

Article

Mechanistic Insights into the Protection Effect of Argonaute–RNA Complex on the HCV Genome

Haiming Zhuang^{1,†}, Dong Ji^{2,†}, Jigang Fan¹, Mingyu Li¹, Ran Tao³, Kui Du³, Shaoyong Lu^{1,*}, Zongtao Chai^{4,5,*} and Xiaohua Fan^{2,*}

¹ Medicinal Chemistry and Bioinformatics Center, School of Medicine, Shanghai Jiao Tong University, Shanghai 200025, China

² Department of Anesthesiology, Changhai Hospital, Naval Medical University, Shanghai 200433, China

³ School of Chemistry and Chemical Engineering, Shaoxing University, Shaoxing 312000, China

⁴ Key Laboratory of Carcinogenesis and Cancer Invasion of Ministry of Education, Department of Liver Surgery and Transplantation, Liver Cancer Institute and Zhongshan Hospital, Fudan University, Shanghai 200032, China

⁵ Department of Hepatic Surgery, Shanghai Geriatric Cancer, Shanghai 201104, China

* Correspondence: lushaoyong@sjtu.edu.cn (S.L.); runout@163.com (Z.C.); fxh95007@126.com (X.F.)

† These authors contributed equally to this work.

Abstract: While host miRNA usually plays an antiviral role, the relentless tides of viral evolution have carved out a mechanism to recruit host miRNA as a viral protector. By complementing miR-122 at the 5' end of the genome, the hepatitis C virus (HCV) gene can form a complex with Argonaute 2 (Ago2) protein to protect the 5' end of HCV RNA from exonucleolytic attacks. Experiments showed that the disruption of the stem-loop 1 (SL1) structure and the 9th nucleotide (T9) of HCV site 1 RNA could enhance the affinity of the Ago2 protein to the HCV site 1 RNA (target RNA). However, the underlying mechanism of how the conformation and dynamics of the Ago2: miRNA: target RNA complex is affected by the SL1 and T9 remains unclear. To address this, we performed large-scale molecular dynamics simulations on the AGO2-miRNA complex binding with the WT target, T9-abasic target and SL1-disruption target, respectively. The results revealed that the T9 and SL1 structures could induce the departing motion of the PAZ, PIWI and N domains, propping up the mouth of the central groove which accommodates the target RNA, causing the instability of the target RNA and disrupting the Ago2 binding. The coordinated motion among the PAZ, PIWI and N domains were also weakened by the T9 and SL1 structures. Moreover, we proposed a new model wherein the Ago2 protein could adopt a more constraint conformation with the proximity and more correlated motions of the PAZ, N and PIWI domains to protect the target RNA from dissociation. These findings reveal the mechanism of the Ago2-miRNA complex's protective effect on the HCV genome at the atomic level, which will offer guidance for the design of drugs to confront the protection effect and engineering of Ago2 as a gene-regulation tool.

Keywords: argonaute; microRNA; molecular dynamics simulations; hepatitis C virus; gene regulation



Citation: Zhuang, H.; Ji, D.; Fan, J.; Li, M.; Tao, R.; Du, K.; Lu, S.; Chai, Z.; Fan, X. Mechanistic Insights into the Protection Effect of Argonaute–RNA Complex on the HCV Genome. *Biomolecules* **2022**, *12*, 1631. <https://doi.org/10.3390/biom12111631>

Academic Editor: Alexander A. Bolshoy

Received: 29 September 2022

Accepted: 31 October 2022

Published: 3 November 2022

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



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

1. Introduction

MicroRNAs (miRNA), which are short non-coding RNAs of ~22 nucleotides, could guide Argonaute (Ago) proteins to interact with target complementary RNAs, yielding the RNA-induced silencing complex (RISC) [1,2]. The formation of the RISC can induce post-transcriptional gene silencing through translation repression or target RNA degradation. This approach plays a key role in regulation of endogenous gene expression and the defense against exogenous RNAs [3]. In targeting the exogenous RNAs from viruses, the host miRNA usually plays an antiviral role. However, the relentless tides of viral evolution have carved out mechanisms to recruit host miRNA as a viral protector [4].

Among them, Hepatitis C Virus (HCV) uniquely utilizes the liver-specific miRNA-122 for replication. MiRNA-122 could interact with the 5' noncoding region of the HCV genome, which was the first direct viral RNA–host miRNA interaction to be discovered [5,6]. An in vitro cell assay showed that this interaction is required to maintain a high viral RNA abundance [5]. HCV infection is a leading cause of liver failure and hepatocellular carcinoma worldwide, with there being over 170 million infected individuals [7,8]. However, the current standard anti-HCV therapy with interferon- α and ribavirin maintains viral clearance in only 50% of patients and lacks specificity [8]. Thus, targeting miRNA-122 to inhibit the HCV replication could be an effective therapeutic strategy with higher specificity, and the miRNA-122-specific inhibitor Miravirsen has shown promising results against HCV infection in phase II clinical trials [8–10].

Two complementary sites (site 1 and site 2) of miRNA-122 at the very end of the HCV genome have been reported [11]. Besides the seed region and the supplementary region that is complementary to the miRNA-122, the site 1 region of the HCV genome contains a special stem-loop structure (SL1) (Figure 1). The host protein Ago2, which acts as the core of RISC, is guided by the miRNA-122 to bind with the site 1 and site 2 regions, protecting the 5' end of HCV genome from enzyme attacks [12]. Ago2 is a bi-lobed protein with a central groove for the binding of miRNAs and target RNAs, and it is composed of four globular domains (N, PAZ, MID and PIWI) and two linker domains (L1 and L2) [13] (Figure 1). The newly reported crystal structure of the Ago2: miRNA-122: HCV site 1 (target RNA) showed that with the SL1 structure propping up the central cleft between the PAZ domain and the PIWI domain, the Ago2 protein adopted an extended conformation [14]. The crystal structure also showed that the 9th nucleotide of the target RNA (T9, guanine) is flipped out and bound in a hydrophobic pocket of Ago2, which was observed for the first time. For the special structure of the SL1 and T9, experiments have shown that the disruption of the SL1 by complete nucleotide mutation to adenine or the depletion of the basic group of the T9 could increase the Ago2 affinity with a lower Ago2-off rate [14], indicating that the SL1 and T9 are detrimental to Ago2 binding. This seems to be contradictory to the fact that the SL1 and T9 are highly conserved in the HCV genome.

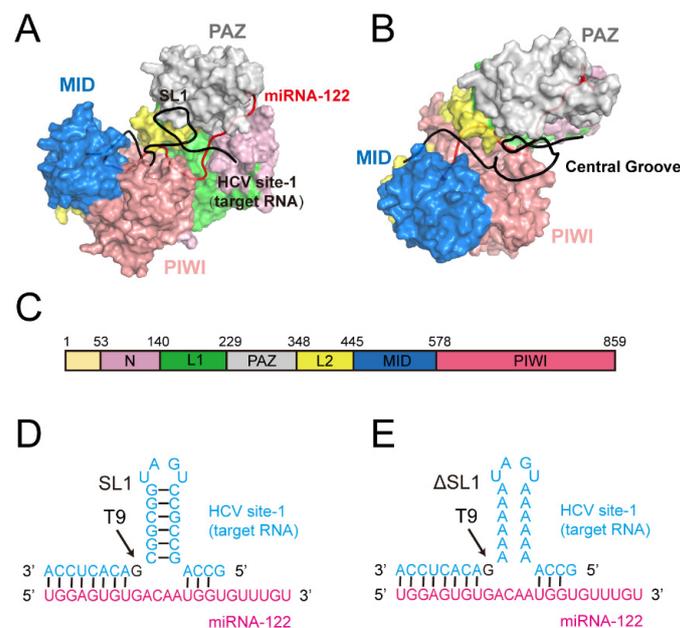


Figure 1. Overview of the Ago2–RNA complex. (A,B) Surface representation of the Ago2: miRNA (red): HCV site 1 RNA (black) complex. The SL1 structure and the central groove were noted. (C) The domain organization of the Ago2 protein. Nucleotide sequences of miR-122 (magenta) and target RNA (blue) in the WT (D) and the Δ SL1 (E) system. The T9 (G21) nucleotide (black) is indicated with arrows.

Although research has shown that this might be related to the unique mechanism of the exonuclease involved in the target RNA degradation [14], the underlying mechanism of how the conformation of Ago2: miRNA: HCV site 1 is affected by the SL1 and T9 remains unclear. Furthermore, although there are many studies on the assembly of the Ago2: miRNA: target RNA complex, the structural basis of how Ago2-miRNA interacting with the target is still poorly understood [15–17]. Thus, with the special structure of the SL1 and T9, the Ago2: miRNA: HCV site 1 complex is a good entry point to study the inner interaction of the Ago–RNA complexes. Additionally, although the crystal structure of the Ago2: miRNA: HCV site 1 complex has been solved, it is challenging to provide enough information for the conformational dynamics of the ternary complex with only one averaged snapshot. A molecular dynamics (MD) simulation, which explores the protein and nucleic acid conformation at the dynamic and atomic levels [18–22], is helpful to investigate the protein binding pocket [23,24] and protein-DNA/RNA recognition [25]. In recent works, we have used the MD simulation to illuminate the structural mechanism of the protein–protein or protein-nucleic acid interactions [26–30]. Previous studies on the Ago2 complexes utilizing the MD simulation have contributed to the understanding of the mechanism of the Ago2–RNA interaction [31–33].

Here, we performed an extensive large-scale MD simulation to investigate how the T9 and SL1 structures of the HCV site 1 RNA affect the target RNA binding affinities. The results showed that the T9 and SL1 structures could induce the departing and less coordinated motion of the PAZ, PIWI and N domains, propping up the mouth of the central groove, which in turn causes the target RNA to be liable to dissociate. Moreover, combining with previous studies, we proposed a new model that the Ago2 protein could adopt a more constraint conformation with the proximity to and the more correlated motion of the PAZ, N and PIWI domains to protect the target RNA from dissociation. This study aims to shed light on the mechanistic understanding of the role of the T9 and SL1 structures, inspiring further research on developing anti-HCV strategies and engineering Ago proteins as gene regulation tools.

2. Materials and Methods

2.1. Structure Preparation

Three systems were performed with the same Ago2: miRNA-122 complex and different target RNAs, including the wild-type target RNA (denoted as “WT”), the T9 base pair-depleted target RNA (denoted as “T9-abasic”) and the SL1-disrupted target RNA (denoted as “ Δ SL1”). The SL1 disruption was achieved by the mutation of all of the paired nucleotides in the stem-loop to adenine. The RNA sequences in the three systems are shown in Figure 1D,E. The crystal structure of the Ago2: miRNA-122: HCV site 1 was selected as the initial structure (PDB ID: 7KI3) [14]. The missing nucleotides were added based on the structure of the Ago2–RNA complex with the PDB ID 4Z4D [34], and the missing residues of the Ago2 protein was added using the MODELLER program. Ten thousand steps of minimization were then performed using Discovery Studio 2019 by the steepest descent algorithm to optimize the complex structure. The minimized structure was selected as the WT system. For the T9-abasic system, the base of T9 was depleted using Discovery Studio 2019, and the non-standard force field of the abasic site was built utilizing the Antechamber program of Amber and the Generalized Amber Force Field. The mutations of the nucleotides in the Δ SL1 system were performed using Discovery Studio 2019.

2.2. MD Simulation

The systems were prepared with the LEaP program using ff14SB and “Rochester” torsions force field to describe the RNA–protein complexes [35–38]. The ternary complex was solvated into a truncated octahedron transferable intermolecular potential three-point (TIP3P) water box [39], and Na⁺ and Cl[−] were added to simulate a normal saline environment. Two rounds of minimization were performed. The first round contained 5000 steps

maximum minimization cycles with the complex being fixed, and the second round contained 10,000 steps maximum minimization cycles with no constraints. Subsequently, all of the systems were heated from 0 to 300 K within 300 ps, which was followed by 700 ps equilibration running in a canonical ensemble (NVT). After all of the preparations were conducted, 5 replicas of independent 1 μ s simulations were performed under random velocities under isothermal isobaric (NPT) conditions. For the long-range electrostatic interactions, the particle mesh Ewald method [40] was used. Covalent bonds involving hydrogens were constrained using the SHAKE method [41].

2.3. Principal Component Analysis (PCA)

The principal component analysis (PCA), which is widely used in describing the kinetic processes that occur during simulations, can transform a series of potentially coordinated observations into orthogonal vectors. Among the vectors, the first two principal component (PC1 and PC2) provide the dominant motions during the simulation [42]. In this analysis, the PCs were generated based on the coordinate covariance matrix of the C α atoms in the Ago protein every 10 frames, and the collected frames were all projected onto the first and second PCs.

2.4. Generalized Correlation Analysis

Generalized correlation (GC_{ij}) analysis was applied to calculate the correlated motion. Compared to the traditional Pearson correlation analysis, the GC_{ij} analysis is independent of the mutual orientations of the atomic fluctuations and can distinguish between the linear and nonlinear contributions [43]. Mutual Information (MI) was introduced to reflect how much information of one atom's position was provided by another:

$$MI[x_i, x_j] = \int \int p(x_i, x_j) \ln \frac{p(x_i, x_j)}{p(x_i)p(x_j)} dx_i dx_j \quad (1)$$

The right side of the equation can be related to the more widely known measure of entropy which is calculated by:

$$H[x] = \int p(x) \ln p(x) dx \quad (2)$$

To calculate based on the correlation between the pairs of atoms:

$$MI[x_i, x_j] = H[x_i] + H[x_j] - H[x_i, x_j] \quad (3)$$

$MI[x_i, x_j]$ was further related to a more intuitive Pearson-like correlation coefficient GC_{ij} , which can be calculated by:

$$GC_{ij} = \left\{ 1 - e^{-\frac{2MI[x_i, x_j]}{d}} \right\}^{-\frac{1}{2}} \quad (4)$$

where d represents the dimensionality of x_i and x_j , which is 3 in this study. The GC_{ij} calculation was applied using `g_correlation` tool in Gromacs 3.3 [44], with the coordinates of C α in each residue. The inter-domain correlation was calculated by the accumulation of the GC_{ij} value between each residue in the respective domains. The accumulated value was normalized based on the maximum and minimum values in the counterpart in all of the three systems.

2.5. Dynamic Network Analysis

A network analysis was performed to reflect the motion connection using VMD [45]. In our analysis, the C α atoms in the Ago2 protein, the N1 atoms in the uracil and cytosine, the N9 atoms in the adenine and guanine and the P atom in the abasic site were chosen as the nodes to represent their corresponding residues. The edges were drawn between the

nodes whose distances were within a cutoff of 4.5 Å for at least 75% of the simulation time. The edge was calculated by:

$$d_{i,j} = -\log(|C_{i,j}|) \quad (5)$$

where i and j represent the two nodes. Additionally, community, which means a group of residues whose connections are stronger was calculated using the Girvan–Newman algorithm [46] through the gcommunities program of VMD.

2.6. Markov State Model (MSM)

The Markov state model (MSM) is a mathematical framework that is used to describe the dynamics of time-series data [47–50]. In our analysis, the MSM was used to differentiate between the conformational states of the energy basin map and to help to extract representative structures from each state. All of the MSM calculations were performed using the PyEMMA software 2.5.7. The MSM transition matrix was first calculated based on the probability of the transition between the different states.

An implied timescale (ITS) test was performed to check the Markovian property and select the proper lag time (τ). The ITS as a function of τ can be calculated by:

$$t_i = -\frac{\tau}{\ln|\lambda_i(\tau)|} \quad (6)$$

where λ_i means the eigenvalue which was obtained from the MSM transition matrix of the i th process. The implied timescale plot can be drawn in which each curve represents the average transition time of one process. When the curve became approximately constant, the corresponding lag time was chosen for the following analysis and the system had a Markovian property. In this study, the lag time was set as 1 ns.

Every frame of each system was projected on the free energy plot according to the characteristic vectors. Then, the K-means algorithm was applied to cluster the two-dimensional conformations into 300 microstates in each system. Based on the 300 microstates, a PCCA+ algorithm was performed to divide them into different clusters. With the divided clusters, the Chapman–Kolmogorov test was applied to further testify the property of the MSM. The structure with the smallest root-mean-square deviation (RMSD) with the frames in the same cluster was chosen as the representative structure.

3. Results

3.1. The Disruption of T9 and SL1 Affects Conformational Dynamics of Ago2–RNA Complex

A total of three systems with different target RNAs were performed, one with the wild-type HCV site 1 RNA (WT), one with the T9 base pair depletion target RNA (T9-abasic), and the other with the SL1-disruption target RNA (Δ SL1). The RNA sequences of the three systems are shown in Figure 1D,E. To investigate the overall conformational dynamics in the three systems, we calculated the C α atoms root-mean-square deviation (RMSD) of the Ago2 protein relative to the initial structure. The RMSD reflects the amplitude, but not the direction of the variation of the simulated structures relative to the initial structure. The results indicated that the three systems reached equilibrium after 200 ns of the simulation (Supplementary Figure S1), and the following analyses were all based on the frames in equilibrium. The RMSD values were 2.06 ± 0.30 Å for the WT system, 2.29 ± 0.42 Å for the T9-abasic system and 2.00 ± 0.21 Å for the Δ SL1 system.

To reflect the local dynamics in the Ago2–RNA complex, the root-mean-square fluctuation (RMSF) of each residue (represented by the C α atoms for Ago2 and P atoms for RNA) was calculated (Supplementary Figure S2). To better show the differences, we subtracted the RMSF value of the T9-abasic and the Δ SL1 system from its counterpart in the WT system and projected the absolute value of the WT system and the difference values of the T9-abasic and the Δ SL1 system on the protein structure (Figure 2). The inspection of the RMSF plot showed that the SL1 region displayed the highest fluctuation in the WT system, and its RMSF value significantly dropped in the T9-abasic and the Δ SL1 system, indicating

that the disruption of T9 and SL1 could stabilize the stem-loop region of the target RNA. In addition, the global fluctuation of the PAZ and PIWI domains generally decreased with the disruption of T9 and SL1, which may be correlated with the tension release of the central groove of Ago2. All of these findings were in line with the fact that compared to the WT system, the disruption of T9 or SL1 could increase the Ago2 affinity to the target.

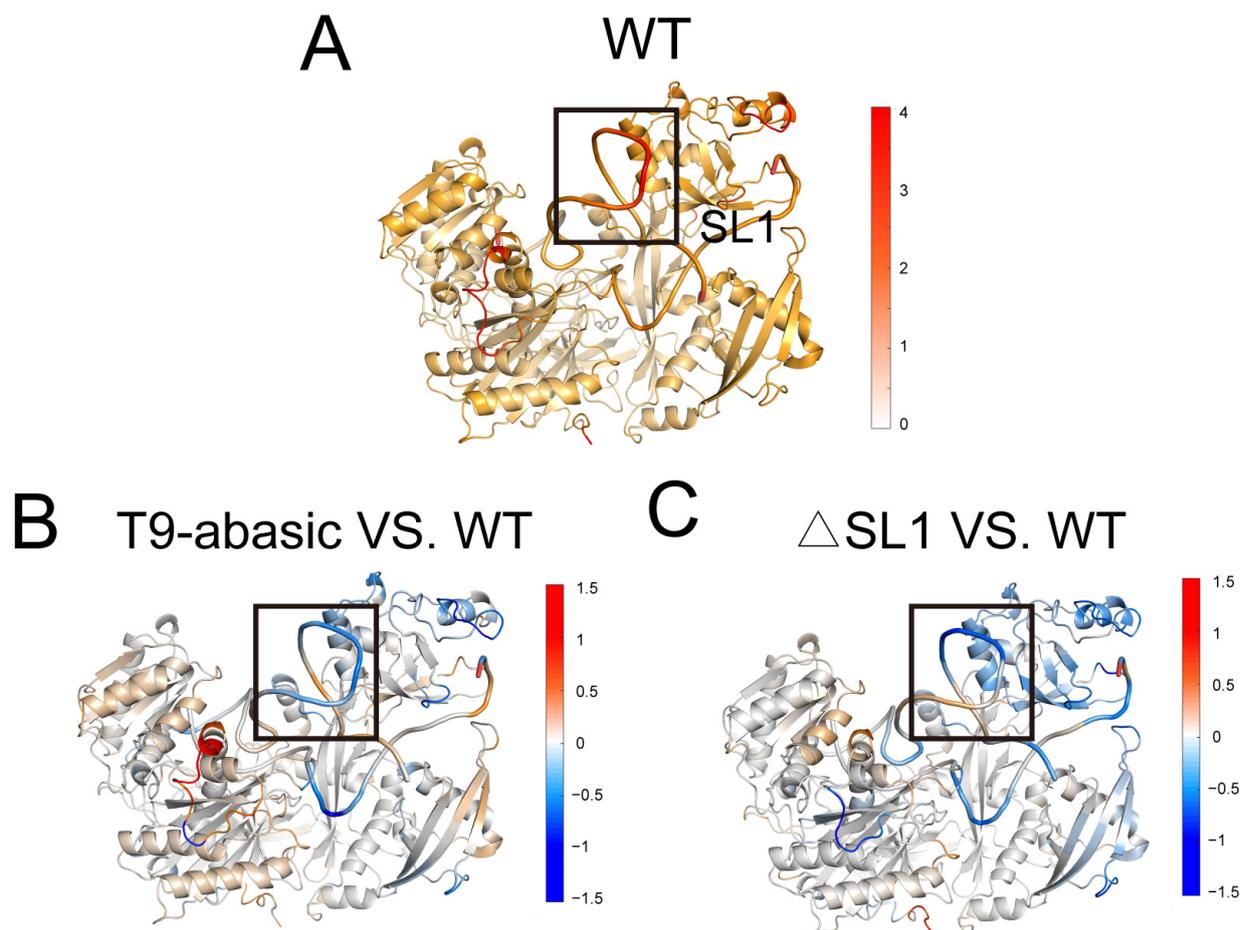


Figure 2. Conformational dynamics of the Ago2–RNA complex. (A) The averaged RMSF values of each residue are plotted on the 3D structures of the Ago2 protein in the WT system. The substations of RMSF values of the T9-abasic and Δ SL1 systems from the counterpart of the WT systems are plotted on the 3D structures (B,C). Positive regions (red) stand for higher RMSF values, whereas negative regions (blue) represent lower RMSF values. The significant fluctuation differences are highlighted with a dashed rectangle in each panel.

3.2. The Disruption of T9 and SL1 Induces the Conformational Transition of Ago2

To characterize the global conformational transition in the three systems, a principal component analysis (PCA) was performed (see in Materials and Methods). We projected the snapshots every 10 frames throughout the simulation of each system on the two-dimensional plot according to the first two principal components (PC1 and PC2) (Figure 3A), and the domain motion along the PC1 and PC2 was shown on the Ago2 protein (Figure 3B,C).

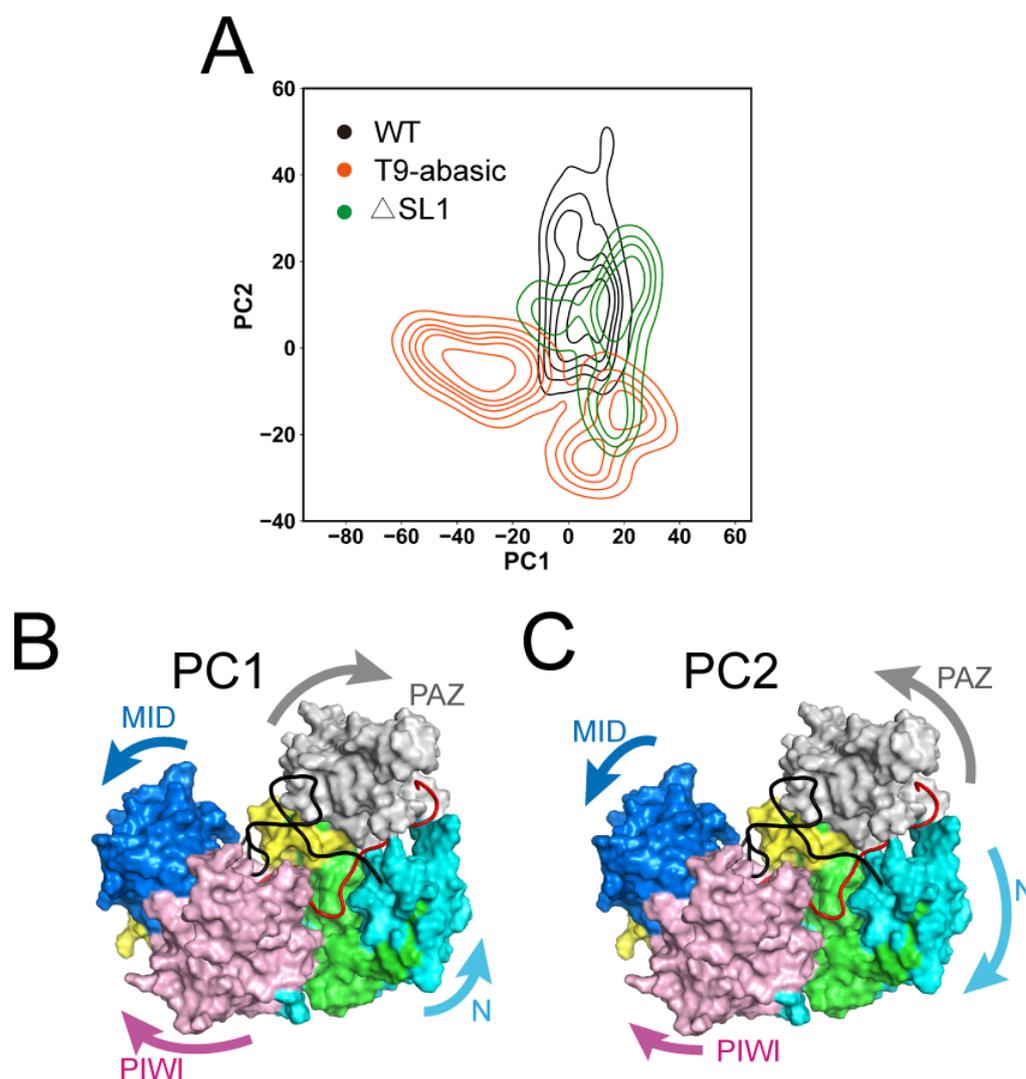


Figure 3. Global conformational transitions of the Ago2–RNA complex. (A) Projections of the first and second principal components (PC1 and PC2) from MD simulations of the Ago2 protein in three systems. The domain motion along PC1 (B) and PC2 (C) was shown.

The T9-abasic system contains one cluster of conformations with a PC1 value that is less than -10 , which was not observed in the WT and Δ SL1 system, indicating a wider conformational space that was induced by the T9 base pair depletion. Along the PC1, the PAZ and N domains underwent a departing motion away from the MID and PIWI domains, widening the central groove transversely. The motion suggested that the Ago2 protein adopted a more constrained conformation in the T9-abasic system. Compared with the WT system, the PC2 value decreased in the T9-abasic system and the Δ SL1 system. Along the PC2, the PAZ and PIWI domains moved away from each other and drove the MID and PIWI domains closer together, yielding the longitudinal opening of the central groove which accommodates the target RNA. This indicates that the disruption of the T9 and SL1 structure could fasten the mouth of the central groove, thus, protecting the target RNA from dissociation. Taking the information provided by PC1 and PC2, we can conclude that the WT T9 and SL1 structures could induce the Ago2 protein in a more extended conformation, leading to a more opening mouth of the central groove which accommodates the target RNA. The opening of the mouth of the central groove could make the target RNA liable to dissociate, which is in line with the experiment results that the T9 or SL1 could decrease the Ago2 affinity to the target RNA.

To further prove the conformational transition that was induced by the T9 and SL1 structure, we projected all of the MD trajectories onto the two-dimensional landscape according to the distances from the E333 C α atom to the S610 C α atom and from the E333 C α atom to the P63 C α atom (Figure 4A–C). $d_{E333-S610}$ represents the distance between the PAZ and PIWI domains, while $d_{E333-P63}$ represents the distance between the PAZ and N domains (Supplementary Figure S3). Thus, the $d_{E333-S610}$ and $d_{E333-P63}$ can reflect the opening of the central groove. The Markov state model (MSM) was used to differentiate the characterized clusters in the free-energy landscape (see in Material and Methods). To testify the Markovian property, the implied timescale test and the Chapman Kolmogorov test were applied (Supplementary Figure S4). The proportion and the representative structures of each state were calculated based on the MSM.

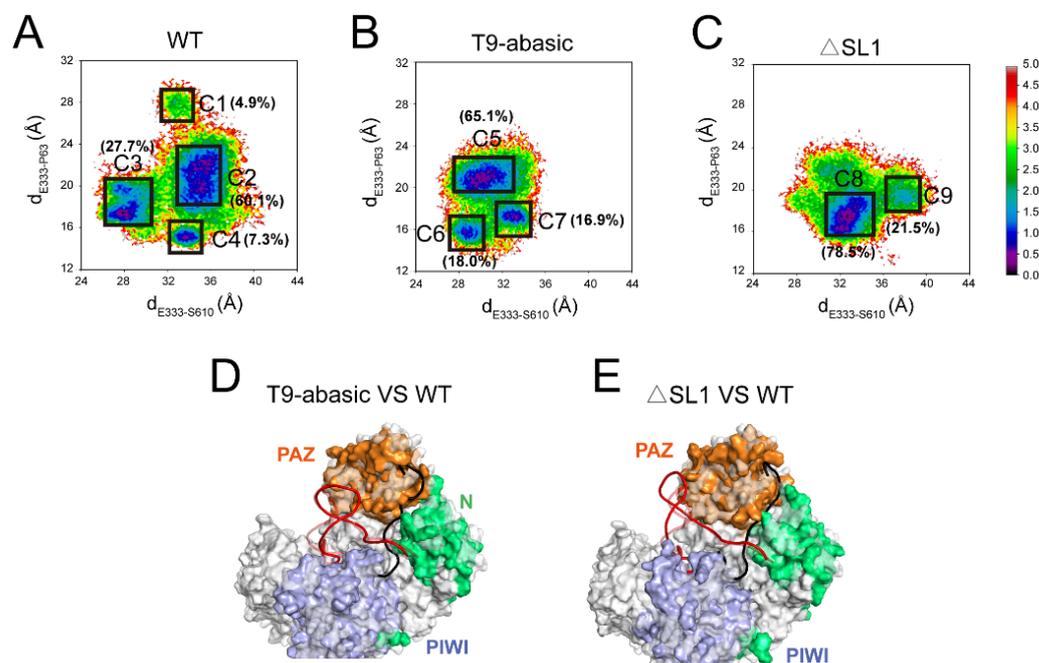


Figure 4. The Ago2 protein adopted a more constrained conformation in the T9-abasic and Δ SL1 systems. (A–C) Conformational landscapes which were generated using the $d_{E333-S610}$ (distance from the E333 C α atom to the P610 C α atom) and $d_{E333-P63}$ (distance from the E333 C α atom to the P63 C α atom) parameters in the WT system (A), the T9-abasic system (B) and the Δ SL1 system (C). Superposition of the dominant conformations from the T9-abasic system (D) and the Δ SL1 system (E) to the WT system (transparency), respectively.

Four states were observed in the WT system with the decrease in the $d_{E333-P63}$ parameter: C1, C2, C3 and C4, and the C2 state was the dominant conformation with it having a proportion of 60.1%. In the T9-abasic system, three clusters (C5, C6 and C7) were observed with the dominant conformation of C5. In the Δ SL1 system, with the increase in the $d_{E333-S610}$ parameter, the dominant C8 state and the minor C9 state were observed. Compared with the WT systems, the states in the T9-abasic system had relatively lower $d_{E333-S610}$ and $d_{E333-P63}$ parameters, indicating that there was a shorter distance between the PAZ, N and PIWI domains. The $d_{E333-P63}$ parameter was reduced in the Δ SL1 system when it was compared with the WT system, reflecting a closer relationship between the PAZ and N domains, which was in accordance with the motion along PC2. In general, the landscape proved the conformational change that was revealed in the PCA analysis and suggested a more proximal relationship between the PAZ, N and PIWI domains with the disruption of the T9 and SL1 structure.

To further illustrate the conformational transition, we superimposed the representative structures of the dominant state in the T9-abasic system (C5) and the dominant state in the

Δ SL1 system (C8) with the dominant state in the WT system (C2), respectively (Figure 4D,E). Though with the mutation to disrupt the SL1 structure in the Δ SL1 system, a stem-loop like structure was preserved during the simulation. This was ascribed to the limited simulation time. To better reflect the conformational changes, the structure of the WT was set to it being transparent. An inspection of the superimposition showed that the proximal of the PAZ, N and PIWI domains that were induced by the disruption of the T9 and SL1 structure caused the shrinkage of the mouth of the central groove, which fastened the target RNA-binding pocket. Taking the data together, the analysis suggested that the WT T9 and SL1 structures could induce a distant relationship between the PAZ, N and PIWI domains, propping up the mouth of the central groove which acts as the target RNA-binding pocket, which in turn makes the target RNA liable to dissociate.

3.3. The Disruption of T9 and SL1 Disturbs the Correlated Motions and Community Networks of the Ago2–RNA Complexes

Using a generalized correlation analysis, an overview of the inter-residue and inter-domain-correlated motions was provided (see in Material and Methods). The results showed that the inter-residue correlation of regions A and B were significantly strengthened with the disruption of the T9 and SL1 (Figure 5A–C). The region a represents the interaction between the residues which are from the PIWI domain and close to the central groove and the residues from the PAZ and MID domains, and thus, its reinforcement suggested that the PIWI domain having a more correlated motion with the PAZ and MID domains is induced by the disruption of the T9 and SL1 structure. Similarly, the enhancement in the region b reflected a more coordinated motion between the residues close to the central groove from the PAZ, PIWI and MID domains. In accordance with the inter-residue correlation, the investigations on the interdependent motions between the different domains also proved that the correlated motion between the PAZ, PIWI and MID domains was strengthened in the T9-abasic system when it was compared with the WT system (Figure 5D,E). However, the variation between the Δ SL1 system and the WT system was not obvious (Figure 5F), which may be related to the confounding influence introduced by the residues that were far from the central groove in each domain.

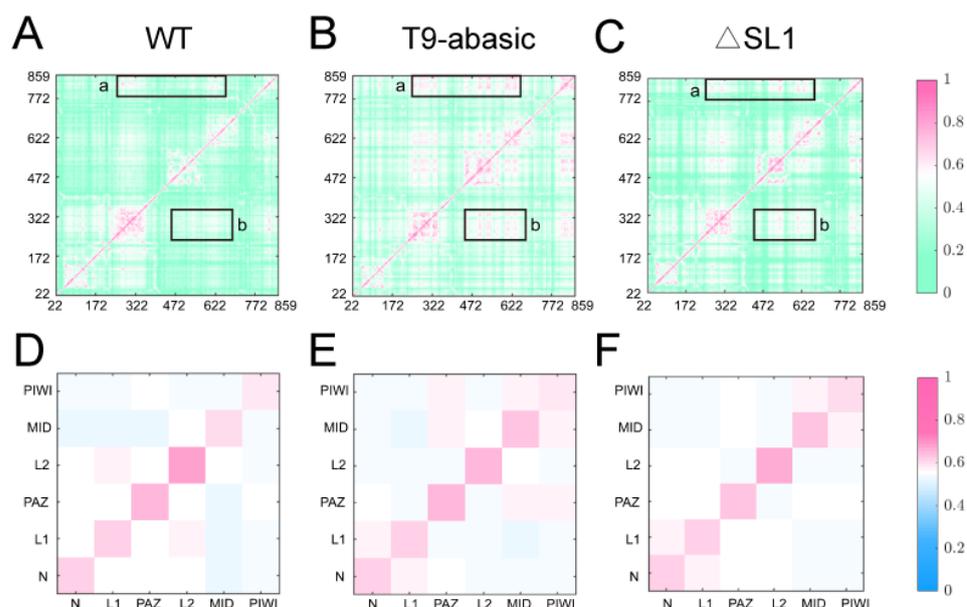


Figure 5. Inter-residue and inter-domain correlation of the Ago2 protein. Generalized residue correlation of the WT system (A), the T9-abasic system (B) and the Δ SL1 system (C). The significant correlation differences are highlighted with a rectangle in each panel. Generalized domain correlation of the WT system (D), the T9-abasic system (E) and the Δ SL1 system (F).

A community network analysis was performed to further uncover the network organization of the Ago2–RNA complex. Each community is shown as circles whose area is proportional to the number of residues that it contains. The strength of the inter-community correlation is represented by the width of the sticks connecting the communities. The community composition and connections are shown in Figure 6. For the constitution of the communities, the SL1 structure exists as a separate community in the WT system, while in the T9-abasic and Δ SL1 system, it belongs to the same community with the PAZ domain, indicating a more correlated relationship between the SL1 and PAZ domain in the T9-abasic and Δ SL1 system, which may strengthen the interaction between the target RNA and Ago2 protein. For the correlation between the communities, the connection between Community L (mainly composed of the PAZ domain) and Community M (mainly composed of the N domain) was intensified in the T9-abasic and the Δ SL1 systems when it was compared with that of the WT system, suggesting a more coordinated motion between the PAZ and N domains which was induced by the disruption of T9 and SL1.

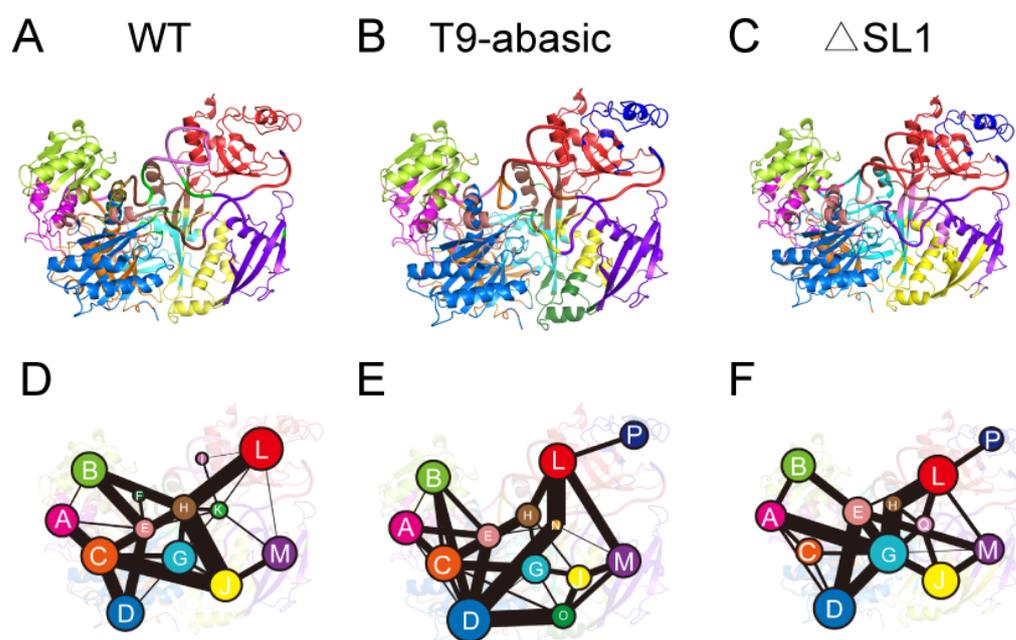


Figure 6. Community network of the Ago–RNA complex. Community composition in the WT system (A), the T9-abasic system (B) and the Δ SL1 system (C). The community network in the WT system (D), the T9-abasic system (E), and the Δ SL1 system (F). Areas of the circles represent the numbers of residues in corresponding communities, and the widths of sticks connecting communities represent the intercommunity connections, respectively.

In conclusion, the disruption of the T9 and SL1 structure induced a more correlated motion between the residues, nucleotides and domains (PAZ, PIWI and N) around the mouth of the central groove, which may be the cause of the shrinkage of the target RNA-binding pocket which was induced by the disruption of the T9 and SL1 structure.

4. Discussion

The Ago2 protein, as the core of RNA-induced silencing complex, plays a critical role in regulating gene expression and defending against exogenous pathogens. However, HCV has evolved a special mechanism which utilizes the complementary sites at the 5' end of the genome to recruit the Ago2–miRNA complex to defend against the enzyme attacks. Here, using large scale MD simulations, we investigated how the special structures, T9 and SL1, in the site 1 HCV RNA affect the target RNA recognition at the atomic level, thereby providing guidance for the development of anti-HCV strategies and the engineering of Ago2 as a gene regulation tool.

The T9 and SL1 could promote the Ago2 protein to adopt a more extended conformation. The PCA analysis and MSM analysis showed that the T9 and SL1 induce the departing motion of the PAZ, PIWI and N domains, propping up the mouth of the central groove which accommodates the target RNA, which may in turn make the target RNA liable to dissociate. Interestingly, we previously found a similar conformational change in another Ago2–RNA complex. To explore the underlying mechanism of how the miRNA extensions increased the target RNA affinity [51], we performed a series of MD simulations on the Ago2: miRNA-122: target RNA complex [52]. The simulation showed that the extension of the miRNA lengths could promote the proximity of the PAZ, PIWI and N domains to tighten the mouth of central groove, which in turn protected the target RNA from water solvent attacking and hindered the target RNAs from escaping from the Ago2–RNA complexes. Moreover, the sum of hydrogen bonds formed between target RNA and the Ago2: miRNA-122 complex was similar with different miRNA lengths. Therefore, we proved that the conformational change, rather than the enhanced interaction is the key factor for the increased target affinity with the miRNA extensions. In this study, we also calculated the number of hydrogen bonds of the target RNA to Ago2 and miRNA in the three systems (Supplementary Figure S5), respectively. In agreement with the results in the previous study, no significant difference was observed between the three systems in addition to the slightly increased hydrogen bonds between the target RNA and the protein in the Δ SL1 system which might be ascribed to the disruption of the inner hydrogen bonds. All of these results indicate that the conformation Ago2 which was adopted played a significant role in catching the target RNA. A more constricted Ago2 conformation which means the shrinkage of the mouth of the central groove would benefit the target RNA binding.

The T9 and SL1 could affect the dynamics and correlated motion of the Ago2–RNA complexes. The RMSF analysis revealed that the SL1 region, PAZ and PIWI domains displayed a higher fluctuation in the WT system when they were compared with those of the T9-abasic and Δ SL1 systems, implying that the instability was induced by the T9 and SL1 structures. According to the GCCM and network analysis, the T9 and SL1 structures also weakened the correlated motion between the PAZ, PIWI and N domains which encompassed the mouth of the central groove. This may be related to the extended conformation which was induced by the SL1 and T9-abasic system.

The modification of the Ago proteins into gene regulation and even editing tools has been a research hotspot [53–55]. However, the structural basis for the recognition and assembly of the Ago–nucleic acid complexes needs to be further investigated for the development of new engineering strategies [56]. Based on the previous and current studies, we proposed a new model that the Ago2 protein could adopt a more constraint conformation with the increased proximity to and more correlated motion of the PAZ, N and PIWI domains to protect the target RNA from dissociation. Thus, the modification of the residues to intensify the interaction between the PAZ, PIWI and N domains may be a feasible strategy to increase Ago's binding affinity.

HCV infection is liable to become chronic due to it having multiple mechanisms of resistance to clearance [57], one of which is the recruitment of Ago–miRNA complexes through the site 1 sequence to protect the terminal of its genome. Our study elucidates the role of the site 1 key structures T9 and SL1 in the conformational dynamics and assembly of the Ago2: miRNA-122: HCV site 1 complex, paving the way for the development of new anti-HCV therapies through the miR-122 pathway.

In summary, the collective sampling of the 15 μ s MD simulations showed that the T9 and SL1 structure of the HCV site 1 RNA could induce the departing motion and the weakened correlation between the PAZ, PIWI and N domains, propping up the mouth of the central groove, which in turn causes the instability of the target RNA. This was in line with the experimental results that the T9 and SL1 structures were detrimental to Ago2 binding.

Moreover, we proposed a new model that the Ago2 protein could adopt a more constrained conformation with the increased proximity to and more correlated motion

of the PAZ, N and PIWI domains to protect the target RNA from dissociation. Thus, the modification of the residues to intensify the interaction between the PAZ, PIWI and N domains may be a feasible strategy to increase Ago's binding affinity. These results shed light on the structural details of the assembly of the Ago–RNA complexes, which will help to offer guidance for the development of anti-HCV strategies and the engineering of Ago2 as a gene regulation tool.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12111631/s1>, Figure S1: Time evolution of the root-mean-square deviation (RMSD) of Ago2; Figure S2: Root-mean-square fluctuations (RMSF) of Ago2-RNA complex; Figure S3: The representation of dE333–V610 and dE333–P63 in the energy landscape; Figure S4: The MSM timescale and Chapman Kolmogorov tests; Figure S5: The number of hydrogen bonds of the target RNA with the Ago2 protein and miRNA.

Author Contributions: S.L., X.F. and Z.C. conceived and supervised the project, designed the experiments, and edited the manuscript; H.Z. performed MD simulation and analysis and drafted the manuscript. H.Z., D.J., J.F., M.L., R.T. and K.D. helped the construction of simulated systems and the analysis of MD trajectories; H.Z. acquired the data and revised the manuscript; Z.C. and X.F. were responsible for funding support. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 82172846).

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

References

1. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. *Cell* **2009**, *136*, 215–233. [[CrossRef](#)] [[PubMed](#)]
2. Hammond, S.M. An overview of microRNAs. *Adv. Drug Deliv. Rev.* **2015**, *87*, 3–14. [[CrossRef](#)] [[PubMed](#)]
3. Lu, T.X.; Rothenberg, M.E. MicroRNA. *J. Allergy Clin. Immunol.* **2018**, *141*, 1202–1207. [[CrossRef](#)] [[PubMed](#)]
4. Gebert, L.F.R.; MacRae, I.J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 21–37. [[CrossRef](#)] [[PubMed](#)]
5. Jopling, C.L.; Yi, M.; Lancaster, A.M.; Lemon, S.M.; Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **2005**, *309*, 1577–1581. [[CrossRef](#)] [[PubMed](#)]
6. Jopling, C.L.; Schütz, S.; Sarnow, P. Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. *Cell Host Microbe* **2008**, *4*, 77–85. [[CrossRef](#)] [[PubMed](#)]
7. Gower, E.; Estes, C.; Blach, S.; Razavi-Shearer, K.; Razavi, H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J. Hepatol.* **2014**, *61* (Suppl. 1), S45–S57. [[CrossRef](#)]
8. Lanford, R.E.; Hildebrandt-Eriksen, E.S.; Petri, A.; Persson, R.; Lindow, M.; Munk, M.E.; Kauppinen, S.; Ørum, H. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* **2010**, *327*, 198–201. [[CrossRef](#)] [[PubMed](#)]
9. Janssen, H.L.A.; Reesink, H.W.; Lawitz, E.J.; Zeuzem, S.; Rodriguez-Torres, M.; Patel, K.; van der Meer, A.J.; Patick, A.K.; Chen, A.; Zhou, Y.; et al. Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **2013**, *368*, 1685–1694. [[CrossRef](#)]
10. Ottosen, S.; Parsley, T.B.; Yang, L.; Zeh, K.; van Doorn, L.-J.; van der Veer, E.; Raney, A.K.; Hodges, M.R.; Patick, A.K. In vitro antiviral activity and preclinical and clinical resistance profile of miravirsin, a novel anti-hepatitis C virus therapeutic targeting the human factor miR-122. *Antimicrob. Agents Chemother.* **2015**, *59*, 599–608. [[CrossRef](#)] [[PubMed](#)]
11. Luna, J.M.; Scheel, T.K.H.; Danino, T.; Shaw, K.S.; Mele, A.; Fak, J.J.; Nishiuchi, E.; Takacs, C.N.; Catanese, M.T.; de Jong, Y.P.; et al. Hepatitis C virus RNA functionally sequesters miR-122. *Cell* **2015**, *160*, 1099–1110. [[CrossRef](#)] [[PubMed](#)]
12. Amador-Cañizares, Y.; Bernier, A.; Wilson, J.A.; Sagan, S.M. miR-122 does not impact recognition of the HCV genome by innate sensors of RNA but rather protects the 5' end from the cellular pyrophosphatases, DOM3Z and DUSP11. *Nucleic Acids Res.* **2018**, *46*, 5139–5158. [[CrossRef](#)] [[PubMed](#)]
13. Schirle, N.T.; MacRae, I.J. The crystal structure of human Argonaute2. *Science* **2012**, *336*, 1037–1040. [[CrossRef](#)] [[PubMed](#)]
14. Gebert, L.F.R.; Law, M.; MacRae, I.J. A structured RNA motif locks Argonaute2:miR-122 onto the 5' end of the HCV genome. *Nat. Commun.* **2021**, *12*, 6836. [[CrossRef](#)]
15. Kuhn, C.-D.; Joshua-Tor, L. Eukaryotic Argonautes come into focus. *Trends Biochem. Sci.* **2013**, *38*, 263–271. [[CrossRef](#)]
16. Hausser, J.; Zavolan, M. Identification and consequences of miRNA–target interactions—Beyond repression of gene expression. *Nat. Rev. Genet.* **2014**, *15*, 599–612. [[CrossRef](#)]
17. Schirle, N.T.; Sheu-Gruttadauria, J.; Macrae, I.J. Structural basis for microRNA targeting. *Science* **2014**, *346*, 608–613. [[CrossRef](#)]
18. Li, M.; Wang, Y.; Fan, J.; Zhuang, H.; Liu, Y.; Ji, D.; Lu, S. Mechanistic Insights into the Long-range Allosteric Regulation of KRAS Via Neurofibromatosis Type 1 (NF1) Scaffold Upon SPRED1 Loading. *J. Mol. Biol.* **2022**, *434*, 167730. [[CrossRef](#)]

19. Zhang, H.; Ni, D.; Fan, J.; Li, M.; Zhang, J.; Hua, C.; Nussinov, R.; Lu, S. Markov State Models and Molecular Dynamics Simulations Reveal the Conformational Transition of the Intrinsically Disordered Hypervariable Region of K-Ras4B to the Ordered Conformation. *J. Chem. Inf. Model.* **2022**, *62*, 4222–4231. [[CrossRef](#)]
20. Shi, Y.; Cao, S.; Ni, D.; Fan, J.; Lu, S.; Xue, M. The Role of Conformational Dynamics and Allostery in the Control of Distinct Efficacies of Agonists to the Glucocorticoid Receptor. *Front. Mol. Biosci.* **2022**, *9*, 933676. [[CrossRef](#)]
21. Lu, S.; Chen, Y.; Wei, J.; Zhao, M.; Ni, D.; He, X.; Zhang, J. Mechanism of allosteric activation of SIRT6 revealed by the action of rationally designed activators. *Acta Pharm. Sin. B* **2021**, *11*, 1355–1361. [[CrossRef](#)] [[PubMed](#)]
22. Jang, H.; Zhang, M.; Nussinov, R. The quaternary assembly of KRas4B with Raf-1 at the membrane. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 737–748. [[CrossRef](#)] [[PubMed](#)]
23. An, X.; Bai, Q.; Bing, Z.; Liu, H.; Yao, X. Insights into the molecular mechanism of positive cooperativity between partial agonist MK-8666 and full allosteric agonist AP8 of hGPR40 by Gaussian accelerated molecular dynamics (GaMD) simulations. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 3978–3989. [[CrossRef](#)] [[PubMed](#)]
24. Maloney, R.C.; Zhang, M.; Jang, H.; Nussinov, R. The mechanism of activation of monomeric B-Raf V600E. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 3349–3363. [[CrossRef](#)] [[PubMed](#)]
25. Shibata, T.; Iwasaki, W.; Hirota, K. The intrinsic ability of double-stranded DNA to carry out D-loop and R-loop formation. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 3350–3360. [[CrossRef](#)]
26. Zhang, H.; Zhu, M.; Li, M.; Ni, D.; Wang, Y.; Deng, L.; Du, K.; Lu, S.; Shi, H.; Cai, C. Mechanistic Insights Into Co-Administration of Allosteric and Orthosteric Drugs to Overcome Drug-Resistance in T315I BCR-ABL1. *Front. Pharmacol.* **2022**, *13*, 862504. [[CrossRef](#)]
27. Fan, J.; Liu, Y.; Kong, R.; Ni, D.; Yu, Z.; Lu, S.; Zhang, J. Harnessing Reversed Allosteric Communication: A Novel Strategy for Allosteric Drug Discovery. *J. Med. Chem.* **2021**, *64*, 17728–17743. [[CrossRef](#)]
28. Wang, Y.; Ji, D.; Lei, C.; Chen, Y.; Qiu, Y.; Li, X.; Li, M.; Ni, D.; Pu, J.; Zhang, J.; et al. Mechanistic insights into the effect of phosphorylation on Ras conformational dynamics and its interactions with cell signaling proteins. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1184–1199. [[CrossRef](#)]
29. Li, X.; Wang, C.; Peng, T.; Chai, Z.; Ni, D.; Liu, Y.; Zhang, J.; Chen, T.; Lu, S. Atomic-scale insights into allosteric inhibition and evolutionary rescue mechanism of *Streptococcus thermophilus* Cas9 by the anti-CRISPR protein AcrIIA6. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 6108–6124. [[CrossRef](#)]
30. Li, X.; Dai, J.; Ni, D.; He, X.; Zhang, H.; Zhang, J.; Fu, Q.; Liu, Y.; Lu, S. Insight into the mechanism of allosteric activation of PI3Kalpha by oncoprotein K-Ras4B. *Int. J. Biol. Macromol.* **2020**, *144*, 643–655. [[CrossRef](#)]
31. Lei, J.; Sheng, G.; Cheung, P.P.; Wang, S.; Li, Y.; Gao, X.; Zhang, Y.; Wang, Y.; Huang, X. Two symmetric arginine residues play distinct roles in *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 845–853. [[CrossRef](#)] [[PubMed](#)]
32. Jiang, H.; Sheong, F.K.; Zhu, L.; Gao, X.; Bernauer, J.; Huang, X. Markov State Models Reveal a Two-Step Mechanism of miRNA Loading into the Human Argonaute Protein: Selective Binding followed by Structural Re-arrangement. *PLoS Comput. Biol.* **2015**, *11*, e1004404. [[CrossRef](#)] [[PubMed](#)]
33. Effects of the PIWI/MID domain of Argonaute protein on the association of miRNAi's seed base with the target. *Rna* **2019**, *25*, 620–629. [[CrossRef](#)] [[PubMed](#)]
34. Schirle, N.T.; Sheu-Gruttadauria, J.; Chandradoss, S.D.; Joo, C.; MacRae, I.J. Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *Elife* **2015**, *4*, e07646. [[CrossRef](#)] [[PubMed](#)]
35. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)] [[PubMed](#)]
36. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T.E.; Lughton, C.A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophys. J.* **2007**, *92*, 3817–3829. [[CrossRef](#)] [[PubMed](#)]
37. Zgarbová, M.; Otyepka, M.; Sponer, J.; Mládek, A.; Banáš, P.; Cheatham, T.E.; Jurečka, P. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902. [[CrossRef](#)]
38. Aytenfisu, A.H.; Spasic, A.; Grossfield, A.; Stern, H.A.; Mathews, D.H. Revised RNA Dihedral Parameters for the Amber Force Field Improve RNA Molecular Dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 900–915. [[CrossRef](#)]
39. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
40. York, D.M.; Wlodawer, A.; Pedersen, L.G.; Darden, T.A. Atomic-level accuracy in simulations of large protein crystals. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8715–8718. [[CrossRef](#)]
41. Ali Hassan, N.Z.; Mokhtar, N.M.; Kok Sin, T.; Mohamed Rose, I.; Sagap, I.; Harun, R.; Jamal, R. Integrated analysis of copy number variation and genome-wide expression profiling in colorectal cancer tissues. *PLoS ONE* **2014**, *9*, e92553. [[CrossRef](#)]
42. Daidone, I.; Amadei, A. Essential dynamics: Foundation and applications. *WIREs Comput. Mol. Sci.* **2012**, *2*, 762–770. [[CrossRef](#)]
43. Lange, O.F.; Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins* **2006**, *62*, 1053–1061. [[CrossRef](#)] [[PubMed](#)]
44. Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* **2001**, *7*, 306–317. [[CrossRef](#)]

45. Eargle, J.; Luthey-Schulten, Z. NetworkView: 3D display and analysis of protein-RNA interaction networks. *Bioinformatics* **2012**, *28*, 3000–3001. [[CrossRef](#)]
46. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [[CrossRef](#)]
47. Husic, B.E.; Pande, V.S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396. [[CrossRef](#)]
48. Lu, S.; Ni, D.; Wang, C.; He, X.; Lin, H.; Wang, Z.; Zhang, J. Deactivation Pathway of Ras GTPase Underlies Conformational Substates as Targets for Drug Design. *ACS Catal.* **2019**, *9*, 7188–7196. [[CrossRef](#)]
49. Lu, S.; He, X.; Yang, Z.; Chai, Z.; Zhou, S.; Wang, J.; Rehman, A.U.; Ni, D.; Pu, J.; Sun, J.; et al. Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design. *Nat. Commun.* **2021**, *12*, 4721. [[CrossRef](#)]
50. Ni, D.; Wei, J.; He, X.; Rehman, A.U.; Li, X.; Qiu, Y.; Pu, J.; Lu, S.; Zhang, J. Discovery of cryptic allosteric sites using reversed allosteric communication by a combined computational and experimental strategy. *Chem. Sci.* **2021**, *12*, 464–476. [[CrossRef](#)]
51. Sheu-Gruttadauria, J.; Xiao, Y.; Gebert, L.F.; MacRae, I.J. Beyond the seed: Structural basis for supplementary microRNA targeting by human Argonaute2. *EMBO J.* **2019**, *38*, e101153. [[CrossRef](#)] [[PubMed](#)]
52. Zhuang, H.; Fan, X.; Ji, D.; Wang, Y.; Fan, J.; Li, M.; Ni, D.; Lu, S.; Li, X.; Chai, Z. Elucidation of the conformational dynamics and assembly of Argonaute-RNA complexes by distinct yet coordinated actions of the supplementary microRNA. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 1352–1365. [[CrossRef](#)] [[PubMed](#)]
53. Hegge, J.W.; Swarts, D.C.; Chandradoss, S.D.; Cui, T.J.; Kneppers, J.; Jinek, M.; Joo, C.; van der Oost, J. DNA-guided DNA cleavage at moderate temperatures by *Clostridium butyricum* Argonaute. *Nucleic Acids Res.* **2019**, *47*, 5809–5821. [[CrossRef](#)] [[PubMed](#)]
54. Hegge, J.W.; Swarts, D.C.; van der Oost, J. Prokaryotic Argonaute proteins: Novel genome-editing tools? *Nat. Rev. Microbiol.* **2018**, *16*, 5–11. [[CrossRef](#)]
55. Cyranoski, D. Replications, ridicule and a recluse: The controversy over NgAgo gene-editing intensifies. *Nature* **2016**, *536*, 136–137. [[CrossRef](#)]
56. Wu, J.; Yang, J.; Cho, W.C.; Zheng, Y. Argonaute proteins: Structural features, functions and emerging roles. *J. Adv. Res.* **2020**, *24*, 317–324. [[CrossRef](#)]
57. Manns, M.P.; Buti, M.; Gane, E.; Pawlotsky, J.-M.; Razavi, H.; Terrault, N.; Younossi, Z. Hepatitis C virus infection. *Nat. Rev. Dis. Prim.* **2017**, *3*, 17006. [[CrossRef](#)]