Envelope protein sequences of dengue virus isolates TH-36 and TH-Sman, and identification of a type-specific genetic marker for dengue and tick-borne flaviviruses

Stephen Y. W. Shiu,1 Wen R. Jiang,2 James S. Porterfield3 and Ernest A. Gould2*

1Department of Microbiology, University of Hong Kong, Pathology Building, Queen Mary Hospital Compound, Hong Kong, 2NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR and 3'Green Valleys', Goodleigh, Barnstaple, Devon EX32 7NH, U.K.

Complementary DNAs were synthesized from the envelope protein genes of two isolates of dengue virus (TH-36 and TH-Sman, previously suggested as possible dengue virus type 5 and dengue virus type 6 respectively) and amplified by the polymerase chain reaction using sense and antisense primers designed from conserved dengue virus gene sequences. The amplified cDNA clones were sequenced in both directions by double-stranded dideoxynucleotide sequencing. Alignment with published dengue virus sequences enabled us to assign these viruses accurately to classified serotypes, confirming that TH-36 and TH-Sman are strains of dengue virus type 2 and dengue virus type 1 respectively. Amino acid changes between the proteins encoded by these two isolates and strains of their respective serotypes may account for the significant antigenic differences observed during previous serological typing of these viruses. Moreover, sequence alignment of flavivirus envelope proteins revealed a hypervariable region, within which members of the dengue and tick-borne virus antigenic complexes show unique peptide sequences. This type-specific hypervariable domain may be useful as a genetic marker for typing dengue and tick-borne flaviviruses.

Dengue viruses, mosquito-borne members of the family Flaviridae (Westaway et al., 1985), are responsible for classic dengue fever and for the associated conditions dengue haemorrhagic fever (DHF) and dengue shock syndrome, each of which has a high level of morbidity and mortality in the tropics. There is evidence suggesting that the more severe forms of dengue fever occur when individuals (usually children) with cross-reactive but non-protective antibodies, acquired maternally or induced by prior infection with a different serotype, are naturally exposed to dengue virus infection (Halstead, 1981, 1988). Despite much research, effective dengue virus vaccines are not yet available, so the antigenic properties of dengue viruses have received much attention.

Currently, four distinct dengue virus serotypes (types 1 to 4; DEN1 to DEN4) can be distinguished by complement fixation (CF), haemagglutination-inhibition and neutralization tests (Porterfield, 1980). Two Thai isolates are of particular interest. Strain TH-36 was isolated in 1958 by Hammon and co-workers from a patient with DHF in Bangkok, and TH-Sman was isolated from a similar patient by Dr Sman Vardhanabhuti. Strain TH-36 was initially identified as DEN2 and TH-Sman as DEN1, but both were shown to differ to a significant degree from the respective prototype strains by CF, plaque neutralization, immunodiffusion and immunoelectrophoresis tests (Hammon et al., 1961; Hammon & Sather, 1964a, b; Ibrahim & Hammon, 1968a, b; Ibrahim et al., 1968). Tests using the soluble complement-fixing antigen also showed that strain TH-Sman differed from the prototype DEN1 virus, but TH-36 was indistinguishable from the prototype DEN2 virus (McCloud et al., 1971). On the basis of these differences, in the first edition of the Arbovirus Catalogue TH-36 strain was registered as DEN5 and TH-Sman as DEN6, but these designations were later withdrawn.

Detection of intratypic strain variation by serological methods is not very efficient because it often fails to define precisely the degree of antigenic difference allowing discrimination between a new type and a new strain. This pitfall can now be circumvented by the application of molecular biology techniques to the study of flaviviruses at the genetic level. Many flaviviruses
have now been sequenced, including dengue viruses of each of the four serotypes (Deubel et al., 1986, 1988; Zhao et al., 1986; Mackow et al., 1987; Mason et al., 1987; Gruenberg et al., 1988; Hahn et al., 1988; Blok et al., 1989; Chu et al., 1989; Osatomi & Sumiyoshi, 1990).

In this report, the nucleotide and deduced amino acid sequences of the antigenic envelope (E) proteins of TH-36 and TH-Sman were determined and compared with published dengue virus sequences with the aim of defining intratypic and intertypic variations at the molecular level.

The isolates TH-36 and TH-Sman used in this study were from Dr Nick Karabatsos (Yale Arbovirus Laboratory, New Haven, Conn., U.S.A.) and had been passaged 17 and 16 times, respectively, through sucking mice before infection of C6/36 Aedes albopictus cells at a multiplicity of 0.1. Infected cell culture medium was harvested 6 days post-infection and virions were purified by precipitation with polyethylene glycol and centrifugation through a 30% to 50% sucrose step gradient (Gould & Clegg, 1987). Viral RNA was extracted from purified virus with SDS and phenol–chloroform. Subsequent molecular manipulations followed standard techniques (Sambrook et al., 1989). First strand cDNAs were primed using the conserved antisense primer 5′ TGTGGTACCTATTCCACTGCCACATTTACAGTGGGGAAGCTGGGGG 3′ (nucleotides 3′ to the E hypervariable domain (amino acids 223 to 230) is highlighted.

Molecular manipulations followed standard techniques (Sambrook et al., 1989). First strand cDNAs were primed using the conserved antisense primer 5′ TGTGGTACCTATTCCACTGCCACATTTACAGTGGGGAAGCTGGGGG 3′ (nucleotides 3′ to the E hypervariable domain (amino acids 223 to 230) is highlighted.

Fig. 1. Nucleotide and amino acid sequences of the E protein gene of dengue virus TH-36.
were subcloned into KpnI-digested and dephosphorylated pUC118 plasmid vector. Two positive subclones of each virus were sequenced in both directions by double-stranded dideoxynucleotide sequencing.

The nucleotide and encoded amino acid sequences of the E protein gene of TH-36 are shown in Fig. 1. Alignment of the E protein sequence of TH-36 with those of other dengue viruses revealed 98.6% amino acid similarity between TH-36 and New Guinea C, the prototype strain of DEN2. This confirms that TH-36 is a strain of DEN2 because the similarity data fall within the range of 97% to 99%, which has been reported for DEN2 isolates (Blok et al., 1989). The amino acid sequences of the E proteins of TH-36 and 16 DEN2 isolates are aligned in Fig. 2. As significant antigenic differences have been reported between TH-36 and the prototype New Guinea C strain (Hammon & Sather, 1964a, b; Ibrahim & Hammon, 1968a, b; Ibrahim et al., 1968), the seven amino acid differences between the two strains were subjected to closer examination. Two of these changes were located in the hydrophobic C-terminal transmembrane anchor and therefore are unlikely to be exposed to antibodies. Of the remaining five amino acid changes, three occur in positions (54, 76 and 385) that encode amino acids conserved between all DEN2 strains (Blok et al., 1989). The amino acid substitutions affecting conserved amino acids (amino acid 385) occurs in or near to the structural element proposed to be an important determinant of flavivirus virulence based upon studies on tick-borne encephalitis (TBE) virus antibody-resistant escape mutants (Holzmann et al., 1990).

The nucleotide sequence and encoded amino acids of the E protein gene of TH-Sman are shown in Fig. 3. Alignment of the E protein sequence of TH-Sman with published dengue virus sequences revealed 97.4% amino acid similarity between TH-Sman and other DEN1 isolates, in contrast to between one and 11 amino acid changes among sequences of DEN1 strains published previously. Hence, TH-Sman is a unique strain of DEN1 which exhibits significantly lower amino acid similarity (96.4% to 97.4%) as opposed to 97.8% to 99.8% with the E protein of other DEN1 isolates. Moreover, alignment of the E protein sequence of TH-Sman with...
those of DEN1 strains showed that there are nine amino acid positions which are conserved in other DEN1 isolates, but are different in TH-Sman (indicated by asterisks in Fig. 2). In particular, three of these occur in a cluster (amino acids 432, 436 and 439). The significance of these amino acid substitutions to the unique antigenic properties reported for TH-Sman remains to be determined. Nevertheless, sequence comparisons do confirm that TH-Sman is a distinctive strain of DEN1.

The E protein sequences of 13 flaviviruses, including the four dengue virus serotypes, were aligned in an attempt to define an antigenic region that may serve as a marker for typing flaviviruses. A type-specific hypervariable domain was identified which comprises unique peptide sequences of the four dengue virus serotypes and members [louping ill (LI), TBE and Langat (LGT) viruses] of the tick-borne complex (Fig. 4). These unique peptide sequences are conserved among strains of dengue (highlighted in Fig. 2), TBE and LI viruses (K. Venugopal, personal communication). In addition, this hypervariable domain of dengue viruses is located within an antigenic region (amino acids 225 to 249) in the E protein of DEN2 which has been defined by the use of synthetic peptides (Roehrig et al., 1990), whereas those of TBE, LI and LGT viruses are likely to constitute part of a neutralization-sensitive region of the E protein (Mandl et al., 1989, 1991; Shiu et al., 1991). Based on these observations, we would like to propose that this type-specific hypervariable domain may be useful as a genetic marker for typing dengue and tick-borne flaviviruses.

Dr S. Y. W. Shiu is a Croucher Foundation Scholar.

References


(Received 19 June 1991; Accepted 29 August 1991)