Supporting Information to

Extensive Molecular Dynamics Simulations Showing That Canonical G8 and Protonated A38H$^+$ Forms Are Most Consistent with Crystal Structures of Hairpin Ribozyme

Vojtěch Mlýnský, Pavel Banáš, Daniel Hollas, Kamila Réblová, Nils G. Walter, Jiří Šponer and Michal Otyepka

1) Parameters of non-standard residues

Four non-standard residues (Fig. S1) were used in presented simulations: N1-deprotonated guanine (RGM), N1-protonated adenine (RAP), guanine-N1,O6-enol tautomer (RGT) and 2'-methoxyadenine (RMA).

The N9-methyl-N1-deprotonated guanine, N9-methyl-N1-protonated adenine and N9-methyl guanine-N1,O6-enol tautomer were fully geometrically optimized at HF/6-31G(d) level of theory using the Gaussian03 program[1] and subsequently used for parameterization of negatively charged deprotonated guanine, protonated adenine and guanine enol tautomer, respectively. The partial atomic charges of all non-standard bases were assigned by the restrained electrostatic potential (RESP) fit procedure at HF/6–31G(d) level of theory according to the scheme of Cornell *et al.*[2,3] The non-standard bases were merged with the corresponding standard parm99 nucleoside residues to obtain topologies for non-standard nucleosides. The N9-methyl atom was replaced by the C1' carbon atom and the residual fractional charge was added to the C1' carbon.

The 2',5'-bismethoxyadenosine-3'-methylmonophosphate model was used to obtain partial atomic charges of 2'-methoxyadenine. The conformation of the methyl-3'-deoxyadenosine-2'-methylmonophosphate modeling 2'-methoxyadenine was derived from the crystal structure (PDB ID 2OUE).[4] The methyl-3'-deoxyadenosine-2'-methylmonophosphate was geometrically optimized at HF/6-31G(d) level of theory using the Gaussian03 program. All sugar-phosphate backbone torsions were constrained during the optimization to retain the conformation of model close to the crystal structure geometry. Subsequently, the partial atomic charges were determined using RESP procedure at HF/6–31G(d) level of theory. The obtained partial RESP charges of C2' and C3' atoms and atoms of their substituents were pasted to parm99 adenine residue and the residual fractional charges were added to C2', C3' carbons.

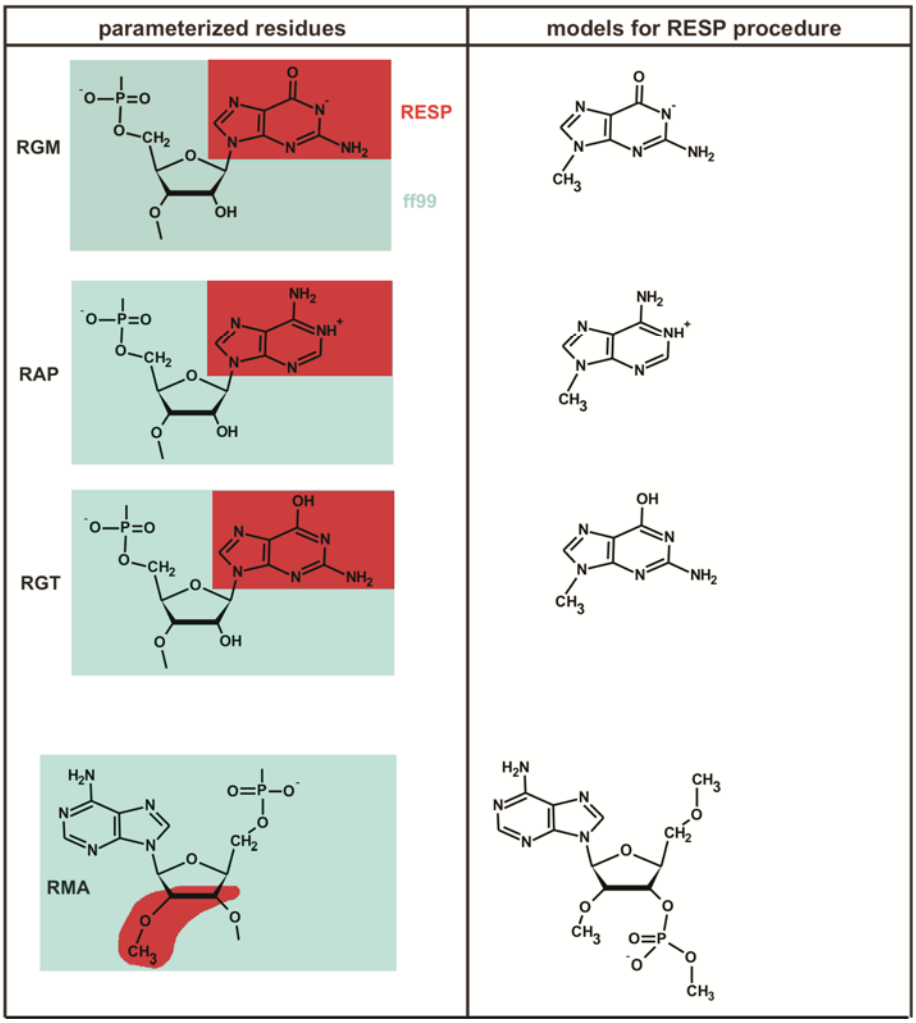

Figure S1: *Left panel shows structures of non-standard residues and right panel models, used for calculation of RESP charges. Whereas cyan insets on the left panel illustrate atoms with default parm99 charges, red insets show atoms, which partial charges were derived from RESP procedure.*

a) *AMBER prep file for 2'-methoxy form of adenine (MRA).*

```
    0    0    2

This is a remark line
molecule.res
MRA      XYZ  0
CHANGE      OMIT DU   BEG
  0.0000
   1   DUMM   DU    M          999.000      999.0      -999.0           .00000
   2   DUMM   DU    M          999.000     -999.0       999.0           .00000
   3   DUMM   DU    M         -999.000      999.0       999.0           .00000
   4   P      P     M            2.314000    1.777000   -0.478000     1.16620
   5   O1P    O2    E            1.933000    3.073000   -1.082000    -0.77600
   6   O2P    O2    E            3.157000    0.872000   -1.292000    -0.77600
   7   O5'    OS    M            2.994000    2.142000    0.923000    -0.49890
   8   C5'    CT    M            2.156000    2.552000    2.020000     0.05580
   9   H5'1   H1    E            1.453000    1.754000    2.258000     0.06790
  10   H5'2   H1    E            1.605000    3.450000    1.741000     0.06790
  11   C4'    CT    M            3.008000    2.848000    3.243000     0.10650
  12   H4'    H1    E            2.447000    3.477000    3.934000     0.11740
```

```
  13  O4'   OS    S          3.228000    1.605000    3.977000    -0.35480
  14  C1'   CT    B          4.559000    1.151000    3.778000     0.03940
  15  H1'   H2    E          5.072000    1.095000    4.737000     0.20070
  16  N9    N*    S          4.539000   -0.297000    3.316000    -0.02510
  17  C8    CK    B          4.337000   -0.846000    2.077000     0.20060
  18  H8    H5    E          4.152000   -0.207000    1.226000     0.15530
  19  N7    NB    S          4.453000   -2.141000    2.049000    -0.60730
  20  C5    CB    S          4.754000   -2.480000    3.363000     0.05150
  21  C6    CA    B          4.998000   -3.708000    3.989000     0.70090
  22  N6    N2    B          4.977000   -4.883000    3.346000    -0.90190
  23  H61   H     E          5.161000   -5.739000    3.851000     0.41150
  24  H62   H     E          4.777000   -4.914000    2.356000     0.41150
  25  N1    NC    S          5.265000   -3.689000    5.302000    -0.76150
  26  C2    CQ    B          5.285000   -2.520000    5.935000     0.58750
  27  H2    H5    E          5.498000   -2.482000    6.993000     0.04730
  28  N3    NC    S          5.073000   -1.309000    5.456000    -0.69970
  29  C4    CB    E          4.806000   -1.356000    4.133000     0.30530
  30  C3'   CT    M          4.416000    3.378000    2.961000     0.02368
  31  H3'   H1    E          4.412000    3.941000    2.028000     0.14322
  32  C2'   CT    B          5.230000    2.096000    2.779000     0.07929
  33  H2    H1    E          5.668000    2.081000    1.781000     0.12625
  34  O2'   OS    S          6.273000    2.050000    3.757000    -0.33043
  35  CM'   CT    3          6.759000    1.235000    3.611000     0.01050
  36  H2'1  H1    E          6.733287    0.539551    4.456873     0.06003
  37  H2'2  H1    E          7.790229    1.577426    3.701673     0.06003
  38  H2'3  H1    E          6.835622    0.719500    2.636731     0.06003
  39  O3'   OS    M          5.030000    4.129000    3.998000    -0.52460


LOOP
   C1'  C2'
   C4   N9
   C4   C5

IMPROPER
   C4   C8   N9  C1'
   H8   N9   C8  N7
   C6   C4   C5  N7
   C5   N6   C6  N1
   C6   H61  N6  H62
   H2   N1   C2  N3
   C5   N9   C4  N3

DONE
STOP
```

b) *AMBER prep file for deprotonated guanine residue (RGM)*

```
    0    0    2

This is a remark line
molecule.res
RGM   XYZ  0
CHANGE     OMIT DU   BEG
  0.0000
   1   DUMM  DU    M    0  -1  -2    0.00       0.00       0.00        0.0000
   2   DUMM  DU    M    1   0  -1    1.00       0.00       0.00        0.0000
   3   DUMM  DU    M    2   1   0    1.00      90.00       0.00        0.0000
   4   P     P     M    3   2   1    1.60     119.04     200.00        1.1662
   5   O1P   O2    E    4   3   2    1.48     109.61     150.00       -0.7760
   6   O2P   O2    E    4   3   2    1.48     109.58      20.00       -0.7760
   7   O5'   OS    M    4   3   2    1.60     101.43     -98.89       -0.4989
   8   C5'   CT    M    7   4   3    1.44     119.00     -39.22        0.0558
   9   H5'1  H1    E    8   7   4    1.09     109.50      60.00        0.0679
  10   H5'2  H1    E    8   7   4    1.09     109.50     -60.00        0.0679
  11   C4'   CT    M    8   7   4    1.52     110.00     180.00        0.1065
  12   H4'   H1    E   11   8   7    1.09     109.50    -200.00        0.1174
  13   O4'   OS    S   11   8   7    1.46     108.86     -86.31       -0.3548
  14   C1'   CT    B   13  11   8    1.42     110.04     105.60       -0.0227
```

```
  15   H1'   H2    E   14  13  11    1.09    109.50   -240.00       0.2006
  16   N9    N*    S   14  13  11    1.49    108.06   -127.70       0.0005
  17   C8    CK    B   16  14  13    1.38    129.20     81.59       0.0125
  18   H8    H5    E   17  16  14    1.08    120.00      0.00       0.1339
  19   N7    NB    S   17  16  14    1.31    114.00   -179.90      -0.5521
  20   C5    CB    S   19  17  16    1.39    103.90      0.00      -0.0301
  21   C6    C     B   20  19  17    1.42    130.40    180.00       0.7707
  22   O6    O     E   21  20  19    1.23    128.80      0.00      -0.7034
  23   N1    NC    S   21  20  19    1.40    111.38    180.00      -0.8744
  24   C2    CA    B   23  21  20    1.38    125.24     -0.10       0.9143
  25   N2    N2    B   24  23  21    1.34    116.02    180.00      -0.9610
  26   H21   H     E   25  24  23    1.01    127.00     -0.82       0.3573
  27   H22   H     E   25  24  23    1.01    116.53   -179.44       0.3573
  28   N3    NC    S   24  23  21    1.33    123.30      0.00      -0.8208
  29   C4    CB    E   28  24  23    1.36    112.20      0.00       0.3334
  30   C3'   CT    M   11   8   7    1.53    115.78   -329.11       0.2022
  31   H3'   H1    E   30  11   8    1.09    109.50     30.00       0.0615
  32   C2'   CT    B   30  11   8    1.53    102.80    -86.30       0.0670
  33   H2'1  H1    E   32  30  11    1.09    109.50    120.00       0.0972
  34   O2'   OH    S   32  30  11    1.43    109.50    240.00      -0.6139
  35   HO'2  HO    E   34  32  30    0.96    107.00    180.00       0.4186
  36   O3'   OS    M   30  11   8    1.42    116.52   -203.47      -0.5246

IMPROPER
 C8   C4   N9   C1'
 C5   N1   C6   O6
 C2   H21  N2   H22
 N7   N9   C8   H8
 N1   N3   C2   N2

LOOP CLOSING EXPLICIT
 C1'  C2'
 C4   C5
 C4   N9

DONE
STOP
```

c) *AMBER prep file for O6H G8 tautomer (RGT)*

```
    0    0    2

R-GUANINE - with 5' - phosphate group and 3' - O(minus) group

RGT  INT     1
CORRECT OMIT DU   BEG
   0.0
   1   DUMM  DU    M    0  -1  -2    0.00      0.00      0.00       0.0000
   2   DUMM  DU    M    1   0  -1    1.00      0.00      0.00       0.0000
   3   DUMM  DU    M    2   1   0    1.00     90.00      0.00       0.0000
   4   P     P     M    3   2   1    1.60    119.04    200.00       1.1662
   5   O1P   O2    E    4   3   2    1.48    109.61    150.00      -0.7760
   6   O2P   O2    E    4   3   2    1.48    109.58     20.00      -0.7760
   7   O5'   OS    M    4   3   2    1.60    101.43    -98.89      -0.4989
   8   C5'   CT    M    7   4   3    1.44    119.00    -39.22       0.0558
   9   H5'1  H1    E    8   7   4    1.09    109.50     60.00       0.0679
  10   H5'2  H1    E    8   7   4    1.09    109.50    -60.00       0.0679
  11   C4'   CT    M    8   7   4    1.52    110.00    180.00       0.1065
  12   H4'   H1    E   11   8   7    1.09    109.50   -200.00       0.1174
  13   O4'   OS    S   11   8   7    1.46    108.86    -86.31      -0.3548
  14   C1'   CT    B   13  11   8    1.42    110.04    105.60       0.0289
  15   H1'   H2    E   14  13  11    1.09    109.50   -240.00       0.2006
  16   N9    N*    S   14  13  11    1.49    108.06   -127.70       0.0020
  17   C8    CK    B   16  14  13    1.38    127.98     81.59       0.0554
  18   H8    H5    E   17  16  14    1.08    120.47      0.00       0.1786
  19   N7    NB    S   17  16  14    1.28    114.44   -179.90      -0.5110
  20   C5    CB    S   19  17  16    1.39    103.90      0.00       0.0332
  21   C6    C     B   20  19  17    1.39    134.97    180.00       0.6339
  22   O6    OH    S   21  20  19    1.32    121.30      0.00      -0.5751
```

```
  23   H1     HO     E    22   21   20     0.95     108.10     180.00        0.4448
  24   N1     NC     S    21   20   19     1.31     120.51     180.00       -0.7821
  25   C2     CA     B    24   21   20     1.34     118.67      -0.10        0.8950
  26   N2     N2     B    25   24   21     1.36     115.62     180.00       -0.9120
  27   H21    H      E    26   25   24     1.00     116.70      -0.82        0.3897
  28   H22    H      E    26   25   24     1.00     116.70    -179.44        0.3897
  29   N3     NC     S    25   23   21     1.32     127.02       0.00       -0.7190
  30   C4     CB     E    29   25   24     1.33     112.30       0.00        0.3634
  31   C3'    CT     M    11    8    7     1.53     115.78    -329.11        0.2022
  32   H3'    H1     E    31   11    8     1.09     109.50      30.00        0.0615
  33   C2'    CT     B    31   11    8     1.53     102.80     -86.30        0.0670
  34   H2'1   H1     E    33   31   11     1.09     109.50     120.00        0.0972
  35   O2'    OH     S    33   31   11     1.43     109.50     240.00       -0.6139
  36   HO'2   HO     E    35   33   31     0.96     107.00     180.00        0.4186
  37   O3'    OS     M    31   11    8     1.42     116.52    -203.47       -0.5246

IMPROPER
 C8    C4    N9    C1'
 C5    N1    C6    O6
 C6    C2    N1    H1
 C2    H21   N2    H22
 N7    N9    C8    H8
 N1    N3    C2    N2

LOOP CLOSING EXPLICIT
 C1'   C2'
 C4    C5
 C4    N9

DONE
STOP
```

d) *AMBER prep file for protonated adenine (RAP)*

```
    0    0    2

This is a remark line
molecule.res
RAP   XYZ  0
CHANGE     OMIT DU   BEG
  0.0000
   1   DUMM   DU     M     0   -1   -2     0.00       0.00       0.00        0.0000
   2   DUMM   DU     M     1    0   -1     1.00       0.00       0.00        0.0000
   3   DUMM   DU     M     2    1    0     1.00      90.00       0.00        0.0000
   4   P      P      M     3    2    1     1.60     119.04     200.00        1.1662
   5   O1P    O2     E     4    3    2     1.48     109.61     150.00       -0.7760
   6   O2P    O2     E     4    3    2     1.48     109.58      20.00       -0.7760
   7   O5'    OS     M     4    3    2     1.60     101.43     -98.89       -0.4989
   8   C5'    CT     M     7    4    3     1.44     119.00     -39.22        0.0558
   9   H5'1   H1     E     8    7    4     1.09     109.50      60.00        0.0679
  10   H5'2   H1     E     8    7    4     1.09     109.50     -60.00        0.0679
  11   C4'    CT     M     8    7    4     1.52     110.00     180.00        0.1065
  12   H4'    H1     E    11    8    7     1.09     109.50    -200.00        0.1174
  13   O4'    OS     S    11    8    7     1.46     108.86     -86.31       -0.3548
  14   C1'    CT     B    13   11    8     1.42     110.04     105.60        0.1297
  15   H1'    H2     E    14   13   11     1.09     109.50    -240.00        0.2007
  16   N9     N*     S    14   13   11     1.52     109.59    -127.70       -0.0532
  17   C8     CK     B    16   14   13     1.37     131.20      81.59        0.1412
  18   H8     H5     E    17   16   14     1.08     120.00       0.00        0.2008
  19   N7     NB     S    17   16   14     1.30     113.93     177.00       -0.5330
  20   C5     CB     S    19   17   16     1.39     104.00       0.00        0.2613
  21   C6     CA     B    20   19   17     1.40     132.42     180.00        0.2748
  22   N6     N2     B    21   20   19     1.34     123.50       0.00       -0.7251
  23   H61    H      E    22   21   20     1.01     120.00     180.00        0.4282
  24   H62    H      E    22   21   20     1.01     120.00       0.00        0.4282
  25   N1     NA     B    21   20   19     1.34     117.43     180.00       -0.2604
  26   H1     H      E    25   21   20     1.00     119.46     180.00        0.3566
  27   C2     CQ     B    25   21   20     1.33     118.80       0.00        0.1964
  28   H2     H5     E    27   25   21     1.08     120.00     180.00        0.1889
```

```
  29   N3     NC     S    27   25   21     1.32     129.17       0.00       -0.4431
  30   C4     CB     E    29   27   25     1.35     110.80       0.00        0.3240
  31   C3'    CT     M    11    8    7     1.53     115.78    -329.11        0.2022
  32   H3'    H1     E    31   11    8     1.09     109.50      30.00        0.0615
  33   C2'    CT     B    31   11    8     1.53     102.80     -86.30        0.0670
  34   H2'1   H1     E    33   31   11     1.09     109.50     120.00        0.0972
  35   O2'    OH     S    33   31   11     1.43     109.50     240.00       -0.6139
  36   HO'2   HO     E    35   33   31     0.96     107.00     180.00        0.4186
  37   O3'    OS     M    31   11    8     1.42     116.52    -203.47       -0.5246

IMPROPER
 C8    C4    N9    C1'
 C6    H61   N6    H62
 N7    N9    C8    H8
 N1    N3    C2    H2
 C5    N1    C6    N6
 C6    C2    N1    H1

LOOP CLOSING EXPLICIT
 C1'   C2'
 C4    C5
 C4    N9

DONE
STOP
```

e) *AMBER parm file non-standard parameters*

```
# force field modification for RMA, RGM, RGT, RAP, IN1, IN2, and L25 residues
MASS
OX     16.00         0.465              based on OS ether and ester oxygen
HX     1.008         0.135              based on HO hydroxyl group

BOND
CT-OX   320.0    1.410       based on CT-OS JCC,7,(1986),230; NUCLEIC ACIDS
HX-OH   553.0    0.960       based on HO-OH JCC,7,(1986),230; SUGARS,SER,TYR
CQ-NA   502.0    1.324       based on CQ-NC JCC,7,(1986),230; ADE

ANGLE
H1-CT-OX       50.0      109.50    based on H1-CT-OS changed based on NMA nmodes
CT-CT-OX       50.0      109.50    based on CT-CT-OS
CB-CA-NA       70.0      117.30    based on CB-CA-NC
NC-CQ-NA       70.0      129.10    based on NC-CQ-NC
H5-CQ-NA       50.0      115.45    based on H5-CQ-NC
CA-NA-CQ       70.0      118.60    based on CA-NC-CQ
CQ-NA-H        50.0      118.00    based on CA-NA-H changed based on NMA nmodes

DIHEDRAL
OX-CT-CT-OS    1    0.144        0.0        -3.    based on OS-CT-CT-OS parm98, TC,PC,PAK
OX-CT-CT-OS    1    1.175        0.0         2.    based on OS-CT-CT-OS Piotr et al.
H1-CT-CT-OX    1    0.25         0.0         1.    based on H1-CT-CT-OS Junmei et al, 1999
OX-CT-CT-OH    1    0.144        0.0        -3.    based on OS-CT-CT-OH parm98, TC,PC,PAK
OX-CT-CT-OH    1    1.175        0.0         2.    based on OS-CT-CT-OH parm98, TC,PC,PAK
X -NA-CQ- X    4    9.60        80.0         2.


NONBON
 OX          1.6837   0.1700            based on OS OPLS ether
 HX          0.6000   0.0157            based on HS W. Cornell CH3SH --> CH3OH FEP
```

2) Behavior of S-turn backbone

The torsion angle analysis of bases making the S-turn region (residues G36, U37, A38, C39, A40) was carried out for the detailed characterization of the S-turn behavior in four chosen simulations (Table S1). Generally, the simulations with canonical A38 showed significantly more dynamical behavior compared to systems with protonated $A38H^+$.

We observed α/γ flips of phosphates of all S-turn nucleobases except C39. Most of these α/γ flips appeared for U37 and $A38/A38H^+$. The most frequent flips were observed in

WT1/G8/A38 simulation, while only short-term flips occurred for U37, A38 bases in OMe/G8/A38 simulation. In contrast, three longer-time flips occurred for U37 nucleobase in OMe/G8/A38H^+ simulation.

The dihedral β was stable and unchanged at ~180° in nearly all described simulations (Table S1), except of U37 nucleobase in WT1/G8/A38 and WT1/G8/A38H^+ systems, where it oscillated between two substates with dihedral β values ~180° and ~80°. Notably, we observed also a 5 ns-long flip of dihedral β (from 180° to 80° and back) for G36 in OMe/G8/A38 simulation.

The δ dihedral was rather unchanged in both simulations with protonated A38H^+, however, residues adjacent to A38 within S-turn changed often their typical δ torsions 150° towards 80° in simulations with canonical A38 (Table S1). Similarly, the sugar puckering of S-turn nucleobases was preferentially C2'-endo in agreement with the initial crystal structures in case of systems with protonated A38H^+, while sugar pucker of U37 changed from C2'-endo to C3'-endo at ~4 ns in WT1/G8/A38, ~20 ns in OMe/G8/A38 simulation.

The ε and ζ dihedrals were stable (fluctuating around ~90°) for A38/A38H^+, and A40 nucleobases. On contrary, ε and ζ of G36, U37 and C39 showed a coupled dynamic behavior with a lot of flips, and were fluctuated typically around values ~90°, ~180° and ~270°. The most ε/ζ flips were observed in both simulation with canonical A38 (WT1/G8/A38 and OMe/G8/A38), while the G36 in WT1/G8/A38H^+ simulations was the only nucleobase carrying ε/ζ flips.

3) Supplementary tables

Table S1: *Summary of changes in dihedral angles of bases (G36, U37, A38, C39 and A40) within the S-turn region. Nucleobases that underwent dynamical flips in torsion angles are listed (see text for detailed description).*

Simulation	dihedrals			
	α – γ	β	δ	ε + ζ
OMe/G8/A38	G36, U37, A38, A40	G36	G36, U37, C39	G36, U37, C39
OMe/G8/A38H$^+$	U37, A38H$^+$			
WT1/G8/A38	G36, U37, A38, A40	U37	U37, C39	G36, U37, C39
WT1/G8/A38H$^+$	U37, A38H$^+$	U37		G36

Table S2: *Table summarizes population of active-site conformation with favorable in-line attack angle (IAA, >130°) and population of A-1(2'-OH) or A-1(2'-OMe) group in the X-ray-like conformation of A-1(2'-OMe) (see Fig. 3A). Populations are averaged over the whole simulation times.*

Simulation	IAA (%)	2'-OMe- like (%)	Simulation	IAA (%)	2'-OMe- like (%)
WT/G8$^-$/A38	2	0	WT2/G8/A38	3	9
WT/G8$^-$/A38H$^+$	3	0	WT1/G8/A38H$^+$	39	2
WT/G8t/A38	7	9	WT2/G8/A38H$^+$	65	0
WT/G8t/A38H$^+$	6	1	WT/G8/A38/ES	2	3
OMe/G8/A38	7	100	WT/G8/A38H$^+$/ES	4	0
OMe/G8/A38H$^+$	10	100	WT/G8/A38/bsc0	68	81
WT1/G8/A38	49	30	WT/G8/A38H$^+$/bsc0	35	0

Table S3: *Populations of the effective reaction conformation (ERC, Fig. S4) averaged over the whole simulation times.*

Simulation	ERC (%)	Simulation	ERC (%)
WT1/G8/A38	1	WT/G8/A38/ES	0
WT2/G8/A38	1	WT/G8/A38H$^+$/ES	4
WT1/G8/A38H$^+$	18	WT/G8/A38/bsc0	0
WT2/G8/A38H$^+$	65	WT/G8/A38H$^+$/bsc0	35

Table S4: *Times when the 'ladder-like' transition in H4 A-RNA duplex appeared.*

Simulation	Time (ns)	Simulation	Time (ns)
WT/G8$^-$/A38	-	WT1/2/G8/A38	-
WT/G8$^-$/A38H$^+$	-	WT1/2/G8/A38H$^+$	-
WT/G8t/A38	-	WT/G8/A38/ES	-
WT/G8t/A38H$^+$	7	WT/G8/A38H$^+$/ES	-
OMe/G8/A38	10	WT/G8/A38/bsc0	29
OMe/G8/A38H$^+$	27	WT/G8/A38H$^+$/bsc0	-

4) Supplementary figures

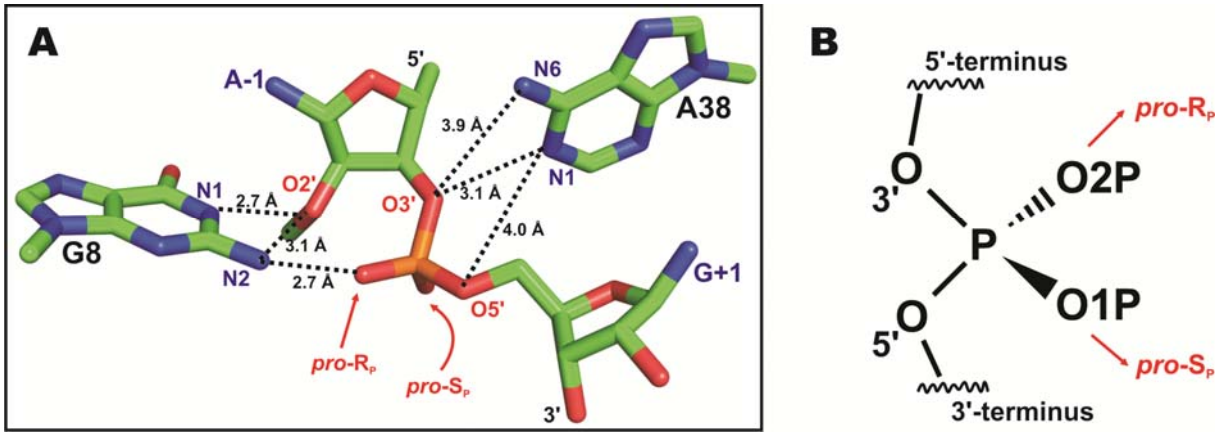

Figure S2: *(A) The active site X-ray geometry (PDB ID 2OUE) was used as a starting geometry for MD simulations. Key residues and atoms participating in H-bonds discussed in the text are labeled and lengths of discussed distances are measured. (B) Definition of pro-R_P/pro-S_P oxygens of the scissile phosphate according to IUPAC terminology.*

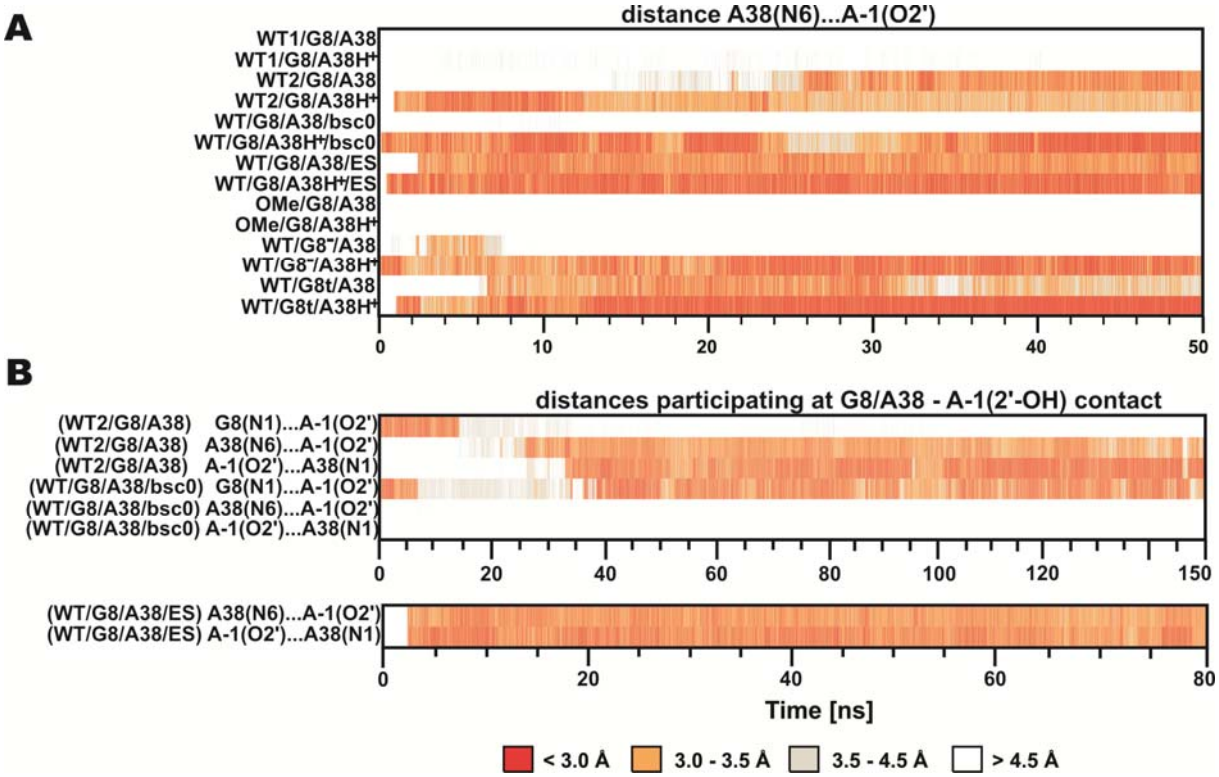

Figure S3: *Interatomic distances within the active site. (A) Time evolution of A38(N6)…A-1(O2') distance in all performed MD simulation. (B) Evolution of specific contacts between A38/G8 and A-1(2'-OH) are shown for the entire simulation times of WT2/G8/A38, WT/G8/A38/bsc0 and WT/G8/A38/ES.*

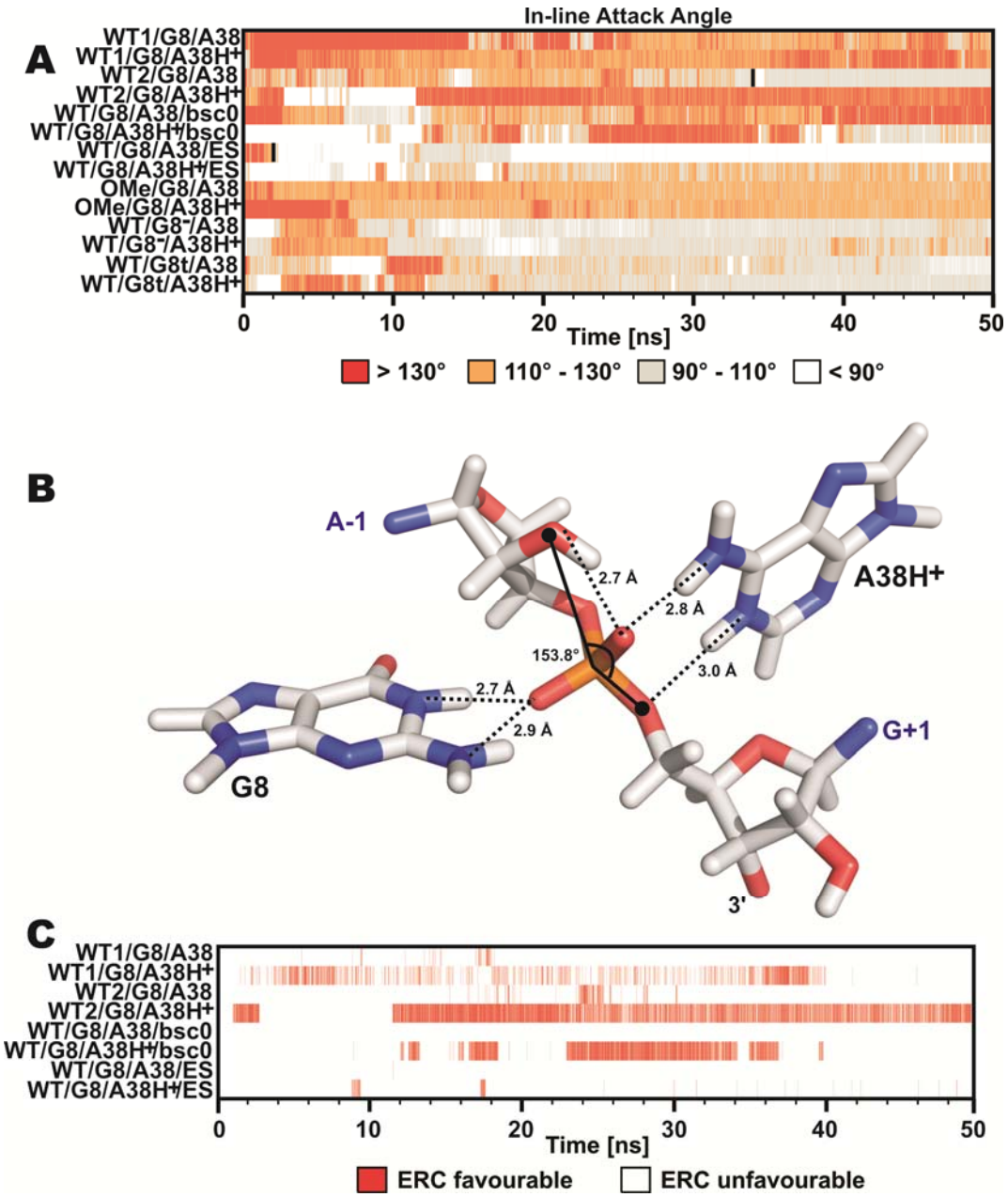

Figure S4: *(A) Evolution of the in-line attack angle (IAA, A-1(2'O)-G+1(P)-G+1(5'O)). IAA is color-coded as indicated in legend. The IAA higher than 130° is considered as a favorable IAA for the nucleophilic attack of A-1(2'-OH) to the scissile phosphate. The two black dashes inside the plot for WT2/G8/A38 and WT/G8/A38/ES simulations indicate times when A-1(2'-OH)…A38(N1) H-bond was formed. (B) The geometry fulfilling a favorable IAA (>130°) and having contacts between the scissile phosphate and both G8 and A38 (G8(N1)…G+1(pro-R_P/pro-S_P) < 4.0Å, A38(N1)…G+1(O5') < 4.0Å) seems to be the effective reactive conformation (ERC). The snapshot was taken from WT/G8/A38H$^+$/bsc0 simulation. (C) Evolution of ERC, the ERC is indicated by red dashes.*

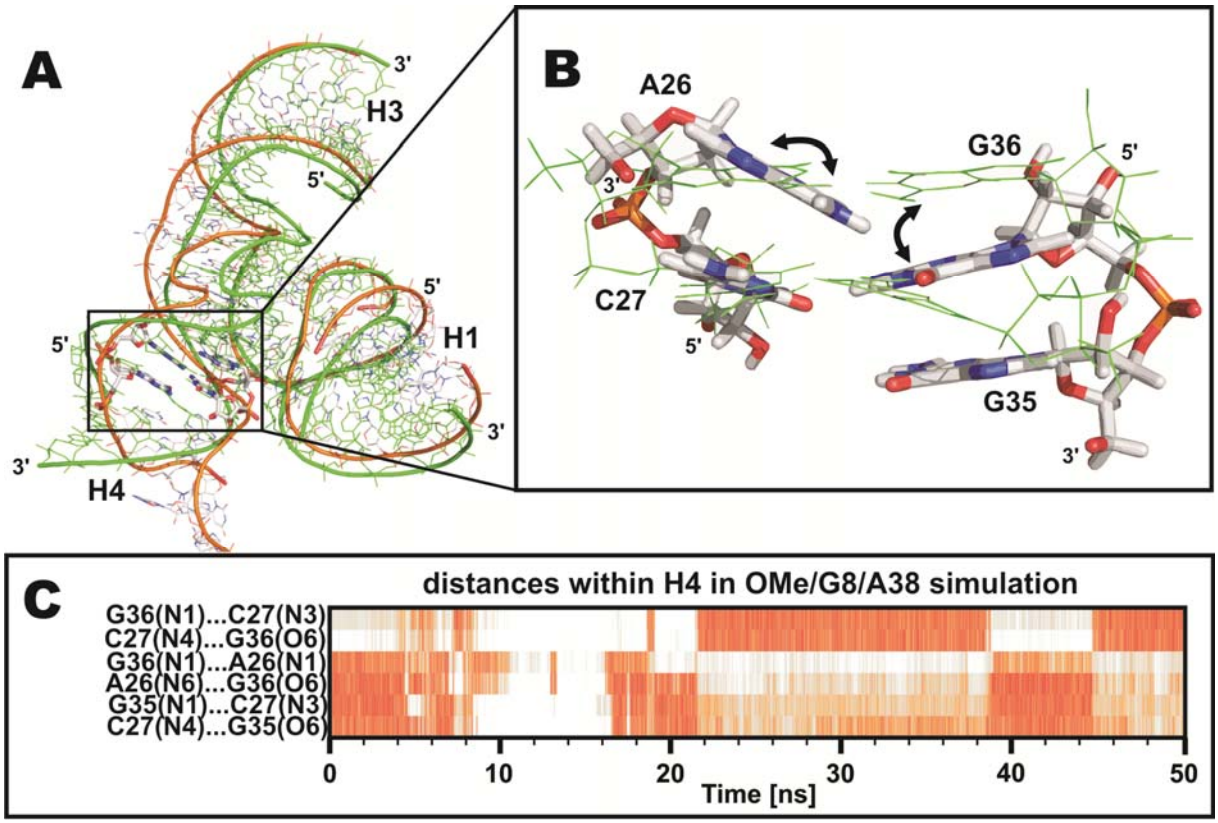

***Figure S5**: (A) Ribbon diagrams show the hairpin ribozyme with labeled helixes H1-H4. (B) The last ns average structure taken from OMe/G8/A38 MD simulations (in sticks) is superimposed with the X-ray geometry (green lines). The G36 nucleobase of H4 interrupts base pair with A26 (black arrows) and creates a new cis Watson-Crick base pair with C27, thus displacing G35 out of pairings. (C) The cis WC-base pair formation between G36-C27 competes with interactions among G35-C27, G36-A26 nucleobases.*

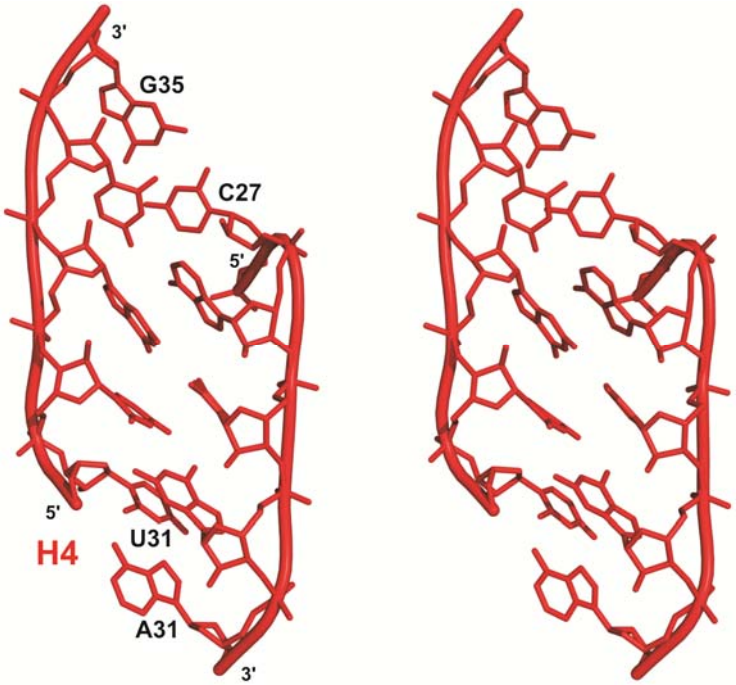

***Figure S6**: Stereo view of the 'ladder-like' structure of H4 A-RNA duplex taken from OMe/G8/A38H$^+$ simulation.*

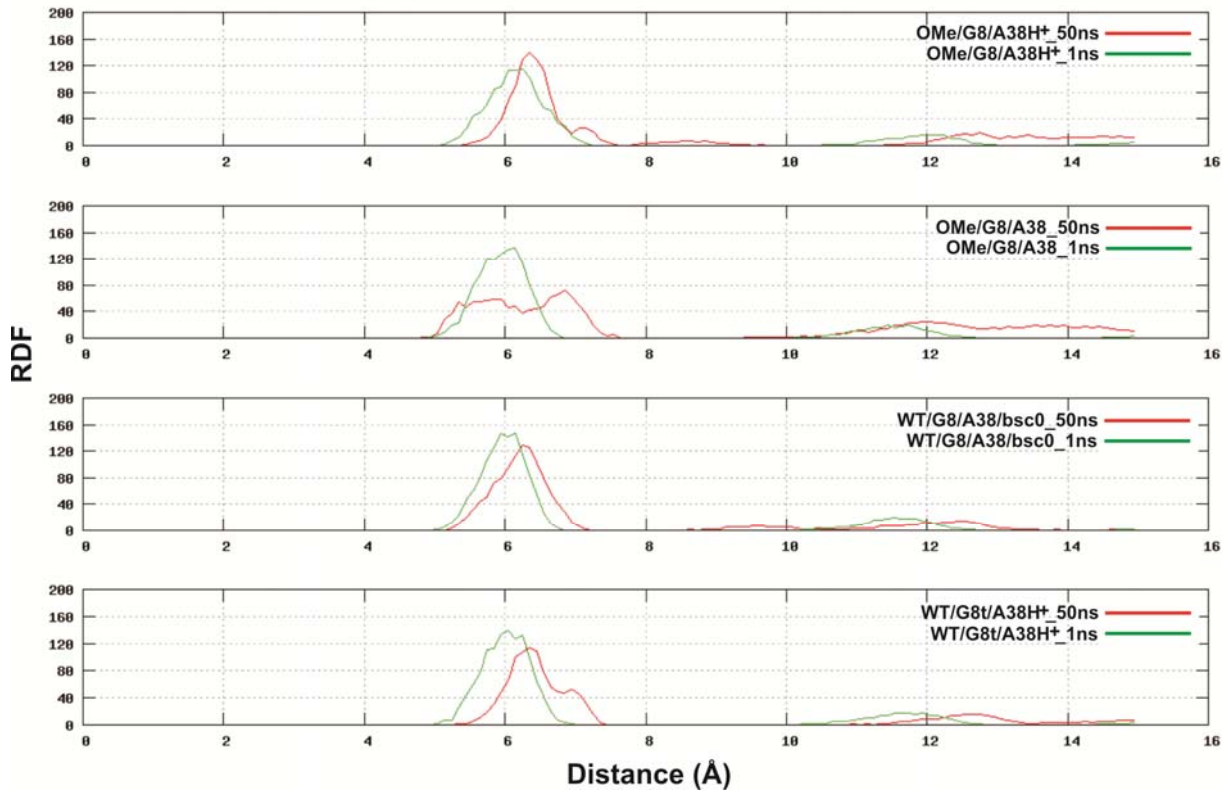

Figure S7*: Peaks in P-P radial distribution function (RDF) of 5 basepairs (C27-G35, A28-U34, C29-G33, G30-C32, A31-U31) of H4 shift towards higher values in the 'ladder-like' structures in OMe/G8/A38H$^+$, OMe/G8/A38, WT/G8/A38/bsc0 and WT/G8t/A38H$^+$ MD simulations. The green line corresponds to the A-RNA conformation and the red one to the 'ladder-like' structure.*

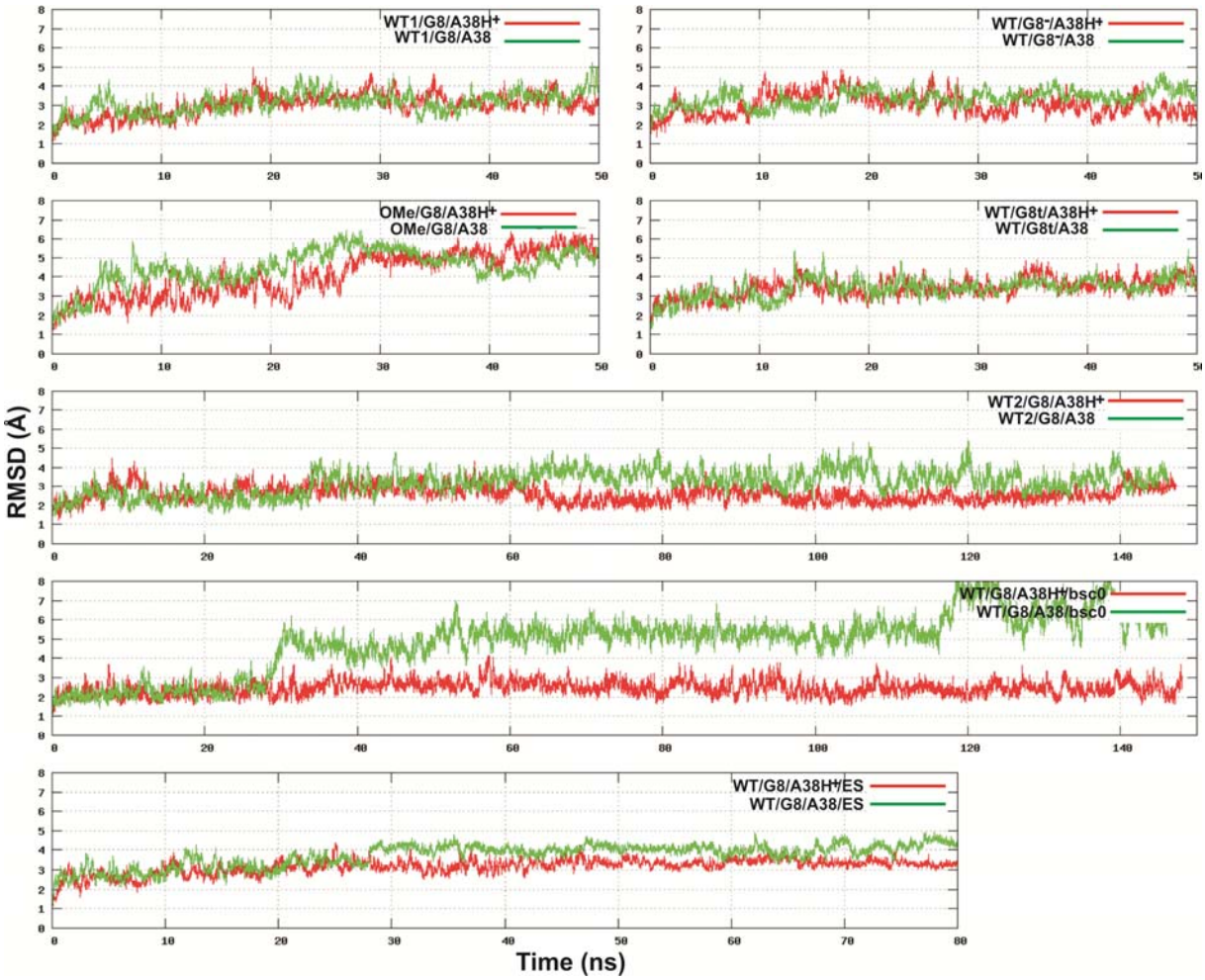

Figure S8*: Root mean square deviation (RMSD) of all heavy atoms from the X-ray structure as a function of time for present MD simulations. The green line represents system with canonical A38 and the red line system with protonated A38H*$^+$ *form.*

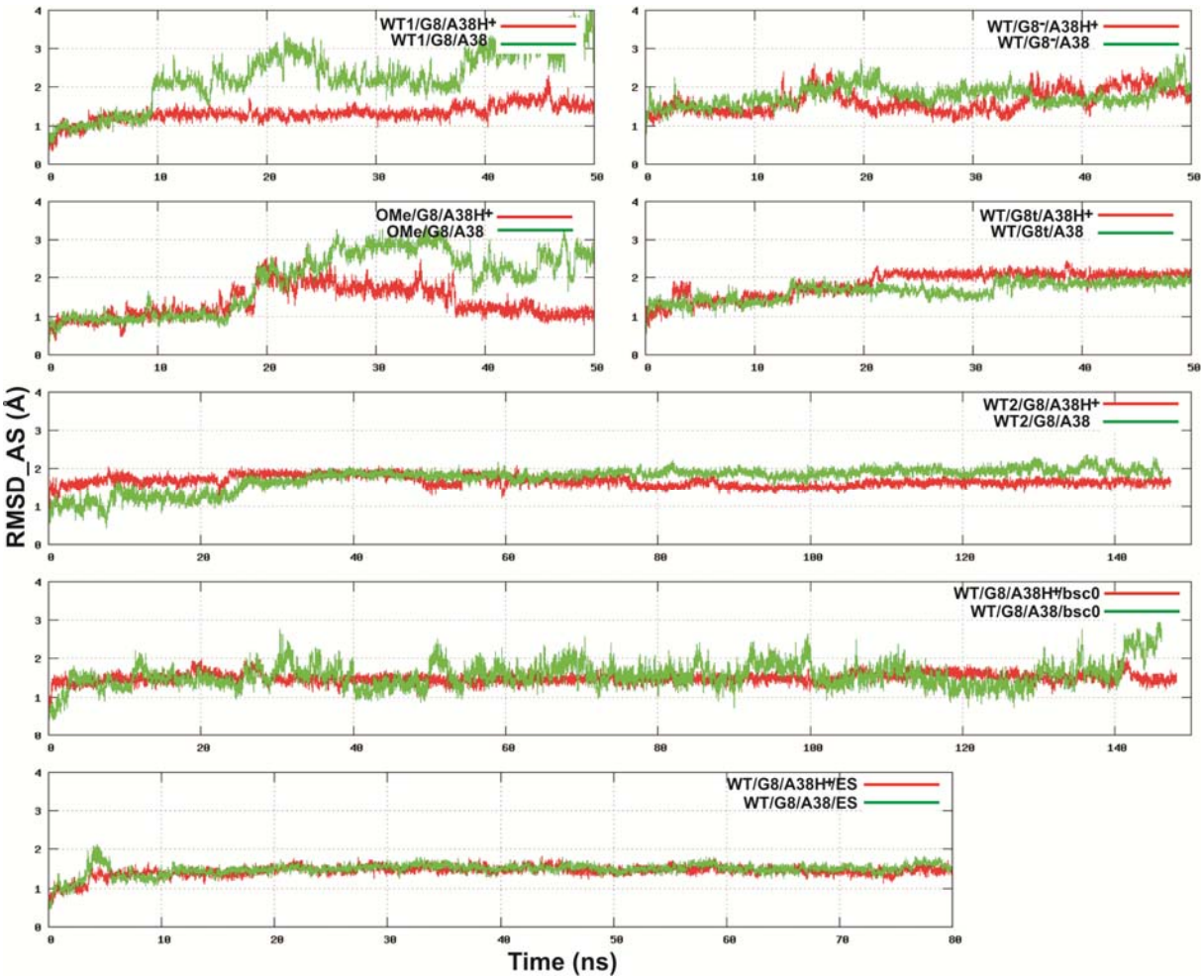

Figure S9. *Root mean square deviation (RMSD) of active site (including A-1, G+1, G8 and A38 nucleobases) heavy atoms from the X-ray structure as a function of time for present MD simulations. The green line represents system with canonical A38 and the red line system with protonated A38H$^+$ form. Higher RMSD values between 18 and 38 ns in OMe/G8/A38H$^+$ simulation (in the system showing best agreement with the X-ray geometry) were caused by temporary expiration of A38H$^+$(N1H)…G+1(O5') H-bond.*

5) Constant pH Molecular Dynamics

a) Parameterization

We considered two titratable groups in the constant-pH MD simulations: guanine G8 and adenine A38. The parm99 van der Waals parameters of canonical forms of guanosine and adenosine were used for deprotonated state of guanosine and protonated state of adenosine. The charges of canonical forms of both residues were taken from parm99 force field, while the RESP charges were calculated for deprotonated guanosine and protonated adenosine, as described in section Parameters of non-standard residues (see page S1).

The relative energies between each protonation state of given nucleotide calibrating the constant-pH MD method were calculated using thermodynamic integration. The modified generalized Born (GB) model developed by Onufriev, Bashford and Case (*igb*=2 in AMBER10 setting) was used in the thermodynamic integration as well as in constant-pH MD simulations. The salt concentration using a modified GB theory based on Debye-Hückel limiting law for ion screening of interactions was set to 0.33 mol·L^{-1}. The twelve-point

Gaussian quadrature was used in thermodynamic integration for evaluation of the free energy difference between protonation forms of titrating residues. In addition to the relative free energies of both protonation forms obtained by thermodynamic integration, the energy of protonated state was increased by pK_aRTln10 correction to account pK_a value in the calibration of constant-pH MD. The relative energy corrections of canonical and deprotonated guanine were 9.2*RTln10 and 99.49 kcal/mol, respectively. Corrections for protonated and canonical adenine were 3.6*RTln10 and 79.94 kcal/mol. Finally, the reference constant-pH MD simulations of model compounds N9-methylguanine/N1-deprotonated-N9-methylguanine and N1-protonated-N9-methyladenine/N9-methyladenine at pH equal to pK_a values of guanosine (pK_a 9.2) and adenosine (pK_a 3.6) were performed to check the set of parameters. The Monte Carlo step was performed every 100 fs in constant-pH MD simulations. We observed a half to half population of protonation states of given nucleobase in the reference constant-pH simulations within precision of 0.1%, as expected for pH equal to pK_a.

b) Reference generalized Born MD simulation

We performed a reference MD simulation utilizing the generalized Born (GB) model to check the stability of RNA structure in the implicit solvent model. Again, the modified GB model (developed by Onufriev, Bashford and Case (*igb*=2 in AMBER10) with the salt concentration using of 0.33 mol/l) was used as a setting of GB implicit model.

The last snapshot of the WT2/G8/A38H$^+$ simulation representing well equilibrated system was used as a starting structure. All ions and water molecules were removed and GB implicit solvation model was turned on. Subsequently, the system was minimized using similar minimization protocol as described in the Methods section of the main text. The periodic boundary condition was turn off, cutoff was increased to 30 Å (as recommended in AMBER10 manual) and the GB model with concentration of salt screening was set as described above. The 1-fs -long integration time integration step was used in thermalization and production phases. Finally, the system was slowly heated to 300 K during 100-ps long thermalization phase.

We observed a disastrous structural degradation of the hairpin ribozyme global structure during the first tens of ps of thermalization phase. The structure of the minimal, junction-less hairpin ribozyme was severely perturbed at the end of 100-ps long thermalization and the unfolding of the structure was observed in the following production phase. Addition of explicit monovalent ions did not influence global structural behavior during the GB simulation.

c) The constant-pH MD simulations

To avoid the structural degradation described in the previous paragraph, we applied the harmonic cartesian restrains to all atoms of the hairpin ribozyme to keep them fluctuating around the starting structure corresponding to equilibrium in explicit solvent condition. Subsequently, the series of constant-pH MD simulation at different pHs were carried out to obtain pK_a constants of the two studied titrating groups. We run 18 simulations with pH ranging from 1 to 31. The simulation protocol was the same as described in the previous paragraph using the modified GB model developed by Onufriev, Bashford and Case (*igb*=2 in AMBER10) with the salt concentration 0.33 mol·L^{-1}. The Monte Carlo step was performed each 5 fs and data were collected for 5 ns after thermalization in each pK_a simulation. Finally, the fit to Henderson-Hasselbach equation was used to estimate pK_a constants (Fig. S10).

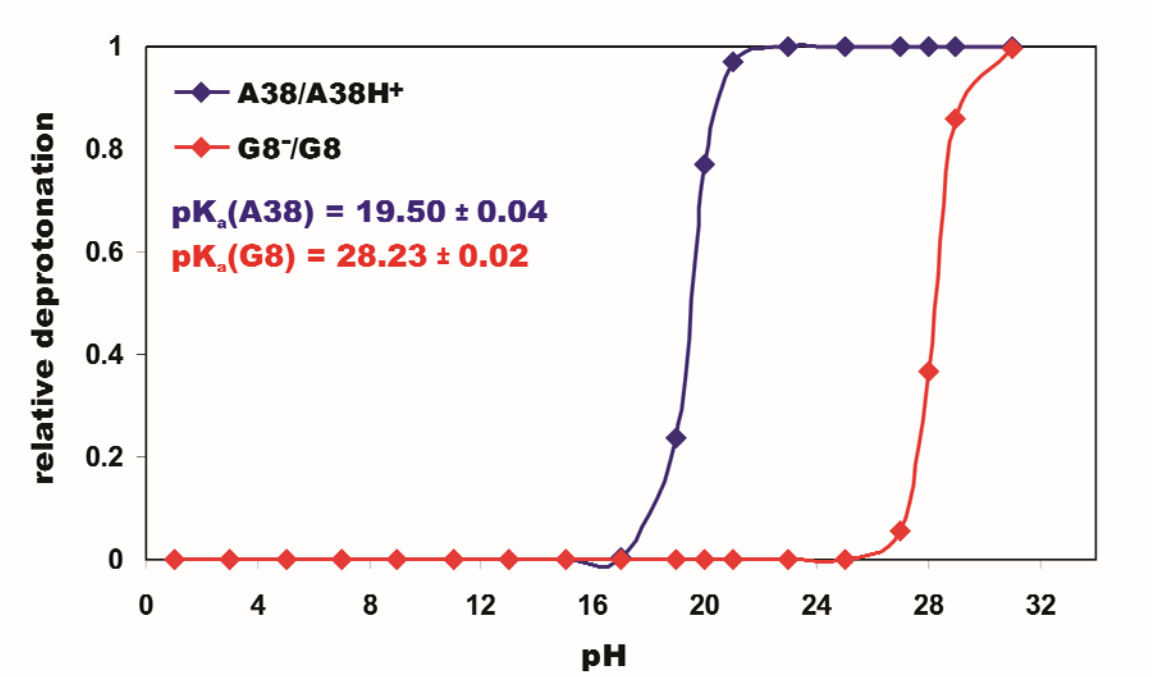

Figure S10. *The titration curves obtained from constant pH MD simulations. The blue and red diamonds stand for relative deprotonation of* $A38H^+$ *and guanine G8 at given pH. Obtained* pK_a *values are indicated in appropriate colors.*

References:

(1) Frisch, M., J.; Frisch, A.; Foresman, J., B. Gaussian 03; Gaussian, Inc.: Pittsburgh, 2003.

(2) 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. *Journal of the American Chemical Society* **1995**, *117*, 5179.

(3) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *Journal of the American Chemical Society* **1993**, *115*, 9620.

(4) Rupert, P. B.; Ferre-D'Amare, A. R. *Nature* **2001**, *410*, 780.