Programmable Nanoscaffolds That Control Ligand Display to a G-Protein-Coupled Receptor in Membranes To Allow Dissection of Multivalent Effects

Andrew V. Dix¹, Steven M. Moss¹, Khai Phan¹, Travis Hoppe², Silvia Paoletta¹, Eszter Kozma¹, Zhan-Guo Gao¹, Stewart R. Durell³, Kenneth A. Jacobson¹, Daniel H. Appella¹*

AUTHOR ADDRESS
¹Laboratory of Bioorganic Chemistry, NIDDK, NIH
²Laboratory of Biochemistry and Genetics, NIDDK, NIH
³Laboratory of Cell Biology, CCR, NCI, NIH

Table of Contents

Materials and Instrumentation.............................................................................................................................................S3
Abbreviations: .....................................................................................................................................................................S3
General Synthesis and Purification Procedures .............................................................................................................S4
PNA Oligomer Structures and Characterization Data.......................................................................................................S6
L-PNA Duplexes ..................................................................................................................................................................S13
Quantification of PNA Oligomer Conjugates....................................................................................................................S13
General Annealing Condition for Formation of PNA:DNA or PNA:PNA Duplexes.....................................................S13
LCMS Analysis of L-PNA:PNA Duplex .............................................................................................................................S13
Experimental Procedures ................................................................................................................................................S15
Cell cultures and membrane preparation .......................................................................................................................S15
Competitive Radioligand Binding to A2A Receptors.......................................................................................................S15
Fluorescent ligand binding experiments with flow cytometry (FCM) ........................................................................S16
Cyclic AMP Accumulation Assay.....................................................................................................................................S18
Molecular Modeling of a PNA duplex bound to an A2AAR homodimer......................................................................S18
MD Simulation of the Complex: ......................................................................................................................................S21
**Materials and Instrumentation**

Commercial-grade reagents and solvents were used without further purification except as indicated. Boc-protected aegPNA monomers were purchased from PolyOrg, Inc. (Leominster, MA, USA). HMBA Resin, 100-200 mesh, 1% DVB was obtained from Advanced Chemtech (Louisville, KY, USA). Boc-mini-PEG was purchased from Peptides International (Louisville, KY, USA). 1-Kγ-PNA Thymine Monomer was synthesized according to published procedures. The radioligands $[^3]$HCGS21680 and $[^{125}]$I-AB-MECA were purchased from PerkinElmer (Waltham, MA, USA), and $[^3]$H-R-PIA was purchased from Moravek Biochemicals (Brea, CA, USA). All other reagents were obtained from Sigma-Aldrich (St. Louis, MO, USA). PNA oligomer synthesis was performed on an Applied BioSystems 433A Automated Peptide Synthesizer. Purification of PNA oligomers was carried out using a X-Bridge Prep BEH 130 C18 5µm (10 x 250 mm) column on an Agilent 1200s HPLC. The typical flow rate was 4 mL/min. HPLC solvents consisted of HPLC grade acetonitrile:MilliQ water (9:1) and 0.10% aqueous TFA. Wavelengths 220 nm, 260 nm, and 315 nm were monitored. High-resolution mass spectra (HRMS) were obtained on a LC/MSD TOF (Agilent Technologies, Santa Clara, CA, USA). DNA oligomers were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA) and used without further purification. UV quantification of PNA and DNA was performed using an Agilent 8453 UV-Vis Spectrophotometer.

**Abbreviations:**

ACN, acetonitrile; (Boc), tert-butoxycarbonyl-; (CGS21680), 2-[p-(2-carboxyethyl)phenyl-ethylamino]-5'-N-ethylcarboxamidoadenosine; (DMEM), Dulbecco’s modified Eagle medium; (DCM), dichloromethane; (DMF), N,N-dimethylformamide; (DMSO), dimethylsulfoxide; (ESI-MS), electrospray ionization mass spectrometry; (HPLC), High Performance Liquid Chromatography; (I-AB-MECA), 4-amino-3-iodobenzyl)adenosine-5'-N-methyl-uronamide; (MBHA resin), 4-methylbenzhydrylamine resin; (NMP), N-methyl-2-pyrrolidinone; (mini-PEG), 8-amino-3,6-dioxaoctanoic acid; (PBS), phosphate buffered saline; (R-PIA), $N^6$-[(R)-phenylisopropyl]adenosine;
(PNA), peptide nucleic acid; (TEA), triethylamine; (TFA), trifluoroactic acid; (TfOH), trifluorosulfonic acid; (TRIS), tris(hydroxymethyl)aminomethane-hydrochloric acid buffered saline; (XAC), Xanthine amine congener; (ZM241385), 4-[2-[7-amino-2-(2-furyl)-1,2,4-triazolo[1,5-a][1,3,5]triazin-5-yl-amino]ethylphenol.

General Synthesis and Purification Procedures

Preparation of PNA Oligomers. MBHA Resin (0.3 mmol/g) was prepared by swelling in DCM and downloading the resin with Boc protected N,N-dimethyl-L-lysine to 0.1 mmol/g capacity. PNA oligomers were made via solid-phase peptide synthesis in accordance with literature procedures on either 5 or 25 µmol scale.\textsuperscript{28, 29, 49}

Sequences. Sequence used for XAC-conjugated PNA: AGT-AGA-TCA-CTG. Complementary antiparallel sequence: CAG-TGA-TCT-ACT. Note, for L-PNAs B_{(2,3)} and B_{(1,14)} the conjugated PNA sequences were modified to AGT-AGA-TCA-TTG and T-AGT-AGA-TCA-CTG-T respectively. The complementary sequences were adjusted accordingly.

General Resin Cleavage. Upon completion of PNA synthesis or solid phase coupling, the PNA-bound resin was transferred to a glass reaction vessel and washed with DCM, then TFA. The resin was swelled in TFA. The solvent was removed and a solution of m-cresol (150 µL), thioanisole (150 µL), TfOH (300 µL), and TFA (900 µL) was added and allowed to sit on the resin for 60 min. The solution was drained into a scintillation vial. This was repeated for a total of 3 washes, each time collecting the eluent in the scintillation vial. The pooled solution was concentrated, transferred to microfuge tubes, and precipitated using diethyl ether at a ratio of 1:10. The resulting flaky off-white solid was washed 3 times with diethyl ether and dried under vacuum. The resulting residue was diluted with 2:1 water:ACN and further purified on reversed phase HPLC.
**General Conjugation Procedures.** The XAC ligand (N-(2-aminoethyl)-2-[4-(2,3,6,7-tetrahydro-2,6-dioxo-1,3-dipropyl-1H-purin-8-yl)phenoxylacetamide, Sigma-Aldrich, St. Louis, MO, USA) was conjugated to the free amino moiety on the PNA oligomer scaffold via squaric acid. This was performed by one of two methods:

1) Resin containing 5 µmol PNA with a free amine was swelled with NMP for 1 h in a glass reaction vessel. Following this, the solvent was drained and replaced with fresh NMP (500 µL). To this, triethylamine (300 µmol, 60 equiv) and 3,4-diethoxy-3-cyclobutene-1,2-dione (150 µmol, 30 equiv) were added. The vessel was sealed and agitated for 3 h. The resin was then washed (2 x NMP, 2 x DCM, and 2 x NMP). A pre-dissolved solution containing XAC (100 µmol, 20 equiv), 1:9 DMSO:NMP and TEA (150 µmol, 30 equiv) was added to the resin. The vessel was sealed and agitated until completion. In general, A-type L-PNAs were reacted for 18 h, while B and C L-PNAs were allowed 36 h to couple. The resin was then washed as previous and then conjugated PNA was cleaved from the resin.

2) In a 2 mL Eppendorf tube, lyophilized cleaved PNA was dissolved in minimal 2:1 anhydrous DMSO:ethanol. To this, triethylamine (60 equiv) and 3,4-diethoxy-3-cyclobutene-1,2-dione (30 equiv) were added. The vessel was flushed with nitrogen, sealed, and agitated for 3 h. The solution was concentrated and the residue was washed with diethyl ether (3 x 2 mL), and dried under vacuum to obtain the squaric acid-conjugated PNA intermediate as an off-white solid. A solution of XAC (20 equiv), 2:1 anhydrous DMSO:ethanol, and triethylamine (30 equiv) was added to the squaric acid-conjugated PNA intermediate. The vessel was flushed with nitrogen, sealed, and agitated until completion. In general, A-type L-PNAs were reacted for 18 h, while B and C L-PNAs were allowed 36 h to react. The solution was purified directly by reversed phase HPLC.
**General HPLC Purification.** PNA and conjugated-PNA residues were purified by reversed-phase HPLC using a 10 x 250 mm Waters XBridge prep BEH130 C18 5µm reverse phase column on an Agilent 1200s HPLC. Wavelengths 220 nm, 260 nm, and 315 nm were monitored. The typical flow rate was 4 mL/min. HPLC solvents consisted of ACN:water (9:1) and 0.15% aqueous TFA.

PNA residues were purified using one of the following methods:

1) Thermostat at 35°C. Gradient hold at 0% ACN 0-2 min, 10% ACN at 5 min, 20% ACN at 20 min, then wash with 100% ACN for 5 min.

2) Thermostat at 35°C. Gradient hold 0% ACN 0-1.9 min, 10% ACN at 2 min, 35% ACN at 25 min, then wash with 100% ACN for 5 min.

3) Thermostat at 50°C. Gradient hold 0% ACN 0-1.9 min, 10% ACN at 2 min, 40% ACN at 35 min, then wash with 100% ACN for 5 min.

**PNA Oligomer Structures and Characterization Data**

Nomenclature: The individual L-PNAs are named in accordance with the nomenclature set out in the main text. Prior to the coupling of the ligand, the PNA oligomer bears a lower case designation. For example, the A-type L-PNA is referred to as PNA a to signify that it does not contain a ligand.

![PNA Structure](image)

**Figure S1.** PNA a. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 19.6 min. ESI-MS for C_{156}H_{211}N_{77}O_{43} expected 3852.83, found 3851.12.
Figure S2. L-PNA A. Using purified PNA, the conjugate was generated using coupling procedure 2 and HPLC method 2 eluting at 23.9 min. ESI-MS for C_{181}H_{237}N_{83}O_{49} expected 4359.35, found 4359.1.

Figure S3. PNA b. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 10.05 min. ESI-MS for C_{172}H_{242}N_{80}O_{49} expected 4214.27, found 4213.40.

Figure S4. L-PNA B. Using purified PNA, the conjugate was generated using coupling procedure 2 and HPLC method 3 eluting at 34.36 min. ESI-MS for C_{222}H_{294}N_{92}O_{61} expected 5227.29, found 5226.30.
**Figure S5.** PNA c. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 17.8 min. ESI-MS for C_{188}H_{273}N_{83}O_{55} expected 4575.70, found 4575.6.

**Figure S6.** L-PNA C. Using purified PNA, the conjugate was generated using coupling procedure 2 and HPLC method 2 eluting at 26.9 min. ESI-MS for C_{263}H_{351}N_{101}O_{73} expected 6095.23, found. 6094.7

**Figure S7.** PNA b_{2,3}. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 14.71 min. ESI-MS for C_{173}H_{243}N_{79}O_{50} expected 4229.28, found 4228.35.

**Figure S8.** L-PNA B_{2,3}. Using purified PNA, the conjugate was generated using coupling procedure 2 and HPLC method 3 eluting at 313.11 min. ESI-MS for C_{223}H_{295}N_{91}O_{62} expected 5242.30, found 5241.16.
**Figure S9.** PNA $b_{6,10}$. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 15.65 min. ESI-MS for $C_{172}H_{242}N_{80}O_{49}$ expected 4214.27, found 4213.53.

**Figure S10.** L-PNA $B_{6,10}$. Using purified PNA, the conjugate was generated using coupling procedure 2 and HPLC method 3 eluting at 31.48 min. ESI-MS for $C_{222}H_{294}N_{92}O_{61}$ expected 5227.29, found 5226.23.

**Figure S11.** PNA $b_{1,14}$. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 15.93 min. ESI-MS for $C_{194}H_{278}N_{88}O_{77}$ expected 4746.77, found 4746.80.
Figure S12. L-PNA B_{1,14}. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 3 eluting at 31.43 min. ESI-MS for C_{244}H_{322}N_{100}O_{69} expected 5759.80 found 5758.61.

Figure S13. Complement PNA1. Me_{2}Lys-TCA-TCT-AGT-GAC-Ac. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 14.59 min. ESI-MS for C_{139}H_{181}N_{69}O_{39} expected 3442.36, found 3441.9.

Figure S14. Complement PNA1_{(1,14)}. Me_{2}Lys-A-TCA-TCT-AGT-GAC-A-Ac The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 14.124 min. The method was modified, increasing the thermostat temperature to 45°C. ESI-MS for C_{161}H_{207}N_{85}O_{43} expected 3992.90, found 3992.06.
Figure S15. Complement PNA1(2,3). Me₂Lys-TCA-TCT-AGT-AAC-Ac. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 13.31 min. ESI-MS for C₁₃₀H₁₈¹N₆₉O₃₈ expected 3424.44, found 3425.51.

Figure S16. Complement PNA2. Me₂Lys-TCA-TCT-AGT-GAC-Me₂Lys-TCA-TCT-AGT-GAC-Me₂Lys-Ac. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 15.59 min. The method was modified, increasing the thermostat temperature to 45°C. ESI-MS for C₂₈₄H₃₇₃N₁₃₉O₇₈ expected 6981.88, found 6980.43.

Figure S17. Complement PNAm3. Me₂Lys-TCA-TCT-AGT-GAC-Me₂Lys-TCA-TCT-AGT-GAC-Me₂Lys-TCA-TCT-AGT-GAC-Me₂Lys-Ac The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 13.73 min. The method was modified, increasing the thermostat temperature to 60°C. ESI-MS for C₄₂₁H₅₄₉N₂₀₇O₁₁₆ expected 10365.18, found 10363.35.
Figure S18. Complement PNA. \( \text{Me}_2\text{Lys-TCA-TCT-AGT-GAC-Me}_2\text{Lys-TCA-TCT-AGT-GAC-Me}_2\text{Lys-TCA-TCT-AGT-GAC-Me}_2\text{Lys-Ac.} \) The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 14.45 min. The method was modified, increasing the thermostat temperature to 70°C. ESI-MS for \( C_{558}H_{725}N_{275}O_{154} \) expected 13748.48, found 13746.6.

Figure S19. L-PNA Ac-C. The PNA was cleaved from the resin upon completion of synthesis and purified by HPLC method 1 eluting at 19.88 min. ESI-MS for \( C_{194}H_{279}N_{85}O_{58} \) expected 4701.89 found 4701.7.
L-PNA Duplexes

Quantification of PNA Oligomer Conjugates

Lyophilized PNA oligomers were dissolved in water. The absorbance of an aliquot was determined by UV-VIS spectroscopy after heating the sample for 5 min at 90°C. This was performed in triplicate. Using the extinction coefficient of the analogous DNA oligomer obtained from Applied Biosystems (Life Technologies, Grand Island, NY), the concentration was determined.

General Annealing Condition for Formation of PNA:DNA or PNA:PNA Duplexes.

RNA/DNAase free microfuge tubes, PNA, DNA and TRIS buffer (pH 7.5) were combined at room temperature. The final TRIS buffer concentration was 100 mM. Equivalents of PNA were added based on the number or repeating 12-mer sequences in the DNA. For example, to generate PNA:DNA multi5, a 5:1 molar ratio of PNA:DNA was used. The solution was heated to 90°C, held for 5 min, then slowly allowed to cool down to 25°C over a period of 3 h.

LCMS Analysis of L-PNA:PNA Duplex

To confirm that the L-PNA:PNA complex was one species, and not an aggregate, mass spectrometry was utilized. Two complexes were analyzed and confirmed by this method: \( B_2P \) and \( B_{(6,10)}4P \).

The PNA-complexes were separated from the monomers by reversed phase HPLC using electrospray ionization mass spectrometry (ESI MS) as the detection method. The HPLC was a Waters 1525u operated at a flow rate of 200 µL per min. Solvent A was 1% acetonitrile in water with 0.2% formic acid and 0.1% TFA. Solvent B was methanol with 20% acetonitrile with 0.2% formic acid and 0.1% TFA. The elution program starts at 0% B and is increased to 100% B in 9 min and finally held for 3 min at 100% B. The HPLC column was a Bruker-Michrom PLRP-S column with internal diameter of 2.1 mm and a length of 150 cm.
The ESI/MS was a Waters LCT Premiere operated in the positive ion V-mode. The ESI capillary voltage was 3.4KV. The multiple charged spectra were deconvoluted with MaxENT1.

B2_P

The components (B_{(2,10)} L-PNA and the complement PNA) were injected individually into the LC/MS system and their respective retention times and multiply charged ESI/MS spectra were recorded. The larger PNA (the complement) had a retention time of 8.99 min, the base peak was the 7+ ion at 975.9 Da and a deconvoluted molecular weight of 6824. The smaller PNA (B_{(2,10)} L-PNA) eluted with a retention time of 8.49 min, the base peak was a 5+ ion at 699.1 Da and a deconvoluted molecular weight of 3490. The PNA-complex was observed with a retention time of 9.07 min and the individual components were simultaneously observed at that retention time. The ESI/MS of the lower mass component again showed a base peak at 1164.5 Da for the 3+ charged ion. The ESI/MS spectrum of the larger component yields the same molecular weight previously observed but the charge distribution is quite different with the base peak becoming the 6+ ion at 1138.4 Da. This change in charge state distribution is consistent with the larger PNA existing in a radically different state in the complex versus the monomeric form.

B_{(6,10)}4_P

The components (B_{(6,10)} L-PNA and the complement PNA) were injected individually into the LC/MS system and their respective retention times and multiply charged ESI/MS spectra were recorded. The larger PNA (the complement) had a retention time of 6.4 min, the base peak was the 8+ ion at 1719.3 Da and a deconvoluted molecular weight of 13764. The smaller monomer (B_{(6,10)} L-PNA) eluted with a retention time of 8.6 min, the base peak was a 4+ ion at 1307.6 Da and a deconvoluted molecular weight of 5226.4. The PNA-complex was observed with a retention time of 10.0 min and the individual components were simultaneously observed at that retention time. The ESI/MS of the lower mass component again showed a base peak at 1307.6 Da for the 4+ charged ion. The ESI/MS spectrum of the
larger component yields the same molecular weight previously observed but the charge distribution is quite different with the base peak becoming the $10^+$ ion at 1375.8 Da. This change in charge state distribution is consistent with the larger PNA existing in a radically different state in the complex versus the monomeric form.

**Experimental Procedures**

**Cell cultures and membrane preparation**

Chinese hamster ovary (CHO) cells stably expressing the recombinant hA1 and hA3ARs, and HEK293 cells stably expressing the hA2AR were cultured in Dulbecco’s modified Eagle medium (DMEM) and F12 (1:1) supplemented with 10% fetal bovine serum, 100 units/mL penicillin, 100 µg/mL streptomycin, and 2 µmol/mL glutamine. In addition, 800 µg/mL geneticin was added to the A2A media, while 500 µg/mL hygromycin was added to the A1 and A3 media. After harvesting, cells were homogenized and suspended in PBS. Cells were then centrifuged at 240 g for 5 min, and the pellet was resuspended in 50 mM Tris-HCl buffer (pH 7.5) containing 10 mM MgCl2. The suspension was homogenized and was then ultra-centrifuged at 14,330 g for 30 min at 4°C. The resultant pellets were resuspended in Tris buffer, incubated with adenosine deaminase (3 units/mL) for 30 min at 37°C. The suspension was homogenized with an electric homogenizer for 10 sec, pipetted into 1 mL vials and then stored at -80°C until the binding experiments. The protein concentration was measured using the BCA Protein Assay Kit from Pierce Biotechnology, Inc. (Rockford, IL).

**Competitive Radioligand Binding to A2A Receptors**

Competition radioligand binding experiments were conducted to determine the binding affinities of PNA conjugates. A range of concentrations of PNA conjugates between 1 nM to 1000 nM was tested in competing for binding to A2A receptors on cell membranes derived from A2A-expressing HEK cells. Assay solutions (200 µL) in binding buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl2) comprised of cell membranes (100 µL), radioligand (50 µL), and PNA conjugates (50 µL) were prepared in test tubes,
which were incubated at 25°C for 1 h in a shaking water bath. Additionally, assay solutions containing either binding buffer instead of the PNA conjugates or 40 µM adenosine-5’-N-ethyluronamide were prepared for determining total and nonspecific radioligand binding to the membranes, respectively. The radioligand agonist [³H]CGS21680 was used for all A₂A experiments. [³H]R-PIA and [¹²⁵I]I-AB-MECA were used for A₁ and A₃ binding experiments, respectively. After incubation, binding was terminated by rapid filtration through glass filter paper. The glass filter paper samples were then read by a scintillation counter (Tri-Carb 2810TR) to determine radioligand binding. The counts per ligand concentration were plotted and curve fit using Prism (GraphPad, San Diego, CA, USA) to obtain IC₅₀ values. Each experiment provided a redundant data set and used 7 different concentrations of duplexed L-PNA. This was repeated in triplicate.

Fluorescent ligand binding experiments with flow cytometry (FCM)

The HEK 293 cells expressing the hA₂AAR were grown in 12-well plates (approximately 200,000 cells/well) and incubated at 37 °C 36 h in the presence of 5% CO₂. When the confluency of the cells reached 80% (approximately 4×10⁵ cells/well), medium was replaced with fresh medium and B₅DF was added in the presence or absence of 10 µM ZM241385, and cells were processed for FCM. Note that B₅DF is was generated using Alexa Fluor 488 labeled DNA purchased from IDT.

HEK 293 cells expressing A₂AARs were incubated with different concentrations of B₅DF ranging from 1 nM to 50 nM for 30 min for a saturation binding experiments. To study binding kinetics, we incubated HEK293 cells expressing A₂AARs with 30 nM B₅DF for different time intervals from 5 min to 3 h. Nonspecific binding was measured in the presence of 10 µM ZM241385.³⁹

At the end of each time interval, the medium was removed and cells were washed two times with DPBS. After washing, 0.5 ml 0.2 % EDTA solution was added to each well, and cells were incubated at 37 °C for 1 min. Following cell detachment, 0.5 ml medium was added to each well to neutralize the EDTA.
The cell suspensions were transferred to polystyrene round-bottom BD Falcon tubes (BD, Franklin Lakes, NJ) and centrifuged for 5 min at 23 °C and 400×g. After centrifugation, the supernatant was discarded, and cells were washed with 2 ml PBS and centrifuged again at 23 °C and 400×g for 5 min. After discarding the supernatant, cells were suspended in 0.3 ml PBS and analyzed by FCM.

The intensity of fluorescence emission of each sample was measured by using FCM. Cell suspensions were vortexed briefly before analysis on a Becton and Dickinson FACSCalibur flow cytometer (BD, Franklin Lakes, NJ) with excitation at 488 nm. Samples were maintained in the dark during the analysis to avoid photobleaching. MFIs were obtained in the FL-1 channel in log mode. Ten thousand events were analyzed per sample. Data were collected using Cell Quest Pro software (BD, Franklin Lakes, NJ).

Association binding results were analyzed by fitting the binding data to a One-phase association equation \( y = y_0 + (\text{Plateau} - y_0)(1 - e^{kx}) \), where \( y_0 \) is the MESF (y value) when time (x value) is zero, Plateau is the MESF at infinite times, and \( K \) is the rate constant, expressed in inverse min.

Saturation binding results were analyzed by fitting a One-site total and nonspecific binding equation to the binding data. Total and nonspecific binding was globally analyzed by fitting the total binding data to the equation \( y = B_{\text{max}}X/(X+K_d) + (NS \times X) \) and \( y = \text{NS} \times X \) for the nonspecific binding data, where \( B_{\text{max}} \) means the maximum specific binding in MESF units, \( K_d \) is the equilibrium binding constant in nM, and \( \text{NS} \) is the slope of nonspecific binding.

The measured fluorescence intensities were corrected with the subtraction of autofluorescence values of HEK 293 cells in the absence of any AR ligand.
Cyclic AMP Accumulation Assay

CHO cells expressing A$_{2A}$AR were seeded in 24-well plates and incubated at 37 °C overnight. The following day, the medium was removed and replaced with DMEM containing 50 mM HEPES, 10 μM rolipram, 3 U/mL adenosine deaminase, and increasing concentrations of a known agonist (CGS21680). The suspected antagonist (B$_{6,10}4p$) was added 20 min before the addition of agonist. The medium was removed, and the cells were lysed with 200 μL of 0.1 M HCl. One hundred microliters of the HCl solution was used in the Sigma Direct cAMP Enzyme Immunoassay following the instructions provided with the kit. The results were interpreted using a BioTek ELx808 Ultra Microplate reader (BioTek, Winooski, VT) at 405 nm.$^{19}$

Molecular Modeling of a PNA duplex bound to an A$_{2A}$AR homodimer

$PNA:PNA$ Duplex:

Atomic-scale, computer models of select PNA-duplexes were developed with the QUANTA (Accelrys) and CHARMM software programs.$^{50}$ The helical conformations were derived from the NMR solution structure of a gamma-methylated PNA-duplex 8-mer (PDB accession code: 2KVJ).$^{43}$ Except for the nitrogenous bases, topologies needed to be developed for all the other molecular components. These were derived from the “all27” set of topologies and parameters provided with the CHARMM program. Finally, all models were energy minimized with CHARMM to eliminate atomic overlap and optimize the bond lengths and angles.

$hA_{2A}$AR model:

To have a more complete 3D structure of the hA$_{2A}$AR, we built a model using the Homology Modeling tool implemented in the MOE suite and the available crystallographic data for this receptor subtype.$^{51}$
The model was based on the highest-resolution hA<sub>2A</sub>AR crystal structure (PDB ID: 4EIY),<sup>41</sup> where the template for the missing IL3 (from Lys209 to Gly218) was another inactive-state hA<sub>2A</sub>AR crystal structure (PDB ID: 3REY).<sup>31</sup> The intracellular C-terminal tail of the receptor (from Leu308 to Ser412) was not modeled, due to the absence of a useful template. Previously published FRET studies showed that the C-term does not participate in the A<sub>2A</sub>AR homodimerization,<sup>45</sup> therefore it seemed reasonable to exclude it from the modeling studies. The AMBER99 force field was used for protein modeling and the Protonate 3D methodology was used for protonation state assignment. The final model was refined through energy minimization until a RMS gradient of 0.1 kcal/mol Å. Model's stereochemical quality was checked using several tools (Ramachandran plot; backbone bond lengths, angles and dihedral plots; clash contacts report; rotamers strain energy report) implemented in the MOE suite.

**Molecular docking of XAC-linker at the hA<sub>2A</sub>AR model:**

The XAC-linker structure was built using the builder tool implemented in the MOE suite and subjected to energy minimization using the MMFF94x force field, until a RMS gradient of 0.05 kcal/mol Å. Molecular docking of the ligand at the hA<sub>2A</sub>AR model was performed by means of the Glide package part of the Schrödinger suite.<sup>52</sup> The docking site was defined using key residues in the binding pocket of the hA<sub>2A</sub>AR model, namely Phe (EL2), Asn (6.55), Trp (6.48) and His (7.43), and a 30Å x 30Å x 30Å box was centered on those residues. Docking of ligand was performed in the rigid binding site using the SP (standard precision) protocol. The top scoring docking conformations were comparable with the crystal pose of XAC at the A<sub>2A</sub>AR (3REY).<sup>31</sup> In particular, the main interactions observed in the crystal between the xanthine scaffold and the receptor were conserved, while the linker was pointing outside of the cavity making contacts with residues in EL2, such as Lys150, Lys153 and Gln157.

**hA<sub>2A</sub>AR homodimer model:**

Homodimers were built starting from our hA<sub>2A</sub>AR model and using the protein-protein docking tool of the ZDOCK server (ZDOCK 3.0.2).<sup>53</sup> From the resulting poses, antiparallel dimers or poses not
compatible with the nature of transmembrane proteins (i.e. excessive inclination or shift along the main axis between the two monomers) were discarded. For selected dimer poses, contact areas between two monomers were refined through energy minimization until a RMS gradient of 0.1 kcal/mol Å, using the AMBER99 forcefield implemented in the MOE suite.

Based on the monomers relative orientation we could identify, among the several reasonable poses returned by the software, two most populated clusters of dimers. The first cluster collected dimers with interface between TM5, TM6 and TM7, the second cluster dimers with interface between TM1, TM2 and helix 8. Both interfaces are comparable with some proposed through computational studies in a previous paper on A$_{2A}$AR homodimerization.$^{46}$ For the majority of the possible dimers, the distance between the binding sites of the two monomers was 30-40Å. Considering that the distance between monomers was quite comparable among different dimers and that an unambiguous identification of the functionally relevant interface is very difficult, we selected one representative dimer belonging to the first cluster as starting point for the following modeling of the A$_{2A}$AR homodimer-PNA duplex construct.

$A_{2A}AR$ homodimer-PNA duplex construct:

To combine the models of the hA$_{2A}$AR homodimer and of the PNA duplex the following procedure was performed. A XAC-linker structure was placed in its docked conformation inside each hA$_{2A}$AR monomer forming the dimer model. Then, the PNA duplex model was manually placed in proximity of the extracellular side of the dimer and the terminal groups of each XAC-linker structure were connected to the PNA chain at positions 6 and 10. Finally, the construct geometry was refined by energy minimization using the software MOE and the Amber12:EHT force field, until a RMS gradient of 0.1 kcal/mol Å. During the minimization the hA$_{2A}$AR dimer and the XAC scaffolds were kept fixed, the linker chains were free to move and the PNA duplex was considered as a rigid body.
MD Simulation of the Complex:

Molecular Dynamics simulation of the above-described complex was performed with the CHARMM program.50 For this, the A<sub>2A</sub>AR homodimer-PNA duplex construct was centered in a cubic cell with edges of 100.0 Å. The A<sub>2A</sub>AR homodimer was embedded in a membrane bilayer of 200 1-palmytoil-2-oleoyl-sn-glycero-3-Phosphatidylcholine (POPC) lipid molecules. The initial coordinates for these were derived from an equilibrated patch provided by Klauda et al.54 The remainder of the space was filled with 18,195 Tip3p water models, for a total of 92,136 atoms. With the cell subjected to periodic boundary conditions, and image-centering for the water and lipid molecules, the system was subjected to extensive energy minimization, and then heating and equilibration dynamics with the Langevin method. This was changed for the production dynamics, which was propagated under conditions of constant pressure (1 ATM) and temperature (300 K). This latter was achieved using the Langevin-piston/Hoover algorithms with the default parameters. The electrostatic and van der Waals interactions were attenuated to zero at 8.5 Å with the respective force-shift and shift functions. Since only the hydrogens of the water molecules were subjected to the SHAKE method, the time-step was set to 1.0 fs. The non-bonded lists were updated every 25 steps, with cutoffs of 11.5 and 12.0 Å for the real and image atoms. The computer simulation was run in parallel on 16 processors, and coordinate snapshots were saved every 1.0 ps for analysis.

Video Captions

**Video S1.** Generation of L-PNA multivalent scaffolds. The video highlights the symmetry between PNA and DNA, as well as the design elements of the L-PNA conjugate. It continues with an extended look at how one L-PNA can generate a family of multivalent constructs through the use of varying the length of the complementary DNA oligomer.
**Video S2.** Modeling all possible bivalent L-PNA:PNA ligand arrangements. Left, the video highlights how all 66 possible combinations of ligand arrangements were generated and theoretically assessed using the model. Additionally, it notes which constructs were prepared and examined experimentally. Right, a table denoting the positions of the linker-ligand modification.

**Video S3.** Trajectory of the Molecular Dynamics simulation of the A$_2$AAR-homodimer/B$_{(6,10)}$1P complex model. The total length is 4 ns, for which snapshots are shown every 1 ps. The receptor monomers are displayed as alpha-carbon traces in yellow and green. For clarity, the water molecules are not shown, and the membrane bilayer is represented simply by white spheres for the phosphorus atoms of the lipid molecules. Spheres are also used for the heavy atoms of the PNA construct. One strand is color-coded blue for the backbone and linkers, and cyan for the bases; while the other is red for the backbone and pink for the bases. The heavy two XAC ligands are shown in magenta.

**Additional Data**

**Linker Length Study**

| mPeg linkers | IC$_{50}$ Values (nM) |
|--------------|------------------------|
|              | A1 | A5 |
| 1            | NS | NS |
| 2            | 440 ± 110 | 120 ± 15 |
| 3            | 2400 ± 190 | 310 ± 28 |

**Chart S1.** Using the same generic PNA 12-mer, a series of four type-A L-PNA oligomers were generated. Each L-PNA contains progressive mini-Peg incorporations, varying the length between the PNA backbone and the XAC ligand. The IC$_{50}$ values of A$_1$ and A$_5$ were determined for each spacer length. Note that a single mini-Peg spacer resulted in a non-soluble (NS) duplex that was not pursued further.
Whole Cell Assay

| Duplex | Ligands | Membrane Assay | Whole Cell |
|--------|---------|----------------|------------|
| B5_D   | 10      | 50 ± 14        | 25 ± 6     |
| C2_D   | 6       | 23 ± 2         | 63 ± 16    |

Chart S2. Binding affinities (IC50 values) of two L-PNA:DNA complexes (Figure 2) were tested in both the radioligand membrane displacement assay19 (columns labeled “Membrane Assay”) and a whole-cell assay adapted to A2AR39 (column labeled “Whole Cell”). For the whole cell assay, the DNA was labeled at the 5’ end with AlexaFluor488 (purchased from IDT). Both labeled and unlabeled B5_D were tested in the Membrane Assay to determine whether the fluorescent group interfered with binding. Binding affinities between the Membrane Assay and Whole Cell assay are consistent, and there is no imapct from the fluorescent label.

Backbone Comparison

Figure S20. The affect of the backbone was examined, wherein the DNA backbone was replaced with a PNA backbone for the A1 complex. An 8x enhancement in binding affinity was observed for the charge neutral L-PNA:PNA complex.
Hill-slope Analysis

| L-PNA:PNA | PNA backbone length |
|-----------|---------------------|
|           | 1       | 2       | 3       | 4       |
| A         | 1.0 ± 0.2 | 0.8 ± 0.1 | 1.2 ± 0.2 | 1.5 ± 0.4 |
| B(6,10)   | 1.0 ± 0.1 | 1.2 ± 0.2 | 1.9 ± 0.3 | 1.5 ± 0.1 |
| B(2,10)   | 1.3 ± 0.1 | 1.6 ± 0.1 | 1.0 ± 0.2 | 0.9 ± 0.1 |
| C         | 0.9 ± 0.1 | 1.1 ± 0.1 | 1.1 ± 0.1 | 1.2 ± 0.1 |

Chart S3. The Hill-slopes for each L-PNA:PNA conjugate are summarized along with the associated error. The global trend supports the interpretation of enhanced binding, specifically in the B_{6,10} family.
RMSD from MD Simulation

Figure S21. Root-Mean-Square Deviations (RMSD) of the $A_{2A}AR$-homodimer/B$_{(6,10)}1P$ model over 5 ns of Molecular Dynamics simulation. The deviations are from the coordinate snapshot after 1 ns of equilibration. Separate plots are presented for A) the $A_{2A}$ receptor protein and B) PNA helix. The results for the protein consider only the backbone atoms of the transmembrane segments (residues 3-23, 46-67,
Individual traces are given for the receptor dimer (black) and each monomer separately (green and blue). The results for the PNA helix consider only the central 8 base-pairs. Individual traces are provided for the backbone (black) and nitrogenous bases (green).

Selectivity and Biological Behavior

![Graph](image)

**Figure S22.** The activity of $B_{(6,10)}4P$ was examined in whole cells against a known agonist (CGS 21680) in whole cells using a standard cAMP assay. The complex is shown to behave as an antagonist, similar to its parent XAC ligand. The $K_B$ value calculated from Schild analysis is 20.9 nM.

**Chart S4.** Multivalent display of XAC increases selectivity for $A_{2A}$ over $A_1$ and $A_3$ compared to monovalent XAC. The binding affinity of $B_{(6,10)}4P$ in $A_{2A}$ overexpressed membranes was compared to AR homologues of $A_1$ and $A_3$. The numerical values in columns 1 and 2 are ratios of $IC_{50}$ values for the indicated receptors. The selectivity of the multivalent construct is compared to recently published literature values of XAC, which itself is nonselective for these receptors.$^{19}$ The numerical values in column 3, which represent the increase in receptor selectivity over XAC, are the ratios of column 2 divided by column 1.

| XAC Literature | $B_{(6,10)}4P$ | Increase |
|----------------|--------------|----------|
| $A_1:A_{2A}$   | 1.4          | 10.2     | 7.2     |
| $A_3:A_{2A}$   | 2.4          | 6.9      | 2.9     |
Thermal Denaturation (Duplex Stability)

| Duplex     | Tm °C | Error |
|------------|-------|-------|
| PNA:DNA    | 57.20 | 0.06  |
| A1_D       | 55.53 | 0.48  |
| B1_D       | 54.54 | 0.45  |
| C1_D       | 64.60 | 0.22  |
| A1_P       | 72.07 | 0.13  |

**Chart S5.** Thermal denaturation experiments were carried out on select complexes. Experiments were performed in 600 μL cuvettes on Agilent 8453 UV-Vis Spectrophotometers. A 5 μM solution (225 μL) of PNA and DNA were prepared in TRIS (100 mM) and added to the cuvette. After warming at 90°C for 5 min, the cuvette was cooled at a rate of 1° per min. The absorbance at 260 nm was plotted versus temperature and curve fit to obtain melting temperatures.

**Control Experiments**

| Compound    | IC₅₀ (nM)      |
|-------------|----------------|
| XAC-miniPeg | 210 ± 8        |
| AcC1_D      | 0% at 10µM     |
| AcC1_P      | 0% at 10µM     |

**Chart S6.** To ensure proper interpretation of the data, numerous control experiments were designed. The XAC ligand was conjugated to mini-PEG in order to properly mimic the monovalent ligand. An acetylated C-type L-PNA (**Figure S19**) was prepared and annealed to both DNA and PNA to form the corresponding complexes. No binding was observed.
**η-values for Multivalent Libraries**

η-values are used to compare adjacent constructs in Figure 2a and Figure 6a in order to illuminate increasing/decreasing trends derived from multivalent effects. While it is possible to use this approach to compare any two non-adjacent complexes, the interpretation of the resulting η-value becomes less clear.

| DNA Backbone Length | A→B | B→C |
|---------------------|-----|-----|
| 1                   | 20  | 1.0 |
| 2                   | 5.2 | 1.6 |
| 3                   | 2.8 | 0.8 |
| 4                   | 2.0 | 1.8 |
| 5                   | 1.7 | 2.3 |

**Chart S7.** η-values for L-PNA:DNA multivalent landscape contrasting increasing ligand content per PNA-12mer. A-Type complexes have 1 ligand per L-PNA 12mer, whereas B-Type contains 2 ligands, and C-Type contains 3 ligands per PNA 12mer. In column 2, “A→B” the ligand content increases by a factor of two for each entry. In column 3, “B→C” the ligand content increases by a factor of 1.5 for each entry.
DNA Backbone Length | A type (L-PNA) | B type (L-PNA) | C type (L-PNA)
---|---|---|---
1→2 | 4.7 | 1.2 | 2.0
2→3 | 1.6 | 0.9 | 0.4
3→4 | 0.9 | 0.6 | 1.5
4→5 | 0.9 | 0.8 | 0.8

**Chart S8.** η-values for L-PNA:DNA multivalent landscape contrasting increasing backbone length. A-Type complexes have 1 ligand per L-PNA 12mer, whereas B-Type contains 2 ligands, and C-Type contains 3 ligands per PNA 12mer. In row 1, where the DNA backbone length increases from 1→2 (12→24 base-pairs), the ligand content increases by a factor of two for each entry. In row 2, where the DNA backbone length increases from 2→3 (24→36 base-pairs), the ligand content increases by a factor of 1.5 for each entry. In row 3, where the DNA backbone length increases from 3→4 (36→48 base-pairs), the ligand content increases by a factor of 1.33 for each entry. In row 4 where the DNA backbone length increases from 4→5 (48→60 base-pairs), the ligand content increases by a factor of 1.25 for each entry.

PNA Backbone Length | A→B₂₁₀ | B₂₁₀→B₆₁₀ | B₆₁₀→C
---|---|---|---
1 | 2.3 | 1.1 | 1.5
2 | 0.6 | 2.8 | 0.5
3 | 0.9 | 2.0 | 0.4
4 | 3.1 | 2.7 | 0.3

**Chart S9.** η-values for L-PNA:PNA multivalent landscape contrasting increasing ligand content per PNA-12mer. A-Type complexes have 1 ligand per L-PNA 12mer, whereas both B-Type contains 2 ligands, and C-Type contains 3 ligands per PNA 12mer. In column 2, “A→B” the ligand content increases by a factor of two for each entry. In column 3, “B→C” the ligand content increases by a factor of 1.5 for each entry.
| PNA Backbone Length | A type (L-PNA) | B₂₁₀ (L-PNA) | B₆₁₀ (L-PNA) |
|---------------------|----------------|--------------|--------------|
| 1→2                 | 1.6            | 0.4          | 1.1          |
| 2→3                 | 0.8            | 1.2          | 0.9          |
| 3→4                 | 0.5            | 1.6          | 2.2          |

**Chart S10.** η-values for L-PNA:PNA multivalent landscape contrasting increasing backbone length. A-Type complexes have 1 ligand per L-PNA 12mer, whereas B-Type contains 2 ligands, and C-Type contains 3 ligands per PNA 12mer. In row 1, where the PNA backbone length increases from 1→2 (12→24 base-pairs), the ligand content increases by a factor of two for each entry. In row 2, where the PNA backbone length increases from 2→3 (24→36 base-pairs), the ligand content increases by a factor of 1.5 for each entry. In row 3, where the PNA backbone length increases from 3→4 (36→48 base-pairs), the ligand content increases by a factor of 1.33 for each entry.

**P-Values for Complex Comparison**

In order to assess the validity in comparing data sets, p-values were determined. These values were determined using Microsoft Excel. It is typically accepted that values < 0.05 are statistically significant, and values < 0.001 are statistically highly significant.
|   | A1   | A2   | A3   | A4   | A5   | B1   | B2   | B3   | B4   | B5   | C1   | C2   | C3   | C4   | C5   |
|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|   | 0.0161 | 0.0131 | 0.0126 | 0.0124 | 0.0122 | 0.0116 | 0.0115 | 0.0115 | 0.0115 | 0.0118 | 0.0113 | 0.0114 | 0.0113 | 0.0113 | 0.0113 |
| 1000'0 | 0.0161 | 0.0099 | 0.0041 | 0.0039 | 0.0030 | 0.0014 | 0.0012 | 0.0013 | 0.0014 | 0.0018 | 0.0011 | 0.0012 | 0.0010 | 0.0010 | 0.0010 |
| 2500'0 | 0.0131 | 0.0099 | 0.0405 | 0.0126 | 0.2022 | 0.0542 | 0.0016 | 0.0011 | 0.0012 | 0.0017 | 0.0029 | 0.0008 | 0.0010 | 0.0007 | 0.0007 |
| 5000'0 | 0.0126 | 0.0041 | 0.0126 | 0.0405 | 0.2022 | 0.1975 | 0.0252 | 0.0131 | 0.0160 | 0.0229 | 0.0945 | 0.0069 | 0.0107 | 0.0059 | 0.0058 |
| 10000'0 | 0.0122 | 0.0030 | 0.0192 | 0.0542 | 0.1975 | 0.2520 | 0.1050 | 0.2071 | 0.3270 | 0.0145 | 0.0106 | 0.0531 | 0.0071 | 0.0064 | 0.0064 |
| 20000'0 | 0.0116 | 0.0014 | 0.0016 | 0.0001 | 0.0030 | 0.1050 | 0.1928 | 0.3160 | 0.0012 | 0.0145 | 0.2285 | 0.0080 | 0.0145 | 0.0080 | 0.0080 |
| 40000'0 | 0.0115 | 0.0013 | 0.0012 | 0.0000 | 0.0036 | 0.2071 | 0.1928 | 0.4709 | 0.0014 | 0.0028 | 0.0648 | 0.0020 | 0.0028 | 0.0648 | 0.0020 |
| 80000'0 | 0.0115 | 0.0014 | 0.0017 | 0.0004 | 0.0058 | 0.3270 | 0.3160 | 0.4709 | 0.0218 | 0.0663 | 0.2033 | 0.0439 | 0.0425 | 0.0663 | 0.2033 |
| 160000'0 | 0.0115 | 0.0014 | 0.0017 | 0.0004 | 0.0058 | 0.3270 | 0.3160 | 0.4709 | 0.0218 | 0.0663 | 0.2033 | 0.0439 | 0.0425 | 0.0663 | 0.2033 |

Chart S11: P-values for L-PNA:DNA multivalent landscape.
|       | $B_{2,3,1}$ | $B_{6,10,1}$ | $B_{2,10,1}$ | $B_{1,14,1}$ |
|-------|-------------|--------------|--------------|-------------|
| $B_{2,3,1}$ | -           | 0.0276       | 0.0197       | 0.4606      |
| $B_{6,10,1}$ | 0.0276      | -            | 0.1937       | 0.0412      |
| $B_{2,10,1}$ | 0.0197      | 0.1937       | -            | 0.0280      |
| $B_{1,14,1}$ | 0.4606      | 0.0412       | 0.0280       | -           |

**Chart S12.** P-values for linker spacing study.
|    | C4  | C3  | C2  | C1  |
|----|-----|-----|-----|-----|
| 0.0010 | 0.0035 | 0.0007 | 0.0113 | 0.1408 |
| 0.0023 | 0.0001 | 0.0010 | 0.0245 | 0.0006 |
| 0.0007 | 0.0003 | 0.0001 | 0.0485 | 0.0040 |
| 0.0066 | 0.0054 | 0.0001 | 0.2062 | 0.0077 |
| 0.0009 | 0.0007 | 0.0013 | 0.0076 | 0.4246 |
| 0.0029 | 0.0001 | 0.0076 | 0.0029 | 0.0010 |
| 0.0007 | 0.0004 | 0.0003 | 0.0001 | 0.4697 |
| 0.0014 | 0.0006 | 0.0001 | 0.0000 | 0.0003 |
| 0.0054 | 0.0025 | 0.0011 | 0.0277 | 0.0001 |
| 0.0010 | 0.0077 | 0.0151 | 0.0116 | 0.0089 |
| 0.0028 | 0.0014 | 0.0040 | 0.0027 | 0.0018 |
| 0.0009 | 0.0025 | 0.0026 | 0.0023 | 0.0115 |
| 0.0029 | 0.0001 | 0.0029 | 0.0031 | 0.0133 |
| 0.0007 | 0.0003 | 0.0001 | 0.0008 | 0.1955 |
| 0.0037 | 0.0018 | 0.0028 | 0.0029 | 0.0146 |
| 0.0018 | 0.0060 | 0.0170 | 0.0160 | 0.0089 |
| 0.0038 | 0.0195 | 0.0077 | 0.0080 | 0.1951 |
| 0.0029 | 0.0014 | 0.0040 | 0.0027 | 0.0018 |
| 0.0009 | 0.0025 | 0.0026 | 0.0023 | 0.0115 |
| 0.0029 | 0.0001 | 0.0029 | 0.0031 | 0.0133 |
| 0.0007 | 0.0003 | 0.0001 | 0.0008 | 0.1955 |
| 0.0037 | 0.0018 | 0.0028 | 0.0029 | 0.0146 |
| 0.0018 | 0.0060 | 0.0170 | 0.0160 | 0.0089 |
| 0.0038 | 0.0195 | 0.0077 | 0.0080 | 0.1951 |
| 0.0029 | 0.0014 | 0.0040 | 0.0027 | 0.0018 |
| 0.0009 | 0.0025 | 0.0026 | 0.0023 | 0.0115 |
| 0.0029 | 0.0001 | 0.0029 | 0.0031 | 0.0133 |
| 0.0007 | 0.0003 | 0.0001 | 0.0008 | 0.1955 |
| 0.0037 | 0.0018 | 0.0028 | 0.0029 | 0.0146 |
| 0.0018 | 0.0060 | 0.0170 | 0.0160 | 0.0089 |
| 0.0038 | 0.0195 | 0.0077 | 0.0080 | 0.1951 |
| 0.0029 | 0.0014 | 0.0040 | 0.0027 | 0.0018 |
| 0.0009 | 0.0025 | 0.0026 | 0.0023 | 0.0115 |
**Theoretical Model**

In this section, we develop a theoretical model to study the dominant interactions of the L-PNA:PNA linker ligand protein system. The model is coarse-grained to allow an efficient evaluation of the density of states and subsequently, equilibrium predictions of protein oligomerization states over the surface the membrane. We show that the placement of the linkers along the PNA can selectively probe the underlying surface distribution of the proteins on the membrane. Comparisons of different configurations of linker placement act as an entropic microscope since the fundamental enthalpic interactions are similar for all B L-PNA systems. The model is broken up into three sections: the distribution proteins on the membrane, the conformational distribution of the PNA:PNA linker-ligand complex (L-PNA:PNA), and the combination of these two ensembles.

**Modeling Membrane Proteins**

Since the protein is mostly buried within the membrane, we model the interaction site as a cylinder whose dimensions are shown in Figure S23. Only the topmost portion of the protein presents a surface to interact with the ligand. The membrane is mostly flat at the length scale of the PNA, hence the distribution of the proteins across the membrane surface can be considered two-dimensional. While the idealized protein surface is slightly elliptical, we model it’s area as a circle whose diameter is $\sigma_p = 32.0\text{Å}$ with a binding pocket of diameter $\sigma_{\text{binding}} = 12.5\text{Å}$.

We postulate that the only coordinated structures formed are monomers and dimers. Higher-order oligomers may exist, but modeling is problematic since they can present themselves through many different local configurations, such as a linear chain or a globular cluster. Since these extended configurations require additional assumptions on the model, for the present work we restrict ourselves to these oligomerization states.

Each individual protein occupies an area of $A_p = \pi\sigma_p^2/4$ over the surface of the membrane. The packing fraction, determined by the surface density $\rho$, is $\phi = A_p\rho$. This quantity
is unknown but can be estimated to be approximately $\phi = 0.20$. The model was tested at higher and lower densities, $\phi = 0.10, 0.30$ and found to be relatively insensitive. The distribution of the proteins is controlled by $f_D$, the fraction of dimers in the total population

$$f_D = \frac{2N_D}{N_M + 2N_D} \quad (1)$$

where $N_M, N_D$ represent the number of monomers and dimers respectively. Over a square representing the membrane of area $A_M$, the packing fraction $\phi = (N_M + 2N_D)A_P/A_M$ and a value of $f_D$ fix the number of monomers and dimers for a numerical simulation.

We are using the L-PNA:PNA complex to probe the surface of the membrane, and indirectly, measure $f_D$. To do so we need to generate ensembles of protein distributions for different values of $f_D$. With a sample area of $A_M = 10\sigma^2$, we subdivide the valid range of
the dimer fraction $f_D \in [0, \ldots, 1]$ into $10^3$ equal intervals and generate $10^6$ ensembles for each interval, see Figure S24. For each set of ensembles we compute the radial distribution function of the proteins as a function of the dimer fraction $g(r; f_D)$ to be used in the next part of the calculation.

Each distribution is assigned an energy proportional to the number of dimers. This dimerization energy $h_D$ contributes to the total energy of the system

$$U_D = \phi h_D f_D$$

At equilibrium, this contribution to the Boltzmann weight will favor particular PNA configurations over others, leading to a ligand-enhanced dimerization.

**L-PNA:PNA complex**

For this work we model only the L-PNA:PNA complex with 12 base-pairs and up to two sites for two mini-peg linkers (8-amino-3,6-dioxaoctanoic acid) and XAC to be attached. These positions along PNA sequence are indexed as $k_1, k_2$ for a B L-PNA system and simply $k_1$ for the A L-PNA system shown in Figure S25. The L-PNA:PNA backbone is modeled as a rigid, all-atom structure with only center of mass translational and rotational degrees of freedom. The PNA duplex interacts with the membrane only through steric repulsion. The location of the PNA duplex is parameterized by $(z, \alpha)$, where $z$ is the height above the membrane and $\alpha$ is a vector of the three Euler angles.

The mpeg linkers themselves are modeled as self-avoiding walks. Starting at the contact site from the PNA backbone, the linker takes 13 steps of approximately 2.5Å at a random angle of no more than 120° degrees measured from the previous step. The walk is rejected if any step comes within 3Å of either the PNA duplex or any previous step (excluding the most recent step). The PNA duplex is volume excluded from itself and the membrane. For each conformation $(z, \alpha)$ of the PNA duplex we sample a series of walks from the starting points of the linkers. In this way we can fully sample the geometric component of the linkers entropic contribution to the system.
Figure S24: Schematic illustration of protein distributions for different values of $f_D$. Values from the upper left to the lower right represent the range of $f_D$ from one to zero respectively. Two green dots signify a pair of proteins in the dimer state while a single dot represents a monomer.

We assume a protein will bind to the ligand only if two conditions are met: the end point of the walk is within 5Å of the membrane and the binding pocket of a protein is present at that location. For a B L-PNA system, there are three situations for the endpoints of the linkers corresponding to the number of linkers found near the membrane. We denote the probability of each of these aggregate conformations as $c_0, c_1, c_2$, corresponding to zero, one or two linkers in contact with the membrane respectively.

For each specific conformation in $c_2$ we need the additional measurement of the linker-linker endpoint separation, $p(r)$. Shown in Figure S26, the linkers which start close along the chain sequence $\Delta k = |k_2 - k_1| < 5$ create a sharp peak at a small distance much less than
Figure S25: Illustration of two possible sidechain locations for a B L-PNA system. In theory sidechains can be built at any of the available $k_1, k_2$ sites, leading to a possible 132 combinations for unique side chains and possible 66 combinations for identical sidechains.

the diameter of the protein. For these small $\Delta k$, there is an entropic penalty for the complex to bind to the dimer since the peak of $p(r)$ is in the range where only a single protein, and hence only a single receptor, can be found due to excluded volume. Conversely when $\Delta k$ is larger, the distribution flattens preferring longer separations and a different distribution of proteins on the membrane. This suggests that the placement of the linkers, $k_1, k_2$, can selectively probe the underlying surface distribution.

Combining the ensembles

Since electrostatic effects are small, the dominant enthalpic interaction of the system will be the interaction of the ligand to the binding site of the protein and the dimerization of the proteins. We let the the strength of the monomeric interaction be $J$, while the interaction energy of two fully complexed ligands to a dimer be $2J + h_B$. Here $h_B$ is a cooperative binding energy reflecting any conformational changes to the binding sites in the dimeric state. For this work, we make the assumption that such allosteric effects are not present, hence we let $h_B = 0$. 
Figure S26: Radial distribution for two linkers $p(r)$ in the B L-PNA system given that both linker endpoints are within contact distance of the membrane. Results are numerically sampled from a Monte-Carlo computation for all values of $k_1, k_2$. The values are colored by $\Delta k = |k_2 - k_1|$ ranging from black to orange over smaller to larger $\Delta k$ values. The dashed lines indicate the range of dimerization.

A B L-PNA system will have three energy levels $U_{L0} = 0$, $U_{L1} = J$, $U_{L2} = 2J$. To complete a statistical mechanical description of the system we need the degeneracy of each state, i.e. the entropic contribution of each energy level codified in the density of states. To do this, we combine the ensembles of the protein distribution and the self-avoiding walks of the PNA-linkers. The model implicitly assumes that the two ensembles, the membrane proteins and the L-PNA:PNA, are separable. In addition, we only consider a single PNA duplex at one time.

When two linkers are in contact range of the membrane, the calculation requires us to consider the distribution of the membrane proteins to properly count the probability of both sites binding, see Figure 4a in the main text. The probability for the first linker to find a protein is proportional to $\phi$. Given that a protein was found at the location of the first linker,
the probability for the second linker to be found at a distance \( r \) is modified by a factor of \( g(r, f_D) \). For a fixed dimer fraction \( f_D \), the probability \( P_n \) that \( n \) linkers will bind is

\[
P_0 = \int [(1 - \phi)(1 - g(r, f_D)\phi)] p(r) \, dr \tag{3}
\]

\[
P_1 = \int 2[\phi(1 - g(r, f_D)\phi)] p(r) \, dr \tag{4}
\]

\[
P_2 = \int [g(r, f_D)\phi^2] p(r) \, dr \tag{5}
\]

remembering that \( p(r) \) is the probability that two linkers will find the membrane and be separated by a distance \( r \) and \( g(r; f_D) \) is the radial distribution of the proteins at a dimer fraction \( f_D \). If the linker separation distance is smaller than the protein diameter \( r < \sigma \), at most only one protein can bind due to excluded volume. This is reflected in Equation 5 where \( g(r < \sigma) = 0 \) making this portion of the integral zero.

The density of states of a B L-PNA system for a particular protein ensemble \( g(r, f_D) \) is given by

\[
w_0 = c_0 + c_2P_0 \tag{6}
\]

\[
w_1 = c_1 + c_2P_1 \tag{7}
\]

\[
w_2 = c_2P_2 \tag{8}
\]

From the density of states we can explicitly compute the partition function using the total energy as \( E_n = U_{Ln} + U_D = U_{Ln} + \phi h_D f_D \) by weighting over each ensemble

\[
Z = \int_0^1 \left[ \sum_{n=0}^2 w_n e^{-E_n/kT} \right] \, df_D \tag{9}
\]

and the fraction of ligands bound for a B L-PNA system

\[
\langle b \rangle_B = (2Z)^{-1} \int_0^1 \left[ \sum n w_n e^{-E_n/kT} \right] \, df_D \tag{10}
\]

The weighting parameters \( w_n \) were determined by Monte-Carlo simulations.
An analogous expression $\langle b \rangle_A$ exists for the A L-PNA system. Here, the expression is simpler since the probabilities are independent of each other. In this case $g(r) = 1$ for all $r$ and $p(r)$ is a constant over the range of integration. The equivalent values for $c_i$, i.e. the probability that zero, one or two linkers will be in contact in the membrane is also separable since the chains are independent. The density of states for the A L-PNA system is $w'_0 = c'_0(1 - \phi)^2$, $w'_1 = 2c'_1c'_0\phi(1 - \phi)$, and $w'_2 = c'_1\phi^2$ where the primes denote values that differ from the B L-PNA system.

**Results**

To relate the results of the numerical model to our experimental results we compare the ratios of the IC$_{50}$ values with the inverse of the ratios of the bound ligand fractions $\langle b \rangle$. These unitless values are defined as

$$R^{\text{exp}}_{ij,i'j'} = \frac{\text{IC}_{50}(i,j)}{\text{IC}_{50}(i',j')}$$ (11)

$$R^{\text{sim}}_{ij,i'j'} = \frac{\langle b \rangle_{(i',j')}}{\langle b \rangle_{(i,j)}}$$ (12)

for the experimental and simulated results respectively.

The free parameters, $J$ and $h_D$, are best-fit using the values of the three IC$_{50}$ ratios of the three B L-PNA systems with $(k_1,k_2) = \{(2,3),(2,6),(2,10)\}$ and the single A L-PNA system at $k_1 = \{(6)\}$. These optimal parameters were obtained $J/kT = -3.2$ and $h_D/kT = -1.4$ corresponding to a fraction of dimerized species of about $f_D \approx 0.9$, see Table S14. The mean error for a given set of system parameters was the average of all ratios of IC$_{50}$ values to those of the corresponding inverse ratios of the bound ligand fractions.

Since $h_D$ directly controls the number of dimerized species, we can examine how changing the value of $h_D$ will change the resulting fit errors. For the three different values of the protein packing fraction $\phi$, we examined how the error changed as a function of the dimer fraction, shown in Figure S27. In all three cases for $f_D > 0.8$, the model was insensitive to a greater
| L-PNA | \((k_1, k_2)\) | IC\(_{50}\) (nM) | \(\langle b \rangle\) |
|-------|-----------------|------------------|-----------------|
| A\(1_p\) | (6,) | 1070 | 0.1788 |
| B\(1_p\) | (2,3) | 383 | 0.5189 |
| B\(1_p\) | (2,6) | 220 | 0.5953 |
| B\(1_p\) | (2,10) | 240 | 0.5955 |

**Chart S14**: Empirical IC\(_{50}\) values for the four different conformations studied. Smaller IC\(_{50}\) values and larger binding fractions \(\langle b \rangle\) indicate stronger binding affinity. The table shows values computed at \(J/kT = -3.2\) and \(h_D/kT = -1.4\) at a dimer fraction of \(f_D = 0.90\).

degree of dimerization. However, for all values less than 0.8 the error rose monotonically, albeit non-linearly.

The results from the model strongly suggest a qualitative prediction that the number of dimerized species must be large. Since the most accurate prediction is still around 20% error, a quantitative prediction require a larger set of measurements. This suggests a new round of experiments for future study. From the view of an entropic microscope, measurements of different values of \((k_1, k_2)\) or varied linker length would modify the density of states but not necessarily the energetics. In addition, measurements of temperature variation for fixed linker configuration would allow for a more accurate estimation of the model parameters.
Figure S27: Mean error (see text) of the model as a function of the dimer fraction for three different packing fractions of proteins on the membrane. In all cases, a large dimer fraction $f_D > 0.8$ was the best predictor of the empirical data.
References and Notes:

(49) Englund, E. A.; Appella, D. H. *Org. Lett.* **2005**, *7*, 3465-3467.

(50) MacKerell, A. D. J.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257-265.

(51) Molecular Operating Environment (MOE), version 2012.10. Chemical Computing Group, Inc. Montreal, Canada.

(52) Schrödinger Suite 2012. Schrödinger, LLC, New York, NY.

(53) Pierce, B. G.; Hourai, Y.; Weng, Z. *PLoS One* **2011**, *6*, e24657.

(54) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D. J.; Pastor, R. W. *J. Phys. Chem. B* **2010**, *114*, 7830-7843.