CHAPTER - 2
REVIEW OF LITERATURE
CHAPTER-2
REVIEW OF LITERATURE

2.1. ALPHAVIRUSES

Genus *Alphavirus* currently enlists 29 species according to the IX edition of the International Committee for the Taxonomy of Viruses (ICTV). Together with *Rubivirus* it forms the family *Togaviridae* [17] of Group IV positive stranded RNA viruses. Primarily arthropod borne, these viruses are spread almost all over the globe except Antarctica and have been divided on the basis of geographic location into old world and new world alphaviruses. Semiliki Forest virus (SFV), Sindbis Virus (SINV), Ross River virus (RRV), O’nyong-nyong Virus (ONNV) and Chikungunya virus (CHIKV) possibly originated in old world countries (Europe, Asia, Africa and Australia) and have therefore been grouped as old world alphaviruses [17-19]. New world alphaviruses, as the name suggests, are distributed throughout new world countries of North and South America; and include Venezuelan Equine Encephalitis virus (VEEV), Western Equine Encephalitis virus (WEEV) and Eastern Equine Encephalitis virus (EEEV) [20]. There exists a close phylogenetic relationship between the old world and new world alphaviruses [21, 22]. This has been attributed to the transoceanic exchanges mediated by commercial trading between America and Africa, as a result the geographic origin of arboviruses in the genus is not completely resolved [17, 19].

On the basis of serological cross reactivity of the sequenced members of the genus; alphaviruses have been categorized into 8 antigenic complexes: Eastern, Western and Venezuelan equine encephalitis, Ndumu, Trocara, Middleburg, Semiliki Forest (SF) and Bramah Forest (BF). Out of these, four sero-complexes that contain the medically most important viruses include- EEE, WEE, VEE and SF serocomplexes. SINV and Ockelbo virus are members of the WEE virus serocomplex while RRV, CHIKV, Mayaro, and Getah viruses constitute the SF serocomplex. EEE and VEE serocomplexes contains EEEV and VEEV, respectively.
2.1.1. Alphaviruses as human pathogens

Alphavirus infection can be broadly categorized into three disease patterns. In humans, infection caused by WEEV and EEEV is asymptomatic or manifested as a nonspecific febrile illness or aseptic meningitis. Viral pathogenesis generally involves introduction of virus by vector per-cutaneously, followed by development of viremia. Initial symptoms usually found in the patients infected with WEEV or EEEV are malaise, headache and fever, followed by nausea and vomiting. Symptoms intensify, and somnolence or delirium may progress into coma over next few days. Seizures, impaired sensorium and paralysis are common [23]. The second disease pattern is the one caused by viruses belonging to VEEV serocomplex. Infection is marked by acute febrile illness with well-defined systemic symptoms. The central nervous disease is less frequent and mild in patients when compared to EEEV and WEEV.

CHIKV a prototype for those causing an acute febrile illness with malaise, rash, severe arthralgias, and arthritis; comprises the third type of alphavirus based disease pathogenesis. ONNV, Mayaro virus and RRV are closely related antigenically to CHIKV and cause similar or identical clinical manifestations. SINV causes similar but milder diseases known as Ockelbo (in Sweden), Pogosta (Finland), or Karelian fever (Russia). Its symptoms include abrupt onset of fever, chills, malaise and joint aches coincident with viremia. The migratory arthralgia, which is common characteristic of these viral diseases, affects extremities (phalanges, wrist and ankles) and occurs more prominently in adults than children. In more severe cases the arthritic pain may persist for several months [24].

Until 2005, old world alphaviruses were of less importance to most government health departments all over the world because of their low mortality and/or morbidity rate in previous localized outbreaks. However, with reemergence of CHIKV as a major epidemic threat in 2006, alphaviruses have taken a center stage as prominent human threat.
2.2. CHIKUNGUNYA VIRUS

Chikungunya virus (CHIKV) belonging to SF serocomplex of alphavirus has been categorized as a Category C Priority Pathogen by National Institute of Allergy and Infectious Diseases (NIAID), USA. Chikungunya disease was first described in the year 1952 in Tanzania [25, 26]. In reference to the stooped posture acquired in the chronic infection owing to severe incapacitating arthralgia the disease was called “Chikungunya; meaning that which bends up,” in Makonde language. Ever since chikungunya fever was first identified in 1952, it has resulted in major epidemics in Africa, India, Myanmar, Sri Lanka and Thailand. The epidemics caused were characterized by intermittent period of about 10 years and the disease was rarely fatal and severe. However, CHIKV caused major epidemics in India, Myanmar, Sri Lanka and Thailand in 1960s. From 2005 till 2015, major and frequent outbreaks of CHIKV have occurred in Indian Ocean islands, Southeast Asia, Europe and also Americas. Rapid urbanization, increased air travel, increased range of vector, inadequate sanitation and microevolution of viral genome are some of the important factors that have contributed to global reemergence of this disease. Despite considerable research efforts over the past decade, effective vaccines or drugs against the disease are not yet successful.

2.2.1. CHIKV- History of outbreaks and Geographic Distribution:

Chikungunya fever, previously neglected arboviral disease, has caught scientific attention as it raves around globe today. Symptoms matching with CHIKV infection have been reported as early as 1779 in Indonesia and Egypt but were documented as Dengue outbreaks [27]. In 1820s a wide spread febrile illness with symptoms consistent with CHIKV infection was proposed to originate from Zanzibar, an archipelago island in Indian Ocean and spread to Caribbean, West Indies and India through slave trade. More than two third of the population was affected in the areas where infection was present but without being fatal. Ninety five percent of the population of Calcutta was affected during the epidemic that lasted 3 months [27]. In 1853 and 1871, Calcutta again witnessed epidemic fever with arthralgia in which more than seventy five percent of the population was affected. After remaining silent for about 50 years India was again hit by a similar outbreak in the year 1923 [27].
CHIKV was isolated for the first time in 1952 from Makonde Plateau in Tanzania during an epidemic of dengue like disease [26, 28]. Serological and antigenic analysis of the virus revealed it to be an alphavirus, closely related to Mayaro and SFV [29]. The virus was isolated from Asia for the first time in 1958 from Bangkok [30]. Since then several confirmed CHIKV outbreaks have been reported from this area. First confirmed CHIKV epidemic from India was reported in the year 1963 in Calcutta (currently Kolkata) [31, 32] followed by infection in Madras (Current Chennai) in which more than 3 million individuals were infected [33, 34]. Viral transmission was apparent in various countries of Western Africa (Senegal, Nigeria, Republic of Guinea Benin and Cote d’Ivoire) and Central and South African countries (Sudan, Uganda, The Central African Republic, Democratic Republic of Congo, Zimbabwe, Malawi, Kenya and South Africa) between 1960 and 2000. Though there were numerous localized outbreaks of Chikungunya in Asia and Africa but these cannot be compared to the current pandemic that spread from La Reunion Island in East Africa to South Asia and from there to South-east Asia since 2005.

![Figure 2.1](https://www.cdc.gov/chikungunya/map/index.html) Countries where human CHIKV infection cases have been reported as of March 2015. This figure is taken from [www.cdc.gov/chikungunya/map/index.html](http://www.cdc.gov/chikungunya/map/index.html). Countries, in which only imported CHIKV cases were reported, have not been considered in this figure.
2.2.2. CHIKV genotypes

On the basis of phylogenetic analysis of cDNA sequences encoding E1 protein, CHIKV isolates have been categorized into three genotypes worldwide: East/Central/South African (ECSA), West African and Asian [8, 35, 36]. The phylogenetic analysis of the Indian Ocean isolates reported after the re-emergence of the virus indicates that the current CHIKV outbreak was caused by the ECSA genotype. Moreover, phylogenetic studies of the isolates revealed that Asian genotypes have a high degree of conservation, whereas African genotypes have higher diversity [8, 36].

![Figure 2.2: Chikungunya virus structure](image)

Figure 2.2: Chikungunya virus structure

Schematic representation of the CHIKV virion cross section, showing capped genomic RNA in the center of the virion encapsulated by the capsid protein. The E1/E2 heterodimeric spikes are embedded in the viral envelope membrane derived from the host cells. (This figure has been adapted from ViralZone:www.expasy.org/viralzone, SIB Swiss Institute of Bioinformatics).

2.2.3. Structure of the Virion

Like all alphaviruses, CHIKV virion has a characteristic icosahedral protein core made of capsid, that encapsidates single-stranded RNA genome [37]. This core is enveloped by a lipid bilayer derived from the infected host cell plasma membrane during budding. The mature virion
which is 70 nm in diameter contains 240 heterodimers of viral glycoproteins E1/E2 arranged as trimeric spikes [20, 38 and 39] on its surface (Figure 2.2)

2.2.4. Genome Organization

The CHIKV genome is a single-stranded, positive-sense RNA of approximately 11.8 kb in size. The viral genome bears significant resemblance to host mRNAs, as it possesses 5’ and 3’ untranslated regions (UTR), with a 5’ terminal methylguanylate cap and 3’ terminal polyadenylated tail (Figure 2.3). The CHIKV genome is comprised of two distinct regions: the 5’ two-thirds of the genome, which encodes the four nonstructural proteins (nsP1, nsP2, nsP3, and nsP4), and the remaining 3’ one-third, which encodes the five structural proteins (capsid, E3, E2, 6K and E1; Figure 2.3 [40]). The structural proteins are expressed at high levels as a subgenomic message from an internal 26S mRNA promoter. The sub-genomic RNA is also capped and polyadenylated like the viral genome [37]. Apart from this, there are four highly conserved regions known as cis-acting conserved sequence elements (CSEs) across alphaviruses. These regions have been found to be involved either in the synthesis of plus stranded RNA from minus stranded template or for synthesis of subgenomic RNA of the virus [20, 41, 42].

![Figure 2.3: Genome organization of CHIKV](image)

The nonstructural proteins of CHIKV (nsP1, nsP2, nsP3 and nsP4) are encoded as a polyprotein from the N terminal open reading frame (ORF 1) while the structural polyproteins (Capsid, E3, 6K, E2 and E1) are encoded by the C terminal ORF 2 [40].
2.3. VIRAL LIFE CYCLE

2.3.1. Entry and disassembly of the virus

Chikungunya exhibits a wide host range starting from mosquito to humans, however, the cellular receptor(s) specific to CHIKV remain elusive. Surface molecules on host membrane like heparan sulfate proteoglycans, Laminin, C-type lectins like Dendritic Cell-Specific Intracellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN) and Liver/lymph node-specific (L)-SIGN have been implicated as alphavirus receptors [43]. Natural resistance-associated macrophage protein (NRAMP) has been recently shown to be involved in binding and entry of SINV in both drosophila and mammalian cells [44]. CHIKV can use one or a combination of these to facilitate viral cell entry by receptor mediated endocytosis. Studies on other members of alphaviruses reveal that regardless of the identity of the host receptor, the viral glycoprotein E2 binds the receptor, as antibodies against E2 neutralize viral infectivity and E2 mutants are unable to bind cells [45, 46]. Further, studies involving disruption of endosomal acidification, cholesterol depletion and functional disruption of endosomal trafficking [47] revealed that CHIKV entry is inhibited in human as well as mosquito cells [48, 49]. Earlier, the mechanism of viral entry known for the alphaviruses was extended to CHIKV, where on binding to receptor the alphaviruses enter the cell by clathrin mediated endocytosis [20]. However, Bernard and colleagues have demonstrated that CHIKV preferentially uses Eps-15 dependent and clathrin independent pathway to infect mammalian cells [47]. On entry, viral glycoproteins undergo a conformational change, as a result of the acidic pH of endosomes, resulting in fusion of viral envelop with the endosomal membrane [20]. This leads to release of nucleocapsid in the cytosol. The released nucleocapsid is highly unstable and rapidly dissociates to release capsid proteins, which then bind to ribosomes to attain stability [50]. Nucleocapsid disassembly releases the viral RNA into the host cell cytoplasm (Figure 2.4). Once the viral genome is uncoated and released into the cytoplasm, it is ready to be translated [37].

2.3.2. Replication of the CHIKV Genome

Once the viral genetic material is released into the host cell, it is directly translated by the host cellular machinery because of its similarity to the host mRNA. The 5’ two third of viral genome is translated as two polyproteins, P123 and P1234 which are further processed to form nonstructural proteins [51, 52]. During initial phase of viral replication, P1234 can be
autoproteolytically cleaved in cis into P123 and nsP4. The presence of an opal termination codon at the C terminus of nsP3 results in the formation of these two types of the polyprotein by read through mechanism of the polymerase (Figure 2.4). In case of SINV the complete polyprotein (P1234) is formed in 10-20% of translations [51-53]. More frequent synthesis of P123 further results in the accumulation of P123 during early infection [20]. P123 is further cleaved in trans at nsP1/nsP2 junction to form mature nsP1 protein. The resulting proteins then readily cleave nsP2/nsP3 in trans, forming nsP2 and nsp3 mature proteins. NsP2 protein contains a papain-like protease and is involved in the processing of precursor polyproteins into mature nonstructural proteins. The presence of proteins capable to cleaving nsP2/nsP3 at this point also results in formation of P12 and P34 from P1234 [20]. In the initial phase of infection the viral genome synthesizes both positive and negative sense RNA. The production of minus strand of viral RNA within the initial hours of infection is critically dependent upon the presence of P123 and P23. The negative sense RNA is synthesized by two complexes comprising of (i) P123 intermediate and nsP4 and (ii) P23 intermediate with nsP1 and nsP4 [53, 54]. The synthesis of negative sense RNA wanes out after 3 hours of infection; in contrast, positive strand synthesis remains steady throughout infection. Alphavirus replication occurs in replication complex in association with cellular membrane and utilizes the negative sense RNA as template for generation of both positive genomic and subgenomic RNAs [20, 55]. The replication complex is comprised of the four mature individual nonstructural proteins. The subgenomic RNA expressed at high levels under the influence of internal 26S promoter serves for amplification of structural proteins in the infected cells.
Figure 2.4 The Alphavirus life cycle

CHIKV enters target cells by endocytosis. Following endocytosis, in the acidic environment of the endosome, viral spike undergoes conformational changes to expose the E1 fusion peptide, which mediates virus–host cell membrane fusion [20, 47]. This leads to cytoplasmic delivery of the core and release of the viral genome. Two precursors of non-structural proteins (nsPs) are translated from the viral mRNA (P123 and P1234), and cleavage of these precursors generates nsP1–nsP4. These proteins assemble to form the viral replication complex, which synthesizes a full-length negative-strand RNA intermediate which serves as the template for the synthesis of both subgenomic (26S) and genomic (49S) RNAs [20]. The subgenomic RNA drives the expression of the C–pE2–6K–E1 polyprotein precursor, which is processed by an autoproteolytic serine protease. The capsid (C) is released, and the pE2 and E1 glycoproteins are generated by further processing through ER. pE2 and E1 associate in the Golgi and are exported to the plasma membrane, where pE2 is cleaved into E2 and E3 [37]. Viral assembly is promoted by binding of the viral nucleocapsid to the viral RNA and the recruitment of the membrane-associated envelope glycoproteins. The assembled alphavirus particle, with an icosahedral core, buds at the cell membrane.
2.3.3. Assembly and Release of Virus

The structural polyprotein is translated from the 26S subgenomic RNA in host cytosol (Figure 2.4). The polyprotein is then processed into individual structural proteins by nascent viral proteins and host proteins in the endoplasmic reticulum (ER). Capsid, the first protein to mature among structural proteins, cleaves itself from the nascent polyprotein by autoproteolytic cleavage [37, 56]. In the cytoplasm, a cis-acting capsid recognizes a packaging signal contained within the genomic RNA to specifically encapsidate full-length genomic RNA to form nucleocapsids [57]. Failure to form proper nucleocapsid, such as encapsidation of incomplete or fragmented genomic viral RNA, ultimately results in the formation of defective interfering (DI) particles. Continued translation of the 26S mRNA produces a nascent polyprotein containing PE2 (E3-E2), E1 and 6K. This polyprotein contains signal sequences directing its insertion into the ER [20]. E3 and 6K are both short proteins that serve as signal peptides. 6K serves as a signal sequence for E1 and remains associated with E1 and E2 in the membrane [37].

The components of the polyprotein move together through the secretory pathway, from the ER to the Golgi and ultimately to the plasma membrane [37]. Oligosaccharides which are later trimmed in the Golgi apparatus are added to PE2 and E1 in the ER. The structural polyprotein are processed into the individual proteins PE2, 6K, and E1 in the ER by host cell enzymes, including signalases and furin-type proteases [37]. A host cell furin-type protease in the Golgi apparatus ultimately cleaves E3, releasing mature E2 [58]. Although the cleavage of E3 is not required for viral assembly, it is required for viral spike maturation and production of infectious progeny [59, 60]. The mature E1/E2 spikes then assemble on the host plasma membrane. Ultimately, the transmembrane glycoproteins interact with the nucleocapsids at the plasma membrane resulting in formation of envelop around the nucleocapsid that buds from the plasma membrane [37, 61]. Capsid-capsid interactions are stabilized by interaction with the glycoproteins, and assembly of nucleocapsid may occur in synchronization with viral budding [62, 63].
2.4. CHIKV Proteins

2.4.1. Nonstructural proteins

Four nonstructural proteins (nsPs) correspond to the four domains resulting from the maturation of nonstructural polyprotein. The nonstructural proteins are critical components of the viral replicase complex. Several functions have been ascribed to these four proteins.

2.4.1.1. Nonstructural protein 1 (nsP1)

NsP1 protein of the alphavirus family has been found to be responsible for capping of viral RNA because of its guanine-7-methyltransferase (MTase) and guanyltransferase (GTase) activities [64]. It is also required for the initiation and/or continuation of minus-strand RNA synthesis [54, 65]. NsP1 is responsible for binding of replicase complex to the membrane through an amphipathic helix located in the middle of the nsP1 sequence. The ability of nsP1 to associate with membrane appears to be a result of addition of palmitates to the protein [66, 67].

2.4.1.2. Nonstructural protein 2 (nsP2)

NsP2 has been proven to be the sole protease responsible for the complete processing of the nonstructural polyprotein [68]. The N-terminal region of nsP2 protein has multiple enzymatic activities [69] including the nucleoside triphosphatase (NTPase) activity, RNA triphosphatase activity and RNA helicase activity [70]. RNA triphosphatase activity is involved in removal of 5’ gamma phosphate of nascent viral RNA to enable it for capping. The C-terminal region of nsP2 is a cysteine protease responsible for viral polyprotein processing [71]. NsP2 has been reported to suppress the interferon response of the host cell, by specific inhibition of JAK/STAT signaling pathway [65]. It has also been suggested that interaction of nsP2 and host factors may be important for cessation of minus strand RNA synthesis.

2.4.1.3. Nonstructural protein 3 (nsP3)

NsP3, the most poorly understood, phosphoprotein [72] comprising of three domains namely, macrodomain, alphavirus unique domain (AUD) and hypervariable domain (HVR). The N-terminal macrodomain is conserved in all life forms and shows ADP-ribose 1’-phosphate phosphatase and RNA binding activities. AUD as the name suggested is conserved in alphaviruses. The C-terminal HVR domain varies from species to species and contains multiple
phosphorylation sites that may be responsible for efficient synthesis of viral RNA. Mutational analysis of nsP3 has been shown to affect negative strand RNA and subgenomic RNA synthesis indicating its role in viral RNA replication [73]. The 2/3 site cleavage has been shown to be the most important factor for switching off the negative strand RNA synthesis [74]. Thus, nsP3 might function in RNA replication together with nsP2 as a structural component responsible for the conformational changes in the replicase, rather than have a direct enzymatic function on the RNA.

2.4.1.4. **Nonstructural protein 4 (nsP4):**

NsP4 protein has been identified as RNA dependent RNA polymerase (RdRp) because of the presence of a GDD motif found in many viral RNAs [75]. The N-terminal region of nsP4 probably interacts with other viral proteins through an unconserved region of its N-terminus [76, 77]. NsP4 has also been shown to possess terminal adenyltransferase activity required for the maintenance of the viral genomic poly(A) tail [78]. Synthesized only when there is read-through of the opal codon, the concentration of nsP4 is well-regulated and it is rapidly degraded because of the presence of conserved destabilizing tyrosine residue at its N-terminus [79, 80].

2.4.2. **Structural proteins**

Five structural proteins are formed by the post-translational modification of the structural polyprotein (capsid-E3-E2-6K-E1) of CHIKV encoded by the subgenomic RNA. E2, E1 and capsid are the major structural proteins that are part of the mature virion and are critical for receptor identification, fusion of viral and host membrane at the time of infection and for encapsulation of viral genome at the time of viral maturation, respectively. E3 and 6K are small peptides involved in the maturation of viral spike and viral assembly and budding.

2.4.2.1. **Capsid**

Capsid protein is the first structural protein formed during the translation of the 26S RNA. Owing to the autocatalytic cleavage property of the capsid protein, it cleaves itself from the amino terminal of the polyprotein (capsid-p62-6K-E1) revealing the signal peptide at the N-terminal region of E3 (then p62), which guides the remaining polyprotein (p62-6K-E1) to the endoplasmic reticulum (ER) membrane [81, 82]. It consists of two domains, an N-terminal RNA
binding domain rich in positively charged amino acids arginine, lysine and proline linked to a C-terminal serine protease domain by a short inter domain peptide. Capsid interacts with the viral genomic RNA through a region of 32 amino acids (amino acids 76 to 107) called the RNA binding domain. This interaction triggers the assembly of nucleocapsids in the host cell cytoplasm [83]. It is interesting that how capsid selects the viral genomic RNA from a pool of cytosolic RNAs in the host. The viral genomic RNA contains a unique encapsidation signal which helps in its recognition, selection and packaging into the nucleocapsids [57]. Within the RNA binding domain of the capsid protein lies an uncharged sequence of 18 amino acid residues which forms the leucine zipper [84]. The leucine zipper has been suggested to mediate the capsid protein dimerization during virus assembly [85]. As soon as one capsid protein binds to the viral genomic RNA, it initiates the binding of other capsid proteins by capsid-capsid interactions and non-specific or charge related capsid-RNA interactions which eventually results in the nucleocapsid assembly [20, 83]. The RNA binding domains form a base on which the protease domain are seen as protrusions forming pentamers and hexamers around 5- and 2-fold icosahedral symmetry axes, respectively [86]. The C-terminal protease domain is folded into two β barrel subdomains with the substrate binding site in between and a hydrophobic pocket close to it for glycoprotein (E2) binding [86]. It is believed that the nucleocapsids diffuse freely to the plasma membrane where the viral envelope glycoproteins have accumulated. The binding of the nucleocapsids to the glycoproteins is proposed to cause a bending of the glycoprotein-containing membrane, eventually leading to the budding of membrane enclosed virus particles [20].

2.4.2.2. **Envelope protein 3 (E3)**

E3 is a small 64-amino-acid, cysteine rich glycoprotein that is released as a result of the cleavage of the precursor protein p62 late in the spike maturation process. It forms a part of the virus spike (non-covalently attached) in some alphaviruses like SFV but is released into the medium of virus infected cells like in the case of other alphaviruses such as SINV and CHIKV [87, 88]. It is therefore not associated with the mature virions in the latter category of viruses. E3 is considered to mediate proper folding of p62 and its subsequent association with E1 [89]. Moreover, cleavage of E3 from the assembled virus spike is critical to make the virus competent for its fusion activity [90, 91].
The amino terminal end of E3 contains a signal sequence that is responsible for the cotranslational translocation of the entire polyprotein containing E3-E2-6K-E1 into the lumen of the ER [92]. In the rough endoplasmic reticulum (RER), the polyprotein is glycosylated by the covalent attachment of the carbohydrate side chains and the protein is further cleaved by signalases into p62 (precursor E3-E2), 6K, and E1 [82]. Early in translation, glycosylation of E3 at the 14th amino acid residue causes the release of E3 from the ER membrane into the lumen [92]. This protects the signal sequence from being cleaved from the E3 protein. The proper folding of the proteins into their native conformation takes place in the RER. Moreover, experiments have proved that E3 is required for the efficient assembly of the virus particles by mediating the proper folding and association of the virus spike proteins [93]. Furthermore, heterodimerization of E1 with pE2 is necessary for it to be transported to the cell surface [93, 94]. Just before the assembled spike appears on the plasma membrane, host furin cleaves and release E3 from the precursor protein p62 in the trans-Golgi network [90, 91]. It has been shown that mutated furin cleavage sites show correct assembly of the virus particles but their infectivity is severely reduced. This occurs presumably because the fusion protein is unable to dissociate from pE2 and initiate fusion [90, 91]. It has been hypothesized that apart from the structural, E3 has an enzymatic and functional role in viral assembly and it is proposed that it acts like a viral protein disulfide isomerase which catalyzes the proper folding and disulfide bond formation in p62/E2 [95]. Thus although E3 may not be required for viral entry into the host it can be considered as a key player in spike assembly and the translocation of spikes to the site of budding.

2.4.2.3. 6K protein

6K is the smallest protein in alphaviruses with a length of about 58-61 amino acids. It is a small hydrophobic peptide acylated with fatty acids and forms a constitutive membrane protein [96, 97]. Based on the prediction of transmembrane regions it was suggested that all mature alphavirus 6K proteins cross the membrane only once [98]. Although the function of 6K in the alphavirus life cycle is not fully understood, it appears to be involved in the processing of virus glycoproteins, membrane permeabilization and virus budding. Because of its structural similarity with other small viral proteins, it was proposed that 6K is a virally encoded ion channel (viroporin) [99]. Since efficient virus maturation requires a high intracellular concentration of
Na\(^+\) ions [100] the role of 6K in ion transport may facilitate the efficient release of progeny virus particles. This explains features of alphavirus infection such as increased permeability of cells to monovalent cations followed by virion budding.

Since 6K protein is not required during early infection, its ion channel activity is unlikely to play a role during membrane fusion [82, 101]. This activity is proposed to play a role in the ER, Golgi, trans Golgi network or plasma membrane. The signal sequence present in 6K protein of SFV has been shown to play a role in the insertion of E1 protein into the lumen of ER [82]. The interactions between E1, E2 and 6K which are required for budding have been investigated in a study using chimeric viruses of SINV and RRV [102]. It has been speculated that 6K is critical for the proper folding of E1 in the heterodimer.

The role of the 6K protein in alphavirus growth is still partially understood. Virus mutants defective in 6K function do replicate, express, and localize their proteins in the infected cells as wild type viruses but have defects in virus release and shut-off of host translation [96, 103]. Since the shutoff of host translation in alphavirus-infected cells is mediated by ionic imbalance in the cytoplasm and the ionic composition of cytoplasm influences efficient budding of viruses, the role of 6K is well supported by these findings. 6K protein has even been pointed as the gene product responsible for the shutoff of host translation through genetic evidence [103].

2.4.2.4. Envelope protein 2 (E2)

E2 is a single pass type I membrane protein which is responsible for viral attachment to target host cell by binding to cell-surface receptors. It is synthesized as a precursor molecule p62 which is processed at the last step, just before virus budding. E2 has an amino terminal end protruding outwards from the viral membrane and a carboxy-terminal membrane spanning anchor. This carboxy-terminus transmembrane is transitory in nature as it is disrupted by palmitoylation, resulting in a re-orientation of the carboxy-terminal tail from the lumen to the cytoplasm [20, 104]. E2 is thus acylated with 3-6 palmitic acid chains resulting in a 35 amino acids cytoplasmic tail in CHIKV. E2 is also glycosylated with two to three oligosaccharide side chains which varies among alphaviruses. Glycosylation takes place in the RER and is thought to play a critical role in increasing the solubility of the protein and preventing its aggregation [105].
After folding into its native conformation in the RER, E2 in the form of p62 interacts with E1 to form a heterodimer, which is then transported to the Golgi-complex [106]. This heterodimerization is highly efficient and is cis directed, which means that the heterodimerization primarily takes place between p62 and E1 that originate from the same translation product (i.e. from the same polyprotein). Since p62 protein is made before E1 during 26S mRNA translation, it is suggested that p62 is retained at its translocation site until E1 has also been synthesized and translocated to the same site [106]. From the Golgi-complex the heterodimer is transported to the plasma membrane via the trans Golgi network. The heterodimers oligomerize into trimers before appearing on the plasma membrane and form the viral spike [107]. The site of oligomerization for E1/E2 heterodimers is still elusive, however, p62 is cleaved into E2 and E3 by a furin like host protease just before its arrival at the plasma membrane [108]. Since p62-E1 heterodimers are more stable and resistant to low pH than E1/E2 heterodimers, uncleaved p62 presumably functions as a molecular chaperone for protection of the viral spike from premature destabilization by the slightly acidic lumen of the trans Golgi network [109, 110].

The release of the E2 carboxy-terminal tail into the cytoplasm is critical since it is involved in budding of the virus particles by interacting with the capsid protein. This step occurs late during the export of proteins to the Golgi-complex and thereby prevents the premature assembly of the virus particles in the endoplasmic reticulum membrane. The amino-acid residues involved in E2-Capsid interactions have been identified by using a structural model of a portion of the 31 amino acid residues of the SFV cytoplasmic tail and finding a capsid binding site in the E2 protein [63]. Though interaction of E2-capsid drives virus budding, it has been shown that E1/E2 interactions is required for the interaction of E2 and capsid [111] Further, recent research on SFV shows that lateral spike-spike interactions are critical for budding of alphaviruses [112]. Thus, it has now been proposed that spike-spike interactions are responsible for virus budding and envelope formation and the role of E2-Capsid interaction is to trigger the spikes to interact laterally [113]. However, it is unclear whether E1/E2 heterodimers form multimeric complexes before binding to nucleocapsid or the heterodimers bind to the nucleocapsid individually during budding.
E2 glycoprotein has also been shown to be responsible for binding to host cell receptors and mutations in E2 alter the virus binding efficiency to cells [114]. Since alphaviruses like SFV and SINV have a wide range of hosts ranging from insects to mammals, they have the ability to infect a wide variety of cell types. This also suggests that different surface molecules may be used by these viruses as receptors for attachment. Major histocompatibility antigens and high-affinity lamin receptors have been recognized as receptors for SFV and SINV, respectively [115]. The high affinity lamin receptor is highly conserved in mammalian cell lines and also has a homolog in mosquito cells. Antibodies directed against the lamin receptor partially inhibited binding of SINV to these cells [115]. Rapid mutations in E2 have been shown to allow alphaviruses to use different kinds of receptors when under selective pressure. For example, cell-surface heparin-sulfate is used by alphaviruses for attachment as a consequence of cell culture adaptation [116]. Three mutation sites in E2 were identified as a consequence of adaptation to growth in BHK cells, namely, E2:1, E2:70 and E2:114 [117]. Other receptors like DC-SIGN, (also known as CD209), L-SIGN (also known as CLEC4M), a 63 kDa protein in chicken cells and 74 kDa and 110 kDa proteins in mouse neuroblastoma cells have been shown as receptors for Sindbis but their precise roles have not been firmly established [116]. It is generally concluded that E2 protein in alphaviruses can bind to multiple receptors on host cells. After receptor binding the virus is internalized by receptor mediated endocytosis.

2.4.2.5. Envelope protein 1 (E1)

E1 glycoprotein is a class II viral fusion protein with its amino-terminal end facing outwards from the membrane and a membrane spanning anchor located near the carboxy-terminal end. It also consists of a short cytoplasmic tail of about 2-3 amino acids [118] and is classified as a single-pass type I membrane protein. Post translational modifications like palmitoylation in or near the transmembrane domain and glycosylation with one to two oligosaccharide side chains have been reported for E1 protein of alphaviruses [20]. A hydrophobic domain located from residues 75 to 97 in E1 is conserved among alphaviruses and acts as the putative fusion peptide domain of the virus [120]. This domain inserts into the host cell membrane and thus plays an important role in the membrane fusion process. E1 is a component of the viral spike which is composed of a trimer of three hetero-trimers (E1/E2/E3)3 or hetero-dimers (E1/E2) [20]. A single virus particle consists of 80 spikes and thus 240 copies
of E1 envelope protein [20]. Crystallographic and cryo-EM studies have shown that E1 glycoprotein of SFV resembles that of flavivirus E protein [118]. Moreover SFV E1 has been structurally characterized into three domains – an amino-terminal β barrel domain located in the center of E1 flanked by a finger-like projecting domain containing the putative fusion peptide and a carboxy terminal Ig-like domain [118]. Before infection E1 has been localized to lie almost parallel to the virus surface in the mature virion [119]. The fusion activity of E1 is inactive as long as E1 is bound to E2 in the mature virion. After virus attachment to target cell and endocytosis, acidification of the endosome triggers the dissociation of the E1/E2 heterodimers and subsequent trimerization of E1. A low-pH trigger for membrane fusion has been demonstrated using both in vitro (virus-liposome systems) and in vivo (cellular assays) systems [109, 121]. Low pH triggers conformational changes in the viral spike protein which eventually causes its fusion with the host endosomal membrane. A fusion activation threshold of pH 6.2 and optimal fusion at pH 5.5 has been shown using SFV [121]. This pH dependence suggests that the alphaviruses fuse from within early endosomes, since the lumen of early endosomes is mildly acidic (having a pH between 6.2 and 5.3). The first spike rearrangement that occurs after the low-pH trigger is the dissociation of the E1/E2 heterodimer. E1 regions masked at neutral pH are exposed and become accessible to monoclonal antibodies (mAbs) [122]. One such acid-specific epitope is the fusion peptide at the tip of the finger-like domain II. E1 moves from its parallel position to the viral membrane to a more upright position, allowing the fusion protein to bind to the opposing target membrane via its fusion peptide. After the exposure of such acid-specific epitopes, E1 monomers rearrange into a homotrimeric configuration which is highly stable [123]. This E1 homotrimer has been shown to be involved in membrane fusion for SFV. A group of five trimers forming a “volcano-like” structure has been proposed to bring the two opposed membranes together for fusion [124].

Very recently, a cryo-EM and crystallization study of the low pH and membrane-inserted conformation of a homotrimer of E1 ectodomains revealed that, compared to the neutral form of the E1 ectodomain, the Ig-like domain (domain III) moves 37 Å toward the fusion loop and interacts with the finger-like domain (domain II), which redirects the polypeptide chain so that its C-terminus points toward the fusion loop at the tip of domain II [125]. At neutral pH, domains II and III lie in a region of extended conformation. The homotrimer is essentially formed by the
central interactions between the bottom β-sheets (domain I) of three E1 subunits and is continued in domain II interacting with the proximal half of domain I of an adjacent E1 subunit. The tips of domain II, the fusion loops, are involved in inter-trimer contacts as they are not exposed [125].

2.5. CHIKV INFECTION AND IMMUNOBIOLOGY

Upon vertebrate transmission via the bite of an infected mosquito, an asymptomatic period of 4 to 7 days is usually observed before the onset of clinical symptoms [126]. The first replicative cycles of alphaviruses in vertebrate hosts are thought to occur in dermal tissue at the site of the mosquito bite. Alphaviruses such as SINV and VEEV have been shown to infect macrophages and other antigen-presenting cells and promote their migration to the lymph nodes from where the virions are disseminated into the bloodstream [127, 128]. High viremias are typical in acute alphavirus infections (10^9 to 10^10 copies of the viral genome per ml of patient sera); the main viral pool being generated in the liver and spleen.

While CHIKV and other alphaviruses have been experimentally shown to infect a variety of mammalian cell types, the animal models for CHIKV- and RRV-induced arthritogenic diseases have provided evidence that fibroblasts, epithelial cells and to lesser extent, macrophages are the main cell types infected in vivo [129-131]. Most alphaviruses are strong inducers of type I interferons (IFN) and lead to a proinflammatory response including the induction of IL-1α, IL-6 and TNF-α [132]. The fever observed in the vast majority of infected individuals is a result of the high levels of proinflammatory cytokines in the bloodstream. The fever is typically high (more than 39 °C) and is often accompanied by maculopapular rash, other skin disorders and gastrointestinal symptoms such as diarrhea and vomiting [133, 134]. In most patients, the acute viremia is cleared from the bloodstream within 10 days [135]. Due to the short time scale of acute infection, the adaptive immune response is not raised against CHIKV at this stage [136]. Anti-CHIKV IgM antibodies are detectable in the blood of most individuals suffering from acute viremia whereas the IgG response is typically seen only after the viremia is cleared. Cross-reactivity of IgG antibodies is usually observed between the viruses of the genus, which may limit the specificity of serological tests as diagnostic tools. The alphaviral species are divided into seven antigenic serocomplexes based on this feature, SFV, CHIKV, RRV and ONNV falling into a same complex [18].
Old World alphavirus infections, particularly CHIKV and RRV, are distinguished from other endemic arboviral diseases by the high incidence of myopathy and polyarthritis [132, 137]. Joint disorders are typically encountered in 10% to 30% of CHIKV-infected individuals. These disorders usually manifest as severe and incapacitating arthralgia that can last up to 1 to 2 years after the acute illness. The arthritic symptoms resemble rheumatoid arthritis, as they are most intense in the joints of extremities, such as fingers, toes, wrists and ankles. The relative importance of the proinflammatory response and direct virus attack in the etiology of connective tissue symptoms are not clear. Alphaviral infection is transient and chronic forms in these or other tissues are not known to exist. Recent studies on CHIKV mouse models have also demonstrated that depletion of the macrophage-derived proinflammatory factors significantly diminishes pathological changes in the skeletal muscle and joints of infected animals [138]. However, satellite cells and fibroblast cell lines have been shown to be infected by CHIKV in vitro [48, 139]. Furthermore, positive immuno-staining for CHIKV antigens has been seen in satellite cells and muscle and synovial fibroblasts in CHIKV-infected mice indicating that the symptomatic tissues are also the sites of virus replication [140]. Pathology induced by different alphaviral species may not be uniform in this respect as fever indicating the proinflammatory response is reported in only approximately half of Ross-River viral infections though most of the patients experience long-lasting and incapacitating arthritis and myopathy [141]. Although Old World alphavirus infections have been described as relatively benign, a significant number of neurological symptoms were observed in a recent epidemic [142]. The CNS symptoms were especially pronounced in small children and led also to deaths and persistent disabilities. In La Réunion, approximately 50% of infected pregnant women transmitted the virus to their newborns and more than half of the infected neonates developed CNS symptoms including seizures, encephalopathy and brain oedema [143, 144]. Permanent disabilities were detected in 10% to 20% of these cases. Similar CNS symptoms, autism and other behavioral disorders as well as peripheral neuropathy were observed in children of all ages and adults, yet at less severe forms. The ability of recent CHIKV isolates to infect neuronal cells is a matter of debate. Suggestions have been made in support of this view, but the mouse models for CHIKV have shown either no CNS staining at all or staining of only cells in the choroid plexus; the latter could be linked with the brain oedema observed in many pediatric CHIKV patients. Attempts to develop a vaccine
against CHIKV have thus far been unsuccessful due to a large number of adverse effects caused by the attenuated virus strains used for this purpose [136]. Medical treatment for alphavirus infection is currently limited to chemotherapy used for symptomatic relief as no chemical agents that suppress alphavirus replication are approved for clinical use. Most patients are treated with nonsalicylate anti-inflammatory drugs, analgesics and dermatological preparation to alleviate the skin lesions [145].

2.6. CHIKV AND ITS HOST

CHIKV infection is maintained in nature by cycling between its invertebrate host mosquito and vertebrate host humans. Indian Ocean epidemic has seen a shift of virus vector from *A. aegyptii* to *A. albopictus* as a result of substitution from A to V at 226 amino acid position in E1 gene. The viral infection results in the alteration of host proteome that can determine the fate of infected host cells which affect disease progression and its outcome. The non-structural proteins of alphaviruses besides their enzymatic activities have been well known to interact with numerous host factors [146-148]. Among the structural proteins of CHIKV, E1 and E2, because of their significant role in viral binding and entry, are key targets for identification of cellular interactors that may facilitate viral infection and pathogenesis.

2.7. CHIKUNGUNYA DIAGNOSIS

Chikungunya infection is diagnosed primarily by detection of the virus, viral RNA by RT-PCR or specific antibodies in the infected patient. The time of sample isolation from patient and volume of sample is the deciding factor for particular technique to be employed for detection. Since CHIKV disease symptoms are similar to dengue fever, blood test is the most reliable way to identify CHIKV infection. Common laboratory tests for CHIKV include serological diagnosis and reverse transcriptase-polymerase chain reaction (RT-PCR). During the early infection stage, normally for 5-10 days after onset of symptoms, as the viral RNA titer is very high, detection of viral RNA level and virus particles in the serum samples are considered for diagnosis. In later phases of infection, serological methods are generally used for the detection of antibody response as the levels of anti CHIKV IgG and IgM are high in serum while the viral titer starts decreasing.
In case the samples are antibody negative, virus is isolated from the patients before appearance of antibodies, which occurs mostly after five days of appearance of symptoms [135]. In addition, a number of molecular assays have been designed for the rapid and sensitive detection of CHIKV infection. These assays are mainly PCR based methods including conventional RT-PCR [149]. Viral RNA can be easily detected by RT-PCR in serum specimens obtained from patients during the acute phase of infection. Chikungunya infections cause high levels of viremia (up to $1 \times 10^6$ plaque-forming units per ml), which typically last for 4–6 days after the onset of illness. Therefore RT-PCR can easily been done within the first 7 days on an acute-phase specimen to confirm CHIKV infection. Other PCR based techniques developed for detection of CHIKV RNA in patient serum include real time loop-mediated RT-PCR [150] and real time TaqMan RT-PCR targeting the envelope E1 gene [151] or non-structural nsP1 gene [152]. Recently, a one-step SYBR-green based real time assay that targets nsP2 gene of CHIKV has been developed for diagnosis [153].

The serological methods used for detection of CHIKV specific antibodies include Enzyme linked immuno assays (ELISA), indirect immunofluorescence (IFA), heamagglutination inhibition (HI) and micro neutralization (MNt). ELISA and IFA are very rapid and sensitive tests for detection of CHIKV infection which can distinguish between CHIKV specific IgM and IgG antibodies. However, the sensitivity of the assay for early detection of antibody response depends on the strain of virus used and the source of antigen.

2.8. TREATMENT AND PREVENTION

CHIKV has caused massive outbreaks during last decade but till date there is no specific treatment for the infection. Currently, the patients are being treated symptomatically with non-steroidal anti-inflammatory drugs (NSAIDs), fluids and medicines to relieve fever and pain like ibuprofen, naproxen and acetaminophen. Other treatment options include Chloroquine, Ribavirin and interferon but are not very effective. Besides these, passive immunization is an appropriate option for prevention and treatment for many viral infections [154]. Human polyclonal antibodies isolated from convalescent phase patient exhibit high neutralization activity in vitro and therapeutic efficacy against CHIKV infection in vivo in mouse model [155]. The polyclonal antibodies have broader reactivity but limited therapeutic value because of shorter viremia in
acute phase and thus only help in reducing viremia faster [136]. To overcome this, human monoclonal antibodies specific for CHIKV can be used. A recent study identified two monoclonal antibodies specific for glycoprotein E1 which strongly and specifically neutralize CHIKV infection in vitro [156].

Various vaccines are being developed for CHIKV till date but none have been licensed yet. The strategies employed for vaccine development involve whole inactivated virus preparation e.g. formaline or tween-ether inactivated whole virus [157, 158], attenuated live vaccines e.g. TSI-GSD-218 [159], recombinant proteins or virus like particles e.g. VLPs expressing structural proteins of CHIKV [160] and DNA vaccination e.g. intramuscular injection of plasmids encoding CHIKV-capsid, E1 and E2 [161, 162]. Since none of these vaccines have been licensed and not available in market, protection against mosquito bite and vector control are the important preventive measures.

2.9. RATIONALE OF THE STUDY

Chikungunya fever, a neglected tropical disease limited to old world countries has become a global threat since it reemergence in 2005. The re-emergent virus has acquired various molecular features that have facilitated the increased disease spread. Characterization of re-emergent CHIKV strain, availability of its proteins, knowledge of intraviral and viral host interactions, especially, in context of the structural components will enable identification of targets for disease inhibition and also enhance the understanding of mechanisms or pathways employed at the time of infection and viral dissemination.
2.10. OBJECTIVES

The aim of the current study is to characterize Indian Ocean isolate of CHIKV and its proteins and generate a viral-viral and viral-host interface of the structural components of the virus. To achieve this, thesis objectives have been defined as follows:

1. Cloning and characterization of CHIKV genes from Indian isolate of 2006 outbreak.

2. Expression and characterization of CHIKV Structural Proteins

3. Interaction analysis of structural proteins of CHIKV.

4. Identification of cellular partners of envelope protein E1 and E2 of CHIKV.