JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

A Refined Model of the HCV NS5A Protein Bound to Daclatasvir Explains Drug-Resistant Mutations and Activity against Divergent Genotypes

Khaled H. Barakat,[*,†,⊥] Anwar Anwar-Mohamed,[†] Jack A. Tuszynski,[‡,§] Morris J. Robins,[∥] D. Lorne Tyrrell,[†,#] and Michael Houghton[†,#]

[†]Li Ka Shing Institute of Virology, University of Alberta, Edmonton, Alberta Canada

[‡]Department of Oncology, University of Alberta, Edmonton, Alberta Canada

[§]Department of Physics, University of Alberta, Edmonton, Alberta Canada

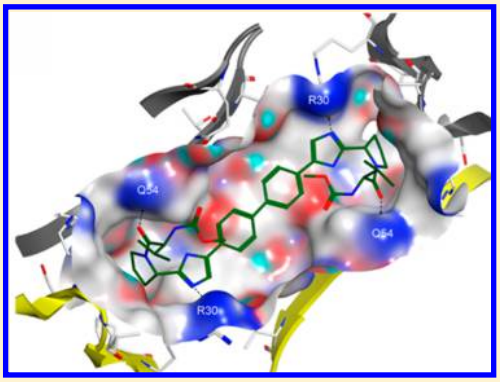
[∥]Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, United States

[⊥]Department of Engineering Mathematics and Physics, Fayoum University, Fayoum, Egypt

[#]Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, Alberta Canada

Supporting Information

ABSTRACT: Many direct-acting antiviral agents (DAAs) that selectively block hepatitis C virus (HCV) replication are currently under development. Among these agents is Daclatasvir, a first-in-class inhibitor targeting the NS5A viral protein. Although Daclatasvir is the most potent HCV antiviral molecule yet developed, its binding location and mode of binding remain unknown. The drug exhibits a low barrier to resistance mutations, particularly in genotype 1 viruses, but its efficacy against other genotypes is unclear. Using state-of-the-art modeling techniques combined with the massive computational power of Blue Gene/Q, we identified the atomic interactions of Daclatasvir within NS5A for different HCV genotypes and for several reported resistant mutations. The proposed model is the first to reveal the detailed binding mode of Daclatasvir. It also provides a tool to facilitate design of second generation drugs, which may confer less resistance and/or broader activity against HCV.



1. INTRODUCTION

Hepatitis C virus (HCV) infection is a major health problem as it can lead to the development of chronic liver diseases such as fibrosis, cirrhosis, and hepatocellular carcinoma in a significant number of infected individuals.[1] The current standard-of-care therapy regimen for chronic HCV infection is a combination of PEGylated interferon (IFN) and ribavirin along with new protease inhibitors in the case of the genotype 1 clade. However, substantial adverse effects together with partial efficacy make it important to develop more potent and safer alternatives.[2] After decades of HCV drug development, an interferon-free oral drug regimen is finally on the horizon. In this regard, most of the current DAAs undergoing clinical trials act on three key viral targets in the HCV replication cycle: NS3/4A protease, NS5A protein, and NS5B RNA polymerase.[3]

Unlike NS3 and NS5B proteins, no enzymatic function has been identified thus far for the NS5A protein,[4] although it is crucial in virus production and has been shown to be involved in modulating host immune response, HCV pathogenicity, and replication.[4] These findings made NS5A a highly attractive target for therapeutic intervention. In this regard, the development of highly potent NS5A inhibitors, which are currently pioneered by Bristol-Myers Squibb (BMS), is serving as a new therapeutic paradigm that also offers broad HCV genotype coverage.

The HCV NS5A protein has 447 amino acids and is active as a homodimer that is organized into three different domains, of which domain I is the most conserved and noticeably the most structured domain.[5] NS5A can exist in either phosphorylated or hyperphosphorylated states.[6] Adding to these complexities is the apparent flexibility of the protein. For example, two distinct crystal structures are currently available for NS5A domain I.[7,8] The structure from the group of Charles Rice has an open conformation,[8] in which two NS5A monomers associate with a large groove in between, while Love's model had a closed and tightly bound conformation.[7] Both structures lack an important highly flexible α-helix formed by the N-terminal residues. Although the two structures represent two distinct dimeric forms of the same protein, there is no evidence that these two structures are readily exchangeable from one form to the other. However, for convenience, we will use the terms closed and open throughout the text below.

Received: October 29, 2013
Published: April 14, 2014

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

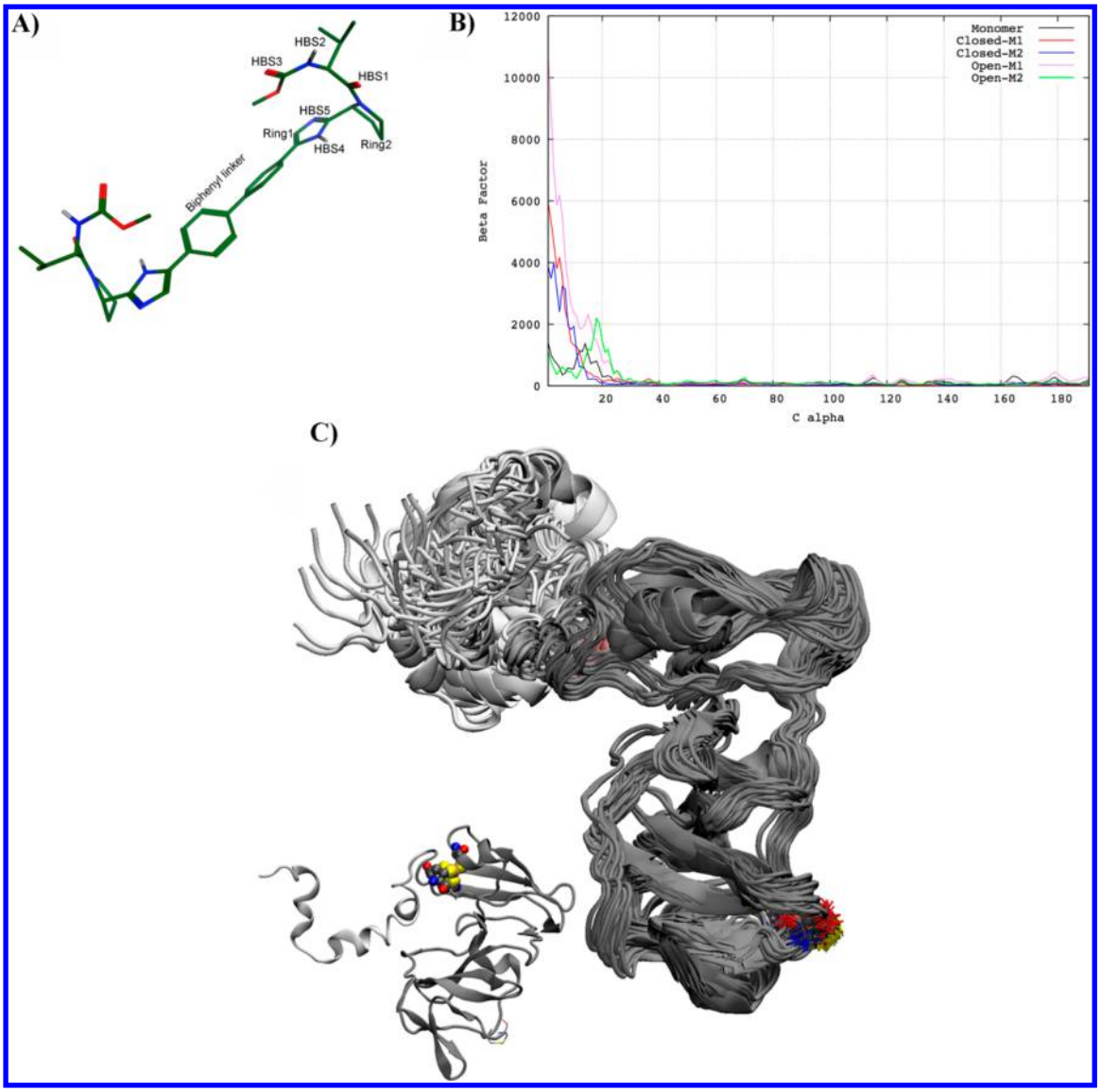


Figure 1. (A) Structure of Daclatasvir. The drug is symmetric with five hydrogen bond interaction sites. (B) Atomic fluctuations (B-factors) of the five monomer structures (two monomers forming the closed conformation, two monomers forming the open conformation, and a single monomer). (C) Single monomer trajectory aligned to investigate the flexibility of the protein structure.

Daclatasvir, developed by BMS, binds selectively in picomolar concentrations to NS5A.[9] The drug (Figure 1A) is currently in phase III clinical trials and has been shown to alter NS5A subcellular localization,[10] block NS5A hyperphosphorylation,[6] and inhibit viral RNA synthesis.[11] Many related NS5A inhibitor analogues have been developed based on the chemical structure of the parent compound.[10,12−14] However, similar to Daclatasvir, they still possess a low resistance barrier to several mutations. This is expected, as such modifications are not directly targeting the protein structural variations due to these mutations. To properly address resistance, a detailed evaluation on where and how Daclatasvir binds to NS5A will greatly assist in future rational drug design of inhibitors against this target.

The present work demonstrates for the first time, at a detailed atomic level, how Daclatasvir and similar dimer pharmacophore compounds bind to NS5A using state-of-the-art molecular modeling methods combined with the massive computational power of the IBM Blue Gene/Q. Our approach involved 28,871 blind docking simulations of Daclatasvir to 125 dominant NS5A conformations generated from unusually long molecular dynamics (MD) simulations. The lowest binding energy model with the best structural fit characteristics correlates remarkably well with the available experimental data. This model thus may be useful in guiding the design of second-generation NS5A inhibitors, which may have less resistance and/or broader activity against this highly diverse virus.

2. RESULTS

2.1. Molecular Dynamics Simulations. Initially, we have run 3 MD simulations for the open and closed conformations as well as for a single monomer structure (see Materials and Methods). The root-mean-square deviations (RMSD) for these three simulations are shown in Figure 2. The closed (Figure 2A) and open (Figure 2B) conformations have similar features that are different from that of the monomer structure (Figure 2C). The closed and open simulations fluctuated below 4 Å, while the trajectory for the monomer structure fluctuated at almost 6 Å. The open conformation experienced larger RMSD

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

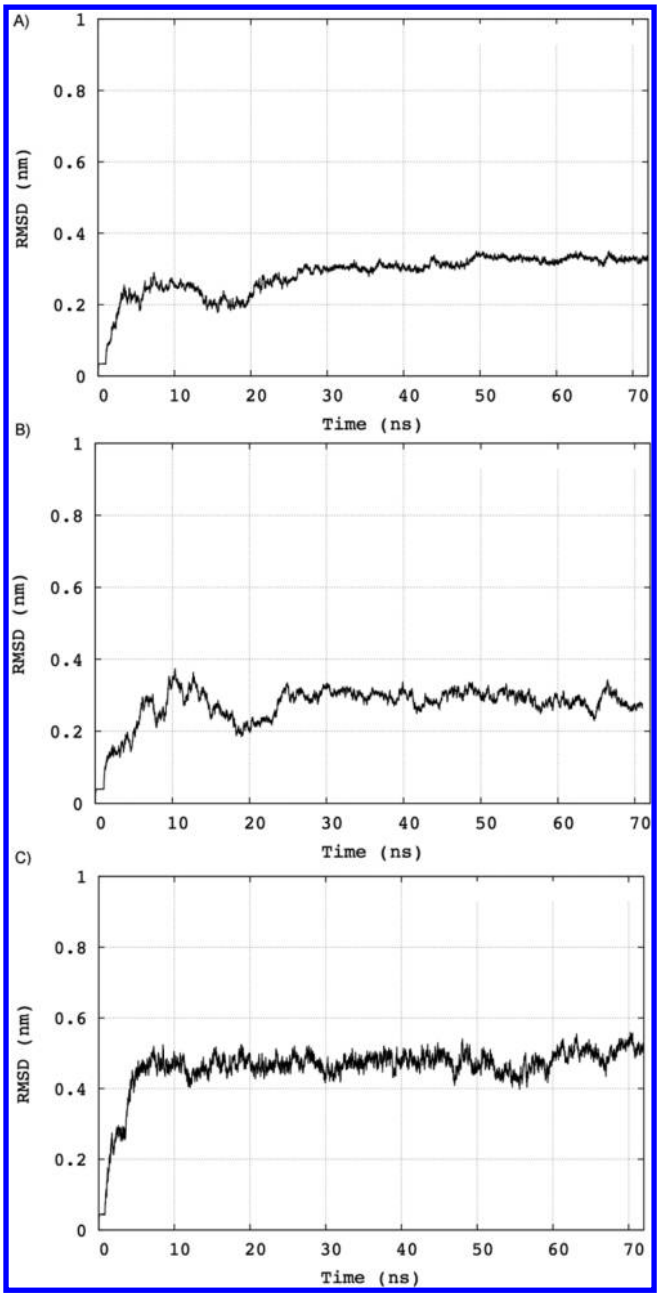


Figure 2. Root-mean-square deviations (RMSD) for the three initial MD simulations for the closed (A), open (B), and monomer (C) structures.

fluctuations during the simulation time compared to the closed structure. In particular, at 10 ns, the open RMSD graph reached its maximum value before descending back to a more rigid conformation followed by a gradual transition to the final equilibrated conformation, which spanned the last 45 ns of the simulation. The same, but minor, behavior was noted in the closed structure. For the monomer simulation, however, the structure seems to have a stable conformation, which was reached progressively with no obvious transitions through intermediate conformations. The high flexibility of the three structures in general can be attributed to the highly flexible N-terminal helices. This is apparent in Figure 1B, which shows the atomic fluctuations per residue for the monomers constituting the closed and open structures compared to the single monomer simulation. For the five monomers (two for the open structure, two for the closed structure, and one for the single monomer), the most flexible region was the N-terminal α-helix. The trajectory for the single monomer simulation is aligned and shown in Figure 1C. Although the α-helix adopted a wide range of different conformations, the rest of the protein structure seems to be rigid and barely fluctuated around a distinct stable conformation.

To investigate whether the RMSD fluctuations for the closed and open systems indicate a possible conformational transition or are due to the enormous flexibility of the N-terminus α-helices, we carried out three additional MD simulations with restraints on the α-helices. We also compared representative structures from three different parts of the closed and open MD simulations representing the initial structure, proposed transition interval, and final equilibrated structures. Figure S1 of the Supporting Information compares the three different states for each system. For the closed system (Figure S1a and b, Supporting Information), it is clear that the α-helices spanned significantly different conformations, while the core protein looks similar to minor shifts at the zinc coordination site, the long loop on the side of the protein, and the β-hairpin at the bottom. Overall, for this system, the NS5A dimer is more open in the transition state compared to the final equilibrated structure. For the open system (Figure S1 c and d, Supporting Information), we noticed almost the same behavior as for the closed system but on a larger scale. In particular, the bottom portion of the protein is more closed in the final equilibrated conformation compared to the initial structure.

Restraining the α-helices had a significant effect on the RMSD graphs for the closed, open, and monomer systems. The RMSD fluctuations were reduced to 0.1, 0.2, and 0.25 nm for the closed, open, and monomer systems (Figure S2, Supporting Information). The RMSD for both the closed and open systems are more stable compared to the monomer, which showed more fluctuations. There was no indication of a possible conformational transition during the MD simulations for the open and closed systems. The B-factor analysis (Figure S3, Supporting Information) for the restrained systems shows that the restrained residues (1−26) have almost zero flexibility as expected. The monomer structure is the most flexible, and the open structure is more flexible than the closed structure (which is almost rigid during the MD simulation). On the basis of these observations, we think that there is no significant conformational transition occurring during the MD simulations for the open or closed systems, and the main fluctuations are due to the high flexibility of the N-terminus α-helices.

Similar observations were obtained by restraining the two α-helices by the membrane lipid bilayer in the closed conformation (Figure S4, Supporting Information). Again, the B-factor (Figure S5, Supporting Information) reveals a great impact of restraining the α-helices on the overall flexibility of the protein. This indicates that the association with the membrane significantly stabilizes the homodimer interactions and considerably reduces its flexibility. Moreover, superimposing the equilibrated models of the membrane-bound and free structures' binding sites revealed important structural differences between the two models (Figure S6, Supporting Information). The clearest differences were located within the hinge region connecting the α-helix to the core of domain I for each monomer. This hinge region includes important residues that have been reported to induce resistance to Daclatasvir such as Leu31, which seems to be rather distant from the surface of the protein in the membrane-bound system compared to the

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

free structure. The same hinge region affects the positioning of the cysteine residues coordinating the zinc ions. Another interesting observation is the sharp structural alternation adopted by loop2 that is part of the α-helix on one of the monomers (Figure S6, Supporting Information). Overall, the upper interface facing the membrane is more closed and compact for the free structure compared to that of the membrane-bound system.

2.2. Clustering Analysis and Preparation for Docking. Over the past decade, many methods and techniques have been developed to introduce protein flexibility into docking simulations. This includes but is not limited to docking ligands to an average structure of a set of different protein conformations, allowing important protein side chains to rotate freely, or docking to an ensemble of protein structures.[15,16] Here, we employed the relaxed complex scheme (RCS) method in our docking simulations.[17] Specifically, Daclatasvir was docked to an ensemble of representative NS5A conformations extracted from MD simulations. The initial MD simulations described above explored the conformational space of the NS5A protein for three different states: closed dimer, open dimer, and single monomer. The number of conformations that one can obtain from these simulations is too large to be included individually during docking simulations. One solution that has been used extensively to reduce the enormous number of protein conformations to a practical representative set of structures that can be used as rigid targets in docking is to cluster these trajectories.[18] Clustering congregates similar conformations into groups in which a representative structure from each group can be used as a rigid target in docking. The objective is to successfully cover the conformational dynamics of the protein structure and expose the dominant conformations to Daclatasvir in order to predict the correct binding location and determine the optimal binding conformation of the drug within NS5A. Figure 3 shows the clustering analysis for the three different MD simulations. To extract all useful information from the three trajectories, we ran the clustering analysis for a wide range of clusters. At every cluster count, two metrics were calculated, namely, DBI and SSR/SST (Materials and Methods). At the optimal number of clusters, one should expect a constant value for the SSR/SST and a local minimum for the DBI values. While many local DBI minima have occurred for the three MD trajectories, the only ones we selected were those that pair with a plateau for the SSR/SST. This correlation took place at cluster counts of 30, 40, and 55 for the closed, open, and monomer simulations, respectively. For each cluster, the centroid structure (the one in the center of the RMSD from the reference structure for each cluster) was taken as the representative conformation for that cluster. All centroids from the 125 different clusters were then used as rigid targets for the blind docking simulations (see below).

2.3. Blind Docking Simulations. In the blind docking procedure, we divided the whole protein surface into independent search boxes that were centered on every 10 surface residues of the protein (see Materials and Methods and Figure S7, Supporting Information). That is, the center of mass of every neighboring 10 residues on the surface of the protein was used as the center of the grid box. This required 28,871 independent docking simulations to cover the 125 different conformations that were obtained from the clustering analysis of the MD trajectories. Each docking box restricted the translational space of Daclatasvir to a smaller search space that

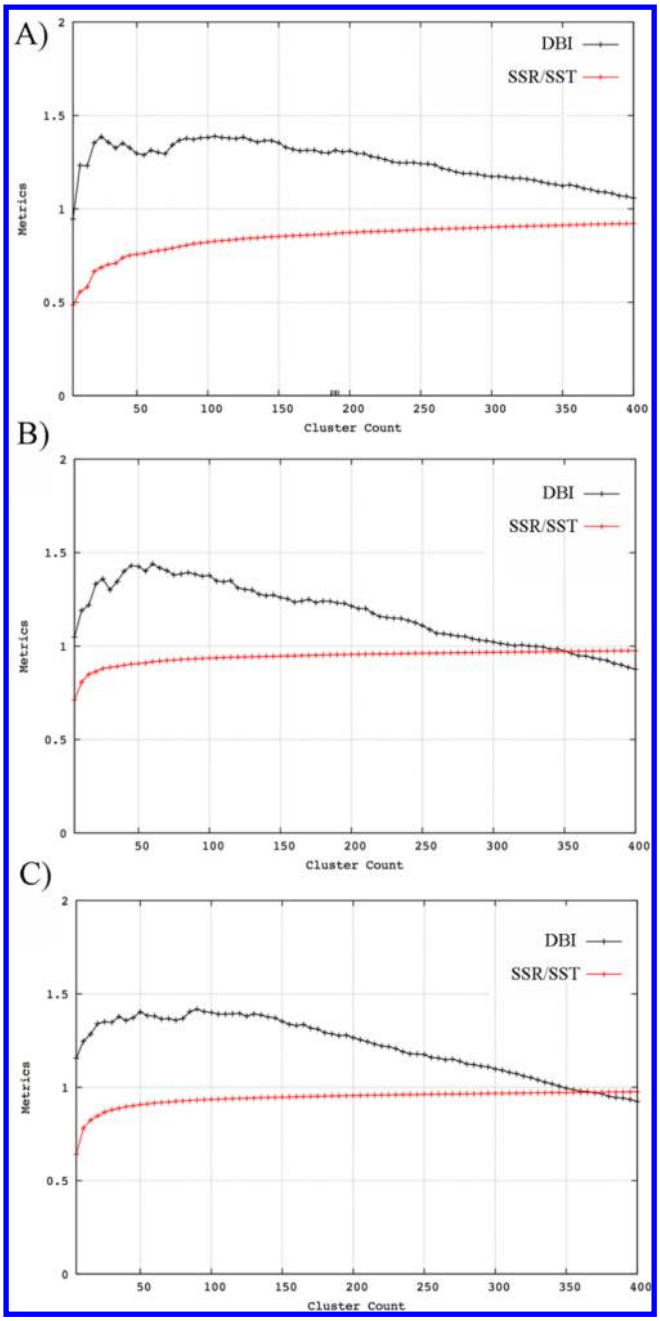


Figure 3. Clustering analysis for the three systems: (A) closed structure, (B) open structure, and (C) monomer structure.

can fit the ligand. Moreover, by docking Daclatasvir to three different ensembles of rigid protein structures, the protein flexibility was partly incorporated within the docking simulation with no need to include side chain flexibility. This was important because we did not know at this point which side chains should be allowed to be flexible during docking as the binding site had not been determined yet.

The major benefit of using the multiple-docking method described above is that it converts the convergence of all possible solutions from clustering of the translational and rotational spaces of the ligand to clustering of the docking boxes. This means that each box represents an independent and focused docking problem that tries to approach the binding site location from a different angle and with a very small increment in the translational space. Each box has twice the dimensions of

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

Daclatasvir to allow the ligand to rotate and move freely within the docking search space. A binding location was identified when a number of boxes were found to be clustered around this location. Figure 4 shows the highest and lowest interaction sites between Daclatasvir and the NS5A protein for the three different studied states. The protein residues are colored with the binding energies where the lowest binding energy sites are shown in red, and those with the highest binding energies are shown in blue. In the closed homodimer (see Figure 4A–C), blind docking identified a symmetric site located at the upper interface between the two monomers. In the monomer structure, however, the drug can bind to two different sites. The first is a binding groove on the middle surface of the protein, opposite to the dimer interface (Figure 4D). The second is on the top of the monomer structure close to the α-helix (Figure 4E). The latter site seems to be too small to accommodate Daclatasvir compared to the closed homodimer conformation (Figure S8a, Supporting Information). Only part of the ligand can fit within a small pocket formed by the upper surface of the monomer and part of the α-helix, while the rest of the drug is protruding from the surface of the protein. Similar to the closed structure, the open homodimer conformations yielded a dominant binding site in the upper region (Figure 4F–H). Although parts of the drug bind well at this position (Figure S8, Supporting Information), the two helices block the site and prevent close association of Daclatasvir.

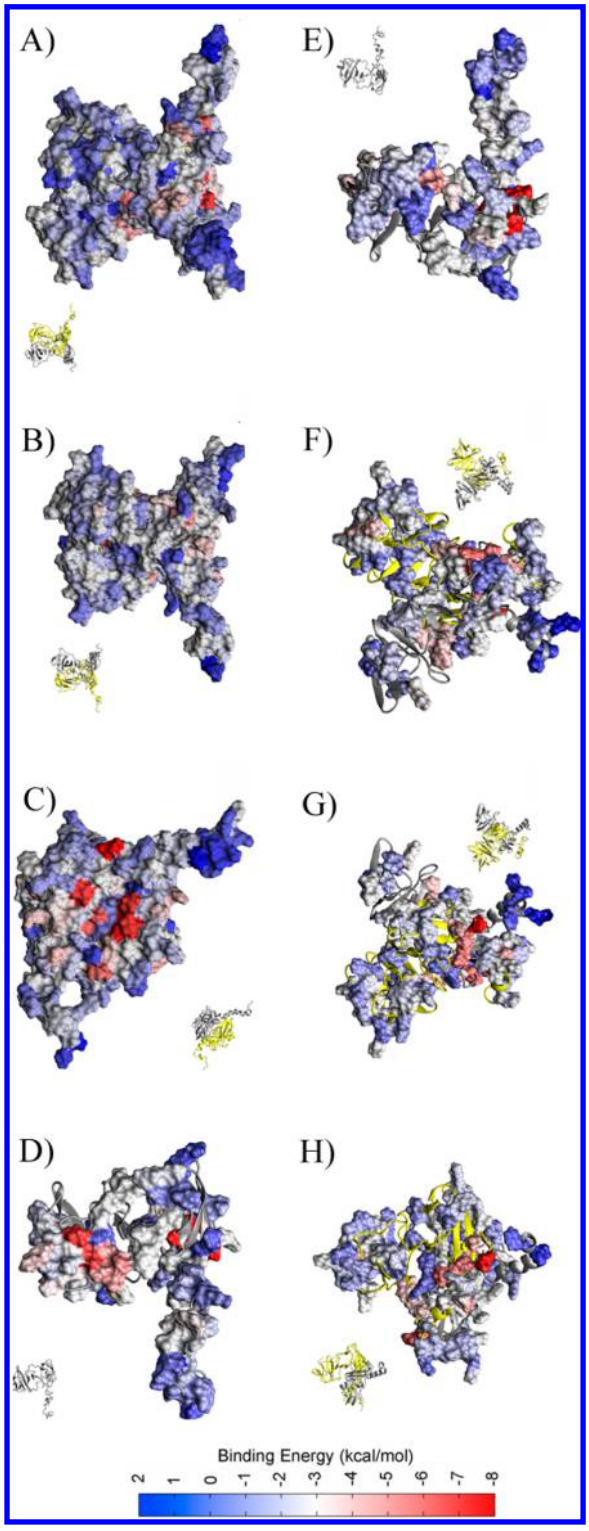


Figure 4. Blind docking analysis. The protein surface is colored with the interaction energy with Daclatasvir with red as the most attractive and blue as the least attractive. Structures for the closed conformation are shown in (A), (B), and (C), for the monomer conformation in (D) and (E), and for the open conformation in (F), (G), and (H).

2.4. Identifying the Daclatasvir Binding Site. Revealing the binding location is a different problem from identifying the mode of binding of the ligand. The former tries to identify the locations in which the ligand can roughly fit into the protein surface, while the latter tries to identify exactly how it fits. Once a binding location was identified, we turned our attention to the problem of determining the most probable mode of binding of the ligand within the discovered binding site. This required more focused docking parameters (mainly increasing the maximum number of energy evaluations and population size (see Materials and Methods) and introducing more flexibility of the target protein by using all clusters' representative structures. Although blind docking limited the potential binding sites for Daclatasvir to a few locations, it was now necessary to identify the best site that is consistent with the available plethora of experimental data and explain the effects of the reported resistance-associated mutations.[19,20] On the basis of this approach, only the top pocket appears viable but with three different scenarios: a symmetrical and complementary site in the closed confirmation (site 1), a blocked pocket with the two helices in the open conformation (site 2), and a small groove in the top part of the monomer (site 3). Although, site 1 appeared to be the most appropriate location for Daclatasvir binding, we thought it would be prudent not to exclude any of the three possibilities before performing more rigorous analysis on these sites. For each pocket, Daclatasvir was redocked with more intensive parameters to the whole cluster representatives of the three systems, relaxed for 10 ns using MD simulations, and finally, its binding affinity to NS5A was estimated using the Molecular Mechanics/Poisson–Boltzmann Surface Area (MM-PBSA) technique.[21] The binding affinities for Daclatasvir within three sites were obtained as -12 ± 2, -7 ± 2, and -6 ± 1 kcal/mol for sites 1, 2, and 3, respectively. On the basis of the binding affinity, the spatial fit of the drug into a pocket, and the correlation with resistance-associated mutations, we predict site 1 to be the binding site.

2.5. Daclatasvir Binding Mode. The drug was then redocked to the whole set of 30 representative structures for the closed conformation. Figure 5 illustrates our refined binding model for Daclatasvir within NS5A. Starting from this model,

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362–373

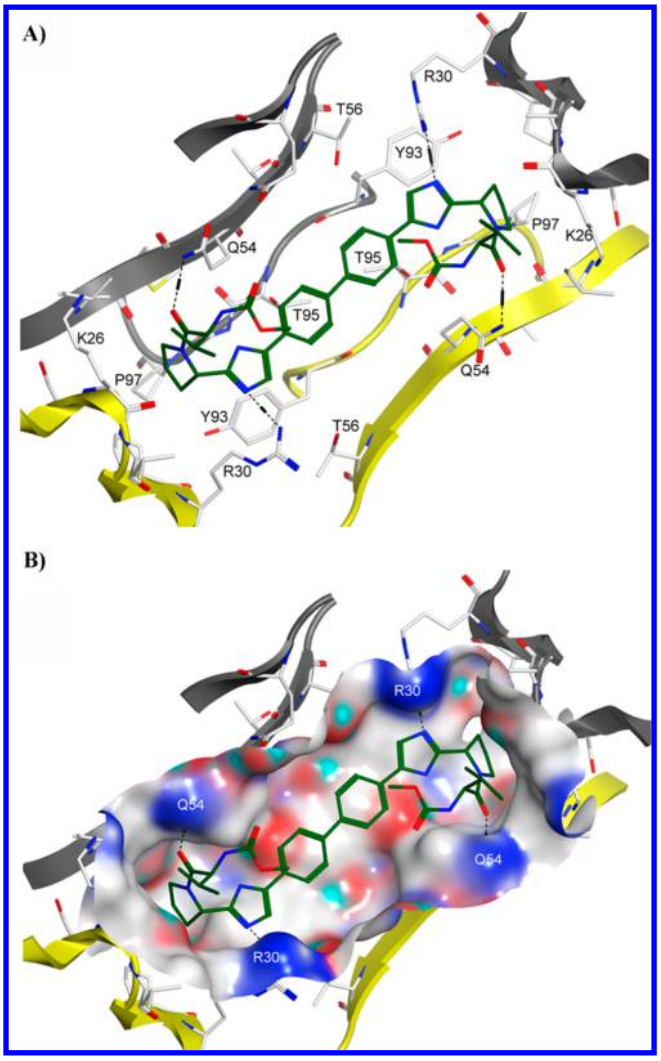


Figure 5. Binding mode for Daclatasvir. (A) Carbon atoms are colored in green for Daclatasvir and in white for the protein residues. Hydrogen bonds are shown as dotted lines, and the protein boundaries of the binding site are shown in cartoon representation, where the first monomer is shown in yellow while the second monomer is shown in gray. (B) The binding site is shown in surface representation. For more details on the interactions between Daclatasvir and the different residues, please refer to the text.

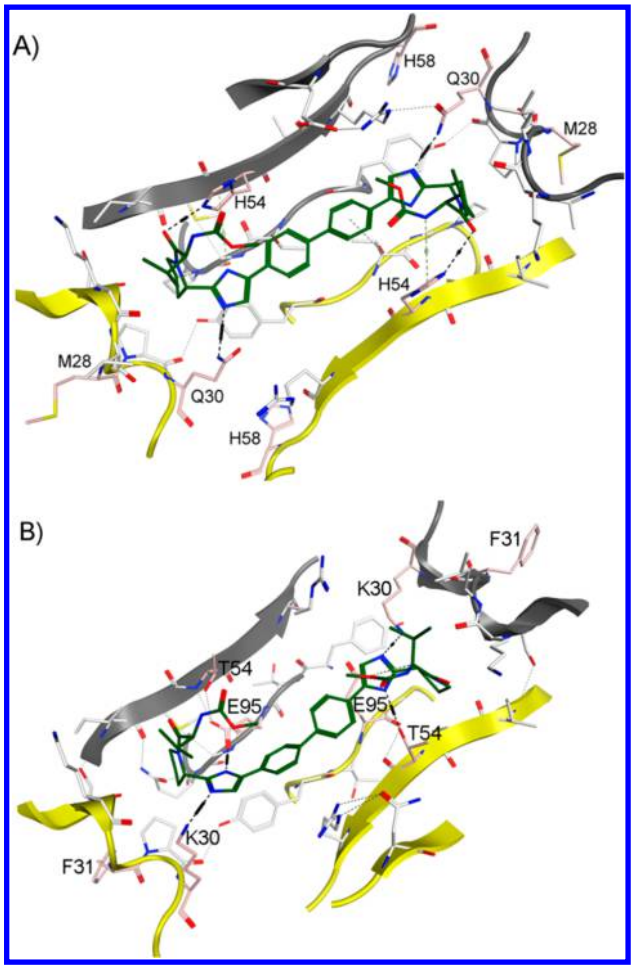


Figure 6. Binding mode for Daclatasvir within genotype 1a (A) and genotype 2a (B). The coloring scheme is same as described in Figure 5.

we created homology models for two different genotypes (GT-1a and GT-2a). We also investigated the effects of three mutations (see below) on the binding mode and affinity of the drug to NS5A (see Discussion for details). For each of the five systems, the solvated NS5A−Daclatasvir systems were subjected to 20 ns MD simulations to relax their interactions and analyze their hydrogen bond networks. Figures 6, 7, and 8 show the binding mode of Daclatasvir to the five systems and explain the effects of the different amino acid substitutions on the binding of the drug. Interestingly, we did not find any water-mediated interactions. This was expected because the interface between Daclatasvir and NS5A shows mainly hydrophobic interactions without possibilities for water molecules to mediate those interactions (see below) with the exception of the four hydrogen bonds described in this study.

2.6. Construction of the Daclatasvir-NS5A Pharmacophore. We constructed a structure-based pharmacophore model using the predicted binding mode of Daclatasvir within NS5A (Figure S9, Supporting Information). Our pharmacophore hypotheses include two hydrophobic regions (preferably aromatic rings) resembling the biphenyl rings and interacting with T95 of NS5A, two hydrophobic features at the edges of the molecule to interact with P97, and finally, two hydrogen bond acceptors to interact with R30. These pharmacophore hypotheses were sufficient to successfully dock previously reported Daclatasvir derivatives within NS5A (see below).

2.7. Pharmacophore-Based Docking of Previously Reported Cases. To validate the binding mode and developed pharmacophore, we carried out pharmacophore-based docking simulations using a set of Daclatasvir analogues. We use the molecular operating environment (MOE) software to construct the pharmacophore and run the pharmacophore-based docking simulation[22] (see Materials and Methods). The docked structures are shown in Figures S10 and S11 of the Supporting Information. Interestingly, regardless of the overall structure of a given compound, the more it adheres to the suggested pharmacophore hypotheses the more it can bind efficiently to NS5A. This is shown in Figure S10 of the Supporting Information for compounds (**1**) dac-28 and (**2**) dac-21 that have been reported to inhibit NS5A for genotype 1b with EC50s of 0.001 nM and ≤1 nM, respectively.[23] This is also illustrated for compounds (**3**) 57610996 and (**4**) 70843595 obtained as similar structures to Daclatasvir from the PubChem database (Figure S10, Supporting Information).

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

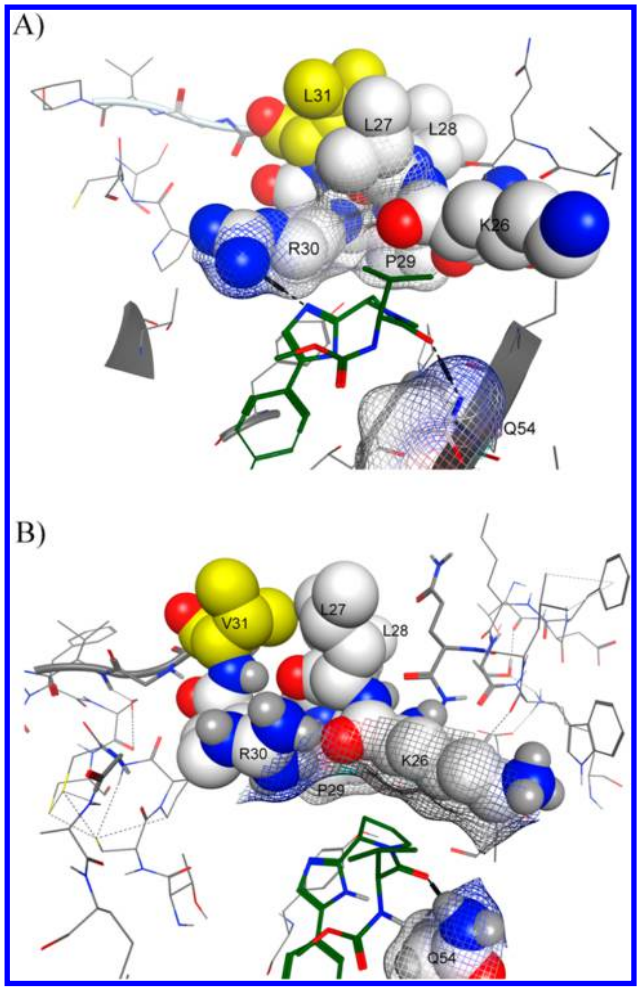


Figure 7. Significance of the L31 to the binding of Daclatasvir to NS5A. (A) L31 interacts with the surrounding hydrophobic residues to construct a hydrophobic cluster that accommodates ring 2 of Daclatasvir. (B) Mutating L31 into valine disrupt this hydrophobic cluster.

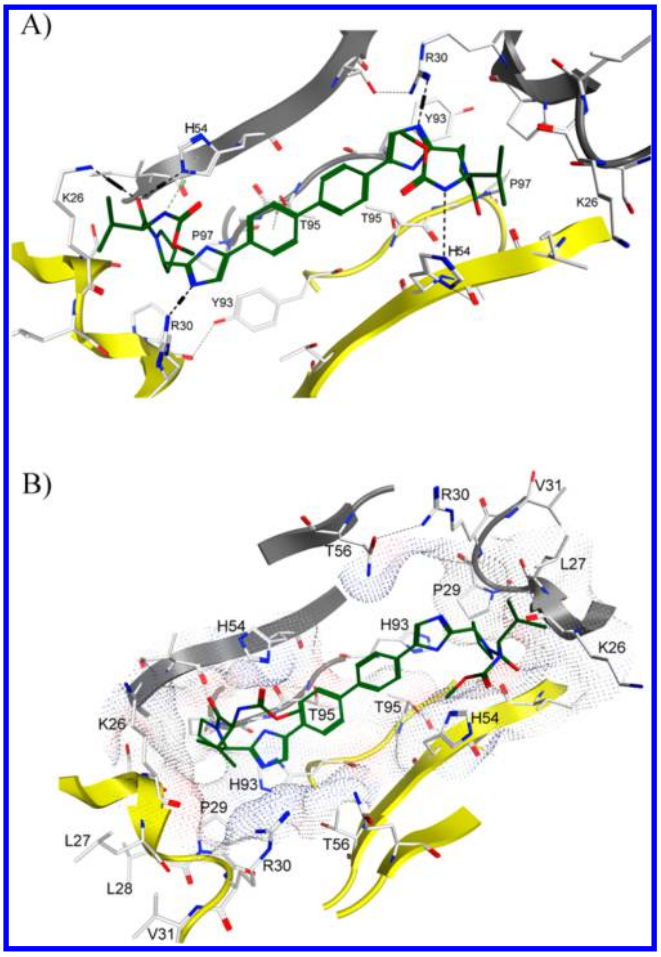


Figure 8. Effects of the Q54H mutation (A) and the triple L31V, Q54H, and Y93H mutations (B) on the binding of Daclatasvir to NS5A.

Compounds that do not satisfy the pharmacophore hypotheses as described above failed to dock properly with the NS5A binding site. Examples include (**5**) dac-27, which has been reported as a weak NS5A inhibitor with an EC50 that is greater than 10 nM[23] (Figure S11, Supporting Information). Likewise, compound (**6**) 58209591 did not dock properly to the binding site. The former lacks the two hydrophobic regions mimicking the biphenyl rings and is extremely flexible in the middle, while the latter has a longer linker between the biphenyl rings preventing them to allocate correctly on the top of the T95 residues from both monomers.

2.8. Predicting Potential Resistance Patterns. The detailed model of Daclatasvir bound to NS5A mentioned above allowed us to predict potential resistance patterns that can develop upon using the drug for HCV treatment. We are primarily concerned with the emergence of new unreported resistant mutations and providing the tools to assist drug developers to design new analogues of Daclatasvir that can be effective in the treatment of HCV with these mutations. In addition to the previously reported mutations discussed below, we also predicted that new mutations might arise generating a conflict with the pharmacophore model described above. For example, changing P97 with a more flexible residue such as glycine would affect the loop region that is holding the two hydrophobic rings at the edges of Daclatasvir. Another example that is shown in Figure S12 of the Supporting Information is an Arg30 mutation, which provides a hydrogen bond to Daclatasvir into a shorter and nonpolar residue such as valine would lead to resistance. In this case, one way to rescue this mutation and develop a new analogue of the drug that would bind efficiently to NS5A would be to replace the nitrogen that acts as a hydrogen bond acceptor in the middle ring of Daclatasvir by a longer hydrophobic moiety that would interact with the valine residue.

3. DISCUSSION

Over 100 patent applications have been filed during the past few years claiming various NS5A inhibitors that are based on dimer pharmacophore structures to which Daclatasvir belongs. To date, all NS5A inhibitors were identified through cell-based replicon screens without knowing where and/or how they bind to NS5A.[24,25] In light of replicon screening data, the most sensitive HCV genotypes to Daclatasvir appear to be GT-1b with an EC50 value of 0.004 nM, while the least sensitive to Daclatasvir is GT-2a (L31M) with an EC50 value of 6.7 nM with an almost 1700-fold difference in potency.[26] Interestingly, the most resistant GT-1a variant is the one with a single amino acid substitution Y93N with an EC50 value of 282 nM, a 14,100-fold difference compared to wild type.[26] These are not

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

the only mutations that show resistance to Daclatasvir. For instance the replicon assay EC50 values of Daclatasvir for GT-1a mutations M28T, Q30E, Q30H, Q30R, L31V, Y93C, and Y93H were 4.1, 150, 8.7, 7.3, 2.1, 20, 11.1, and 32 nM, respectively.[6,9,27]

An essential step to predict an accurate binding mode of Daclatasvir to NS5A was to determine the exact location in which the drug binds to the protein. Although mutations that confer resistance to Daclatasvir have been localized within domain I of the protein,[19] on which we focused in this study, this domain is still large enough to accommodate many binding sites for the drug. Other studies also show that Daclatasvir interferes with RNA binding to NS5A, which may indicate a competition for binding between the two molecules. Here, we employed a technique that we used previously to aid the discovery of the binding location of Laulimalide to tubulin.[28] We used blind docking in which the drug was docked extensively to the surface of the protein and the binding energies and binding conformations were monitored and ranked throughout the simulation. It is important to mention that blind docking is a recently developed technique that is developing as a robust tool. Conventional methods employ a single large box that covers the whole surface of the protein to search for putative binding sites. Conventional methods also use a rigid protein structure as the target for docking.[29−32] Although using a single box is adequate for small proteins where the spacing between the grid points is small enough to provide sufficient grid resolution for docking (given the fact that there is a maximum limit of 126 grid points for AUTODCOK[31−33]), the same method does not work well for larger proteins. Here, we employed a new methodology, developed in-house, to account for the protein flexibility and to provide the highest grid resolution possible for the docking simulations. The same procedure helped in identifying the binding location of the well-known microtubule stabilizing agent, Laulimalide, on the surface of tubulin.[28]

Strikingly, Daclatasvir has a unique symmetrical binding mode with a symmetrical binding site (Figure 5). The biphenyl rings rest on top of the two methyl groups of Thr95 in the two monomers (Figure 5A). The rings of the two Tyr93 residues extend the hydrophobic channel to accommodate the remainder of the biphenyl linker and ring 1 of Daclatasvir. The long chains of Lys26, Ile27, and Ile28 from the two α-helices and Pro97 from each monomer complete the hydrophobic corners of the pocket. Ring 2 of the drug fits snuggly into the hydrophobic groove formed at the corner of the pocket (Figure 5B). Arg30 and Gln54 form four symmetrical hydrogen bonds with Daclatasvir. Depending on the orientation of Gln54 and Daclatasvir, one of three possible hydrogen bonds can be formed with HBS1, HBS2, or HBS3 (Figure 1A). On the other hand, Arg30 can only form a hydrogen bond with HBS5. Interaction of the drug with Tyr93 helps in maintaining these hydrogen bonds by properly carrying ring 1 of Daclatasvir to the right position. As described below, these hydrogen bonds seem to be essential for the activity of the drug and provide a rationale for the observed resistant mutations.

These findings allowed us to construct a pharmacophore model that involves four essential features necessary for a successful Daclatasvir-based structure (Figure S9, Supporting Information). Docking of structures similar to that of Daclatasvir within the identified binding site and the resulting correlation of their fitting with reported activities (see Results) provided an important validation of our pharmacophore model.

Restraining the two α-helices had a great impact on the overall flexibility of the protein structure and, most notably, on the Daclatasvir binding site and its surrounding residues. This was observed in two different scenarios. In the first, we induced restraints on the α-helices for the closed, open, and monomer systems during MD simulations. In the second, the two α-helices were submerged in the membrane bilayer. In both cases, this restriction on the movement of α-helices tremendously increased the rigidity of the protein dimer and affected the topology of the Daclatasvir binding site by making it more open in the membrane-bound case compared to the free structure. This indicates that Daclatasvir may only bind to the free form of the closed NS5A dimer. This selectivity for binding to the closed conformation has an important implication for the binding of RNA to the NS5A dimer, which seems to bind in between the two monomers in the open conformation. By binding only to the closed structure, Daclatasvir stabilizes the NS5A into this conformation, shifting the equilibrium to more closed dimers and reducing the ability for RNA to bind to NS5A and thus inhibiting replication.

Our model is in agreement with the available experimental data for the different genotypes and reported mutations. Figure 6A illustrates the mode of binding of the drug to genotype 1a. Within the Daclatasvir binding site, there is a very high degree of similarity between the G1b and G1a in NS5A (Figure S13, Supporting Information). The main differences are in changing Arg30 in G1b to glutamine and changing Gln54 to histidine. This residue change forms a perfect fitting space for Daclatasvir to form the observed four hydrogen bonds. Pro58 is far from the binding location of Daclatasvir and the structural effect of its mutation to histidine is not clear. Leu31 forms a hydrophobic cluster with Leu27 and Leu28 (Figure 7A) that restricts the orientation of Gln30 (R30 in GT-1b) to establish a strong hydrogen bond with the HBS5 site of Daclatasvir, which complements the hydrophobic corner of the binding pocket for ring 2 of the drug. Mutating Ile31 into methionine still maintains this hydrophobic cluster with no major effect on the affinity of the ligand. Analogously, Figure 6B illustrates the binding mode of Daclatasvir within genotype 2-a. Changing Arg30 into lysine and Gln54 into threonine sustained the four hydrogen bonds required for Daclatasvir activity. The three hydrogen-bonding sites (HBS1−HBS3) in the drug provided several alternatives to keep these hydrogen bonds. As shown in Figure 6B, they can either form a direct hydrogen bond with the hydroxyl group of Thr54 or have a bridged hydrogen bond with Thr54 through an interaction with Glu95.

The model described here also provides a clear picture of the effects of many drug-resistant mutations described in the literature. For example, mutating Leu31 into valine in GT-1a has been reported to significantly increase the EC50 of the drug from 0.02 to 20 nM, a 1000-fold difference.[19] This mutation (Figure 7B) replaces leucine with a shorter hydrophobic residue, which reduces the effect of the hydrophobic cluster required to coordinate a strong hydrogen bond with Arg30. It also alters the hydrophobic groove for ring 2 of Daclatasvir making it wider than it should be to properly fit this ring. As a result, the binding site is clearly distorted at this location resulting in a direct effect on the affinity of the ligand. Similar to the Q54H mutation described above for GT-1a, mutating Gln54 into asparagine does not affect the hydrogen bond network (Figure 8A). The most disruptive mutation that has

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

been reported is a combination of three mutations namely, L31V, Q54H, and Y93H (Figure 8B).[19] Changing Tyr93 into histidine displaced the hydrophobic linker and ring 1 in the drug from the proper position to form the hydrogen bonds with Arg30 and His54 (which is already shorter than the wild type Q54). In addition to that, the substitution of Leu31 to valine deforms the hydrophobicity of the pocket as described above. This consistency of our model with the available experimental data described in the literature strongly indicates that our model accurately predicts the binding location of Daclatasvir and its mode of binding within NS5A. As such, our model can be used to design new drugs active against a drug-resistant virus as well as drugs with greater resistance barriers and with broader activities (against the multiple highly variable HCV genotypes) as compared with Daclatasvir (work in progress).

4. MATERIALS AND METHODS

4.1. Preparation of the Initial Structures. This study focused on three HCV genotypes (1b isolate Con1, 1a isolate H77, and 2a isolate JFH-1) along with mutations (L31V, Q54N, and the triple mutation L31V, Q54H, and Y93H). To represent the core structure of NS5A, two crystal structures were used, namely, the closed conformation (PDB code 3FQM)[7] and the open conformation (PDB code 1ZH1)[8]. The monomers of the closed structure comprised residues 32−191, although one of the monomers lacked the Leu32 residue. The open structure spanned residues 36−198. To add the N-terminal α-helices, we used the available Penin's NMR structure (PDB code 1R7D),[34] which spanned residues 1−31. The software Swiss-PdbViewer (http://spdbv.vital-it.ch/) was used to ligate the backbone atoms of the core and the added N-terminal α-helix for each monomer. To make the open and closed structures equal in length, we added residues 32 and 192 to 198 to each monomer of the closed structure (Figure S13, Supporting Information for the sequence alignment of the used structures in this study). In adding the missing residues, we also used the Swiss-PdbViewer software. Rotamer libraries were used to avoid any steric clashes of the newly added amino acid with its environment and to maximize its hydrogen bonding with the surrounding amino acids. As this α-helix belongs to a 1a genotype, homology modeling was used to build a 1b genotype version of the helix before adding it to the 1b genotype structures. That is, following the modeling of the 1b version of the α-helix, it was integrated within the two monomers of the closed conformation and the two monomers of the open conformation (Figure S14, Supporting Information). Thereafter, we selected the monomer structure that originally included residue Leu32 from the closed homodimer after adding the missing residues as described above to prepare a monomer structure for subsequent MD simulations and blind docking analysis (see below). This was important to determine if Daclatasvir could bind to the interface between the two proteins forming the homodimer. Because the optimal binding site for Daclatasvir was observed only in the closed structure (see Results), we modified the closed conformation to generate the three mutants and the other two genotypes. To generate structures for the other variants, their sequences were obtained from the Protein Knowledge Database (UniProtKB: http://www.uniprot.org/) and were aligned to the sequence of the used structure using the alignment utility of the same database. This was followed by mutational analysis using the Swiss-PdbViewer software (http://spdbv.vital-it.ch/) to replace every amino acid from 3FQM with its counterpart that represents the targeted genome or mutation. For each monomer, there were four cysteine residues coordinating a zinc ion (residues C39, C57, C59, and C80) and a disulfide bond formed by residues C142 and C190.

Co-crystallized water molecules required careful consideration in our study. It is well known that water molecules that are located close to or within the binding site can mediate several interactions with the ligands. However, a conserved water-bridged interaction requires the presence of the investigated ligand during the experimental structure determination process. As the structures that we used in this study had no Daclatasvir bound, we decided to include all co-crystallized water molecules only during the MD simulations in order to help refine the protein structures and extract the most realistic set of protein conformations for docking. During blind docking simulations and similar to many reported cases in the literature,[30,31] all water molecules were removed from the equilibrated structures assuming that they should be displaced by the ligand during the binding process. The importance of water molecules in Daclatasvir binding was investigated during the 20 ns explicit solvent all atom MD simulations that were performed on the refined ligand complexes systems (see below). All co-crystallized ligands were removed as they are used merely for crystallization purposes.[7]

4.2. Molecular Dynamics (MD) Simulations. For all generated structures, MD simulations were carried out using the NAMD program,[35] at a mean temperature of 300 K and physiological pH (pH 7). The simulation protocol followed the same procedure as described in previous work.[18,36−38] In brief, we carried out three initial MD simulations on the closed, open, and monomer structures using the all-hydrogen AMBER99SB force field[39] simulated in a 12 Å-wide buffer of water molecules. Protonation states of all ionizable residues were calculated using the program PDB 2PQR[40] followed by adding the proper concentrations of sodium and chloride ions to neutralize the systems. Each solvated protein was then minimized, heated with heavy restraints on the backbone atoms, equilibrated for 100 ps with a gradual removal of the restraints, and finally run for 70 ns of MD simulation. Following the same protocol, we carried out an additional 13 MD simulations. The first set comprises three MD simulations with 10 ns on the top hits from docking simulations as a post-processing analysis to the blind docking simulations to rank the identified pockets. The second set includes six MD simulations on the bound ligand in the selected binding site for 20 ns each to analyze the hydrogen bond network. These simulations were also used to generate an ensemble for binding energy calculations by storing the trajectories every 10 ps. That is, 2000 snapshots from each system were used for subsequent binding energy calculations using the MM-PBSA method. However, due to the high computational cost for the entropy calculations, we used only 50 snapshots from the whole 2000 structures to calculate the entropy contributions for each system (see below). The third set included three 50 ns MD simulations in which residues 1−26 from each monomer were restrained in motion. The final set comprised only one 30 ns MD simulation for a membrane-bound dimer in the closed conformation. To construct a model with the two α-helices bound to the membrane bilayer, we used the CHARMM-GUI Lipid Builder.[41] Overall, 256 lipid molecules were used to submerge and restrain the two α-helices (128 on the lower leaflet and 128 on the upper leaflet). Water molecules were used to solvate the membrane-bound system with a 15 Å buffer surrounding the system. Sodium and chlorine ions were added to neutralize the system and provide a 150 mM ionic concentration. The lipid 11 AMBER force field was used to model the membrane,[42] and the AMBER99SB force field[39] was used for the protein. Initially, the protein backbone and membrane atoms were restrained during minimization and heating, followed by gradually removing the restraints. A 5 ns equilibration phase was then used with very weak restraint on the lipid atoms (0.1 kcal/mol), followed by 25 ns of production run where the whole system was free to move.

For the ligand-bound systems, the ligand parameters were obtained using the generalized AMBER force field (GAFF).[43] For each ligand, partial charges were calculated with the AM1-BCC method using the Antechamber module of AMBER 10.

Root-mean-square deviations (RMSD) and B-factors were computed over the duration of the simulation time using the PTRAJ utility. Hydrogen bond analyses were performed by computing the average distance between donor and acceptor atoms. A hydrogen bond was defined by a heavy donor−heavy acceptor distance $\leq$3.4 Å, a light donor-heavy acceptor distance $\leq$2.5 Å, and a deviation of less than $\pm$40° from linearity.

4.3. Clustering Analysis. Here, we followed the same clustering methodology we developed in previous work. The RMSD conforma-

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

tional clustering was performed using the average-linkage algorithm using cluster counts ranging from 5 to 500 clusters. Clustering analysis was performed on the three initial MD simulations for the monomer and open and closed conformations. Structures were extracted at 10 ps intervals over the entire 70 ns simulation times. All C_α-atoms were RMSD fitted to the minimized initial structures in order to remove overall rotation and translation. The clustering quality was anticipated by calculating two clustering metrics, namely, the Davies−Bouldin index (DBI)[44] and the "elbow criterion".[45] A high-quality clustering scheme is expected when DBI experiences a local minimum versus the number of clusters used. On the other hand, using the elbow criterion, the percentage of variance explained by the data is expected to plateau for cluster counts exceeding the optimal number of clusters.[45] Using these metrics, by varying the number of clusters, one should expect for adequate clustering, a local minimum for DBI, and a horizontal line for the percentage of variance exhibited by the data (see Results).

4.4. Blind Docking. Representatives of all clusters (125 distinct NS5A conformations) were used as rigid targets for the blind docking simulations. These conformations represent clusters from the monomer (55), closed (30), and open (40) structures. For each conformation, the surface residues were detected using Sanner's molecular surface prediction software,[46] and the center of mass of every neighboring 10 surface-exposed amino acids was used as the center of a docking box. The average number of docking boxes used was 156, 304, and 271 for the monomer, open, and closed structures, respectively. That is, a total of 28,871 independent docking simulations for Daclatasvir on NS5A. Figure S1 of the Supporting Information shows an example of a single blind docking simulation using a structure from the closed representatives ensemble. All docking runs were performed using AUTODOCK,[33] version 4.028. Each docking box expanded for 85 grid points in each direction with spacing of 0.35 Å between every two adjacent points, enough to cover twice the size of Daclatasvir. The docking method and parameters were similar to the ones used in our previous work.[38,47] Using the Lamarckian Genetic Algorithm (LGA), the docking parameters included an initial population of 150 random individuals, a maximum number of 10,000,000 energy evaluations, 100 trials, 30,000 maximum generations, a mutation rate of 0.02, a crossover rate of 0.80, and the requirement that only one individual can survive into the next generation. For subsequent focused docking upon identifying potential binding site, the maximum number of energy evaluations and population number were doubled.

4.5. Pharmacophore Modeling and Pharmacophore-Based Docking. A pharmacophore is a straightforward model that describes the essential interactions (features) behind the binding of a ligand to its target. Using the molecular operating environment (MOE), we constructed a structure-based pharmacophore model with Daclatasvir bound to NS5A using the pharmacophore quarry editor in MOE.[22] The model elucidated the essential features for Daclatasvir to bind to NS5A.

We then used the docking module in MOE to dock similar structure to Daclatasvir within the identified binding site. The pharmacophore features described in the pharmacophore model were used as restraint on the docking simulation to rapidly align the docked structures within the binding site. The pose rescoring docking protocol was used as the method to finally rank the docked structure. All receptor atoms were included in the energy evaluation, while the docking space was restricted to the Daclatasvir binding site. The pharmacophore was used as the placement method, and two scoring functions were used to evaluated the binding interaction of the docked ligands to the binding site, namely, the London method and the GBVI/WSA method, where only the top five structures were retained for visual inspection and further analysis.[22]

4.6. Binding Energy Analysis. We used the MM-PBSA technique[21] to predict the binding energies. Similar to the work described previously in the literature,[37,47−49] the total free energy for each system was estimated as the sum of the average molecular mechanical gas-phase energies (E_{MM}), solvation free energies (G_{solv}), and entropy contributions (TS_{solute}) of the binding reaction

$$G = E_{MM} + G_{solv} - TS_{solute} \quad (1)$$

The molecular mechanical (E_{MM}) energy of each snapshot was calculated using the SANDER module of AMBER10. The solvation free energy (G_{solv}) was estimated as the sum of electrostatic solvation free energy, calculated by the finite difference solution of the Poisson−Boltzmann equation in the Adaptive Poisson−Boltzmann Solver (APBS) and nonpolar solvation free energy calculated from the solvent-accessible surface area (SASA) algorithm. The solute entropy was approximated using the normal mode analysis. For each protein−ligand complex, the binding free energy was approximated by the difference between the bound and free systems

$$\Delta G^o = G_{gas}^{\text{NS5A-Daclatasvir}} + \Delta G_{solv}^{\text{NS5A-Daclatasvir}} + \{\Delta G_{solv}^{\text{Daclatasvir}} + \Delta G_{solv}^{\text{NS5A}}\} \quad (2)$$

The parameters used include a dielectric constant for the protein−ligand complex of 1, a dielectric constant for the water of 80, an ionic concentration of 0.15 M, and a surface tension of 0.005 with a zero surface offset to estimate the nonpolar contribution of the solvation energy.

We selected 2000 snapshots from each trajectory to predict the molecular mechanics and solvation contributions; however, for the entropy, we only used 50 snapshots for each system. Our selection of the snapshots' frequency was based on estimating the correlation time similar to the work described by Genheden and Ryde.[50] That is, the delta MM-PBSA energy points from the whole MD trajectory (X) have been divided into blocks (Y_i) of equal time spaces (τ). The function Φ is then calculated according to the following equation

$$\Phi = \frac{\tau \times \sigma^2(Y)_\tau}{\sigma^2(X)} \quad (3)$$

where $\sigma^2(X)$ is the variance of the whole trajectory delta MM-PBSA energy points and $\sigma^2(Y)_\tau$ is the variance of the averages of the energy data points within the blocks of length τ (i.e., for each block, the average delta energy is calculated and then the variance of the n blocks generated is then used in eq 3 as $\sigma^2(Y)_\tau$ for a certain τ). The length of the block (τ) is then varied. The values of Φ are expected to be constant when the block averages are statistically independent, and at this point, the time correlation can be estimated.

5. CONCLUSION

Hepatitis C virus (HCV) infection is a major health problem.[1] The current standard-of-care therapy (PEGylated interferon and ribavirin along with new protease inhibitors) is not fully effective and induces substantial adverse side effects. Therefore, alternative direct-acting antivirals (DAAs) for HCV treatment are currently under development.[2] Most of the DAAs are focused on targeting three viral proteins involved in the HCV replication cycle, namely, NS3/4A protease, NS5A protein, and NS5B RNA polymerase.[3]

Daclatasvir is the first in the class of NS5A inhibitors ever developed and the most potent HCV inhibitor so far. The drug binds selectively in picomolar concentrations to NS5A.[9] It has been shown to alter NS5A subcellular localization,[10] block NS5A hyperphosphorylation[6], and inhibit viral RNA synthesis.[11] Despite its extraordinary activity, Daclatasvir has a low resistance barrier to several mutations and is only active against few HCV genotypes. To overcome resistance and to develop more broadly effective derivatives of the drug, a detailed evaluation of where and how Daclatasvir binds to NS5A is warranted.

The present work demonstrates at a detailed atomic level how Daclatasvir and similar dimer pharmacophore compounds bind to NS5A. Here, we used state-of-the-art molecular modeling methods combined with the massive computational

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

power of the IBM Blue Gene/Q to search for the binding location and refine the mode of binding of Daclatasvir. Our approach involved 28,871 blind docking simulations of Daclatasvir to 125 dominant NS5A conformations. Our model correlates remarkably well with the available experimental data and will facilitate the design of next generation NS5A inhibitors that would be associated with less resistance and broader activity against various HCV genotypes.

■ ASSOCIATED CONTENT

Supporting Information

Fourteen figures including description of the structural differences between different conformational states for open and closed MD simulations, RMSD of restrained MD simulations, B-factor for restrained simulations, membrane-bound closed conformation, B-factor for membrane bound system, comparison of membrane-bound system to free equilibrated closed structure, blind docking simulation; illustration of binding mode of Daclatasvir within open and monomer conformations; pharmacophore model; pharmacophore-based docking simulations, R30 V mutation, sequence alignment, and full models for open and closed conformations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Kbarakat@ualberta.ca.

Author Contributions

The manuscript was written through contributions of all authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We wish to thank Dr. Kirk E. Jordan and Dr. Constantinos Evangelinos (both from IBM Thomas J. Watson Research Center) for their outstanding assistance in setting up the working environment in the Blue Gene/Q super computer facility to carry out this work. Funding for this research was provided through grants from Canada Excellence Research Chair (CERC) to M.H., Canadian Institutes of Health Research (CIHR) postdoctoral fellowship to K.B., Alberta Innovates-Health Solutions (AIHS) postdoctoral fellowship to K.B. and A.A-M., and Canadian Association of Gastroenterology (CAG).

■ REFERENCES

(1) Jacobson, I. M.; McHutchison, J. G.; Dusheiko, G.; Di Bisceglie, A. M.; Reddy, K. R.; Bzowej, N. H.; Marcellin, P.; Muir, A. J.; Ferenci, P.; Flisiak, R.; George, J.; Rizzetto, M.; Shouval, D.; Sola, R.; Terg, R. A.; Yoshida, E. M.; Adda, N.; Bengtsson, L.; Sankoh, A. J.; Kieffer, T. L.; George, S.; Kauffman, R. S.; Zeuzem, S.; Team, A. S. Telaprevir for previously untreated chronic hepatitis C virus infection. *N. Engl. J. Med.* **2011**, *364*, 2405−2416.

(2) Di Bisceglie, A. M.; McHutchison, J.; Rice, C. M. New therapeutic strategies for hepatitis C. *Hepatology* **2002**, *35*, 224−231.

(3) Abe, H.; Hayes, C. N.; Hiraga, N.; Imamura, M.; Tsuge, M.; Miki, D.; Takahashi, S.; Ochi, H.; Chayama, K. A translational study of resistance emergence using sequential direct-acting antiviral agents for hepatitis C using ultra-deep sequencing. *Am. J. Gastroenterol.* **2013**, *108*, 1464−1472.

(4) Wong, K. A.; Worth, A.; Martin, R.; Svarovskaia, E.; Brainard, D. M.; Lawitz, E.; Miller, M. D.; Mo, H. Characterization of hepatitis C virus resistance from a multiple-dose clinical trial of the novel NS5A inhibitor GS-5885. *Antimicrob. Agents Chemother.* **2013**, *57*, 6333−6340.

(5) Fukuma, T.; Enomoto, N.; Marumo, F.; Sato, C. Mutations in the interferon-sensitivity determining region of hepatitis C virus and transcriptional activity of the nonstructural region 5A protein. *Hepatology* **1998**, *28*, 1147−1153.

(6) Fridell, R. A.; Qiu, D.; Valera, L.; Wang, C.; Rose, R. E.; Gao, M. Distinct functions of NS5A in hepatitis C virus RNA replication uncovered by studies with the NS5A inhibitor BMS-790052. *J. Virol.* **2011**, *85*, 7312−7320.

(7) Love, R. A.; Brodsky, O.; Hickey, M. J.; Wells, P. A.; Cronin, C. N. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *J. Virol.* **2009**, *83*, 4395−4403.

(8) Tellinghuisen, T. L.; Marcotrigiano, J.; Rice, C. M. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature* **2005**, *435*, 374−379.

(9) Gao, M.; Nettles, R. E.; Belema, M.; Snyder, L. B.; Nguyen, V. N.; Fridell, R. A.; Serrano-Wu, M. H.; Langley, D. R.; Sun, J. H.; O'Boyle, D. R., 2nd; Lemm, J. A.; Wang, C.; Knipe, J. O.; Chien, C.; Colonno, R. J.; Grasela, D. M.; Meanwell, N. A.; Hamann, L. G. Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature* **2010**, *465*, 96−100.

(10) Targett-Adams, P.; Graham, E. J.; Middleton, J.; Palmer, A.; Shaw, S. M.; Lavender, H.; Brain, P.; Tran, T. D.; Jones, L. H.; Wakenhut, F.; Stammen, B.; Pryde, D.; Pickford, C.; Westby, M. Small molecules targeting hepatitis C virus-encoded NS5A cause subcellular redistribution of their target: insights into compound modes of action. *J. Virol.* **2011**, *85*, 6353−6368.

(11) Guedj, J.; Dahari, H.; Rong, L.; Sansone, N. D.; Nettles, R. E.; Cotler, S. J.; Layden, T. J.; Uprichard, S. L.; Perelson, A. S. Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 3991−3996.

(12) Amblard, F.; Zhang, H.; Zhou, L.; Shi, J.; Bobeck, D. R.; Nettles, J. H.; Chavre, S.; McBrayer, T. R.; Tharnish, P.; Whitaker, T.; Coats, S. J.; Schinazi, R. F. Synthesis and evaluation of non-dimeric HCV NS5A inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2031−2034.

(13) DeGoey, D. A.; Betebenner, D. A.; Grampovnik, D. J.; Liu, D.; Pratt, J. K.; Tufano, M. D.; He, W.; Krishnan, P.; Pilot-Matias, T. J.; Marsh, K. C.; Molla, A.; Kempf, D. J.; Maring, C. J. Discovery of pyrido[2,3-d]pyrimidine-based inhibitors of HCV NS5A. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3627−3630.

(14) Zhang, H.; Zhou, L.; Amblard, F.; Shi, J.; Bobeck, D. R.; Tao, S.; McBrayer, T. R.; Tharnish, P. M.; Whitaker, T.; Coats, S. J.; Schinazi, R. F. Synthesis and evaluation of novel potent HCV NS5A inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 4864−4868.

(15) Barakat, K. H.; Mane, J. Y.; Tuszynski, J. A. Virtual Screening: An Overview on Methods and Applications. In *Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications*; Liu, L. A., Wei, D., Li, Y., Lei, H., Eds.; Medical Information Science Reference: Hershey, PA, **2011**.

(16) Yuriev, E.; Ramsland, P. A. Latest developments in molecular docking: 2010−2011 in review. *J. Mol. Recognit.* **2013**, *26*, 215−239.

(17) Barakat, K.; Tuszynski, J. Relaxed complex scheme suggests novel inhibitors for the lyase activity of DNA polymerase beta. *J. Mol. Graph. Model.* **2011**, *29*, 702−716.

(18) Jordheim, L. P.; Barakat, K. H.; Heinrich-Balard, L.; Matera, E. L.; Cros-Perrial, E.; Bouledrak, K.; El Sabeh, R.; Perez-Pineiro, R.; Wishart, D. S.; Cohen, R.; Tuszynski, J.; Dumontet, C. Small molecule inhibitors of ERCC1-XPF protein−protein interaction synergize alkylating agents in cancer cells. *Mol. Pharmacol.* **2013**, *84*, 12−24.

(19) Fridell, R. A.; Wang, C.; Sun, J. H.; O'Boyle, D. R., 2nd; Nower, P.; Valera, L.; Qiu, D.; Roberts, S.; Huang, X.; Kienzle, B.; Bifano, M.; Nettles, R. E.; Gao, M. Genotypic and phenotypic analysis of variants resistant to hepatitis C virus nonstructural protein 5A replication complex inhibitor BMS-790052 in humans: In vitro and in vivo correlations. *Hepatology* **2011**, *54*, 1924−1935.

(20) Pawlotsky, J. M. NS5A inhibitors in the treatment of hepatitis C. *J. Hepatol.* **2013**, *59*, 375−382.

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373

(21) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd Calculating structures and free energies of complex molecules: ombining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(22) MOE: Molecular Operating Environment. http://www.chemcomp.com.
(23) Desai, M. C.; Meanwell, N. A. *Successful Strategies for the Discovery of Antiviral Drugs*; Royal Society of Chemistry: U.K., 2013.
(24) Schmitz, U.; Tan, S. L. NS5A–From obscurity to new target for HCV therapy. *Recent Pat. Antiinfect. Drug Discovery* **2008**, *3*, 77−92.
(25) Belda, O.; Targett-Adams, P. Small molecule inhibitors of the hepatitis C virus-encoded NS5A protein. *Virus Res.* **2012**, *170*, 1−14.
(26) Gao, M. Antiviral activity and resistance of HCV NS5A replication complex inhibitors. *Curr. Opin. Virol.* **2013**, *3*, 514−520.
(27) Fridell, R. A.; Qiu, D.; Wang, C.; Valera, L.; Gao, M. Resistance analysis of the hepatitis C virus NS5A inhibitor BMS-790052 in an in vitro replicon system. *Antimicrob. Agents Chemother.* **2010**, *54*, 3641−3650.
(28) Bennett, M. J.; Barakat, K.; Huzil, J. T.; Tuszynski, J.; Schriemer, D. C. Discovery and characterization of the laulimalide-microtubule binding mode by mass shift perturbation mapping. *Chem. Biol.* **2010**, *17*, 725−734.
(29) Kuznetsov, A.; Faustova, I.; Jarv, J. Computational simulation of ligand docking to L-type pyruvate kinase subunit. *Comput. Biol. Chem.* **2013**, *48C*, 40−44.
(30) Mi, R.; Hu, Y. J.; Fan, X. Y.; Ouyang, Y.; Bai, A. M. Exploring the site-selective binding of jatrorrhizine to human serum albumin: Spectroscopic and molecular modeling approaches. *Spectrochim. Acta, Part A* **2014**, *117*, 163−169.
(31) Correa-Basurto, J.; Bello, M.; Rosales-Hernandez, M. C.; Hernandez-Rodriguez, M.; Nicolas-Vazquez, I.; Rojo-Dominguez, A.; Trujillo-Ferrara, J. G.; Miranda, R.; Flores-Sandoval, C. A. QSAR, docking, dynamic simulation and quantum mechanics studies to explore the recognition properties of cholinesterase binding sites. *Chem. Biol. Interact.* **2013**, *209C*, 1−13.
(32) Di Micco, S.; Renga, B.; Carino, A.; D'Auria, M. V.; Zampella, A.; Riccio, R.; Fiorucci, S.; Bifulco, G. Structural insights into Estrogen Related Receptor-beta modulation: 4-methylenesterols from *Theonella swinhoei* sponge as the first example of marine natural antagonists. *Steroids* **2013**, *80C*, 51−63.
(33) Osterberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: Incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* **2002**, *46*, 34−40.
(34) Penin, F.; Brass, V.; Appel, N.; Ramboarina, S.; Montserret, R.; Ficheux, D.; Blum, H. E.; Bartenschlager, R.; Moradpour, D. Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *J. Biol. Chem.* **2004**, *279*, 40835−40843.
(35) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **1999**, *151*, 283−312.
(36) Barakat, K.; Issac, B.; Stepanova, M.; Tuszynski, J. Effects of Temperature on the p53-DNA Binding Interactions and Their Dynamical Behavior: Comparing the Wild Type to the R248Q Mutant. Poster presented at the 17th Canadian Symposium on Theoretical Chemistry, Edmonton, AB, Canada, 2010.
(37) Barakat, K.; Mane, J.; Friesen, D.; Tuszynski, J. Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. *J. Mol. Graph. Model.* **2010**, *28*, 555−568.
(38) Barakat, K. H.; Jordheim, L. P.; Perez-Pineiro, R.; Wishart, D.; Dumontet, C.; Tuszynski, J. A. Virtual screening and biological evaluation of inhibitors targeting the XPA-ERCC1 interaction. *PLoS One* **2012**, *7*, e51329.
(39) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712−725.
(40) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **2007**, *35*, W522−W525.
(41) CHARMM-GUI. http://www.charmm-gui.org/?doc=input/membrane.
(42) Skjevik, A. A.; Madej, B. D.; Walker, R. C.; Teigen, K. LIPID11: A modular framework for lipid simulations using AMBER. *J. Phys. Chem. B* **2012**, *116*, 11124−11136.
(43) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(44) Davies, D. L.; Bouldin, D. W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *1*, 224.
(45) Shao, J.; Tanner, S.; Thompson, N.; Cheatham, T. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* **2007**, 2312−2334.
(46) Sanner, M. F.; Olson, A. J.; Spehner, J. C. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **1996**, *38*, 305−320.
(47) Barakat, K. H.; Huzil, J. T.; Jordan, K. E.; Evangelinos, C.; Houghton, M.; Tuszynski, J. A computational model for overcoming drug resistance using selective dual-inhibitors for aurora kinase A and its T217D variant. *Mol. Pharmacol.* **2013**, *10*, 4572−4589.
(48) Barakat, K. H.; Law, J.; Prunotto, A.; Magee, W. C.; Evans, D. H.; Tyrrell, D. L.; Tuszynski, J.; Houghton, M. Detailed computational study of the active site of the hepatitis C viral RNA polymerase to aid novel drug design. *J. Chem. Inf. Model.* **2013**, *53*, 3031−3043.
(49) Friesen, D. E.; Barakat, K. H.; Semenchenko, V.; Perez-Pineiro, R.; Fenske, B. W.; Mane, J.; Wishart, D. S.; Tuszynski, J. A. Discovery of small molecule inhibitors that interact with gamma-tubulin. *Chem. Biol. Drug Des.* **2012**, *79*, 639−652.
(50) Genheden, S.; Ryde, U. How to obtain statistically converged MM/GBSA results. *J. Comput. Chem.* **2010**, *31*, 837−846.

DOI: 10.1021/ci400631n
J. Chem. Inf. Model. 2015, 55, 362−373