptide, (PS) positive strand primer and (NS) negative strand primer. The restriction endonuclease recognition sequences (EcoRI, Bsp106I, Clal) within each primer are underlined.
**Left Arm of the SurfZAP™ Vector** Subsequent to Cloning of Insert SurfVP1 DNA:
Oligonucleotide primer SurfVP1PS2 and SurfVP1PS1 were designed such that, following PCR amplification and ligation into the left arm of the SurfZAP vector, the 5' end of insert HAV VP1 would include a complete *NotI* site and the remainder of the pelB leader sequence downstream of the *NotI* site.

5'-AATTAACCCCTCATAAGAGGAACAAAGCTGGAGCTTGAATTCTTAACTCGC
  T3 Promoter
  EcoRI
  -----------------------------------------------pelB leader
  M K Y L L P T A A A G L L
  CAAGGAGACATCATTAATG AAA TAC CTA TTG CCT ACG GCG GCC GCA GGT CTC CTC
  RBS
  *NotI*
  sequence-----------------------------------------------
  L L A A Q P A M A
  CTC TTA GCA GCA CAA CCA GCA ATG GCC GTT GGA GAT GAT TGT GGA
  RBS
  *NotI*
  HAV VP1

**Right Arm of the SurfZAP™ Vector** Subsequent to Cloning of Insert SurfVP1 DNA:
Oligonucleotide primer SurfVP1NS was designed such that, following PCR amplification and ligation into the right arm of the SurfZAP vector, the 3' end of insert HAV VP1 (shown upstream of the *SpeI* site) would include a complete *SpeI* site and would be in frame with the downstream sequences encoding the carboxyl terminal domain (amino acids 198-406) of phage cpIII protein.

5'-............................CCT AGA TCA GAG GAA GAT AAA AGA TTT GAG
  G G G G S
  ACT AGT GGA GGT GGA GGT AGC CCAATTCGTGGTGAACCCGGGGGGGAAGTT
  *SpeI*
  198
  of phage cpIII protein---
  )............................TCTTAAATCATGCCAGTTCAAAAGGTTATTCATTTCTAGGTTAA
  406  Stop

GCGGCCGTCGGAGGCGGCGGGCATACCAATTCCGCCATA TAGTAGTCTGATTAT-3'
  T7 Promoter

**Figure 24.** A schematic diagram representing the design of insert DNA for cloning into the SurfZAP™ vector.
synthesized by beta-cyanoethyl phosphoramidite chemistry (Genelab, Dept. of Vet. Med. & Parasitol., School of Veterinary Medicine, LSU, Baton Rouge, LA).

**PCR Amplification Purification and Preparation of Insert DNA for SurfZAP and pT7-7 Cloning:** PCR used to amplify insert HAV VP1 DNA for cloning into the SurfZAP vector was performed in two steps. A 100 µl reaction containing 200 µM of dATP, dCTP, dTTP and dGTP; 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl$_2$, 0.01% gelatin, 2.5 units of AmpliTag DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT), 1.0 µM each of SurfVP1PS1 and SurfVP1NS primers and 500 ng template pGEM-HAV DNA was set up in step 1. This PCR reaction mix was then subjected to 30 cycles of amplification in a Perkin-Elmer model 9600 Programmable Thermal Cycler as follows: Denaturation at 95°C for 30 sec followed by 94°C for 1 min, annealing at 55°C for 90 sec, extension at 72°C for 90 sec, with a cycle extension of 10 sec per cycle and then cooling to 4°C. Following amplification, the PCR reaction mixture was separated from the upper oil layer, and 10 µl of the reaction mixture was subjected to a 2% agarose gel electrophoretic analysis to detect amplified products. The remaining reaction mixture was diluted with 2 ml of TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA), transferred into a Centricon-100 device (Amicon, Beverly, MA) and subjected to centrifugation at 1000 X g for 20 min to remove non-incorporated primers and dNTPs. The PCR product concentrated in the retentate was recovered by inverting the Centricon device and centrifuging at 500 X g for 2 min. The concentration of DNA calculated after measuring the $A_{260}$ of diluted samples. A second PCR reaction was then set up under the same conditions except that 100 ng of PCR product from Step 1 was
used as template and 1 \( \mu M \) of SurfVP1PS2 primer was used instead of SurfVP1PS1 primer. Following completion of thermal cycling, the products were recovered by using the Centricon-100 device. A 2\% agarose gel electrophoresis and \( A_{260} \) were used to analyze the second PCR reaction. The concentrated PCR product from the Step 2, SurfVP1 insert DNA, was further purified by the GeneClean II method to remove remaining pGEM-HAV template DNA. DNA concentration of the purified insert DNA was measured and 5 \( \mu \)g of the DNA was digested with the restriction enzymes, \( N_{orl} \) and \( Spel \) (Stratagene, La Jolla, CA), in separate but successive reactions to generate compatible, cohesive ends for forced directional cloning into the SurfZAP vector. The double-digested insert DNA was purified and concentrated by centrifuging in a Centricon-100 device to remove oligonucleotide fragments which could interfere in the subsequent ligation reaction. The DNA concentration was determined as before.

PCR reactions for amplifying insert DNA for pT7-cloning were performed by the same approach followed for the pAX-cloning. Plasmid pGEM-HAV DNA was used as the template for HAV genes and specific primer sets for VP1 or VP3. A 794 gene segment from Transposon 5 (Tn5) encoding resistance to Kanamycin was chosen as a positive control to be used throughout the course of the pT7-cloning experiments. Plasmid pSup2021 (provided by Dr. Alan Biel, Dept. of Microbiology, LSU, Baton Rouge, LA) which carries the Tn5 insert was used as the template in the control PCR reaction with the specific primer set, T7KANPS and T7KANNS (Figure 24). Depending on the outcome of these PCR reactions and where it was deemed necessary, a structured trial-and-error optimization of the PCR buffer conditions for each individual template-
primer set combinations was performed by using the Opti-Prime kit (Stratagene, La Jolla, CA) as outlined previously under Materials & Methods in Chapter 2. Following successful PCR amplification, as verified by agarose gel analysis, and subsequent Centricon-100 purification, DNA concentrations were determined for each insert DNA sample. Five \( \mu g \) of each insert DNA was subjected to successive restriction enzyme digestions with \( \text{EcoRI} \) and \( \text{Bsp}1061 \) enzymes to generate cohesive ends compatible for ligation into a similarly digested vector pT7-7 plasmid. The double digested insert DNA samples were purified and concentrated by using Centricon-100 devices and DNA concentrations were determined by \( A_{260} \) measurements. Each digested insert DNA sample was stored frozen at \(-20^\circ\text{C}\) until their further use.

**Plasmid Extraction and Preparation of Vector pT7-7 DNA for Cloning:**

Plasmid pT7-7 DNA was extracted from \( E.\text{coli} \) host strain DH1 by a modification of the alkaline lysis method (Birnboim and Doly, 1979) and then purified by the Geneclean II method. Competent \( E.\text{coli} \) host strain BL21 (Stratagene, La Jolla, CA) was transformed with the purified plasmid pT7-7 DNA by the calcium chloride method, described earlier (Chapter I: Materials & Methods). Transformants were used in a large scale plasmid extraction. Briefly, plasmid DNA was extracted from 1 liter LB broth culture of \( E.\text{coli} \) BL21 cells carrying pT7-7 by the modified alkaline lysis method. The DNA pellet from the final step of the procedure was treated as follows. DNA pellets were dissolved in 10 ml of acetate-MOPS buffer (0.1M Sodium Acetate, pH 8.0, 3-[N-MOrpholino]Propane-Sulfonic acid), precipitated by addition of 20 ml cold absolute ethanol and incubated at \(-20^\circ\text{C}\) for 10 min. The precipitate was pelleted by centrifugation at 10,000 rpm for 10
min and the supernatant discarded. The pellet was resuspended in 2 ml of TE buffer and to this an equal volume of a Lithium Chloride solution (5M LiCl, 0.05M MOPS) was added. After an incubation at 0°C for 15 min, the resulting precipitate was removed by centrifugation at 10,000 rpm for 10 min. The supernatant fluid was reprecipitated with absolute ethanol and the resulting pellet was dissolved in TE buffer by heating to 65°C for 10 min. Next, the DNA preparation was treated with DNase-free RNase (at a final concentration of 100 μg of RNase/ml of DNA solution) and incubated at 37°C for 30 min followed by 65°C for 5 min. Then, 0.6 ml of 10% SDS and 0.6 ml of 5M NaCl were added to the DNA solution, mixed and the solution was extracted one time with an equal volume of Phenol-Chloroform solution (Phenol: Chloroform: Isoamyl alcohol = 25:24:1) and then with an equal volume of Chloroform. The resulting aqueous phase was mixed with 720 μl of 3M Sodium acetate at pH 5.2 and precipitated with two volumes of absolute ethanol as before. The solution was incubated at -80°C for 30 min, centrifuged at 10,000 rpm for 20 min to harvest the precipitate. The resultant DNA pellet was washed one time with 70% ethanol, dried under vacuum at room temperature and dissolved in 1 ml of TE buffer. DNA concentration of this plasmid preparation was estimated by measuring A_{260}. Approximately 250 μg of plasmid DNA was subjected to sequential restriction enzyme digestions with EcoRI and Clal (Gibco BRL, Gaithersburg, MD) and purified by GeneClean II. The DNA concentration of purified, double-digested vector DNA was calculated following A_{260} measurement of sample dilutions. The double-digested DNA and residual undigested plasmid DNA in TE buffer were stored separately at -20°C until needed.
Transformation of Host *E. coli* BL-21 Cells with Recombinant pT7 DNA by Electroporation: Individual ligation reactions were set up to join restriction digested insert DNA with vector pT7 DNA following procedures similar to those described previously (Materials and Methods, Chapter 1). Competent *E. coli* BL21 cells were prepared as follows. *E. coli* host cells were grown in a salt-free growth medium, YENB broth (Bacto yeast extract [0.75%], Bacto Nutrient Broth [0.8%]; Difco Laboratories, Detroit, MI) at 30°C with shaking. These steps were taken to decrease the incidence of arcing during electroporation and if possible, to improve the efficiency of electroporation.

Fresh or frozen cells (that had been gently thawed) were used in the electrotransformation reaction as outlined previously. Following electroporation, cells were immediately resuspended in 1 ml SOC medium (2% Bacto Tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 MgSO₄ and 20 mM glucose; Bacto tryptone was from Difco laboratories, Detroit, MI and the other chemicals were from EM Science, Gibbstown, NJ). The cell suspensions were transferred into Falcon sterile polypropylene tubes (17 X 100 mm, Becton Dickinson and Company, Lincoln Park, NJ) and incubated at 30°C for 1 h. Following incubation, the cells were transferred to microfuge tubes, sedimented at 3,000 X g for 5 min and then gently suspended in 300 µl fresh SOC medium which contained 100 µg/ml ampicillin. They were plated on selective SOC-amp plates in triplicate. The SOC-amp plates were incubated overnight at 30°C and then examined for recombinant colonies, and selected recombinants were chosen at random for further studies.
Plasmid DNA isolation was performed on selected recombinants to identify potential candidates for HAV protein production. Plasmid DNA was extracted from 3 ml cultures of each recombinant following overnight incubation in SOC-amp broth, by the PERFECTprep™ plasmid DNA kit (5 Prime-3 Prime, Inc., Boulder, CO). Cells were pelleted in microfuge tubes from 3 ml overnight cultures by centrifugation at 12,000 X g for 30 sec. The cellular pellets were suspended in 100 μl of Solution I (50 mM Tris-Cl, pH 7.6, 10 mM EDTA, pH 8.0, 100 μg/ml RNase A), the cells lysed by addition of 100 μl of Solution II (0.2 N NaOH, 1.0% SDS) and mixed well by repeated gentle inversion. Each lysate was neutralized by addition of 100 μl of Solution III (1.32 M Potassium Acetate, pH 5.2) while mixing by vigorous inversion. The lysates were subsequently centrifuged at 16,000 X g for 2 min to remove cell debris and other insolubles. The supernatant fluids were transferred to individual PERFECTprep spin column-collection tube assemblies, 450 μl of a thoroughly resuspended PERFECTprep DNA binding matrix in Guanidine-HCl was added, mixed and centrifuged as before. After discarding the filtrates, 400 μl of a diluted Purification Solution consisting of ethanol, Tris-Cl, NaCl and EDTA (NOTE: individual concentrations were not provided by manufacturer) was added to each spin column which were shaken briefly and then centrifuged for 60 sec. After discarding the filtrate, the columns were sedimented again and transferred to fresh collection tubes. Plasmid DNA were eluted by suspending the DNA binding matrix in 50 μl of TE buffer and vortexing briefly followed by centrifugation. Eluted plasmid DNA was recovered in the filtrates. About 5 μl of each plasmid DNA sample was digested with EcoRI and ClaI and analyzed by 0.8% agarose
gel electrophoresis. DNA concentrations of positive recombinant DNA preparations were determined by $A_{260}$ measurements and stored at -40°C as described before. Linearized recombinant plasmid DNA was used as the template along with the corresponding sets of VP1 or VP3 or KAN primer sets for individual PCR reactions done to confirm the presence of the relevant insert DNA. The reaction conditions for each template-primer set were the same as previously optimized. Following confirmation of the presence of insert DNA, representative bacterial clones were chosen for the studies which involved the induction of HAV capsid protein expression from the recombinant pT7-7 vectors.

**Expression of HAV Capsid Proteins from Recombinant pT7-7 Vector:**
Recombinant pT7 plasmids carrying HAV VP1 or VP3 inserts and KAN insert DNA were used to transform freshly prepared, electroporation competent *E.coli* BL-21 (DE3) pLysS cells (Table 7). Following transformation and isolation of recombinants on SOC-amp plates, representative clones were confirmed by plasmid DNA isolation and PCR verification for insert DNA. Frozen stocks of each recombinant, for each insert DNA, were prepared. In addition, three recombinant *E.coli* BL-21 bacterial clones: T7-VP1#1e, T7-VP3#3a and T7KAN#6b were chosen for studying the IPTG-induction of recombinant protein expression by methods elaborated upon earlier with one modification. Cells were grown at 30°C throughout the experiment. Samples representing glucose induced catabolite repression of protein expression from the lac promoter (gr), time 0 ($T_0$; i.e. at the time of IPTG addition to cells removed previously from under glucose repression), and one, two, three and four hours post-induction ($T_1$, $T_2$, $T_3$ and $T_4$) were obtained.
Samples for SDS-PAGE analysis were prepared from cell pellets of IPTG-induced cells and stored at -20°C until needed.

Whole cell lysates were analyzed by 12% SDS-PAGE and Coomassie blue staining for expression of recombinant proteins using polyacrylamide gels of 1.5 mm thickness along with a Modular Mini-PROTEAN II Electrophoresis System (Bio-Rad Laboratories, Hercules, CA). Depending on the outcome of the IPTG-induction experiments, Western Blots were done on selected recombinants by methods described earlier.

**Cloning of Insert HAV DNA into λSurfZAP Vector**: *E.coli* XL-1 Blue was selected as the host strain for the initial transduction experiments. A broth culture of the host cells was started in 10 ml LB broth by inoculating with a single colony from a fresh LB-tetracycline (LB medium containing tetracycline at 12.5 μg/ml) streak plate. Following overnight incubation at 37°C with vigorous shaking, the 10 ml culture was transferred into 100 ml of fresh, sterile LB medium with 0.2% (v/v) maltose and 10mM MgSO₄ and incubated for 2-3 hours or until the culture attained an optical density at 600 nm (OD₆₀₀) of 0.2-0.3. The cells were sedimented at 1000 X g for 10 min and suspended in sterile 10 mM MgSO₄ to a final OD₆₀₀ of 1.0 and were used immediately in the transduction experiment.

Ligation reactions to generate concatameric, recombinant λSurfZAP DNA were set up with NorI-SpeI double-digested insert HAV SurfVP1 and λSurfZAP DNA based on the manufacturer's recommendations (Section IV: Instruction Manual for SurfZAP™ Cloning Kit, Stratagene, La Jolla, CA). Insert to vector molar ratios were calculated by
using the following formula; picomole ends/μg of DNA = 10⁶ ÷ (number of base pairs 
X 660). Insert to vector ratios of 1:1, 2:1 and 4:1 were employed to set up individual 
ligation reactions in a final volume of 5 μl (consisting of 1.0 μl SurfZAP vector arms at 
1 μg/μl concentration, 0.5 μl of 10X ligase buffer [500 mM Tris-HCl, pH 7.5, 70 mM 
MgCl₂ and 10 mM DTT], 0.5 μl of 10 mM rATP, 0.5 μl of T4 DNA ligase enzyme [0.2 
U] in addition to insert DNA and dH₂O). A mock ligation reaction without insert DNA 
was included as a negative control. All ligation reactions were incubated at 4°C 
overnight. Next, one μl of each ligation reaction was packaged into viable lambda 
bacteriophage particles using Gigapack® Gold packaging extracts (Stratagene, La Jolla, 
CA) in accordance with the manufacturer's suggested protocol. The packaging extracts 
were removed from from storage at -80°C and placed on dry ice. Each set of packaging 
extract contained a yellow colored tube with the Sonic extract and a red colored tube with 
the Freeze-Thaw extract. The Sonic extract was thawed on wet ice at 4°C while the 
Freeze-Thaw extract was thawed between fingers. One μl of λSurfZAP DNA from a 
ligation reaction was added immediately to the Freeze-Thaw extract as soon as it began 
to thaw. Then, fifteen μl of thawed Sonic extract was quickly added to the freeze-thaw 
extract containing the DNA and the tube contents were mixed well with a pipet tip. The 
reaction mixture was centrifuged for 3-5 sec and then incubated at 22°C for 1.5 to 2 h. 
Next, 500 μl of sterile SM buffer (each liter containing 5.8 g of NaCl, 2.0 g of MgSO₄ 
·H₂O, 50.0 mL of 1M Tris-HCl at pH 7.5, 5.0 mL of 2% (w/v) gelatin in addition to 
distilled water to make up the rest of the volume) and 20 μl chloroform were added 
successively to each packaging reaction mix, the reaction mix was centrifuged briefly for
3-5 sec to sediment debris and the supernatant fluid was transferred to a new tube prior to setting up appropriate dilutions in SM buffer to titer lambda phage and thereby evaluate the success of the ligation and packaging reactions.

Serial dilutions of lambda phage, 1:10³, 1:10⁶ and 1:10⁹, were prepared in sterile microfuge tubes in SM buffer. About 1-10 µl of each dilution was mixed with 200 µl of freshly prepared *E.coli* XL1-Blue cells (OD₆₀₀ = 0.5). The phage-bacteria mixture was incubated with occasional shaking, at 37°C for 15 min to permit phage adsorption onto host cells. Next, the phage-bacteria mixture was transferred to tubes which contained 3 ml of melted and cooled NZY top agar at 50°C (each liter containing 5 g of NaCl, 2 g of MgSO₄·7H₂O, 5 g of yeast extract, 10 g of NZ amine [casein hydrolysate; Difco laboratories, Detroit, MI] and 7 g agarose; pH of the medium adjusted to 7.5 with NaOH) and vortexed briefly to ensure even mixing of contents. The top agar was immediately poured onto prewarmed (to 37°C) NZY agar plates (same formulation as NZY top agar except that the final amount of agarose added was 15 g/liter) and the plates were incubated at 37°C overnight. The NZY plates were examined for plaques and the phage titer [in plaque-forming units per milliliter (pfu/ml)] was calculated using the following formula: (Number of plaques X dilution factor X 1000) ÷ Volume of phage solution added to host cells in milliliters. The plates were sealed with parafilm and stored at 4°C until their use in PCR to verify the presence of cloned insert DNA.

**PCR Screening and Verification of Plaques Containing Recombinant λ phage carrying SurfVPI DNA**: A simple miniprep procedure was used to isolate and purify λSurfVPI DNA from multiple phage lysates prior to rapid PCR screening and
identification of recombinant plaques. A 10 ml culture of lambda sensitive host *E. coli* XL1-Blue cells was grown overnight in LB broth (containing 0.2% maltose and 10mM MgSO₄) at 37°C with vigorous shaking. Five ml of the culture was inoculated into 60 ml of fresh LB-Maltose-MgSO₄ medium in a 250-ml Erlenmeyer flask and incubated at 37°C with shaking for 2 hours. Three ml of the log phase culture was aliquoted into each of 20 sterile 17 X 150 mm test tubes. Next, plaques selected from the lambda plates from the phage titer experiment were toothpicked individually and inoculated separately into 3 ml cultures of *E. coli* XL1-Blue cells. Each phage inoculated cultures was incubated in shaker-incubator or on a rotating wheel in an incubator at 37°C for 4h or until lysis of bacterial cells was evident. Fifty µl of chloroform was added to each tube, and the contents were vortexed for 1 min to lyse the remaining cells. Then, 1.5 ml of each lysed culture was transferred into sterile, labeled microfuge tubes. Two µl each of DNAse (5mg/ml) and DNAse-free RNAse (10mg/ml) were added to each tube, mixed, and the cultures were incubated for 15 min in a 37°C water bath. The cultures were then centrifuged at 10,000 X g for 10 min to pellet cell debris and the supernatant fluid from each culture was aliquoted into two microfuge tubes. Next, 750 µl of ice-cold isopropanol was added to each tube, the contents mixed thoroughly by vortexing and the tubes were centrifuged as before. The supernatant from each tube was carefully drained and the tubes were laid inverted on absorbent paper or kimwipes for 5 min. Each translucent phage pellet was resuspended in 100 µl of TE buffer. The resuspended pellet from both tubes corresponding to a single phage plaque originally were pooled into one tube. Then, 200 µl of buffered phenol was added to each pooled phage suspension,
vortexed vigorously for 10 min and centrifuged at 14,000 rpm for 2 min. The resulting aqueous layer was reextracted with buffered phenol and the aqueous layer from the second phenol extraction step was transferred into a fresh microfuge tube. Next, 200 µl of chloroform was added to each tube, mixed well and centrifuged as before. The aqueous layer was carefully removed into a fresh tube and the chloroform extraction repeated one more time. The aqueous layer from the second extraction was treated with 20 µl of 3 M Na-acetate at pH 4.8 and mixed well. 500 µl of ice-cold 100% ethanol was added to each tube to precipitate phage DNA at room temperature for 15 min. Then, the tubes were centrifuged at 14,000 rpm for 10 min to pellet DNA. After discarding the supernatants, the pellets were washed once with 1 ml of 70% ethanol and centrifuged again. The supernatants were carefully drained and the phage DNA pellets were dried under vacuum for 5 min prior to resuspending each in 100 µl of sterile TE buffer. Two µl of each phage DNA sample was used as template DNA in the subsequent individual PCR reactions while the remaining sample was stored frozen until needed.

Individual PCR reactions to verify the successful cloning of insert SurfVPl into the λSurfZAP vector were set up in 100 µl volumes (10 µl of 10 X PCR buffer, 8 µl of 25 mM MgCl₂, 2 µl of each of dNTPs, 0.5 µl of AmpliTaq DNA Polymerase, 2µl of SurfVP1PS1 primer, 2µl of SurfVP1NS primer and 67.5 µl of sterile distilled H₂O). Thermal cycling was performed for 30 cycles as follows: Cycles 1-3, 95°C/2 min followed by 55°C/1 min followed by 72°C/90 sec; Cycles 4-30, 94°C/90sec followed by 50°C/1 min followed by 72°C/90°C with a cycle extension of 10 sec/cycle. Following the PCR, 10 µl of each reaction mix was mixed with 5 µl of DNA loading dye and
electrophoresed on a 2% agarose gel, stained with ethidium bromide and photographed. A positive amplification of a ~950 bp DNA fragment was considered as evidence of successful cloning and the corresponding recombinant λSurfVP1 plaques identified. Ten such recombinant plaques were chosen for generation of 50 ml lysates for short term storage at 4°C, for long term storage at -80°C and for the single clone in vivo excision experiments to produce the plasmid equivalents (i.e. pSurfVP1) of recombinant DNA from their corresponding λ phage forms.

For the long term storage of the individual phage lysates, 930 µl aliquots of each phage lysate were added to several microfuge tubes and 70 µl of filter-sterilized DMSO (DiMethylSulfOxide; Mallinckrodt Inc., Paris, KY) was added to each tube, mixed well by pipetting and the tubes were stored frozen at -80°C.

**Single Clone in vivo Excision of Recombinant pSurfVP1 from λSurfVP1 clones:** Transduction competent host *E.coli* XL1-Blue and SOLR strains were prepared as described in a previous section and used in the in vivo excision experiments to convert recombinant DNA from the initial lambda phage DNA form to a final plasmid form via a single-stranded filamentous phage DNA form. XL1-Blue cells and SOLR cells were freshly prepared each time from log-phase cultures and the cells were resuspended in sterile MgSO₄ to the required OD₆₀₀ just prior to their use.

The ExAssist™ interference-resistant helper phage, which is a derivative of the single-stanned f1 coliphage, was part of the SurfZAP™ cloning kit and was used in the in vivo excision following the manufacturer's suggested protocol. Prior to the in vivo excision experiment, the helper phage was titered with XL1-Blue cells. Phage dilutions
(10^4 to 10^7) in TE buffer were prepared and 1 µl of each dilution was mixed with 200 µl of XL1-Blue cells (OD_{600}). The phage-bacteria mix was incubated for 15 min at 37°C to allow for phage adsorption, mixed with premelted and precooled NZY top agar at 50°C, poured onto prewarmed (37°C) NZY plates and incubated at 37°C overnight. The next day, the plates were screened for turbid plaques and the phage titer determined as before.

For \textit{in vivo} excision of pSurfVP1, ten individual recombinant phage lysates that had been identified previously by PCR were selected. Each recombinant phage was combined with 40 µl of transduction competent XL1-Blue cells (OD_{600}=5.0) in separate, sterile 15 ml capped conical centrifuge tubes (Corning Costar Corporation, Cambridge, MA) at a multiplicity of infection (MOI) of 1:80 lambda phage-to-host cell ratio (assuming that a OD_{600} of 5.0 for XL1-Blue cells = 4 X 10^9 cells/ml). Next, the ExAssist helper phage was added to the lambda phage-bacteria mix at a MOI of 2:1 helper phage-to-host cell ratio and the reaction mix was incubated at 37°C for 15 min. Then, 5 ml of LB broth was added, the contents of the tube mixed well by gentle pipetting and the culture incubated at 37°C overnight with gentle shaking. The following day, the cultures were placed in a water bath at 70°C for 30 min to lyse the lambda particles and XL1-Blue cells and then centrifuged at 2500 x g for 10 min to pellet the cell debris. The supernatants containing the excised recombinant single-stranded phage were recovered into fresh sterile tubes and stored at 4°C until further use.

\textit{E.coli} SOLR cells harboring pSurfVP1 were produced from the supernatants of the \textit{in vivo} excision experiment by preparing dilutions in TE buffer of the supernatants
as was done in the helper phage titration experiment. One to ten µl of each dilution was then mixed with 200 µl of SOLR cells (OD₆₀₀ = 1.0) and preincubated at 37°C for 15 min. Then, the SOLR cells were spread uniformly on the surface of LB-Amp agar (with 50µg/ml ampicillin) with a sterile glass hockey stick and incubated overnight at 37°C. The next day, the plates were checked for Ampicillin resistant (Amp') bacterial colonies. The titer of recombinant phagemid was estimated [in colony-forming units per milliliter (cfu/ml) using the formula, (Number of colony-forming-units X Dilution factor X 1000) ÷ Volume of SOLR cells plated. Several of the Amp' colonies were selected for plasmid DNA extraction by the PERFECTprep procedure described previously. The plasmid DNA obtained were screened for presence of SurfVP1 insert by restriction enzyme analysis with EcoRI and by PCR following procedures outlined already. Frozen stocks of bacterial cells that were confirmed as recombinants were generated for long term storage at -80°C. The pSurfVP1 plasmid DNA samples were stored at -20°C until needed.

Recombinant Lambda DNA Isolation and Purification: The ten recombinant λSurfVP1 phage isolates that had been identified previously by PCR were chosen for the extraction and purification of recombinant phage DNA. A λQUICK!™ kit (Bio 101, Inc., La Jolla, CA) was used for a rapid isolation and purification of phage DNA from a 100 ml culture of each infected/lysed host bacteria as follows. A fresh overnight 5 ml culture of lambda sensitive *E.coli* XL1-Blue cells was grown at 37°C in LB-Maltose-MgSO₄. The next morning, the 5 ml culture was infected with 10¹² pfu of a recombinant phage lysate and incubated for 20 min in a 37°C water bath. Next, the 5 ml culture was
added to 50 ml of fresh, sterile LB-Maltose-MgSO$_4$ medium in a 250 ml Erlenmeyer flask and incubated on a shaking platform in a 37°C incubator for 4-5 hours. Then, 500µl of chloroform was added and the culture was shaken vigorously at 37°C for 15 minutes. Phage lysate was obtained by centrifugation of the lysed culture at 10,000 X g for 10 min to pellet cell debris. The supernatant containing phage particles in suspension was transferred into a sterile, polypropylene centrifuge bottle. An equal volume of isopropanol was added, mixed well and the mixture was centrifuged as before to obtain a translucent phage pellet. The supernatant was removed carefully to the maximum extent possible and the pellet was resuspended in 1.2 ml TE buffer. The phage suspension was then removed into a microfuge tube and centrifuged at 14,000 rpm for 2 min to pellet any remaining cell debris. The supernatant containing the phage particles was filtered through a 0.45µM Nalgene™ syringe filter (Nalge company, Rochester, NY) and collected in a fresh microfuge tube. Then, 0.5 µl of NUCLEASE MIXX (supplied with the λQuick! kit) was added to the phage suspension and incubated at 37°C for 15 min. Next, 0.4 ml of λPHAGE DROP BUFFER (supplied with the kit) was added to the suspension, mixed well and left at 37°C for 20 min to precipitate phage. Phage particles were pelleted by centrifuging at 10,000 X g for 5 min and discarding the supernatant. After resuspending the pellet in 150 µl of TE buffer, 150 µl of λPHAGE LYSIS/BINDING BUFFER (supplied with kit) was added, mixed well until the solution became clear and incubated at 70°C for 10 min to lyse the phage particles. Next, 10 µl of λGLASSMILK™ (supplied as part of kit) was added to the solution, mixed gently but thoroughly and incubated for 5 min at room temperature to allow for phage DNA binding.
to the glassmilk. The solution was then spun for 30 sec in a microcentrifuge at 14,000 rpm and the resulting supernatant removed. The λDNA/GLASSMILK™ pellet was washed with 300 μl of λDNA wash (supplied with kit) and spun for 30 sec to remove the supernatant. Without resuspending the pellet, the washing step was repeated. Then, the washed pellet was resuspended gently in 20 μl of TE buffer and incubated at 37°C for 5 minutes before centrifuging again at 14,000 rpm for 30 seconds. The resulting recombinant λ DNA was transferred into a new microfuge tube, DNA concentration of the solution measured by determining OD_{260} of diluted samples and the undiluted solution was stored frozen at -20°C until later.

The recombinant λ DNA samples were used as a source of template DNA in PCR reactions to verify the presence of insert SurfVP1 DNA as well as in experiments concerning the \textit{in vitro} excision of pSurfVP1 by PCR.

\textbf{\textit{In vitro} Excision-PCR to Generate Recombinant Phagemid DNA with Insert HAV VP1 DNA from Recombinant λSurfVP1 Clones:} The essential steps of this procedure are depicted schematically in Figure 25. Oligonucleotide primers, positive-strand (PS) primer F1ORIPS with the sequence 5'-CCC GAT ATC CAA CAC TCA ACC CTA TCT C-3' and negative-strand (NS) primer F1ORINS with the sequence 5'-CCC GAT ATC TTC CAG TTT GGA AC A AGA G-3' were synthesized by beta cyanoethyl phosphoramidite chemistry (Genelab, Dept. of Vet. Med. & Parasitol., School of Veterinary Medicine, LSU, Baton Rouge, LA). Primers F1ORIPS and F1ORINS are derived from the nucleotide sequence of the bacteriophage f1 functional origin of replication (Dotto \textit{et al.}, 1984), parts of which are represented in the
in vitro Excision-PCR to Rescue Recombinant pSurfVP1 from λSurfZAP Carrying HAV VP1 cDNA Insert

Figure 25a. A schematic representation of the essential steps in the in vitro-Excision PCR procedure employed in the amplification of the entire pSurf-VP1 phagemid DNA from within the larger λSurf-VP1 DNA, as an alternative to the E.coli and phage dependent in vivo excision and rescue.
Restriction Enzyme Digested pSurfVP1

Ligation with T4 DNA Ligase

4378, XspI
MaeI, 132

4054, ScaI
3943, PvuI

Circular pSurfVP1 DNA

Transform E. coli SOLR™ Host Cells [Electroporation]

Screen Recombinant E. coli Colonies by Plasmid DNA
Isolation and by PCR with SurfVP1 Primers

Figure 25b. A schematic representation of the essential steps in the in vitro-Excision PCR procedure employed in the amplification of the entire pSurf-VP1 phagemid DNA from within the larger λSurf-VP1 DNA, as an alternative to the E. coli and phage dependent in vivo excision and rescue.
terminator (T) and initiator (I) sequences that constitute the boundaries of the phagemid (pSurfscript) portion of the λSurfZAP™ vector. The λSurfZAP vector also contains the lacZ promoter for transcription of cloned DNA inserts, multiple cloning site (MCS) with flanking T3-T7 promoters for insertion of foreign DNA, Ampicillin resistance gene (Amp') as a selective marker and the colE1 replication origin (Figure 22) within the region flanked by the T and I sequences. In the recombinant λSurfVP1 phage DNA, the insert HAV VP1 DNA is cloned between the NotI and SpeI sites at the multiple cloning site. All of the phagemid DNA located between and including the T and I sequences is essential for successful in vivo or in vitro excision of recombinant phagemids and for their subsequent replication and maintenance as double-stranded plasmid DNA within bacterial host cells. Both primers were designed to include a recognition site for the restriction enzyme, EcoRV, towards their respective 5'ends to assist in the recircularization of the plasmid following a successful amplification of phagemid DNA.

The entire recombinant phagemid DNA between the T and I sequences of λSurfVP1 was amplified by PCR using the Expand™ Long Template PCR System (Boehringer Mannheim corporation, Indianapolis, IN). Since this resulted in a double-stranded plasmid form of the recombinant DNA that could be easily circularized subsequently, this procedure was termed in vitro Excision-PCR. The PCR reaction itself was set up as follows by preparing Master mix 1 and 2 in separate microfuge tubes. Master mix 1 contained 1 μl of λSurfVP1 DNA (from theλQuick! procedure) as template, 300 nM each of the PS and NS primers and sterile redistilled water to make up a total final volume of 25 μl in a thin-walled PCR reaction tube (Perkin-Elmer Cetus,
Norwalk, CT) on ice. Master mix 2 consisted of 350 μM each of dATP, dTTP, dGTP and dCTP, 5 μl of 10X PCR buffer with MgCl₂ (500 mM Tris-HCl, pH 9.2, 160 mM (NH₄)₂SO₄ and 17.5 mM MgCl₂), 0.5-1 μl of enzyme mix (a mixture of Pwo and Taq DNA polymerases) and sterile redistilled water to a final volume of 25 μl in a regular microfuge tube on ice.

A modified "hot start" protocol (Chou, et al., 1992) was employed for the thermal cycling steps of the Excision-PCR using a Gene Amp System 9600 Thermal Cycler (Perkin-Elmer Cetus, Norwalk, CT). Master mix 2 was heated to 94°C for 5 min to denature the λSurfVP1 template DNA completely and then cooled to 65°C to allow for annealing of primers before adding the master mix 2 to initiate the first elongation step at 68°C for 10 min. Thereafter, thermal cycling was continued with denaturation at 94°C for 30 sec followed by annealing at 62°C for 30 sec followed by elongation at 68°C for 6 min for the next 30 cycles. A cycle elongation step of 20 sec per cycle was added to the elongation step beginning with cycle #11 and extending through cycle #30. After a final extension step at 68°C for 10 min, the reactions were cooled to 4°C and thermal cycling was stopped. The PCR product was separated from the upper oil overlay and collected into a fresh, sterile microfuge tube. A five μl volume of the product was subjected to a 0.6 % agarose gel electrophoretic analysis to confirm if the amplified product was of the anticipated size. The remaining PCR product was purified by using Centricon-100 filter devices as described earlier. DNA concentration of the purified PCR product was determined following OD₂₆₀ measurements of diluted samples. A PCR with SurfVP1PS2 and SurfVP1NS primers and 100 ng of purified product from
Excision-PCR was performed under the same thermal cycling conditions as was mentioned for amplification of insert SurfVP1 DNA from pGEM-HAV to confirm that the template DNA was indeed a PCR-amplified linear form of pSurfVP1 plasmid. The products of this reaction were analyzed by a 2% agarose gel electrophoresis as before. The remaining purified pSurfVP1 preparation was stored at -80°C until needed.

Restriction enzyme digestion with EcoRV (Gibco BRL, Gaithersburg, MD) of 5 μg of in vitro excised and purified pSurfVP1 was performed and the enzyme digested DNA was purified by GENECLEAN II procedure as outlined earlier. After estimating the DNA concentration, individual ligation reactions were set up with 125 ng, 250 ng and 500 ng of EcoRV digested pSurfVP1 to circularize the plasmid DNA. The ligation reaction conditions were the same as was described earlier for cloning into plasmid vectors. Next, the recircularized pSurfVP1 was used to transform freshly prepared host E.coli BL-21 cells by electroporation and several successful recombinants that resulted were chosen at random for further studies.

Plasmid DNA were extracted by the PERFECTprep™ procedure from several recombinants that resulted from transformation experiments involving in vitro excised and recircularized pSurfVP1. Restriction analysis with EcoRI and PCR with SurfVP1PS2 and SurfVP1NS primers were used to positively identify recombinants carrying pSurfVP1 as was done earlier. Frozen stocks of such recombinants were generated for long term storage at -80°C and for future experiments.

Expression of recombinant HAV SurfVP1 protein: Representative bacterial recombinants were selected from amongst the pool of in vivo and in vitro excision
derived clones for studies aimed at expression of recombinant HAV VP1 protein. Log-
phase cultures of recombinant cells were produced and induced with IPTG according to
a protocol similar to the one described already for IPTG-induction of pT7-7 vector
derived recombinants carrying HAV genes. Timed samples before and following the
IPTG-induction were collected and whole cell lysates prepared for analysis by 12 %
SDS-PAGE as described earlier. Any unused whole cell lysate samples were stored at
-20°C until needed later.

**Codon Usage Preference Frequencies in *E.coli* for HAV Genes:** A computer-
based analysis of frequency of preferred codon usage in *E.coli* for HAV capsid protein
genes VPO, VP1 and VP3 was performed using a Gateway 2000 P5-75 personal
computer (Gateway2000, Inc., N. Sioux city, SD). The codon usage database, an
extended world wide web (WWW) version of Codon Usage Tabulated from Genbank
(CUTG) genetic sequence database, was utilized for this purpose by accessing the WWW
The database consists of all of the complete sequenced and reported protein coding
sequences (CDS). Codon frequency tables showing frequency (per thousand) and count
for each codon as a sum of all CDS's for a given species were obtained and used for
comparison between HAV, *E.coli* (Host cells used in our study for expression of
recombinant HAV proteins), *Homo sapiens* (natural host for HAV replication) and
Human Poliovirus (a member of Picornaviridae related to HAV whose proteins have
been successfully expressed in *E.coli*) sequences.
RESULTS

Oligonucleotide primers were designed to enable successful PCR amplification of HAV VP1 insert DNA for cloning into the λSurfZAP and pT7-7 vectors (Figures 21 and 22). The primers for λSurfZAP cloning ensured the inclusion of the NotI and SpeI restriction enzyme sites and the essential portions of the pelB DNA leader sequence that is lacking in the double-digested vector DNA supplied by the manufacturer (Figure 25). All positive-strand (PS) primers for pT7-cloning also included the essential recognition site for EcoRI enzyme and all negative-strand (NS) primers included a site for Bsp1061/ClaI enzyme.

Insert SurfVP1 DNA was generated by successful amplification by two separate PCR reactions performed in a pre-determined sequence (Figure 26). The first PCR reaction which included plasmid template pGEM-HAV, primers SurfVP1PS1 and SurfVP1NS, resulted in a 945 bp DNA fragment that included the downstream SpeI site and part of the upstream pelB leader sequence. The subsequent PCR reaction which employed the 945 bp PCR amplified DNA from the previous reaction as template and with SurfVP1PS2 and SurfVP1NS primers, yielded a 978 bp complete SurfVP1 insert DNA that now contained the NotI restriction enzyme site in addition to the pelB leader sequence (Figure 24).

Insert DNA for pT7-cloning were generated by PCR in separate reactions as described under material and methods and resulted in successful amplification of a 929 bp T7VP1 and 766 bp T7VP3 DNA from template pGEM-HAV and a 823 bp T7KAN.
Figure 26. Agarose gel analysis of insert Surf VP1 DNA amplified by PCR for cloning into the SurfZAP vector. Lane 1: BioMarker low molecular weight DNA standard, Lane 2: VP1 DNA (945 bp) following the first round of amplification using pGEM-HAV as template, Lane 3: Insert SurfVP1DNA (978 bp) amplified during the second round of PCR using VP1DNA (945 bp) from the first round as template, and Lane 5: BioMarker low molecular weight DNA standard.
DNA from template pSup2021. The PCR conditions were optimized for each insert DNA amplification using the Opti-Prime kit.

Each insert DNA for pT7-cloning was subjected to sequential digestion with the appropriate restriction enzymes and purified by Geneclean II procedure. Plasmid vector, pT7-7, was also digested and purified in a similar fashion. Ligation reactions with insert:vector ratios of 1:1, 2:1 and 4:1 were set up initially. Prior to using *E. coli* BL21 host cells, we attempted without much success, electro-transformation in several *E. coli* strains like UT2300, DH1 (*recA1 endA1 gyrA96 thi-1 hsdR17(riK mK+) supE44 relA1*), JM101 (*sup E thi-1Δ(lac-proAB) [F' traD36 proAB lacZΔM15]*) and HB101 (*supE44 ara14 galK2 lacY1Δ(gpt-proA)62 rpsL20 (Str') xyl-5 mtl-1 recA13 Δ(mcrC-mrr) HsdS* (r* m*)). No stable transformants were recovered by varying insert:vector ratios or conditions of temperature, enzyme source, or duration of ligation reactions. However, when *E. coli* BL21 host cells were used, the efforts were successful and stable recombinants were obtained with a transformation frequency of ~10⁷ transformants/µg of vector DNA with respect to HAV VP1 and VP3 inserts and ~10⁸ transformants/µg of vector DNA with respect to KAN insert. Control reactions which were done with undigested pT7-7 vector DNA yielded ~10⁹ transformants/µg of vector DNA while those with cut (with *EcoRI*) and religated vector produced ~10⁸ transformants/µg of vector DNA. Several recombinants were chosen as candidates and examined by plasmid DNA isolation, digestion with *EcoRI* and *ClaI* enzymes, and agarose gel electrophoretic analysis (Figure 27). The presence of specific insert DNA in each of the template pT7-VP1, pT7-VP3 and pT7-KAN recombinant plasmids was independently confirmed by
Figure 27. Agarose gel analysis of restriction enzyme digests of recombinant pT7-derived plasmids carrying HAV genes, VP1 (~900 bp) or VP3 (~750 bp), or Tn5 derived KAN (~800 bp) gene insert. The plasmid DNA samples were digested with EcoRI and Clal enzymes in separate but successive reactions. Lanes 1,7,11: Vector pT7-7 without insert, Lanes 5, 10, 15: BioMarker low molecular weight DNA standard, Lanes 2-4: pT7VP1 DNA samples from different clones, Lanes 8-9: pT7VP3 DNA samples, Lanes 12-14: pT7KAN DNA samples.
PCR. Subsequently, *E. coli* BL-21(DE3)pLysS cells were transformed with these plasmids by electroporation and individual colonies were selected for induction of protein expression. The BL-21 (DE3) pLys host cells, but not the BL21 host cells, contain a λ lysogen, DE3. Phage DE3 encodes the T7 RNA polymerase under the control of *lacUV5* promoter. Controlled protein expression of specific genes placed downstream of the T7 RNA polymerase binding site (promoter) on vector pT7-7 is possible through IPTG induction of the *lacUV5* promoter.

Several recombinant *E. coli* BL-21 (DE3) pLysS clones, [T7-VP1 clones 1a, 1e and 2b; T7-VP3 clones 3a, 4b, and 4c; and T7-KAN clones 6a and 6b] were subjected to IPTG induction of recombinant protein expression from cloned genes carried on the recombinant pT7 vector. Induction of log-phase cells of each recombinant culture in concert with control bacterial cultures (BL-21 (DE3) pLysS cells without any recombinant pT7 plasmid and BL-21 (DE3) pLysS with pT7 plasmid without any insert) was performed with addition of IPTG to a final 1.0 mM concentration in each culture. Hourly samples were collected from each culture. Each sample was prepared for SDS-PAGE analysis on 12% discontinuous polyacrylamide gels. Gels were stained with Coomassie blue and examined for evidence of successful recombinant protein expression. Although the SDS-PAGE gel corresponding to T7-VP3 clones showed the appearance of a new protein band at ~30 kD (Figure 28), there were no new protein bands corresponding to the expected size of VP1 (~33 kD) on the SDS-PAGE gel representing T7-VP1 clone (not shown).
Figure 28. SDS-PAGE analysis for recombinant HAV VP3 protein expression from pT7-VP3 carrying \textit{E.coli} BL-21 (DE3) pLysS cells. Whole cell lysates prepared from various recombinant clones, 3 h after IPTG induction of T7 RNA polymerase expression in host cells, are shown. Lane 1: Molecular weight marker. Lane 2: T7-VP3 # 3a. Lanes 3 and 4: Host cells carrying pT7 vector without any insert DNA. Lanes 5 and 6: T7-VP3 # 4b and 4c. Lane 7: Host cells carrying recombinant pT7-KAN (Clone # 6; Positive control). Lane 8: Molecular weight marker.
Cloning of insert SurfVP1 DNA into the SurfZAP vector was performed successfully. PCR amplified insert SurfVP1 DNA (978 bp) was digested sequentially with the restriction enzymes, NotI and SpeI. Ligation reactions to generate concatameric recombinant lambda DNA suitable for packaging into phage particles were set up at insert:vector ratios of 1:1, 2:1 and 4:1 using the double-digested insert and the vector DNA arms, supplied predigested by the manufacturer. Following overnight incubation at 4°C, one µl of each ligation reaction was packaged into infective λ phage particles by using the Gigapack® II Gold packaging extracts. Next, each packaged reaction was used in preparation of serial dilutions in SM buffer. About 1-10 µl of each dilution was then used to infect transduction competent E.coli XL1-Blue host cells to the phage titer from each packaging reaction. After incubation of each phage-bacteria mix, the phage infected bacteria were plated onto individually labeled NZY plates and incubated overnight to allow for development of recombinant phage plaques. Plaque counts were performed and phage titers ranging from ~ 1 X 10^5 - 1 X 10^7 recombinant plaques or plaque forming units (pfu)/µg of vector arms were obtained. The highest titer of recombinant plaques resulted from the ligation reaction containing a calculated insert SurfVP1:vector SurfZAP ratio of 2:1 and the lowest, at the ratio of 4:1. A negative control reaction corresponding to the vector self-ligation, which had been set up along with the test ligation and packaging reactions to help determine the potential background resulting from any undigested vector, resulted in fewer than 1 X 10^2 plaques.

Next, several recombinant λSurfVP1 plaques were selected for PCR verification of the recombinant nature of the corresponding phage particles. A miniprep procedure
was designed to extract \( \lambda \) DNA from 3 ml lysates of host \( E. coli \) XL1-Blue cells following infection with phage toothpicked from the isolated plaques. The isolated \( \lambda \) DNA was used as template DNA in individual PCR reactions to confirm the presence of insert SurfVP1 DNA. A total of 150 plaques were screened by this approach and 138 were identified as recombinants based on successful amplification of the SurfVP1 DNA. The results of eight such PCR reactions with eight different DNA templates derived from recombinant \( \lambda \) plaques by the lambda miniprep procedure and as analysed by a 2% agarose gel electrophoresis are shown in Figure 29.

Ten recombinant \( \lambda \)SurfVP1 isolates, RJS3, RJS4, RJS5, RJS7, RJS8, RJS10, RJS11, RJS12, RJS13 and RJS 16, each identified by PCR as mentioned previously, were chosen for single clone \textit{in vivo} excision of recombinant phagemid, pSurfVP1 from within the larger \( \lambda \)SurfVP1 DNA. Recombinant \( E. coli \) SOLR colonies on LB-Amp agar medium resulting from the first round of this experiment were counted to determine the titer of phagemid lysates following \textit{in vivo} excision. Titers which ranged from 0.5-1.0 \( \times 10^5 \) cfu/ml were obtained from individual phagemid lysates derived from corresponding \( \lambda \)SurfVP1 phage progenitors. Since each \( \lambda \)SurfVP1 phage was used at a concentration of \(~2 \times 10^6\) pfu/\textit{in vivo} excision reaction to generate phagemid lysates from \( E. coli \) XL1-Blue cells, the resulting phagemid titer represented one excised and rescued phagemid for every 20 to 40 input \( \lambda \)SurfVP1 phage or an \textit{in vivo} excision efficiency of 2.5-5.0 %.

Several Amp\(^{\prime}\) colonies of recombinant \( E. coli \) SOLR cells were selected for plasmid DNA extraction to isolate recombinant pSurfVP1 DNA as a prelude to
Figure 29. Agarose gel analysis of PCR products to verify the presence of insert SurfVP1 DNA during screening for recombinant λSurfVP1 plaques. Lane 1: BioMarker low molecular weight DNA standard, Lanes 2-6 and 8-9: insert SurfVP1 DNA (945 bp) amplified from individual recombinant λ DNA templates using HAVVP1 specific primers, SurfVP1PS1 and SurfVP1NS. Lane 7 did not reveal any positive amplification of insert from the corresponding λ DNA template.
restriction analysis and PCR studies to confirm the presence of insert DNA. Plasmid 
DNA was isolated by two different miniprep procedures, alkaline lysis (Birnboim, and 
Doly, 1979) and PERFECTprep™ (5 Prime → 3 Prime, Inc., Boulder, CO). A 2% 
agarose gel electrophoretic analysis of plasmid DNA isolated by the alkaline lysis method 
is shown in Figure 30. Ethidium bromide stained bands on the agarose gel revealed 
DNA of higher than the anticipated molecular weight for pSurfVP1 (~4.5 kb). Figure 
31 shows a 2% agarose gel electrophoretic analysis of plasmid DNA samples isolated by 
the PERFECTprep™ procedure. The PERFECTprep™ procedure yielded plasmid DNA 
that was relatively free of contaminating genomic DNA as compared with the alkaline 
lysis method. However, the molecular size of the plasmid DNA samples (>12 kb) 
isolated by this procedure was several fold greater than that expected. The precise 
reasons for this phenomenon remained unclear at this juncture. Therefore, while it was 
declared to repeat the in vivo excision experiment again, alternatives were also 
considered.

Recombinant λ phage DNA were isolated from 100 ml lysates of ten λSurfVP1 
isolates by using the λQUICK!™ kit. The ten isolates were the same as those used in the 
earlier in vivo excision experiment namely, RJS3, RJS4, RJS5, RJS7, RJS8, RJS10, 
RJS11, RJS12, RJS13 and RJS16. Using the DNA samples isolated as template DNA, 
a general strategy of employing long range PCR with a mixture of thermostable DNA 
polymerases was developed to precisely amplify and subclone the entire phagemid DNA 
sequences in vitro from within the larger λSurfZAP or similar vectors, referred to as in 
vitro Excision-PCR. Primers complementary to the f1 phage initiator and terminator
Figure 30. Agarose gel analysis of plasmid pSurfVP1 DNA extracted from recombinant *E. coli* SOLR cells by the alkaline lysis method following the *in vivo* excision procedure. Lanes 1, 11, 21 and 31 contained the one KB DNA ladder while lanes 2-10, 12-20, 22-30 and 32-39 contained plasmid DNA samples from individual recombinant clones.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 31. Agarose gel analysis of plasmid pSurfVP1 DNA extracted from recombinant *E. coli* SOLR cells by the PERFECTprep™ procedure following the *in vivo* excision experiment. Lane 1: One KB DNA ladder, Lanes 2-8: Plasmid DNA isolated from different recombinant clones.
sequences that constitute the essential and terminal flanking sequences of phagemid pSurfscript DNA in its linear form within the λSurfZAP genome were designed. These primer sequences, the positive strand FIORIPS and the negative strand FIORINS, are shown in Figure 21. Both primers contain the recognition sequence for the restriction enzyme, EcoRV on the 5' side of the f1 sequences so that the double-stranded phagemid DNA, if successfully amplified using these primers, would contain identical restriction enzyme sites at both ends. Such a PCR amplified product could then be subjected to EcoRV digestion and self-igation and thereby recircularized into a double-stranded plasmid form capable of transforming appropriate host E.coli. We chose the DNA from the λSurfVP1 clone RJS10 as template DNA to optimize the conditions for the in vitro Excision-PCR reaction using the Expand™ Long Template PCR system. Three separate in vitro Excision-PCR reactions were set up differing only in the total amounts of the enzyme mix consisting of thermostable Taq and Pwo DNA polymerases. The volumes of the thermostable enzyme mix tested were 0.5 μl, 0.75 μl and 1.0 μl. A control PCR in which the Amplitaq DNA polymerase replaced the thermostable enzyme mix was also set up. The results from this pilot experiment were analyzed by 0.6% Agarose gel electrophoresis (Figure 32). The PCR reactions containing 0.75 μl and 1.0 μl of the thermostable enzyme mix as well as the control reaction with Amplitaq alone yielded products within the anticipated size (≈4.5 kb). However, the reactions performed with the thermostable enzyme mix gave better yields when compared to the reaction with Amplitaq alone. In addition, we also noted that there was little amplification of the desired product with 0.5 μl of the enzyme mix. Therefore, 0.75 μl of the thermostable
Figure 32. Agarose gel analysis of pSurfVP1 DNA amplified by \textit{in vitro}-Excision PCR using \textlambda Surf-VP1 template DNA. Lane 1: One KB DNA ladder, Lanes 2-4: PCR amplified DNA samples from reactions containing 0.5 \textmu l, 0.75 \textmu l and 1.0 \textmu l of the Expand\textsuperscript{TM} thermostable enzyme mix of Pwo and Taq DNA polymerases, and Lane 5: DNA from reaction containing 0.5 \textmu l of AmpliTaq DNA polymerase (ref. Materials and Methods).
enzyme mix was used as a standard amount in the *in vivo* Excision-PCR reactions involving other \( \lambda \)SurfVP1 DNA templates. Subsequently, pSurfVP1 DNA was successfully amplified from all ten \( \lambda \)SurfVP1 DNA templates. Next, the PCR-generated linear pSurfVP1 DNA was purified and used as template DNA along with SurfVP1 specific primers in PCR reactions to verify that the DNA from the *in vitro* Excision-PCR did contain the SurfVP1 insert and it was in fact the recombinant pSurfVP1 DNA. The results of the PCR verification of pSurfVP1 generated by the *in vitro* Excision-PCR are shown in Figure 33.

The purified, linear pSurfVP1 DNA samples were subjected to restriction digestion with *EcoRV* and purified again by the GENECLEAN II procedure. Then, ligation reactions were set up to circularize the plasmid and the products from the ligation reaction used to transform *E.coli* BL-21 host cells by electroporation. Several recombinants were chosen randomly for plasmid DNA extraction by the PERFECTprep™ procedure prior to restriction analysis and PCR to confirm their recombinant status and also the success of the *in vitro* excision experiment. One clone, # 26, was chosen for further studies.

The *in vivo* excision experiment was repeated again, during the time that we were engaged in developing the *in vitro* Excision-PCR experiment as a viable alternative if necessary. The plasmid DNA samples extracted from the various recombinant *E.coli* SOLR clones that were obtained from this round of *in vivo* excision and rescue of the pSurfVP1 phagemid, were analysed by 0.8% agarose gel electrophoretic analysis following restriction digestion with *EcoRI*. The results of a comparative analysis of
Figure 33. Agarose gel analysis of PCR products to verify the presence of insert SurfVP1 DNA in pSurfVP1 following in vitro-Excision PCR. Lane 1: 100 bp DNA marker, Lanes 2-5: SurfVP1 DNA amplified using different in vitro amplified and genecleaned pSurfVP1 DNA templates.
pSurfVP1 DNA from clone #2 obtained by the in vivo excision method and clone #26 obtained by the in vitro Excision-PCR method are shown in Figure 34. The recombinant nature of each plasmid was subsequently confirmed by PCR with SurfVP1 specific primers.

Recombinant E.coli clones which carry pSurfVP1 were subjected to 1.0 mM IPTG induction for production of recombinant VP1 fusion protein, according to procedures described earlier for expression of β-galactosidase fusions (Chapter 1, this dissertation). Whole cell lysates were prepared and analyzed by 12% discontinuous SDS-PAGE. No new protein bands corresponding to the expected size of recombinant Surf-VP1 protein (~57 kD) were detected following Coomassie blue staining of the SDS-PAGE gel.

The HAV nucleotide sequences encoding VPO, VP3 and VP1 capsid proteins were analysed for presence and frequency of rare codons with respect to the protein synthetic machinery of the host E.coli cells which may have resulted in a biased codon usage, since protein expression from cells containing pT7-VP1 or pSurf-VP1 plasmids was poor even after IPTG induction. The results of the computer-aided analysis of codon usage are tabulated in Table 8 and Table 9. The HAV sequences for VP1 and VP3 display a preponderence of codons corresponding to the rarer tRNA species in E.coli for amino acids Ser, Pro, Ala, Arg, Leu, Val, Thr, Glu and Gly. In the sequence for VP3 (246 aa), there are a total of 134 codons representing the nine amino acids listed previously, of which 114 represent rarer codons. Similarly, in case of VP1(294 aa), out of a total 183 codons for the same set of nine amino acids, 163 are represented by rarer
Figure 34. Agarose gel analysis of plasmid pSurfVP1 DNA extracted from recombinant clones # 2 and # 26. Clone # 2 was obtained by *in vivo* excision and Clone # 26, by *in vitro* excision and subsequent transformation of appropriate host cells. Lane 1 and 2: Undigested and *EcoRI* digested pSurfVP1 DNA from clone # 2 respectively, Lane 3 and 4: Undigested and *EcoRI* digested pSurfVP1 DNA from clone # 26. Lane 5: One KB DNA ladder.
Table 8. A comparison of codon usage for selected amino acids in HAV VP genes

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon*</th>
<th>Frequency of Codons used per thousand*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E. coli</td>
</tr>
<tr>
<td>Leucine</td>
<td>UUA</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td>UUG</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>CUU</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>CUC</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>CUA</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>CUG</td>
<td>51.8</td>
</tr>
<tr>
<td>Valine</td>
<td>GUU</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td>GUC</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>GUA</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>GUG</td>
<td>25.4</td>
</tr>
<tr>
<td>Serine</td>
<td>UCU</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td>UCC</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>UCA</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>UCG</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>AGU</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>15.4</td>
</tr>
<tr>
<td>Proline</td>
<td>CCU</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>CCC</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>CCA</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>CCG</td>
<td>22.8</td>
</tr>
<tr>
<td>Threonine</td>
<td>ACU</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>ACA</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>ACG</td>
<td>13.4</td>
</tr>
<tr>
<td>Alanine</td>
<td>GCU</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td>GCC</td>
<td>24.9</td>
</tr>
<tr>
<td></td>
<td>GCA</td>
<td>20.4</td>
</tr>
<tr>
<td></td>
<td>GCC</td>
<td>32.3</td>
</tr>
<tr>
<td>Glutamate</td>
<td>GAA</td>
<td>40.3</td>
</tr>
<tr>
<td></td>
<td>GAG</td>
<td>19.1</td>
</tr>
<tr>
<td>Arginine</td>
<td>CGU</td>
<td>21.3</td>
</tr>
<tr>
<td></td>
<td>CGC</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>CGA</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>CGG</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>AGA</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>AGG</td>
<td>1.7</td>
</tr>
<tr>
<td>Glycine</td>
<td>GGU</td>
<td>25.6</td>
</tr>
<tr>
<td></td>
<td>GGC</td>
<td>29.0</td>
</tr>
<tr>
<td></td>
<td>GGA</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>GGG</td>
<td>10.7</td>
</tr>
</tbody>
</table>

*Preferred codons in E. coli are underlined; *Codons for a given amino acid with the highest frequency of usage are depicted in bold.
Table 9. A comparison of codon usage frequency in *E. coli* for select amino acids encoded in the capsid protein genes of HAV.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th># of Preferred codons/Total # of codons for the given amino acid in the entire gene</th>
<th>Frequency (per thousand) of usage for preferred codons/Total frequency (per thousand) of all codons for the given amino acid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HAV VP0 (245 aa)</td>
<td>HAV VP3 (246aa)</td>
</tr>
<tr>
<td>Leucine</td>
<td>4/24</td>
<td>1/15</td>
</tr>
<tr>
<td>Valine</td>
<td>4/20</td>
<td>5/20</td>
</tr>
<tr>
<td>Serine</td>
<td>0/19</td>
<td>0/17</td>
</tr>
<tr>
<td>Proline</td>
<td>0/10</td>
<td>0/13</td>
</tr>
<tr>
<td>Threonine</td>
<td>2/20</td>
<td>4/23</td>
</tr>
<tr>
<td>Alanine</td>
<td>0/14</td>
<td>0/17</td>
</tr>
<tr>
<td>Glutamate</td>
<td>4/12</td>
<td>5/7</td>
</tr>
<tr>
<td>Arginine</td>
<td>0/8</td>
<td>1/11</td>
</tr>
<tr>
<td>Glycine</td>
<td>10/18</td>
<td>4/11</td>
</tr>
</tbody>
</table>

Note: Data pertaining to those amino acids for which the preferred codon usage in HAV VP1 or VP3 is zero are depicted in bold.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
codons. A complete codon frequency table for comparison of codon usage for all amino acids in *E.coli* is provided in Appendix.

It is interesting to note that only 111 of the 294 codons for VP1, 111 of 246 for VP3 and 104 of 245 for VP0 are represented by the preferred codons of *E.coli* translational machinery. The codon bias indices (codon bias index is defined as fraction of total codons in a given gene that are preferred codons; de Boer and Kastelstein, 1986) were calculated for HAV VP1, VP3 and VP0 and corresponded to 0.38, 0.45 and 0.43 respectively. Genes which are highly expressed in *E.coli* such as those encoding elongation factor Tu (EF-Tu; 200,000 to 1,000,000 copies per cell), and ribosomal proteins L7 and L12 (140,000 to 300,000 copies per cell) have a codon bias index of 0.84 whereas that encoding lac repressor (10 copies per cell) has a codon bias index of 0.18 (Bennetzen and Hall, 1982).

**DISCUSSION**

Our earlier studies, as outlined in chapters I and II, involving expression of HAV capsid proteins as *E.coli* β-galactosidase fusions were only moderately successful. Attempts to improve the transcriptional efficiency in the gene region concerning the cloned VP1 segment on the recombinant plasmid vector, pAXBoxAVP1 yielded unexpected results incompatible with the overall objectives of our study. Subsequently, we decided to attempt cloning of HAV genes in a non-fusion format into appropriate *E.coli* plasmid vectors and try to express the viral proteins as close as possible to their native form in both sequence and structure. At the same time, we also have explored approaches to find a suitable, inexpensive and safe alternative to the time-consuming and
costly tissue-culture method to produce HAV particles for use in antibody production, to develop assays for fast and reliable detection from clinical and environmental specimens and for potential use in vaccine production.

Several attempts were made to clone and express HAV capsid genes from vector pSE420 (Invitrogen Corporation, San Diego, CA) (i) as proteins without any fusion component and (ii) as epitope-labelled proteins carrying multiple myc tags (Roth, et al., 1991) without much success. The frequency of transformation was very low (greater than a 100-fold reduction as compared with transformantion with vector only) and no stable recombinants were recovered following transformation of host E. coli strains, GW1000, UT2300 or Top10. We also tried to clone a XbaI-Xhol restriction enzyme digested fragment from pGEM-HAV representing the complete HAV ORF, except portions of VP4 and 3Dpol genes, into vector pBC (Stratagene, La Jolla, CA). hoping that if successful, this would permit us to generate ssHAV RNA for use in assays to develop protocols for detection of HAV from shellfish. Again, we observed that the frequency of transformation was very low and any recombinants recovered following several individual cloning and transformation experiments were not stable. Cultures inoculated with the few recombinants from these experiments showed poor and very slow growth as seen by the naked eye and upon phase contrast microscopy, revealed signs such as increased filamentation, inclusion body formation and cell death. All attempts to isolate the recombinant plasmids from these cells were futile. These findings strongly suggested the possible toxicity of cloned HAV genes or their products, prematurely expressed, to host E. coli.
Therefore, we decided to employ a phage T7-promoter based vector-host system which would enable a controlled exclusive expression of the cloned HAV genes and thereby address the question if premature expression of native HAV capsid proteins was indeed responsible for the earlier findings. We also have used a phage λ-derived vector-host system in this study in parallel with the phage T7 system as an alternative approach to circumvent problems in cloning, to successfully amplify and recover the recombinant DNA prior to attempts at finding appropriate \textit{E. coli} host strains that would enable their stable maintainence and expression. The latter approach had the extended appeal of allowing \textit{in vivo} excision of the recombinant DNA into a plasmid form, expression of recombinant HAV proteins on \textit{E. coli} surface and of generating single-stranded, recombinant filamentous phage particles that display the HAV capsid proteins on their surface while simultaneously containing the corresponding gene sequence within.

\textit{Escherichia coli} BL-21 host cells were transformed successfully with recombinant pT7-VP1 and pT7-VP3 plasmids at a reasonable frequency of transformation. Since there is was known source of T7 RNA polymerase within these host cells, any premature expression of the HAV genes was not possible and may have aided in a more efficient transformation and stable maintainence of the recombinant pT7-VP1 and pT7-VP3 plasmids. Host \textit{E. coli} cells with these recombinant plasmids, upon phase contrast microscopy, displayed no appreciable differences in their morphology as compared to the control cell populations (i.e. cells without and with vector pT7-7). Also, plasmid DNA extraction from the recombinant cells was fruitful in contrast to our earlier attempts with other vector/host systems used for expression of HAV capsid proteins in a non-fusion
format. The presence of the cloned insert DNA in several of the recombinant plasmids was confirmed by restriction digestions and by PCR.

The recombinant plasmids were then transformed into host BL-21(DE3)pLysS cells which carried the gene for T7 RNA polymerase under the inducible transcriptional control of the lacUV5 promoter and thereby could allow for a controlled expression of the HAV genes. Subsequently, cells carrying pT7-VP3 plasmids displayed a new protein band corresponding in size to that of HAV VP3, following an IPTG induction of T7 RNA polymerase expression in the host cells. However, none of the recombinant clones screened similarly for expression of VP1 showed any new bands on SDS-PAGE which was intriguing. Also, Western blot analyses of IPTG-induced bacterial cell lysates of host cells carrying pT7-VP3 and pT7-VP1, did not show any recognition by the anti-HAV rabbit polyclonal or mouse monoclonal antibodies of protein bands at the expected size (data not shown). Possible explanations for these observations include poor translation and post-translational degradation by host proteases in the case of VP1, and lack of proper conformation due to improper folding in vivo or due to denaturation induced by SDS during the analysis in the case of VP3.

A poor translation of mRNA encoding HAV VP1 may have been due to nonsense and missense mutations introduced into the DNA encoding VP1 during PCR amplification with Taq DNA polymerase prior to cloning. Instability of target mRNA could be another reason for poor expression of protein of interest although previous reports suggest that in a T7 RNA polymerase dependent expression system, amounts of stable mRNA produced are not only sufficient but substantial enough to actually
overwhelm the normal mRNA degradation mechanisms of host *E. coli* (Studier, and Moffatt, *et al.*, 1986). Still, factors such as an unfavorable distribution of rare codons and interfering secondary structures within the target mRNA which can result in ribosomal stalling, and translational frameshifting which can lead to premature termination of translation, cannot be ruled out and may contribute to poor translation of the target mRNA. In addition, the possibility exists that VPI protein, in whatever amounts it was expressed, may have been toxic leading to an induction of a heat-shock like response in host cells resulting in rapid breakdown of the expressed protein. The host cells, although deficient in *lon* and *ompT* proteases, still possess other proteases capable of degrading abnormal and toxic proteins (Goldberg, and Goff, 1986).

Cloning of HAV VP1 DNA sequence into the λSurfZAP™ vector and the subsequent subcloning into the phagemid pSurf-VP1 form were accomplished concurrently with our attempts to clone HAV capsid genes into the pT7-7 vector. During the process, we developed a miniprep procedure for screening of recombinant plaques and an alternative to the *in vivo* excision and rescue of phagemid pSurfVPI DNA as outlined in the materials and methods earlier. Both procedures also involved different applications of the ubiquitous PCR.

The miniprep procedure was designed to enable simultaneous extraction and preparation of template recombinant phage λ DNA from several recombinant plaques following inoculation of multiple, small scale (3 ml) broth cultures of phage-sensitive host cells. The standard protocols in common use for preparing λ DNA from small scale liquid lysates require at least 50-ml liquid phage lysates per plaque screened, and involve
prolonged ultracentrifugation steps (Davis, et al., 1980; Ausubel, et al., 1992; Sambrook, et al., 1989), both of which were avoided in our miniprep procedure. Recombinant phage DNA was successfully recovered from 20 recombinant phage lysates at a time by the miniprep procedure. A small volume of each DNA preparation (two μl out of a final 100 μl obtained from each three ml lysate) was found sufficient to act as template DNA in the PCR reactions that were carried out subsequently to confirm the presence of the cloned insert DNA. Thus, the combined miniprep procedure and PCR not only enabled a rapid and specific screening for recombinant phage but also helped reduce the amounts of reagents needed significantly.

Following confirmation of the recombinant nature of several λSurf-VP1 plaques, we attempted in vivo excision and rescue of recombinant phagemid pSurf-VP1 DNA (Short, et al., 1988; Short and Sorge, 1992). In our first attempt at excision and rescue of the recombinant phagemids, we utilized ten different λSurf-VP1 phage lysates and obtained several recombinant E.coli SOLR™ colonies on LB-amp medium. However, plasmid DNA extracted from these cells revealed DNA bands at higher (> 12 kb) than anticipated (4.5 kb) molecular size for the recombinant phagemid. Also, the PERFECKTprep™ procedure for plasmid DNA extraction yielded DNA samples that appeared relatively free of protein and chromosomal DNA contamination when compared with those obtained by alkaline lysis procedure (Figure 30 and Figure 32). Since the reasons for the apparent increase in the observed molecular weight of the recombinant plagemid DNA were unclear, we decided to repeat the experiment and also to develop alternative protocols for recovering phagemid DNA successfully. Upon repeating the in
vivo excision and rescue procedure, we recovered several recombinants which were verified to contain pSurfVP1 by restriction analysis and by PCR. Selected recombinants were then used in induction experiments to study expression of surface-directed fusion protein containing HAV VP1.

The in vitro-Excision PCR procedure, as outlined schematically in Figure 25a and b, was developed as a feasible alternative to the in vivo excision and rescue procedure. The entire pSurf-VP1 DNA (~4.5 kb) including the f1 phage-derived, flanking initiator (I) and terminator (T) sequences was amplified as double-stranded linear DNA by "long PCR" (Cohen, 1994). A mixture of thermostable DNA polymerases, Taq polymerase (without 3'→5' exonuclease or proofreading activity) and Pwo polymerase (with active proofreading capability), was used to ensure a high yield of target DNA without compromising the fidelity of DNA replication (Barnes, 1994; Cheng, et al., 1994). The restriction enzyme site, EcoRV, was engineered into the flanking primers used in this PCR so that restriction digestion and ligation reactions could be performed at a later time on the resulting linear DNA to generate a covalently closed circular pSurfVP1 DNA. Neither the insert VP1 nor the parent phagemid pSurfscript DNA inherently contain an EcoRV site. Following a successful amplification, the pSurfVP1 DNA was circularized by EcoRV digestion and ligation with T4 DNA ligase and used to transform host E.coli cells by electroporation. Plasmid DNA extracted from the resulting recombinants was confirmed by restriction analysis and PCR for carrying the previously cloned insert VP1.
DNA. Thus, this procedure enabled the subcloning of DNA of interest from within the larger λSurfVP1 (~45 kb) into the smaller pSurfVP1 phagemid (~4.5 kb). The intricate essential steps of in vivo excision and rescue procedure, such as the need to use of two or more different E.coli host strains (XL-Blue for excision and SOLR™ for single-stranded rescue of phagemid DNA) and ExAssist helper phage, and the necessity to transform the recombinant plasmids following rescue into new host strains later on since SOLR cells do not permit single-standed phagemid replication, are bypassed completely by the in vitro-Excision procedure which is based on the long PCR. In addition, the in vitro-Excision PCR also allows amplification of large amounts of the complete phagemid portion of the popular λZAP vectors or its derivatives including λSurfZAP regardless of the size of any cloned foreign DNA inserts they might carry, in a relatively inexpensive and time-saving manner and employing a common set of primers directed towards the phage f1 I and T sequences. Although the in vivo excision and rescue may be very efficient for mass excision of libraries prepared in lambdaZAP vector or its derivatives, the in vitro-PCR may be more convenient and advantageous in situations where a precise subcloning of only a single or few λ-derived clones is desired or where the protein product of the cloned sequence may be too toxic to certain host E.coli cells. In the latter instance when subcloning based solely on in vivo methods may be difficult if not impossible, the in vitro-Excision PCR method also will be useful for generating large amounts of recombinant phagemid DNA in vitro from the parent λ-derived clones prior to a simultaneous transformation and screening of multiple host E.coli strains to select the ideal one for an optimal expression of the target gene on the relevant phagemid DNA.
Recombinant *E. coli* clones containing pSurf-VP1, obtained by either one of the two approaches discussed earlier, were subjected to induction of target protein by IPTG. However, our efforts have not been successful so far. The recombinant BL-21 host cells under IPTG induction were examined by phase contrast microscopy and did not exhibit any discernible changes in morphology when compared with controls that did not carry any plasmids. Western blot analysis of whole cell lysates of recombinant cells subjected to IPTG induction did not reveal any bands of recognition with the various anti-HAV antibodies tested (Rabbit, mouse and human). The reasons for such a poor expression concerning HAV VP1 remain obscure although it is noteworthy that a similar problem was encountered earlier during our attempts to express VP1 from other prokaryotic expression vector systems such as the T7 promoter/T7RNA polymerase expression system. Hence, we decided to explore the likelihood that an inefficient codon usage pattern concerning HAV gene expression in *E. coli* may have been responsible for poor levels of expression. However, it is possible that other factors such as rapid mRNA degradation, potential toxicity of VP1 to host cells leading to induction of host proteases, and errors introduced into the coding sequences by PCR may also produce similar effects. It has been suggested previously that the *Taq* polymerase (which does not possess any 3'-5' exonuclease activity) may have an error rate of ~1 mismatched base per 10-50,000 bases added (Cheng, *et al.*, 1994) and may significantly affect the fidelity of amplification of sequences longer than 10 kb when used alone.

The study of codon usage profile for HAV gene expression in *E. coli* was based on Grantham's genome hypothesis, which states that taxonomically related organisms
have similar codon usage (de Boer and Kastelein, 1986). Previously reported observations have shown that each gene in a given organism conforms to a distinct codon usage pattern of the species and exhibits a bias regarding the choice of synonymous codons consistent with the other genes from the same or closely related genomes (Grantham, et al., 1980a, b). Highly expressed genes (such as membrane and ribosomal proteins) use only a limited number of available codons as compared to poorly expressed genes (regulatory proteins like lac repressor) and the degree of codon usage bias for many genes in various organisms, including E.coli, is directly related to the cellular levels of their protein products (Gouy and Gautier, 1982; Benetzen and Hall, 1982).

In E.coli, there is a strong and interesting correlation between the codon usage profiles of highly expressed genes and the relative abundance of each tRNA species. The tRNA levels corresponding to codons preferred frequently are higher than the tRNA species for rarer codons and highly expressed genes display a strong preference for codons corresponding to major tRNA species and avoid the use of minor tRNAs. Conversely, genes expressed at low levels tend to exhibit a relatively frequent use of minor codons and the availability of the corresponding minor tRNA species may be a rate-limiting factor in the translation of such genes (Ikemura, 1981a, b).

The codon usage database, searchable by name of organism through the World Wide Web, contains all reported data pertaining to codon usage in individual genes from various species. Codon usage data contained in this database is for complete genes only and is calculated using the nucleotide sequence data obtained from the GenBank Genetic Sequence Database. In addition, the compilation is synchronized with each major release
of Genbank (Nakamura, et al., 1997). Using the codon usage database, we obtained individual codon usage frequency tables for HAV, E.coli, Human poliovirus 3 and Human protein coding sequences as a sum of the codon usage for each species. The study of codon usage profile in E.coli for HAV sequences revealed several interesting features (Table 8 and 9). In the HAV VP1 sequence, only 111 of 294 codons are codons recognized by major tRNA species in E.coli. Similarly, the VP3 sequence utilizes only 111 preferred codons in a total of 246 for the protein. There are no preferred codons corresponding to amino acids, Ser, Pro, Ala and Arg, in either VP1 or VP3 sequences. In addition, the amino acids, Leu, Val, Thr, Glu, and Gly are encoded mostly by rarer codons in both genes. The codon bias indices calculated for VP1 (0.38) and VP3 (0.45) correspond to genes that are expressed at low levels in E.coli such as the RNA polymerase β subunit (codon bias index=0.53; cellular levels estimated to be 7000 to 15000 protein molecules per cell). However, it has to be noted that in case of HAV VP1 and VP3 sequences that are carried on multicopy expression vectors such as pT7, mRNA levels following induction will be several fold higher compared with those of the E.coli genes and may contribute significantly to the total levels of protein that can be expressed. Still, if tRNA availability is the rate-limiting factor during translation, the expression of HAV genes could be adversely affected despite an abundance of target mRNA within the E.coli, since shortage of appropriate, charged tRNA species can lead to ribosomal pausing, rapid mRNA degradation in the regions not protected by the translating ribosomes and also an increase in the misincorporation rate during translation. This could result in prematurely truncated proteins, or proteins with miscorporated amino
acids, both of which may assume improper conformation and can be recognized as abnormal by the host cell protein degradation machinery. In addition, the abnormal proteins could provoke a host cell stress response similar to that observed during heat shock, resulting in their selective degradation (Goldberg and Goff, 1986).

A comparison of the codon usage profile of *E.coli*, HAV, Human poliovirus and Human protein coding sequences for the selected amino acids which are preferentially encoded by rarer codons in HAV (Table 8), shows that the biased codon usage by Human Poliovirus 3 is more closely related (such as for amino acids Leu, Val, Thr, Ala, and Gly) to those of human and *E.coli* profiles than that exhibited by HAV. Although the significance of this finding needs further clarification, it is tempting to speculate the correlation between the relative ease of propagation of poliovirus in tissue culture and the expression of its proteins in *E.coli* as compared to the difficulties faced with HAV.
CONCLUSIONS

Cloning and Expression of the Capsid Proteins of Hepatitis A Virus in *Escherichia coli* as β-Galactosidase Fusion Products: The cDNA regions encoding Hepatitis A Virus (HAV) capsid proteins VP1, VP2 and VP3 were individually amplified by PCR and cloned successfully into prokaryotic expression vector, pAX4b+, for expression of recombinant HAV proteins as β-galactosidase fusions in host *Escherichia coli* cells. Recombinant protein expression, upon IPTG-induction, was comparatively better in host cells deficient in Lon protease (UT2300 strain) than in those that were not (GW1000 or TOP10 strains). Microscopic studies of IPTG-induced host cells revealed the formation of intracellular inclusion bodies. The inclusion bodies containing recombinant proteins were extracted and partially purified by employing a French Press lysis-Sarkosyl extraction procedure. The recombinant fusion proteins, VP1III, VP2B2 and VP3T3, exhibited antigenic properties when assessed by ELISA and Western Blot assays with anti-HAV polyvalent and monoclonal antibodies generated in animals immunized with intact HAV HM-175 particles.

However, the expression of recombinant HAV proteins and the intended affinity-based purification of these proteins (~150 kD) were complicated by an apparent co-expression of β-galactosidase (~125 kD). Factors that needed to be studied as possibilities for the simultaneous appearance of both β-galactosidase and the recombinant β-gal-VP fusion proteins in host cells following their exposure to IPTG included premature termination of transcription, rapid decay of recombinant mRNA, inefficient translation due to a biased codon usage, autoproteolytic or host-protease...
mediated cleavage of nascent recombinant fusion proteins. Some of these concerns were addressed in the later studies. Further, the anti-HAV antibodies cross-reacted with the β-galactosidase during Western blot analysis, suggesting a putative sharing of epitopes in the SDS-mediated denatured state between HAV capsid proteins and *E. coli* β-galactosidase which are apparently unrelated to one another.

**Effects of an Anti-terminator Sequence on Recombinant HAV Protein Expression in *Escherichia coli***: Premature termination of transcription was considered as one of the chief factors that may have been responsible for the observed phenomenon of co-expression of *E. coli* β-galactosidase and the β-gal-VP fusion proteins in recombinant host cells. Hence, the hypothesis that inclusion of an anti-terminator sequence upstream of HAV VP coding sequences might result in full-length transcripts and thereby render translation of fusion protein more efficient to yield the recombinant VP1 protein (~150 kD) without a simultaneous expression of the non-recombinant *E. coli* β-galactosidase (~125 kD), was tested. The BoxA antiterminator sequence from the *E. coli* *rrn* operon was incorporated upstream of the HAV VP1 and cloned into pAX4b+ vector. However, induction of fusion protein expression from host cells carrying pAXBoxAVP1 did not lead to any discernible expression of HAV VP1 fusion protein. Rather, in addition to the β-galactosidase (~150 kD), a novel, 33 kD protein band was observed upon SDS-PAGE analysis of IPTG-induced recombinant whole cell lysates. A limited N-terminal aminoacid sequence determination of the ~33 kD protein was obtained, and a computerized search and comparison with previously reported sequences in the protein databases was performed. The results revealed that the ~33 kD protein

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
shared varying degrees of homology with a host of protein sequences from various sources. Significant homology was discovered between the ~33 kD protein, N-terminal sequences from the leading regions of conjugative plasmids, F and R1, and putative open reading frames (gene YFJQ and Rhs elements) found in *E.coli* chromosomal DNA which display characteristics of DNA-binding proteins. However, the gene responsible for expression of the 33 kD protein in the recombinant *E.coli* cells carrying plasmid pAXBoxAVP1 was obscure and remains to be identified. In addition, the reasons for an apparent lack of expression of recombinant HAV VP1 fusion protein in the same cells were unclear. While it is tempting to speculate that the expression of the 33 kD protein with putative DNA-binding properties may have been a result of a generalized stress (SOS) response in host cells involving events such as overexpression and degradation of seemingly toxic recombinant HAV proteins, further investigations are essential to definitively resolve this mystery.

**Alternative Approaches for Cloning of HAV Capsid Genes for Expression in *Escherichia coli*:** Since our earlier efforts at expression of HAV capsid proteins as *E.coli* β-galactosidase fusions were only moderately successful, we considered alternate strategies such as (i) cloning and expression of HAV capsid genes in a non-fusion format from a phage T7-promoter based vector-host system and (ii) to clone HAV genes for surface expression of recombinant proteins on host *E.coli* membrane and to obtain chimeric, single-stranded, filamentous coliphages which could potentially replace the tissue-culture grown HAV particles in diagnostic assays and in immunological studies.
Recombinant pT7-VP1 and pT7-VP3 plasmids were generated by following standard molecular biological techniques and host *E. coli* BL-21 cells were transformed successfully to yield several stable recombinant clones. Plasmid DNA samples from several of these recombinants were then utilized to transform host *E. coli* BL-21(DE3)pLysS cells as a prelude to induction of recombinant protein expression. However, subsequent experiments revealed that only those recombinants that carried pT7-VP3 plasmids expressed a new protein corresponding in size to that of HAV VP3 while recombinants carrying pT7-VP1 plasmids did not show any protein product corresponding to HAV VP1 upon SDS-PAGE and Western Blot analyses. It is possible that poor translation and/or post-translational degradation by host proteases could have contributed to the results observed. Nonsense or missense mutations could have been introduced into the cDNA encoding the HAV VP proteins by the error-prone *Taq* DNA polymerase (~1 mismatch per 10-50,000 bases added) during the PCR amplification step prior to cloning, resulting in a poor translation, premature termination of protein synthesis and formation of aggregates of truncated protein fragments that would trigger the host SOS response and subsequent protease-mediated degradation. Factors such as an unfavorable distribution of rare codons and translational frameshifting were also considered to potentially contribute towards a premature termination of translation of target mRNA.

Cloning HAV VP1 sequence into a phage λ derived vector, λSurfZAP™, was performed to obtain recombinants with potential for expression of HAV proteins on *E. coli* and filamentous coliphage surfaces. Novel procedures were designed and tested
successfully to meet the following objectives: (i) rapid preparation of template λ DNA templates from multiple, small scale (3 ml) broth cultures of phage-sensitive host cells as a prelude to a PCR-based verification of their recombinant nature and, (ii) subcloning phagemid DNA carrying the HAV VP1 insert as an alternative to the in vivo excision and rescue procedure. We have termed the latter procedure as in vitro Excision-PCR. Following subcloning of λSurf-VP1 clones, several recombinant host E.coli BL-21 cells carrying the plasmid pSurf-VP1 were obtained and verified by PCR to contain the HAV VP1 DNA. However, subsequent experiments to induce expression of recombinant protein were unsuccessful. This problem was similar to what we had encountered earlier in our efforts to express HAV VP1 from various prokaryotic expression systems. Hence, we investigated further the likelihood of an inefficient codon usage pattern as a possible reason for the apparent lack of recombinant protein expression.

Codon usage profiles in E.coli for HAV sequences were obtained by a computer-assisted search of the codon usage databases and compared with those for E.coli, Human poliovirus 3 and human protein coding sequences. The HAV VP1 and VP3 sequences showed a preponderance of rare codon usage for several amino acids, suggesting that codon usage bias in these sequences could have adversely affected their successful translation. A shortage of appropriate, charged tRNA species which occurs when rare codons are used at a higher than usual frequencies, can lead to ribosomal pausing, rapid mRNA turnover and an increase in the misincorporation rate of non-cognate amino acids during translation. All of these would result in prematurely truncated, abnormal proteins which would provoke a host cell response akin to that observed during heat
shock, resulting in their selective degradation. Hence, it appears that if HAVcapsid sequences have to be expressed efficiently in \textit{E.coli}, one has to seriously consider modifying the DNA sequences for VP1 and VP3 by replacing rarer codons with synonymous codons that are frequently employed in \textit{E.coli} and for which the corresponding tRNA species are present in relative abundance.
REFERENCES


162

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Plunkett, G. 1995. Hypothetical 31.3 KD Protein in ALPA-GABD Intergenic Region (O273). (Direct Sequence Submission to EMBL/GENBANK/DDBJ Data Banks)


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Sturrock, S.S. and Collins, J.F. 1993. MPsrch version 1.3., Biocomputing Research Unit, University of Edinburgh, UK


## APPENDIX

A Comparison of Codon Usage for HAV VP genes with that of *Escherichia coli* genes.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon(^a)</th>
<th>Frequency of Codons used per thousand(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E.coli</td>
<td>HAV VP</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>UUU</td>
<td>21.2</td>
</tr>
<tr>
<td></td>
<td>UUC</td>
<td>17.1</td>
</tr>
<tr>
<td>Leucine</td>
<td>UUA</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td>UUG</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>CUU</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>CUC</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>CUA</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>CUG</td>
<td>51.8</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>AUU</td>
<td>29.0</td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>25.4</td>
</tr>
<tr>
<td></td>
<td>AUA</td>
<td>5.2</td>
</tr>
<tr>
<td>Methionine</td>
<td>AUG</td>
<td>27.0</td>
</tr>
<tr>
<td>Valine</td>
<td>GUU</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td>GUC</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>GUA</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>GUG</td>
<td>25.4</td>
</tr>
<tr>
<td>Serine</td>
<td>UCU</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td>UCC</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>UCA</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>UCG</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>AGU</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>15.4</td>
</tr>
<tr>
<td>Proline</td>
<td>CCU</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>CCC</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>CCA</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>CCG</td>
<td>22.8</td>
</tr>
<tr>
<td>Threonine</td>
<td>ACU</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>ACA</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>ACG</td>
<td>13.4</td>
</tr>
<tr>
<td>Alanine</td>
<td>GCU</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td>GCC</td>
<td>24.9</td>
</tr>
<tr>
<td></td>
<td>GCA</td>
<td>20.4</td>
</tr>
<tr>
<td></td>
<td>GCG</td>
<td>32.3</td>
</tr>
</tbody>
</table>

\(^a\)Preferred codons in *E.coli* are underlined.

\(^b\)Codons for a given amino acid with the highest frequency of usage are depicted in bold
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Frequency of Codons used per thousand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>E. coli</em> HAV HAV HAV HAV Polio Human</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VP0 VP3 VP1 ORF virus</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>UAU</td>
<td>16.3 28.6 28.5 30.6 28.4 16.3 12.3</td>
</tr>
<tr>
<td></td>
<td>UAC</td>
<td>12.7  12.2 12.2  6.8  6.9  29.3 17.0</td>
</tr>
<tr>
<td>Histidine</td>
<td>CAU</td>
<td>12.3  32.7 20.3  13.6 19.8  8.4  9.7</td>
</tr>
<tr>
<td></td>
<td>CAC</td>
<td>10.0  8.2  0.0   3.4   4.3  14.1 14.6</td>
</tr>
<tr>
<td>Glutamine</td>
<td>CAA</td>
<td>14.1 40.8 24.4  6.8  21.5 22.1 11.5</td>
</tr>
<tr>
<td></td>
<td>CAG</td>
<td>29.2 24.5 28.5 17.0 23.3 16.4 33.7</td>
</tr>
<tr>
<td>Asparagine</td>
<td>AAU</td>
<td>18.0 28.6 48.8 37.4 37.9 20.9 16.7</td>
</tr>
<tr>
<td></td>
<td>AAC</td>
<td>22.1  8.2  4.1   6.8   6.9  27.5 20.9</td>
</tr>
<tr>
<td>Lysine</td>
<td>AAA</td>
<td>35.1 24.5 28.5 20.4 36.2 32.7 23.4</td>
</tr>
<tr>
<td></td>
<td>AAG</td>
<td>12.1  12.2 12.2 20.4 18.9 27.8 33.9</td>
</tr>
<tr>
<td>Aspartate</td>
<td>GAU</td>
<td>32.1 24.5 44.7 54.4 44.8 26.8 22.2</td>
</tr>
<tr>
<td></td>
<td>GAC</td>
<td>20.2  16.3 8.1  10.2  9.5  27.8 26.9</td>
</tr>
<tr>
<td>Glutamate</td>
<td>GAA</td>
<td>40.3 16.3 20.3 13.6 26.7 23.9 28.1</td>
</tr>
<tr>
<td></td>
<td>GAG</td>
<td>19.1 32.7 8.1  37.4 26.7 29.0 40.6</td>
</tr>
<tr>
<td>Cysteine</td>
<td>UGU</td>
<td>5.0  4.1  24.4  6.8 14.6  9.8  9.6</td>
</tr>
<tr>
<td></td>
<td>UGC</td>
<td>6.4  4.1  4.1  0.0   2.6  7.8 12.7</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>UGG</td>
<td>13.7 12.2 16.3 13.6 16.4 12.3 13.3</td>
</tr>
<tr>
<td>Arginine</td>
<td>CGU</td>
<td>21.3  0.0  4.1  0.0   0.9  3.6  4.7</td>
</tr>
<tr>
<td></td>
<td>CGC</td>
<td>21.5  0.0  0.0  0.0   0.0  5.4 11.0</td>
</tr>
<tr>
<td></td>
<td>CGA</td>
<td>3.5   0.0  4.1  0.0   0.9  2.0  6.0</td>
</tr>
<tr>
<td></td>
<td>CGG</td>
<td>5.4   0.0  0.0  0.0   0.0  1.8 11.3</td>
</tr>
<tr>
<td></td>
<td>AGA</td>
<td>2.7   24.5 20.3  37.4 30.1 22.8 10.9</td>
</tr>
<tr>
<td></td>
<td>AGG</td>
<td>1.7   8.2 16.3 10.2 11.2  8.9 10.9</td>
</tr>
<tr>
<td>Glycine</td>
<td>GGU</td>
<td>25.6 32.7 16.3 10.2 16.4 19.2 11.1</td>
</tr>
<tr>
<td></td>
<td>GGC</td>
<td>29.0  8.2  0.0 10.2  6.0  11.1 23.7</td>
</tr>
<tr>
<td></td>
<td>GGA</td>
<td>8.2   24.5 16.3  47.6 29.3 18.2 16.9</td>
</tr>
<tr>
<td></td>
<td>GGG</td>
<td>10.7  8.2 12.2  6.8   9.5  17.1 16.6</td>
</tr>
</tbody>
</table>

*aPreferred codons in *E. coli* are underlined.*

*bCodons for a given amino acid with the highest frequency of usage are depicted in bold.
VITA

Lingaiah Chandrashekar was born on September 13, 1965, in Bangalore, Karnataka, India. He graduated from National High School and College, Basavanagudi, Bangalore, in 1983. He attended Bangalore Medical College, Fort, Bangalore, between 1983 and 1989, and received his bachelor in medicine and bachelor in surgery (M.B.,B.S.) degree in December, 1989. In the fall of 1990, he joined the doctoral program in Microbiology at the Department of Microbiology, Louisiana State University, Baton Rouge, Louisiana, where he is presently a candidate for the doctor of philosophy degree in Microbiology with a minor in Biochemistry.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Lingaiah Chandrashekar

Major Field: Microbiology

Title of Dissertation: Cloning and Expression of Capsid Proteins of Hepatitis A Virus in Escherichia coli

Approved:

[Signatures]

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

April 3, 1997