CONTRACT NO: DAMD17-90-C-0069

TITLE: STRUCTURE AND EXPRESSION OF GENES FOR FLAVIVIRUS IMMUNOGENS

PRINCIPAL INVESTIGATOR: Maurille J. Fournier
Thomas L. Mason

CONTRACTING ORGANIZATION: University of Massachusetts
Department of Biochemistry and Molecular Biology
Amherst, Massachusetts 01003

REPORT DATE: January 20, 1992

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION University of Massachusetts Department of Biochemistry & Molecular Biology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Amherst, MA 01003	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Medical Research and Development Command 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Contract No. DAMD17-90-C-0069
8c. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, MD 21702-5012	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61102A PROJECT NO. 3M1 61102BS13 TASK NO. AA WORK UNIT ACCESSION NO. 052

11. TITLE (Include Security Classification)

Structure and Expression of Genes for Flavivirus Immunogens

12. PERSONAL AUTHOR(S)

Maurille J. Fournier and Thomas L. Mason

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Report	FROM 9/20/90 to 9/30/91	January 20, 1992	9

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Flaviviruses; Recombinant DNA; PCR amplification; subunit vaccine; biotechnology, RA I, Dengue Virus
06	01		
06	13		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

As the first step toward expression of full-length copies of the DEN-1 E and PrM proteins in E. coli, cDNA copies of these proteins were amplified by PCR technology and cloned into T7 RNA polymerase expression vectors. Analysis revealed that while the PrM protein was successfully placed into the expression plasmid, and expressed under inducing conditions, an intact copy of the E protein could not be successfully cloned under any of a wide variety of protocols. Rather, plasmids containing all of the 5' half of the E coding sequence, and an undetermined portion of the 3' end of the sequence were repeatedly obtained. Our results indicated that even the truncated version of E could only be obtained in a low copy-number vector. The truncation event does not occur at the level of PCR amplification of the E protein sequence. The explanation for our failure to obtain full-length clones of the E protein is not readily apparent, as the expression vector employed was designed for expression of otherwise toxic genes, and no expression of cloned sequences should be taking place in the host strains used for plasmid constructions.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
☐ UNCLASSIFIED/UNLIMITED ☑ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Mrs. Virginia Miller	22b. TELEPHONE (Include Area Code) 301-663-7325	22c. OFFICE SYMBOL SGRD-RMI-S

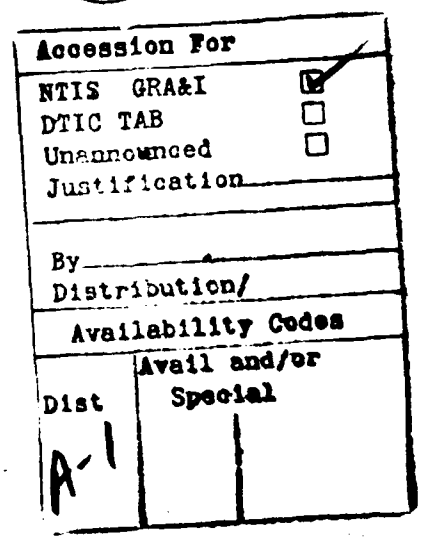
DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

FOREWORD

In conducting research using animals, the investigators adhered to the "Guide for the Care and Use of Laboratory Animals", prepared by the Committee on Care and Use of Laboratory Animals of the Institue of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

For the protection of human subjects, the investigators have adhered to policies of applicable Federal Law 45CFR46.

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.



Final Report - DAMD 17-90-C-0069
15 September 1990 - 14 September 1991

Principal Investigators: Maurille J. Fournier and Thomas L. Mason
Department of Biochemistry and Molecular
and Cellular Biology
Lederle Graduate Research Center
University of Massachusetts
Amherst, MA 01003

A. Other Contributing Personnel.

Name	Position	Dates of Service
Cheryl Thomas, Ph.D.	Postdoctoral Fellow (part-time)	5/90 - 9/91
Laurie (Palmer) Smith, B.S.	Research Assistant	1/90 - 9/91

B. Project Aims

It has been the goal of this research group to develop genetically engineered vaccines for the dengue and encephalitis viruses. We regret that a severe shortage of research personnel during the contract period covered in this report forced us to narrow our goals considerably from those of the previous contract period (June 1986-September 1990; contract No. DAMD 17-86-C-6156). Our recent efforts were concentrated on developing a system whereby full-length, essentially unaltered, copies of the Dengue-1 E and PrM proteins could be expressed in *Escherichia coli*. While the potential advantages of bacterial-based expression of viral immunogens are numerous--e.g., time, expense, procedural simplicity--most research groups working in this field have shifted their attention away from *E. coli* in favor of various eukaryotic systems. The prejudice against expression in bacteria derives in part from the fear that recombinant viral proteins cannot be processed or recovered in a sufficiently "natural" form to elicit a protective immune response. However, two lines of evidence generated by the earlier work of this group and our collaborators led us to believe that it was worthwhile to further investigate the possibility of producing viral immunogens in *E. coli*.

First, as reported in the Final Report for contract DAMD-17-86-C-6156, we found that *E. coli*-expressed Domain II of the DEN-1 E protein elicited antibodies that exhibited a positive response in *in vitro* plaque reduction assays. Second, it was recently demonstrated by Mason *et al.* (1991) that the protein products of recombinant vaccinia-JEV structural protein constructs are capable of generating a protective immune response in mice; the authors note that cultured cells infected with the recombinant viruses produce an extracellular "empty JEV particle" consisting of assembled E and PrM proteins, and suggest that it is those particles that are responsible for the high level of neutralizing antibodies produced by the immunized mice (see also Konishi *et al.*, 1991).

With these findings in mind, we sought to assemble bacterial expression plasmids capable of producing high levels of intact DEN-1 E or PrM proteins,

with the intention of assembling the purified proteins, *in vitro*, into particles analogous to those observed by Mason and colleagues. While this line of research would also afford a valuable opportunity to study the assembly of viral particles, our ultimate goal was to use the assembled particles as immunogens.

C. Program

Expression of Recombinant DEN-1 PrM and E Proteins in Bacteria

The plasmid vector chosen for expression of both dengue antigens under study was pET3a, one of the T7 RNA polymerase expression vectors constructed by Studier and co-workers (1990). This vector is designed for easily inducible, high-level expression of heterologous recombinant proteins; otherwise toxic gene products have reportedly been successfully expressed in this system, an important consideration in our choice of the plasmid. The design of the vector is such that the cloned gene is placed in-frame behind a short phage leader sequence, and thus, expressed as a fusion protein. To address our concern that the presence of the phage amino acids might interfere with proper folding of the E or PrM proteins, or with their interaction with each other in the *in vitro* assembly process discussed above, we also designed alternate cloning protocols in which the dengue genes would be placed in-frame directly behind a start codon.

1. PCR Amplification of the DEN-1 PrM Protein Coding Sequence

Our aim was to express a full-length PrM protein containing as few sequence or structure deviations from the natural protein as practicable. We decided to obtain the DNA fragment to be cloned into pET3a via PCR amplification of a cDNA clone (p1p45) encoding the DEN-1 structural proteins. Precise tailoring of the 5′ and 3′ termini of the PrM sequence was accomplished by designing PCR primers that incorporated the requisite restriction sites for placement of the fragment into pET3a, as either a fusion or non-fusion construction (Figure 1). The fusion protein construction requires in-frame placement of the cloned fragment into a *Bam*HI site; therefore, both the upstream and downstream primers were designed to include *Bam*HI sites. In addition, the upstream primer included an *Nde*I site immediately adjacent to the *Bam*HI site; as the expression vector contains an *Nde*I site at the beginning of the phage leader sequence, incorporation of an *Nde*I site at the 5′ end of the PrM sequence allows for complete removal of the phage leader sequence from the fusion construction, while maintaining the correct reading frame of the cloned insert with respect to the vector start codon. The downstream primer also encodes a nonsense codon at the appropriate location to maintain the exact carboxyl end of the PrM (M) protein.

The amplification and cloning schemes outlined here result in only a few minor alterations of the PrM sequence: in the fusion construction, two amino acids (his and met) are added to the 3′ end of the phage leader sequence; and, the coding sequence of the non-fusion derivative initiates with met rather than with phe. Finally, the upstream PCR primer deliberately incorporates one sequence change at the DNA level: the rare (for *E. coli*) CGA codon for arg at amino acid number six of PrM was changed to CGT.

2. PCR Amplification of the DEN-1 E Protein Coding Sequence

As in the case of the PrM protein, our aim was to express an E protein that was essentially identical to the natural Dengue-1 envelope protein. The protocol for obtaining the E coding sequence to be expressed from pET3a is very similar to that outlined in the previous paragraph for PrM. The target DNA for PCR amplification of the DEN-1 E protein sequence was also carried on plasmid plp45. The PCR primers encoded *Bam*HI sites for cloning E into the expression vector behind the phage leader sequence, and the upstream primer also encoded an *Nde*I restriction site for subsequent in-frame excision of the phage leader sequence (Figure 2). (It should be noted that the latter manipulation is not as straightforward in this case, as there is an additional, internal, *Nde*I site in the E protein sequence at aa298 [Mason *et al.*, 1987].) The exact carboxyl terminus of the E protein is retained by inclusion of a nonsense codon in the downstream primer.

The design of the primers ensures that changes to the natural E sequence are minimal: three rare (*E. coli*) codons are deliberately altered at the DNA level to codons that are more efficiently utilized in an *E. coli* system; and, one additional amino acid (his) has been added to the 3' end of the phage leader sequence in order to construct the *Nde*I site at the amino terminus of the E coding sequence.

3. Placement of the E and PrM Protein Sequences in Expression Vector pET3a

a. The PrM Protein. The PCR amplification reactions worked well for both the E and PrM sequences. In both cases, plentiful amounts of a single reaction product were obtained; restriction enzyme digest analysis of samples of the amplified DNA indicated that the length and structure of the products were apparently correct. Several full-length PrM (fusion) constructions were obtained with relatively little difficulty. For expression of the cloned sequence, the recombinant T7 plasmids were placed in the *E. coli* host strain BL21 (DE3)/plysS, and expression from the T7 promoter induced by addition of IPTG to the growth medium. Preliminary SDS-PAGE analysis of total protein from the induced cells indicates that the cultures were enriched for a protein of the appropriate size to be PrM. Further analysis of the PrM-expressing strains was postponed until E-expressing constructions were obtained.

b. The E Protein. Unfortunately, isolation of T7-recombinant plasmids containing intact, full-length E coding sequences proved to be quite difficult. Ampicillin-resistant colonies invariably contained plasmids that had suffered rearrangements or a variety of deletion events, the latter apparently in both the parent vector sequences and in the inserted E sequence. The explanation for these phenomena is not apparent, as the host strains used--HMS174, HB101, DH5αF', SURE (Stratagene, Inc.)--do not encode T7 polymerase, thus the inserted E sequence would be expected to be "phenotypically silent". Numerous strategies were employed to try to circumvent the difficulties encountered in trying to obtain a full-length E product. The various approaches tried, and the results of each, are summarized below.

i. We suspected that a cryptic bacterial promoter may have been inadvertantly inserted upstream of the E protein coding region as a result of the introduction of restriction enzyme recognition sites into the PCR primer, and fusion of the E sequence with the phage leader sequence of pET3a. To circumvent that potential problem, the PCR amplification product was placed in a pET3a vector from which the phage leader sequence had been excised. In addition, we attempted to place the E sequence into plasmids PUC19 and pBR322, as intermediate cloning steps. Neither of these approaches was successful, as the plasmids obtained from these cloning experiments also contained deletions or appeared to be rearrangements of the parent vectors.

ii. The 3'-terminal 120 nucleotides of the E protein sequence encode two hydrophobic domains. It is likely that the hydrophobic amino acids serve as a membrane anchor and are not directly involved in the association between the M and E proteins; furthermore, hydrophobic domains frequently impede expression of heterologous proteins in *E. coli*. While it is not obvious why these sequences should render our constructions inviable, given the apparent absence of a translational promoter, we decided to try to clone a truncated form of the E protein, produced by partial digestion of the PCR-amplified E product with the restriction enzyme *Ava*II, which cleaves the E sequence immediately upstream of the hydrophobic region. Discouragingly, no viable recombinant plasmids were obtained by this method.

iii. We thought it might be significant that all three of the plasmid systems used--pET3a, PUC19, and pBR322--share a common high copy-number origin of replication. With that in mind, we attempted to place the E sequence into an alternative T7 expression plasmid (pRG), in which the pBR-based *ori* was replaced with the low copy-number (4-5 plasmids per cell) origin of replication of plasmid pSC101. (This plasmid does not fuse a phage leader sequence to the translated product.) In addition, new PCR primers that encoded alternative restriction sites at the 3'-end of the fragment were synthesized; the new primers facilitated directional cloning into either pET3a or pET-RG.

As with all previous approaches, a variety of plasmid constructs were obtained in these experiments. Most of the plasmids appeared to contain inserts with the terminal (5'and 3') restriction sites intact, but the length of the fragments was only ca. 750-800 nucleotides (full-length is approx. 1500 nucleotides). Restriction digest analysis of the short E fragments seemed to indicate that the coding sequence was intact at least as far as the *Xba*I site at position 659 of the E sequence, but restriction sites that fall between that position and the 3' end of the gene are missing. Unfortunately, restriction analysis alone is not adequate to define the borders of the apparent internal deletion with accuracy, and it is not clear at what position the E sequence resumes in these clones.

Two of these specimens were placed in the host strain BL21(DE3/plysS) for expression, but it was not determined before the end of the funding period whether the truncated proteins were expressed from the T7 promoter. The simplest explanation for the presence of an internal deletion is that the PCR amplification process went awry, perhaps because the template assumed a complicated secondary structure which the *Taq* polymerase could not negotiate faithfully. However, every PCR reaction was sized on an agarose

gel, and there is no question that the PCR amplification products were the correct size. Furthermore, detailed restriction enzyme analysis of the PCR products indicated no internal deletions at that level.

D. References.

Konishi, E., Pincus, S., Fonseca, B, Shope, R., Paoletti, E. and Mason, P.W. (1991) Comparison of protective immunity elicited by recombinant vaccinia viruses that synthesize E or NS1 of Japanese encephalitis virus. Virology *185*:401-410.

Mason, P.W., McAda, P.C., Fournier, M.J. and Mason, T.L. (1987) Sequence of the Dengue-1 virus genome in the region encoding the three structural proteins and the major nonstructural protein NS1. Virology *161*:262-267.

Mason, P.W., Pincus, S., Fournier, M.J., Mason, T.L., Shope, R.E. and Paoletti, E. (1991) Japanese encephalitis virus-vaccinia recombinants produce particulate forms of the structural membrane proteins and induce high levels of protection against lethal JEV infection. Virology *180*:294-305.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct the expression of cloned genes. Methods Enzymol. *185*:60-89.

D. Publications stemming from contract work.

Mason, P.W., Pincus, S., Fournier, M.J., Mason, T.L., Shope, R.E. and Paoletti, E. (1991) Japanese encephalitis virus-vaccinia recombinants produce particulate forms of the structural membrane proteins and induce high levels of protection against lethal JEV infection. Virology *180*:294-305.

NdeI
UPSTREAM PRIMER: (5′) attaaggatcccatatgTTC/CAT/CTG/ACC/ACC/CGt/GGG (3′)
BamHI

BamHI
DOWNSTREAM PRIMER: (3′) ACC/ATT/GAG/GTA/GGT/TAC/CGGatcctaggtagga (5′)
STOP

Figure 1. Structure of the primers used for PCR amplification of the DEN-1 PrM protein coding sequence. *Bam*HI sites incorporated into both the upstream and downstream primers allow for in-frame placement of the amplified sequence into expression plasmid pET3a (Studier *et al.*, 1990). The *Nde*I site in the upstream primer is designed to facilitate subsequent removal of the phage leader sequence--derived from the expression vector--for expression of the DEN-1 PrM protein as a non-fusion product. The C-terminus of PrM is defined by incorporation of a TAG nonsense codon in the downstream primer, immediately 5′ to the *Bam*HI cloning site. (Smaller case letters represent bases that are not complementary to the DNA sequence being amplified.)

NdeI

UPSTREAM PRIMER: (5') tattaggatcccatATG/CGt/TGC/GTG/GGt/ATc/GGC/AAC (3')

*Bam*HI

BamHI

DOWNSTREAM PRIMER: (3') GAT/CCT/CAG/TAC/CAA/GTC/CGCatcctaggttcca (5')

STOP

Figure 2. Structure of the primers used for PCR amplification of the DEN-1 E protein coding sequence. *Bam*HI sites incorporated into both the upstream and downstream primers allow for in-frame placement of the amplified sequence into expression plasmid pET3a (Studier *et al.*, 1990). The *Nde*I site in the upstream primer is designed to facilitate subsequent removal of the phage leader sequence--derived from the expression vector--for expression of the DEN-1 PrM protein as a non-fusion product. The C-terminus of PrM is defined by incorporation of a TAG nonsense codon in the downstream primer, immediately 5' to the *Bam*HI cloning site. (Smaller case letters represent bases that are not complementary to the DNA sequence being amplified.)