II. LITERATURE REVIEW

The genus *Flavivirus* within the family *Flaviviridae* comprises about 73 enveloped RNA viruses that are transmitted by either mosquitoes or ticks or without a known vector (Burke and Monath, 2001). For these viruses, a single-stranded RNA genome of about 11 kb encodes a polyprotein, which is cleaved by cellular and viral enzymes into three structural proteins (C, prM, and E) and seven nonstructural proteins (Lindenbach and Rice, 2001). Virion consists of two envelope proteins, E and prM, and an internal C protein, which binds genomic RNA. Differences in antigenicities of E allow the subdivision of flaviviruses into eight antigenic complexes and a number of unclassified viruses, which include the prototype yellow fever virus (YFV) (Calisher et al., 1989). More recent assignments based on nucleotide sequence variations of the nonstructural gene 5 generally agree with antigenic classifications (Kuno et al., 1998).

A. Biology of dengue virus

Dengue virus generally causes a mild febrile illness, DF. A much more severe form of the disease, DHF or dengue shock syndrome (DSS) occurs less frequently. The virus is transmitted by the *Aedes aegypti* mosquitoes, which are most active during the day and can usually be found near or in human dwellings. There are four distinct serotypes of dengue virus, consisting of a single-stranded linear RNA molecule, which is capped at the 5'end, but lacks a poly $(A)^+$ tract at the 3'end (Henchal and Putnak, 1990). The four serotypes are almost indistinguishable in terms of the clinical and pathological symptoms caused by them, but they could be identified by neutralization tests utilizing monoclonal antibodies and polymerase chain reaction (Cardosa, 1998). The RNA genome is about 10-11 kilobases in length, surrounded by an icosahedral nucleocapsid of about 30 nm in diameter and covered with a lipid envelope of about 10 nm in thickness. The complete virion is about 50 nm in diameter (Henchal and Putnak, 1990). Complete nucleotide sequences of the genomic RNA of the dengue viruses have been described. The order of the coding sequences is 5' C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 3'. The C-prM-E sequence

encodes the structural proteins, whereas the non-structural proteins are encoded in the region of NS1 to NS5. The genomic RNA has a type I cap at its 5'end (m^7GpppAmp) followed by the conserved dinucleotide sequence AG and ends with the conserved dinucleotide CU (Rice, 1996). The protein derived from translation of the viral genomic RNA is a single, long-precursor polyprotein that is successively cleaved by both host and viral-encoded proteases to produce structural and non-structural proteins (Henchal and Putnak, 1990; Rice, 1996).

1. Viral proteins

Dengue virus proteins consist of three viral structural (C, prM and E) and seven non-structural (NS) proteins.

1.1. Viral structural proteins: The mature virion contains three structural proteins: C protein, the nucleocapsid or core protein; M protein, a membrane-associated protein; and E protein, the envelope protein. The immature virions are mainly located intracellularly and contain prM protein, which is an M protein precursor.

The C protein, 112 to 127 amino acids in length and a molecular weight of 9 to 12 kDa, is the first viral polypeptide synthesized during translation. This protein has a high positive charge content because of the large numbers of lysine and arginine residues. This highly basic property is believed to allow C protein to associate with the negatively charged viral RNA molecule, forming the nucleocapsid (Hahn et al., 1988). The C protein lacks an N-terminal, hydrophobic signal sequence, which suggests that its synthesis occur on the non-membrane-bound ribosome. A hydrophobic stretch of amino acids at the C protein C-terminus probably acts as the transmembrane secretion signal for prM. This hydrophobic domain may serve to transiently anchor the C proteins to a membrane at the replication site after cleavage, probably by a host cell "signalase" at the N-terminus of the prM protein (Mason, 1989).

The prM protein (18.1-to-19.1 kDa) is the glycosylated precursor of the structural protein M and is contained within the intracellular immature virions. It forms a heterodimeric molecule with the E protein (Wengler and Wengler, 1989).

The E-prM interaction is believed to protect the E protein from irreversible inactivation during transport to the cell surface through acidic post-Golgi vesicles (Konishi and Mason, 1993). The prM protein contains six cysteine residues that form three disulfide bridges. However, during viral release, this protein is generally cleaved at acidic post-Golgi vesicles to form the M and the N-terminal pr segment. This cleavage occurs shortly before or simultaneously with virion release (Heinz and Allison, 2001).

The structural protein M (7 to 9 kDa) is located in the C-terminal portion of prM and is present in mature virions (Rice, 1996; and Henchal and Putnak, 1990). The cleavage is mediated by the host enzyme furin or other enzymes of similar specificity (Randolph et al., 1990; Stadler et al., 1997). This cleavage, which may occur in the acidic post-Golgi vesicle, appears to precede the virus release from the cell. The formation of M from prM appears to be the crucial, terminal event in virion morphogenesis. It results in a large increase in viral infectivity and reorganization of the viral surface structure that is composed of E-prM heterodimers in immature virions (Wengler and Wengler, 1989; Allison et al., 1995). The association of prM with the E protein in immature virions appears to make the latter protein resistant to the low pH, which induces the conformational change required for fusion. This association has been related to the stability of pH-sensitive epitopes of the E protein that prevents viral aggregation at an acidic pH (Heinz et al., 1994). The prM protein cleavage by furin *in vitro* results in a 100-fold increase in specific infectivity and the acquisition of membrane fusion and hemagglutination activity. In the mature virions the envelope protein E can undergo low-pH induced structural rearrangement that is characteristic of mature virions (Stadler et al., 1997). Cleavage of the prM can be blocked by the treatment of infected cells with ammonium chloride or bafilomycin A1, resulting in the release of virus particles containing the uncleaved prM precursor protein. These immature virions are less infectious than mature ones (Wengler and Wengler, 1989; Randolph et al., 1990; Heinz et al., 1994) and do not induce cell-cell fusion at an acidic pH (Guirakhoo et al., 1991; Guirakhoo et al., 1993).

The E glycoprotein (495 amino acids) is the major protein component of the mature virion surface. The E protein plays a role in several important biological activities including receptor binding, hemagglutination of goose erythrocytes, fusion

activity in acidic endosomes, virus assembly, and induction of protective immunity (Heinz et al., 1994; Raviprakash et al., 2000). The structure of the 395 residue N-terminal tryptic fragment of the tick-borne encephalitis virus (TBEV) E glycoprotein (Heinz et al., 1991; Allison et al., 1995) has been determined by X-ray crystallography (Rey et al., 1995). This E fragment crystallizes as a dimer, consistent with the expectation that TBEV assembly results in the formation of homodimers. Each monomer consists of three domains: the structurally central N-terminal domain I (residues 1-51, 137-189, 285-302), followed by the dimerization domain II (residues 52-136, 190-284) and finally the C-terminal Ig-like domain III (residues 303-395). The dimerization domain contains the hydrophobic 'fusion' peptide (residues 98-110) essential for virus-cell fusion (Allison et al., 2001). It has been proposed that domain III functions as the binding site for cellular receptors (Rey et al., 1995; Bhardwaj et al., 2001). In addition, domain III has been recognized as the receptor attachment site in competition experiments with monoclonal antibodies (Beasley and Aaskov, 2001; Crill and Roehrig, 2001; Lok et al., 2001). The 101-residue C-terminal end of the TBEV E glycoprotein, which was not part of the X-ray structure, is called the 'stem anchor' region, with the stem being composed of residues 396 to 449 and the hydrophobic transmembrane anchor region composed of residues 450-496. The stem and transmembrane anchor regions have each been predicted to consist of two α-helices (Stiasny et al., 1996; Allison et al., 1999).

1.2. Viral non-structural proteins: Seven viral non-structural (NS) proteins encoded in 3' terminus of polyprotein follow the E protein. The order of the non-structural proteins coding from the 5' end to 3' end is 5' NS1 - NS2A- NS2B - NS3 - NS4A - NS4B - NS5 3'.

NS1, the first non-structural protein, is a 42 to 52 kDa protein. NS1 includes 12 strictly conserved cysteine residues, one to three N-linked glycosylation sites, and regions of high sequence conservation. The function of NS1 in viral replication is unknown. The NS2 consists of two proteins, NS2a and NS2b. NS2a (18 to 22 kDa) is required for proper proteolytic processing of the C-terminus of NS1 (Falgout and Markoff, 1995). NS2b (13 to 15 kDa) is involved in the protease function of the NS2b- NS3 complex (Preugschat et al., 1990). The NS3, a hydrophilic protein (68 to

70 kDa), is highly conserved among flaviviruses and is believed to be an enzymatic component of the RNA replication machinery. NS3 is a large trifunctional protein, containing protease, helicase, and RNA triphosphatase activities. The N-terminal region of NS3 contains conserved amino acids shared by trypsin-like serine protease, whereas, the C-terminal domain is homologous to helicases (Bazan and Fletterick, 1989; Gorbalenya et al., 1998). RNA-stimulated NTPase activity has been demonstrated in the purified NS3 (Wengler and Wengler, 1991). The presence of proteinase and helicase domains within a single viral protein is a common structural motif found in alphavirus, pestivirus, potyvirus, and coronavirus proteins as well as flaviviruses. The primary function of a viral proteinase is in the regulated posttranslational production of individual gene products from a polyprotein precursor. The viral proteinase can also produce processing intermediates, which may themselves be functional components of the viral life cycle (Preugschat et al., 1990). The NS4a and NS4b are hydrophobic proteins, with the molecular weights of 16 kDa and 27 kDa, respectively. They might be RNA replication complex cofactors along with the putative viral RNA-dependent RNA polymerase, NS5. Finally, the NS5 protein (103 to 104 kDa) is the last protein encoded in the long open reading frame, the largest and most highly conserved flavivirus protein. NS5 is a basic protein, lacking of long hydrophobic stretches, and is believed to be the flavivirus RNA-dependent RNA polymerase. The N-terminal domain of NS5 (between residues 60 and 145) is homologous to a region of methyltransferases implicated in S-adenosylmethionine binding. It has been suggested that this domain may be involved in methylation of the 5'cap structure (Rice, 1996). The NS5 protein is at least a bifunctional protein, possessing both methyltransferase and RNA polymerase activities.

2. Dengue attachment and entry into host cell

Dengue virus attaches to the host cell surface via cellular receptors in order to infect a cell. The attachment is considered a major determinant of viral host-range and tissue tropism, and diverse cell surface molecules are identified as virus receptors. Dengue viral E protein (495 amino acid protein with two potential glycosylation sites) is the attachment protein for the virus (Kuhn et al., 2002; Modis et al., 2003, Rey,

2003). E protein contains multiple neutralization epitopes involved in the fusion of the virus membrane and in virus binding to cellular receptor molecules (Bielefeldt-Ohmann, 1998; Kuhn et al., 2002; Modis et al., 2003). The identification of dengue virus receptor(s) on target cells has been unclearly definitive. A cell surface protein on human monocytes, as well as some glycoproteins and glycosaminoglycans, has been speculated to play a role as virus receptors in past studies. Bielefeldt-Ohmann stated (1998) that the dengue virus binding entities on the cell surface membrane vary between cell types and so it is not clear how the dengue virus-binding molecules are related to the dengue virus binding entities. Previous studies indicated that cell surface heparan sulfates (HS) are involved in the attachment of dengue virus to mammalian cells including Vero, CHO, and human hepatoma cells (Chen et al., 1997; Germi et al., 2002; Hilgard and Stockert, 2000; Hung et al., 1999; Lin et al., 2002). Recently, DC-SIGN, a dendritic cell surface lectin, was also shown to mediate dengue virus infection to primary dendritic cells (Navarro-Sanchez et al., 2003; Tassaneetrithep et al., 2003). The molecular mechanism by which dengue virus enters cells (C6/36 mosquito cell and BHK21 cells) has been investigated. The E domain III plays an important role in binding of dengue virus type 2 with host cells (Hung et al., 2004). In addition, the E domain III interacts with heparan sulfates on BHK21 cells, and a loop region containing amino acid 380-389 of E domain III may participate in dengue virus type 2 binding to C6/36 cells (Hung et al., 2004).

At physiological pH, E proteins form dimers on the surface of the virion. Upon the exposure to pH below 6.5, a conformational shift occurs. This leads to a rearrangement of E to form trimers, allowing the fusion between the virion and the host cell membrane (Cardosa, 1998, Schalich et al., 1996; Allison et al., 2001). Being a positive sense RNA virus, dengue RNA itself is directly read by the host cell's ribosomes, functioning like a normal mRNA present in the cell. It stimulates the host cell to reproduce all the viral components through the processes of replication, transcription and translation. In order to create new viral RNA, the viral RNA creates an "anti-sense" version (negative sense RNA) of itself as a template. The newly formed viruses then leave the host cell by budding process. The dengue virus has an affinity for monocyte-macrophages, B lymphocytes and bone marrow cells in humans, causing an abrupt onset of high fever, abdominal flush, vomiting, headache and in

extreme cases extremely low platelet counts and circulatory collapse (Kurane et al., 1994; Rothman and Ennis, 1999).

3. Structure of mature virion and viral assembly

Mature infectious flaviviruses are icosahedral, ~500 Å in diameter (Kuhn et al., 2002) and contain a positive strand RNA genome of 10.7 kb with a single long open reading frame that is translated as a polyprotein of about 3,388 amino acids. Signal and stop-transfer sequences direct the translocation of the polyprotein back and forth across the membrane. The structural proteins are in the N-terminal region and are anchored in the endoplasmic reticulum by multiple membrane-spanning amino acid sequences. The polypeptide is subsequently cleaved by cellular and virally encoded proteinases and glycosylated by cellular glycosyltransferases to yield three structural proteins. For dengue virus, anchored capsid (anC) consisting of 113 residues, prM protein consists of 166 residues and glycoprotein E consists of 495 residues. The nucleocapsid core of the mature virion consists of the genomic RNA surrounded by multiple copies of the capsid protein C. This core is enveloped by a 40Å thick lipid bilayer derived from the endoplasmic reticulum of the host cell. Outside the membrane envelope is a layer of 180 copies of the E glycoprotein organized into a herringbone pattern (Kuhn et al., 2002) plus 180 copies of the M protein. Both the E and M proteins are anchored in the membrane by their C-terminal domains (Lindenbach and Rice, 2001).

For dengue virus, the structure of mature virus has been determined to 24 A° resolution (Kuhn et al., 2002) using cryoelectron microscopy (cryoEM) and image reconstruction techniques. The cryoEM density was interpreted in terms of 90 copies of an E:E homodimer, homologous with the TBEV E:E dimer crystal structure (Rey et al, 1995). The dengue virus structure does not have the anticipated T = 3 symmetry (Caspar and Klug, 1962) expected when there are 3×60 monomers in the glycoprotein shell. However, it was proposed (Kuhn et al., 2002) that acid pH in the endosome would weaken E-E adhesion within the E:E homodimer and allow the reorganization of the E molecules to generate a particle with the classical T = 3 symmetry. By extrapolation (Kuhn et al., 2002) from the cryoEM structure of a recombinant fusogenic T = 1 subviral particle (Ferlenghi et al., 2001), such a T = 3 structure would

have E:E:E homotrimers, that consists of E arranged in a similar way to those found in fusogenic subviral particles (Heinz and Allison, 2000), and contain sufficiently exposed membrane areas to allow fusion to the host cell.

Recently, structures of prM-containing immature dengue and yellow fever virus particles were determined to 16 and 25 Å resolution, respectively, by cryoelectron microscopy and image reconstruction techniques (Zhang et al., 2003). The closely similar structures show 60 icosahedrally organized trimeric spikes on the particle surface. Each spike consists of three prM:E heterodimers, where E is an envelope glycoprotein and prM is the precursor to the membrane protein M. The pre-peptide components of the prM proteins in each spike cover the fusion peptides at the distal ends of the E glycoproteins in a manner similar to the organization of the glycoproteins in the alphavirus spikes (Zhang et al., 2003). Each heterodimer is associated with an E and a prM transmembrane density. These transmembrane densities represent either an E:E or prM:prM antiparallel coiled coil by which each protein spans the membrane twice, leaving the C-terminus of each protein on the exterior of the viral membrane (Zhang et al., 2003), consistent with the predicted membrane-spanning domains of the unprocessed polyprotein.

The assembly of flaviviruses in the endoplasmic reticulum is followed by the modification of the two envelope proteins, E and prM, and virion export through the secretory pathway (Lindenbach and Rice, 2001). In addition to N-glycosylation and subsequent modifications, prM (approximately 19 to 23 kDa) is cleaved into a soluble pr peptide and a virion-associated M protein (approximately 8 to 8.5 kDa) by *trans*-Golgi resident furin (Stadler et al., 1997). This results in two different forms of virions; i.e. E- and prM containing intracellular virions and E- and M-containing extracellular virions. The spatial arrangement of E and prM/M is known for both forms and also for a recombinant subviral particle. On the surface of extracellular virions of dengue virus, head-to-tail dimers of E lie parallel to the lipid bilayer and 30 sets of three parallel dimers are arranged in a herringbone pattern over a layer of non-exposed M protein (Kuhn et al., 2002). A different arrangement (T=1) of E dimers is found on the icosahedral surfaces of subviral particles of another flavivirus, TBEV (Ferlenghi et al., 2001). In contrast, three prM-E heterodimers are organized into each of the 60 icosahedrally arranged spikes that project the pr portion of prM out of the

surfaces of intracellular dengue virions (Zhang et al., 2003). The pr portion of prM also covers the fusion peptide at the tip of domain II of E in intracellular virions (Zhang et al., 2003). Functionally, E binds cell surface receptors, involves in envelope fusion to the cellular membrane, and serves as the major target for neutralization by antibodies. The direct role of prM/M during the early phase of infection is not known (Burke and Monath, 2001; Lindenbach and Rice, 2001; Roehrig et al., 1998). In the case of dengue virus, some anti-prM monoclonal antibodies cause a low level of neutralization *in vitro*, but this activity may be due to their cross-reactivities with E (Falconar, 1999; Kaufman et al., 1989). On the other hand, it is well established that dengue virus prM and M actively induce a protective immune response, and passively administered anti-prM antibodies protect mice against a lethal challenge (Bray and Lai, 1991; Falconar, 1999; Kaufman et al., 1989; Vazquez et al., 2002). The protective effect of a non-neutralizing anti-prM monoclonal antibody was also observed for a Langat virus challenge (Iacono-Connors et al., 1996).

The final step in virion assembly is cleavage of the prM glycoprotein by furin or other furin-like proteases into an N-terminal, 91-amino acid long, pr-peptide and a 75- residue long M protein (for dengue virus). Furin is a cellular subtilisin-like endoproteinase that recognizes an (R/K)-X-(R/K)-R motif (Randolph et al., 1990; Stadler et al., 1997; Thomas, 2002). Cleavage leads to the dissociation of prM:E heterodimers (Wengler and Wengler, 1989; Allison, et al., 1995) and the formation of E:E homodimers (Rey et al., 1995; Stiasny et al., 1996). The glycosylated pre-peptide is released from the maturing particle, altering the accessibility of domain II of the E glycoprotein to antibody binding (Stiasny et al., 1996). Unlike the mature virion, the prM containing particles are unable to fuse with cells in acidic conditions (Guirakhoo et al., 1992). In contrast, mature particles become fusogenic when transported through acidified endosomes during the normal infectious process (Heinz et al., 1994). They undergo conformational changes that alter the antigenic properties once again, resulting in the formation of E:E:E homotrimers (Stiasny et al., 1996; Zhang et al., 2003). Thus the pr-peptide probably functions to prevent the E protein from prematurely undergoing conformational changes that trigger fusion in endosomal

vesicles (Heinz et al., 1994), analogous to the function of PE2 in alphaviruses (Strauss and Strauss, 1994).

B. Subtilisin-like proprotein convertases family

1. Members of subtilisin-like proprotein convertase (SPCs) family in mammalian cells

The processing of precursor proteins via limited proteolysis is an important and widely used cellular mechanism for the generation of biologically active proteins and peptides in appropriate cellular compartments. The major endoproteolytic processing enzymes of the secretory pathways are the endoproteases homologous to bacterial subtilisin or the processing protease, Kexin (Kex2p), in yeast (Nakayana, 1997; Steiner, 1998; Zhou et al., 1999: Thomas, 2002). Furin in mammalian cells is encoded by the *fur* gene on chromosome 15. Subsequently, seven mammalian Kexin-homologue enzymes were identified. Although each enzyme had been independently named by the discoverers, a simplified nomenclature for the group of mammalian processing proteases has been proposed (Chan et al., 1992; Bergeron et al., 2000; Thomas, 2002), using the term subtilisin-like proprotein convertases (SPCs):

SPC1 [Furin/pair amino acid convertase enzyme (PACE)]

SPC2 [Prohormone convertase 2 (PC2)]

SPC3 [Prohormone convertase 1 (PC1)/Prohormone convertase 3 (PC3)]

SPC4 [Pair amino acid convertase enzyme 4 (PACE4)]

SPC5 [Prohormone convertase 4 (PC4)]

SPC6A [Prohormone convertase 5 (PC5)/Prohormone convertase 6A(PC6A)

SPC6B[Prohormone convertase 5 (PC5)/Prohormone convertase 6B(PC6B)

SPC7 [Lymphoma prohormone convertase (LPC)/Prohormone convertase 7 (PC7)/Prohormone convertase 8 (PC8)].

2. General structure of subtilisin-like proteins convertases

The subtilisin-like catalytic domain that extends approximately over 330 amino acids is highly conserved among the eukaryotic proprotein convertases. In particular, the active-site residues of the Asp, His and Ser catalytic triad and an Asn

residue which stabilizes the oxyanion hole in the transitional state (Bryan et al., 1986) are present at corresponding positions in all members; except for PC2, where the Asn residue is replaced by an Asp. The sequences flanking these residues are conserved as well. In addition, about 140-amino-acid region following the catalytic domain, which has been variously referred as the 'Homo B', 'P' or 'middle' domain, is well conserved among eukaryotic convertases, including yeast Kex2p, but is absent in bacterial subtilisins. The Homo B domain is essential for catalytic activity (Zhong et al., 1996; Hatsuzawa et al., 1992; Takahashi et al., 1993; Takahashi et al., 1995). Within this domain, there is a conserved Arg-Gly-Asp sequence that is reminiscent of the recognition sequence for integrins. Mutation of these three residues in PC1/PC3 results in loss of catalytic activity and mis-sorting of neuroendocrine convertase towards the constitutive secretory pathway (Lusson et al., 1997; Nakayama, 1997). Another conserved region is the propeptide, which is autocatalytically removed through cleavage at an Arg-Xaa-Lys-Arg site during maturation of the convertases. Towards the C-terminus, furin, PACE4 and PC5/PC6 A and B have a Cys-rich domain. The Cys topography is well conserved (Nakagawa et al., 1993; Roebroek et al., 1992), although its role is currently unknown. Furin, PC5/PC6B and LPC/PC7/PC8/SPC7 also have a transmembrane domain near the C-terminus (Nakayama, 1997; Thomas, 2002).

Recently, a three-dimensional structure of mouse furin has been demonstrated (Henrich et al., 2003). Similar to subtilisin, the core of the catalytic domain consists (Siezen et al., 1994; Ueda et al., 2003) of a highly twisted β-sheet composed of seven parallel and one antiparallel β-strands, flanked by five helices adjacent and two peripheral helices and by two β-hairpin loops (Henrich et al., 2003). However, a number of extended surface loops, several of them surrounding and shaping the active site cleft or mediating interaction with the P domain, show quite different lengths and conformations when compared with those observed in bacterial subtilisins (Henrich et al., 2003). With the typical catalytic triad of subtilisin-related serine proteases, the catalytic domain of mouse furin has asparagine, histidine, aspartic acid and serine residues at the active site (Bergeron et al., 2000; Henrich et al., 2003). The P domain is organized as a separate eight-stranded β-sandwich and the complete folding of P domain appears to be regulatory of the catalytic activity. It is also possibly involved

in the stabilizing effect on this binding cleft of catalytic domain of SPCs members (Zhou et al., 1998; Ueda et al., 2003, Henrich, et al., 2003). It may do this by helping to balance the surface charge asymmetry in the substrate-binding region of the catalytic domain. The roles of the substrate recognition of subtilisin-like proprotein convertases have been identified by mutational analysis. The results confirmed that the negative charged residues of the catalytic pockets interact with positive charged residues of the substrate (Creemers et al., 1993; Henrich et al., 2003). The consensus substrate sequence is Arg(P4)-Xaa(P3)-Lys/Arg(P2)-Arg(P1)↓-Xaa(P1')-, where Xaa can be almost any amino acid and ↓ represents the cleavage site where the peptide bone is hydrolyzed. In addition, the mouse furin crystallized structure can also explain furin's preference for the basic residues at P3, P5 and P6 sites (Henrich et al., 2003).

Autoactivation of SPC1 (furin) serves as a model for the other subtilisin-like proprotein convertases, with the exception of SPC2 (PC2). Intramolecular cleavage of the propeptide (Pro) allows furin to exit the endoplasmic reticulum (Nakayama, 1997). However, the propeptide remains attached noncovalently until the cleaved inactive proenzyme reaches the *trans*-Golgi network, where an increased acidic ($pH \leq 6.5$) and calcium-enriched environment facilitates dissociation of the propeptide. A second cleavage within the propeptide then precludes further inhibitory interactions, which results in full activation. A similar mechanism of activation has been demonstrated for SPC3 (PC1/PC3), SPC5 (PC4) and SPC7 (LPC/PC7/PC8). Propeptide of SPC4 (PACE4) is autoclaved slowly, but it can probably also occur prior to its exit from the endoplasmic reticulum (Zhou et al., 1999).

3. Furin/SPCs localization and trafficking

SPC1 (furin/PACE) is a secretory pathway endoprotease that catalyses the maturation of a strikingly diverse group of proprotein substrates, ranging from growth factors and receptors to pathogen proteins in multiple compartments within the *trans*-Golgi network/endosomal system. Human furin is initially synthesized as 100 kDa core-glycosylated pro-furin, which is later converted into 94 kDa forms by the removal of propeptide at the Arg-Ala-Lys-Arg site, at the residues 104-107 (Denault and Luduc, 1995; Nakayama, 1997; Thomas, 2002).

SPC2 (PC2) is more complex in transport and activation. This convertase, which is unique, requires the acidic conditions of a late post-Golgi compartment for activation. In the ER, proSPC2 interacts with 7B2, a 27-kDa neuroendocrine secretory protein that is coexpressed with SPC2 in many neuroendocrine tissues. The 7B2 protein is required for the production of active SPC2, which might function as a chaperone protein to assist in proSPC2 folding. The 7B2 protein is cleaved at a polybasic site toward the C-terminus, similar to the cleavage of furin or related TGN protease, which results in the release of an N-terminal 21-kDa form and an inhibitory C-terminal fragment (Steiner, 1998; Bergeron et al., 2000).

Several SPC members of subtilisin-like convertases, SPC3 (PC1/PC3), SPC4 (PACE4), SPC5 (PC5/PC6), SPC7 (LPC/PC7/PC8), have sequence specificity similar to that of SPC1 (furin) (Henrich et al., 2003). The mammalian subtilisin-like proprotein convertases can be classified into three groups based on their tissue distribution. SPC1 (furin), SPC4 (PACE4), SPC6 (PC5/PC6) and SPC7 (LPC/PC7/PC8) are expressed in a broad range of tissues and cell lines (Thomas, 2002). In contrast, the expression of SPC2 (PC2) and SPC3 (PC1/PC3) is limited to neuroendocrine tissues, such as pancreatic islets, pituitary, adrenal medulla and many brain areas (Thomas, 2002). The expression of SPC5 (PC4) is highly restricted to testicular spermatogenic cells (Thomas, 2002). Within cells, SPC1 (furin) and SPC7 (LPC/PC7/PC8), both of which have a transmembrane domain, are localized in the *trans*-Golgi network (TGN) (Nakayama, 1997; Steiner, 1998; Thomas, 2002). Another SPC with a transmembrane domain, SPC6B (PC5/PC6B), is also localized in the Golgi area, although it appears not to concentrate in the *trans*-Golgi network. The neuroendocrine-specific convertases, SPC2 (PC2) and SPC3 (PC1/PC3), are mainly localized in secretory granules. SPC6A (PC5/PC6A) has been reported to be localized to secretory granules (Nakayama, 1997; Seidah and Chretien, 1994).

4. Biochemical and enzymatic feature of furin

Consistent with furin as the first and so forth the best-characterized enzyme of subtilisin-like proprotein convertase, furin's enzymatic activity and specificity has been evaluated *in vitro* and *in vivo*. In vitro assays, relying on the use of either the shed form of furin and/or the soluble furin measure, the cleavage of various substrates

such as the fluorogenic tetrapeptides or use of internally quenched fluorogenic peptidyl substrates (Krysan et al., 1999). Krysan and his coworker (1999) found that furin could cleave such substrates with kinetics comparable to those observed with extended peptides and physiological substrates. With the best of these hexapeptidyl methylcoumarinamides, furin displayed k_{cat}/K_M values greater than 10^6 $M^{-1}s^{-1}$. Furin exhibited striking substrate inhibition with hexapeptide but not with tetrapeptide substrates, an observation of significance to the evaluation of peptide-based furin inhibitors. Quantitative comparison of furin and Kexin recognition at P1, P2, and P4 demonstrates that whereas interactions at P1 make comparable contributions to catalysis by the two enzymes, furin and Kexin exhibited a 10-fold lesser dependence on P2 recognition. However, a 10-100 fold greater dependence on P4 recognition was observed. Furin has recently been shown to exhibit P6 recognition. It was demonstrated that this interaction contributes 1.4 kcal/mol toward catalysis independence of the nature of the P4 residue. They have also shown that those favorable residues at P2 and P6 compensate for fewer optimal residues at either P1 or P4. The quantitative analysis of furin specificity sharply distinguishes the nature of substrate recognition by the processing and degradative members of subtilisin-related proteases.

The heterogeneity of SPC in insects appears to be less extensive than in mammals. Analysis of the whole genome sequence of *Anopheles gambiae* reveals only three members of the furin-like proprotein convertase family (furin 1, dfurin 2, and amontillado) in anopheles mosquito (Riehle et al., 2002). In the *Aedes aegypti* mosquito, the deduced amino acid sequence of pro-vitellogenin convertase has a high similarity to a domain structure characteristic of subtilisin-like convertases. Chen and Raikhel (1996) identified and characterize a 115-kDa pro-vitellogenins (pro-Vg) processing enzyme, pro-vitellogenin convertase (pro-VC), from a vitellogenic female fat body cDNA library of *Aedes aegypti* mosquito. The amino acid sequence of vitellogenin convertase reveals a high structural similarity to furin-like convertase. It has a typical structure of furin-like convertases which comprised of pre-propeptide, catalytic, cysteine-rich, C-terminal transmembrane, and cytoplasmic domains and recognizes the same cleavage sites that contain paired basic amino acid motif, Arg (P4)-Xaa(P3)-Lys/Arg(P2)-Arg(P1)↓.

5. Implication of furin cleavage and flaviviruses infectivity

Current evidence implicates prM as a chaperone for E during intracellular virion assembly and maturation (Lindenbach and Rice, 2001). After cleavage of the flaviviral polyprotein at the C-prM, prM-E, and E-NS1 junctions by host signalase, prM and E noncovalently associate in the endoplasmic reticulum; prM-E heterodimers are subsequently incorporated into immature virions (Wengler and Wengler, 1989). The heterodimeric interaction between prM and E is important for the proper folding and transport of E (Konishi and Mason, 1993), whereas folding of prM does not require the presence of E (Lorenz et al., 2002). Based on TBEV, sites located within the predicted α-helical regions of the stem and the membrane-spanning region of E are required for stabilization of the prM-E heterodimer (Allison et al., 1999). During the transport through the secretory pathway, the prM prevents E from undergoing premature conformational changes and oligomeric rearrangements induced by the acidic pH of the *trans*-Golgi network (Allison et al., 1995; Guirakhoo et al., 1992; Heinz et al., 1994; Stiasny et al., 1996). In a recently study, the infectivity of TBEV prM-enriched particles that were produced by the addition of NH_4Cl to infected cells could be enhanced ~100-fold after furin treatment (Stadler et al., 1997). Furthermore, about a third of the particles produced by furin treatment of immature dengue virus had a smooth surface and diameters consistent with that of mature particles as opposed to the bumpy surface and larger diameter of the immature particles (Zhang et al., 2003). In contrast, there were no smooth-surfaced particles in untreated immature dengue virus samples (Zhang et al., 2003). The correlation coefficients relating the furin-treated particle projections to the native particle reconstruction (Kuhn et al., 2002) also indicated that the immature particles had been changed to mature infectious virions. Therefore it is apparent that the structure of the prM-containing particles (immature particles) was converted to that of infectious virions by the furin treatment and that the immature particles serve as intermediates in the assembly pathway (Zhang et al., 2003). The furin treatment results in the conversion of the prM:E heterodimer to an E:E homodimer and releases the pre-peptides from the top of the spikes, thus allowing the E glycoproteins to lie parallel to the membrane.

The effect of enhancement of the dengue prM cleavage on infectivity was studied recently by the construction of full-length cDNA clones in which the 13-amino acid sequence just proximal to dengue pr-M junction was replaced with the homologous region derived from the pr-M junction of Japanese encephalitis (JE) virus (Keelapang et al., 2004). The 13-amino acid residues proximal to the pr-M junction of JEVpr/16681 chimeric virus shared the furin consensus sequence of dengue parent strain, but differs by the presence of three additional positively charged residues, resulting in an increase the positive charges from four to seven residues and the absence of two negatively charged residues (Table 1). In addition, it contained one additional minimal furin motif, Arg-Xaa-Xaa-Arg, which is known to be cleaved by furin in a few precursors (Brennan and Nakayama, 1994; Hatsuzawa et al., 1992; Klimpel et al., 1992; Molloy et al., 1992; Morsy et al., 1994). Moreover, the side chains of two non-charged amino acids at the amino acid positions P9 (197) and P6 (200) of JEVpr/16681 pr-M junction also differ from those in the 16681 parent virus. Analysis of the resultant JEVpr/16681 chimeric virus revealed that the cleavage of prM in JEVpr/16681 was clearly enhanced as compared with the parent dengue virus strain 16681. Unexpectedly, while the prM protein of JEVpr/16681 is cleaved to the greater extent than the parent strain, the replication of this chimeric virus is substantially reduced when compared with the parent virus (Keelapang et al., 2004). However, it is not known whether the enhanced prM cleavage is due to each or combination of changes of the amino acid sequence at the13-amino acid region proximal to the pr-M junction. It also remains unclear how or what sequence variations of amino acid sequence proximal to the pr-M junction of JEVpr/16681are involved in the reduction of dengue replication.

Table 1. Comparison of pr-M junction of insect-borne flaviviruses and pr-M junction chimeras

Virus[a]	Amino acid sequence[b]	P1 to P13 charge[d]				Accession no.[e]
		+	-	Net	Set[c]	
Dengue virus type 1	YGTCS QTG**E**H **RRD**KR↓SVALA PHVGL	4	2	+2	1	P33478
Dengue virus type 2	YGTCT TMG**E**H **RRE**KR SVALV PHVGM					U87411
Dengue virus type 3	YGTCN QAG**E**H **RRD**KR SVALA PHVGM					NC001475
Dengue virus type 4	YGTCT QSG**ER** **RRE**KR SVALT PHSGM					M14931
TBEV	YG**R**CG **K**Q**E**GS -RTRR SVLIP SHAQG	5	1	+4	1	U27495
Langat	YG**R**CG **RRE**GS -RSRR SVLIP SHAQR					P29837
Sofjin	YG**R**CG **K**Q**E**GS -RTRR SVLIP SHAQG					X03870
KFDV	YG**R**CG **K**PAGG -RNRR SVSIP VHAHS					X74111
Louping ill	YG**R**CG **K**Q**E**GS -RTRR SVLIP THAQG					M59376
Powassan	YG**R**CG **R**QAGS -RGKR SVVIP THAQK					L06436
Yellow fever	YG**KCD** SAG**R**S **RR**SRR AIDLP THENH	6	1	+5	1	AF094612
JEV	YG**R**CT **RTRH**S **K**RSRR SVSVQ THGES	7	0	+7	2	M55506
Kunjin	YG**R**CT **K**TRHS **RR**SRR SLTVQ THGES					D00246
MVEV	YG**R**CT **RARH**S **K**RSRR SITVQ THGES					AF161266
SLEV	YG**R**CT **RMGH**S **RR**SRR SISVQ HHGDS					M16614
West Nile	YG**R**CT **K**TRHS **RR**SRR SLTVQ THGES					NC001563
16681Nde(+)	YGTCT TMG**E**H **RRE**KR SVALV PHVGM	4	2	+2	1	NA
TBEVpr/16681	YG**R**CG **K**Q**E**GS -RTRR SVALV PHVGM	5	1	+4	1	NA
YFVpr/16681	YG**KCD** SAG**R**S **RR**SRR SVALV PHVGM	6	1	+5	1	NA
JEVpr/16681	YG**R**CT **RTRH**S **K**RSRR SVALV PHVGM	7	0	+7	2	NA
16681pr(+7, -2)	YG**R**CT **RMREH** **RR**EKR SVALV PHVGM	7	2	+5	3	NA
16681pr(+4, -0)	YGTCT TMG**HH** **RR**SKR SVALV PHVGM	4	0	+4	1	NA
16681pr(+7, -0)	YG**R**CT **RMRHH** **RR**SKR SVALV PHVGM	7	0	+7	3	NA
16681pr(+9, -0)	YG**RCR** **RSRRS** **RR**SKR SVALV PHVGM	9	0	+9	3	NA
Cleavage position	15 11 6 1 -1 -6 -10					

a KFDV, Kyasanur Forest disease virus; MVEV, Murray Valley encephalitis virus; SLEV, St. Louis encephalitis virus.

b Amino acids in bold represent additional charged residues between positions P1 and P13. The arrow indicates the prM cleavage site.

c Set of non-overlapping R-X-K/R-R and R-X-X-R consensus sequences (when X can be any amino acid residue).

d For antigenic complexes with known sequences from multiple members, charge contents represent those of the consensus sequence.

e NA, not applicable.