Biochemical and biophysical characterization of the RNA-dependent RNA polymerase from Chikungunya virus and discovery of a novel inhibitor as a potential drug candidate

Marjorie Freire
University of Sao Paulo

Luis Basso
State University of Northern Rio de Janeiro Darcy Ribeiro

Luis Mendes
University of Sao Paulo

Nathalya Mesquita
University of Sao Paulo

Melina Mottin
Universidade Federal de Goiás

Rafaela Fernandes
University of Sao Paulo

Lucca Policastro
University of Sao Paulo

Andre Godoy
University of Sao Paulo  https://orcid.org/0000-0002-0613-9164

Igor Santos
Federal University of Uberlandia

Uriel Ruiz
Federal University of Uberlandia

Ícaro Caruso
Sao Paulo State University

Bruna Sousa
Universidade Federal de Goias

Ana Jardim
Federal University of Uberlandia

Fabio Almeida
Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro

Laura Gil
FIOCRUZ
Carolina Andrade
Glaucius Oliva (✉ oliva@ifsc.usp.br)
University of Sao Paulo  https://orcid.org/0000-0003-2719-0302

Article

Keywords: CHIKV, nsP4, RNA-dependent RNA-polymerase, biophysical characterization, inhibitor discovery

Posted Date: January 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1280888/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Chikungunya virus (CHIKV) is the causative agent of Chikungunya fever, an acute febrile and arthritogenic illness with no effective treatments available. The development of effective therapeutics strategies could be significantly accelerated with detailed knowledge of the molecular components behind CHIKV replication. However, drug discovery is hindered by our incomplete understanding of their main components. The RNA-polymerase RNA-dependent (nsP4-CHIKV) is considered the key enzyme of the CHIKV replication complex and a suitable target for antiviral therapy. Herein, the nsP4-CHIKV was extensively characterized through experimental and computational biophysical methods. In the search for new molecules against CHIKV, a compound designated LabMol-309 was identified and mapped to bind to the nsp4-CHIKV active site. The inhibitory activity of LabMol-309 was evaluated in cellular-based antiviral assays using a CHIKV replicon system and a reporter virus. In conclusion, this study highlights biophysical features of nsP4-CHIKV and identified a new compound as a promising antiviral agent against CHIKV infection.

Introduction

The Chikungunya virus (CHIKV) belongs to the Togaviridae family and is the causative agent of Chikungunya fever. The main transmission route occurs through the bite of infected female mosquitoes of the Aedes sp. Genus. After CHIKV infection, the proportion of individuals who develop clinical and debilitating symptoms is considered the highest compared to other arboviruses, with an average of 80% of symptomatic cases (1, 2). The control of the mosquito vector remains the best prophylaxis since there are no licensed vaccines or efficient antivirals available (3). In this scenario, the infection caused by CHIKV has a high social impact and constitutes a serious public health issue (3).

Chikungunya fever presents an acute phase characterized by high fever, arthralgia, myalgia, headaches, edema, periorbital pain and cutaneous rash (4, 5). Later, some patients progress to the so-called chronic phase, mainly characterized by persistent arthralgia and musculoskeletal pain for months and even years (4, 6–8). Due to the lack of specific antiviral drugs, the infection is treated symptomatically, using analgesics and antipyretics in the acute phase and also with non-steroidal anti-inflammatory drugs (NSAIDs) and corticosteroids in the chronic phase (5).

CHIKV is a spherical, enveloped, and positive single-stranded RNA virus. As a member of the Alphavirus genus, its genome has approximately 12 kb and codes for two distinct polyproteins: structural and non-structural (9). The first one is cleaved and gives rise to five structural proteins – E1, E2, E3, C and 6k – which are part of the structure and viral assembly (9). The envelope proteins, specially E2 and E1, are responsible for virus anchoring, receptor interaction and membrane fusion, promoting virus entry in the host cell (10, 11). Due to its location on the viral surface, envelope proteins are targets of the humoral immune response and thus become targets for the vaccine's development against CHIKV (12).
On the other hand, the non-structural polyprotein is cleaved in four non-structural proteins (nsP1, nsP2, nsP3 and nsP4) that form the viral replication complex and also have some functions in the infection process, such as interaction with host factors (9, 13, 14). The nsP1 has methyltransferase and guanylyltransferase activities, promoting the capping of viral RNA. In addition, this protein also anchors the replication complex in the cell membrane (15). The nsP2 is a multifunctional protein with NTPase and helicase activities in the N-terminal region. Its C-terminus has a cysteine-protease activity, responsible for processing the non-structural polyprotein (15). The nsP3 is an accessory protein that recruits host cell factors that participate and optimize the replicative process (16). Finally, nsP4 is the RNA-dependent RNA polymerase (RdRp), which is considered the key enzyme of the CHIKV replication complex and acts principally by promoting the synthesis and elongation of viral RNA (17).

Therefore, due to their vital participation in the viral life cycle, non-structural proteins emerge as potential targets for the antiviral drugs’ development, aiming to interrupt the replication process and, consequently, the viral elimination (18). Among these proteins, nsP4 is an attractive target due to its central role in viral genome replication, transcription and some acting in genome repair (17, 19). To date, no three-dimensional structure of the nsP4 of CHIKV (nsP4-CHIKV) have been reported. This fact generates limitations and challenges when the final goal is applying a structure-based drug design against this target.

Due to the lack of high-resolution structural information of the nsP4-CHIKV and the need to search for new drugs to treat the infection caused by this virus, this work presents a detailed biophysical characterization of this protein. Size exclusion chromatography coupled with multi-angle light scattering was used to infer the oligomeric state of the nsp4-CHIKV protein in solution. A high prevalence of ordered helical secondary structures was observed by circular dichroism, which also showed that the nsP4-CHIKV unfolds under a cooperative transition during thermal denaturation. The thermal denaturation was further studied using differential scanning calorimetry, indicating a kinetically controlled process.

Moreover, in the search for ligands towards the development of new inhibitors, 12 compounds were screened using differential scanning fluorimetry experiments. One compound, LabMol-309, showed a significant effect on nsP4-CHIKV stability and was submitted to MicroScale Thermophoresis experiments to confirm its interaction with the protein. Furthermore, this interaction was also analyzed by nuclear magnetic resonance, monitoring chemical shift perturbation. To complete, in silico studies of molecular docking and molecular dynamics simulations were performed with the nsP4-CHIKV 3D model and the compound LabMol-309, to suggest the possible binding mode and maintenance of LabMol-309 at the protein active site.

Furthermore, the inhibitory activity of LabMol-309 was studied in cellular-based antiviral assays using a CHIKV replicon system and a reporter virus, and the results showed that this molecule caused inhibition in these cellular assays and has the potential to be further evaluated as a CHIKV inhibitor.

In this way, this work provides novel structural features of nsP4-CHIKV and identified a new compound that interacts with this protein, generating perspectives in the drug development field to treat the infection...
caused by CHIKV and, potentially other alphaviruses.

**Results And Discussion**

**nsP4-CHIKV purification and SEC-MALS analysis**

The nsP4-CHIKV was expressed and purified using chromatography systems and the purity was confirmed by acrylamide gel electrophoresis (Supplementary Figure 1). The nsP4-CHIKV is formed by 492 amino acids and has a theoretical molecular mass of 54.54 kDa. This construction covers the full-length RNA-dependent RNA polymerase (RdRp) domain, responsible for the nsP4-CHIKV function and where the catalytic aspartic acid residues (Asp253 and Asp348) are located (20). The Asp348 is in the GDD motif, a highly conserved sequence of viral polymerases (20).

To determine the oligomeric state of nsP4-CHIKV in the working buffer solution, SEC-MALS was employed to estimate the accurate molecular mass (MM) of the protein without the need to use globular proteins as standards for mass determination (21). The SEC-MALS data showed a low polydispersity index and yielded a MM of $(60 \pm 1) kDa$ for nsP4-CHIKV (Figure 1A). The proximity of the experimental value with the theoretical molecular mass of the protein suggests that the monomeric state is the most populated oligomer under the evaluated conditions.

**Evaluation of nsP4-CHIKV secondary structure profile**

The CD spectroscopy was used to estimate the secondary structure content of nsP4-CHIKV in solution (22). The CD spectrum of nsP4-CHIKV is characteristic of an α-helical rich protein, with two negative minima at 208 and 222 nm and one positive maximum around 195 nm (Figure 1B). This profile corroborates with structural features described for viral polymerases and other polymerase structures solved experimentally (20).

The Dichroweb (23) was used in the quantitative analysis of the CD spectrum. In this way, the CDSSRT method (24) presented an NMRSD of 0.009 for the SP175 database and 0.012 for Sets 4 and 7, considered the most adequate fit and the most representative of the secondary structure content of nsP4-CHIKV. Secondary structure content estimation from the CD spectrum confirmed the higher percentage of α-helix in the nsP4-CHIKV (60%). The remaining fractions were: 6.3% of sheets, 9.6% of turns and 24% of disordered structure.

**nsP4-CHIKV thermal stability profile through CD spectroscopy**

The CD technique was also applied to study the thermal behavior of this protein when subjected to temperature variations (25). Due to its helical profile, the folding transitions of nsP4-CHIKV was studied by monitoring the ellipticity at 222 nm as a function of the temperature variation (25). The results showed that the thermal denaturation process occurred cooperatively, presenting the transition from folded to unfolded states in a defined way (Figure 1C).
From this result, the thermodynamic parameters of nsP4-CHIKV unfolding were obtained with the adjustment of a two-state equilibrium model to the experimental data. Changes in thermal capacity were not considered and the adjustment was made taking into account the linear changes in pre- and post-transition ellipticity, as a function of temperature variation (25). Thus, it was possible to obtain the \( T_m \) and the apparent enthalpy variation (\( \Delta H_{\text{app}} \)) for the same protein concentration at two different heating rates (Table 1). The variation of the apparent entropy of the unfolding transition at each \( T_m \) was calculated as \( \Delta S_{\text{app}} = \Delta H_{\text{app}} / T_m \), considering \( \Delta G = 0 \) for each \( T_m \).

### Table 1

**Thermodynamic parameters of nsP4-CHIKV unfolding by CD spectroscopy.** The melting temperature \((T_m)\) and the apparent enthalpy change \((\Delta H_{\text{app}})\) were determined by fitting the CD data to a two-state equilibrium model (25). The apparent entropy change \((\Delta S_{\text{app}})\) of the unfolding transition at \( T_m \) was calculated as \( \Delta S_{\text{app}} = \Delta H_{\text{app}} / T_m \) since \( \Delta G = 0 \) at \( T_m \).

| Rate (°C/min) | \( T_m \) (°C) | \( \Delta H_{\text{app}} \) (kcal/mol) | \( \Delta S_{\text{app}} \) (cal/mol.K) |
|--------------|---------------|---------------------------------|---------------------------------|
| 0.5          | 41.8 ± 0.2    | 122 ± 12                        | 387 ± 38                        |
| 1.0          | 43.7 ± 0.2    | 104 ± 8                         | 331 ± 26                        |

A variation in the \( T_m \) of nsP4-CHIKV, as a function of the heating rate, was observed during the experiments. The dependence of the thermodynamic parameters on the heating rate was confirmed through the DSC technique, discussed in the next section. Regarding the values of enthalpy and entropy change, both agree with the values observed for globular proteins (26).

### Differential Scanning Calorimetry (DSC)

The thermal behavior of nsP4-CHIKV was also evaluated using the calorimetric method DSC (27). The DSC experiments were performed at different protein concentrations and different heating rates. Figure 1D shows the temperature-dependence of the heat capacity profile \((C_p - \text{sample minus reference})\) of nsP4-CHIKV along with the instrumental buffer baseline, acquired with a scan rate of 10 °C/h. As a globular protein, the thermogram comprises three events: the pre-transition baseline, corresponding to the heat capacity of the native state \((C_{\text{pn}})\), the endothermic unfolding transition with a melting temperature \((T_m)\) at approximately the maximum value of the \( C_p \) curve, and the post-transition baseline, representing the \( C_p \) of the denatured state \((C_{\text{pd}})\).

DSC experiments were performed at different scan rates to test its effect on the DSC profile and the reversibility of the transitions. The temperature- and scan rate-dependence of the excess heat capacity
profile of nsP4-CHIKV are illustrated in Figure 1E. The nsP4-CHIKV undergoes an irreversible thermal denaturation since the second heating scans obtained at all scan rates used do not show any signal. Moreover, the transition peak shows a clear scan rate dependence, characteristic of kinetically-dependent irreversible processes (28, 29). Indeed, all thermodynamic parameters associated with nsP4-CHIKV thermal denaturation exhibit a scan rate dependence (Table 2). The dependences of the thermodynamic parameters on the heating rate are markedly non-linear (Supplementary Figure 2). This feature illustrates the non-equilibrium character of its denaturation processes. The dependence on the heating scan might also suggest that the rates of the conformational transitions are rather low compared with the heating rates.

Table 2

**Thermodynamic parameters associated with the nsP4-CHIKV thermal denaturation by DSC.** The melting or unfolding temperature, $T_m$, represents the temperature where the excess heat capacity curve reaches its maximum value. The calorimetric enthalpy change, $\Delta H_{cal}$, was calculated as the area under the DSC trace. $\Delta T_{1/2}$ corresponds to the linewidth at half the height of the transition peak. The entropy change of the unfolding transition at $T_m$, $\Delta S$, was calculated as $\Delta S = \Delta H_{cal} / T_m$. The van't Hoff enthalpy change, $\Delta H_{vH}$, was calculated as $4RT_m^2C_{p,max}/\Delta H_{cal}$. Analyses of the thermograms were performed with MicroCal Origin software.

| Concentration (µM) | Rate (°C/h) | $T_m$ (°C) | $\Delta H_{cal}$ (kcal/mol) | $\Delta T_{1/2}$ (°C) | $\Delta S$ (cal/mol) | $\Delta H_{vH}$ (kcal/mol) | $\Delta H_{vH}/\Delta H_{cal}$ |
|-------------------|-------------|------------|-----------------------------|----------------------|-------------------|-----------------------------|-----------------------------|
| 9.3               | 8           | 36.3       | 205                         | 4.8                  | 662               | 142                         | 0.69                         |
| 11.9              | 13          | 37.0       | 102                         | 4.4                  | 330               | 151                         | 1.48                         |
| 7.4               | 33          | 38.8       | 86                          | 4.3                  | 275               | 165                         | 1.92                         |
| 7.4               | 64          | 39.8       | 66                          | 4.5                  | 210               | 150                         | 2.28                         |

Uncertainties: $T_m$ (± 0.2 °C), $\Delta H_{cal}$ (± 1–3 kcal/mol), $\Delta T_{1/2}$ (± 0.2 °C)

The dependence of the $T_m$ on the scan rate was used to calculate the kinetic activation energy, $E_a$, for the irreversible nsP4-CHIKV thermal denaturation. According to Sanchez-Ruiz et al. (29) the $T_m$ shifts induced by different heating scan rates, $v$, for an irreversible two-state process can be modeled by the following equation:
where \( A \) is the pre-exponential factor in the Arrhenius equation and \( R \) is the gas constant. Thus, by plotting \( \ln(\nu/T_m^2) \) against \( 1/T_m \), the apparent activation energy can be determined from the slope of the curve. The Arrhenius plot showing the scan rate-dependent changes in the \( T_m \) is illustrated in Figure 1F, from which \( E_a \) was determined as (110 ± 4) kcal/mol, similar to other proteins (30).

**Evaluation of nsp4-CHIKV interaction with compounds**

In the search for new compounds able to interact and inhibit the nsp4-CHIKV in solution, an initial experimental screening with a series of 12 compounds was performed using DSF (or ThermoFluor). The DSF can detect changes in the apparent \( T_m \) of a protein in the presence and absence of ligands, resulting in an estimate of the thermal shift (\( \Delta T_m \)) derived from these differences. This shift can be used as a primary indicator of complex formation (31).

For nsp4-CHIKV, the \( T_m \) in the absence of compounds (only with DMSO) was 37.7 ± 0.4. Among the compounds, one called LabMol-309 (Figure 2A) caused the highest thermal shift (~4°C). Thus, this data suggests the occurrence of interaction between LabMol-309 and nsp4-CHIKV and this compound was chosen to proceed with the other assays.

The interaction between LabMol-309 and nsp4-CHIKV was later analyzed by MST and solution-NMR. The MST data showed an affinity curve with the occurrence of well-defined bound and unbound states (Figure 2B). From that, the estimated dissociation constant (\( K_D \pm \Delta K_D \)) for the interaction of nsp4-CHIKV with LabMol-309 was (6 ± 1) µM.

The interaction of LabMol-309 with nsp4-CHIKV was further evaluated by solution-NMR, monitoring the chemical shift perturbation (CSP). Figure 2C shows the spectra obtained for the compound LabMol-309 in the presence (in red) and absence of the protein (in blue). These spectra were superimposed, and the chemical shift differences were identified and mapped according to the respective positions of the proton resonances (Figure 2C) previously identified in LabMol-309 assignment (Supplementary Figure 3).

Therefore, the detection of these chemical shifts perturbations is additional evidence for the interaction between nsp4-CHIKV and compound LabMol-309, reinforcing the results obtained through DSF and MST experiments.
The nsP4-CHIKV three-dimensional model and structural analysis

A combined analysis of several biophysical techniques suggested that the nsP4-CHIKV is a monomeric α-helical rich protein capable of forming a complex with the compound LabMol-309 within a moderate dissociation constant (10^{-6} M). The nsP4-CHIKV 3D structural model was obtained using AlphaFold2 (32). The model obtained showed a very high per-residue confidence score (pLDDT) for more than 90% of the covered sequence, and only for the N and C-terminal regions pLDDT were low (Figure 3A). The full-length nsP4-CHIKV structural model is illustrated in Figure 3.

In the nsP4-CHIKV structure, it was possible to identify the regions corresponding to the fingers (residues 93-315), palm (residues 316-502) and thumb (residues 503-611) domains, which are characteristics of viral polymerases (Figure 3B). In the catalytic site, located in the palm domain, it is possible to identify the catalytic aspartic acid dyad (Asp371 and Asp466) separated by 7.5 Å. Besides, it is also possible to notice that nsP4-CHIKV contains an extra N-terminal domain (residues 1-92) forming a coiled coil. A similar motif is observed in the N-terminal region of the nsp8 from the SARS-CoV-2 RdRp complex, likely responsible for proper positioning of the nascent strand of RNA during polymerization.

In the model of nsP4-CHIKV, it is also possible to notice that the electrostatic surface potential surface of the active site cavity is remarkably positive, a signature feature of nucleic acids-interacting motifs (Figure 3C). More than that, the analysis using ConSurf reveals that the catalytic site region and its surroundings are highly conserved (Figure 3D). All this serves as corroboration for the robustness of the AlphaFold2 model of nsP4-CHIKV.

Molecular docking and molecular dynamics of LabMol-309 against nsP4-CHIKV

To investigate the binding mode of LabMol-309 at the nsP4-CHIKV, docking calculations were performed using the DockThor server. The docking results suggest that LabMol-309 binds to the nsP4-CHIKV active site, interacting with residues of the GDD catalytic triad (Asp466 and Asp467), with a docking score of -7.15 Kcal/mol. LabMol-309 makes H-bonds with Glu369, Asp466, Asp467, Gly507, Arg573 and cation-π interactions with Lys295 residue (Figure 4). The nitrogen atom of the indole group and amine of the pyridine group of LabMol-309 make relevant interactions with Asp466 and Asp467, respectively. Additionally, the indole group makes cation-π interaction with Lys295.

The structural stability of the structural model of the nsP4-CHIKV/LabMol-309 complex calculated by docking was evaluated using 100 ns molecular dynamic (MD) simulation. Figure 5A presents the values of root mean square deviation (RMSD) for the backbone atoms of the protein and non-hydrogen atoms of the ligand from the initial structure. It is possible to observe that RMSD values are stable after 5 ns of simulation and reach plateaus around 0.3 and 0.5 nm for the protein and ligand, respectively. Figure 5B
shows that the number of contacts < 0.6 nm between nsP4-CHIKV and LabMol-309 do not drop down to zero throughout the MD simulation, indicating the ligand interacts with the protein all the time. The number of hydrogen bonds is stable throughout the MD simulations and presents an average value of three (Figure 5C). An evaluation of the hydrogen bonds with significant percentage occupancy (< 5%, Supplementary Table 2) during the MD trajectory reveals that Glu369, Asp371, Asp466, Asp467, Asn468, Lys501, and Arg573 are important for the stabilization of the protein-ligand complex. It is worth noting that Glu369 and Asp371 presented percentage occupancies higher than 70%. Taking together all MD analyze, it can conclude that the structural model of the nsP4-CHIKV/LabMol-309 complex is stable throughout the simulation.

### Inhibition of CHIKV replication by LabMol-039 through replication-based and viral infection assays

Although the data involving the complex formation between nsP4-CHIKV and LabMol-309 are solid, it still not possible to conclude that this compound would have inhibitory activity against this enzyme, or even against CHIKV-infection. The biochemical characterization of Alphavirus nsP4 polymerase activity has also been challenging because this protein cannot perform its function on its own (33). Different parts of nsP4 recognize the promoters for minus and plus strands. However, the binding requires the presence of the other non-structural proteins (nsP1, nsP2 e and nsP3) to form the replication complex and enable the de novo RNA synthesis (14, 33, 34). The interactions between these proteins within the complex during replication are limited and mainly uncharacterized (14). The main reason for this lack of information is that no high-resolution structure is available for the polymerase or the entire replication complex (33).

Given these limitations, it was not possible to evaluate the effect of the interaction with LabMol-309 directly on the enzymatic activity of purified nsP4-CHIKV. Therefore, the inhibitory activity of the compound LabMol-039 was evaluated through the replicon-based screenings in a dose-dependent manner to determine its effective and cytotoxic concentrations (EC$_{50}$ and CC$_{50}$, respectively). Replicon cells were incubated with 2-fold serial dilutions of compound (from 20 to 0.03 µM for EC$_{50}$ and from 100 to 0.30 µM for CC$_{50}$) and luciferase signals or cell viability was evaluated after 48 h. LabMol-039 displayed an EC$_{50}$ value of 10.0 ± 0.07 µM and showed high toxicity to the cells with a CC$_{50}$ of 17.1 ± 0.6 µM (Figure 6). As a result, the obtained selectivity index (CC$_{50}$/EC$_{50}$) of LabMol-039 was 1.7.

To confirm the antiviral activity of the LabMol-309, HEK-293T cells were infected with CHIKV-nanoluciferase, a recombinant CHIKV that express the Nanoluciferase reporter, at a multiplicity of infection (MOI) of 0.1 in the presence or absence of the compound at 5 µM. Nanoluciferase activity levels, which is proportional to viral replication, were assessed 16 hours post-infection (adapted from (35, 36)).
Viability assays were performed in parallel and DMSO 0.1% was used as untreated control (Figure 6C). As a result, these assays demonstrated that the LabMol-309 decreased CHIKV replication by 73.5% with cell viability of 79.4% (Table 4; Figure 6C).

| Compound        | Cell Viability (%) | CHIKV Replication (%) | CHIKV Replication Inhibition (%) |
|-----------------|--------------------|-----------------------|----------------------------------|
| LabMol-309 5 µM | 79.4               | 26.5                  | 73.5                             |
| DMSO 0.1%       | 100                | 100                   | 0                                |

Altogether, the results obtained from the antiviral assays suggest that LabMol-309 is a potential molecule to be further evaluated as an inhibitor of CHIKV replication, even though it showed high toxicity to the replicon cells. This isolated effect observed only in the replicon cells can be explained by the differences in the cell lines (BHK-21 versus 293T) and incubation periods used in the antiviral activity experiments (48 h for replicon-based screenings compared to 16 h for the viral infection assays). Nevertheless, LabMol-309 could be further optimized to reduce its cytotoxicity and increase the selectivity index in cell-based antiviral assays.

**Conclusions**

Given the absence of specific treatments for CHIKV infection, it is important to study, at the molecular level, the viral replication proteins and their use as targets for antivirals development. In this work, the polymerase nsP4 CHIKV was characterized through experimental and computational biophysical methods, generating unprecedented structural and dynamic information about this protein. A high prevalence of ordered helical secondary structures was observed by circular dichroism, which also showed that the nsP4-CHIKV unfolds under a cooperative transition during thermal denaturation. The thermal denaturation was further studied using differential scanning calorimetry, indicating a kinetically controlled process.

Considering nsP4-CHIKV as a promising target, a screening of compounds was performed through the DSF technique, and a compound designated LabMol-309 was identified. This biomolecular interaction was validated through MST, resulting in an interaction at a low micromolar range. The NMR was also used to confirm the interaction of LabMol-309 and nsP4-CHIKV, demonstrated the occurrence of chemical shifts perturbations and in silico molecular dockings and molecular dynamics simulations suggested the possible binding mode of this compound and its maintenance at the active site. Finally, the inhibitory activity of LabMol-309 was evaluated in cellular-based antiviral assays using a CHIKV replicon system and a reporter virus, demonstrating the inhibitory potential of this compound.
In summary, this study highlights biophysical features of nsP4-CHIKV, contributing to basic research on alphaviruses polymerase, and identified a new compound as a promising antiviral agent against CHIKV infection. These findings could contribute to development of novel candidates targeting nsP4-CHIKV and support the progress in the field of therapeutic strategies for CHIKV and other alphavirus infections.

Methods

CHIKV nsP4 cloning and overexpression

The coding region of nsP4-CHIKV (GenBank KP164572.1; PROT-ID AJY53709.1 – residues 118 to 611) was subcloned in the pET-SUMO vector using the LIC methodology (37). This construct encodes an nsP4-CHIKV with a N-terminal 6xHis-tag followed by a TEV protease cleavage site (ENLYFQ;GAM) and the fusion protein tag SUMO. For protein expression, this plasmid construction was transformed into *E. coli* Rosetta and cultured in Terrific Broth (TB) medium at 37°C and 200 RPM until an OD$_{600}$ of 1.0. The protein expression was induced with 1 mM of isopropyl-$\beta$-thiogalactoside and incubation at 18°C, 200 RPM for 16 h. The cells were harvested by centrifugation at 5000 x g for 30 min at 4°C and resuspended in buffer A (50 mM Tris pH 8.0, 500 mM NaCl, 10% glycerol). Cells were disrupted by sonication and clarified by centrifugation at 12,000 x g for 30 min at 4°C.

nsP4-CHIKV purification

nsP4-CHIKV was purified using an AKTA Purifier System (GE Healthcare). The first step was affinity chromatography, using a HisTrap HP 5.0 mL column (GE Healthcare) pre-equilibrated with buffer A (50 mM Tris pH 8.0, 500 mM NaCl, 10% glycerol). The elution was performed using 50 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole, 10% glycerol. The buffer was exchanged through dialysis to eliminate the imidazole excess. The 6xHis-tag-SUMO was cleaved by TEV protease during overnight incubation at 4°C. A second affinity chromatography step was performed using the same system to collect the HisTag-less protein obtained after TEV treatment. A final purification step was done using size-exclusion chromatography on an XK 26/1000 Superdex 75 column (GE Healthcare) pre-equilibrated in gel filtration buffer (50 mM Tris pH 8.0, 200 mM NaCl and 5% glycerol). The eluted fractions were collected and analyzed by SDS-PAGE to confirm their purity and mass spectrometry was performed to confirm the presence of nsP4-CHIKV. The final protein sample was concentrated using 30 kDa MWCO centrifugal concentrators (Vivaspin, Sartorius). Protein concentrations were determined spectrophotometrically in a Nanodrop 1000 spectrophotometer, using the measured absorbances at 280 nm and the theoretical extinction coefficient of 36,495 M$^{-1}$cm$^{-1}$.

Size Exclusion Chromatography coupled with Multi-angle Light Scattering (SEC-MALS)

The oligomeric state of the purified nsP4-CHIKV was evaluated by size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) in running buffer composed of 50 mM Tris-HCl pH 8.0 and 200 mM NaCl. For that, 50 μL of purified nsP4-CHIKV at a concentration of 1.5 mg/mL was
injected in a Waters 600 HPLC system (Waters) coupled in-line with a UV detector, a mini DAWN TREOS multi-angle light scattering apparatus (Wyatt Technology), a column Superdex 75 Increase 10/300 GL (GE Healthcare), and a refractive index detector Optilab T-rEX (Wyatt Technology). The light scattering detectors were normalized with bovine serum albumin (Sigma-Aldrich) and the flow rate used was 0.5 mL/min. The data were processed using ASTRA7 software (Wyatt Technology) with the following parameters: refractive index of 1.331, 0.890 cP for the viscosity of the solvent, and a refractive index increment of 0.1850 mL/g. Protein solutions were centrifuged for 10 minutes at 10,000 x g at a controlled temperature of 4°C immediately before use.

**Circular Dichroism (CD)**

Far UV-CD spectra (195-280 nm) were measured in a Jasco J-810 spectrometer (Jasco Corporation, Japan) equipped with a Peltier control system and using a quartz cell with a 1 mm pathlength. The spectra were recorded from 280 to 195 nm, with a scanning speed of 100 nm/min, a spectral bandwidth of 1 nm and a response time of 0.5 s. All the protein samples were in a final concentration of 2.5 µM diluted in water. Spectral deconvolution was applied to estimate the secondary structure content using the DICHROWEB web server (23). Three different methods were combined with three different databases to improve the reliability of the results. The detailed analysis of the results generated by these combinations is provided in Supplementary material (Supplementary Table 1). The estimated values of secondary structure fractions were averaged from each database used. The best fit was determined from the analysis of the NRMSD parameter, which was considered satisfactory when closer to 0 (38). Thermal denaturation experiments were performed by monitoring the ellipticity at 222 nm in the range from 20 to 80°C using heating rates of 0.5°C/min and 1.0°C/min.

**Differential Scanning Calorimetry (DSC)**

DSC measurements were carried out with the purified protein solution at 7.4 µM, 9.3 µM and 12 µM, diluted in buffer 50 mM Tris-HCl (pH 8.0), 200 mM NaCl and 5% glycerol. Protein and reference samples (buffer) were degassed for 5 min prior to measurements. The experiments were performed on a VP-DSC MicroCal MicroCalorimeter (Microcal, Northampton, MA, USA) using heating rates of 8°C/h, 13°C/h, 33°C/h and 64°C/h. The thermograms were recorded from 10 to 70°C, at a controlled pressure of 1.6 atm. Instrumental buffer baselines were recorded before the protein unfolding experiments to register the thermal history of the calorimeter. The raw DSC traces were subtracted with the buffer baseline and then normalized by protein concentration. The thermogram analysis, as well as the subtraction of the buffer calorimetric response, baseline correction and integration of the calorimetric peaks referring to the phase transitions, were performed using the MicroCal Origin software.

**Differential Scanning Fluorimetry (DSF)**

In the search for new binders to nsp4-CHIKV, a series of 12 compounds from the OpenZika project were tested (39, 40). The compounds were purchased from Chembridge Library (https://www.chembridge.com/) with minimum purity of 90%. For an initial assessment, DSF assays were conducted in a qPCR system Mx3000P (Agilent). Protein melting temperatures ($T_m$), assuming a
two-state transition model, were determined by monitoring the fluorescence intensity variation as a function of temperature for the extrinsic probe SYPRO Orange (Invitrogen). The protein solutions were at a final concentration of 8 µM, diluted in gel filtration buffer. Compounds were added at the final concentration of 80 µM and standard samples were prepared only with the addition of DMSO. The thermal variations were in the range 25-75°C in a stepwise increment of 1°C/min. The $T_m$ values obtained for each compound were subtracted from the values of the standard samples to identify compounds that caused significant $T_m$ changes. For the next steps, the compound that presented highest thermal shifts ($\Delta T_m$) was selected, considering the deviations of the triplicates (41). Data were analysed using the software Origin Pro 9.5.1. All experiments were conducted in triplicate.

**MicroScale Thermophoresis (MST)**

Experiments were performed on a Monolith® NT.115 instrument (Nanotemper technologies). The nsP4-CHIKV was labelled on cysteine residues with NT-647-Maleimide dye (Nanotemper Technologies) using the Monolith NTTM Protein Labeling Kit RED-MALEIMIDE as per manufacturer’s instructions. Samples of 25 nM of cys-labelled nsP4-CHIKV with 5% DMSO were used. An initial binding test was carried out with compound at the concentration of 100 µM, to check the occurrence of interaction between the protein and the compound. Then, a serial dilution of the compound from 200 µM to 0.012 µM (12 nM) was performed to obtain the binding curve. The dissociation constant ($K_D$) was obtained by fitting the binding curve with the Hill function, using GraphPad Prism 8 (Graph Pad Software).

**Chemical shift perturbation**

The LabMol-309 resonance assignment was performed using a Bruker Avance III 600 MHz. 1H-13C-HSQC, COSY and TOCSY were acquired at 298 K using 1 mM of LabMol-309 in D$_2$O. The interaction between LabMol-309 and nsP4-CHIKV was studied using a Bruker Avance IIIHD 500 MHz in a solution of 20 mM (2H)$_{11}$-Tris/HCl pH 7.5, 200 mM NaCl and 250 µM of LabMol-309. One-dimensional $^1$H spectra in the presence and absence of 3 µM NSP4-CHIKV were acquired, and the chemical shift difference was analyzed. The data processing and analysis were performed using TopSpin 4.09.

**The nsP4-CHIKV tridimensional model and structural analysis**

The nsP4-CHIKV sequence (residues 1 to 611) was used to generate the 3D model by AlphaFold2, developed by DeepMind (https://alphafold.ebi.ac.uk/) (32). For docking calculations, the nsP4-CHIKV model was structurally refined at GalaxyRefine server (42). Surface charge was calculated using APBS (43) and residues conservation was analyzed with ConSurf, following the default parameters (44). Pymol (45) was used to render the 3D images.

**Molecular docking of nsP4-CHIKV and LabMol-309**

The docking calculations were performed using the DockThor VS web (46, 47), focusing on the active binding site (Asp371 and Asp466 residues). The nsP4-CHIKV and LabMol-309 structures were prepared
using the Protein Preparation Wizard tool (48) and LigPrep tool (49). The docking grid was centered at the active binding site; grid size 20 Å; and grid coordinates x, y and z of -27.84 Å, 12.89 Å and 28.25 Å, respectively. The search algorithm precision mode was set up in the standard configuration of genetic algorithm parameters, with the soft docking mode activated. The PLIP server (50) was used to analyze the protein-ligand patterns (hydrogen bonds, hydrophobic interaction, cation-π, π-stacking, water and salt bridge interactions). Poseview server (51, 52) was used to generate 2D interaction diagram and VMD program was used to render the 3D images (53).

**Molecular dynamic simulations**

The initial positions of the nsP4-CHIKV-bound LabMol-309 for the molecular dynamic (MD) simulations were obtained by the molecular docking results and its topology parameterizations (Molid 814093) were obtained from the ATB server (54). The MD simulations were performed by using GROMACS package version 5.0.7 (55). The molecular system of the protein-ligand complex was modeled with the GROMOS54A7 force field (56) and SPC water model (57), using a cubic box solvated with 200 mM NaCl. The simulation was realized in ensemble NPT at 25°C and 1.0 bar using modified Berendsen thermostat with $\tau_T = 0.1$ ps and Parrinello-Rahman barostat with $\tau_P = 2.0$ ps and compressibility = $4.5 \times 10^{-5}$·bar$^{-1}$. A cutoff value of 12 Å was used for both Lennard-Jones and Coulomb potentials and long-range electrostatic interactions were calculated using the Particle Mesh Ewald algorithm (PME) (58). Energy minimization was performed with a steepest descent integrator and conjugate gradient algorithm, using 1000 kJ·mol$^{-1}$·nm$^{-1}$ as the maximum force criterion. One hundred thousand steps of molecular dynamics were performed for each NVT and NPT equilibration, applying force constants of 1000 kJ·mol$^{-1}$·nm$^{-2}$ to all heavy atoms of the protein-ligand complex. At the end of preparation, 100 ns MD simulation of the structural model of the protein-ligand complex was carried out for data acquisition. Next, the trajectory was aligned and analyzed according: RMSD of backbone atoms for protein and nonhydrogen atoms for the ligand, number of hydrogen bounds (cutoff distance of 3.5 Å and maximum angle of 30°) between protein and ligand, and number of contacts < 0.6 nm between all atoms of the protein and of the ligand.

**Cells and Virus**

HEK-293T cells were purchased from The Global Bioresource Center (ATCC) and maintained in Dulbecco’s modified Eagle’s medium (DMEM, Sigma-Aldrich) supplemented with 100U/mL of penicillin (Hyclone Laboratories), 100 mg/mL of streptomycin (Hyclone Laboratories), 1% dilution of stock of non-essential amino acids (Hyclone Laboratories) and 10% of fetal bovine serum (FBS, Hyclone Laboratories) in a humidified 5% CO2 incubator at 37°C. BHK-21-Gluc-nSP-CHIKV-99659 cell line, harboring a replicative CHIKV replicon expressing *Gaussia* luciferase (Gluc) as a reporter gene, was maintained in DMEM 10% FBS with 500 µg/ml G418 (Sigma-Aldrich). The development of this replicon cell line will be described elsewhere. The CHIKV expressing *nanoluciferase* reporter (CHIKV-nanoluc) used for the antiviral assays is based on the CHIKV isolate LR2006OPY1 (East/Central/South African genotype) and was produced, rescued, and titrated as previously described (35, 36).
CHIKV replicon-based screenings

LabMol-309 at 200 mM in 100% DMSO was diluted with assay media to a final concentration of 1% (v/v) DMSO and was evaluated in a dose-dependent manner to determine its effectiveness (EC$_{50}$) and cytotoxic (CC$_{50}$) concentrations, as described in (59). Approximately 2 x 10$^4$ replicon cells/well in DMEM 10% FBS were seeded in a 96-well plate. After 16 h of incubation at 37°C 5% CO$_2$, the medium was replaced with fresh DMEM supplemented with 2% FBS and compound was added to the cells at 2-fold serial dilutions. After a 48 h-incubation, 40 µL of the cells’ supernatant containing secreted Gluc were mixed with 50 µl of Renilla luciferase Assay Reagent (Promega). The Gluc activity was measured using SpectraMax i3 Multi-mode Detection Platform (Molecular Devices). Replicon cells in 1% DMSO were used as negative control (0% inhibition). The compound concentration required to inhibit 50% of the Gluc activity (EC$_{50}$) was estimated using the OriginPro 9.0 software. The cytotoxicity was evaluated through a cell proliferation-based MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay, as described in (60). The compound concentration required to cause 50% cytotoxicity (CC$_{50}$) was estimated using the OriginPro 9.0 software. The dose-response curves were performed twice in duplicates. The EC$_{50}$ and CC$_{50}$ values were used to determine the selectivity index (SI = CC$_{50}$/EC$_{50}$) of the compound.

Infection assays using CHIKV-nanoluc

HEK-293T cells were seeded at a density of 1 x 10$^5$ cells per well into 48-well plates for 24 h and infected with CHIKV-nanoluc at a multiplicity of infection (MOI) of 0.1 PFU/cell in the presence or absence of LabMol-309 at 5 µM. Samples were harvested using Renilla-luciferase lysis buffer (Promega®) at 16 hours post-infection (h.p.i.), and the levels of virus replication were quantified by measuring nanoluciferase activity using the Renilla luciferase Assay System (Promega®). Data were analyzed for normal distribution to demonstrate the applicability of the parametric test. Then, a Two-way ANOVA test was employed to compare the treatment of LabMol-309 with the DMSO 0.1%, with a p < 0.05 (adapted from (35, 36)). The cell viability was measured by MTT assay, as described previously (35, 36). Cell viability was calculated according to the equation (T/C) × 100%, where T and C represent the mean optical density of the treated (cells incubated with LabMol-309 at 5 µM) and untreated (cells in DMSO 0.1%) control groups, respectively. The results were calculated using GraphPad Prism 8.0.0 for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com).

Declarations

Acknowledgments

This project was supported by the São Paulo Research Foundation – FAPESP (grants to G.O, 2013/07600-3, M.C.L.C.F 2018/17095-8, R.S.F. 2018/05130-3 and L.R.P. 2021/01706-0). We thank DeepMind Technologies Ltd for generating the AlphaFold2 ab initio prediction of the CHIKV nsP4 structure before public availability of their platform.
Author contributions

M.C.L.C.F, L.G.M.B, L.F.S.M and G.O. contributed to the ideation and conceptualization of the project. M.C.L.C.F, L.G.M.B, L.F.S.M, F.C.L.A. and N.C.M.R.M. performed the biochemical, biophysical and NMR experiments and analysis. C.H.A., M.M. and B.K.P.S provided the compounds. A.S.G performed structural analysis of 3D model. M.M performed the molecular docking and I.P.C. performed molecular dynamics simulations. L.H.V.G.G. provided the replicon system. R.S.F. and L.R.P. performed the cell-based replicon assays and analysis. I.A.S, U.E.A.R. and A.C.G.J. performed the viral assays and analysis. M.C.L.C.F., L.F.S.M., L.G.M.B., F.C.L.A, A.S.G., C.H.A. and G.O. contributed to the editing and revision of the manuscript.

Competing interests

The authors declare no competing interests

References

1. Chastel C. Infections inapparentes chez l’Homme: Un cheval de Troie pour l’introduction et la diffusion des arbovirus transmis par des moustiques dans les régions non endémiques? Bull la Soc Pathol Exot. 2011 Aug;104(3):213–9.

2. Weaver SC, Lecuit M. Chikungunya virus and the global spread of a mosquito-borne disease. N Engl J Med. 2015;372(13):1231–9.

3. Silva JVJ, Ludwig-Begall LF, Oliveira-Filho EF d., Oliveira RAS, Durães-Carvalho R, Lopes TRR, et al. A scoping review of Chikungunya virus infection: epidemiology, clinical characteristics, viral co-circulation complications, and control [Internet]. Vol. 188, Acta Tropica. Elsevier B.V.; 2018 [cited 2021 Jun 17]. p. 213–24. Available from: /pmc/articles/PMC7092809/

4. Thiberville S-D, Moyen N, Dupuis-Maguiraga L, Nougairede A, Gould EA, Roques P, et al. Chikungunya fever: Epidemiology, clinical syndrome, pathogenesis and therapy. Antiviral Res [Internet]. 2013 Sep 1 [cited 2019 Jul 8];99(3):345–70. Available from: https://www.sciencedirect.com/science/article/pii/S0166354213001666?via%3Dihub

5. Vairo F, Haider N, Kock R, Ntoumi F, Ippolito G, Zumla A. Chikungunya: Epidemiology, Pathogenesis, Clinical Features, Management, and Prevention. Vol. 33, Infectious Disease Clinics of North America.
6. Thiberville S-D, Boisson V, Gaudart J, Simon F, Flahault A, de Lamballerie X. Chikungunya Fever: A Clinical and Virological Investigation of Outpatients on Reunion Island, South-West Indian Ocean. Tesh RB, editor. PLoS Negl Trop Dis [Internet]. 2013 Jan 17 [cited 2019 Nov 21];7(1):e2004. Available from: https://dx.plos.org/10.1371/journal.pntd.0002004

7. Simon F, Javelle E, Cabie A, Bouquillard E, Troisgros O, Gentile G, et al. French guidelines for the management of chikungunya (acute and persistent presentations). November 2014. Med Mal Infect. 2015 Jul;45(7):243–63.

8. Marimoutou C, Vivier E, Oliver M, Boutin JP, Simon F. Morbidity and impaired quality of life 30 months after chikungunya infection: Comparative cohort of infected and uninfected french military policemen in reunion island. Med (United States). 2012 Jul;91(4):212–9.

9. Rausalu K, Utt A, Quirin T, Varghese FS, Žusinaite E, Das PK, et al. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. Sci Rep. 2016 Nov 15;6.

10. Voss JE, Vaney MC, Duquerroy S, Vonrhein C, Girard-Blanc C, Crublet E, et al. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. Nature. 2010 Dec 2;468(7324):709–12.

11. Melton J V., Ewart GD, Weir RC, Board PG, Lee E, Gage PW. Alphavirus 6K proteins form ion channels. J Biol Chem. 2002 Dec 6;277(49):46923–31.

12. Powers AM. Vaccine and therapeutic options to control chikungunya virus. Vol. 31, Clinical Microbiology Reviews. American Society for Microbiology; 2018.

13. Lum FM, Ng LFP. Cellular and molecular mechanisms of chikungunya pathogenesis. Vol. 120, Antiviral Research. Elsevier B.V.; 2015. p. 165–74.

14. Rupp JC, Sokoloski KJ, Gebhart NN, Hardy RW. Alphavirus RNA synthesis and non-structural protein functions. J Gen Virol. 2015 Sep 1;96(9):2483–500.

15. Silva LA, Dermody TS. Chikungunya virus: Epidemiology, replication, disease mechanisms, and prospective intervention strategies [Internet]. Vol. 127, Journal of Clinical Investigation. American Society for Clinical Investigation; 2017 [cited 2021 Jun 21]. p. 737–49. Available from: https://pubmed.ncbi.nlm.nih.gov/28248203/

16. Cunha MS, Costa PAG, Correa IA, de Souza MRM, Calil PT, da Silva GPD, et al. Chikungunya Virus: An Emergent Arbovirus to the South American Continent and a Continuous Threat to the World [Internet]. Vol. 11, Frontiers in Microbiology. Frontiers Media S.A.; 2020 [cited 2021 Jun 17]. p. 1297. Available from: www.frontiersin.org

17. Rashad AA, Mahalingam S, Keller PA. Chikungunya virus: Emerging targets and new opportunities for medicinal chemistry. J Med Chem. 2014 Feb 27;57(4):1147–66.

18. Bakar FA, Ng LFP. Nonstructural proteins of alphavirus—potential targets for drug development. Vol. 10, Viruses. MDPI AG; 2018.
19. Thiboutot MM, Kannan S, Kawalekar OU, Shedlock DJ, Khan AS, Sarangan G, et al. Chikungunya: A Potentially Emerging Epidemic? Brooker S, editor. PLoS Negl Trop Dis [Internet]. 2010 Apr 27 [cited 2019 Nov 22];4(4):e623. Available from: https://dx.plos.org/10.1371/journal.pntd.0000623

20. Choi KH. Viral polymerases. Viral Mol Mach. 2012;267–304.

21. Some D, Amartely H, Tsadok A, Lebendiker M. Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). J Vis Exp. 2019;148:e59615.

22. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc [Internet]. 2007 Jan [cited 2020 Nov 12];1(6):2876–90. Available from: /pmc/articles/PMC2728378/?report=abstract

23. Whitmore L, Wallace BA. DICHRORWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res [Internet]. 2004 Jul 1 [cited 2020 Nov 11];32(WEB SERVER ISS.). Available from: https://pubmed.ncbi.nlm.nih.gov/15215473/

24. Johnson WC. Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins Struct Funct Bioinforma. 1999;35(3):307–12.

25. Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat Protoc. 2007 Jan;1(6):2527–35.

26. Khechinashvili NN, Janin J, Rodier F. Thermodynamics of the temperature-induced unfolding of globular proteins. Protein Sci. 1995;4(7):1315–24.

27. Durowoju IB, Bhandal KS, Hu J, Carpick B, Kirkitadze M. Differential scanning calorimetry — A method for assessing the thermal stability and conformation of protein antigen. J Vis Exp [Internet]. 2017 Mar 4 [cited 2020 Nov 15];2017(121):55262. Available from: /pmc/articles/PMC5409303/?report=abstract

28. Karantzeni I, Ruiz C, Liu C-C, LiCATA VJ. Comparative thermal denaturation of Thermus aquaticus and Escherichia coli type 1 DNA polymerases. Biochem J. 2003;374(3):785–92.

29. Sanchez-Ruiz JM, Lopez-Lacomba JL, Cortijo M, Mateo PL. Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. Biochemistry. 1988;27(5):1648–52.

30. Lyubarev A, Kurganov B. Analysis of DSC data relating to proteins undergoing irreversible thermal denaturation. J Therm Anal Calorim. 2000;62(1):51–62.

31. Scott DE, Spry C, Abell C. Differential scanning fluorimetry as part of a biophysical screening cascade. Fragm Drug Discov Lessons Outlook. 2016;139–72.

32. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9.

33. Pietilä MK, Hellström K, Ahola T. Alphavirus polymerase and RNA replication. Virus Res. 2017;234:44–57.

34. Rubach JK, Wasik BR, Rupp JC, Kuhn RJ, Hardy RW, Smith JL. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. Virology. 2009 Feb 5;384(1):201–8.
35. de Oliveira DM, de Andrade Santos I, Martins DOS, Gonçalves YG, Cardoso-Sousa L, Sabino-Silva R, et al. Organometallic complex strongly impairs chikungunya virus entry to the host cells. Front Microbiol. 2020;11.

36. Santos IA, Shimizu JF, de Oliveira DM, Martins DOS, Cardoso-Sousa L, Cintra ACO, et al. Chikungunya virus entry is strongly inhibited by phospholipase A2 isolated from the venom of Crotalus durissus terrificus. Sci Rep. 2021;11(1):1–12.

37. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 1990 Oct 25;18(20):6069–74.

38. Mao D, Wachter E, Wallace BA. Folding of the mitochondrial proton adenosine triphosphatase proteolipid channel in phospholipid vesicles. Biochemistry. 1982;21(20):4960–8.

39. Ekins S, Perryman AL, Horta Andrade C. OpenZika: an IBM world community grid project to accelerate Zika virus drug discovery. PLoS Negl Trop Dis. 2016;10(10):e0005023.

40. Mottin M, Borba JVVB, Braga RC, Torres PHM, Martini MC, Proenca-Modena JL, et al. The A–Z of Zika drug discovery. Drug Discov Today. 2018;23(11):1833–47.

41. Wang B, Shi W, Miao Z. Confidence analysis of standard deviational ellipse and its extension into higher dimensional Euclidean space. PLoS One. 2015;10(3):e0118537.

42. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012;40(W1):W294–7.

43. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. Protein Sci. 2018 Jan;27(1):112–28.

44. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 2016;44(W1):W344–50.

45. The PyMOL Molecular Graphics System, Version 1.8, Schrödinger L. The PyMOL Molecular Graphics System, Version 1.8. 2015.

46. Santos KB, Guedes IA, Karl ALM, Dardenne LE. Highly flexible ligand docking: benchmarking of the DockThor program on the LEADS-PEP protein–peptide data set. J Chem Inf Model. 2020;60(2):667–83.

47. Guedes IA, Barreto AMS, Marinho D, Krempser E, Kuenemann MA, Sperandio O, et al. New machine learning and physics-based scoring functions for drug discovery. Sci Rep. 2021 Dec;11(1):3198.

48. Schrödinger Release 2015-2: Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2015.

49. Schrödinger Release 2015-2: LigPrep, Schrödinger, LLC, New York, NY, 2015.

50. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein–ligand interaction profiler. Nucleic Acids Res. 2015 Jul;43(W1):W443–7.

51. Stierand K, Maaß PC, Rarey M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. Bioinformatics. 2006;22(14):1710–6.
52. Fricker PC, Gastreich M, Rarey M. Automated drawing of structural molecular formulas under constraints. J Chem Inf Comput Sci. 2004;44(3):1065–78.

53. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996 Feb 1;14(1):33–8.

54. Malde AK, Zuo L, Breeze M, Stroet M, Poger D, Nair PC, et al. An Automated force field Topology Builder (ATB) and repository: Version 1.0. J Chem Theory Comput [Internet]. 2011 Dec 13 [cited 2020 Sep 10];7(12):4026–37. Available from: https://pubs.acs.org/doi/abs/10.1021/ct200196m

55. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1–2:19–25.

56. Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur Biophys J. 2011;40(7):843–56.

57. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction Models for Water in Relation to Protein Hydration. In Springer, Dordrecht; 1981 [cited 2020 Sep 10]. p. 331–42. Available from: https://link.springer.com/chapter/10.1007/978-94-015-7658-1_21

58. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys. 1995;103(19):8577–93.

59. Fernandes RS, de Godoy AS, dos Santos IA, Noske GD, de Oliveira KIZ, Gawriljuk VO, et al. Discovery of an imidazonaphthyridine and a riminophenazine as potent anti-Zika virus agents through a replicon-based high-throughput screening. Virus Res. 2021 Apr;299:198388.

60. Li JQ, Deng CL, Gu D, Li X, Shi L, He J, et al. Development of a replicon cell line-based high throughput antiviral assay for screening inhibitors of Zika virus. Antivir Res. 2017/12/31. 2018;150:148–54.

Figures
Biophysical characterization of nsP4-CHIKV.  

A) SEC-MALS analysis of nsP4-CHIKV in solution. The proximity of the estimated molecular weight of (60 ± 1) kDa with the molecular weight of the protein suggests that nsP4-CHIKV is mostly at a monomeric state under the evaluated conditions. 

B) Secondary structure profile of nsP4-CHIKV by CD spectroscopy. The nsP4-CHIKV spectrum suggests the predominance of helical secondary structures. 

C) Thermal stability of nsP4-CHIKV probed by CD spectroscopy. The ellipticity at 222 nm as a function of temperature for nsP4-CHIKV was recorded at 0.5 °C/min (left) and 1.0 °C/min (right) and transformed to the protein unfolded fraction. Solid lines are best fits to the CD data using a two-state equilibrium model. The thermodynamic parameters of the protein unfolding transition are summarized in Table 1. 

D) Scan-rate normalized (10 °C/h) DSC data of the irreversible thermal denaturation of 12 µM of nsP4-CHIKV and the corresponding instrumental buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl e 5% glycerol) baseline. 

E) Excess heat capacity of nsP4-CHIKV at four different scan rates, as indicated, obtained after normalization by protein concentration and subtraction of the buffer baseline. 

F) Arrhenius-type plot showing the scan rate dependence of the nsP4-CHIKV unfolding temperature, $T_m$. The slope yields the activation energy for the irreversible denaturation of nsP4-CHIKV.
Figure 2

Evaluation of nsp4-CHIKV interaction with the compound LabMol-309. A) LabMol-309 chemical structure. B) Binding affinity curve of nsP4-CHIKV interacting with LabMol-309 by MST. The compound LabMol-309 was titrated in a concentration range of 200 µM to 0.012 µM. The curve was fitted to the Hill function and the estimated $K_D \pm \Delta K_D$ was (6 ± 1 µM). C) Chemical shift perturbation analysis based on the superimposition of LabMol-309 one-dimensional 1H spectra obtained in the presence (red line) and in and absence of nsP4-CHIKV (blue line). On the right, chemical shift differences were measured (in Hz) for each proton of the LabMol-309 structure.
Figure 3

nsP4-CHIKV three-dimensional model by AlphaFold2. A) nsP4-CHIKV AlphaFold2 model colored according to pLDDT. At the top right is shown the pLDDT color scales. B) Model of nsP4-CHIKV colored according to domain assignment. The N-terminal helixes, Finger, Catalytic and Thumb domains are colored in purple, dark blue, green and red, respectively. In the inset, a detailed view of predicted catalytic aspartic residues is presented. C) Electrostatic potential projected on the surface charge of nsP4-CHIKV, calculated with APBS. Positive regions are colored in blue and negative regions in red. D) ConSurf analysis of nsP4-CHIKV model. The amino-acids are colored by their conservation grades using the color-coding bar, with green-through-purple indicating variable-through-conserved.
Figure 4

Molecular interactions of LabMol-309 and nsP4-CHIKV, predicted by docking calculations. **A)** 3D interactions of LabMol-309 and nsP4-CHIKV residues. Hydrogen bonds are shown as gray dashed lines, cation-p interactions are shown in green dashed lines. Oxygen, nitrogen and hydrogen atoms are shown in red, blue and white, respectively. Carbon atoms of LabMol-309 and protein residues are shown in purple and gray, respectively. Catalytic residues of the GDD triad are highlighted in bold. **B)** 2D interaction diagram of LabMol-309 and nsP4-CHIKV residues. Hydrogen bonds are shown as gray dashed lines, cation-p interactions are shown in green dashed lines. Catalytic residues of the GDD triad are highlighted in blue.
Figure 5

Evaluation of stability of the structural model of the nsP4-CHIKV/LabMol-309 complex from 100 ns MD simulation. A) RMSDs of backbone atoms of the protein (black line) and non-hydrogen atoms of the ligand (gray line). B) Number of contacts between atoms of the protein and ligand for distances less than 0.6 nm. C) Number of hydrogen bonds formed between the protein and ligand.
Figure 6

Antiviral activity of LabMol-309. The EC$_{50}$ and CC$_{50}$ curves from replicon-based screenings are shown in (A) and (B), respectively. CHIKV replicon cells were incubated with the compound at 2-fold serial dilutions (from 20 µM to 0.03 µM for EC$_{50}$ and from 100 to 0.3 µM for CC$_{50}$) for 48 hours, and Gluc activity/cell viability were measured from cells’ supernatant. Representative results from two independent experiments performed in duplicates. Error bars represent the standard deviations. C) The effect of the compound on HEK-293T viability and CHIKV infection. HEK-293T cells were infected with CHIKV-nanoluciferase, and simultaneously treated with LabMol-309 at 5 µM. After 16 hours post-infection (h.p.i), cells were lysed, and a Renilla-luciferase assay was performed to assess CHIKV replication. Cell viability was accessed by treating HEK-293T cells with the compound at 5 µM. Mean values of three independent experiments, each measured in triplicate including the standard deviation, are shown. DMSO 0.1% was used as untreated control. P values < 0.05 were considered significant (*** P < 0.001, **** P < 0.0001).
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Papernsp4SuppMaterial.docx