GENETIC AND MOLECULAR STUDIES OF THE PHLEBOTOMUS FEVER GROUP OF VIRUSES

ANNUAL REPORT

PROFESSOR DAVID H.L. BISHOP

FEBRUARY 10, 1988

Supported by

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

Grant No. DAMD17-84-G-4035

Institute of Virology
Mansfield Road, Oxford OX1 3SR, UK

Approved for public release; distribution unlimited

Findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents
Evidence has been obtained that Punta Toro (PT) virus, a member of the Phlebotomus fever serogroup of viruses (Phleboviruses, Bunyaviridae) has 3 virion RNA species (large, L, medium, M, and small, S), which are comparable in number to those of other members of the Bunyaviridae family, but different in size by comparison to the size of the S RNA of Bunyavirus genus members of the family. Studies supported by this grant report the coding strategies of the PT viral M and S RNA species and the characterization of their mRNA species. The results were obtained by cDNA cloning and sequencing the viral S and MRNA species, as well as by direct analyses of the virion proteins and viral specified mRNA species. The studies have indicated that the PT nucleocapsid protein, N, has a size of 26.9 kilodaltons. The N protein is coded in a viral-complementary S mRNA species. A second S RNA open reading frame has been identified in the S viral-sense strand. It codes for a 29.1 kilodalton non-structural protein, NSc. The function of this putative protein is not known. There are two subgenomic S mRNA species (viral-complementary and...
Item 19 (Cont'd)

viral-sense) corresponding to the 2 open reading frames. Each mRNA is approximately half the size of the viral S RNA. The transcription termination sites for both S mRNA species have been mapped to an A and U rich intergenic region of the S RNA. Non-viral primer sequences have been identified at the 5' termini of the mRNA species.

The PT M RNA species codes, in its viral-complementary sequence, for a 146.4 kilodalton precursor to the G1 and G2 glycoproteins and a second non-structural protein, a 30 kilodalton NSM. The amino terminal sequences of the G1 and G2 proteins have been mapped to the viral RNA sequence and the data compared to the M RNA sequence information of Rift valley fever (RVF) virus obtained by USAMRIID staff and other contractees (Dr M. Collett).
TABLE OF CONTENTS

I. SUMMARY .................................................. 2

II. REPORT .................................................... 3
   A. Introduction ........................................... 3
   B. Results from this Reporting Period ................. 4
   C. Summary of Progress Report ......................... 6
   D. Publications emanating from this Grant .......... 6
   D. Other Literature Cited .............................. 7
1. SUMMARY

We have undertaken the biochemical and genetic characterization of members of the Phlebotomus fever (PHL) group viruses that constitute the Phlebovirus genus of the negative sense RNA virus family Bunyaviridae (Bishop et al., 1980). This genus of viruses includes members that are known human pathogens and that are of consequence to military and civilian personnel in particular regions of the world (Bishop and Shope, 1979). Rift valley fever (RVF) virus has caused epidemic of infection including several hundred deaths in a single outbreak in Egypt in 1977-1978. The sandfly fever virus isolates, Sicilian and Naples (SFS, SFN), originally were recovered from American troops in 1943-1944 during epidemics of sandfly fever among military personnel. Isolates representing New World phleboviruses have been obtained from military personnel stationed in Panama (Punta Toro, PT, virus). Other New World isolates, some of which are the etiologic agents of human infection, include Aguagata (AGU), Anhanga (ANH), Bujaru (BUJ), Icoaraci (ICO), Chagres (CHG), Buenaventura (BUE), Candiru (CDU) and Itaporanga (ITP) viruses. Other Old World isolates include Karmimabad (KAR) virus.

Our analyses have been directed towards developing a strategy for phlebovirus vaccine development. Initial studies, supported by prior U.S. Army Medical Research and Development contracts, were therefore aimed at delineating the structural components of phleboviruses. With the demonstration of a tripartite RNA genome for phleboviruses, genetic analyses were undertaken to determine the coding strategies of the 3 RNA species and to determine if recombinant viruses could be obtained and used for vaccine purposes. From analyses of intertypic reassortant PT viruses we showed that the viral 7x10^5 dalton small (S) RNA species codes for the viral 26.9 kilodalton nucleocapsid (N) protein. Although not proven, based on the precedent of the bunyavirus coding strategy, it was probable that the middle (M) size phlebovirus RNA (2x10^6 daltons) coded for the 2 viral glycoproteins (50-70x10^3 daltons) leaving the 3x10^6 dalton large (L) RNA species to code for the 200x10^3 dalton, large L protein (putative transcriptase component) found in viruses. The genetic studies, including interference assays, indicated that although intertypic reassortant viruses could be obtained, heterotypic virus interactions were not demonstrable (i.e., heterotypic phleboviruses did not interfere with each other and did not reassort their genomes in dual virus infections). Although not all phleboviruses were tested for genetic interactions, the results did not hold out much hope for using this approach for vaccine development. Consequently, we
have taken in the work supported by this grant, an alternative approach, that of developing subunit vaccines. Central to this new approach is the determination of the coding strategy of the three viral RNA species and the characterization of the epitopes on the viral antigens that elicit and interact with neutralizing antigens. Such goals can be realized by cloning and sequencing studies of the viral RNA in concert with the antigenic epitope analyses being undertaken at the U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) (Drs. J. Smith and J. Dalrymple). The results of cloning and sequencing the S RNA of PT phlebovirus indicated an open reading frame in the viral-complementary S RNA sequence (S mRNA, coding for N). An open reading frame in the viral-sense strand was identified which codes for a non-structural protein, NS₅.

II REPORT

A. Introduction

The objectives of this contract are to determine the infection strategy of members of the PHL group viruses (Phlebovirus genus, Bunyaviridae) and to develop protocols for vaccine development. Since this group of exotic viruses includes agents that cause illnesses in epidemic proportions (e.g., RFV) in different parts of the world, and therefore are of military importance, our objectives relate to the question of deriving vaccines that will be useful in protecting military personnel against virus infections.

To realize these objectives, prior studies involved determining if reassortant PHL viruses could be used for preparing vaccines. This was studied by determining which PHL group viruses are genetically compatible (capable of producing reassortant viruses with other members of the genus, or members of other genera of the family). Studies conducted in our laboratories have demonstrated that certain virus serotypes of the CAL serogroup of the Bunyavirus genus (another genus of the Bunyaviridae) are capable of producing reassortant viruses with other CAL group members (Gentsch & Bishop, 1976; Gentsch et al., 1977; Gentsch et al., 1979; Rozhon et al., 1981). However, not all CAL group bunyavirus crosses yield reassortants, indicating that their gene pool is limited. Also no crosses involving viruses representing different bunyavirus serogroups (e.g., CAL and the Group C, or Bunyamwera, serogroups) have yet yielded reassortant viruses. Analyses of reassortant bunyaviruses have shown that the S RNA codes for the nucleocapsid (N) protein and a non-structural
protein (NS$_S$) (Gentsch and Bishop, 1978; Cash et al., 1979; Fuller and Bishop, 1982). The M RNA codes for the two viral glycoproteins, G1 and G2, as well as a second non-structural protein, designated NS$_M$ (Fuller et al., 1983; Eshita and Bishop, 1984).

Results reported previously from prior contracts have documented that PHL group viruses have a tripartite RNA genome with RNA segments that are designated L, M and S. From dual PT temperature sensitive mutant, or wild-type, virus coinfections recombinant intertypic PT viruses were obtained. However, no heterotypic reassortants were obtained using different PHL group viruses. Although not all PHL virus combinations were tested, this negative result suggested (but does not prove) that there are restrictions at the genetic level between different PHL gorup viruses. DI interference studies also suggest that there are restrictions on genetic interactions between many of the PHL group viruses.

An alternative strategy for vaccine development is to identify and prepare adequate quantities of the viral antigen(s) that elicit protective antibodies. Central to this issue is knowledge of the RNA segment coding assignments and characterization of the viral antigen(s) that interact with such antibodies (see Gentsch and Bishop, 1978, 1979). We have demonstrated that the S RNA segment codes for N protein and a non-structural protein designated NS$_S$ (Ihara et al., 1984; Bishop 1985a,b). The coding arrangement has been termed an ambisense arrangement (Bishop, 1985a,b). The PT coded S mRNA species have been characterized (Ihara et al., 1985a; Bishop, et al., 1985). Proof that the PT M RNA codes for the glycoproteins G1 and G2 as well as a non-structural protein, NS$_M$, has been obtained (Ihara et al., 1985b). Studies by USAMRIID personnel have provided confirmatory data as well as additional evidence that the PT glycoproteins interact with neutralizing sera.

B. Results from this Reporting Period.

The coding arrangement of PT S RNA

The cloning and sequencing data for PT virus (Ihara et al., 1984) established that the S RNA coded for two proteins, N and NS$_S$ in an ambisense arrangement (Bishop, 1985a,b; Bishop, 1986a,b). This coding strategy raised the question of how the proteins were made, viz: from full length viral-complementary (N) and viral-sense (NS$_S$) mRNA species, or from subgenomic mRNAs specific for each gene product. The results of our studies demonstrated that subgenomic (half-
length) mRNAs were involved. The question was answered by identifying the sizes and sequence representation of the S coded mRNAs using Northern analyses with cDNA and appropriate synthetic oligonucleotide probes (Ihara et al., 1985a; Bishop et al., 1985). The results suggest that the U- and A-rich intergenic regions of S RNA species (viral and full length viral-complementary) serve as transcription terminators for both the N and NS5 mRNA species.

Additional studies to characterize the 5' termini of the PT S coded mRNAs established that they have non-viral primer sequences at the ends, presumably originating from cellular mRNAs that are scavenged by the viral transcriptase (Ihara et al., 1985a).

**The coding arrangement of the PT M RNA**

The M RNA of PT virus was cloned and sequenced (Ihara et al., 1985b). The RNA is 4330 nucleotides long (mol. wt. 1.46x10^6, base composition: 26.7% A, 33.6% U, 18.5% G, 21.2% C) and has 3' - and 5'-terminal sequences that, depending on the arrangement, are complementary for some 15 residues. The viral RNA codes in its viral-complementary sequence for a single primary gene product (the viral glycoprotein precursor) that is comprised of 1313 amino acids (146,376 Da) and is abundant in cysteine residues but has few potential asparagine-linked glycosylation sites. The 5'-noncoding region of the PT M viral-complementary RNA is short (16 nucleotides); the 3'-noncoding sequence is much longer (372 nucleotides). The latter is rich in short stretches of adenylate residues, like the 3'-noncoding regions of the PT S mRNA species (Ihara et al., 1984). No other large open reading frame has been identified in either the viral, or viral-complementary, M RNA sequences. Limited amino-terminal sequence analyses of the two viral glycoproteins have indicated that gene order and potential cleavage sites in the glycoprotein precursor. The data suggest the existence of a 30x10^3-Da polypeptide (designated NSM) in the glycoprotein precursor that precedes the G1 protein (i.e., gene product order: NSM-G1-G2). Examination of the sequence of the PT M gene product reveals the presence of multiple hydrophobic sequences including a 19-amino acid, carboxy-proximal, hydrophobic region (G2). This hydrophobic sequence is followed by a 13-amino acid-terminal sequence rich in charged amino acids. The size and constitution of the carboxy-terminal region is consistent with a transmembranal and anchor function for the glycoprotein in the viral envelope. Other regions of the glycoprotein precursor contain sequences of amino acids with a predominantly hydrophobic character (23, 50, and 20 amino acids in length). Their functions are
unknown. the amino terminus of the G1 protein is located near the end of the 23 amino acid-long hydrophobic sequence of the presumptive precursor, the hydrophobic 50 amino acid sequence lies within G1, and the amino terminus of G2 is located in the middle of the 20 amino acid-long hydrophobic sequence. From the recently determined sequence of the M RNA gene product of (Collett et al., 1985) considerable homology with the PT M RNA gene product has been detected, except for the indicated NSM protein. No homology has been identified between snowshoe hare bunyavirus (Bunyavirus, Bunyaviridae) glycoprotein precursor and that of PT virus.

C. Summary of Progress Report

The coding strategies of the S and M RNA species of PT virus have been determined. Sequence analyses of the PT S RNA postulates that the PT S viral-complementary mRNA codes for a 26.9 kilodalton protein. This protein is N. In the viral-sense strand there is an open reading frame that codes for a 29.1 kilodalton non-structural protein, designated NS5.

Our analyses have shown that the size ranges of phlebovirus glycoproteins and their surface arrangement are similar to those of uukuviruses, and unlike those of Bunyavirus genus viruses. The viral M RNA codes for a 146.4 kilodalton precursor to the viral glycoproteins and a second non-structural protein, the 30 kilodalton NSM. The gene order is NSM-G1-G2. The M mRNA species have been characterized.

D. Publications emanating from this grant


**E. Other Literature Cited.**


DISTRIBUTION LIST

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 & Development Command
ATTN: SGRD-RMI-S
Fort Detrick, Frederick, Maryland 21701-5012

2 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 20014-4799

1 copy
Commandant
Academy of Health Sciences, US Army
ATTN: AHS-CDM
Fort Sam Houston, TX 78234-6100
Findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.
<table>
<thead>
<tr>
<th>22a NAME OF RESPONSIBLE INDIVIDUAL</th>
<th>22b TELEPHONE (include Area Code)</th>
<th>22c OFFICE SYMBOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARY FRANCES BOSTIAN</td>
<td>301-563-7325</td>
<td>SCRD-RM1-G</td>
</tr>
</tbody>
</table>

DD FORM 1473, 84MAR

83 APR edition may be used until exhausted.

All other editions are obsolete.