

Molecular Docking Investigation of the Binding Interactions of Macrocyclic Inhibitors with HCV NS3 Protease and its Mutants (R155K, D168A and A156V)

Ahmed A. Ezat · Nihal S. El-Bialy ·
Hamdy I. A. Mostafa · Medhat A. Ibrahim

Published online: 28 December 2013
© Springer Science+Business Media New York 2013

Abstract Hepatitis C Virus (HCV) non-structural protein 3 (NS3) protease drug resistance poses serious challenges on the design of an effective treatment. Substrate Envelope Hypothesis, “the substrates of HCV NS3/4A protease have a consensus volume inside the active site called substrate envelope” is used to design potent and specific drugs to overcome this problem. Using molecular docking, we studied the binding interaction of the different inhibitors and protein and evaluated the effect of three different mutations (R155K, D168A and A156V) on the binding of inhibitors. P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences have the best scores against the wild-type protein -204.506 and -206.823 kcal/mole, respectively. Also, charged P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences have low scores with the wild-type protein -200.467 and -203.186 kcal/mole, respectively. R155K mutation greatly affects the conformation of the compounds inside the active site. It inverts its orientations, and this is because the large and free side chain of K155 which restricts the conformation of the large P2–P4 macrocycle. The conformation of charged P2–P4 macrocycle of 3/4A hexapeptide sequence in wild-type, A156V and D168A proteins is nearly equal; while that of charged P2–P4 macrocycle of 4A/4B hexapeptide sequence is different. Nevertheless, these compounds have a slight

increase of Van der Waals volume compared to that of substrates, they are potent against mutations and have good scores. Therefore, the suggested drugs can be used as an effective treatment solving HCV NS3/4A protease drug resistance problem.

Keywords HCV · Drug resistance · Substrate envelope · Molecular docking · Macrocyclic and NS3 protease

Abbreviations

HCV	Hepatitis C Virus
NS3	Non-structural protein 3
NS4A	Non-structural protein 4A
FDA	Food and drug administration
PEG IFN- α	Pegylated interferon α
HIV-1	Human immunodeficiency virus-1
PM3	Parameterization model, version 3
MM3	Molecular mechanics 3
MO-G	Molecular orbital package
NS	Non-structural protein
PMF	Potential of mean force

1 Introduction

Drug resistance poses serious challenges on the design of effective drugs to treat quickly evolving diseases such as Hepatitis C Virus (HCV) [1–3]. Resistance emerges under the selection pressure of the drug, making the drug-resistant population to dominate and eventually the drug becomes ineffective [4]. There are different mechanisms that cause drug resistance such as mutation of the target which reduces its affinity for the drug [4]. This is a

A. A. Ezat (✉) · N. S. El-Bialy · H. I. A. Mostafa
Biophysics Department, Faculty of Science, University of Cairo,
Giza 11757, Egypt
e-mail: ezat@sci.cu.edu.eg

M. A. Ibrahim
Spectroscopy Department, National Research Centre, Dokki,
Cairo 12311, Egypt

common mechanism as the viruses replicate rapidly and their polymerases lack proof-reading activities [5]. These facts must be taken into account while designing effective drugs to avoid the emergence of resistance. One important strategy to design effective drugs is the “Substrate Envelope Hypothesis” of the target protein. According to this hypothesis, inhibitors fitting within the volume occupied by the enzyme’s natural substrate should tend to resist mutations. So, any mutations affect inhibitor binding would in turn affect substrate binding, rendering the enzyme ineffective [4, 5]. This hypothesis can be applied in the design of effective treatment to solve HCV NS3/4A protease drug resistance problem [5, 6].

The study of the structural, dynamic and energetic basis of drug resistance at the molecular level has validated this hypothesis and motivated its use in the design of specific and effective inhibitors of drug-resistant mutants [7–9].

So, the design of inhibitors sharing steric properties like those of enzyme’s natural substrates as well as having high binding affinity will contribute to overcome the drug resistance problem.

HCV, the etiological agent of the non-A non-B hepatitis, was identified at the molecular level in 1989. It causes chronic liver disease and develops into cirrhosis in 15–20 % of the infected individuals and leading eventually to liver failure and hepatocellular carcinoma [10]. HCV contains 3 structural and 7 non-structural proteins. One of the most important HCV enzymes is NS3, which forms a non-covalent complex with its 54-peptide NS4A cofactor [11]. NS3 comprises two functional domains, N-terminal protease and C-terminal helicase. NS3 protease domain is responsible for cleavage of the non-structural region of polyprotein to produce the mature non-structural proteins. It cleaves at four downstream regions (NS3/4A, NS4A/4B, NS4B/5A and NS5A/5B) [11]. A decapeptide substrate for the HCV NS3 protease, with 6 residues on the N-terminal side and 4 residues on the C-terminal side, would be described as NH₂-P6-P5-P4-P3-P2-P1-P1’-P2’-P3’-P4’-OH, with the scissile bond located between the P1 and P1’ residues [12]. Several specific NS3/4A inhibitors of different scaffolds have been designed to treat HCV. The most two important categories are linear ketoamides and macrocyclic. An example of linear ketoamides are telaprevir and boceprevir which have been recently accepted by food and drug administration (FDA) to be used in the treatment of HCV in combination with the current available treatment PEG IFN- α and ribavirin [13, 14]. They contain an electrophilic moiety forming a covalent complex with the protease and dissociate slowly to inhibit its function [13]. Examples of the other category (macrocyclic) are TMC 435, MK-5172 and MK-7009 which depend on rigid and non-covalent binding to inhibit the enzyme function [15, 16]. Macrocyclic inhibitors share a rigid scaffold that have

a pre-organized conformation and fit tightly within the active site of protein, increasing their binding. They are also not susceptible to hydrolysis by the cell peptidases like the other linear peptide drugs which are easily hydrolyzed and show high conformational degrees of freedom affecting their proper binding [15, 16].

Several mutations have been found to render the designed drugs ineffective and cause what is called drug resistance. Examples of these mutations are R155K/Q, D168A/V/E, and A156V/T/S. The study of the role played by these amino acids and their contributions to the binding interactions with drugs is essential to know the effect of these mutations.

Arginine 155 (R155) is one of several residues forming the S2 pocket, and thus has the potential to form interactions with the large, aromatic P2 substituents of most macrocyclic inhibitors such as TMC435 and ITMN-191. In the wild type enzyme, this basic R155 helps to stabilize the S4–S2 region by salt bridge interactions with aspartic acid 168 (D168) [7–9]. Two substitutions are commonly found in this position, either for glutamine (Q) or lysine (K).

Alanine 156 (A156) lies close to the P2 residue, with possible hydrophobic interactions with P2 and its substituent [7–9]. The most commonly observed substitutions are for threonine (T), serine (S) and valine (V). The underlying mechanism of resistance in all three mutations is increased bulk when the methyl in alanine is replaced by the larger side-chains of threonine, valine or serine. Consequently, there is steric repulsion between the enzyme and the P2 and P4 residues of the inhibitors [7–9].

D168 is not directly involved in any enzyme–inhibitor interactions. It is located between the S2 and the S4 pockets, and forms a stabilizing link between the two pockets by forming salt bridges to both R123 (S4 pocket) and R155 (S2 pocket) [7–9]. The most commonly observed substitutions for D168 are alanine or valine (i.e. D168A or D168V).

The replacement of charged D168 with alanine or valine cannot restore the salt bridges to the two arginines. This makes the two arginines in the D168A/V variants more flexible and affects the interactions of inhibitors [7–9].

Molecular modeling techniques are effective tools and widely used to study the mechanisms of reactions, compare the performance of different drugs and predict the molecular and physicochemical properties of molecules [17, 18]. These techniques have been also used in designing promising ligands inhibiting viral enzymes such as HIV-1 and HCV proteases [18–20].

Molecular docking is an important technique to study protein–ligand interactions and facilitate the design of potent drugs. Docking is a computational tool that places a small molecule (ligand) in the binding site of its macromolecular target (receptor) and estimates its binding

affinity. In molecular docking, based on the protein structures, thousands of possible poses of association between the ligand and its receptor are tried and evaluated using an energy scoring function; the pose with the lowest energy score is predicted as the “best match”, i.e., the binding mode [21].

According to these considerations, docking simulation is applied to study the binding mode of HCV-NS3 protease inhibitors and the effect of three different mutations (R155K, D168A and A156V) on the binding of inhibitors.

2 Computational Methods

2.1 System Preparation

The investigated compounds (Ligands) are built using SCIGRESS 3.1 molecular modeling software [22–24], the geometries of the investigated compounds are defined by performing an optimize geometry calculation using MO-G, implemented in SCIGRESS software, at PM3 level of theory. All the computations are carried out on Dell precision T3500 workstation.

HCV NS3/4A full protein was retrieved from Protein Data Bank (PDB code: 4A92) [25] and opened using SCIGRESS. Chain A has been chosen and other chains have been deleted. Hydrogen atoms have been added to the protein. The geometry of all protein has been classically optimized using Augmented MM3 force field implemented in SCIGRESS with conjugate gradient energy minimizer and convergence energy (0.001 kcal/mole) in order to prepare protein for docking simulations [24].

The three mutants have been prepared by replacing the residues with the other ones using the same software and the three protein mutant (A156V, D168A, R155K) structures have also been optimized classically using Augmented MM3 force field implemented in SCIGRESS with conjugate gradient energy minimizer and convergence energy (0.001 kcal/mole) [24] and become ready for docking simulations.

Active site residues (Q41, F43, H57, G58, D81, R109, K136, G137, S138, S139, G140, G141, F154, R155, A156, A157, D168, M485, V524, Q526 and H528) [24] have been selected for both the wild-type enzyme and mutant structures in order to study their interaction with investigated compounds. The capital letter represents the code of amino acid and the number represents its position in the protein. Docking is performed against the full HCV NS3/4A protein not the protease domain only in order to take advantage of the helicase domain which makes favorable interactions with ligands and this increases their binding [25]. The PM3—optimized geometries of compounds are considered as ligands to be prepared for docking.

2.2 Molecular Docking Simulations

Molecular docking has been used to investigate the interactions between the compounds and protein at the molecular level. Docking calculations have been performed for the wild-type protein and three mutations. Docking simulations are performed using FASTDOCK implemented in SCIGRESS [23, 24]. The ligands and side chains of active site residues are allowed to be flexible to simulate the induced—fit binding of the ligands in the active site [26], no main chain movements are taken into account [24].

FASTDOCK uses genetic algorithm to predict the correct binding mode of ligand in the active site “pose” and knowledge—based scoring function (PMF) [27] to score these poses and rank the different compounds according to their scores. After docking calculations are performed, the best poses are further optimized using Augmented MM3 force field in SCIGRESS with Conjugate gradient energy minimizer and convergence energy (0.001 kcal/mole) [24].

3 Results

Figures 1, 2 and 3 present the studied compounds for the three groups (1st, 2nd and 3rd), respectively.

Tables 1, 2 and 3 present the sequence structure of studied compounds for the three groups (1st, 2nd and 3rd), respectively.

Table 4 present the docking scores of the first group compounds against wild-type protein and three different mutations (R155K, A156V, and D168A). 3/4A hexapeptide sequence (DLEVVT) and its charged counterpart show the best scores against the wild type protein -192.449 and -194.143 kcal/mole, respectively. They also have moderately good scores against different mutations compared to other hexapeptide sequences. The best scores is bold.

Tables 5 present the docking scores of the second group compounds against wild-type protein and three different mutations (R155K, A156V, and D168A). P1–P3 macrocycle of 5A/5B hexapeptide sequence (EDVVCC) has the best scores against the wild type protein -181.584 kcal/mole. The best scores is bold.

Tables 6 present the docking scores of the third group compounds against wild-type protein and three different mutations (R155K, A156V, and D168A). P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences have the best scores with the wild-type protein -204.506 and -206.823 kcal/mole, respectively.

Also, charged P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences have low scores with the wild-type protein -200.467 and -203.186 kcal/mole, respectively. They show good binding scores with the three different mutations (R155K, A156V, and D168A) compared to that

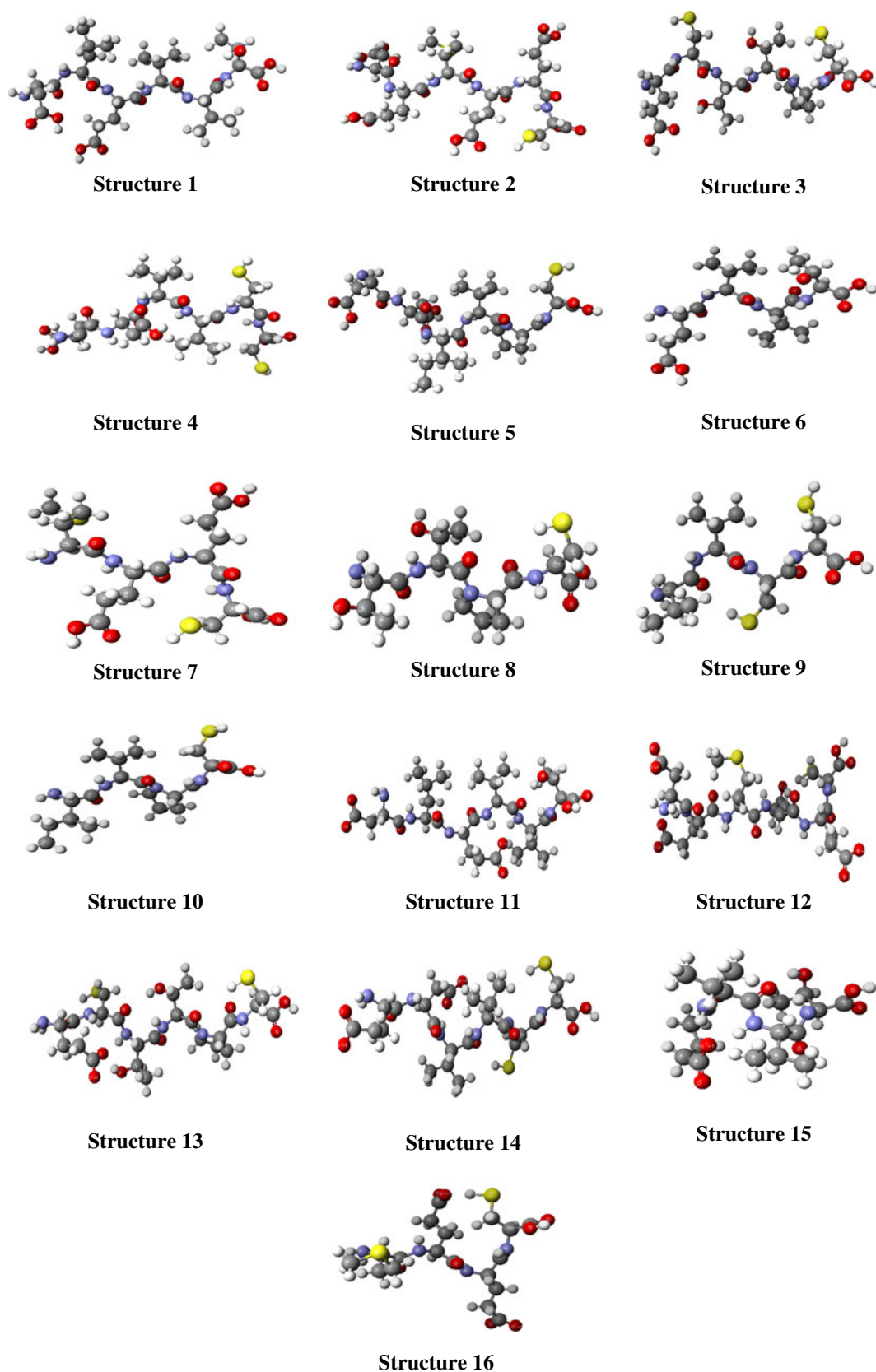


Fig. 1 The PM3—Optimized geometries of first group compounds. Compounds are colored with the atoms as carbon: *gray*, hydrogen: *white*, nitrogen: *blue* and oxygen: *red* (ball and stick model)

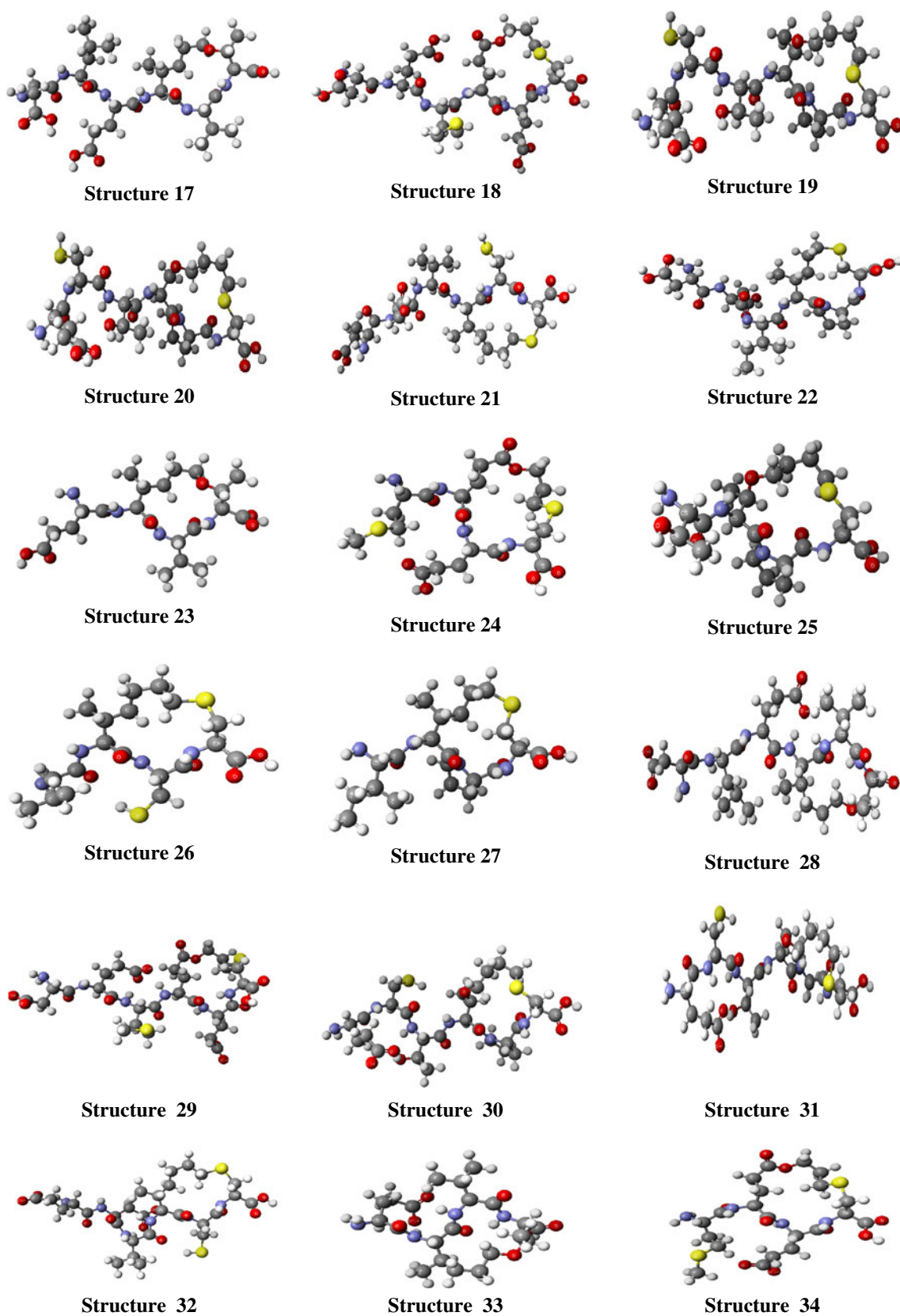


Fig. 2 The PM3—Optimized geometries of second group compounds. Compounds are colored with the atoms as carbon: *gray*, hydrogen: *white*, nitrogen: *blue* and oxygen: *red* (ball and stick model)

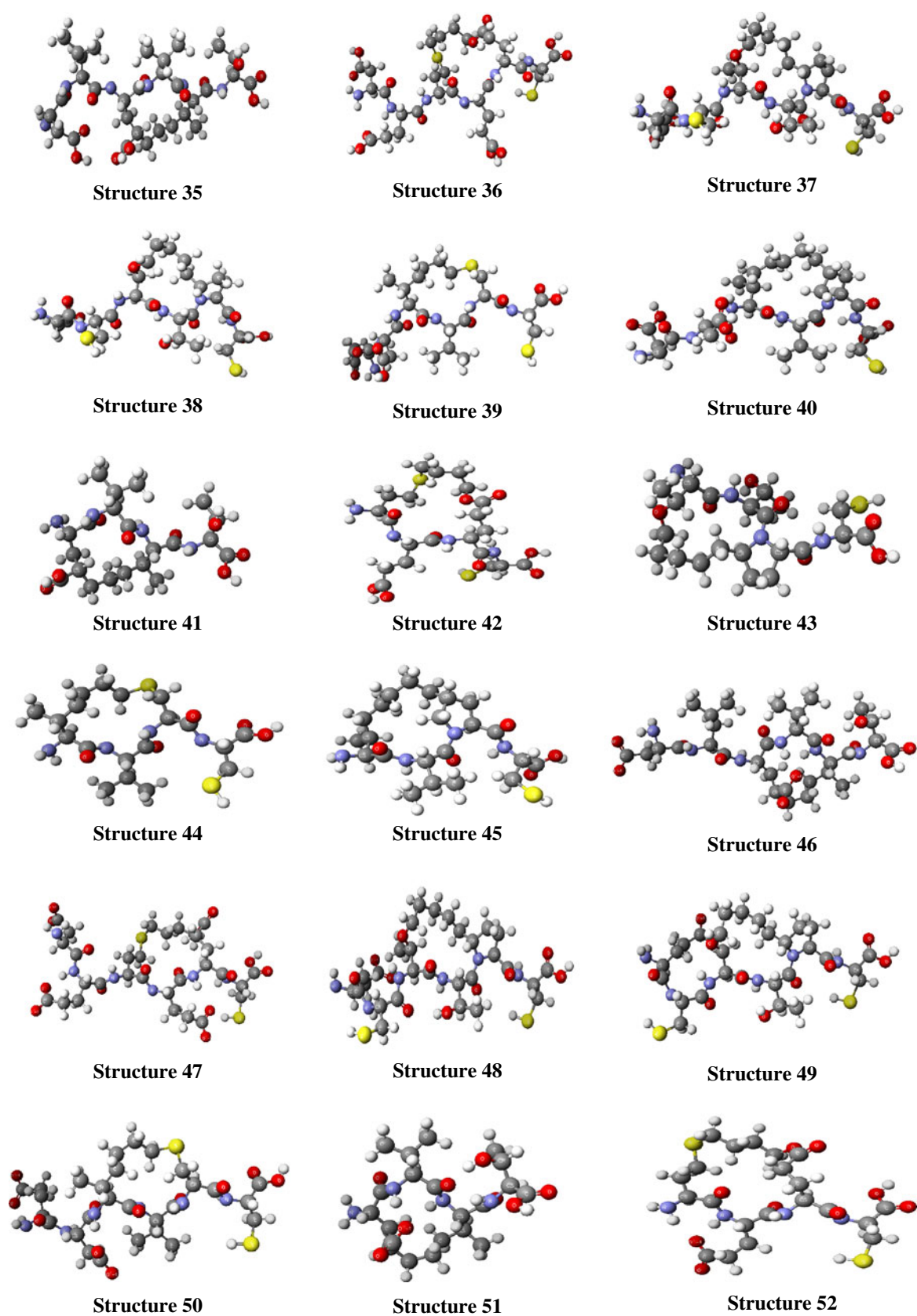


Fig. 3 The PM3—Optimized geometries of third group compounds. Compounds are colored with the atoms as carbon: *gray*, hydrogen: *white*, nitrogen: *blue* and oxygen: *red* (ball and stick model)

Table 1 The sequence structure of studied first group (linear peptides) containing neutral and charged compounds

Compound	Sequence
1	3/4A hexapeptide sequence (DLEVVT)
2	4A/4B hexapeptide sequence (DEMEEC)
3	4B/5A hexapeptide sequence (ECTTPC)
4	5A/5B hexapeptide sequence (EDVVCC)
5	Modified 5A/5B hexapeptide sequence (DDIVPC)
6	3/4A tetrapeptide sequence (EVVT) ^a
7	4A/4B tetrapeptide sequence (MEEC) ^a
8	4B/5A tetrapeptide sequence (TTPC) ^a
9	5A/5B tetrapeptide sequence (VVCC) ^a
10	Modified 5A/5B tetrapeptide sequence (IVPC) ^a
11	Charged 3/4A hexapeptide sequence (DLEVVT) ^b
12	Charged 4A/4B hexapeptide sequence (DEMEEC) ^b
13	Charged 4B/5A hexapeptide sequence (ECTTPC) ^b
14	Charged 5A/5B hexapeptide sequence (EDVVCC) ^b
15	Charged 3/4A tetrapeptide sequence (EVVT) ^{a,b}
16	Charged 4A/4B tetrapeptide sequence (MEEC) ^{a,b}

“E, D, V, C, I, T, L, M and P are one letter codes for glutamic acid, aspartic acid, valine, cysteine, isoleucine, threonine, leucine, methionine and proline amino acids, respectively.”

^a Tetrapeptides are formed by truncating the last two residues

^b The acidic residues are charged

of P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences. The best scores is bold.

Figure 4a presents the docked conformation of P2–P4 macrocycle of 5A/5B hexapeptide sequence with the wild-type protein active site, using Scigress software. Active site residues are defined previously in computational methods section. For all docked conformations of inhibitors, active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: gray, hydrogen: white, nitrogen: blue and oxygen: red (ball and stick model).

The ligand forms one intra hydrogen bond and steric clashes (bumps) with Q41, H57, K136 and R155. These bumps appear between atoms that are within 0.87 times the sum of their Van Der Waals radii and that is not bonded or in the same bond angle.

Figure 4b, d present molecular model of docked conformation of P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences with the wild-type protein active site, created by pose view website (<http://poseview.zbh.uni-hamburg.de/poseview>) [28].

Figure 4c presents the docked conformation of P2–P4 macrocycle of modified 5A/5B hexapeptide sequence with the wild-type protein active site, using Scigress software.

Table 2 The sequence structure of studied second group (P1–P3 macrocyclic structures) containing neutral and charged compounds

Compound	Sequence
17	P1–P3 macrocycle of 3/4A hexapeptide sequence (DLEVVT) ^a
18	P1–P3 macrocycle of 4A/4B hexapeptide sequence (DEMEEC) ^a
19	P1–P3 macrocycle of 4B/5A hexapeptide sequence (ECTTPC) ^a
20	P1–P3 macrocycle of 4B/5A hexapeptide sequence (ECTSPC) ^a
21	P1–P3 macrocycle of 5A/5B hexapeptide sequence (EDVVCC) ^a
22	P1–P3 macrocycle of modified 5A/5B hexapeptide sequence (DDIVPC) ^a
23	P1–P3 macrocycle of 3/4A tetrapeptide sequence (EVVT) ^{a,b}
24	P1–P3 macrocycle of 4A/4B tetrapeptide sequence (MEEC) ^{a,b}
25	P1–P3 macrocycle of 4B/5A tetrapeptide sequence (TTPC) ^{a,b}
26	P1–P3 macrocycle of 5A/5B tetrapeptide sequence (VVCC) ^{a,b}
27	P1–P3 macrocycle of modified 5A/5B tetrapeptide sequence (IVPC) ^{a,b}
28	P1–P3 macrocycle of charged 3/4A hexapeptide sequence (DLEVVT) ^{a,c}
29	P1–P3 macrocycle of charged 4A/4B hexapeptide sequence (DEMEEC) ^{a,c}
30	P1–P3 macrocycle of charged 4B/5A hexapeptide sequence (ECTTPC) ^{a,c}
31	P1–P3 macrocycle of charged 4B/5A hexapeptide sequence (ECTSPC) ^{a,c}
32	P1–P3 macrocycle of charged 5A/5B hexapeptide sequence (EDVVCC) ^{a,c}
33	P1–P3 macrocycle of charged 3/4A tetrapeptide sequence (EVVT) ^{a,b,c}
34	P1–P3 macrocycle of charged 4A/4B tetrapeptide sequence (MEEC) ^{a,b,c}

“E, D, V, C, I, T, L, M and P are one letter codes for glutamic acid, aspartic acid, valine, cysteine, isoleucine, threonine, leucine, methionine and proline amino acids, respectively.”

^a P1–P3 macrocycles are formed through linking the side chains of P1–P3 amino acids

^b Tetrapeptides are formed by truncating the last two residues

^c The acidic residues are charged

The ligand forms four intra hydrogen bond and one hydrogen bond with R123. It also forms steric clashes (bumps) with H57, K136, R155, A156 and Q526.

Figure 5a presents the docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, using Scigress software.

The ligand forms one hydrogen bond with K136 and steric clashes (bumps) with D81, H57 and A156. The

Table 3 The sequence structure of studied third group (P2–P4 macrocyclic structures) containing neutral and charged compounds

Compound	Sequence
35	P2–P4 macrocycle of 3/4A hexapeptide sequence (DLEVVT) ^a
36	P2–P4 macrocycle of 4A/4B hexapeptide sequence (DEMEEC) ^a
37	P2–P4 macrocycle of 4B/5A hexapeptide sequence (ECTTPC) ^a
38	P2–P4 macrocycle of 4B/5A hexapeptide sequence (ECTSPC) ^a
39	P2–P4 macrocycle of 5A/5B hexapeptide sequence (EDVVCC) ^a
40	P2–P4 macrocycle of modified 5A/5B hexapeptide sequence (DDIVPC) ^a
41	P2–P4 macrocycle of 3/4A tetrapeptide sequence (EVVT) ^{a,b}
42	P2–P4 macrocycle of 4A/4B tetrapeptide sequence (MEEC) ^{a,b}
43	P2–P4 macrocycle of 4B/5A tetrapeptide sequence (TTPC) ^{a,b}
44	P2–P4 macrocycle of 5A/5B tetrapeptide sequence (VVCC) ^{a,b}
45	P2–P4 macrocycle of modified 5A/5B tetrapeptide sequence (IVPC) ^{a,b}
46	P2–P4 macrocycle of charged 3/4A hexapeptide sequence (DLEVVT) ^{a,c}
47	P2–P4 macrocycle of charged 4A/4B hexapeptide sequence (DEMEEC) ^{a,c}
48	P2–P4 macrocycle of charged 4B/5A hexapeptide sequence (ECTTPC) ^{a,c}
49	P2–P4 macrocycle of charged 4B/5A hexapeptide sequence (ECTSPC) ^{a,c}
50	P2–P4 macrocycle of charged 5A/5B hexapeptide sequence (EDVVCC) ^{a,c}
51	P2–P4 macrocycle of charged 3/4A tetrapeptide sequence (EVVT) ^{a,b,c}
52	P2–P4 macrocycle of charged 4A/4B tetrapeptide sequence (MEEC) ^{a,b,c}

“E, D, V, C, I, T, L, M and P are one letter codes for glutamic acid, aspartic acid, valine, cysteine, isoleucine, threonine, leucine, methionine and proline amino acids, respectively.”

^a P2–P4 macrocycles are formed through linking the side chains of P2–P4 amino acids

^b Tetrapeptides are formed by truncating the last two residues

^c The acidic residues are charged

conformation of the ligand in the active site facilitates the formation of three intramolecular hydrogen bonds in ligand.

Figure 5b, d present molecular model of docked conformation of P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences with the wild-type protein active site, created by pose view website.

Figure 5c presents the docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, using Scigress software.

Table 4 Docking scores of first group compounds against wild-type NS3/4A protein and three different mutations (R155K, A156V, and D168A)

Compound	Score WT (kcal/mole)	R155K (kcal/mole)	D168A (kcal/mole)	A156V (kcal/mole)
1	−192.449	−183.225	−152.144	−170.165
2	−180.657	−178.085	−150.682	−134.391
3	−143.997	−151.041	−178.943	−165.93
4	−150.217	−147.243	−161.004	−171.46
5	−157.092	−152.455	−204.336	−180.024
6	−135.252	−146.629	−133.587	−139.269
7	−152.934	−151.773	−163.06	−151.168
8	−161.907	−118.168	−137.594	−124.508
9	−122.967	−162.156	−139.64	−103.17
10	−154.467	−146.392	−127.095	−129.124
11	−194.143	−173.282	−190.414	−140.02
12	−184.818	−136.547	−170.615	−148.068
13	−152.667	−152.088	−146.05	−158.009
14	−137.261	−141.155	−136.848	−140.193
15	−140.28	−159.201	−131.424	−156.992
16	−139.605	−137.463	−162.356	−157.952

The best scores are bold

Table 5 Docking scores of second group compounds against wild-type NS3/4A protein and three different mutations (R155K, A156V, and D168A)

Compound	Score WT (kcal/mole)	R155K (kcal/mole)	D168A (kcal/mole)	A156V (Kcal/mole)
17	−155.468	−169.177	−172.455	−191.25
18	−146.257	−146.576	−140.317	−168.706
19	−157.238	−147.055	−155.162	−152.984
20	−131.629	−171.138	−137.631	−174.517
21	−181.584	−146.32	−130.043	−166.219
22	−176.922	−133.174	−134.687	−131.111
23	−165.423	−182.209	−139.034	−168.516
24	−141.675	−139.494	−143.181	−164.86
25	−164.385	−130.157	−165.102	−138.717
26	−123.332	−148.047	−122.627	−114.142
27	−129.389	−156.141	−123.993	−154.491
28	−156.229	−143.702	−137.662	−139.818
29	−157.988	−185.658	−172.699	−160.157
30	−178.519	−149.533	−194.437	−138.954
31	−130.568	−116.995	−204.717	−127.448
32	−138.321	−184.398	−161.012	−133.231
33	−132.157	−177.429	−116.59	−136.186
34	−156.665	−185.922	−121.406	−134.165

The best scores is bold

The ligand adopts two intramolecular hydrogen bonds, two hydrogen bonds with K136 and one hydrogen bond

Table 6 Docking scores of third group compounds against wild-type NS3/4A protein and three different mutations (R155K, A156V, and D168A)

Compound	Score WT (kcal/mole)	R155K (kcal/mole)	D168A (kcal/mole)	A156V (kcal/mole)
35	-174.994	-176.643	-197.968	-173.664
36	-168.67	-160.43	-135.593	-115.132
37	-173.619	-179.836	-191.275	-124.253
38	-159.982	-144.464	-162.352	-168.516
39	-204.506	-167.426	-194.678	-170.707
40	-206.823	-198.291	-165.58	-109.806
41	-138.028	-115.734	-135.469	-171.559
42	-117.007	-126.764	-125.203	-124.332
43	-126.501	-172.205	-128.271	-133.686
44	-128.536	-118.722	-142.063	-121.881
45	-112.36	-112.89	-126.839	-126.754
46	-200.467	-198.024	-186.843	-173.654
47	-203.186	-197.22	-191.477	-180.623
48	-144.133	-162.485	-198.409	-128.564
49	-172.564	-179.165	-167.678	-133.338
50	-167.383	-165.408	-164.053	-169.41
51	-128.157	-150.427	-137.821	-144.302
52	-159.632	-154.094	-144.82	-115.368

The best scores is bold

with H57. It makes several steric clashes with several amino acid residues like K136, H57, D81, R155 and Q526.

Figure 6a presents the docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the active site of A156V mutant, using Scigress software.

The ligand adopts four intramolecular hydrogen bonds and one hydrogen bond with R155 and Q41. The bulk amino acid valine 156 makes two steric clashes with the ligand.

Figure 6b, d present molecular model of docked conformation of P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences with the active site of A156V mutant, created by pose view website.

Figure 6c presents the docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the active site of A156V mutant, using Scigress software.

The ligand adopts three intramolecular hydrogen bonds and two hydrogen bonds with K136. It makes several steric clashes with several amino acids like K136, V156, Q526 and H528.

Figure 7a presents the docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the active site of D168A mutant, using Scigress software.

The ligand forms two intramolecular hydrogen bonds and two hydrogen bonds with H57 and Q41. It also makes a steric clash with K136.

Figure 7b, d present molecular model of docked conformation of P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences with the active site of D168A mutant, created by pose view website.

Figure 7c presents the docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the active site of D168A mutant, using Scigress software.

The ligand forms three intramolecular hydrogen bonds and one hydrogen with K136, H57 and A157. It also makes several bumps with K136 and H57.

Figure 8a presents the docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the active site of R155K mutant, using Scigress software.

The ligand forms four intramolecular hydrogen bonds and two hydrogen bonds with K155 and one hydrogen bond with H57. It also makes steric clashes with K136 and D81.

Figure 8b, d present molecular model of docked conformation of P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences with the active site of R155K mutant, created by pose view website.

Figure 8c presents the docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the active site of R155K mutant, using Scigress software.

The ligand forms one intramolecular hydrogen bond. It makes one hydrogen bond and steric clash with K136.

4 Discussion

Drug resistance is a major obstacle in the treatment of quickly evolving diseases. HCV NS3/4A protein is an essential protein in the viral life cycle. It is a multifunctional enzyme possessing serine protease activity in the N-terminal third of the protein and RNA helicase/NTPase activity in the C-terminal portion. The X-ray crystal structure of the full-length NS3/4A revealed that the protease active site was situated in the interface between the helicase and the protease domains, creating a well-defined binding cleft. Protease and helicase residues of the NS3/4A participate in the interactions with the substrates and inhibitors. One important strategy to design effective drugs of HCV NS3/4A mutants is the “Substrate Envelope Hypothesis” of the target protein. According to this hypothesis, inhibitors that fit within the volume occupied by the enzyme’s natural substrate should tend to resist mutations. So, any mutations affect inhibitor binding would in turn affect substrate binding, rendering the enzyme ineffective [4, 5]. Recent studies of the analysis of the NS3/4A substrates and inhibitor envelopes reveal the areas where the consensus inhibitor volume extended beyond the substrate envelope [7–9]. These areas correspond to drug resistance mutations including Arg155,

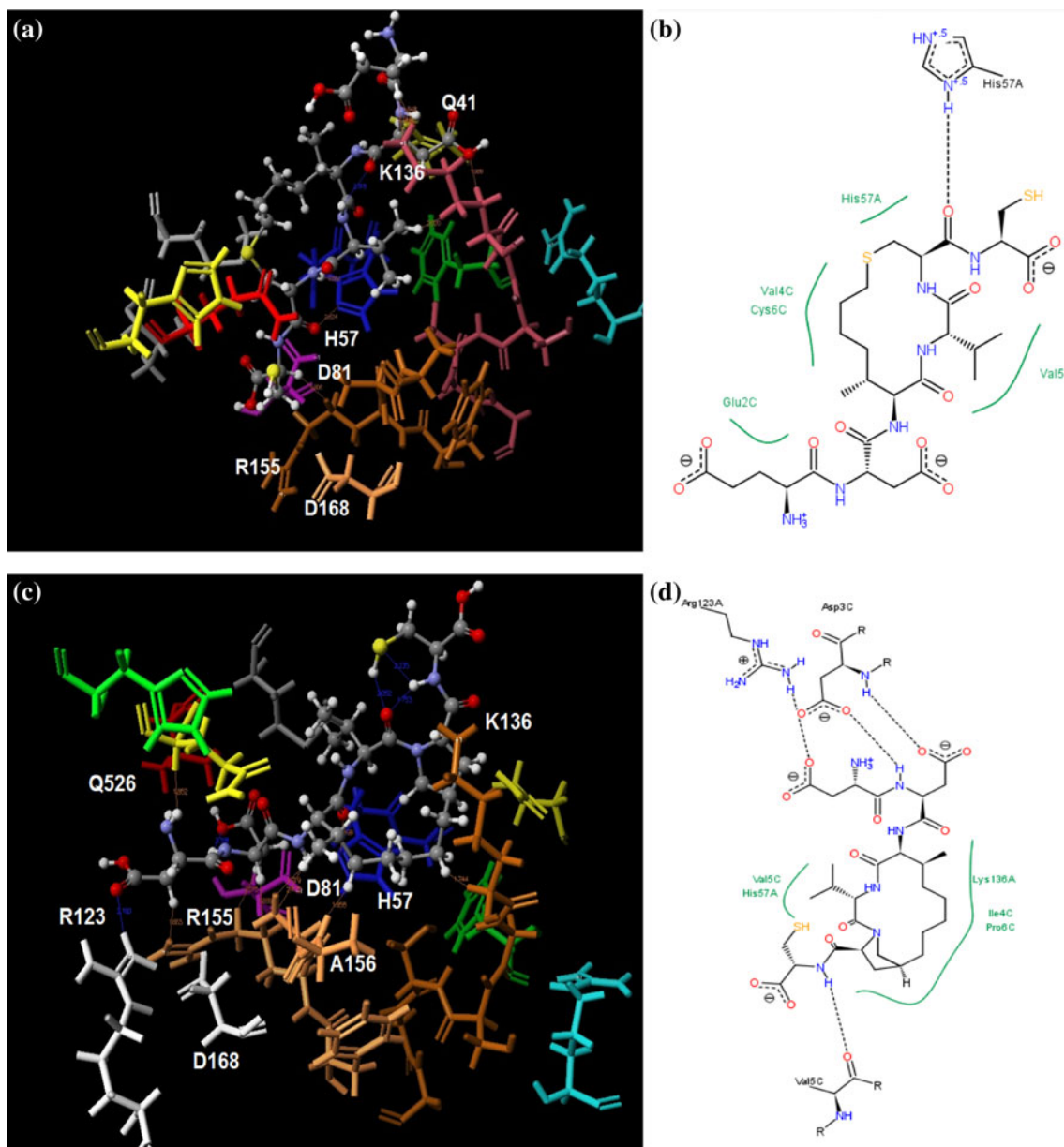


Fig. 4 **a** Docked conformation of P2–P4 macrocycle of 5A/5B hexapeptide sequence with the wild-type protein active site, using Scigress software. **b** Molecular model of docked conformation of P2–P4 macrocycle of 5A/5B hexapeptide sequence with the wild-type protein active site, created by pose view website. (<http://poseview.zbh.uni-hamburg.de/poseview>). **c** Docked conformation of modified P2–P4 macrocycle of 5A/5B hexapeptide sequence with the wild-type protein active site, using Scigress software. **d** Molecular model of

docked conformation of modified P2–P4 macrocycle of 5A/5B hexapeptide sequence with the wild-type protein active site, created by pose view website. Active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: *gray*, hydrogen: *white*, nitrogen: *blue* and oxygen: *red* (ball and stick model)

Ala156 and Asp168 at the protease active site as well as the two conserved helicase residues Gln526 and His528 that strongly interact with the inhibitors [9, 23].

In this study, docking simulation has been used to rank the suggested compounds according to their binding scores with the wild-type and mutant proteins. It is also used to analyze the conformation of the best compounds inside the active site of proteins and show how the mutations affect the binding of

compounds. The design of the suggested compounds is based on the Substrate Envelope Hypothesis for HCV NS3 protease. The hypothesis states that “the substrates of HCV NS3/4A protease corresponding to NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B junctions have consensus Van der Waals volume inside the active site of protease and this determines the specificity of recognition of substrates by protease”. The designed compounds have comparable steric

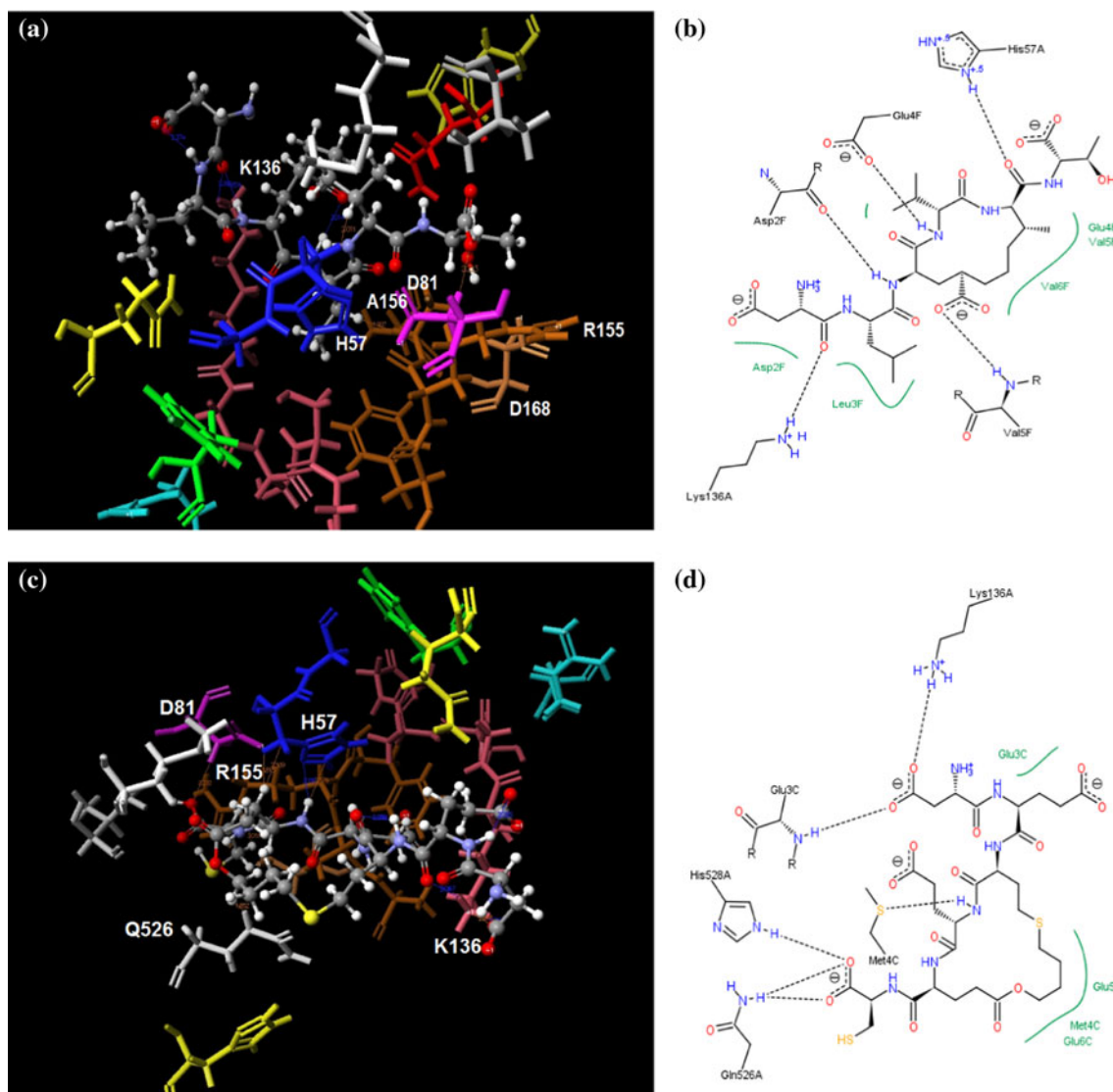


Fig. 5 **a** Docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, using Scigress software. **b** Molecular model of docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, created by pose view website. (<http://poseview.zbh.uni-hamburg.de/poseview>). **c** Docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, using Scigress software. **d** Molecular model of docked

conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, created by pose view website. Active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: gray, hydrogen: white, nitrogen: blue and oxygen: red (ball and stick model)

parameters to the four different substrates as a trial to solve the drug resistance issue.

We approximate the substrate envelope by calculating the Van der Waals volume and surface area at semi-empirical PM3 level of theory for all HCV substrates. Then, we compare the Van der Waals volume and surface area for the designed ligands and find that the parameters of the designed ligands have comparable VDW volume and surface area to these substrates. Although some of the best compounds have little increase in volume than that of substrates (unpublished data).

Tables 6 present the docking scores of the P2–P4 macrocyclic compounds against wild-type protein and three different mutations (R155K, A156V, and D168A). P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences have the best scores with the wild-type protein -204.506 and -206.823 kcal/mole, respectively. Also, charged P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences have low scores with the wild-type protein -200.467 and -203.186 kcal/mole, respectively. They show good binding scores with the three different mutations (R155K, A156V, and D168A) compared to that of

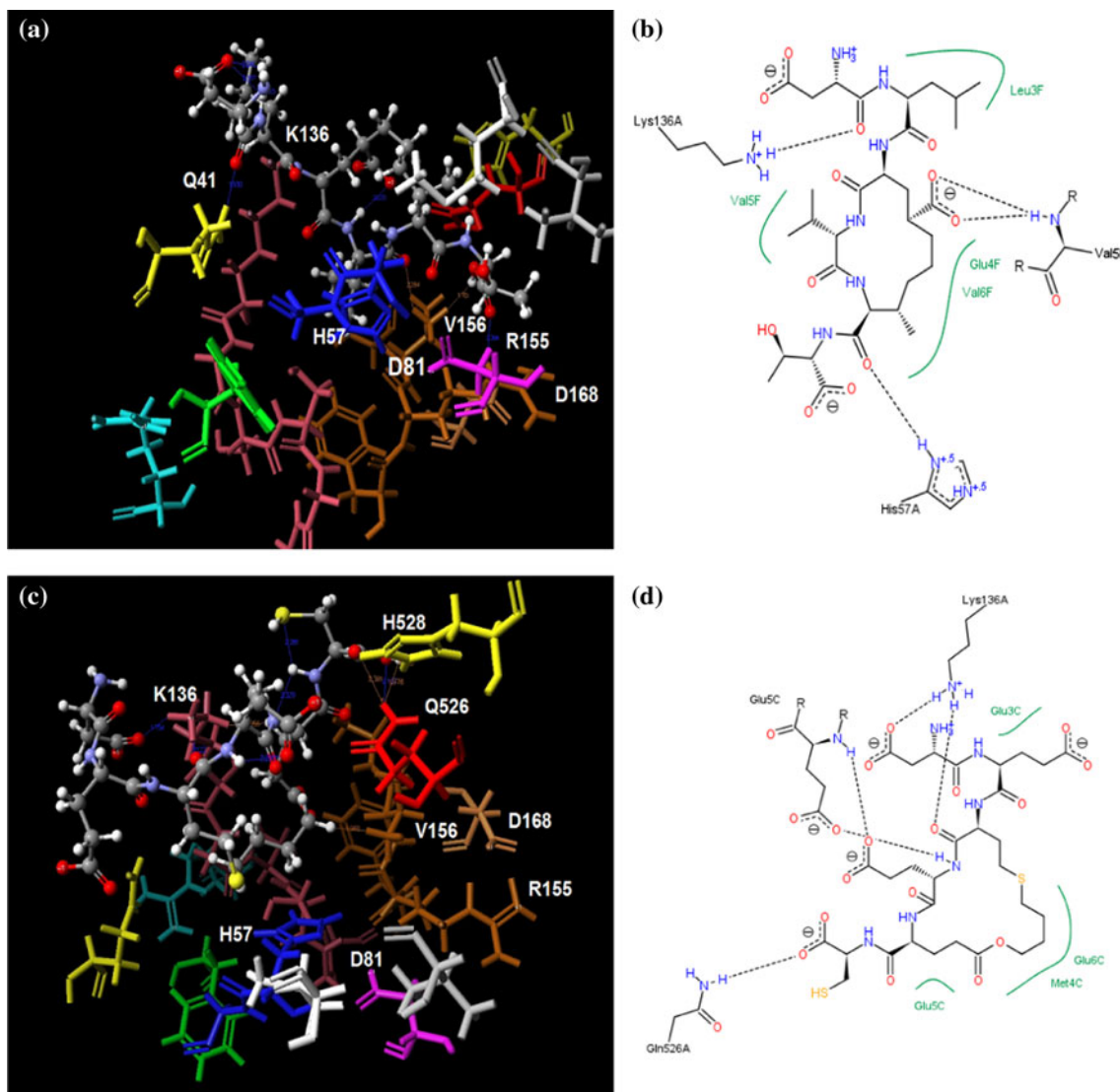


Fig. 6 **a** Docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the NS3/4A protease A156V mutant active site, using Scigress software. **b** Molecular model of docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, created by pose view website. (<http://poseview.zbh.uni-hamburg.de/poseview>). **c** Docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the NS3/4A protease A156V mutant active site, using Scigress software.

d Molecular model of docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, created by pose view website. Active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: gray, hydrogen: white, nitrogen: blue and oxygen: red (ball and stick model)

P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences.

4.1 Binding Mode of P2–P4 Macrocycle of 5A/5B Hexapeptide Sequence with Wild-type Protein

The conformation of the ligand in the active site is presented in Fig. 4a; the P1 cysteine residue lies beyond R155 which contributes to forming the S2 binding pocket of P2 residue [7–9]. It also lies near two of the catalytic triad H57 and D81. P2–P4 macrocycle lies near helicase domain

residues H528, Q528 and M485. The P3 valine, P5 and P6 acidic residues form steric clashes with Q41 and K136.

4.2 Binding Mode of P2–P4 Macrocycle Modified 5A/5B Hexapeptide Sequence with Wild-type Protein

The conformation of the ligand in the active site is presented in Fig. 4c. The whole ligand is impressed in the binding cleft. The large side chains of residues especially P2, P3 and P4 make several steric clashes with the surrounding residues. P2–P4 macrocycle lies near catalytic triad and K136. The P5 and

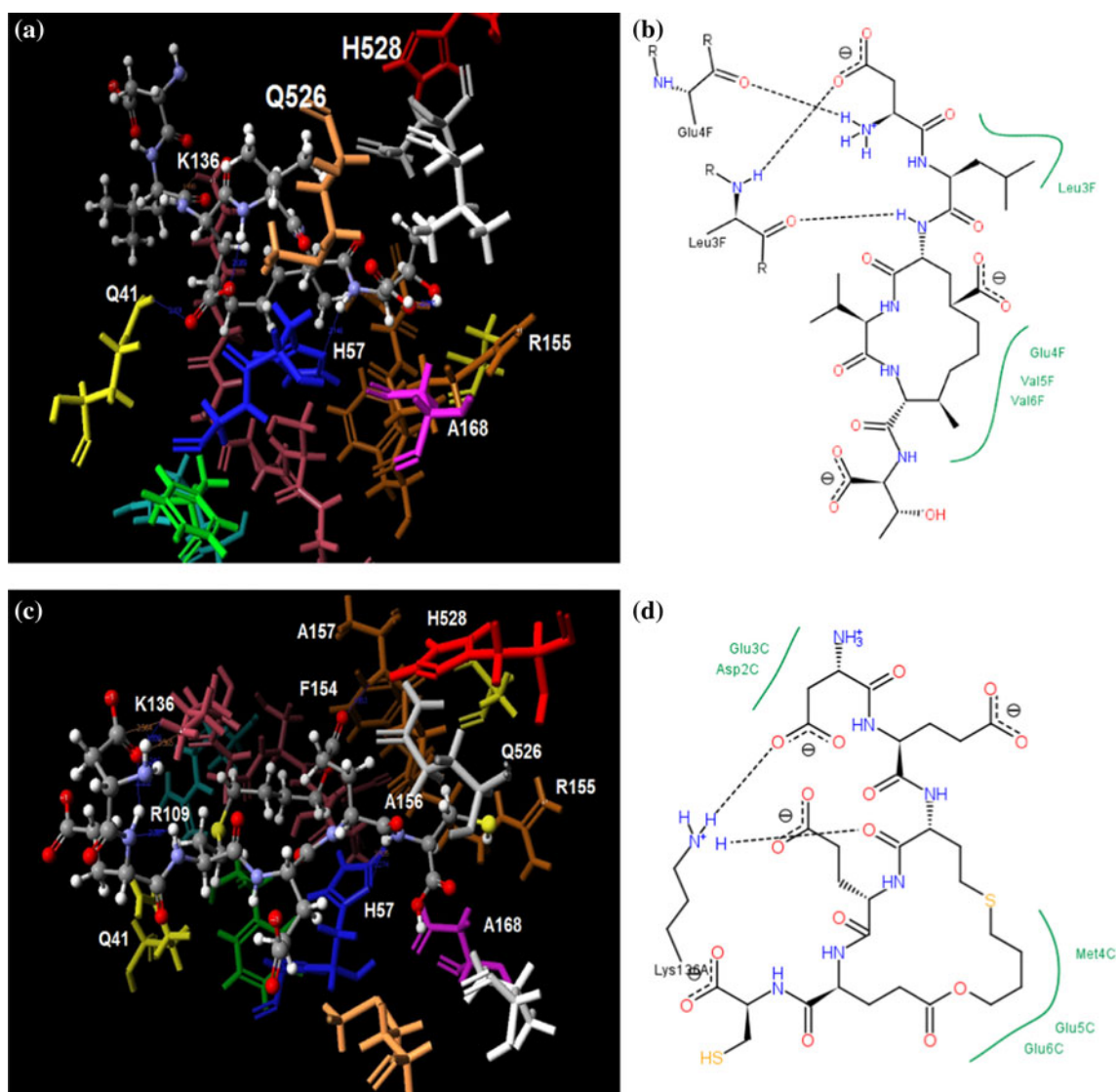


Fig. 7 **a** Docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the NS3/4A protease D168A mutant active site, using Scigress software. **b** Molecular model of docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, created by pose view website. (<http://poseview.zbh.uni-hamburg.de/poseview>). **c** Docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the NS3/4A protease D168A mutant active site, using Scigress software.

d Molecular model of docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, created by pose view website. Active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: gray, hydrogen: white, nitrogen: blue and oxygen: red (ball and stick model)

P6 acidic residues interact with R123 and R155 that form S4 and S2 binding pockets. P6 also make steric clash with Q526.

4.3 Binding Mode of Charged P2–P4 Macrocycle of 3/4A Hexapeptide Sequence with Wild-type and Mutant Proteins

4.3.1 Wild-type Protein

The conformation of the ligand in the active site is presented in Fig. 5a. The P1 residue makes steric clash with

D81, and the large side chain of P4 residue makes steric clash with A156. Carbonyl group of P6 negatively—charged aspartate forms hydrogen bond with positively—charged K136.

4.3.2 A156V Mutant

The conformation of the ligand in the active site is presented in Fig. 6a. A156 lies close to the P2 residue, with possible hydrophobic interactions with P2 and its substituent [7–9]. The underlying mechanism of resistance is

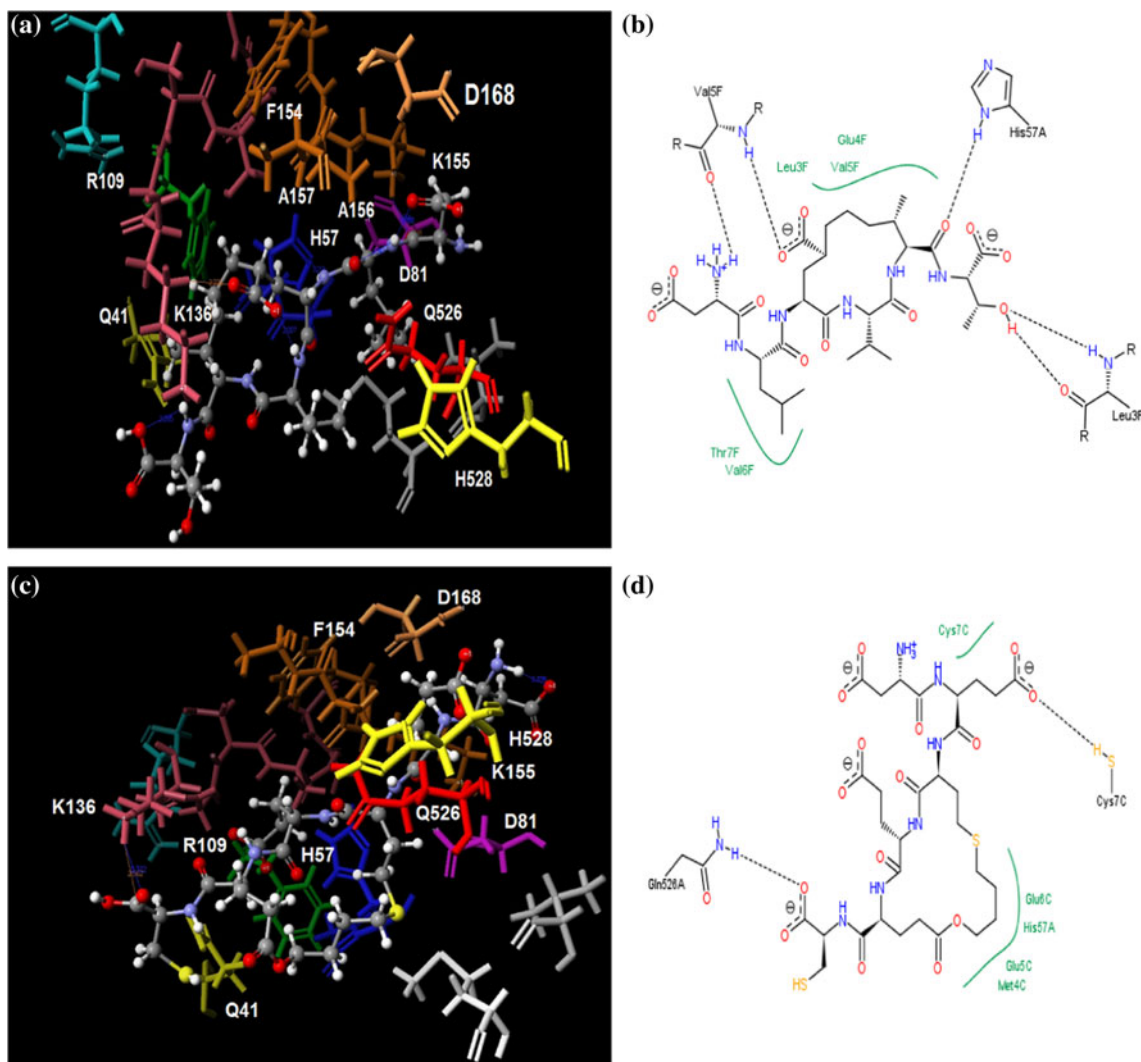


Fig. 8 **a** Docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the NS3/4A protease R155K mutant active site, using Scigress software. **b** Molecular model of docked conformation of P2–P4 macrocycle of 3/4A hexapeptide sequence with the wild-type protein active site, created by pose view website. (<http://poseview.zbh.uni-hamburg.de/poseview>) **c** Docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the NS3/4A protease R155K mutant active site, using Scigress software.

d Molecular model of docked conformation of P2–P4 macrocycle of 4A/4B hexapeptide sequence with the wild-type protein active site, created by pose view website. Active site amino acid residues are represented as tubes, colored according to residue type (Sequence protocol- Karplus and Schultz Flexibility). Inhibitor is colored with the atoms as carbon: *gray*, hydrogen: *white*, nitrogen: *blue* and oxygen: *red* (ball and stick model)

increased bulk when the methyl in alanine is replaced by the larger side-chain of valine. Consequently, there is steric repulsion between the enzyme and the P2 and P4 residues of the inhibitors [7–9]. The mutation effect is manifested in the lowering of docking score compared to the wild-type protein.

The conformation of the ligand in the active site in A156V mutant greatly resembles that of the ligand in the active site in wild-type protein. The large side chain of valine compared to alanine makes steric repulsion with P1 and P3 residues and forms steric clashes with it. The P1

residue lies near R155 and forms one hydrogen bond with it. Carbonyl group of P6 residue forms hydrogen bond with Q41.

4.3.3 D168A Mutant

The conformation of the ligand in the active site is presented in Fig. 7a. D168 is not directly involved in any enzyme–inhibitor interactions. It is located between the S2 and the S4 pockets, and forms a stabilizing link between the two pockets by forming salt bridges to both R123 (S4

pocket) and R155 (S2 pocket) [7–9]. The replacement of charged D168 with alanine cannot restore the salt bridges to the two arginines. This makes the two arginines in the D168A mutant more flexible and affects the interactions of inhibitors [7–9]. The conformation of the ligand in the active site in D168A mutant resembles that of the ligand in the active site in wild-type protein. This effect is clear through the flexibility of arginine155 and P1 residue lies far from it and makes no interaction with it compared to the wild-type and A156V mutant. P4 side chain forms hydrogen bond with Q41.

4.3.4 R155K Mutant

The conformation of the ligand in the active site is presented in Fig. 8a. R155 is one of several residues forming the S2 pocket, and thus has the potential to form interactions with the large, aromatic P2 substituents of most macrocyclic inhibitors. In the wild type enzyme, this basic R155 helps to stabilize the S4–S2 region by salt bridge interactions with aspartic acid 168 (D168) [7–9]. The conformation of the ligand in the active site in R155K mutant is totally different from that in the active site in wild-type protein. The molecule is inverted with P5 and P6 residues lies near K155 compared to wild-type protein. P5 makes steric clash with D81. P4 residue forms hydrogen bond with H57.

4.4 Binding Mode of Charged P2–P4 Macrocycle of 4A/4B Hexapeptide Sequence with Wild-type and Mutant Proteins

4.4.1 Wild-type Protein

The conformation of the ligand in the active site is presented in Fig. 5c. The P1 residue makes steric clash with D81. P1 and P2 residues lie near R155 residue. P3 forms hydrogen bond and steric clash with H57. P2–P4 macrocycle orients toward helicase residue Q526. P4 residue forms hydrogen bond with K136. P6 makes steric clash with K136.

4.4.2 A156V Mutant

The conformation of the ligand in the active site is presented in Fig. 6c. The bulky side chain of valine makes steric repulsion with the large P2–P4 macrocycle. P2–P4 macrocycle lies far from H57 and D81. This causes the conformation of the ligand in the active site in A156V mutant is different from that in the active site in wild-type protein. P1 cysteine lies near helicase residues Q526 and H528. It makes steric clashes with A157. P4 and P6 form hydrogen bonds with K136.

4.4.3 D168A Mutant

The conformation of the ligand in the active site is presented in Fig. 7c. P1 residue lies near A168 and far from the more flexible R155. It forms one hydrogen bond and makes a steric clash with H57. P2 residue forms hydrogen bond with A157. P6 conformation and interaction with K136 is conserved.

4.4.4 R155K Mutant

The conformation of the ligand in the active site is presented in Fig. 8c. The conformation of the ligand in the active site in R155K mutant is totally different from that in the active site in wild-type protein. The molecule is inverted with P5 and P6 residues lies near K155 and P1 lies near K136 compared to wild-type protein. This also happens with P2–P4 macrocycle of 3/4A hexapeptide sequence with this mutation. This may be because the large and free side chain of K155 which restricts the conformation of the large P2–P4 macrocycle.

According to the binding interactions of the compounds, we observed that K136 interacts mainly with the compounds in wild-type and mutants. This observation is consistent with mutagenesis experiments that confirmed the importance of this amino acid as a determinant of ligand binding [29].

5 Conclusion

Using molecular modeling techniques especially molecular docking, specific inhibitors to HCV NS3 protease and its mutants are designed. Also, the interactions of these inhibitors with the active site of proteins are studied. The effect of mutations on the binding of these inhibitors is elucidated. P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences have the best scores against the wild-type protein -204.506 and -206.823 kcal/mole, respectively. Also, charged P2–P4 macrocycles of 3/4A and 4A/4B hexapeptide sequences have low scores with the wild-type protein -200.467 and -203.186 kcal/mole, respectively. They show good binding scores with the three different mutations (R155K, A156V, and D168A) compared to that of P2–P4 macrocycles of 5A/5B and modified 5A/5B hexapeptide sequences. R155K mutation greatly affects the conformation of the compounds inside the active site. It inverts its orientations, and this is because the large and free side chain of K155 which restricts the conformation of the large P2–P4 macrocycle. The conformation of charged P2–P4 macrocycle of 3/4A hexapeptide sequence in wild-type, A156V and D168A proteins is nearly equal; while that of charged P2–P4 macrocycle of

4A/4B hexapeptide sequence is different. Nevertheless, these compounds have a slight increase of Van der Waals volume compared to that of substrates, they are potent against mutations and have good scores. So, they are expected to be good drugs for HCV NS3 protease and its mutants and solve the drug resistance issue as well.

References

1. Pawlotsky JM (2012) Is hepatitis virus resistance to antiviral drugs a threat? *Gastroenterology* 142:1369–1372
2. Vermehren J, Sarrazin C (2012) The role of resistance in HCV treatment. *Best Pract Res Clin Gastroenterol* 26(4):487–503
3. Welsch C, Zeuzem S (2012) Clinical relevance of HCV antiviral drug resistance. *Curr Opin Virol* 2(5):651–655
4. Kairys V, Gilson MK, Lather V, Schiffer CA, Fernandes MX (2009) Toward the design of mutation-resistant enzyme inhibitors: further evaluation of the substrate envelope hypothesis. *Chem Biol Drug Des* 74:234–245
5. Romano KP, Ali A, Royer WE, Schiffer CA (2010) Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding. *Proc Natl Acad Sci* 107(49):20986–20991
6. Romano KP, Laine JM, Deveau LM, Cao H, Massi F, Schiffer CA (2011) Molecular mechanisms of viral and host cell substrate recognition by hepatitis C virus NS3/4A protease. *J Virol* 85(13):6106–6116
7. Pan D, Xue W, Zhang W, Liu H (1820) Yao X (2012) Understanding the drug resistance mechanism of hepatitis C virus NS3/4A to ITMN-191 due to R155 K, A156 V, D168A/E mutations: a computational study. *Biochim Biophys Acta* 10:1526–1534
8. Romano KP, Ali A, Aydin C, Soumana D, Ozen A, Deveau LM, Silver C, Cao H, Newton A, Petropoulos CJ, Huang W, Schiffer CA (2012) The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors. *PLoS Pathog* 8(7):e1002832. doi:10.1371/journal.ppat.1002832
9. Xue W, Wang M, Jin X, Liu H, Yao X (2012) Understanding the structural and energetic basis of inhibitor and substrate bound to the full-length NS3/4A: insights from molecular dynamics simulation, binding free energy calculation and network analysis. *Mol Bio Syst* 8:2753–2765
10. Manns MP, Foster GR, Rockstroh JK, Zeuzem S, Zoulim F, Houghton M (2007) The way forward in HCV treatment—finding the right path. *Nat Rev Drug Discov* 6:991–1000
11. Raney KD, Sharma SD, Moustafa IM, Cameron CE (2010) Hepatitis C virus non-structural protein 3 (HCV NS3): a multi-functional antiviral target. *J Biol Chem* 285(30):22725–22731
12. Lin C (2006) HCV NS3-4A Serine Protease a book chapter in *Hepatitis C viruses: genomes and molecular biology*. In: Tan S (ed), Horizon Bioscience, p 163–206
13. Chen KX, Njoroge FG (2010) The journey to the discovery of Boceprevir: an NS3-NS4 HCV protease inhibitor for the treatment of chronic hepatitis C. *Prog Med Chem* 49:1–36
14. Chevaliez S, Pawlotsky JM (2007) Interferon-based therapy of hepatitis C. *Adv Drug Deliv Rev* 59:1222–1241
15. Avolio S, Summa V (2010) Advances in the development of macrocyclic inhibitors of hepatitis C virus NS3-4A protease. *Curr Top Med Chem* 10:1403–1422
16. Venkatraman S, Njoroge FG (2009) Macrocyclic inhibitors of HCV NS3 protease. *Expert Opin Ther Patents* 19(9):1277–1303
17. Elfiky AA, Elshemey WM, Gawad WA, Desoky OS (2013) Molecular modeling comparison of the performance of NS5b polymerase inhibitor (PSI-7977) on prevalent HCV genotypes. *Protein J* 32(1):75–80
18. Saleh NA, Elfiky AA, Ezat AA, Elshemey WM, Ibrahim M (2014) The electronic and QSAR properties of modified Telaprevir compounds as HCV NS3 protease inhibitors. *J Comp Theo NanoSci* 11:1–5
19. Ibrahim M, Saleh NA, Elshemey WM, Elsayed AA (2012) Hexapeptide functionality of cellulose as NS3 protease inhibitors. *Med Chem* 8:6–10
20. Ibrahim M, Saleh NA, Elshemey WM, Elsayed AA (2012) Fullerene derivative as anti-HIV protease inhibitor: molecular modeling and QSAR approaches. *Mini Rev Med Chem* 12(6):447–451
21. Huang SY, Zou X (2010) Advances and challenges in protein-ligand docking. *Int J Mol Sci* 11:3016–3034
22. Stewart JJP (2009) SCIGRESS, Version 2.9.0, Fujitsu Limited, United States
23. Ganguly S, Bahare RS (2013) Molecular docking studies of novel thiazolidinedione analogs as HIV-1-RT inhibitors. *Med Chem Res* 22:3350–3363
24. Mostafa HIA, El-bialy NS, Ezat AA, Saleh NA, Ibrahim MA, QSAR analysis and molecular docking simulation of suggested Peptidomimetic NS3 Protease Inhibitors. *Curr Comp aided Drug Des* (in press)
25. Schiering N, Arcy AD, Villard F, Simić O, Kamke M, Monnet G, Hassiepen U, Svergun DI, Pulfer R, Eder J, Raman P, Bodendorf U (2011) A macrocyclic HCV NS3/4A protease inhibitor interacts with protease and helicase residues in the complex with its full-length target. *Proc Natl Acad Sci* 108(52):21052–21056
26. Kroemer RT (2007) Structure-based drug design: docking and scoring. *Curr Prot Pep Sci* 8:312–328
27. Muegge I (2006) PMF scoring revisited. *J Med Chem* 49:5895–5902
28. <http://poseview.zbh.uni-hamburg.de/poseview>. Accessed 22 Nov 2013
29. Steinkühler C, Biasiol G, Brunetti M, Urbani A, Koch U, Cortese R, Pessi A, Francesco RD (1998) Product inhibition of the hepatitis C virus NS3 protease. *Biochemistry* 37:8899–8905