JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

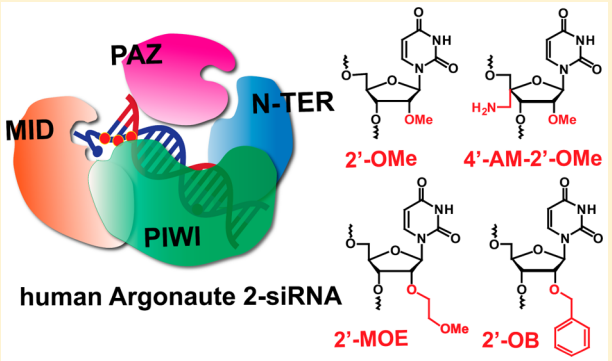
Probing the Binding Interactions between Chemically Modified siRNAs and Human Argonaute 2 Using Microsecond Molecular Dynamics Simulations

S. Harikrishna* and P. I. Pradeepkumar*

Department of Chemistry, Indian Institute of Technology Bombay, Mumbai−400076, India

Supporting Information

ABSTRACT: The use of chemical modifications in small interfering RNAs (siRNAs) is warranted to impart drug-like properties. However, certain chemical modifications especially those on the sugar have deleterious effects on the RNA interference (RNAi) when they are placed at key positions in the seed region of an siRNA guide strand. In order to probe the effect of chemically modified siRNAs [(2′-*O*-methyl, 4′-*C*-aminomethyl-2′-*O*-methyl, 2′-*O*-(2-methoxyethyl), and 2′-*O*-benzyl] on human Argonaute 2 (hAGO2), the catalytic engine of RNAi, we have developed a model of its open conformation. Results from microsecond MD simulations of 15 different siRNA−hAGO2 complexes provide insights about how the key noncovalent interactions and conformational changes at the seed region are modulated, depending upon the nature and position of chemical modifications. Such modification induced structural changes can affect siRNA loading into hAGO2, which may influence RNAi activity. Our studies show that microsecond MD simulations can provide useful information for the design of therapeutically relevant siRNAs.



INTRODUCTION

Small interfering RNAs (siRNAs) comprise 21−22 nucleotides (nt) of which 19 nt are base paired to form a duplex with 2 nt overhangs present at the 3′-ends of both the strands.[1] The key endonuclease protein, human argonaute 2 (hAGO2), in the RNA induced silencing complex (RISC), cleaves target mRNA that has base-complementarity to the sequence of the guide strand of siRNA. This leads to gene silencing termed as RNA interference (RNAi).[2−5] Therefore, siRNAs possess high therapeutic potentials; however, the canonical RNAs have several limitations that prohibit their use as drugs.[6] Chemical modifications have been introduced into siRNAs to enhance nuclease resistance, improve RNAi potency, reduce off-target effects, and to facilitate intracellular delivery.[7−9] However, it has been observed that the effect of chemical modifications (Figure 1) on RNAi is position-dependent, especially at the seed region (position 2−8 from 5′-end) of the siRNA guide strand.[9−11] For example, the 2′-deoxy-2′-fluoro (2′-F) and 2′-*O*-methyl (2′-OMe) modifications are well tolerated at most of the positions in the guide strand,[11−13] whereas the bulkier 2′-*O*-(2-methoxyethyl) (2′-MOE) modification is not tolerated in the seed region.[11] siRNAs containing 2′-*O*-benzyl (2′-OB) modification in the guide positions g5, g8, g15, and g19 enhances the RNAi activity compared to the unmodified counterpart.[14] On the other hand, the 2′-OB group at position g2 is detrimental to the RNAi activity.[15] The bifunctional 4′-*C*-aminomethyl-2′-*O*-methyl (4′-AM-2′-OMe) modification at position g3 affects the gene silencing activity.[16] A nucleotide walk-through (replacing every nucleotide in the siRNA) that is required to identify the optimal positions for introducing modifications is cost prohibitive and labor intensive. Therefore, a structure based rationalization on the position-dependent tolerance of chemical modifications utilizing a model of siRNA−hAGO2 complex could help in the design of therapeutically useful siRNAs. In this direction, recent reports highlighted the use of computational approaches to design chemically modified siRNAs.[17−19]

One of the key steps involved in the RNAi is the loading of siRNA into hAGO2, where, the passenger strand in the siRNA has the same sequence as the target mRNA. Thus, the siRNA−hAGO2 binary complex is the mimic of hAGO2-guide strand/target mRNA ternary complex.[2,20] Structural studies on the archaeal, yeast and human AGOs shed light on the domain architecture, siRNA binding, conformational changes associated with guide strand loading, target mRNA binding as well as the mechanism of target cleavage.[21−29] The AGO family in both prokaryotes and eukaryotes shares a similar bilobed four domain architecture (N-terminal, PAZ, MID, and PIWI along with two linkers L1 and L2).[20,22,29,30] The MID domain is involved in the recognition and binding of 5′-phosphorylated end of the guide strand.[31] The PIWI domain is responsible for the cleavage of passenger strand/target mRNA.[32] PAZ domain binds and

Received: December 19, 2016
Published: March 13, 2017

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

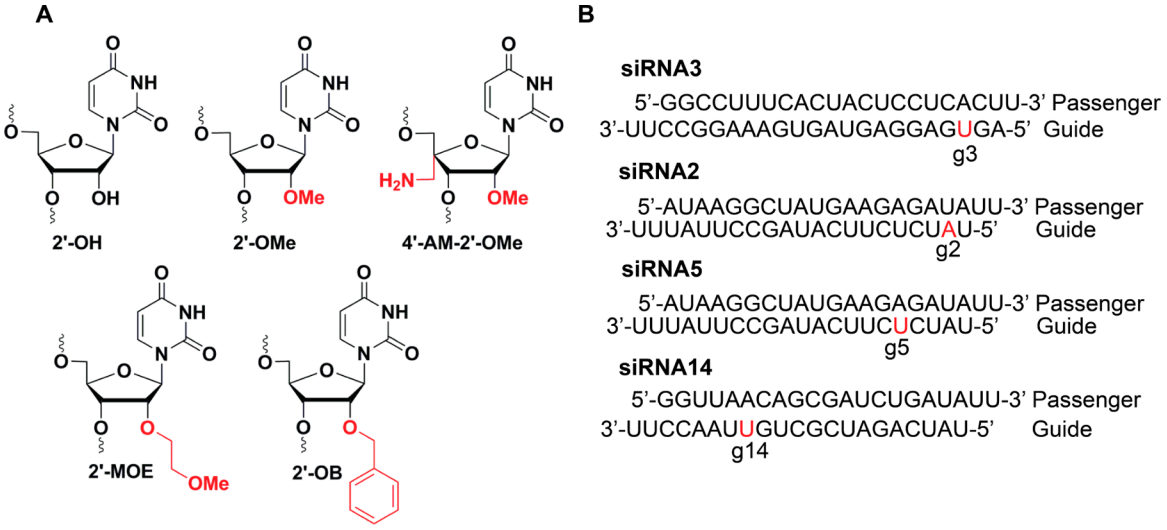


Figure 1. Chemical structure of sugar modifications and siRNA sequences used in our MD simulations study. (**A**) Structure of canonical (2′-OH) and 2′-sugar modifications: 2′-*O*-methyl (2′-OMe), 4′-*C*-aminomethyl-2′-*O*-methyl (4′-AM-2′-OMe), 2′-*O*-(2-methoxyethyl) (2′-MOE), and 2′-*O*-benzyl (2′-OB) uridine. (**B**) siRNA sequences in which the modifications are present in the guide strand positions: siRNA3 (position g3), siRNA2 (position g2), siRNA5 (position g5), and siRNA14 (position g14). The modified positions are highlighted in red color. The g1 nt in the siRNA3 guide strand starts base pairing with a 3′-overhang nt of the passenger strand; therefore, the numbering of position is different from the typical siRNA. All the siRNA sequences employed in the MD simulations are listed in Table S1 of the Supporting Information.

releases the 3′-end of the guide strand after the cleavage of passenger strand or target mRNA.[33] The N-terminal domain is required for the unwinding of siRNA duplex at the 5′-end of the passenger strand.[34] The noncovalent contacts of the siRNA guide strand (sugar−phosphate backbone) with hAGO2 domains are critical for the activity of this enzyme.[21,30] A comprehensive siRNA−hAGO2 molecular dynamics (MD) simulations in microsecond time scale is required to probe the structural effects of various chemical modifications on hAGO2 binding. Only a few MD studies have been reported so far on the siRNA−*Thermus thermophilus* argonaute (*Tt*AGO) and siRNA−hAGO2 complexes.[35−39] But none of these studies investigated the role of chemically modified siRNAs in RNA−hAGO2 binding. While this manuscript was under preparation, a crystal structure of siRNA guide strand alternatively modified with 2′-OMe and 2′-F modification in complex with hAGO2 was reported.[40] This structure shows that these modifications do not affect the noncovalent interactions between siRNA and hAGO2 severely. It should be noted here that 2′-OMe and 2′-F are less bulky modifications, and they are reported to be tolerated throughout the siRNA.[11,12] Another crystal structure of the siRNA guide strand containing an unnatural triazolyl modification at the g1 position in complex with hAGO2 was also reported.[41] This structure revealed that the modification interacts strongly with the 5′-phosphorylated binding site and controls the base pairing of the seed region and thereby provides target specificity. Though these structures provide useful insights on the effects of chemically modified guide strand binding to hAGO2, structures of modified siRNA duplex in complex with hAGO2 are yet to be reported.

The crystal structure reported for the hAGO2 with unmodified miRNA−target mRNA duplex reveals only 8 nt base pairs.[22] Therefore, modeling an open conformation of hAGO2 is required to study the binding interactions of the hAGO2 and siRNA having 19 base pairs of guide/target complex.[23,37,39] Since the hAGO2 extensively makes contacts with the sugar−phosphate backbone of the siRNA in particular at the seed region[22] and the presence of chemical modifications at these region is sensitive toward RNAi, we studied the effect of sugar modifications located at those sites. Modifications (2′ and 2′-4′ dual) and positions in the guide strand of the siRNA (g2, g3, g5 and g14) were chosen (Figure 1) based upon their reported effect on RNAi activity as revealed by experiments.[11,14−16,42] Recent achievements in the molecular dynamics (MD) simulations of large biomolecules[43,44] and force field developments[45−49] permit us to study the time dependent conformational changes in the siRNA−hAGO2 complex at a microsecond scale.[50−52] Here, we report the first MD simulations of hAGO2 in the presence of canonical siRNAs and a number of chemically modified siRNAs (Figure 1) in microsecond time scale. The effects of chemical modifications on hAGO2 binding were probed[53] to unravel the position-dependent tolerance of chemical modifications in the guide strand of siRNAs that could assist in the design of therapeutic siRNAs

MATERIALS AND METHODS

Modeling the Open Conformation of hAGO2. The accommodation of the double stranded siRNA into the hAGO2 structure was performed through comparative modeling (PDB entries: 4NCB, 3HJF, and 3HVR)[27,29] using MODELER,[54] and the domain reorientations were performed using AIDA.[55] The missing amino acid residues in the PDB entries 4W5O, 3HJF, 4NCB, and 3HVR were added using MODELER. Nucleic acids in the *Tt*AGO structures were removed and used as a template to model the open conformation of hAGO2. The orientation of N-terminal domain is different in the case of *Tt*AGO bound guide RNA and hAGO2 bound guide RNA (Figure M1, Supporting Information). Therefore, in the present model the orientation is retained as found in the hAGO2 crystal structure. This could facilitate the unobstructed guide-target pairing in hAGO2, as reported for the *Kluyveromyces polysporus* AGO (*Kp*AGO) structure (PDB entry: 4F1N).[23] Sequence alignment of hAGO2 and *Tt*AGO was gathered from a previous report.[23] Since the loops in the PIWI domain largely hinder the binding of the target RNA, the PIWI domain was reoriented and modeled using MODELER to accommodate the guide-target base pairing

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

beyond 8 nt as suggested elsewhere.[35,56] Afterward, the orientation of the linker and PAZ domains with respect to the MID and the PIWI domains was modeled (AIDA program) using the distance restraints between the domains as seen in the *Tt*AGO open conformation (Figure M1, Supporting Information). The TM-Score (>0.8) was used to filter the final model. A similar kind of domain−domain distance increment approaches to develop the open conformation of hAGO2 has been reported earlier.[35,57] The model obtained from AIDA was subjected then to 2000 steps of conjugate and steepest descent minimization using SANDER module of AMBER 14.[58] PROCHECK was utilized to validate the model.

Molecular Docking of siRNA and hAGO2. ModeRNA[59] was used to model different sequences of siRNA (6−18 nt) based on the guide DNA templates in the *Tt*AGO crystal structure (PDB entry: 4NCB). Then the 2 nt overhangs at the 3′-end of the passenger and guide strand were removed. The backbone coordinates of the guide-target RNA from 1 to 5 nt were directly adopted from the hAGO2 crystal structure (PDB entry: 4W5O), because the conformational changes in the RNA were required beyond 6 nt.[56] The backbone of the guide RNA was aligned with the guide DNA in the *Tt*AGO, then the open conformation of hAGO2 was aligned with the *Tt*AGO to yield a superimposed primary model.[57] The guide-target base pairs were docked with hAGO2 using the Hex docking package by maintaining the default parameters.[60] From the 500 RNA conformers generated by Hex, the final conformation was selected based upon the RMSD between the reference structure and the conformers obtained (RMS cutoff ∼1.0 Å). The crystal structures of *Tt*AGO (PDB entries: 4NCB, 3HM9, and 3HVR) were used as reference for the validation of docked poses. For the unmodified, and chemically modified siRNA3 and siRNA14 (Table S1, Supporting Information), precision docking was carried out using Hex and Glide. Using the Hex generated conformer, a grid inside the protein was created manually from the backbone of a nucleotide to a distance of 4 Å. The side chains of the amino acids selected under the grid was kept flexible during docking, along with the backbone atoms in the nt using Glide (SP and XP).[61] The input partial charges were assigned using the AMBER force field.[62] The scaling cutoff for van der Waals radii was set to 0.1 Å. Final output was set as 100 pose for the single docking. All other parameters were set as default in Glide. The docked poses were filtered based upon the best docking score and maximum number of intermolecular interactions. The resulting binary siRNA−hAGO2 complex was subjected minimization to eliminate the close contacts. For the chemically modified siRNA5 and siRNA2 (Table S1, Supporting Information), precision docking was carried out using Autodock 4.2.[63] During each round, only two base pairs were docked[36] based upon the specific grid in the binding site obtained from the superimposition of the reference structure and model of hAGO2. The maximum number of generations were set as 5000 and the number of iterations in the local search were set as 1000. And the number of evaluations were set as 2 500 000. All other parameters were set as default in Autodock 4.2.

Protonation State. The protonation state of the amino acids ARG, LYS, GLU and ASP side chains were assigned using PDB 2PQR program[64] at pH 7.0 in AMBER force field. The protonation states of the HIS side chains were manually assigned by examining the structures. Protonation state of histidines utilized were as follows: (HID−δ, HIE−ε, HIP−double) HID56, HID77, HIE81, HID144, HIE 168, HID 203, HID 271, HIE 291, HIE 316, HID 324, HIE 336, HID 443, HID 466, HIE 506, HIE 600, HIE 621, HIP 634, HID 681, HIE 682, HIE 711, HID 712, HIP 742, HID 753, HID 764, HIP 766, HIE 788, HIE 807, HID 816, HID 822, HID 829, HID 839, HIP 849.

Preparation of Protein−RNA Complex. The topology and coordinate inputs of protein and RNA complex were prepared using the xleap module in AMBER 14.[58] The system was neutralized using K^+, Cl^- ions and excess ions (100 mM) were added to mimic the physiological conditions. The net charge of the system was neutral. Mg^{2+} ions at the cleavage site were parametrized using a reported procedure.[65] The conformation of 5′-phoshphate end in the guide strand is retained as in the crystal structure and the parameters were assigned using an earlier report.[46] The partial charges for the 2′-*O*-benzyl modifications was calculated at the nucleotide level using Gaussian 09 (HF/6-31G*) package.[66] The calculated charges were then fitted using the RESP algorithm (Figure S1, Supporting Information).[67] The force field parameters for 2′-OMe, 4′-AM-2′-OMe, and 2′-MOE were derived from the previous reports.[16,46,67,68] The force fields used for the RNA and the protein are bsc0χ_{OL3}[45,48,62,69] and ff12SB,[58] respectively, as suggested in a recent study.[51] Using TIP3P water molecules, the system was solvated up to 8 Å from any of the solute atoms.[70] The number of water molecules included were ∼55500.

MD Simulation Protocol. The equilibration and MD simulations were performed as reported earlier.[51,71] Briefly, the water molecules and ions in the solvated complex were minimized in 2500 steps of the steepest descent, and in 2500 steps of the conjugate gradient with 50 kcal/mol restraint on the solute atoms. Solvent molecules were relaxed using short MD simulations (250 ps) with 30 kcal/mol restraint on the solute atoms at a temperature of 100 K. The system was then heated from 100 to 300 K in 150 ps. Followed by this, 10 stages of minimization (2000 steps) and 10 stages of MD simulations (30 ps) were carried out with restraints of 20, 16, 12, 10, 6, 5, 4, 3, 2, and 1 kcal/mol on solute atoms. Afterward, a 300 ps of restrained (0.05 kcal/mol on solute atoms) MD simulations was performed. Finally, unrestrained production MD simulations were performed for 1 μs using CUDA version of pmemd[72] in a GPU accelerated version[73,74] of AMBER 14.[58] Particle mesh ewald method was utilized for calculating the contributions from the nonbonded interactions with a cutoff of 10 Å. SHAKE was used to treat the bonds involving hydrogen atoms. The unrestrained MD simulations were performed in NPT ensemble of 2 fs time step. A constant pressure of 1 atm was maintained using Bendersen weak-coupling barostat with a time constant of 1 ps.[75] The MD simulation temperature (300 K) was maintained by Bendersen thermostat with a time constant of 4 ps. MD trajectories were saved for every 10 ps and then extracted at every 100 ps time interval for further analysis.

Structural Analysis. Root mean square deviations (RMSDs) were calculated for the backbone heavy atoms of protein (CA, C, and N) and RNA (P, O5′, C5′, C4′, C3′, and O3′) using CPPTRAJ module in AMBER 14.[76] All the frames were superimposed to the initial structure after removing water molecules from the MD trajectories. Root mean square fluctuations (RMSF) of protein during MD simulations were then calculated for the heavy atoms using CPPTRAJ. MD snapshots depicting the averaged structures were calculated using CPPTRAJ. The X3DNA package was utilized to compute the RNA helical parameters and backbone dihedral angles.[77] van der Waals interaction energies between the 2′-OB modification in siRNA5 and the hydrophobic binding sites in the protein, and electrostatic energies between 4′-AM-2′-OMe and the AAs in the

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

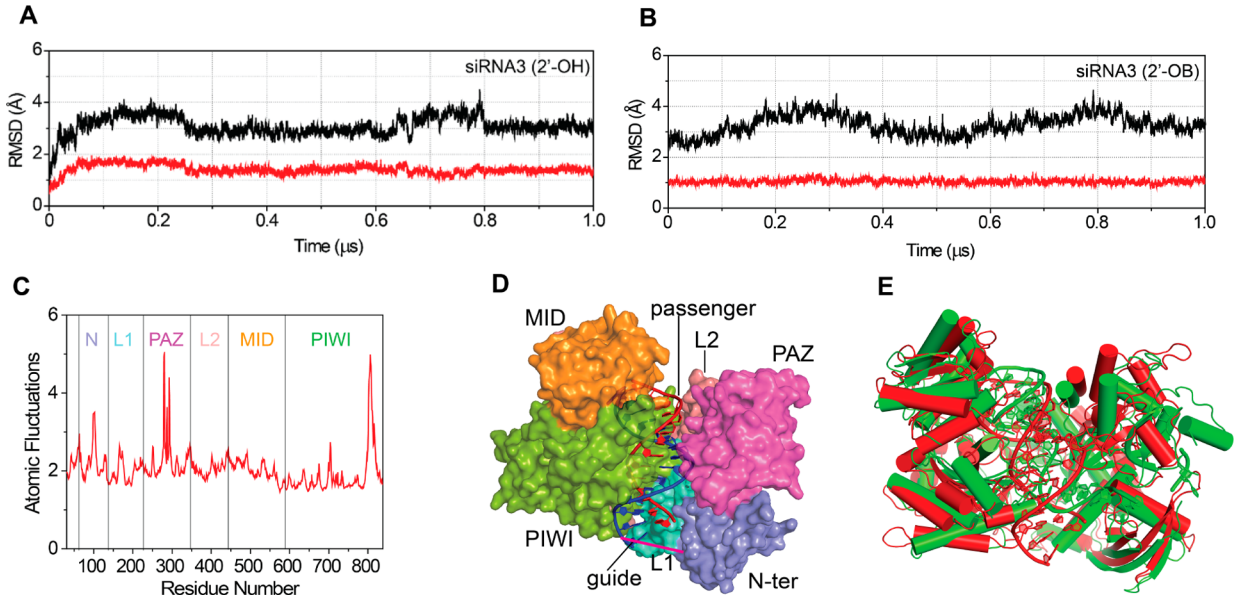


Figure 2. RMSD graphs, RMSF plot, and model of the siRNA3−hAGO2 binary complex, in which the siRNA3 duplex has a 19-mer guide and 18-mer passenger RNA (17 base pair). (**A**) Root mean square deviation of the backbone atoms in the protein (black) and siRNA (red) during the MD simulations of unmodified siRNA3−hAGO2 complex. (**B**) Root mean square deviation of the backbone atoms in the protein (black) and siRNA (red) during the MD simulations of 2′-OB modified siRNA3−hAGO2 complex. (**C**) Root mean square fluctuations (RMSF) of the backbone atoms in the protein during the MD simulations of siRNA3−hAGO2 complex calculated from the last 800 ns of MD simulations. The domains are marked in the graph on the basis of the residue numbers. (**D**) MD snapshot of hAGO2 averaged from the last 200 ns of the 1 μs MD simulations. The magenta line between the N-terminal and PIWI domain indicate the nucleic acid binding channel, which facilitates the unobstructed guide-target base pairs. Protein is represented in surface (20% transparency). RNA is represented in cartoon. (**E**) Structural alignment (RMSD ∼ 2.4) of the siRNA3−hAGO2 model (red; 17 base pair) and hAGO2 crystal structure (green; 9 base pair; PDB entry: 4W5O). Protein and RNA are represented in cartoon.

protein were computed using MMPBSA pairwise energy decomposition analysis using a reported procedure.[78] Hydrogen bonds were considered based upon the heavy atom distance (donor−acceptor) cutoff of ≤3.3 Å and an angle cutoff of ≥135°. The distance between the two heavy atoms were calculated using PTRAJ module. All the structural alignments were carried out using combinatorial extension algorithm.[79] Two residues are considered to be stacked, if the distance between the center of masses is less than 5 Å and the angle between the planes of the residues is less than 30°. Distances were computed using PLUMED plugin.[80] MD trajectories were visualized using UCSF Chimera.[81]

Principal Component Analysis. Principal component analysis (PCA) of the protein motion was calculated from the last 800 ns of the MD simulation trajectories of the three unmodified siRNA−hAGO2 complexes. The translational and rotational motions of the system was removed by least-squares fit on the average MD structure. The eigen vectors were calculated as

$$C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$$

x_i and x_j are the coordinates of the ith and jth Cα atoms in the protein. The total number of Cα atoms in the system was 838. The $\langle x_i \rangle$ and $\langle x_j \rangle$ is the time average over all the configuration of the atoms during the MD simulations.[82] This calculation was carried out using parallel version of the CPPTRAJ module in AMBER 14. The principal motions of the protein are shown in graphical representations using VMD tools.[83]

Cluster Analysis. Average-linked hierarchical clustering algorithm was used for the cluster analysis and performed based on a previous report.[84] The representative structures were obtained from the major cluster emerged from the MD trajectories. The RMSD cutoff for the neighboring cluster set was given as 2 Å. This analysis was performed using CPPTRAJ module.

Binding Free Energy Calculation and Decomposition. For this calculation, 1000 snapshots were extracted from the last 500 ns MD simulations at a time interval of 500 ps. The free energy of siRNA−hAGO2 complex binding was computed using Molecular Mechanics/Poisson−Boltzmann Surface Area (MM/PBSA) method.[85] The ΔG is given by the following equation

$$\Delta G = G_{\text{hAGO2+siRNA}} - (G_{\text{hAGO2}} + G_{\text{siRNA}})$$

where $G_{\text{hAGO2+siRNA}}$, G_{hAGO2}, and G_{siRNA} are the free energies of the siRNA−hAGO2 complex, the isolated hAGO2 and siRNA, respectively.[78] In the MM/PBSA method, each free energy term is calculated as the sum of bond energy (bond, angle and dihedrals), van der Waals energy, electrostatic energy, polar and nonpolar contributions to the solvation energy, absolute temperature and entropy. The contributions from the polar and nonpolar terms were calculated using Adaptive Poisson−Boltzmann Solver (APBS) method. The external or the dielectric constant of the solvent was set to 80, the internal or the dielectric constant of the solute (protein and RNA) was set to 1. The nonpolar solvation energy was assumed to be proportional to the surface area constant, which is equal to 0.0072 kcal·mol^{-1}·Å^{-2}. Normal-mode analysis was used to estimate the entropy of the siRNA−hAGO2 complexes. Water molecules were removed from the trajectories. The extracted structures were then minimized with no cutoff for nonbonded interactions by conjugate gradient and then Newton−Raphson minimizations. The cutoff for root-mean-square of the elements of gradient vector was set to attain <10^{-4} kcal/mol.[86] The per-residue energy

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

Table 1. Binding Free Energy Components siRNA−hAGO2 Complexes Calculated from the Last 500 ns of the 1 μs MD Simulations[a]

	ΔE_{ELEC}	ΔE_{VDW}	ΔE_{MM}	ΔG_{PB}	ΔG_{NP}	ΔG_{SOLV}	ΔH_{PB}	$T\Delta S$	ΔG ($\Delta\Delta G$)
siRNA3 (2′-OH)	−2446 ± 21	−316 ± 9	−2762 ± 23	2472 ± 22	−13 ± 0.6	2459 ± 21	−303 ± 9	−57 ± 3	−246 ± 8
siRNA3 (2′-OMe)	−2413 ± 20	−293 ± 12	−2706 ± 22	2443 ± 21	−12 ± 0.4	2431 ± 21	−287 ± 11	−53 ± 4	−234 ± 12 (+12 ± 10)
siRNA3 (4′-AM-2′-OMe)	−2410 ± 23	−274 ± 13	−2684 ± 22	2420 ± 21	−11 ± 0.6	2409 ± 22	−275 ± 8	−51 ± 4	−224 ± 11 (+21 ± 10)
siRNA3 (2′-MOE)	−2411 ± 22	−268 ± 11	−2679 ± 21	2418 ± 24	−12 ± 0.6	2406 ± 23	−273 ± 10	−56 ± 3	−217 ± 16 (+28 ± 12)
siRNA3 (2′-OB)	−2409 ± 22	−263 ± 10	−2672 ± 22	2411 ± 23	−11 ± 0.7	2400 ± 23	−272 ± 10	−53 ± 4	−219 ± 13 (+27 ± 10)
siRNA2 (2′-OH)	−2445 ± 23	−309 ± 11	−2754 ± 23	2469 ± 25	−14 ± 0.8	2455 ± 24	−299 ± 7	−55 ± 3	−244 ± 9
siRNA2 (4′-AM-2′-OMe)	−2391 ± 27	−243 ± 13	−2634 ± 24	2388 ± 21	−12 ± 0.6	2376 ± 20	−258 ± 9	−49 ± 4	−209 ± 14 (+35 ± 11)
siRNA2 (2′-OB)	−2388 ± 26	−247 ± 11	−2635 ± 23	2391 ± 25	−10 ± 0.7	2381 ± 22	−254 ± 8	−50 ± 5	−204 ± 13 (+39 ± 11)
siRNA5 (2′-OH)	−2445 ± 23	−309 ± 11	−2754 ± 23	2469 ± 25	−14 ± 0.8	2455 ± 24	−299 ± 7	−55 ± 3	−244 ± 9
siRNA5 (2′-OMe)	−2452 ± 23	−301 ± 13	−2753 ± 23	2469 ± 24	−13 ± 0.6	2456 ± 20	−297 ± 9	−55 ± 5	−242 ± 10 (+2 ± 9)
siRNA5 (4′-AM-2′-OMe)	−2450 ± 18	−296 ± 12	−2746 ± 24	2464 ± 23	−12 ± 0.8	2452 ± 22	−294 ± 8	−55 ± 4	−239 ± 7 (+5 ± 8)
siRNA5 (2′-MOE)	−2404 ± 23	−253 ± 11	−2657 ± 22	2399 ± 21	−10 ± 0.7	2389 ± 21	−268 ± 10	−49 ± 4	−219 ± 14 (+24 ± 12)
siRNA5 (2′-OB)	−2445 ± 24	−301 ± 13	−2746 ± 23	2462 ± 22	−13 ± 0.7	2449 ± 21	−297 ± 11	−54 ± 5	−243 ± 8 (+1 ± 8)
siRNA14 (2′-OH)	−2450 ± 19	−298 ± 12	−2748 ± 23	2464 ± 24	−14 ± 0.6	2450 ± 23	−298 ± 9	−56 ± 3	−242 ± 7
siRNA14 (2′-OMe)	−2455 ± 21	−296 ± 11	−2751 ± 21	2469 ± 22	−13 ± 0.8	2456 ± 24	−295 ± 9	−55 ± 4	−240 ± 8 (+2 ± 7)
siRNA14 (4′-AM-2′-OMe)	−2452 ± 21	−293 ± 10	−2745 ± 22	2466 ± 21	−13 ± 0.8	2453 ± 23	−292 ± 8	−55 ± 5	−237 ± 10 (+5 ± 9)

[a]All the values are mentioned in kilocalories per mole. ΔE_{ELEC}, electrostatic energy; ΔE_{VDW}, van der Waals energy; ΔE_{INT} internal energy is negligible in all the cases; $\Delta E_{MM} = \Delta E_{ELEC} + \Delta E_{VDW} + \Delta E_{INT}$; ΔG_{PB}, polar solvation free energy; ΔG_{NP}, nonpolar solvation free energy; $\Delta G_{SOLV} = \Delta G_{PB} + \Delta G_{NP}$; $\Delta H_{PB} = \Delta E_{MM} + \Delta G_{SOLV}$; $T\Delta S$, total entropy contribution; $\Delta G = \Delta H_{PB} - T\Delta S$, total binding free energy; $\Delta\Delta G = \Delta G_{modified} - \Delta G_{unmodified}$, difference in the binding free energy. Averaged over 1000 snapshots. All these values are determined using MM-PBSA approach,[85] entropy contribrutions using normal-mode analysis.

decomposition was carried out using previously reported methods.[78]

RESULTS AND DISCUSSION

The global structural differences in the open and closed conformations of the structurally well studied prokaryotic AGO are due to the orientations of the L1 and L2 linkers, PAZ, and N-terminal domains with respect to the MID and PIWI domains.[20,24,25] The closed conformation of hAGO2 cannot be docked with the complete double stranded siRNA. Hence, using an iterative refinement of domain orientations[55] (AIDA) and comparative modeling[54] (MODELER 9v7), an energetically favorable open hAGO2 conformation was developed.[20,23] The open conformation of hAGO2 was then utilized to dock with the unmodified and chemically modified (2′-OMe, 4′-AM-2′-OMe, 2′-MOE, and 2′-OB, Figure 1A) 21−22 nt siRNAs (siRNA2, siRNA3, siRNA5, and siRNA14, Figure 1B and Table S1, Supporting Information) using Hex and Glide precision docking protocols.[61] The crystal structure of the *Kp*AGO[23] revealed that the orientation of N-terminal domain allows the extended nucleic acid binding channel having unobstructed guide−target pairs. Exploiting this feature, we could dock hAGO2 with 22 nt siRNA. However, the MD simulations (200 ns) revealed that this complex was unstable at the 3′-end of the guide strand (beyond 17 nt) and N-terminal domain of the protein. This was reflected in the RMSD graph of the siRNA−hAGO2 complex (Figure S2B, Supporting Information). To avoid this fluctuation, the siRNA was trimmed to 17 nt in further studies. The positions (g2, g3, g5, and g14) of the four different chemical modifications and the siRNA sequences utilized in this study are shown in Figure 1 and Table S1, Supporting Information. Each complex was subjected to equilibrations, followed by this an unrestrained MD simulations (particle mesh ewald)[72] was performed for 1 μs using GPU accelerated version[73,74] of AMBER 14.[58]

Structural Stability. The structural stability of the complexes was assessed by calculating the RMSD of the backbone atoms of the protein and RNA. The complexes were found to be globally stable, which were evident from the RMSD graphs acquired from the MD trajectories (Figure 2A and B and Supporting Information). The small fluctuations observed in the RMSD graphs indicate sufficient sampling and time scale employed in the simulations. To understand the protein-RNA recognition at the atomistic level, the dynamic features of the amino acids were calculated. The RMSF values of the protein backbone reveal that the PAZ domain is most flexible domain the hAGO2 followed by the PIWI and N-terminal domains. MID domain was found to be the rigid during the course of MD simulations (Figure 2C).

Model of hAGO2 and Unmodified siRNA with 17 Base Pairs. The averaged structure obtained from MD simulations of siRNA−hAGO2 model (open conformation) is shown in Figure 2D and Figure S3 of the Supporting Information. Structural alignment of the siRNA−hAGO2 complex and the crystal structure of hAGO2 (PDB entry: 4W5O) showed notable differences in the helix-7 and the orientation of the PAZ domain (Figure 2E). At the domain level, the structural alignment showed similar secondary structural features (Figure S4, Supporting Information). The model showed that the binding pocket of the 5′-phosphate in the guide RNA is recognized by number of amino acids (AAs)[21,22,30] with >98% occupancies during the course of MD simulations. For the three unmodified siRNA−hAGO2 complexes studied, the overall siRNA−hAGO2 interactions were found to be similar (Figure S5, Supporting Information). The RNA base pair parameters and the intramolecular contacts in the siRNA were analyzed using X3DNA[77] and PTRAJ[76] modules. The results showed that 2′-OH of g1 make H-bond with the phosphate backbone of g2, and also the 2′-OHs of other nts at various positions make H-bonds with O4′ in the sugar as reported in the crystal structures (Figure S6,

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

Table 2. Per-Nucleotide Decomposition Energy (Electrostatic and van der Waals) of Guide Strand Nucleotides in siRNA−hAGO2 Complexes Calculated from the Last 500 ns of the 1 μs MD Simulations[a]

	g1	g2	g3	g4	g5	g6	g13	g14	g15
siRNA3 (2′-OH)	−133 ± 7	−72 ± 3	−96 ± 7	−87 ± 4	−74 ± 2	−59 ± 3	−36 ± 4	−41 ± 4	−33 ± 4
siRNA3 (2′-OMe)	−130 ± 9	−73 ± 4	**−91 ± 5**	−82 ± 5	−73 ± 4	−61 ± 5	−34 ± 7	−40 ± 6	−31 ± 6
siRNA3 (4′-AM-2′-OMe)	−106 ± 6	−57 ± 4	**2 ± 10**	−66 ± 6	−69 ± 2	−59 ± 4	−34 ± 7	−40 ± 5	−32 ± 6
siRNA3 (2′-MOE)	−92 ± 7	−54 ± 3	**−21 ± 11**	−39 ± 4	−67 ± 2	−58 ± 5	−35 ± 4	−41 ± 6	−30 ± 5
siRNA3 (2′-OB)	−98 ± 7	−55 ± 4	**−23 ± 8**	−44 ± 7	−67 ± 4	−58 ± 4	−34 ± 5	−40 ± 3	−33 ± 4
siRNA2 (2′-OH)	−135 ± 8	−71 ± 5	−95 ± 9	−89 ± 5	−71 ± 5	−58 ± 6	−35 ± 6	−41 ± 5	−31 ± 6
siRNA2 (4′-AM-2′-OMe)	−89 ± 11	**−47 ± 7**	−69 ± 6	−82 ± 4	−68 ± 3	−53 ± 7	−33 ± 8	−38 ± 5	−28 ± 4
siRNA2 (2′-OB)	−76 ± 9	**−41 ± 9**	−66 ± 7	−74 ± 5	−67 ± 6	−54 ± 7	−28 ± 8	−42 ± 4	−31 ± 4
siRNA5 (2′-OH)	−133 ± 9	−66 ± 7	−90 ± 7	−85 ± 4	−68 ± 6	−55 ± 8	−31 ± 9	−40 ± 9	−34 ± 9
siRNA5 (2′-OMe)	−128 ± 5	−64 ± 6	−91 ± 5	−84 ± 9	**−62 ± 4**	−52 ± 6	−33 ± 6	−42 ± 4	−28 ± 4
siRNA5 (4′-AM-2′-OMe)	−124 ± 7	−67 ± 7	−87 ± 7	−81 ± 4	**−56 ± 11**	−51 ± 8	−32 ± 8	−41 ± 6	−32 ± 3
siRNA5 (2′-MOE)	−129 ± 5	−66 ± 7	−85 ± 5	−56 ± 5	**−7 ± 2**	−17 ± 8	−32 ± 11	−39 ± 8	−28 ± 8
siRNA5 (2′-OB)	−126 ± 11	−67 ± 5	−94 ± 8	−83 ± 6	**−76 ± 9**	−57 ± 7	−35 ± 7	−38 ± 7	−29 ± 6
siRNA14 (2′-OH)	−133 ± 9	−70 ± 6	−91 ± 11	−88 ± 5	−69 ± 7	−61 ± 7	−36 ± 6	−42 ± 7	−32 ± 4
siRNA14 (2′-OMe)	−134 ± 6	−71 ± 6	−92 ± 7	−87 ± 7	−66 ± 7	−59 ± 9	−29 ± 9	**−37 ± 6**	−33 ± 7
siRNA14 (4′-AM-2′-OMe)	−134 ± 7	−70 ± 7	−90 ± 8	−90 ± 7	−70 ± 6	−55 ± 5	−28 ± 9	**−31 ± 9**	−26 ± 8

[a]All the values are mentioned in kilocalories per mole and determined using the MM-PBSA per-residue energy decomposition method.[78] The positions in bold are the modified positions in the correspoding siRNA−hAGO2 complex.

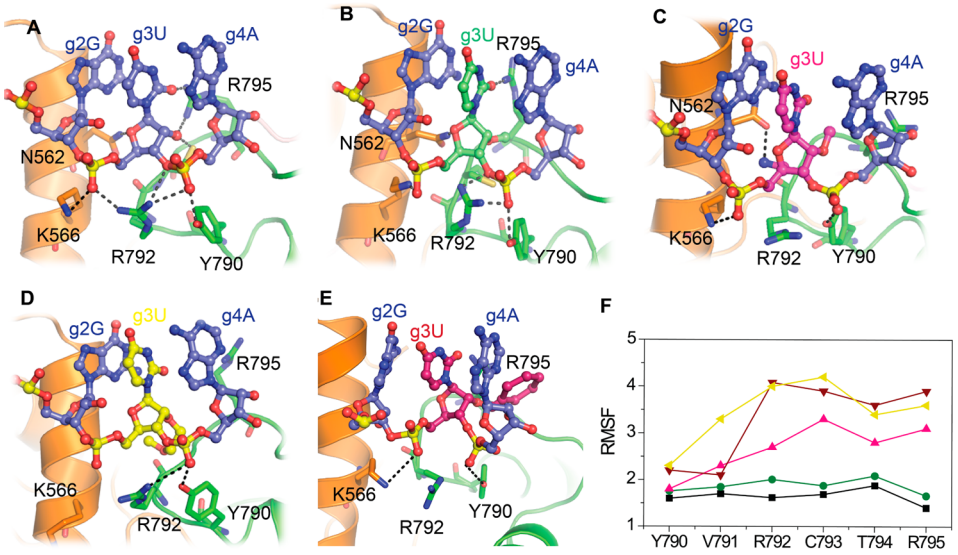


Figure 3. Effect of sugar modifications at position g3 of siRNA3−hAGO2 complexes. Best representative structures from the last 200 ns of the MD simulations of (**A**) unmodified, (**B**) 2′OMe (green), (**C**) 4′-AM-2′-OMe (magenta), (**D**) 2′-MOE (yellow) and (**E**) 2′-OB (brown) modified siRNA3−hAGO2 complexes. (**F**) Comparison of RMSFs (Å) from ensemble average of key residues (790−795) in hAGO2 which is interacting at the position g3 of the unmodified (black) and modified siRNAs. Protein is represented in cartoon, interacting AAs residues are represented in sticks. RNA (blue) atoms and bonds are represented in spheres and sticks, respectively. The black dashed lines indicate the noncovalent interactions between nucleotides of siRNA and AAs residues of hAGO2.

Supporting Information).[22] The helical parameters of the siRNAs were similar to those of a typical A-form RNA duplex (Figure S7, Supporting Information). The motions of the domains in hAGO2 during the dynamics were analyzed using principal component analysis of MD trajectories, which revealed considerable fluctuations in the PAZ domain.[87,88] In addition, two different stretches of loops including one from the PAZ (N283−G301) and the other from PIWI (H816−R837) domains were found to fluctuate more as revealed from RMSF values and cluster analysis (Figures 2C and S8, Supporting Information).[35,39,84] The MM/PBSA energy calculations[85] were performed for all the complexes to find the energy penalties for the formation of modified RNA−hAGO2 complexes. It was evident that the electrostatic (ΔE_{ELEC}) and van der Waals interactions (ΔE_{VDW}) are the major factors involved in the complex formation (Table 1). The polar solvation (ΔG_{SOLV}) energies were unfavorable in all the cases. To further confirm the modification induced perturbations in the noncovalent interactions, per-residue energy decompositions[78] were estimated for guide strand nucleotide positions (Table 2). The key results and their structural and functional implications are discussed below.

Bulky 2′-Sugar Modifications at Position g3 Disturbs hAGO2 Binding. We have employed four different chemical modifications at position g3 in the seed region of the guide strand (Figure 3 and Table S1, Supporting Information). The MD results of the unmodified siRNA3−hAGO2 complex showed that, this position is recognized by R792, R795 and Y790, which make noncovalent interactions with the phosphate backbone and

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

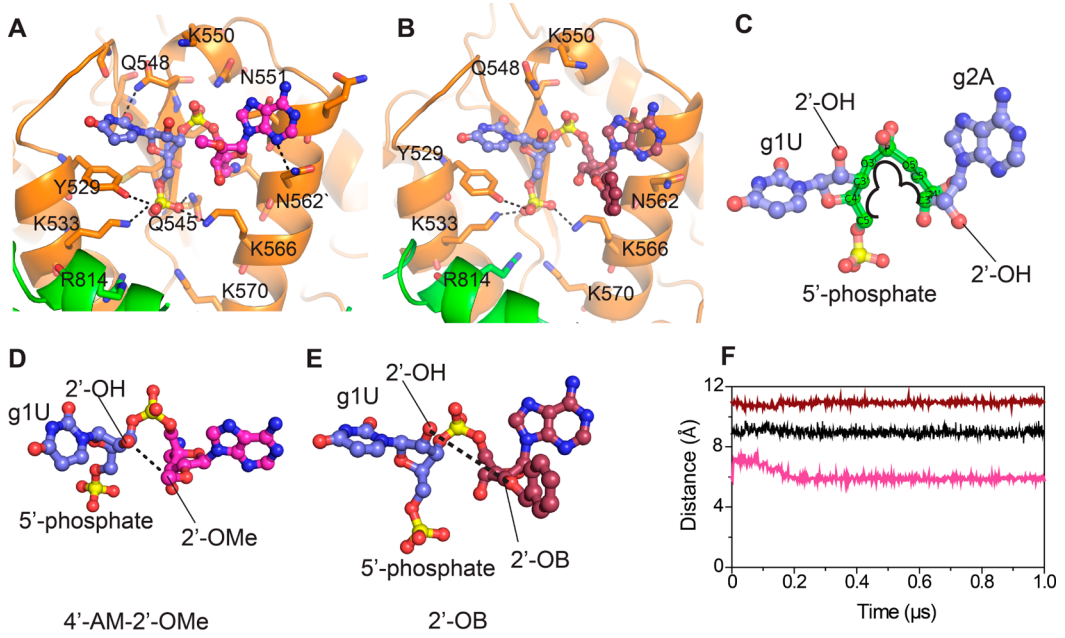


Figure 4. MD snapshots showing the binding interactions and conformational effects of 2′-modifications in the clover leaf (position g1 and g2) junction present at the 5′-guide of siRNA2. The 5′-guide RNA binding pocket in the MID domain (orange) averaged from the last 100 ns of the MD simulations of (**A**) 4′-AM-2′-OMe (magenta) (**B**) 2′-OB (ruby) modified siRNA2−hAGO2 complexes. The black dashed lines indicate the noncovalent interactions between nucleotides of siRNA and AAs residues of hAGO2. Protein is represented in ribbon and interacting atoms are represented in sticks. RNA (light blue) is represented in stick and sphere. The conformation of dinucleotide junction in the (**C**) unmodified, (**D**) 4′-AM-2′-OMe (magenta), and (**E**) 2′-OB (ruby) modified siRNA2−hAGO2 complexes. The dashed line indicates the distance between the 2′-oxygen atoms in position g1 and g2. (**F**) Time dependent variations of distance between 2′-oxygen atoms in the g1U and that in g2A of the unmodified (black), 4′-AM-2′-OMe (magenta), and 2′-OB modified (brown) siRNA2−hAGO2 complexes.

with the 2′-OH group as shown in Figure 3A. Introduction of 2′-OMe modification at g3 reorients the R795 side chain, however, all other major interactions are retained. In the case of 4′-AM-2′-OMe modification, due to the presence of positive charge in the sugar, electrostatic interactions of both R792 and R795 with the RNA backbone are severely affected (Figure 3B and C, Figure S9, Supporting Information). The 4′-*C*-aminomethyl group of 4′-AM-2′-OMe nt make a H-bond with N562 (OD1) (Figures 3C and S9, Supporting Information). The modification leads to the deviation in the backbone torsions, loss of one W−C base pair due to the *C2′-endo* sugar conformation[16] and also leads to electrostatic repulsions with arginine side chains (Figure S10, Supporting Information). These observations were also supported by the MM-PBSA free energy calculations (Table 1) and the per-residue energy decomposition analysis (Table 2). This could indeed affect the seed region recognition and loading of modified siRNA into hAGO2, which can lead to the reduced RNAi activity as observed in experiments.[16] Bulky modifications such as 2′-MOE and 2′-OB disturbs the electrostatic interactions of R792 and R795 with the backbone of the siRNA as well (Figure 3D and E, Tables 1 and 2). In the modified siRNA3−hAGO2 complexes, the amino acids (AA) V791−R795 move away (∼3 Å) in comparison to the unmodified siRNA−hAGO2, to avoid the steric clash between AAs (V791 and C793) and MOE or OB moieties (Figures S10 and S11, Supporting Information). These observations were also supported by the fluctuations observed in the RMSF values for the interacting AAs (Figure 3F).

In addition to the above-described perturbations between modified siRNA and hAGO2, the 2′-OB and 2′-MOE modifications also disturb the key electrostatic contacts between the residue K566 with the phosphate backbone of g3 and with the 5′-phosphate of g1.[20] The distance between g1 and g3 P atoms in the unmodified complex was ∼6.5 Å (Figure S12, Supporting Information). However, in the modified complexes, the corresponding distances were found to be ∼7.2 Å (2′-OB) and ∼7.6 Å (2′-MOE). These enlarged distances are responsible for the loss of RNA contacts with K566. Overall, these results show that bulky and charged sugar modifications at g3 adversely affect the hAGO2 interactions with siRNA.

Modifications at Position g2 Destabilize 5′-Phosphate Binding. Depending upon the chemical nature, modifications at position g2 of the siRNA have deleterious effects on RNAi activity.[14,15] We have performed MD simulations of the hAGO2−siRNA2 complex, in which the modification (2′-OB or 4′-AM-2′-OMe) is placed at g2 (Table S1, Supporting Information). The MD results show that the 5′-phosphate backbone of g1 and g2 in unmodified siRNA2 make multiple H-bonds/electrostatic contacts with Q545, Y529, R814, K570, K566, N551, K550, T526, and Q548, which are important for the 5′-recognition of siRNA by hAGO2 (Figures S4 and S13 Supporting Information). In the presence of the 4′-AM-2′-OMe modification, noncovalent interactions of siRNA with R814, K570, N551, K550, and T526 were lost during the course of MD simulations (Figures 4A and S13, Supporting Information). In addition, the noncovalent interactions between g1U and the nucleotide specific loop (P523−P527)[31] present in the MID domain are lost. The 2′-OB modification disturbs the contacts between g2 and R814, K570, N551, K550, T526 and Q548 AA residues (Figures 4B and S13, Supporting Information). The loss of multiple noncovalent interactions, which are necessary for the 5′- phosphate recognition, can lead to inefficient loading of siRNA into hAGO2, which in turn could affect the RNAi activity.[14] These effects are also reflected in the binding free

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

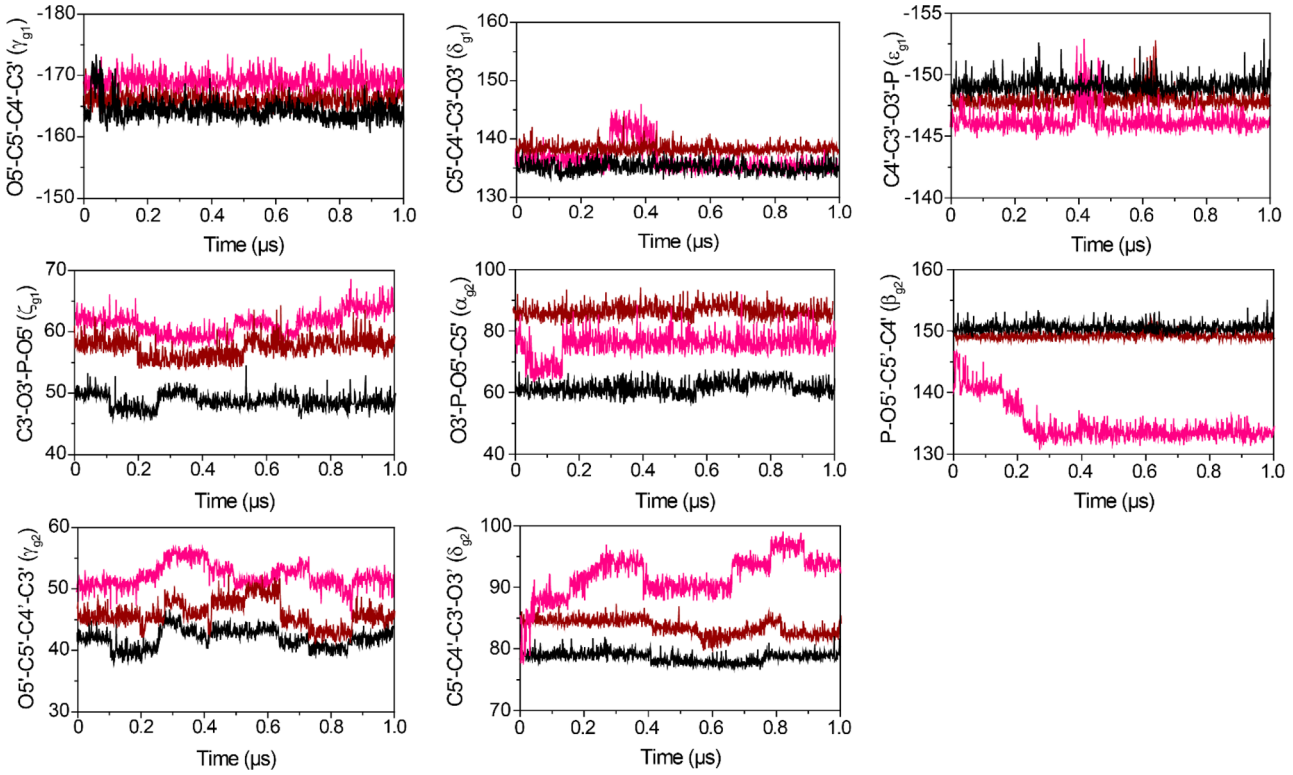


Figure 5. Time-dependence of backbone dihedral angles in the clover leaf geometry of the 5′-dinucleotide during the MD simulation of siRNA2−hAGO2 complexes. The dihedral angle is measured as the angle between the four heavy atoms as depicted on the plots. The dihedrals in the unmodified (black), 4′-AM-2′-OMe (magenta), and 2′-OB (brown) modified siRNA2 are highlighted using different colors.

energies of the modified siRNA2−hAGO2 complexes ($\Delta\Delta G \sim$ +35 kcal/mol, Table 1), and the loss of noncovalent interactions is also clearly visible from the per-residue energy decomposition values (Table 2).

The loss of noncovalent interactions between RNA and hAGO2 at position g2 can be ascribed to the changes in the clover leaf shaped unique sugar−phosphate backbone conformation formed by position g1 and g2 (Figure 4C).[30,89,90] The clover leaf junction comprises of eight sugar−phosphate backbone torsions. These torsions in the unmodified siRNA did not fluctuate throughout the course of MD simulations (Figure 5). In presence of 4′-AM-2′-OMe, deviations were observed in the β_{g2}, γ_{g2}, and δ_{g2} torsions. These torsional deviations can be attributed to the *C2′-endo* sugar conformation induced by 4′-AM-2′-OMe modification.[16] Furthermore, deviations were observed in ζ_{g1} and α_{g2} torsions, due to the presence of this modification. In case of the 2′-OB, larger deviations were observed in ζ_{g1} and α_{g2} (Figure 5) to minimize the steric clashes between the modification and the interacting AAs of hAGO2. As a result the orientation of the nucleotide at the 5′-end of the guide strand is perturbed (Figure 4B and E). These observations indicate that studies on the backbone conformations of the modified nucleotides in an siRNA−hAGO2 complex can be helpful to place appropriate modification at position g2. These results also illustrate the possible reasons behind the tolerance of less bulky modifications having C3′-*endo* sugar conformation such as 2′-OMe and 2′-F at position g2.[11] It should be noted here that the ribose sugar in position g1 adopts DNA-like C2′-*endo* conformation and from position g2 onward the siRNA adopt sugar C3′-*endo* conformation.[22,24,25,30,31] Hence, a preorganized C2′-*endo* sugar conformation like 2′-deoxy-2′-fluoroarabino nucleic acid (2′-FANA) at the 5′-end of siRNA is tolerated at g1 without affecting RNAi.[91]

Along with backbone torsional changes, the distance between the 2′-oxygen atoms of g1 and g2 increases or decreases due to the presence of modifications (Figure 4F). The average distance between these atoms in the unmodified siRNA was found to be 9.3 ± 0.3 Å. In the case of 4′-AM-2′-OMe, this distance decreased (6.1 ± 0.2) due to the *C2′-endo* sugar conformation at positions g1 and g2. Whereas in the case of 2′-OB, the distance increased (11.2 ± 0.5 Å), due to the torsional changes in the clover leaf geometry induced by modification (Figure 4F). These results clearly show that depending upon the chemical nature of 2′-modification, the geometry of clover leaf junction varies, which could in turn affect the loading of the siRNA into the MID domain of hAGO2. Therefore, 5′-phosphate modifications which could impart conformational preorganization at the g1 and g2 positions would facilitate hAGO2 recognition without sacrificing RNAi activity. This is in fact supported by the enhanced RNAi activity reported for the 5′-phosphate modifications such as R-C5′-*Me* and -C5′-malonyl.[92,93]

2′-*O*-Benzyl Moiety at Position g5 Creates a Hydrophobic Binding Site in the PIWI Loop. The bulky 2′-MOE modification in the seed region of siRNA drastically reduces the RNAi activity.[11] Surprisingly, 2′-OB modification and its derivative (2′-*O*-methyl-4-pyridine) were tolerated and enhances the *in vivo* RNAi activity, when they are placed at positions g5, g8, g15, and g19.[14,15] To unravel the contrasting role of these bulky modifications, the 2′-OB and 2′-MOE were placed at position g5 in the siRNA5. In the unmodified siRNA5−hAGO2 complex, the loop of PIWI domain (H753−S760) interacts with the RNA backbone (Figures 6A and C and S14, Supporting Information). The 2′-OB modification at g5 had a striking effect on the

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

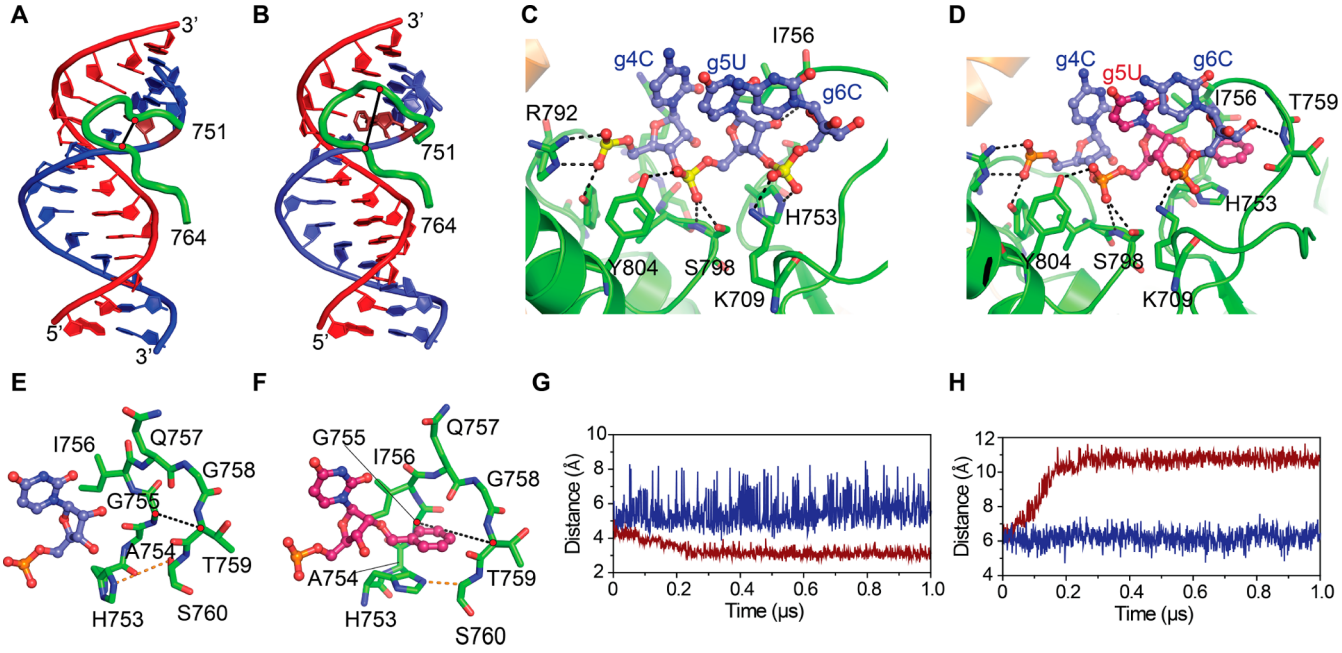


Figure 6. MD snapshots and distance graphs depicting the hydrophobic binding site in the PIWI loop of hAGO2 formed in the presence of 2′-OB modification. Best representative structure from the last 200 ns MD simulations of the (**A**) unmodified siRNA5 and (**B**) 2′-OB modified siRNA5 and the PIWI loop (751−764) in hAGO2. Protein (PIWI loop) is represented in cartoon (green, with loop radius setting of ∼0.6). RNA is represented in cartoon; guide (blue) and passenger (red) strands are highlighted. The approximate position of G755 and T759 residues are highlighted using red dots, and black line indicate the distance between these residues. The best representative structures from the last 200 ns of the MD simulations of (**C**) unmodified and (**D**) 2′-OB (ruby) modified siRNA5−hAGO2 complexes. The black dashed lines indicate the noncovalent interactions between nucleotides of siRNA and AAs residues of hAGO2. MD snapshot of the (**E**) unmodified (blue) and (**F**) 2′-OB (ruby) modified nucleotide at position g5 along with the loop (green) from the PIWI domain. The orange line indicates distance between H753 (NE2) and S760 (OG). Black dashed line indicates the distance between C-alpha atoms in the G755 and T759 residues. AAs in the protein is represented in sticks, RNA is represented in spheres and sticks. (**G**) Time-dependence of H-bond distance fluctuations between the heavy atoms of H753 (NE2) and S760 (OG) in the 2′-OB-siRNA5−hAGO2 (brown) complexes. (**H**) Time-dependence of distance between the C-alpha atoms in G755 and T759 (highlighted as red circle in parts E and F) from the simulations of unmodified siRNA5 and 2′-OB-siRNA5−hAGO2 complex.

orientation of this loop (Figure 6B and 6D). As a result, the size of the loop (distance between the C_α atoms in G755 and T759) widens to a maximum of ∼11 Å, compared to the size of ∼5 Å in the unmodified siRNA (Figure 6E and F). In this enlarged loop, H753 restricts the mobility of the benzyl moiety of 2′-OB modification by forming stacking interactions as shown in Figure 6F. While the loop creates the hydrophobic binding site for the modification (Figure S15 Supporting Information), a H-bond between H753 and the RNA backbone is broken. However, the enlarged loop is stabilized by the formation of internal H-bond between H753 (NE2) and S760 (OG) (Figure 6G). Stacking of H753 with the benzyl moiety and the van der Waals interactions (A754, I756 and Q757) between the loop residues and the benzyl moiety facilitate the formation of a hydrophobic binding pocket to accommodate the modification (Figure S15, Supporting Information). On the other hand, the 2′-MOE modification at g5 disturbs the H- bond between the H753 (NE2) and RNA backbone. Due to the rigid gauche conformation adapted by the MOE group,[68] the hydrophobic loop of the PIWI domain moves ∼2.5 Å to avoid the steric clash with the MOE moiety (Figures S16 and S17, Supporting Information). The contrasting effects of these two modifications were also confirmed by the free energy and per-residue energy decomposition calculations (Tables 1 and 2).

To compare the effects of the two bulky modifications (2′-OB and 2′-MOE) with the less bulky ones, we introduced 2′-OMe and 4′-AM-2′-OMe in position g5 of siRNA5. Noncovalent interactions were preserved during the course of dynamics without any substantial disturbances with these modifications (Figure S14 and S16, Supporting Information). The positively charged 4′-*C*-aminomethyl group in 4′-AM-2′-OMe does not disturb the electrostatic interaction between K709 and phosphate backbone, because K709 points away from the *C*4′ in sugar moiety (6.2 Å). These observations were further supported by corresponding free energies obtained from MM-PBSA and per-residue energy decomposition calculations (Table 1 and 2). Based on these findings, though experimental data is not available, we speculate the dual modifications such as 4′-AM-2′-OMe and 4′-*C*-aminomethyl-2′-deoxy-2′-fluoro (4′-AM-2′-F)[94] will be tolerated at position g5. It is interesting to note here that an amide modification in the backbone is tolerated at position g5 of the siRNA.[95] This may be attributed to the orientation of the C═O bond in the modification, which is similar to that of the P−O bond in A-type RNA. Furthermore, only small changes in the minor and major groove widths were observed in the central part of the amide modified RNA in comparison to a typical A form RNA.[95]

Altogether, the results presented here suggest that structure sensitive chemical modifications in appropriate positions of siRNA could modulate the atomic level interactions between siRNA and hAGO2. Also, our results underscore the deficiencies in using static computational models to map the effect of chemical modifications in siRNA on hAGO2 binding. For example, based upon a docked structure, one could envision that the introduction of bulky 2′-OB moiety at the g5 position of siRNA would affect the noncovalent contacts with hAGO2.

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

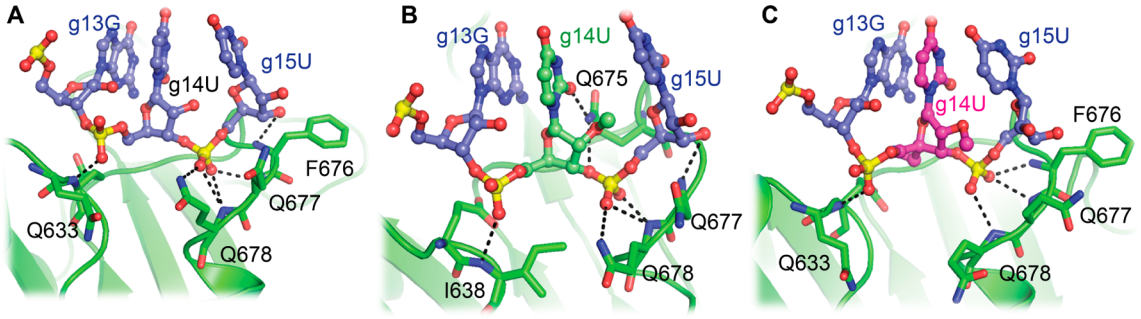
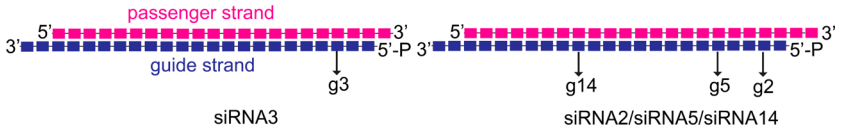


Figure 7. MD snapshots depicting the siRNA14-hAGO2 interactions in the presence of modifications at position g14 of siRNA. Best representative structure from the last 200 ns of the MD simulations of (**A**) unmodified, (**B**) 2′-OMe (light green), and (**C**) 4′-AM-2′-OMe (magenta) modified siRNA14- hAGO2 complexes. The black dashed lines indicate the noncovalent interactions between nucleotides of siRNA and AAs residues of hAGO2. The representations of protein and RNA are described as in Figure 3.

Table 3. Summary of Tolerance of the Chemical Modifications at Various Positions in siRNA Observed in the Experiments (exp) and from the MD Simulation (model) Study



	2′-OH		2′-OMe		4′-AM-2′-OMe		2′-MOE		2′-OB	
mod. positions	model	exp	model	exp	model	exp	model	exp	model	exp
g2	+	+	nd[a]	−	−	−	nd	−	−	−
g3	+	+	+	+	−	−	−	−	−	−
g5	+	+	+	+	+	nd	−	−	+	+
g14	+	+	+	−	+	nd	nd	−	nd	−

[a]nd denotes not determined. The + represents the chemical modification in siRNA is tolerated. And − represents the modification is poorly or not tolerated. The g*n* represents the *n*th position of the nucleotide in the guide strand. The position g3 in siRNA3 and the positions g2, g5, and g14 in siRNA2, siRNA5, and siRNA14, respectively, are shown above. The sequences of all siRNAs are shown in Figure 1B.

However, the MD simulations presented here give unique insights on the accommodation of this modification by the formation of hydrophobic pocket.

Modification at Position g14 Retains the Noncovalent Interactions with hAGO2. A recent report suggested that the incorporation of chemical modifications including 2′-OMe and 2′-F units at position g14 of siRNA decreases RNAi efficiency.[42] A static model of siRNA−hAGO2 complex was used to rationalize these findings. Based on this model, it was suggested that the steric clashes between the glutamine rich region in the PIWI domain, and the modification is responsible for decrease in the RNAi activity.[42] This observation led us to explore the binding interactions between 2′-OMe and 4′-AM-2′-OMe modifications at g14 using our model. MD results showed that Q633, F676, Q677, and Q678 make H-bonds with the backbone of the unmodified siRNA14 (Figures 7 and S18, Supporting Information). Presence of 2′-OMe at g14 leads to loss of noncovalent interactions with Q633, however, I638 makes H-bond with the backbone. All other interactions were retained without any substantial deviations. In the case of 4′-AM-2′-OMe, the interaction with Q678 with the backbone is lost after 200 ns of simulations, and all other interactions were retained as in the unmodified siRNA14 (Figure S18, Supporting Information). In addition, there were no significant difference in the $\Delta\Delta G$ values (Table 1 and 2) in the presence of modifications at position g14, which indicate the favorable binding of modified siRNA with hAGO2. The effect of modifications at position g14 on RNAi may be due to the unfavorable interactions between the siRNA and hAGO2 during the other steps involved in the RNAi mechanism.[2,96,97] It should also be noted here that a number of other studies showed moderate to high tolerance of sugar modifications at position g14.[11,98,99]

SUMMARY AND CONCLUSIONS

The hAGO2, endonuclease protein, involved in the RNAi machinery was modeled to accommodate the double stranded siRNA (2−17 nt base pairs, guide-passenger/target) using recent crystal structures and modeling results.[22,23,30] This model mimics the hAGO2−siRNA or the hAGO2-guide-target mRNA complex involved in the RNAi pathway. Using this model and microsecond MD simulations, modification induced structural changes of hAGO2 were studied using a set of 2′-sugar modified (2′-OMe, 2′-MOE, 2′-OB, and 4′-AM-2′-OMe) siRNAs. At positions g2, g3, and g5, of the siRNA guide strand, we were able to delineate how the key siRNA−hAGO2 noncovalent interactions are modulated depending upon the nature of chemical modifications (Table 3). The structural perturbations on hAGO2 imparted by chemical modifications, when they are present at the seed region of siRNA guide strand, can affect the siRNA loading and which in turn may affect RNAi activity.

At position g3, as in siRNA3, the 2′-bulky or positively charged modifications (2′-MOE, 2′-OB, 4′-AM-2′-OMe) disturb the key noncovalent interactions of siRNA and hAGO2. Sugar modifications (4′-AM-2′-OMe and 2′-OB) at position g2 (siRNA2), which is next to the 5′-end of the guide strand lead to the loss of key noncovalent interactions with the hAGO2, and induce deviations in the sugar−phosphate backbone torsions (clover leaf geometry). Hence, at positions g2/g3 less bulkier and

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

neutral chemical modifications such as 2′-deoxy-2′, 4′-difluoro[100] and 2′-F, 4′-Me could be used to retain RNAi activity. At position g5, to accommodate bulky aromatic modification such as 2′-OB, formation of a hydrophobic binding pocket was observed. Therefore, we predict that incorporation of modification like 2′-*O*-imidazolylmethyl at this position could enhance the stacking of imidazole ring of the histidine side chain with the modification and thereby facilitate the siRNA loading into hAGO2. At position g14, our model show that modifications (2′-OMe and 4′-AM-2′-OMe) do not alter the key siRNA−hAGO2 contacts.

The model of hAGO2−siRNA complex presented here could be used for the structure based design of novel therapeutically appealing siRNAs. Importantly, the conformational changes of hAGO2 in the presence of chemical modifications in siRNA were observed only after 200−300 ns of MD simulations. This warrants the use long time-scale simulations in the microsecond range to probe the precise effect of chemical modifications on hAGO2 binding. Altogether, the present study offers a benchmark for the MD simulations of hAGO2 with chemically modified siRNAs. Structural evidence of the modified siRNA−hAGO2 complexes is required to further validate the defined conformational changes and atomic contacts observed from the MD simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00773.

Analysis of domain orientations in AGO; docked structures and RMSD graphs of siRNA and hAGO2 complexes; energy optimized structure of 2′-OB nucleotide; model of siRNA−hAGO2 complex; structural comparison of model and hAGO2 crystal structures; schematic representation of unmodified siRNA−hAGO2 interactions; H-bond interactions and local base parameters of siRNA3; principal component and cluster analysis of hAGO2; MD snapshots of siRNA−hAGO2 complexes; electrostatic and van der Waals interaction energies between 2′-OB of siRNA5 and amino acids in the hAGO2; time-dependence of key H-bond distances in siRNA−hAGO2 complexes; and the siRNA sequences and modifications used in the MD simulations (PDF)

Interactive rotatable model of complex (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: harikrishna.s@iitb.ac.in (S.H.).

*E-mail: pradeep@chem.iitb.ac.in (P.I.P.).

ORCID

P. I. Pradeepkumar: 0000-0001-9104-3708

Funding

This work is supported by grants from IRCC-IIT Bombay and Department of Biotechnology, Government of India (RNAi-Technologies Platform, BT/PR10693/AGR/36/586/2008).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Computer center, IIT, Bombay, and National PARAM Supercomputing Facility (NPSF)−PARAM YUVA, Centre for Development of Advanced Computing (CDAC), Pune, are gratefully acknowledged for providing high performance computing facilities. We also thank Raghu Rangaswamy, Schrödinger India, for providing access to Glide 5, and Saurja Dasgupta for critically reading the manuscript. S.H. thanks Department of Atomic Energy−Board of Research in Nuclear Sciences (DAE-BRNS), Government of India, and IRCC-IIT Bombay for the Ph.D. fellowships.

■ REFERENCES

(1) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806−811.

(2) Wilson, R. C.; Doudna, J. A. Molecular Mechanisms of RNA Interference. *Annu. Rev. Biophys.* **2013**, *42*, 217−239.

(3) Hutvagner, G.; Simard, M. J. Argonaute Proteins: Key Players in RNA Silencing. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 22−32.

(4) Kawamata, T.; Tomari, Y. Making RISC. *Trends Biochem. Sci.* **2010**, *35*, 368−376.

(5) Carthew, R. W.; Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642−655.

(6) Kurreck, J. RNA Interference: From Basic Research to Therapeutic Applications. *Angew. Chem., Int. Ed.* **2009**, *48*, 1378−1398.

(7) Peacock, H.; Kannan, A.; Beal, P. A.; Burrows, C. J. Chemical Modification of siRNA Bases To Probe and Enhance RNA Interference. *J. Org. Chem.* **2011**, *76*, 7295−7300.

(8) Deleavey, G. F.; Damha, M. J. Designing Chemically Modified Oligonucleotides for Targeted Gene Silencing. *Chem. Biol.* **2012**, *19*, 937−954.

(9) Shukla, S.; Sumaria, C. S.; Pradeepkumar, P. I. Exploring Chemical Modifications for siRNA Therapeutics: A Structural and Functional Outlook. *ChemMedChem* **2010**, *5*, 328−349.

(10) Bramsen, J. B.; Laursen, M. B.; Nielsen, A. F.; Hansen, T. B.; Bus, C.; Langkjær, N.; Babu, B. R.; Højland, T.; Abramov, M.; Van Aerschot, A.; Odadzic, D.; Smicius, R.; Haas, J.; Andree, C.; Barman, J.; Wenska, M.; Srivastava, P.; Zhou, C.; Honcharenko, D.; Hess, S.; Müller, E.; Bobkov, G. V.; Mikhailov, S. N.; Fava, E.; Meyer, T. F.; Chattopadhyaya, J.; Zerial, M.; Engels, J. W.; Herdewijn, P.; Wengel, J.; Kjems, J. A Large-Scale Chemical Modification Screen Identifies Design Rules to Generate siRNAs with High Activity, High Stability and Low Toxicity. *Nucleic Acids Res.* **2009**, *37*, 2867−2881.

(11) Prakash, T. P.; Allerson, C. R.; Dande, P.; Vickers, T. A.; Sioufi, N.; Jarres, R.; Baker, B. F.; Swayze, E. E.; Griffey, R. H.; Bhat, B. Positional Effect of Chemical Modifications on Short Interference RNA Activity in Mammalian Cells. *J. Med. Chem.* **2005**, *48*, 4247−4253.

(12) Allerson, C. R.; Sioufi, N.; Jarres, R.; Prakash, T. P.; Naik, N.; Berdeja, A.; Wanders, L.; Griffey, R. H.; Swayze, E. E.; Bhat, B. Fully 2′-Modified Oligonucleotide Duplexes with Improved in Vitro Potency and Stability Compared to Unmodified Small Interfering RNA. *J. Med. Chem.* **2005**, *48*, 901−904.

(13) Manoharan, M.; Akinc, A.; Pandey, R. K.; Qin, J.; Hadwiger, P.; John, M.; Mills, K.; Charisse, K.; Maier, M. A.; Nechev, L.; Greene, E. M.; Pallan, P. S.; Rozners, E.; Rajeev, K. G.; Egli, M. Unique Gene-Silencing and Structural Properties of 2′-Fluoro-Modified siRNAs. *Angew. Chem., Int. Ed.* **2011**, *50*, 2284−2288.

(14) Kenski, D. M.; Butora, G.; Willingham, A. T.; Cooper, A. J.; Fu, W.; Qi, N.; Soriano, F.; Davies, I. W.; Flanagan, W. M. siRNA-optimized Modifications for Enhanced In Vivo Activity. *Mol. Ther.–Nucleic Acids* **2012**, *1*, e5.

(15) Butora, G.; Kenski, D. M.; Cooper, A. J.; Fu, W.; Qi, N.; Li, J. J.; Flanagan, W. M.; Davies, I. W. Nucleoside Optimization for RNAi: A High-Throughput Platform. *J. Am. Chem. Soc.* **2011**, *133*, 16766−16769.

(16) Gore, K. R.; Nawale, G. N.; Harikrishna, S.; Chittoor, V. G.; Pandey, S. K.; Höbartner, C.; Patankar, S.; Pradeepkumar, P. I. Synthesis, Gene Silencing, and Molecular Modeling Studies of 4′-C-Aminomethyl-2′-O-methyl Modified Small Interfering RNAs. *J. Org. Chem.* **2012**, *77*, 3233−3245.

(17) Harrison, J. G.; Zheng, Y. B.; Beal, P. A.; Tantillo, D. J. Computational Approaches to Predicting the Impact of Novel Bases on RNA Structure and Stability. *ACS Chem. Biol.* **2013**, *8*, 2354−2359.
(18) Onizuka, K.; Harrison, J. G.; Ball-Jones, A. A.; Ibarra-Soza, J. M.; Zheng, Y.; Ly, D.; Lam, W.; Mac, S.; Tantillo, D. J.; Beal, P. A. Short Interfering RNA Guide Strand Modifiers from Computational Screening. *J. Am. Chem. Soc.* **2013**, *135*, 17069−17077.
(19) Xu, L.; Wang, X.; He, H.; Zhou, J.; Li, X.; Ma, H.; Li, Z.; Zeng, Y.; Shao, R.; Cen, S.; Wang, Y. Structure-Based Design of Novel Chemical Modification of the 3′-Overhang for Optimization of Short Interfering RNA Performance. *Biochemistry* **2015**, *54*, 1268−1277.
(20) Kuhn, C.-D.; Joshua-Tor, L. Eukaryotic Argonautes Come into Focus. *Trends Biochem. Sci.* **2013**, *38*, 263−271.
(21) Elkayam, E.; Kuhn, C.-D.; Tocilj, A.; Haase, A. D.; Greene, E. M.; Hannon, G. J.; Joshua-Tor, L. The Structure of Human Argonaute-2 in Complex with miR-20a. *Cell* **2012**, *150*, 100−110.
(22) Schirle, N. T.; Sheu-Gruttadauria, J.; MacRae, I. J. Structural Basis for microRNA Targeting. *Science* **2014**, *346*, 608−613.
(23) Nakanishi, K.; Weinberg, D. E.; Bartel, D. P.; Patel, D. J. Structure of Yeast Argonaute with Guide RNA. *Nature* **2012**, *486*, 368−374.
(24) Wang, Y.; Juranek, S.; Li, H.; Sheng, G.; Tuschl, T.; Patel, D. J. Structure Of An Argonaute Silencing Complex with a Seed-Containing Guide DNA and Target RNA Duplex. *Nature* **2008**, *456*, 921−926.
(25) Wang, Y.; Sheng, G.; Juranek, S.; Tuschl, T.; Patel, D. J. Structure of the Guide-Strand-Containing Argonaute Silencing Complex. *Nature* **2008**, *456*, 209−213.
(26) Boland, A.; Huntzinger, E.; Schmidt, S.; Izaurralde, E.; Weichenrieder, O. Crystal Structure of the MID-PIWI Lobe of a Eukaryotic Argonaute Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10466−10471.
(27) Wang, Y.; Juranek, S.; Li, H.; Sheng, G.; Wardle, G. S.; Tuschl, T.; Patel, D. J. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **2009**, *461*, 754−761.
(28) 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.
(29) Sheng, G.; Zhao, H.; Wang, J.; Rao, Y.; Tian, W.; Swarts, D. C.; van der Oost, J.; Patel, D. J.; Wang, Y. Structure-Based Cleavage Mechanism of Thermus thermophilus Argonaute DNA Guide Strand-Mediated DNA Target Cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 652−657.
(30) Schirle, N. T.; MacRae, I. J. The Crystal Structure of Human Argonaute2. *Science* **2012**, *336*, 1037−1040.
(31) Frank, F.; Sonenberg, N.; Nagar, B. Structural Basis for 5[Prime]-Nucleotide Base-Specific Recognition of Guide RNA by Human AGO2. *Nature* **2010**, *465*, 818−822.
(32) Song, J.-J.; Smith, S. K.; Hannon, G. J.; Joshua-Tor, L. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* **2004**, *305*, 1434−1437.
(33) Jung, S.-R.; Kim, E.; Hwang, W.; Shin, S.; Song, J.-J.; Hohng, S. Dynamic Anchoring of the 3′-End of the Guide Strand Controls the Target Dissociation of Argonaute−Guide Complex. *J. Am. Chem. Soc.* **2013**, *135*, 16865−16871.
(34) Kwak, P. B.; Tomari, Y. The N domain of Argonaute Drives Duplex Unwinding During RISC Assembly. *Nat. Struct. Mol. Biol.* **2012**, *19*, 145−151.
(35) 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.
(36) Wang, Y.; Li, Y.; Ma, Z.; Yang, W.; Ai, C. Mechanism of MicroRNA-Target Interaction: Molecular Dynamics Simulations and Thermodynamics Analysis. *PLoS Comput. Biol.* **2010**, *6*, e1000866.
(37) Xia, Z.; Clark, P.; Huynh, T.; Loher, P.; Zhao, Y.; Chen, H.-W.; Rigoutsos, I.; Zhou, R. Molecular Dynamics Simulations of Ago Silencing Complexes Reveal a Large Repertoire of Admissible 'Seed-Less' Targets. *Sci. Rep.* **2012**, *2*, 569.
(38) Nam, S.; Ryu, H.; Son, W.-j.; Kim, Y. H.; Kim, K. T.; Balch, C.; Nephew, K. P.; Lee, J. Mg^{2+} Effect on Argonaute and RNA Duplex by Molecular Dynamics and Bioinformatics Implications. *PLoS One* **2014**, *9*, e109745.
(39) Zhu, L.; Jiang, H.; Sheong, F. K.; Cui, X.; Gao, X.; Wang, Y.; Huang, X. A Flexible Domain-domain Hinge Promotes an Induced-fit Dominant Mechanism for the Loading of Guide-DNA into Argonaute Protein in Thermus thermophilus. *J. Phys. Chem. B* **2016**, *120*, 2709−2720.
(40) Schirle, N. T.; Kinberger, G. A.; Murray, H. F.; Lima, W. F.; Prakash, T. P.; MacRae, I. J. Structural Analysis of Human Argonaute-2 Bound to a Modified siRNA Guide. *J. Am. Chem. Soc.* **2016**, *138*, 8694−8697.
(41) Suter, S. R.; Sheu-Gruttadauria, J.; Schirle, N. T.; Valenzuela, R.; Ball-Jones, A. A.; Onizuka, K.; MacRae, I. J.; Beal, P. A. Structure-Guided Control of siRNA Off-Target Effects. *J. Am. Chem. Soc.* **2016**, *138*, 8667−8669.
(42) Zheng, J.; Zhang, L.; Zhang, J.; Wang, X.; Ye, K.; Xi, Z.; Du, Q.; Liang, Z. Single Modification at Position 14 of siRNA Strand Abolishes its Gene-Silencing Activity by Decreasing Both RISC Loading and Target Degradation. *FASEB J.* **2013**, *27*, 4017−4026.
(43) Whitford, P. C.; Blanchard, S. C.; Cate, J. H. D.; Sanbonmatsu, K. Y. Connecting the Kinetics and Energy Landscape of tRNA Translocation on the Ribosome. *PLoS Comput. Biol.* **2013**, *9*, e1003003.
(44) Bock, L. V.; Blau, C.; Schröder, G. F.; Davydov, I. I.; Fischer, N.; Stark, H.; Rodnina, M. V.; Vaiana, A. C.; Grubmüller, H. Energy Barriers and Driving Forces in tRNA Translocation Through the Ribosome. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1390−1396.
(45) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.
(46) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of Polyphosphate Parameters for use with the AMBER Force Field. *J. Comput. Chem.* **2003**, *24*, 1016−1025.
(47) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(48) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886−2902.
(49) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99sb Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(50) Šponer, J.; Banáš, P.; Jurečka, P.; Zgarbová, M.; Kührová, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J. Phys. Chem. Lett.* **2014**, *5*, 1771−1782.
(51) Krepl, M.; Havrila, M.; Stadlbauer, P.; Banas, P.; Otyepka, M.; Pasulka, J.; Stefl, R.; Sponer, J. Can We Execute Stable Microsecond-Scale Atomistic Simulations of Protein−RNA Complexes? *J. Chem. Theory Comput.* **2015**, *11*, 1220−1243.
(52) Pérez-Villa, A.; Darvas, M.; Bussi, G. ATP Dependent NS3 Helicase Interaction with RNA: Insights from Molecular Simulations. *Nucleic Acids Res.* **2015**, *43*, 8725−8734.
(53) Ditzler, M. A.; Otyepka, M.; Šponer, J.; Walter, N. G. Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe. *Acc. Chem. Res.* **2010**, *43*, 40−47.
(54) Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(55) Xu, D.; Jaroszewski, L.; Li, Z.; Godzik, A. AIDA: ab initio Domain Assembly Server. *Nucleic Acids Res.* **2014**, *42*, W308−W313.
(56) Ipsaro, J. J.; Joshua-Tor, L. From Guide To Target: Molecular Insights into Eukaryotic RNA-Interference Machinery. *Nat. Struct. Mol. Biol.* **2015**, *22*, 20−28.

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

(57) Gan, H. H.; Gunsalus, K. C. Assembly and Analysis of Eukaryotic Argonaute−RNA Complexes in MicroRNA-Target Recognition. *Nucleic Acids Res.* **2015**, *43*, 9613−9625.
(58) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 14*; University of California, San Francisco, 2014.
(59) Rother, M.; Milanowska, K.; Puton, T.; Jeleniewicz, J.; Rother, K.; Bujnicki, J. M. ModeRNA Server: an Online Tool For Modeling RNA 3D structures. *Bioinformatics* **2011**, *27*, 2441−2442.
(60) Macindoe, G.; Mavridis, L.; Venkatraman, V.; Devignes, M.-D.; Ritchie, D. W. HexServer: an FFT-based Protein Docking Server Powered By Graphics Processors. *Nucleic Acids Res.* **2010**, *38*, W445−W449.
(61) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177−6196.
(62) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(63) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−2791.
(64) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. PDB2PQR: Expanding and Upgrading Automated Preparation of Biomolecular Structures for Molecular Simulations. *Nucleic Acids Res.* **2007**, *35*, W522−W525.
(65) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion−Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 1493−1502.
(66) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2013.
(67) Aduri, R.; Psciuk, B. T.; Saro, P.; Taniga, H.; Schlegel, H. B.; SantaLucia, J. AMBER Force Field Parameters for the Naturally Occurring Modified Nucleosides in RNA. *J. Chem. Theory Comput.* **2007**, *3*, 1464−1475.
(68) Lind, K. E.; Ferguson, D. M.; Mohan, V.; Manoharan, M. Structural Characteristics of 2′-O-(2-Methoxyethyl)-Modified Nucleic Acids From Molecular Dynamics Simulations. *Nucleic Acids Res.* **1998**, *26*, 3694−3699.
(69) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform In Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(70) 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.
(71) Krepl, M.; Réblová, K.; Koča, J.; Šponer, J. Bioinformatics and Molecular Dynamics Simulation Study of L1 Stalk Non-Canonical rRNA Elements: Kink-Turns, Loops, and Tetraloops. *J. Phys. Chem. B* **2013**, *117*, 5540−5555.
(72) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(73) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(74) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374−380.
(75) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to An External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(76) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(77) Lu, X.-J.; Olson, W. K. 3DNA: A Versatile, Integrated Software System for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic-Acid Structures. *Nat. Protoc.* **2008**, *3*, 1213−1227.
(78) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein−Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras−Raf and Ras−RalGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891−913.
(79) Shindyalov, I. N.; Bourne, P. E. A Database and Tools for 3-D Protein Structure Comparison and Alignment using the Combinatorial Extension (CE) Algorithm. *Nucleic Acids Res.* **2001**, *29*, 228−229.
(80) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(81) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(82) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412−425.
(83) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(84) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **2007**, *3*, 2312−2334.
(85) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69−82.
(86) Reyes, C. M.; Kollman, P. A. Structure and Thermodynamics of RNA-Protein Binding: Using Molecular Dynamics and Free Energy Analyses to Calculate the Free Energies of Binding and Conformational Change. *J. Mol. Biol.* **2000**, *297*, 1145−1158.
(87) Ming, D.; Wall, M. E.; Sanbonmatsu, K. Y. Domain Motions of Argonaute, the Catalytic Engine of RNA Interference. *BMC Bioinf.* **2007**, *8*, 1−11.
(88) Rashid, U. J.; Paterok, D.; Koglin, A.; Gohlke, H.; Piehler, J.; Chen, J. C.-H. Structure of Aquifex aeolicus Argonaute Highlights Conformational Flexibility of the PAZ Domain as a Potential Regulator of RNA-induced Silencing Complex Function. *J. Biol. Chem.* **2007**, *282*, 13824−13832.
(89) Mládek, A.; Banáš, P.; Jurečka, P.; Otyepka, M.; Zgarbová, M.; Šponer, J. Energies and 2′-Hydroxyl Group Orientations of RNA

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

Backbone Conformations. Benchmark CCSD(T)/CBS Database, Electronic Analysis, and Assessment of DFT Methods and MD Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 463−480.
(90) Sponer, J.; Mladek, A.; Sponer, J. E.; Svozil, D.; Zgarbova, M.; Banas, P.; Jurecka, P.; Otyepka, M. The DNA and RNA Sugar-Phosphate Backbone Emerges as the Key Player. An Overview of Quantum-Chemical, Structural Biology and Simulation Studies. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15257−15277.
(91) Deleavey, G. F.; Frank, F.; Hassler, M.; Wisnovsky, S.; Nagar, B.; Damha, M. J. The 5′ Binding MID Domain of Human Argonaute2 Tolerates Chemically Modified Nucleotide Analogues. *Nucleic Acid Ther.* **2013**, *23*, 81−87.
(92) Zlatev, I.; Foster, D. J.; Liu, J.; Charisse, K.; Brigham, B.; Parmar, R. G.; Jadhav, V.; Maier, M. A.; Rajeev, K. G.; Egli, M.; Manoharan, M. 5′-C-Malonyl RNA: Small Interfering RNAs Modified With 5′-Monophosphate Bioisostere Demonstrate Gene Silencing Activity. *ACS Chem. Biol.* **2016**, *11*, 953−960.
(93) Prakash, T. P.; Lima, W. F.; Murray, H. M.; Li, W.; Kinberger, G. A.; Chappell, A. E.; Gaus, H.; Seth, P. P.; Bhat, B.; Crooke, S. T.; Swayze, E. E. Identification of Metabolically Stable 5′-Phosphate Analogs that Support Single-Stranded siRNA Activity. *Nucleic Acids Res.* **2015**, *43*, 2993−3011.
(94) Gore, K. R.; Harikrishna, S.; Pradeepkumar, P. I. Influence of 2′-Fluoro versus 2′-O-Methyl Substituent on the Sugar Puckering of 4′-C-Aminomethyluridine. *J. Org. Chem.* **2013**, *78*, 9956−9962.
(95) Mutisya, D.; Selvam, C.; Lunstad, B. D.; Pallan, P. S.; Haas, A.; Leake, D.; Egli, M.; Rozners, E. Amides are Excellent Mimics of Phosphate Internucleoside Linkages and are Well Tolerated in Short Interfering RNAs. *Nucleic Acids Res.* **2014**, *42*, 6542−6551.
(96) Jo, M. H.; Shin, S.; Jung, S.-R.; Kim, E.; Song, J.-J.; Hohng, S. Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. *Mol. Cell* **2015**, *59*, 117−124.
(97) Sashital, D. G.; Doudna, J. A. Structural Insights into RNA Interference. *Curr. Opin. Struct. Biol.* **2010**, *20*, 90−97.
(98) CHIU, Y.-L.; RANA, T. M. siRNA Function in RNAi: A Chemical Modification Analysis. *RNA* **2003**, *9*, 1034−1048.
(99) Bumcrot, D.; Manoharan, M.; Koteliansky, V.; Sah, D. W. Y. RNAi Therapeutics: A Potential New Class of Pharmaceutical Drugs. *Nat. Chem. Biol.* **2006**, *2*, 711−719.
(100) Martinez-Montero, S.; Deleavey, G. F.; Martin-Pintado, N.; Fakhoury, J. F.; Gonzalez, C.; Damha, M. J. Locked 2′-Deoxy-2′,4′-Difluororibo Modified Nucleic Acids: Thermal Stability, Structural Studies, and siRNA Activity. *ACS Chem. Biol.* **2015**, *10*, 2016−202.

DOI: 10.1021/acs.jcim.6b00773
J. Chem. Inf. Model. 2017, 57, 883−896

Supporting Information

Probing the Binding Interactions between Chemically Modified siRNAs and Human Argonaute 2 Using Microsecond Molecular Dynamics Simulations

S. Harikrishna* and P. I. Pradeepkumar*

Department of Chemistry, Indian Institute of Technology Bombay, Mumbai – 400076 India

*E-mail: harikrishna.s@iitb.ac.in or pradeep@chem.iitb.ac.in

Table of Contents

Analysis of domain orientations in AGO protein. This analysis was performed in order to gather the domain distance parameters required to generate an open conformational model of hAGO2. The crystal structure of hAGO2 reported in complex with only guide RNA/miRNA (PDB entry: 4F3T) in which the 3'-end of the RNA interacts with the PAZ domain (Figure M1A). In this structure, the distance between the PIWI and PAZ (central cleft) domain (24.9 Å) is measured as a distance between the C-alpha atoms of Q632 and K263. The distance between the MID and PAZ (21.7 Å) domain is measured as a distance between the C-alpha atoms of R527 and T251. The distance between the N-terminal and PAZ domain (18.4 Å) is measured as a distance between the C alpha atoms of P67 and V330. Structure of hAGO2 reported in complex with guide-target (2-9 nt base pairing; PDB entry: 4W5O), in which the 3'-end of the RNA is interacting at the PAZ domain (Figure M1B). The domain distances were calculated using the same C-alpha atoms as in the hAGO2-guide RNA/miRNA complex structures. The distance between the MID and PAZ, PIWI and PAZ domains were increased from 8 to 10 Å upon the binding of target RNAs. In *Thermus thermophilus* AGO (*Tt*AGO) in complex with guide: target DNA (2-19 base pairing; PDB entry: 4NCB), the distance between the PIWI and PAZ (central cleft) domain (24.9 Å) is measured as a distance between the C-alpha atoms of S572 and T201 (Figure M1C). The distance between the MID and PAZ domain (27 Å) is measured as a distance between the C-alpha atoms of S572 and T201. The distance between the N-terminal and PAZ domain (33 Å) is measured as a distance between the C-alpha atoms of G61 and K230. The accommodation of duplex in *Tt*AGO increases the distances between the N-terminal and PAZ around 15 Å, but the accommodation of guide: target DNA duplex in hAGO2 increases the distance between PIWI and PAZ domain around 12 Å (Figure M1D). Structural alignment of hAGO2 in complex with guide-target (2-9 nt base pairing; PDB entry: 4W5O) and *Tt*AGO in complex guide: target DNA (2-19 nt base pairing; PDB entry: 4NCB) shows that the major conformational changes are in the PAZ domain (Figure M1E). The RMSD between the two structure is 4.5 Å. The significant difference was seen in the orientation of PAZ and N-terminal domains in the enzymes (Figure M1F). All these structural alignments were carried out using DALI algorithm.[1]

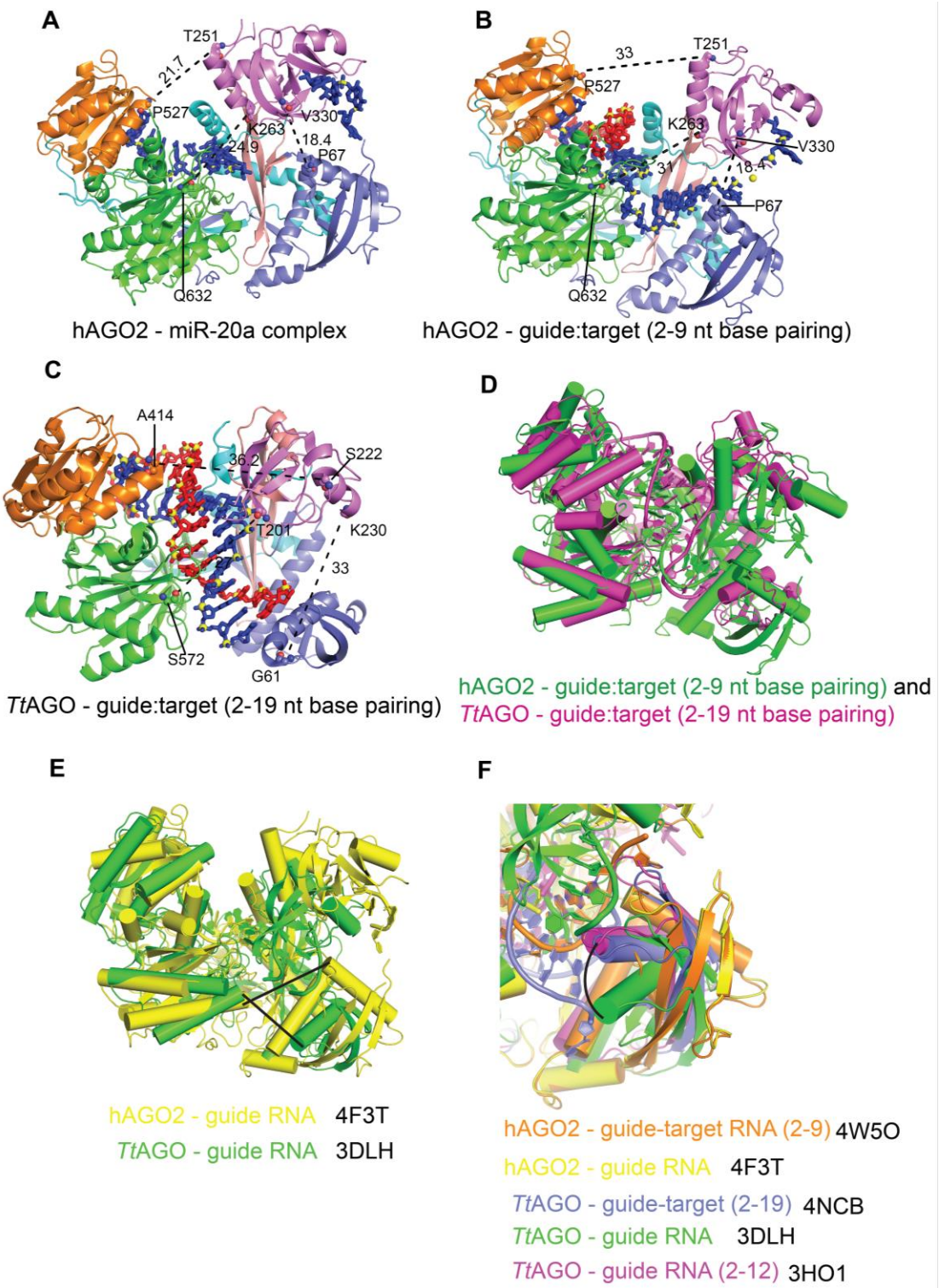
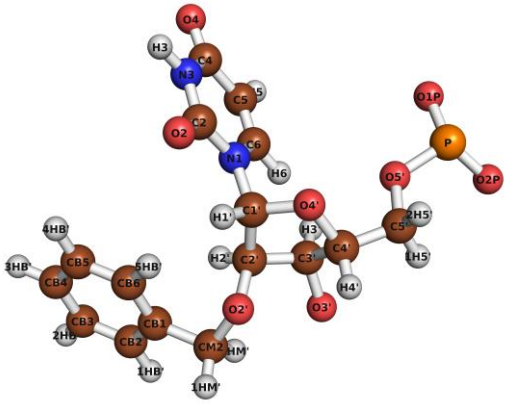


Figure M1. Distances between the domains and structural alignment of the argonaute proteins, N-terminal (blue), MID (orange), PIWI (green), PAZ (magenta), L1 (wheatish), L2 (cyan) domains and the guide or miRNA (blue) and the target RNA (red). The protein is represented in cartoon with ribbon or cylindrical helices. RNA is represented in spheres and sticks. (**A**) The guide RNA/miRNA-hAGO2 complex (PDB entry: 4F3T), in which the 3'-end of the RNA is interacting at the PAZ domain. (**B**) The guide-target RNA-hAGO2 complex (2-9 nt base pairing; PDB entry: 4W5O), in which the 3'-end of the RNA is interacting at the PAZ domain. (**C**) The guide: target DNA-*Thermus thermophilus* AGO complex (2-19 nt base pairing; PDB entry: 4NCB). (**D**) Structural alignment of guide:target-hAGO2

complex (2-9 nt base pairing; green; PDB entry: 4W5O) and guide: target DNA-*Tt*AGO complex (2-19 base pairing; magenta; PDB entry: 4NCB). (**E**) Structural alignment of guide RNA-hAGO2 complex (yellow; PDB entry: 4F3T) and guide RNA-*Tt*AGO complex (green; PDB entry: 3DLH). Black lines indicate the distance between the PIWI and N-terminal domain in hAGO2 and *Tt*AGO. (**F**) Structural alignment of *Tt*AGO and hAGO2 to show the conformational changes (black line) observed upon target RNA binding to the guide RNA in *Tt*AGO. Similar conformational changes were not observed in the hAGO2 during target mRNA binding.[2]

Energy optimized structure and RESP charge of 2'-*O*-benzyl nucleotide



Atom name	ESP charge	Atom name	ESP charge
P	1.104552	O2'	− 0.261245
O1P	− 0.679190	CM2	− 0.078544
O2P	− 0.679190	1HM'	0.051256
O5'	− 0.286067	2HM'	0.051256
C5'	− 0.166710	CB1	0.154247
1H5'	0.052807	CB2	− 0.188974
2H5'	0.052807	1HB'	0.140662
C4'	0.352368	CB3	− 0.159608
H4'	0.045827	2HB'	0.131426
O4'	− 0.540710	CB4	− 0.155290
C1'	0.248076	3HB'	0.122372
H1'	0.112758	CB5	− 0.159608
N1	0.121107	4HB'	0.131426
C6	− 0.363940	CB6	− 0.188974
H6	0.235136	5HB'	0.140662
C5	− 0.302491	C3'	0.065419
H5	0.169742	H3'	0.146812
C4	0.602168	O3'	− 0.421742
O4	− 0.574387		
N3	− 0.364529		
H3	0.319754		
C2	0.453765		
O2	− 0.543213		
C2'	0.015673		
H2'	0.092313		

Figure S1. Energy optimized geometry and calculated RESP charges for the 2'-*O*-benzyl (2'-OB) nucleotide using HF/6-31G* basis set in Gaussian 09 program.[3-5]

Docked structure and RMSD graph of the siRNA3-hAGO2 complex

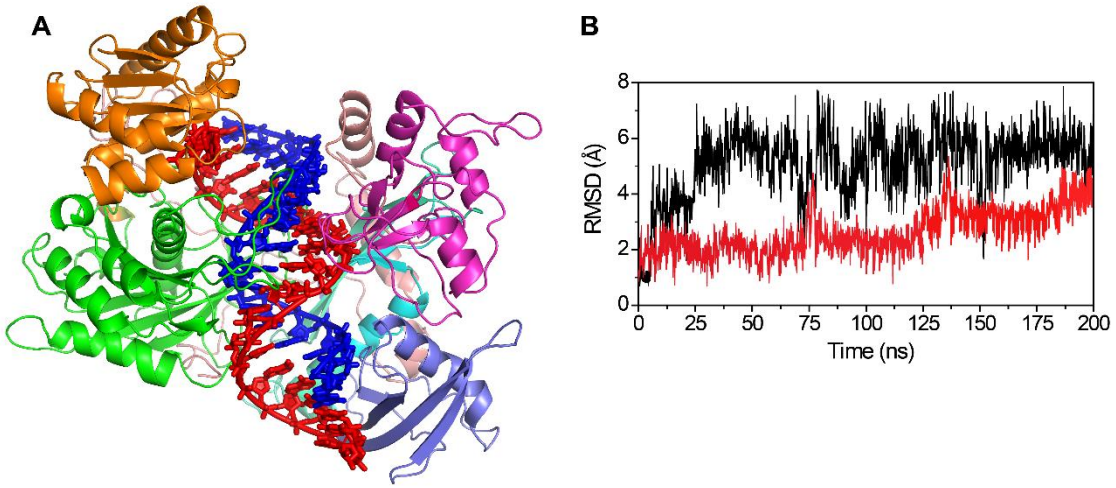


Figure S2. Docked structure and RMSD graph from the MD simulations of siRNA3 (22 nt) and hAGO2 complex. **(A)** Docked structure of the unmodified siRNA3-hAGO2 complex. Protein is represented in cartoon, RNA backbone in cartoon and atoms in sticks. siRNA guide strand is shown in red and the passenger strand is shown in blue. (**B**) RMSD graph of hAGO2 (black) and siRNA (red) from the MD simulations. RMSD is calculated using the equilibrated structure as the reference.

Model of siRNA3-hAGO2 complex emerged from the MD simulations

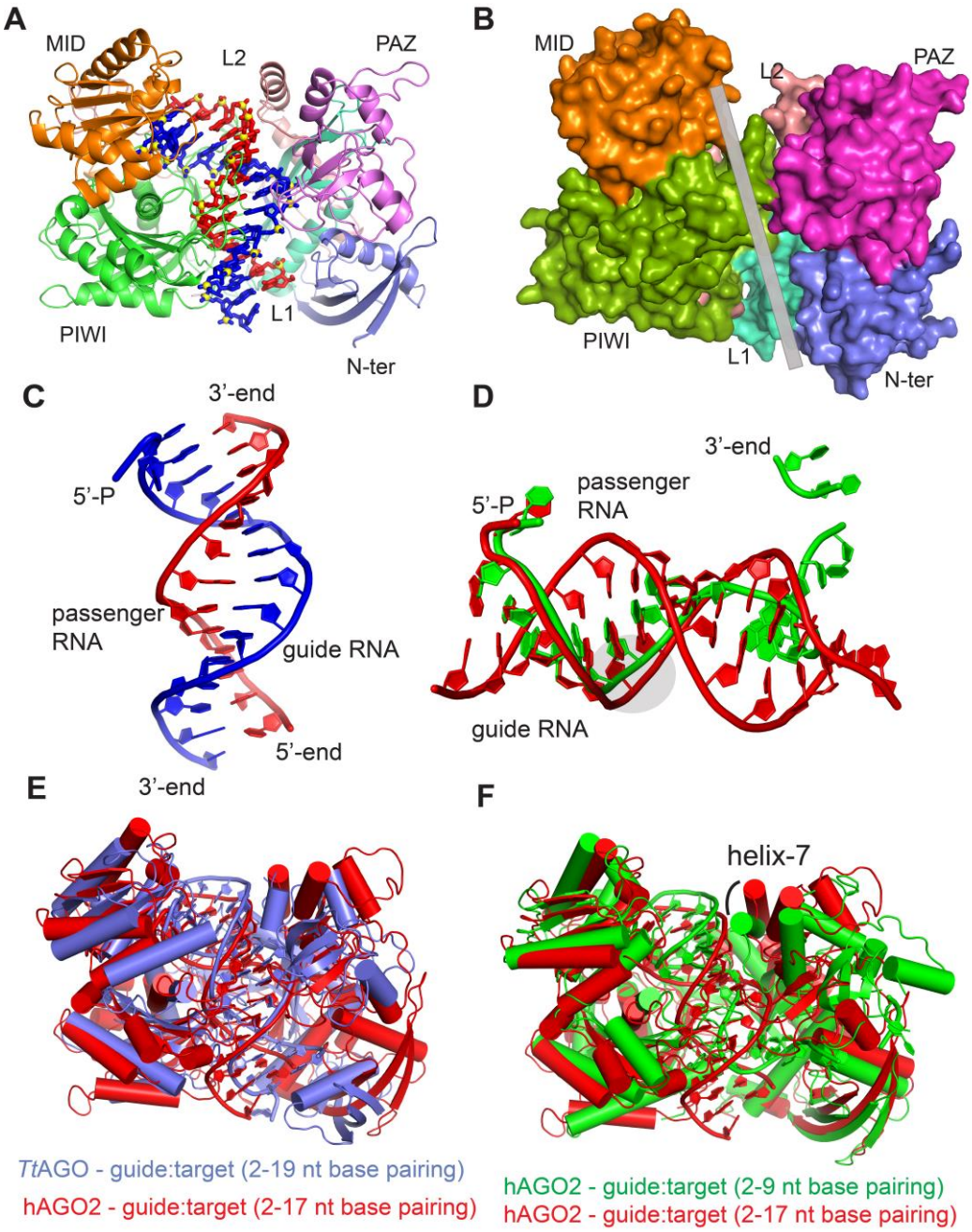


Figure S3. Model of siRNA3-hAGO2 complex emerged from the MD simulations. (**A**) Overall structure of siRNA3-hAGO2 complex. Protein is represented in cartoon and siRNA is represented in sticks and spheres. (**B**) Best representative structure of the hAGO2 protein, in which the siRNA was removed to show the central nucleic acid binding channel. The tunnel is highlighted in grey. Protein is represented in surface. (**C**) Best representative structure of siRNA3 to show the A-type RNA duplex. siRNA is represented in cartoon (**D**) Superimposed crystal structure of siRNA guide strand (green, PDB entry: 4W5O) and the siRNA3 (red) emerged from the MD simulations. (**E**) Structural alignment of *Tt*AGO crystal structure (blue, PDB entry: 4NCB) and hAGO2 model (red). (**G**) Structural alignment of hAGO2 crystal structure (green, PDB entry: 4W5O) and hAGO2 model (red).

Structural comparison between the hAGO2 model and the crystal structure

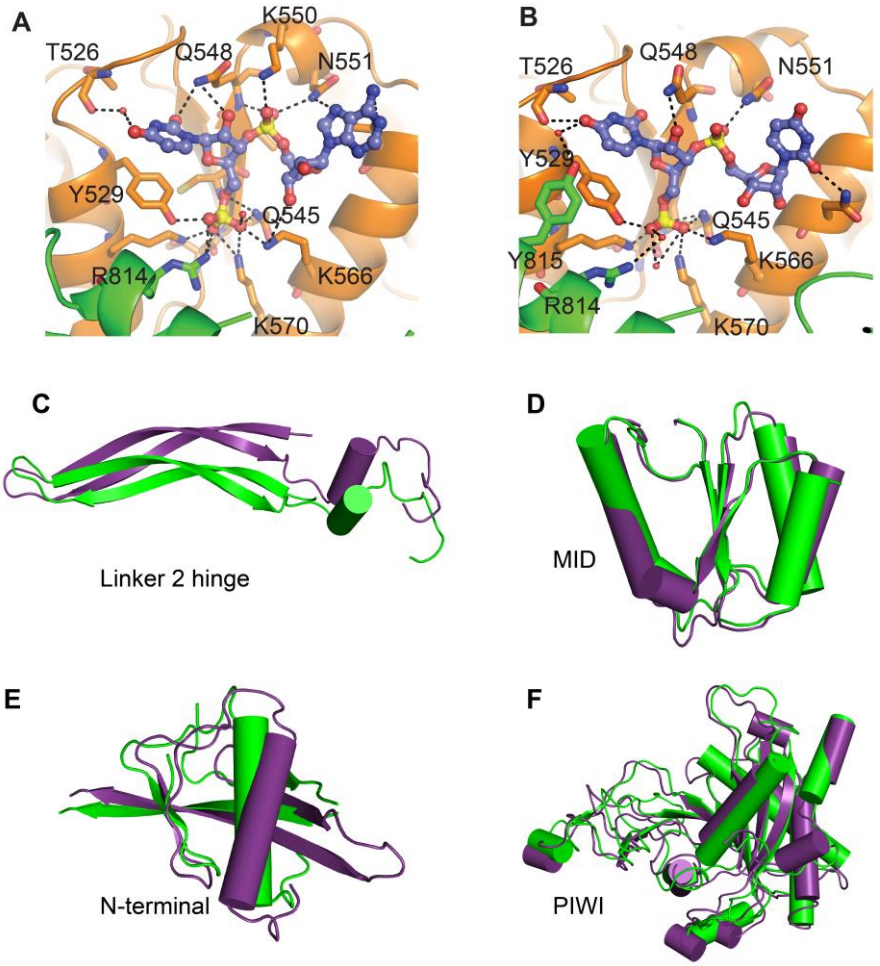


Figure S4. Architecture of the 5'-binding pocket in the MID domain and the structural alignment of domains between the crystal structure of hAGO2 (green, 2-9 nt base pairing, PDB entry: 4W5O) and the hAGO2 model (purple) emerged from the MD simulations. The 5'-guide RNA binding pocket in the MID domain of the protein averaged from the last 200 ns of the MD simulations of (**A**) siRNA2-hAGO2 complex and (**B**) the hAGO2 crystal structure. Protein is represented in ribbon, interacting atoms are represented in sticks. RNA is represented in stick and spheres. Water molecules are shown as small non-bonded red coloured spheres. Inter molecular H-bonds are shown in dashed lines. Water mediated contacts made by R814 with the 5'-phosphate in which the side chain is significantly rotated. All the intermolecular contacts are ≤ 3.3 Å and are present during 98 % of the total simulation time. (**C**) Structural alignment of the linker 2, which acts as a hinge and shifts around 4 Å to accommodate the target RNA binding. Protein domains are represented in cartoon. (**D**) Alignment of the MID domain (RMSD ~1.5 Å) shows the rigidity of the domain indicated by the less conformational changes. (**E**) Alignment of N-terminal domain (RMSD ~2.3 Å) shows the moderate flexibility of the domain. (**F**) Alignment of PIWI domain(RMSD ~3 Å).

Schematic representation of the noncovalent interactions in unmodified siRNA-hAGO2 complex

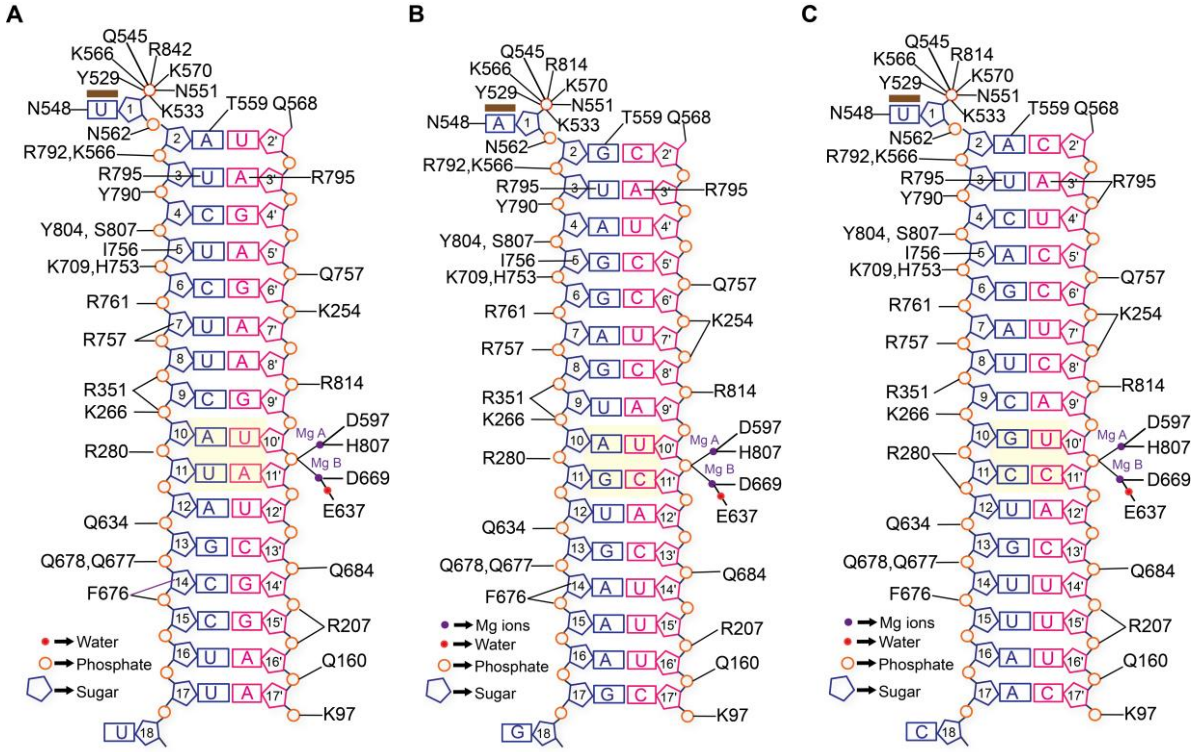


Figure S5. Schematic representation of the siRNA-hAGO2 noncovalent interactions observed from the MD simulations of (**A)** siRNA3-hAGO2, (**B**) siRNA2/siRNA5-hAGO2 and (**C**) siRNA14-hAGO2 complexes. The interactions represented here are present >40 % of the last 800 ns MD simulations. All other interactions having < 25 % occupancies were discarded.

H-bond interactions in the siRNA3

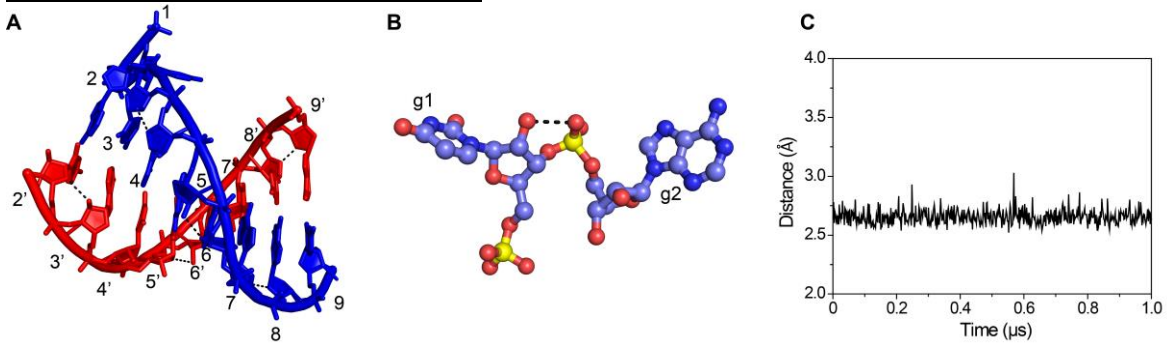


Figure S6. Intramolecular H-bond interactions in the siRNA3 during the course of MD simulations of siRNA3-hAGO2 complex. (**A**) H-bonds between 2'-OH and O4' in the consecutive nucleotides represented by black dashed lines. siRNA is represented in cartoon, guide strand in blue and the passenger strand in red. (**B**) H-bond interaction between 2'-OH and oxygen (OP1) between g1 and g2 indicated by black dashed lines. (**C**) Time-dependence of H-bond distance between g1 (2'-OH) and g2 (5'-phosphate) as shown in Figure S6B. The average H-bond distance is 2.6 Å.

Local base pair parameters of siRNA2/5

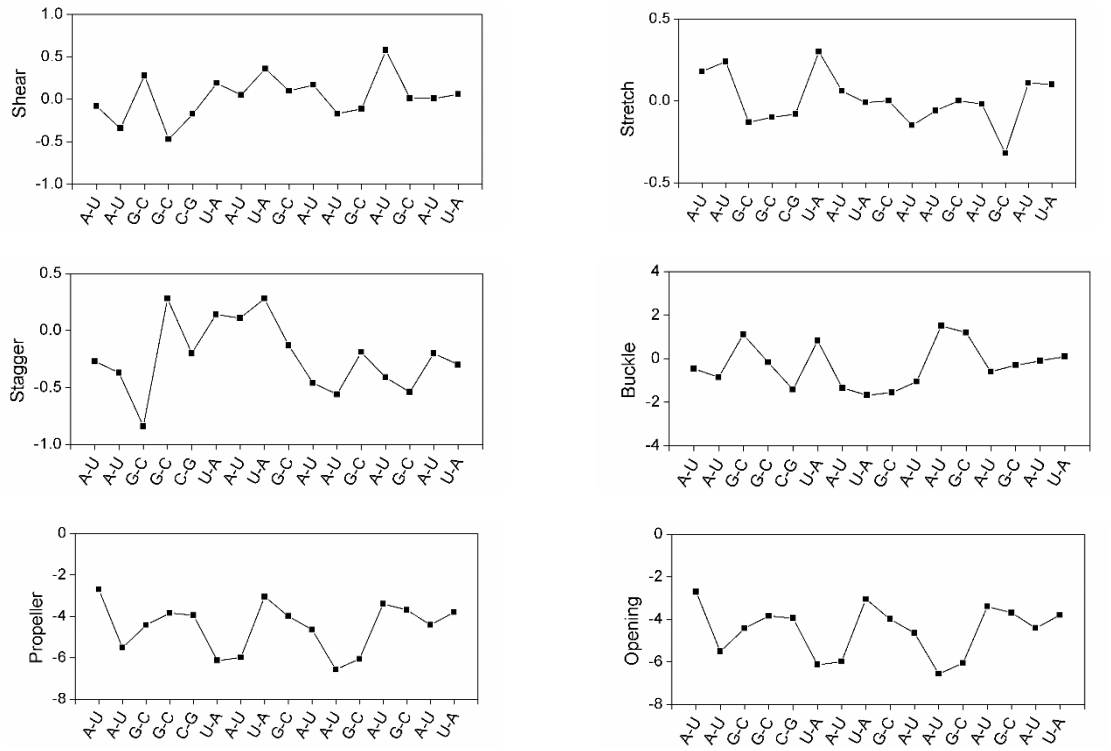


Figure S7. Local RNA base pair parameters from the MD simulations of the unmodified siRNA2/siRNA5-hAGO2 complex. The parameters are calculated from the ensemble averaged structure from the last 200 ns of the MD simulations using X3DNA. These parameters for the unmodified siRNA3 and siRNA14 were found to be similar with those of the unmodified siRNA5. The translational intra base pair helical parameters such as stretch, shear and stagger are mentioned in Å. And the rotational intra base pair helical parameters such as propeller, buckle and opening are mentioned in degrees.

Principal component and cluster analysis of the hAGO2

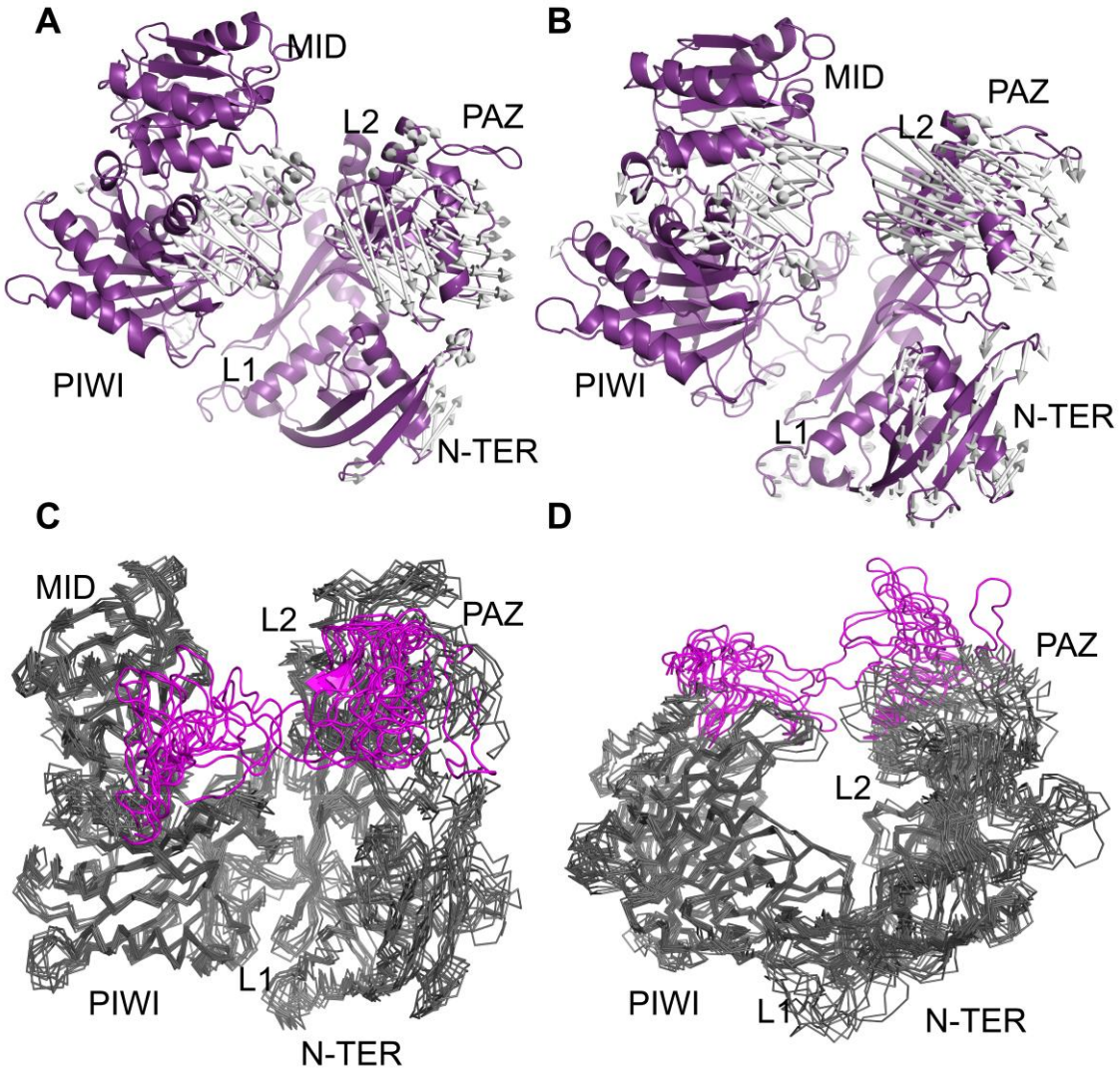


Figure S8. Principal component and cluster analysis of the hAGO2 and the unmodified siRNAs from the MD simulations. (**A**) First principal component mode of motion for the protein. Similar results were seen in RMSF of the protein (Figure 2C, main text). Arrows indicate direction of the motion and the length of the arrows indicate the amplitude of the motion. The arrows in L1 and PAZ domain indicates that 50% motion in the first principal component. Protein is represented in cartoon. (**B**) Second principal component mode of motion of the protein. The first two principal component of the protein motion corresponds to the 80% motion of the MD trajectory. The most rigid domains are MID and PIWI. (**C**) Best representative structures from the 10 clusters are shown to depict the stability and motion of the protein domains. Protein is represented in ribbon. The disordered loops in PAZ and PIWI domains are highlighted in magenta. (**D**) Rotated (55°) view of the best representative structures from the 10 clusters, which reveal the nucleic acid binding channel.

Time-dependence of key H-bond/electrostatic interaction distances in the siRNA3-hAGO2 complexes

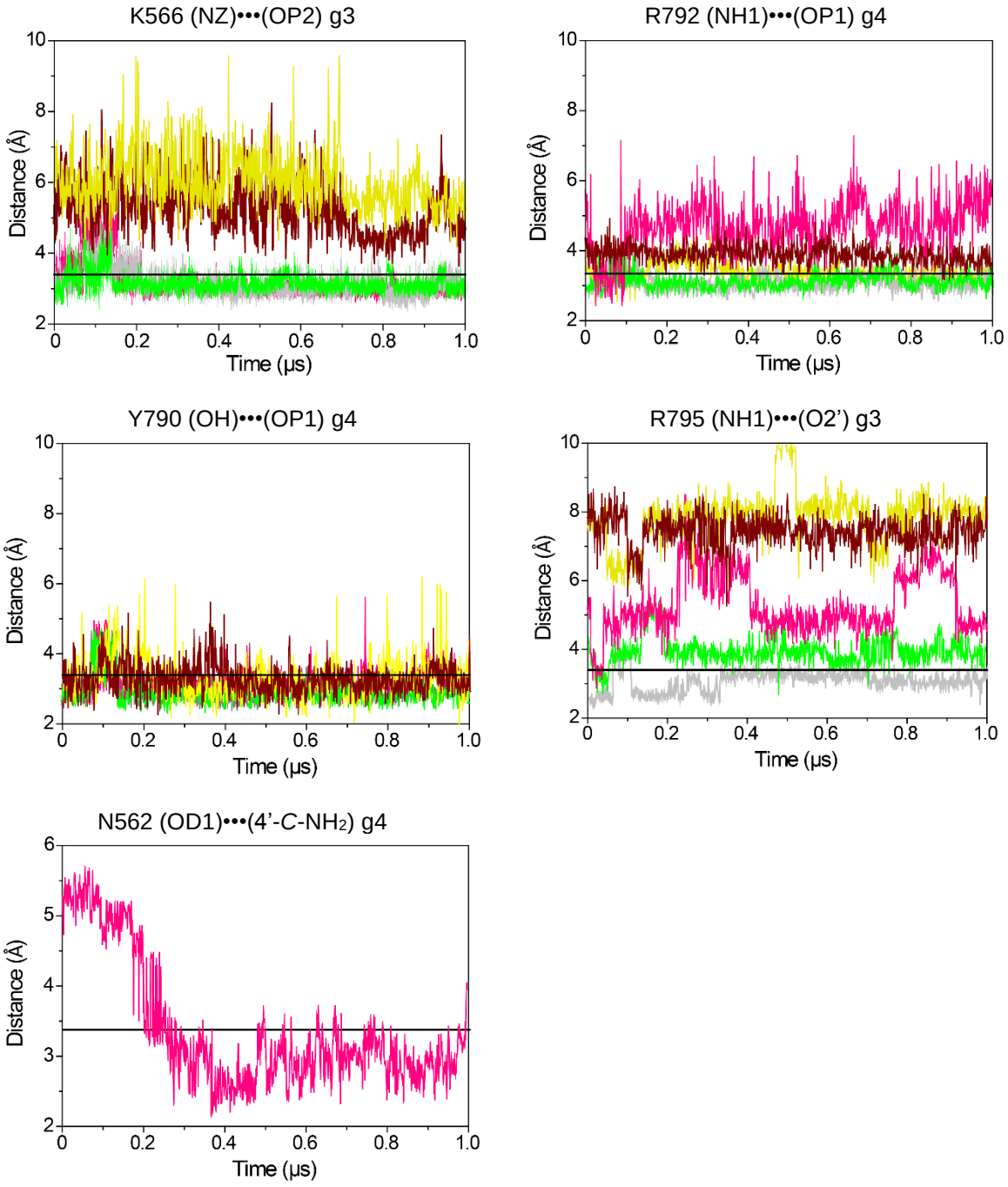


Figure S9. Time-dependence of key H-bond distances during the course of MD simulations of the siRNA3-hAGO2 complex. The H-bonds are measured as the distances between the heavy atoms as mentioned above the graph. Black line indicates the H-bond distance with a cut-off value of 3.3 Å. The distances in the unmodified (grey), 2'-OMe (green), 4'-AM-2'-OMe (magenta), 2'-MOE (yellow), and 2'-OB (brown) modified siRNAs are highlighted using different colours.

Electrostatic and van der Waals energies of the siRNA3-hAGO2 complexes

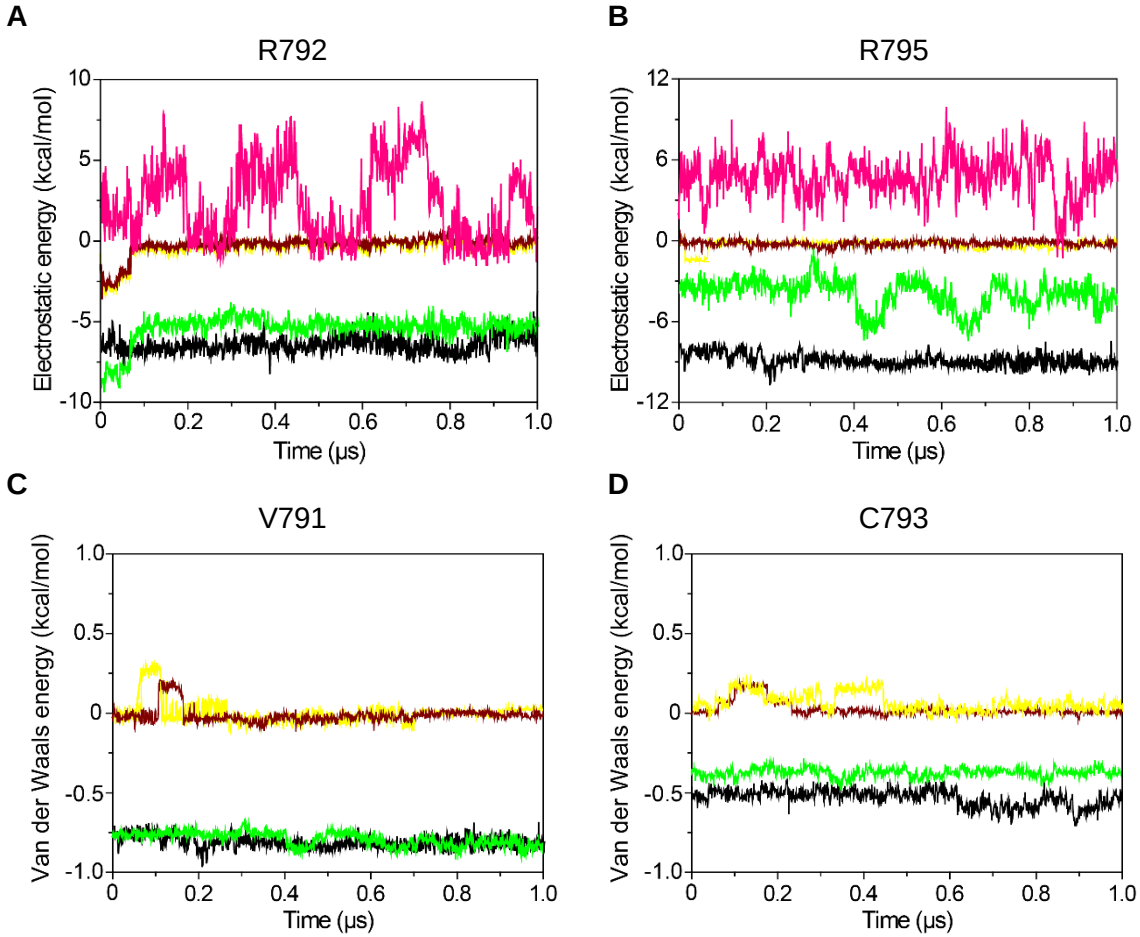


Figure S10. Electrostatic and van der Waals interaction energies betweeen g3 in siRNA3 and AAs in hAGO2 during the course of MD simulations. (**A**) Electrostatic interaction energies between g3 of siRNA and R792 in the PIWI domain. (**B**) Electrostatic interaction energies between g3 of siRNA and R795 in the PIWI domain. (**C**) Van der Waals interaction energies between g3 of siRNA and V791 in the PIWI domain. (**D**) Van der Waals interaction energies between g3 of siRNA and C793 in the PIWI domain of hAGO2. The energies in the unmodified (grey), 2'-OMe (green), 4'-AM-2'-OMe (magenta), 2'-MOE (yellow), and 2'-OB (brown) modified siRNAs are highlighted using different colours. The interacting AAs are mentioned above the graph.

MD snapshots around g3 of the unmodified and the modified siRNA3-hAGO2 complexes

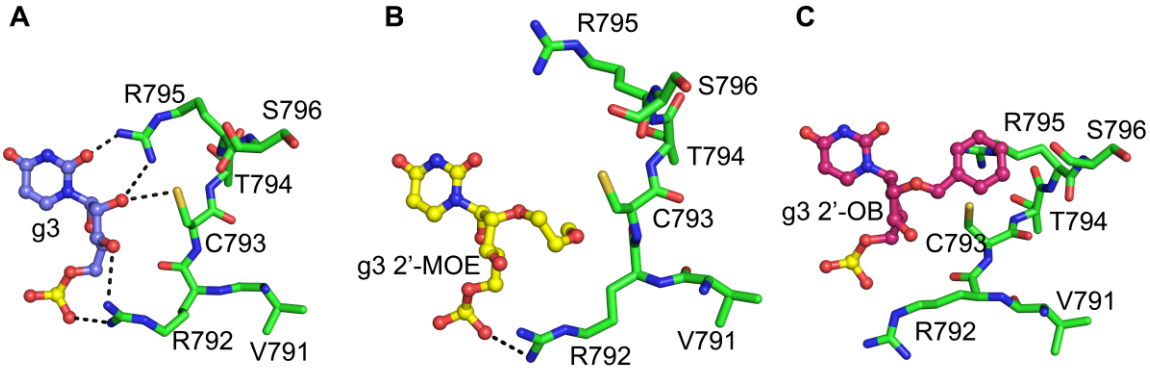


Figure S11. Averaged MD snapshots from the last 200 ns of simulations depicting the key interactions between and position g3 in the siRNA3 and the AAs in the PIWI domain. (**A**) Unmodified (blue), (**B**) 2'-MOE (yellow), and (**C**) 2'-OB (brown) modified siRNAs. Only the interacting AAs are shown for clarity. AAs in the protein are represented in green sticks. RNA atoms and bonds are represented in spheres and sticks, respectively. The unmodified and the modified nucleotides are highlighted using different colours. The black dashed lines indicate the non-covalent interactions.

Phosphate distance between g1 and g3 in the siRNA3-hAGO2 complexes

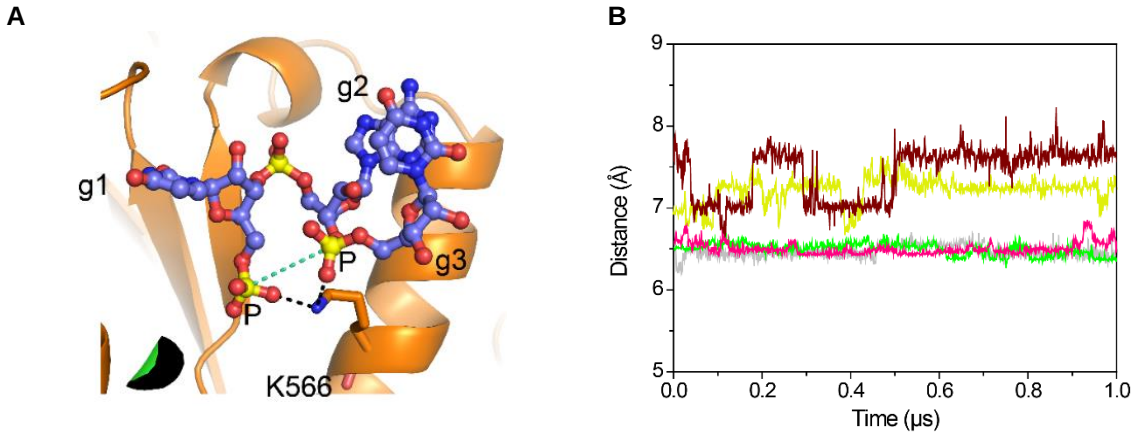


Figure S12. Phosphate distance between the guide strand nucleotides (g1 and g3) in the unmodified and in the modified siRNA3-hAGO2 during the course of MD simulations. (**A**) Electrostatic interaction between lysine (K566) in the MID domain and the phosphate backbone in g1 and g3 of siRNA. Black dashed lines indicate the electrostatic contacts. Cyan dashed line indicates the distance between the two phosphate atoms in g1 and g3. Protein is represented in cartoon and RNA is represented in spheres and sticks. (**B**) Time-dependence of distance between the phosphate atoms in g1 and g3 during the course of MD simulations. The distances in the unmodified (grey), 2'-OMe (green), 4'-AM-2'-OMe (magenta), 2'-MOE (yellow), and 2'-OB (brown) modified siRNAs are highlighted using different colours.

Time-dependence of key H-bond/electrostatic interaction distances in siRNA2-hAGO2 complexes

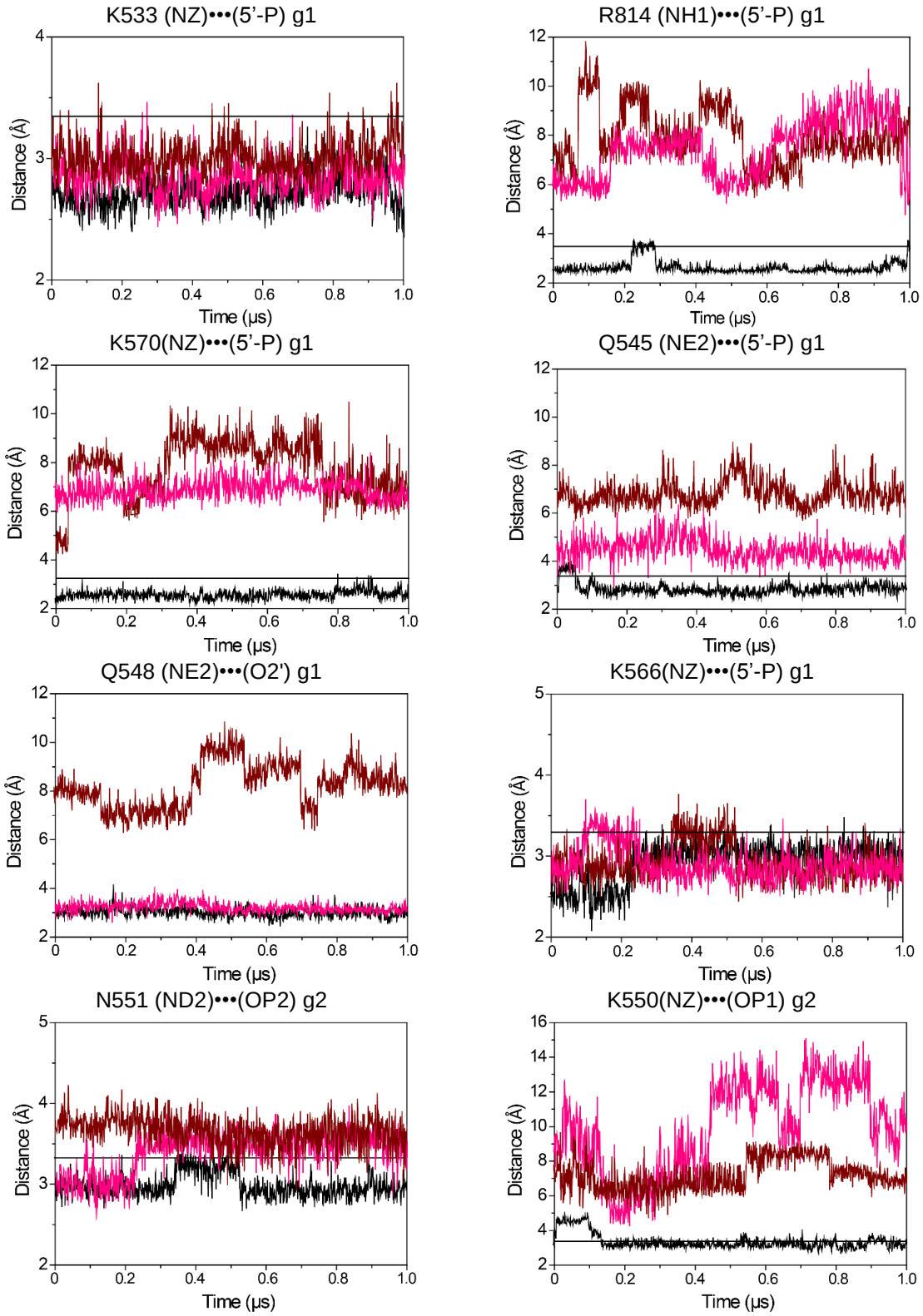


Figure S13. Time-dependence of key H-bond distances during the course of MD simulations of siRNA2 -hAGO2 complexes. The H-bond is measured as the distance between the heavy atoms as mentioned above the graphs. Black line indicates the H-bond distance with a cut-off value of 3.3 Å. The distances in the unmodified (black), 4'-AM-2'-OMe (magenta), and 2'-OB (brown) modified siRNAs are highlighted using different colours.

Time-dependence of key H-bond/electrostatic interaction distances in the siRNA5-hAGO2 complexes

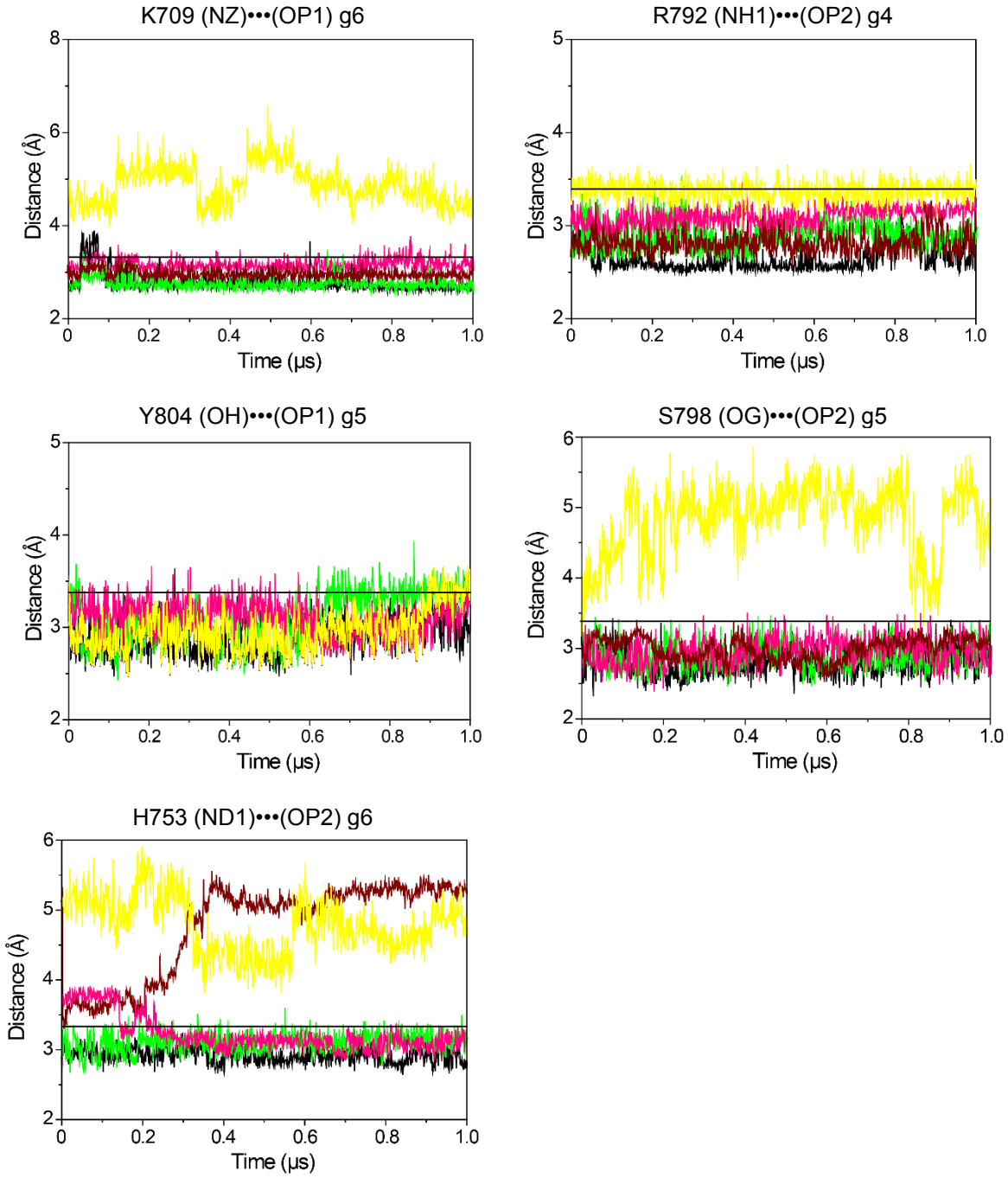


Figure S14. Time-dependence of key H-bond distances during the MD simulations of siRNA5-hAGO2 complexes. The H-bond is measured as the distance between the heavy atoms as mentioned above the graphs. Black line indicates the H-bond distance with a cut-off value of 3.3 Å. The distances in the unmodified (black), 2'-OMe (green), 4'-AM-2'-OMe (magenta), 2'-MOE (yellow), and 2'-OB (brown) modified siRNAs are highlighted using different colours.

Van der Waals interaction energies between 2'-OB modification in the siRNA5 and amino acids in the PIWI domain of hAGO2

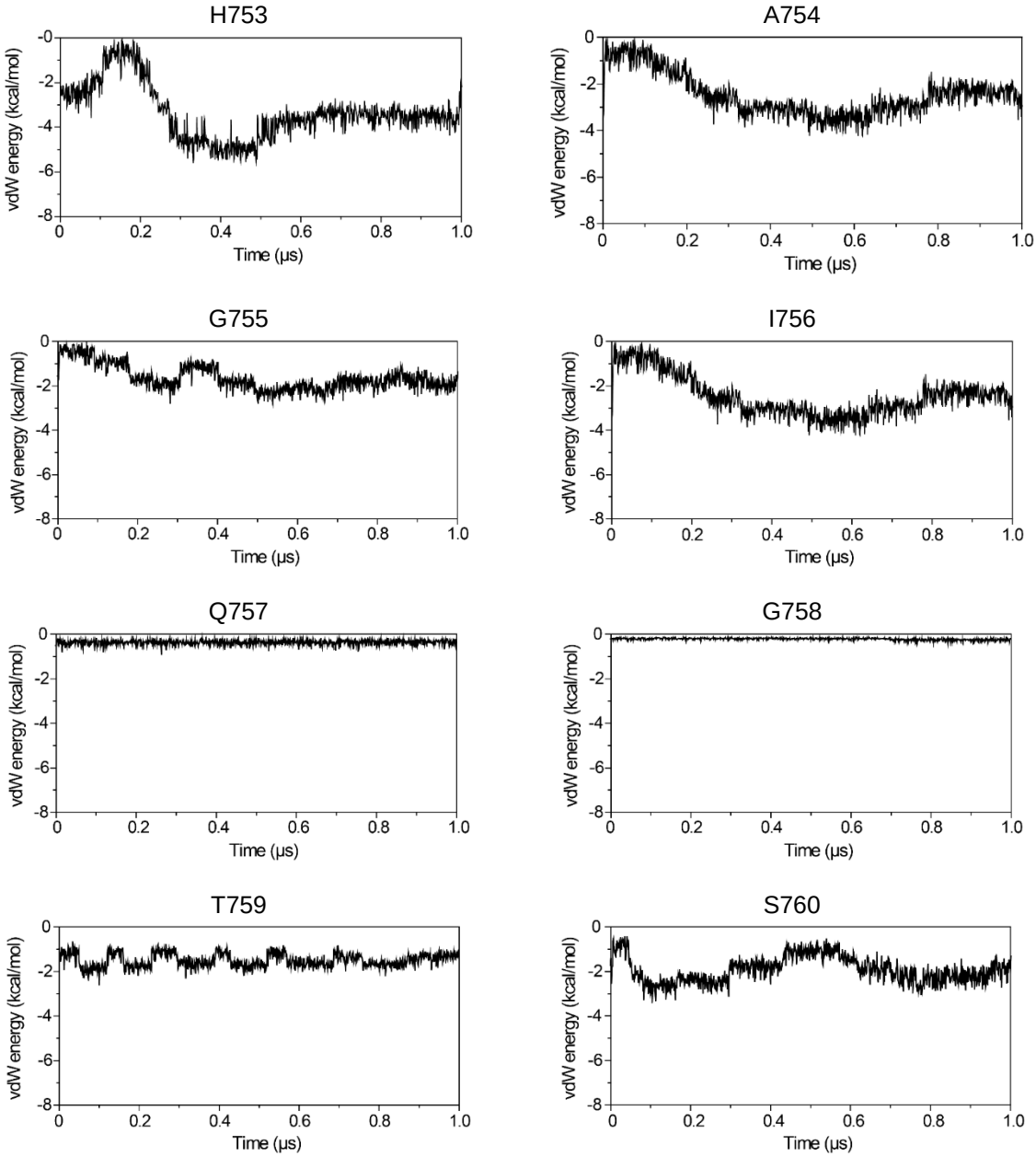


Figure S15. Van der Waals interaction energies between 2'-OB and the AAs in the PIWI domain during the course of MD simulations of siRNA5-hAGO2 complex. The interacting AAs are mentioned above the graphs.

MD snapshots of position g5 in the unmodified and in the modified siRNA5 -hAGO2 complexes

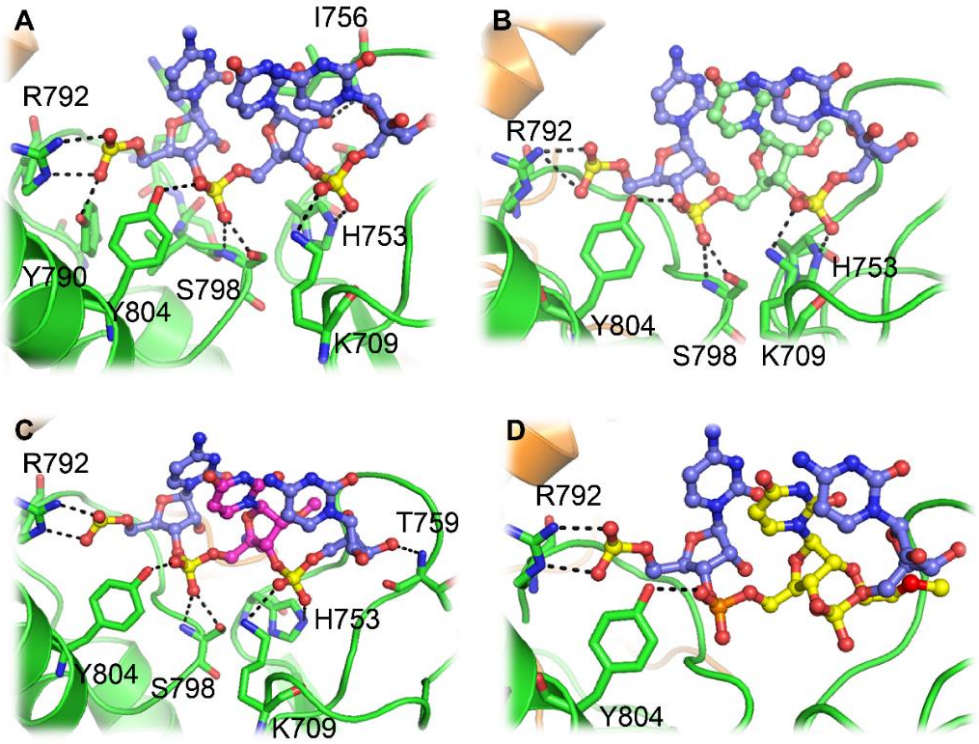


Figure S16. MD snapshots depicting the binding interactions around position g5 of the unmodified and the modified siRNA5-hAGO2 complexes. Best representative structures from the last 200 ns of the MD simulations of (**A**) unmodified, (**B**) 2'-OMe (green), (**C**) 4'-AM-2'-OMe (magenta) and (**D**) 2'-MOE (yellow) modified siRNA5-hAGO2 complexes. Protein is represented in cartoon, interacting AAs are represented in sticks, RNA is represented in spheres and sticks. The black dashed lines indicate the non-covalent interactions between nucleotides of the siRNA and the AAs of hAGO2.

MD snapshots of the loops around g5 in the unmodified and in the 2'-MOE modified siRNA5-hAGO2 complexes

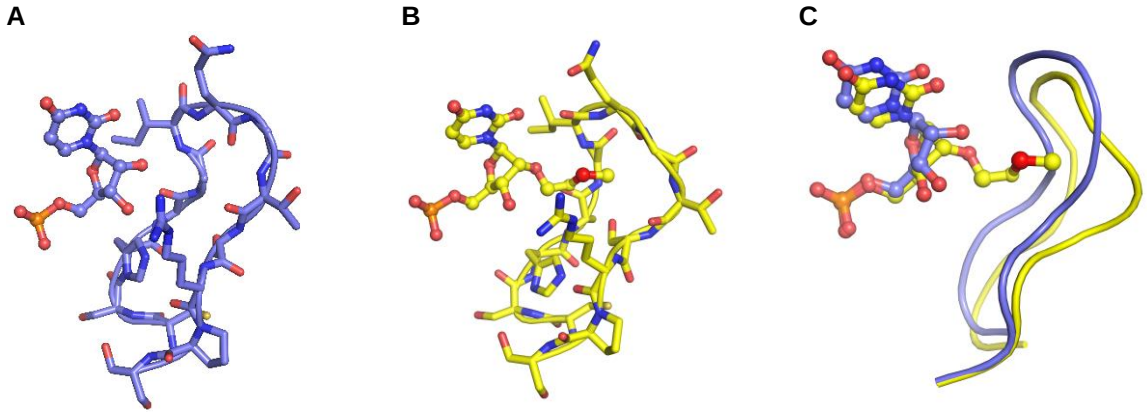


Figure S17. MD snapshots depicting the orientation of the hydrophobic loop around position g5 in the unmodified and 2'-MOE modified siRNA5-hAGO2 complexes. Best representative structures from the last 200 ns of the MD simulations of (**A**) unmodified and (**B**) 2'-MOE modified siRNA5-hAGO2 complexes. (**C**) Superposition of the averaged structure of unmodified (blue) and 2'-MOE (yellow) modified siRNA5-hAGO2 complexes.

Time-dependence of key H-bond distances in the siRNA14-hAGO2 complexes

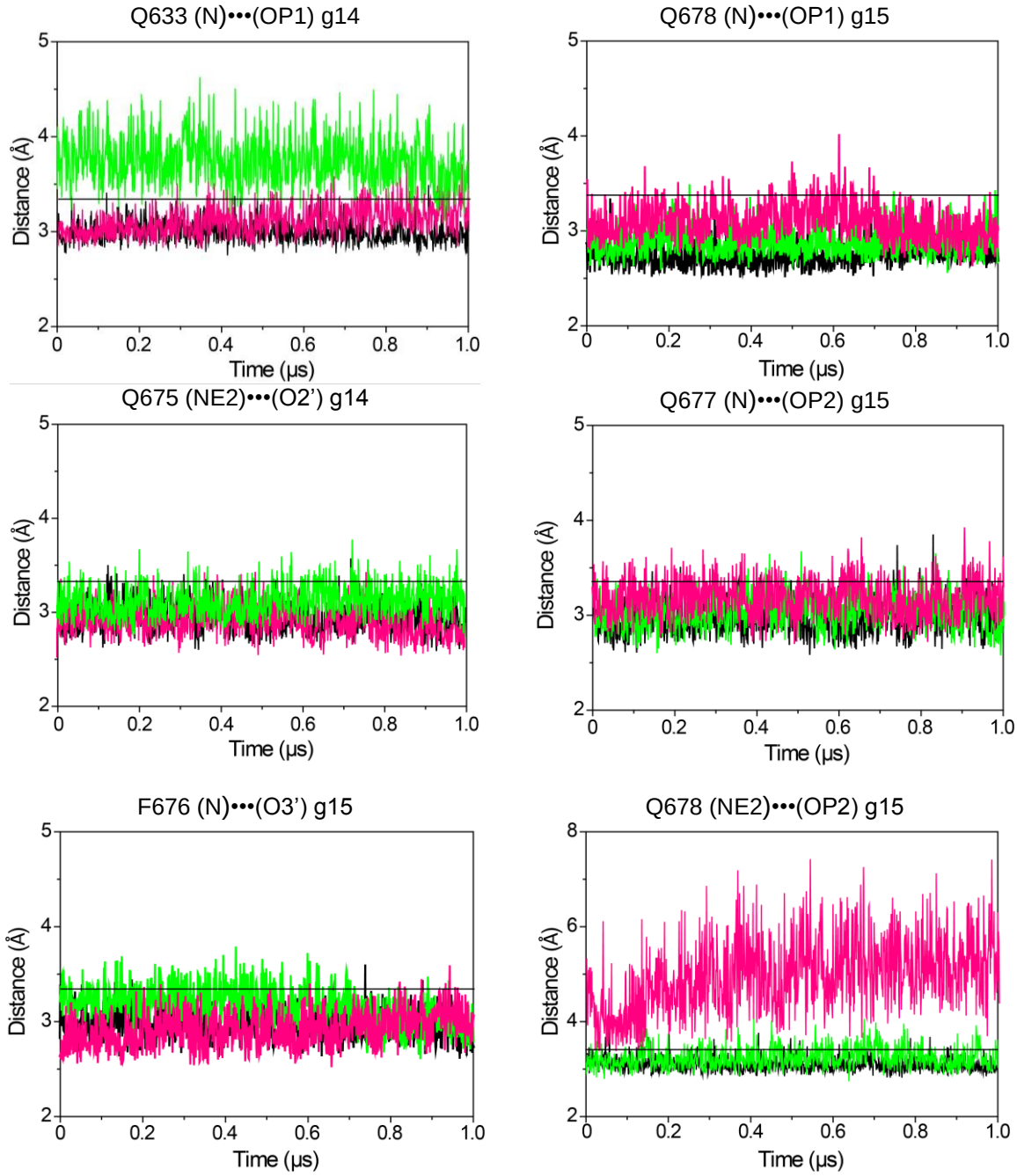
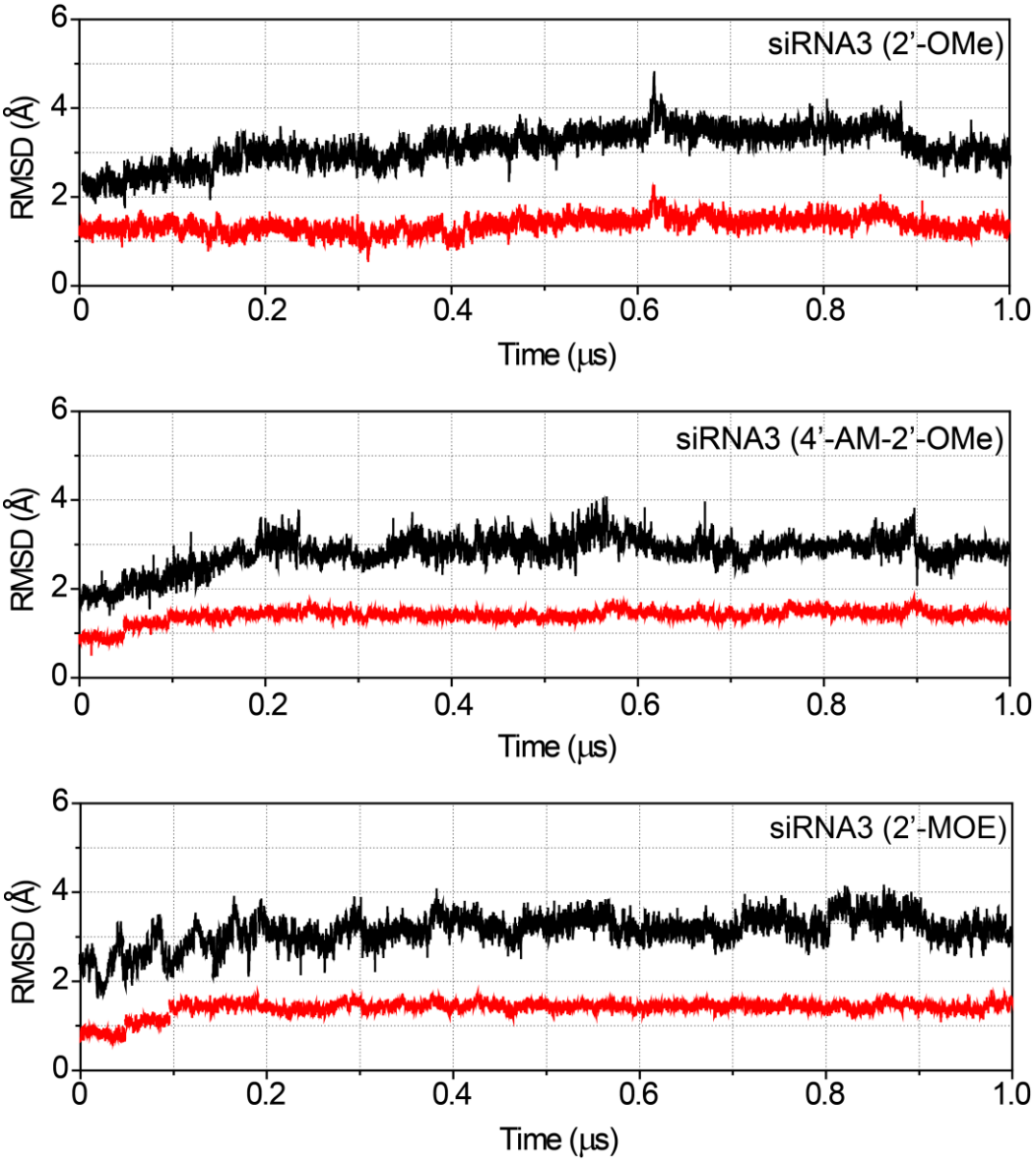
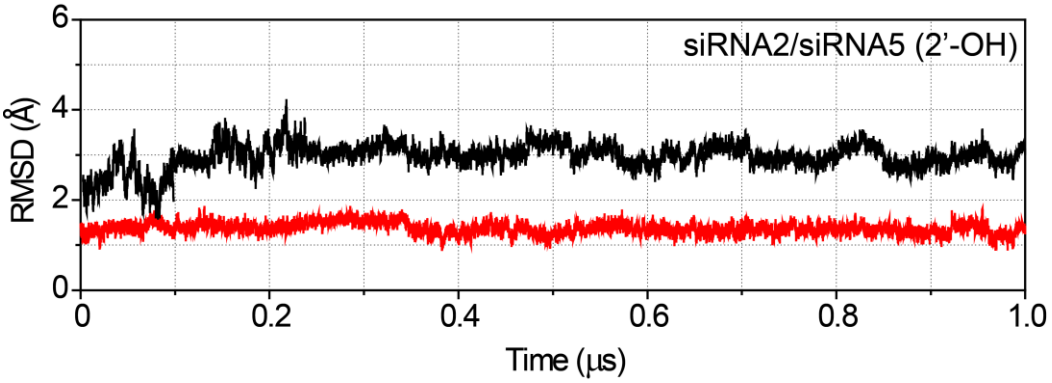


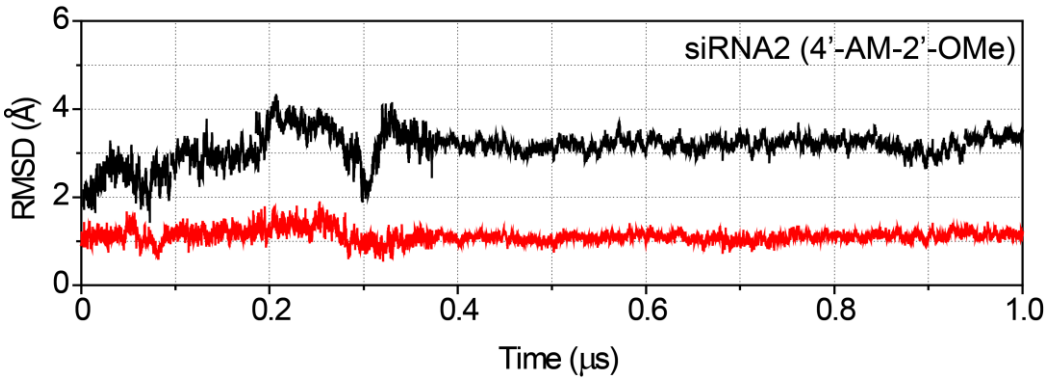
Figure S18. Time-dependence of the key H-bond distances during the MD simulations of siRNA14 and hAGO2 complexes. The H-bond is measured as the distances between the heavy atoms as mentioned above the graphs. Black lines indicate the H-bond distance with a cut-off value of 3.3 Å. The distances in the unmodified (black) and 2'-OMe (green), 4'-AM-2'-OMe (magenta) modified siRNAs are highlighted using different colours.

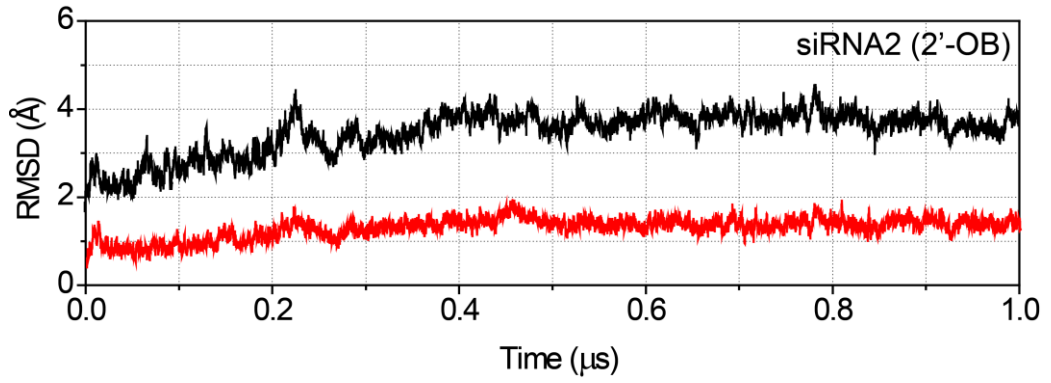
Time-dependent RMSD graphs of the siRNAs and the hAGO2

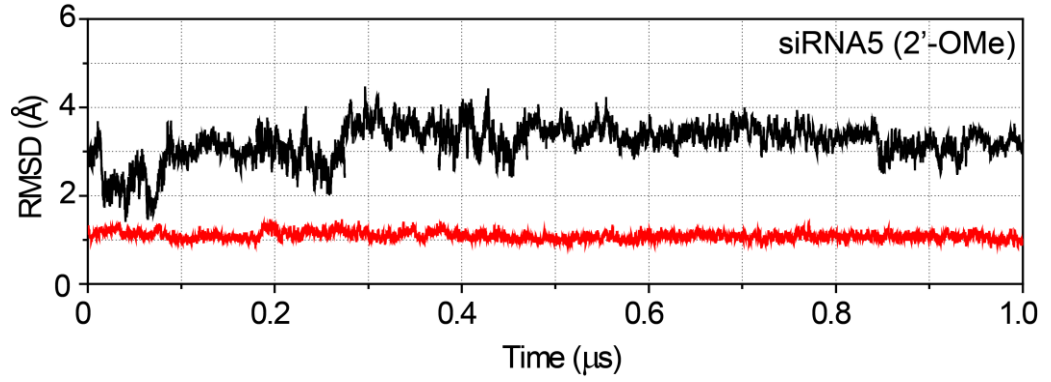
Time-dependent root-mean-square deviation (RMSD) graphs of backbone heavy atoms in the siRNA (P, O5', C5', C4', C3', and O3'; red) and in the hAGO2 (CA, C, and N; black) calculated with respect to the initial structure after equilibrations. Each graph is labelled with the specific name of the complex used (Table S1, Supporting Information)

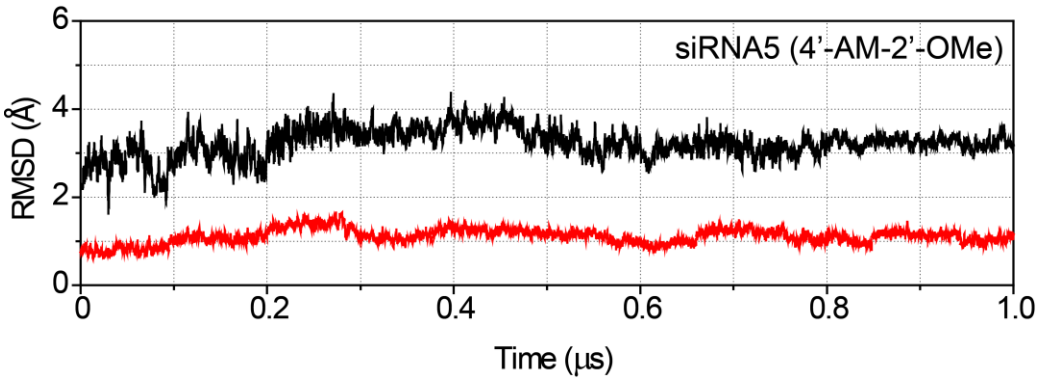


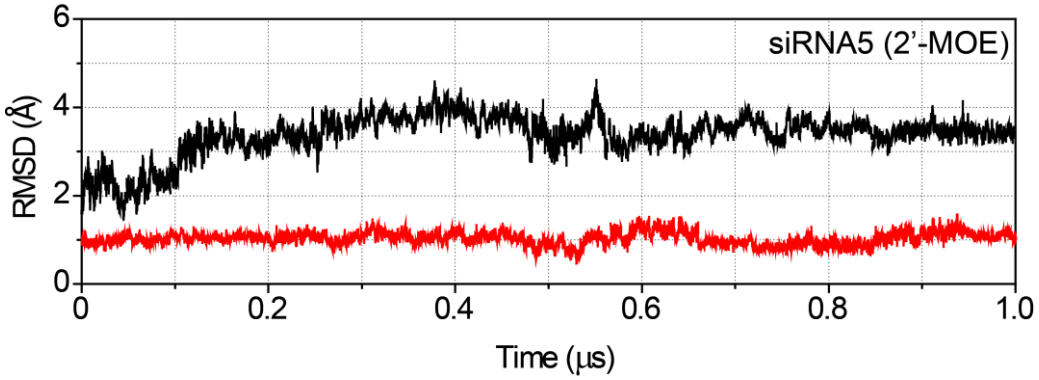

siRNA2/siRNA5 (2'-OH)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0

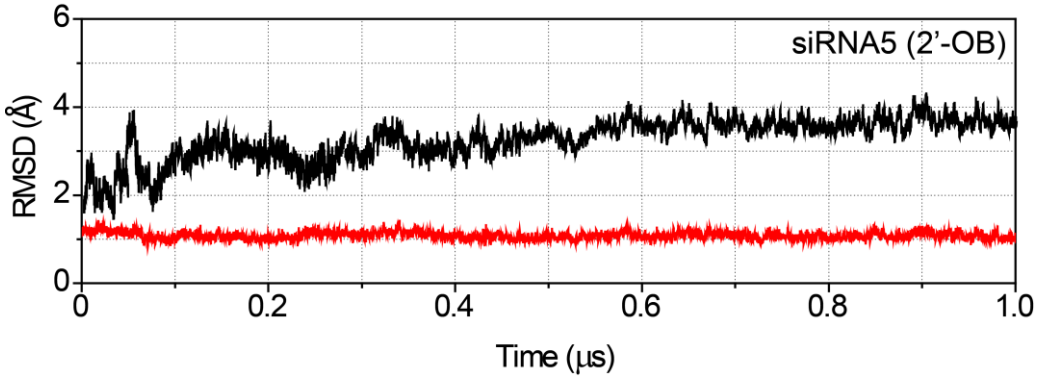

siRNA2 (4'-AM-2'-OMe)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0

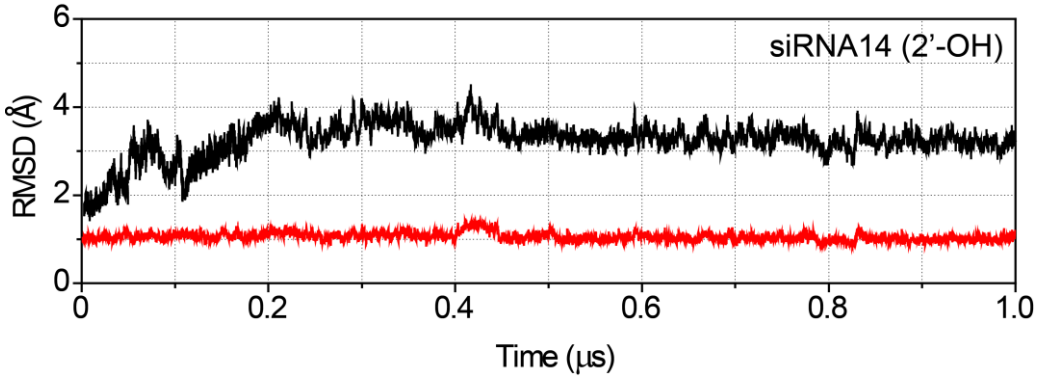

siRNA2 (2'-OB)
RMSD (Å)
Time (μs)
0
2
4
6
0.0
0.2
0.4
0.6
0.8
1.0

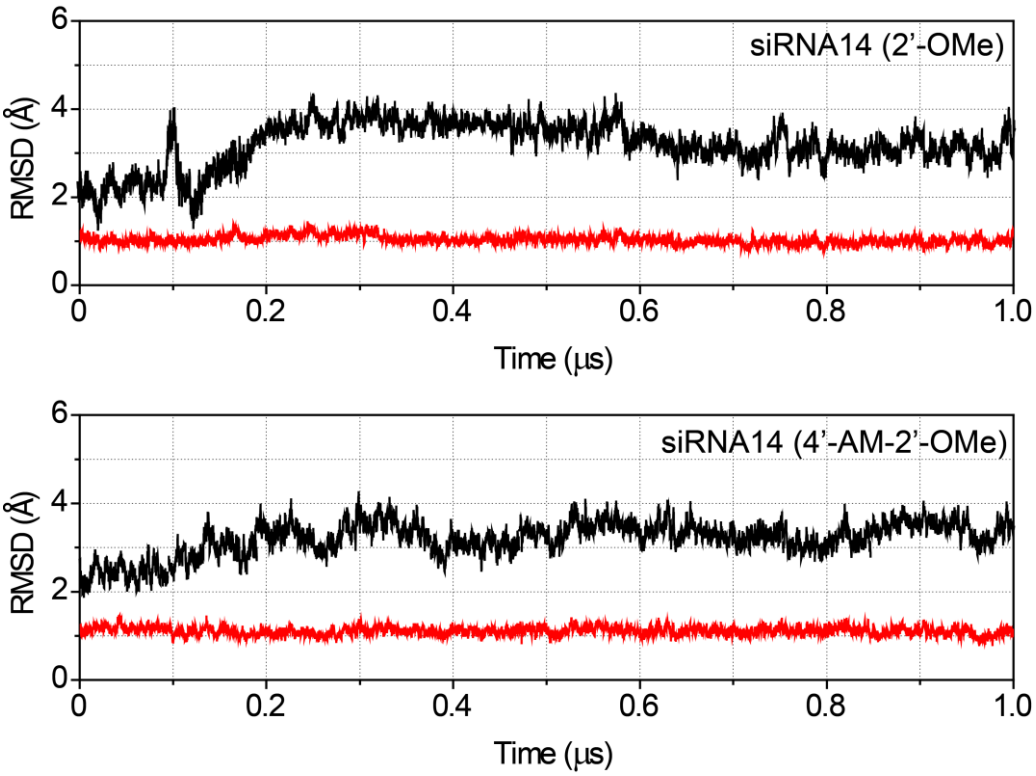

siRNA5 (2'-OMe)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0


siRNA5 (4'-AM-2'-OMe)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0


siRNA5 (2'-MOE)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0


siRNA5 (2'-OB)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0


siRNA14 (2'-OH)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0


siRNA14 (2'-OMe)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0

siRNA14 (4'-AM-2'-OMe)
RMSD (Å)
Time (μs)
0
2
4
6
0
0.2
0.4
0.6
0.8
1.0

siRNA sequances used in the MD simulations

No	Sequence (passenger: guide)	Modification	siRNA code
1	5'-GGCCUUUCACUACUCCUACUU-3' 3'-UUCCGGAAAGUGAUGAGGAUGA-5'	Unmodified	siRNA3 (2'-OH)
2	5'-GGCCUUUCACUACUCCUACUU-3' 3'-UUCCGGAAAGUGAUGAGGAUGA-5'	2'-OMe	siRNA3 (2'-OMe)
3	5'-GGCCUUUCACUACUCCUACUU-3' 3'-UUCCGGAAAGUGAUGAGGAUGA-5'	4'-AM-2'-OMe	siRNA3 (4'-AM-2'-OMe)
4	5'-GGCCUUUCACUACUCCUACUU-3' 3'-UUCCGGAAAGUGAUGAGGAUGA-5'	2'-MOE	siRNA3 (2'-MOE)
5	5'-GGCCUUUCACUACUCCUACUU-3' 3'-UUCCGGAAAGUGAUGAGGAUGA-5'	2'-OB	siRNA3 (2'-OB)
6	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	Unmodified	siRNA2/siRNA5 (2'-OH)
7	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	4'-AM-2'-OMe	siRNA2 (4'-AM-2'-OMe)
8	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	2'-OB	siRNA2 (2'-OB)
9	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	2'-OMe	siRNA5 (2'-OMe)
10	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	4'-AM-2'-OMe	siRNA5 (4'-AM-2'-OMe)
11	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	2'-MOE	siRNA5 (2'-OMe)
12	5'-AUAAGGCUAUGAAGAGAUAUU-3' 3'-UUUAUUCCGAUACUUCUCUAU-5'	2'-OB	siRNA5 (2'-OB)
13	5'-GGUUAACAGCGAUCUGAUAUU-3' 3'-UUCCAAUUGUCGCUAGACUAU-5'	Unmodified	siRNA14 (2'-OH)
14	5'-GGUUAACAGCGAUCUGAUAUU-3' 3'-UUCCAAUUGUCGCUAGACUAU-5'	2'-OMe	siRNA14 (2'-OMe)
15	5'-GGUUAACAGCGAUCUGAUAUU-3' 3'-UUCCAAUUGUCGCUAGACUAU-5'	4'-AM-2'-OMe	siRNA14 (4'-AM-2'-OMe)

Table S1. siRNA sequences and position of the modifications used in the MD simulations studies. The sequences of the siRNA are shown completely (22nt), but only 17 nt (from 5'-end of guide RNA) base pairing was used in MD studies. The position of modification is highlighted in red color.

References

(1) Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2001**, *29*, 228.

(2) Schirle, N. T.; Sheu-Gruttadauria, J.; MacRae, I. J. *Science* **2014**, *346*, 608.

(3) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M. J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Gusarov, S., Kovalenko, A., and Kollman, P. A. *AMBER 14, University of Californai, San Francisco* **2014**.

(4) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.

(5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.