Supporting Information

Title: Engineering Cooperative Tecto-RNA Complexes Having Programmable Stoichiometries

Authors: Irina V. Novikova[1], Bachar H. Hassan [1§], Marina Mirzoyan[1], Neocles B. Leontis [1,2*]

[1] Department of Chemistry and Center for Photochemical Sciences, Bowling Green State University, Bowling Green, Ohio 43403

[2] Center for Biomolecular Sciences, Bowling Green, State University, Bowling Green, Ohio 43403

[§]**Present address:** Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[*]**To whom correspondence should be addressed.** Neocles B. Leontis, phone: 1-419-372-8663; fax: 1-419-372-9809; email: leontis@bgsu.edu or nleontis@nsf.gov

Supplemental Material S-1. List of RNA sequences used in the study.

Sequences used for homo-assembly:

1c (8/3//6/5):gggauaugguaccaccugccugugugcuucggcaaucuggcaaccgggggaacuugguacggggaaacucgugguggguaccuaaguccuu
1b (7/4//6/5):gggauaugguccaccugaccugugugcuucggcaaucuggucaaccgggggaacuugguacggggaaacucgugguggaccuaaguccuu
1a (6/5//6/5):gggauaugguccaccugacccugugugcuucggcaaucugggucaaccgggggaacuugguacggggaaacucguggugaccuaaguccuu
8 (5/6//6/5):gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccgggggaacuugguacggggaaacucgugguaccuaaguccuu
2a (5/6//5/6):gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccgggggaacugguacggggaaacuucgugguaccuaaguccuu
2b(5/6//4/7):gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccgggaacgguacggcgggaaacugucgugguaccuaaguccuu
2c (5/6//3/8):gggauaugguaccugacuccugugugcuucggcaaucuggagucaacgggaacguacggccgggaaacuggucgugguaccuaaguccuu
2d (5/6//2/9):gggauaugguaccugacuccugugugcuucggcaaucuggagucaacggaaguacggcucgggaaacugagucgugguaccuaaguccuu
3a (7/4//5/6):gggauaugguccaccugaccugugugcuucggcaaucuggucaaccgggaacugguacggggaaacuucgugguggaccuaaguccuu
3 (6/5//5/6):gggauauggucaccugacccugugugcuucggcaaucugggucaaccgggaacugguacggggaaacuucguggugaccuaaguccuu
3b (6/5//4/7):gggauauggucaccugacccugugugcuucggcaaucugggucaaccgggaacgguacggcgggaaacugucguggugaccuaaguccuu
3c (6/5//3/8):gggauauggucaccugacccugugugcuucggcaaucugggucaacgggaacguacggccgggaaacuggucguggugaccuaaguccuu
4 (7/4//4/7): gggauaugguccaccugaccugugugcuucggcaaucuggucaaccgggaacgguacggcgggaaacugucgugguggaccuaaguccuu

Sequences used for hetero-assembly:

1bDA:gggauaugguccaccugaccugugugcuucggcaaucuggucaaccgggguaacuugguacggggaacucgugguggaccuaaguccuu
1bDB:gggaaucugguccaccugacgcuacuugcuucggcaaaggcgucaaccgggggaacuugguacggggaaacucgugguggaccugugucuu
8TA: gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccgggguaacuugguacggggaacucgugguaccuaaguccuu
8TB: gggaaucugguaccugacucgcuacuugcuucggcaaaggcgagucaaccgggguaaacuugguacgggguaacucgugguaccugugucuu
8TC:gggaaaggcguaccugacuccuaagugcuucggcauauggagucaaccgggggaacuugguacggggaaacucgugguacgcuacuuccuu
1bDBmut:gggaaucugguccaccugaaaaggcggcuucggccgcuacuuucaaccgggggaacuugguacggggaaacucgugguggaccugugucuu
8TBmut: gggaaucugguaccugacuaaaggcggcuucggccgcuacuuagucaaccgggggaaacuugguacgggguaacucgugguaccugugucuu
8TCmut:gggcgcuacuuuaccugacuccuaagugcuucggcauauggagucaaccgggggaacuugguacggggaaacucgugguaaaaggcgccuu
8DA/8TA/8HA: gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccgggguaacuugguacggggaacucgugguaccuaaguccuu
8DB: gggaaucugguaccugacucgcuacuugcuucggcaaaggcgagucaaccgggggaacuugguacggggaaacucgugguaccugugucuu
8TetC/8HC:gggaaaggcguaccugacuccuaagugcuucggcauauggagucaaccgggcuaagugcuucggcauaugcuugguacgggggaa
cucgugguacgcuacuuccuu
8TetD/8PA: gggaaucugguaccugacucgaaagagucaaccgggggaacuugguacggggaaacucgugguaccugugucuu
8PB/8HD: gggaaucugguaccugacucgaaagagucaaccggggcuacuugcuucggcaaaggccuugguacggggaaacucgugguaccugugucuu
8PC/8HE: gggauaugguaccugacucguaagagucaaccggggcuacuugcuucggcaaaggccuugguacgggggaacucgugguaccuaaguccuu
8PD/8HF: gggaaucugguaccugacucguaagagucaaccgggggaacuugguacggggaaacucgugguaccugugucuu
8PE: gggauaugguaccugacuccugugugcuucggcaaucuggagucaaccggggaaacuugguacgggguaacucgugguaccuaaguccuu
8TB/8HB: gggaaucugguaccugacucgcuacuugcuucggcaaaggcgagucaaccggggaaacuugguacgggguaacucgugguaccugugucuu
B(7): gggauauggucguggucuccuaagugcuucggcauauggagaccugagggaaacuucaaccggggaaacuugguacgaccuaaguccuu

Supplemental Material S-2: Determination of equilibrium constants for dimer and trimer formation.

Here, we derive the equations used to determine equilibrium constants for dimer and trimer formation, treating hetero- and homo-complex formation separately.

1.Heterodimer formation

First, we consider association of two monomers, M1 and M2, to form a dimer D. For convenience, we write this as a dissociation process:

$$D \rightarrow M1 + M2,\ K_d(het-\mathrm{dim}er) = \frac{[M1]\cdot[M2]}{[D]} \quad (1)$$

For equilibrium constant determination using gel mobility shifts, one molecule is radiolabeled and used in much lower concentration than the other. We will take M1 to be the radiolabeled molecule so that [M1] << [M2]. Because of this, we can assume [M2] >> [D] and then we can take [M2] equal to the amount added to the solution. Let T to be the total concentration of M1, in monomer and dimer form:

$$T = [M1] + [D] \quad (2)$$

The dimer concentration is equal to T, the total concentration of M1, times the percentage of M1 that is in the dimer form. This can be quantified by integrating the radiolabel counts of the monomer and dimer bands on scanned gel. The total counts (C_T) equal the sum of the counts measured for monomer (C_M) and for dimer (C_D). Thus, the percentage of dimer, *r*, is:

$$r = \frac{C_D}{C_T} = \frac{C_D}{(C_{M1} + C_D)} \quad (3)$$

We can express [D] in terms of T and *r*:

$$[D] = T \cdot \frac{C_D}{(C_{M1} + C_D)} = T \cdot r \quad (4)$$

Finally, we can express K_d (het-dimer) in terms of T, [M2] and r:

Thus,

$$K_d(het-\mathrm{dim}er) = \frac{[M1]\cdot[M2]}{[D]} = \frac{(T-[D])\cdot[M2]}{[D]} = \frac{(T-T\cdot r)\cdot[M2]}{T\cdot r} = \frac{(1-r)\cdot[M2]}{r} \quad (5)$$

To find K_d (het-dimer), we carry out non-linear fitting of Eq. 5 using the experimental data for *r* and [M2].

2. **Homodimer formation**

Written as a dissociation process, we have for homodomerization:

$$D \rightarrow 2 \cdot M, \quad K_d(\mathrm{homo-dimer}) = \frac{[M]^2}{[D]} \qquad (6)$$

The total concentration of M is given by:

$$T = [M] + 2[D] \qquad (7)$$

Typically a fixed and small amount (<0.5 nM) of radiolabeled monomer, M* is mixed with varying amounts of unlabeled monomer M and equilibrium is determined by gel mobility shift analysis. We assume [M*] << [M] and that any dimer formed contains at most one molecule of M*. In other words, there are no significant amounts of dimer consisting of two molecules of M* and so [D] is proportional to the measured counts of M* in the dimer electrophoresis band. Again we have $C_T = C_M + C_D$. Because two monomers combine to make each dimer, $[D]_{max} = T/2$. Combining these equations we can express [D] in terms of T and the measured counts, C_M and C_D or T and *r* where is defined in Eq. 3

$$[D] = \frac{T}{2} \cdot \frac{C_D}{(C_M + C_D)} = \frac{T}{2} \cdot \frac{C_D}{C_T} = \frac{T}{2} \cdot r \quad (8)$$

Finally:

$$K_d(\mathrm{homo-dimer}) = \frac{[M] \cdot [M]}{[D]} = \frac{\left(T - 2[D]\right)^2}{[D]} = \frac{2 * \left(T - T \cdot r\right)^2}{T \cdot r} = \frac{2 \cdot T^2 \cdot \left(1 - r\right)^2}{T \cdot r} = \frac{2 \cdot T \cdot \left(1 - r\right)^2}{r}, \quad (9)$$

To find K_d (homo-dimer), we carry out non-linear fitting of Eq. 9.

3. Homo-trimer formation

We write homo-trimer equilibrium as dissociation directly to monomer units:

$$Tri \rightarrow D + M \rightarrow 3 \cdot M, \; overall \; K_d(\mathrm{homo-trimer}) = \frac{[M]^3}{[Tri]} \quad (10)$$

The total concentration is given:

$$T = [M] + 2[D] + 3[Tri] \quad (11)$$

For cooperatively forming homo-trimers, we do not observe any dimer band so we take [D]= 0 so Eq. 11 reduces to:

$$T = [M] + 3[Tri] \quad (12)$$

As for homo-dimers, we assume that the concentration of labeled M is very small compared to added non-radioactive M, [M*] << [M]. Thus, each trimer contains at most one molecule of M* and C_{Tri} is proportional to the measured counts of the monomer band. Also, the maximum concentration of trimer that can be produced is T/3. We redefine *r* to be the percentage of trimer:

$$r_{Tri} = \frac{C_{Tri}}{(C_M + C_{Tri})} = \frac{C_{Tri}}{C_T} \quad (13)$$

With these assumptions we have:

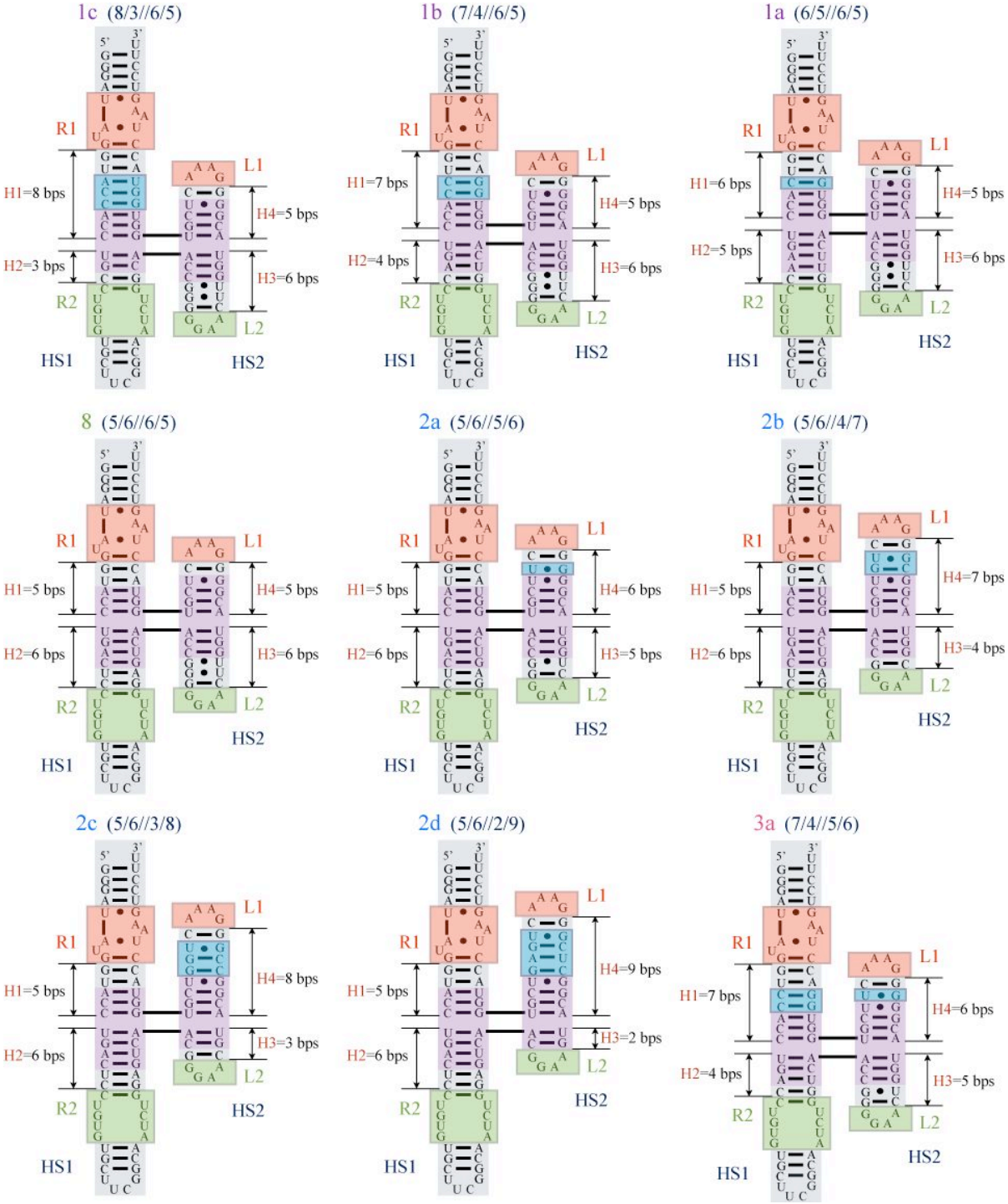
$$[Tri] = \frac{T}{3} \cdot r_{Tri} \quad (14)$$

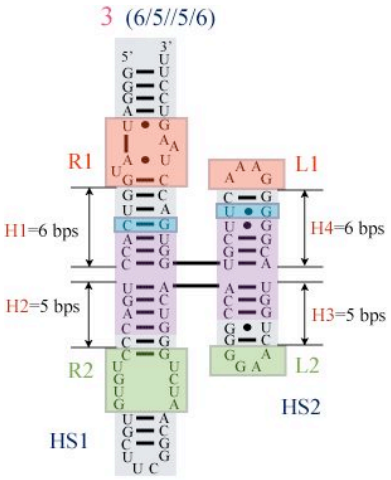
Combining Eq. 10 , 12 and 14 we obtain an expression for K_d (homo-trimer) in terms of T and r_{Tri}:

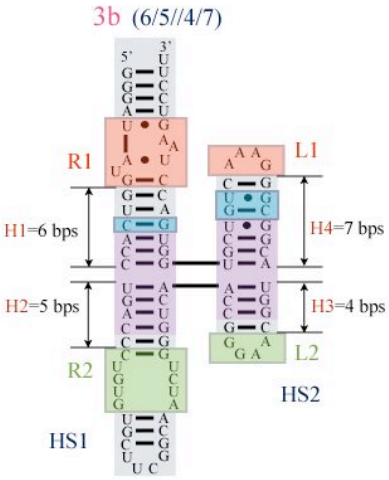
$$K_d = \frac{[M]^3}{[Tri]} = \frac{\left(T - 3 \cdot [Tri]\right)^3}{[Tri]} = \frac{3 \cdot \left(T - 3[Tri]\right)^3}{T \cdot r_{Tri}} = \frac{3 \cdot \left(T - T \cdot r_{Tri}\right)^3}{T \cdot r_{Tri}} = \frac{3 \cdot T^3 \cdot \left(1 - r_{Tri}\right)^3}{T \cdot r_{Tri}} = \frac{3 * [T]^2 * \left(1 - r\right)^3}{r}, \quad (15)$$

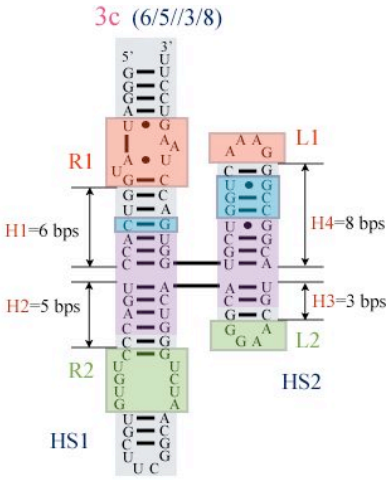
We use non-linear fitting to obtain K_d from T vs *r* data.

Figure S-1. Detailed molecular design of tecto-RNA units used in this work. Color-coding of RNA regions is consistent with their 2D diagrams in Figure 1 and 3.




3 (6/5//5/6)
R1
L1
H1=6 bps
H4=6 bps
H2=5 bps
H3=5 bps
R2
L2
HS2
HS1


3b (6/5//4/7)
R1
L1
H1=6 bps
H4=7 bps
H2=5 bps
H3=4 bps
R2
L2
HS2
HS1


3c (6/5//3/8)
R1
L1
H1=6 bps
H4=8 bps
H2=5 bps
H3=3 bps
R2
L2
HS2
HS1

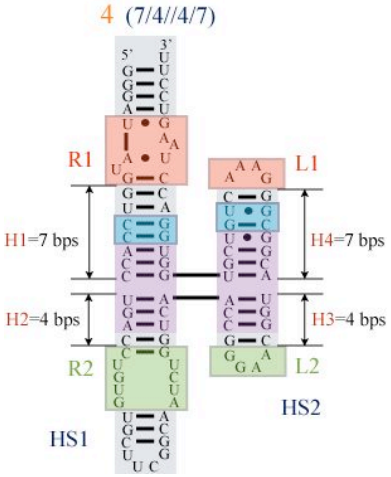

4 (7/4//4/7)
R1
L1
H1=7 bps
H4=7 bps
H2=4 bps
H3=4 bps
R2
L2
HS2
HS1

Figure S-2. Detailed molecular design of hetero-multimeric systems, **8TA-8TB-8TC** and **1bDA-1bDB** and their mutants. Arrows represent tertiary contacts. RNA mutants such as **1bDBmut**, **8TBmut** and **8TCmut** have receptor regions (R3) that are being rotated by 180° from the original design of **1bDB**, **8TB** and **8TC** (highlighted in yellow).

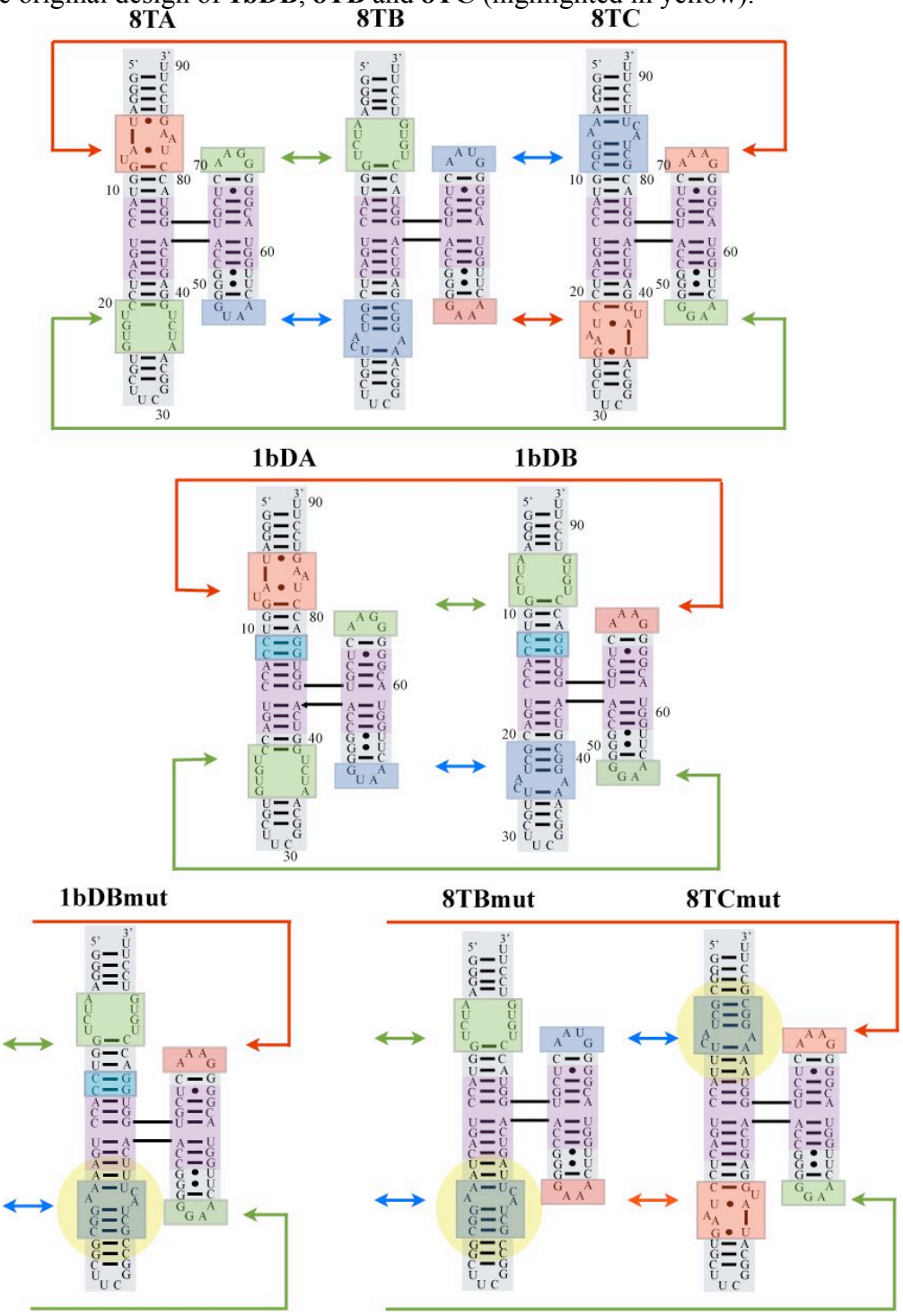


Figure S-3. Dilution of pre-assembled molecule **8*** from 250 nM down to picomolar concentrations. Control RNA - molecule **B** containing the same 4WJ with two GAAA loops and two GAAA-receptors placed on its arms - creates a ladder of multimeric complexes with the dimer as a major product (34).

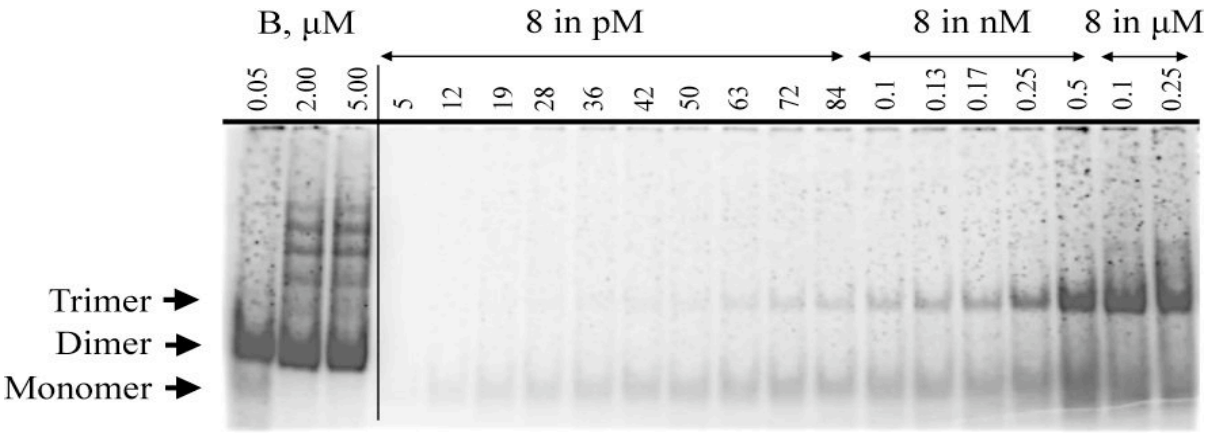


Figure S-4. Exchange kinetics of molecule **8** for the incorporation into pre-assembled complexes. A. Representation of the exchange reaction. From the top-view, each monomer can be visualized as two circles connected via 4WJ crossover. Stars correspond to radiolabel. B. Native gel electrophoresis. Unlabeled molecule **8** was pre-assembled at 1 μM concentration at 30°C in the presence of 15 mM magnesium ions by a standard procedure described in 'Materials and Methods" section of this work. After self-assembly, a small amount of radiolabeled **8*** was added to the mixture and divided into 10 μL aliquots. At certain time intervals, samples were quickly mixed with 10 μL of loading buffer and placed on dry ice. The samples were loaded in reverse order to minimize the experimental error. C. Percentage of trimer versus time. The data was analyzed by nonlinear fitting to a single exponential curve. K_{ex} was found from the following equation: $K_{ex}=\ln 2/t(1/2)$, where t(1/2) is time in seconds at which 50% of the exchange between the free monomers in solution and the units in the closed complexes occur. A pseudo-first order rate constant (K_{ex}) of 3.43×10^{-4} s^{-1}, corresponding to t(1/2)=21±2 min (1260±120s) was obtained by fitting the exchange curve.

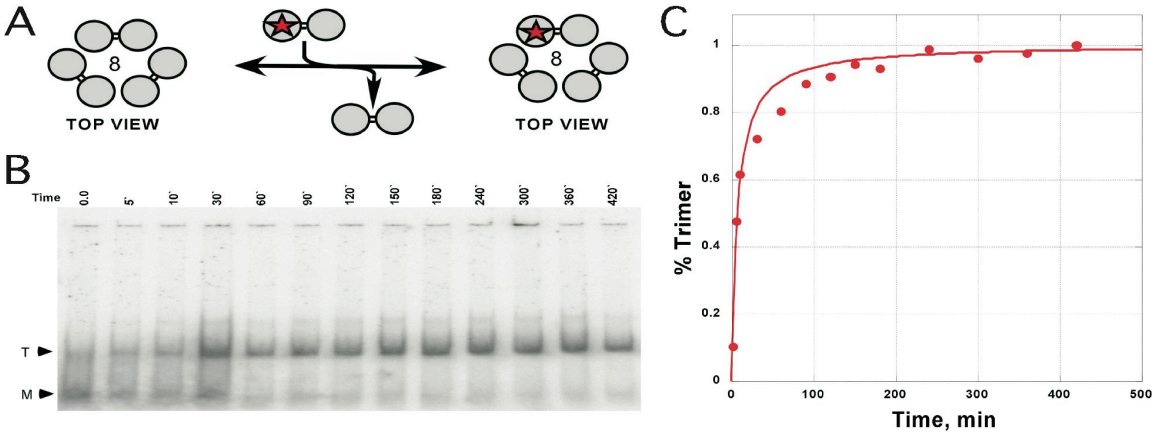


Figure S-5. Schematic diagrams of hetero-multimeric complexes derived from molecule **8** to study the stoichiometry of the complex formation. GAAA-receptor/loop motifs are illustrated in red, green indicates GGAA-receptor/loops, and GUAA-receptor/loop interactions are presented in blue. Various loops and receptors are placed in a way to avoid self-assembly of individual units, and the monomer designs were optimized in a way that HS2 of the n^{th} subunit is complement to the HS1 interaction interface of the first subunit.

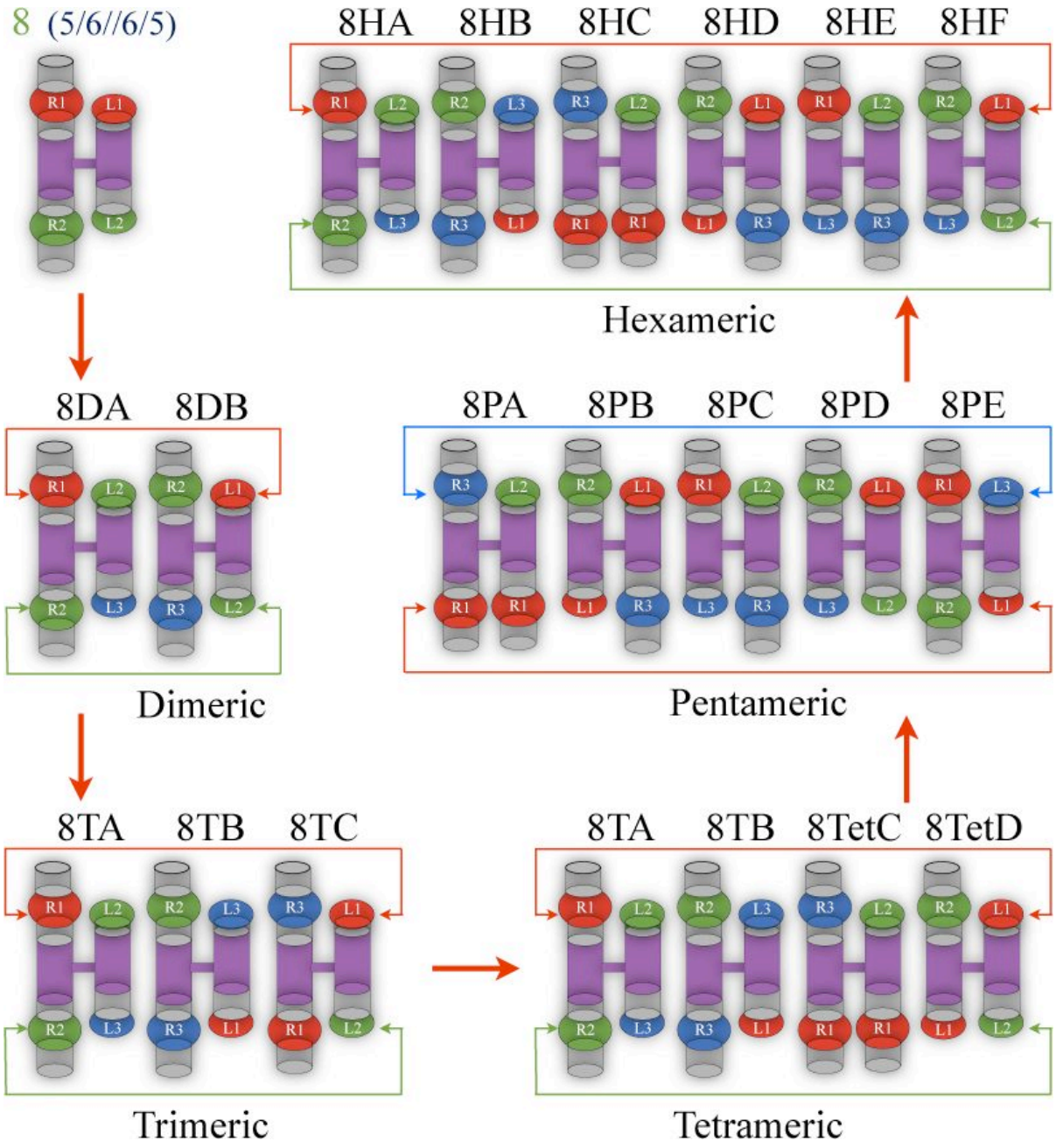


Figure S-6. Native gel analysis to establish the stochiometry of molecule **8**. Radiolabeled molecules are denoted with an asterisk (*). The concentration of molecules is indicated either in the vertical lanes or between parentheses. A. **8DB-8DA** system. Appearance of bands of similar mobility to molecule **8** complex could be explained by the formation of **8DA-8DB*-8DA** or **8DB-8DA*-8DB** complexes. B. Assembly of heterotetrameric, pentameric and hexameric systems. C. Self-assembly of hetero-trimeric complex **8TA-8TB-8TC**.

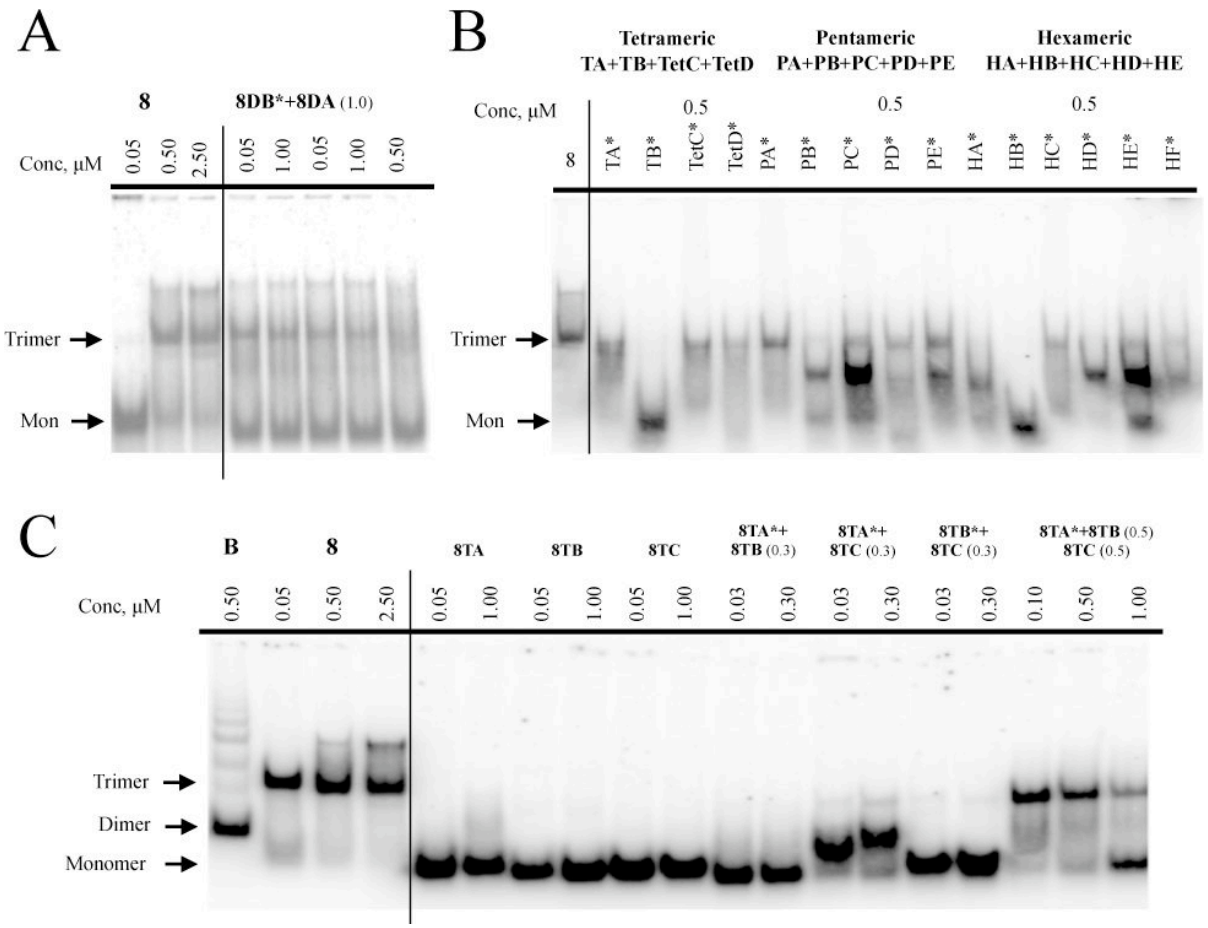


Figure S-7. Comparative native gel study between the unmodified heteromultimeric sets with mutant forms. A. Native gel of hetero-assembly between **8TA*** with **8TBmut and 8TCmut**. B. Native gel electrophoresis of unmodified hetero-dimeric and -trimeric sets. Detailed sequence information can be found in Figure S2.

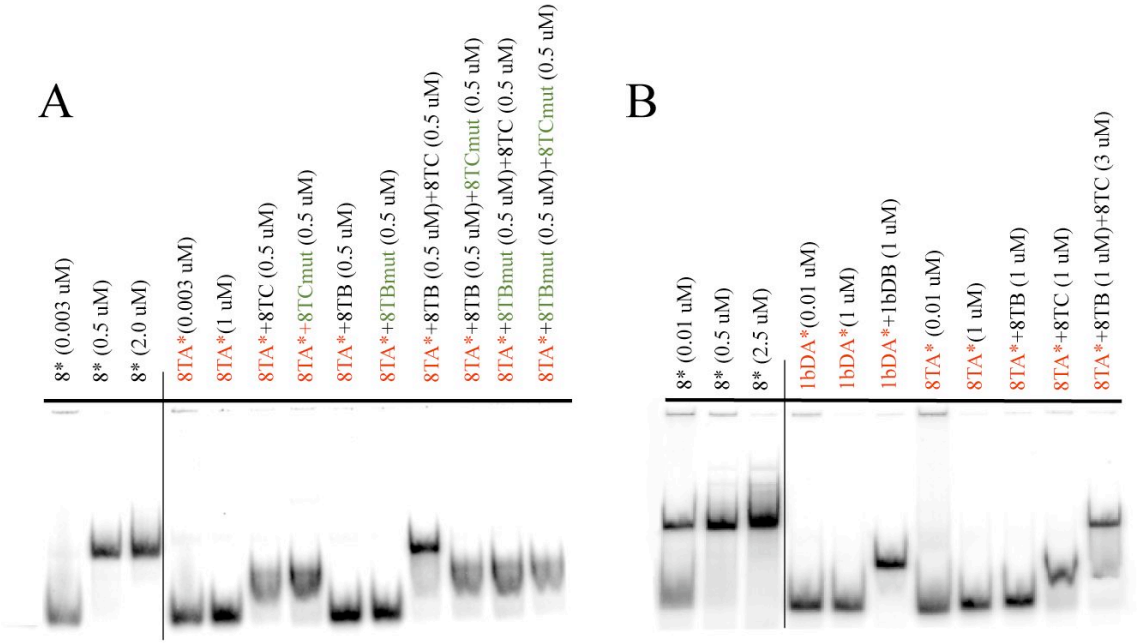


Figure S-8. TEM images of molecule **8**, **2a**, **2b**, **2c** and **2d** complexes. RNA assembly was performed as described in Material and Methods section. The final concentration of RNA in each sample is 2.5 µM. A 3 µl volume of RNA sample was placed on carbon-coated grid purchased from Electron Microscopy Sciences (www.emsdiasum.com) for 2 minutes, washed with increasing concentrations of ethanol (from 25% to 100%) and air-dried. Tungsten-shadowing of the sample was carried under high vacuum at ~30 mA for a period of 10-15 min. The obtained RNA constructs were visualized on the grid on a Transmission Electron Microscope (TEM JEOL100CX) at the Bowling Green State University EM Center. The white scale bar on each TEM image corresponds to 200 nm.

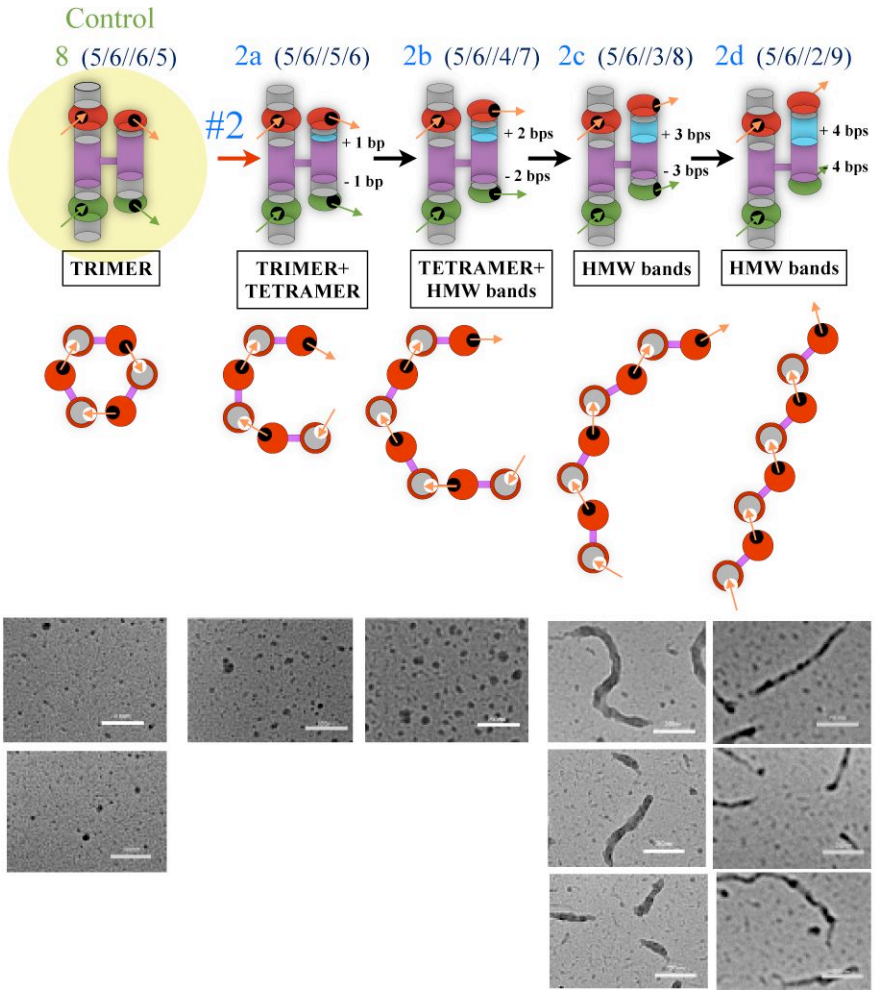


Figure S-9. Denaturing gels (8% acrylamide) of **1bDA*** and **1bDB*** performed to study lead-induced cleavage in the 20-50 nts regions. The areas of interest such as receptor or loop positions are outlined in dotted squares A. Structure probing of radiolabeled **1bDA*** in monomer and dimer states. B. Lead-induced cleavage of radiolabeled **1bDB*** in monomer and dimer states.

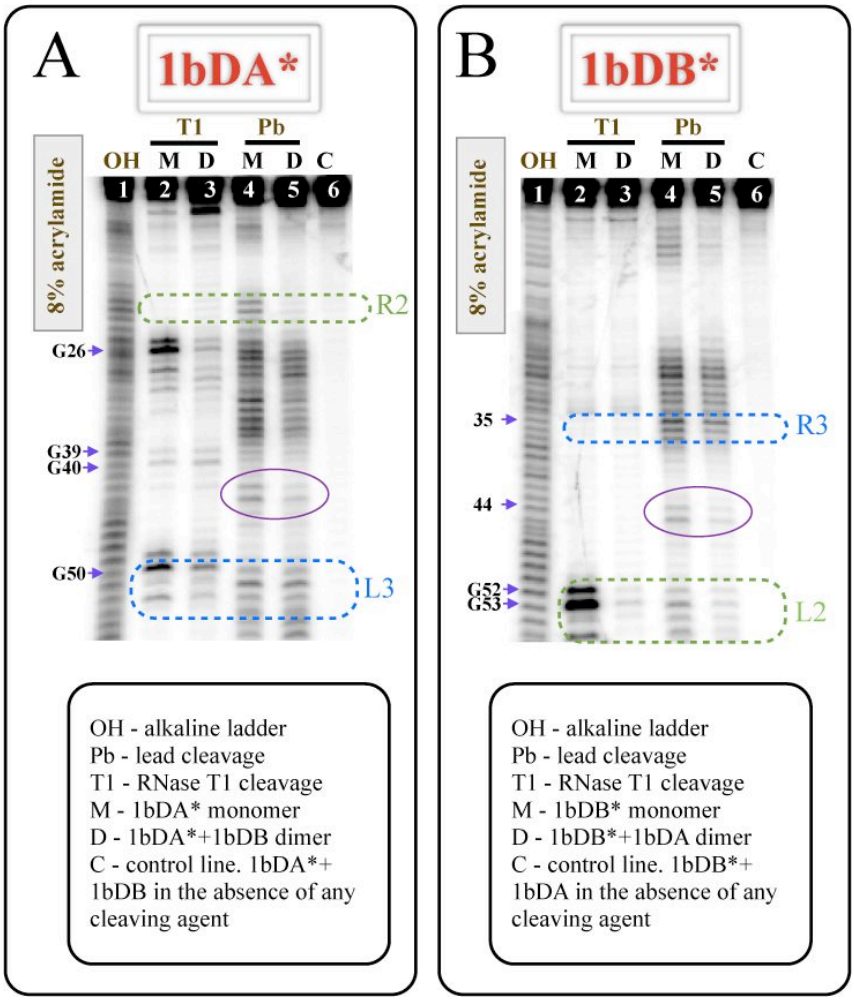


Figure S-10. Dissociation constants measurements. The K_d experiments were performed by titrating a fixed amount of radiolabeled molecule with its partner molecule. In the case of **1b** self-association, fixed amount of **1b*** was titrated with increasing concentrations of cold **1b**. The same procedure was performed to find K_8 for molecule **8** homo-assembly. Dissociation constants were determined as described above in the Supporting Information and in Materials and Methods of the article.