

# New MicroRNAs Candidates to Treat Human Colorectal Cancer; Molecular Dynamic Simulations Found a Major Down-Regulator of CLCA4 Tumor Suppressor Gene

**Fariborz Asghari Alashti**

University of Tehran

**Bahram Goliaei** (✉ [goliaei@ut.ac.ir](mailto:goliaei@ut.ac.ir))

University of Tehran

**Leila Karami**

Kharazmi University

**Serguei Vassiliev**

University of New Brunswick

**Najmeh Jooyan**

University of Shiraz: Shiraz University

**Zarrin Minucheher**

NIGEB: National Institute for Genetic Engineering and Biotechnology

---

## Research Article

**Keywords:** microRNA, Molecular dynamics simulation, binding free energy, CLCA4, Colorectal cancer, GEO datasets

**Posted Date:** April 14th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-381601/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Introduction

Colorectal cancer (CRC) is one of the most common malignancies worldwide. The expression of *CLCA4*, a tumor suppressor gene, decreases significantly in cancer cells of CRC. In this study, we identified miRNAs target the mRNA of the *CLCA4* gene.

## Objective

The aim of this study was the identification of miRNAs involved in CRC.

## Material and methods

We predicted miRNA(s) that target *CLCA4* mRNA applying TargetScan v.7. Then through analysis of Gene Expression Omnibus (GEO) datasets, among them, miRNA(s) over-expressed in CRC cells were determined. To identify miRNAs with the highest potential to down-regulate *CLCA4* through binding, we calculated the binding free energies of the candidate miRNA- mRNA complexes using the molecular mechanics energies combined with several solvation models: The Poisson–Boltzmann (MM/PBSA), the generalized Born (MM/GBSA), and the three-dimensional reference interaction site model with Kovalenko–Hirata closure relation (3D-RISM-KH).

## Results

Our TargetScan analysis predicted that 106 miRNAs could bind to *CLCA4* 3' UTR mRNA. Hsa-miR-934, hsa-miR-574-5p, hsa-miR-377-3p, hsa-miR-5580-3p, hsa-miR-4775, hsa-miR-590-3p and hsa-miR-501-5p showed increased expression in CRC samples compared to normal cells. MD results found the lowest free energy changes in three hsa-miR-377-3p, hsa-miR-574-5p and hsa-miR-501-5p miRNAs.

## Conclusion

This research beside introducing a new fast and low cost plan to find best candidate of miRNAs to bind their targets, suggested miR-501-5p as a biomarker for early diagnosis of CRC. As well, preventing of down regulation of the *CLCA4* expression through interrupting in the expression of miR-574-5p and miR-377-3p and more effectively miR-501-5p probably treat or slow down the development of colorectal cancer.

## 1. Introduction

Colorectal cancer (CRC) is the third most common cancer in the world and the third diagnosed cancer among men and women in the United States [1]. This tumor is usually diagnosed at the average age of 69 years, in which 60% are over the age of 65 and 36% are over 75 years [2]. A Mortality from this cancer

has been reduced due to the knowledge of risk factors, early detection, and prevention. However, CRC is still one of the leading health problems all over the world.

*CLCA4* significantly downregulated in colon and breast cancer cells. The protein encoded by *CLCA4* belongs to the family of calcium-sensitive chloride conductance proteins expressed mainly in the colon and inhibits the proliferation of cancer cells, so plays a critical role in suppression of CRC [3, 4]. The decreased expression of this protein could change intracellular and extracellular acidity. Due to the effects of extracellular pH on the function of the immune system, and the weakening of many cellular responses at lowered extracellular pH [5], it can be suggested that the decreased expression of *CLCA4* prevents the immune system from killing cancer cells. Messenger RNA (mRNA) destabilization and translational silencing are among the most factors leading to the repression of protein production.

Up to third of the genes encoding proteins in the human genome regulate by miRNAs through binding to the 3'-UTR (untranslated region) of target mRNA. More than 50 percent of these genes are in the sensitive and vulnerable genomic cancer-related regions. miRNAs are short non-coding RNAs, approximately 19–22 nucleotides long. They regulate gene expression at the post-transcriptional level [6, 7]. Genes regulated by miRNAs controls important biological processes including response to stress, cell proliferation, apoptosis, cell division, differentiation, and cancer [8]. Some miRNAs are oncomirs while others are tumor suppressors. The overexpression of the former and underexpression of the latter leads to cancer development [9].

The role of miRNAs in tumor suppressing and oncogenesis of colorectal cancer has been studied and proven. For instance, up-regulation of miR-378 suppress proliferation, migration and invasion of colon cancer cells by inhibiting *SDAD1* [10] and inversely down-regulation of miR-21 decreased tumor grow through reduction in  $\beta$ -catenin, *SOX9*, expression of some proinflammatory and procarcinogenic cytokines [11]. Therefore, miRNAs scanning (especially circulating ones) are recommended as highly sensitive and specific noninvasive test to diagnose and control cancers including colorectal cancer [12]. So finding appropriate miRNAs associating in carcinogenicity is a crucial step not only to early diagnosis of cancer but also to treat that through blocking or overexpression related miRNA.

Given the importance of the *CLCA4* gene in CRC and the serious role of miRNAs in gene expression regulation, in this study we focused on identifying miRNAs that could be the cause of decreased *CLCA4* expression in CRC patients. Identifying such miRNAs along with other identified miRNAs will be an important step to design a treatment for CRC. Increasing the number of miRNAs as well as finding the best ones can decrease false negatives diagnosis. For finding these relevant miRNAs, molecular dynamics (MD) simulations of the miRNA-mRNA (binary) complexes were performed. The purpose of these extensive simulations is to better understand the miRNA molecular recognition mechanism by providing atomic-level details that are unreachable in experiments because of resolution restrictions. Moreover, since the stability of miRNA-mRNA duplex structure can be predicted by calculating thermodynamic parameters such as binding affinities of nucleic acids strands [13] thermodynamic analyses were also done to profoundly investigate the interactions between miRNA and its target

applying MM/PBSA (molecular mechanics Poisson-Boltzmann surface area), MM/GBSA (Molecular Mechanics/Generalized Born Surface Area) and 3DRISM-KH (three-dimensional reference interaction site model with the Kovalenko–Hirata).

## 2. Material And Methods:

### 2.1. Finding target for miRNAs

Target sequences of miRNA in this study were predicted using TargetScan v.7, 1 web server [14]. The TargetScan is one of the most reliable tools for predicting biological targets of miRNAs. This tool searches the existence of conserved 8mer, 7mer, and 6mer sites that match the seed region of each miRNA using an extensive set of 14 robust features related to miRNA, mRNA and site [14].

### 2.2. Analysis of Gene Expression Omnibus (GEO) datasets

Although a reduction in the level of CLCA4 gene expression in CRC has been proven [15], we downloaded GEO databases to confirm this decline. We studied five GEO datasets of GSE21510 [16], GSE25071 [17], GSE15781 [18], GSE32323 [19] and GSE8671 [20] related to the mRNA and protein expression of genes in CRC and control profiles by GEO2R analyzer. Also, in order to investigate whether the miRNAs detected by the Targetscan program were upregulated or downregulated in CRC, at first, we carried out library studies, then downloaded and analyzed the miRNAs expression data of four GEO datasets of GSE35602 [21], GSE35982 [22], GSE39833 [23] and GSE77380 [24] in CRC.

### 2.3. Detecting miRNAs targeting CLCA4 mRNA gene

To detect miRNA targets, a computational prediction is one of the most efficient methods [25]. To predict the miRNAs that have conserved sites on CLCA4, we used TargetScan v.7.1 [14]. The TargetScan is one of the most robust and reliable tools for predicting biological targets of miRNAs [14]. While the basic features of the miRNA-mRNA interactions, such as seed match, site accessibility, free energy, and conservation, are considered in all miRNA target prediction tools [26], TargetScan implements an extensive set of 14 features related to miRNA, mRNA, and site [14]. It also deliberates the degree of repression and the cooperative miRNA function [14]. According to the identifier of CLCA4 corresponds, we used the most prevalent transcript (ENST00000370563.3) with 409nt, 3' UTR length for CLCA4 (ENSG00000016602.8) (Fig. 1).

Figure 1. The sequences of 409 nt of 3' UTR of CLCA4 and highly upregulated miRNAs in CRC tissues in comparison with normal tissues which could be bound to.

### 2.4. Molecular models of miRNA-mRNA complexes

Because crystallographic structures of miRNA-mRNA complexes are unavailable, we constructed their 3D models. First we created unimolecular complexes by joining the strands of miRNAs and their mRNA targets by a GAAA tetramer loop (Fig. 2).

Then we used these single-stranded RNA sequences in SimRNAWeb [27], an RNA modeling tool, to generate 3D structure models. Afterward, the GAAA tetraloop was removed to disperse the two strands (Fig. 2).

Figure 2. Schematic representation of introducing a GAAA tetraloop between miRNA and mRNA strands.

## 2.5. MD simulations

All molecular dynamics simulations were carried out using the Amber18 [28] package with the ff99bsc0\_chiOL3 force field [29]. To neutralize the system, sodium counterions were used. After neutralization  $\text{Na}^+$  and  $\text{Cl}^-$  ions were added to mimic 0.15M salt solution. All model systems of miRNA-mRNA complexes were solvated in a box of TIP3P water molecules with a spacing between the edge of the box and the RNA molecules of  $15 \text{ \AA}$ . Particle Mesh Ewald (PME) was engaged to treat electrostatic interactions with the default settings. The steepest descent energy minimization with 500 steps was carried out to relax any structural bad contacts in the solvated systems. It was followed by 50ps of heating to 300K in the NVT ensemble with restrained RNA heavy atoms (constraint force constant =  $2 \text{ kcal}/(\text{mol} \cdot \text{\AA}^2)$ ), 50 ps of density equilibration in NPT ensemble at 300K with the same constraints, and 10 ns equilibration without restraints in the NPT ensemble at 300K with PMEMD program from Amber18. All simulations were performed with a 2 fs time step, Langevin thermostat and SHAKE algorithm to constrain bonds with hydrogen atoms. Each of the eight solvated systems was simulated for 200 ns.

We used the CPPTRAJ module of Amber18 [30] to analyze MD trajectory files. Hydrogen bonds were identified with a distance cutoff of  $3.5 \text{ \AA}$  and an angle cutoff of  $135^\circ$ .

## 2.6. Binding free energy calculation

Hybridization of miRNA with mRNA makes miRNA-mRNA complex, accordingly the free energy difference between complex and two dissociated RNA strands can be calculated as:

$$\Delta G_{\text{bind}} = \langle G_{\text{miRNA-mRNA}} \rangle - \langle G_{\text{miRNA}} \rangle - \langle G_{\text{mRNA}} \rangle$$

Where the angular brackets denote average over all MD trajectory frames.

The binding free energy is composed of the gas-phase MM energetic contribution  $\Delta E_{\text{MM}}$ , entropic contribution, and the solvation free energy:

$$\Delta G_{\text{bind}} = \langle \Delta E_{\text{MM}} \rangle + \langle \Delta G_{\text{sol}} \rangle - T \langle \Delta S_{\text{MM}} \rangle$$

Molecular mechanics and solvation contributions can be further decomposed into the following terms:

$$E_{\text{MM}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdW}}$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{pol}} + \Delta G_{\text{nonpol}}$$

Here  $E_{MM}$  (the total gas phase energy) which is composed of  $E_{int}$  (internal energy of bonded interactions),  $E_{ele}$  (electrostatic energy) and  $E_{vdW}$  (van der Waals energy) between the strands of miRNA-mRNA complex was acquired using Amber18.

The calculation of  $\Delta G_{sol}$  for all miRNA-mRNA complexes was carried out by three methods (MM/GBSA, MM/PBSA and MM/3D-RISM-KH) on 200 ns MD trajectories using single-trajectory protocol.

The solvation free energy  $\Delta G_{sol}$  is comprised of  $\Delta G_{pol}$  (polar) and  $\Delta G_{np}$  (nonpolar) which were calculated using methods of Poisson-Boltzmann and generalized Born continuum electrostatics molecular theory of solvation as implemented in Amber18. In 3D-RISM-KH method, thermodynamic calculations, electronic properties and molecular solvent structure of a soluble molecule in a solvent with medium computational cost are possible and the solvent and soluble chemical properties such as hydrogen bonds and hydrophobic forces are provided by a three-dimensional site of solvent density distribution [31]. We have not decomposed solvation free energy into polar and nonpolar terms in calculations using 3D-RISM-KH method due to high computational cost.

Entropic contribution  $-T < \Delta S_{MM} >$  composed of the vibrational, rotational and translational components was calculated by normal mode analysis (nmode module of the Amber18). We used several solvation models to assess robustness of free energy calculations.

All simulations were carried out on Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)) high-performance clusters.

## 3. Results

### 3.1. Down-regulation of CLCA4 in CRC

Although downregulation of CLCA4 in CRC was proved in previous studies, our analysis on the data of CLCA4 mRNA expression in five GEO datasets (GSE21510, GSE25071, GSE15781, GSE32323 and GSE8671), verified the downregulation. According to the analysis of this data, we respectively found an 80-fold, 180-fold, 19-fold, 22-fold and 38-fold downregulation of CLCA4 mRNA expression in CRC tissues in comparison with normal tissues (Table 1).

Table 1  
Result of comparison CRC patients' tissue samples with control groups by GEO2R analyzer

P.Value	log <sub>2</sub> FC	Fold up/down	ID Dataset
4.50e-18	-6.59	-96	GSE21510 [16]
5.39e-16	-5.98	-63	
1.48e-04	-7.11	-138	GSE25071 [17]
6.22e-04	-6.54	-93	
1.76e-05	-8.27	-304	
3.08e-04	-4.25	-19	GSE15781 [18]
2.28e-06	-4.91	-30	GSE32323 [19]
8.59e-03	-3.86	-14	
4.05e-17	-5.26	-38	GSE8671 [20]

### 3.2. Prediction of miRNAs targeting CLCA4.

Applying TargetScan v.7,1 to predict miRNAs which could bind to CLCA4 3' UTR [14], we identified 106 miRNAs (Table 2).

Table 2

MiRNAs bound with position of CLCA4 3' UTR (repeated miRNAs show binding at different position)

<b>miRNAs bound with position of CLCA4 3' UTR</b>								
No	Name of miRNAs	Site type	No	Name of miRNAs	Site type	No	Name of miRNAs	Site type
<b>1</b>	hsa-miR-1206	7mer-A1	<b>37</b>	hsa-miR-520g-3p	7mer-A1	<b>73</b>	hsa-miR-4330	7mer-m8
<b>2</b>	hsa-miR-1236-3p	7mer-A1	<b>38</b>	hsa-miR-520h	7mer-A1	<b>74</b>	hsa-miR-4635	7mer-m8
<b>3</b>	hsa-miR-1252-3p	7mer-A1	<b>39</b>	hsa-miR-524-5p	7mer-A1	<b>75</b>	hsa-miR-4643	7mer-m8
<b>4</b>	hsa-miR-1277-5p	7mer-A1	<b>40</b>	hsa-miR-526b-5p	7mer-A1	<b>76</b>	hsa-miR-4658	7mer-m8
<b>5</b>	hsa-miR-138-1-3p	7mer-A1	<b>41</b>	hsa-miR-548a-3p	7mer-A1	<b>77</b>	hsa-miR-4717-3p	7mer-m8
<b>6</b>	hsa-miR-16-1-3p	7mer-A1	<b>42</b>	hsa-miR-548a-3p	7mer-A1	<b>78</b>	hsa-miR-548az-3p	7mer-m8
<b>7</b>	hsa-miR-19a-5p	7mer-A1	<b>43</b>	hsa-miR-548ar-3p	7mer-A1	<b>79</b>	hsa-miR-548az-3p	7mer-m8
<b>8</b>	hsa-miR-19b-1-5p	7mer-A1	<b>44</b>	hsa-miR-548ar-3p	7mer-A1	<b>80</b>	hsa-miR-548e-3p	7mer-m8
<b>9</b>	hsa-miR-19b-2-5p	7mer-A1	<b>45</b>	hsa-miR-548f-3p	7mer-A1	<b>81</b>	hsa-miR-548e-3p	7mer-m8
<b>10</b>	hsa-miR-2052	7mer-A1	<b>46</b>	hsa-miR-548f-3p	7mer-A1	<b>82</b>	hsa-miR-548g-3p	7mer-m8
<b>11</b>	hsa-miR-2681-5p	7mer-A1	<b>47</b>	hsa-miR-5580-3p	7mer-A1	<b>83</b>	hsa-miR-548l	7mer-m8
<b>12</b>	hsa-miR-300	7mer-A1	<b>48</b>	hsa-miR-5582-3p	7mer-A1	<b>84</b>	hsa-miR-5692a	7mer-m8
<b>13</b>	hsa-miR-302a-5p	7mer-A1	<b>49</b>	hsa-miR-5688	7mer-A1	<b>85</b>	hsa-miR-5692b	7mer-m8
<b>14</b>	hsa-miR-3606-3p	7mer-A1	<b>50</b>	hsa-miR-5691	7mer-A1	<b>86</b>	hsa-miR-5692c	7mer-m8
<b>15</b>	hsa-miR-362-5p	7mer-A1	<b>51</b>	hsa-miR-574-5p	7mer-A1	<b>87</b>	hsa-miR-590-3p	7mer-m8
<b>16</b>	hsa-miR-362-5p	7mer-A1	<b>52</b>	hsa-miR-578	7mer-A1	<b>88</b>	hsa-miR-590-3p	7mer-m8
<b>17</b>	hsa-miR-3658	7mer-A1	<b>53</b>	hsa-miR-600	7mer-A1	<b>89</b>	hsa-miR-643	7mer-m8

miRNAs bound with position of CLCA4 3' UTR								
18	hsa-miR-375	7mer-A1	54	hsa-miR-6515-3p	7mer-A1	90	hsa-miR-6728-5p	7mer-m8
19	hsa-miR-381-3p	7mer-A1	55	hsa-miR-6739-3p	7mer-A1	91	hsa-miR-6790-5p	7mer-m8
20	hsa-miR-382-3p	7mer-A1	56	hsa-miR-6805-3p	7mer-A1	92	hsa-miR-6844	7mer-m8
21	hsa-miR-3942-3p	7mer-A1	57	hsa-miR-6853-3p	7mer-A1	93	hsa-miR-6845-3p	7mer-m8
22	hsa-miR-4282	7mer-A1	58	hsa-miR-7112-3p	7mer-A1	94	hsa-miR-892c-5p	7mer-m8
23	hsa-miR-4645-3p	7mer-A1	59	hsa-miR-7151-5p	7mer-A1	95	hsa-miR-934	7mer-m8
24	hsa-miR-4666a-3p	7mer-A1	60	hsa-miR-7856-5p	7mer-A1	96	hsa-miR-944	7mer-m8
25	hsa-miR-4684-3p	7mer-A1	61	hsa-let-7a-3p	7mer-m8	97	hsa-miR-98-3p	7mer-m8
26	hsa-miR-4709-3p	7mer-A1	62	hsa-let-7b-3p	7mer-m8	98	hsa-miR-3646	8mer
27	hsa-miR-4711-5p	7mer-A1	63	hsa-let-7f-1-3p	7mer-m8	99	hsa-miR-3659	8mer
28	hsa-miR-4738-3p	7mer-A1	64	hsa-miR-138-2-3p	7mer-m8	100	hsa-miR-3973	8mer
29	hsa-miR-4775	7mer-A1	65	hsa-miR-205-3p	7mer-m8	101	hsa-miR-4699-5p	8mer
30	hsa-miR-495-3p	7mer-A1	66	hsa-miR-3163	7mer-m8	102	hsa-miR-501-5p	8mer
31	hsa-miR-5094	7mer-A1	67	hsa-miR-34c-3p	7mer-m8	103	hsa-miR-501-5p	8mer
32	hsa-miR-510-3p	7mer-A1	68	hsa-miR-3671	7mer-m8	104	hsa-miR-5187-5p	8mer
33	hsa-miR-512-5p	7mer-A1	69	hsa-miR-367-5p	7mer-m8	105	hsa-miR-7-1-3p	8mer
34	hsa-miR-513a-3p	7mer-A1	70	hsa-miR-377-3p	7mer-m8	106	hsa-miR-7-2-3p	8mer
35	hsa-miR-513c-3p	7mer-A1	71	hsa-miR-409-3p	7mer-m8			
36	hsa-miR-520d-5p	7mer-A1	72	hsa-miR-4328	7mer-m8			

### 3.3. miRNAs up-regulated in CRC

We have performed library studies based on laboratory experiments about 106 miRNAs were predicted targeting CLCA4, to identify those miRNAs which have been reported on their increase in CRC tissues. Our results showed that 6 miRNAs including hsa-miR-4775, hsa-miR-5580-3p, hsa-miR-574-5p, hsa-miR-590-3p and hsa-miR-934 and hsa-miR-377-3p, were highly expressed in CRC tissues in comparison with normal tissues [32–37].

### 3.4. Analysis of GEO datasets for identifying upregulated miRNAs

106 miRNAs were investigated for the extent of their expression in accordance with experimental data of the gene expression of GEO. So, we analyzed the downloaded GEO datasets of GSE35602 [21], GSE35982 [22], GSE39833 [23] and GSE77380 [24]. The results showed that hsa-miR-501-5p, which is among the miRNAs with 8mer site type seeds (potentially are expected to be the most forceful [38] and has the high value of repression [38, 39] ) is more highly expressed in CRC than normal (Table 3). This miRNA could bind to the two sequences of CLCA4 mRNA; so, CLCA4 mRNA with two binding sites for this miRNA could have more chance to be adhered.

Table 3  
Results of comparison expression of hsa-miR-501-5p in CRC patients' tissue samples with normal samples by GEO2R analyzer.

P.Value	log <sub>2</sub> FC	Fold up/down	ID Dataset
0.0071185	1.18	2.3	GSE35602 [21]
3.72e-02	1.09	2.2	
0.0300922	1.87	3.7	
0.0091826	1.57	2.9	GSE35982 [22]
2.72e-02	1.43	2.7	GSE39833 [23]
3.97e-02	3.73	13.3	GSE77380 [24]

### 3.5. Choosing miRNAs upregulated in CRC samples for studies with MD simulations.

According to analysis of GEO datasets of CRC samples and library studies of experimental results, eight miRNAs (hsa-miR-934, hsa-miR-574-5p, hsa-miR-377-3p, hsa-miR-5580-3p, hsa-miR-4775, hsa-miR-590-3p, hsa-miR-501-5p with mRNA target position of 3' UTR 69–76 and hsa-miR-501-5p with mRNA target position of 3' UTR 144–151) were chosen to be analyzed with MD simulations. There are five site types in total: 8mer, 7mer-m8, 7mer-A1, 6mer, and 6mer offset (listed in the order from the strongest binding to the weakest). All 106 miRNAs with the ability to bind to the mRNA sequence of the CLCA4 gene predicted by

TargetScan are binding to 8mer, 7mer-m8, or 7mer-A1 site types. Thus miRNAs selected for MD simulations belong to these three groups as shown in Fig. 3.

Figure 3. Secondary structure topology derived from 3D RNA models. Base-pairs identified by 3DNA-DSSR v2.0.0-2020aug01 [<https://x3dna.org>]. Images generated by VARNA v.3.93 [40][<http://varna.lri.fr>]. Seed regions are surrounded with dashed lines.

To compare simulated structures with the initial structure of the duplex, we plotted the root-mean-square deviation (RMSD) of RNA backbone atoms C3', C4', C5', O5' and P as a function of time (Fig. 4). The results show that all complexes except miR-590 reached equilibrium state during the first 10 ns of simulation. RMSD of miR-590 exhibited a modest increase at about 120 ns of simulation which persisted up to 200 ns. A longer simulation is required to determine whether it is a temporary fluctuation or persistent conformational change. The largest fluctuations were observed in miR-501(144) complex.

Figure 4. Root-mean-square deviation (RMSD) of miRNAs-mRNA complex backbone atoms with respect to the initial conformation as a function of time. RMSDs of miRNAs-mRNA complexes in three binding site groups (8mer, 7mer-A1, and 7mer-m8) are shown in panels A, B, and C respectively. Time starts from the beginning of production run after 10 ns long- equilibration.

Also to study the differences of structure, we drew snapshots during simulation time of 50, 100, 150, and 200 ns (Fig. 5).

Figure 5. Snapshots of 8 miRNA-mRNA complexes at simulation times of 50, 100, 150 and 200 ns. In the last column all 4 snapshots are superimposed.

## 3.6. Lifetime of hydrogen bonds

Reports on the lifetime of hydrogen bonds were extracted and summarized in the table below. These results indicate that hsa-miR-501-5p, hsa-miR-377-3p, and hsa-miR-574-5p complexes have 65 (two sites), 37 and 34 hydrogen bond respectively (27) in the seed region. These miRNAs not only have more cases of Hbonds in the seed region but also these bonds are more stable and lasting for the most time of simulation. As well miR-501(144) and miR-574 showed longer HB lifetime in comparison to others for the most time of simulation. Inversely miR-5580 complex presented both the least number of HB and HB life time. All data has been presented in Table 4.

Table 4  
Information of HB analysis in seed and out of seed area

miRNA-mRNA complementray	NumHB	Present HB over 60% of simulation time			Max lifetime in seed region	Max lifetime out of seed region	
		Cases in seed region	Average in seed region	Max out of seed region			
miR-501144	37	9 out of 21	0.915	0.723	0.915	42	46
miR-574	34	11 out of 21	0.905	0.735	0.78	35	20
miR-377	37	11 out of 20	0.865	0.742	0.86	29	26
miR-934	31	8 out of 18	0.9	0.7281	0.91	31	29
miR-50169	28	9 out of 18	0.88	0.721	0.88	24	49
miR-590	30	6 out of 15	0.875	0.668	0.80	21	33
miR-4775	32	3 out of 18	0.655	0.638	0.91	12	64
miR-5580	27	0	0.18	0	0.92	10	41

### 3.7. Binding Free energy calculations

The data from the calculation of binding free energy of all miRNA-mRNA complexes by MM / GBSA) demonstrate that hsa-miR-377-3p, hsa-miR-574-5p and hsa-miR-501-5p, respectively, have the lowest  $\Delta E_{total}$  and  $\Delta G_{bind}$  in the 8 studied complexes with more than - 100 and - 76 Kcal / mol respectively (Fig. 6). Also, the results of data analysis of free energy and their physical components by the three studied methods are summarized in Table 5.

Figure 6. Results of calculation of  $\Delta G_{bind}$  as a function of number of frames with MM/GBSA method. 501(144) and 501(69) are labeled as 501a and 501b, respectively. Shaded areas in the plots indicate the range of the 95% confidence intervals for the mean. (Plot using full quasi-harmonic entropy vibr + rot + trans).

The data in the table shows that  $-T < \Delta S_{MM} >$ , the average contribution of entropy to the binding free energy obtained by normal mode analysis is the highest in hsa-miR-574-5p, hsa-miR-501-5p and hsa-miR-377-3p. The most negative  $\Delta E_{total}$  calculated in all three methods are related to the hsa-miR-574-5p,

hsa-miR-501-5p and hsa-miR-377-3p complexes. Also free binding energy for these three miRNAs are -81.1, -80.5 and -76.8.

Although there is a little difference in the results of  $\Delta G$  bind in the order of the most negative with three methods, miRNAs of 377, 574, 501(144), 934, 4775, 5580, 501(69), overall considering entropy contribution binding energies of most complexes are similar and 501(69) seems to be slightly better than others and these miRNAs seem to be almost equally good candidates to bind with their CLCA4 mRNA target and miR-590 cannot be considered in this case.

Table 5

The binding free energy and its physical components of all complexes with three different used methods of MM/GBSA, MM/PBSA, and MM/3D-RISM-KH calculated using 20000 samples from 200 ns long trajectories.

Energy Terms	Mir-377	MIR_4775	miR-574	miR-934	miR-5580	miR-501(69)	miR-501(144)	miR-590
$\Delta E_{ele}$	6621	5884	6730	4318	3830	2713	6570	5984
$\Delta E_{vdW}$	-149	-131	-162	-113	-106	-88	-174	-139
$\Delta E_{MM}$	6471	5753	6568	4205	3724	2625	6396	5845
$-T\Delta S_{MM}$	41.56	31.49	43.98	28.02	13.74	5.42	45.37	29.26
$-T\Delta S_{MM}$ by nmode	60.54	50.51	63.29	40.40	31.98	26.20	60.82	52.85
MM/GBSA								
$\Delta G_{sol}$	-6594	-5857	-6693	-4301	-3805	-2697	-6518	-5931
$\Delta G_{pol}$	-6580	-5845	-6678	-4290	-3795	-2689	-6503	-5917
$\Delta G_{np}$	-14	-12	-14	-10	-10	-8	-16	-14
$\Delta E_{pol} + \Delta E_{ele}$	41	39	52	28	35	24	68	67
$\Delta E_{total}^*$	-123	-104	-124	-96	-80	-72	-122	-86
$\Delta G_{bind}^{**}$	-81.1	-72.9	-80.5	-68.0	-66.8	-67.0	-76.8	-56.3
MM/PBSA								
$\Delta G_{sol}$	-6551	-5856	-6655	-4282	-3786	-2688	-6485	-5867
$\Delta G_{pol}$	-6538	-5845	-6642	-4273	-3777	-2681	-6470	-5855
$\Delta G_{np}$	-13	-11	-13	-10	-9	-8	-14	-12
$\Delta E_{pol} + \Delta E_{ele}$	83	39	89	45	53	32	100	130
$\Delta E_{total}$	-80	-75	-87	-78	-62	-63	-88	-22
$\Delta G_{bind}$	-38.3	-43.4	-43.1	-49.7	-48.5	-58.0	-43.0	7.5
MM/3D-RISM-KH								

\*  $\Delta E_{total} = \Delta E_{MM} + \Delta G_{sol}$ . \*\* Binding free energy  $\Delta G_{bind} = \Delta E_{total} - T\Delta S_{MM}$

Energy Terms	Mir-377	MIR_4775	miR-574	miR-934	miR-5580	miR-501(69)	miR-501(144)	miR-590
$\Delta G_{sol}$	-6516	-5815	-6604	-4243	-3751	-2663	-6441	-5875
$\Delta E_{total}$	-46	-34	-37	-38	-27	-38	-45	-29
$\Delta G_{bind}$	-4.44	-2.71	7.08	-9.58	-12.86	-32.18	0.57	-0.04
* $\Delta E_{total} = \Delta E_{MM} + \Delta G_{sol}$ . ** Binding free energy $\Delta G_{bind} = \Delta E_{total} - T\Delta S_{MM}$								

## 4. Discussion

CRC is a major health problem and the third cause of death due to cancer. Altered expression of some genes, including tumor-suppressors or oncogenes, are the most important factors in tumorigenicity and cancer development [41]. CLCA4 is one of the tumor suppressor genes and its encoded protein is a calcium-dependent chloride channel. Because of its importance in regulation of cell acidity [42, 43] and immune system [4, 44], we checked CLCA4 expression in cancerous cells. Using five GEO datasets, our results showed down expression of CLCA4 mRNA in CRC tissues from 19 to 80 fold in comparison with normal tissues. These data is in accordance to previous data which reported down-regulation CLCA4 expression in CRC [15]

MiRNAs through binding to the target mRNA act as an epigenetic mechanism of gene regulation and are one of effective factors at gene down-expression [45, 46], so the decrease of expression level of CLCA4 in CRC prompted us to investigate miRNAs could target CLCA4 and determine the mechanism of their action. The role of some miRNAs has been reported in developing and repressing colorectal cancer. In this study, we considered identifying miRNAs which could repress CLCA4 expression. Applying TargetScan v.7, it was predicted that 106 miRNAs could bind to CLCA4 3' UTR MR. miRNAs bind to their mRNA target through 5 types of miRNA binding sites, including 8mer site (the highest affinity), 7mer-m8 (matched at position 8), 7mer-A1 site (adenosine at position 1), 6mer site and offset 6mer site (the least affinity). All determined miRNAs were binding to 8mer, 7mer m8, or 7mer-A1 site types. The 8mer site type (a precise match to positions 2–8 of the mature miRNA, followed by an 'A') of miRNA binding site, has more effects on gene repression and advises the robust suppression [38], probably because of the increased binding site complementarity [39].

Of this 106 recognized miRNAs, laboratory studies and our analysis through GEO datasets of CRC samples defined seven miRNAs (with eight sites) are highly expressed, so were selected to be analyzed with MD simulations. These miRNAs included hsa-miR-934, hsa-miR-574-5p, hsa-miR-377-3p, hsa-miR-5580-3p, hsa-miR-4775, hsa-miR-590-3p, hsa-miR-501-5p with mRNA target position of 3' UTR 69–76 and hsa-miR-501-5p with mRNA target position of 3' UTR 144–151. Snapshot structures obtained from the MD trajectory of system was used to calculate the average interaction energies of the miRNA and mRNA [47]. Results showed that three hsa-miR-377-3p, hsa-miR-574-5p and hsa-miR-501-5p miRNA, respectively, have the lowest entropy and free energy changes which is in line with the process of changes of in

observed hydrogen bond lifetime. These features along with the higher number of H bonds in the seed area, has made these tree complexes more stable and lasting. The lower free energy represents higher binding affinity that are more effective in gene suppression [48]. Therefore, hsa-miR-377-3p, hsa-miR-574-5p and hsa-miR-501-5p have major role in CLCA4 down-regulation and causing CRC. Considering these finding, these three miRNAs are recommended as potential targets to treat CRC. Inhibiting their binding to CLCA4 mRNA can prevent its down-regulation. Of this miRNAs, hsa-miR-501-5p bind to two sequences of CLCA4 mRNA with the highest number of H bonding. Also, as it was mentioned hsa-miR-501-5p is among the miRNAs with 8mer site type seeds which is the potentially most forceful and high value of repression miRNA [38, 39]. Therefore, CLCA4 mRNA with two binding sites for this miRNA have more chance to make hsa-miR-501-5p- CLCA4 mRNA complex. As a miRNA with high ability to bind to CLCA4 mRNA and reduce its expression, which up-regulates in CRC, miR-501-5p can be a biomarker for early detection of this cancer. As well, interrupting in conforming of hsa-miR-501-5p- CLCA4 mRNA complex may be an effective way to prevent or slow-down development of CRC. These findings shed light on future experimental studies.

In conclusion, with the help of bioinformatics algorithms and prediction tools, we identified miRNAs which could target CLCA4 and found that their expression was upregulated in CRC tissues (according to library studies and by analyzing the GEO database). By studying and analyzing the structural condition and thermodynamic data of these eight miRNAs using molecular dynamics simulations, we have shown that seven miRNAs including miR-4775, miR-934, miR-5580 and mostly miR-574-5p, miR-377-3p and specially miR-501-5p with two binding positions can play the oncomir role in regulating of CLCA4 gene expression in CRC. Although more research is needed to understand whether downregulation of these miRNAs led to increasing CLCA4 gene expression, elevating intracellular pH and decreasing extracellular pH, these miRNAs could be potential candidates for colorectal cancer diagnosis and therapy.

## Declarations

### Acknowledgements:

We thank Burton Yang at Sunnybrook Health Sciences Centre, Department of Laboratory Medicine and Pathobiology, University of Toronto, for support of this research. This research was enabled in part by support provided by Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)).

### Conflicts of interests:

The authors have no conflicts of interest to declare that are relevant to the content of this article.

## References

1. Siegel et al (2017) Colorectal cancer statistics, 2017. *CA Cancer J Clin* 67: 177-193.
2. Falzone L, Scola L, Zanghì A, Biondi A, Di Cataldo A, et al. (2018) Integrated analysis of colorectal cancer microRNA datasets: Identification of microRNAs associated with tumor development. *Aging*

(Albany NY) 10: 1000.

3. Yu et al. (2013) Loss of CLCA4 promotes epithelial-to-mesenchymal transition in breast cancer cells. *PLoS One* 8: e83943.
4. Yang et al. (2015) Low expression of chloride channel accessory 1 predicts a poor prognosis in colorectal cancer. *Cancer* 121: 1570-1580.
5. Lardner A (2001) The effects of extracellular pH on immune function. *J Leukoc Biol* 69: 522-530.
6. Ambros V (2004) The functions of animal microRNAs. *Nature* 431: 350-355.
7. Felekis K, Touvana E, Stefanou C, Deltas C (2010) microRNAs: a newly described class of encoded molecules that play a role in health and disease. *Hippokratia* 14: 236.
8. Esquela-Kerscher A, Slack FJ (2006) Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 6: 259-269.
9. Shah et al. (2012) Mechanistic insights into the role of microRNAs in cancer: influence of nutrient crosstalk. *Front Genet* 3: 305.
10. Zeng et al. (2017) miR-378 suppresses the proliferation, migration and invasion of colon cancer cells by inhibiting SDAD1. *Cell Mol Biol Lett* 22: 12.
11. Shi et al. (2016) Novel evidence for an oncogenic role of microRNA-21 in colitis-associated colorectal cancer. *Gut* 65: 1470-1481.
12. Fesler A, Jiang J, Zhai H, Ju J (2014) Circulating microRNA testing for the early diagnosis and follow-up of colorectal cancer patients. *Molecular diagnosis & therapy* 18: 303-308.
13. Breslauer et al. (1986) Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A* 83: 3746-3750.
14. Agarwal et al. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4.
15. Chu et al. (2014) Gene expression profiling of colorectal tumors and normal mucosa by microarrays meta-analysis using prediction analysis of microarray, artificial neural network, classification, and regression trees. *Dis Markers* 2014: 634123.
16. Tsukamoto et al. (2011) Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin Cancer Res* 17: 2444-2450.
17. Agesen et al. (2011) CLC and IFNAR1 are differentially expressed and a global immunity score is distinct between early- and late-onset colorectal cancer. *Genes Immun* 12: 653-662.
18. Snipstad et al. (2010) New specific molecular targets for radio-chemotherapy of rectal cancer. *Mol Oncol* 4: 52-64.
19. Khamas et al. (2012) Screening for epigenetically masked genes in colorectal cancer Using 5-Aza-2'-deoxycytidine, microarray and gene expression profile. *Cancer Genomics Proteomics* 9: 67-75.
20. Sabates-Bellver et al. (2007) Transcriptome profile of human colorectal adenomas. *Mol Cancer Res* 5: 1263-1275.
21. Nishida et al. (2012) Microarray analysis of colorectal cancer stromal tissue reveals upregulation of two oncogenic miRNA clusters. *Clin Cancer Res* 18: 3054-3070.

22. Fu et al. (2012) Identifying microRNA-mRNA regulatory network in colorectal cancer by a combination of expression profile and bioinformatics analysis. *BMC Syst Biol* 6: 68.
23. Ogata-Kawata et al. (2014) Circulating exosomal microRNAs as biomarkers of colon cancer. *PLoS One* 9: e92921.
24. Yoshimoto et al. (2018) Pulmonary Carcinoids and Low-Grade Gastrointestinal Neuroendocrine Tumors Show Common MicroRNA Expression Profiles, Different from Adenocarcinomas and Small Cell Carcinomas. *Neuroendocrinology* 106: 47-57.
25. Riffo-Campos et al. (2016) Tools for Sequence-Based miRNA Target Prediction: What to Choose? *Int J Mol Sci* 17.
26. Peterson et al. (2014) Common features of microRNA target prediction tools. *Front Genet* 5: 23.
27. Magnus et al. (2016) SimRNAweb: a web server for RNA 3D structure modeling with optional restraints. *Nucleic Acids Res* 44: W315-319.
28. Case et al. (2018) Amber 2018. University of California, San Francisco.
29. Zgarbova et al. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* 7: 2886-2902.
30. Roe DR, Cheatham TE, 3rd (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* 9: 3084-3095.
31. Joshy et al. (2014) Calculation of binding free energy of short double stranded oligonucleotides using MM/3D-RISM-KH approach. *Journal of Molecular Liquids*.
32. Zhao et al. (2017) miR-4775 promotes colorectal cancer invasion and metastasis via the Smad7/TGFbeta-mediated epithelial to mesenchymal transition. *Mol Cancer* 16: 12.
33. Wang et al. (2017) Novel circulating microRNAs expression profile in colon cancer: a pilot study. *Eur J Med Res* 22: 51.
34. Ji et al. (2013) miR-574-5p negatively regulates Qki6/7 to impact beta-catenin/Wnt signalling and the development of colorectal cancer. *Gut* 62: 716-726.
35. Sun et al. (2017) MiR-590-3p promotes proliferation and metastasis of colorectal cancer via Hippo pathway. *Oncotarget* 8: 58061-58071.
36. Zhang et al. (2014) Genome-wide uncovering of STAT3-mediated miRNA expression profiles in colorectal cancer cell lines. *Biomed Res Int* 2014: 187105.
37. Liu et al. (2018) miR-377-3p drives malignancy characteristics via upregulating GSK-3beta expression and activating NF-kappaB pathway in hCRC cells. *J Cell Biochem* 119: 2124-2134.
38. Steinkraus et al. (2016) Tiny giants of gene regulation: experimental strategies for microRNA functional studies. *Wiley Interdiscip Rev Dev Biol* 5: 311-362.
39. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215-233.
40. Darty K, Denise A, Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25: 1974-1975.

41. Liu et al. (2017) miR-19a promotes colorectal cancer proliferation and migration by targeting TIA1. Mol Cancer 16: 53.
42. Webb et al. (2011) Dysregulated pH: a perfect storm for cancer progression. Nat Rev Cancer 11: 671-677.
43. Lardner A (2001) The effects of extracellular pH on immune function. Journal of leukocyte biology 69: 522-530.
44. Hou et al. (2017) CLCA4 inhibits bladder cancer cell proliferation, migration, and invasion by suppressing the PI3K/AKT pathway. Oncotarget 8: 93001-93013.
45. Wuet al. (2006) MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A 103: 4034-4039.
46. Kamińska K, Nalejska E, Kubiak M, Wojtysiak J, Żoła Ł, et al. (2019) Prognostic and predictive epigenetic biomarkers in oncology. Molecular diagnosis & therapy 23: 83-95.
47. Ouyang D, Zhang H, Herten D-P, Parekh HS, Smith SC (2010) Structure, Dynamics, and Energetics of siRNA- Cationic Vector Complexation: A Molecular Dynamics Study. The Journal of Physical Chemistry B 114: 9220-9230.
48. Cao Y, Xu Q, Jin ZD, Zhou Z, Nie JH, et al. (2011) Induction of adaptive response: pre-exposure of mice to 900 MHz radiofrequency fields reduces hematopoietic damage caused by subsequent exposure to ionising radiation. Int J Radiat Biol 87: 720-728.

## Figures



Figure 1

The sequences of 409 nt of 3' UTR of CLCA4 and highly upregulated miRNAs in CRC tissues in comparison with normal tissues which could be bound to.

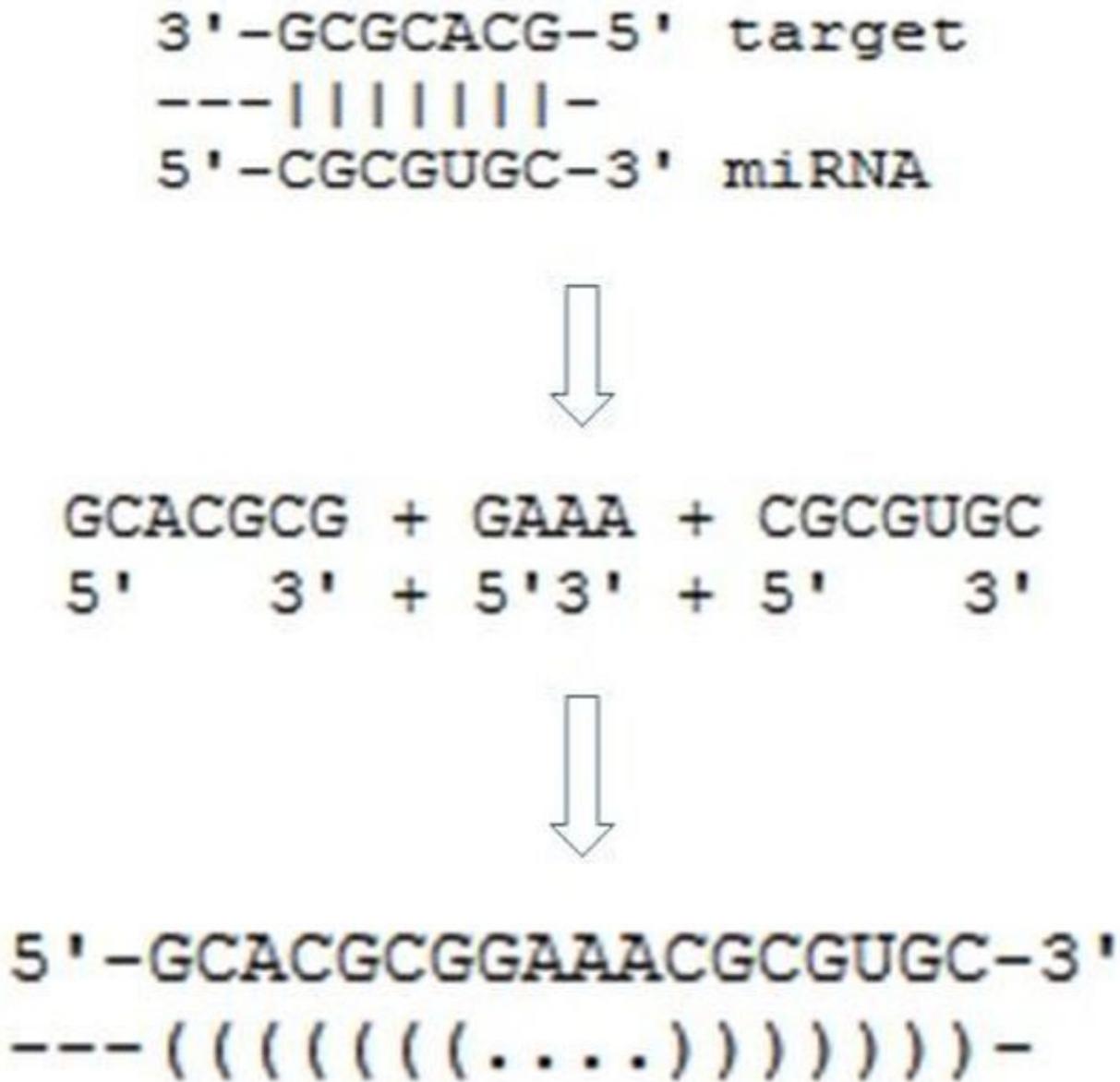
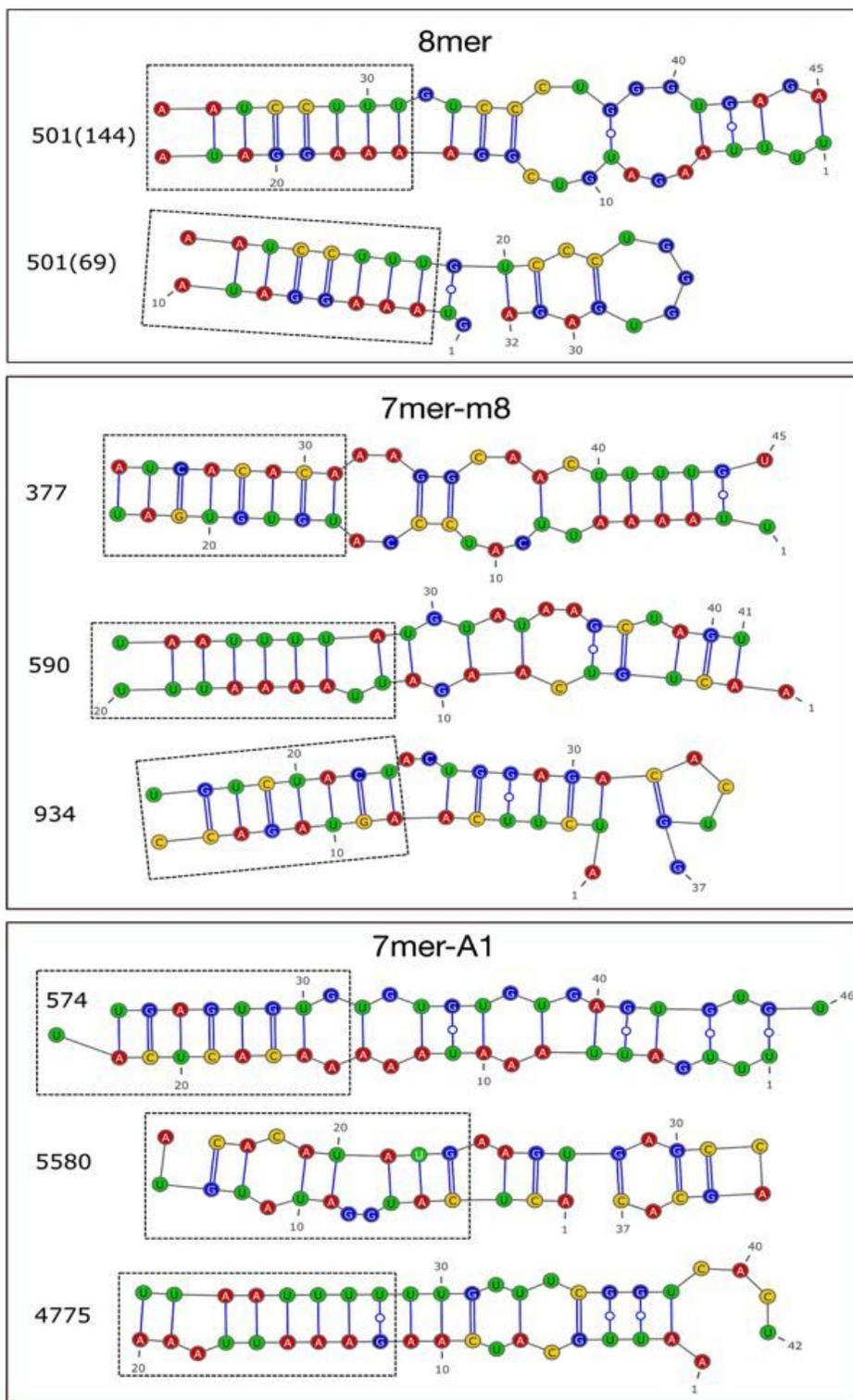


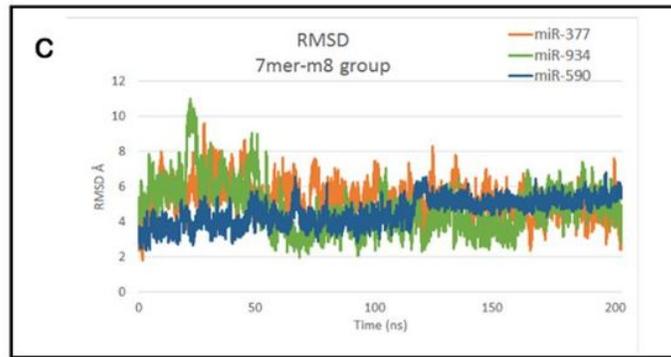
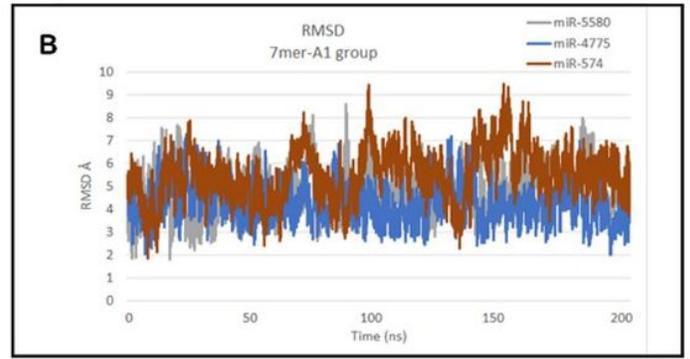
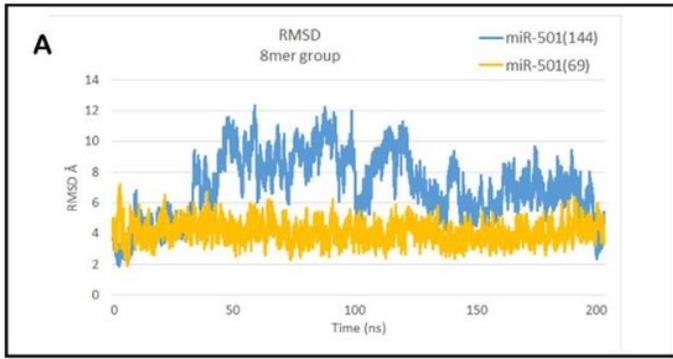
Figure 2

Schematic representation of introducing a GAAA tetraloop between miRNA and mRNA strands.



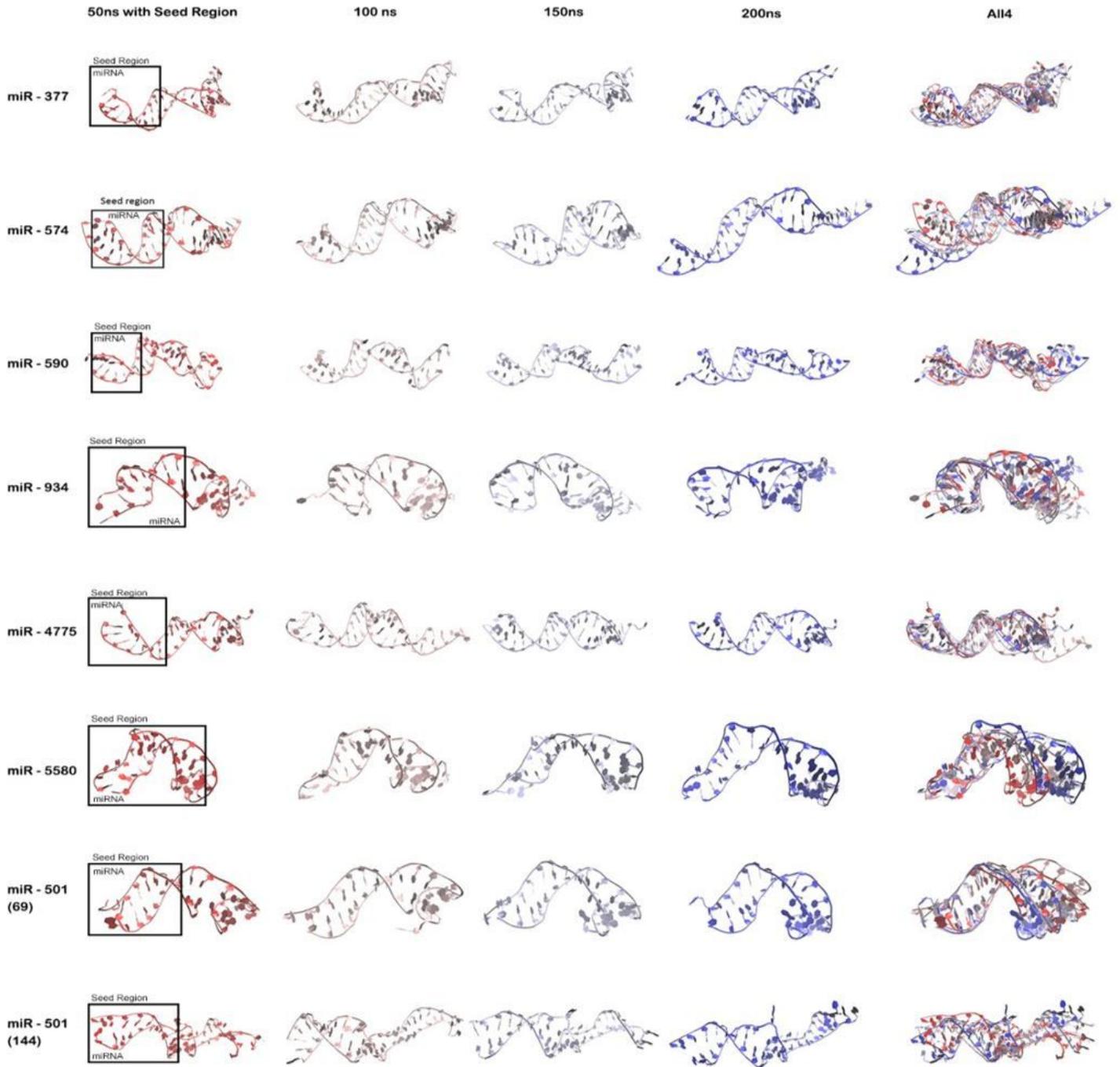
**Figure 3**

Secondary structure topology derived from 3D RNA models. Base-pairs identified by 3DNA-DSSR v2.0.0-2020aug01 [<https://x3dna.org>]. Images generated by VARNA v.3.93 [40][<http://varna.lri.fr>]. Seed regions are surrounded with dashed lines.



**Figure 4**

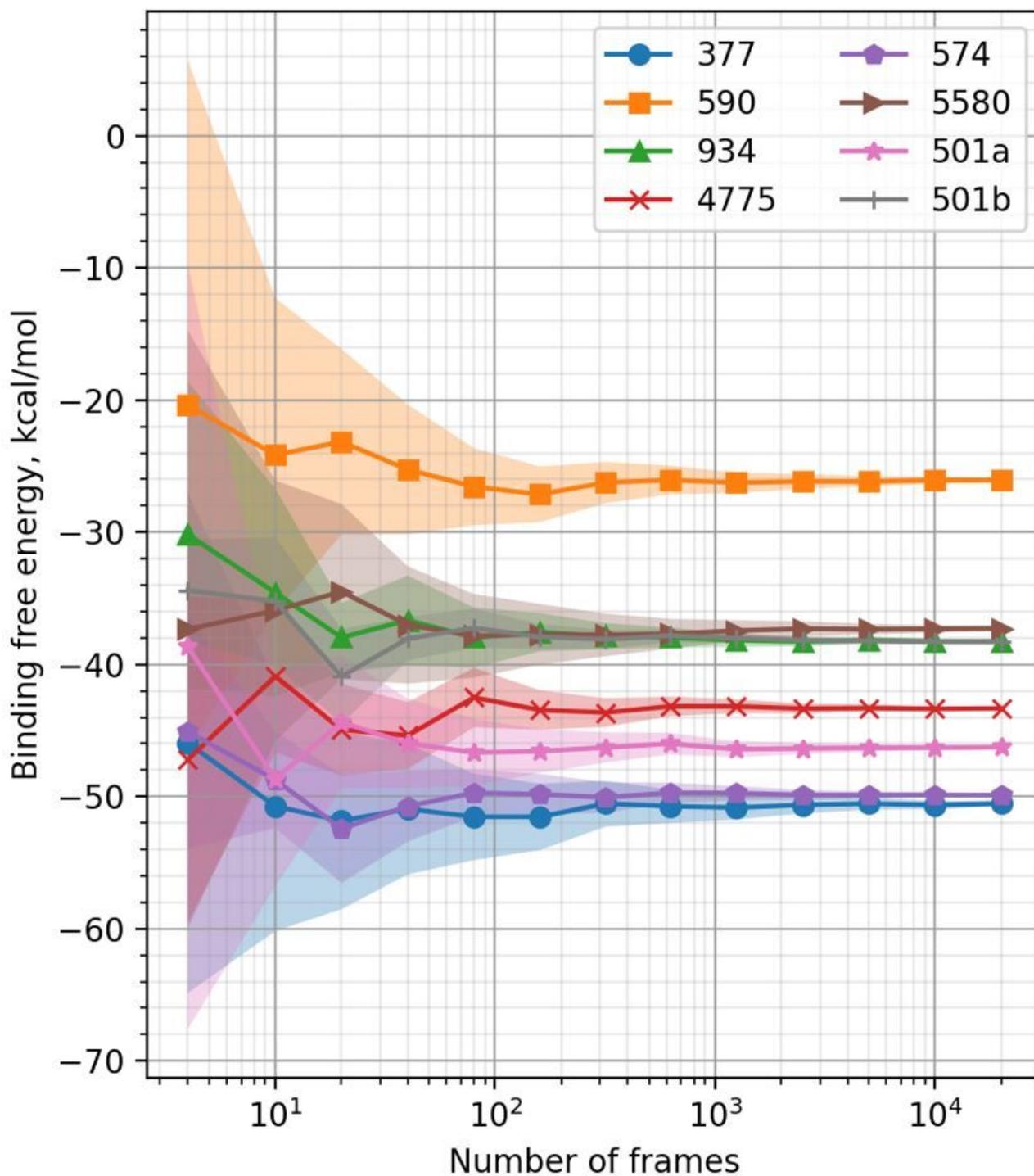
Root-mean-square deviation (RMSD) of miRNAs-mRNA complex backbone atoms with respect to the initial conformation as a function of time. RMSDs of miRNAs-mRNA complexes in three binding site groups (8mer, 7mer-A1, and 7mer-m8) are shown in panels A, B, and C respectively. Time starts from the beginning of production run after 10 ns long- equilibration.



**Figure 5**

Snapshots of 8 miRNA-mRNA complexes at simulation times of 50, 100, 150 and 200 ns. In the last column all 4 snapshots are superimposed.

## MM/GBSA



**Figure 6**

Results of calculation of  $\langle \Delta G_{\text{bind}} \rangle$  as a function of number of frames with MM/GBSA method. 501(144) and 501(69) are labeled as 501a and 501b, respectively. Shaded areas in the plots indicate the range of the 95% confidence intervals for the mean. (Plot using full quasi-harmonic entropy vibr+rot+trans).