Peptide inhibitors against SARS-CoV-2 2'-O-methyltransferase involved in RNA capping: A computational approach

Mainak Dutta a,*, Eldhose Iype b,**

a Department of Biotechnology, BITS Pilani, Dubai Campus, Academic City, Dubai, United Arab Emirates
b Department of Chemical Engineering, BITS Pilani, Dubai Campus, Academic City, Dubai, United Arab Emirates

ARTICLE INFO

Keywords:
SARS-CoV-2
Nsp16
Peptide inhibitors
Therapeutics
Molecular dynamics simulation

ABSTRACT

The COVID-19 pandemic is still evolving and is caused by SARS-CoV-2. The 2'-O-methyltransferase (nsp16) enzyme is crucial for maintaining the stability of viral RNA for effective translation of viral proteins and its life cycle. Another protein, nsp10, is important for enzymatic activity of nsp16. Any disturbance in the interaction between nsp16 and nsp10 may affect viral replication fidelity. Here, five peptide inhibitors, derived from nsp16, were designed and assessed for their effectivity in binding to nsp10 using molecular dynamics simulation. The inhibitors were derived from the nsp10/nsp16 binding interface. Post-simulation analysis showed that inhibitors 2 and 5 were stable and bind to the nsp16 interacting region of nsp10 which could potentially prevent the interaction between the two proteins. The proposed peptides are useful starting points for the development of therapeutics to manage the spread of COVID-19.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) caused the global outbreak of coronavirus disease 2019 (COVID-19). The novel virus belongs to the Coronaviridae family and Betacoronavirus genus [1]. Two other coronaviruses, SARS-CoV and MERS-CoV, also belong to the same genus which had caused two major worldwide outbreaks in the past two decades. COVID-19 was declared a pandemic by the World Health Organization on March 2020 [2]. As of June 2021 >170 million cumulative cases were reported globally with more than 3.7 million deaths [3].

The replication and transcription of the genome of coronaviruses is mediated by the replicase gene. It is increasingly becoming clear that the proteins encoded by the replicase gene of SARS-CoV and SARS-CoV-2 show high functional similarity. The gene expresses two polyproteins, pp1a and pp1ab, which are cleaved into several non-structural proteins (nsps). Two proteases, main protease, M^pro^ (nsp5) and papain-like protease, PR^pro^ (nsp3) cleave the polyproteins to form other nsps. Many of them such as RNA-dependent RNA polymerase (nsp12) [4], RNA helicase and triphosphatase (nsp13) [5], exoribonuclease and N7-methyltransferase (nsp14) [6], endonuclease (nsp15) [7], and 2'-O-methyltransferase (nsp16) [8] are key enzymes which form the replicase–transcriptase complex (RTC) or act as accessory proteins involved in viral RNA replication and sub-genomic RNA transcription [9].

One of the key mechanisms by which the SARS-CoV-2 RNA escapes the cellular innate immunity guardians is by capping its 5' end [10]. This helps it to mimic the host mRNA, thus maintaining its stability and integrity for effective translation. Since the virus replicates in the cytoplasm, it is unable to access the capping machinery of the host located in the nucleus. They have thus evolved to encode their own capping enzymes. Several nsps such as nsp14 and nsp16 are involved in viral RNA capping [11]. Nsp14 generates the cap-0 structure by methylating 5' guanine of the Gppp-RNA at the N7 position (m7Gppp-RNA) [6]. Nsp16 selectively 2'-O-methylates the ribose of the following nucleotide to convert the cap-0 to cap-1 structure (m7GpppN^m^-RNA) [12]. This 2'-O-methylation is critical and an indispensable step for replication of the coronavirus genome [13]. Nsp16 has a binding pocket for 5'-adenosyl-l-methionine (SAM) which acts as methyl group donor for the 2'-O-methylation reaction. Biochemical and crystallographic studies have revealed that another nsp, nsp10, stabilizes the SAM binding pocket and extends the substrate RNA binding groove by inducing conformational changes in nsp16 [8,14,15]. It is therefore clear that binding of nsp10 is important for the enzymatic activity of
Thus, the nsp10/nsp16 interface is considered as a drug target [15]. For past several months computational approaches have been used to screen for potential small molecule therapeutics against the nsp10/nsp16 complex of SARS-CoV-2. However, they target either the SAM binding site or the RNA groove [16]. In the current study we attempt to identify therapeutics which may prevent the binding of nsp10 with nsp16. However, since the nsp10-nsp16 interface is not clearly defined and extends over relatively large surface area, lack of small and defined binding site may prove challenging for design of small molecules targeting the interface. Peptide therapeutics are therefore, an alternative approach to inhibit such protein-protein interactions [17]. Moreover, peptide therapeutics are highly specific to the target proteins and show limited interference with normal human biological processes leading to faster FDA approval [18,19]. Henceforth, in the present study we have designed and simulated peptide inhibitors from the nsp10 interacting region of nsp16. The proposed peptides can be used as therapeutics to manage COVID-19.

Table 1

| Inhibitor | Sequence | Important Conformations |
|-----------|----------|-------------------------|
| 1         | (6876) VAPGTAVLRQWL (6887) | α4 helix |
| 2         | (6876) VAPGTAVLRQWLPTGTLLVDSLDNDFVSDADSTLIG (6911) | α4 helix, β3 and β4 strands |
| 3         | (6836) KGRIMMVAYTQLCQYLNTLTLAVPYMNVRHIHAGSDKGVAPGTAVLRQWLPTGTLLVDSLDNDFVSDADSTLIG (6911) | α3 and α4 helices, β2, β3 and β4 strands |
| 4         | Sequence 1: (6836) KGRIMMVAYTQLCQYLNTLTLAVPYMN (6862) | α3 and α4 helices, β3 and β4 strands |
| 5         | Sequence 1: (6836) KGRIMMVAYTQLCQYLNTLTLAVPYMN (6862) | α3 and α4 helices, β3 and β4 strands |

nsp16. Thus, the nsp10/nsp16 interface is considered as a drug target [15].

For past several months computational approaches have been used to screen for potential small molecule therapeutics against the nsp10/nsp16 complex of SARS-CoV-2. However, they target either the SAM binding site or the RNA groove [16]. In the current study we attempt to identify therapeutics which may prevent the binding of nsp10 with nsp16. However, since the nsp10-nsp16 interface is not clearly defined and extends over relatively large surface area, lack of small and defined binding site may prove challenging for design of small molecules targeting the interface. Peptide therapeutics are therefore, an alternative approach to inhibit such protein-protein interactions [17]. Moreover, peptide therapeutics are highly specific to the target proteins and show limited interference with normal human biological processes leading to faster FDA approval [18,19]. Henceforth, in the present study we have designed and simulated peptide inhibitors from the nsp10 interacting region of nsp16. The peptides were designed based on a recently published crystal structure of the nsp10/nsp16 complex (PDB: 6W4H). Some of the inhibitors were found to be structurally stable and bind strongly to nsp10 which could potentially prevent the latter’s interaction with nsp16. The proposed peptides can be used as therapeutics to manage COVID-19.
2. Methods

Design of peptide inhibitors: The crystal structure of nsp10/nsp16 complex (PDB: 6W4H) was used to design the peptide inhibitors. A detailed analysis was performed using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC to study the interacting regions of nsp10 and nsp16. Five inhibitors were designed from the nsp16 region which interacts with nsp10 (Details can be found in the Results and Discussion section).

Molecular dynamics simulation: Molecular dynamics simulation of protein (nsp10)-peptide inhibitor and control (nsp16) complexes were performed using GROMACS package [20] under physiological conditions with known initial position obtained from the crystal structure (PDB: 6W4H). Coordinates and topology files were generated using pdbs2gmx program. Charmm36 all atom force field [21] was used to generate parameters for the simulation. For water, TIP3P [22] model is used, and the system is neutralized using Na+ or Cl− ions. The amount of water molecules added to the system depends on the size of the simulation box, which depends on the size of the peptide-protein complex. A tolerance gap of 10 Å from the extreme ends of the complex to the box boundaries are provided on each side, filling them with water molecules corresponding to the liquid density. The initial amount of the ions is zero. In order to emulate physiological conditions, we have added Na+ or Cl− ions to the simulation box such that the whole system is charge neutral. The number of ions added for each peptide-protein complex are as follows: inhibitor 1: 1 Cl−, inhibitor 2: 4 Na+, inhibitor 3: No ions added, inhibitor 4: 1 Na+ , inhibitor 5: 2 Na+ and control: 3 Cl−. First, the system is minimized using steepest descent algorithm with a maximum 50000 steps. Next, an NVT simulation is performed at 300 K for 50000 steps, followed by an NPT simulation for the same number of steps. The non-bonded interactions were cut-off at 10 Å, and the long-range electrostatics were approximated using Particle Mesh Ewald method. Time step was chosen to be 2 fs throughout. Bonds involving H-atom were constrained using LINCS algorithm. The temperature coupling is done separately for protein + peptide and water + ions groups using modified Berendsen [23] thermostat. Pressure under NPT ensemble is maintained at 1 bar using Parrinello-Raham barostat. The production run was performed after equilibration for 150 ns for each of the complexes.

Post-simulation analysis-Interaction energy, binding energy, linear interaction energy, root mean square deviation of gyration:

Post-simulation analysis-Interaction energy, binding energy, linear interaction energy, root mean square deviation of gyration:

Interaction energies were computed between the peptides and nsp10. The ensemble averaged electrostatic and Lennard-Jones (LJ) interactions between the protein and peptide were computed by scanning through the coordinates written during the production run. The electrostatic and LJ interactions are calculated using the equations,

\[
U_{\text{elec}} = \sum_{\text{non bonded pairs}} \frac{q_i q_j}{r_{ij}^n}
\]

Here, \(q_i\) is the partial charge on atom i, \(r_{ij}\) is the distance between atoms i and j, \(ϵ\) is the dielectric constant of the medium, \(r_{ij}^{\text{min}}\) is the distance at which the LJ potential is at its minimum.

In addition, the binding free energies of peptides complexed with nsp10 (only last 25 ns) were computed using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) analysis using g_mmpbsa package [24]. For this, one frame at every nanosecond (ns) simulation starting from 125 to 150 ns were extracted and the binding free energies were calculated as the difference between free energy of the complex minus sum of the free energies of the peptide and protein. The solute and solvent dielectric constants were taken to be 4 and 80, respectively.

\[
\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - \left( \Delta G_{\text{peptide}} + \Delta G_{\text{protein}}(\text{nsp10}) \right)
\]

Here, \(\Delta G_{\text{bind}}\) is the binding free energy, \(\Delta G_{\text{complex}}\) is the free energy of the complex, \(\Delta G_{\text{peptide}}\) is the free energy of the peptide, and \(\Delta G_{\text{protein}}(\text{nsp10})\) is the free energy of nsp10.

To gain further insight into the peptide-protein interaction, we have performed the linear interaction energy (LIE) analysis using gmx lie method in GROMACS. For this, two sets of simulations were used: one with peptide and receptor (bound state), and second with peptide alone (unbound state). Then the van der Waals (LJ) and electrostatic interactions between the peptides and surroundings were estimated to calculate LIE using the below equation.

\[
\Delta G_{\text{LIE}} = \alpha \Delta V^{\text{rel}} + \beta \Delta V^{\text{el}}
\]

Here, \(\Delta V^{\text{rel}}\) and \(\Delta V^{\text{el}}\) are the differences in van der Waals and electrostatic interactions of peptides between the bound and unbound states. Also, \(\alpha = 0.181\), and \(\beta = 0.5\).

The Root Mean Square Displacement (RMSD) is calculated using the rms program within GROMACS for the entire production run as well as for the last 25 ns of the simulation. RMSD for a simulation trajectory consisting of N frames is defined using the equation

\[
\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_0)^2}
\]

Here, \(\mathbf{r}_i\) is the coordinates of all atoms in the ith frame and \(\mathbf{r}_0\) is the coordinates of a reference structure (time, \(t = 0\)). Finally, the radius of gyration of peptides were also computed using the grgyrate program in GROMACS for the entire production run.

3. Results and Discussion

Design of peptide inhibitors: While nsp10 is part of pp1a and pp1ab polyproteins, nsp16 is cleaved from pp1ab [25]. Nsp16 is made of 298 residues (polyprotein residues 6799–7096). Upon analyzing the nsp10/nsp16 interface we found nine residues including Lys 6836, Gly 6837, Met 6839, Lys 6874, Ala 6881, Arg 6884, Glu 6885, Ser 6903 and...
Asp 6904 from nsp16 to interact with nsp10 with either H bonds or water bridges. Lys 6836, Gly 6837 and Met 6839 are part of a long loop connecting the $\alpha_2$ and $\alpha_3$ helices. Lys 6874 is part of a loop which connects $\alpha_3$ helix and $\beta_3$ strand. Ala 6881, Arg 6884 and Gln 6885 on the other hand are present in the $\alpha_4$ helix, whereas Ser 6903 and Asp 6904 come from a linker connecting the $\beta_3$ and $\beta_4$ strands. These nine residues were considered as critical for nsp10/nsp16 interaction. Furthermore, a portion of nsp10 is immersed into the hydrophobic pocket formed by $\alpha_3$ and $\alpha_4$ of nsp16 [15].

While designing the peptide inhibitors we considered their probable stability, size and binding efficiency with nsp16. Peptide stability is imparted by the secondary structures including $\alpha$-helices and sheets. Though most of the FDA approved peptides are of less than 20 residues long [26], several recent reports suggest larger peptides of up to 100 residues show desirable in vitro and in vivo results [19,27]. Binding efficiency of peptide-protein interaction is governed by the critical residues via H-bond interactions and also by the hydrophobic pocket. Inhibitor 1 was designed comprising of the $\alpha_4$ helix which included three critical residues (Fig. 1a). Inhibitor 2 is composed of $\alpha_4$ helix, $\beta_3$ and $\beta_4$ strands which included five critical residues (Fig. 1b). Both $\beta_3$ and $\beta_4$ strands were included in this inhibitor as, according to the crystal structure, they stabilize each other. Inhibitor 3 spanned from Lys 6836 to Gly 6911. It is composed of $\alpha_3$, $\beta_2$, $\alpha_4$, $\beta_3$ and $\beta_4$ conformations. In addition to all the nine critical residues, it also included majority of the hydrophobic pocket which interacts with nsp10 (Fig. 1c). We also designed two other inhibitors, 4 and 5, which are shortened forms of inhibitor 3. Fig. 1d and f shows inhibitor 4, where residues Lys 6836 to Leu 6855 included the $\alpha_3$ helix and Lys 6874 to Gly 6911 included $\alpha_4$, $\beta_3$ and $\beta_4$. Lys 6836 (N-term of $\alpha_3$) was connected to Lys 6874 (N-term of $\alpha_4$) with the help of an extra peptide bond by bringing the corresponding

![Fig. 3. Post-simulation quantitative analysis of binding conformations of the inhibitors with nsp10. (A) Time averaged interaction energy between the inhibitors and nsp10. (B) Binding energies between the inhibitors and nsp10. Overall RMSD plot of (C) whole peptide and (D) critical amino acids and. (E) Average RMSD of the whole peptide and critical amino acids and generated from the last 25 ns frames. (F) Radius of gyration plot for the whole peptides. Nsp16 was used as a control. Standard Deviations are represented as error bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)](image-url)
sequences closer and incorporating an additional carbonyl group. Inhibitor 5 also comprised of all the four secondary structures, however, unlike inhibitor 4, no additional functional group was inserted to connect the two helices. Here, Asn 6862 (Fig. 2e–f), ß3 and ß4 strands lost their conformation in inhibitor 2 (Fig. 2c), however the hydrogen bond interactions between the strands remained intact indicating minimal fluctuation in the structure. Conformations of ß3, ß4, ß2 and ß3 were retained in inhibitor 3 (Fig. 2d). ß4 strand however, lost its conformation. It is important to note here that none of the critical residues reside in this strand. Fig. 2e shows that none of the secondary structure conformations except ß4 were maintained in inhibitor 4 indicating that incorporating an additional carbonyl group alters the dihedral angle constraints leading to loss of original conformation. Fig. 2f shows that the ß sheet and both the a helices were stable in inhibitor 5 (Fig. 2f). In summary, all the critical conformations were retained in inhibitors 1, 2, 3 and 5 and they bind in a manner which is similar to that seen in the crystal structure.

4. Conclusion

In the present study we use molecular dynamics simulation approach to explore the possibility of inhibiting nsp10/nsp16 interaction using peptides derived from nsp16. Our results indicate that though inhibitors 1, 2, 3 and 5 were stable post-simulation, inhibitors 2 and 5 showed comparable binding energy and acceptable RMSD compared to the control. They also cover the nsp10 surface which interacts with nsp16. It is however, worth mentioning that the current work is an initial proof-of-principle study and the stability of the inhibitors requires validation in vivo and in vivo settings. Once the stability and efficacy of the peptides are validated, they could be delivered into the cell using nanoparticle carriers to improve their bioavailability. This peptide therapeutic approach has potential to inhibit 2’-O–MTase activity of nsp16, thereby preventing viral RNA capping, and finally translation of viral proteins.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A.E. Gorbalygina, et al., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat. Microbiol. 5 (4) (2020) 536-544, https://doi.org/10.1038/s41556-020-0695-z.
[2] G. Pascarella, et al., COVID-19 diagnosis and management: a comprehensive review, J. Intern. Med. (2020), https://doi.org/10.1111/joim.13091.
[3] Johns Hopkins Coronavirus Resource Center. https://coronavirus.jhu.edu/.
[4] A.J. te Velthuis, J.J. Arnold, C.E. Cameron, S.H. van den Worm, E.J. Snijder, The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent, Nucleic Acids Res. 38 (1) (2010) 203–214, https://doi.org/10.1093/nar/gkp904.
[5] K.A. Ivanov, V. Thié, J.C. Doble, Y. van der Meer, E.J. Snijder, J. Ziebuhr, Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase, J. Virol. 78 (11) (2004) 5619-5632, https://doi.org/10.1128/JVI.78.11.5619-5632.2004.
[6] Y. Chen, et al., Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase, Proc. Natl. Acad. Sci. U. S. A. 106 (9) (2009) 3484-3489, https://doi.org/10.1073/pnas.0808790106.
[7] Y. Kim, et al., Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2, Protein Sci. 29 (7) (2020) 1596-1605, https://doi.org/10.1002/pro.3873.
[8] Y. Chen, et al., Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2’-O-methylation by nsp16/nsp10 protein complex, PLoS Pathog. 7 (10) (2011), https://doi.org/10.1371/journal.ppat.1002294.e1002294.
[9] A.R. Fehr, S. Perlman, Coronavirus: an overview of their replication and pathogenesis, Methods Mol. Biol. 1282 (2015) 1-23, https://doi.org/10.1007/978-1-4969-2438-7_1.
[10] R. Zür, et al., Ribose 2’-O-methylation provides a molecular signature for the distinction of self and non-self RNA dependent on the RNA sensor Mda5, Nat. Immunol. 12 (2) (2011) 137-143, https://doi.org/10.1038/ni.1979.
[11] E.J. Snijder, E. Decroly, J. Ziebuhr, The nonstructural proteins directing coronavirus RNA synthesis and processing, Adv. Virus Res. 96 (2016) 59-126, https://doi.org/10.1016/bs.avir.2016.08.008.
[12] M. Bouvet, et al., In vitro reconstitution of SARS-coronavirus mRNA cap methyltransferase, PLoS Pathog. 6 (4) (2010), e1000863, https://doi.org/10.1371/journal.ppat.1000863.
[13] S. Daffis, et al., “2’O-methylation of the viral mRNA cap evades host restriction by IFIT family members, Nature 468 (7322) (2010) 452-456, https://doi.org/10.1038/nature09489.
[14] M. Rosol-Lema, et al., The crystal structure of nsp10/nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine, bioxiv (2020), https://doi.org/10.1101/2020.04.17.047498.
[15] P. Krafiková, J. Šilhán, R. Nencka, E. Boura, Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinuefungin, Nat. Commun. 11 (1) (2020) 3717, https://doi.org/10.1038/s41467-020-17495-9.
[16] S.K. Maurya, A.K. Maurya, N. Mishra, H.R. Siddleq, Virtual screening, ADME/T, and binding free energy analysis of anti-viral, anti-protease, and anti-infectious compounds against NSP10/NSP16 methyltransferase and main protease of SARS-CoV-2.
[17] Y. Han, P. Krnl, Computational design of ACE2-based peptide inhibitors of SARS-CoV-2, ACS Nano 14 (4) (2020) 5143–5147, https://doi.org/10.1021/acsnano.0c02857.

[18] B. Leader, Q.J. Baca, D.E. Golan, Protein therapeutics: a summary and pharmacological classification, Nat. Rev. Drug Discov. 7 (1) (2008) 21–39, https://doi.org/10.1038/nrd2399.

[19] I. Inamoto, J.A. Shin, Peptide therapeutics that directly target transcription factors, Pept. Sci. 111 (1) (2019), e24048, https://doi.org/10.1002/pep2.24048.

[20] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1) (1995) 43–56, https://doi.org/10.1016/0010-4655(95)00042-E.

[21] R.B. Best, et al., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone $\phi$, $\psi$ and side-chain $\chi_1$ and $\chi_2$ dihedral angles, J. Chem. Theor. Comput. 8 (9) (Sep. 2012) 3257–3273, https://doi.org/10.1021/ct300400x.

[22] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (4) (Jul. 1983) 926–935, https://doi.org/10.1063/1.445869.

[23] G. Busi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (1) (Jan. 2007) 14101, https://doi.org/10.1063/1.2408420.

[24] R. Kumari, R. Kumar, A. Lynn, g.mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (7) (Jul. 2014) 1951–1962, https://doi.org/10.1021/ci500020m.

[25] A. Wu, et al., Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, Cell Host Microbe 27 (3) (2020) 325–328, https://doi.org/10.1016/j.chom.2020.02.001.

[26] B.J. Bruno, G.D. Miller, C.S. Lim, Basics and recent advances in peptide and protein drug delivery, Ther. Deliv. 4 (11) (Nov. 2013) 1443–1467, https://doi.org/10.4155/tde.13.104.

[27] A. Sorolla, et al., Precision medicine by designer interference peptides: applications in oncology and molecular therapeutics, Oncogene 39 (6) (2020) 1167–1184, https://doi.org/10.1038/s41388-019-1056-3.