Supplementary_A for:

Tiny miRNA in Complex with Human Argonaute-2 and mRNA Target: Molecular Dynamics Study in Pure Water and Water-Methanol Mixture.

D.Schiumarini, A. Casavola , F. Porcelli , A. Grottesi and L. Avaldi

THE ENVIRONMENT

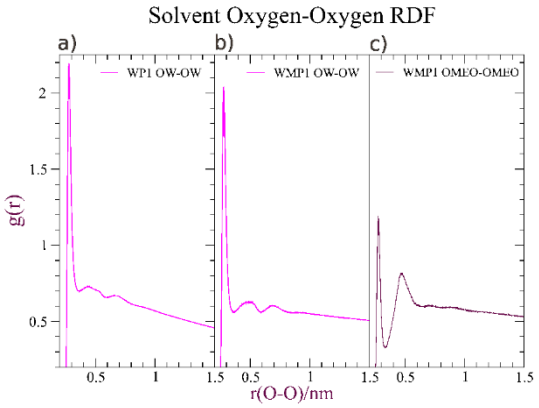

Figure S1 Oxygen correlation length for pure and mixed solvent: a) water oxygen correlation length for sample WP1; b) water oxygen and and c) methanol oxygen correlation length for sample WMP1.

To note that the solution water-methanol, 50% in volume, presents, for the same space occupancy, a molecule number that is ¾ of that one in pure water.

THE COMPLEX: STRUCTURAL ANALYSIS

Table S1 shows the average backbone RMSD values relative to protein, miRNA and mRNA: protein and miRNA display comparable values.
In fig. S3a the RMSD trends of the complex trajectories are plotted. Looking to the graph, related to protein-RNA system backbone, it is noted a smooth oscillating behaviour of mixed solvent simulations, moving away and getting closer again to the other trajectories, in a time interval of around 400 ns.
In Fig. S4, the plots of RMS fluctuations of the protein are presented and in table S1 are listed the residue intervals involving RMSF picks greater than 0.35 nm^2 (Y) or 0.4 nm^2 (Y+), the mixed solvent samples showing regions with more intensively structured RMS fluctuations; the intervals listed are mostly portions of floating loops binding antiparallel beta strands and alpha helices.

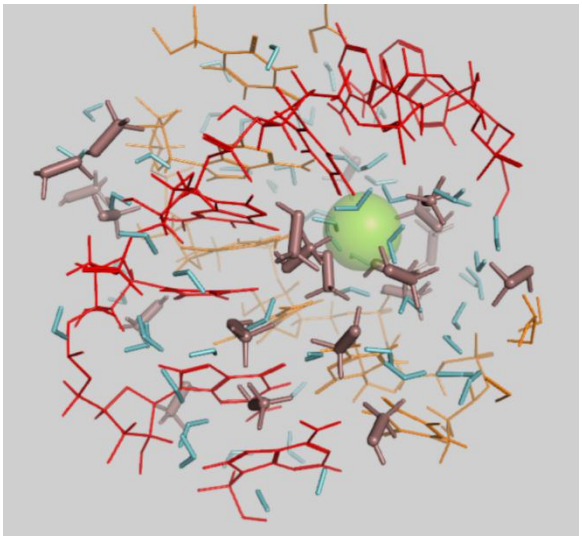

Figure S2 Image of the 'seed-box', the region within 1.2 nm from the centre of mass of the RNA seed (g2-g8 paired to tt8-t2), for WMP0 sample: in light blue are represented water molecules, in brown methanol molecules, in red and orange miRNA and mRNA seed sequence, in green the magnesium atom.

The trajectories RMSD reported in fig s3a are relative to the complete backbone RISC structure: in fig S3b the RMSD data reported are relative to the original RISC structure not considering the sequence listed in table S2. That let a rough estimate of the weight in RMSF, and the complex dynamics, of the sequences ignored in the reduced structure. in the case of the complete structures the RMSF sums/nm^2 over all RISC Backbone atoms are 70, 59, 74, 105, for WP0, WP1, WMP0 and WMP1 respectively and 22,24, 33, 28 in the case of the truncated ones: even if not

considering the contribution of the intervals reported in table S1, the RMSF of mixed solvent samples remains greater than that of the systems in pure water.

Table S1. Backbone RMSD data (in nm) for protein, miRNA and mRNA, for the different samples (WP0 and WP1 in water, WMP0 and WMP1 in water-methanol 70:30 molar ratio).

	Protein	miRNA	mRNA
WP0	0.25(+/-0.03)	0.34(+/-0.05)	022(+/-0.02)
WP1	0.26(+/-0.03)	0.33(+/-0.02)	0.21(+/-0.05)
WMP0	0.38(+/-0.03)	0.41(+/-0.06)	0.30(+/-0.06)
WMP1	0.40(+/-0.05)	0.37(+/-0.05)	0.26(+/-0.04)

Table S2. RMS fluctuations intervals with high values in intensity (RMSF >= 0.35 nm) and/or structured profile: Y+ indicates RMS fluctuation interval particularly significant for the sample(RMSF >= 0.5 nm).

	Res.Int.	WP0	WP1	WMP0	WMP1
N	62:66	Y		Y	Y
	106:113	Y		Y	y
	119:126	Y	Y	Y+	Y
L1	140:160				Y
	183:191	Y+	Y+		Y
PAZ	238:250	Y	Y	Y	Y+
	270:279	Y	Y	Y	Y
	292:308	Y+	Y+	Y+	Y+
	311:319	Y	Y	Y	Y
	330:339	Y+		Y+	Y
L2	387:398				Y+
	421:427	Y	Y+	Y+	Y
MID	457:464	Y	Y	Y+	Y
	495:500			Y	Y
	511:512		Y		
PIWI	601:607		Y		Y
	671:678				Y
	817:848	Y+	Y+	Y+	Y+

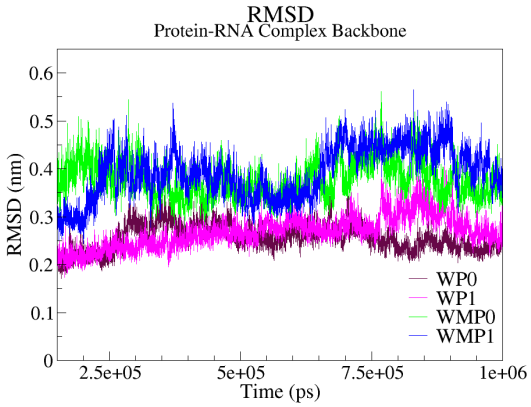

Figure S3a. Backbone RMSD of protein-ligand complex trajectories in pure water (samples WP0, WP1) and in water-methanol mixture, 70:30 molar ratio (samples WMP0, WMP1).

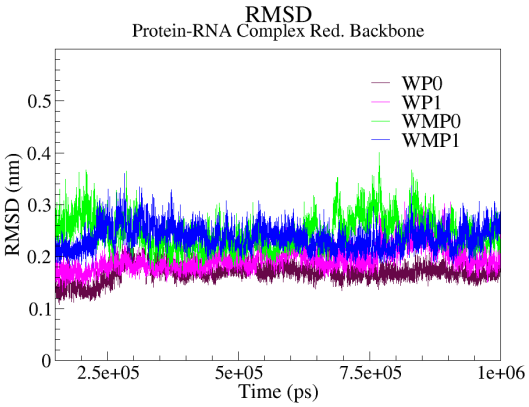

Figure S3b. Reduced backbone RMSD of protein-ligand complex trajectories in pure water (samples WP0, WP1) and in water-methanol mixture, 70:30 molar ratio (samples WMP0, WMP1).

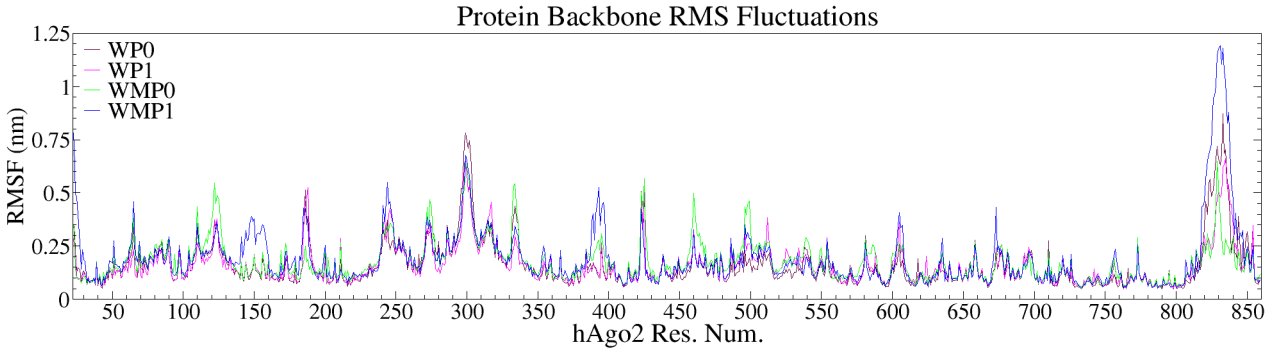

Figure S4. Protein RMS fluctuations as a function of residue number for water and mixed solvent samples

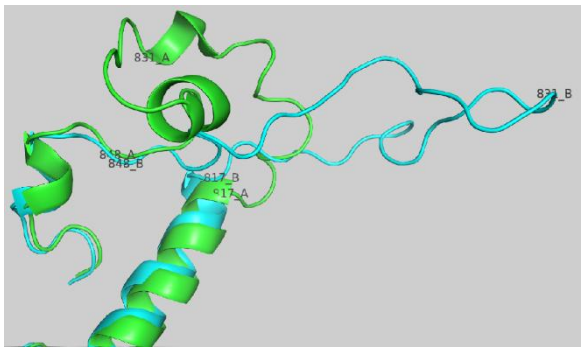

Figure S5. PIWI sequence around residue interval 817-148 of h-AGO2 for sample WMP1: the green sequence is relative to simulation time t=430 ns and the distance between residue 817_A and residue 831_A is 1.52 nm; the cyan protein sequence is relative to simulation time t=1000 ns and the distance between residue 817_B and residue 831_B is 3.9 nm: it is evident for this interval, the completely loss of the secondary structure.

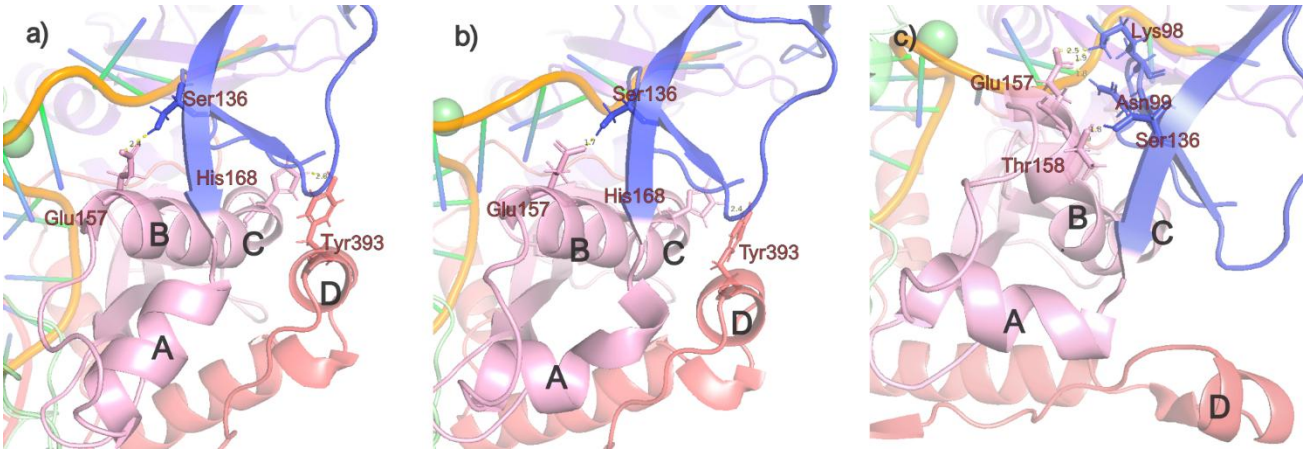

Figure S6. A(res. 140-148), B(res.156-163), C(res.390-400) sequence configuration at 200 ps (a) 400 ps (b) and 900 ps (c) of the WMP1 simulation. In b) it can be noted the relaxation of the A sequence; in c, sequence B is climbed on N with the formation of bonds Thr158-Ser136, Glu157-Lys98 and Glu157-Asp199 (in a and b the dominant bond involving B sequence and N domain is Glu157-Ser136); again in c, sequence D lies away from C, the bond His168-Tyr393 being broken.

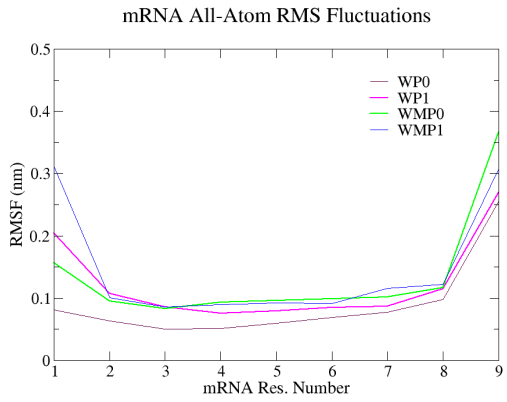

Figure S7. All atom rmsf for mRNA in water sample (WP0, WP1) and in mixed solvent (WMP0, WMP1). (WMP0, WMP1).

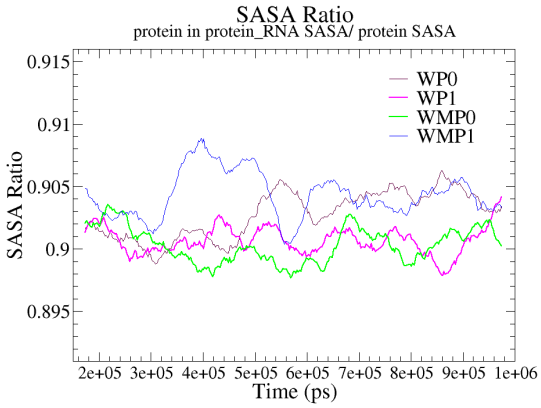

Figure S8a. Ratio of protein-RNA SaSA on protein SASA

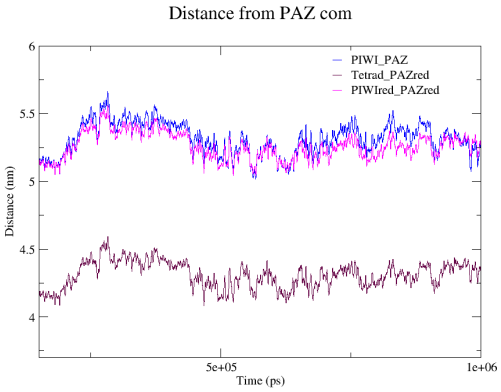

Figure S8b. WMP1: PAZ com distance for PIWI com, DEDH tetrad com.

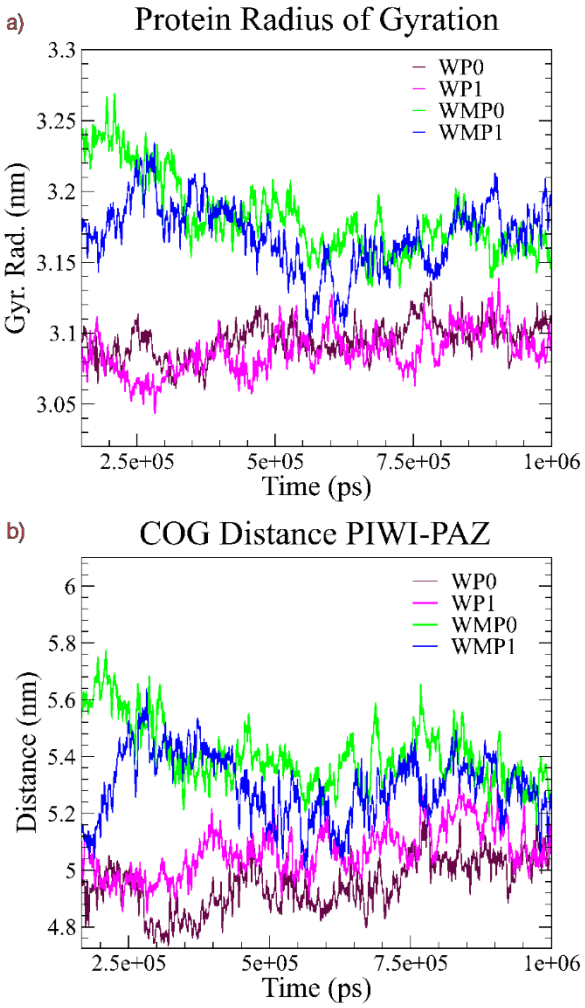

Figure S9. a) Radius of gyration and b) PIWI-PAZ distance (centre of geometry) as a function of time for the four samples.

HYDROGEN BOND ANALISYS: miRNA-PROTEIN

To show the different components of the interaction between miRNA and protein residues, in figure S10a,b the frequency intensity of hydrogen bond vs protein residue number involved is reported for WP0 and WMP1, in figure S11 for WP1 and WMP0: the bar colours distinguish the occupancy value (octane Occ>0.9, mauve 0.75<fc<0.9, magenta 0.4<fc<0.75). and in figure S10 are shown maximum hydrogen bond occupancy versus miRNA nucleotides for the various samples. In tables S3 and S4 the details for the couples donor-acceptor involved in the hydrogen bonds, in the case of nucleotide 17 and 18, are given and differences emerge depending on the environment for the miRNA strand tail.

Table S3. For the various samples, A, B and C columns show the number of hydrogen bonds protein-miRNA with occupancy greater than 0.9, greater than 0.4, less than 0.02 respectively. D column shows, for each sample, the sum of occupancy values relative to persistent hydrogen bonds (Occ greater than 0.4) that denotes the strength of protein-ligand bond.

	A nhb Occ> 0.9	B nhb Occ 0.4	>	C nhb Occ<0.02	D Cum-Occ for Occ>0.4
WP0	15	40		365	31.7
WP1	10	40		372	28.6
WMP0	8	32		469	23.4
WMP1	11	28		375	21,1

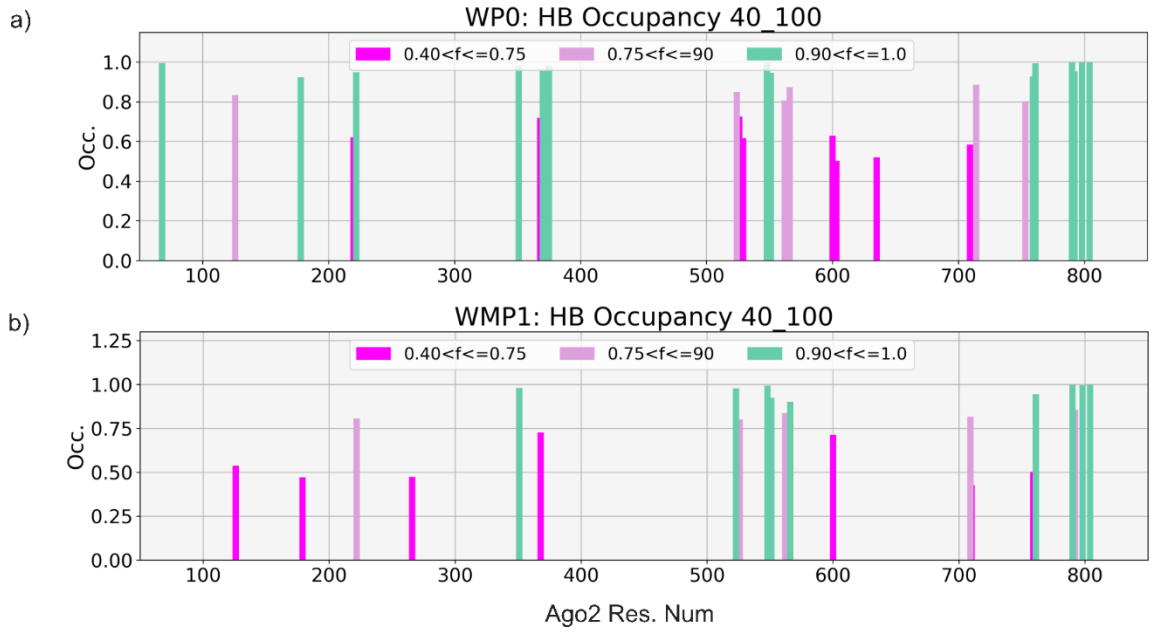

Figure S10. Protein-miRNA **h**ydrogen bond Occupancy vs number of Ago2 residues (a: WP0 sample, b: WMP1 sample. As stated in the legend, occupancy values are grouped by color as follows: Octane Occ>0.9, plume 0.75<fc<0.9, magenta 0.4<fc<0.75.

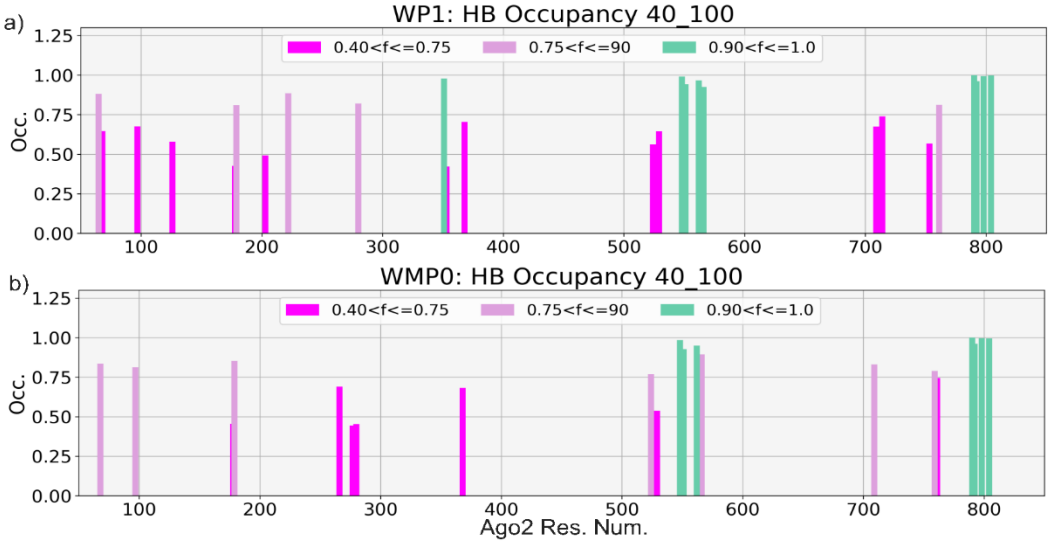

Figure S11. Occupancy of hydrogen bonds protein-miRNA, vs protein residues for WP1 and WMP0.

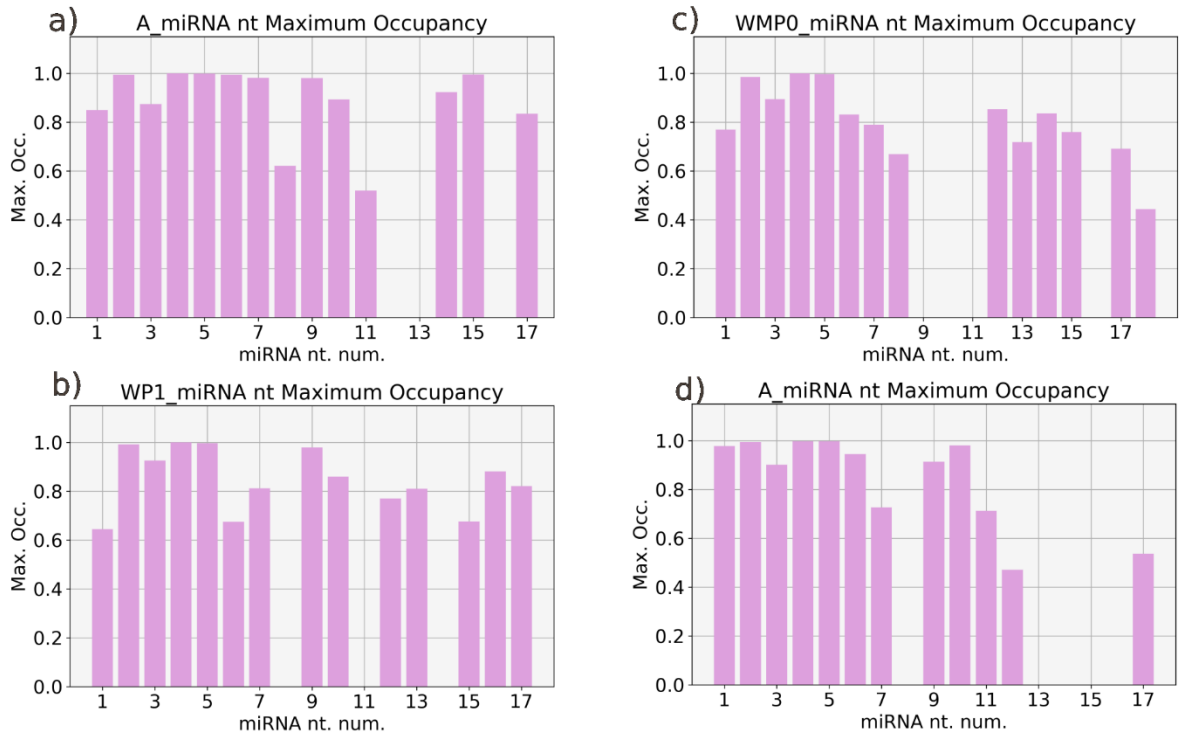

Figure S12. Maximum protein-miRNA hydrogen bond Occupancy versus miRNA nucleotides for samples WP0 (a), WP1 (b), WMP0 (c), WMP1 (d).

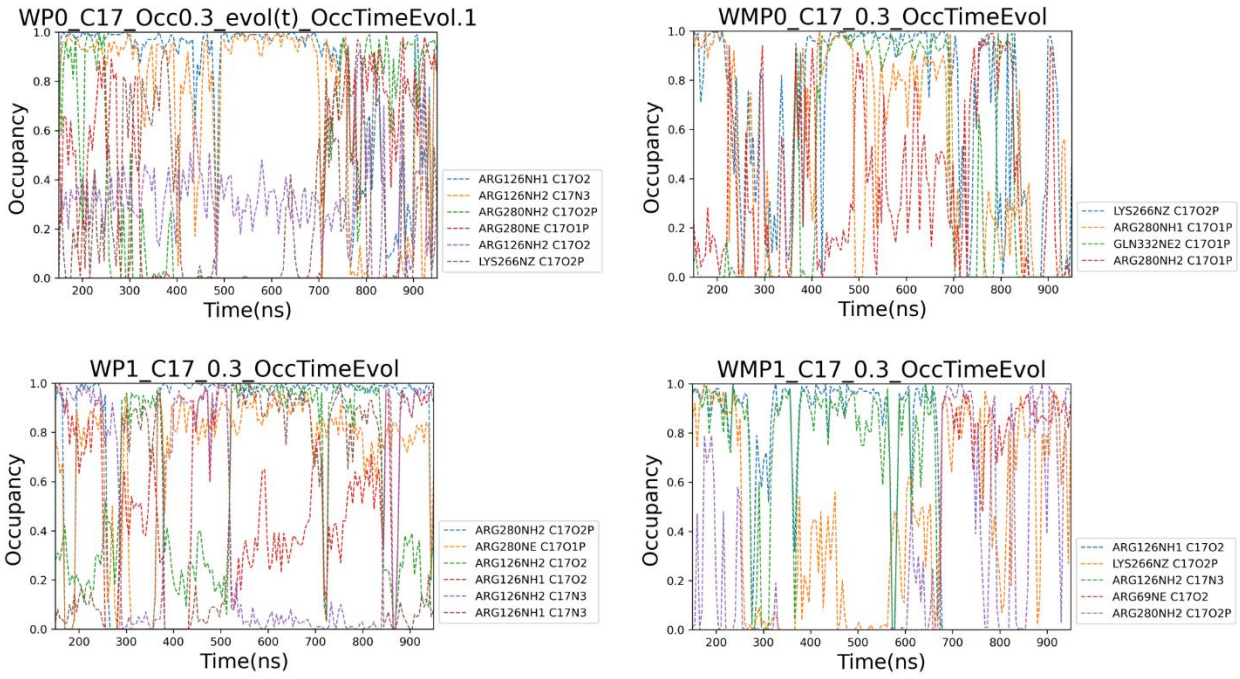

Figure S13. Hydrogen bond time evolution for nucleotide C17 in samples WP0 (a), WP1 (b), WMP0 (c), WMP1 (d): to note the simultaneous bonds PAZ-miRNA , L2-miRNA for WMP0 sample and PAZ-miRNA , N-miRNA for the other ones

Table S4 a,b,c,d. nt 17 hydrogen bonds for the various samples (Occ.>=0.2).

Table 4_a. WP0, nt 17.

Donor	Hydrogen	Acceptor	Occ.	mRNA
ARG280NH2	ARG280HH21	C17O2P	0.393818	022(+/-0.02)
ARG280NE	ARG280HE	C17O1P	0.359435	0.21(+/-0.05)
LYS266NZ	LYS266HZ1	C17O2P	0.301207	0.30(+/-0.06)
LYS266NZ	LYS266HZ1	C17O1P	0.247100	
ARG126NH2	ARG126HH21	C17O2	0.309921	
ARG126NH2	ARG126HH21	C17N3	0.710509	
ARG126NH1	ARG126HH11	C17O2	0.834207	0.26(+/-0.04)

Table 4_b. WP1, nt 17.

Donor	Hydrogen	Acceptor	Occ.
ARG280NH2	ARG280HH21	C17O2P	0.820893
ARG280NE	ARG280HE	C17O2P	0.223222
ARG280NE	ARG280HE	C17O1P	0.716134
ARG126NH2	ARG126HH21	C17O2	0.578378
ARG126NH2	ARG126HH21	C17N3	0.467031
ARG126NH1	ARG126HH11	C17O2	0.552967
ARG126NH1	ARG126HH11	C17N3	0.448327

Table 4_c. WMP0, nt17

.

Donor	Hydrogen	Acceptor	Occ.
ARG277NH1	ARG277HH11	U18O1P	0.442844
ARG277NH2	ARG277HH21	U18O2P	0.324591
GLN332NE2	GLN332HE21	U18O4'	0.284311
U18N3	U18H3	LYS65O	0.230213
ARG277NH2	ARG277HH21	U18O1P	0.202870

Table 4_d. WMP1, nt17

Donor	Hydrogen	Acceptor	Occ.
ARG280NH2	ARG280HH21	C17O2P	0.310982
LYS266NZ	LYS266HZ1	C17O2P	0.473090
LYS266NZ	LYS266HZ1	C17O1P	0.230339
ARG126NH2	ARG126HH21	C17N3	0.473031
ARG126NH1	ARG126HH11	C17O2	0.536557
ARG69NH2	ARG69HH21	C17O2	0.279337
ARG69NE	ARG69HE	C17O2	0.325510
ARG69NE	ARG69HE	C17N3	0.231045

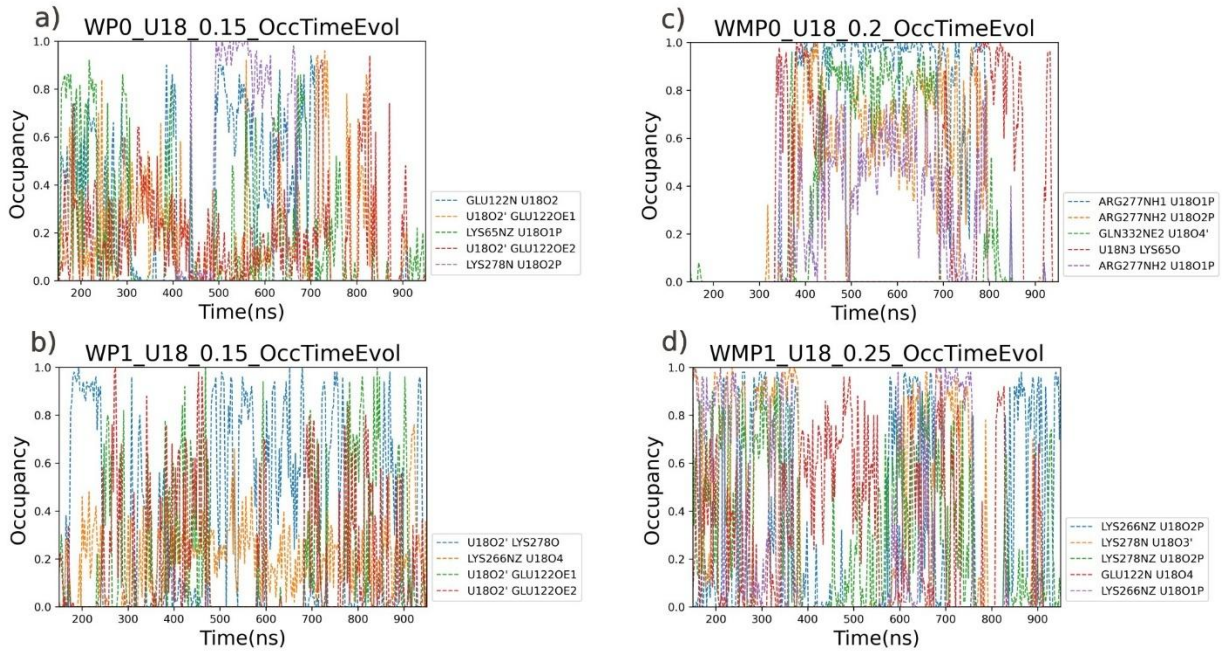

Figure S14. Hydrogen bond time evolution for nucleotide C18 in samples WP0 (a), WP1 (b), WMP0 (c), WMP1 (d): to note the rapid bond state change and the simultaneous N-miRNA and PAZ-miRNA bonds for all samples.

Table S5 a, b, c, d. nt 18 hydrogen bonds for the various samples (Occ.>=0.2).

Table S5_a. WP0, nt 18, (1 over 112)

Donor	Hydrogen	Acceptor	Occ.
GLU122N	GLU122H	U18O2	0.248572

Table S5_b. WP1, nt 18, (2 over 63)

Donor	Hydrogen	Acceptor	Occ.	mRNA
U18O2'	U18HO'2	LYS278O	0.368684	0.30(+/-0.06)
LYS266NZ	LYS266HZ1	U18O4	0.216870	0.26(+/-0.04)

Table S5_c. WMP0, nt 18, (5 over 176)

Donor	Hydrogen	Acceptor	Occ.
U18N3	U18H3	LYS65O	0.230281
GLN332NE2	GLN332HE21	U18O4'	0.284395
ARG277NH2	ARG277HH21	U18O2P	0.324687
ARG277NH2	ARG277HH21	U18O1P	0.202929
ARG277NH1	ARG277HH11	U18O1P	0.442974

Table S5_d. WMP1, nt 18, (7 over 115)

Donor	Hydrogen	Acceptor	Occ.
U18O2'	U18HO'2	GLU122OE2	0.237692
U18O2'	U18HO'2	GLU122OE1	0.212223
LYS278NZ	LYS278HZ1	U18O2P	0.290924
LYS278N	LYS278H	U18O3'	0.354744
LYS266NZ	LYS266HZ1	U18O2P	0.364920
LYS266NZ	LYS266HZ1	U18O1P	0.261867
GLU122N	GLU122H	U18O4	0.273219

Table S6. Protein – miRNA hydrogen bonds (Occ. > 0.4) grouped on protein domains: the same colour highlights nucleotides bonded to residues of different domains: yellow colour is used for nucleotides bonded to N and L1, red for nucleotides bonded to N and Paz……

	WP0	WP1	WMP0	WMP1
N	C17-R126 G15-R68 A14-R68	C17-R126 U16-R68 U16-R97 U16-Lys65 G15-R68 G15-R97	A13-R97 A14-R97 A14-R68 G15-R68	C17-R126
L1	G9-Thr222 A14-GLY178 U8-Ser220	G15-Gly178 A13-R179 C12-R179 C12-His203 G9-Thr222	A14-GLY178 C12-R179	C12-R179 G9-THR222
PAZ		C17-R280	C17-R280 C17-Lys266 U18-R277	C17-Lys266
L2	U7-R375 U7-R370 U7-Thr368 G9-R351 C10-R351	C10-R351 G9-R351 G9-ILE353 U7-THR368	U7-thr368	C10-R351 G9-R351 U7-THR368
MID	U1-Thr526 U1-Gly524 U1-Tyr529 U2-GLN548 U2-ASN551 U2-ASN562 C3- Lys566	C3-Lys566 U2-GLN548 U2-ASN562 U2-ASN551 U1-Tyr529 U1-Thn526 U1_GLY524	U2-GLN548 U2-ASN562 U2-ASN551 C3-Lys566 U1_Tyr529 U1-Thr526 U1_GLY524	C3-Lys566 U2-GLN548 U2-ASN551 U2-ASN562 U1-PRO523 U1-THR526
PIWI	C10-His600 U7-R714 U7-THR759 A6-His753 A6-R761 C5-Ser798 C5-TYR804 A4-TYR790 A4-R792	U7-R761 U7-R714 A6-Tyr790 A6-R792 A6-Lys709 A6-His753 C5-Ser798 C5-Tyr804 A4-Tyr790 A4-R792	U7-R761 U8-R761 U7-Thr759 A6-R761 A6-Lys709 C5-Ser798 C5-Tyr804 A4-R792 A4-Tyr790	C11-His600 C11-R710 U7-R761 A6-R761 A6-Lys709 A6-THR759 C5-TYR804 C5-SER798 A4-R792

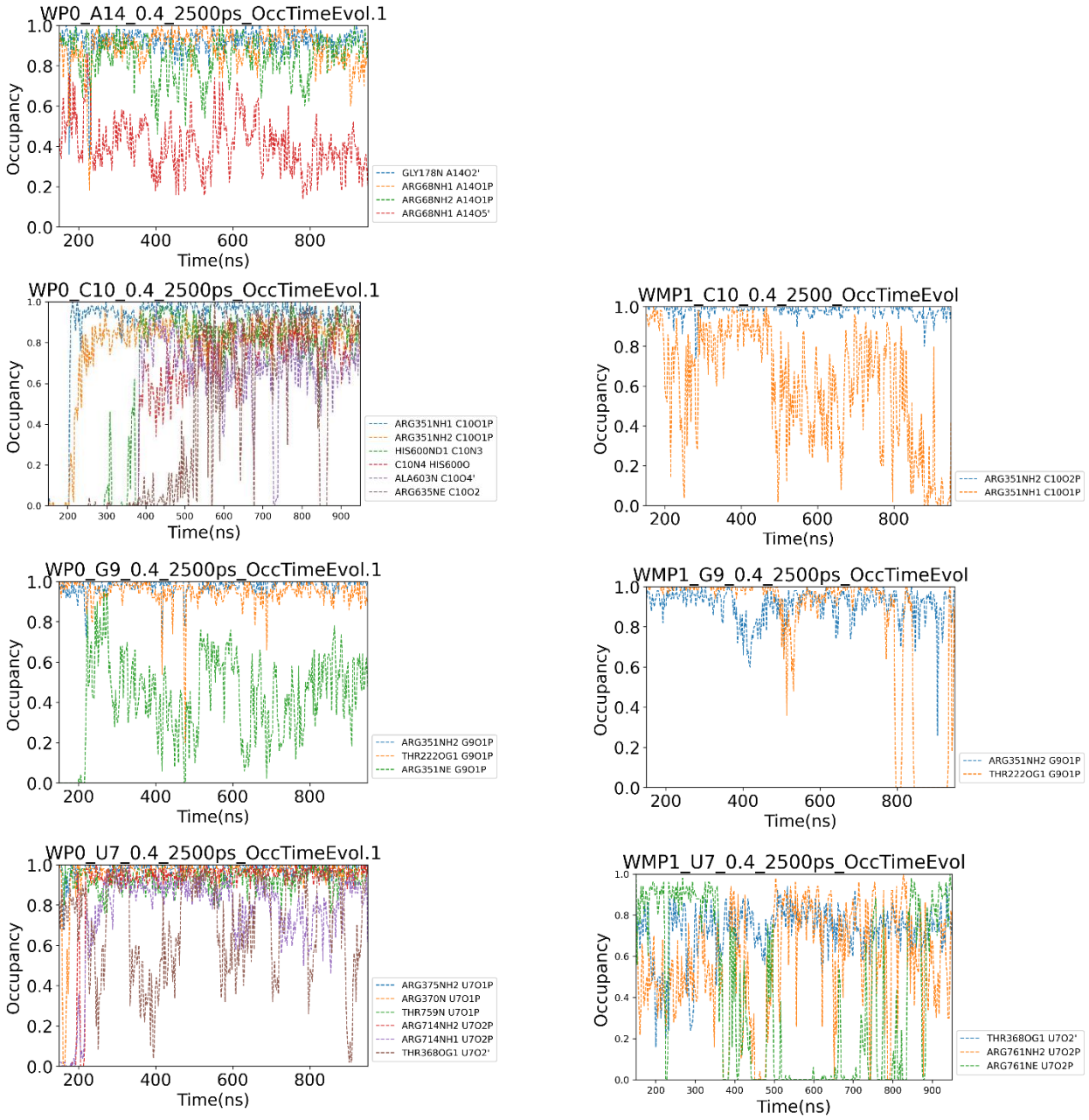

Figure S15 a. Hydrogen bond time evolution for U7, G9,C10 and A14 nucleotides of samples WP0 and WMP1: it is highlighted that the miRNA nucleotides in the range nt 7-14 are more strongly bound in the case of the pure water system.

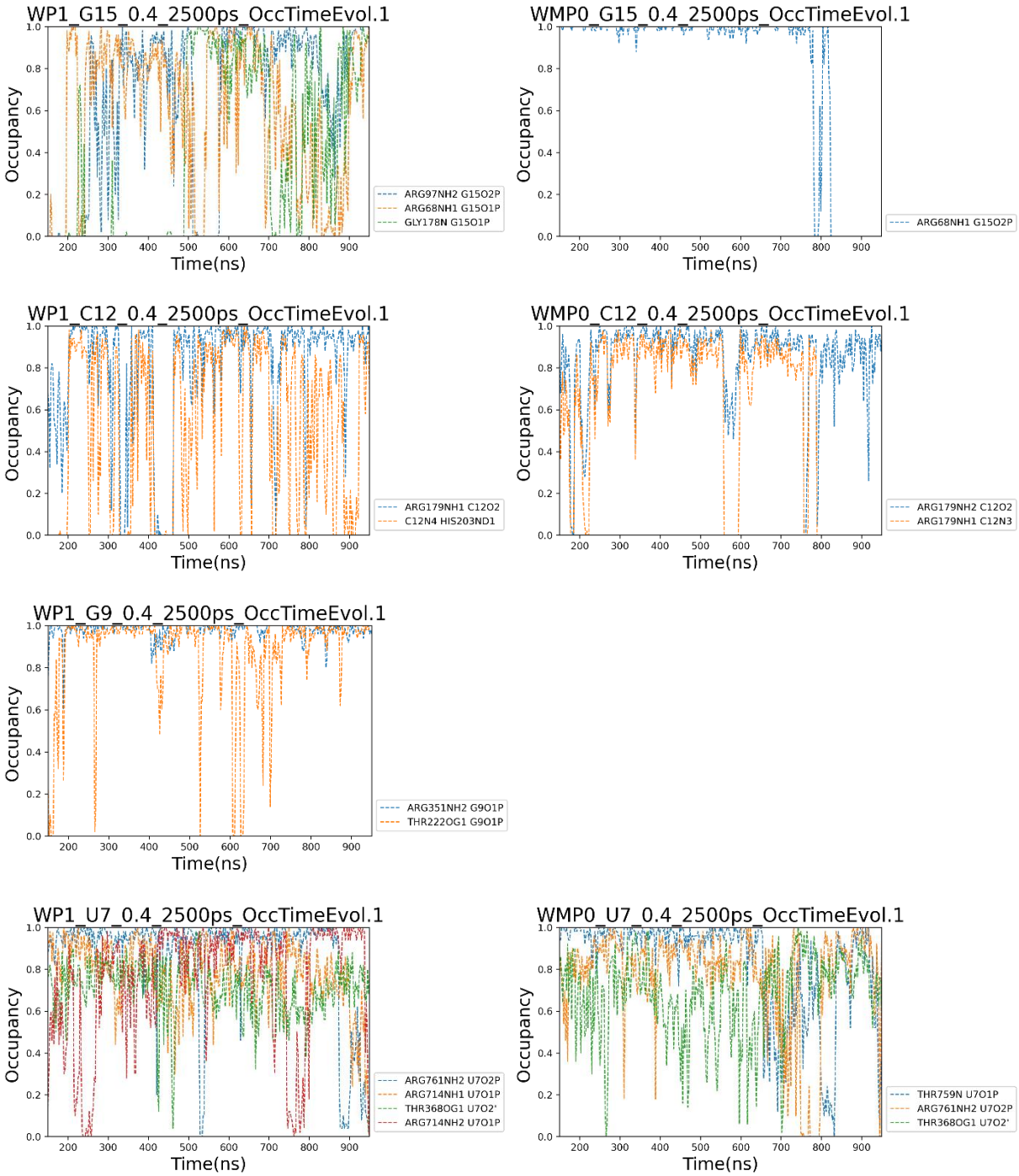

Figure S15 b. Hydrogen bond time evolution for U7, G9,C10 and A14 nucleotides of samples WP1 and WMP0: it is highlighted that the miRNA nucleotides in the range nt 7-14 are more strongly bound in the case of the pure water system.

HYDROGEN BOND ANALYSIS: mRNA-PROTEIN

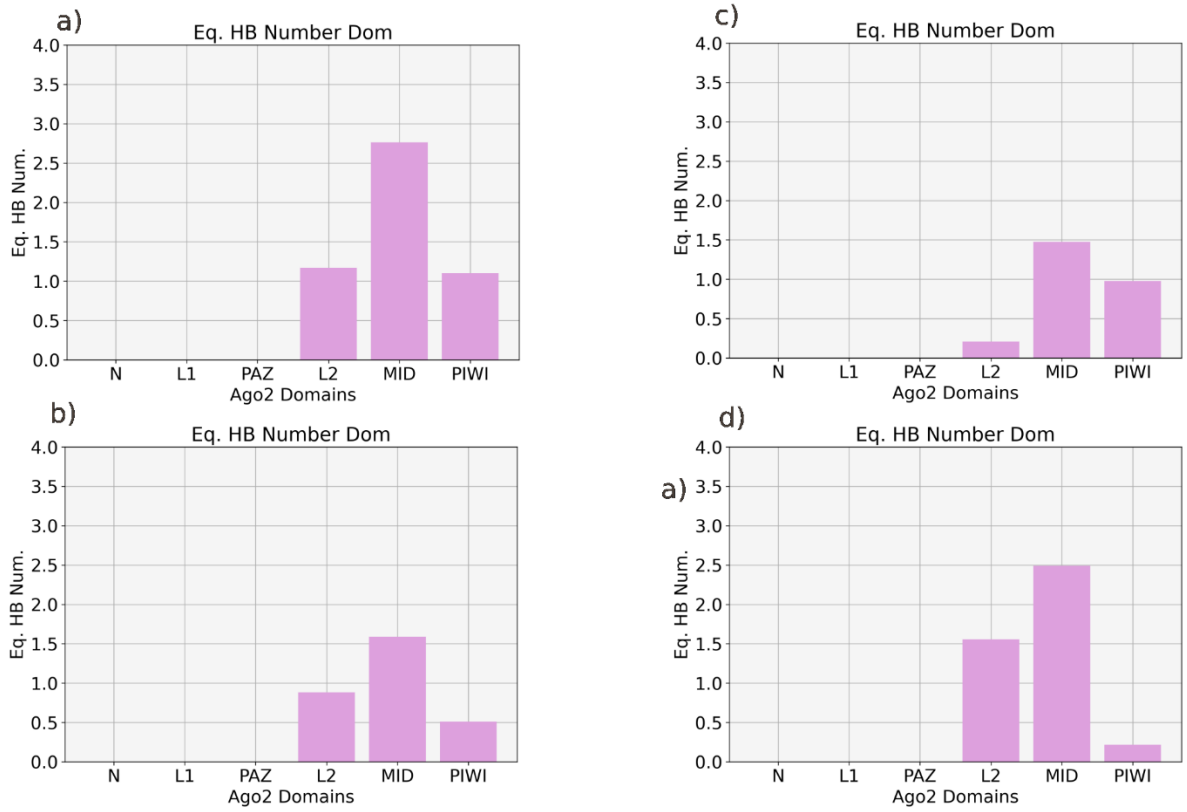

Figure S16. mRNA equivalent hydrogen bond number versus protein domains for samples WP0 (a), WP1 (b), WMP0 (c), WMP1 (d).

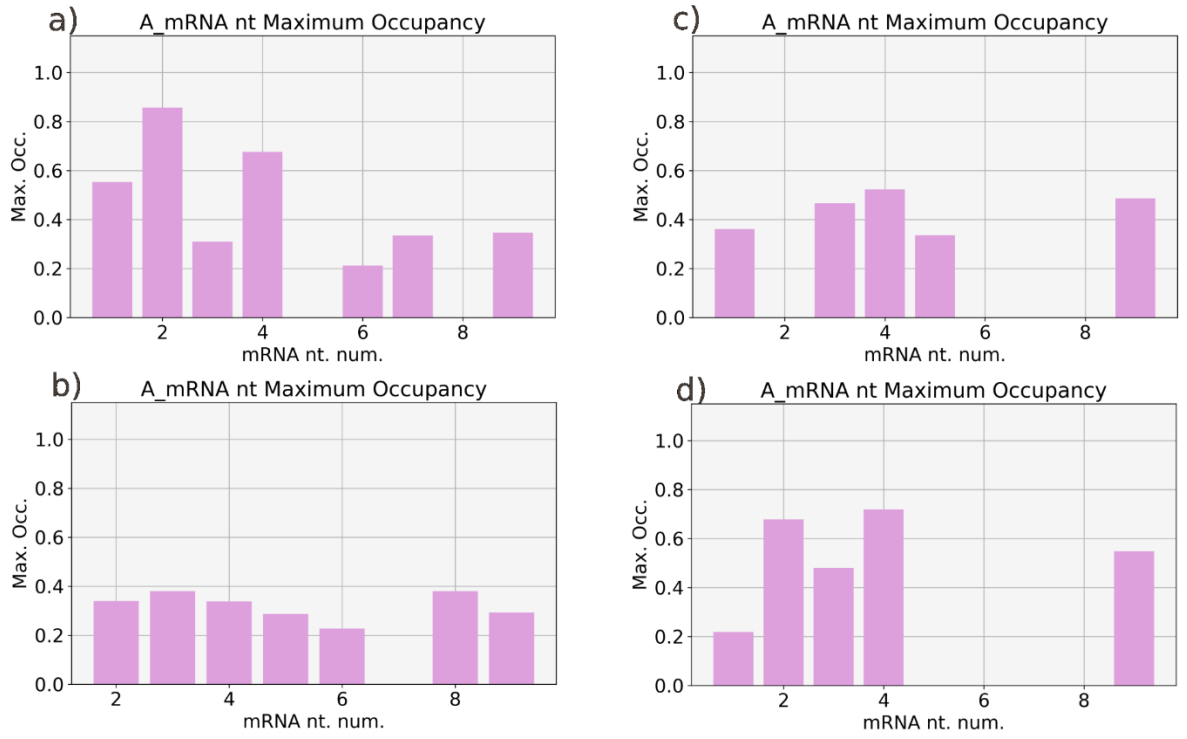

Figure S17. Maximum protein-mRNA hydrogen bond Occupancy versus nucleotides for samples WP0 (a), WP1 (b), WMP0 (c), WMP1 (d).

Table S7. Protein – mRNA hydrogen bonds (Occ. > 0.3) grouped on protein domains.

	WP0	WP1	WMP0	WMP1
L2	T9-R438* T9-Met437* T3_ASP358*	T3-T361*		T3-Lys355* T3-Thr361
MID	T9-GLN558* T4-Lys525 T2-Lys533 T2-Tyr529	T2-Tyr529* T4-Lys525* T8-GLN558*	T9-Ser561 T4-Lys525 T3-Lys525	T9-GLN558 T4-Lys525 T2-Lys533
PIWI	T7-Lys726* T1-HIS807		T5-GLN757* T1-His807*	

Protein-Protein Hydrogen Bonds

in Tables S8 and S9, the time averaged hydrogen bonds inter and intra domain are shown: intra domain hydrogen bond number (hbn) are lower (3%) for mixed solvent systems; greater, for the same systems, the intra domain ones (3%), above all in the case of PIWI unit (hbn increment 9%).

Table S8. Protein-solvent, protein-protein, protein-miRNA average hydrogen bond number. nm)..

	Prot-Solv[a]	Prot-Solv /mol_solv[b]	Prot-Prot	Prot-miRNA
WP0	1611	4 *10-2	615	49
WP1	1642	4.1*10-2	610	39
WMP0	1048 w 405 met	5 *10-2	631	40
WMp1	1056 w 411 met	5 *10-2	629	37

[a] w term stays for protein-water hydrogen bonds, met for protein methanol ones in the mixed solvent samples.

Table S9. In each cell, the time averaged hydrogen bonds inter and intra domain are reported (WP0,WP1 in the first line, WMP0, WMP1 in the second one): in the domain column appears also the domain atom number. number.

	N.	L1	PAZ	L2	MID	PIWI
N (1444)	60, 60 63, 61	3, 3 2, 4	1, 1 2, 2	3, 3 2, 2	0, 0 0, 0	0, 0 0, 0
L1 (1395)	3, 3 2, 4	59, 59 60, 60	8,7 7, 7	5, 5 6, 6	0, 0 0, 0	5, 5 5, 4
PAZ (1959)	1, 1 2, 2	8, 7 7, 7	81,83 84,85	3, 3 4, 3	0, 0 0, 0	2,<< 0, >0
L2 (1553)	3, 3 2, 2	5, 5 6, 6	3, 3 4, 3	50,51 54,51	9, 8 8, 10	12,12 8,12
MID (2132)	0, 0 0, 0	0, 0 0, 0	0,0 0,0	9, 8 8, 10	97,94 96,99	10,10 11,11
PIWI (4437)	0, 0 0, 0	5, 5 5, 4	2, >0 0, >0	12, 8 12,12	10,10 11,11	209, 208 221,216

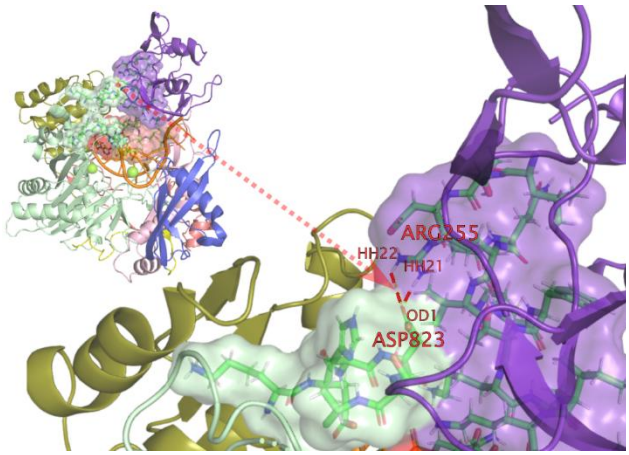

Figure S18. Detail of the interaction between ARG255 (atoms HH21, HH22) and ASP823(OD1) in WP0 sample at time 380 ns.

SOLVENT-SOLVENT HYDROGEN BONDS

The number of water-water hydrogen bonds for single molecule is rhw_w=1.75 in pure water; in mixed solvent this value is greater and stands at rhw_wm=1.85: however in mixed solvent the ratio total hydrogen bond number on total molecule number is rhwm_wm= 1.44, in reason of the lower ability of methanol to make methanol-methanol and methanol-water hydrogen bonds (respectively rhmm_wm=0.27 and rhmm_wm=1.77), especially as donor.

RNA-SOLVENT HYDROGEN BONDS

For miRNA the average number of solvent hydrogen bonds in pure water is greater than the one in mixed solvent (128 vs 118 and in the mixed solvent 25 H-bonds over 118 are with methanol); for mRNA, again, the hydrogen bond number is greater for samples in pure water (74 vs 68 and in the mixed solvent 15 over 68 are with methanol).

MAGNESIUM CATION COORDINATION

As expected, all three magnesium cations coordinate six atoms each. The magnesium atom lying on the central miRNA region strongly binds the phosphate oxygens of nucleotides g12 and g13, in sample WP0 also N7 of adenosine g14; remaining molecules are solvent molecules, one of them a methanol molecule in mixed solvent sample.

The magnesium cation close to the center of mass of the seed box is connected to U4O4 of mRNA, and five solvent molecules (in mixed solvent two are methanol molecules).

The cation surrounded by the DEDH tetrad is firmly bound to residues ASP597 (D) and coordinates five solvent molecules (two are methanol molecules in mixed solvent samples). In the WP0 case the residues 669 (D) lies at the same distance of residues 807(H), inferior to 0.5 nm, while GLU637 (E) stays at 1 nm from the ion; for WP1 and WMP0, ASP669 is strictly close to the cation, GLU637 and HIS807 are respectively at 1 nm and at 0.4 nm from it; sample WMP1 shows ASP669 strictly close to the cation, ASP637 and HIS807 respectively at 0.45 nm and at 0.8 nm from it: the distance HIS807-MG cation presents the maximum standard deviation.

PRINCIPAL COMPONENT ENERGY AND COVARIANCE MATRICES

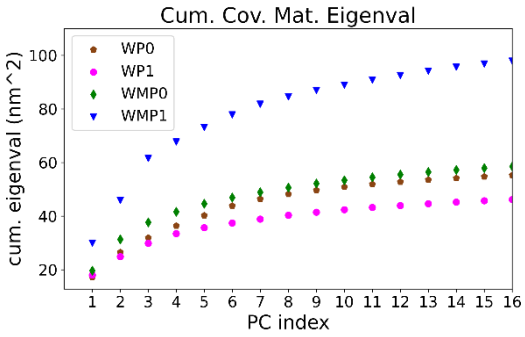

Figure S19. Cumulative Covariance Matrix Eigenvalue trends: for all the samples the ratio (sum of first ten eigenvalues)/(total eigenvalues sum) is greater then 0.85.

In figure S17 the trajectory covariance matrices, are reported, the minimum and the maximum value (/nm^2) for each sample being: -0.120902 0.279955 for WP0, -0.0803914 0.257289 for WP1, -0.0763909 0.235881 for WMP0, -0.40801 0.605551 for WMP1. PAZ and PIWI intra domain covariance data sums (/nm^2) are also presented: 7287.7 and 5057.0 for WP0, 5826.5 and 6357.7 for WP1, 7372.5 and 6655.6 for WMP0, 6461.0 and 11292.3 for WMP1.

Sample WMP1 shows very high covariance values relative to PIWI interval 820-840, both for intra and inter domain covariance data: other intra and iter domain regions are less relevant showing some more significant values along matrix diagonal around sequences 283-335, 388-397, 503-511, 143-154 (PAZ, L2, MID, N).

WMP0 is the most correlated sample with the less intense PIWI region 817-844 and strong correlation N-PIWI for the N sequence 117-123. Also intense are intra ed iter domain covariance PAZ data (287-337, 241-256), above all, and MID (494-512). On the matrix diagonal line to be noted the spot around residues 392, 425, 461 (not visible in the figure).

Sample WP1 shows inter and intra domain relevant covariance values for intervals 815-855, 454-547, 247-340; again inter domain relevant data for diagonal spots at residues 187, 425, 582.

WP0 presents more intense intra and inter domain values for the whole PAZ and PIWI 820-840 regions and for diagonal spots at 187, 424, 540.

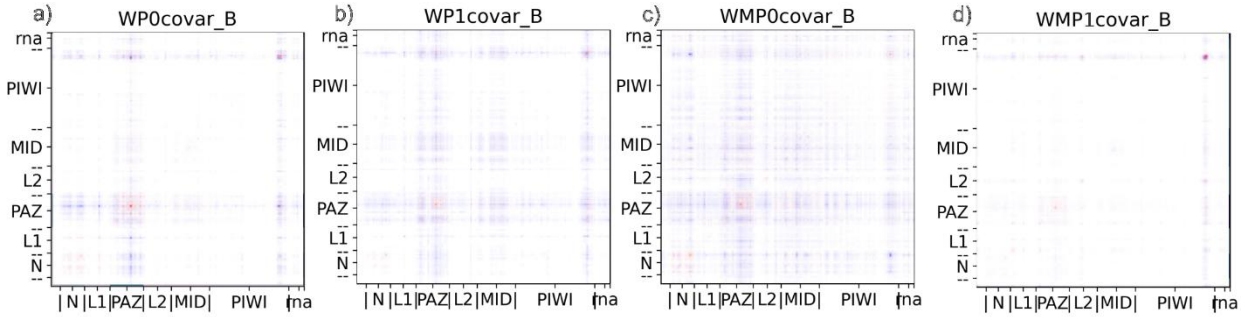

Figure S20. Covariance matrices of the full complex with positive (blue) and negative (red) value; into 'rna' section, the auxiliary internal tick separates miRNA from mRNA.