Effect of Hairpin Loop Structure on Reactivity, Sequence Preference and Adduct Orientation of a DNA-Interactive Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Antitumour Agent

David E. Thurston*[1], Higia Vassoler[2], Paul J.M. Jackson[1], Colin H. James[2], Khondaker M. Rahman*[1]

[1]*Institute of Pharmaceutical Science, Britannia House, 7 Trinity Street, London SE1 1DB*

[2]*UCL School of Pharmacy, University College London, 29/39 Brunswick Square, London WC1N 1AX*

Supporting Information

1.0 Determination of Rate Constants

The area under the curves (AUCs) for the decreasing oligonucleotides peaks ((*Seq-1 to seq-8*) throughout the time course study were obtained through the Chromquest software program, and AUCs were obtained to calculate the reaction rate constants.

$$[\mathbf{2}] + [\text{DNA}] \rightarrow [\mathbf{2}/\text{DNA adduct}] \quad \text{(Equation 1)}$$

Determination of the rate constant for the reaction of **2** with each oligonucleotide was based on the stoichiometry for the reaction (Equation 1), which implies that the rate of [DNA] consumption must be equal to the rate of [drug-DNA adduct] formation if no intermediate species are formed. Thus, the instantaneous rate of consumption of DNA (v) at a given time is:

$$v = -\mathrm{d}\,[\text{DNA}]/\mathrm{d}t \quad \text{(Equation 2)}$$

The instantaneous rate of formation of the drug-DNA adduct at a given time is:

$$v = \mathrm{d}[\text{drug-DNA adduct}]/\mathrm{d}t \quad \text{(Equation 3)}$$

It becomes possible to write the following equation:

$$v = -\mathrm{d}[\text{DNA}]/\mathrm{d}t = \mathrm{d}[\text{drug-DNA adduct}]/\,\mathrm{d}t \quad \text{(Equation 4)}$$

The rate law equation that describes the relationship between rate of reaction (v) and the rate constant (k) cannot be predicted from the stoichiometric equation, so it is assumed to adopt the general form described by (Equation 5). To determine the actual rate law equation describing the kinetics of DNA consumption in the reaction (Equation 1), the order of the reaction (a and b, Equation 5) with respect to each one of the reactants involved must be determined experimentally.

$$v = k\,[\mathbf{2}]^{a}.[\mathrm{DNA}]^{b} \quad \text{(Equation 5)}$$

In this instance, extensive experimental methods to determine a and b values were avoided

instead, the "isolation" and "half-life methods" were employed. The order of reaction with respect to [DNA] concentration (b value) was forced into pseudo-first order using the isolation method in order to allow calculation of rate constants. For the "isolation method", the concentration of **2** was considered in excess relative to the concentration of DNA in a 4:1 ratio, and concentration of **2** was assumed to remain relatively constant as adduct formation proceeded. It is possible to assume that the rate of drug-DNA adduct formation was not affected by consumption of **2**, and so the order of reaction with respect to GWL-78 concentration becomes dispensable, and the general rate law (Equation 5) can be written in the form (Equation 6):

$$v = k'\,[\mathrm{DNA}]^{b} \quad \text{(Equation 6)}$$

If, $k' = k\,[\mathbf{2}]^{a}$ and $[\mathbf{2}]_{t(x)} = [\mathbf{2}]_{t(0)}$

The order of reaction with respect to DNA concentration was assumed to be pseudo-first order due to the shape of the exponential decay curve obtained for peak areas plotted against incubation time, which then enable the concentration of [DNA] as a function of reaction time to be established, as described by (Equation 7):

$$\mathrm{d[DNA]}/\mathrm{d}t = -k'[\mathrm{DNA}] \quad \text{(Equation 7)}$$

From this differential equation (Equation 7), the following integral was established:

$$\ln([\mathrm{DNA}]/[[\mathrm{DNA}]_{0}) = -k't \quad \text{(Equation 8)}$$

Therefore: $[DNA]_t = [DNA]_0 e^{-k't}$ (Equation 9)

where *k'* is the rate constant determined from the experimental HPLC measurements.

Therefore, Equation 9 indicates that for pseudo-first order consumption of DNA in the reaction with excess GWL-78, the rate constant for the reaction, *k*', is obtained from the slope of the exponential curve (**Figure S1**).

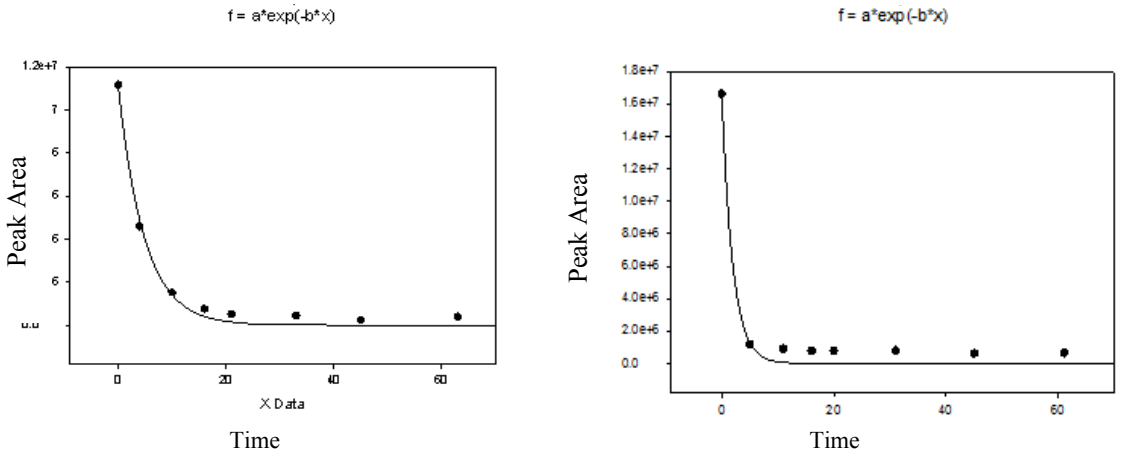
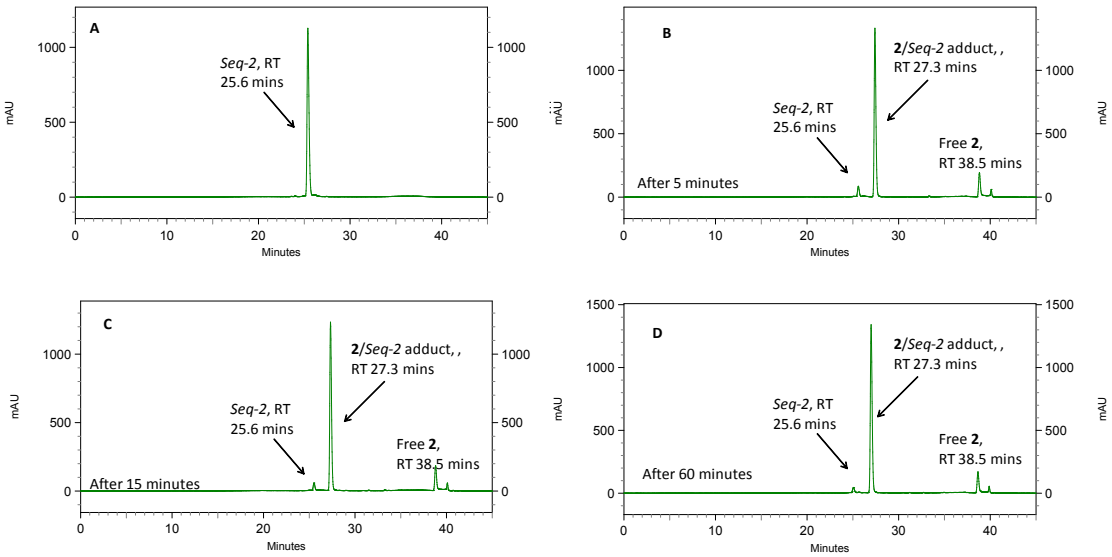


Figure S1: The integrated peak areas for unreacted oligonucleotide plotted against incubation time. The exponential decay curve suggests a pseudo-first-order reaction with respect to oligonucleotide concentration. The slope of the curve varied for different oligonucleotide sequences with faster reactions producing steeper slopes.

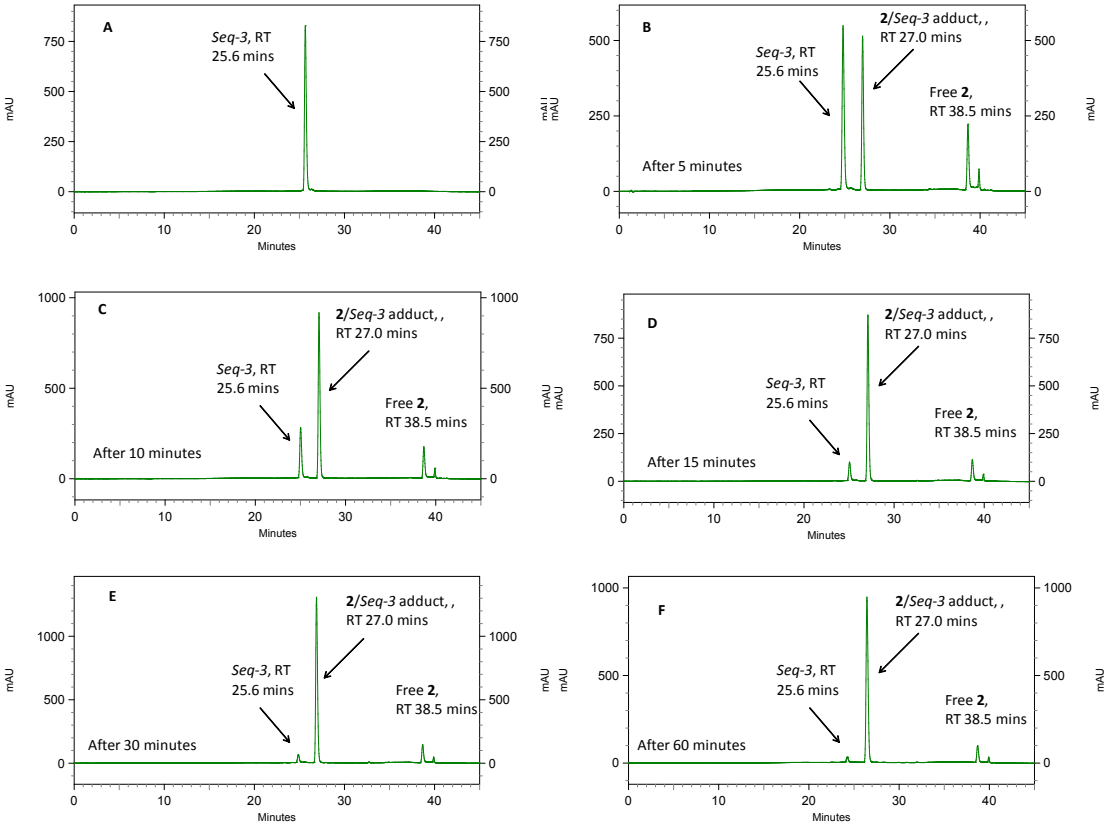
Figure S2: HPLC traces of the time course study for reaction of the 16 different oligonucleotides with 2 (GWL-78)

Please see main text for HPLC trace of time course study of **2** with Seq-1

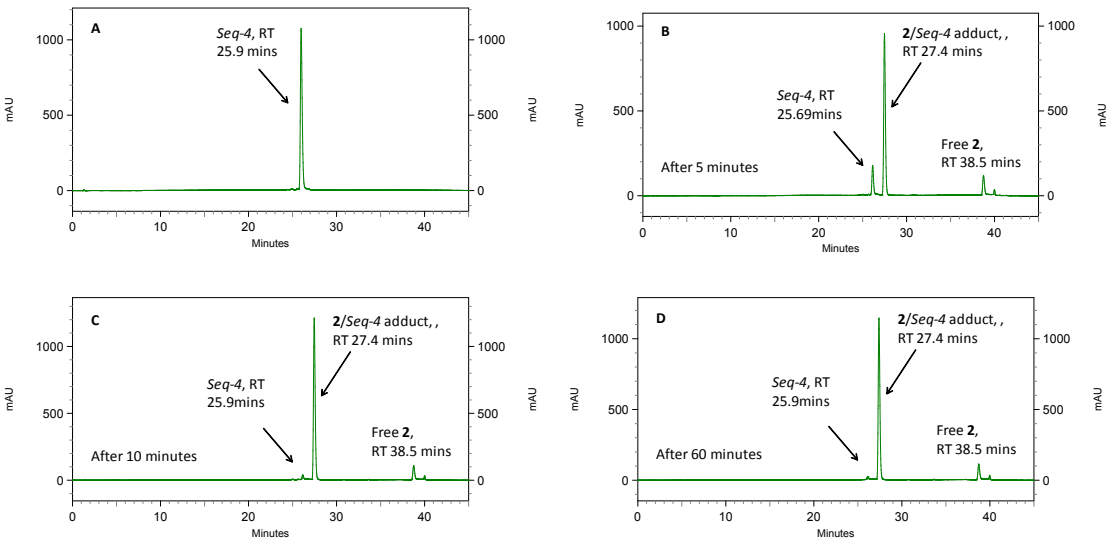
A. Time Course Study of reaction of 2 and *Seq-2*



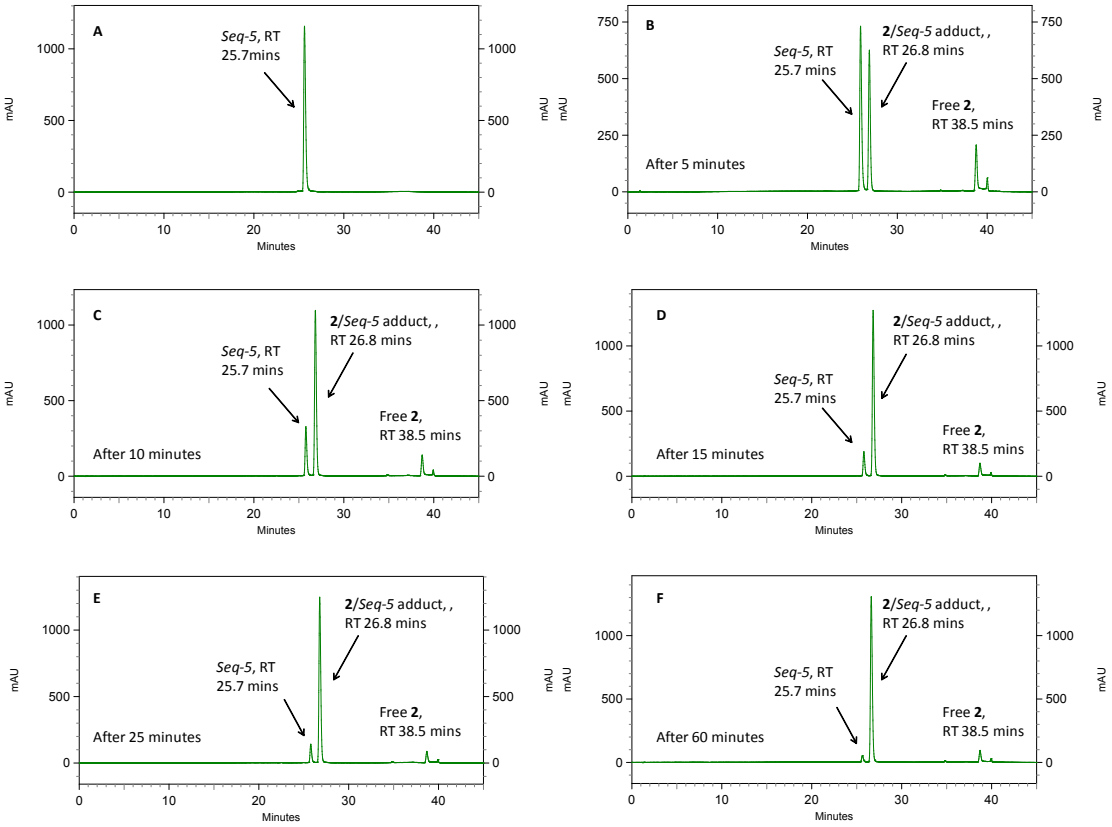
B. Time Course Study of reaction of 2 and *Seq-3*



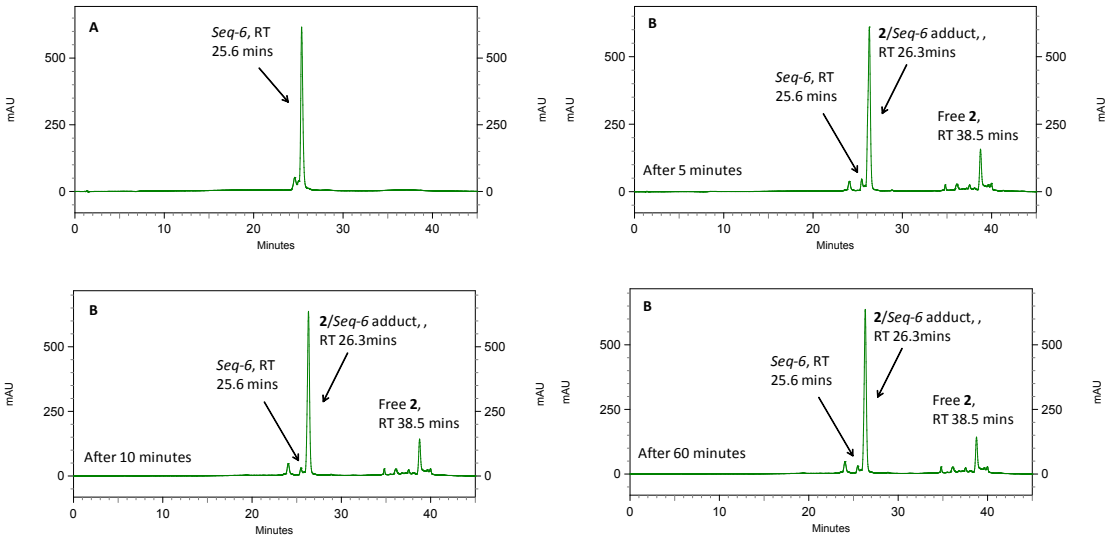
C. Time Course Study of reaction of 2 and *Seq-4*



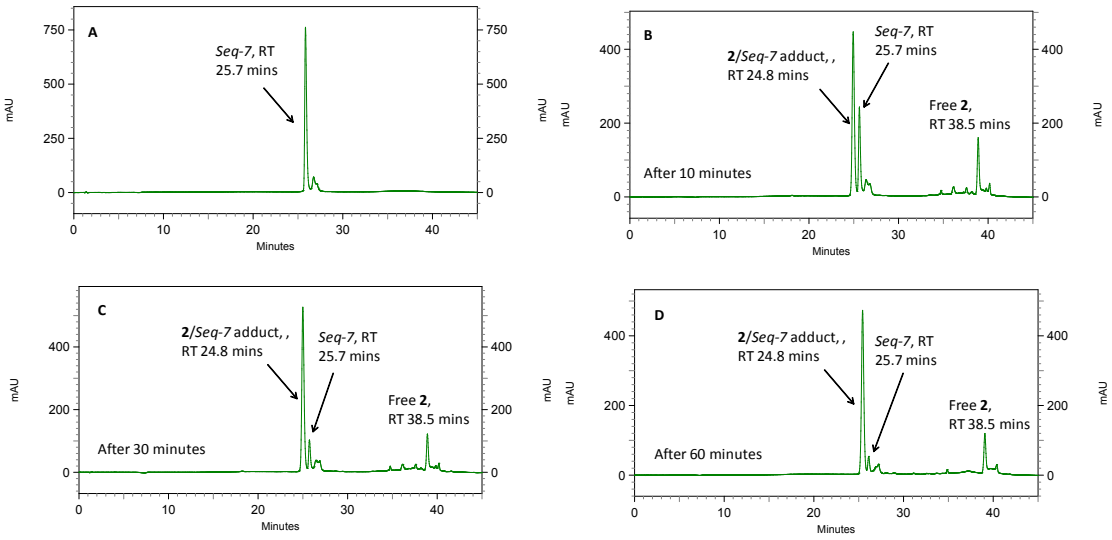
D. Time Course Study of reaction of 2 and *Seq-5*



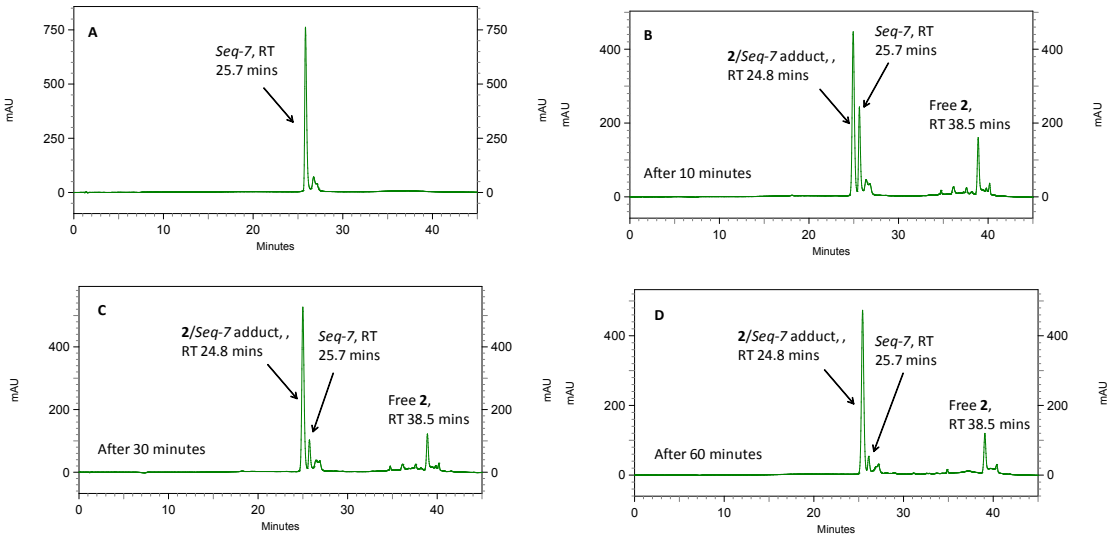
E. Time Course Study of reaction of 2 and *Seq-6*



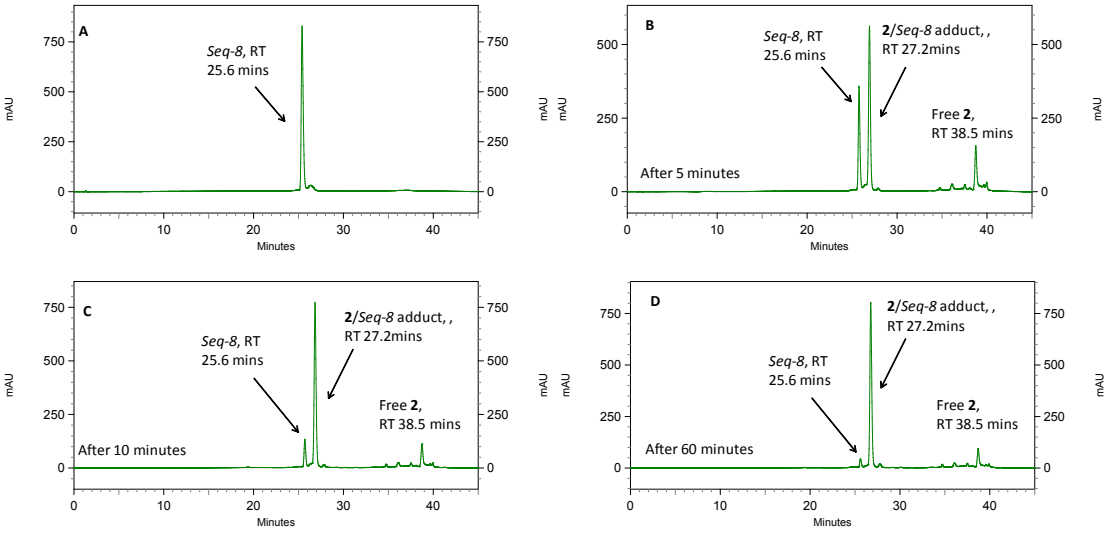
F. Time Course Study of reaction of 2 and *Seq-7*



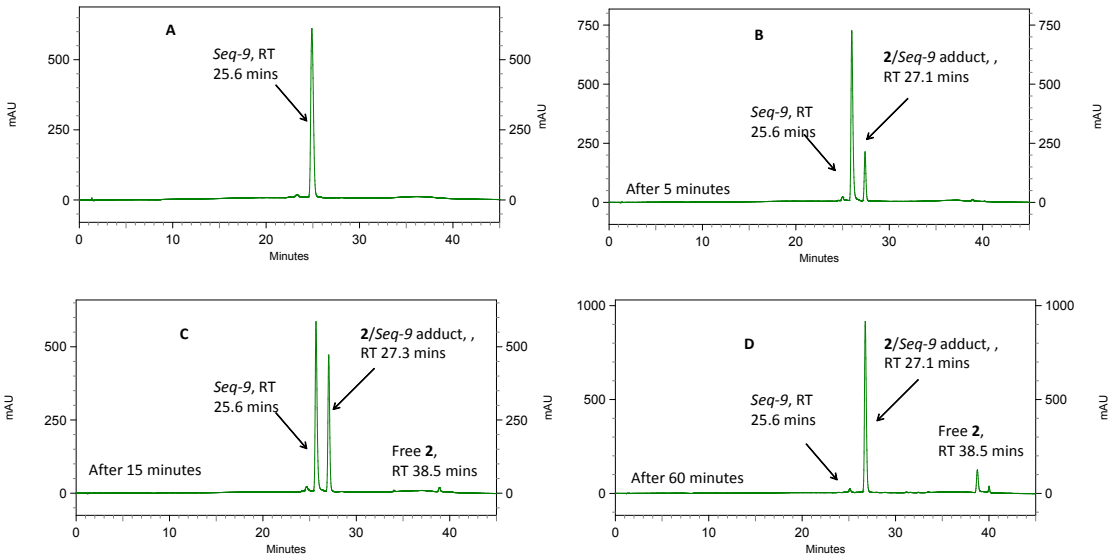
G. Time Course Study of reaction of 2 and *Seq-7*



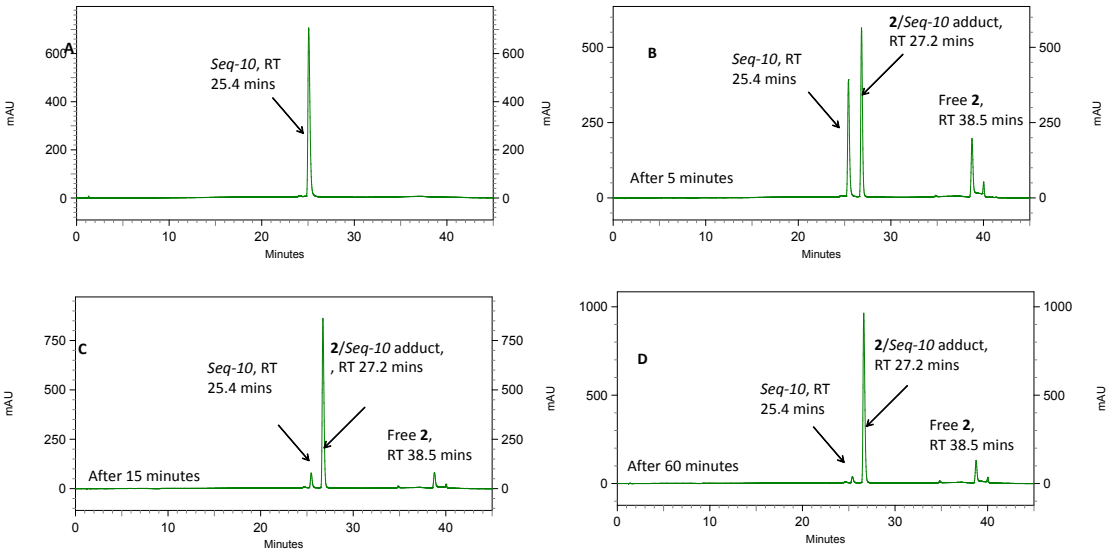
H. Time Course Study of reaction of 2 and *Seq-8*



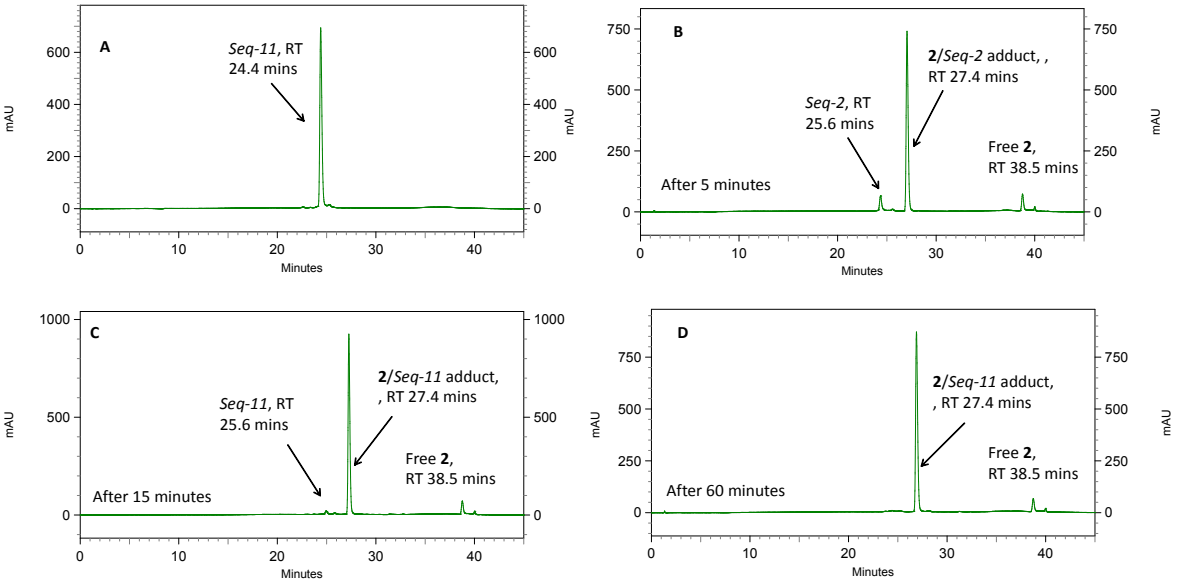
I. Time Course Study of reaction of 2 and *Seq-9*



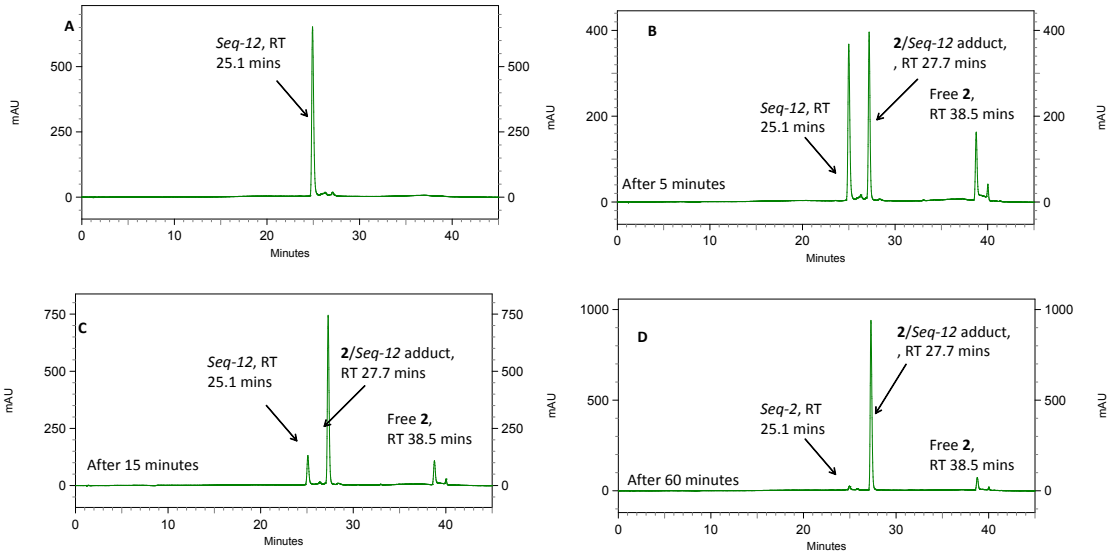
J. Time Course Study of reaction of 2 and *Seq-10*



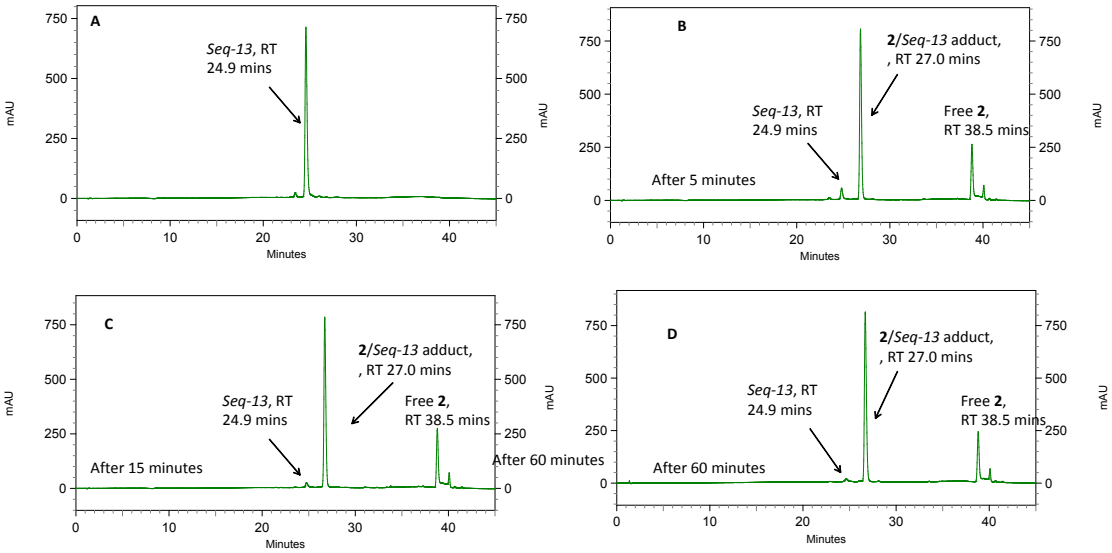
K. Time Course Study of reaction of 2 and *Seq-11*



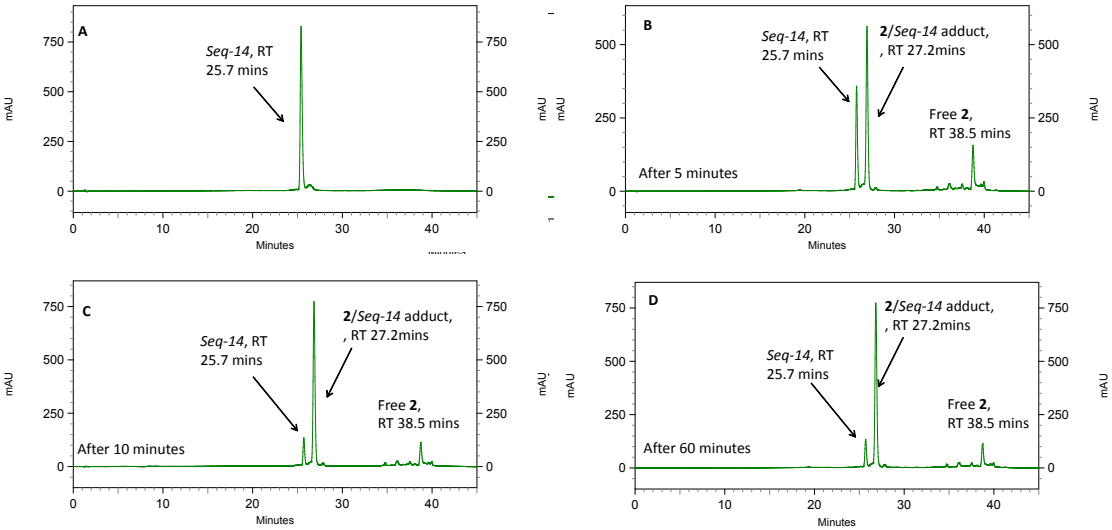
L. Time Course Study of reaction of 2 and *Seq-12*



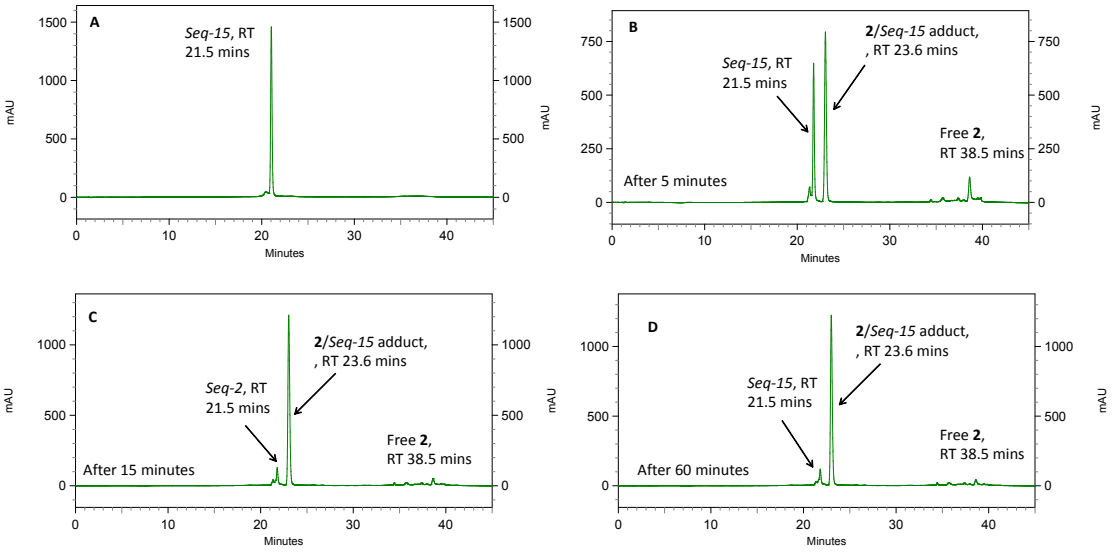
M. Time Course Study of reaction of 2 and *Seq-13*



N. Time Course Study of reaction of 2 and *Seq-14*



O. Time Course Study of reaction of 2 and *Seq-15*



E. Time Course Study of reaction of 2 and *Seq-16*

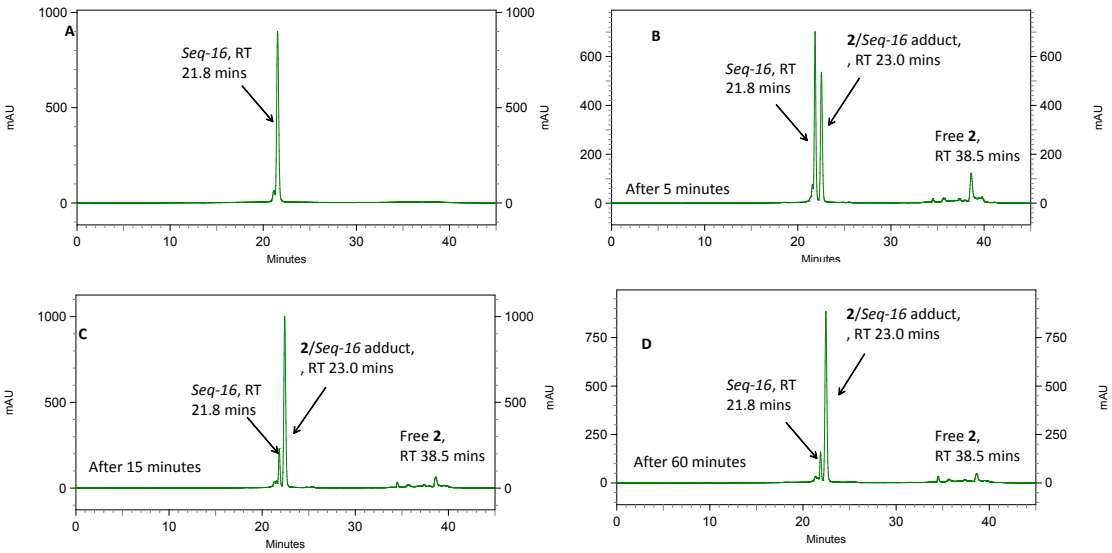
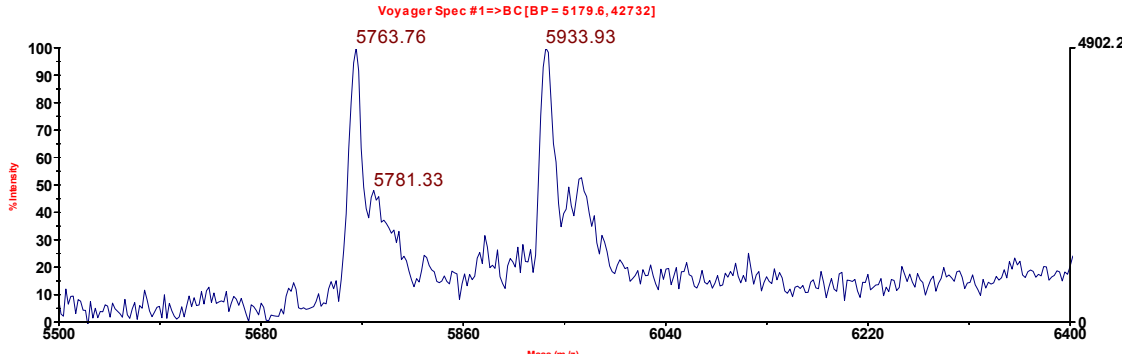


Figure S3: MALDI-TOF Mass spectra of adducts of 2 and the Oligonucleotides (Seq-1 to Seq-16)

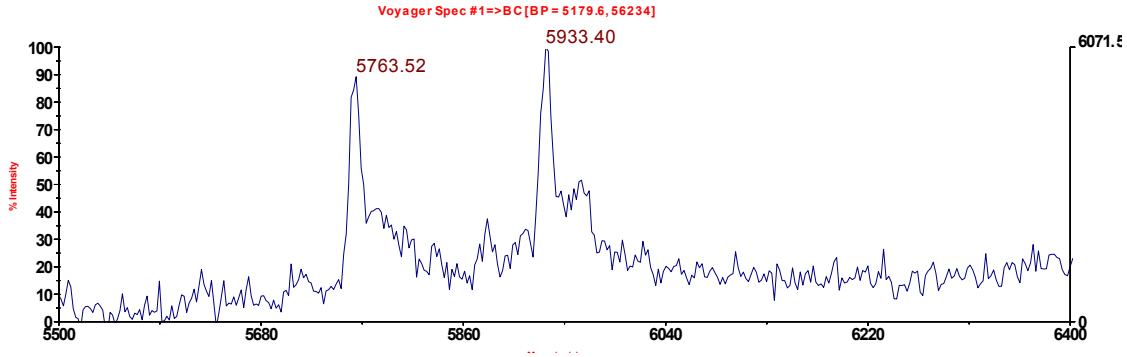
A.

2/*Seq-1* adduct	**Observed Mass**	**5763.76**



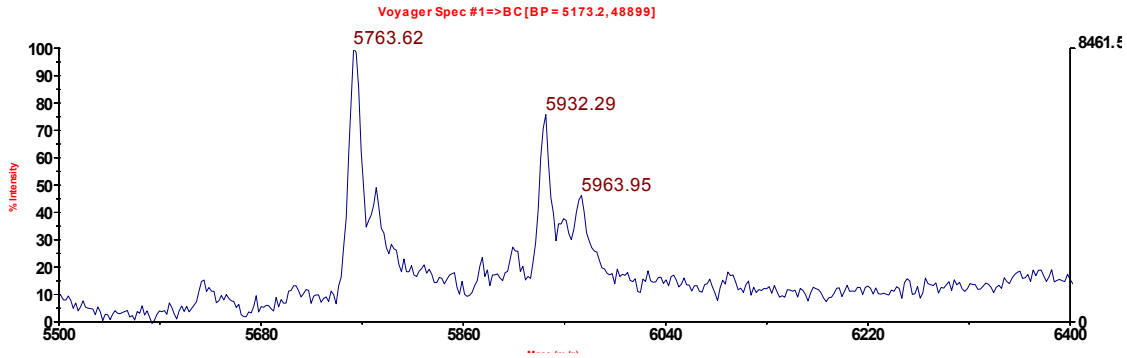
B.

2/*Seq-2* adduct	**Observed mass**	**5763.52**



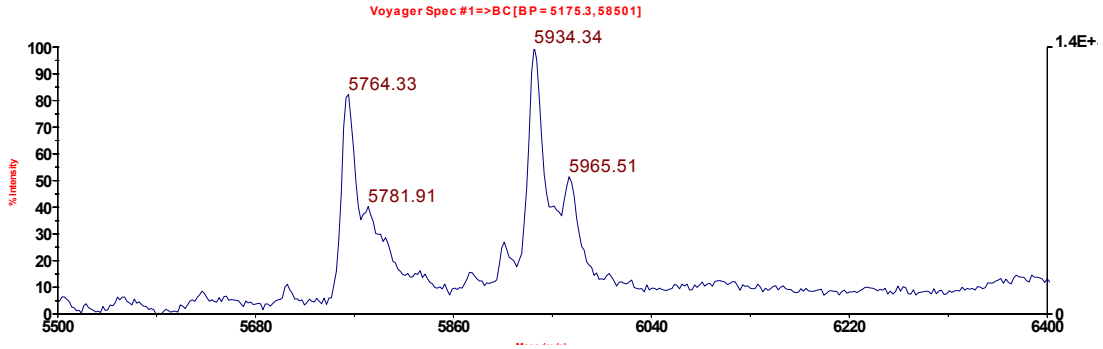
C.

2/*Seq-3* adduct	**Observed Mass**	**5763.62**



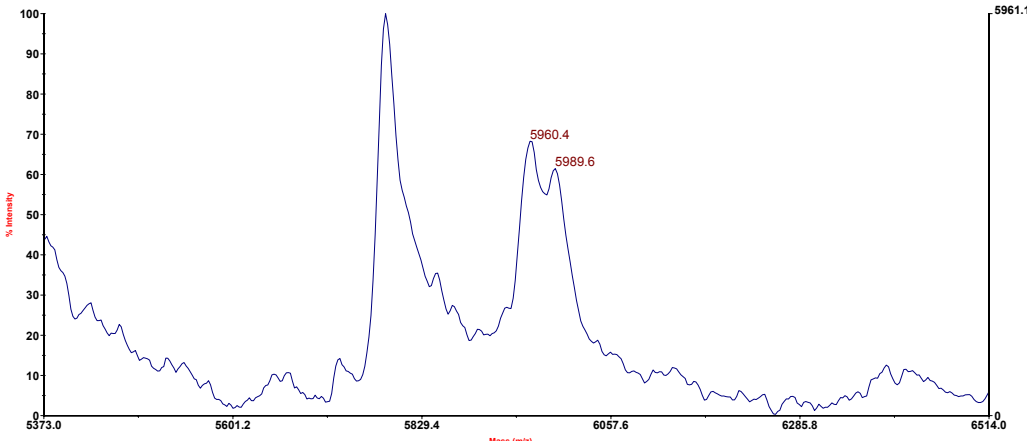
D.

2/*Seq-4* adduct	**Observed Mass**	**5764.33**



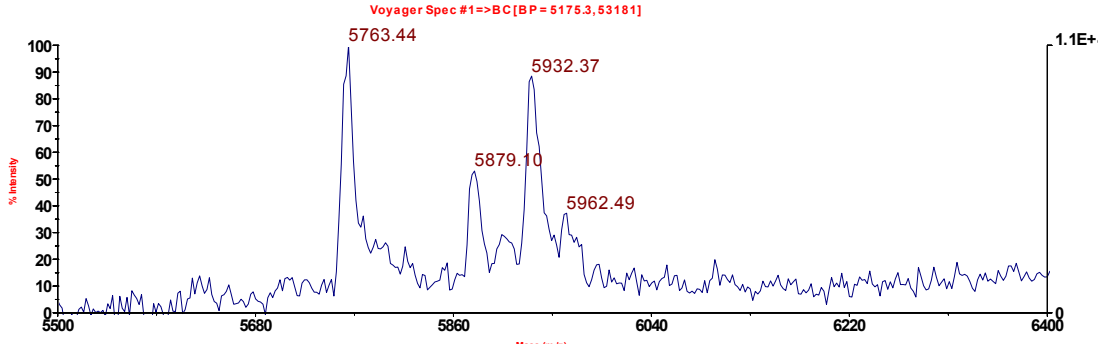
E.

2/*Seq-5* adduct	**Observed Mass**	**5787.9**



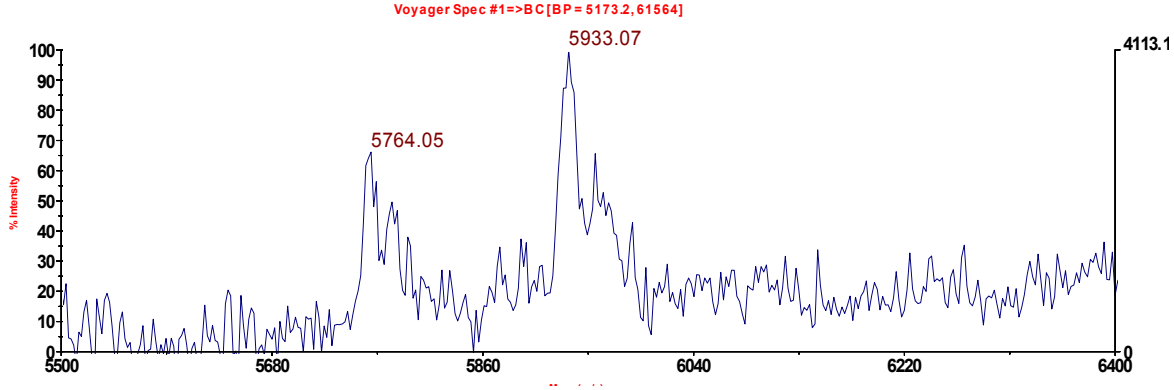
F.

2/*Seq-6* adduct	**Observed Mass**	**5763.44**



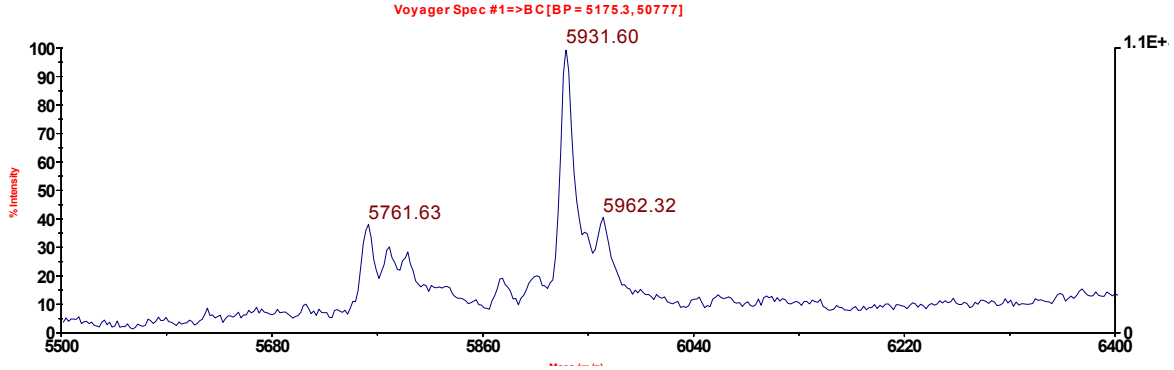
G.

2/*Seq-7* adduct	**Observed Mass)**	**5764.05**



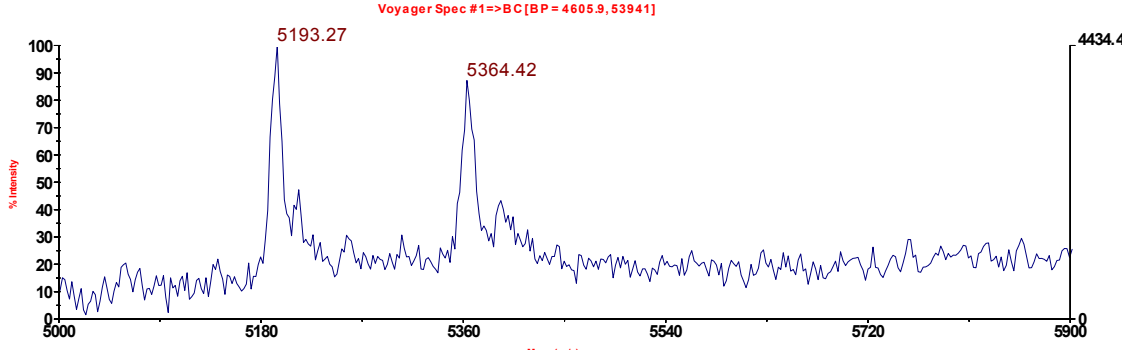
H.

2/*Seq-8* adduct	**Observed Mass**	**5761.63**



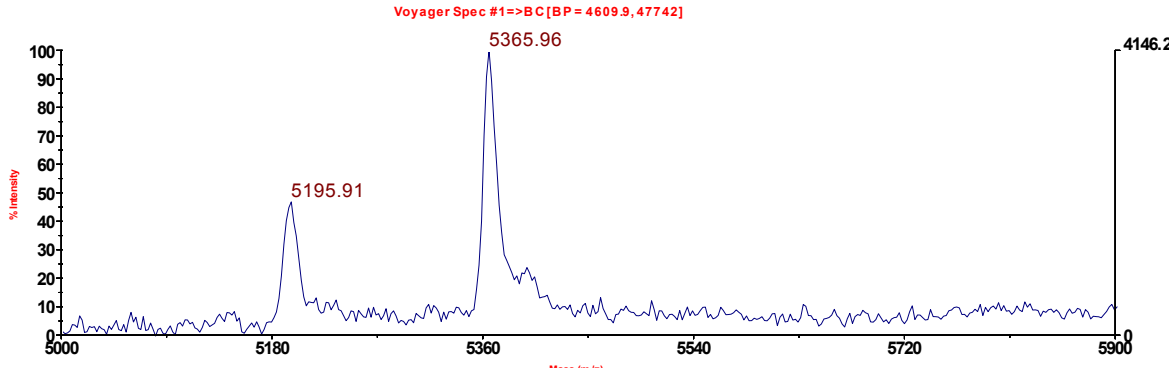
I.

2/Seq-9 adduct	**Observed Mass**	**5193.27**



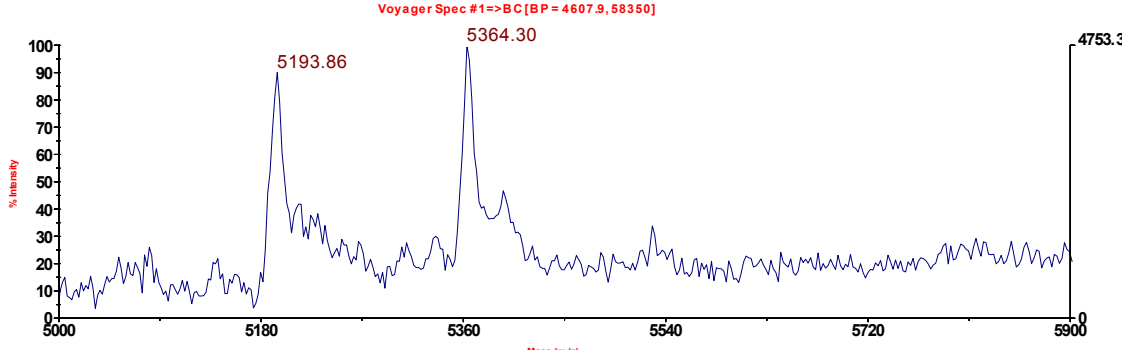
J.

2/Seq-10 adduct	**Observed Mass**	**5195.91**



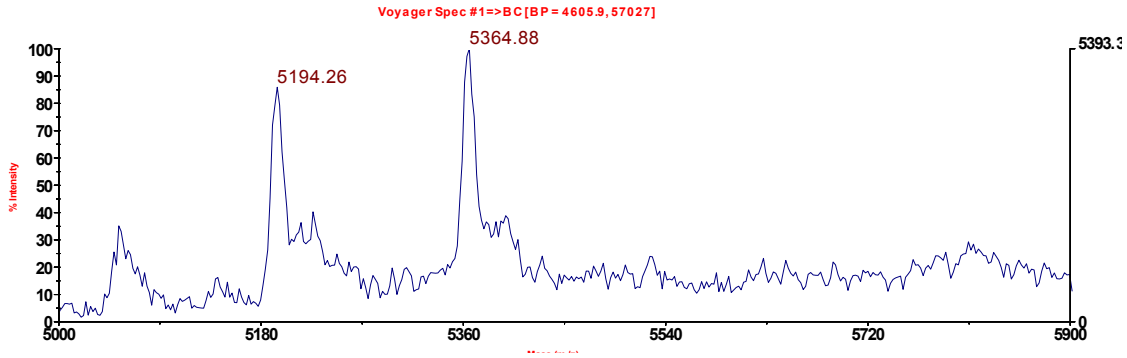
K.

2/Seq-11 adduct	**Observed Mass**	**5193.86**



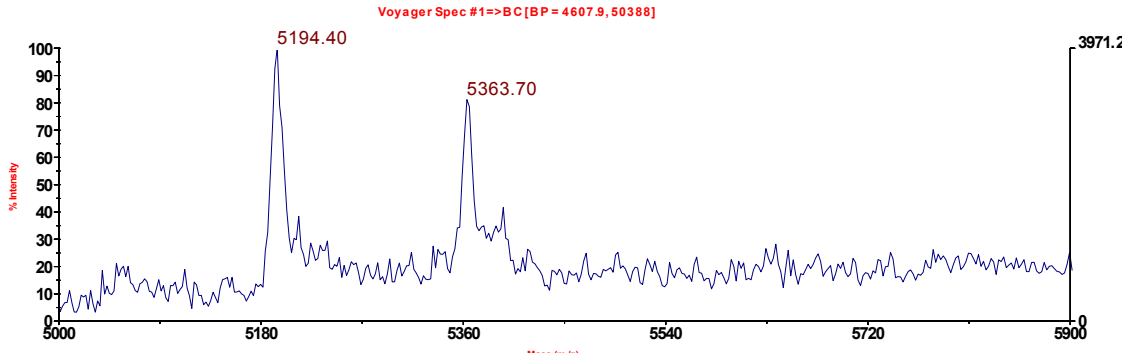
L.

2/Seq-12 adduct	**Observed Mass**	**5194.26**



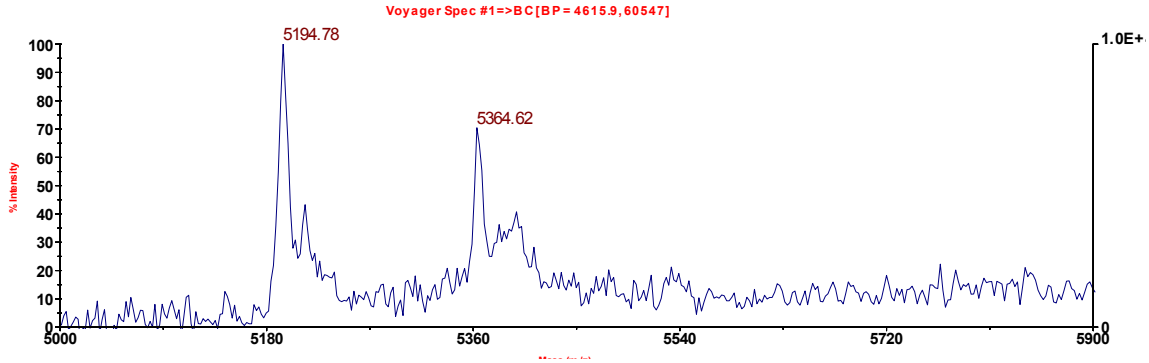
M.

2/Seq-13 adduct	**Observed Mass**	**5194.40**



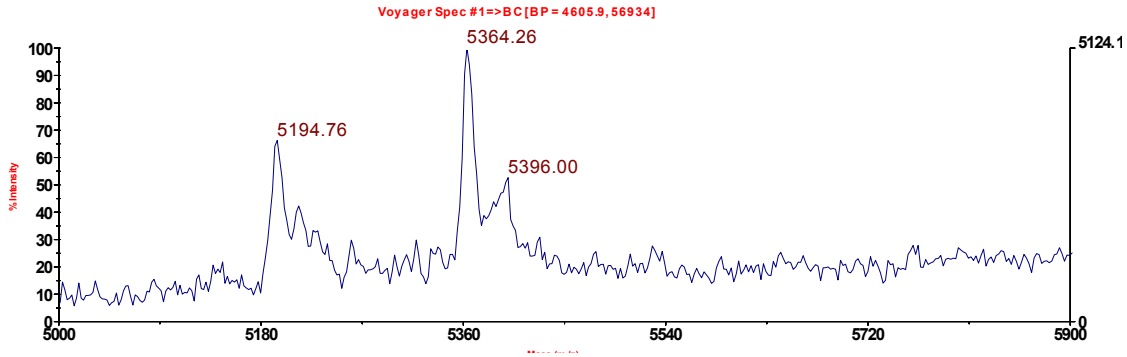
J.

2/Seq-14 adduct	**Observed Mass**	**5194.78**



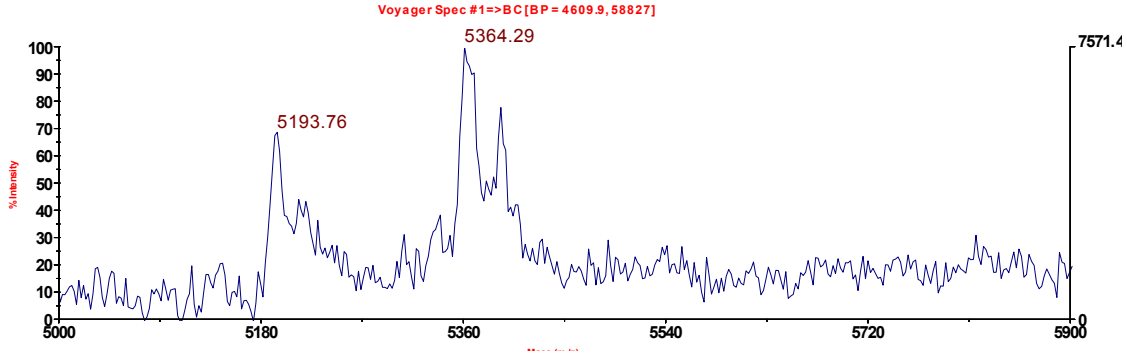
K.

2/Seq-15 adduct	**Observed Mass**	**5194.76**



L.

2/Seq-16 adduct	**Observed Mass**	**5193.76**



Molecular Modelling

Figure S4: *Plot of the simulation of GWL-78 (**2**) covalently bound to G2 showing the formation of a consistent hydrogen bond (black) between N10-H of the ligand and carbonyl of an adjacent thymine T3 of Seq 2 (5-TGT-TATA-TTT-TATA-ACA-3'). Distances vary between 3 and 5Å for the hydrogen bond observed.*

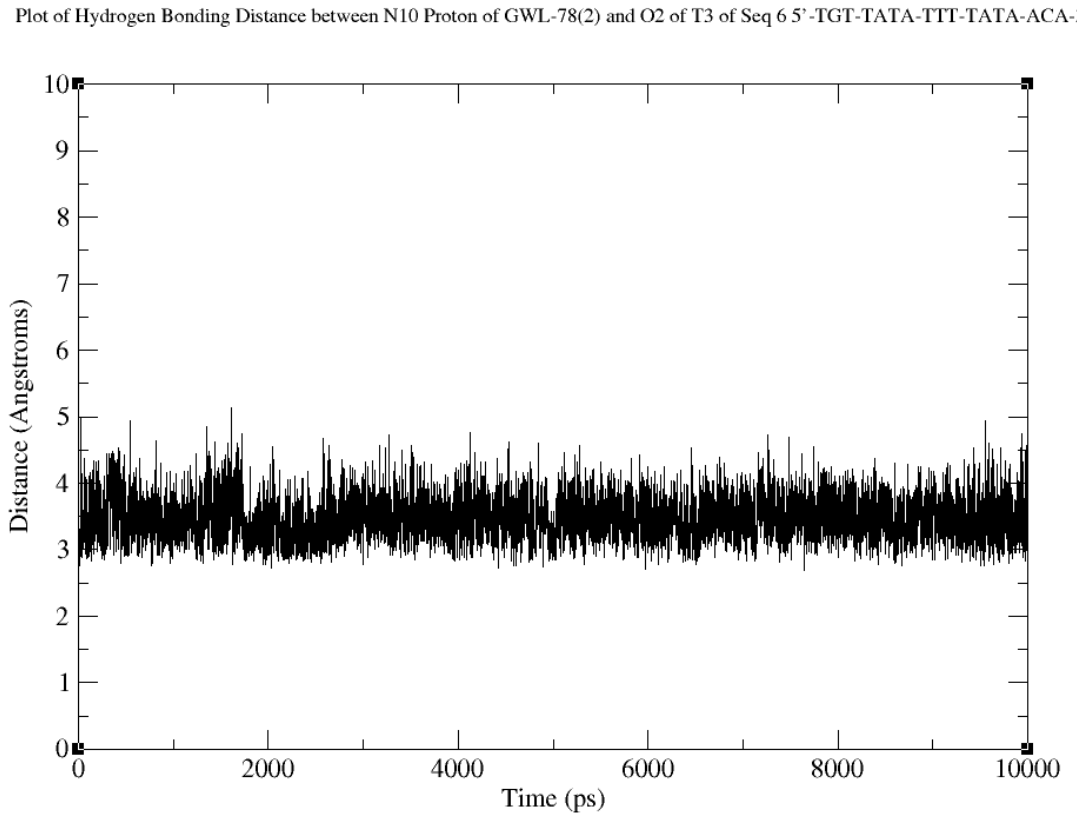


Figure S5: *Plot of the simulation of GWL-78 (**2**) covalently bound to G6 (oriented in A-ring 5' direction) showing the formation of a consistent hydrogen bond (black) between N10-H of the ligand and ring nitrogen of an adjacent thymine A13 of Seq 2 (5-TATA-TGT-TTT-ACA-TATA-3'). Distances vary between 3 and 5Å for the hydrogen bond observed.*

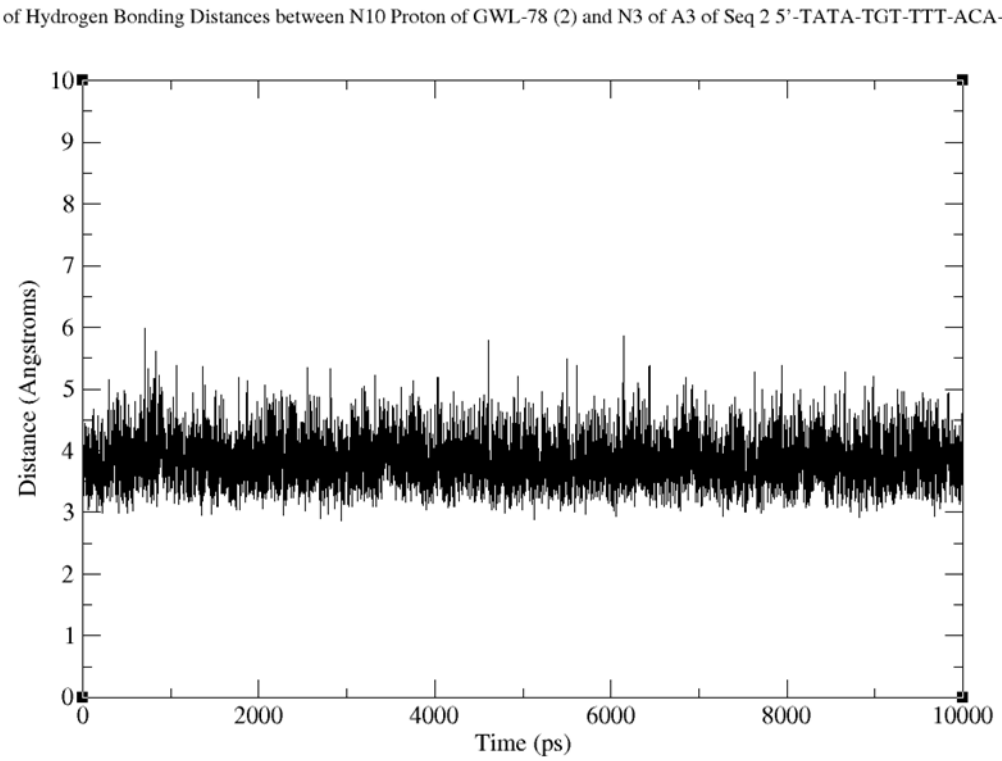


Figure S6: *RMSD plot of GWL-78 (2) covalently attached to G2 (in A-ring 3' orientation) of Seq 6 (5'-TGT-TATA-TTT-TATA-ACA-3') showing the simulation was at equilibrium for its duration.*

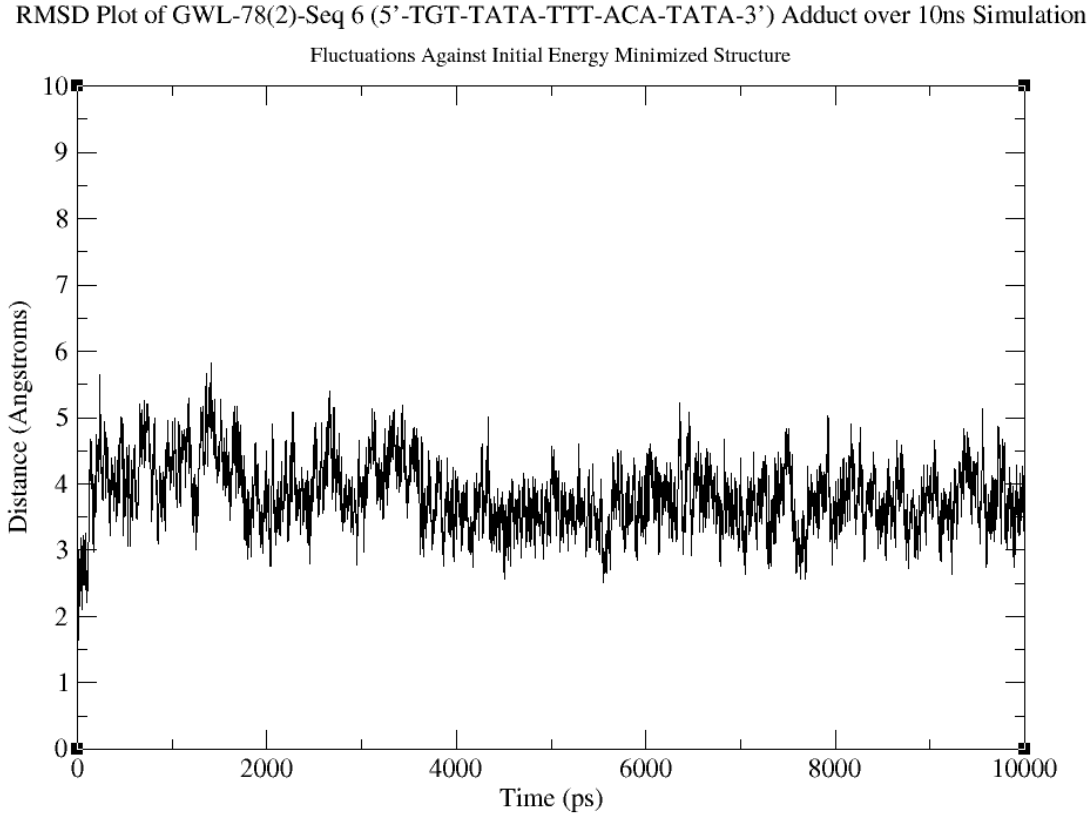


Figure S7*: RMSD plot of GWL-78 (2) covalently attached to G6 (in A-ring 5' orientation) of Seq 2 (5'-TATA-TGT-TTT-ACA-TATA-3') showing the simulation was at equilibrium for its duration.*

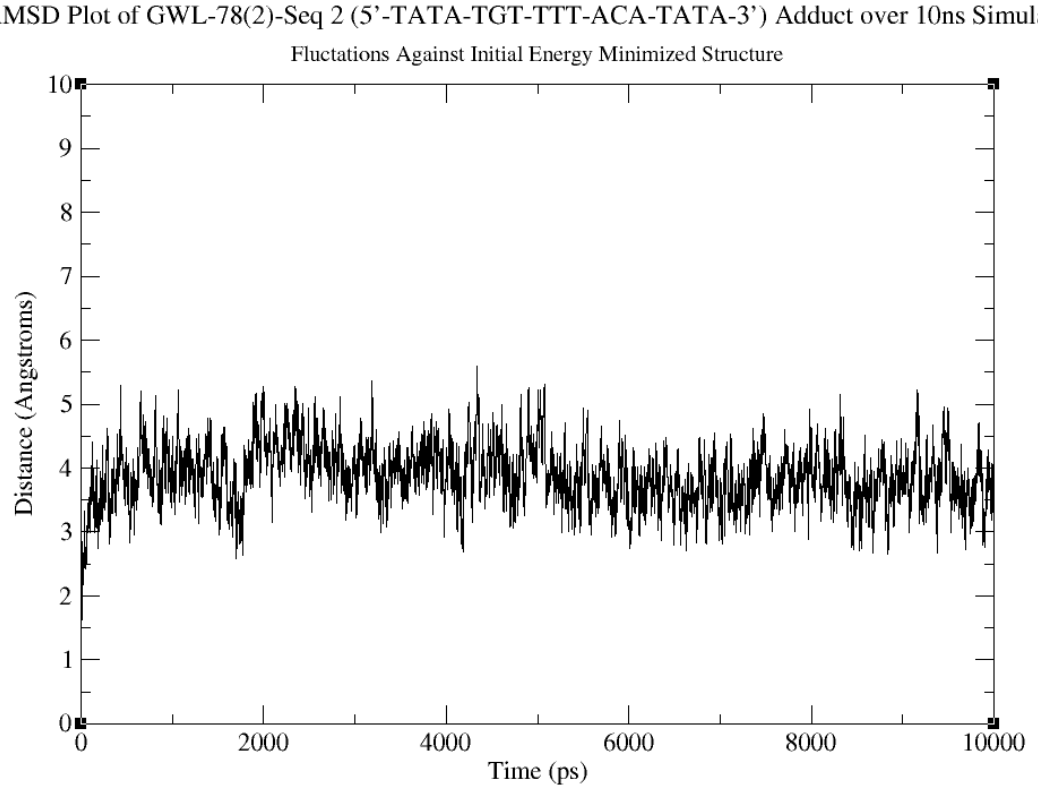


Figure S8: *RMSD plot of GWL-78 (2) covalently attached to G2 (in A-ring 3' orientation) of Seq 16 (5'-TGA-TATA-HEG-TATA-ACT-3') showing the simulation was at equilibrium for its duration.*

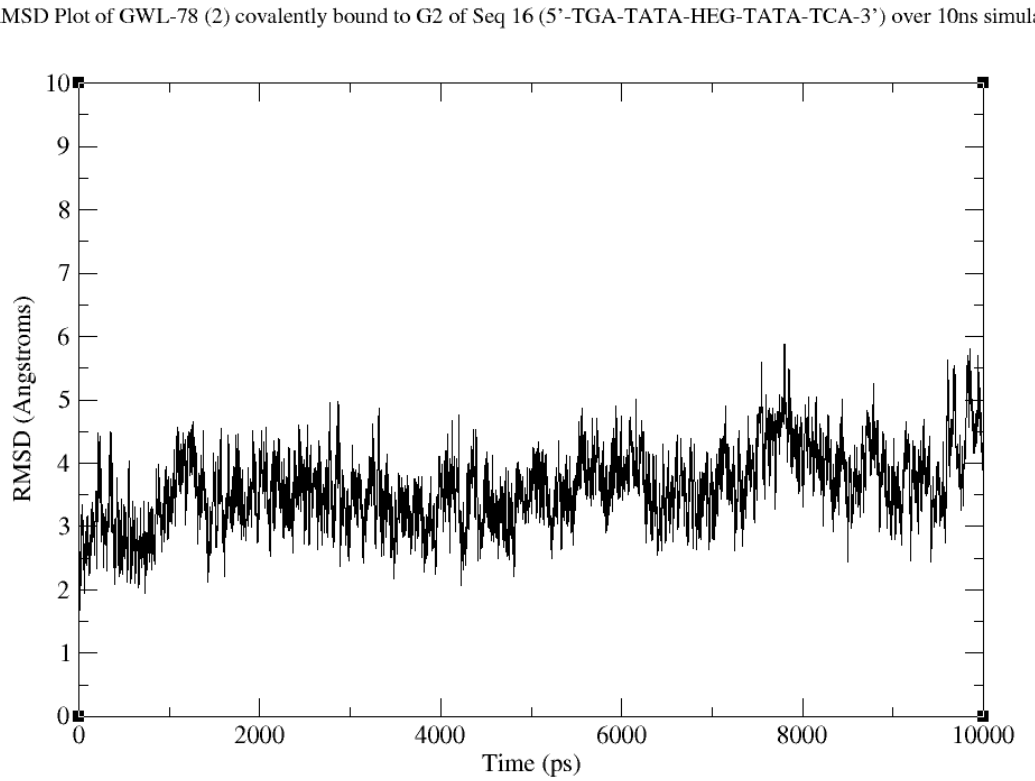


Figure S9: *RMSD plot of GWL-78 (2) covalently attached to G6 (in A-ring 3' orientation with the polyamide moiety traversing the loop) of Seq 10 (5'-TATA-TGT-HEG-ACA-TATA-3') showing the simulation was at equilibrium for its duration. Although at equilibrium, the RMSD varies between 2 and 6.5Å suggesting large variations in conformation of the complex.*

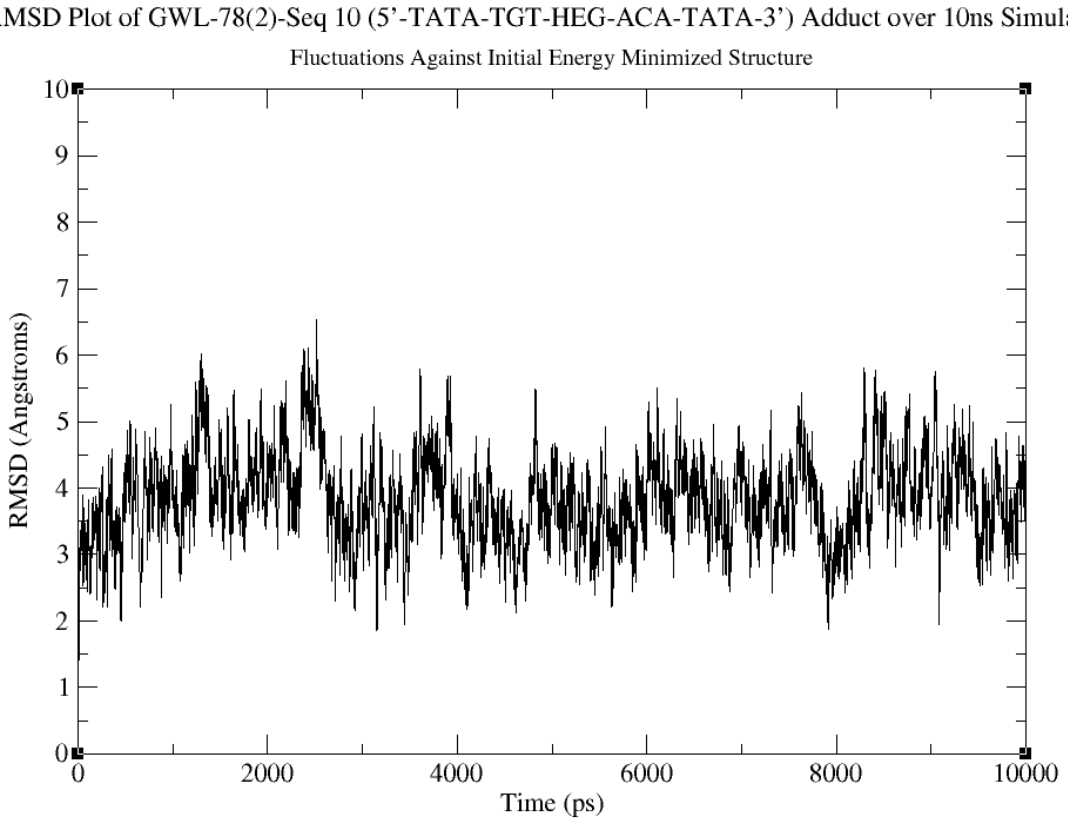


Figure S10: *Plot of hydrogen bonding interactions in the simulation of GWL-78 (**2**) covalently bound to G6 (oriented in A-ring 3' direction) with the polyamide chain traversing the flexible HEG loop. The graph shows the formation of a transient hydrogen bond (black) between the amide NH of the first pyrrole of the ligand and oxygen of the HEG loop of Seq 10 (5-TATA-TGT-HEG-ACA-TATA-3'). Distances vary between 3 and 12Å for the interaction, indicating the flexibility of the loop.*

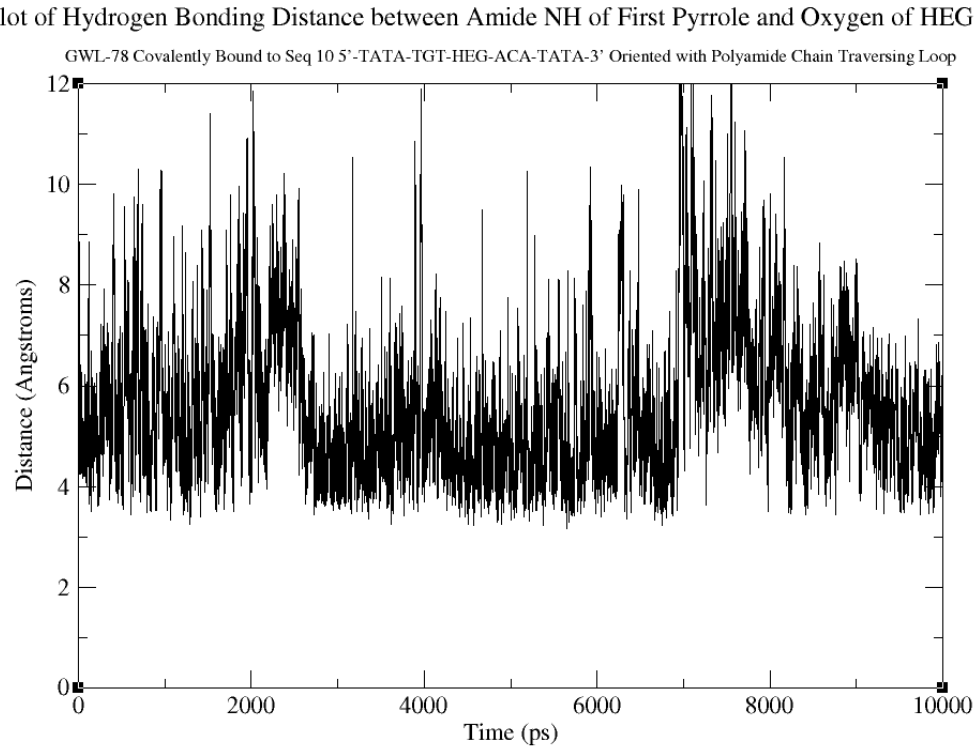


Figure S11: *RMSD plot of GWL-78 (2) covalently attached to G2 (in A-ring 3' orientation) of Seq 6 (5'-TGT-TATA-TTT-TATA-ACA-3') showing the simulation was at equilibrium for its duration.*

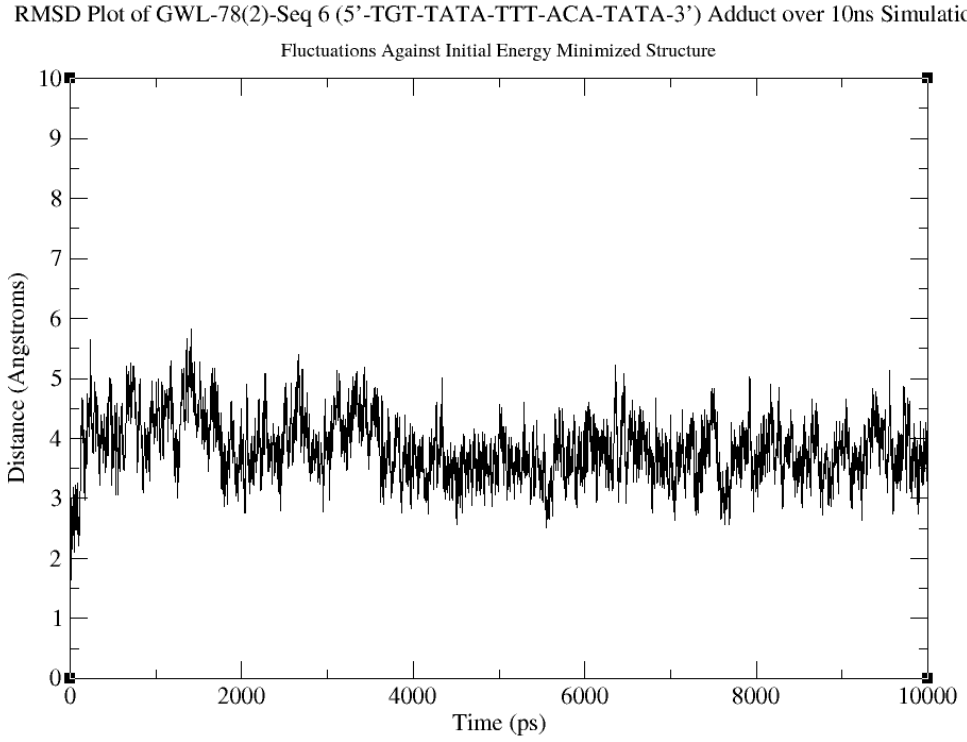


Figure S12: *Plot of the simulation of GWL-78 (**2**) covalently bound to G2 showing the formation of a consistent hydrogen bond (black) between N10-H of the ligand and carbonyl of an adjacent thymine T3 of Seq 14 (5-TGT-TATA-HEG-TATA-ACA-3'). Distances vary between 3 and 5Å for the hydrogen bond observed.*