Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supplementary data

An Investigation on the Interaction Modes of a Single-Strand DNA Aptamer and RBP4 Protein: A Molecular Dynamic Simulations Approach

Raheleh Torabi[a], Kowsar Bagherzadeh[b,c], Hedayatollah Ghourchian[a], Massoud Amanlou[c*]

[a]Laboratory of Microanalysis, Institute of Biochemistry & Biophysics, University of Tehran, P.O. Box 13145-1384, Tehran, Iran.

[b]Razi Drug Research Center, Iran University of Medical Sciences, Tehran, Iran.

[c]Department of Medicinal Chemistry, Faculty of Pharmacy and Medicinal Plants Research Center, Tehran University of Medical Sciences, Tehran, Iran.

* Address correspondence to this author 3Department of Medicinal Chemistry, Faculty of Pharmacy and Medicinal Plants Research Center, Tehran University of Medical Sciences, Tehran, Iran; Tel: +98-21-66959067; Email: amanlou@tums.ac.ir

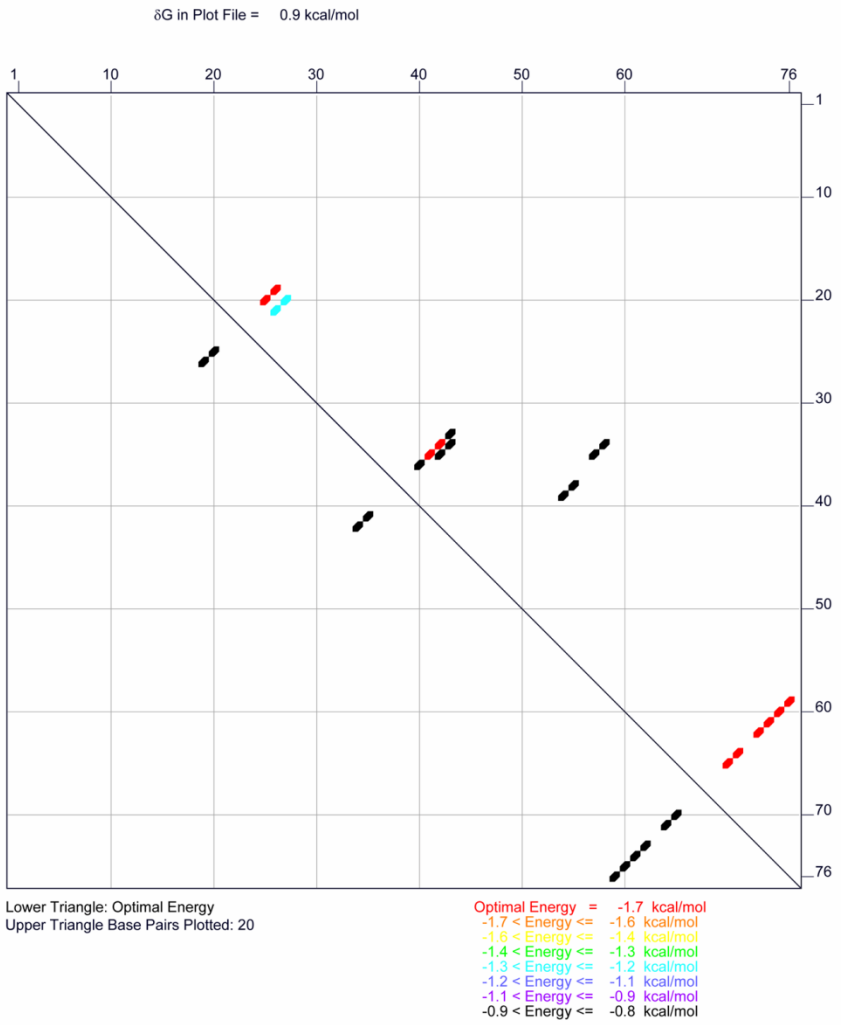


Figure S1. Fold of RBA at 37 C◦, obtained from mfold web server

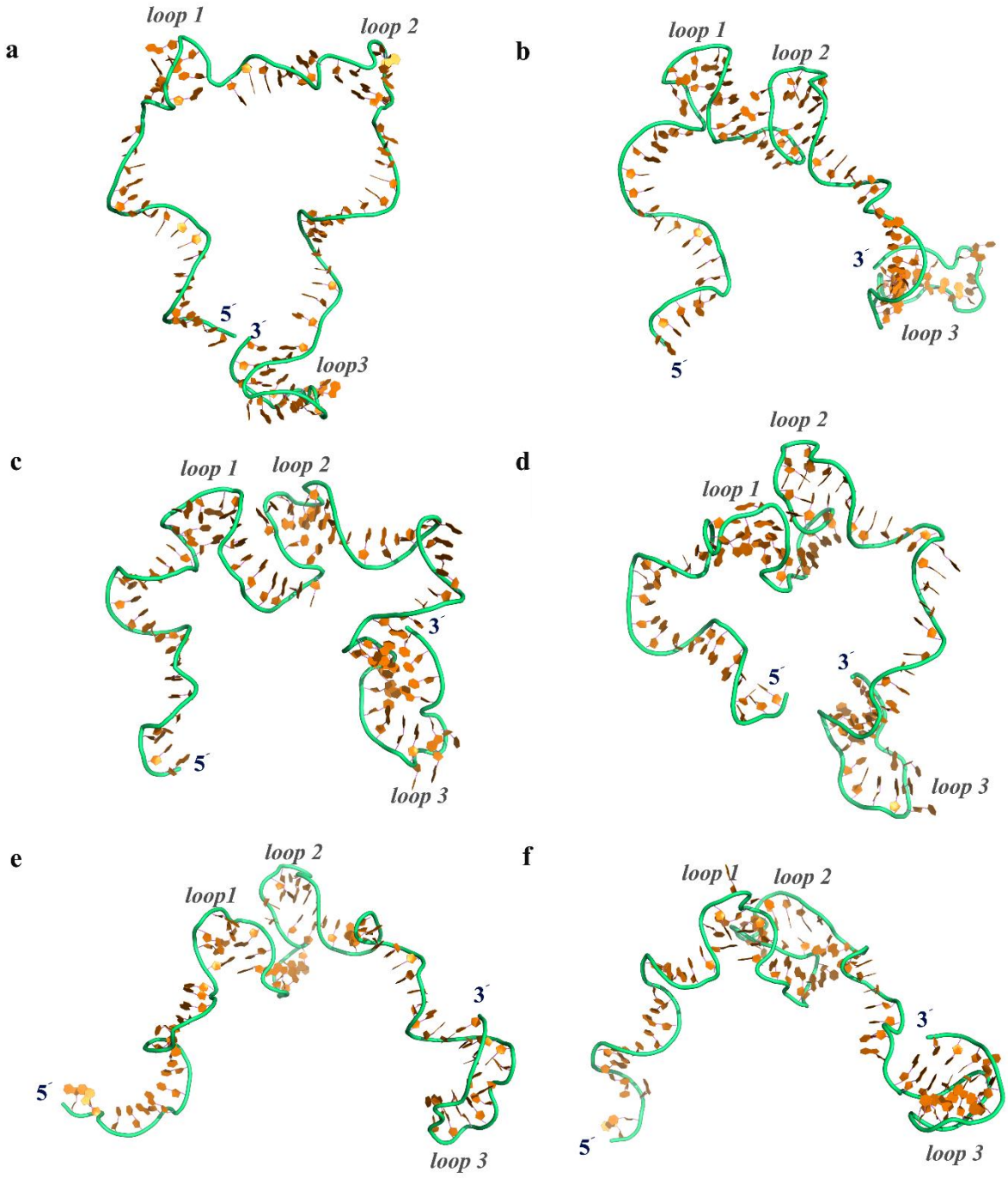


Figure S2. RBA evolution during 50 ns of MDs; **a**. the energy minimized structure, **b**. 10 ns, c. 20 ns, **d**. 30 ns (**RBA-a**), e. 40 ns, **f**. 50 ns (**RBA-b**)

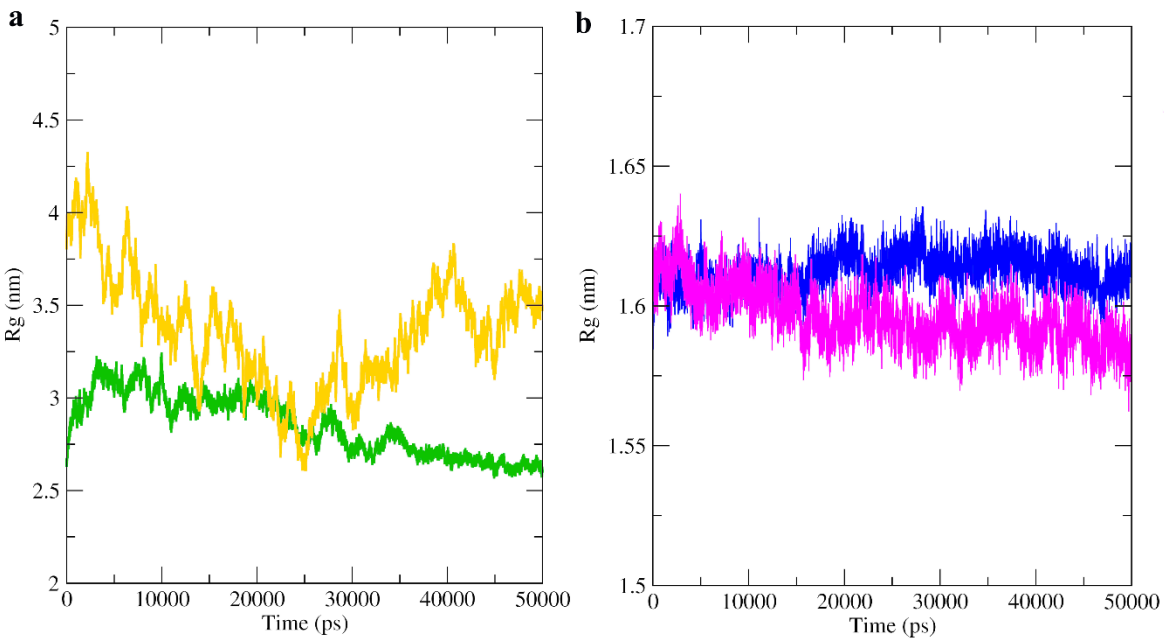
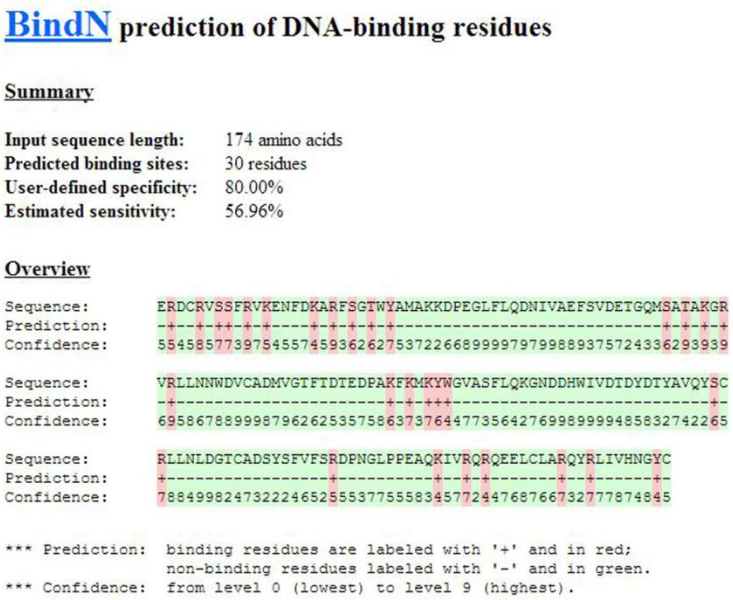


Figure S3. Plots of radius of gyration as a function of simulations time, **a**. lone RBA is shown in orange and RBA in complex with RBP4 is shown in green, **b**. lone RBP4 is shown in magenta and RBP4 in complex with RBA is shown in blue



BindN prediction of DNA-binding residues

Summary

Input sequence length: 174 amino acids
Predicted binding sites: 30 residues
User-defined specificity: 80.00%
Estimated sensitivity: 56.96%

Overview

```
Sequence:       ERDCRVSSFRVKENFDKARFSGTWYAMAKKDPEGLFLQDNIVAEFSVDETGQMSATAKGR
Prediction:     -+---+++-+-+----+-+-+-+-+------------------------------+-+-+-+
Confidence:     554585773975455745936262753722668999979799889375724336293939

Sequence:       VRLLNNWDVCADMVGTFTDTEDPAKFKMKYWGVASFLQKGNDDHWIVDTDYDTYAVQYSC
Prediction:     -+----------------------+-+-+++------------------------------+-
Confidence:     695867889998796262535758637376447735642769989999485832742265

Sequence:       RLLNLDGTCADSYSFVFSRDPNGLPPEAQKIVRQRQEELCLARQYRLIVHNGYC
Prediction:     +------------------+----------+---+-+-------+---+------+-
Confidence:     788499824732224652555377555834577244768766732777874845

*** Prediction:  binding residues are labeled with '+' and in red;
                 non-binding residues labeled with '-' and in green.
*** Confidence:  from level 0 (lowest) to level 9 (highest).
```

Figure S4. RBP4 residues predicted as surface ones which are liable of binding with the aptamer predicted by BindN webserver (http://bioinformatics.ksu.edu/bindn/)

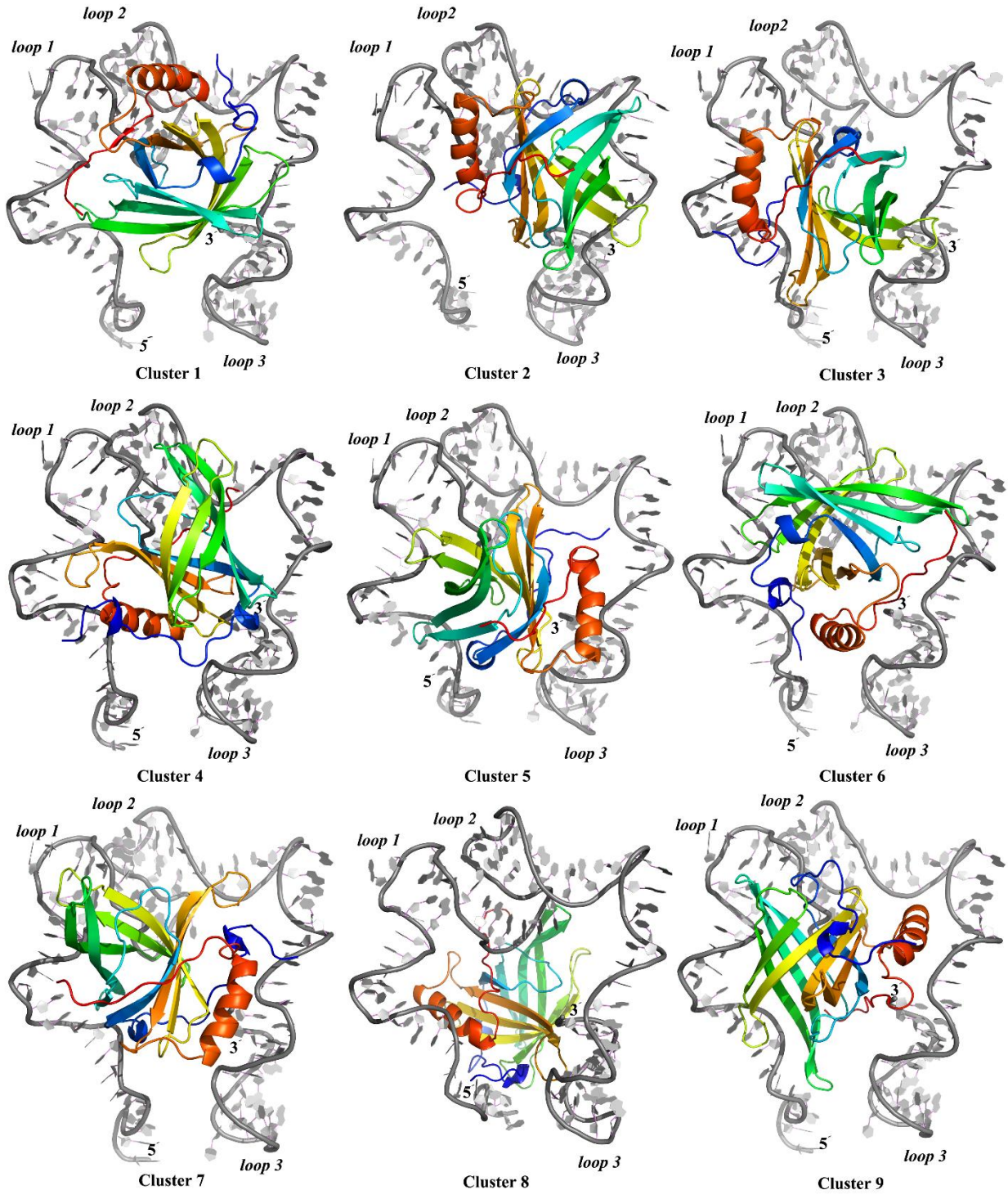


Figure S5. RBP4-(RBA-a) modes of interactions obtained from HADDOCK

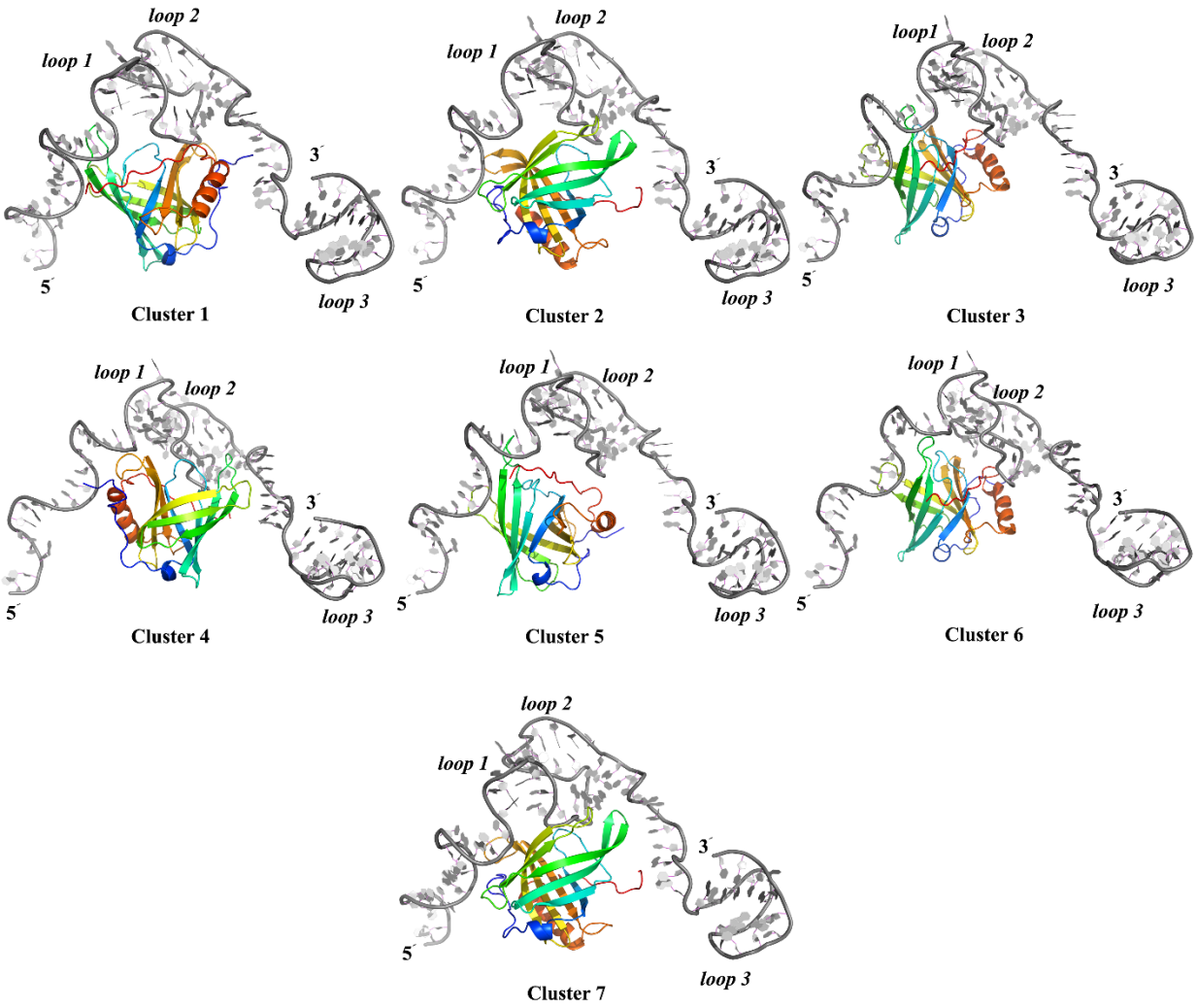


Figure S6. RBP4-(RBA-b) modes of interactions obtained from HADDOCK

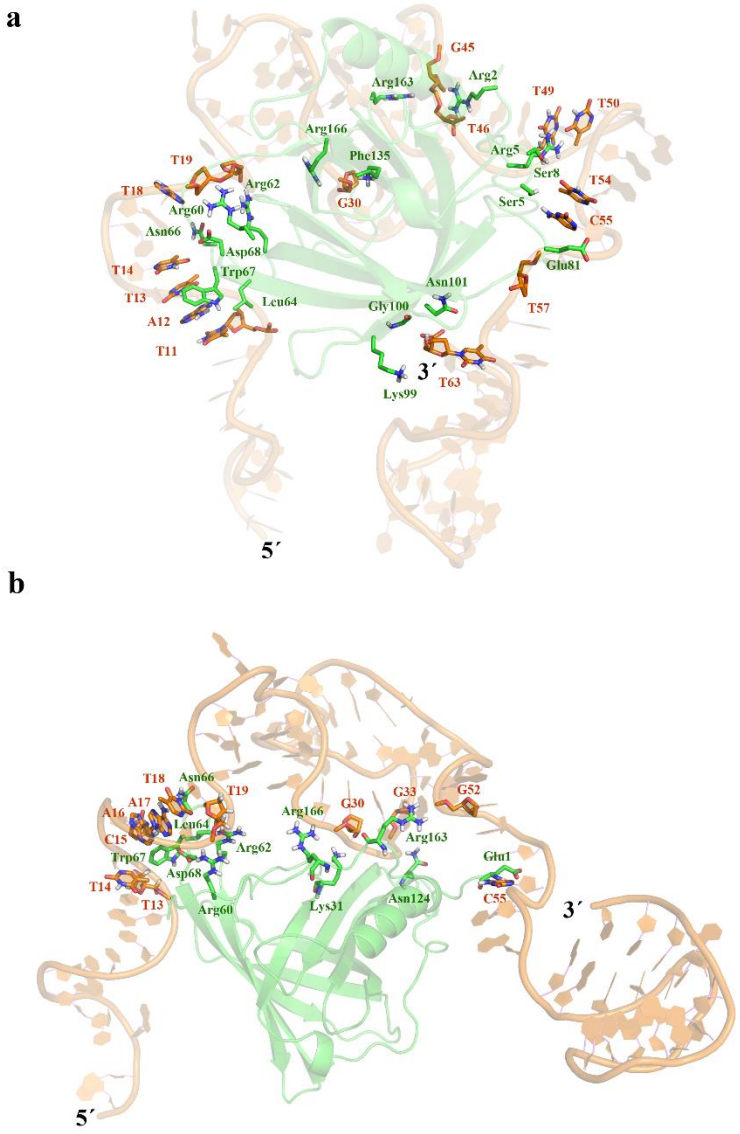


Figure S7. The structures obtained from HADDOCK for RBP4-RBA mode of interactions with the least HADDOCK score, a. RBP4 in complex with RBA-a, and b. RBP4 in complex with RBA-b.

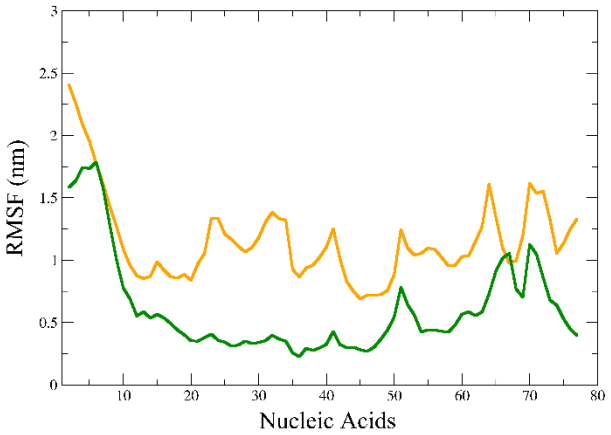


Figure S8. Root mean square fluctuations of the lone RBA (in orange) and the one in complex with RBP4 (green) nucleic acids durin 50 ns of simulations

Table S1. The statistics of the top 9 clusters of RBP4-RBA HADDOCK. The top cluster is the most reliable according to HADDOCK. Its Z-score indicates how many standard deviations from the average this cluster is located in terms of score (the more negative the better).

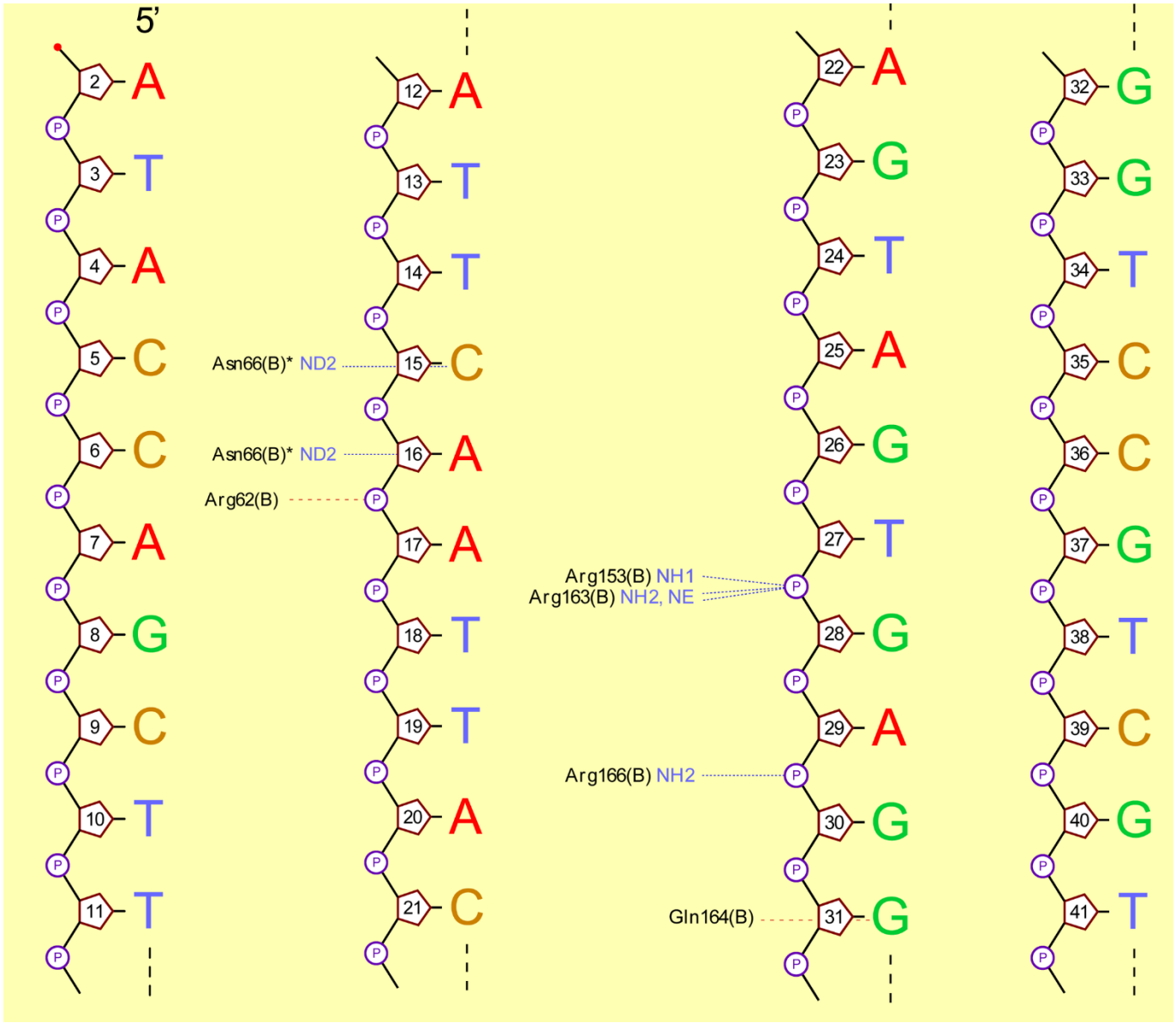
	Cluster 1	**Cluster 5**	**Cluster 2**	**Cluster 4**	**Cluster 6**	**Cluster 3**	**Cluster 7**	**Cluster 8**	**Cluster 9**
HADDOCK score	**-53.4 ±11.4**	**18.1 ± 18.6**	**25.3 ± 11.8**	**25.9 ± 7.8**	**39.3 ± 18.9**	**43.7 ± 13.6**	**60.6 ± 8.8**	**64.4 ± 19.4**	**81.9 ± 27.3**
Cluster size	23	6	19	7	6	11	6	4	4
RMSD from the overall lowest-energy structure	1.3 ± 0.8	17.0 ± 0.5	14.7 ± 0.1	18.7 ± 0.5	19.5 ± 0.0	17.7 ± 0.1	17.4 ± 0.4	24.2 ± 0.6	16.6 ± 0.4
Van der Waals energy	-83.5 ± 7.4	-64.5 ± 14.0	-79.2 ± 4.0	-68.5 ± 6.4	-67.7 ± 3.6	-42.7 ± 8.2	-47.0 ± 2.1	-66.3 ± 11.5	-61.2 ± 12.1
Electrostatic energy	-945.4 ± 33.8	-756.8 ± 105.2	-626.4 ± 60.3	-645.6 ± 93.2	-560.0 ± 132.7	-687.0 ± 80.1	-597.9 ± 92.7	-433.5 ± 88.2	-368.5 ± 28.7
Desolvation energy	56.4 ± 9.5	66.3 ± 7.7	57.2 ± 4.6	54.4 ± 6.6	52.3 ± 6.5	53.0 ± 7.2	52.0 ± 10.5	42.7 ± 11.2	36.2 ± 6.7
Restraints violation energy	1627.4 ± 78.7	1677.0 ± 149.3	1725.1 ± 128.1	1692.4 ± 140.3	1667.4 ± 97.5	1708.6 ± 78.2	1751.9 ± 63.1	1746.6 ± 45.5	1806.8 ± 129.1
Buried Surface Area	3463.6 ± 107.9	2749.1 ± 289.8	2723.6 ± 111.8	2754.4 ± 290.1	2701.9 ± 109.3	2075.8 ± 275.1	2165.5 ± 49.8	1998.2 ± 319.4	2330.3 ± 209.3
Z-Score	-2.4	-0.4	-0.2	-0.2	0.1	0.3	0.7	0.8	1.3

Table S2. The statistics of the top 7 clusters of RBP4-RBA HADDOCK. The top cluster is the most reliable according to HADDOCK. Its Z-score indicates how many standard deviations from the average this cluster is located in terms of score (the more negative the better).

	Cluster1	**Cluster2**	**Cluster7**	**Cluster3**	**Cluster6**	**Cluster4**	**Cluster5**
HADDOCK score	**22.2 ± 7.4**	**50.4 ± 11.1**	**50.6 ± 15.8**	**54.0 ± 8.4**	**62.4 ± 20.9**	**82.1 ± 9.7**	**91.7 ± 11.0**
Cluster size	16	11	4	8	4	5	4
RMSD from the overall lowest-energy structure	23.1 ± 0.5	23.8 ± 0.3	24.2 ± 0.4	19.4 ± 0.1	18.8 ± 0.2	24.8 ± 0.2	20.7 ± 0.1
Van der Waals energy	-61.6 ± 2.5	-62.3 ± 5.3	-66.7 ± 6.1	-60.9 ± 7.6	-61.3 ± 5.8	-44.4 ± 5.7	-42.7 ± 9.6
Electrostatic energy	-413.1 ± 41.8	-251.6 ± 19.3	-238.0 ± 51.9	-189.8 ± 39.3	-211.7 ± 65.6	-210.6 ± 29.2	-101.8 ± 9.4
Desolvation energy	23.5 ± 6.0	21.0 ± 6.9	22.4 ± 2.4	2.2 ± 4.6	11.1 ± 6.0	28.2 ± 5.5	-5.3 ± 1.7
Restraints violation energy	1428.8 ± 21.7	1420.3 ± 23.6	1424.8 ± 53.6	1507.4 ± 54.4	1549.3 ± 19.69	1404.0 ± 40.3	1600.6 ± 13.5
Buried Surface Area	2280.9 ± 117.7	2001.2 ± 125.5	2035.4 ± 149.8	1711.5 ± 143.7	1858.8 ± 97.4	1563.1 ± 106.8	1319.9 ± 142.8
Z-Score	**-1.7**	-0.4	-0.4	-0.2	0.2	1.1	1.5

RBP4 and RBP interactions during 50 ns of MD simulations

In all Figures 9, 10, 11, 12, and 13, the pentagon shows backbone sugar and base-number, the P shows phosphate group, the star shows residue/water on plot more than once, the blue dashes show Hydrogen bond to ss-DNA, the red dashes show non-bonded contact to ss-DNA (< 3.35A). Figures are generated by NUCPLOT v.1.1.4 [1] .


5'
2 A
3 T
4 A
5 C
6 C
7 A
8 G
9 C
10 T
11 T
12 A
13 T
14 T
15 C
16 A
17 A
18 T
19 T
20 A
21 C
22 A
23 G
24 T
25 A
26 G
27 T
28 G
29 A
30 G
31 G
32 G
33 G
34 T
35 C
36 C
37 G
38 T
39 C
40 G
41 T
Asn66(B)* ND2
Asn66(B)* ND2
Arg62(B)
Arg153(B) NH1
Arg163(B) NH2, NE
Arg166(B) NH2
Gln164(B)

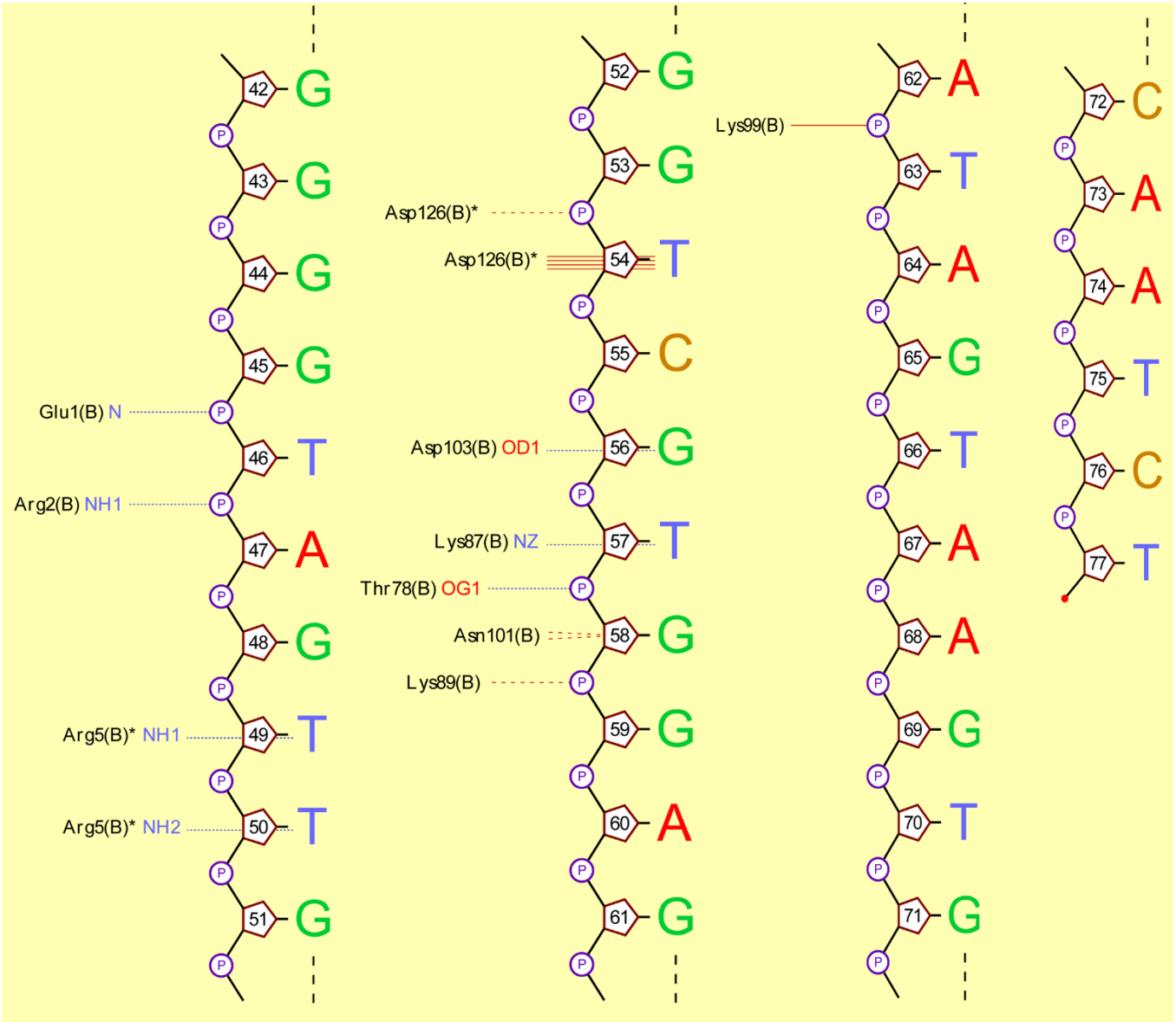
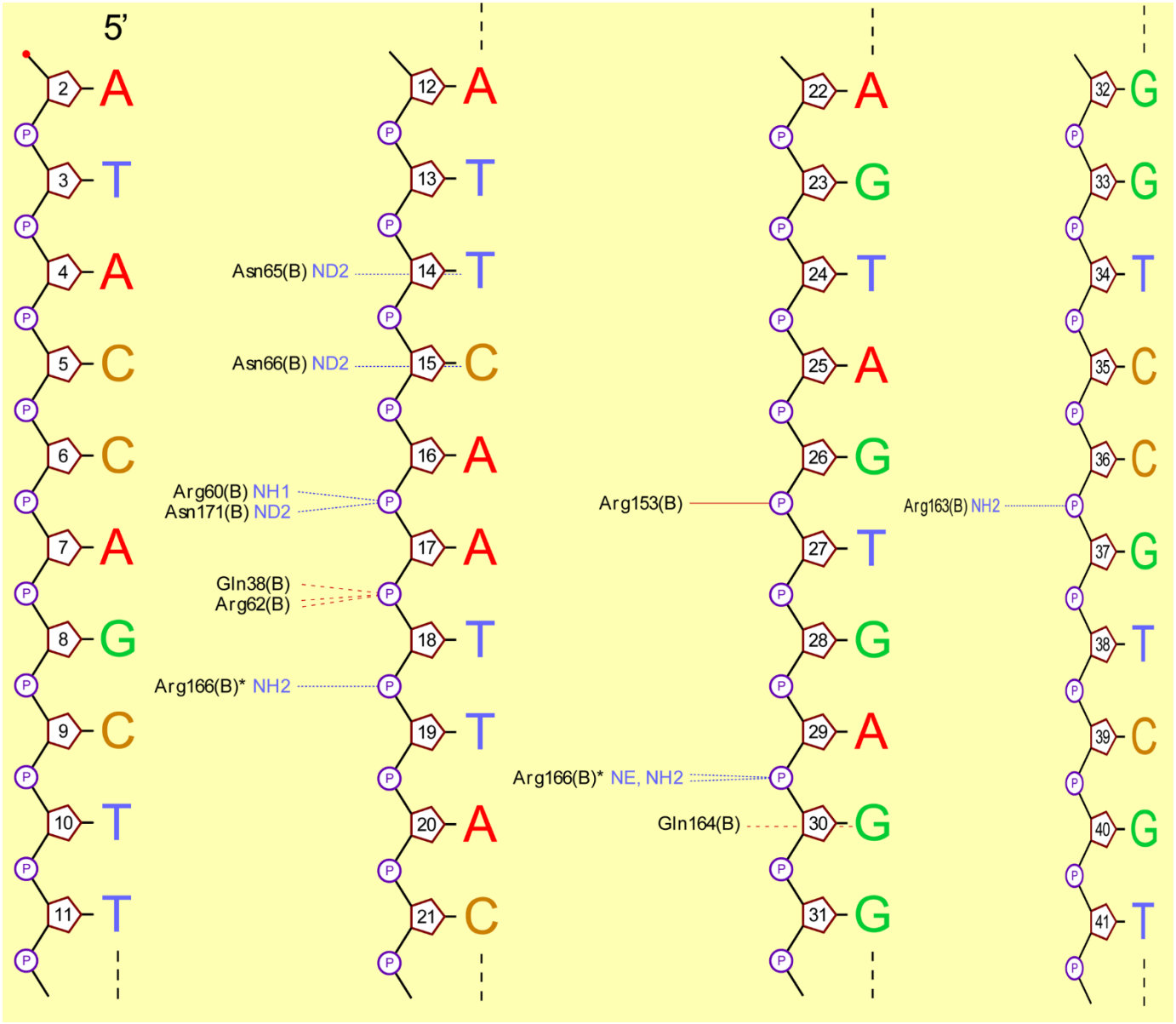


Figure S9. Interactions between RBP4 and RBA after **10ns** of MD simulations


5’
2 A
3 T
4 A
5 C
6 C
7 A
8 G
9 C
10 T
11 T
12 A
13 T
14 T
15 C
16 A
17 A
18 T
19 T
20 A
21 C
22 A
23 G
24 T
25 A
26 G
27 T
28 G
29 A
30 G
31 G
32 G
33 G
34 T
35 C
36 C
37 G
38 T
39 C
40 G
41 T
Asn65(B) ND2
Asn66(B) ND2
Arg60(B) NH1
Asn171(B) ND2
Gln38(B)
Arg62(B)
Arg166(B)* NH2
Arg153(B)
Arg166(B)* NE, NH2
Gln164(B)
Arg163(B) NH2

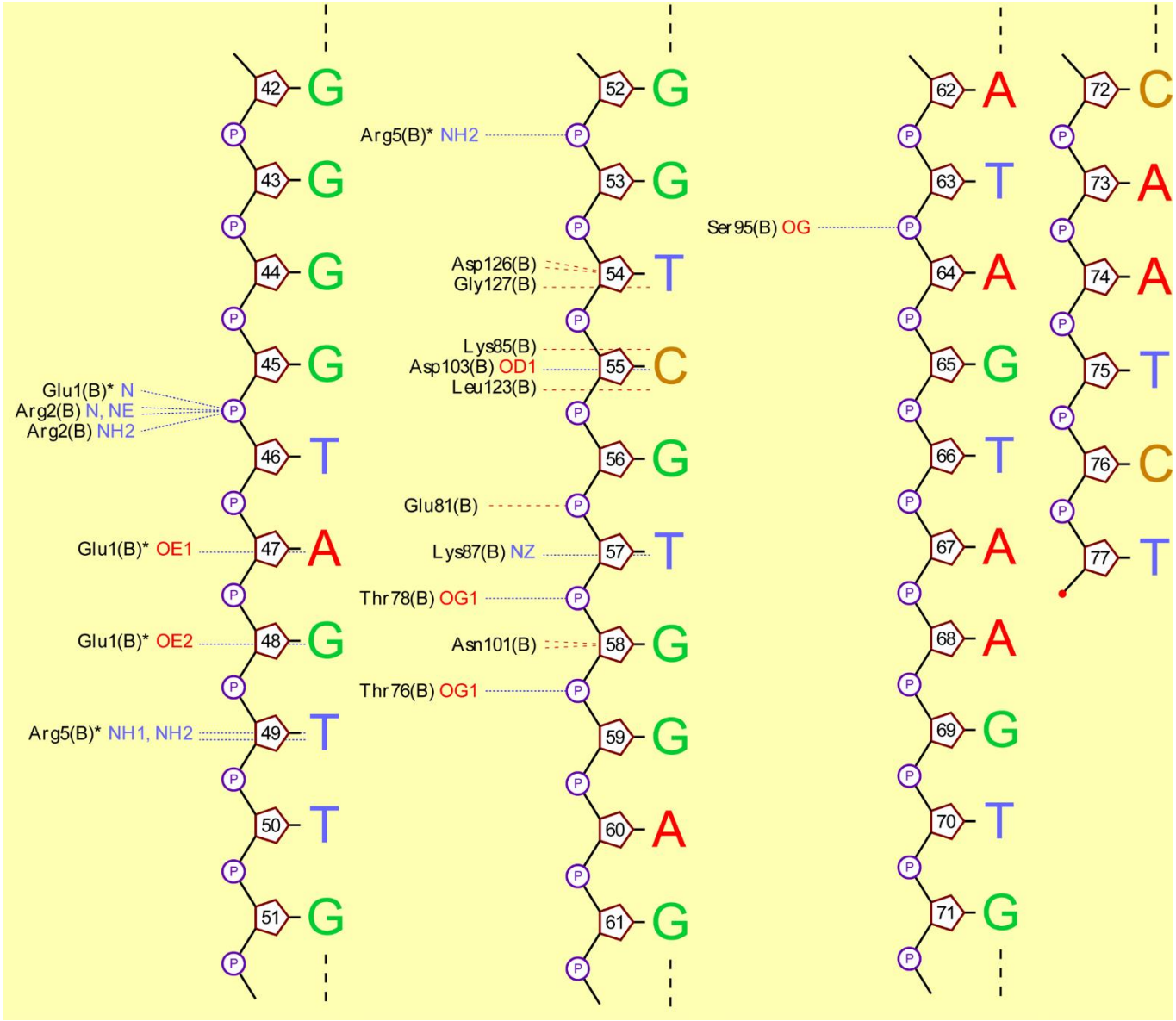
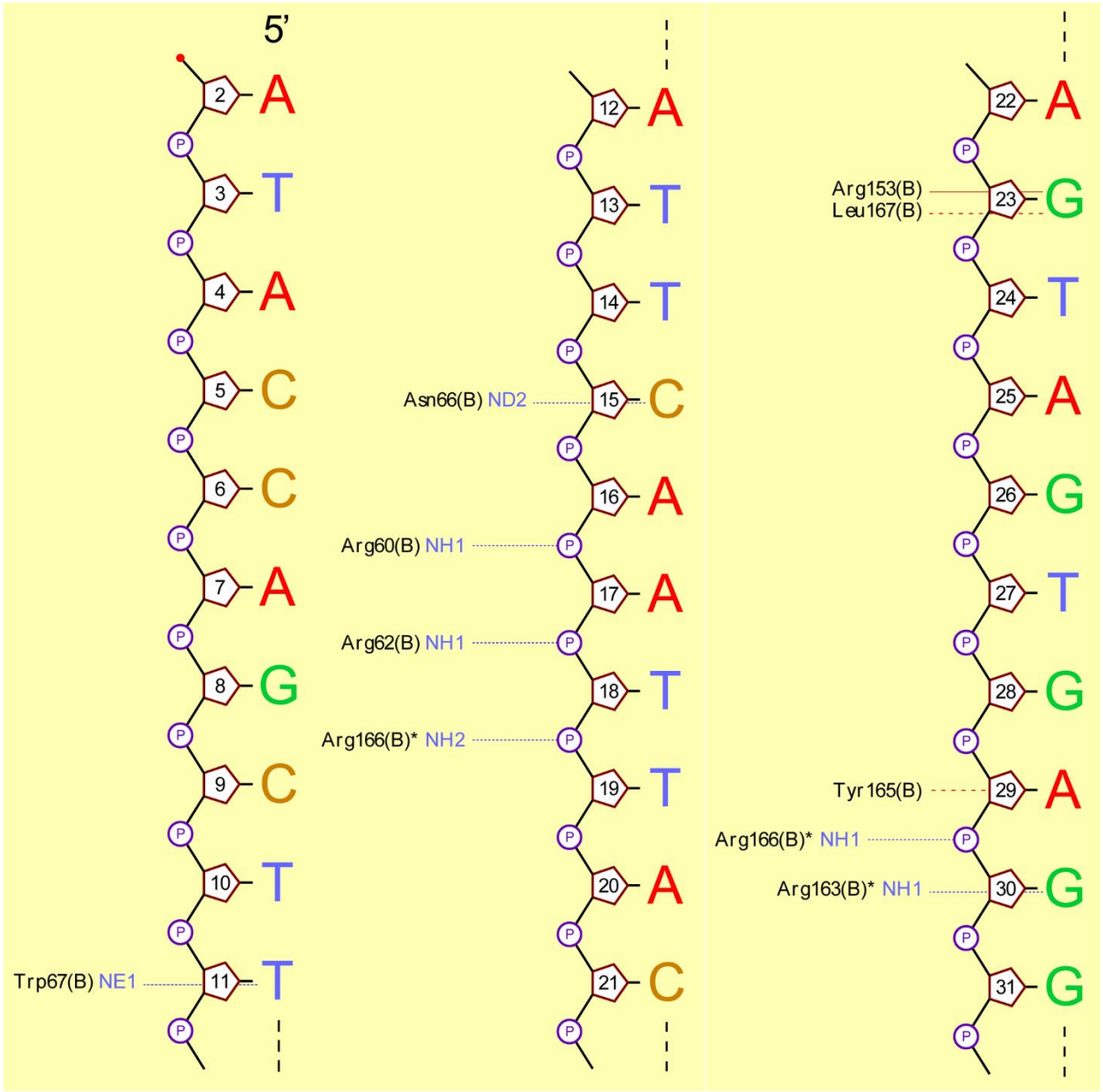
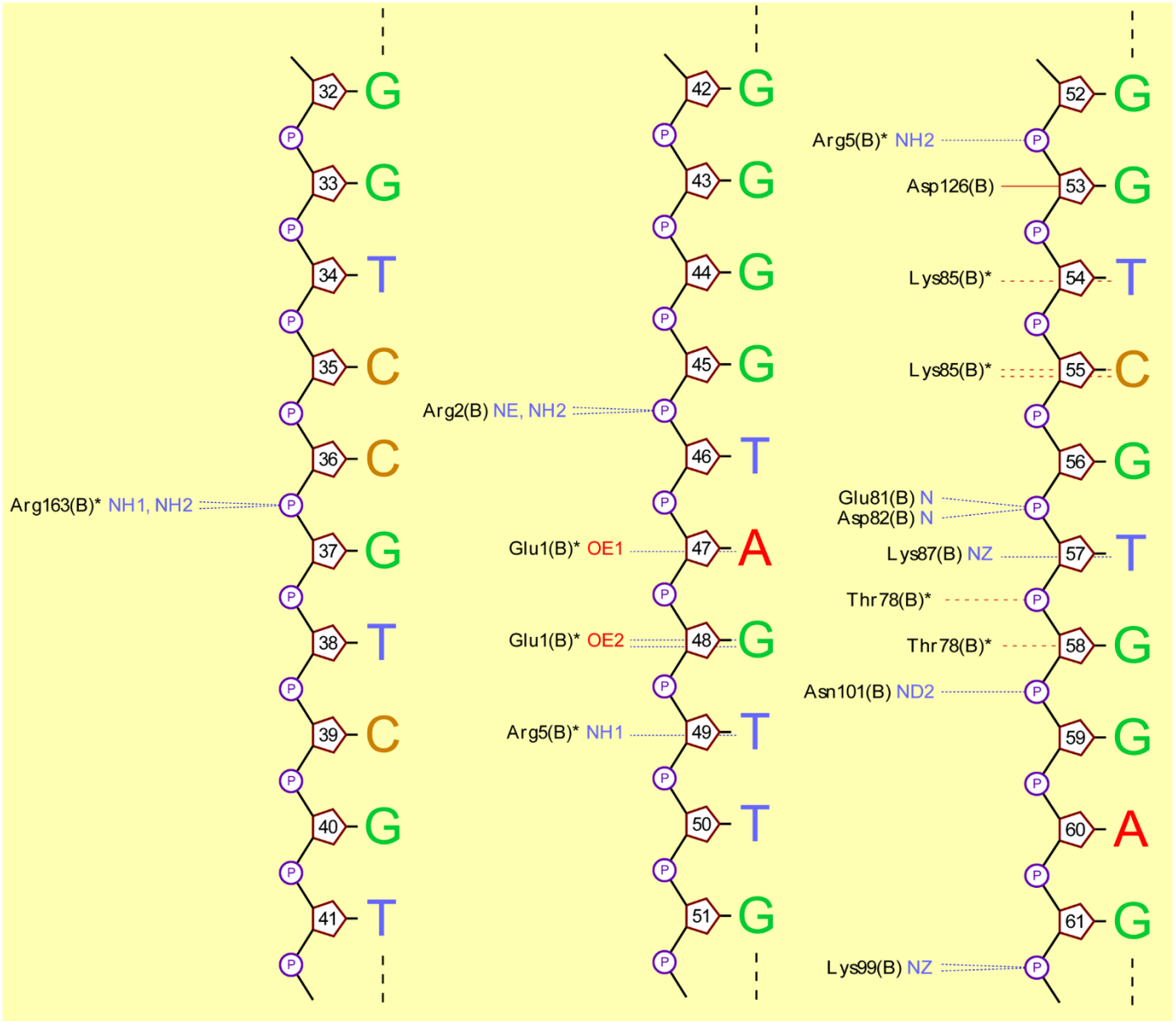


Figure 10. Interactions between RBP4 and RBA after **20ns** of MD simulations


5'
2 A
3 T
4 A
5 C
6 C
7 A
8 G
9 C
10 T
11 T
Trp67(B) NE1
12 A
13 T
14 T
15 C
Asn66(B) ND2
16 A
Arg60(B) NH1
17 A
Arg62(B) NH1
18 T
Arg166(B)* NH2
19 T
20 A
21 C
22 A
Arg153(B)
Leu167(B)
23 G
24 T
25 A
26 G
27 T
28 G
Tyr165(B)
29 A
Arg166(B)* NH1
Arg163(B)* NH1
30 G
31 G


32 G
33 G
34 T
35 C
36 C
Arg163(B)* NH1, NH2
37 G
38 T
39 C
40 G
41 T
42 G
43 G
44 G
45 G
Arg2(B) NE, NH2
46 T
Glu1(B)* OE1
47 A
Glu1(B)* OE2
48 G
Arg5(B)* NH1
49 T
50 T
51 G
52 G
Arg5(B)* NH2
Asp126(B)
53 G
Lys85(B)*
54 T
Lys85(B)*
55 C
56 G
Glu81(B) N
Asp82(B) N
Lys87(B) NZ
57 T
Thr78(B)*
Thr78(B)*
58 G
Asn101(B) ND2
59 G
60 A
61 G
Lys99(B) NZ

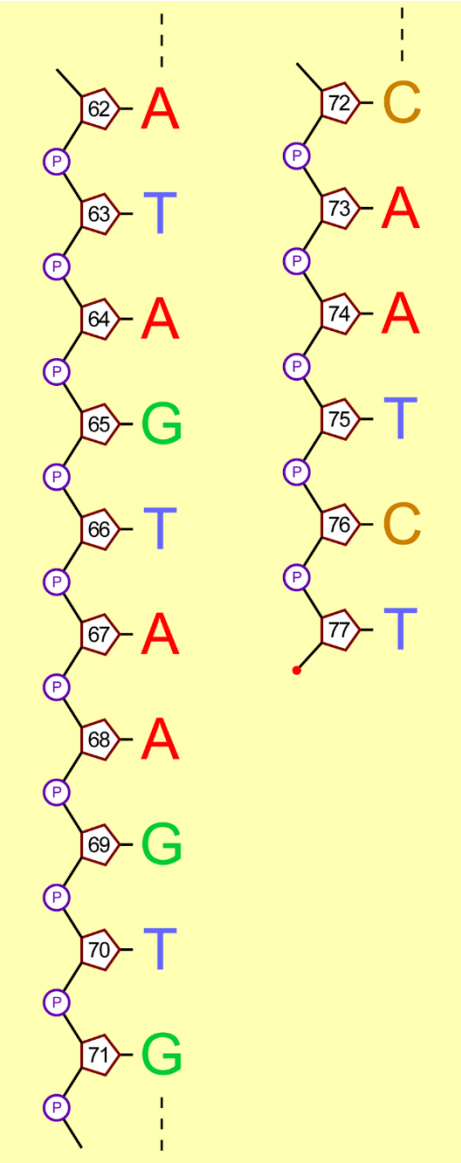
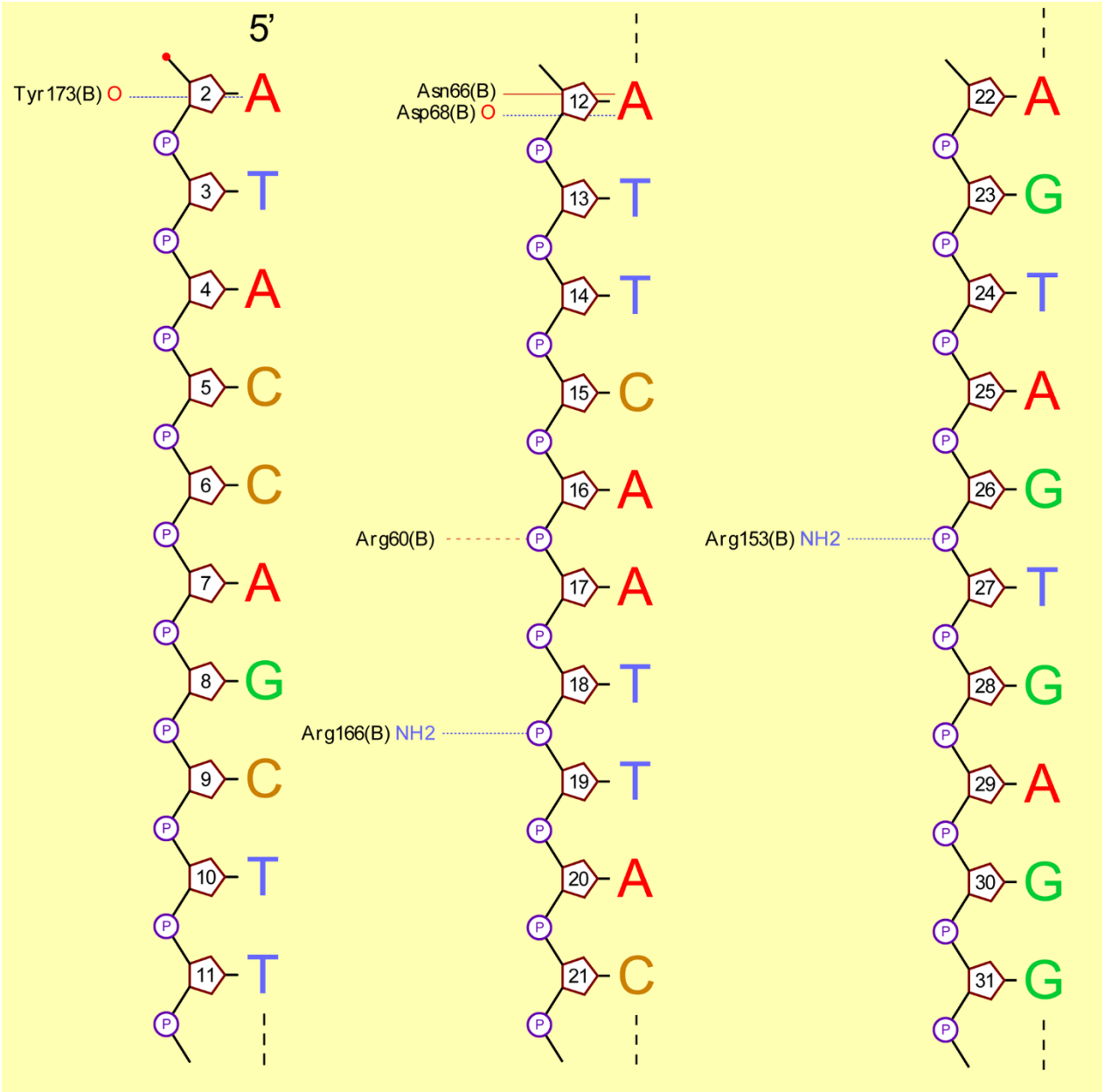
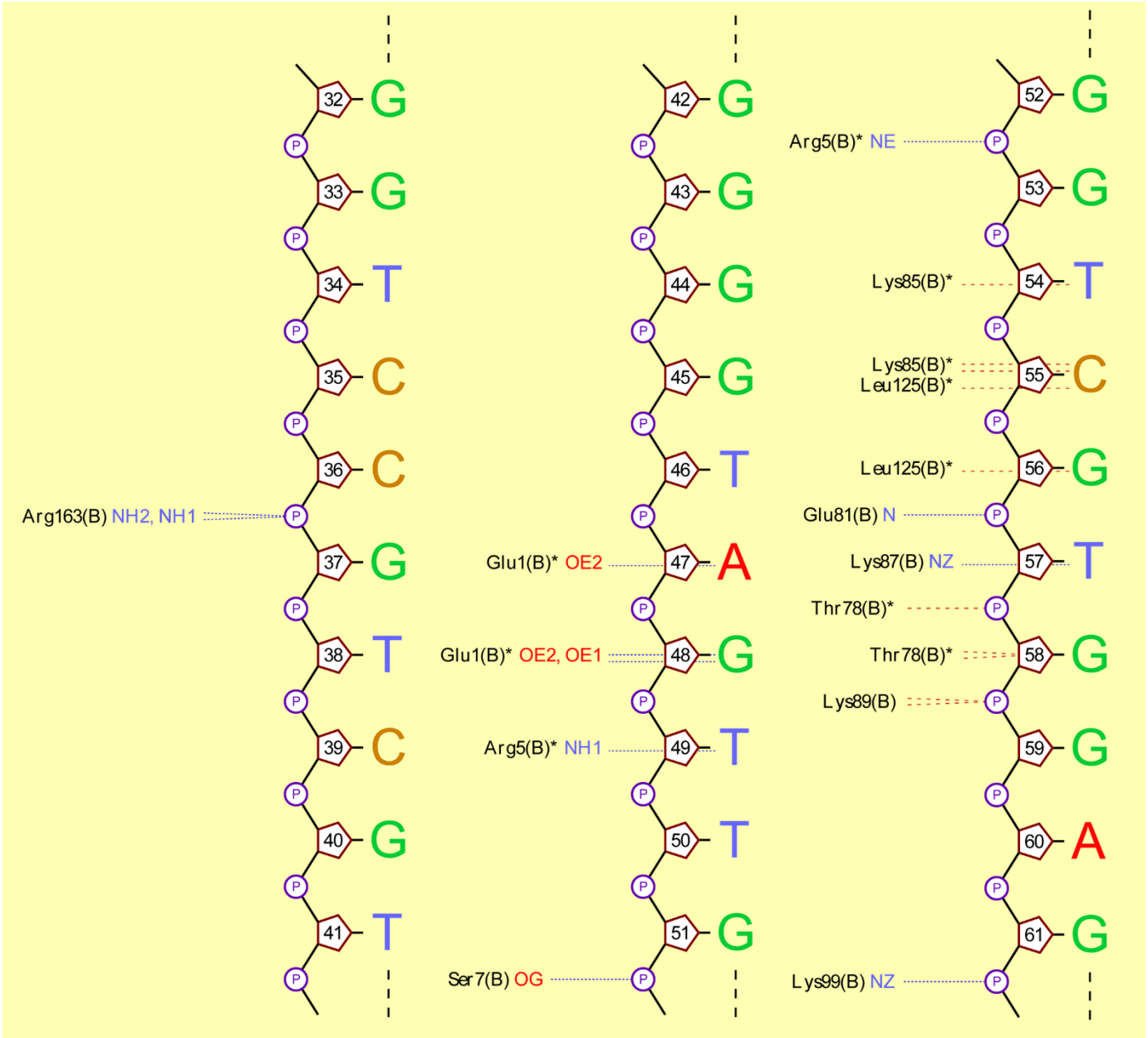


Figure S11. Interactions between RBP4 and RBA after **30ns** of MD simulations


5’
Tyr173(B) O
Asn66(B)
Asp68(B) O
Arg60(B)
Arg166(B) NH2
Arg153(B) NH2
2 A
3 T
4 A
5 C
6 C
7 A
8 G
9 C
10 T
11 T
12 A
13 T
14 T
15 C
16 A
17 A
18 T
19 T
20 A
21 C
22 A
23 G
24 T
25 A
26 G
27 T
28 G
29 A
30 G
31 G
P


32 G
33 G
34 T
35 C
36 C
Arg163(B) NH2, NH1
37 G
38 T
39 C
40 G
41 T
42 G
43 G
44 G
45 G
46 T
Glu1(B)* OE2
47 A
Glu1(B)* OE2, OE1
48 G
Arg5(B)* NH1
49 T
50 T
51 G
Ser7(B) OG
52 G
Arg5(B)* NE
53 G
Lys85(B)*
54 T
Lys85(B)*
Leu125(B)*
55 C
Leu125(B)*
56 G
Glu81(B) N
Lys87(B) NZ
57 T
Thr78(B)*
Thr78(B)*
58 G
Lys89(B)
59 G
60 A
61 G
Lys99(B) NZ

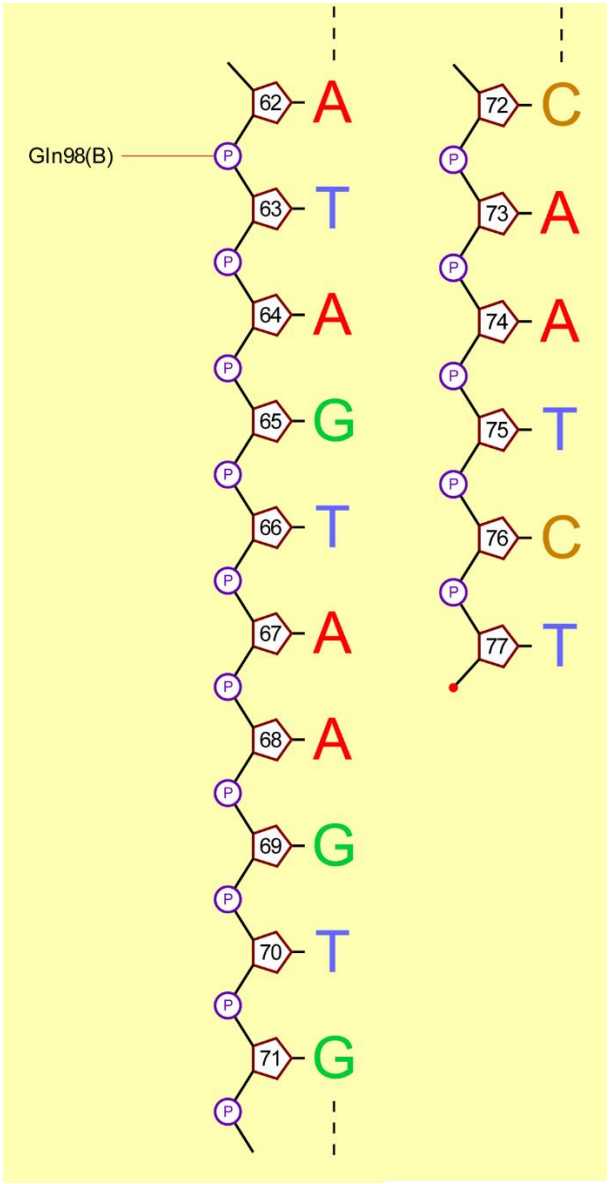
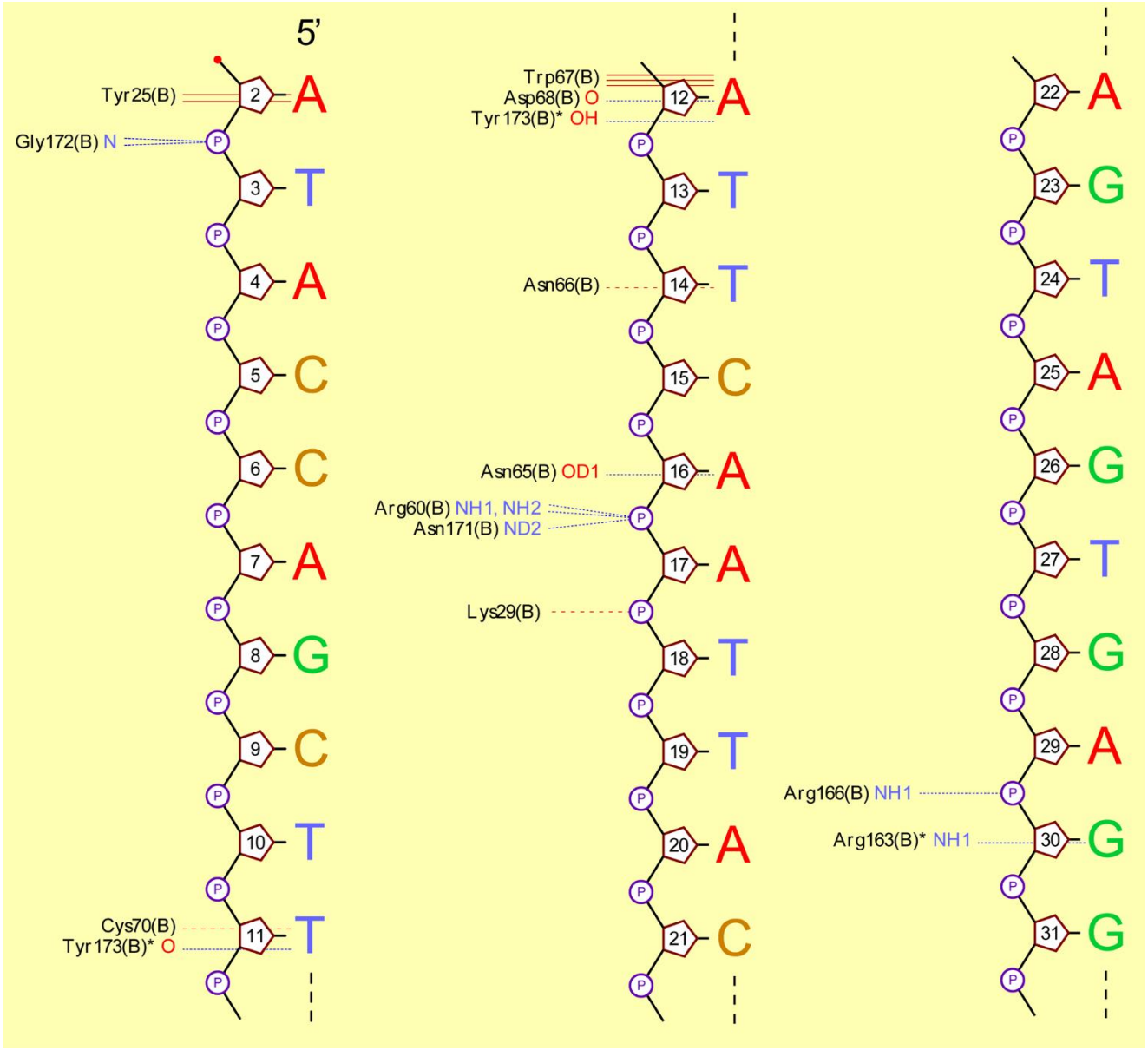
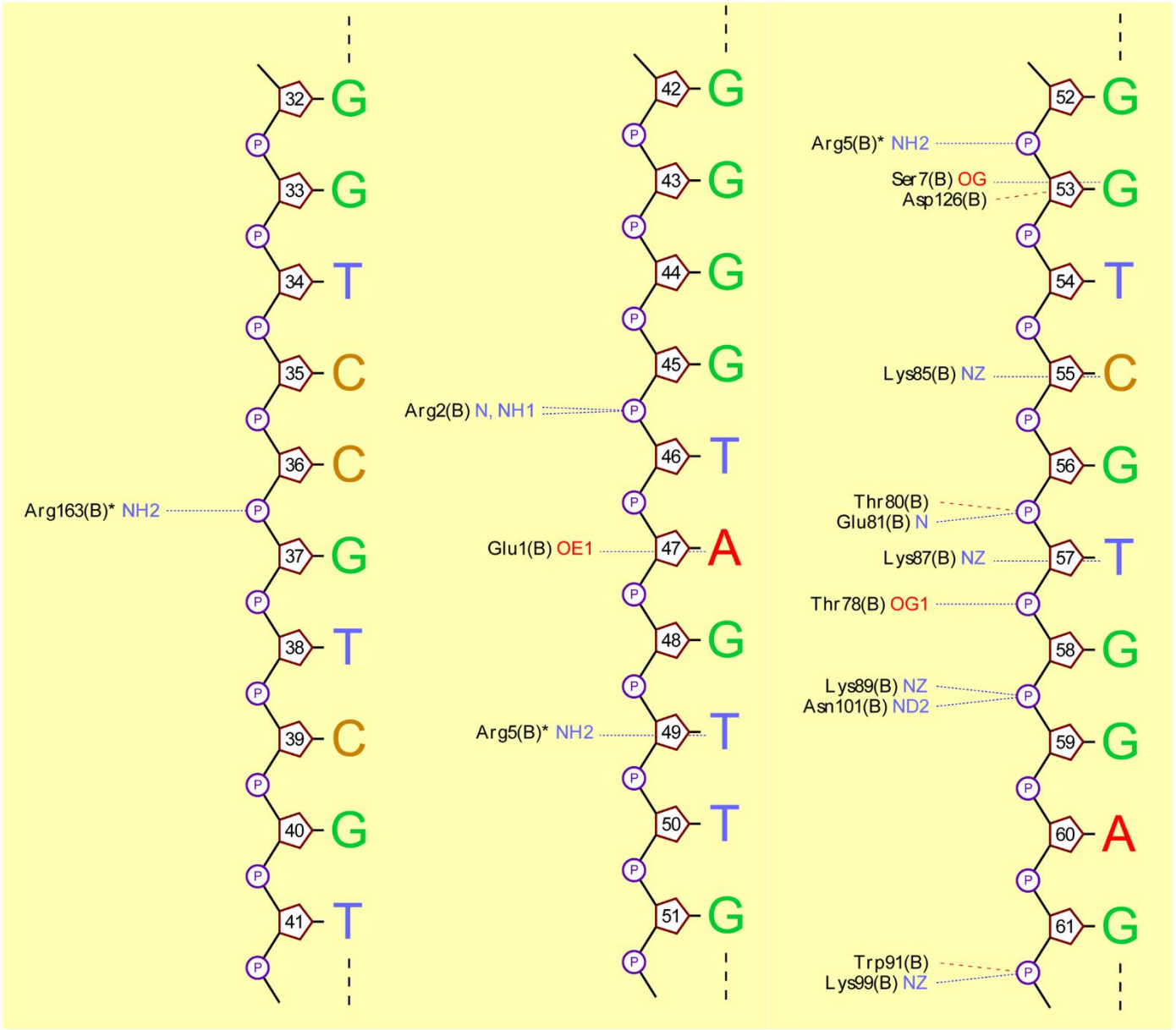


Figure S12. Interactions between RBP4 and RBA after **40ns** of MD simulations


5'
Tyr25(B)
Gly172(B) N
Cys70(B)
Tyr173(B)* O
Trp67(B)
Asp68(B) O
Tyr173(B)* OH
Asn66(B)
Asn65(B) OD1
Arg60(B) NH1, NH2
Asn171(B) ND2
Lys29(B)
Arg166(B) NH1
Arg163(B)* NH1
A T A C C A G C T T A T T A C A G T A G T A G G
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
P


32 G
33 G
34 T
35 C
36 C
Arg163(B)* NH2
37 G
38 T
39 C
40 G
41 T
42 G
43 G
44 G
45 G
Arg2(B) N, NH1
46 T
Glu1(B) OE1
47 A
48 G
Arg5(B)* NH2
49 T
50 T
51 G
52 G
Arg5(B)* NH2
Ser7(B) OG
Asp126(B)
53 G
54 T
Lys85(B) NZ
55 C
56 G
Thr80(B)
Glu81(B) N
Lys87(B) NZ
57 T
Thr78(B) OG1
58 G
Lys89(B) NZ
Asn101(B) ND2
59 G
60 A
61 G
Trp91(B)
Lys99(B) NZ

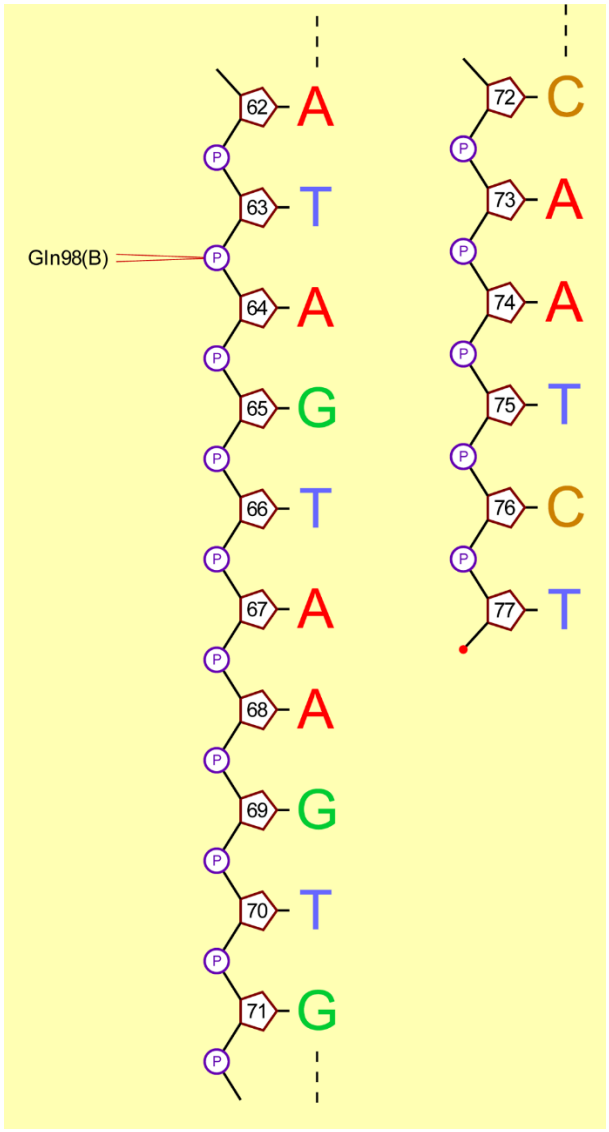


Figure S13. Interactions between RBP4 and RBA after 50ns of MD simulations

Reference

1. N. M. Luscombe, R. A. Laskowski and J. M. Thornton, *Nucleic Acids Res*, 1997, **25**, 4940-4945.