The molecular basis of virulence of the encephalitogenic flaviviruses

Peter C. McMinn

Department of Microbiology, The University of Western Australia, Nedlands, WA 6907, Australia

Introduction

Flaviviruses are small (50 nm), lipid enveloped RNA viruses (Westaway et al., 1986), most of which are arthropod-transmitted, and several of which cause clinically significant human diseases including yellow fever, dengue fever, Japanese encephalitis and tick-borne encephalitis (reviewed in Monath & Heinz, 1996). Members of the genus Flavivirus within the family Flaviviridae have been grouped into eight antigenic complexes on the basis of cross-reactivity in neutralization assays (Calisher et al., 1989); all except two of the viruses which cause encephalitis belong to either the Japanese encephalitis virus (JE) antigenic complex (mosquito-transmitted) or to the tick-borne encephalitis virus (TBE) antigenic complex. The two exceptions are Rocio virus, a mosquito-transmitted virus not currently classified within an antigenic complex but which cross-reacts with viruses of the JE complex and the non-arthropod-borne Modoc virus, which has been associated with a single case of aseptic meningitis in a child (Reeves, 1990).

Many flaviviruses, including dengue viruses (DEN), yellow fever virus (YF) and members of the JE and TBE serocomplexes cause encephalitis in laboratory rodents (mice and hamsters) after peripheral or intracerebral inoculation and provide an excellent model of arbovirus-mediated encephalitis in humans (Monath, 1986). Significant advances have been made recently in understanding the molecular biology and biochemistry of flavivirus replication (see Chambers et al., 1990; Rice, 1996), the antigenic structure and function of flavivirus proteins (see Heinz & Roehrig, 1990; Monath & Heinz, 1996), and in the molecular genetics of virulence and the pathogenesis of encephalitis (see Monath & Heinz, 1996). The purpose of this review is to provide an overview of the current literature on the molecular genetics of flavivirus encephalitis and to demonstrate the emergence of common themes in the molecular pathogenesis of encephalitis caused by the neurotropic arboviruses.

The flavivirus genome

Genome structure

Genomic RNA of viruses within the genus Flavivirus is single stranded and of positive polarity, and is typically ~ 11 kb in length (see Chambers et al., 1990; Rice, 1996). The most notable feature of the flavivirus genome is the presence of a single open reading frame (ORF) of ~ 10 kb (spanning more than 90% of the genome) which codes for 3386–3433 amino acids (see Chambers et al., 1990; Rice, 1996). Untranslated regions flank the ORF at the 5′ and 3′ termini. The 5′ untranslated region (UTR) is capped (type I) and is 95–132 nt in length, commences with the conserved dinucleotide AG and contains conserved elements potentially involved in secondary structure formation (Rice, 1996). The nucleotide sequence of the 5′ UTR is conserved within members of specific antigenic serocomplexes (Brinton & Dispoto, 1988). The 3′ UTR lacks a poly(A) tract, terminates with the conserved dinucleotide CU and contains several conserved sequences specific for mosquito- or tick-borne viruses (Rice et al., 1986; Rice, 1996). A conserved region in the 3′ UTR of mosquito-borne viruses (CS2) is complementary to a conserved sequence within the C gene and may be involved in circularization of the viral genome. Recent studies indicate that the 3′ UTR is of variable length (114–624 nucleotides in length) in both tick-borne (Wallner et al., 1995) and mosquito-borne (Wang et al., 1996) flaviviruses, due to the presence of deletions and/or reiterated sequences.

A common coding order of flavivirus proteins has been established: the structural genes encoding components of the virion, namely the capsid protein (C), the membrane protein (prM/M) and the major envelope glycoprotein (E) occupy ~ 25% of the ORF at the 5′ end. The rest of the ORF (~ 75%) encodes non-structural (NS) proteins found in infected cells and are named in order: NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. NS1 is a membrane-associated glycoprotein of unknown function but thought to be involved in the early stages of virus replication (see Rice, 1996). NS3 is the second largest viral protein; sequence comparisons and biochemical analyses suggest that NS3 may be trifunctional, with protease, helicase and RNA triphosphatase activities (see Rice, 1996). The N-terminal region of NS3 contains significant homology...
P. C. McMinn

Fig. 1. The flavivirus RNA genome is shown. The single long open reading frame (ORF) spans the genome and commences at an AUG codon and terminates in a UAA codon. The polyprotein encoded by the ORF is represented by an open bar. The ORF is flanked by untranslated regions (UTRs) at the 5′ and 3′ ends; the 5′ terminus is capped and the 3′ UTR terminates in the conserved dinucleotide CU (Rice et al., 1986). Viral proteins produced by proteolytic cleavage of the polyprotein are shown below the open bar and the approximate amino acid lengths are indicated. The genome organization is deduced from the complete nucleotide and deduced amino acid sequences of MVE (Dalgarno et al., 1986; Lee et al., 1990), YF (Rice et al., 1985), KUN (Coia et al., 1988; Speight et al., 1988) and JE (Sumiyoshi et al., 1987).

Infectious cDNA clones

Full-length infectious cDNA clones have been developed for YF (Rice et al., 1989), DEN-2 (Pryor et al., 1996), DEN-4 (Lai et al., 1991), JE (Sumiyoshi et al., 1992), Kunjin virus (KUN; Khromykh & Westaway, 1994) and TBE (Mandl et al., 1997). The cDNA of these clones has been fully sequenced, and the virulence and antigenic properties of derived virus populations determined and compared to wild-type virus (Table 1). Gritsun & Gould (1995) developed a new approach to the production of infectious RNA transcripts of TBE, in which uncloned, full-length cDNA was derived by RT–PCR amplification and subsequently used as an in vitro transcription template. Although viable virus populations resulted from direct inoculation of suckling mouse brain with full-length RNA transcripts of the linear cDNA, a major disadvantage of this approach is that the cDNA template population is heterogeneous (due to the random incorporation of nucleotide changes during RT–PCR) and therefore the genetic background of the progeny virus population is undefined. Con-

Table 1. Flavivirus infectious cDNA clones

<table>
<thead>
<tr>
<th>Virus and strain</th>
<th>Derivation</th>
<th>Vector full-length or partial clone</th>
<th>RNA polymerase</th>
<th>Specific infectivity of derived RNA</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>YF 17D</td>
<td>Cloning of cDNA library</td>
<td>pMT21* partial</td>
<td>SP6</td>
<td>2–3% of parent virus</td>
<td>Rice et al. (1989)</td>
</tr>
<tr>
<td>DEN-4 814669</td>
<td>Cloning of cDNA library</td>
<td>pBR322 full length</td>
<td>SP6</td>
<td>2–3% of parent virus</td>
<td>Lai et al. (1991)</td>
</tr>
<tr>
<td>JE</td>
<td>Cloning of cDNA library</td>
<td>pBR322 partial</td>
<td>T7</td>
<td>&lt; 1% of parent virus</td>
<td>Sumiyoshi et al. (1992)</td>
</tr>
<tr>
<td>KUN JaOAr982</td>
<td>Cloning of cDNA library</td>
<td>pBR322 full length</td>
<td>SP6</td>
<td>&lt; 1% of parent virus</td>
<td>Khromykh &amp; Westaway (1994)</td>
</tr>
<tr>
<td>MRM61C</td>
<td>RT–PCR of 5′ and 3′ termini</td>
<td>pBR322 full length</td>
<td>T7</td>
<td>Similar to parent virus</td>
<td>Mandl et al. (1997)</td>
</tr>
<tr>
<td>TBE</td>
<td>Cloning of cDNA library</td>
<td>pBR322 full length</td>
<td>T7</td>
<td>Similar to parent virus</td>
<td>Mandl et al. (1997)</td>
</tr>
<tr>
<td>Neudoerfl</td>
<td>RT–PCR of 5′ and 3′ termini</td>
<td>pBR322 partial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBE</td>
<td>Cloning of cDNA library</td>
<td>pBR322 full length</td>
<td>T7</td>
<td>Similar to parent virus</td>
<td>Mandl et al. (1997)</td>
</tr>
<tr>
<td>Hypr</td>
<td>RT–PCR of 5′ and 3′ termini</td>
<td>pBR322 partial</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* pMT21 is a 1.9 kb low copy number plasmid derived from pBR322 (Rice et al., 1988).
sequently, this method is unsuitable for rigorous examination of the effect of single mutations within the genome on the phenotypic characteristics of derived virus populations.

There have been several recent reports of the use of infectious cDNA clones to characterize regions of the flavivirus genome responsible for virulence by: (1) construction of heterotypic chimeras of TBE and DEN-4 (Pletnev et al., 1992) and intratypic chimeras of DEN-4 (Kawano et al., 1993), (2) plaque purification of infectious cDNA clone-derived JE populations (Sumiyoshi et al., 1995) and (3) site-directed mutagenesis of full-length TBE/DEN-4 chimeric cDNA (Pletnev et al., 1993) and YF cDNA (Muylajert et al., 1996). Site-directed mutagenesis of infectious cDNA clones appears to be a promising technique for analysis of flavivirus replication and for characterization of virulence determinants.

Structure of the envelope glycoprotein

The flavivirus major envelope glycoprotein (E) is the dominant antigen in eliciting neutralizing antibodies and protective immune responses in the host (Heinz, 1986; Heinz & Roehrig, 1990; Mason et al., 1991; Konishi et al., 1992). Protein E is also the putative cell receptor binding protein and mediator of membrane fusion and cell entry (Marsh & Helenius, 1989) and appears to be an important determinant of neurovirulence and neuroinvasiveness in animal models (reviewed in Monath & Heinz, 1996).

The deduced protein E sequences of DEN-1, -2 and -4 (Mason et al., 1987; Deubel et al., 1986; Zhao et al., 1986), JE (Sumiyoshi et al., 1987), KUN (Coia et al., 1988), MVE (Dalgarno et al., 1986), TBE (western subtype; Mandl et al., 1988), West Nile virus (WN; Wengler et al., 1985) and YF (17D strain; Rice et al., 1985) have been determined, and protein E found to comprise between 493 and 501 amino acids. The most conserved feature of these sequences is 12 cysteine residues at corresponding positions in all flaviviruses which form disulphide bridges at defined locations (Nowak & Wengler, 1987). The antigenic structure of several flavivirus E proteins has been analysed through the use of monoclonal antibodies (MAbs), and topological epitope maps derived from competitive binding studies have indicated that most epitopes cluster into three distinct antigenic domains, A, B and C (Heinz et al., 1983; Roehrig et al., 1983; Kimura-Kuroda & Yasui, 1983, 1986; Cecilia et al., 1988; McMinn et al., 1995b). This information, together with the recognition that all 12 cysteine residues are conserved (Nowak & Wengler, 1987) and that hydrophilicity profiles are very similar (McAda et al., 1987), led to the proposal of a two-dimensional model of protein E (Mandl et al., 1989).

The X-ray crystallographic structure of the external domain of TBE protein E (derived by limited tryptic cleavage from purified virions) has been determined (Heinz et al., 1991; Rey et al., 1995), and its relationship to the antigenic structure of TBE protein E (Heinz et al., 1983) analysed by examination of the deduced amino acid sequences of anti-E protein MAb escape variants (Mandl et al., 1989; Holzmann et al., 1997). A three-dimensional model for the TBE protein E external domain, based on these studies, is presented in Fig. 2. The external domain forms a head-to-tail dimer (170 Å in length) with a convex external curvature (conforming to the 50 nm diameter of the virion) and is anchored to the membrane at its distal ends. Each monomer is composed of three separate domains. The C-terminal 100 amino acids form a β-barrel composed of seven antiparallel β-strands resembling an immunoglobulin constant region (domain III) which corresponds to antigenic domain B (Mandl et al., 1989). This is connected by a flexible region to the central domain (I) which includes the N-terminal 120 amino acids folded into an eight-stranded anti-parallel β-barrel and which corresponds to antigenic domain C. Two long loops which extend from domain I are primarily involved in dimer contacts (domain II) and correspond to antigenic domain A. It has been suggested that the lateral surface of domain III is involved in binding to an as yet unidentified receptor (Rey et al., 1995) and that a highly conserved, hydrophobic peptide between residues 98 to 111 located in each of the domain II loops may be involved in fusion activity (Roehrig et al., 1989). As indicated previously, conservation of the disulphide bridges and antigenic structure of protein E suggests that the three-dimensional structure determined for the TBE protein E external domain (Fig. 2) is common to other flaviviruses.

Molecular determinants of virulence

The pathogenesis of flavivirus encephalitis after peripheral inoculation of laboratory rodents has been extensively reviewed (Albrecht, 1968; Monath, 1986; Monath & Heinz, 1996). After subcutaneous inoculation, virus replication is first detected within draining lymph nodes (Malkova & Frankova, 1959; Huang & Wong, 1963; McMinn et al., 1996a) and this is followed by development of a plasma viraemia (Malkova & Frankova, 1959; Huang & Wong, 1963; McMinn et al., 1996a). Many extraneural tissues are infected during viraemia and release of virus from these tissues allows the continuance of viraemia for several days (see Albrecht, 1968; Monath, 1986; Monath & Heinz, 1996). Virus entry into the brain occurs during the viraemic phase (Malkova & Frankova, 1959; Huang & Wong, 1963; McMinn et al., 1996a). The ability of a virus to replicate in peripheral tissues, induce viraemia and invade the central nervous system (CNS) is referred to as ‘neuroinvasiveness’. Clinically apparent encephalitis develops in mice when virus titres in the CNS surpass ~10^7 p.f.u./g (McMinn et al., 1996a) and most neurons contain significant quantities of virus antigen (Albrecht, 1968) or RNA (McMinn et al., 1996a). The ability of virus to initiate cytopathic infection in the CNS and to cause encephalitis is referred to as ‘neurovirulence’.

The complete nucleotide and deduced amino acid sequences
Fig. 2. Ribbon diagram representing the structure of the tryptically cleaved envelope protein dimer isolated from TBE (Rey et al., 1995); domains I, II and III are indicated. Coordinates for the figure were derived from the Brookhaven Protein Structure Database (accession number 1SVB). (A) Amino acid residues associated with mutations which result in attenuation of neuroinvasiveness in mice are numbered within the open circles. (B) Amino acid residues associated with mutations which result in attenuation of neurovirulence in mice are numbered within the open circles.

of flavivirus strains of high or low neurovirulence in mice are known for YF-Asibi and its attenuated vaccine derivative 17D (Rice et al., 1985; Hahn et al., 1987), and for the wild-type JE SA-14 strain and its attenuated derivative SA-14-14-2 (Nitayaphan et al., 1990; Ni et al., 1994). Comparison of the nucleotide sequences of the prototype TBE strain ‘Neudoerfl’ and the more neuroinvasive strain ‘Hypr’ has also been reported (Wallner et al., 1996). These studies provide little insight into the molecular basis of neurovirulence or neuroinvasiveness due to the large number of nucleotide and amino acid differences involved and their distribution throughout the genome – for example there are 67 nucleotide and 31 amino acid differences between YF-Asibi and YF-17D and these are distributed across the entire genome. Recently, more detailed information on the location of YF neurovirulence determinants has come from comparison of the complete nucleotide and deduced amino acid sequences of several other YF wild-type and vaccine strains (Jennings et al., 1993; Santos et al., 1995) – only 13
Amino acid substitutions are specific to the attenuated strains, of which five are located within conserved regions of protein E.

Attenuation of neurovirulence or neuroinvasiveness during limited cell culture passage has proved to be a more informative approach to the molecular analysis of virus virulence than comparison of wild-type and highly passaged attenuated viruses. Attenuation during limited cell passage has been reported for the mosquito-borne flaviviruses MVE (Lobigs et al., 1990; McMinn et al., 1995a), WN (Dunster et al., 1990), YF (Hardy, 1963; Barrett et al., 1990) and JE (Cao et al., 1995). Attenuation of the passaged viruses was consistently associated with amino acid changes in protein E, determined either by examination of changes in the reactivity patterns of anti-E protein MAb (Barrett et al., 1990; Dunster et al., 1990) or, more specifically, by determination of the nucleotide and deduced amino acid sequences in the structural protein genes (Lobigs et al., 1990; McMinn et al., 1995a; Cao et al., 1995). Although the high frequency of amino acid changes in protein E of these attenuated viruses suggests that this protein contains neurovirulence and/or neuroinvasiveness (collectively termed ‘virulence’) determinants and has a major role in the pathogenesis of encephalitis, comparison of nucleotide and deduced amino acid sequences of parental and attenuated strains was limited to the structural protein genes in these studies and it is possible that attenuation was due to unidentified mutations in other regions of the genome.

A more rigorous approach to the mapping of virulence determinants on protein E has come from sequencing the E genes of attenuated neutralization escape variants selected with anti-E protein MAb of several flaviviruses, including TBE (Holzmann et al., 1990, 1997), JE (Cecilia & Gould, 1991; Hasegawa et al., 1992), louping-ill virus (LI; Jiang et al., 1993) and MVE (McMinn et al., 1995b). Engineering of infectious cDNA clones has also resulted in precise mapping of virulence determinants on protein E (Sumiyoshi et al., 1995) and on the NS1 protein (Muylaert et al., 1996) of derived virus populations. These studies confirm that protein E plays a major role in determination of virulence phenotype and that single amino acid substitutions in protein E are sufficient to cause loss of neurovirulence or neuroinvasiveness. In most cases the attenuated variants were not significantly altered in their ability to replicate in vitro compared to wild-type virus, suggesting that these mutations in protein E altered virus tropism for critical cell-types within the host. Evidence is also emerging that other regions of the genome (such as the NS1 gene) may influence virulence, and that the specific virulence phenotype of a viral variant may be determined by several regions of the genome.

**Neuroinvasiveness**

The means by which encephalitogenic flaviviruses cross the blood–brain barrier and gain entry into the CNS has been debated for many years (reviewed in Johnson, 1982) and is still unresolved. Possible mechanisms include (1) passive diffusion or transcytosis of virus particles across cerebral capillary endothelial cells or (2) replication of virus in endothelial cells followed by budding of virus on the parenchymal side. Alternatively, virus may bypass the blood–brain barrier by (3) invading the olfactory epithelium from the bloodstream followed by infection of olfactory neurones and entry into the CNS. Animal studies indicate that different flaviviruses may enter the CNS by different routes. Mims (1957) isolated YF from both the anterior and posterior parts of the murine CNS at the same time after peripheral inoculation, suggesting haematogenous spread consistent with (1) or (2). By contrast, Monath et al. (1983) established that Saint Louis encephalitis virus infected the olfactory epithelium of peripherally inoculated hamsters and then passed via olfactory nerve fibres to the olfactory lobes of the CNS, consistent with (3). McMinn et al. (1996a) showed that both wild-type MVE and an attenuated neutralization escape variant entered the murine CNS via the olfactory pathway and spread through the brain in a rostral to caudal direction over 3 to 4 days. However, the replication and distribution of the attenuated variant within the CNS were more restricted than for the parental virus. Interestingly, invasion of the murine CNS by the neurotropic alphavirus Venezuelan equine encephalitis virus has also been shown to occur via the olfactory pathway (Charles et al., 1995). Taken together, these data suggest that the olfactory pathway is the most likely mode of entry into the murine CNS by the encephalitogenic arboviruses. It is probable that virus invades the olfactory epithelium during the viraemic phase and enters the olfactory lobes of the CNS by retrograde spread along olfactory nerve axons.

A major function of the flavivirus E protein is in receptor binding to target cells and it is possible that specific mutations in protein E which alter virus binding to target cells within peripheral tissues may result in loss of neuroinvasiveness. A study by Lobigs et al. (1990) found that MVE variants of low neuroinvasiveness selected during serial passage in SW13 cells were consistently altered at residue 390 (G→R) in protein E, a residue located on an exposed loop of the lateral surface of domain III (Fig. 2A). Position 390 in the MVE E protein is part of an RGD sequence, the cell binding motif of several adhesion proteins recognized by members of the integrin receptor superfamily (Ruoslahti & Pierschbacher, 1987). RGD tripptides are also found at the same location on the primary amino acid sequence of the JE and YF E protein and it has been suggested that this motif forms part of the binding site for an as yet unidentified receptor (Rey et al., 1995). An amino acid substitution within the lateral surface of domain III of the TBE protein E at position 384 (Y→H; Holzmann et al., 1990) and in domain III at position 368 (G→R; Holzmann et al., 1997) are both associated with loss of neuroinvasiveness in mice (Fig. 2A). Although these studies show that mutations within the putative receptor binding site of protein E may lead to loss of neuroinvasiveness, firm evidence for a link between altered
receptor binding and loss of neuroinvasiveness has not been reported to date. It is likely that precise mapping of the flavivirus receptor binding site on protein E, and its role in virulence, will be undertaken in the near future by site-directed mutagenesis of infectious cDNA clones.

As flaviviruses enter cells by receptor-mediated endocytosis (Marsh & Helenius, 1989), post-receptor binding events, including endocytosis and low pH-induced fusion activity, are necessary for entry of virus into cells and the successful initiation of virus replication. Hasegawa et al. (1992) selected a JE neutralization escape variant, altered at position 52 in protein E (Q → R or K) and of low neuroinvasiveness in mice, associated with a reduced rate of virus uptake into the cell (i.e. removal from the cell surface) following the initial receptor binding event. Residue 52 is located in the element that links domains I and II of the E protein tertiary structure (Rey binding event. Residue 52 is located in the element that links domains I and II (Rey et al., 1995; Fig. 2A), a region suggested as a potential ‘hinge’ involved in low pH-induced conformational change. Mutations at position 52 in protein E may either directly or indirectly change the conformation of the receptor binding site, with the effect of altering the interaction between virions and the cell receptor such that the rate of endocytosis of bound virions is reduced. Alternatively, mutations at position 52 in protein E may reduce the rate of endocytosis of virus without causing direct alteration of the receptor binding site, such as by decreasing the movement of virus–receptor complexes into clathrin-coated pits.

Several flavivirus variants of low neuroinvasiveness and with altered pH-dependent fusion activity have been described. A JE strain defective in cell-to-cell fusion activity and of low neuroinvasiveness in mice was found to have different anti-E MAb reactivity profiles compared to a neuroinvasive and fusion competent JE strain (Higgs & Gould, 1991), suggesting that differences in the fusion activity and virulence of the two strains were causally linked to the changed epitope structure of protein E. Neutralization escape variants of MVE selected with an anti-E MAb (McMinn et al., 1995b) and changed at position 277 (S → I), were of low neuroinvasiveness in mice and had grossly abnormal haemagglutination activity and cell-to-cell fusion activity compared to parental virus (McMinn et al., 1996b). The S277I mutants also showed delayed growth in single-step growth curves in Vero cells compared to parental virus (McMinn et al., 1995b), although final infectivity titres were similar. By contrast, neutralization escape variants selected with the same MAb and altered at position 277 (S → N) were of high neuroinvasiveness and had identical haemagglutination, fusion and growth properties to parental virus, indicating that the altered phenotype of E-277 mutants is specific to certain residues at this site (such as isoleucine). The S277I mutation also appears to cause a major change in tissue tropism in host mice after peripheral inoculation (McMinn et al., 1996c). Attenuated virus fails to replicate within lymph nodes draining the inoculation site and produces delayed and diminished viraemia and CNS invasion compared to wild-type virus. Cecilia & Gould (1991) selected a JE neutralization escape variant of low neuroinvasiveness in mice and of low haemagglutination activity, altered in protein E at residue 270 (I → S), a site close to the S277I mutation in MVE which resulted in similar phenotypic changes. It is notable that residues 270 to 277 in protein E are in the putative ‘hinge’ element linking domains I and II (Rey et al., 1995; Fig. 2A). Furthermore, residues 270 to 277 are spatially close to the highly conserved putative ‘fusion peptide’ between residues 98 and 111 (Roehrig et al., 1989). It is tempting to speculate that mutations in this region of protein E alter neuroinvasiveness by reducing the efficiency of fusion between the viral envelope and cell membrane within the acidified endosome. However, as other amino acid changes within this ‘hinge’ element (Q52R,K) appear to influence steps in virus entry that occur prior to the initiation of fusion (Hasegawa et al., 1992), the specific function of this region of protein E remains unclear.

In summary, several studies have established that single amino acid changes in critical determinants of protein E are sufficient to cause loss of neuroinvasiveness without significantly altering neurovirulence or the ability of virus to replicate in vitro. Attenuation of these variants is probably due to a change in the ability of protein E to mediate virus entry by receptor-mediated endocytosis into critical cell-types within peripheral tissues of the host. Studies which provide evidence for the association of mutations in protein E with loss of neuroinvasiveness are presented in Fig. 2(A) and in Table 2. There are currently no available data on the presence of neuroinvasiveness determinants within structural proteins other than protein E, or within the non-structural proteins or untranslated regions. However, it seems likely that several regions of the genome may act simultaneously to produce a particular neuroinvasiveness phenotype. Comparison of the complete nucleotide sequences of the highly neuroinvasive JE strain P3 with the non-neuroinvasive strains SA14/USA and S892 has shown that the different virulence phenotypes of these strains are associated with several amino acid differences in the structural proteins C and E and in the non-structural proteins NS2B and NS5 (Ni & Barrett, 1996), suggesting that the particular virulence phenotype of these variants is controlled by determinants in at least four viral proteins.

**Neurovirulence**

Several studies on the molecular genetics of neurovirulence have been reported recently, and have identified determinants of this phenotype in both structural and non-structural regions of the genome. Methods used for the selection of attenuated variants include serial passage of virus in mice (Schlesinger et al., 1996), plaque purification of uncloned virus populations (Sanchez & Ruiz, 1996), neutralization escape selection (Liang et al., 1993) γ-irradiation mutagenesis (Chen et al., 1996), construction of intratypic (Kawano et al., 1993; Sumiyoshi et al., 1995) and heterotypic (Pletnev et al., 1992) chimeras from
infectious cDNA clones and site-directed mutagenesis of infectious clones (Pletnev et al., 1993; Muylaert et al., 1996).

Chimeric TBE and DEN-4 viruses constructed by incorporation of TBE structural protein gene cDNA (prM, E) into the DEN-4 infectious cDNA clone (Pletnev et al., 1992) had a high neurovirulence phenotype identical to that of the parental TBE strain and in marked contrast to the low neurovirulence phenotype of the background DEN-4 virus. Thus neurovirulence determinants of TBE appear to be located within the prM or E gene region. Site-directed mutations within the protein E glycosylation site [at residue 154 (N → L) in domain I; Fig. 2B] and the prM/M cleavage site significantly decreased neurovirulence of the TBE/DEN-4 chimeras in mice and restricted virus replication in cell culture (Pletnev et al., 1993); both mutations led to a major reduction in virus glycoprotein expression (Pletnev et al., 1993), indicating that failure to synthesize mature M and E proteins resulted in defective virus replication in vitro and in low neurovirulence in mice. Intratypic chimeras of DEN-4 were constructed by incorporating the structural protein genes (C, prM, E) of a low neurovirulence variant or a mouse-adapted variant of high neurovirulence into the DEN-4 infectious cDNA clone (Kawano et al., 1993). Virus populations derived from chimeras containing structural protein genes from the high or low neurovirulence variant had an identical neurovirulence phenotype to the parental viruses. These studies show clearly that TBE and DEN neurovirulence determinants are located within the structural protein genes.

As has been found for neuroinvasiveness, the envelope glycoprotein appears to play a major role in the control of neurovirulence, and attenuation of this phenotype has been associated with specific mutations in protein E. Serial passage of JE in HeLa cells resulted in the selection of variants of low neurovirulence (Cao et al., 1995); however, as the attenuated viruses had an average of six amino acid changes in protein E it was not possible to associate attenuation with changes at specific amino acid residues. However, the HeLa passaged viruses bound much less efficiently to mouse brain membrane receptor preparations than did parental virus (Cao et al., 1995), suggesting that mutations in protein E directly or indirectly altered the conformation of the receptor binding site and that this was associated with attenuation. Differences in neurovirulence and in virus replication in mouse brain between the attenuated YF-17D and the neurovirulent mouse brain-passaged (Porterfield) strain are associated with three mutations in protein E, at positions 52 (R → G), 305 (F → V) and 380 (R → T) (Schlesinger et al., 1996). Attenuation of LI neurovirulence is associated with single amino acid changes in protein E neutralization escape variants at positions 308 (D → N) or 310 (S → P) (Jiang et al., 1993). Intratypic chimeras of JE with a single amino acid difference at residue 138 (E → K) in protein E were constructed from full-length infectious cDNA (Sumiyoshi et al., 1992, 1995) or by γ-irradiation mutagenesis (Chen et al., 1996); high neurovirulence is associated with glutamine at this position and low neurovirulence is associated with lysine. Lysine at residue 138 also resulted in a small plaque phenotype (Sumiyoshi et al., 1995; Chen et al., 1996) and in failure of virus to bind to a 57 kDa cell membrane protein from BHK-21 cells (Chen et al., 1996), suggesting that altered

**Table 2. Single amino acid changes in protein E associated with loss of neuroinvasiveness**

<table>
<thead>
<tr>
<th>Virus</th>
<th>Amino acid change in protein E*</th>
<th>Position on protein E tertiary structure†</th>
<th>Effect on protein E function</th>
<th>Selection method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBE</td>
<td>181 (D → Y)</td>
<td>Domain I</td>
<td>Not known</td>
<td>Neutralization-escape selection</td>
<td>Holzmann et al. (1997)</td>
</tr>
<tr>
<td>TBE</td>
<td>123 (A → K)</td>
<td>Domain II</td>
<td>Not known</td>
<td>Neutralization-escape selection</td>
<td>Holzmann et al. (1997)</td>
</tr>
<tr>
<td>JE</td>
<td>270 (I → S)</td>
<td>Base of domain II</td>
<td>Altered haemagglutination</td>
<td>Neutralization-escape selection</td>
<td>Cecilia &amp; Gould (1991)</td>
</tr>
<tr>
<td></td>
<td>52 (Q → R, K)</td>
<td>Base of domain II</td>
<td>Delayed penetration into Vero cells</td>
<td>Neutralization-escape selection</td>
<td>Hasegawa et al. (1992)</td>
</tr>
<tr>
<td>MVE</td>
<td>277 (S → I)</td>
<td>Base of domain II</td>
<td>Altered fusion activity and haemagglutination</td>
<td>Neutralization-escape selection</td>
<td>McMinn et al. (1995b)</td>
</tr>
<tr>
<td>MVE</td>
<td>390 (G → R)</td>
<td>Within RGD sequence on lateral surface of domain III</td>
<td>Not known</td>
<td>Serial passage in SW13 cells</td>
<td>Lobigs et al. (1990)</td>
</tr>
<tr>
<td>TBE</td>
<td>384 (Y → H)</td>
<td>Lateral surface of domain III</td>
<td>Not known</td>
<td>Neutralization-escape selection</td>
<td>Holzmann et al. (1990)</td>
</tr>
<tr>
<td></td>
<td>386 (G → R)</td>
<td>Lateral surface of domain III</td>
<td>Increased pH threshold for conformational change</td>
<td>Neutralization-escape selection</td>
<td>Holzmann et al. (1997)</td>
</tr>
</tbody>
</table>

* Numbering is from the N-terminal amino acid of the E protein of the specific virus.
† According to Rey et al. (1995).
Table 3. Single amino acid changes in protein E associated with loss of neurovirulence

<table>
<thead>
<tr>
<th>Virus</th>
<th>Amino acid change in protein E*</th>
<th>Position on protein E</th>
<th>Effect on protein E function</th>
<th>Selection method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBE/DEN-4 chimera</td>
<td>154 (N → L)</td>
<td>Domain I</td>
<td>Loss of glycosylation site on protein E</td>
<td>Site-directed mutagenesis of infectious cDNA clone</td>
<td>Pletnev et al. (1993)</td>
</tr>
<tr>
<td>JE</td>
<td>138 (E → K)</td>
<td>Base of domain II</td>
<td>Altered binding to 57 kDa BHK cell membrane protein</td>
<td>Intratypic chimera of infectious cDNA</td>
<td>Sumiyoshi et al. (1995)</td>
</tr>
<tr>
<td>LI</td>
<td>308 (D → N)</td>
<td>Lateral surface of domain III</td>
<td>Not known</td>
<td>Neutralization-escape selection</td>
<td>Chen et al. (1996)</td>
</tr>
<tr>
<td>LI</td>
<td>310 (S → P)</td>
<td>Lateral surface of domain III</td>
<td>Not known</td>
<td>Neutralization-escape selection</td>
<td>Jiang et al. (1993)</td>
</tr>
<tr>
<td>DEN-2</td>
<td>390 (H → N)</td>
<td>Lateral surface of domain III</td>
<td>Not known</td>
<td>Plaque purification of mixed virus population</td>
<td>Sanchez &amp; Ruiz (1996)</td>
</tr>
</tbody>
</table>

* Numbering is from the N-terminal amino acid of the E protein of the specific virus.
† According to Rey et al. (1995).
‡ Five mutations in the conserved tripeptide 383–385 (EPG) of the DEN-2 E protein all result in loss of neurovirulence in mice (Hiramatsu et al., 1996).

receptor binding by the E138K form of protein E was causally linked to attenuation. Several individual amino acid substitutions in a conserved tripeptide (EPG) at positions 383 to 385 in the E protein of DEN-2/DEN-4 chimeras (Hiramatsu et al., 1996) derived from the DEN-2 infectious cDNA clone (Lai et al., 1991) led to loss of neurovirulence in mice. Sanchez & Ruiz (1996) compared the plaque size and neurovirulence of 11 plaque purified clones of DEN-2, and showed that the cloned virus populations could be divided into three groups on the basis of heterogeneity in plaque size, neurovirulence and deduced amino acid sequences at position 390 in protein E. A large plaque, high neurovirulence group was associated with histidine, an intermediate plaque, intermediate neurovirulence group was associated with asparagine, and a small plaque, low neurovirulence group was associated with aspartate. Unfortunately, the animals used in these studies were not challenged with virus by peripheral inoculation and it is currently unclear whether loss of neurovirulence due to mutations in protein E is always associated with loss of neuroinvasiveness. Such a link would suggest that loss of neurovirulence represents a higher degree of attenuation than loss of neuroinvasiveness alone. Studies providing evidence for the association of specific mutations in protein E with changes in neurovirulence in mice are presented in Fig. 2(B) and in Table 3.

According to the structural model of protein E (Rey et al., 1995), residues 52 of YF and 138 of JE are predicted to be located at the base of domain II (Rey et al., 1995; Fig. 2B). Mutations at position 138 appear to lead to altered virus binding to membrane proteins, suggesting that amino acid changes at this site may influence receptor binding in addition to membrane fusion. Residues 308 and 310 of LI (Jiang et al., 1993; Fig. 2B), 305 and 380 of YF (Schlesinger et al., 1996), and residues 383 to 385 (Hiramatsu et al., 1996; Fig. 2B) and 390 (Sanchez & Ruiz, 1996; Fig. 2B) of DEN-2 map to the lateral surface of domain III. The change at residue 380 (R → T) of the YF E protein is of particular interest as it occurs within the conserved RGD sequence, similar to non-neuroinvasive strains of MVE (Lobigs et al., 1990), providing further evidence for the role of this motif in the virulence of mosquito-borne flaviviruses.

It is notable that mutations at position 52 on the JE E protein (Hasegawa et al., 1992) or at position 390 in the RGD sequence of the MVE E protein (Lobigs et al., 1990) are associated specifically with attenuation of neuroinvasiveness, whereas combined amino acid changes at corresponding positions in the YF E protein are associated with attenuation of neurovirulence (Schlesinger et al., 1996). It is possible that the combined effect of the three amino acid changes (R52G, F305V and R380T) in YF protein E is to alter the ability of virus to replicate both in peripheral tissue and in the CNS, producing a low neuroinvasiveness and neurovirulence phenotype. Alternative explanations include, firstly, that the amino acid change at position 305 alone may be responsible for the low neurovirulence of YF-17D or, secondly, that the type of attenuation linked to specific mutations in protein E is characteristic for a particular virus. As Schlesinger et al. (1996) did not specify the neuroinvasiveness phenotype of the two YF strains, the relationship between determinants of neuroinvasiveness and neurovirulence on the YF envelope protein remains unclear.
Several studies have shown that neurovirulence may be altered by mutations in the non-structural or untranslated regions of the flavivirus genome. Site-directed mutagenesis within the first glycosylation site of the DEN-4 (Lai et al., 1991) and the YF (Muylaert et al., 1996) NS1 protein results in a dramatic reduction in virus replication and NS1 expression in infected cells, accompanied by a large reduction in neurovirus activity compared to parental virus. By contrast, mutagenesis of the second glycosylation site of the DEN-4 and YF NS1 protein has no effect on virus replication or virulence (Lai et al., 1991; Muylaert et al., 1996). Although these studies show that adequate expression of NS1 in infected cells is necessary for efficient virus replication and for expression of the neurovirulent phenotype, they do not provide evidence for specific NS1 neurovirulence determinants in the absence of grossly abnormal virus replication. An MVE variant of low neurovirulence in mice, selected by serial passage in Vero cells (McMinn et al., 1991), was reduced in its ability to grow in mouse brain and in mouse neuroblastoma (N18) cells compared to parental virus, whereas replication in other vertebrate and invertebrate cell lines was unaffected. Nucleotide sequencing of the 5′ UTR and structural protein genes showed that the attenuated variant had an identical sequence to parental virus in this region of the genome, indicating that neurovirulence determinant(s) are located within non-structural genes and/or the 3′ UTR. Interestingly, the neuroinvasiveness of this variant was found to be identical with that of parental virus [the 50% lethal dose (LD₅₀) after intraperitoneal inoculation was very similar to the LD₅₀ after intracerebral inoculation], providing clear evidence that neurovirulence may be a separate phenotype from neuroinvasiveness. It is likely that the neurovirulence determinants within the non-structural gene region of this virus are involved in neurone-specific replication. Comparison of the full nucleotide sequences of a neuroviral strain of JE with three attenuated variants derived by serial cell or suckling mouse passage or by UV irradiation (Ni et al., 1995) revealed the presence of mutations within highly conserved sites in NS2B (E63D) and NS3 (A105G) of the attenuated strains. As mutations within the YF (Chambers et al., 1993) and DEN-4 (Falgout et al., 1993) NS2B protein which correspond to NS2B-63 of JE reduce or eliminate the cleavage activity of the virus encoded serine protease, it is possible that these mutations attenuate JE by altering viral protease activity and polyprotein processing. However, it is unclear what selective advantage mutations within the NS2B–NS3 protease may confer on the virus.

The mechanism by which neurovirulent flaviviruses induce encephalitis in the host is unknown; however, the mouse-nervoivirulent alphavirus Sindbis (SIN) has been shown to cause encephalitis in mice by inducing apoptosis in infected neurons (Lewis et al., 1996). Furthermore, induction of apoptosis in SIN-infected neurons is thought to be due to specific inactivation of the cellular oncogene bcl-2 by the viral glycoprotein E2 (Levine et al., 1993; Ubol et al., 1994). Recently, Despres et al. (1996) showed that the neurovirulence of DEN-1 strains in mice was associated with the ability to induce apoptosis in infected neuroblastoma cells and that the induction of apoptosis correlated with the quantity of expressed DEN-1 proteins. The possibility that the induction of apoptosis underlies the pathogenesis of flavivirus encephalitis is highlighted by recent evidence that the severity of encephalitis in MVE-infected mice correlates with the degree of neuronal apoptosis induced (T. Kendrick, V. B. Matthews & P. C. McMinn, unpublished data). Taken together, these data suggest that induction of apoptosis in infected neurons may cause flavivirus-induced encephalitis in mice, and thus the molecular pathogenesis of flavivirus encephalitis may be similar to that caused by SIN.

Summary

Studies of flavivirus virulence have focussed on variants with differing neurovirulence or neuroinvasiveness in animal models, with the majority based on determining the contribution of viral structural proteins, in particular the envelope glycoprotein. Specific virulence determinants on protein E have recently been identified by the application of several experimental techniques such as selection and characterization of neutralization escape variants in the presence of anti-E protein MAbs or by site-directed mutagenesis of the E gene of infectious cDNA clones (see Fig. 2, Tables 2 and 3). However, the association of particular amino acid substitutions in protein E with loss of neurovirulence and/or neuroinvasiveness appears to be dependent on the virus system under study and thus the mechanism by which protein E controls these two virulence phenotypes is unclear. Existing data show that neurovirulence and neuroinvasiveness determinants map to identical regions of protein E, including the lateral surface of domain III (the putative receptor binding site) and the base of domain II (thought to be involved in pH-dependent fusion activity). Attenuation of some domain II and III variants is associated with changes in binding to cell membrane proteins or in pH-dependent fusion activity, suggesting that defects in the early events of virus replication are responsible for virulence attenuation. A small number of amino acid changes in domain I are associated with attenuation of virulence, including mutations in the protein E glycosylation site which result in significant reductions in viral glycoprotein expression. The mechanism of attenuation associated with other mutations in domain I is unknown. Several studies have identified putative neurovirulence determinants within other regions of the genome (e.g. non-structural protein genes, 3′ UTR), and in one example the mutation attenuates neurovirulence without alteration in neuroinvasiveness. Further work is necessary to map neurovirulence determinants within these regions. It is likely that significant advances in this field will come from the application of infectious cDNA clone technology.
Sincere thanks to Professor C. J. Burrell, Dr N. Davis-Poynter, Dr M. Lindsay and Dr L. Sammels for critical appraisal of the manuscript. Thanks also to Dr F. X. Heinz for help with the preparation of Fig. 2. The author is supported by grants from the Raine Foundation of the University of Western Australia, the National Health and Medical Research Council of Australia and the Australian Research Council.

References


Lee, E., Feron, C., Simpson, R., Weir, R. C., Rice, C. M. & Dalgarno, L. (1990). Sequence of the 3’ half of the Murray Valley encephalitis virus genome and mapping of the nonstructural proteins NS1, NS3 and NS5. Virus Genes 4, 201–217.


