STRUCTURE AND EXPRESSION OF GENES FOR FLAVIVIRUS IMMUNOGENS

Annual Report

1 September 1985

Maurille J. Fournier and Thomas L. Mason

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD 17-82-2237

Department of Biochemistry
University of Massachusetts
Amherst, Massachusetts 01003

DOD 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.
(U) Structure and Expression of Genes for Flavivirus Immunogens

Maurille J. Fournier and Thomas L. Mason

Annual Report FROM 9/1/84 TO 8/31/85

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

Banks of cloned cDNA and specific antibodies are being used to characterize the structure and expression of the genomes of two flaviviruses, the Japanese encephalitis virus (JEV) and the dengue-type 1 virus (DEN-1). Major aims of the program are to gain insight into the mechanism of viral replication, to characterize the immunopathological response of the mammalian host and to identify peptide immunogens that could be appropriate for use as an antiviral vaccine and for viral diagnosis. Approximately 10.25 kb or 93% of the estimated 11 kb JEV genome has been cloned and is represented on five overlapping clones of 2.4-3.5 kb; the uncloned portion corresponds to about 450 bases at the 5'-terminus and 180 bases at the 3'-end. A bank of DEN-1 cDNA has also been developed and physical and expression analyses are in progress.

Physical and functional maps of the JEV genomes are being developed by DNA sequencing (60% complete) and in vivo expression and immunological screening analyses, the latter in E. coli with the Xgt11 vector (Young and Davis, Proc. Natl. Acad. Sci. USA 80, 1194, 1983).
The results have identified the genomic segments encoding the capsid (C), membrane (M) and envelope (E) structural proteins (S) and the non-structural (NS) proteins NS1, NS3 and NS5. The JEV genome is strikingly similar in size and organization to that of the yellow fever (YFV) virus, the first flavivirus genome to be completely sequenced (Rice et al., Science 229, 726, 1985) and the sequenced C-M portion of the West Nile virus (Castle et al., Virol. 145, 227, 1985). The six major JEV cistrons occur in the same order in YFV i.e., C-M-E-NS1-NS3-NS5 and an open reading frame extends at least through the C-M-E-NS1 coding region, consistent with possible co-expression of a S-NS polyprotein precursor. The encoded proteins also exhibit a high degree of relatedness. A segment of E protein regions determined to be rich in epitopes for both virus-neutralizing and flavivirus cross-reactive monoclonal antibodies shows 45% sequence homology with the corresponding YF protein and an essentially identical hydropathy profile. Together, these results suggest the possibility that other, perhaps even all flavivirus genomes will be organized in similar fashion and that the genetic relatedness will also extend to the proteins encoded.

The Xgt11 expression vector has also been used to develop an initial epitope map of the E protein, the major target of the virus neutralizing and protective antibodies that are elicited on infection. Antibody binding assays with subgenic viral E-lacZ fusion proteins have shown that a 140 amino acid segment in the carboxy-terminal half of the E protein contains epitopes for all nine mAbs tested thus far as well as binding sites for a portion of the anti-E antibodies present in HMAF. Among the E-specific JEV antibodies that bind to this particular E protein derivative are mAbs that have strong virus-neutralizing activity in a plaque reduction assay, one that cross-reacts with the dengue-3 virus and one mAb that also cross-reacts with dengue-1, -2, and -3. These results, obtained with filter-bound segments of the E protein, support the view that it will be possible to map with high resolution individual submolecular domains that define neutralizing and protective epitopes. Immunogens encompassing these elements are obvious candidates for use in vaccine development. The results also demonstrate that virus-specific and conserved epitopes defined at the virion level are also evident at the protein molecular level. This situation offers the promise of developing both specific and broadly cross-reactive antigens and antibodies for use as productive immunogens and viral diagnosis.
STRUCTURE AND EXPRESSION OF GENES FOR FLAVIVIRUS IMMUNOGENS

Annual Report

1 September 1985

Maurille J. Fournier and Thomas L. Mason

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD 17-82-2237

Department of Biochemistry
University of Massachusetts
Amherst, Massachusetts 01003

DOD 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.
SUMMARY

Banks of cloned cDNA and specific antibodies are being used to characterize the structure and expression of the genomes of two flaviviruses, the Japanese encephalitis virus (JEV) and the dengue-type 1 virus (DEN-1). Major aims of the program are to gain insight into the mechanism of viral replication, to characterize the immunopathological response of the mammalian host and to identify peptide immunogens that could be appropriate for use as an antiviral vaccine and for viral diagnosis. Approximately 10.25 kb or 93% of the estimated 11 kb JEV genome has been cloned and is represented on five overlapping clones of 2.4-3.5 kb; the uncloned portion corresponds to about 450 bases at the 5'-terminus and 180 bases at the 3'-end. A bank of DEN-1 cDNA has also been developed and physical and expression analyses are in progress.

Functional mapping of the viral genomes is being carried out through identification of individual protein coding regions by in vivo expression and immunological screening analyses coupled with DNA hybridization and sequencing assays to position the immunopositive coding segments on the genome. In this strategy random cDNA segments are cloned into the E. coli expression vector λgt11 to yield genetically fused viral-E.coli β-galactosidase hybrid proteins (Young and Davis, 1983 a,b). The immunological probes used to identify the viral protein moiety include assorted monoclonal antibodies (mAbs) specific for the major envelope (E) and membrane (M) proteins and antibodies present in hyperimmune mouse ascitic fluid (HMAF) that recognize most and perhaps all of the virus specified proteins; class-specific antibodies are selected from the HMAF by immunoadsorption-desorption from the recombinant viral-lacZ fusion proteins.

Mapping results obtained thus far for the JEV genome have yielded the relative positions of nucleotide segments encoding the capsid (C), (M) and (E) proteins and the non-structural (NS) proteins NS1, ns2a, NS3, and NS5. Higher resolution information about the structure of the genome has been derived from DNA sequence results obtained for: 1) 650 bases corresponding to a portion of the C, pre-M and M coding segments, 2) A 2400 base segment commencing in the midst of the E protein coding region and extending toward the 3’-terminus, 3) 830 bases of the NS3 cistron and 4) 1200 bases comprising a major portion of the NS5 coding region. Altogether, about 50% of the genome has been sequenced thus far.

Interestingly, the mapping and sequence results show the JE viral genome to be strikingly similar to that of the yellow fever (YF) virus, the first flavivirus to be completely sequenced (Rice et al., 1985) and the West Nile (WN) virus for which the sequence of a portion of the structural gene region has been determined (Castle et al., 1985). Comparison of these results shows that 1) the six major JEV genes mapped thus far occur in the same order in YFV i.e., C-M-E-NS1-(ns2a)-NS3--NS5; 2) the 2.4 kb long segment of sequenced JEV cDNA corresponds to a continuous translational open reading frame (ORF), consistent with the structure of the YF genome which encodes one ORF of over 10 kb and, a single ORF of at least 0.9 kb in WNV encoding the C, and pre-M proteins and 3) the predicted structures of the E, NS1 and ns2a proteins of JEV and YFV exhibit a high degree of relatedness. The E proteins have virtually identical hydrophobicity profiles for the
carboxy-half segments with strong (46%) but less conservation at the amino acid sequence level. The NS1 and NS2a protein regions also show good relatedness in hydrophobicity profiles and amino acid sequence conservation (36% and 33%, respectively). Taken together, these results suggest the possibility that other, perhaps even all flavivirus genomes will be organized in a similar fashion and that the envelope and perhaps other proteins will turn out to be closely related among the 70 or so known flaviviruses.

The λgt11 expression vector has also been used to develop an initial epitope map for the E protein, the major target of the virus neutralizing and protective antibodies that are elicited on infection. Antibody binding assays with subgenomic viral E-lacZ fusion proteins have shown that a 140 amino acid segment in the carboxy-terminal half of the E protein contains epitopes for all nine mAbs tested thus far as well as binding sites for a portion of the anti-E antibodies present in HMAF. Among the E-specific JEV antibodies that bind to this particular E protein derivative are mAbs that have strong virus-neutralizing activity in a plaque reduction assay, and other that cross-react with the dengue-3 virus and one mAb that also cross-reacts with dengue-1, -2, and -3. These results, obtained with filter-bound segments of the E protein, support the view that it will be possible to map with high resolution individual submolecular domains that define neutralizing and protective epitopes. Immunogens encompassing these elements are obvious candidates for use in vaccine development. The results also demonstrate that virus-specific and conserved epitopes defined at the virion level are also evident at the protein molecular level. This situation offers the promise of developing both specific and broadly cross-reactive antigens and antibodies for use as protective immunogens and viral diagnosis.

Taken together, the proposed research will lead to a more detailed understanding of the complex molecular biological and pathobiological properties of flaviviruses. This work will also assess the potential of flavivirus subunit vaccines and provide new protein and nucleic acid probes for rapid viral diagnosis.

FOREWORD

The investigators have abided by the National Institutes of Health Guidelines for Research Involving Recombinant DNA Molecules (April 1982) and the Administrative Practices Supplements.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td>6</td>
</tr>
<tr>
<td><strong>II. PROGRESS REPORT</strong></td>
<td>13</td>
</tr>
<tr>
<td>A. Objectives and Period Covered by the Report</td>
<td>13</td>
</tr>
<tr>
<td>B. Research Personnel Comprising the Project Team</td>
<td>14</td>
</tr>
<tr>
<td>C. Detailed Progress Report</td>
<td>15</td>
</tr>
<tr>
<td><strong>CLONING OF THE JEV AND DEN-1 GENOMES</strong></td>
<td>15</td>
</tr>
<tr>
<td>1. Development of the cDNA banks</td>
<td>15</td>
</tr>
<tr>
<td>2. Definition of the genomic library and verification that the cDNA is viral in origin</td>
<td>17</td>
</tr>
<tr>
<td>3. Use of the cDNA as a diagnostic probe</td>
<td>17</td>
</tr>
<tr>
<td><strong>FUNCTIONAL MAPPING OF THE JEV GENOME</strong></td>
<td>17</td>
</tr>
<tr>
<td>1. Mapping strategy</td>
<td>17</td>
</tr>
<tr>
<td>2. Characterization of phage clones encoding viral antigens</td>
<td>18</td>
</tr>
<tr>
<td>3. Nucleotide sequence analysis of the JEV genome</td>
<td>19</td>
</tr>
<tr>
<td>4. Comparison of the JE and yellow fever virus genomes and proteins</td>
<td>20</td>
</tr>
<tr>
<td><strong>SUMMARY AND IMPLICATIONS OF THESE RESULTS</strong></td>
<td>21</td>
</tr>
<tr>
<td>D. Objectives for the Remainder of the Current Contract Period</td>
<td>22</td>
</tr>
<tr>
<td>E. Presentations and Publications</td>
<td>23</td>
</tr>
<tr>
<td><strong>III. LITERATURE CITED</strong></td>
<td>25</td>
</tr>
<tr>
<td><strong>IV. TABLES AND FIGURES</strong></td>
<td>29</td>
</tr>
<tr>
<td>Figure 1. Preparation and Cloning of Viral cDNA</td>
<td>29</td>
</tr>
<tr>
<td>Figure 2. cDNA Clones Comprising the JEV Genomic Bank</td>
<td>30</td>
</tr>
<tr>
<td>Figure 3. Verification that the Cloned cDNA is of Viral Origin</td>
<td>31</td>
</tr>
<tr>
<td>Figure 4. Strategy for Antigenic Mapping of the Viral Genome</td>
<td>32</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (cont'd)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>Partial Functional Map of the JEV Genome</td>
<td>33</td>
</tr>
<tr>
<td>6.</td>
<td>Immunoblot Analysis of JEV-\lambda tll</td>
<td>34</td>
</tr>
<tr>
<td>7.</td>
<td>Deletion of Certain E-Coding Segments Increases the Accumulation of E-lacZ Fusion Proteins in E.coli</td>
<td>35</td>
</tr>
<tr>
<td>8.</td>
<td>Affinity Purification of Antibodies from Hyperimmune Ascitic Fluid</td>
<td>36</td>
</tr>
<tr>
<td>9.</td>
<td>Identification of JEV Coding Regions with Affinity Purified Antibodies</td>
<td>37</td>
</tr>
<tr>
<td>10.</td>
<td>Hydropathy Profiles for the Major Envelope Proteins of the Japanese Encephalitis and Yellow Fever Viruses</td>
<td>39</td>
</tr>
<tr>
<td>11.</td>
<td>Hydropathy Patterns for the NS1 Proteins of JEV and YFV</td>
<td>40</td>
</tr>
<tr>
<td>II.</td>
<td>Properties of JEV-\lambda tll Recombinants</td>
<td>41</td>
</tr>
<tr>
<td>III.</td>
<td>Properties of the JEV-mAbs Used in Immunological Screening of \lambda tll Recombinants</td>
<td>43</td>
</tr>
<tr>
<td>V.</td>
<td>DISTRIBUTION LIST</td>
<td>44</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

The flaviviridae constitute one of the four genera in the recently defined family of togaviruses (i.e. togaviridae). Before reclassification, the flaviviruses corresponded to the Group B arboviruses; some new members are non-arboviruses. Group classification is currently based on serology but other features also distinguish each genus. The flaviviruses account for some two-thirds of the 90 or so known togaviruses, and among arthropod-borne viruses, they are the most important disease-causing agents in man. Twenty-eight (46%) of the known flaviviruses have been associated with human diseases that include yellow fever, dengue fever, and encephalitis (Porterfield, 1980; Shope, 1980). Because they occur in both tropical and temperate zones, the flaviviruses are a significant threat to global public health.

Among the flaviviruses, there is a complex diversity in the virus and host factors that determine virulence and pathogenesis. For example, different viruses exhibit markedly different tissue specificities; many, like Japanese encephalitis virus are neurotropic, others such as the dengue viruses replicate primarily in host macrophages, and yellow fever virus is viscerotropic. It is not surprising that the corresponding diseases have equally diverse and complex pathologies. Clearly, the understanding of flavivirus molecular biology, pathobiology, and epidemiology offers a formidable challenge in biomedical research.

Previous work has shown that flaviviruses, like all togaviruses, are enveloped and have an icosahedral ribonucleoprotein core. The viral genome is an infectious, single-stranded (plus-stranded) RNA approximately 11 kilobases (kb) in length (see reviews by Westaway, 1980, Russell et al., 1980 and Strauss and Strauss, 1983). At least some flavivirus genomes, including that of dengue-serotype 2 (DEN-2), are capped at the 5' terminus, but unlike the alphaviruses, the 3' terminus lacks or only has a small poly A tract. The lipid envelope contains two proteins, a major envelope or E protein (Mr=50-60 kDa), which is usually glycosylated, and a nonglycosylated membrane (M) protein (Mr=7-8 kDa). The nucleocapsid contains the RNA and a capsid (C) protein (Mr=14 kDa). In addition to the structural proteins of the virion, three to five poorly defined nonstructural (NS) proteins are specified by the viral genome. Throughout this proposal we use the flavivirus protein nomenclature proposed by Rice et al. (1985). A list of the polypeptides specified by the Japanese encephalitis virus is
shown in Table 1 below. The table is fashioned after a listing developed by Westaway (1973) and identifies each protein by both the new and pre-established designations (Westaway et al., 1980).

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (KD)</th>
<th>Glycosylated</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C (V2)</td>
<td>13.7</td>
<td>No</td>
<td>Nucleocapsid protein</td>
</tr>
<tr>
<td>prM (NV2)</td>
<td>~20.0</td>
<td>Yes</td>
<td>Precursor to M</td>
</tr>
<tr>
<td>M (V1)</td>
<td>~8.5</td>
<td>No</td>
<td>Virion envelope protein</td>
</tr>
<tr>
<td>E (V3)</td>
<td>53.0</td>
<td>Yes</td>
<td>Major virion envelope protein</td>
</tr>
<tr>
<td><strong>Non-Structural Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS1 (NV3)</td>
<td>46.0</td>
<td>Yes</td>
<td>Soluble complement fixing antigen</td>
</tr>
<tr>
<td>ns2a(NV2 ½) (NV2)</td>
<td>19-21</td>
<td>No</td>
<td>Function unknown</td>
</tr>
<tr>
<td>ns2b(NV1 ½)</td>
<td>~14.0</td>
<td>No</td>
<td>Function unknown</td>
</tr>
<tr>
<td>NS3 (NV4)</td>
<td>69.0</td>
<td>No</td>
<td>Function unknown</td>
</tr>
<tr>
<td>ns4a(NVX) (NV2 ½)</td>
<td>31.0</td>
<td>No</td>
<td>Function unknown</td>
</tr>
<tr>
<td>ns4b(NV1)</td>
<td>10.3</td>
<td>No</td>
<td>Function unknown</td>
</tr>
<tr>
<td>NS5 (NV5)</td>
<td>96.0</td>
<td>No</td>
<td>Function unknown</td>
</tr>
</tbody>
</table>
The sequence of the entire RNA genome of the vaccine strain of the yellow fever virus was recently determined by Rice et al. (1985). It reveals a single long open reading frame of 10,044 nucleotides with the structural proteins encoded within the first 780 amino acid residues at the N-terminus of the predicted polyprotein; the nonstructural proteins are encoded by the remainder of the open reading frame. This genomic organization suggests that the mature viral proteins are produced by proteolytic cleavage of a polyprotein precursor. Partial sequences for the genomes of West Nile virus (Castle et al., 1985) and Japanese encephalitis virus (see Progress Report - page 13) reveal genomic organizations similar to that of the yellow fever virus. Sequence comparisons also show that regions of these three flavivirus genomes are highly conserved (see Progress Report - page 13 - for detailed discussion).

Despite a large body of excellent research, several important gaps remain in our understanding of flavivirus replication and assembly. Studies of replication and structure have been hindered by the low yield of virus from cells in culture, the long latent period—often of the order of 10-12 hours, and the fact that host cell protein and RNA synthesis is not shut off during infection (Westaway, 1980). Fortunately, many of these obstacles can now be readily overcome through the use of specific, sensitive assays with nucleic acid hybridization probes for the viral genome and with immunological probes for the viral proteins. Recombinant DNA technology and new immunological methods have made it possible to generate large quantities of probes suitable for such applications. In addition, DNA sequence, and thus, protein sequence information is extremely valuable in the design of critical experiments. With these technical advances it is now possible to develop effective strategies to answer important questions about the mode of replication and translation of the viral RNA, the pathways and subcellular sites of virus assembly, maturation, and shedding, and the molecular architecture of the virion. The next few years will bring rapid progress in these areas of research.

The tools developed for research at the molecular and cellular levels will also be valuable in studies at the level of the intact organism and in populations of organisms in nature. Nucleic acid and antibody probes will provide the means to establish the molecular bases for variation in virulence of virus strains, and they should also be useful in diagnostic tests, epidemiological studies, and in evaluating vector and host responses to infection. In addition, the production of relatively large amounts of individual viral proteins in recombinant microorganisms will allow the potential efficacy of antiviral subunit vaccines to be evaluated and the epitope profile of the host immune response to be determined at higher levels of precision than was previously possible.

The focus of this proposal is on two flaviviruses, the dengue-1 (DEN-1) and Japanese encephalitis (JE) viruses. Both viruses share the distinction of being major problems in clinical medicine. Dengue hemorrhagic fever (DHF) is an immunopathologic disease associated with infections by four different serotypes of the dengue virus. In the severe form of dengue infection,
dengue shock syndrome (DSS), there is a five to ten percent case fatality rate (Monath, 1985a). DHF/DSS is prominent in Southeast Asia (20,000 cases/year), but a recent epidemic in Cuba (10,000 cases) shows that the disease is a threat in the Western Hemisphere (Monath, 1985). Six confirmed cases of dengue fever were recorded in the United States in 1984/85 (AP news release, the U.S. Centers for Disease Control, Atlanta, CA, Sept., 1985). Japanese encephalitis is an endemic and epidemic problem throughout much of Asia. A recent study in Thailand has shown an annual incidence in the childhood population of 20-50 cases per 100,000 (Shope et al., 1985). The morbidity rate is approximately 20% and many survivors suffer permanent neuropsychiatric damage.

Currently, there are no fully acceptable vaccines for either Japanese encephalitis or DHF/DSS. This point was strongly emphasized in a recent evaluation of available flavivirus vaccines (cf. Russell et al., 1984). Although inactivated JE vaccines have been produced and used in Asia (Japan, China, Taiwan, and Korea) for several years live attenuated JE vaccines were being tested in humans in China, (Shope, 1984). They are not likely to be approved for general use in the United States nor are they fully suitable for widespread immunization in Southeast Asia or other developing regions. For dengue, the development of live virus vaccines for the four serotypes has been difficult. Experimental DEN vaccines have thus far given variable and generally disappointing results in preliminary human tests (Eckels, 1984). At this point it seems essential to systematically explore the potential of new strategies for the development of safe, effective JE and DEN vaccines.

Any vaccine strategy should be based as much as possible on knowledge of the viral pathologies, which are quite different for JE and DHF/DSS. A significant body of information shows that a primary infection by any one of the four dengue serotypes rarely leads to the severe symptoms of DHF/DSS, which are found almost exclusively in patients with a secondary infection by a heterologous serotype (Halsted, 1980, 1982). This phenomenon is most likely the result of IgG-dependent enhancement of viral replication in Fc receptor-bearing monocytes (Halsted, 1980; Collins and Porterfield, 1984). The enhancement mechanism involves the formation of a complex between the virus and non-neutralizing antibodies induced by a previous heterologous infection. The complex then binds to and promotes the infection of the target monocytes, resulting in a progressive amplification of viral replication. Although this brief description is an oversimplification of a complex phenomenon (see Halsted, 1980, 1982 for reviews), it emphasizes that any dengue vaccine must elicit a strongly cross-protective immune response in order to avoid the problem of immune enhancement.

Immune enhancement does not appear to be of great importance in the pathogenesis of JE, although immunopathological processes may be a factor in Japanese encephalitis among the elderly in Japan (Monath, 1985a). Three pathways of pathogenesis have been defined for flaviviral encephalitis (see Monath, 1985a): 1) fatal encephalitis usually preceded by extensive, early viremia and extraneural replication; 2) subclinical encephalitis usually preceded by low viremia and late brain infection; and 3) inapparent infection with trace viremia and no neuroinvasion. There are probably
several pathways for neuroinvasion by JEV, but with the possible exception of access through olfactory neurons (Monath et al., 1983), there appears to be a good correlation with the extent of viremia and subsequent neural penetration. It is not surprising, therefore, that a strong humoral immune response appears to play a major protective role in flaviviral encephalitis (see Monath, 1985 for references). Burke et al. (1985a,b) have established a strong correlation between the survival of JE patients and an early, quantitatively strong serum IgG response to the virus. Local immune response in the central nervous system was also found to be important in clearance of the virus and recovery. The presence of IgM against JEV in the CNS long after recovery suggests the possibility of persistent JEV infections. While immune responses in the CNS are clearly important after viral neuropenetration in JE patients strong humoral immunity should provide sufficient protection against peripheral invasion of JEV from arthropod vectors.

The protective immune response to flavivirus infection appears to be directed toward the E and NS1 proteins with E being the most important target. Epitope maps determined by various functional assays (hemagglutination inhibition, virus neutralization, cross-reactivity, and ELISA) and by competitive binding assays with monoclonal antibodies have identified several different domains on the E protein of DEN and JEV (Gentry, et al., 1982; Kimura-Kuroda and Yasui, 1983; Yasui, 1984; Burke et al., 1984; Henchal et al., 1985) and other flaviviruses (Heinz, et al., 1983, 1984; Schlesinger et al., 1983, 1984a, 1984b; Monath et al., 1984; Gould et al., 1983). Several of these mAbs were strongly positive in in vitro virus neutralization assays. Another particularly important observation is that passive immunization of mice with a complement-dependent cytotoxic antibody to a nonstructural protein (NS1) provides protection against a lethal challenge of yellow fever encephalitis virus (Schlesinger, 1984b). We are not aware of similar studies with other flaviviruses, but the protective potential of antibodies to the NS1 protein of either DEN or JEV deserves careful testing. When considered together, these observations suggest that a strategy for a subunit vaccine based on the E and NS1 proteins has a high probability of success. Evaluation of this potential constitutes one of the major goals of our research program.

There is good reason to be optimistic about prospects for DEN and JE vaccines based on nonreplicating antigens. Promising advances have been made in the application of recombinant DNA technology to vaccine development for several viral diseases. These include hepatitis, influenza, feline leukemia, gastroenteritis, herpes simplex, polio, rabies, and foot-and-mouth disease (see Brown et al., 1984, Norrby, 1984, 1985; Murphy et al., 1984 for references). A hepatitis-B vaccine made by recombinant DNA in yeast is in clinical trials and will apparently be the first genetically engineered vaccine to become available for use in humans. While the self-assembly property of the hepatitis-B surface antigen may offer unique immunogenic advantages, the prospects are good that other envelope proteins can be engineered to have similar properties. Ongoing research to optimize the mode of presentation of synthetic component or subcomponent immunogens will certainly provide new methods to enhance the efficacy of nonreplicating vaccines. The various candidates for vaccine immunogens, with the exception of hepatitis-B, are only in the very early stages of testing, so it would be
premature to predict success for any of them. Nevertheless, the potential benefits of safe, efficient, inexpensive vaccines derived from recombinant DNA fully supports the prevailing consensus that further research in this area should have high priority (cf. Norrby, 1985 and Monath, 1985a,b).

Heterologous virus vectors, such as vaccinia virus, have been demonstrated to be effective vehicles for the production of protective antigens encoded by isolated genes. Dramatic successes to date include protection of vaccinated laboratory animals against challenge by hepatitis B virus (Paoletti et al., 1984), herpes simplex virus (Paoletti et al., 1985), a human influenza virus (Panicali et al., 1983; Smith et al., 1983b) and rabies (Wiktor et al., 1984). In theory this approach could be used for a wide variety of different immunogens including the E and NS1 proteins of DEN and JEV. Multiple immunizations with different vaccinia recombinants may even be feasible (Perkus et al., 1985). In practice, however, there must be a real question of whether immunization with live heterologous vaccinia will be approved for human use in the near future. Since in this context vaccinia is really a cloning/expression vehicle, extensive testing will certainly be required to prove that each new heterologous DNA insert does not confer undesirable properties to the vaccinia genome. For example, while vaccinia DNA is cytoplasmic and by itself is not integrative, with the foreign DNA insert the genome could conceivably become recombinogenic and integrative. It is also possible that the expression of new envelope proteins could alter the tropism of vaccinia. Adequate testing to rule out such possibilities will certainly require a costly, long-term commitment. It appears, therefore, that vaccinia offers immediate benefits as a delivery vehicle for animal vaccines and as an experimental tool, but its enormous potential for use in human vaccines is not likely to be realized in the near future. This argues strongly for the importance of efforts to develop effective nonreplicating immunogens.

Synthetic peptides provide an attractive alternative to the biosynthetic production of antigens for use in the identification of neutralization and protective epitopes and in the dissection of the immune response (Berzofsky, 1985). There has also been considerable enthusiasm for the use of peptides as components in vaccines (Norrby, 1985; Ada and Skehel, 1985). In several cases, synthetic peptides have been used to induce neutralizing antibodies (for recent references see Lerner et al., 1985), but so far their ability to provide a strong protective immune response has not been as clearly documented. In the case of foot-and-mouth disease virus (FMDV), immunization of guinea pigs with a 20-residue peptide corresponding to the 141-160 segment of the VP1 protein protected the animal against an experimental challenge with virus (Bittle, et al., 1982). Similar peptide representatives of the VP1 protein of polio virus elicited only low titers of neutralizing antibody (Emini et al., 1983; Chow et al., 1985), but they were effective in priming animals to produce high neutralization titers when subsequently injected with a subimmunizing dose of the intact virion (Emini et al., 1983). Incomplete in vivo protection was provided by immunization with peptides for antigenic determinants of herpes simplex virus glycoprotein D (Dietzschold et al., 1985) and for the hepatitis-B surface antigen (Gerin et al., 1985).

Several lines of evidence suggest that the limited size of these
peptides is the major limitation to their general applicability as potent immunogens. This is due to factors that are intrinsic to the properties of the antigen and to extrinsic factors that reflect the host immune response (see Berzofsky, 1985 for review). The success of a single synthetic peptide as an immunogen will depend on the overall balance of these factors, which will vary for each peptide and for any given person or animal being immunized. Thus, it is difficult to accurately predict which synthetic peptide antigens will be effective immunogens. One can only select candidate peptides on the basis of the immunological and structural properties of the native protein, i.e., make intelligent guesses, and then proceed to extensive testing. In general, an effective peptide vaccine will probably require the use of a mixture of peptides of at least 15 to 20 residues in length and the effective use of carrier antigens and adjuvants. The peptide mixture is needed to circumvent the problem of variation in the ability of an individual to respond to a specific antigen and relatively long peptides will enhance the expression of antigenic sites that are dependent on the stabilization of secondary structures.

One distinct advantage of synthetic immunogens is that the peptides may be effective in targeting an immune response to sites that are not immunogenic in the native protein. This is a feature with important implications for flavivirus vaccine strategies. For example, the protection provided by a whole virus or even a subunit vaccine might be highly specific for a given virus whereas immunization with peptides that express highly conserved regions of viral proteins could result in significant cross-protection.

In summary, several viable strategies are available for the development of flavivirus vaccines, and at this stage it would be foolish to totally dismiss any of them. However, it would be equally foolish not to set priorities. Our primary focus will be to thoroughly test the vaccine potential of genetically engineered derivatives of the E and NS1 proteins of the JEV and DEN-1 produced by recombinant DNA strategies with prokaryotic and eukaryotic expression systems. This priority is based on the belief that nonreplicating component or subcomponent immunogens provide the most feasible route for applying the power of recombinant DNA technology to the development of vaccines for use in humans. At the same time, we will not ignore the potential of heterologous virus vectors and synthetic peptides as alternatives. Finally, it should be emphasized that regardless of the success of these "state of the art" strategies, the information base generated from the proposed work will be extremely valuable in the development and evaluation of live attenuated vaccines.
II. PROGRESS REPORT

A. Objectives and Period Covered by the Report.

The major, early goals of the program are listed below. The following report details the progress made during the first 18 months of the present two-year contract period (1 March, 1984 to 28 Feb, 1986). The stated objectives were:

1. To develop banks of cloned cDNA fragments corresponding to the genomes of the Japanese encephalitis and dengue-I viruses.

2. To identify DNA sequences in the gene libraries that correspond to specific viral proteins.

3. To identify specific antigenic domains in the major envelope protein, E, that elicit virus neutralizing antibody and establish the essential molecular features of these elements.

4. To determine the amino acid sequences of the E proteins through DNA sequencing.

5. To determine the extent to which the DNA encoding the E protein of JEV is homologous to the corresponding genes from other flaviviruses and to exploit useful homologies to clone other E cDNAs; emphasis will be on the JE-related viruses (JE-complex) especially the dengue serotypes.

6. To prepare E protein and specific E protein fragments for in vivo immunological testing.

An assortment of secondary goals and longer range objectives were also listed in the previous proposal. The more important of these constitute specific objectives for the next phase of the program.

All of the major aims listed above have been addressed during the contract period, however work up to the present has focused primarily on JEV. Our strategy is to conduct the technically intensive developmental work with the JEV materials and then extend the program to the DEN-1 virus. Genetic, biochemical and immunological relationships between the JE and DEN-1 viruses are to be exploited whenever possible.
C. Detailed Progress Report

CLONING OF THE JEV AND DEN-1 GENOMES

Cloning of the JEV genome was initiated in the previous contract period and has been brought to a state of near completion during the present contract. Between 90 and 95% of the estimated 11 kb genome has been cloned and it is expected that the remaining portions will be isolated before the end of the current contract year. A bank of dengue-1 cDNA clones has been developed and characterization of this collection is in progress.

1. Development of the cDNA banks. Our approach to developing cDNA banks from the RNA genomes of the flaviviruses is based on now-conventional cDNA synthesis and cloning strategies. We have however, invested considerable effort in systematically optimizing each of the various procedures involved. This has been done with a view to obtaining simultaneously large-as-possible cDNA clones with full-as-possible representation of the genome, all with maximal efficiency. As our cloning expertise with the viral genomes improved, we were able to replace an earlier collection of smaller, statistically less manageable JEV cDNA clones (inserts 0.1 to 1.5 kb) with a set of five larger clones which overlap extensively. The cloned cDNAs that comprise our current working bank range in size from 2.4 to 3.5 kb. Collectively, these fragments account for 10.25 kilobases of unique information. Two strategies were used in developing these clones and both are shown in schematic form in Figure 1 (FIGURES AND TABLES). The structures of the five cloned cDNAs comprising our JEV genomic bank are shown in Figure 2. Full details of the development and early characterization of these clones will be provided in two manuscripts to be supplied under separate cover. Only the major features of this work are described here.

Approximately 9 kb of the unique JEV cDNA was derived by reverse transcription from a synthetic DNA primer complementary to the last 12 nucleotides at the 3'-end of the viral RNA. (The RNA sequence was provided by Dr. C. Schmaljohn of the U. S. Army Medical Research Institute of Infectious Diseases). An overlapping 5’-proximal clone containing an additional 1.25 kb was produced via reverse transcription from a second internal primer. The new primer corresponds to a 17-base sequence that occurs at about the 2.5 kb point in the genome. The estimated 0.5 and 0.2 kb segments missing from the 5’- and 3’-termini will be cloned via primer extension synthesis. The sizes of the missing segments were estimated from primer extension cDNA synthesis assays and correspondence with the YF genome. Cloning of these last segments is in progress and should be completed within a few weeks.
The strategy used in developing the isolates shown in Figure 2, involved cloning of double stranded cDNA, (Efstratidis et al., 1976; Efstratiadis and Villa-Komaroff, 1979; P. McAld et al., in preparation); some earlier clones were prepared by cloning of cDNA:RNA hybrids. Except for the 5'-terminal clone (PM-7 in Figure 2) double stranded cDNA was generated through self-primed extension synthesis. To minimize loss or damage to the 5'-most portion of the 5'-terminal clone (through S1 nuclease treatment) second strand cDNA synthesis of this segment was initiated from a poly dG primer annealed to a poly dC tail added to the RNA:DNA hybrid (Reichoudhury, et al., 1980). All cDNA preparations were introduced into the plasmid vector pBR322 (Covarrubias et al., 1981) through the use of complementary homopolymer tailing; cloning was at the PstI site of the β-lactamase gene.

The ends of the individual clones quite likely occur in the midst of protein coding regions or other important functional domains. To minimize the disadvantage this might pose in our functional analyses, new subgenomic clones are being developed that are contiguous through the regions in question. This is done through fusion of the existing clones at common restriction sites. Thus far, the overlapping PM-2 and PM-4 pair (see Figure 2) have been joined in this way. Construction of fusions of the unique sequences in the other clones is in progress. For reasons relating to biosafety and analytical convenience no effort will be made to assemble a genomic-length cDNA clone.

The JEV genome has received first attention with each new strategical and technical phase of our program. The successful cloning methodology refined with the JEV genome has now also been applied to the development of a corresponding bank of DEN-1 cDNA. Interestingly, attempts to use the two synthetic JEV DNA oligonucleotides to prime DEN-1 cDNA synthesis were successful, yielding apparent full-length transcripts in the case of the 3'-terminal oligomer and the expected shorter product for the internal primer (reverse transcription of both viral RNAs under study shows very strong dependence on added primer). With these primers and the aforementioned procedures for obtaining double-stranded cDNA, a collection of some 400 putative DEN-1 clones has been established. Screening of the resulting plasmids has revealed a broad range of insert sizes, as expected. Twenty-four clones with inserts in excess of 1 kb have been selected for further characterization. Two clones possess inserts of 5 and 8 kb, corresponding in size to about 40% and 75% of the 11-12 kb DEN-1 genome. Cross-hybridization and restriction analyses are in progress to determine the relatedness of the DEN-1 clones to each other and to the JEV genome.

In addition to promoting cDNA synthesis from the DEN-1 RNA, it has also been shown that the 3'-proximal primer also initiates synthesis on a DEN-4 RNA template (C.-J. Lai, personal communications). The presence in all four
dengue genomes of common sequence features such as the conserved 3'-terminal element would offer an important advantage in cloning these and other related cDNAs.

2. Definition of the genomic library and verification that the cDNA is viral in origin. The individual clones used to define the current genomic bank of JEV DNA were identified by cross-hybridization (Grunstein and Hogness, 1975; Kafatos et al., 1979) and restriction enzyme screening assays (Maniatis et al., 1975; Smith and Birnsteil, 1976). cDNAs related through sequence overlap were first identified by hybridization probing of the initial bank. This analysis was carried out with several of the larger cDNAs derived from the same collection. A partial physical map of the resulting unique DNA was then developed by restriction analysis.

The results from the hybridization screening suggested that perhaps up to 30% of the insert-containing clones in the initial bank were unrelated to the JEV genome. The presence of non-viral DNA sequences probably results from contaminating host RNA or DNA in the initial viral RNA preparation. The viral cDNAs were identified by a series of northern hybridization assays, first with preparations of full-length viral RNA and then blots of electrophoretically fractionated RNA from infected Vero cells (Alwine et al., 1979; Thomas, 1980). Total RNA isolated from virions (Repik et al., 1983) and RNA from uninfected cells served as controls in these assays. The results obtained with one of the putative viral cDNA clones are shown in Figure 3. Final proof that the genomic bank corresponds to JEV sequences came from eventual DNA sequencing and immunological analysis of proteins encoded.

3. Use of the cDNA as a diagnostic probe. The cDNA probes that define our current genomic bank are now being used successfully to detect JEV nucleic acid in a variety of biological materials including clinical samples. Dr. Erik Henchal, working in the Armed Forces Research Institute of Medical Sciences in Bangkok, has thus far been able to identify in hybridization assays the presence of viral nucleic acid in infected cells in culture, brain tissue, mosquitoes and human cerebrospinal fluids (see attached letter in Appendix).

FUNCTIONAL MAPPING OF THE JEV GENOME

1. Mapping strategy. Segments of the viral genome corresponding to the individual protein coding regions are being indentified by in vivo expression and immunological screening methods. The strategy, shown in Figure 4, involves subcloning of randomly produced cDNA fragments into the
phage vector λgt11 (Young and Davis, 1983a,b) to yield a fusion between JEV cDNA and the E. coli gene encoding β-galactosidase (lacZ). When incorporated in the correct orientation and translational reading frame the JEV cDNA segments can be coexpressed with the lacZ gene to yield a lacZ-JEV fusion protein. Bacterial cells infected with hybrid phage are screened for JEV antigen using polyclonal hyperimmune mouse ascitic fluids (HMAF) or monoclonal antibodies (mAbs) derived from virus-infected mice. Prior characterization of the HMAF in western blot analyses with protein from infected Vero or mosquito cells showed it to contain strong immunological reactivity to the E and M (and pre-M) structural proteins, and the non-structural proteins NS1 and NS3; weak reactions to proteins corresponding in size to C and NS5 were also detected. A collection of some 75 anti-JEV monoclonal antibodies supplied to us by Dr. Joel Dalrymple (Department of Viral Biology, USAMRID) round out the repertoire of immunological reagents available for analysis. The collection of mAbs has been only partially characterized thus far, but is known to contain a large number of independently derived antibodies to the E-protein including several that are virus neutralizing and many that display varying degrees of cross-reactivity to other flaviviruses including the dengue serotypes (J. Dalrymple, M. K. Gentry and D. Burke, unpublished results and Gentry et al., 1982).

2. Characterization of phage clones encoding viral antigens. The feasibility of our mapping strategy has been nicely demonstrated with the detection and partial characterization thus far of an extensive number of immunologically positive λgt11-JEV hybrid clones encoding portions of four and probably five virus-specified proteins and to precursor forms of three of these. Summaries of the results obtained for a subset of these recombinants are shown in Figure 5 and Table II. While analysis of the recombinants is still in progress, the results available clearly document the power of the λgt11 expression system for the systematic identification and mapping of viral epitopes (see also Mocarski et al., 1985).

The availability of mAbs to the E and M proteins has permitted the rapid and direct identification of clones that express immunoreactive epitopes of these proteins. λgt11 libraries prepared from either the 5’-proximal clone PM-7, or generated directly from cDNA fragments from the reverse transcription reaction were probed directly with these mAbs. An example of the reactivity observed for three clones derived from the PM-7 bank is shown in Figure 6. Results from similar western blots for clone J7-1 have shown that it reacts with all nine anti-E mAbs that have been extensively tested to date.

Efforts to identify the molecular positions of the epitopes that react with the nine E protein mAbs (shown in Table II) were undertaken by dissecting the cDNA insert in J7-1 with restriction enzymes and subcloning the resulting segments into λgt11. The parent clone has 720 bp of viral DNA.
while the two recombinants contain 580 and 420 bp, the latter corresponding to a JEV ORF of 140 amino acids. Interestingly, the fusion proteins from both recombinants retained the binding of all nine mAbs (Table II) and displayed plaque immunoreactivity that was much stronger than that observed for the larger parent clone J7-1. As is evident in Figure 7, the fusion proteins derived from the shorter clones are very substantially more abundant, thus accounting for the observed enhancement of immunoreactivity. It is particularly noteworthy that the enhanced stability of the fusion proteins specified by the shorter E sub-clones is correlated with the deletion of a pair of C-terminal hydrophobic elements that may correspond to the membrane anchoring domains (see also Figure 10 below).

Identification of protein antigens that react with antibodies in the heterogeneous HMAF but not to the mAb probes has been accomplished by an indirect immunological analysis. The procedure takes advantage of the high level of expression of the fusion proteins in the λgt11 system and the ability of these proteins to specifically bind a subset of antibodies in the immune ascitic fluids. This procedure, which is outlined in Figure 8, has permitted us to identify the protein coding regions for a large number of recombinants. The reactivities of the affinity purified antibodies on Western blots of virion proteins (Figure 9A) and of total protein from JEV-infected mosquito cells (Figure 9B) were used to establish the identity of HMAF-reactive recombinants and to confirm the identity of recombinants isolated with mAb probes (Table II).

With the affinity purification procedure it has also been possible to identify several clones that express epitopes contiguous on the genome. Clones J7-7 and J7-8 (Figure 5, Table II), which were originally detected with a mAb to the M protein, also contain epitopes to the E protein that are recognized by antibodies in HMAF but not by the nine mAbs used thus far to identify E protein epitopes. These results show that these clones express a polypeptide bridge between the M and E proteins and that there are epitopes on the E protein in addition to those recognized by the nine mAbs. Clone J101 appears to express a bridge between the E and NS1 proteins since its fusion protein reacts, although weakly, with antibodies to the E and NS1 proteins. The weak immunoactivity of this clone is consistent with the relatively low level of the antigen in the E.coli lysogen (Table II). The coexpression of segments from two proteins provides a direct demonstration of open reading frames between the coding segments for E, M and NS1. These functional mapping studies together with DNA sequence information results (see below) show that the JEV genome encoding the structural and major non-structural proteins is organized exactly like that of the yellow fever virus genome.

3. Nucleotide sequence analysis of the JEV genome. DNA sequencing has been largely concentrated on the cDNA encoding the antigenic domains of M, E and NS1 identified in the expression analyses described above. At this time
we have developed sequence information for about 50% of the genome. Included are: 1) 660 bases corresponding to a small portion of the E coding segment, pre-M and M; 2) 2400 contiguous bases extending from the middle of the E coding region through NS1 and ns2a, (1.6-3.9 kb on the genomic map); 3) 830 nucleotides of the NS3 segment and 4) approximately 1.7 kb of the 3’ proximal NS5 coding region.

Perhaps the most striking feature of these results is the finding that the entire 2.4 kb E-NS1-ns2a DNA segment can be read as a single translatable sequence in one open reading frame (ORF). This result, verified by the aforementioned coexpressed proteins in E. coli, is consistent with the possibility that the JEV genome might be organized into a single ORF like that of the sequenced yellow fever virus (Rice et al., 1985). New sequence results for 0.9 kb of DNA comprising the upstream C-M region in the West Nile virus show this segment to also consist of a single ORF (Castle et al., 1985).

4. Comparison of the JE and yellow fever virus genomes and proteins. Comparison of the available JEV sequences to those of YFV show that there is considerable conservation between these two flavivirus genomes. The degree of relatedness is greater yet at the protein level and is particularly striking for the two E proteins. The very evident sequence homology allowed ready identification and alignment of the corresponding genomic segments. Evidence that the assignments are correct is provided by the finding that the functional maps also correspond, although each was derived from different criteria. The sequenced West Nile segment can also be aligned with the ‘consensus’ JEV and YFV map.

Taken together the physical and functional mapping results show excellent correspondence. Because some of the sequence data for the C, M and NS5 regions is still preliminary we will confine our detailed comparisons to the 2.4 kb segments encoding E, NS1 and ns2a; the JEV sequence results start at the mid-point of the E protein. When the two viral segments encoding E are aligned at the left-most conserved amino acid domains, a pair of strongly hydrophobic domains some 180 amino acids away at the carboxyl terminus are within several residues of being in exact register. The NS1 and ns2a sequences show similar size correspondence when aligned in the same way. The three protein segments exhibit overall amino acid sequence homologies of 46%, 36% and 33% when alignment is optimized by occasional insertions or deletions of one or two amino acids to accommodate the slight differences in sequence length (results not shown). Absolutely conserved clusters of two to eight amino acids occur over the full length of these proteins. Similar homologies also occur for the deduced M, NS3 and NS5 proteins. At the nucleotide level, the E protein coding sequences show about 50% homology.

Interestingly, comparison of the hydrophobicity profiles for the E-NS1-
SUMMARY AND IMPLICATIONS OF THESE RESULTS.

Three major gains have come from our research program during the current contract period. They are: 1) establishment of a cDNA bank that accounts for approximately 95% of the JEV genome; 2) demonstration of an effective strategy for functional mapping of the genome and antigenic sites encoded within the viral proteins; and 3) development of DNA sequence information that allows direct comparison of the immunologically important E and NS1 proteins of two different flavivirus to be made at the primary structure level. The results from the functional and sequence analyses reveal that major portions of the genomes of the yellow fever, JE and probably West Nile virus genomes are organized in the same way, suggesting the possibility that the genetic similarity may extend to the entire genome. The similarities of the JEV and YFV genomes includes: 1) the order and relative genomic positions of at least six major protein coding regions (C, M, E, NS1, ns2a, NS3 and NS5); 2) the theoretical possibility that the contiguously encoded structural and non-structural proteins are co-translated in their natural hosts as a polyprotein precursor; in neither genome is there a stop codon in the junction sequence between the E and NS1 coding segments, and 3) several features of the primary and higher structures have been conserved for the dominant immunogen, the E protein and also for the non-structural proteins NS1 and ns2a.

The structural relatedness of the two flavivirus genomes presumably also reflects functional conservation at several levels. The sequence and immunological mapping results described here have special implications from the dual perspectives of vaccine development and viral diagnosis. The discovery that a 140 amino acid segment in the carboxy-half of the E protein contains epitopes for every one of the nine monoclonal antibodies tested thus
far, most with different specificities, indicates that this is an important and potentially valuable antigenic region. Supporting this view are the findings that the collection of mAbs tested includes three known JE virus-neutralizing antibodies, antibodies that cross-react with one to three of the dengue viruses and other flaviviruses, and an antibody that is typespecific for the JE virus.

The primary sequence information and secondary folding patterns that can be predicted from this information constitute critical first steps in our approach to identify and characterize the various epitopic domains. These results provide encouragement that the goals of eventually developing both specific and broadly cross-reactive vaccines and diagnostic reagents can indeed be realized.

D. Objectives for the Remainder of the Current Contract Period

Work during the remainder of the contract period will focus on the following specific aims:

Japanese encephalitis virus:

1. To complete the cloning of the genome. With the possible exception of troublesome terminal segments this aim should be readily attained with the aid of newly developed 5'- and 3'-proximal synthetic primers. It can be noted that the occurrence of unforeseen difficulties does not preclude the determination of nucleotide sequence information for the missing regions; this can be done by direct sequencing of the primary cDNA transcript.

2. To complete the genomic mapping. For all virus-specified proteins for which antibody probes are currently available. There is still uncertainty about the eventual number of proteins that can be detected with the antibody preparations on hand. However, at the least, suitable probes exist for the three structural proteins, C, M and E and the major non-structural proteins.

3. To extend the DNA sequence information to include all of the structural protein coding region and as much as possible of the remaining nonstructural region of the E and NS-1 protein.

4. To characterize the antigenically important domains in the E and
NS1 proteins. Special emphasis will be placed on defining epitopes to which virus neutralizing antibody reacts.

5. To develop first generation recombinant E and NS1 immunogens. Evaluation will involve immunogenicity and possibly animal protection analyses.

Dengue-type 1 virus:

1. To define a physical map of the genome.

2. To identify and sequence the E protein coding region.

3. To initiate comparative studies of the genetic relatedness of the DEN-1 and JE viruses. Preliminary assessments will be made of the degree of relatedness of a) the genomes and b) the E proteins in particular. This will be achieved through the use of nucleic acid hybridization analyses, immunological cross-reactivity assays with viral lacZ fusion proteins, and DNA sequence comparisons of the E coding segments. The immunological testing will be done with JEV and DEN monoclonal antibodies and hyperimmune ascitic fluids.

E. Presentations and Publications

Accounts of our early progress were presented at three international meetings during the present contract period. These papers are identified below as are titles of two manuscripts currently in preparation; the manuscripts will be forwarded when available.


**Anticipated Publications.** During the remaining half year of the present contract we expect to extend our current findings to the point where additional manuscripts can be prepared. Likely topics for these anticipated works include:


2. The Immunogenic Potential of Recombinant Derivatives of the Envelope Protein of the Japanese Encephalitis Virus.

3. Nucleotide Sequences of the Capsid and Membrane Protein Coding Segments of the Japanese Encephalitis Virus Genome.

4. Sequence and Antigenic Domains of the Major Envelope Protein of the Dengue Type-1 Virus.
III. LITERATURE CITED


IV. TABLES AND FIGURES

PREPARATION AND CLONING OF VIRAL cDNA

Figure 1. Two strategies have been used to clone JEV cDNA. One involves direct cloning of cDNA:RNA hybrids, the other cloning of double-stranded cDNA. The initial banks of cDNA were derived by reverse transcription of denatured viral RNA from a synthetic DNA primer complementary to the 3' terminus (sequence provided by G. Schmaljohn). All cDNA was cloned into the beta-lactamase gene of pBR322 via homopolymer tailing.
Figure 2. Bank of genomic cDNA clones for the Japanese encephalitis virus (JEV), Nakayama strain. Viral cDNA was prepared and cloned into the plasmid vector pBR322 as outlined in Figure 1. Recombinant plasmids containing JEV cDNA were identified by hybridization assays with viral RNA and the cloned DNAs related to each other by cross-hybridization dot-blot and restriction enzyme analyses. Clones PM-2, -4, -5 and -6 were developed from cDNA transcripts initiated from a primer complementary to the 3' terminus of the viral RNA; the cDNA contained in PM-7 was derived from a primer positioned near the 5'-end of PM-6. All clones shown were derived from double-stranded cDNA. The five clones comprising the genomic bank contain cDNA inserts of 2.5-3.5 kb and together account for 10.25 kb (93%) of the estimated 11 kb viral genome. Approximately 730 nucleotides remain to be cloned, corresponding to about 450 bases at the 5'-terminus and 180 bases at the 3'-end. The cloned portion of the viral RNA is shown as a solid line in the genomic map (top) while the uncloned terminal segments are designated with broken lines.
Figure 3. Verification that the cloned cDNA is of viral origin. Evidence that the putative viral cDNA clones shown in Figure 2 corresponds to JEV genomic material was developed by northern blot hybridization analysis. Total cellular RNA from JEV-infected and uninfected Vero cells and RNA from isolated virions was fractionated electrophoretically and transferred to nitrocellulose filters. Replicate filters were incubated under hybridizing conditions with $^{32}$P-labeled cDNA from clones PM-2-PM-6, washed and analyzed by autoradiography. Lane A corresponds to crude virion RNA (partially degraded), lane B contained RNA from infected cells and lane C was control RNA from uninfected cells. The arrow at the left edge shows the position of full-length viral RNA. The positions of the two large ribosomal RNA species are also shown. The results shown are for the cDNA clone PM-4, essentially identical results were observed with each of the other cDNAs tested. The nature of the rRNA-like positive material present in infected cell RNA but not in control RNA from uninfected cells is unknown.
Figure 4. Strategy for antigenic mapping of the viral genome. Protein coding regions in the flavivirus genome are identified by subcloning of random fragments of cDNA into the lacZ gene of the λgt11 expression vector (Young and Davis, 1983a,b; Nunberg et al., 1984; Mocarski et al., 1985). Translatable sequences inserted in the correct orientation and reading frame will be coexpressed with lacZ to yield a hybrid lacZ-viral fusion protein. Viral antigens are detected by immunological screening with polyclonal and monoclonal antibodies. The procedure can be used with either previously cloned or uncloned cDNA.

In the case of existing plasmid clones, one or more recombinants is subjected to digestion with DNase I in the presence of Mn⁺⁺ to yield a mix of fragments ranging from full-length (~8 kb) to ~500 base pairs; uncloned cDNA is used directly after second strand synthesis following S1-nuclease treatment. The DNA fragments are treated with EcoRI methylase to protect all RI sites from subsequent digestion, and the ends blunted by fill-in synthesis with phage T-4 DNA polymerase. Synthetic EcoRI linkers are ligated to the fragments and the products separated from excess unligated linkers.

Following EcoRI digestion the fragments are resolved by chromatography on Sephacryl S-1000, yielding different size populations of fragments with RI ends. The RI fragments are then ligated into phosphatase-treated λgt11 DNA, packaged into phage heads, and plated on E.coli cells on LAC indicator plates. The resulting plaques are screened immunologically with anti-JEV hyperimmune ascitic fluid or monoclonal antibodies. Immunopositive recombinants are plaque purified for further characterization.
Antigenic Map of the JEV Genome

Figure 5. Partial functional map of the JEV genome. The map positions of 17 antigen coding cDNAs in λgt11-JEV subclones have been determined by a combination of hybridization, restriction enzyme and DNA sequencing analyses. The identities of the proteins encoded were determined by reaction with E- and M-specific monoclonal antibodies or identification by western blot assays of antibodies present in polyclonal ascitic fluids that bind to the recombinant viral protein. Additional properties of these and other subclones are given in Table II. The identifying mAb probes are listed in Table III and the procedure used to affinity purify and identify reactive antibodies in mouse hyperimmune ascitic fluid is shown in Figure 8. The shaded regions on the RNA genome show the four antigenic coding regions identified to date, corresponding to the M, E, NS1 and NS3 proteins. Clones J7-8 and J-9 express joint M-E and E-NS proteins. Clone J7-15 encodes a 140 amino acid segment of the E protein that binds several different mAbs, including virus neutralizing Ab; this clone was identified by DNA deletion analysis. A 1.4 kb segment encoding a major portion of the NS5 protein has been identified by sequence analysis and positioned at 9-10.2 kb on the map. DNA sequence results have been developed for about 5 kb including: a 660 base region encoding a small portion of C, pre-M and M; a 2.4 kb segment that encodes the carboxy-half of E, all of NS1 and ns2a, 830 bases of NS3 and 1.2 kb of NS5. The E-NS1-ns2a cluster corresponds to a single open reading frame.
Figure 6. Immunoblot analysis of JEV-gt11 recombinants that express lacZ fusion proteins containing segments of the E and M proteins. Recombinant phage J7-1, -2, and -3 were isolated as immunoreactive plaques from an expression library created by random subcloning of pHM-7 DNA into the unique Eco RI site in the lacZ gene of gt11. Phage-infected cells (E. coli strain Y-1089) grown under conditions that maximize the expression of fusion proteins were used to prepare protein extracts. The induced cells were collected by centrifugation, resuspended in SDS sample buffer (Laemmli, 1970) containing 1 mM phenyl-methyl-sulfonyl-fluoride (PMSF), heated at 70 °C for 15 min, and then lysed by sonication. The lysate was clarified by centrifugation and an equal aliquot of each sample containing approximately 75 μg of protein was resolved by SDS-polyacrylamide gel electrophoresis in a 7.5% gel. Part of the gel was stained with Coomassie blue and the proteins in the remainder of the gel were electrophoretically transferred to nitrocellulose paper (Towbin, et al., 1979). Sections of the protein blot were incubated with mAbs to either β-galactosidase or the E or M proteins of JEV, washed, and then incubated with 125I-labeled rabbit anti-mouse IgG second antibody. The resulting autoradiographs are shown. The corresponding antibodies are: second panel, IE7, anti-β-galactosidase; third panel, a pool of anti-E mAbs quick code numbers 42, 46, 54, 57, 63, 70, 71, 72; fourth panel, anti-M, quick code number 37. The protein standards are: β-galactosidase (E. coli), 116 kDa; phosphorylase B (rabbit muscle), 96 kDa; bovine serum albumin, 66 kDa.
Figure 7. Deletion of certain E-coding segments increases the accumulation of E-lacZ fusion protein in E.coli. Deletions at the 3' end of the cDNA insert in J7-1 were created by digesting the phage DNA with the Apa I or Sst I restriction enzymes followed by blunt ending, Eco RI linker ligation, digestion with Eco RI, and recloning into AgtII. This procedure regenerates the same lacZ-JEV junction that was originally present at the 5' end of the cDNA insert in J7-1, but new 3' junctions are formed. The immunopositive recombinants J7-1A and J7-1S were isolated and shown to contain shorter cDNA inserts (see Table II and Figure 5). The fusion proteins expressed by these recombinants were analyzed exactly as described in the legend for Figure 6 except that HMAF instead of the anti-M mAb was used as the immunological probe in the fourth panel. Based on the intensity of the Coomassie blue stain and the immunological reactivity, the relative abundance of the fusion proteins in the infected cells is estimated to be J7-1S > J7-1A > J7-1. The same pattern of immunoreactivity was observed in plaque assays of these recombinants (not shown).
Figure 8. Affinity purification of antibodies from HMAF by binding to JEV-lacZ fusion proteins. In order to identify the protein antigens that react with antibodies in the polyclonal HMAFs, we have taken advantage of the high level of fusion protein expression in the λgt11 system to affinity purify subsets of anti-JEV antibodies, which can then be used to probe western blots of virion proteins or of protein extracts from JEV-infected cells. Briefly, this procedure involves the preparation of filters from plates with a high density of phage plaques (approximately 2,000 plaques per 9 cm Petri plate). The filter is blocked in Tris-buffered saline (TBS), pH 8.1, 3% BSA and then incubated for 2 h at room temperature with gentle agitation in approximately 3 mL of a solution containing a 1:30 dilution of the HMAF in TBS + BSA. Excess antibody is then removed by washing once for 15 minutes in TBS, twice in TBS, 0.1% NP40, and once more in TBS. In the final step, specifically bound antibody is eluted according to the procedure of Smith and Fischer (1985). The filter is incubated for 60 seconds in 2 mL of 50 mM glycine-HCl, 500 mM NaCl, 0.5% Tween 20, 0.01% BSA, pH 2.3, and the eluate is rapidly neutralized with an equal volume of ice-cold 100 mM NaHPO₄. The elution and neutralization steps are repeated, the two eluates are combined, diluted with one-half volume of TBS, and BSA is added to a final concentration of 1%. The eluted antibodies are stored at 4°C before use.
Figure 9. Identification of JEV coding regions with affinity-purified antibodies. Antibodies specific for individual JEV-\textsuperscript{agtl1} recombinants were recovered from HMAF by immunoabsorption (see Figure 8). These antibodies were then used to probe western blots of proteins derived from either purified virions (panel A) or homogenates of JEV-infected mosquito cells (panel B). The immunoreactivity of the affinity-purified antibodies with \textit{bona fide} viral proteins reveals the genetic origin of the open reading frames that are expressed in the individual \textit{agtl1} recombinants. The clone designations for the recombinants are listed beneath each filter strip. The results of these analyses are summarized in Table II. The abbreviation pNS1 is used to designate a protein that appears to contain part or all of NS1 plus an additional ~10 kDa polypeptide extension at the C terminus (data not shown). The other abbreviations are: D\textsubscript{HMAF}, anti-\textit{DEN-1} HMAF, AF\textsuperscript{+}, control ascitic fluid.
Figure 10. Hydropathy profiles for the major envelope proteins of the
Japanese encephalitis and yellow fever viruses. The patterns shown were
developed from the DNA and deduced protein sequences (Rice et al. 1985 and
this study) using a variation of the computer program of Kyte and Doolittle
(1982). In this program (Pustell and Kafatos, 1982; International
Biotechnology, Inc.) a positive hydropathy index denotes a region of
hydrophobic character; negative values denote hydrophilic character. The
search length used was nine amino acids and the patterns shown are for the
carboxy-proximal 180 amino acids. This segment accounts for approximately
40% of the length of the mature protein. The pair of strongly hydrophobic
domains at the carboxy terminus in both proteins resemble known membrane-
anchoring regions (Kyte and Doolittle, 1982).
Figure 11. Hydropathy patterns for the NS1 proteins of JEV and YFV. The hydrophobicity profiles were developed as described in Figure 10 and aligned to allow direct comparison. The YFV sequence is from Rice et al. (1985); the JEV sequence was developed in the current study.
<table>
<thead>
<tr>
<th>λ-clone</th>
<th>rel. abundance</th>
<th>MW (kDa)</th>
<th>cDNA Size (bp)</th>
<th>mAb Binding</th>
<th>Affinity purif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>λgt11</td>
<td>(+++)</td>
<td>(0)</td>
<td>(0)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>J7-4</td>
<td>+/-</td>
<td>25</td>
<td>550</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>J7-3</td>
<td>+</td>
<td>21</td>
<td>430</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>J7-7</td>
<td>+/-</td>
<td>29</td>
<td>750</td>
<td>M</td>
<td>M,E</td>
</tr>
<tr>
<td>J7-8</td>
<td>+</td>
<td>20</td>
<td>650</td>
<td>M</td>
<td>M,E</td>
</tr>
<tr>
<td>J7-5</td>
<td>+/-</td>
<td>20</td>
<td>850</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J7-6</td>
<td>+</td>
<td>39</td>
<td>1,300</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J7-2</td>
<td>+/-</td>
<td>36</td>
<td>1,150</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J7-1</td>
<td>+</td>
<td>25</td>
<td>720</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J7-1A</td>
<td>++</td>
<td>20</td>
<td>580</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J7-1S</td>
<td>+++</td>
<td>13</td>
<td>420</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Jc-5</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>J-101</td>
<td>+</td>
<td>18f</td>
<td>1,750</td>
<td>E</td>
<td>E,NS1</td>
</tr>
<tr>
<td>J-102</td>
<td>+</td>
<td>18</td>
<td>480</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>J-9</td>
<td>+++</td>
<td>5</td>
<td>130</td>
<td>-</td>
<td>NS1</td>
</tr>
<tr>
<td>J-55</td>
<td>nd</td>
<td>nd</td>
<td>200</td>
<td>-</td>
<td>NS1</td>
</tr>
<tr>
<td>J-64</td>
<td>nd</td>
<td>nd</td>
<td>200</td>
<td>-</td>
<td>NS1</td>
</tr>
<tr>
<td>Jc-3</td>
<td>+++</td>
<td>2</td>
<td>300</td>
<td>-</td>
<td>NS1</td>
</tr>
<tr>
<td>Jc-8</td>
<td>+++</td>
<td>2</td>
<td>250</td>
<td>-</td>
<td>NS1</td>
</tr>
<tr>
<td>J-15</td>
<td>+/-</td>
<td>30</td>
<td>840</td>
<td>-</td>
<td>NS3</td>
</tr>
<tr>
<td>J-21</td>
<td>+++</td>
<td>20</td>
<td>600</td>
<td>-</td>
<td>NS3</td>
</tr>
<tr>
<td>J-2</td>
<td>++</td>
<td>16</td>
<td>230</td>
<td>-</td>
<td>NS3</td>
</tr>
<tr>
<td>J-12</td>
<td>+</td>
<td>31</td>
<td>610</td>
<td>-</td>
<td>NS3</td>
</tr>
<tr>
<td>J-14</td>
<td>+</td>
<td>31</td>
<td>970</td>
<td>-</td>
<td>NS3</td>
</tr>
</tbody>
</table>
Table II (cont'd)

a Established by the intensity of Coomassie blue staining of fusion proteins. The amount of β-galactosidase expressed by λgt11 infected cells was assigned a relative value of +++ (see legend of Figure 6 for details).

b Molecular weight estimate for the JE-encoded portion of the β-galactosidase fusion protein.

c The identity of the virus-encoded epitopes established by reactivity with mAbs of known specificity.

d Identity of the viral coding regions as established by the reactivity of affinity purified antibodies (see legend of Figure 9 for description of this assay).

e Not determined.

f The disparity of fusion protein size and cDNA insert size probably reflects the fact that the stained fusion protein band is a degradation product of the fused translation product.
<table>
<thead>
<tr>
<th>Quick Code</th>
<th>Hybridoma</th>
<th>Viral Protein Specificity</th>
<th>Neutralization Activity</th>
<th>Flavivirus Cross-Reactivity&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Smallest Immunoreactive Agtll Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>JEV2-2F1-2</td>
<td>M</td>
<td>-</td>
<td>-</td>
<td>37-3</td>
</tr>
<tr>
<td>42</td>
<td>JEV2-5A11-1</td>
<td>E</td>
<td>++++</td>
<td>TNB</td>
<td>37-1S</td>
</tr>
<tr>
<td>43</td>
<td>JEV2-5F1-6</td>
<td>E</td>
<td>++++</td>
<td>TMB</td>
<td>37-1S</td>
</tr>
<tr>
<td>46</td>
<td>JEV2-7F12-3</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>37-1S</td>
</tr>
<tr>
<td>54</td>
<td>JEV3-10E1-4</td>
<td>E</td>
<td>+</td>
<td>-</td>
<td>37-1S</td>
</tr>
<tr>
<td>57</td>
<td>JEV3-11B9-3</td>
<td>E</td>
<td>-</td>
<td>DEN-1,-2,-3</td>
<td>37-1S</td>
</tr>
<tr>
<td>62</td>
<td>JEV3-11G5-4</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>37-1S</td>
</tr>
<tr>
<td>70</td>
<td>JEV3-12H11-9</td>
<td>E</td>
<td>-</td>
<td>DEN-3</td>
<td>37-1S</td>
</tr>
<tr>
<td>71</td>
<td>JEV3-14E6-1</td>
<td>E</td>
<td>+</td>
<td>-</td>
<td>37-1S</td>
</tr>
<tr>
<td>72</td>
<td>JEV3-14H5-2</td>
<td>E</td>
<td>-</td>
<td>-</td>
<td>37-1S</td>
</tr>
</tbody>
</table>

<sup>a</sup>Determined by hemagglutination inhibition assays (D. Burke and J. Dalrymple, personal communications).

The viruses tested were: Dengue serotypes 1,-2,-3,-4; Tembusu (TMB); Wesselsbron, Langat.
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Recipient</th>
</tr>
</thead>
</table>
| 5 copies | Commander  
US Army Medical Research Institute of Infectious Diseases  
ATTN: SGRD-UIZ-M  
Fort Detrick, Frederick, MD 21701-5011 |
| 1 copy  | Commander  
US Army Medical Research and Development Command  
ATTN: SGRD-RMI-S  
Fort Detrick, Frederick, MD 21701-5012 |
| 12 copies | Defense Technical Information Center (DTIC)  
ATTN: DTIC-DDAC  
Cameron Station  
Alexandria, VA 22304-6145 |
| 1 copy  | Dean  
School of Medicine  
Uniformed Services University of the Health Sciences  
4301 Jones Bridge Road  
Bethesda, MD 20814-4799 |
| 1 copy  | Commandant  
Academy of Health Sciences, US Army  
ATTN: AHS-CDM  
Fort Sam Houston, TX 78234-6100 |
END

DTC 3-88

FILED DATE