CHAPTER 2

REVIEW OF LITERATURE
2.1 Chikungunya Epidemics

Chikungunya virus (CHIKV) is an arthritic Old World alphavirus, transmitted by Aedes sp. of mosquitoes, mainly Aedes aegypti and Aedes albopictus. CHIKV was first isolated in 1952 from the Makonde plateau in Tanzania [7] and gets its name from Kamakonde vernacular language which means “that contorts or bends up” – the classical symptom observed in patients suffering from chikungunya disease. Alike other arboviruses, CHIKV outbreaks are mainly reported during monsoon season as the vector density increases with heavy rains. The first significant outbreak was documented in 1958 in Bangkok, Thailand [14] and then in 1963 in Kolkata, India [15, 16]. Following these, CHIKV was responsible for several outbreaks in different parts of India till 1973 [15, 16]. Since then only a few minor outbreaks of CHIKV were reported from Asian and African countries for almost three decades [17, 18]. The long absence of CHIKV epidemic in the endemic Indian Ocean region put forth an impression that CHIKV had disappeared from these regions [19, 20] and hence chikungunya disease was categorized as an uncommon tropical infection until 2004, when the virus re-emerged in Kenya with massive epidemic outbreaks affecting 75% of the population [21, 22]. Since then, there has been an alarming increase in the incidences of CHIKV infection with case reports as high as five million from India, South East Asia and Africa [23-25].

The CHIKV outbreak of 2005 in La Reunion indicated that the virus was no longer confined to developing countries. Furthermore, it was during this epidemic when neurological complications and mother to fetal transmission with significant mortality was reported for the first time to be associated with CHIKV infection [26, 27]. In 2007, the virus gained global attention when it established an autochthonous transmission causing a significant outbreak in Italy infecting more than 250 people [28, 29]. This was the first reported incident of local transmission following an imported case and highlighting the potential of this virus to invade Europe, North America and Australia. The drastic changes in global travel and vector distribution might be responsible for increased number of imported cases of infection and hence the risk of epidemics.

Thus far CHIKV infection has been reported from nearly 40 countries throughout the world. In 2008, the US National Institute of Allergy and Infectious Diseases (NIAID) categorized CHIKV as category C priority pathogen [18, 30]. This category include those
pathogens which either could be used as a source of mass dissemination in future or causes major public health impacts due to high morbidity and mortality. The mortality rate of CHIKV during Reunion and Mauritius outbreaks has been estimated to be 3:10 to 1:1000 and 47:1000, respectively, mainly in neonates, adults with underlying conditions such as hypertension, diabetes or heart diseases and elderly patients [23, 31, 32].

2.2 Taxonomy

CHIKV belongs to the genus Alphavirus (family Togaviridae) which contains nearly thirty recognised viruses including a number of medically significant pathogens such as Western, Eastern and Venezuelan Equine Encephalitis viruses (WEEV, EEEV and VEEV, respectively), Chikungunya virus (CHIKV), Sindbis virus (SV) and Semliki Forest virus (SFV), with wide geographical distribution [Fig. 2.1; 33].

Based on their very wide geographical distribution the alphaviruses are classified into two groups; the Old World alphaviruses and the New World alphaviruses [6, 34]. The Old World alphaviruses mainly include Semliki Forest virus (SFV), Sindbis virus (SINV), Chikungunya virus (CHIKV), O’nyong-nyong virus (ONNV) and Ross River virus (RRV) which are distributed throughout the Old World countries of Europe, Asia, Australia and Africa while the New World alphaviruses mainly include Venezuelan Equine Encephalitis virus (VEEV), Eastern Equine Encephalitis virus (EEEV) and Western Equine Encephalitis virus (WEEV) distributed throughout the New World countries of North and South America [6]. The alphaviruses which have been sequenced are categorised into three large groups on the basis of their serological cross reactivity: the VEE/EEE group, the SF group and the SIN group. The VEE/EEE group is exclusively New World in its distribution while the SF and SIN groups are primarily distributed in the Old World countries [6]. WEE is a New World alphavirus, but due to its serological cross reactions it is included in the SIN group (Table 2.1). Along with differences in their geographic locations these groups also differ in pathogenesis, tissue and cellular tropism, cytotoxicity and interference with virus-induced immune responses.

The place of origin, the rate of divergence and the mode of distribution of alphaviruses has been debated over the years. Levinson and co-workers in 1990 proposed that the
alphaviruses originated in the New World and were spread to the Old World twice, once to form the SIN group and then to form SF group [35]. Later in 1993, Weaver and colleagues hypothesised that the alphaviruses originated either in the New World or the Old World over 2,000 to 3,000 years ago [36]. It has been postulated in these reports that migratory birds and humans may have contributed to the spread of the viruses. The Old World alphaviruses cause diseases characterised mainly by fever, rash and arthralgia evolving into arthritis in humans while the New World alphaviruses cause encephalitis in humans and horses [34].

Figure 2.1: Classification of Chikungunya virus
Table 2.1: Classification of alphaviruses based on their serological cross reactivity and geographical distribution.

<table>
<thead>
<tr>
<th>Serological group</th>
<th>Virus</th>
<th>Distribution</th>
</tr>
</thead>
</table>
| VEE/EEE group     | Venezuelan Equine Encephalitis virus (VEEV)  
Eastern Equine Encephalitis virus (EEEV) | New World |
| SIN group         | Western Equine Encephalitis virus (WEEV)  
Sindbis virus (SINV) |     |
| SF group          | Semliki Forest virus (SFV)  
Chikungunya virus (CHIKV)  
O'nyong-nyong virus (ONNV)  
Ross river virus (RRV) | Old World |

2.3 Transmission Cycle

CHIKV is transmitted by bite of infected mosquito. CHIKV is endemic to both Asia and Africa but the transmission cycle vary considerably in these continents (Fig. 2.2). In Africa, CHIKV is maintained via zoonotic/sylvatic cycle relying on nonhuman primates and *Ae. furcifer* and *Ae. africanus* mosquitoes [37]. The African profile is characterized by occasional small epidemics in rural areas with little impact on public health. On the contrary, in Asia primary host of CHIKV are humans and primary vector is *Ae. aegypti* in most of the urban epidemics [18]. The expansion in vector range to include *Ae. albopictus* occurred because of the A226V mutation in E1 protein of CHIKV which increased the fitness of CHIKV for *Ae. albopictus* and improved transmissibility to vertebrate host through this species [38, 39]. The mutated strain colonized by *Ae. albopictus* is highly efficient in causing epidemics in tropical and temperate countries. Epidemics in rural Africa occurs at much smaller scale in comparison to Asian outbreaks where it infected thousands to millions of people because of the lower human population densities and more stable herd immunity. The ecological differences between zoonotic/sylvatic and urban cycles are due to the seasonality of vector larval habitat, vertebrate host abundance, herd immunity and vector host preferences [40].
As CHIKV is a zoonotic arbovirus, it is maintained in nature both by sylvatic (particularly in Africa) and urban cycle. In the sylvatic cycle, the virus is transmitted by forest dwelling mosquito species to non-human primates while the urban cycle involves transmission of the virus between urban mosquito species and humans. Non-human primates, birds and rodents might serve as CHIKV reservoir for its maintenance outside the epidemic periods.

### 2.4 Dissemination

The viral infection starts with the bite of infected mosquito. Following transmission, virus directly enters into the subcutaneous capillaries with some of the viral particles infecting susceptible skin cells. The local replication takes place in skin cells such as macrophages or fibroblasts and endothelial cells for a limited period after which the virion produced spreads rapidly to secondary lymphoid organs close to the inoculation site (Fig. 2.3). The virus then
enters into the lymphatic circulation after release from draining lymph nodes and then into the blood circulation through thoracic duct. The type of cell and tissues infected by CHIKV might affect the disease progression and clinical manifestations. Studies have reported that CHIKV can infect a number of cell types involving fibroblasts, epithelial and endothelial cells [41]. The primary sites of infection are bone and joint associated connective tissues and skeletal muscles due to the presence of specific receptors and/or hyposensitivity to type I interferon response [41, 42]. The virus is disseminated to other target organs like liver, muscle, joints and remote lymphoid organs by blood either as free virions or in the form of infected monocytes. These cells along with other immune cells like lymphocytes, natural killer cell and dendritic cells have been shown to be non-permissive to CHIKV infection [43]. However, recent studies have shown that monocytes are early targets of CHIKV [44]. In liver, spleen and synovial tissues macrophages have been reported to be infected by CHIKV [45, 46]. Both monocytes and macrophages have low levels of virus producing non-infectious virions [43, 44, 47] but these are attractive targets for CHIKV since these cells remain in circulation and help in dissemination of virus to other susceptible tissues by infiltration of these cells. The pathological events associated with the infection of liver (hepaticyte apoptosis) and lymphoid organs (adenopathy) are mostly subclinical, while in muscles and joints infiltration of mononuclear cell and viral replication cause severe pain, resulting in arthritis.
Figure 2.3: Dissemination of Chikungunya in vertebrates

Following mosquito bite, CHIKV replicates in skin cells and monocytes. Then the virus disseminates to the liver, muscle, joints, lymphoid tissue (lymph nodes and spleen) and brain. The characteristic features of infection in different tissues are also indicated.

2.5 Clinical manifestations

CHIKV usually causes a nonlethal, self-limiting and febrile illness in humans. The incubation period of CHIKV generally ranges from 2–4 days but in some cases it can reach 12 days [7, 48]. CHIKV infection mainly persists in two phases: acute phase and chronic phase. The acute phase is characterized by classical triad of symptoms involving high fever, polyarthralgia and rash. Other symptoms, although rarely observed, include myalgia, headache, retro-orbital pain, pharyngitis, lymphadenopathy and dysgeusia [26, 49]. During acute phase the viremic level can reach up to $10^8$ viral particles per ml of blood and the plasma concentration of type I interferons can range from 0.5-2 ng/ml which further induces other cytokines and chemokines.
The persistent chronic phase is characterized by polyarthralgia that can last from weeks to several years beyond the acute phase [53]. Silent infection/asymptomatic infections (infections without illness) also occur in some patients but very rarely. Serosurveys indicate that 3-25% patients with CHIKV antibodies have asymptomatic infection [54-56].

Generally the associated symptoms resolve in 7-10 days except for joint pain and stiffness [57-58]. The hallmark of CHIKV infection is erratic, relapsing and incapacitating arthralgia which may persist for several months. CHIKV mainly attacks fibroblasts, which explains the involvement of muscles, joints and skin connective tissue. The pain associated with the disease is due to the high number of nerve endings present within these tissues. Generally the joint pain and inflammation associated with CHIKV infection affect small joints involving fingers, wrist and tarsus but may also affect large joints like knees and shoulders [55]. Simultaneous involvement of several joints for a prolonged period can lead to polyarthritis or polyarthralgia [59]. Neurological complications involving meningo-encephalitis, peripheral neuropathy, myelopathy, myeloneuropathy and myopathy were observed during 2006 outbreaks in Indian Ocean region [60]. Some rare complications involving hemorrhage, myocarditis and hepatitis have also been reported after CHIKV infection [61]. CHIKV disease was not considered as life threatening till 2006 outbreak of La Reunion, where it caused a significant number of deaths among infected population [62]. During the recent outbreaks, cases of mother to child transmission with a vertical transmission rate of 49% have also been reported [27]. Intrapartum transmission resulted in neonatal complications such as neurologic disease, hemorrhage, and myocardial disease. Abortions during the first trimester following maternal CHIKV infection have been rarely reported [63]. Moreover, cases of multiple organ failure and eye infection (CHIKV neurorenitis) have also been reported [64-67].

2.6 Host Immune response

During CHIKV infection, robust cellular and molecular innate responses have been reported to be capable of controlling the infection [45]. Type I interferon (IFN) immune response play central role in combating viral infections at a very early phase by controlling viral replication and its dissemination towards target tissues [68-70]. IFN-α is detected on the first day
of CHIKV infection and its plasma concentration correlates with the viral load [26]. This response is achieved through rapid activation of NK cells, plasmacytoid dendritic cells (pDCs) and T cells [45]. CHIKV does not directly infect these cells, but is expected to directly activate hematopoietic cells, especially plasmacytoid DCs by ssRNA [43]. The acute IFN response is short lived and levels of IFN-α drops rapidly within 12-24 hrs following maximum viremia [70, 71]. The viral load during the acute phase is related to the concentration of IFN-α, IL-1-RA, IL-6, MCP-1/CCL-2, IL-12, IP-10/CXCL-10 [69], IL-18 and IL-18BP [52]. This robust inflammatory response during the early stage of infection contributes to the elimination of virus from the blood and clinical recovery. The studies on type I IFN deficient mouse models have shown the viral dissemination to central nervous system besides livers, muscle and joints [41]. The imperfect type I IFN response during infection is responsible for longer disease persistence leading to chronic phase variable levels of myalgia and arthralgia, resulting in debilitating arthritic syndrome. There exists a recovery phase demarcating the acute and chronic phases during which CCR-2 downstream signaling is inhibited to prevent monocytes-macrophage recruitment to the tissues. This intermediary stage plays a critical role in the establishment of chronic phase that follows.

During chronic phase, mainly monocytes, macrophages and T lymphocytes infiltrate into the synovium and target tissues which causes muscular and articular damage due to viral replication in target cells and indirect immune response activation [46]. CHIKV persistence leads to the production of cytokines, chemokines and pro-inflammatory markers e.g. IFN-α, IL-6, MCP-1/CCL-2, IL-8 and MMP2, that can be detected in the synovial fluids from the patients with chronic infection [46]. The infected macrophages are main contributors in disease establishment due to the maintenance of viral replication and synthesis of inflammatory mediators in infected tissues, leading to the chronic local inflammation and joint pains [46]. Moreover, the secretion of soluble factors from infected macrophages enhances the inflammatory response recruiting and activating target tissue lymphocytes and NK cells [72]. Massive infiltration of inflammatory cells leads to the apoptosis and necrosis of muscle fibroblast [73] and this leads to cartilage destruction due to the phagocytosis of apoptotic cells by macrophages [Fig. 2.4; 74, 75].
The inflamed joints attract monocytes, T cells, and natural killer (NK) where they become activated. The infected macrophages in joints cause local inflammation and produce cytokines, chemokines, and pro-inflammatory effectors, such as MCP-1/CCL-2, IL-8, IL-6, IFN-α, and MMP2. The phagocytosis of apoptotic bodies from infected cells contributes to viral persistence. This causes arthritis which is accompanied by high rates of fibroblast apoptosis and cartilage destruction. The damage caused to the muscles and the pain associated is probably due to the chronic inflammation.

2.7 Virion structure and Genome organization

As a typical alphavirus, CHIKV is an icosahedral enveloped virus of about 70 nm in diameter. The single-stranded RNA genome is encapsidated by capsid that is arranged in a T=4 icosahedral protein shell. The virion has a host cell acquired lipid membrane which envelops the icosahedral shell and embedded with 80 spikes of viral envelope proteins. Each spike is a trimer of E1 and E2 heterodimers i.e., the envelope proteins initially organise as heterodimers followed by their association into trimers (trimer of heterodimers) to form the spike protrusions on this host acquired membrane (Fig. 2.5).
Figure 2.5: Structure of Chikungunya virus

The genome of CHIKV is composed of a positive sense single strand RNA molecule which is organized as- 5′ cap-nsp1-nsp2-nsp3-nsp4-(junction region)-C-nspE2-6k-nspE1-poly(A)-3′ (Fig. 2.6). The genome is of 11,805 nucleotides, excluding the 5′ cap nucleotide, an I-poly(A) tract (an internal poly(A) tract in the 3′ NTR of highly passage stains; 76) and the 3′ poly(A) tail, which encodes two open reading frames (ORFs). The 5′ non-translatable region (NTR) is composed of 76 nucleotides and 3′ poly(A) tail is of 526 nucleotides. The 5′ ORF constituting two-thirds of the viral genome codes for non-structural polyprotein whereas 3′ ORF codes structural polyprotein. These polyproteins are post-translationally cleaved into mature proteins by viral and host proteases. The non-structural polyprotein is cleaved by viral protease into four non-structural proteins (nspS) which constitutes the replicase complex of the virus. The structural polyprotein is processed by both viral and host protease into five structural proteins, the capsid, envelope proteins E1 and E2, and two small cleavage products E3 and 6K, which are required for viral encapsidation and budding. The untranslated junction regions present between two ORFs, a
conserved sequence of 21 nucleotides, acts as a promoter site for transcription of subgenomic mRNA [77].

Figure 2.6: Schematic representation of Chikungunya genome

CHIKV genome has a 5’ cap and a 3’ poly(A) tail. The genomic RNA contains two open reading frames (ORFs). The N-terminus ORF constituting two-third of the genome encode for non-structural polyprotein while C-terminus ORF produces structural polyprotein. The non-structural polyprotein is processed into four non-structural proteins (nsP1-4) while structural produces five proteins (Capsid, E3, E2, 6K and E1).

2.8 Replication cycle

The molecular mechanism of CHIKV replication in host cell is a less explored area. The basic replication cycle is very much similar to other alphaviruses. The lifecycle of CHIKV in host cell starts with the identification of specific receptors on the host cell surface followed by endocytosis. After successful entry, viral genome is released into the cytoplasm by nucleocapsid disassembly for replication and translation. The newly synthesized RNA is encapsidated by
capsid proteins and the nucleocapsid cores get entrapped inside the lipid membrane of host embedded with viral spikes releasing the progeny virions (Fig. 2.5).

2.8.1 CHIKV entry and release of nucleocapsid into host cell

The cellular and tissue tropism of CHIKV is not known completely. As CHIKV is an enveloped virus it can infect multiple types of cell within its mammalian host involving epithelial, endothelial, fibroblasts and monocyte derived macrophages but not all the cell types indicating that it identifies receptors not ubiquitously present on all cell types [43]. Enveloped viruses utilize membrane associated receptors for identification and enter into specific target cells [78] such as DC-SIGN, laminin receptors and glycosaminoglycans which are used by SFV and SINV [6, 78-80]. But the receptors for CHIKV are still unknown.

Alike other members of genus Alphavirus, CHIKV starts its life cycle through pH dependent endocytosis in clathrin coated vesicles via receptor mediated interactions, but the exact mechanism is not known. The envelope proteins E1 and E2 are required for the internalization of virus into the host cell. The process of virion entry is initiated by the identification and interaction of E2 protein with cell surface receptors [78, 82]. The prohibitin (PHB) protein expressed on microglial cells has been identified as CHIKV binding protein [83]. This protein is formed by oligomerisation of two highly homologous proteins PHB1 and PHB2 [84]. Only PHB1 has been confirmed as CHIKV E2 interactor and play role in virus internalization. This is one way of viral entry into host cells yet other pathways have to be explored.

As the virus containing endosomes mature, the pH of endosomal vesicle becomes slightly acidic. The acidic environment of endosome causes a conformational change in the viral spikes, leading to the dissociation of E2-E1 heterodimers and formation of E1 homotrimerers. The trimerisation of E1 protein exposes the hydrophobic fusion peptide which refolds to form hairpin-like structure and inserts into the late endosomal membrane [85]. The insertion results in the fusion of viral and host membrane, yielding a fusion pore and releases the nucleocapsid in the cytoplasm [86]. This process of internalisation depends on the low pH of endosomal compartment and cholesterol profile of host cell membrane, which is also required for budding process during alphavirus infection [87, 88].
2.8.2 Intracellular replication, transcription and translation of viral genome

Till date no experimental work has been done on the intracellular processes of CHIKV replication, transcription and translation but the general information regarding these steps can be deduced from other alphaviruses. The intact form of nucleocapsid (NC) is released into host cytoplasm, which disassembles within 5 minutes [89]. After the release of nucleocapsid into host cell cytoplasm, the capsid protein binds to the large ribosomal subunit through a conserved domain in its C-terminal region [90]. This binding probably activates the disassembling process of the nucleocapsid and hence the release of viral RNA into the cytoplasm [91]. The replication of alphaviruses is a two-step pathway. The synthesis of both minus-strand and plus-strand RNA molecules is under the control of non-structural proteins which constitute the replicase complex of the virus. After the release of viral RNA into cytoplasm, the 5’ ORF is translated into non-structural polyprotein (P1234) using host translational machinery. This is cleaved by viral protease present within the C-terminal domain of nsP2 into P123 and mature nsP4. The polyprotein P123 forms a complex with free nsP4 along with some host proteins which act as minus-RNA strand replicase complex (also called early replicase complex; 92). The formation of minus-strand RNA takes place only during the early infection, later on when the concentration of P123 increases in the cytoplasm and can perform a bimolecular reaction; the precursor is further processed into mature nsPs. The mature nsPs in association with some host proteins act as a plus strand RNA replicase or late replicase complex, presumably in a conformation different from early replicase complex [6]. An intermediate complex composed of nsP1, P23 and nsP4 may exist transiently which is capable of both plus- and minus- strand RNA synthesis [93-95]. The late replicase complex utilize negative sense RNA for amplification of sub-genomic (26S) RNA and genomic (49S) RNA [96]. The 26S RNA is translated into structural polyprotein which is cleaved cotranslationally and posttranslationally.

2.8.3 Processing of structural polyprotein

The structural proteins are synthesized in the order of C-pE2-6k-E1 (pE2 is the precursor for E3 and E2; 97). The capsid protein present on the N- terminus of the polyprotein is cleaved by its serine protease activity that acts in cis to release it from the polyprotein. The new N-terminus of polyprotein contains a signal sequence for translocation and insertion of pE2-6k-E1 to the endoplasmic reticulum (ER) membrane [98]. This signal sequence also contains a
carbohydrate attachment site which is responsible for the retention of pE2-6k-E1 in ER. Following insertion into ER, this signal sequence plays no role in further translocation to plasma membrane. While in ER lumen the C-terminus of pE2 functions as a stop transfer signal and membrane anchor for E2. The anchor sequence is followed by a signal for insertion of 6k in ER membrane which is cleaved from the polyprotein by ER signalases during processing. The C-terminus of 6k protein contains a signal sequence for translocation of E1 and gets cleaved by signalase after translocating E1 to the ER membrane.

The translocated pE2 and E1 are processed and undergo post-translational modification to form pE2 –E1 complex. Precursor pE2 and E1 undergo a complex series of folding intermediates requiring chaperons and disulphide bond formation [99]. The glycosylation starts in the lumen of ER and the carbohydrate moieties are subsequently modified during transport through the Golgi network. These include both simple and complex carbohydrates moieties. The simple carbohydrates are processed by cellular enzymes into complex type.

During its transport through the trans-golgi network the precursor pE2 is cleaved by furin (host protease) to E3 and E2 before its arrival at the plasma membrane [100]. This cleavage is required for entry and fusion activation in the new cells [101, 102]. Uncleaved pE2 can also be efficiently transported and incorporated into budding virion but the progeny viruses are non-infectious [103].

2.8.4 Nucleocapsid assembly

The alphavirus nucleocapsid (NC) is composed of single copy of RNA genome and 240 copies of capsid protein (C). This assembly process is very specific and only viral genomic RNA is packaged because of the presence of packaging signal which is required for efficient encapsidation. The assembly of NC is a multistep process which starts with the interaction of C dimer with the newly synthesized RNA genome [104]. NC is formed from a nucleation event in which the amino acid region from 81 to 112 of C identifies the packaging signal and leads to the encapsidation by interacting between these regions [105-107]. The dimerisation of C takes place through coiled-coil interaction of helix I from individual molecules to stabilize the dimer [108, 109]. After initiation, additional molecules of C are recruited into NC which involves lateral interactions with proteins already partially assembled and electrostatic interaction between
highly charged N-terminal domain of CP and genomic RNA. The lateral interactions are specific while an electrostatic interaction is nonspecific and does not require any specific protein structure.

2.8.5 Virus Budding

The last step of alphavirus lifecycle is budding of progeny virions from the host cell plasma membrane and release into the extracellular space for transmission of infection to the neighbouring healthy cells. This process requires effective interaction between newly synthesized NCs and the glycoproteins present on cell surface [110, 111]. NCs assembled in the cytoplasm diffuse freely to the plasma membrane and interact with the viral spike through the cytoplasmic tail of E2 [112, 113]. This interaction triggers the maturation event of NC during which it starts contracting. Also, the interaction provides ample amount of free energy to propel the NC through the host cell plasma membrane resulting in the complete complementation of glycoprotein spike [114]. The efficient assembly and budding of virion particles requires interaction of transmembrane domains of E1 and E2 glycoproteins [115]. The assembly of virus particles also involves the lateral interactions between the glycoprotein molecules [116, 117] and also requires cholesterol in the membrane for efficient budding process [118].
Figure 2.7: Replication cycle of Chikungunya Virus

The virus enters into the host cell by receptor mediated endocytosis (A). The low pH of the endosome causes the fusion of viral and host membrane releasing the nucleocapsid into the host cell cytoplasm (C). In the early stages, the P1234 is translated from the viral genome (D) which is cleaved by nsP2 protease activity into P123 and free nsP4 which associate with host proteins to form the replication complex (RC). The RC proceeds to produce the full length minus strand. In the later stages, P123 is cleaved to yield mature non-structural proteins nsP1, nsP2, nsP3 and nsP4. These non-structural proteins along with host cell proteins act as the plus strand replicase (E), which produces the plus strand 49S RNA and 26S sub-genomic mRNA (F). The 26S sub-genomic RNA encodes the polyprotein precursor for structural proteins, which is cleaved to yield Capsid protein and pE2-6K-E1 (G). The structural proteins are further processed in the endoplasmic reticulum (I) and Golgi complex and are then transported to the plasma membrane (J). The viral RNA is packed into the nucleocapsid and assembles with lipid membrane embedded with spikes of envelope protein (K) releasing the mature virions from the plasma membrane.
2.9 Chikungunya proteins

The proteome of CHIKV is composed of nine proteins - four non-structural and five structural proteins.

2.9.1 Non-structural proteins

The non-structural polyprotein, its cleavage intermediates and fully cleaved mature nsPs play important role in viral replication and transcription. Each individual protein plays a specific role in this process and associate together to form a functional replicase complex.

2.9.1.1 Non-structural protein 1 (nsP1)

NsP1 is required for the capping and cap methylation of newly synthesized viral genomic and subgenomic RNAs as it possess both guanine-7-methyltransferase and guanyltransferase activities [119-121]. Alphaviruses possess a novel cytoplasmic RNA capping pathway which differs from the nuclear RNA capping of eukaryotes [122]. Studies based on phenotypic analysis of nsP1 mutants have demonstrated that the protein may be required for the specific initiation and maintenance of minus-strand replicative intermediates, along with its enzymatic activities [123].

NsP1 is the sole protein of the viral replicase complex which exhibits membrane affinity. The protein is localized to the cytoplasmic surface of the plasma membrane rich in phosphotidylserine [124, 125]. The membrane association is mediated by a binding peptide present in the latter half of the N-terminus conserved domain containing cystein palmitoylation as well as distinct patches of basic and hydrophobic residues [5, 122, 126]. This peptide also acts as a localization signal that translocates the replicase complex to the cytoplasmic vacuoles (CPVs) and remain tightly bound due to the interaction of nsP1 with membrane lipids. This attachment serves to concentrate the viral RNAs and the enzymes on membrane surface which enhance the efficiency of replication and protect double stranded viral replication intermediates degradation by antiviral system of the host cell [126].
2.9.1.2 Non-structural protein 2 (nsP2)

The nsP2 of alphaviruses possess multiple enzymatic activities and plays various roles in viral replication cycle. This protein has two main domains- N-terminal helicase domain and C-terminal protease domain. The helicase domain is required for unwinding of RNA duplex during replication and transcription [127]. This domain also contains RNA triphosphatase and nucleoside triphosphatase (NTPase) activities [122, 128]. RNA triphosphatase activity is required for removal of 5’ gamma phosphate from nascent viral RNA enabling the capping reaction [5]. NTPase activity coupled with nucleotide hydrolysis is essential for viral replication and in association with certain virus specific proteins is responsible for inhibition of host DNA replication [128]. A peptide sequence in the N-terminal of nsP2 behaves as nucleotide binding motif which is highly conserved among alphaviruses and mutation in this region render the protein deficient of its all three enzymatic activities [122, 129].

The C-terminus of nsP2 act as a cysteine protease similar to papain like proteases required for the conversion of non-structural polyprotein into intermediate polyproteins and mature nsPs [130, 131]. The nsP2 has been shown to be the sole protease responsible for complete processing of polyprotein [131]. The cleavage site present at the junctions of individual nsP is similar in amino acid sequence [130]. Both forms of nsP2 i.e. mature protein and in polyprotein form, are proteolytically active with different substrate preferences. The polyprotein P1234 is cleaved in cis only at the junction of nsp3/nsP4 to form mature nsP4 and P123 while cleavage at other sites occurs in trans [132].

Besides its role in viral replication cycle, nsP2 also affect key cellular processes such as transcription and translation through distinctly independent events [133-135]. These events do not affect the viral replication, instead downregulate the antiviral response and circumvent the host translational machinery for rapid production of viral proteins [133, 134]. NsP2 has also been proven to cause cytopathic effects in host cell during viral infection which depends on its presence in mature form [134, 135]. Till date the molecular functions of this protein are not known completely. There are several undefined functions of this multifunctional enzyme mediated through interactions with the host factors that affect the outcome of viral infection.
2.9.1.3 Non-structural protein 3 (nsP3)

NsP3 is composed of three distinct domains: N-terminal macrodomain conserved among prokaryotes and eukaryotes; central alphavirus unique domain (AUD) conserved among alphaviruses and a C-terminal hyper variable region [HVR; 136]. Till date, not much information is available about the functions of this protein. The macro domain possesses both ADP-ribose 1’-phosphate phosphatase and RNA binding activities [137]. Mutational studies have shown that both forms of the protein, mature as well as P123, is required for minus-strand and subgenomic RNA synthesis [96, 138, 139]. NsP3 is highly phosphorylated at multiple serine and threonine residues mainly located in the hypervariable region of the protein. The phosphorylation pattern is complex with a total of sixteen possible sites [5, 136]. The actual purpose of phosphorylation is not known but its absence in SFV potentially modulates the efficiency of minus strand RNA synthesis [136]. A fraction of nsP3 has been shown to localize to the nuclear envelope within infected mammalian cells which is distinct from nsP3 present in cytoplasmic replicase complex [140].

2.9.1.4 Non-structural protein 4 (nsP4)

The alphavirus nsP4 has been identified as RNA dependent RNA polymerase [RdRp; 5, 141-143]. The N-terminus contains a small region which is not conserved among other viral RdRps while C-terminus maintains the homology with other viral RdRps [138]. The N-terminus region may be involved in the interaction with other viral proteins, mainly nsP1 and several unidentified host proteins [144, 145]. NsP4 has also been shown to possess terminal adenyltransferase activity via addition of non-template adenine to the 3’ end of acceptor RNA [146]. This activity is required for the maintenance of the viral genomic poly(A) tail. The cellular concentration of nsP4 is low in comparison to other nsPs because of the employment of opal codon readthrough in some alphavirus species and secondly due to the presence of a conserved destabilizing tyrosine residue in the N-terminus region responsible for rapid degradation of the protein [147]. Although this tyrosine residue causes rapid degradation of nsP4, yet its replacement with nonaromatic residue results in poor RNA replication [148].
2.10 Chikungunya and its Host

CHIKV is maintained in nature by cycling between its invertebrate host mosquito and vertebrate host humans. 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 are known to interact with numerous host factors. Among these, non-structural protein 2 (nsP2) plays major role in inhibiting interferon response of host produced during CHIKV infection [149] and interact with the proteins of translation pathway and RNA splicing factors to enhance its own genome replication [150]. The identification of cellular interactors will help in understanding the mechanisms by which virus hijacks the cellular machinery and blocks the host antiviral response for its replication and production of progeny virions.

2.11 Role of nsP2 in host cell responses

The entry of virus into host cells is followed by replication of viral genome in infected cells. The efficiency of replication not only depends on the expression of viral proteins and replication machinery but also on the ability of virus to interfere with the antiviral response triggered in host cells after infection. Different viruses have evolved different mechanisms to inhibit host antiviral responses by interfering with cellular signaling pathways. The alphaviruses combat with cellular responses by inhibiting transcription of antiviral proteins without affecting viral replication and egress [135]. The nsP2 protein of Old World alphaviruses has been reported to play role in the transcriptional shutoff through different mechanisms [135, 151, 152]. The C-terminus of nsP2 contains a nuclear localization signal [153, 154] and mutation in this region affects the cytopathogenesis of Semliki Forest virus [SFV; 155]. The accumulation of nsP2 within the nucleus inhibit host cellular mRNA transcription by causing degradation of DNA directed RNA polymerase II (RPBII) subunit [156].

Another mechanism employed by alphavirus nsP2 to inhibit host cell responses involve antagonism of interferon (IFN) response in two ways- beta interferon (IFN-β) transcription via global host shutoff and downstream type I/II IFN induced janus kinase/signal transducer and
activators of transcription (JAK-STAT) signaling [149, 152, 157, 158]. The activation of JAK-STAT signaling results in translocation of phosphorylated STAT1/2 to the nucleus which further activates the translation of antiviral genes [159]. The inhibition of phosphorylated STAT1/2 nuclear transport is the major factor affecting alphavirus virulence [158, 160]. The CHIKV nsP2 can individually inhibit the translocation of pSTAT1/2 to the nucleus inhibiting the IFN downstream signaling. Mutational studies on CHIKV nsP2 has shown that mutation in NLS abolish the inhibition of JAK-STAT signaling [161].

Taken together, various studies performed on nsP2 suggest that it is a multifunctional protein which is involved in regulation of RNA replication, cytopathic effects and inhibition of host antiviral responses. Hence, nsP2 is an important target protein of CHIKV to identify its cellular counterparts which will help in understanding the molecular mechanisms behind its multiple functions and further in designing the antivirals against CHIKV.

2.12 Diagnosis

Chikungunya infection is mainly diagnosed on the basis of three parameters- clinical, epidemiological and laboratory criteria. However, the clinical and epidemiological similarities among Dengue and CHIKV make its diagnosis difficult and many times misdiagnosed as Dengue.

Generally, onset of high fever and severe arthralgia or arthritis is considered as the possible cases of CHIKV infection but laboratory confirmation is crucial for its distinction from Dengue and other arthritic alphaviruses. The laboratory findings are based on the viremic kinetics and host antibody response. During the early infection stage normally for 5-10 days after onset of symptoms, the detection of viral RNA level and virus particles in the serum samples are considered for diagnosis when the viral RNA titre is very high. While during the later phases of infection, serological methods are being generally used for the detection of antibody response when the levels of anti CHIKV IgG and IgM are high in serum and viral titer starts decreasing.

Viral isolation from infected patients is used for CHIKV detection in antibody negative samples which are obtained on or before 2 days of illness because after the generation of
antibodies, the virus isolation from serum samples is a difficult task [162]. Besides this, various 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 [163, 164], real time loop-mediated RT-PCR [165] and real time TaqMan RT-PCR targeting the envelope E1 gene [166] or non-structural nsP1 gene [167]. Recently, a one step SYBR-green based real time assay has been developed for CHIKV diagnosis which targets the non-structural nsP2 gene [168]. The serological methods used for detection of CHIKV specific antibodies include Enzyme linked assays (ELISA), indirect immunofluorescence (IFA), hemagglutination inhibition (HI) and micro neutralization (MNt). ELISA and IFA are very rapid and sensitive tests for infection detection 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.

2.13 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 these are not very effective for treatment options. Besides all these options passive immunization is an appropriate option for prevention and treatment for many viral infections [169]. 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 [49]. 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 [48]. To overcome this problem, 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 [170].

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. fomaline or tween-ether inactivated whole virus [50, 171-173], attenuated live vaccines e.g. TSI-GSD-218 [174, 175], recombinant proteins or virus like particles e.g. VLPs expressing structural proteins of CHIKV [44] and DNA vaccination e.g. intramuscular injection of plasmids encoding CHIKV-capsid, E1 and E2 [176, 177]. Since none of these vaccines have been licensed and not available in market, protection against mosquito bite and vector control are the main preventive measures.

2.14 Rationale of the study

Although CHIKV has been identified in 1962, there is dearth of knowledge considering the interaction among its proteins and with the host cell at a molecular level. The analysis of viral-viral and viral-host protein interactions is an important step towards understanding the molecular mechanisms involved in viral replication and infection. Among the CHIKV encoded proteins, the non-structural proteins play important role in determining the rate of viral infection, as they form the replicase complex of the virus and also participate in regulating host responses generated against viral infection. The main objective of this study is to identify the interactions among non-structural (nsPs) proteins that make up the replicase complex (intra-viral) and those with the host (viral-host) which may play role in viral replication and persistence. The intraviral protein interactions were identified by using Yeast two-hybrid (Y2H) system, pulldown and ELISA. The host proteins targeted by this virus were identified by in silico method based on a structural similarity based approach. Further, cDNA Brain library was screened interactors of CHIKV nsP2 using Y2H system and the identified proteins were confirmed by in vitro assays. The interactions identified will help in understanding the viral biology and for designing inhibitory molecules against these interactions to combat CHIKV infection.