Normal Mode Analysis with Gaussian Network Modeling of Hepatitis C Virus NS5A Protein Characterizes Dimer Interfaces

Aaron Agarunov

Department of Chemistry, Hunter College, 695 Park Avenue, New York, New York 10065, United States

Professor Lei Xie

Department of Computer Science, Hunter College, The City University of New York

Modern treatments for Hepatitis C infections involve various combinations of direct-acting antiviral agents (DAA). These agents inhibit nonstructural proteins encoded by the virus for managing viral replication and assembly. Hepatitis C NS5A protein is targeted by several potent DAAs with great success. However, many of the functions and mechanisms of NS5A and its associated ligands remain elusive. An ensemble of available NS5A domain I structures are discussed and observed. Although NS5A inhibitor pseudosymmetry is thought to complement the dimeric NS5A protein, the dimeric interfaces found among the crystal structures differ. Protein-ligand binding is explored to identify likely binding poses and the activity at a large groove defining an interface termed the *CD cleft*. Normal mode analysis (NMA) is used to predict and characterize structural dynamics by identifying the meaningful deformational motions associated with the protein. Coarse-grained elastic network models then introduce simplifications to capture large-scale conformational dynamics more efficiently. Gaussian network modeling (GNM) of NS5A allows investigation of the protein's great flexibility and likely communication network between its phosphorylated and dephosphorylated states. The role of binding at the CD cleft is explored for understanding requirements for NS5A inhibition by DAA and function by RNA. Low-frequency modes from GNM and ANM are used to characterize atomic residue fluctuations and directions, and suggest a strong role for NS5A dimerization and multimerization for binding.

INTRODUCTION

Hepatitis C is a communicable disease that targets the liver. While acute infection by the hepatitis C virus (HCV) leaves patients mostly asymptomatic, up to 80% of patients develop chronic hepatitis C infection.[1] As of 2015, an estimated 142 million people worldwide are infected with the hepatitis C virus.[2] Long-term HCV-infected patients typically suffer from liver disease, hepatocellular carcinoma, cirrhosis, or experience liver failure.

HCV is a single-stranded, enveloped RNA virus found with one of seven possible genotypes. Each genotype is further described with a series of subtypes, and genotypes 1a and 1b are the most common.[1] The efficacy of therapeutic treatments targeting HCV depends heavily on the patient's genotype.

The HCV RNA genome has 9600 nucleotide bases in a single open reading frame encoding a large precursor protein. The precursor is cleaved to yield structural proteins (core protein, E1, and E2) and nonstructural proteins (NS1, NS2, NS3, NS4A, NS4B, NS5A, NS5B).[3] While there is no direct cure to HCV infection, there are numerous FDA-approved

direct-acting antiviral (DAA) drugs that successfully target HCV nonstructural proteins. HCV NS proteins have critical roles in viral replication and proliferation. Common treatments target these protein products to disrupt HCV replication and assembly.

The viral nonstructural protein 5A (NS5A) is a proline-rich phosphoprotein anchored to the intracellular endoplasmic reticulum (ER) membrane bilayers by its N-terminal amphipathic α-helix. NS5A has three domains with functional roles in genome replication, viral nucleic acid encapsidation, host cell modulation, and virion assembly.[4,5,27] The first domain is dimeric and has a zinc-binding motif. Domain II is unstructured but required for replication, likely due to phosphorylation sites in the disordered region that interact with protein kinases CK2, CK1α, PKA, and Plk1.[6] Domain III is also inherently unstructured but indispensable due to significant roles in assembly with the core proteins.[4,7]
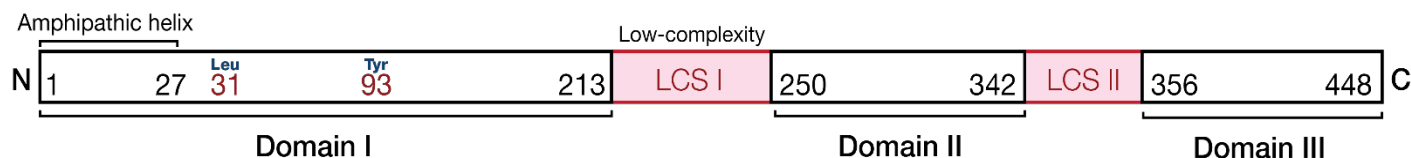

Figure 1. Structure of HCV NS5A domain I in genotype 1a.[14] Accession number NP_751927 has 99% identity with PDB accession code 4CL1 (structure of interest) for defining domain I in genotype 1a. 4CL1 chain A has 177 amino acids and describes residue positions 31 to 207 in the NP_751927 NS5A sequence. LCS denotes low-complexity sequences with polyproline motifs necessary for replication. AHH would denote the amphipathic helix that anchors domain I to the ER membrane. L31 and Y93 are common sites for resistance mutations naturally prevalent in HCV, and are particularly important for studies characterizing changes in DAA potencies when mutated.

NS5A interacts with several cellular proteins and affects other replicon centers upon protein-ligand binding at multiple sites. Its structural flexibility is likely due to the disordered domains I and II. NS5A can be basal-phosphorylated or hyper-phosphorylated, a mechanism facilitated by the protein for cell communication. CK2 phosphorylation, for instance, has been shown to modulate long-range electrostatic interactions between NS5A.[6] Inhibition requires adequate interference with this communication network in addition to binding of the inhibitor, as shown when previous compounds of clinical interest successfully bound HCV 1b Y93H mutants without affecting NS5A function.[5]

NS5A inhibitors are integral components of combination therapies against HCV infection. NS5A DAAs like daclatsvir (DCV) and ledipasvir (LDV) are joined with drugs like sofosbuvir, a nucleotide analog inhibitor of the RNA polymerase NS5B.[8] Daclatasvir is a prototype, first-generation direct-acting antiviral protease inhibitor used in therapy with sofosbuvir or ribavirin and heavily studied for its functions in potent NS5A inhibition. DCV binds tightly to domain I where it reduces NS5A affinity for RNA.[9] DCV has also been shown to disrupt late steps in phosphatidylinositol 4-kinase alpha (PI4KA) gene activation, which is crucial for NS5A phosphorylation.[10] By altering the phosphorylation state, the replication life cycle of HCV is threatened. Daclatasvir has also been shown to have two modes of action, impacting virion assembly and release of infectious particles in addition to disrupting replication.[11] Research focused on first-generation antivirals like daclatasvir has enabled the introduction of more potent drugs in HCV combination therapies, such as the second-generation agent velpatasvir.

Most studies investigate the mutations that confer the most resistance against popular protease inhibitors. Previous research with HCV recombinants has revealed the necessity of NS5A structural sites, noting that the amphipathic α-helix, domain I, low-complexity sequences (LCS), and domain II were universally essential for viral replication in all genotypes.[12] In particular, LCS I and II were found to be proline-rich regions critical for function as their removal left HCV non-functional.[6]

NS5A function varies based on genotype-specific residue variants in domain I. DCV and LDV, which both bind at NS5A domain I, have similar resistance profiles, displaying lower potency in variants with mutations at Tyr93.[13] Substitutions at amino acids Leu31, Gln54, and Tyr93 showed significant compound resistance for inhibition in HCV

genotype 1a.[5] This is in agreement with further studies on first-generation DAAs that expanded on the resistance profiles and suggested weak binding to resistance mutants L31V and Y93H. Single mutations of L31V or Y93H allows up to 28-fold resistance of NS5A to daclatasvir.[14] Genotype 1a, a subtype comprising 31% of 83.4 million genotype 1 cases spread across the world,[15] is particularly susceptible to single amino acid changes near NS5A's N-terminal region.[16] L31V/M, Y93H/N, and Q30E/H/R are resistance mutations for genotype 1a, and L31V/F, P32L, and Y93H/N for genotype 1b,[16] all occurring within 20 Å of the N-terminus. Combining NS5A DAAs with NS3 protease or NS5B polymerase inhibitors significantly reduces the risk associated with developing these resistant mutations, leading to the principal use of combination therapy in HCV treatment.
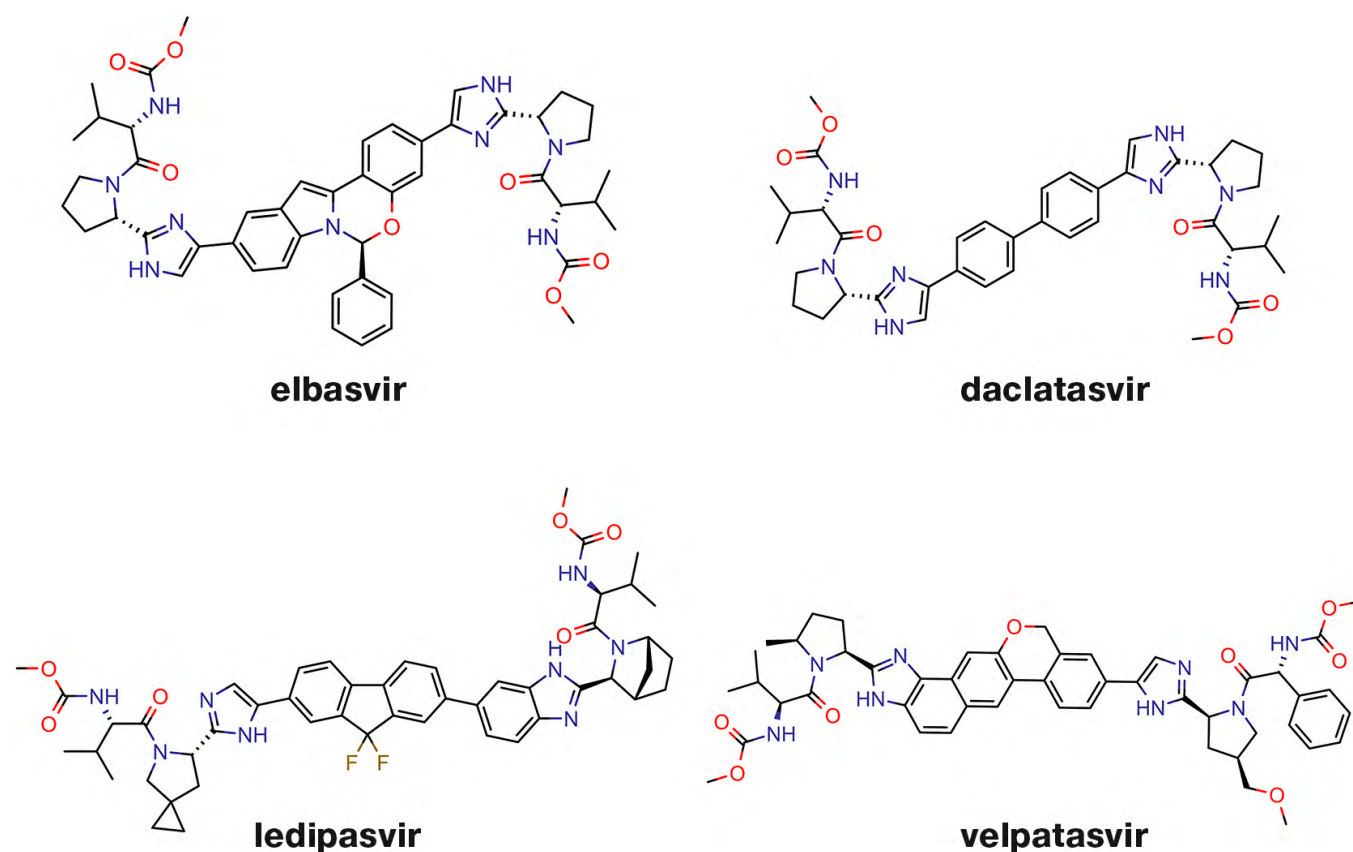

Figure 2. Four NS5A inhibitors currently approved for clinical use in HCV treatments. These drugs are common components of HCV combination therapies, typically seen in conjunction with NS5B inhibitors like sofosbuvir and dasabuvir. The velpatasvir-sofosbuvir combination has resulted in cure rates up to 90% for patients with genotype 1-6.

From recent crystallography and modeling studies, it is known that NS5A is dimeric in nature. Highly potent NS5A inhibitors also display symmetry or pseudosymmetry to correspond with the protein dimer (Figure 2). However, the dimeric interface and binding sites are not clear. Binding of daclatasvir was *not* shown to affect NS5A dimerization,[10] and instead shown to likely occur at domain I near the zinc-binding motif and away from phosphorylation sites.[17] Evidence for the formation of multimeric complexes of NS5A implied the high potency of inhibitors is facilitated by conformational changes to adjacent multimers that take place upon protein-ligand binding.[5] Conformational receptor modeling studies have demonstrated how DCV can have several binding modes and interfere with protein-protein interactions at the relevant membrane interfaces,[14] and potentially bind to two different conformations of the protein.[16] This conformation-dependent communication network is crucial for inhibition replication and assembly.

Normal mode analysis.

Normal mode analysis (NMA) is a harmonic simulation technique that can predict and characterize large-scale motions of biomolecules near their native state. NMA requires a set of coordinates (e.g. as stored in a PDB file) and a force field for evaluation. NMA is calculated in a vacuum, and assumes that all atoms move in phase, where the classic NMA model is performed on all atoms.[18]

In stage (1), the conformational potential energy is minimized in Cartesian coordinate space. Common force field implementations (e.g. CHARMM, AMBER) used for MD simulations are employed for energy minimization. In stage (2), the second derivatives of the potential energy function with respect to the weighted Cartesian coordinates are stored in a square matrix. In stage (3), this Hessian matrix is weighted and diagonalized to yield the eigenvalues and the eigenvectors.[18]

$$\mathcal{H} = \begin{bmatrix} H_{11} & \cdots & H_{1N} \\ \vdots & & \vdots \\ H_{N1} & \cdots & H_{NN} \end{bmatrix} \qquad H_{ij} = \begin{bmatrix} \dfrac{\partial^2 V_{ij}}{\partial x_i\, \partial x_j} & \dfrac{\partial^2 V_{ij}}{\partial x_i\, \partial y_j} & \dfrac{\partial^2 V_{ij}}{\partial x_i\, \partial z_j} \\ \dfrac{\partial^2 V_{ij}}{\partial y_i\, \partial x_j} & \dfrac{\partial^2 V_{ij}}{\partial y_i\, \partial y_j} & \dfrac{\partial^2 V_{ij}}{\partial y_i\, \partial z_j} \\ \dfrac{\partial^2 V_{ij}}{\partial z_i\, \partial x_j} & \dfrac{\partial^2 V_{ij}}{\partial z_i\, \partial y_j} & \dfrac{\partial^2 V_{ij}}{\partial z_i\, \partial z_j} \end{bmatrix} \qquad \Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j,\ R_{ij} \leq R_c \\ 0 & \text{if } i \neq j,\ R_{ij} > R_c \\ -\sum\limits_{i,\ i \neq j}^{N} \Gamma_{ij} & \text{if } i = j \end{cases}$$

Hessian matrix (left). The Hessian ($\mathcal{H}$) describes the force constant, where each element is described by a sub-matrix (middle). Each element H_{ij} of the Hessian is a square matrix ($3N \times 3N$) with second-order partial derivatives of the potential function V_{ij}, where i and j are nodes (atoms). In the anisotropic network model, each node is a residue's C^α atom and the Hessian describes the orientation of the nodes. The total *potential* describes the interactions between nodes.[21] **Kirchhoff matrix** (right). The Kirchhoff (Γ) is a symmetric matrix that enumerates all subsets of nodes i and j. R_{ij} is the distance between the ith and jth C^α atoms and R_c is the cut-off distance describing the range of interaction between non-bonded C^α to define connectivity. In the Gaussian network model, the inverse of the Kirchhoff matrix is decomposed to yield the modes since $|\Gamma| = 0$.[22]

An eigenvalue of the Hessian provides a frequency that characterizes an eigenvector. The eigenvectors of the Hessian are the independent harmonic patterns of motion, the *normal modes*. A normal mode of an oscillating system is described by uniform sinusoidal motion at a fixed frequency. In practice, the normal modes characterized with the lowest frequencies are typically robust and functionally relevant (i.e. contributing most to motions and changes in the structure), given that the associated R_c is not too large or small.[18,19]

Elastic network model.

Elastic network models (ENM) employ simplifications of NMA for more efficient computational simulations. ENMs require unique parameters; primarily the cut-off distance for spatial interactions, R_c, and the uniform force constant for elastic bonds, γ.[20] Reasonable cut-off distances are typically between 7 and 8 Å. Energy minimization is no longer needed because the minimized distances between elastic connections are reflected in the potential energy function of the network.

Coarse-grained models of elastic networks consider only a subset of atoms. Despite the elasticity, the model may still retain the meaningful, slowest normal modes, as shown by Tirion.[23] ENMs are built using harmonic potentials at γ describing interactions between a protein's alpha carbons within R_c.[24]

Two common coarse-grained implementations of these elastic spring networks are the widely used Gaussian network model (GNM) and the anisotropic network model (ANM). These ENMs employ NMA with less degrees of freedom and less resources to predict large-scale conformational dynamics in a much smaller timeframe. This approach is made possible by a coarse-grained implementation such as denoting nodes (residues) by their C^α atoms. The energy potential

in such a coarse-grained ENM that describes how alpha carbons interact is described, where r_{ij} is the instantaneous distance between C^α atoms i and j, and r_{ij}^0 is the distance at equilibrium (natively),

$$V_{ENM} = \frac{\gamma}{2}\left[\sum_{i,j}^{N}(r_{ij} - r_{ij}^0)^2\right]$$

Gaussian network model.

The Gaussian network model is a unidimensional model proposed by Bahar and Erman et al.[22] GNMs are built at the residue level, describing each residue uniformly with Gaussian-distributed fluctuations about their mean positions. Only C^α atoms are considered, and interactions between these junctions within R_c behave like linear springs.[22]

GNMs make use of the Kirchhoff (Laplacian) matrix to describe the network topology by observing the connectivity of inter-atomic contacts between the nodes. GNMs assume the protein is near its equilibrium state when evaluating the mean squared displacement for residue fluctuations. The Kirchhoff matrix is defined by a Heaviside step function. If nodes i and j are within the cut-off distance, the ijth element in Γ is -1; if outside the cut-off distance, the ijth element is 0. Diagonal elements of this symmetric matrix are defined where $i = j$, and so the iith element reflects the connectivity of residue i.[22] Normal mode decomposition is accomplished by diagonalization and inversion of the Kirchhoff matrix.[20]

GNMs are coarse-grained, low-resolution models that identify residues by their alpha carbons to build the Kirchhoff matrix. The R_c distance for defining bonded or non-bonded interactions between residue pairs can be much smaller in GNMs than in all-atom models. To prevent yielding additional trivial eigenvalues, 8 Å was used for all calculations (by default, ProDy uses 10 Å and Bio3D uses 8 Å).

When considering a single spring between nodes (residues) i and j, the potential energy function can be described, where ΔR_i is the distance between the equilibrium and instantaneous position vectors of residue node i (and similarly ΔR_j for node j), and Γ_{ij} is the ijth element of the Kirchhoff matrix.

$$V_{GNM} = \frac{\gamma}{2}\left[\sum_{i,j}^{N}\Delta R_i \Gamma_{ij} \Delta R_j\right]$$

Modes are obtained following eigenvalue decomposition of Γ using the non-zero frequencies in the inverse of the Kirchoff matrix, Γ^{-1}. The slowest modes contribute the most to expected fluctuations, potentially revealing mechanisms of cooperative motions. Minima might reveal residues that contribute to correlated motions, such as a site for hinge-bending.[24] The GNM partitions can be built, and the mean-square fluctuations and cross-correlations (below) can be produced by constructing the covariance matrix $\frac{k_BT}{\gamma}\Gamma^{-1}$, where k_B is the Boltzmann constant and T is the temperature.

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_BT}{\gamma}(\Gamma^{-1})_{ij}$$

GNM and the anisotropic network model (ANM).

GNMs were extended to the anisotropic network model by Atilgan and Doruker et al. The anisotropic network model depends on the scalar product of the fluctuation vectors[24] whereas the GNM calculates independent fluctuation vectors from the X, Y and Z components of the fluctuations $\Delta \mathbf{R}_i$ and $\Delta \mathbf{R}_j$ of residues i and j.[20,24] The potential function for ANMs uses scalars calculated from their components yielding anisotropic fluctuations when considering the Hessian sub-matrixes wherein partial derivatives are taken with respect to the X, Y, and Z axes. When the distances (springs) between

residues change in a GNM, direction is immediately accounted for by taking weighted averages for the potential function, and thus directional effects (changes in the vector) are ultimately lost. In ANMs, the scalar magnitude of the instantaneous vectors is taken in each calculation and fluctuations are *not* isotropic. Thus, when GNM calculations account for orientation independently, the fluctuations between C^α atoms are isotropic, whereas ANM calculations based on all distances retain directional preferences.[5] Anisotropy denotes the different properties observed based on the directional axes, and is useful in spatial motions when conformational structures are available. GNMs predict the *magnitudes* induced by deformation vectors describing fluctuations with each mode; in practice, this allows GNM models an advantage in describing describe motion distributions and implications of distinct residues more accurately.[24]

METHODS AND RESULTS

Crystal structures and recent studies of HCV NS5A have revealed that domain I exists as a dimer.[27] Additionally, NS5A is thought to exist in higher polymeric structures that associate with an elaborate communication network between HCV replication centers when binding highly symmetric ligands.[5] Phosphorylation and subsequent conformational changes are thought to be highly associated with the communication network for successful inhibition. In total, there have been four crystal structures of HCV NS5A domain I (NS5A-D1) elucidated on RCSB PDB.

In 2005, Tellinghuisen et al. described NS5A's three domains, characterized the N-terminal domain (NS5A-D1) by crystallization (PDB 1ZH1), and noted a single zinc-binding motif per protein coordinated by four cysteine residues.[17,25] The study proposed the charge distribution in the molecule could lead to the deep claw-like groove between the monomers and force the arms to extend out from the groove.

In 2009, Love et al. discovered a novel dimeric form of HCV NS5A domain I from HCV genotype 1b and characterized two NS5A-D1 crystal structures. Love et al. grew the crystals in the presence of two different detergents, FOS-choline-9 (PDB 3FQM) and a maltopyranoside (PDB 3FQQ).[26] Despite the difference in methodology, alignment of 3FQM and 3FQQ monomers via Bio3D yields a root-mean-square deviation of atomic positions of 0. 3FQQ was chosen for analysis and visualization.

Additionally, there are 2 crystal structures of peptides from NS5A protein in complex with the c-Src and Fyn SH3 domain (PDB 4QT7, 3UA7), and 4 NMR structures of NS5A's membrane anchor domain (1-31, N-terminal alpha helices) (PDB 1R7{D, E, F, G}). SH3 binding was found to impose transient structural changes to NS5A-D2 and D3.[6] The C-terminal region of NS5A is largely unstructured but highly flexible. An NMR study found multiple peptide regions in D2 and D3 with helical propensities and binding partners, and functionally relevant beta-turns near the C-terminal end that interact with NS5B polymerase.[6] Domain III is known to have phosphorylation sites that promote long-range interactions. There are no particular crystal structures available of the C-terminal domains in either state of phosphorylation, or any structures of NS5A-D1 bound to a ligand. However, the structure of the membrane anchor domain (1R7E, minimized average structure of 51 structures in 1R7D) was appended to precede residues when visualizing latter structures (PDB 4CL1).

Successful NS5A inhibitors targeting several genotypes have been large (> 750 Da), aromatic, and have had dimeric symmetry or pseudosymmetry to accommodate domain I. The interfaces observed in the X-ray structures differ meaningfully. Many studies of potent inhibitors towards NS5A-D1 have revealed symmetric binding modes, although asymmetric binding modes at poly-proline regions have also been demonstrated by docking prior to the most recent crystal structures.[14]
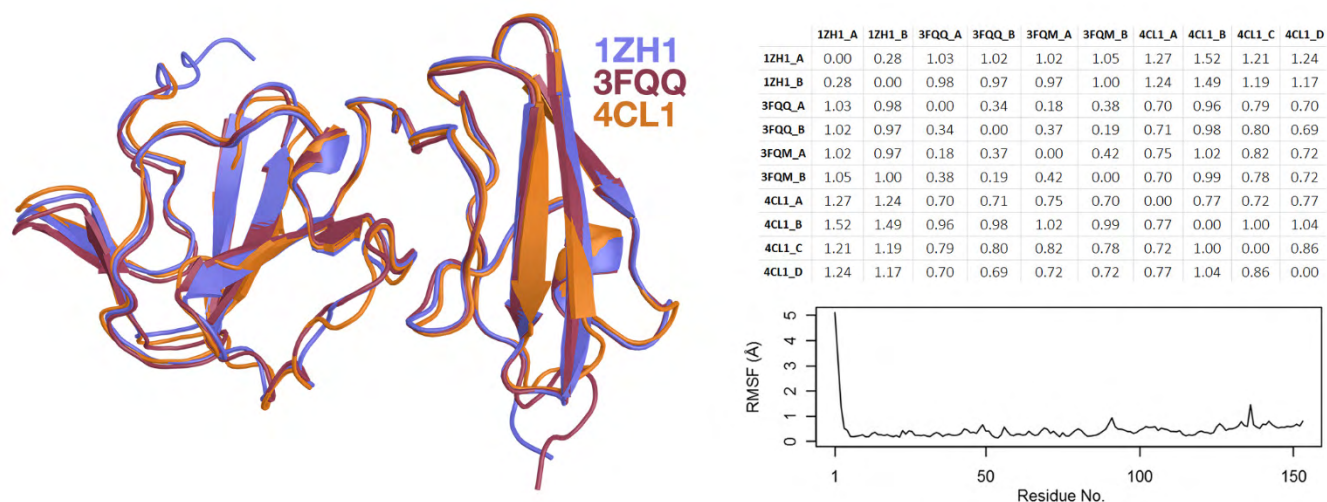
	1ZH1_A	1ZH1_B	3FQQ_A	3FQQ_B	3FQM_A	3FQM_B	4CL1_A	4CL1_B	4CL1_C	4CL1_D
1ZH1_A	0.00	0.28	1.03	1.02	1.02	1.05	1.27	1.52	1.21	1.24
1ZH1_B	0.28	0.00	0.98	0.97	0.97	1.00	1.24	1.49	1.19	1.17
3FQQ_A	1.03	0.98	0.00	0.34	0.18	0.38	0.70	0.96	0.79	0.70
3FQQ_B	1.02	0.97	0.34	0.00	0.37	0.19	0.71	0.98	0.80	0.69
3FQM_A	1.02	0.97	0.18	0.37	0.00	0.42	0.75	1.02	0.82	0.72
3FQM_B	1.05	1.00	0.38	0.19	0.42	0.00	0.70	0.99	0.78	0.72
4CL1_A	1.27	1.24	0.70	0.71	0.75	0.70	0.00	0.77	0.72	0.77
4CL1_B	1.52	1.49	0.96	0.98	1.02	0.99	0.77	0.00	1.00	1.04
4CL1_C	1.21	1.19	0.79	0.80	0.82	0.78	0.72	1.00	0.00	0.86
4CL1_D	1.24	1.17	0.70	0.69	0.72	0.72	0.77	1.04	0.86	0.00

Figure 3. Rough PDB alignment and superimposition of known NS5A-D1 structures. The root-mean-square deviation of atomic positions (RMSD) values are tabulated, and there is very high sequence identity between the structures. Root mean-square fluctuations are also visualized (RMSF). Gaps were accounted for. Variance near the first few residues is due to different lengths of the structures for characterizing NS5A-D1 following the AHH. The structures are inherently similar, so the dimeric interfaces and interacting residues will be important. Sequence alignment was done via the Bio3D R library, with visualization in PyMol.

The initial structure by, Tellinghuisen et al. (1ZH1) had a major groove formed under the point of contact, with long, acidic arms extending outwards. The following structure study by Love et al. faced difficulties, as the authors were not able to resolve the oligomeric state of NS5A unambiguously, instead producing likely oligomerization models.[26, 28] However, the monomers were almost identical (Figure 3). They obtained two non-overlapping dimerization interfaces on opposite faces, implying NS5A-D1 polymerization at physiological conditions. Of note, the C-termini face outwards in a fashion that could potentially situate the unstructured domains and accommodate this polymerization. The model is consistent with the observation that NS5A inhibitors bind NS5A directly[9] and subsequently block NS5A oligomer formation and endogenous RNA binding.[28] The interface observed in PDB accession code 3FQQ was used for the study.

Further crystallography studies faced similar models and also resolved crystal structures with two novel dimerization interfaces (Figure 4). In 2014, Lambert et al. determined the crystal structure of genotype 1a NS5A-D1. The final structure was composed of 4 monomers (A, B, C, D) arranged into T-shaped asymmetric units, suggesting D1 monomers associate in multimer assemblies (PDB 4CL1).[27] The PDB entry has four monomers, two dimeric interfaces, one zinc ion per monomer in their respective motifs, and stabilizing sulfate ions derived from the crystallization. Monomers A and B associate in a head-to-tail dimer, whereas monomers C and D form a head-to-head dimer. Lambert et al. aligned and superimposed Love's 3FQQ (1b) structure with Lambert's 4CL1 structure (1a). They discovered that the 4CL1 AB dimer associates identically to Love's 3FQQ dimer, but in an antiparallel fashion. Lambert et al.'s model provides further support for the generation of higher order oligomeric structures that could extend towards the membranous web.[27]

The existence of bound sulfate ions in the structure derived from the crystallization buffer indicate a possible RNA binding site and a potential function in preventing NS5A hyper-phosphorylation due to the local region.[27] Studies suggest that NS5A-D1, LCS I and D2 bind RNA to promote dimerization.[29] While both Lambert et al. and Love et al. depict possible sites that could accommodate viral RNA, the binding site is unclear. When Lambert et al.'s monomers (PDB 4CL1) are structurally superimposed with the monomers in the 2005 study, the RNA-binding groove proposed by Tellinghuisen et al. is sufficiently aligned. This sulfate ion suggests a stabilizing role and could be accommodated by the large binding groove observed in the CD dimer.
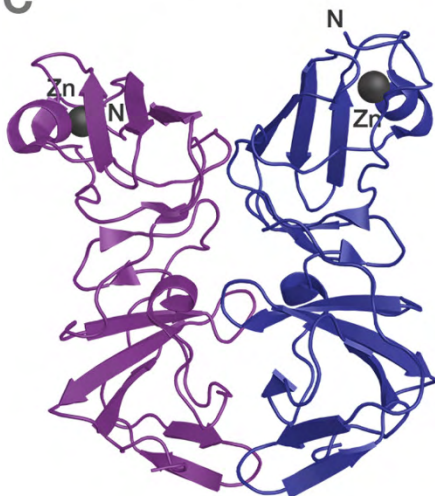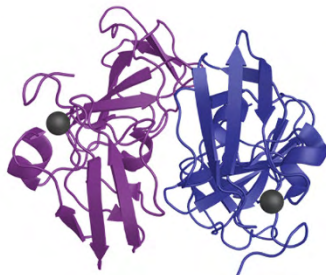
4CL1 Lambert et al.

A
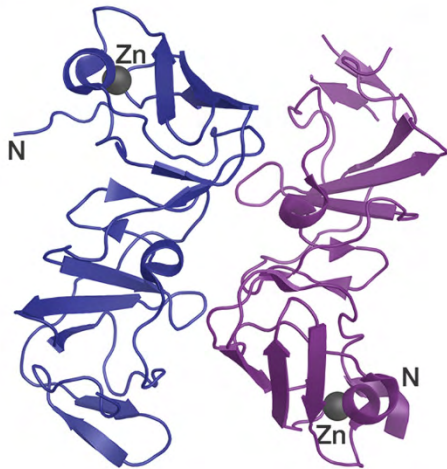
B
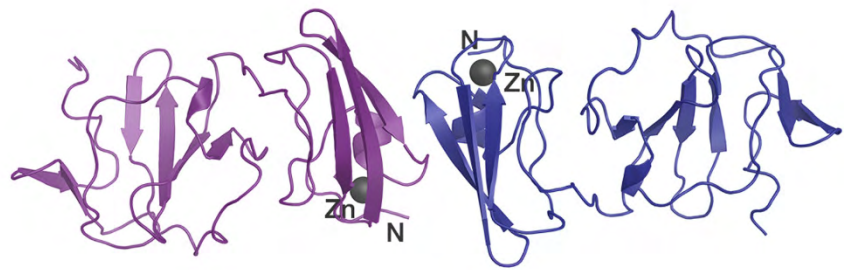
3FQQ Love et al.
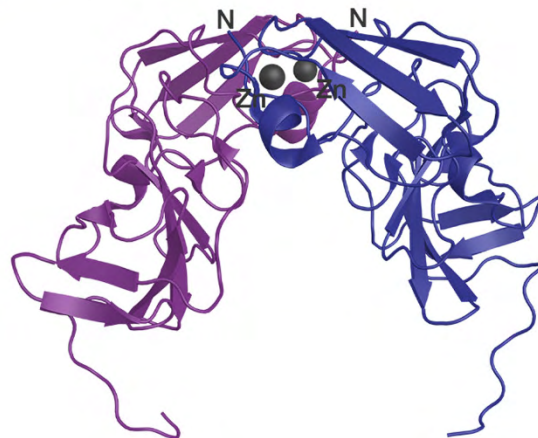
C

D

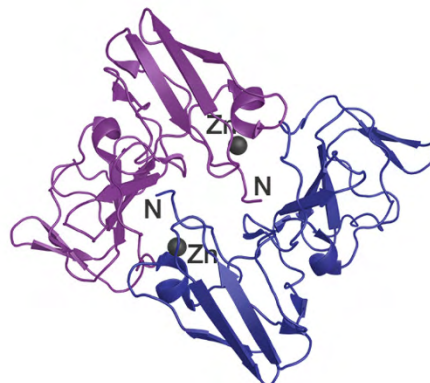**1ZH1** Tellinghuisen et al.

E

F

Figure 4. All X-ray crystal structures elucidated from studies on HCV NS5A genotypes 1a and 1b. The zinc ion at the zinc-binding motif is shown in grey, and the N-terminal end is labeled in each structure. Monomers are shown in purple and blue. **(A)** PDB 4CL1 monomers A and B form a head-to-tail dimer, much like the structure seen in (C) but in an antiparallel fashion. **(B)** 4CL1 monomers C and D form a head-to-head dimer. **(C)** PDB 3FQQ (maltopyranoside detergent) dimer. **(D)** Alternative top-down view of (C), rotated 90 degrees. **(E)** PDB 1ZH1 dimer with a claw-like shape, from the first X-ray study of NS5A-D1 by Tellinghuisen et al. **(F)** Alternative top-down view of (E), rotated 90 degrees.

Poses at the CD cleft.

The CD cleft was shown to accommodate daclatasvir in a favorable binding mode via docking.[27] Y93H and E62H are located about 20 Å from the cleft, but are both about 6.6 Å from R56, suggesting a possible communication network through this residue (via PyMol). Lambert et al. note that other binding modes may be more beneficial to complement the symmetry seen in prior dimers.[27] The CD dimer has less residue interactions at the surface, namely between residues 74-78 and 83-84, whereas the AB dimer interface has contact surfaces between residues 97-99, 112-115, 149, and 160-161 of the monomers.[27] In addition, Lambert et al. noted that both dimers exhibit packing in similar fashions. The head-to-head CD dimer with twofold symmetry was particularly interesting due to close proximity between the N-terminal domains and zinc-binding motifs. Simple docking procedures onto this dimer were performed to determine how current NS5A inhibitors could accommodate this open conformation.
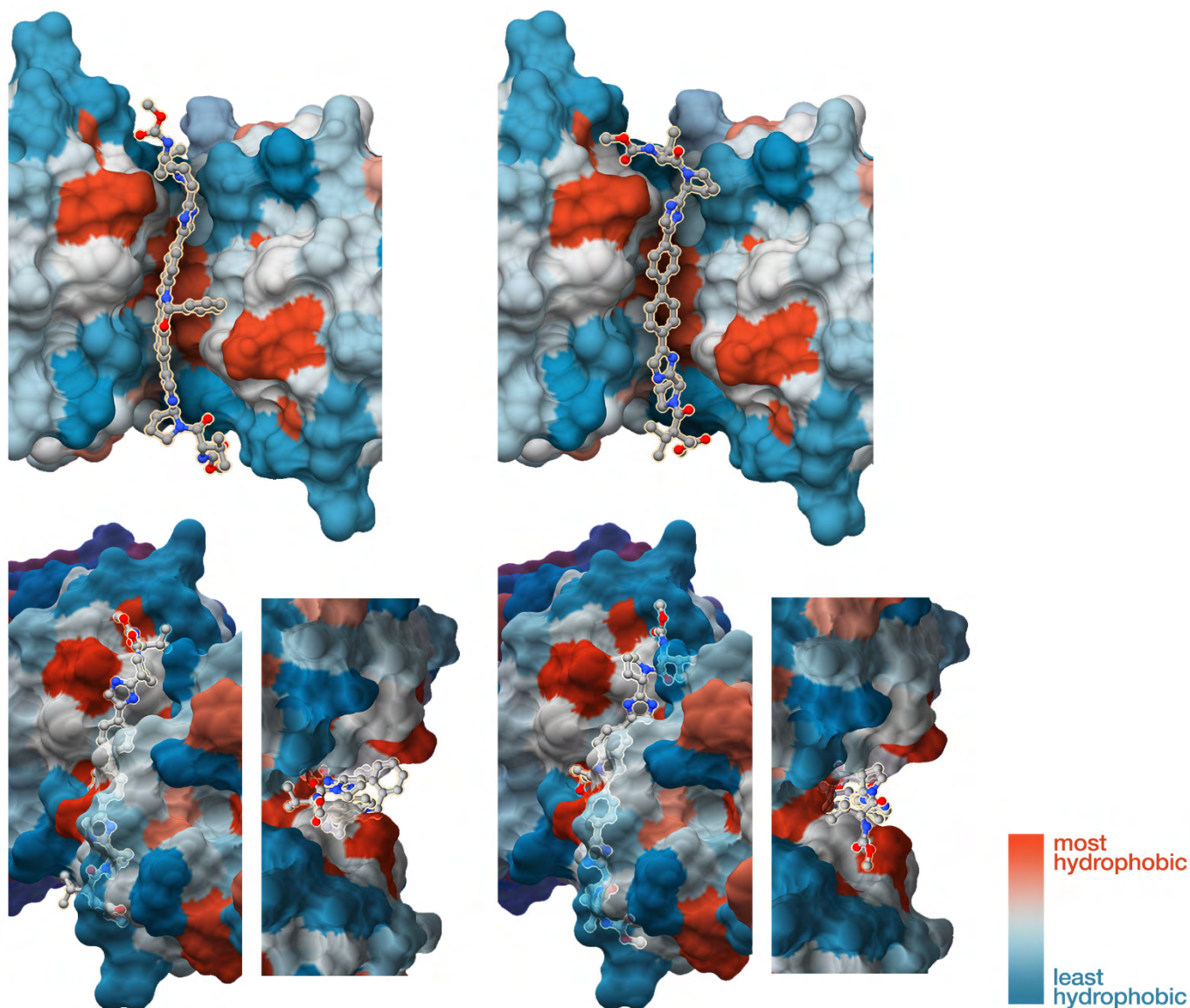
Figure 5. Poses of **daclatasvir** (left) and **elbasvir** (right) docked onto the CD dimer of HCV NS5A-D1 (PDB 4CL1). The surface of NS5A-D1 is shown, colored by relative hydrophobicity using the Kyte-Doolittle scale. Cyan indicates high hydrophobicity and orange indicates low hydrophobicity. These inhibitors were docked with the SwissDock web service; The modes here represented the highest ranked clusters (by SwissDock FullFitness) for both drugs.

Chimera Dock Prep was used to add charges to residues and ions (zinc, sulfate), add hydrogens, and complete side chains via the Dunbrack rotamer library where necessary.[30] SwissDock was used to rigidly dock elbasvir and daclatasvir to the CD dimer.[31] The most meaningful ΔG for the top-ranking cluster for daclatasvir was visualized (-8.9397 kcal/mol, although the FullFitness, the SwissDock metric, was ranked second). A binding mode identical to that observed by Lambert et al. was found (Figure 5). Elbasvir, an FDA-approved NS5A inhibiting agent used with NS3/4A protease inhibitors, was also docked in the same manner to the CD dimer via SwissDock because of high structural similarity. Its top-ranked cluster displayed a similar binding mode.
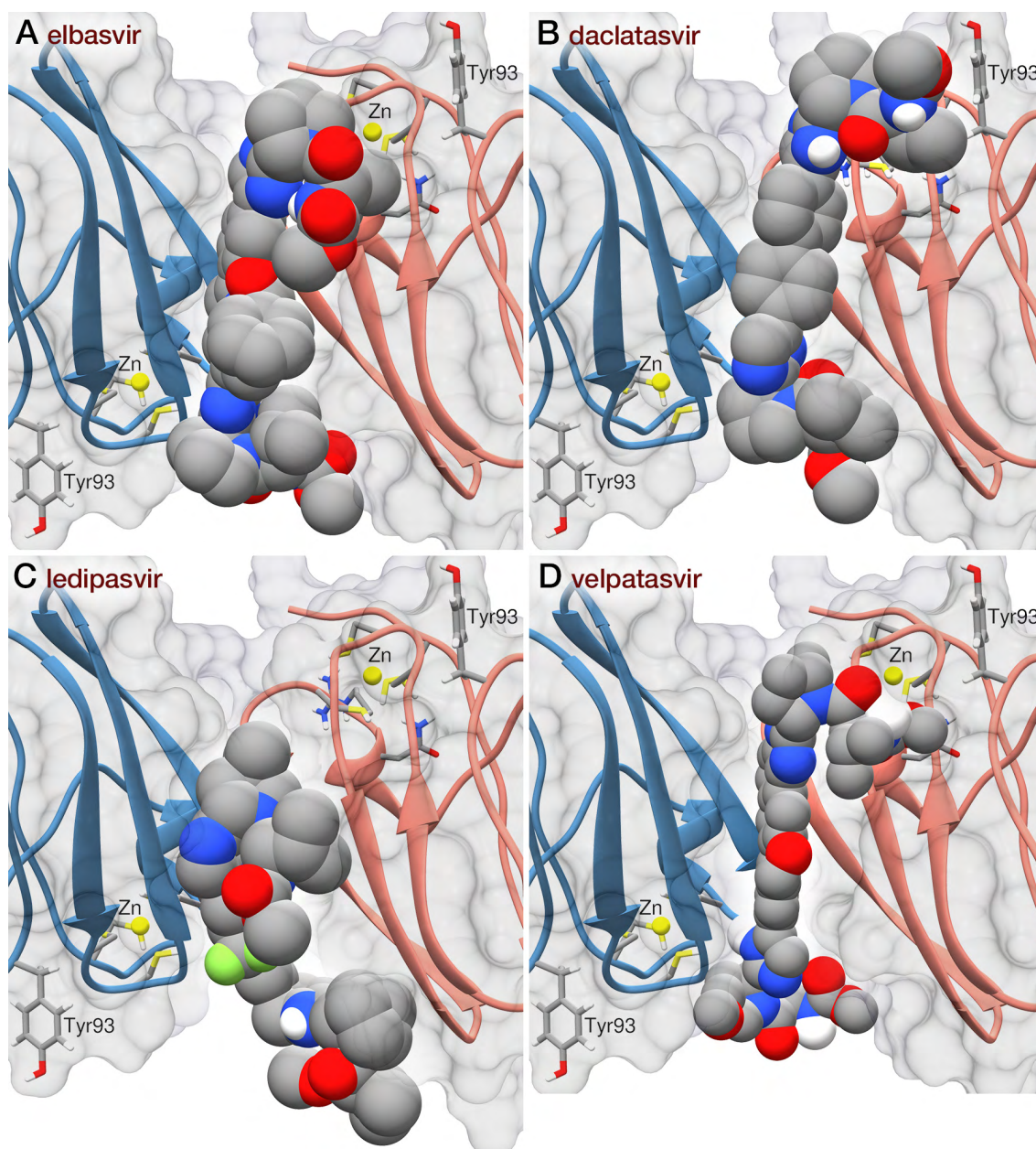
Figure 6A-D. The best-scoring poses for each of **(A)** elbasvir, **(B)** daclatasvir, **(C)** ledipasvir, and **(D)** velpatasvir docked onto the CD dimer of HCV NS5A-D1 (PDB 4CL1). Nine binding modes were predicted for each ligand after restricting the search volume to one-third of the total surface volume, centered on the cleft. All compounds preferred this side of the cleft. The common residue for resistance mutations, Y93, is labeled. The zinc-binding motif is also visible. Important residues in A-D are shown as sticks, and the ligand is shown as spheres. The topological surface is also provided. Visualized in UCSF Chimera.
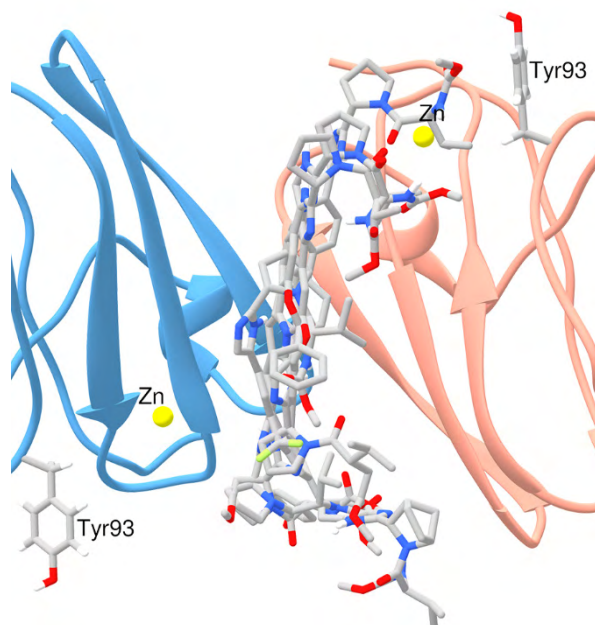
Figure 6E. A superimposition of all compounds in the cleft is presented. Ligands are shown as sticks. Monomers are blue and tan. Zinc ion and Y93 are labeled. Visualized in UCSF Chimera.

Clinically useful inhibitors (Figure 2) were docked with AutoDock Vina to explore the role of the CD cleft.[32] Correct 3-D structures were obtained for daclatasvir, elbasvir, ledipasvir, and velpatasvir, or otherwise built from isomeric SMILES and adjusted accordingly when necessary. Locally, AutoDock Vina was similarly prepared with Dock Prep and given a search volume one-third of the total, centered on the CD cleft (Figure 6). The purpose was to understand the binding preference in the CD cleft. While one side of the cleft has a much larger groove, the opposite face is situated closely to potentially coordinate with very hydrophobic residues situated at the surface, namely M72, I74 and V75. The chosen ligands rarely preferred the opposite face, and the top-ranked binding poses were always in the large groove, which is also the likely binding site for RNA given the presence of the sulfate ion after crystallization. Note the slightly asymmetric modes as seen in the ledipasvir pose, which is not entirely uncommon especially when using the Love et al. *back-to-back* dimer explored previously.[33] Importantly, Q54 and L31 were absent, but Y93 was present (labeled), indicating its paramount interactivity considering the compatibility of the binding modes.

Elastic network models.

Cutoff values were 8 Å for GNM modeling and 15 Å for ANM modeling. Theoretical B-factors were calculated using elastic network models and compared to crystallographic experimental data contained in the PDB as isotropic B-factor values for atom records. GNM severely overestimated fluctuations for residues, especially around residues 40-45, 80-83, 90-93, 140 and 156. A slight trend was visible but the data largely displayed inaccurate prediction of thermal motions.[34] Adjusting residue numbers for the gaps (i.e. considering AHH comprises the first 30 residues), we can see that the motions are associated with charged or polar residues thought to be interacting upon ligand-binding, and other residues located in front of common resistance mutations that participate in NS5A's coordination network (Y93). This might be due to residue absence or mobility considering the nature of HCV. FirstGlance in Jmol revealed each chain had about 22 missing residues including many charged amino acids, in addition to a poor resolution, which might explain the overestimation by GNM. Previous studies have shown that GNM calculations produce high correlation coefficients of

experimental B-factors given that contacts between neighboring molecules are accounted for. Considering the unaccounted existence of higher order multimeric NS5A structures, this is likely the cause for a low correlation coefficient. In comparison, B-factors for Love et al.'s structure (deposited as 3FQQ) displayed an identical coefficient (0.47) following overlap with experimental and theoretical B-factors, and GNM was still prone to overestimation for this protein. Of note, the resolution for 3FQQ was significantly better (2.20 Å compared to 3.50 Å), so it is likely inherent to the structure.
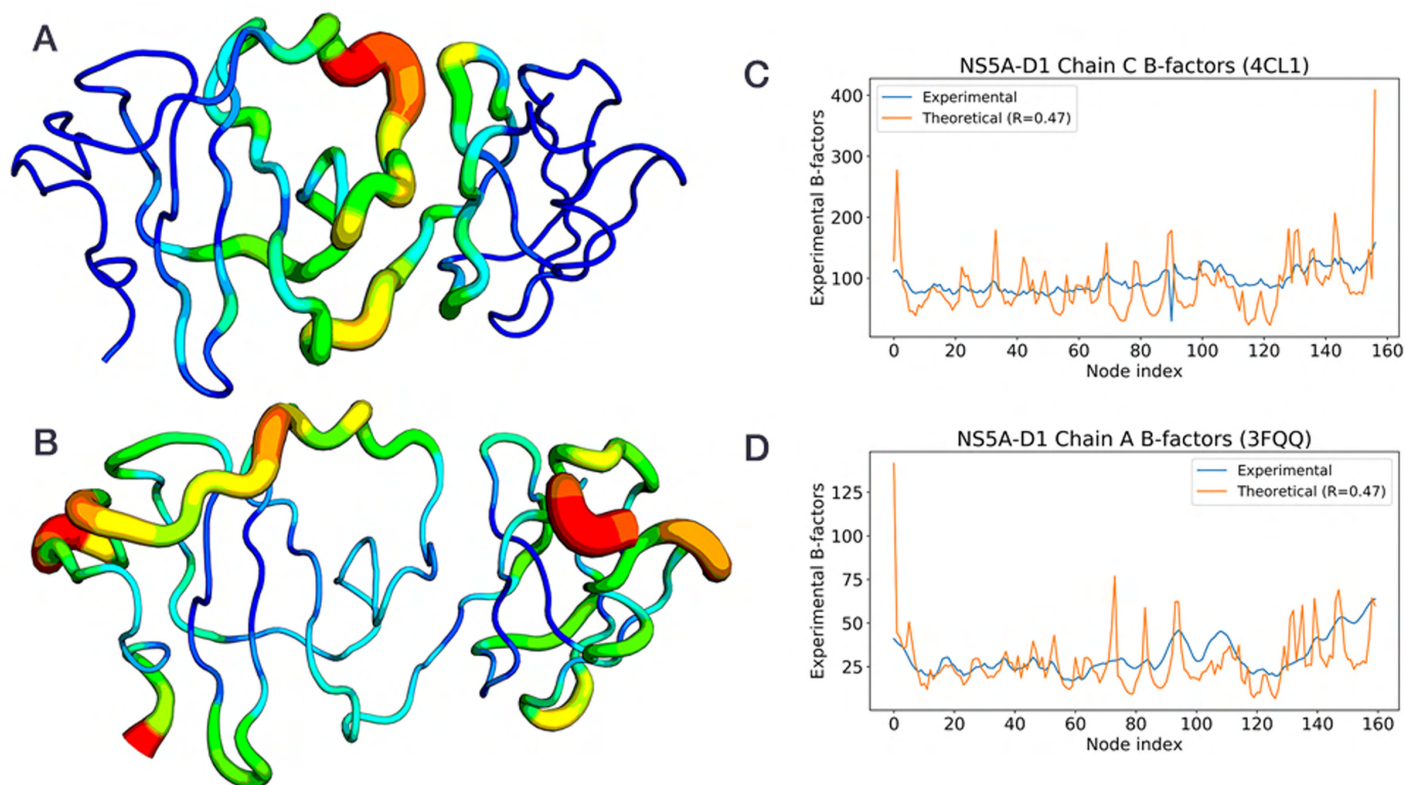
Figure 7. (A) Deformations calculated for NS5A-D1 by NMA on PDB 4CL1. Deformations are at the sites of relevant residues participating in the coordination network with NS5A if the CD cleft is the desired binding site. **(B)** Residue fluctuations. Importantly, the model predicts fluctuations at the dimer site that accommodates towards CD interface more than the AB interface. **(C)** Experimental and theoretical B-factors do not correspond for 4CL1 monomer A. GNM B-factors have been proven to show more accurate overlaps of data in previous studies than ANM modeling, but neither provided a meaningful coefficient with overlaps. **(D)** GNM with experimental and theoretical B-factors for 3FQQ. Deformations and fluctuations calculated in R Bio3D. B-factors analyzed with python ProDy. Visuals in PyMol.

Visualizing the residue fluctuations suggested by this model indicates flexibility at the C-terminal and N-terminal regions. This describes the anchoring of the amphipathic α-helix from the N-terminal region and the positioning of unstructured domains at the C-terminal region, where phosphorylation is also thought to occur and increase based on inhibitor activity. Deformations centered to the structure are at sites of important interacting residue Y93 (Figure 7). While the 3FQQ structural monomers included Y93, L61, and Q54, the Lambert et al. structure (4CL1) only retained Y93 at a site where fluctuations are most evident. Nonetheless, GNM with PDB 3FQQ yielded similar residue fluctuations (not shown).
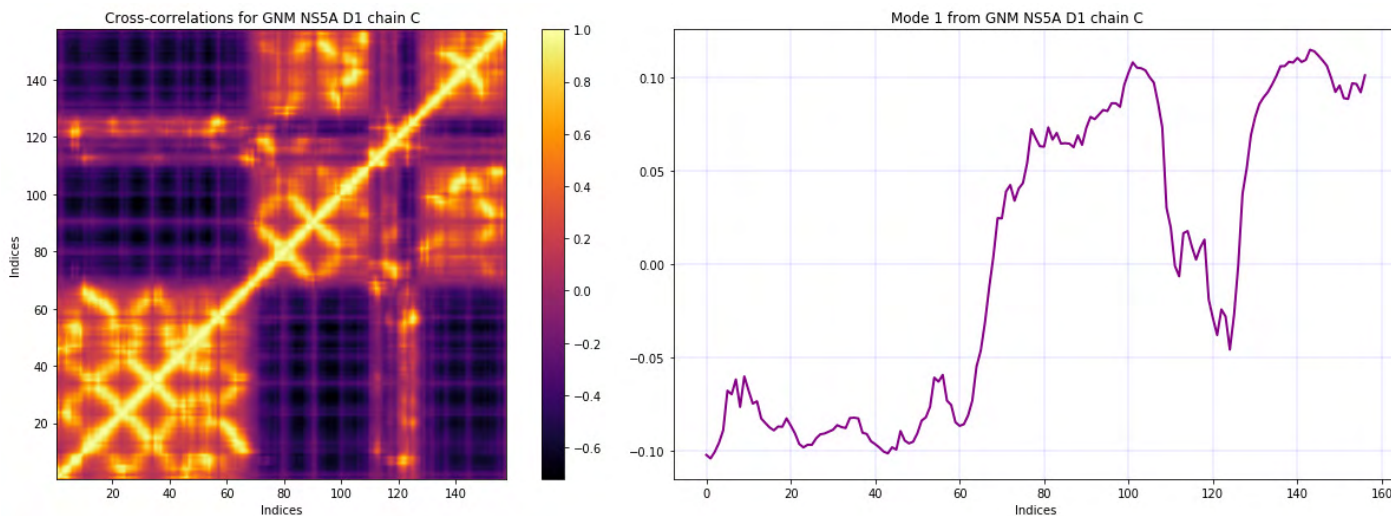
Figure 8. GNM cross-correlations between residue fluctuations (left). Perfect concerted motions and colored dark yellow for visibility, and anti-correlated motions are colored black. The shape of the slowest, most meaningful mode as characterized by its eigenvectors (right). Trivial modes are not retained from calculations, so this is the first mode.

The shape of the most meaningful, lowest-frequency mode as calculated by GNM is shown (Figure 8). Additionally, cross-correlations of residue pairs between all modes was computed. There are strong correlations near the binding site residues, particularly near the dimeric interfaces for AB and CD dimers. The most frequently observed mutations observed in genotype 1a include Y93 and several residues following the AHH, namely positions 28 and 30-31.[35] These positions are not reflected in 4CL1. In addition, common mutation sites in genotype 1b (as explored in Love's 3FQQ structure) includes Leu31, which also does not align from H77 with the 4CL1 FASTA sequence. The peaks surrounding the local minima around index 120 implicate cooperative motions near the C-terminal domain. These could be related to the phosphorylation state of NS5A, which is thought to be mediated through unstructured domains following from the C-terminus of domain I.

Correlation network analysis.

The availability of crystal structures for NS5A genotypes is limited. While there are dozens of NS5B ligand-bound structures available, NS5A has never been elucidated in a ligand co-crystallization study. This prevents comparative investigations of conformations for NS5A without computationally intensive generation of meaningful trajectories.

One of the several mechanisms by which NS5A inhibitors are thought to affect domain I incorporates a network between the monomers of a multimeric organization of NS5A proteins. Hyperphosphorylation in domains II and III may be directly affected by NS5A inhibitors binding at domain I, as well as by the associated changes with phosphorylated residues at the C-terminus.[27] If other drugs behave like daclatasvir, they would interfere with this signaling network and induce conformational changes that disrupt RNA synthesis and virion assembly.[11] Dimerization is likely central to both stability, the replication life cycle, and inhibition. However, these interactions following the exposure of the NS5A surface for inhibitor binding are transient, and *in vitro* studies of NS5A interactions are uncommon. In agreement with the hypothesis that NS5A inhibitors have multiple modes of action and reduce both viral replication and assembly, Love et al. concluded that phosphorylation could be a modulation switch between replication and assembly.[26] Furthermore, NS5A inhibitors that have shown multiple binding modes would indicate their ability to bind either phosphorylation state, which provides evidence for their broad activity profiles among many HCV genotypes.[14]

HCV NS5B is an RNA-dependent RNA polymerase.[36] It has an extremely efficient viral replication complex with a catalytic core that yields billions of new virions daily in infected patients. The mechanism by which NS5B carries out replication is not completely clear, but many studies have produced an extended understanding of its functions. Sofosbuvir is a nucleotide analog prodrug NS5B inhibitor heavily used in HCV treatment regimens as part of combination therapies with drugs targeting NS5A and NS3/4A. After metabolism, its converted to its active form, a triphosphate antiviral agent. There are over 150 crystal structures of HCV NS5B available crystallized with ligands and without (apo). Two structures were chosen from a 2015 crystallography study of HCV.[36] The study deposited 12 structures into PDB at an average resolution of 2.7 Å. Eleven of these structures were co-crystallized with a ligand and symmetrical RNA primer template where applicable. The remaining structure presents NS5B isolated in its apo form (PDB 4WT9). Due to its clinical relevance, a structure bound with sofosbuvir diphosphate, Mn^{2+} and the RNA primer was chosen (PDB 4WTG). Direct comparison of these structures' conformations could reveal motions associated with NS5B binding in the absence of NS5A data.
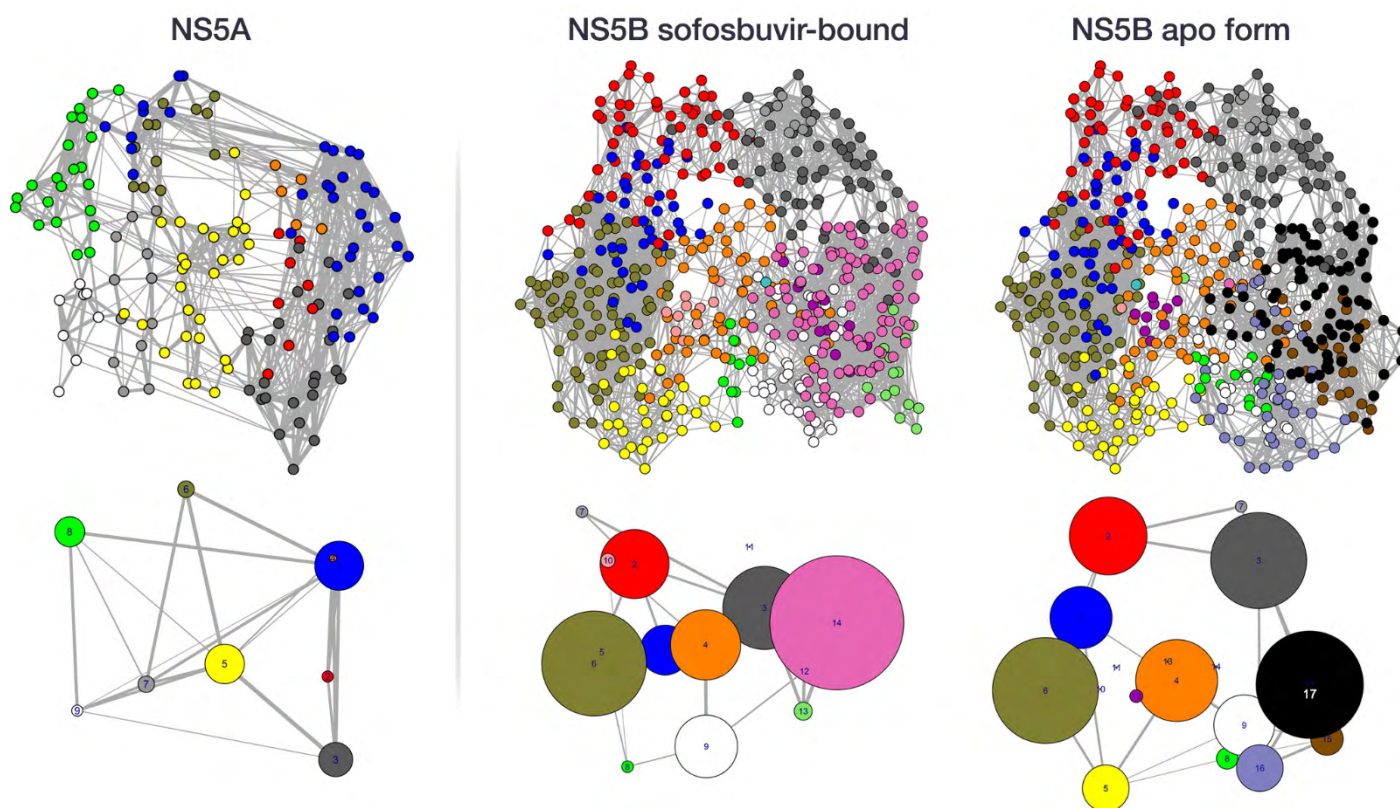

Figure 9. Correlation network analysis via NMA & Gaussian network modeling of NS5A (4CL1) and NS5B in apo form (4WT9) and ligand-bound form (4WTG) with a residue interaction cut-off of 8 Å. The threshold for each individual correlation value was 0.35 when ignoring values for defining network edges. GNM calculation with NS5A showed very few community nodes and thousands of edges (not shown). The NS5A network map shown is derived from standard NMA mode calculation to understand correlated motions between domains IA and IB, indicating an aggregation bias near the CD cleft and C-terminal region. NS5B modes were calculated by GNM. There are 14 community nodes for NS5B with sofosbuvir diphosphate bound and 17 community nodes for NS5B in its apo form (without any ligands). The coarse-grained representation is shown below each all-atom representation. R Bio3D was used for generation and visualization.

All-residue (C^α) and coarse-grained cluster networks were constructed for NS5A and NS5B structures using NMA, GNM, and the Girvan-Newman hierarchical algorithm for clustering.[37] In the all-residue graphs, each C^α atom is a node and the weight of their connecting line represents the residues' associated cross-correlation value. Correlation matrices are developed for each structure after obtaining modes following the Kirchhoff matrix (GNM). Edge generation relies

on the coupling strength (correlation). A cutoff value of 0.35 is assigned, wherein values lower than the cutoff are not satisfactory for constructing edges. Adjusting the cutoff value even marginally yields remarkably different networks. Manual inspection of the data suggest 0.35 was suitable to include relevant nodes and retain visible edges, otherwise the network is too busy due to high connectivity for both proteins. Thick, dense areas indicate high correlation between residues, which may depict cooperative motions in a structure. The Girvan-Newman approach incorporates vertex betweenness centrality to calculate the shortest path between vertex pairs. The nodes are separated by modularity, partitioning them into communal clusters. The betweenness clustering gives rise to dense communities with few edges to other dense neighborhoods.

The sofosbuvir-bound NS5B and apo form NS5B networks are similar, if not identical in most regions. There is significantly greater node packing in the ligand-bound conformation. Sofosbuvir-bound NS5B nodes 9, 13, and 14 aggregate at the binding site, moving closer to other communities but leaving them almost entirely untouched. This movement also generates additional edges for nodal cluster 8. The coarse-grained cluster network illustrates the difference in connectivity by the collection and contraction at nodes 12, 13, and 14. In comparison, the NS5A coarse-grained network has less correlated clusters, despite having significantly less residues than NS5B. There is extremely high connectivity near the CD cleft among the β-hairpins. C^{α} NMA with this clustering method for an NS5A monomer showed communal clustering towards the N-terminal and C-terminal regions. Of note, cluster 1 displayed long-range edge connections between the CD cleft interface and near the C-terminus. Local packing among five large nodes corresponds with NS5A-D1 flexibility. The range observed in cluster 1 suggests a putative hinge site along this axis, likely between domains IA and IB with strongly correlated motions (Figure 10). Cluster 5 denotes a spot where many charged and hydrophobic residues reside.
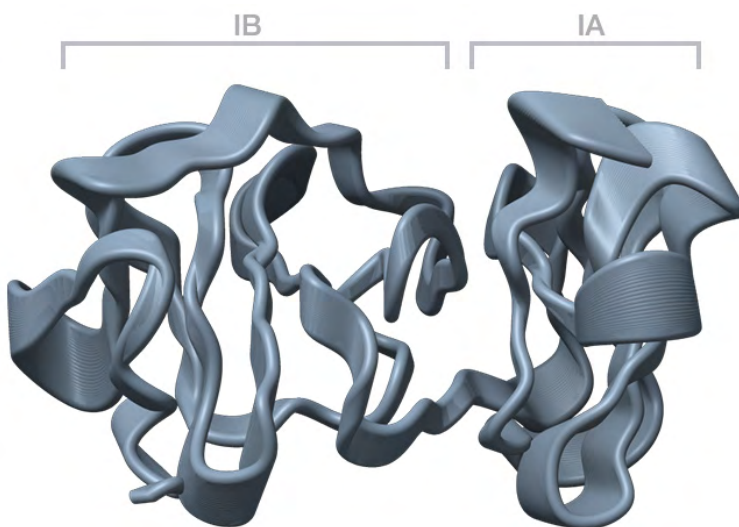

Figure 10. Structure of NS5A-D1 depicting atomic displacements according to directional vectors calculated via ANM. NS5A-D1 subdomains IA and IB are labeled.

Directionality overlaps.

ANM calculations and single-chain deformation analysis measured local flexibility in NS5A atoms relative to neighbors within a 15 Å cutoff. The vectors from slowest mode of the ANM model (Figure 11B) were illustrated to indicate the predicted motions. Principal component analysis examines the relationships between equivalent residues to yield experimental calculations of motions (Figure 11C). Principle components are linearly uncorrelated modes of

variation. The python ProDy interface in combination with the Normal Mode Wizard (NMWiz) VMD plugin was used. ANM was performed on monomer C of the CD dimer in Lambert et al.'s recent structure (PDB 4CL1). PCA was performed on an ensemble of ten available NS5A-D1 monomers. These monomers were previously shown to be in nearly identical conformations with high sequence similarity, so we expect an ideal overlap between the vectors. The lowest-frequency normal mode was overlapped with the first principal component, and overlay was similarly observed with the second mode and component. Overlap analysis revealed comparable directionality towards the center of the protein, but slightly orthogonal arrows in the peripheral areas. The predicted movement of residues 90-93 towards the C-terminal (including Tyr93) and residues 100-103 towards the N-terminal domain might be relevant to the apo conformation of NS5A.

It would have also been very useful to develop conformations of the proteins along ENM modes. While RMSF calculation provides fluctuations magnitudes, deformation analysis can provide deformation vectors and associated energies for each modes obtained from the Hessian. Furthermore, revisiting NS5B apo and ligand-bound conformations would have been very useful, as deformation vectors can be approximated between the different conformations. For this, a custom ensemble of NS5B protein structures would need to be created to separate ligand-bound structures, clean structures, select the chains without primers, and similarly locate the ideal reference compound. The modes could then be inspected to determine which contribute to specific conformational changes upon binding. Obtaining a trajectory from a simulation of NS5B activity would be even more suitable, and VMD was well-equipped to perform this analysis.
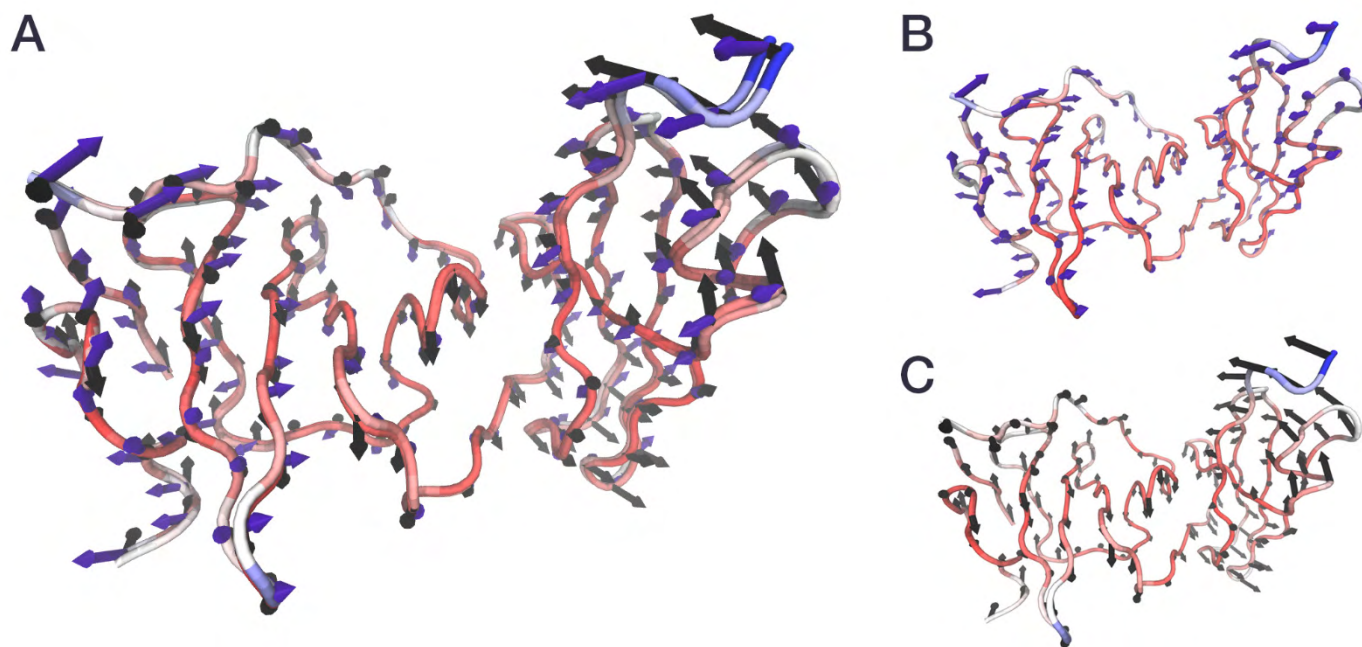

Figure 11. ANM and PCA overlap analysis of first meaningful normal mode and principal components. **(A)** Overlap of the ENM predictions and PCA experimental components based on PDB structures. **(B)** ANM mode calculation yields directionality arrows. **(C)** The relevant principal component presents a slightly different result. Exported from R Bio3D and visualized in VMD.

DISCUSSION

Hepatitis type C infection is a major health burden worldwide. Combination therapies are an invaluable approach that have made use of NS5A and NS5B inhibitors in highly efficacious treatments with incredibly impressive cure rates. Through the use of drugs like sofosbuvir (NS5B) and velpatasvir (NS5A), resistance mutation development is often circumvented and lasting interference of viral replication and assembly is sustained.

Common NS5A inhibitors in clinical use display high structural similarity. They are large, aromatic structures with pseudosymmetry in the context of their target protein. Docking at the most recent crystal structure obtained from NS5A's first domain suggests a likely role in binding for a large groove termed the CD cleft. The study that deposited the relevant structure found two separate dimeric interfaces, evidence for multimerization, and coincided heavily with previous studies that suggested how drugs like daclatasvir modulate NS5A function in binding via affecting phosphorylation. NS5A inhibitors are thought to have two modes of action, bind to both phosphorylation states of NS5A, cause long-range electrostatic interactions in a communication network between NS5A dimers, and affect both the viral replication cycle and NS5A's role in virion assembly. More recently developed drug generations that have expanded on daclatasvir, such as velpatasvir, are thought to function in a similar manner due to their high potency.

Docking at the CD dimer cleft revealed elbasvir, the most similar NS5A ligand to daclatasvir, also prefers to bind at the groove when considering all possible poses. The binding poses are not necessarily symmetric, but necessitate binding in a favorable range to the indispensable Tyr93 residue and zinc-binding motif, or towards hydrophobic residues or possible pi-stacking interactions near aromatic rings. Imperfect binding will likely accommodate binding of RNA, NS5A's endogenous ligand. It has also been suggested that the groove can accommodate RNA due to the presence of a sulfate ion following Lambert et al.'s crystallization. Further docking at the cleft revealed that most clinically useful ligands bind on the surface facing the large groove. The CD dimer is a good candidate for mediating NS5A inhibition upon protein-ligand binding, explaining the high potency following conformational changes that communicate with nearby dimers. The exposed pocket is sufficiently close to Tyrosine-93, an aromatic residue essential for NS5A inhibitor activity.

Correlation network analysis via GNM and NMA reveals that the region near the putative binding interface of the CD dimer is a site of high connectivity that spans across the protein. The N-terminus, C-terminus, and site of Tyr93 are dense, isolated communities unlike the high correlations seen in other relevant proteins like NS5B polymerase (RdRp). There is considerable aggregation bias towards the N-terminal region by the zinc-binding motif and CD interface, but adequate association with other community nodes that depicts the proteins range of flexibility if the groove was to move following binding. Similarly, all analyses display deformations at the C-terminal region where the unstructured domains are typically found, further suggesting NS5A-D1 motions and D2/D3 activity are inter-dependent. The network graph implies high flexibility in the protein, and suggests a hinge region between the subdomains. PCA of NS5A-D1 monomers agrees that this loop exhibits anticorrelated motions. Further exploration of NS5A-D1 in a ligand-bound conformation is necessitated for clarifying the mechanism. In particular, characterizing different NS5A genotypes may reveal why DAAs have broad efficacy profiles despite variance in key resistance-conferring mutations across subtypes.[18]

Co-crystallization of NS5A with a ligand has proven to be a difficult challenge. Studies that consider NS5A-D2 and D3 phosphorylation and provide comparative investigations between ligand-bound and apo form proteins would further elucidate inhibitor activity at this interface by characterizing the phosphorylation state of NS5A.

All-atom molecular dynamics (MD) simulations will be extremely relevant and useful to this protein problem. Unlike NMA-based models, MD simulations can explore anharmonic fluctuations.[39] As conformational sampling times improve, long-range signaling of NS5A towards NS5B and neighboring dimers as related to inhibitor binding can be explored. It is not certain whether MD simulation is feasible in this context as the utility of plausible time scales in studying protein-ligand binding at NS5A is ambiguous, but several binding modes can be captured by MD simulations and explored for their accuracy.[38] Previously, MD simulation has been used with success in NS5B,[47] but made minimal conclusions about NS5B conformations for RNA replication. Considering NS5A has been shown to be greatly flexible,

the protein is a good candidate for simulating conformational transitions and observing results for overlap with ENM results. Functional transitions between conformational forms can be characterized as to demonstrate the extent of the network involved with NS5A upon protein-ligand binding.[24] Additionally, characterization of mechanisms at domains I and II will include the role of phosphorylation in inhibition. Even so, ENMs like the GNM are invaluable tools to characterize NS5A-D1 flexibility. Future studies should characterize NS5A-D1 dimerization, multimerization structures, communication network activity mediated by phosphorylation, and the specific interfaces where DAAs bind to induce inhibition.

SOFTWARE

The ProDy Python Package for protein structural dynamics analysis.[40,41]
The Bio3D R package for protein sequence, structure and trajectory analysis.[42]
The Visual Molecular Dynamics (VMD) software for displaying large biomolecules.[43]
The Normal Mode Wizard (NMWiz) VMD plugin for visual comparative analysis of normal modes.
PyMol, an open-source molecular visualization software.[44]
MarvinSketch, a chemical structure editor built in Java.
UCSF Chimera, a software for molecular modeling. The Chimera Dock Prep tool was used prior to docking with AutoDock Vina.[30,32]
SwissDock, a web service for molecular docking.[31]

ABBREVIATIONS

HCV Hepatitis C Virus; **NS** Nonstructural; **D1** Domain I; **LCS** Low-complexity sequence;
GNM Gaussian Network Model; **ANM** Anisotropic Network Model; **NMA** Normal Mode Analysis;
MD molecular dynamics; **DAA** direct-acting antiviral; **DCV** daclatasvir; **LDV** ledipasvir

REFERENCES

[1] Li, H.-C. (2015). Hepatitis C virus: Virology, diagnosis and treatment. World Journal of Hepatology. Baishideng Publishing Group Inc. doi:10.4254/wjh.v7.i10.1377

[2] Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., … Murray, C. J. L. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. Elsevier BV. doi:10.1016/s0140-36(16)31678

[3] KATO, N. (2000). Genome of Human Hepatitis C Virus (HCV): Gene Organization, Sequence Diversity, and Variation. Microbial & Comparative Genomics. Mary Ann Liebert Inc. doi:10.1089/omi.1.2000.5.129

[4] Ross-Thriepland, D., & Harris, M. (2014). Hepatitis C virus NS5A: enigmatic but still promiscuous 10 years on! Journal of General Virology. Microbiology Society. doi:10.1099/jgv.0.000009

[5] O'Boyle II, D. R., Sun, J.-H., Nower, P. T., Lemm, J. A., Fridell, R. A., Wang, C., … Gao, M. (2013). Characterizations of HCV NS5A replication complex inhibitors. Virology. Elsevier BV. doi:10.1016/j.virol.2013.06.032

[6] Sólyom, Z., Ma, P., Schwarten, M., Bosco, M., Polidori, A., Durand, G., … Brutscher, B. (2015). The Disordered Region of the HCV Protein NS5A: Conformational Dynamics, SH3 Binding, and Phosphorylation. Biophysical Journal. Elsevier BV. doi:10.1016/j.bpj.2015.06.040

[7] Bartenschlager, R., Lohmann, V., & Penin, F. (2013). The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. Nature Reviews Microbiology. Springer Nature. doi:10.1038/nrmicro3046

[8] Sofosbuvir (Sovaldi). (n.d.). Retrieved May 25, 2017, from hepatitisc.uw.edu

[9] Ascher, D. B., Wielens, J., Nero, T. L., Doughty, L., Morton, C. J., & Parker, M. W. (2014). Potent hepatitis C inhibitors bind directly to NS5A and reduce its affinity for RNA. Scientific Reports. Springer Nature. doi:10.1038/srep04765

[10] Chukkapalli, V., Berger, K. L., Kelly, S. M., Thomas, M., Deiters, A., & Randall, G. (2015). Daclatasvir inhibits hepatitis C virus NS5A motility and hyper-accumulation of phosphoinositides. Virology. Elsevier BV. doi:10.1016/j.virol.2014.12.018

[11] Guedj, J., Dahari, H., Rong, L., Sansone, N. D., Nettles, R. E., Cotler, S. J., … Perelson, A. S. (2013). Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. Proceedings of the National Academy of Sciences. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1203110110

[12] Scheel, T. K. H., Prentoe, J., Carlsen, T. H. R., Mikkelsen, L. S., Gottwein, J. M., & Bukh, J. (2012). Analysis of Functional Differences between Hepatitis C Virus NS5A of Genotypes 1–7 in Infectious Cell Culture Systems. (A. Siddiqui, Ed.), PLoS Pathogens. Public Library of Science (PLoS). doi:10.1371/journal.ppat.1002696

[13] Nakamoto, S. (2014). Hepatitis C virus NS5A inhibitors and drug resistance mutations. World Journal of Gastroenterology. Baishideng Publishing Group Inc. doi:10.3748/wjg.v20.i11.2902

[14] Nettles, J. H., Stanton, R. A., Broyde, J., Amblard, F., Zhang, H., Zhou, L., … Schinazi, R. F. (2014). Asymmetric Binding to NS5A by Daclatasvir (BMS-790052) and Analogs Suggests Two Novel Modes of HCV Inhibition. Journal of Medicinal Chemistry. American Chemical Society (ACS). doi:10.1021/jm501291c

[15] Messina, J. P., Humphreys, I., Flaxman, A., Brown, A., Cooke, G. S., Pybus, O. G., & Barnes, E. (2014). Global distribution and prevalence of hepatitis C virus genotypes. Hepatology. Wiley-Blackwell. doi:10.1002/hep.27259

[16] Issur, M., & Götte, M. (2014). Resistance Patterns Associated with HCV NS5A Inhibitors Provide Limited Insight into Drug Binding. Viruses. MDPI AG. doi:10.3390/v6114227

[17] Lim, P. J., Chatterji, U., Cordek, D., Sharma, S. D., Garcia-Rivera, J. A., Cameron, C. E., … Gallay, P. A. (2012). Correlation between NS5A Dimerization and Hepatitis C Virus Replication. Journal of Biological Chemistry. American Society for Biochemistry & Molecular Biology (ASBMB). doi:10.1074/jbc.m112.376822

[18] Hayward S, de Groot BL. Normal Modes and Essential Dynamics, in: Kukol A, editor. Methods in Molecular Biology, vol. 443, Molecular Modeling of Proteins, New Jersey, Humana Press, 2008, 89–106. link

[19] Y.H. Sanejou and, Elastic Network Models: Theoretical and Empirical Foundations, in: L. Monticelli, E. Salonen (Eds.), Biomolecular Simulations, in: The Series Methods in Molecular Biology, vol. 924, Humana Press, 2012. arXiv:1102.2402

[20] Jun Koo Park. Normal mode analysis and gaussian network model. Technical report, Iowa State University, 2008. link

[21] Atilgan, A. R., Durell, S. R., Jernigan, R. L., Demirel, M. C., Keskin, O., & Bahar, I. (2001). Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. Biophysical Journal. Elsevier BV. doi:10.1016/s0006-3495(01)76033-x

[22] Haliloglu, T., Bahar, I., & Erman, B. (1997). Gaussian Dynamics of Folded Proteins. Physical Review Letters. American Physical Society (APS). doi:10.1103/physrevlett.79.3090

[23] Tirion, M. M. (1996). Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. Physical Review Letters. American Physical Society (APS). doi:10.1103/physrevlett.77.1905

[24] Rader AJ, Chennubhotla C, Yang LW, Bahar I, The Gaussian Network Model: theory and applications. In: CuiQ, BaharI, Eds. 2006. Normal mode analysis. Theory and applications to biological and chemical systems. Boca Raton, FL: CRC Press, Taylor & Francis Group, 41–64. link

[25] Tellinghuisen, T. L., Marcotrigiano, J., & Rice, C. M. (2005). Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. Nature. Springer Nature. doi:10.1038/nature03580

[26] Love, R. A., Brodsky, O., Hickey, M. J., Wells, P. A., & Cronin, C. N. (2009). Crystal Structure of a Novel Dimeric Form of NS5A Domain I Protein from Hepatitis C Virus. Journal of Virology. American Society for Microbiology. doi:10.1128/jvi.02352-08

[27] Lambert, S. M., Langley, D. R., Garnett, J. A., Angell, R., Hedgethorne, K., Meanwell, N. A., & Matthews, S. J. (2014). The crystal structure of NS5A domain 1 from genotype 1a reveals new clues to the mechanism of action for dimeric HCV inhibitors. Protein Science. Wiley-Blackwell. doi:10.1002/pro.2456

[28] Lim, P. J., Chatterji, U., Cordek, D., Sharma, S. D., Garcia-Rivera, J. A., Cameron, C. E., … Gallay, P. A. (2012). Correlation between NS5A Dimerization and Hepatitis C Virus Replication. Journal of Biological Chemistry. American Society for Biochemistry & Molecular Biology (ASBMB). doi:10.1074/jbc.m112.376822

[29] Hwang, J., Huang, L., Cordek, D. G., Vaughan, R., Reynolds, S. L., Kihara, G., … Cameron, C. E. (2010). Hepatitis C Virus Nonstructural Protein 5A: Biochemical Characterization of a Novel Structural Class of RNA-Binding Proteins. Journal of Virology. American Society for Microbiology. doi:10.1128/jvi.01319-10

[30] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--A visualization system for exploratory research and analysis. Journal of Computational Chemistry. Wiley-Blackwell. doi:10.1002/jcc.20084

[31] Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Research. Oxford University Press (OUP). doi:10.1093/nar/gkr366

[32] Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry. Wiley-Blackwell. doi:10.1002/jcc.21334

[33] Berger, C., Romero-Brey, I., Radujkovic, D., Terreux, R., Zayas, M., Paul, D., … Bartenschlager, R. (2014). Daclatasvir-Like Inhibitors of NS5A Block Early Biogenesis of Hepatitis C Virus–Induced Membranous Replication Factories, Independent of RNA Replication. Gastroenterology. Elsevier BV. doi:10.1053/j.gastro.2014.07.019

[34] Trueblood, K. N., Bürgi, H. B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., … Abrahams, S. C. (1996). Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. Acta Crystallographica Section A Foundations of Crystallography. International Union of Crystallography (IUCr). doi:10.1107/s0108767396005697

[35] Sun, J.-H., O'Boyle II, D. R., Zhang, Y., Wang, C., Nower, P., Valera, L., … Gao, M. (2012). Impact of a baseline polymorphism on the emergence of resistance to the hepatitis C virus nonstructural protein 5a replication complex inhibitor, BMS-790052. Hepatology. Wiley-Blackwell. doi:10.1002/hep.25581

[36] Appleby, T. C., Perry, J. K., Murakami, E., Barauskas, O., Feng, J., Cho, A., … Edwards, T. E. (2015, February 12). Structural basis for RNA replication by the hepatitis C virus polymerase. Science. American Association for the Advancement of Science (AAAS). doi:10.1126/science.1259210

[37] Girvan, M., & Newman, M. E. J. (2002, June 11). Community structure in social and biological networks. Proceedings of the National Academy of Sciences. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.122653799

[38] Chen, Y.-C. (2015). Beware of docking! Trends in Pharmacological Sciences. Elsevier BV. doi:10.1016/j.tips.2014.12.001

[39] Doruker P, Atilgan AR, Bahar I. (2000). Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor. Proteins. 40(3):512-24. pmid:10861943

[40] Bakan, A., Meireles, L. M., & Bahar, I. (2011). ProDy: Protein Dynamics Inferred from Theory and Experiments. Bioinformatics. Oxford University Press (OUP). doi:10.1093/bioinformatics/btr168

[41] Bakan, A., Dutta, A., Mao, W., Liu, Y., Chennubhotla, C., Lezon, T. R., & Bahar, I. (2014). Evol and ProDy for bridging protein sequence evolution and structural dynamics. Bioinformatics. Oxford University Press (OUP). doi:10.1093/bioinformatics/btu336

[42] Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., & Caves, L. S. D. (2006). Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics. Oxford University Press (OUP). doi:10.1093/bioinformatics/btl461

[43] Humphrey, W., Dalke, A. and Schulten, K., VMD - Visual Molecular Dynamics. J. Molec. Graphics. (1996). vol. 14, pp. 33-38.

[44] The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

[45] Li,H., Chang,YY, Lee,JY, Bahar,I. and Yang,LW. (2017) DynOmics: dynamics of structural proteome and beyond. Nucleic Acids Res., online published.

[46] Zheng, W., Liao, J.-C., Brooks, B. R., & Doniach, S. (2007). Toward the mechanism of dynamical couplings and translocation in hepatitis C virus NS3 helicase using elastic network model. Proteins: Structure, Function, and Bioinformatics. Wiley-Blackwell. doi:10.1002/prot.21326

[47] Asafi, M. S., Yildirim, A., & Tekpinar, M. (2016). Comparative Investigation of Normal Modes and Molecular Dynamics of Hepatitis C NS5B Protein. Journal of Physics: Conference Series. IOP Publishing. doi:10.1088/1742-6596/707/1/012036

[48] Lezon, T. R., Shrivastava, I. H., Yang, Z. & Bahar, I. (2009) Elastic Network Models For Biomolecular Dynamics: Theory and Application to Membrane Proteins and Viruses. Handb. Biol. Networks. 129–158. link