ORIGINAL ARTICLE

Identification of a conserved neutralizing epitope in the G-protein of Chandipura virus

Daya V. Pavitrakar[1] · Rekha G. Damle[1] · Anuradha S. Tripathy[1] · Pratip Shil[1]

Abstract

Chandipura virus (CHPV), associated with an encephalitic illness in humans, has caused multiple outbreaks with high mortality in central and western India in recent years. The present study compares surface glycoprotein (G-protein) from prototype and recent outbreak strains using *in silico* tools and *in vitro* experiments. *In silico* epitope predictions (B-cell and T-helper cell) for the sequences, 3D structure prediction and comparison of the G-proteins of the strains: I653514 (Year 1965), CIN0327 (Year 2003) and 148974 (Year 2014) revealed that the CHPV G-protein is stable and antigenic determinants are conserved. A monoclonal antibody developed against strain CIN0327 (named NAbC) was found to neutralize prototype I653514 as well as the currently circulating strain 148974. *In silico* antigen-antibody interaction studies using molecular docking of predicted structures of NAbC and G-proteins of various CHPV strains led to the identification of a conserved neutralizing epitope in the fusion domain of G-protein, which also contained a putative T-helper peptide. The identification of a conserved neutralizing epitope in domain IV (fusion domain amino acids 53 to 172) of CHPV G-protein is an important finding that may have the scope towards the development of protective targets against CHPV infection.

Introduction

Taxonomically, Chandipura virus (CHPV) belongs to the *Vesiculovirus* genus of the *Rhabdoviridae* family. CHPV causes acute encephalitis in human paediatric populations [1]. This virus was responsible for encephalitis outbreaks reported in Andhra Pradesh/Telengana and Maharashtra, India during the years 2003 and 2004 respectively [1, 2]. It is a negative sense, single stranded RNA virus which consists of five structural proteins in the canonical order N-P-M-G-L: a nucleoprotein (N), a nucleocapsid-associated phospho-protein (P), a matrix protein (M), a glycoprotein (G) and the viral polymerase (RdRp, L) located between 3' leader and 5' trailer sequences.

The G-protein is a trimeric trans-membrane glycoprotein that enables virus adsorption, assembly, budding and also elicits an antibody response thus acting as a major antigenic determinant [3]. Depending on the state of infection it adopts different reversible conformational states: i) native state present on the virus surface and stable above pH 7.0 [4] ii) activated state that fuses with the target membrane [5] iii) A fusion inactive post-fusion state which is stable under low pH conditions [6]. In earlier studies, the structure of a pre-fusion form of vesicular stomatitis virus (VSV) G-protein was analysed by the molecular replacement of different domains. The G-protein was divided into different domains, domain I: Lateral domain (1 to 17 and 310 to 382), domain II: Trimerization domain (18 to 35, 259 to 309, and 383 to 405), domain III: PH domain (36 to 46 and 181 to 258), domain IV: Fusion domain (53 to 172), Cter C-terminal part (406 to 413) and RbI-II domain: Rigid block (1 to 25 and 273 to 382) [7].

Being the major antigenic determinant, the G-protein of CHPV requires in depth studies to better understand its potential as a target for different preventive and diagnostic approaches. The emergence of massive CHPV outbreaks in 2003 with case fatality rates as high as 56-75 % [2] as well as isolated outbreaks thereafter from newer geographical

✉ Pratip Shil
    shil.p@gov.in; shilpratip@gmail.com

[1] ICMR-National Institute of Virology, 130/1 Sus Road, Pashan, Pune 411021, India

regions of India have generated interest in understanding the virus and its antigenic potential.

Pioneering studies have been performed to understand the biological mechanisms underpinning the neuropathology caused by CHPV. Genetic characterization and phylogenetic analysis have revealed changes in the G-protein of CHPV, however, the impact of these mutations on protein structure needs further in-depth analysis [8–10].

CHPV outbreaks have also been reported in south central India covering a few states: Andhra Pradesh (Telengana) and Maharashtra. Over recent years outbreaks have been reported from new areas like Gujarat, which is in the extreme western part of India. This raised many queries including, whether recent outbreaks were caused by altered antigenicity or altered surface protein properties of the virus. This prompted us to undertake comparative analyses of G-proteins from recent strains and the established prototype using bioinformatics tools along with supportive laboratory experiments. In the present investigation, we have compared the G-proteins of recent outbreak (2014) strains with those from 2003 and the established prototype strain (1965) using bioinformatics techniques to understand sequence, structural and antigenic variability. Additionally we have developed a neutralizing monoclonal antibody (MAb) against CHPV and have assessed its neutralizing ability against the emerging/recent CHPV strains.

Materials and methods

Cells and virus strains

Vero cells and CHPV strains used in the study were obtained from ATCC and the virus repository of the National Institute of Virology, Pune, respectively. The accession number and CHPV strain names are enlisted in Supplementary Table 1.

Dataset

The dataset comprised seven Indian isolates of CHPV covering the outbreaks 1965-2014 from different geographical locations. During analysis, the G-protein sequences were subjected to Multiple Sequence Alignment (MSA) using the CLUSTALW algorithm as implemented in MEGA 6.0 [11]. A phylogenetic tree was constructed using the maximum likelihood algorithm and as a test, a bootstrap with 1000 replications was used. Pair-wise comparison of sequences (all possible combinations) were performed using the ALIGN algorithm as implemented in the ISHAN package [12]. Three representative strains (I653514, CIN0327 and 148974) isolated at different time points were selected for further structure-based analysis and docking studies.

Antigenic variability

In-silico prediction of antigenic determinants was performed for three selected G-protein sequences using methods described earlier [13, 14] as implemented in the B-cell epitope prediction tool at the Immune Epitope Database (IEDB) available at http://www.iedb.org/. For implementation of the Kolaskar method of antigenicity prediction [13] a moving window of seven amino acids and threshold antigenicity of 1.0 were considered as ideal conditions. ElliPro based predictions of linear and conformational B-cell epitopes were also carried out [15].

Monoclonal antibody development, characterization and sequencing

The MAb against Indian CHPV strain (CIN0327), isolated during the 2003 Andhra Pradesh outbreak, was developed using conventional mouse hybridoma technologies, described earlier [16]. In brief, infant BALB/c mice were inoculated intraperitoneally with 100 PFU of CHPV. Brains of infected sick mice were harvested at 24 hours (hrs) post inoculation and a 10% mouse brain suspension was prepared in normal saline. The clarified supernatant was used for immunization of 6-8 wks old BALB/c mice. The spleen from immune mice was harvested and fusion of spleen cells with Sp2/0 myeloma cells at a proportion of 5:1 was carried out using 50% polyethylene glycol HybriMax (MW 1450, Sigma Aldrich, USA) as a fusogen. The hybrids secreting anti-CHPV antibodies were selected on the basis of reactivity in an in-house antibody detection ELISA using purified CHPV as a coating antigen. The positive hybrids were cloned by limiting dilution and individual clones were screened for the secretion of anti-CHPV MAbs in ELISA. The positive clones were propagated and characterized for further use. The study was approved by the Institutional Biosafety and Animal Ethics Committee and the experiments were conducted in accordance with the guidelines.

Characterization of MAbs was done with respect to isotype analysis, determination of protein specificity and ability to neutralize CHPV. The isotyping of clones was carried out using a Rapid ELISA Mouse mAb Isotyping Kit (Thermo Scientific™ Pierce™, USA) as per manufacturer's instructions. The protein specificity of the clone was determined by western blot analysis, as per the established protocol [16] with suitable modifications for CHPV. In brief, 10 μg of a purified preparation of CHPV was electrophoresed through a 12% polyacrylamide gel. The separated proteins were electrophoretically transferred to a nitrocellulose membrane (Bio-Rad, USA) at a constant voltage

of 25 V for 30 min in electrophoretic transfer cell (Trans-Blot SD, Bio-Rad, USA). Non-specific protein binding sites were blocked by using 5% skimmed milk powder in Tris-buffered saline (TBS) overnight at 4°C. Culture supernatant of MAb clones (1:50) was used as a source of primary antibody. During the assay, an anti-CHPV mouse polyclonal antibody was used as a known positive control. Antibody bound to CHPV antigen was probed with HRP conjugated anti-mouse IgG (Sigma, USA) and 3,3′-diaminobenzidine tetrahydrochloride hydrate (Sigma, USA) was added as a chromogenic substrate for the development of an antigen-antibody reaction.

The ability of MAbs to neutralize CHPV strains from different geographical regions and time was tested using an *in vitro* cytopathic effect (CPE) based micro-neutralization assay (MN-CPE method) as described earlier with a few modifications [17]. The assay was performed using Vero cells. The MAb culture supernatant was serially two-fold diluted in Minimum Essential Medium with 2% foetal calf serum and 100 TCID$_{50}$ of the pre-titrated homologous CHPV strain was used as a source of virus. The anti-CHPV immune mouse serum and normal mouse serum, diluted serially as above, were used as positive and negative assay controls respectively. The assay was terminated at 48 hrs post infection by staining with amido black stain. Fifty percent neutralization dose (ND$_{50}$) titer >10 was considered as a positive reaction.

For antibody sequencing the MAb clone was propagated and total cellular RNA was extracted from 5×10^6 cells using a Qiagen RNeasy mini kit (Qiagen, Germany) as per manufacturer's instructions. This was used for cDNA synthesis using SSIII RT (Invitrogen, ThermoFischer scientific, USA). PCR amplification of the variable region (Fv) of the heavy and light chain was carried out using Platinum taq DNA polymerase high-fidelity (Invitrogen, ThermoFischer scientific, USA) with primers and cycling conditions as reported earlier [16]. The PCR amplified products were gel purified using QIAquick gel extraction kit (Qiagen, Germany). The heavy and light chain variable region was sequenced by Sanger's method using Big Dye terminator V3.1 cycle sequencing kit (Applied biosystem, ThermoFischer scientific, USA) as per manufacturer's instructions and sequencing was carried out using a ABI3100 Genetic Analyzer. The coding amino acid (AA) sequences for the antibody nucleotide sequence was generated using the "Translate tool" available on the ExPASy proteomics server (http://expasy.org/tools/dna.html).

3D structure prediction of the G-protein and the Fv of MAb

The 3D structures of the G-proteins from the strains AHA42525, ADO63668 and ATB17675 were predicted using the Swiss Model online workstation with the VSV pre-fusion structure as a template (PDB ID: 5i2S) [18].

3D structures for the Fv region of the MAb were predicted using the ABodyBuilder tool available from the Structural Antibody Prediction Server [19] http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/WelcomeSAbPred.php. For all predicted structures energy minimization was performed using the GROMOS96 force field application in the Swiss PDB-Viewer (SPDBV) [18]. Predicted structures were also subjected to a PROCHECK analysis [20]. Visualizations of protein structures and imagery were generated using the Discovery Studio Visualizer, a product of BIOVIA software available at http://3dsbiovia.com/products/.

Molecular docking

A rigid body docking protocol was used to determine the binding site and interacting residues between the antigen and antibody. *In-silico* unconstrainted docking of the MAb with the G-protein of AHA42525, ADO63668 and ATB17675 strains was carried out using the ZDOCK server with default parameters [21, 22]. From the 10 different solutions returned by ZDOCK in each case, the best solution was selected based on: i) the complementarity determining regions (CDR) of NAbC interacting with the antigen at the antigen-antibody interface, and ii) the value of minimized energy of the complex being the least [16].

Prediction of T-helper epitopes

T-helper epitopes (MHC Class II epitopes) were predicted for G-protein sequences using MHC class II epitope prediction tools available from IEDB (IEDB; http://www.iedb.org/) [23]. The most prevalent HLA class II alleles from both the Andhra Pradesh-Telangana and Gujarat region, where the CHPV outbreaks were reported, were considered for prediction [24, 25]. The Andhra Pradesh-Telangana region located in southern India had the following dominant Human Leukocyte Antigen - antigen D Related (HLA DRB1) alleles: HLA DRB1-04, HLA DRB1-07 and HLA DRB1-15, whereas the Gujarat region had: HLA DRB1-03, HLA DRB1-11, HLA DRB1-13 and HLA-DRB1-15 alleles.

Study design

CHPV G-protein sequences from seven strains were compared by MSA. Of these, two strains (CIN0327 and 148974) having the maximum differences in amino acid composition with respect to the prototype strain of 1965, were considered for detailed analysis. The 3D structure for these two G-proteins and the 1965 strain was predicted. The antigenic variability between these three strains was determined by comparing the predicted linear and conformational B-cell epitopes. The G-protein sequences were also screened for the existence of any conserved T-helper epitope using MHC class II epitope prediction tools for the available dominant alleles in the reported outbreak regions. A MAb was developed against the CIN0327 strain of CHPV (G-protein, accession number: ADO63668) and tested for its neutralizing ability against currently circulating CHPV strains representing different years of isolation. The sequencing of the Fv of individual MAbs was carried out and the corresponding amino acid sequences were used to predict 3D structures. To identify the existence of any conserved neutralizing B-cell epitopes, *in silico* docking of the MAb with the predicted structures of the G-proteins from various CHPV strains was carried out.

Results

Sequence data

The analysis of the amino acid sequence of the G-protein of seven Indian isolates of CHPV revealed amino acid substitutions at four positions compared to the prototype strain (Supplementary Table 2). These were located in domain II (Trimerization domain) at positions L19S and Y22S; in domain III (PH domain) at S41N and G222A/S and in the RblII (Rigid block II) domain at P367M. Apart from these, a few additional mutations were observed (Supplementary Table 2) but all were favourable substitutions having similar

physico-chemical properties and hence did not alter the back-bone fold of the protein. A phylogenetic tree (supplementary Figure 1) for the CHPV G-protein sequences was constructed with Isfahan virus (ACCN: YP_007641385) as the outlier. All the CHPV strains were found to be highly conserved.

Antigenic variability

The linear B-cell epitope prediction of the three selected strains revealed that most of the epitope were conserved. Supplementary Table 3 lists all the predicted B-cell epitopes in the ecto-domain, from amino acid 1-413 of the CHPV G-protein. All twenty-one predicted epitopes were conserved including the longest epitope: 122-SGTLVSPGFPPESCG-YASVTDSEFLVIMITPHHVGVD-158. Mutation G222A observed in domain III in strain CIN0327 and G222S in strain 148974 did not alter the antigenicity of the epitope covering amino acids 220-230. The prototype strain has a unique potential B-cell epitope occurring at 38-VTKSTRY-CPM-47, located in domain III which was not present in strains from 2003 onwards, due to a S41N mutation. Results from the ElliPro revealed the existence of conserved linear and discontinuous B-cell epitopes as detailed in supplementary Table 4. The ElliPro analysis identified segment 122-146 as strongly antigenic (both as a linear as well as a conformational epitope).

MAb development and characterization

In the *in vitro* neutralization assay, the MAb at dilution 1:12800 neutralized homologous strain CIN0327, showing 100% neutralization of 100 $TCID_{50}$ virus, indicating a very high ND50 titer (Figure 1). This MAb was named as NAbC (Neutralizing Antibody against CHPV). The testing of heterologous neutralization by NAbC also showed its ability to neutralize CHPV strains e.g. I653514, 1015405, 12588, 1210586 and 148974.

Fig. 1 *In vitro* neutralization of CHPV by MAb, NAbC. Row-1: MAb (NAbC) neutralized 100 $TCID_{50}$ virus with ND50 titer >12800. Row-2: Polyclonal antibody against CHPV. Row-3: Normal mouse serum. Row-4: Well (1-5) - Virus control and well (6-8) - Cell control
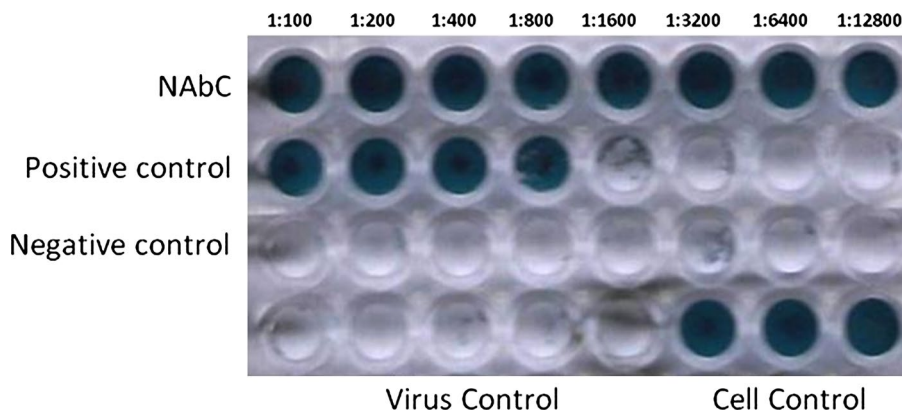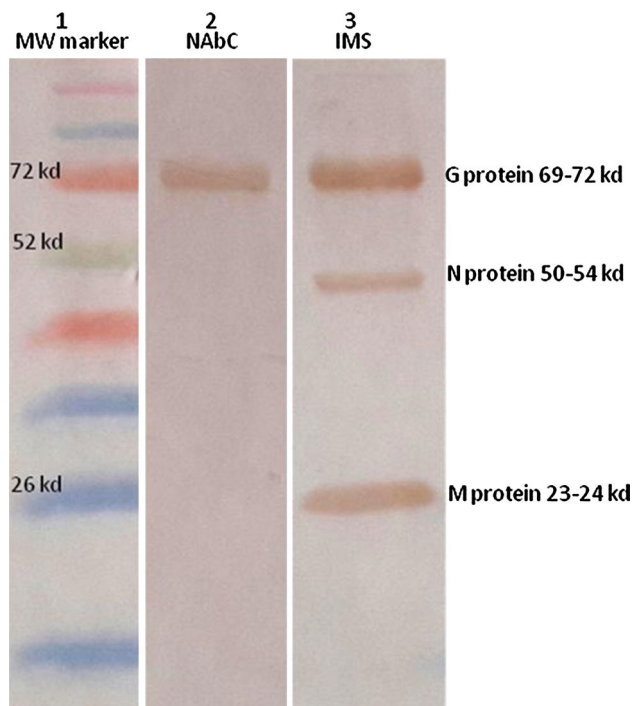
Fig. 2 Western blot (WB) analysis of anti-CHPV MAb, NAbC. Purified CHPV, 10 μg /well was subjected to electrophoresis through a 12% SDS-PAGE gel and proteins were transferred to a nitrocellulose membrane and probed with anti-CHPV MAb. The MAb, NAbC specifically detected the G-protein of CHPV in WB. Lane 1: MW-Protein molecular weight marker, Lane 2: NAbC – CHPV antigen probed with MAb culture supernatant, showing a band at 69-72 kd corresponding to the G-protein of CHPV. Lane 3: CHPV antigen probed with anti-CHPV polyclonal antibody collected from the immunized mice

A hybrid, secreting antibodies to CHPV, was obtained which yielded a further MAb clone, which was characterized. The isotype of the MAb was IgG2a with a kappa light chain. In Western blot, the MAb recognized the surface glycoprotein G of CHPV (Figure 2).

Sequencing of the variable regions of NAbC

The variable region of heavy (V_H) and light (V_L) chains of NAbC were amplified by PCR. Bands of 409 bp and 340 bp corresponding to the V_H and V_L respectively were observed on a 1.5% agarose gel. The PCR products were gel purified and sequenced from both ends using sense and anti-sense primers. Before undergoing antibody structure prediction the sequences were analyzed using the NCBI protein BLAST tool. The sequence data generated in this study has been submitted to GenBank: NAbC-V_H (GenBank accession number: MH395742) and NAbC-V_L (GenBank accession number: MH395743).

Table 1 Structural properties of the predicted structures

Protein structures	Occupancy* in Ramachandran plot (%)	Minimized energy (kJ/mol)
AHA42525	99.2	-46718.444
ADO63668	99.5	-49839.891
ATB17675	99.5	-47359.129
5i2s	100	-54915.047
NAbC	99.4	-9792.784

*Includes most favoured region, generously allowed region and additionally allowed region

3D structure predictions of the G-protein and Fv of MAb

The 3D structure prediction of the ecto-domain of CHPV G-proteins from the strains I653514, CIN0327 and 148974 were carried out using the crystal structure of VSV G-protein (PDB ID- 5i2s) as a template. The minimized energy for the predicted structures is summarized in Table 1. The occupancy of amino acids in the Ramachandran plot for the G-protein of all three strains was found to be 99.2%, 99.5% and 99.5 % respectively (favourable and additionally allowed regions).

The optimized models of NAbC_Fv regions and CDR for both variable light and variable heavy chains were identified. The identified CDR regions are listed in Table 2. The minimized energy for the predicted structure was found to be -9792.784 kJ/mol. The PROCHECK analysis revealed that the occupancy of amino acids in the Ramachandran plot was 99.4 % (favourable and additionally allowed regions).

Molecular docking

The unconstrained docking of the predicted 3D structure of NAbC with modeled structures of the CHPV G-protein from three different strains revealed docking foot prints covering the residues stretching from 122-SGTLVSPGFP-PESCG-136. This peptide is part of the predicted B-cell

Table 2 CDR regions (Kabat classification) in NAbC Fv

CDR regions of heavy and light chains of NAbC

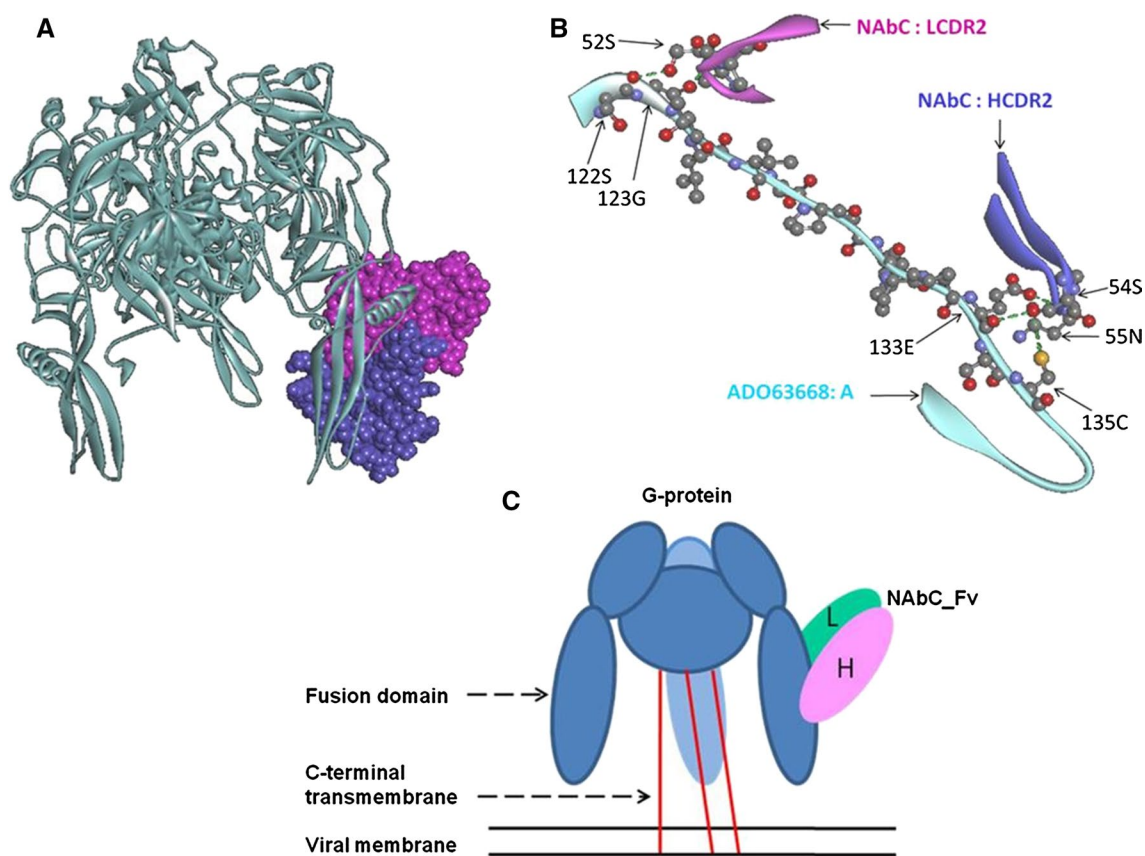| CDR	NAbC sequence
HCDR1	31-SYYIS-35
HCDR2	50-EINPSNGGTNFNEKFKS-65
HCDR3	95-LGHDGYYGPSYYAMDY-110
LCDR1	24-KASQYVGTAVA-34
LCDR2	50-WASARHT-56
LCDR3	89-QQYSSFPRT-97

Fig. 3 A) NAbC Fv with heavy chain (blue) and light chain (magenta) complexed with the A chain of the CHPV G-protein trimer (ribbon mode, Cyan color). B) Close-up view of epitope CG122 of the G-protein forming H-bonds with complementarity-determining regions (CDRs) of NAbC Fv. C) Schematic representation of the NAbC Fv binding to the fusion domain of the G-protein (pre-fusion state) as expected under natural physiological conditions

epitope 122-SGTLVSPGFPPESCGYASVTDSEFLVIM-ITPHHVGVD-158, which is conserved in all three strains. A representative image is presented in Figure 3A. In each case, residues of NAbC formed stable hydrogen bonds with residues in the G-protein (Figure 3B). This region was found to have a conserved linear epitope by the Kolaskar method. The residues of the G-protein (strain CIN0327) Ser122 and Gly123 formed H-bonds with the residue Ser52 of LCDR2 (second complementarity determining region of V_L) in NAbC. H-bonds were also noted between residue Glu133 of the G-protein with residue Ser54 and Asn55 of HCDR2 (second complementarity determining region of V_H) and Cys135 of the G-protein with Ser54 of HCDR2. These H-bonds held true for the docked NAbC G-protein complexes with strains I653514 and 148974 as well. We named the antigenic determinant 122-SGTLVSPGFPPESCG-136 as CG122. In the pre-fusion conformation of the G-protein, CG122 occurs as a linear beta strand in the fusion domain and is fully exposed under natural physiological conditions.

In order to evaluate the specificity of NAbC to the antigenic determinant CG122, we carried out *in silico*

mutagenesis by introducing the following mutations in the G-protein: E133M and E133A separately. The mutated sequences were modelled using the G-Protein structure of VSV as the template (PDB ID: 5i2S). The root mean square deviation (RMSD) of the models with mutations E133M and E133A was found to be 0.05Å. Docking of mutated G-proteins to NAbC revealed a loss of recognition by the antibody. In each case the CDRs of NAbC did not bind to the mutated antigenic regions. The docking outputs returned by the ZDOCK analysis are physically unrealistic and not feasible under natural physiological conditions.

T-helper epitopes

Considering the fact that outbreaks in the Andhra Pradesh-Telangana region have affected southern Indian populations, we considered HLA DRB1 alleles prevalent in human populations in that geographical area. These are HLA DRB1-04 (12-14%), HLA DRB1-07 (14-19%) and HLA DRB1-15 (19-28%) [24]. Predictions revealed the existence of overlapping T-helper peptides in three different segments of

the G-protein from aa 10-24, 141-160 and 479-493, considering the lowest percentile value (cut off <5), indicating highest affinity [23]. Details of the predicted T-helper epitopes are listed in supplementary Table 5. Of the three segments the MHC class II peptides with the highest affinity occur between aa 141-160 for all the alleles present in the Andhra Pradesh region (supplementary Table 5). This T-helper peptide correspondeds to the neutralizing B-cell epitope (122-SGTLVSPGFPPESCGYASVTDSEFLVIMIT-PHHVGVD-158).

HLA DRB1 alleles prevalent in the geographical area of Gujarat which were available in the IEDB reference allele list are: HLA DRB1-03, HLA DRB1-11, HLA DRB1-13 and HLA DRB1-15. For all these HLA DRB1 alleles except DRB1-03, predictions revealed the existence of overlapping T-helper peptides (with highest affinity and lowest percentile score of < 5) in three different segments of the G-protein, those identified in the DRB1 alleles of Andhra Pradesh-Telangana, i.e. aa 10-24, 141-160, and 479-493. For the DRB1-03 allele no predictions were obtained in the region 141-160. The high affinity predicted T helper epitopes in the antigen-antibody interaction domain among the different alleles are summarized in supplementary Table 5.

Discussion

Understanding the underlying factors leading to recurrent Chandipura outbreaks with high mortality deserves special attention. The present study is an attempt to determine the amino acid changes in the G-protein of Chandipura virus occurring during recent CHPV outbreaks in new areas of India and associate these with functional and structural changes to the virus. Previous reports [26, 27] of putative roles for the CHPV G-protein prompted us to characterize G-proteins from various strains in terms of: i) their sequence, ii) a prediction and comparison of 3D structures, and iii) predictions of their antigenicity involving: a) B-cell linear and conformational epitopes and b) T-cell epitopes (MHC class II), using appropriate bioinformatics tools.

Bioinformatic-based analyses of the sequences revealed that the G-proteins of the strains CIN0327 (Andhra pradesh/Telangana) and 148974 (Gujarat) had seven and six amino acid mutations respectively compared to the I653514 prototype strain, though these mutations did not alter the 21 major epitopes. The favourable substitutions did not alter the physico-chemical properties significantly and did not have any effect on the 3D fold of the backbone chain of the G-protein. A recent study has also identified mutations in one of the circulating CHPV strains [28], the majority of these amino acid changes were either favourable substitutions or occurred at non-antigenic region of the protein. The region comprising aa 4- 27 was predicted (*in silico*) to have antigenic properties

in all the strains in spite of the variations in amino acid composition. Since this region remains embedded in the trimer interface in the native pre-fusion state of the G-protein and did not interact with the host immune system, it could not be considered as a potential B-cell epitope. However, since these were *in silico* predictions only, in parallel, laboratory experiments were carried out where MAbs against CHPV were developed. The MAb NAbC neutralized the CIN0327 strain and other strains I653514 and 148974, leading us to hypothesize the existence of a conserved and immunodominant neutralizing epitope.

The recognition of an antigenic region by any antibody depends on key factors such as surface electrostatics that determines the formation of H-bonds or salt bridges and shape compatibility [16, 29]. To delineate the epitopes for NAbC, we sequenced the Fv region of the antibody, predicted the 3D structure and performed molecular docking with the G-proteins using standard bioinformatics tools. Docking studies revealed that NAbC binds to the G-protein at the antigenic site 122-136 which is part of the B-cell epitope 122-SGTLVSPGFPPESCGYASVTDSEFLVIM-ITPHHVGVD-158. Five H-bonds were formed between residues in the CDRs of NAbC and residues 122-136 of the G-protein. Experimental data indicated that NAbC binds to both denatured (Western blot) and native (neutralization assay) forms of the G-protein. This corroborated our bioinformatic analyses that the antigenic region 122-136 (CG122) is a conserved linear epitope.

Specificity evaluation of antigen-antibody binding in terms of docking of NAbC with two different mutants of the G-protein (e.g. E133M and E133A) created by *in silico* mutagenesis showed that NAbC failed to recognize or bind to the mutated versions of the G-proteins. The docking output generated complexes which were not feasible under natural physiological conditions (non-feasible solutions). Altered local surface contour and surface electrostatics in the mutants could account for the failure of NAbC to effectively recognize and bind (Figure 4).

Considering the HLA class II alleles prevalent in the population of the affected areas, predictions of T-helper epitopes in the G-proteins were conducted using standard IEDB tools. Considering the demographics of the districts where this virus traditionally causes outbreaks, we considered the available HLA DRB1 alleles of the native population for prediction of the T-helper epitopes. Also, in order to account for the recent outbreak in Gujarat (2014 strain 148974), we predicted T-helper epitopes for the said strain using the HLA information available for populations in Gujarat [25] and selected only the high affinity (lowest percentile) T-helper peptides. Only three regions on the G-protein, that account for overlapping T-helper peptides, were identified. One of them coincided with the antigenic determinant: 122-SGTLVSPGFPPESCGYASVTDSEFLVIMITPHH-

Fig. 4 Surface electrostatic analyses of the CHPV G-protein strain CIN0327 (Accession no. ADO63668) as well as mutants E133M and E133A
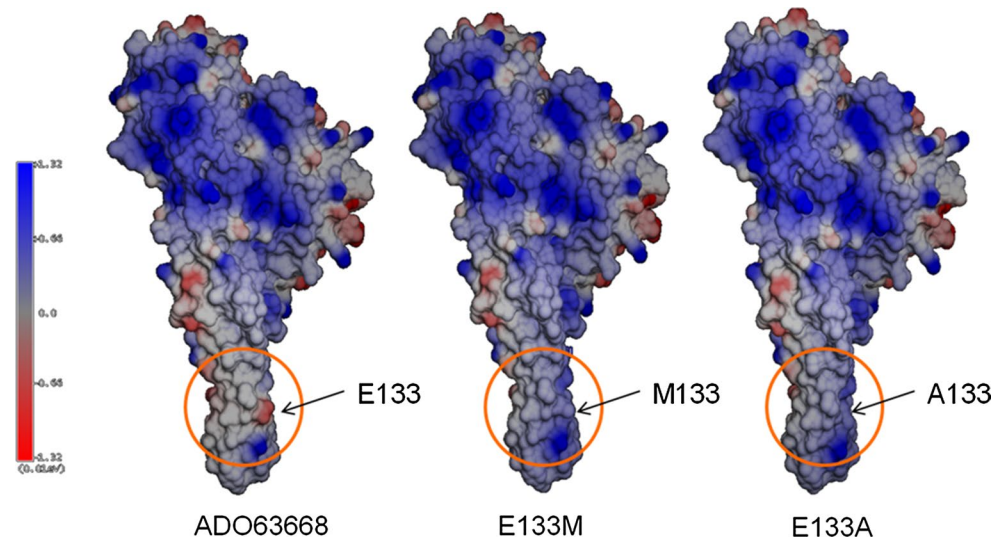
VGVD-158. The T-helper peptide in this region is 141-160. Thus, we have identified an antigenic region 122-158 which offers a conserved B-cell neutralizing epitope as well a T-helper peptide 141- VTDSEFLVIMITPHHVGVDDY-160.

In our opinion, the answer to the question, "Is there any change in the surface protein properties of the virus that led to the spread to newer areas?" is no. The mutations (or amino acid differences) in the G-protein of recent isolates did not alter the antigenicity of exposed epitopes nor the 3D structure of the surface protein drastically. Spread to new areas could either be due to changes in other proteins of the virus or due to environmental factors which need appropriate epidemiological and zoonotic investigation. However, the present study helped us in achieving the following: i) antigenic characterization of the CHPV G-protein of different strains, ii) development of a neutralizing MAb i.e. NAbC, iii) delineation of a conserved neutralizing epitope 122- SGTLVSPGFPPESCG-136 and conserved T-helper peptides (covering the amino acids: 141 - 160) in the fusion domain of the G-protein.

In a nutshell, the current study reveals that the CHPV G-protein is stable and antigenic determinants on the protein surface are conserved. The identification of a conserved neutralizing epitope in domain IV (fusion domain amino acid -53 to 172) on CHPV G-protein is an important finding which may lead towards the development of therapeutic tools. Research to this effect may be undertaken in future.

Compliance with ethical standards

Data availability The NAbC variable region sequences generated during the study are available in the GenBank: GenBank Accession numbers MH395742 and MH395743.

Ethical approval All animal experiments were approved by Institutional Animal Ethical committee and were carried out according to guidelines of IAEC.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Rao BL, Basu A, Wairagkar NS, Gore MM, Arankalle VA, Thakare JP, Jadi RS, Rao KA, Mishra AC (2004) A large outbreak of acute encephalitis with high fatality rate in children in Andhra Pradesh, India, in 2003, associated with Chandipura virus. Lancet 364(9437):869–874. https://doi.org/10.1016/S0140-6736(04)16982-1

2. Sudeep AB, Gurav YK, Bondre VP (2016) Changing clinical scenario in Chandipura virus infection. Indian J Med Res 143(6):712–721. https://doi.org/10.4103/0971-5916.191929

3. Neumann G, Whitt MA, Kawaoka YA (2002) Decade after the generation of a negative-sense RNA virus from cloned cDNA— what have we learned? J Gen Virol 83:2635–2662. https://doi.org/10.1099/0022-1317-83-11-2635

4. Clague MJ, Schoch C, Zech L, Blumenthal R (1990) Gating kinetics of pH-activated membrane fusion of vesicular stomatitis virus with cells: stopped-flow measurements by dequenching of octadecylrhodamine fluorescence. Biochemistry 29(5):1303–1308. https://doi.org/10.1021/bi00457a028

5. Durrer P, Gaudin Y, Ruigrok RW, Graf R, Brunner J (1995) Photolabeling identifies a putative fusion domain in the envelope glycoprotein of rabies and vesicular stomatitis viruses. J Biol Chem 270:17575–17581. https://doi.org/10.1074/jbc.270.29.17575

6. Yao Y, Ghosh K, Epand RF, Epand RM, Ghosh HP (2003) Membrane fusion activity of vesicular stomatitis virus glycoprotein

G is induced by low pH but not by heat or denaturant. Virology 310(2):319–332. https://doi.org/10.1016/S0042-6822(03)00146-6

7. Roche S, Rey FA, Gaudin Y, Bressanelli S (2007) Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. Science 315(5813):843–848. https://doi.org/10.1126/science.1135710

8. Basak S, Mondal A, Polley S, Mukhopadhyay S, Chattopadhyay D (2007) Reviewing Chandipura: a vesiculovirus in human epidemics. Biosci Rep 27(4–5):275–298. https://doi.org/10.1007/s10540-007-9054-z

9. Cherian SS, Gunjikar RS, Banerjee A, Kumar S, Arankalle VA (2012) Whole genomes of Chandipura virus isolates and comparative analysis with other rhabdoviruses. PLoS One 7(1):e30315. https://doi.org/10.1371/journal.pone.0030315

10. Verma AK, Ghosh S, Pradhan S, Basu A (2016) Microglial activation induces neuronal death in Chandipura virus infection. Sci Rep 6:22544. https://doi.org/10.1038/srep22544

11. Tamura Koichiro, Stecher Glen, Peterson Daniel, Filipski Alan, Kumar Sudhir (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197

12. Shil P, Dudani N, Vidyasagar PB (2006) ISHAN: sequence homology analysis package. In Silico Biol 6(5):373–377

13. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276(1–2):172–174. https://doi.org/10.1016/0014-5793(90)80535-Q

14. Gangwar RS, Shil P, Cherian SS, Gore MM (2011) Delineation of an epitope on domain I of Japanese encephalitis virus Envelope glycoprotein using monoclonal antibodies. Virus Res 158(1–2):179–187. https://doi.org/10.1016/j.virusres.2011.03.030

15. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinform 9:514. https://doi.org/10.1186/1471-2105-9-514

16. Damle RG, Jayaram N, Kulkarni SM, Nigade K, Khutwad K, Gosavi S, Parashar D (2016) Diagnostic potential of monoclonal antibodies against the capsid protein of chikungunya virus for detection of recent infection. Arch Virol 6:1611–1622. https://doi.org/10.1007/s00705-016-2829-4

17. Damle RG, Patil AA, Bhide VS, Pawar SD, Sapkal GN, Bondre VP (2017) Development of a novel rapid micro-neutralization ELISA for the detection of neutralizing antibodies against Chandipura virus. J Virol Methods 240:1–6. https://doi.org/10.1016/j.jviromet.2016.11.007

18. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22(2):195–201. https://doi.org/10.1093/bioinformatics/bti770

19. Dunbar J, Krawczyk K, Leem J, Marks C, Nowak J, Regep C, Georges G, Kelm S, Popovic B, Deane CM (2016) SAbPred: a structure-based antibody prediction server. Nucleic Acids Res 44(W1):W474–W478. https://doi.org/10.1093/nar/gkw361

20. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK-a program to check the stereochemical quality of protein structures. J App Cryst 26:283–291. https://doi.org/10.1107/S0021889892009944

21. Chen R, Li L, Weng Z (2003) ZDOCK: an initial-stage protein-docking algorithm. Proteins 52(1):80–87. https://doi.org/10.1002/prot.10389

22. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z (2014) ZDOCK server:interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics 30(12):1771–1773. https://doi.org/10.1093/bioinformatics/btu097

23. Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B (2008) A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 4(4):e1000048. https://doi.org/10.1371/journal.pcbi.1000048

24. Dedhia L, Gadekar S, Mehta P, Parekh S (2015) HLA haplotype diversity in the South Indian population and its relevance. Indian J Transplant 9(4):138–143. https://doi.org/10.1016/j.ijt.2015.10.016

25. Singh A, Sharma P, Kar HK, Sharma VK, Tembhre MK, Gupta S, Laddha NC, Dwivedi M, Begum R, Gokhale RS, Rani R, Indian Genome Variation Consortium (2012) HLA alleles and amino-acid signatures of the peptide-binding pockets of HLA molecules in vitiligo. J Invest Dermatol 132(1):124–134. https://doi.org/10.1038/jid.2011.240

26. Jadi RS, Sudeep AB, Barde PV, Arankalle VA, Mishra AC (2011) Development of an inactivated candidate vaccine against Chandipura virus (Rhabdoviridae: Vesiculovirus). Vaccine 29(28):4613–4617. https://doi.org/10.1016/j.vaccine.2011.04.063

27. Venkateswarlu CH, Arankalle VA (2010) Evaluation of the immunogenicity of a recombinant glycoprotein-based Chandipura vaccine in combination with commercially available DPT vaccine. Vaccine 28(6):1463–1467. https://doi.org/10.1016/j.vaccine.2009.11.072

28. Damle RG, Sankararaman V, Bhide VS, Mahamuni SA, Walimbe AM, Cherian SS (2017) Genetic characterization of the glycoprotein G of Chandipura viruses in India with emphasis on an outbreak of 2015. Infect Genet Evol 55:112–116. https://doi.org/10.1016/j.meegid.2017.09.002

29. Shil P, Chavan S, Cherian S (2011) Molecular basis of antigenic drift in Influenza A/H3N2 strains (1968–2007) in the light of antigen-antibody interactions. Bioinformation 6(7):266–270. https://doi.org/10.6026/97320630006266