



Article

Identifying Different Mutation Sites Leading to Resistance to the Direct-Acting Antiviral (DAA) Sofosbuvir in Hepatitis C Virus Patients from Egypt

Aly Atef Shoun ^{1,*}, Rania Abozahra ², Kholoud Baraka ², Mai Mehrez ³ and Sarah M. Abdelhamid ²¹ Microbiology and Immunology Department, Faculty of Pharmacy, Sinai University, El Arish 45518, Egypt² Microbiology and Immunology Department, Faculty of Pharmacy, Damanhour University, Damanhour 22511, Egypt; rania.ragab@pharm.dmu.edu.eg (R.A.); kholoud.baraka@pharm.dmu.edu.eg (K.B.); sara.magdy@pharm.dmu.edu.eg (S.M.A.)³ National Hepatology and Tropical Medicine Research Institute (NHTMRI), Cairo 11511, Egypt; maymehrezmd@yahoo.com

* Correspondence: aly.shoun@su.edu.eg

Abstract: The hepatitis C virus (HCV) is a major global health challenge and a leading cause of morbidity and mortality. Many direct-acting antivirals (DAAs) target essential macromolecules involved in the virus' life cycle. Although such DAAs achieve great success in reducing the viral load in genotype 1 infections, other genotypes demonstrate different levels of response. This study focused on mutation sites associated with patients with genotype 4a infections that failed to respond to treatment with sofosbuvir. The genotyping of HCV samples from patients with virological failure, and responder patients, was conducted using Geno2Pheno webserver-based full NS5B sequences. We constructed 3D structural models for all the samples and used structural analysis to investigate the effect of amino acid substitution on the observed resistance to SOF-based treatment, and the docking of sofosbuvir into the active sites of the 10 models was performed. Finally, 10 molecular dynamic (MD) simulation experiments were conducted to compare the stability of the 3D models of the resistant samples against the stability of the 3D models of the responder samples. The results highlighted the presence of HCV subtype 4a in all ten samples; in addition, an amino acid (aa) substitution in the palm region may hinder HCV polymerase activity. In this study, we provide evidence that a mutation in the NS5B gene that induces resistance to sofosbuvir in patients with the S282T/C/R mutant virus is present in the Egyptian population. Overall, the docking and MD results support our findings and highlight the significant impact of the identified mutations on the resistance of HCV NS5B RNA-dependent RNA polymerase to direct-acting antivirals (DAAs).



Citation: Shoun, A.A.; Abozahra, R.; Baraka, K.; Mehrez, M.; Abdelhamid, S.M. Identifying Different Mutation Sites Leading to Resistance to the Direct-Acting Antiviral (DAA) Sofosbuvir in Hepatitis C Virus Patients from Egypt. *Microorganisms* **2022**, *10*, 679. <https://doi.org/10.3390/microorganisms10040679>

Academic Editor: Isabelle Chemin

Received: 22 February 2022

Accepted: 17 March 2022

Published: 22 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: HCV; RNA-dependent RNA polymerase; RNA sequencing; resistance; molecular dynamics

1. Introduction

HCV is a global health threat, with more than 71 million cases annually, and is the primary cause of death for nearly 399,000 people each year [1]. Chronic HCV infections are associated with the development of liver cirrhosis, hepatocellular carcinoma, liver failure, and death [1,2].

HCV is an enveloped single-stranded, positive-sense RNA virus (+ssRNA virus) belonging to the Flaviviridae family and the *Hepacivirus* genus [3]. Its genome is approximately 9.6 kb in length, and it shows genetic diversity, with eight confirmed genotypes, numbered 1 to 8, and 86 subtypes [4,5]. HCV genotyping plays a crucial role in molecular epidemiology and aids in determining appropriate DAAs and treatment durations [6]. Genotype 1 of HCV is the most common type in the United States, Europe, and Japan, accounting for approximately 60% of total global infections [7]. A high incidence of genotype

2 is found in Western Africa [7], whereas HCV GT3 predominates in Southeast Asia. Genotype 4 is broadly disseminated in Egypt, with limited infections elsewhere in the world. Genotypes 5, 6, and 7 are mainly prevalent in South Africa, Southeast Asia, and Central Africa, respectively. Egypt shows the highest prevalence of HCV globally, with nearly 15% of the total population suffering from HCV, and genotype 4a is the most common subtype [8,9]; this is attributable to improper parenteral injections during antischistosomiasis campaigns that continued from the 1920s to 1980s [10].

The HCV genome encodes a polyprotein of nearly 3000 amino acids, divided into 10 viral proteins. Seven of these proteins are non-structural: P7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B. The three structural proteins are Core, envelope E1, and E2 [11]. HCV non-structural 5B (NS5B), an RNA-dependent RNA polymerase, comprises ~591 amino acids and located at the C-terminus of the polyprotein chain; it is responsible for the replication of the positive-strand genomic RNA [12–14]. Many inhibitors that target this enzyme are available. The pan-genotypic inhibitor sofosbuvir initiated a new era in the treatment of HCV infections after its approval in 2013 [15]. RNA-dependent RNA polymerase (RdRp) lacks proofreading ability, which may be the main source of mutations in the HCV genome [16]. HCV NS5B has no match in mammalian cells and is thus a validated target for antiviral therapy with few target-related side effects. HCV NS5B is 65 KDa, with characteristic palm, thumb, and finger domains [17]. The palm domain contains the active site, in addition to four allosteric sites: palm I, palm II, thumb I, and thumb II [17].

Several DAAs for treating HCV have been developed in the last eight years as an effective, safe, and well-tolerated combination with very high sustained virological responses (SVRs)—defined as HCV RNA becoming undetectable 12 or 24 weeks after the end of treatment [18]. After the successful treatment of HCV-infected patients with oral DAA, the level of HCV RNA remains undetectable [19]. However, in some patients, the DAA does not completely eradicate the HCV RNA [20]. A DAA's activity varies among HCV genotypes due to structural differences among the protein targets [20]. The variable response is mainly due to drug-resistance-associated substitution (RAS). RAS is found in quasi-species in some HCV-infected patients. These variants may not be sensitive to DAAs [21]. Most data on HCV drug resistance relate to genotypes 1 and 3; however, genotype 4 infects approximately 18 million people in the Middle East and North Africa and is not sufficiently studied [22,23].

The purpose of this study was to identify the mutations responsible for resistance to sofosbuvir-based treatment for HCV genotype 4a, and to characterize their role in resistance to sofosbuvir using homology modeling, docking, and molecular dynamic simulations.

2. Patients and Methods

2.1. Blood Sampling

Thirty-four samples were collected from the National Hepatology and Tropical Medicine Research Institute. Informed consent was obtained from each patient before taking the sample. All the available samples and patient data were collected with informed ethical consent and Institutional Review Board (IRP) approval. The Damanhur Local Ethics Committee approved the study (code 1121PM24F) and administered the ethical guidelines from the 1975 Declaration of Helsinki.

Initially, 14 males and 20 females, aged 41–79 years, with viral loads ranging from 307,000 to 18,969,000 IU/L determined using a COBAS® AmpliPrep/Quantitative Test (v2.0), following the manufacturer's instructions, were included. Other serum samples from the same patients were collected and were centrifuged at 6000 RPM within 2 h of sample collection; 0.5–1 mL of serum was transferred using a sterile pipette into a clean, sterile, previously labeled 1.5 mL Eppendorf tube.

2.2. Viral RNA Extraction

The RNA extraction used an Invitrogen Viral RNA kit (Invitrogen, ThermoFisher Scientific, Waltham, MA, USA), following the manufacturer's instructions. The RNA was subsequently reverse transcribed. For the cDNA preparation, a High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific, Waltham, MA, USA) was used, following the manufacturer's instructions, except for adding primers. We used 2 μ L of a random primer, 2 μ L of forward primer region 1, and 2 μ L of reverse primer region 2.

2.3. Primer Sequences and PCR Program

The NS5B region size is approximately 1773 bp. Therefore, it was impossible to amplify and sequence the entire region as one fragment. Instead, we implemented an overlapping amplicon strategy to cover the entire sequence of this region (Table 1).

Table 1. The primers used in this work.

Primer Name	Sequence	Reference
G4R1F	TWGTAACACCCTGTGCDGCTGAAG	This study
G4R1R	CCGCAAACCATAGTGCACT	This study
G4R2F	CTGAGAGACTGCACYATGCTYGT	This study
G4R2R	CCCTARGGTCGGAGTGTTAAGCT	This study

F, forward primers; R, reverse primers.

2.4. PCR Amplification of the NS5B Region

A specific PCR product was designed to amplify the entire NS5B region from the synthesized cDNA. The PCRs included 25 μ L of a DreamtaqHot Start PCR master mix kit (Thermo Scientific, Waltham, MA, USA), 2 μ L of a random primer, 2 μ L of the forward primer, 2 μ L of the reverse primer, and 4 μ L of each cDNA sample. The reaction was made up to a final volume of 50 μ L with water. The reactions were loaded into a thermal cycler (Multigene, Labnet, Edison, NJ, USA) under the following conditions: for the Maxima Hot Start PCR master mix kit, an initial enzyme activation for 5 min at 95 °C was followed by 38 cycles of 45 s at 95 °C for denaturation, 45 s at a selected annealing temperature, and 1–2 min at 72 °C for extension. The final extension was performed for 10 min at 72 °C. The PCR amplification product was electrophoresed on an agarose gel before PCR purification with a 1.8% agarose gel in TAE buffer. A 100 bp DNA ladder was used (Figure S1).

2.5. Multiple Sequence Alignment

A total of 2202 HCV whole-genome sequences were aligned using the MUSCLE MSA tool, and the resulting file was used in the design of primers. We focused only on genotype 4a, the genotype that displays the highest incidence in Egypt.

2.6. Sanger Sequencing

NS5B fragments were amplified using a GeneJet PCR purification kit (K0691, Thermo Scientific, Waltham, MA, USA). The purified amplicons were sequenced on both strands via bi-directional Sanger sequencing using a 3730 xl DNA analyzer.

2.7. Data Analysis

Data analysis was performed using the Geno2Pheno webserver.

2.8. Phylogenetic Analysis

The evolutionary distances were calculated using the Kimura 2-parameter method with a discrete gamma distribution using the MEGA-X software, and a phylogenetic tree was constructed using the neighbor-joining method. One thousand replicates were used to calculate the bootstrap values. A phylogenetic tree was built using around 50 HCV sequences from other HCV isolates obtained from the HCV and GenBank databases.

2.9. Homology Modeling

The sequences of the NS5B polymerases of HCV GT4a were generated by the Sanger sequencing of 10 samples from responder and resistant patients, in addition to a reference (wild-type); these were used to construct 10 corresponding homology models. The sequence of the wild-type form of HCV GT4a polymerase was retrieved from the UniProtKB database (entry 4a (Y11604)) (www.uniprot.org, accessed on 8 September 2021). Eleven sequences (ten samples and a reference wild-type) were incorporated into the identification of templates with high identity, using the BLASTp server [24] against PDB databases. The search conditions for high-identity templates were optimized by setting an E-value cutoff of 0.0001. Based on the degree of similarity and the identity between the query sequences and templates, the most similar template from the search was selected to build models. The BLOSUM 62 weight matrix algorithm (alignment score), with a gap penalty and extension of 10 and 0.5, respectively, was implemented as the pairwise alignment algorithm. Target–template aligned sequences were uploaded to the HHpred server to generate target–template alignment profiles (<http://toolkit.tuebingen.mpg.de/hhpred>, accessed on 8 September 2021). Finally, the ten profiles, and the crystal structure of the selected template, were fed into the modeling software to generate the required 3D model structures. The final refinement of the models was completed using the Modrefiner server (<https://zhanglab.ccmb.med.umich.edu/ModRefiner/>, accessed on 9 September 2021).

2.10. 3D Structure Validation

The final refined 3D models of the HCV GT4a NS5B polymerase were validated using several servers, including Procheck to generate Ramachandran plots and the Verify protein 3D server to score the models [25].

2.11. Docking

The 10 optimized HCV NS5B polymerase models' structures were used to perform the subsequent docking step. All the docking simulations were conducted using the MOE software [26]. The 10 enzymes and the sofosbuvir were prepared using the standard structure optimization protocol of the software. Moreover, the 10 enzymes were energy-minimized and the active sites were set as reported in the literature. The docking of sofosbuvir into the active sites of the 10 models was performed using triangular matcher as a placement method and London dg as a scoring function. The DS visualizer, available from Biovia Inc., was used to visualize and analyze the docking results through the generation of 2D interaction images.

2.12. Molecular Dynamics

2.12.1. Molecular Dynamics for the Unbound Models

All the molecular dynamics in this work were assessed using the NAMD/MOE molecular interface. After solvating all 10 3D models using the TIP3P solvation model, the models were energy-minimized using the steepest descend algorithm under the AMBER12, extended Huckel theory (EHT) force field for a maximum of 50,000 steps [27]. The typical workflow of the NAMD simulation was applied for all of the MD simulations, starting by equilibrating every model for 5 ns, and all the well-equilibrated models were subjected to a production stage for 50 ns without any constraints using a time step of 1 fs. The generated trajectories were saved every 10 ps [28]. Furthermore, a constant temperature of 310 K and a constant pressure of 1 atm were maintained by Langevin dynamics and a Nose Hoover Langevin piston, respectively [29]. The particle mesh Ewald method with a grid point density of 0.92 Å was used to calculate the long-range electrostatics on a cubic box with periodic boundary conditions of 65 Å [30]. A cutoff of 12 Å for van der Waals interactions and a switching distance of 10 Å were used for production runs [30]. Finally, all the recorded trajectories were used to calculate the RMSD for all of the residues to judge the model's stability.

2.12.2. Molecular Dynamics for Sofosbuvir with Samples 10 (Responder) and 3 (Resistant)

The same conditions for simulating the unbound models were applied to simulate sofosbuvir in complex with samples 10 (responder) and 3 (resistant) for 50 ns.

3. Results and Discussion

3.1. Sample Collection and Sequence Analysis

The presence of an HCV NS5B region was confirmed in 10 (29.4%) of the 34 samples using PCR. Five samples were from patients with virological failure, and the other five were from responder patients that were treated with a SOF-based regimen. After RNA extraction and cDNA synthesis, purification and sequencing were performed. For accurate genotyping, we sequenced the full-length NS5B, and the DNA Data Bank of Japan (DDBJ) was used for the submission of all of the ten positive HCV samples (Table 2).

Table 2. Accession numbers for ten novel genotypes.

Accession Number	Sample No.
LC632938	Sample 1
LC632939	Sample 2
LC632940	Sample 3
LC632941	Sample 4
LC632942	Sample 5
LC632943	Sample 6
LC699305	Sample 7
LC632944	Sample 8
LC632945	Sample 9
LC632946	Sample 10

The Geno2Pheno webserver was implemented for detecting susceptible mutation sites. All the samples belonged to genotype 4a, with a similarity of nearly 90% to the reference sequence 4a (Y11604). This finding is consistent with other studies reporting from different governorates in Egypt, such as Damietta, Sharkia, Ismailia, and Alexandria [23,31–35]. HCV 4a is the most common subtype in south Egypt, followed by subtypes 1g, 4l, 4n, and 4o [32]. Subtype 4a is the most common subtype in the governorate of Ismailia, followed by subtypes 1g and 4o [34]. The most common subtype in the Sharkia governorate is 4a, followed by 4o, 1g, and 4n [31]. In the Alexandria governorate, the most common subtype is 4a, followed by 4m, 4o, 4n, and 4p [33]. Because of the small number of patients in the current study, subtypes other than 4a were not detected.

SOF, which targets the NS5B region, is extensively used in treating HCV patients in Egypt [36]. Due to the short period of SOF use, it is not yet known if treated patients may show post-treatment relapse [6]. Thus, we selected patients who failed to achieve the required sustained virologic responses (SVRs) after full therapeutic courses of treatment with SOF. The 3D structure of HCV NS5B polymerase has three domains: fingers, a palm, and a thumb. The active site of the palm domain comprises the residues 188–227 and 287–370 [37]. The important residues in the palm region contain D220, D225, G317, D318, and D319. The motif B residues G283, T286, T287, and N291, which participate in sugar selection by RNA-dependent RNA polymerase, are also highly conserved [37]. For the right palm, residues 360–370 are important for maintaining a proper secondary structure, forming motif E of the palm domain [37]. The presence of an amino acid substitution in these residues may hinder HCV polymerase activity [38]. First, one sample displayed the amino acid substitution D318Y, and two samples demonstrated a D319V mutation. Second, one sample had amino acid substitutions of G283S and T286H, and two samples showed amino acid substitutions of T287P/C and N291Y/T. Third, one sample showed the following amino acid substitutions: L360P, E361Q, L362E, I363T, T364R, S365G, C366T, S367P, S368V, N369E, and V370S. Two samples displayed L320G/R polymorphism (Table 3). This site confers low resistance to SOF treatment. Moreover, C316A/L polymorphism was present

in two samples. Residues T532–R570 are responsible for regulating polymerase activity and lie in the NS5B C-terminus [38]. Sample 3 displayed eight amino acid substitutions, whereas sample 4 showed 21 amino acid substitutions: I5389L, A540P, G543A, R544K, T552N, G554A, Y555C, S556P, G557V, D559S, I560L, Y561T, H562S, S563K, V564P, S565L, H566G, A567E, R568I, P569R, and R570K. The presence of 21 amino acid substitutions indicates high variability in this region. A membrane anchor is formed from amino acids 571–591 in the NS5B C-terminus [38]. We recorded amino acid substitutions in three of the five samples, with 13 amino acid substitutions in this region: W571E, F572K, W573R, F574G, C575I, L576M, L577R, L578N, L579R, A580 Q/S, A581V, Y586F, and N590A. Most recent research has concentrated on a single catalytic site that can be made inactive by mutations affecting the S282 amino acid position, where S is replaced by T or, often, by C, which is mostly found in genotypes 1 and 2 [39]. Moreover, four of the five samples displayed substitutions T/C/R in position S282 (S282T/C/R) (Table 3).

Table 3. Amino acid variance in the NS5B region of nine HCV isolates.

4a ED43 Y11604	LC632938	LC632939	LC632940	LC632941	LC632942	LC632943	LC632944	LC632945	LC632946
235	A	V	V	V	V	V	T	V	T
241	Q	Q	H	H	H	H	R	R	R
254	T	T	T	T	T	T	A	A	A
261	Y	Y	F	Y	F	F	Y	Y	Y
267	H	H	L	H	H	H	H	H	H
270	R	R	L	T	E	K	K	K	K
275	G	G	G	G	G	L	G	G	G
282	S	T	R	R	R	C	S	S	S
285	Y	Y	Y	D	Y	I	Y	Y	Y
293	L	L	L	L	L	G	L	L	L
303	I	I	I	F	I	F	I	I	I
307	G	G	G	G	G	A	G	R	R
309	R	R	R	E	R	R	R	R	R
322	V	V	V	K	V	H	V	V	V
327	D	D	G	D	D	Y	D	D	D
333	N	N	N	A	N	N	N	N	N
336	L	L	L	L	L	K	L	L	L
360	L	L	L	L	L	P	L	L	L
361	E	E	E	E	E	Q	E	E	E
363	I	I	I	I	I	T	I	I	I
365	S	S	S	S	S	G	S	S	S
367	H	S	A	S	S	P	S	S	S
370	V	V	V	V	V	S	V	V	V
552	T	T	K	T	T	T	T	M	M
555	A	A	A	A	A	A	A	A	S
557	G	G	G	G	G	G	G	Q	G
559	D	D	D	D	D	D	D	H	D
561	Y	Y	Y	Y	Y	Y	Y	S	Y
566	H			Q			A	Q	A
569	P			T			R	T	R
570	R			P		P	Y	P	Y
573	L			R		L	L	R	L
574	L			P		Q	L	G	L
577	L			Q				R	
579	P			C					R
580	T			S					
590	A			L			A		

The therapeutic outcomes from targeting the HCV NS5B polymerase region are significantly affected by amino acid substitution. We constructed 3D structure models for all the samples, and these models were subjected to structural analysis to investigate the effects of amino acid substitution on the resistance to SOF.

3.2. Phylogenetic Analysis

Phylogenetic analysis was conducted on the NS5B sequences from the 10 analyzed samples. These isolates' full-length NS5B sequences were associated with 50 HCV subtype 4 sequences, including the reference sequences and all the sequences from the Egyptian HCV isolates, to establish a neighbor-joining phylogenetic tree (Figure 1), with high bootstrap values.

In terms of genotyping, gene sequencing analysis revealed that the 10 HCV isolates were all connected to HCV genotype 4, with variation in the subtypes. The phylogenetic tree showed two main branches: The first main branch included all our sequences, whereas the other sequences obtained from GenBank in addition to five sub-branches (all of them nearly genotype 4) showed differences when resistance analysis was conducted. The second main branch was divided into two sub-branches, including the genotype 4a (with a similarity to the closest reference above 90%), which included analysis of the full-length hepatitis C virus sequences obtained from GenBank.

The sequences of the ten positive HCV samples were genotyped using the NCBI BLAST software as follows:

The BLAST results for almost all of the samples, with accession numbers LC632938–LC632946, revealed that they were all HCV genotype 4a. Some of the most related sequences for those samples were MG454451, MG454460, MG60280, and MG453394, and all of these deposited samples were used in resistance analysis for patients with genotype 1–6 HCV infection treated with sofosbuvir/velpatasvir in phase III studies (ASTRAL-1, ASTRAL-2, and ASTRAL-3), with an identity of 92% [40]. KY608670: Hepacivirus C isolate 105, from France with genotype 4o NS5B gene, partial cds with an identity equal to 93% [41]. KY608675: Hepacivirus C isolate 1428 NS5B gene, partial cds, from USA with genotype 4o NS5B gene, with a 93% identity [42]. KY283130: Hepacivirus C strain EJ1, complete genome, with genotype 4a, isolate from Egypt with an 88% identity.

3.3. Homology Modeling

Homology modeling is critical because it is the only technique in the field of drug discovery that provides a 3D protein structure when a crystal or NMR structure is absent. Thus, the effect of sequence mutations in each sample on the resistance to SOF was assessed with 11 3D models constructed using the NS5B polymerase of genotype 1b PDB (3 hkw) as a template. This polymerase shows high identity and similarity with the target sequences (Figure 2 and Table 4). All of the servers used for validation indicated high-quality models. Ramachandran diagrams generated by Procheck [42] showed acceptable stereochemical properties for all the models (Figure 3). Furthermore, validation was ensured using Verify protein 3D, which was used to (1) assess the model's accuracy by analyzing the compatibility of an atomic 3D model with its own amino acid sequence (1D), (2) calculate the percentage of residues with an average score > 0.2 (a residue score > 0.2 is considered reliable), and (3) conduct a final assessment of the 3D structure (pass, warning, or fail) [42]. The results of the verification were all "pass" for each of the generated 3D models.

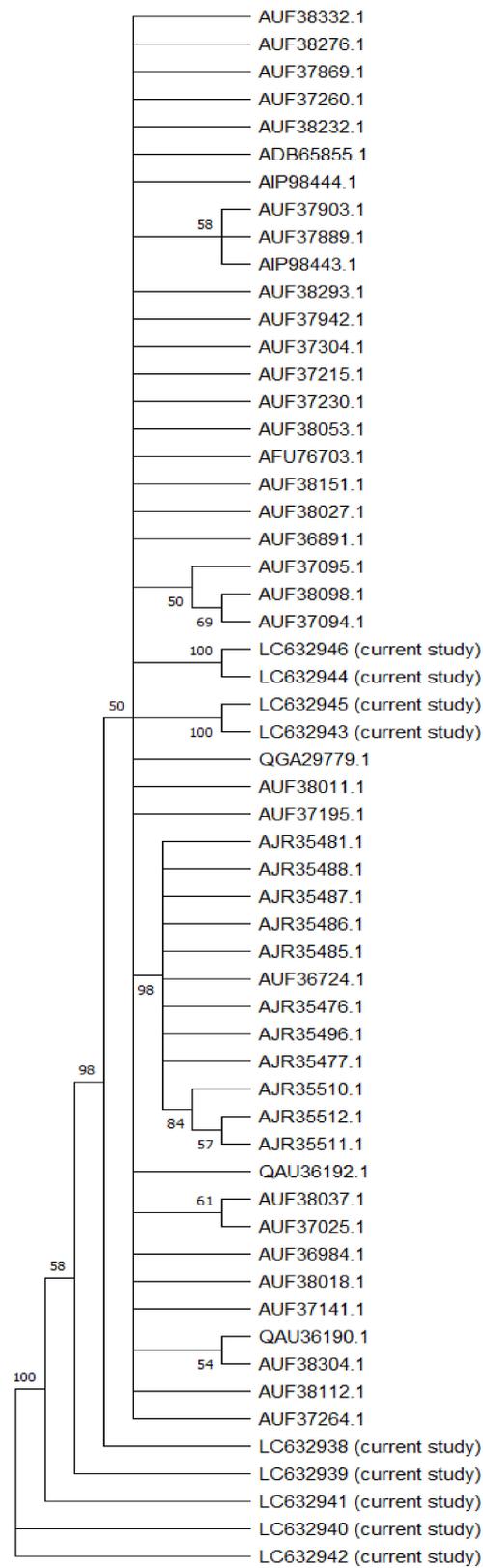


Figure 1. Neighbor-joining phylogenetic tree of the full-length sequence of NS5B of several HCV genotype 4 representatives. Bootstrap values based on 1000 replicates are shown next to the branches; bootstrap values of more than 50% only are shown. Sequences are labeled to the right of each branch in the order of GenBank accession number and isolate name.

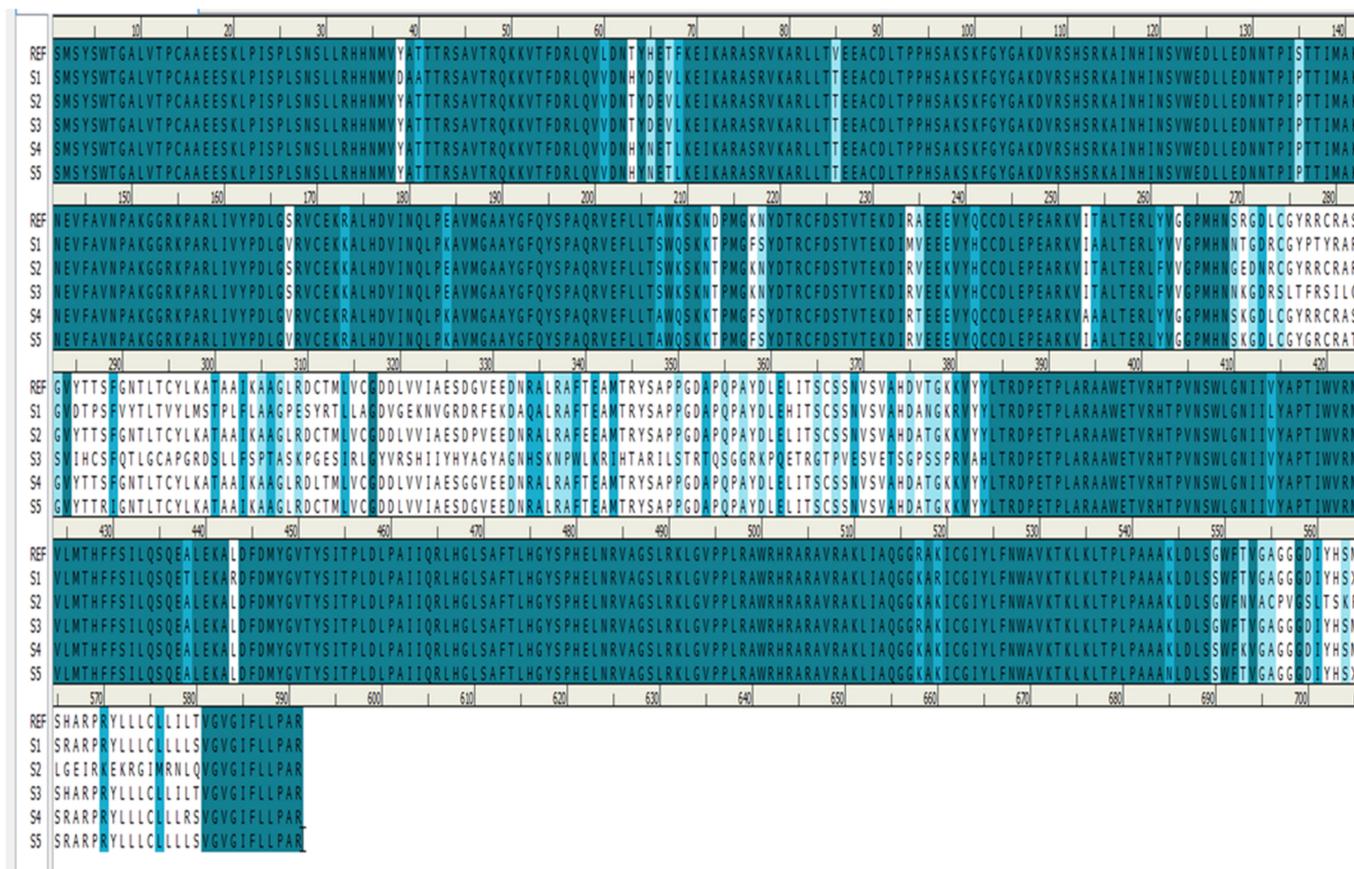


Figure 2. Sequence alignment between the reference (R) (Y11604) and the five resistant samples (S1–S5).

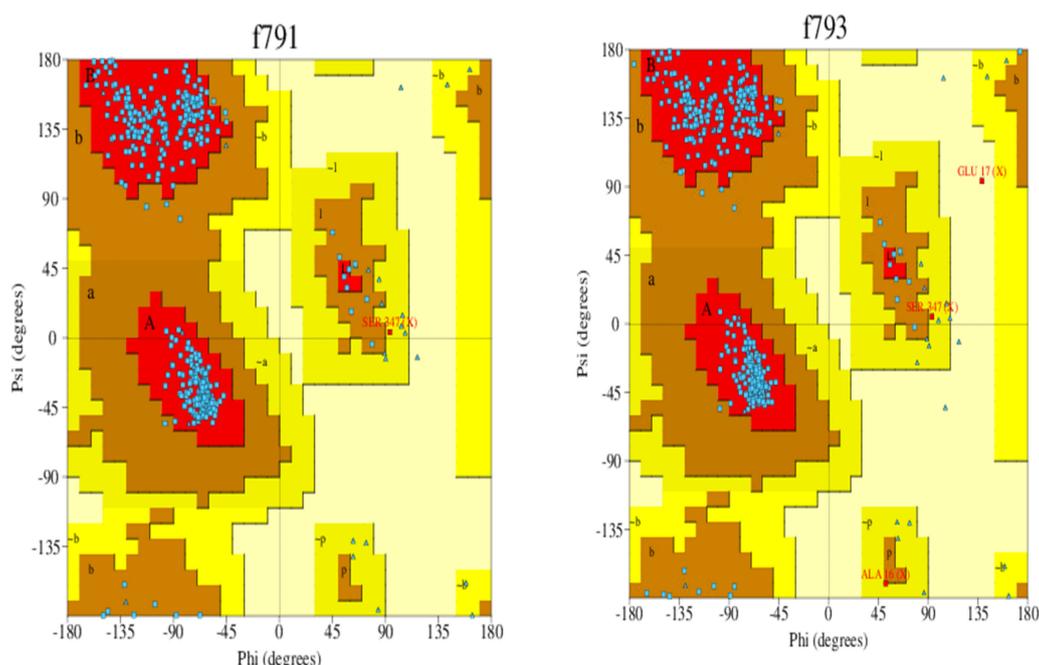


Figure 3. Cont.

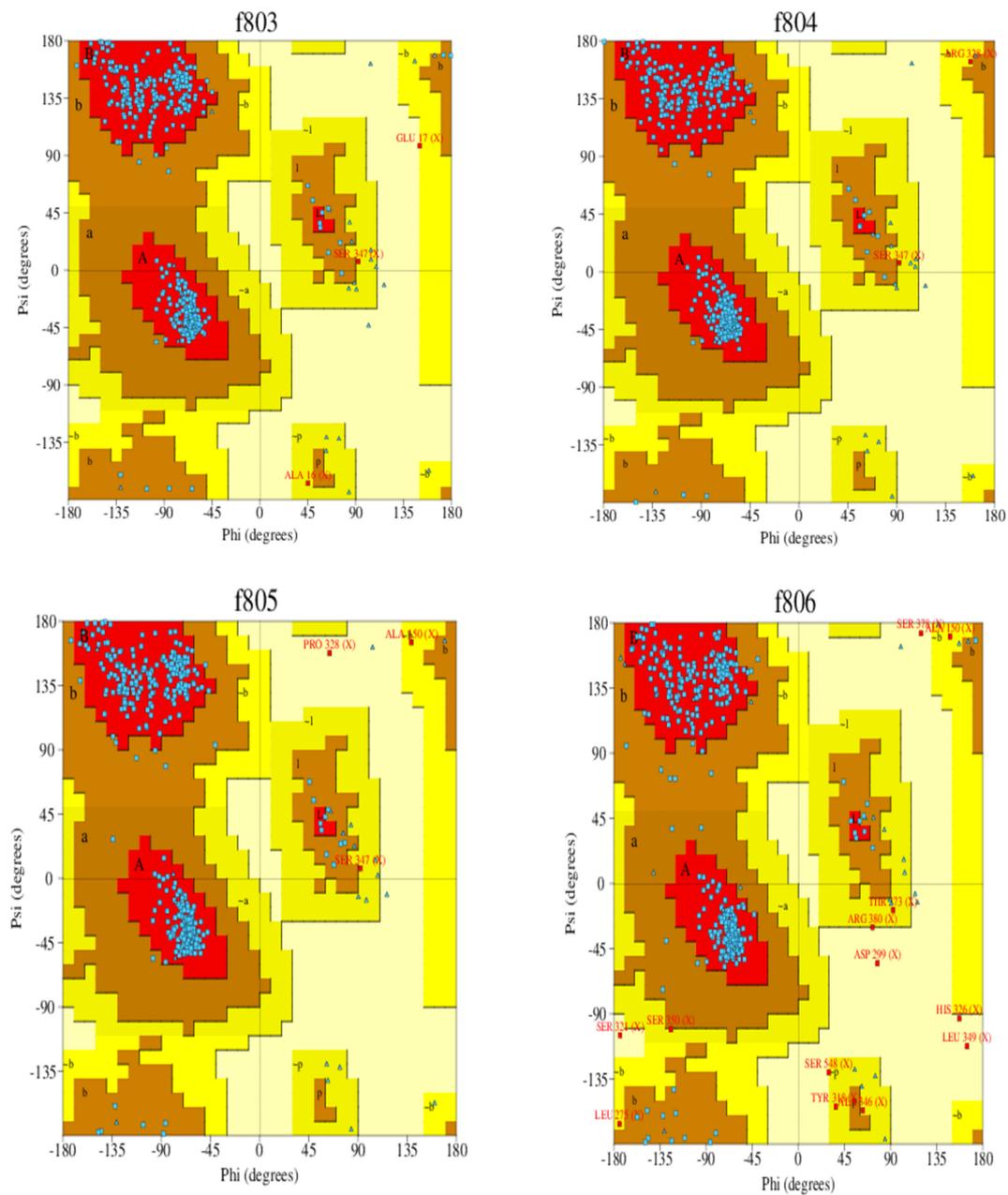


Figure 3. Ramachandran plot of the 3D models of HCV NS5B genotype 4a-resistant samples and reference exceeding 91.4% of the residues in the favored regions and around 5% of residues in the allowed regions. F791 (reference), f793 (sample 1), f803 (sample 2), f804 (sample 3), f805 (sample 4), and f806 (sample 5).

Table 4. Sequence identity and similarity between the template and five resistant samples.

Name	Identity	Similarity
Reference 4a (Y11604)	78%	1.28
Sample 1	78%	1.28
Sample 2	76%	1.26
Sample 3	72%	1.17
Sample 4	75%	1.25
Sample 5	67%	1.07

3.4. Structural Analysis of the Generated Models

Polymerase enzymes replicate genomes into multiple copies. This activity requires certain structural features that enable the enzyme to perform its function. For instance, the enzyme must accommodate both the template and the replicate in its active site. Thus, certain dimensions in the active site are required. Such structural features and dimensions affect the enzyme activity and impact the binding of drugs to their targets. We reported distances between Glycine 188 and Threonine 227, and Threonine 287 and Valine 370, of 22.96 and 26.04, respectively, in the wild-type model (Figure 4 and Table 5). Variations in these distances may weaken the binding of SOF. For example, in sample 5, the distances between Gly188 and Thr227, and Thr287 and Val370, were 27.13 and 28.85, respectively (Figure 5). This increase in active site size probably prevents SOF from strongly interacting with the active site and consequently reduces its activity. Thus, we found that some amino acid substitutions in our samples potentially underlie the resistance to SOF via the alteration and disruption of the polymerase's binding sites and structural features.

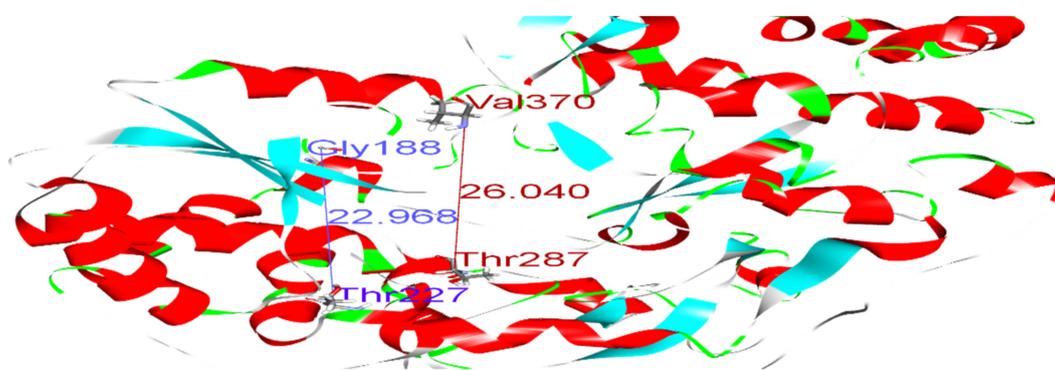


Figure 4. Distances between Gly188 and Thr227 (blue), and between Thr287 and Val370 (dark red), in the wild-type model.

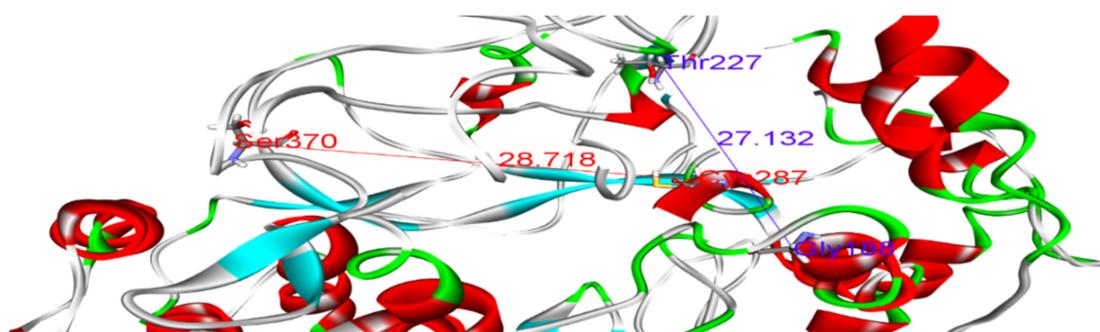


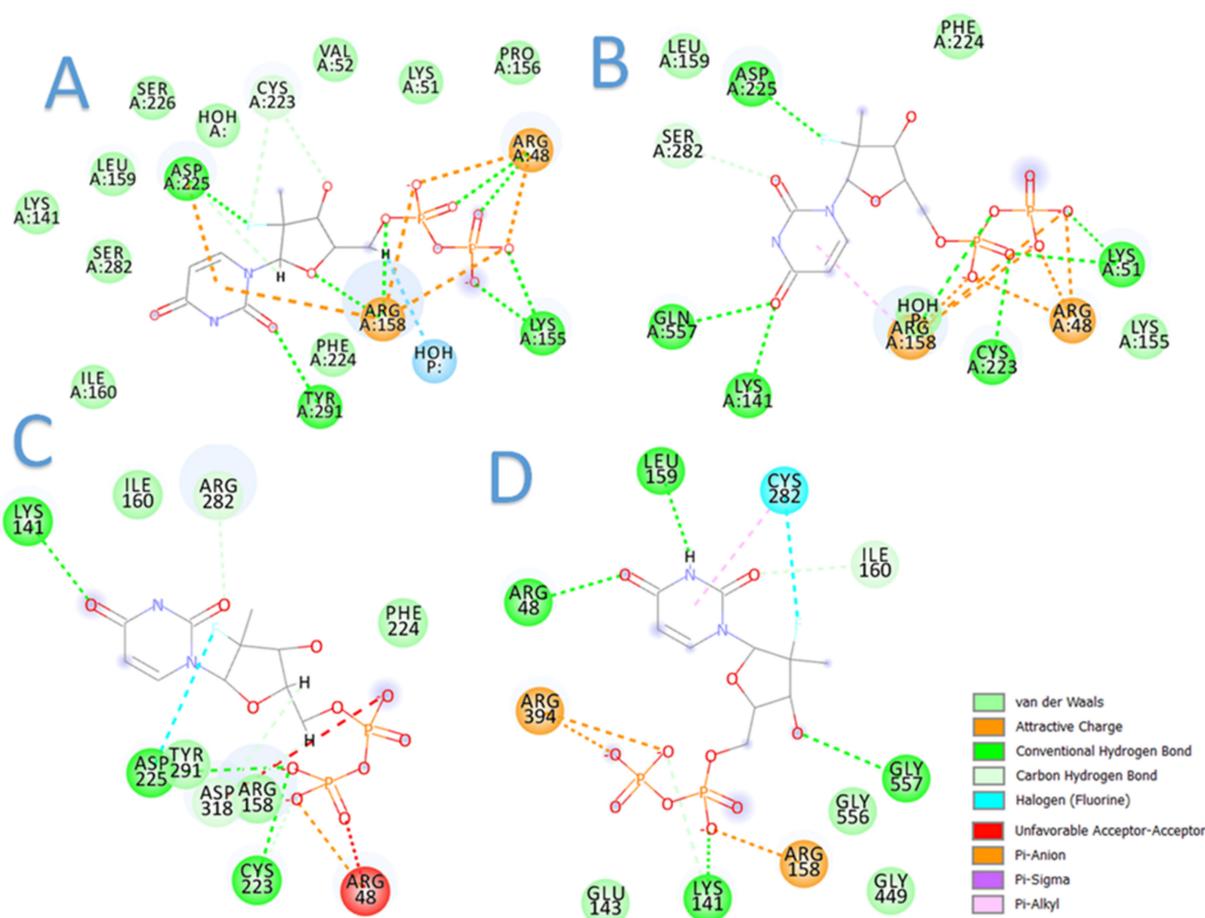
Figure 5. Distances between Gly188 and Thr227 (blue), and between Cys287 and Ser370 (dark red), in the sample 5 model.

3.5. Docking

The current docking studies were conducted to provide insights into plausible mechanisms of sofosbuvir resistance in terms of binding interactions. In this context, sofosbuvir was docked in the active sites of the models. The results are highly consistent with the clinical outcomes, in which the binding of sofosbuvir with the resistant samples was tempered by the found mutations (Figure 6). This was evidenced by the lower docking scores achieved with sofosbuvir compared to the higher docking scores of sofosbuvir with the responder samples. Table 6 summarizes the docking results for sofosbuvir with the 10 samples.

Table 5. Structural analysis and distances measured in each generated model.

Name	Distance from Gly188 to Thr227	Distance from Thr287 to Val370
Reference 4a (Y11604)	22.96 Å	26.04 Å
Sample 1	25.22	27.01
Sample 2	23.99	25.76
Sample 3	24.57	24.70
Sample 4	24.55	26.19
Sample 5	27.13	28.85

**Figure 6.** Two-dimensional diagrams of sofosbuvir interaction with HCV GT4a NS5B: (A) sample 9 (responder), (B) sample 10 (responder), (C) sample 3 (resistant), and (D) sample 5 (resistant).

3.6. Dynamic Simulation

3.6.1. Molecular Dynamics for the Unbound Models

Dynamic simulation is of great value for drug discovery and biomedical studies, such as for finding new drugs for attractive targets [25], evaluating inhibitor strength, and explaining the effects of certain mutations on drug-resistance profiles. The latter use was targeted to achieve our primary goal of MD simulation. Accordingly, 10 molecular dynamic experiments of 50 ns were performed. Five resistant and five responder samples were assessed. Significant effects of the mutations on the response of HCV NS5B RNA-dependent RNA polymerase to DAAs were found. The MD simulations showed that the 3D models of the responder samples had greater stability than those of the resistant samples. The calculated RMSD ranged from 1.55 to 2.23 Å and from 3.20 to 3.87 Å for the

responder and resistant samples, respectively (Figures 7 and 8 and Figure S2 and Table 7). The significant difference in the calculated RMSD for all of the residues of the most stable responder, sample 5 (1.8 Å), and the highest variable non-responder, sample 3 (3.4 Å), is depicted in Figure 9. The calculated RMSD of the responder sample was less than 2 Å throughout the 50 ns simulation, unlike the RMSD of the non-responder, which exceeded 2 Å. The reported mutations played a direct role in the increased dynamicity and flexibility of the HCV NS5B RNA-dependent RNA polymerase and diminished the ability of DAAs to form a stable complex with the enzyme binding site. In turn, this poor binding resulted in a failed virological response. Conversely, the 3D models from the responder samples showed relatively stable conformations, facilitating DAA–enzyme binding, yielding the desired efficacy.

Table 6. The docking scores of sofosbuvir with the 10 models of HCV GT4a NS5B.

Sample No.	Docking Score (KCal/Mole)
Sample 1 (resistant)	−16.5
Sample 2 (resistant)	−17.3
Sample 3 (resistant)	−15.8
Sample 4 (resistant)	−17.4
Sample 5 (resistant)	−16.2
Sample 6 (responder)	−22.7
Sample 7 (responder)	−21.9
Sample 8 (responder)	−22.3
Sample 9 (responder)	−22.9
Sample 10 (responder)	−23.6

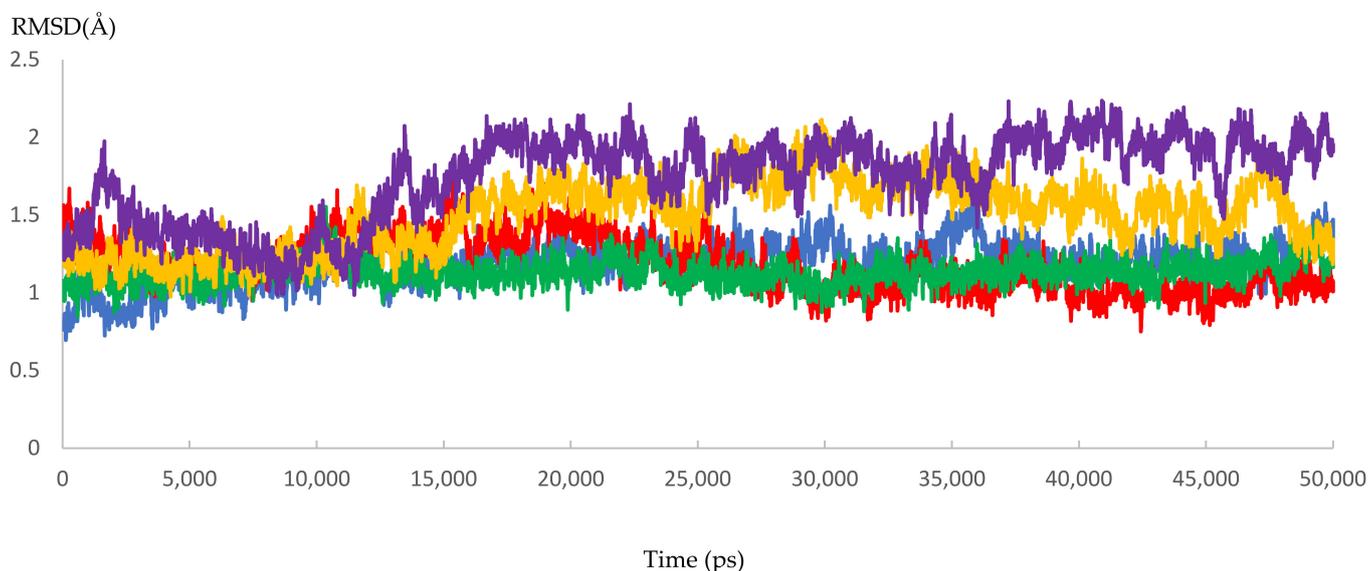


Figure 7. The RMSD of five dynamic simulation experiments for the HCV 4a NS5B responder samples. The green color represents sample 6, red represents sample 7, blue represents sample 8, yellow represents sample 9, and purple represents sample 10.

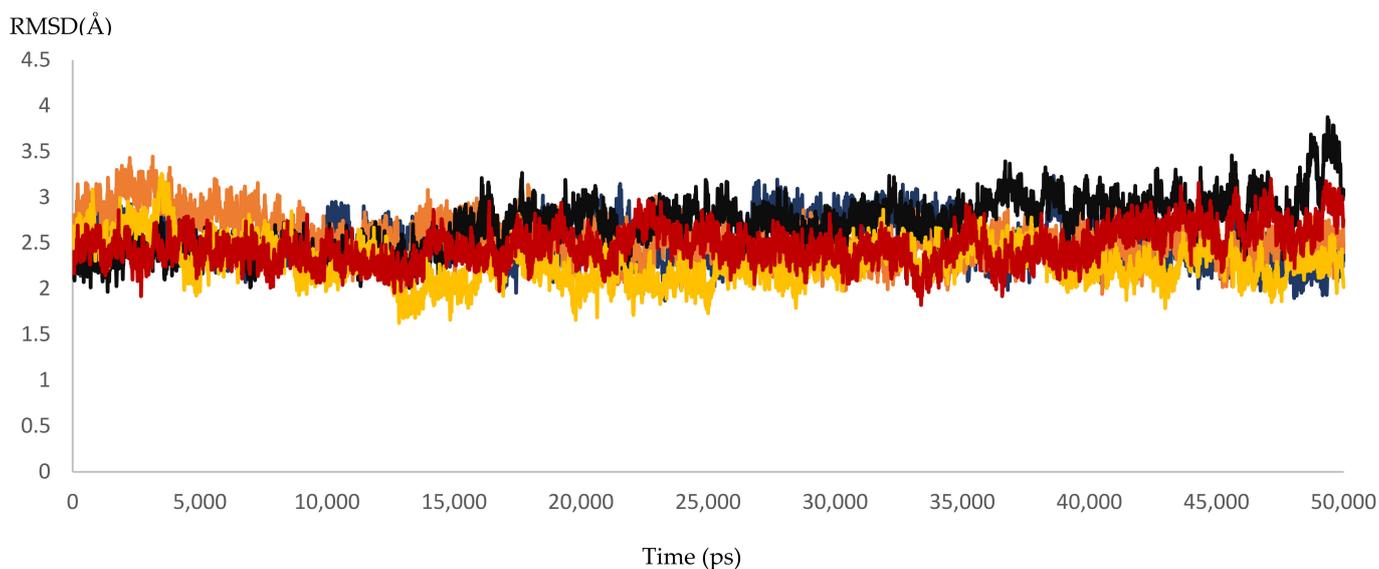


Figure 8. The RMSD of five dynamic simulation experiments for HCV 4a NS5B resistant samples. The blue color represents sample 1, brown represents sample 2, black represents sample 3, yellow represents sample 4, and red represents sample 5.

Table 7. Average RMSD for 10 molecular dynamic samples for 50 ns (five resistant and five responder samples).

Sample Number	Average (Å) \pm SD	Status
S1	2.524207 \pm 0.1	Resistant
S2	2.55801 \pm 0.11	Resistant
S3	2.700728 \pm 0.08	Resistant
S4	2.293318 \pm 0.09	Resistant
S5	2.471093 \pm 0.12	Resistant
S6	1.189008 \pm 0.06	Responder
S7	1.190665 \pm 0.07	Responder
S8	1.132385 \pm 0.07	Responder
S9	1.505865 \pm 0.05	Responder
S10	1.745415 \pm 0.08	Responder

3.6.2. Molecular Dynamics for Sofosbuvir with Sample 10 (Responder) and Sample 3 (Resistant)

To reinforce the docking results, the predicted binding poses of sofosbuvir with samples 10 (responder) and 3 (resistant) were simulated for 50 ns. The results were perfectly aligned with the clinical and docking results, as indicated by the RMSD calculations (Figure 10). The RMSD for sofosbuvir in complex with sample 10 was less than 2 Å throughout the simulation, indicating stable binding and highlighting the ability of sofosbuvir to inhibit the responder samples. By contrast, the RMSD of sofosbuvir in complex with sample 3 exceeded 3 Å, highlighting unstable binding, which explains the resistance to therapy with sofosbuvir.

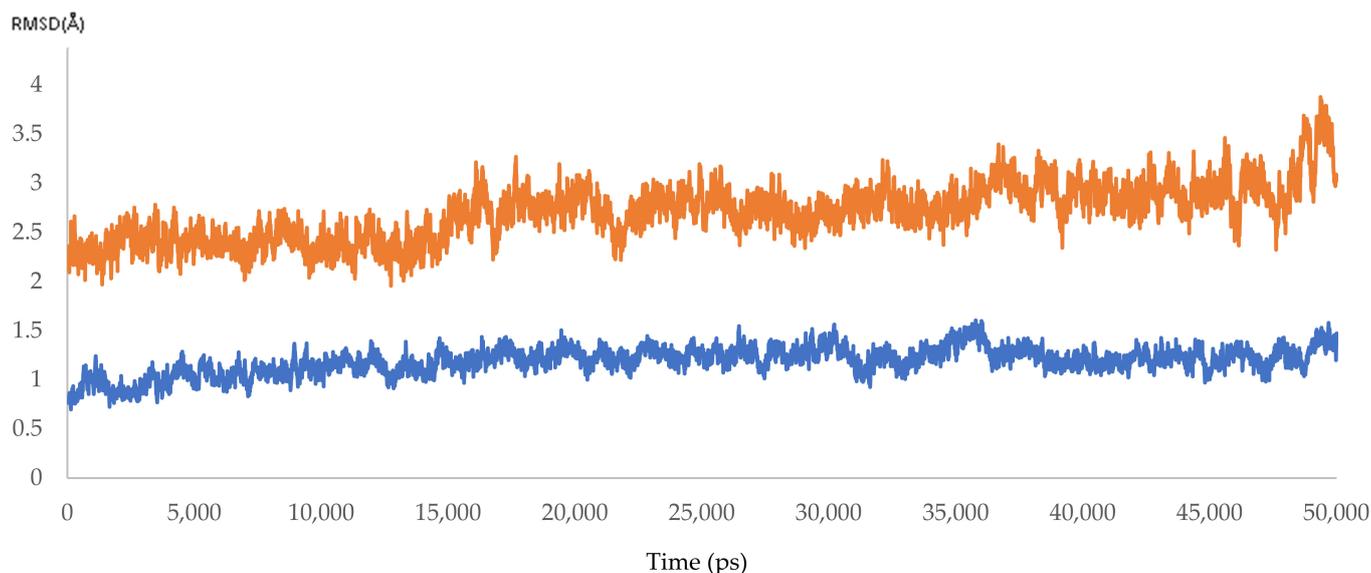


Figure 9. The RMSD of HCV 4a NS5B, comparing the highest resistant and responder samples (3.4 and 1.8 Å).

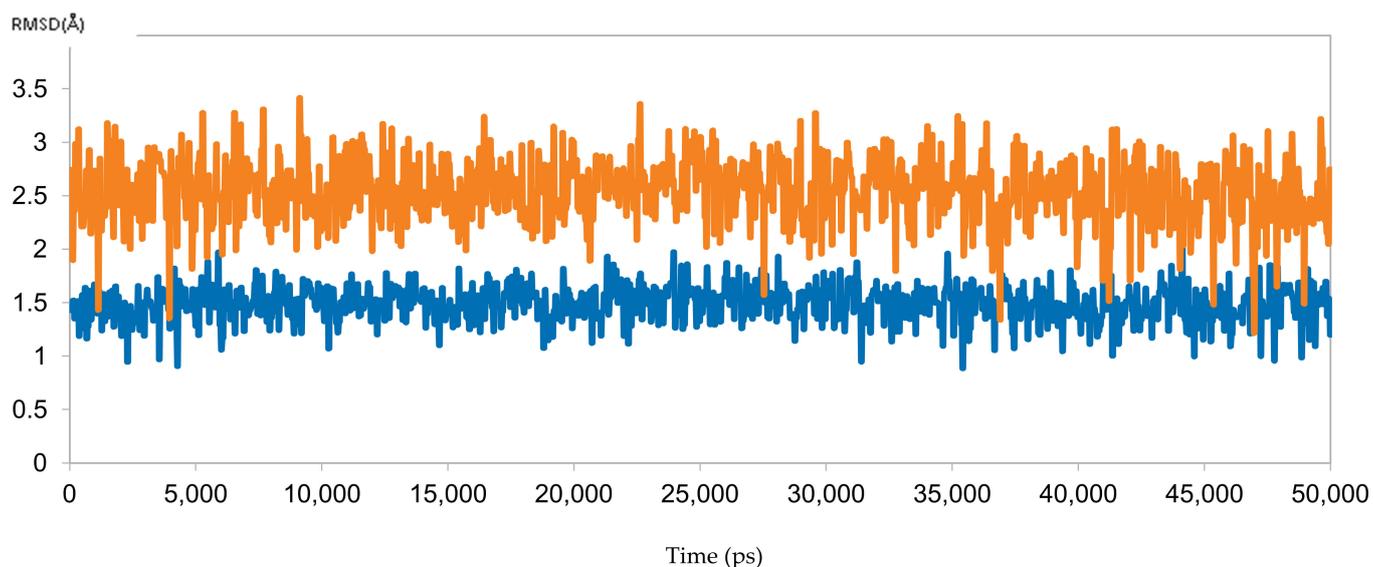


Figure 10. The RMSD of two dynamic simulations of sofosbuvir in complex with HCV 4a NS5B, comparing responder sample 10 (blue) and resistant sample 3 (brown).

4. Conclusions

We analyzed Egyptian HCV patients with virological failure by sequencing the NS5B region and recording amino acid substitutions. According to our findings, the most common HCV genotype among Egyptian patients was 4a. In addition, the findings demonstrate the relationship between the amino acid substitution and resistance to a SOF-based regimen in the NS5B of HCV genotype 4a-infected patients. To the best of our knowledge, this is the first study to report the application of molecular dynamic simulations to assess the resistance to sofosbuvir attributable to mutations in NS5B polymerase in genotype 4a. Collectively, the findings of this study encourage designing a specific treatment protocol for HCV genotype 4a-infected patients, an infection that is dominant in Egypt, to avoid resistance. Finally, the limitation of this study was the minimal sample size due to funding constraints. More research with larger sample size is required to verify the current findings.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/microorganisms10040679/s1>, Figure S1: Gel electrophoresis represents sample no.1 and sample no.2 with a total size of around 1800 bp. (R1) Region 1 with size 822 bp, and (R2) Region 2 with size 922 bp. Figure S2: Histogram showing the average RMSD of ten dynamic simulation experiments for HCV 4a NS5B resistant (S1:S5) and responder samples (S6:S10).

Author Contributions: All the authors contributed to the study's conception and design. The material preparation, data collection, and analysis were performed by S.M.A., A.A.S., K.B., M.M. and R.A. The first draft of the manuscript was written by A.A.S., and all the authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was self-funded.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by (Damanhur Local Ethics Committee, protocol code 1121PM24F).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Accession numbers for ten novel genotypes: LC632938, LC632939, LC632940, LC632941, LC632942, LC632943, LC699305, LC632944, LC632945, LC632946.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Centre for Disease Analysis. Web Annex B. WHO Estimates of the Prevalence and Incidence of Hepatitis C Virus Infection by WHO Region, 2015. In *Global Hepatitis Report 2017*; WHO: Geneva, Switzerland, 2018; p. 55. Available online: <https://apps.who.int/iris/bitstream/handle/10665/277005/WHO-CDS-HIV-18.46-eng> (accessed on 8 September 2021).
2. Cooke, G.S.; Lemoine, M.; Thursz, M.; Gore, C.; Swan, T.; Kamarulzaman, A.; DuCros, P.; Ford, N. Viral hepatitis and the Global Burden of Disease: A need to regroup. *J. Viral Hepat.* **2013**, *20*, 600–601. [[CrossRef](#)] [[PubMed](#)]
3. Lyles, D.; Rupprecht, C. Rhabdoviridae: The Viruses and their Replication. In *Fields Virology*; Lippincott-Raven: Philadelphia, PA, USA, 2006; pp. 1363–1408.
4. Khattab, M.A.; Ferenci, P.; Hadziyannis, S.J.; Colombo, M.; Manns, M.P.; Almasio, P.L.; Esteban, R.; Abdo, A.A.; Harrison, S.A.; Ibrahim, N.; et al. Management of hepatitis C virus genotype 4: Recommendations of An International Expert Panel. *J. Hepatol.* **2011**, *54*, 1250–1262. [[CrossRef](#)] [[PubMed](#)]
5. Hedskog, C.; Parhy, B.; Chang, S.; Zeuzem, S.; Moreno, C.; Shafran, S.D.; Borgia, S.M.; Asselah, T.; Alric, L.; Abergel, A.; et al. Identification of 19 novel hepatitis C virus subtypes-further expanding HCV classification. *Open Forum Infect. Dis.* **2019**, *6*, ofz076. [[CrossRef](#)] [[PubMed](#)]
6. Wiessing, L.; Ferri, M.; Grady, B.; Kantzanou, M.; Sperle, I.; Cullen, K.J.; Hatzakis, A.; Prins, M.; Vickerman, P.; Lazarus, J.V.; et al. Hepatitis C virus infection epidemiology among people who inject drugs in Europe: A systematic review of data for scaling up treatment and prevention. *PLoS ONE* **2014**, *9*, e103345. [[CrossRef](#)]
7. Manos, M.M.; Shvachko, V.A.; Murphy, R.C.; Arduino, J.M.; Shire, N.J. Distribution of hepatitis C virus genotypes in a diverse US integrated health care population. *J. Med. Virol.* **2012**, *84*, 1744–1750. [[CrossRef](#)] [[PubMed](#)]
8. Kouyoumjian, S.P.; Chemaitelly, H.; Abu-Raddad, L.J. Characterizing hepatitis C virus epidemiology in Egypt: Systematic reviews, meta-analyses, and meta-regressions. *Sci. Rep.* **2018**, *8*, 1661. [[CrossRef](#)]
9. Waheed, Y.; Saeed, U.; Anjum, S.; Afzal, M.S.; Ashraf, M. Development of global consensus sequence and analysis of highly conserved domains of the HCV NS5B protein. *Hepat. Mon.* **2012**, *12*, e6142. [[CrossRef](#)] [[PubMed](#)]
10. Hajarizadeh, B.; Grebely, J.; Dore, G.J. Epidemiology and natural history of HCV infection. *Nat. Rev. Gastroenterol. Hepatol.* **2013**, *10*, 553–562. [[CrossRef](#)]
11. Ahmad, I.; Holla, R.P.; Jameel, S. Molecular virology of hepatitis e virus. *Virus Res.* **2011**, *161*, 47–58. [[CrossRef](#)]
12. El-Sokkary, M.M.A.; Gotina, L.; Al-Sanea, M.M.; Pae, A.N.; Elbargisy, R.M. Molecular characterization of hepatitis C virus for developed antiviral agents resistance mutations and new insights into in-silico prediction studies. *Infect. Drug Resist* **2020**, *13*, 4235–4248. [[CrossRef](#)]
13. Krekulová, L.; Řehák, V.; Riley, L.W. Structure and functions of hepatitis C virus proteins: 15 years after. *Folia Microbiol.* **2006**, *51*, 665–680. [[CrossRef](#)] [[PubMed](#)]
14. Fatima, K.; Azhar, E.; Mathew, S.; Damanhour, G.; Qadri, I. Comparison of structural architecture of HCV NS3 genotype 1 versus Pakistani genotype 3a. *Biomed. Res. Int.* **2014**, *2014*, 749254. [[CrossRef](#)] [[PubMed](#)]
15. Lin, M.V.; Chung, R. Recent FDA approval of sofosbuvir and simeprevir. Implications for current HCV treatment. *Clin. Liver Dis.* **2014**, *3*, 65–68. [[CrossRef](#)] [[PubMed](#)]

16. Di Maio, V.C.; Cento, V.; Mirabelli, C.; Artese, A.; Costa, G.; Alcaro, B.S.; Perno, C.F.; Ceccherini-Silberstein, F. Hepatitis c virus genetic variability and the presence of ns5b resistance-Associated mutations as natural polymorphisms in selected genotypes could affect the response to ns5b inhibitors. *Antimicrob. Agents Chemother.* **2014**, *58*, 2781–2797. [[CrossRef](#)] [[PubMed](#)]
17. Lesburg, C.A.; Cable, M.B.; Ferrari, E.; Hong, Z.; Mannarino, A.F.; Weber, P.C. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Biol.* **1999**, *6*, 937–943. [[CrossRef](#)] [[PubMed](#)]
18. European Association for the Study of the Liver. EASL Recommendations on Treatment of Hepatitis C 2018. *J. Hepatol.* **2018**, *69*, 461–511. [[CrossRef](#)] [[PubMed](#)]
19. Stedman, C. Sofosbuvir, a NS5B polymerase inhibitor in the treatment of hepatitis C: A review of its clinical potential. *Ther. Adv. Gastroenterol.* **2014**, *7*, 131–140. [[CrossRef](#)]
20. Wyles, D.L.; Luetkemeyer, A.F. Understanding hepatitis C virus drug resistance: Clinical implications for current and future regimens. *Top Antivir. Med.* **2017**, *25*, 103–109.
21. Yamashita, T.; Takeda, H.; Takai, A.; Arasawa, S.; Nakamura, F.; Mashimo, Y.; Hozan, M.; Ohtsuru, S.; Seno, H.; Ueda, Y.; et al. Single-molecular real-time deep sequencing reveals the dynamics of multi-drug resistant haplotypes and structural variations in the hepatitis C virus genome. *Sci. Rep.* **2020**, *10*, 2651. [[CrossRef](#)]
22. Mahmud, S.; Chemaitelly, H.S.; Kouyoumjian, S.P.; Al Kanaani, Z.; Abu-Raddad, L.J. Key associations for hepatitis C virus genotypes in the Middle East and North Africa. *J. Med. Virol.* **2020**, *92*, 386–393. [[CrossRef](#)]
23. Abdel-Ghaffar, T.Y.; Sira, M.M.; El Naghi, S. Hepatitis C genotype 4: The past, present, and future. *World J. Hepatol.* **2015**, *7*, 2792–2810. [[CrossRef](#)] [[PubMed](#)]
24. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]
25. El-Hassab, M.A.E.M.; El-Bastawissy, E.E.; El-Moselhy, T.F. Identification of potential inhibitors for HCV NS5b of genotype 4a by combining dynamic simulation, protein–ligand interaction fingerprint, 3D pharmacophore, docking and 3D QSAR. *J. Biomol. Struct. Dyn.* **2020**, *38*, 4521–4535. [[CrossRef](#)]
26. El Hassab, M.A.; Fares, M.; Amin, M.K.A.H.; Al-Rashood, S.T.; Alharbi, A.; Eskandrani, R.O.; Alkahtani, H.M.; Eldehna, W.M. Toward the identification of potential α -ketoamide covalent inhibitors for sars-cov-2 main protease: Fragment-based drug design and mm-pbsa calculations. *Processes* **2021**, *9*, 1004. [[CrossRef](#)]
27. Mackerell, A.D.; Feig, M.; Brooks, C.L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J. Comput. Chem.* **2004**, *25*, 1400–1415. [[CrossRef](#)]
28. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)] [[PubMed](#)]
29. Feller, S.E.; Zhang, Y.; Pastor, R.W.; Brooks, B.R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621. [[CrossRef](#)]
30. Bhardwaj, V.K.; Singh, R.; Sharma, J.; Rajendran, V.; Purohit, R.; Kumar, S. Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors. *J. Biomol. Struct. Dyn.* **2020**, *39*, 3449–3458. [[CrossRef](#)]
31. Fakhr, A.E.; Pourkarim, M.R.; Maes, P.; Atta, A.H.; Marei, A.; Azab, M.; Van Ranst, M. Hepatitis C Virus NS5B Sequence-Based Genotyping Analysis of Patients from the Sharkia Governorate, Egypt. *Hepat. Mon.* **2013**, *13*, 12706. [[CrossRef](#)]
32. Elkady, A.; Tanaka, Y.; Kurbanov, F.; Sugauchi, F.; Sugiyama, M.; Khan, A.; Sayed, D.; Moustafa, G.; Abdel-Hameed, A.-R.; Mizokami, M. Genetic variability of hepatitis C virus in south Egypt and its possible clinical implication. *J. Med. Virol.* **2009**, *81*, 1015–1023. [[CrossRef](#)]
33. Genovese, D.; Dettori, S.; Argentini, C.; Villano, U.; Chionne, P.; Angelico, M.; Rapicetta, M. 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.* **2005**, *43*, 1902–1909. [[CrossRef](#)] [[PubMed](#)]
34. Youssef, A.; Yano, Y.; Utsumi, T.; Abd El-Alah, E.M.M.; Abd El-Hameed, A.E.E.S.; Serwah, A.E.H.A.; Hayashi, Y. Molecular Epidemiological Study of Hepatitis Viruses in Ismailia, Egypt. *Intervirology* **2009**, *52*, 123–131. [[CrossRef](#)] [[PubMed](#)]
35. El-Tahan, R.R.; Ghoneim, A.M.; Zaghoul, H. 5'â€2 UTR and NS5B-based genotyping of hepatitis C virus in patients from Damietta governorate, Egypt. *J. Adv. Res.* **2018**, *10*, 39–47. [[CrossRef](#)] [[PubMed](#)]
36. Goma, A.; Allam, N.; Elsharkawy, A.; El Kassas, M.; Waked, I. Hepatitis C infection in Egypt: Prevalence, impact and management strategies. *Hepat. Med.* **2017**, *9*, 17–25. [[CrossRef](#)] [[PubMed](#)]
37. Ago, H.; Adachi, T.; Yoshida, A.; Yamamoto, M.; Habuka, N.; Yatsunami, K.; Miyano, M. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure* **1999**, *7*, 1417–1426. [[CrossRef](#)]
38. El-Tahan, R.R.; Ghoneim, A.M.; Zaghoul, H. Dissection of two drug-targeted regions of Hepatitis C virus subtype 4a infecting Egyptian patients. *Virus Genes* **2020**, *56*, 564–581. [[CrossRef](#)]
39. Donaldson, E.F.; Harrington, P.R.; O'Rear, J.J.; Naeger, L.K. Clinical evidence and bioinformatics characterization of potential hepatitis C virus resistance pathways for sofosbuvir. *Hepatology* **2015**, *61*, 56–65. [[CrossRef](#)]
40. Hezode, C.; Reau, N.; Svarovskaia, E.S.; Doehle, B.P.; Shanmugam, R.; Dvory-Sobol, H.; Hedskog, C.; McNally, J.; Osinusi, A.; Brainard, D.M.; et al. Resistance analysis in patients with genotype 1-6 HCV infection treated with sofosbuvir/velpatasvir in the phase III studies. *J. Hepatol.* **2018**, *68*, 895–903. [[CrossRef](#)]

41. Welzel, T.M.; Bhardwaj, N.; Hedskog, C.; Chodavarapu, K.; Camus, G.; McNally, J.; Brainard, D.; Miller, M.D.; Mo, H.; Svarovskaia, E.; et al. Global epidemiology of HCV subtypes and resistance-associated substitutions evaluated by sequencing-based subtype analyses. *J. Hepatol.* **2017**, *67*, 224–236. [[CrossRef](#)]
42. Laskowski, R.A.; Rullmann, J.A.C.; MacArthur, M.W.; Kaptein, R.; Thornton, J.M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **1996**, *8*, 477–486. [[CrossRef](#)]