Life Sciences 238 (2019) 116958

Contents lists available at ScienceDirect

Life Sciences

journal homepage: www.elsevier.com/locate/lifescie

Molecular dynamics and docking reveal the potency of novel GTP derivatives against RNA dependent RNA polymerase of genotype 4a HCV

Abdo A. Elfiky*, Alaa Ismail

Biophysics Department, Faculty of Science, Cairo University, Giza, Egypt

ARTICLE INFO

Keywords:
Molecular modeling
Molecular dynamics simulation
HCV genotype 4a
Polymerase
Nucleotide inhibitors
Protein-ligand docking

ABSTRACT

Aim: To work on Hepatitis C Virus (HCV), one of the major causes of liver cirrhosis and hepatocellular carcinoma, polymerase of genotype 4a that have no solved structures deposited in the protein data bank (PDB) yet. Understanding the dynamics and testing some novel inhibitors are also covered.
Materials and methods: Molecular Dynamics Simulation (MDS) is performed for a period of 1 μs on comparatively modeled then validated NS5b of subtype 4a. Following MDS analysis, molecular docking is performed to test the inhibitory performance of eight novels suggested guanosine derivatives using 181 different conformations of the protein model gathered during the MDS run after the equilibration period.
Key findings: The results yield that the eight modified, at position 2', GTP derivatives (fluorine, Hydroxyl, and sulphonyl oxydanyl) have binding energies comparable to the parent molecule, GTP. Besides, the eight suggested compounds have lower binding energies (and hence better in binding) compared to sofosbuvir (a drug approved by FDA in 2013 against HCV) and ribavirin (a wide range acting antiviral drug used before against HCV).
Significance: Combined molecular dynamics and molecular docking are able to test the hypothesis of HCV polymerase dynamics doesn't affect the nucleotides (or nucleotide inhibitors) binding to its active site. Despite the reported highly dynamic subtype 4a of HCV; all the nucleotide inhibitors under the study are able to, tightly, bind to NS5b of genotype 4a. This behavior is reported before for the Zika virus polymerase, as well.

1. Introduction

Hepatitis C Virus identified for the first time 30 years ago [1]. HCV is one of the confirmed reasons for liver complications such as cirrhosis and loss of liver function. Besides, hepatocellular carcinoma may develop in chronically infected patients [2–7].

HCV is an RNA virus and characterized by massive mutations that result in the differentiation of many quasi-species inside the individual patient. World Health Organization (WHO) classified HCV to seven genotypes with about 70% sequence identity. Egypt has a vast number of infections that reach more than 14% of the population while the dominant genotype (~90%) is 4, subtype a [8–10].

The primary therapeutic regimen was interferon and ribavirin until eight years ago [11–13]. In the year 2011 and the following years, the Food and Drug Administration (FDA) approved different drugs, called Direct Acting Antivirals (DAAs). These drugs inhibit specific viral proteins, mainly NS3 serine protease, NS5b RNA dependent RNA polymerase (RdRp) and NS5a [3,5–7,11,14].

IDX-184 is a guanosine derivative that proved excellent inhibitory performance against HCV NS5b. It gave better results compared to Ribavirin [2,10,15]. Due to complications associated with the regimen in which IDX-184 was involved, the clinical trials were halted in March 2013. On the other hand, the precursor of this drug (Guanosine nucleotide) may be used as a seed for other modified compounds that could be safe and give better inhibitory performance against NS5b.

Eight novel modified compounds generated, in this study, from the physiological guanosine, in its triphosphate form, (GTP). Molecular docking technique was utilized to test protein/ligand binding, at different dynamics states, and to score the binding free energies for the different ligands to the protein active site at different conformations (every 5 ns) [13,16–18]. Besides, some docked structures were investigated to analyze the resulting complexes.

No solved protein structures for genotype 4a are deposited in the protein data bank (PDB) yet [9,19–21]. For this reason, comparative protein modeling was performed to obtain a 3D model that equilibrated for a period of 1 μs using Molecular Dynamics Simulation (MDS). The equilibrated protein system is analyzed and discussed in detail in the next sections.

The results revealed high potency of the suggested compounds against HCV polymerase. Besides, the most movable secondary

* Corresponding author. Biophysics Department, Faculty of Science, Cairo University, Giza, Egypt.
E-mail addresses: abdo@sci.cu.edu.eg, aelfiky@ictp.it (A.A. Elfiky).

https://doi.org/10.1016/j.lfs.2019.116958
Received 8 August 2019; Received in revised form 8 October 2019; Accepted 11 October 2019
Available online 16 October 2019
0024-3205/

structure element (β-hairpin 1) doesn't affect nucleotide-binding even at the movement events. N-terminal is less stable compared to C-terminal; in addition, it involved in the nucleotide-binding (through H-bonds (D1) and salt bridges (R3) with D99 and D100) suggesting a role of nucleotide-binding on the degradation rate of the polymerase through the N-degron pathway.

2. Materials and methods

2.1. Model building

Genotype 4a HCV NS5b RdRp sequence (GB: AET98770.1) was retrieved from the National Center for Biotechnology Information (NCBI) protein database. The downloaded sequence was used to construct the all atoms 3D model using comparative protein modeling server PHYRE 2 [20] using HCV RdRp of genotype 1a (HC-J4 strain PDB code: 1NB4) as a template. The protein model was validated using **S**tructure **A**nalysis and **Ve**rification **S**erver (SAVES webserver). Three software used in the validation of the model; PROCHECK, PROVE, and ERRAT. These software test stereo-chemical properties, atomic volumes, and overall error factor of the generated protein model, respectively [22–26].

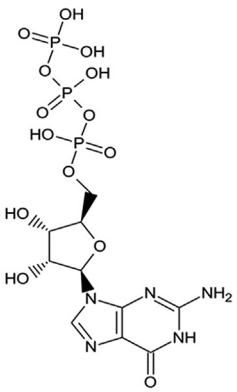
2.2. Molecular Dynamics Simulation

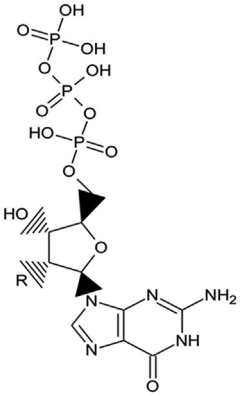
Genotype 4a HCV NS5b all atoms 3D model was solvated with 12,759 water molecules. The default solvate function of VMD is used to generate a water box of size 81.6 × 77.2 × 67.5 $Å^3$ centered at 59.5, 61.5, 62.4 Å such that the protein coordinates lie in the center of the box. The padding is set to be between 10 and 25 Å in the 3 directions to ensure that the protein is confined in the cell without interacting with other images. The density of the water in the system is guaranteed to be 1 gm/cm^3 using an in-home script. The total polymerase charge was −3, so three sodium ions (Na^+) are added to the system to neutralize the charges. The solvent was minimized for 100 ps using NAMD software and utilizing CHARMM27 force field. After that, Normal Pressure and Temperature (NPT) ensemble is used to relax the protein for a total period of 5 ns. The protein volume was fluctuating around 396100 $Å^3$ that used to generate a box size of (73.43 × 73.43 × 73.43) Å for the periodic boundary condition. A production run for 1 μs was performed using Normal Volume and Temperature (NVT) ensemble. Explicit solvation scheme is maintained for all the simulations using the TIP3P water model. The protein coordinates were retrieved every 5 ns for the next step of our work, the docking study. All MDS were conducted on Cy-Tera supercomputer facility of Cyprus Institute of Science under the project number "pro15b114s1".

2.3. Molecular docking

SCIGRESS 3.0 software was used to conduct the docking calculations [13,17,27,28]. Ligands structure optimization was performed in three steps; using molecular mechanics force field MM3 [29] then semi-empirical quantum mechanical parameterization method 6 (PM6) [30] and finally using density functional theory (DFT) B3LYP functional [15,17,31]. The calculations were performed using SCIGRESS 3.0 software installed on Dell precession T3600 workstation.

After the removal of all water molecules from the protein model, the optimized ligands were docked into the active site of NS5b (D99 and D100, conserved active site amino acids) using the default setting of SCIGRESS [17]. The potential of mean force (PMF) scoring function is used with 0.25 grid spacing. A genetic algorithm is utilized with 50 population size, 3000 maximum generation, 0.8 cross rate, and 0.3 mutation rate. The maximum iteration for local search is set to 20 with a 0.06 rate. The active site was treated as flexible during the docking experiment. The selection of the consecutive aspartates (D99 and D100) for docking is based on the literature and the solved structures of NS5b from other genotypes. Docking scores were retrieved, and the docking





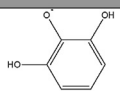
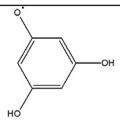
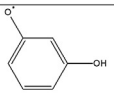
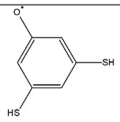
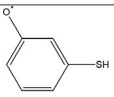
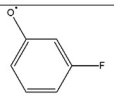
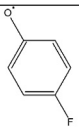
compound	Substitution name	Chemical structure of R group
Compound 1	(2,6-dihydroxyphenyl)oxidanyl	
Compound 2	(2-hydroxyphenyl)oxidanyl	
Compound 3	(3,5-dihydroxyphenyl)oxidanyl	
Compound 4	(3-hydroxyphenyl)oxidanyl	
Compound 5	(3,5-disulfanylphenyl)oxidanyl	
Compound 6	(3-sulfanylphenyl)oxidanyl	
Compound 7	(3-fluorophenyl)oxidanyl	
Compound 8	(4-fluorophenyl)oxidanyl	

Fig. 1. The structures of active guanosine nucleotide (GTP) and The modified compounds (2′ carbon of the ribose ring where the OH group is replaced by the R group. The structures are sketched by the freeware version 2012 of ChemSketch.

poses examined for established interactions using Protein-Ligand Interaction Profiler (PLIP), PyMOL, and Maestro software [32–34].

3. Results and discussion

3.1. Generation of the inhibitor molecules

Fig. 1 shows the structures of GTP and the suggested modifications. The modifications are introduced in the 2′ carbon of the ribose ring of GTP in which we replace the hydroxyl group with different aromatic substitutions. These aromatic substitutions could make steric hindrance once it has attached to the newly formed RNA strand, the primer, and blocking the polymerization process. In this study, a total of 8 suggested compounds are tested; four hydroxyl-substituted oxidanyl (mono- and di- ortho/meta), two sulfanyl-substituted oxidanyl (mono- and di-meta) and two fluorine-substituted oxidanyl (mono-meta/para) compounds. These modifications are based on other nucleotide inhibitors in the market and under clinical trials in which bulky group or halide substitutions are introduced in the 2' carbon of the ribose ring of nucleotides [2,3,10,11,19]. All the docking calculations are done for real compounds (no transition states) after optimizing its 3D structures quantum-mechanically as mentioned in the materials and methods section.

3.2. HCV RdRp of genotype 4a model building

Figure S1 A presents the sequence alignment between the query model for genotype 4a and the sequence of template solved protein structure (PDB ID: 1NB4) from which PHYRE 2 built the 108-residues model. The sequence identity was 79%, and the confidence of the built model was 100%. The all atoms root mean square (RMS) deviation of the superimposed model to the template structure is 0.36 Å (figure S1 B). Besides, the model was valid according to SAVES software outputs. Ramachandran plot gives (93.4% in the preferred region, 3.8% in the allowed area and only 2.8% in the disallowed region). ERRAT (overall quality factor is 79.798%) and PROVE software (Total number of buried outlier protein atoms was 3.1% of the scored atoms) outputs are suitable for the protein model.

Secondary structures of the model are in good agreement with that of the solved structure, the template. NS5b RdRp of genotype 4a model consists of three α-helices (α1, α2, and α3) and two β-hairpins (β-hairpin 1 and 2), one of these hairpins (β-hairpin 1) located between α2 and α3 while the other one (β-hairpin 2) lies after α3 just before the C-terminal (see figure S1 B). The conserved active site amino acids (D99 and D100) are in the beta-turn structure of the β-hairpin 2. These two amino acids are very conservatives and are exposed to the solvent (see figure S1 B).

3.3. HCV RdRp of genotype 4a model dynamics

Molecular Dynamics Simulation for 1 μs was performed on the HCV RdRp Genotype 4a model to equilibrate it and explore the different possible conformations it possesses to test these conformations against anti-polymerase drugs. Fig. 2 A through C show snapshot structures of NS5b RdRp of HCV genotype 4a model in three different conformations. The first snapshot (A) is taken after 150 ns (just after equilibration of the protein-solvent system) while the second (B) and third (C) snapshots are taken in the middle (500 ns) and near the end of the simulation (890 ns). From the figures, the most mobile secondary structural element of the model is the β-hairpin 1. It rotates about 180^0 during the MD run. It moves from a plane in the z-axis, perpendicular to the plane of the α3 helix (see Fig. 2A), to near the x-axis plane (still perpendicular to α3) at 500 ns, while finally, at 890 ns it reaches the same plane of α3 (y-axis plane). Interestingly, it is noticed that the β-hairpin 1 motif loses its secondary structure and become a loop from the beginning of the MDS (see Fig. 2 A and S1 B), while near the end of the simulation (Fig. 2C and E), part of this motif is folded back but into a small α-helix (5-residues only). β-hairpin 2 also unfold into a loop but stepwise during the simulation while the overall shape of it remains as the hairpin due to some maintained H-bonds as we will show later (see Fig. 2A-C and F).

Consequently, the Root Mean Square Deviation (RMSD) versus simulation time (Fig. 2D) exhibit dramatic changes during the 1 μs MD run. Both backbone (blue line) and all heavy atoms (orange line) RMSDs are shown in Fig. 2D. The protein system is equilibrated in about 70 ns followed by about 0.5 μs of fluctuations around 4 Å. A small increase in the RMSDs (up to ~ 6 Å) is found in the period 260–390 ns. On the other hand, a rapid jump in the RMSDs that reach 10 Å characterize the period starting from 640 ns and until 870 ns when it reduced to ~8 Å until the end of the simulation run. This jump in the RMSD is suggested to be due to the rotation of the β-hairpin 1 motif (see Fig. 2A–C).

Fig. 2 D also shows the surface accessible surface area (SASA) (yellow line) and the radius of gyration (RoG) (gray line) calculated for each frame of the simulation time. SASA values are stable during the MD run with an average value of ~8000 $Å^{2,}$ which reduced to ~7000 $Å^2$ after the big jump (after 720 ns and until the end of the simulation). This illustrates a conformational change that resulted in more backing of the protein, which will be discussed when talking about H-bonds and distances later. Consequently, when the protein backing is enhanced the protein radius is reduced, and that behavior can be deduced from the values of the calculated RoG for our protein model (RoG values reduced from ~16 Å to ~ 14.5 Å after the big jump at 720 ns and ongoing).

This unusual behavior of the model encourages us to follow some distances that may explain such jump in the RMSD and increasing the packing revealed from both SASA and RoG. Three distances are recorded during the MD simulation. These distances are calculated between the α^{carbon} of the residues to check its pattern during the simulation, especially at the big jump appeared in the RMSD (Fig. 2D). D6-D53, D6-D99, and V20–K79 distances are shown in Fig. 2E, where it represented in blue, red, and brown lines, respectively. An enlarged panel is showing the big jump region reported in the RMSD (600–800 ns), for further illustration, on the top center of Fig. 2E. We have two events (dashed green ovals 1 and 2 at 650 and 720 ns, respectively) of dramatic changes in distances. These quick changes are due to massive conformational changes of the underlying secondary structural elements. Snapshots are taken before and after the big jump at 560 and 750 ns, respectively (Fig. 2E, top corners).

D6-D53 (blue line) represents the distance between the D6 residue, at the beginning of the α1-helix, and the D53 residue, near the center of the β-hairpin 1 element. D53 residue lies in the center of the formed small α-helix after the big jump. The changes in this distance are related to the movement of the β-hairpin 1 element relative to the α1-helix. D6-D53 exhibits a reduction in both 1 and 2 events, as shown in the two green dashed ovals of Fig. 2E enlarged panel. The decrease of D6-D53 is accompanied by an increase of the D6-D99 distance (red line) at 650 ns (event 1) and also an expansion of the V20–K79 distance (brown line) at 720 ns (event 2). D6-D99 represents the distance between the active site aspartic acid residue D99 to the beginning of the α1-helix element. It increases from ~5 up to ~30 Å at 650 ns (event 1). On the other hand, V20–K79 distance represents the distance between the centers of α1 and α3 helices. It increased from ~7 up to ~9 Å at the second event (at 720 ns) and persisted until the end of the simulation. From the observed changes in these three distances, we conclude that at event 1 the α1 helix is moved apart from the β-hairpin 2 (carrying the active residues D99 and D100) and gets close to the β-hairpin 1. While during event 2 again α1 helix is moved toward the β-hairpin by making a conformational unfolding as the α1-α3 distance increased. This is accompanying the formation of small α-helix in the unfolded β-hairpin 1 element that could help in maintaining the conformational stability of the protein by getting much more amino acids in contact through H-

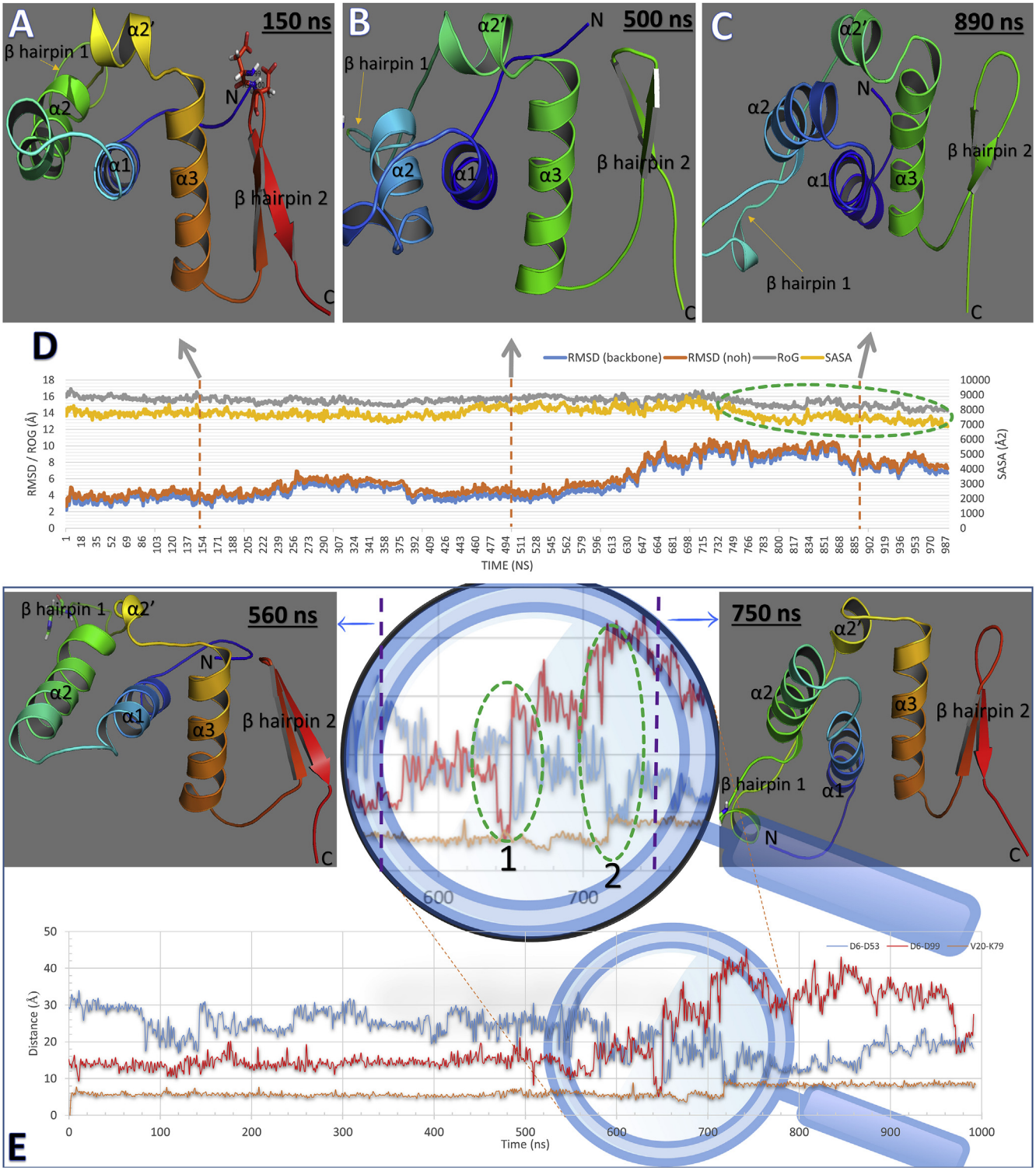


Fig. 2. Structures of HCV NS5b RdRp of genotype 4a modeled by PHYRE 2 web server and equilibrated by MDS for 150 ns (A), 500 ns (B) and 890 ns (C) in NVT ensemble using the TIP3P water model shown in colored cartoon representation. Structural elements are labeled (D) Backbone (blue) and heavy atoms (orange) Root Mean Square Deviation (RMSD), Radius of Gyration (RoG) (gray), and Surface Accessible Surface Area (SASA) (yellow) versus time of the 1μs MDS run. (E) D6-D53 (blue), D6-D99 (red), and V20–K79 (brown) distances versus time. The big jump is shown in the enlarged panel showing the two events 1 and 2 (green dashed ovals). Two snapshots at 560ns and 750 ns are prepared like (A). (F) Per residue Root Mean Square Fluctuations (RMSF). The mobile elements are contained between two dashed blue or green lines. Two snapshots are taken at the beginning and the end of the simulation and prepared like (A). (G) H-bonds number versus time (top) and some of the salt bridges that established in the protein model versus time (bottom). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

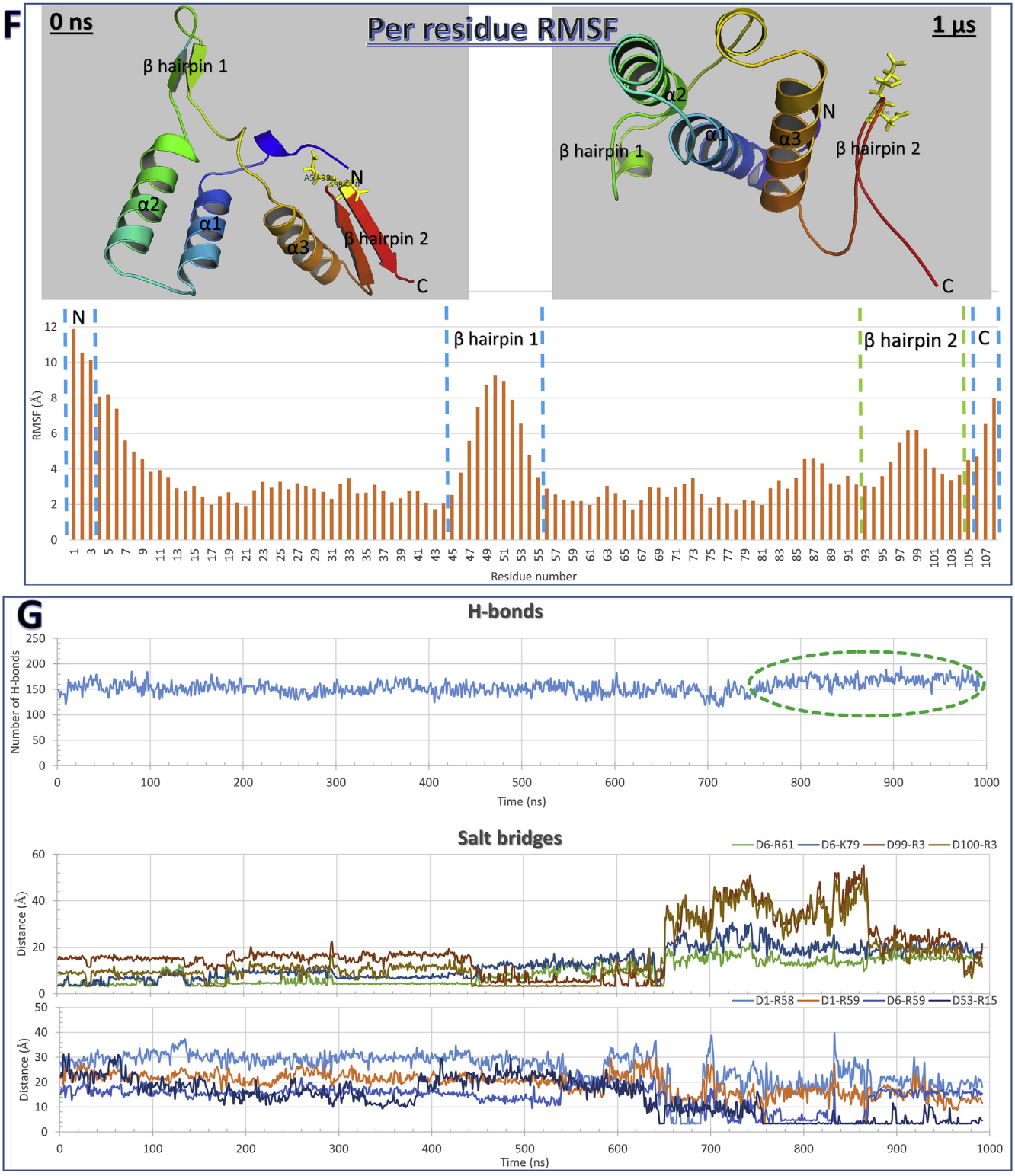


Fig. 2. (*continued*)

bonding. N-terminal movement can play a significant role in these rearrangements due to its very high mobility.

To get much understanding of what happened during the simulation, especially at the big jump region, we performed per residue Root Mean Square Fluctuation (RMSF) analysis (Fig. 2F). As expected, the two terminals of the protein model (N and C termini) are flexible. Noticeable that the N-terminal is more flexible (RMSF of 12 Å) compared to the C-terminal (RMSF of 8 Å). For the core part of the protein, β-hairpin 1 shows the most flexible region (RMSF of 9 Å). β-hairpin 2 region (between the two green dashed lines) shows moderate fluctuations during the MDS (RMSF of 6 Å) while other parts of the core protein model are stable during the simulation. Two snapshots are taken (Fig. 2F top) at the beginning and the end of the simulation showing the changes that happened upon simulating the protein dynamics. The most noticeable changes are in the terminals (N and C), and β-hairpins 1 and 2 unfolded into loops. Besides, β-hairpins 1

Table 1
H-bonds formed between the residues of the RdRp model over the simulation period of 1 μs

Donor	Acceptor	Occupancy %	Structural element
ARG59-Side	GLU11-Side	91.53%	α1 - β-hairpin 1
VAL16-Main	LYS12-Main	95.06%	α1
LEU41-Main	LEU37-Main	97.58%	α2
ARG40-Main	ALA36-Main	93.75%	α2
GLY18-Main	ILE14-Main	95.56%	α1
ARG61-Side	GLU17-Side	90.83%	α1 - β-hairpin 1
LYS79-Side	GLU17-Side	61.90%	α1- α3
GLU17-Main	ASP13-Main	93.65%	α1
THR38-Main	ILE34-Main	98.19%	α2
THR38-Side	ILE34-Main	92.64%	α2
ALA80-Main	CYS76-Main	97.58%	α3
ALA83-Main	LYS79-Main	96.37%	α3
ALA86-Main	ALA82-Main	96.27%	α3
GLN81-Main	TYR77-Main	96.67%	α3
ARG85-Main	GLN81-Main	92.24%	α3
VAL102-Main	LEU95-Main	71.37%	β-hairpin 2
MET94-Main	VAL102-Main	53.23%	β-hairpin 2
LEU95-Main	VAL102-Main	97.88%	β-hairpin 2
ILE104-Main	THR93-Main	90.42%	β-hairpin 2
CYS92-Main	ILE104-Main	23.89%	β-hairpin 2
THR93-Main	ILE104-Main	45.46%	β-hairpin 2
SER107-Main	ASP91-Main	14.72%	β-hairpin 2
LEU37-Main	VAL33-Main	98.29%	α2
THR75-Main	GLY71-Main	94.05%	α3
THR75-Side	GLY71-Main	93.25%	α3
LYS79-Main	THR75-Main	98.39%	α3
CYS76-Main	ASN72-Main	94.86%	α3
TYR77-Main	THR73-Main	93.95%	α3
ALA82-Main	LEU78-Main	97.48%	α3
LEU78-Main	LEU74-Main	98.39%	α3
LYS79-Side	ASP13-Side	51.21%	α1- α3
ARG61-Side	ASP13-Side	41.43%	α1 - β-hairpin 1
ARG58-Side	GLY44-Main	17.14%	β-hairpin 1
ARG58-Side	ARG40-Main	29.03%	β-hairpin 1 – α2
ARG90-Side	GLU106-Side	29.44%	C-terminal
ARG85-Side	ASP91-Side	41.53%	α3- β-hairpin 2
ARG90-Side	ASP108-Main	28.53%	β-hairpin 2
ALA105-Main	THR93-Main	14.11%	β-hairpin 2
CYS55-Main	LYS51-Main	25.30%	β-hairpin 1
LEU54-Main	LYS51-Main	23.99%	β-hairpin 1
ASP1-Main	ASP99-Side	12.00%	N-terminal - β-hairpin 2
THR93-Main	ALA105-Main	18.45%	β-hairpin 2
TYR66-Side	THR38-Side	42.44%	α2- α2′
TYR42-Side	ILE14-Main	14.92%	α1- α2
THR38-Main	TYR66-Side	11.19%	α2- α2′
LEU54-Side	LYS51-Main	11.49%	β-hairpin 1
CYS55-Side	LYS51-Main	10.69%	β-hairpin 1
ARG15-Side	ASP53-Side	21.77%	α1- β-hairpin 1
ARG90-Main	GLU106-Main	18.45%	β-hairpin 2
GLU106-Main	ARG90-Main	20.46%	β-hairpin 2
LYS51-Main	MET47-Main	17.54%	β-hairpin 1

changed its direction dramatically, and small α-helix centered at D53 is formed after the big jump events, while β-hairpin 2 has little changes.

Fig. 2 G (top) shows the number of H-bonds that are formed in the protein model against time. As can be noticed, the number of H-bonds before the big jump region is ~150, but upon the jump (after 750 ns), it rises a little bit to ~ 170. This is consistent with both SASA and RoG, which shows an increase in the protein backing after 750 ns. Table 1 lists some of the H-bonds and its % occupancy during the MDS run. From the table, we can conclude that α-helices (especially α3 helix that have > 90% occupancies) are the most stable secondary structural elements in the model. On the other hand, β-hairpin 1 is the most unstable element in terms of the occupancies of its H-bonds. C-terminal (29.4%) is more stable than N-terminal (12%) as reflected from the % occupancies.

Besides, Fig. 2G (bottom) shows some of the salt bridges that present in the protein model versus time. We have here four salt bridges (D6-R61, D6-K79, D99-R3, and D100-R3) that were stable until 650 ns when it starts to dissociate. On the other hand, D1-R58, D1-R59, D6-R59, and D53-R15 show the reverse behavior as it formed after 650 ns. This means that a conformational change occurred in the protein model with increasing protein backing. From the salt bridges, we can see that the most movable parts are the N-terminal residues (D1, R3, and D6) and the β-hairpin 1 region and its preceding residues (R53, R58, R59, and R61). The N-terminal residue D1 makes salt bridges with the region just after the β-hairpin 1 (R58 and R59) only at the jump region (650–700 ns).

Interestingly, The N-terminal of the protein (R3) interacted with the active site residues (D99 and D100) between 450 and 650 ns of the simulation. This causes a question to arises here; Is the change of the structural elements relative to the active site has an effect on the binding of nucleotides or its derivative inhibitor? In the next section, we will answer this question by performing molecular docking at different conformations of the protein during the MDS run and see how is the binding energy during the entire MDS. Also, does N-terminal flexibility and its involvement in the GTP binding, as we will see later, affect the recognition of the polymerase by its recognin E3 ligase for proteasome degradation?

Nucleotide inhibitors against HCV RdRp of genotype 4a:

Fig. 3 shows the relative average binding energies (docking scores relative to GTP score, where the error bars represent the standard deviation) for the docking of the eight suggested compounds (blue) along with the parent compound, GTP (green) and the drugs, sofosbuvir (orange), which approved by FDA in December 2013, ribavirin (brown), a wide-range acting antiviral, and IDX-184 (yellow), which was under clinical trials. Docking study was performed using 181 different conformations represent the protein every five ns of MDS starting from 100 ns. From the figure, we can conclude that the binding energies for all of the suggested compounds are comparable with that of the parent compound and hence competitive binding, against GTP, for the protein active site is suggested for the compounds and thus polymerase inhibition.

As a comparison between the relative binding energies of the suggested compounds and IDX-184, sofosbuvir and ribavirin to NS5b active site; the proposed compounds have better values compared to sofosbuvir and for less extent to that of ribavirin. On the other hand, no significant difference between the relative binding energies of the suggested compounds and IDX-184 is reported. Fig. 4 shows the relative average binding energies represented in dot plots as a function of time for IDX-184 (A), sofosbuvir (B), ribavirin (C), and modification 3 (D), which is the best-suggested compound based on its relative average docking score. A comparison between GTP and each compound illustrates the effectiveness of using these compounds against the NS5b of genotype 4a (see Figure S2). Sofosbuvir and to less extent ribavirin are less in their relative binding energies compared to GTP during the MDS run. On the other hand, IDX-184 has more or less relative binding energies as GTP. These results are in good agreement with our previous *in silico* study made on different HCV genotypes 1a, 2b, and 3b.

3.4. Sofosbuvir versus suggested modified nucleotide inhibitors

We noticed some relative docking score values that deviate from the average values. To quantify the binding mode of these cases, we further analyzed the binding patterns utilizing Protein-Ligand Interaction Profiler (PLIP) for GTP and Sofosbuvir. The best ten and the worst ten values are used in this analysis, and the results are tabulated in Tables 2 and 3 for GTP and Sofosbuvir, respectively. Two significant interactions are established between the ligands and the protein binding site upon docking, which are H-bonds and salt bridges. Other interactions still occur but less frequently, such as hydrophobic interactions and π-cation interaction.

As shown from Table 2, the number of H-bonds and salt bridges formed between GTP and the protein is higher in the best scores complexes compared to the lowest scored complexes. This pattern is also

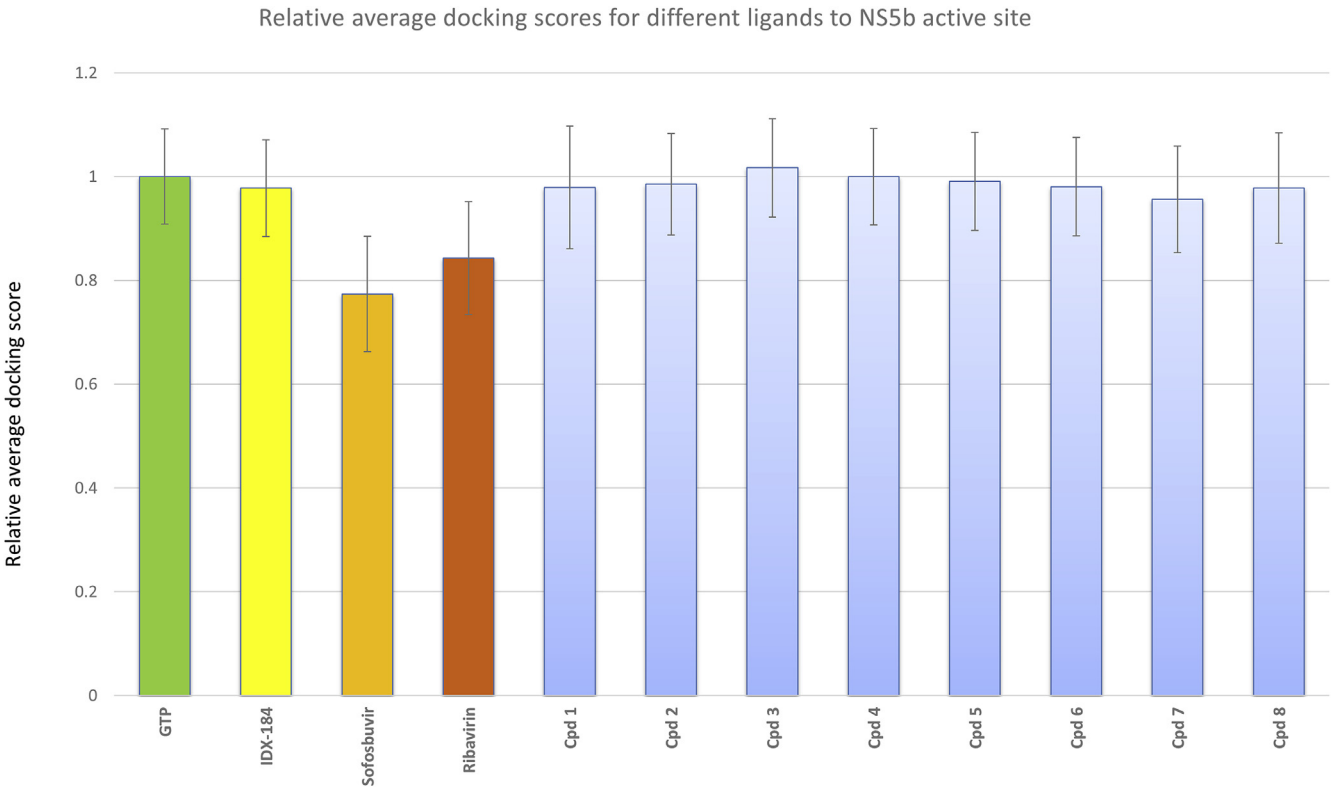


Fig. 3. Bar graph representing the relative average docking scores for GTP (green), IDX-184 (yellow), Sofosbuvir (orange), Ribavirin (brown), and the suggested modifications (blue) into NS5b of HCV subtype 4a. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

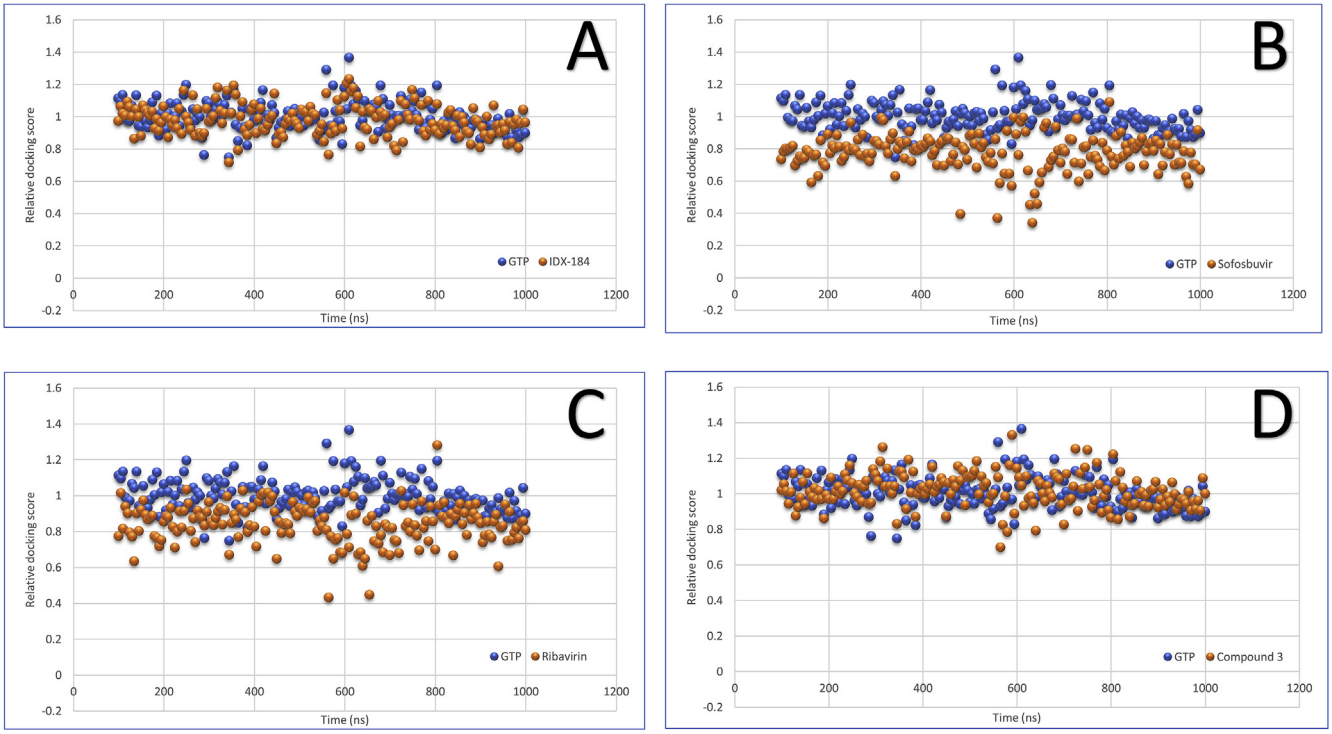


Fig. 4. Relative docking scores as a function of time (ns). GTP values are compared to a) IDX-184, b) Sofosbuvir, c) Ribavirin, and d) suggested compound 3 with (3,5-dihydroxyphenyl)oxidanyl as R group instead of the hydroxyl group in the parent compound. The time course is from 100 ns to 1000 ns.

Table 2
Analysis of the best and worst docking score values obtained at different time steps of the MDS for GTP utilizing PLIP web server.

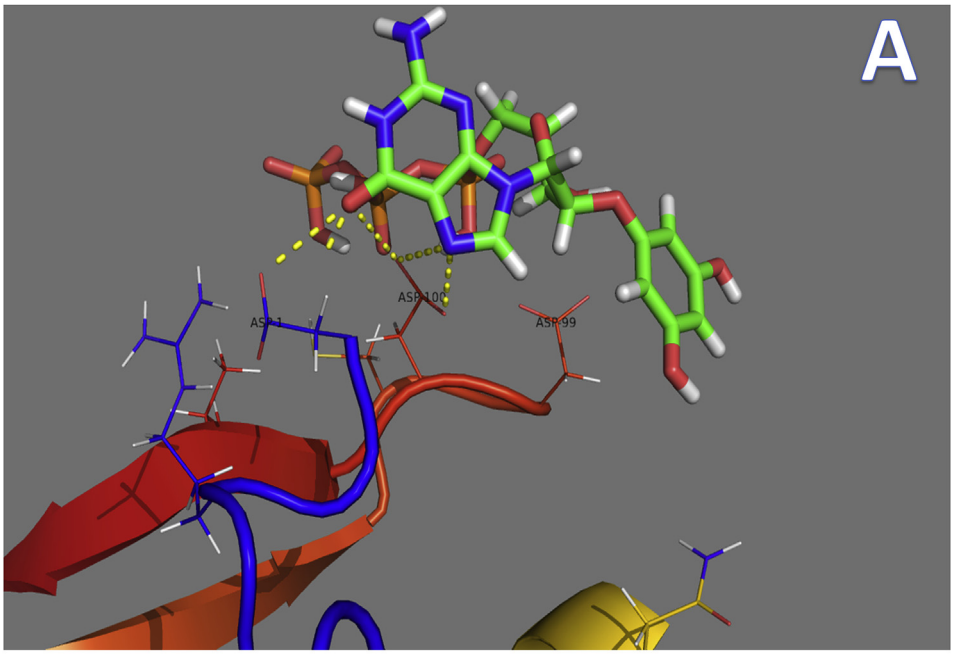
GTP best 10 docking values								
MDS time at which docking performed	H-bonds		Salt bridges		π-cation interaction		Hydrophobic interaction	
	No.	Amino acids involved	No.	Amino acids involved	No.	Amino acids involved	No.	Amino acids involved
250	2	D1 and D100	1	D99				
420	5	T2, D99 (3) and D100	1	D99				
560	6	D1, R3 (3) and D100 (2)	4	D1, R3, D99 and D100	1	R3		
575	3	R3 (2) and D99	3	R3 (2) and D100				
600	3	R3, D99 and D100	2	R3 and D100				
610	6	D1, R3, D99 (3) and D100	3	R3 (2) and D100				
625	5	R3 (2), D99 and D100 (2)	3	D1, R3 and D99				
630	2	R3 (2)	2	R3 and D100				
680	1	R3	3	R3 (2) and D100	1	R3		
805	4	K79, D99 and D100 (2)	2	K79 and D99				
GTP worst 10 docking values								
285	3	D1, R3 and D99	1	D100				
290			1	D100				
345			1	D99				
365	2	D99 and D100	1	D99				
385	2	D1 and D100	1	D99				
450	1	D99	1	D100				
545			1	D100				
595	5	D1, R3, D99 (2) and D100			1	R3		
845	4	D99 (3) and D100	1	D100				
900	2	D99 and D100	1	D99				

concluded from Table 3, but the situation is somehow different. Sofosbuvir form a fewer number of H-bonds and salt bridges with the protein compared to GTP, and this is reflected in its relative docking scores values (see Figs. 3 and 4). The primary interacting amino acids in the best GTP complexes are D1, R3, D99, and D100, while for the worst complexes, the interactions mainly occur through the active site aspartates only (D99 and D100). This means that the best interaction of GTP to the polymerase is when the N-terminal residues D1 and R3 are in close vicinity of the active site aspartates. For Sofosbuvir, the number of interactions is less in the worst complexes, but no constant pattern of amino acids could be concluded. This could be due to the chemistry of Sofosbuvir (a uridine derivative) that differ from that of GTP.

As an example, Fig. 5 A shows the docked structure of (3,5-dihydroxyphenyl)oxidanyl guanosine triphosphate (compound 3 in Table 1) to the NS5b RdRp of HCV genotype 4a model. The relative docking score to GTP is 1.18. This docking study was performed after 150 ns of the MDS run. During ligand docking, different polar contacts established, from which one H-bond is formed (Fig. 5 B) between the amino acid D1 and the hydroxyl group of guanine heterocycles. The same occurred with other modified compounds (1, 2, 4–8) but with different

Table 3
Analysis of the best and worst docking score values obtained at different time steps of the MDS for Sofosbuvir utilizing PLIP web server.

Sofosbuvir best 10 docking values								
MDS time at which docking performed	H-bonds		Salt bridges		π-cation interaction		Hydrophobic interaction	
	No.	Amino acids involved	No.	Amino acids involved	No.	Amino acids involved	No.	Amino acids involved
250	3	D1, D99 and D100						
315	3	D99 (2) and D100						
320	2	D100 (2)						
600	3	D99 and D100 (2)						
615	1	R3	1	R3				
620	4	R3, D99 and D100 (2)	1	R3				
665	6	R3 (3) and D100 (3)						
670	3	R3, G98 and D99	1	R3	1	R3		
735	4	T2, K79 and D99 (2)	2	R3 (2)				
805	2	R61 and N72	2	K79 (2)				
Sofosbuvir worst 10 docking values								
485	1	D99						
565	2	D1 and R3						
570	2	R3 and N72						
595	5	D1, R3 (2), C97 and D99					1	D100
635	3	R3, D99 and D100	1	R3				
640	5	N72 (3), T73 and D99						
645	2	G98 and D99	1	K79				
650	4	R3, D99 and L101 (2)					1	D100
655	3	R3, D100 and L101						
975								



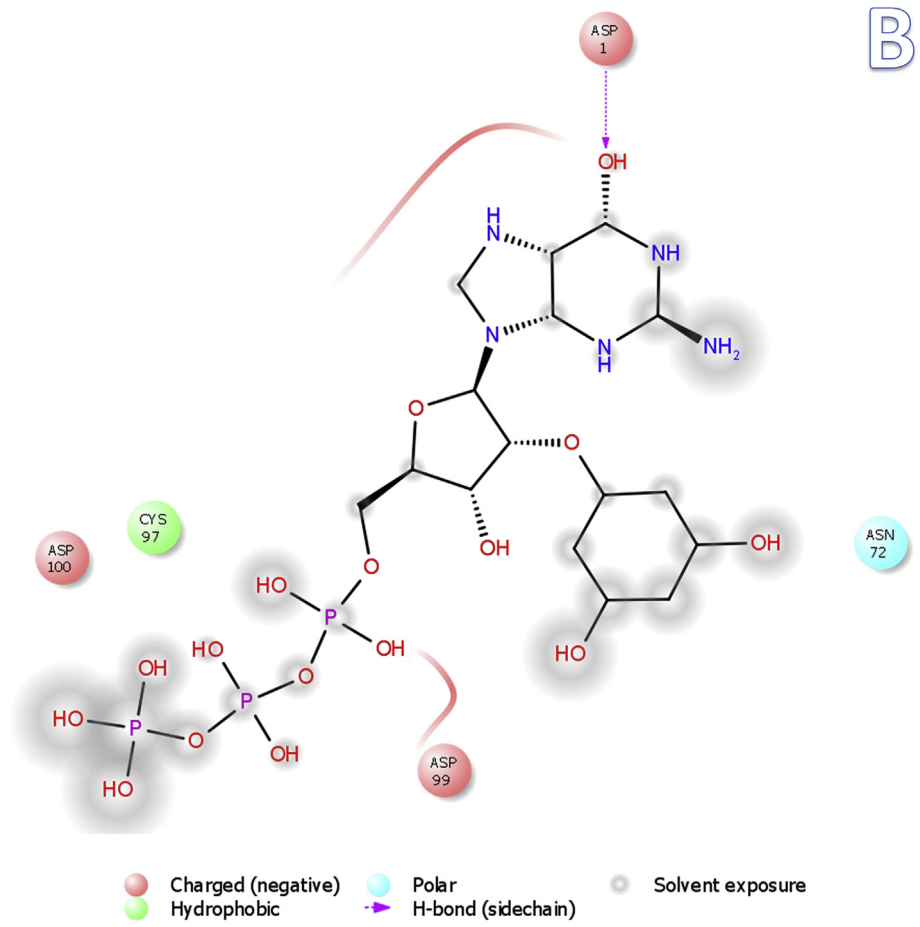


Fig. 5. (A) Structure of (3,5-dihydroxyphenyl) oxidanyl guanosine triphosphate (compound 3 in Fig. 1) docked into NS5b model showing some polar contacts (dashed yellow lines) between the ligand and D1 and D100 of the protein. A representation made by PyMOL software. (B) 2D representation for the same docked compound to the protein active site showing the residues surround the ligands and only one H-bond formed between D1 and the hydroxyl group of the heterocycle of the nucleotide inhibitor. Gray smudges represent water accessible atoms or groups. Active site pocket amino acids are represented in one letter code and by different colors according to the scheme in the figure bottom. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

number of H-bonds and various amino acids involved.

Figure S3 shows the 2D representation for all compounds docked into the NS5b active site at 150 ns of the MDS. It shows how H-bonds were formed and the polarity of the amino acids around the ligands after 150 ns of MDS run. Besides; gray smudges represent water accessible atoms or groups. At least one H-bond is formed between the suggested compounds and the binding site. Compound 4 is an exception which interacts with the hydrophobic residues C97 and V102 through its modified substitution.

A comparison between IDX-184 and other drugs used against HCV NS5b is made using our model representing genotype 4a. The relative average docking scores in the course of 1 μs of MDS is 0.98 ± 0.09 for

IDX-184. This is better than that of the approved drug, sofosbuvir (0.77 ± 0.11), the drug R7128, which is under clinical trials, (0.69 ± 0.20), and ribavirin (0.84 ± 0.11). All the suggested compounds can bind effectively to RdRp of HCV genotype 4a model without any effect of conformational flexibility. Besides, the binding energies are as good as the physiological molecule GTP; hence, a competitive inhibition is suggested.

4. Conclusion

Finding a potent drug is the primary goal of drug designers. Understanding the dynamical behavior of a target protein is an important task that should be tackled in drug design. The most mobile elements of the polymerase model we introduce here don't affect nucleotide inhibitor binding. In this study, we present eight novel guanosine-derivative compounds as potent as IDX-184 in inhibiting genotype 4a HCV and more efficiently than sofosbuvir and ribavirin. In a previous study, IDX-184 proved to be a successful candidate against all HCV genotypes. Further experimental validation is suggested as a next step for developing, novel, potent nucleotide inhibitors against the Egypt prevalent subtype of HCV.

Declaration of competing Interest

All the authors declare that no conflict of interest for this work.

Acknowledgment

Dr. Ali Hassanali is appreciated for his guidance and his revision. Prof. Dr. Wael Elshemey is appreciated for performing the docking calculations on his computational facility. MDS was carried out on the Cy-Tera supercomputer facility of Cyprus Institute of Science under the production run of the project “pro15b114s1”. This work is done during the Junior associate award granted to Abdo Elfiky 2016–2021 of the Abdus Salam International Center for Theoretical Physics (ICTP), Trieste, Italy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lfs.2019.116958.

References

[1] Q.L. Choo, G. Kuo, A.J. Weiner, L.R. Overby, D.W. Bradley, M. Houghton, Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome, Science 244 (1989) 359–362.
[2] R. De Francesco, L. Tomei, S. Altamura, V. Summa, G. Migliaccio, Approaching a new era for hepatitis C virus therapy: inhibitors of the NS3-4A serine protease and the NS5B RNA-dependent RNA polymerase, Antivir. Res. 58 (2003) 1–16.
[3] A.A. Elfiky, W.M. Elshemey, W.A. Gawad, 2′-Methylguanosine prodrug (IDX-184), phosphoramidate prodrug (sofosbuvir), diisobutyryl prodrug (R7128) are better than their parent nucleotides and ribavirin in hepatitis C virus inhibition: a molecular modeling study, J. Comput. Theor. Nanosci. 12 (2015) 376–386.
[4] D. Genovese, S. Dettori, C. Argentini, U. Villano, P. Chionne, M. Angelico, et al., Molecular epidemiology of hepatitis C virus genotype 4 isolates in Egypt and analysis of the variability of envelope proteins E1 and E2 in patients with chronic hepatitis, J. Clin. Microbiol. 43 (2005) 1902–1909.
[5] S.M. Lemon, J.A. McKeating, T. Pietschmann, D.N. Frick, J.S. Glenn, T.L. Tellinghuisen, et al., Development of novel therapies for hepatitis C, Antivir. Res. 86 (2010) 79–92.
[6] C. Sarrazin, C. Hézode, S. Zeuzem, J.-M. Pawlotsky, Antiviral strategies in hepatitis C virus infection, J. Hepatol. 56 (2012) S88–S100.
[7] P.L. Yang, M. Gao, K. Lin, Q. Liu, V.A. Villareal, Anti-HCV drugs in the pipeline, Curr. Opin. Virol. 1 (2011) 607–616.
[8] M.M. Bahgat, A.A. Ibrahim, D.N. Abd-Elshafy, A.A. Mesalam, H.E. Gewaid, A.A. Ismaeil, et al., Characterization of NS3 protease from an Egyptian HCV genotype 4a isolate, Arch. Virol. 154 (2009) 1649–1657.
[9] A.A. Elfiky, A.G.W., M. EW, Hepatitis C viral polymerase inhibition using directly acting antivirals, a computational approach, in: A.M. (Ed.), Software and Techniques for Bio-Molecular Modeling, Austin publishing group, USA, 2016, p. 197.
[10] D.J. Graham, M. Stahlhut, O. Flores, D.B. Olsen, D.J. Hazuda, R.L. LaFemina, et al., A genotype 2b NS5B polymerase with novel substitutions supports replication of a chimeric HCV 1b:2b replicon containing a genotype 1b NS3-5A background, Antivir. Res. 69 (2006) 24–30.
[11] T. Asselah, Daclatasvir plus sofosbuvir for HCV infection: an oral combination therapy with high antiviral efficacy, J. Hepatol. 61 (2014) 435–438.
[12] R. Gonzalez-Grande, M. Jimenez-Perez, C. Gonzalez Arjona, J. Mostazo Torres, New approaches in the treatment of hepatitis C, World J. Gastroenterol. 22 (2016) 1421–1432.
[13] N.A. Saleh, A.A. Ezat, A.A. Elfiky, W.M. Elshemey, M. Ibrahim, Theoretical study on modified boceprevir compounds as NS3 protease inhibitors, J. Comput. Theor. Nanosci. 12 (2015) 371–375.
[14] S. Yan, G. Larson, J.Z. Wu, T. Appleby, Y. Ding, R. Hamatake, et al., Novel thiazolones as HCV NS5B polymerase allosteric inhibitors: further designs, SAR, and X-ray complex structure, Bioorg. Med. Chem. Lett 17 (2007) 63–67.
[15] A.A. Elfiky, W.M. Elshemey, IDX-184 is a superior HCV direct-acting antiviral drug: a QSAR study, Med. Chem. Res. 25 (2016) 1005–1008.
[16] R. De Francesco, A. Carfí, Advances in the development of new therapeutic agents targeting the NS3-4A serine protease or the NS5B RNA-dependent RNA polymerase of the hepatitis C virus, Adv. Drug Deliv. Rev. 59 (2007) 1242–1262.
[17] A.A. Elfiky, Zika viral polymerase inhibition using anti-HCV drugs both in market and under clinical trials, J. Med. Virol. (2016) n/a-n/a.
[18] S. Yan, T. Appleby, G. Larson, J.Z. Wu, R. Hamatake, Z. Hong, et al., Structure-based design of a novel thiazolone scaffold as HCV NS5B polymerase allosteric inhibitors, Bioorg. Med. Chem. Lett 16 (2006) 5888–5891.
[19] A.A. Elfiky, W.M. Elshemey, W.A. Gawad, O.S. Desoky, Molecular modeling comparison of the performance of NS5b polymerase inhibitor (PSI-7977) on prevalent HCV genotypes, Protein J. 32 (2013) 75–80.
[20] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. 10 (2015) 845–858.
[21] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods 12 (2015) 7–8.
[22] W.M. Elshemey, A.A. Elfiky, W.A. Gawad, Correlation to protein conformation of Wide-angle X-ray Scatter parameters, Protein J. 29 (2010) 545–550.
[23] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[24] R. Luthy, Bowie JU, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[25] J. Pontius, J. Richelle, S.J. Wodak, Deviations from standard atomic volumes as a quality measure for protein crystal structures, J. Mol. Biol. 264 (1996) 121–136.
[26] Zhang C, Liu S, Zhou H, Zhou Y. The dependence of all-atom statistical potentials on structural training database. Biophys. J..86:3349-3358.
[27] N.A. Saleh, A.A. Elfiky, A.A. Ezat, W.M. Elshemey, M. Ibrahim, The electronic and quantitative structure activity relationship properties of modified telaprevir compounds as HCV NS3 protease inhibitors, J. Comput. Theor. Nanosci. 11 (2014) 544–548.
[28] K.L. Summers, A.K. Mahrok, M.D. Dryden, M.J. Stillman, Structural properties of metal-free apometallothioneins, Biochem. Biophys. Res. Commun. 425 (2012) 485–492.
[29] J.H. Lii, N.L. Allinger, Molecular mechanics. The MM3 force field for hydrocarbons. 3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons, J. Am. Chem. Soc. 111 (1989) 8576–8582.
[30] J.J.P. Stewart, Optimization of parameters for semiempirical methods. III Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi, J. Comput. Chem. 12 (1991) 320–341.
[31] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[32] W.L. DeLano, PyMOL (2002).
[33] S. Release, 1 Maestro. Schrödinger, LLC, New York, NY, USA, 2016.
[34] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, M. Schroeder, PLIP: fully automated protein–ligand interaction profiler, Nucleic Acids Res. 43 (2015) W443–W447.