Targeting VEGF with LNA-Stabilized G-rich Oligonucleotide for Efficient Breast Cancer Inhibition

Stacey L. Edwards,[1,2] Vasanthanathan Poongavanam,[3] Jagat R. Kanwar,[4] Kislay Roy,[4] Kristine M. Hillman,[1,2] Neerati Prasad,[5] Rikke Leth-Larsen,[6] Michael Petersen,[3] Maja Marušič,[7] Janez Plavec,[7] Jesper Wengel,[3] and Rakesh N. Veedu[1,3,8]*.

[1]School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia 4072; [2]QIMR Berghofer Medical Research Institute, Brisbane, Australia 4006; [3]Nucleic Acid Center and Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense M, Denmark 5230; [4]Nanomedicine-Laboratory of Immunology and Molecular Biomedical Research (NLIMBR), School of Medicine (SoM), Molecular and Medical Research (MMR) Strategic Research Centre, Faculty of Health, Deakin University, Pigdons Road, Waurn Ponds, Geelong, Victoria 3217; [5]Department of Pharmacology, Kakatiya University, Warangal 506009 Andhra Pradesh, India. [6]Institute of Molecular Medicine, Department of Cancer and Inflammation Research, University of Southern Denmark, Odense C, Denmark 5000; [7]Slovanian NMR Institute, National Institute of Chemistry, Ljubljana, Slovenia SI-1001; [8]Center for Comparative Genomics, Murdoch University, Murdoch, Perth 6150 Australia.

Cell lines and transient transfections: MCF-7, MDA-MB-231 and HS578T breast cancer cell lines and HeLa cervical cancer cell lines were grown in RPMI medium (Life Technologies) with 10% fetal calf serum (FCS) and antibiotics. MCF10A normal breast epithelial cell lines were grown in DMEM/F12 medium (Life Technologies) with 5% horse serum (HS), 10 µg/ml insulin, 0.5 µg/ml hydrocortisone, 20 ng/ml epidermal growth factor (EGF), 100 ng/ml cholera toxin and antibiotics. HUVEC (Human Umbilical Vein Endothelial Cells) were grown in EGM-2 media (Lonza). For transient transfection assays, aptamers were reverse transfected into cells using Lipofectamine 2000 (Life Technologies) in accordance with the manufacturer's protocol. Forty-eight hours after transfection, cell viability was measured by MTT assay or replated for colony formation assays.

VEGF ELISA assay: The supernatant from transiently transfected MCF-7 cells was collected, centrifuged, and VEGF levels were measured using the Human VEGF Quantikine ELISA Kit (R&D Systems) according to the manufacturer's instructions. The assay was conducted in biological triplicates and as technical duplicates.

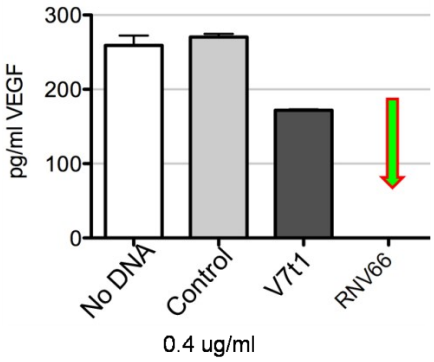


Figure S1. VEGF-ELISA using RNV66 and MCF-7 breast cancer cells *in vitro*.

Colony formation assays: Transiently transfected MCF-7 cells were seeded in 6 well plates at a concentration of 5,000 cells per well, then normal media replaced every four days. After 14 days, colonies were visualized by crystal violet staining.

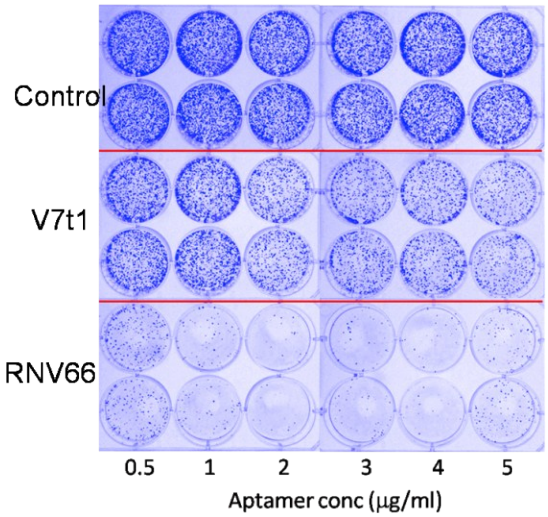


Figure S2. Colony formation assay using RNV66 and MCF-7 breast cancer cells *in vitro*.

Cell viability assay using MDA-MB-231and HS578T: MDA-MB-231 and HS578T (1.5×10^5 cell/ml) cells were transiently transfected as described above in 96 well plates in quadruplicate. Forty-hours after transfection, 10 µl of the 12 mM MTT stock reagent was added to each well and the plate incubated at 37 ^{o}C for 4 hours. 100 µl of SDS-HCl solution was then added the each well, mixed thoroughly and the plate incubated at 37 ^{o}C for 4 hours. Each sample was mixed again and the absorbance at 570 nm was measured to determine cell viability.

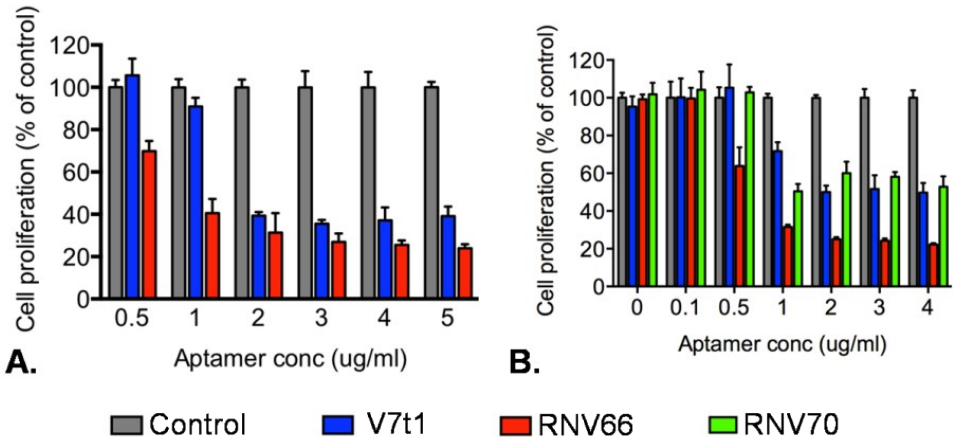


Figure S3. MTT-based cell proliferation assay using RNV66 against HS578T (A) and MDA-MB-231 (B) cells.

Confocal microscopy: Cells were seeded in 6 well plates, once confluent they were transfected using the transfection reagent (Santa Cruz) with 10 nmol/ml concentration of

RNV66 and 10 nmol/ml of scrambled aptamer. An untreated well was maintained as control. The cells were incubated for a period of 6h with the aptamers and then washed with phosphate buffer saline (PBS), fixed using 4% PF for 20 min at 37 ^{o}C. Cells were permeabilized using 0.01% triton-X100 for 5 min. Cells were further blocked with 3% BSA for 30 min. The cells were washed and incubated with primary antibody (1:100) (mouse anti-VEGF, Santa Cruz) for 1h at 37 ^{o}C. Post washing thrice with PBS the cells were incubated with fluorescein isothiocyanate (FITC) conjugated secondary antibody (1:100) (anti-mouse-FITC, Sigma Aldrich) for 1h in dark. The cells were washed and mounting media with 4',6-diamidino-2-phenylindole (DAPI) was added to the slide. The slide was analysed in Leics Tcs SP5 laser immunoconfocal microscope.

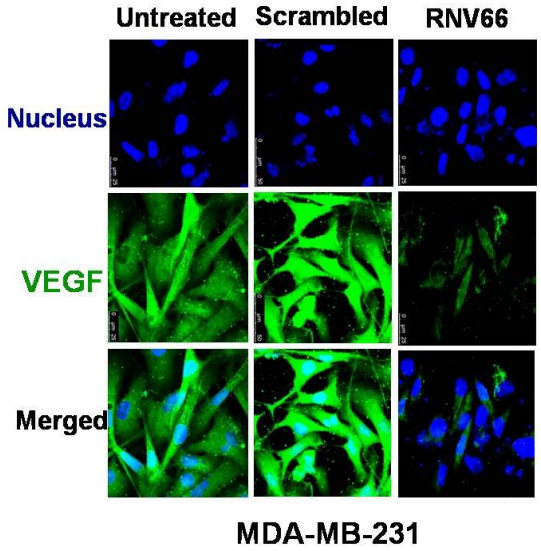


Figure S4. VEGF binding specificity on MDA-MB-231 cells after treating with RNV66.

RNV66 in the media without transfection: Cell viability assay using MCF7, MDA-MB-231 and HUVEC: 1x10^{3} or 5x10^{3} (HUVEC) cells were plated in 96 well plates in quadruplicate in culture medium containing heat-inactivated FCS (62^{o}C for 30 mins) and 0-40 µg/ml aptmers. Forty-hours after aptamer treatment, 10 µl of the 12 mM MTT stock reagent was added to each well and the plate incubated at 37^{o}C for 4 hours. 100 µl of SDS-HCl solution was then added the each well, mixed thoroughly and the plate incubated at 37^{o}C for 4 hours. Each sample was mixed again and the absorbance at 570 nm was measured to determine cell viability.

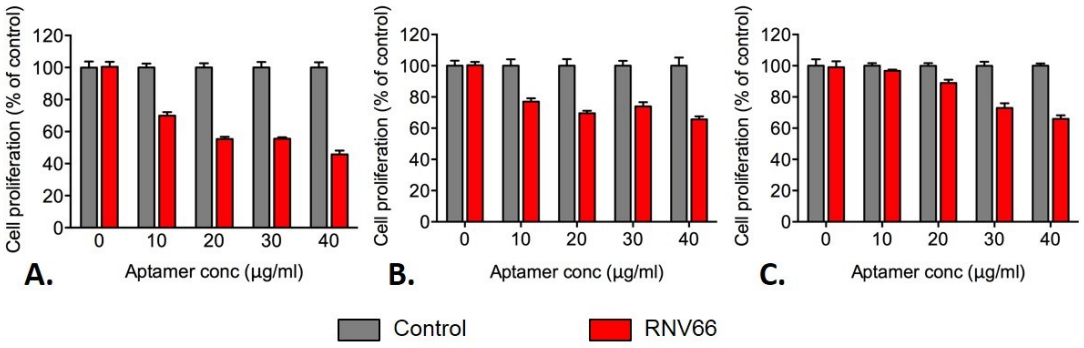


Figure S5. MTT-based cell proliferation assay using RNV66 in media against MCF7 (1), MDA-MB231 (B), and HUVEC (C) cell lines.

HeLa Cell Proliferation: HeLa Cell Proliferation: HeLa ($7.5x10^4$ cell/ml) cells were transiently transfected or aptamers added to the media as described above in 96 well plates in quadruplicate. Forty-hours after transfection, 10 µl of the 12 mM MTT stock reagent was added to each well and the plate incubated at 37 ^oC for 4 hours. 100 µl of SDS-HCl solution was then added the each well, mixed thoroughly and the plate incubated at 37 oC for 4 hours. Each sample was mixed again and the absorbance at 570 nm was measured to determine cell viability.

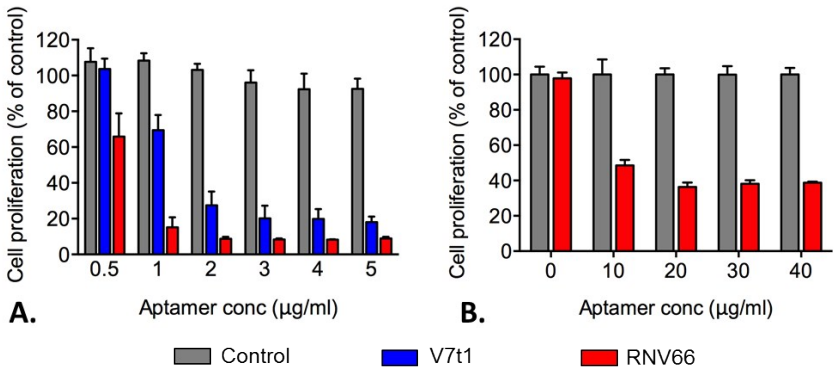


Figure S6. MTT-based cell proliferation assay using RNV66 transfected into (A) or added to the media of (B) HeLa cells.

Migration assay:
Cells were plated in culture inserts in 24 well plates and treated with RNV66 (10 nmol/ml) and scrambled aptamer (10 nmol/ml) for a period of 12h. An untreated well was maintained as a control. After incubation the treatment was removed and insert was fixed with 4% PF for 20 minutes followed by washing with 1×PBS and staining with 0.2% crystal violet for 20 minutes followed by washing again. The inserts were then viewed using inverted microscope for observing the migratory efficacy of the cells.

LDH release assay: Cells were seeded in 96 well plates. Once the cells were confluent they

were treated with RNV66 (10 nmol/ml) and scrambled control aptamer (10 nmol/ml) for a period of 24h. An untreated well was maintained as a control. The LDH reagent was added to cells as guided by the cytotoxicity detection kit (Invitrogen). The OD was measured at 595 nm and percentage cytotoxicity was calculated.

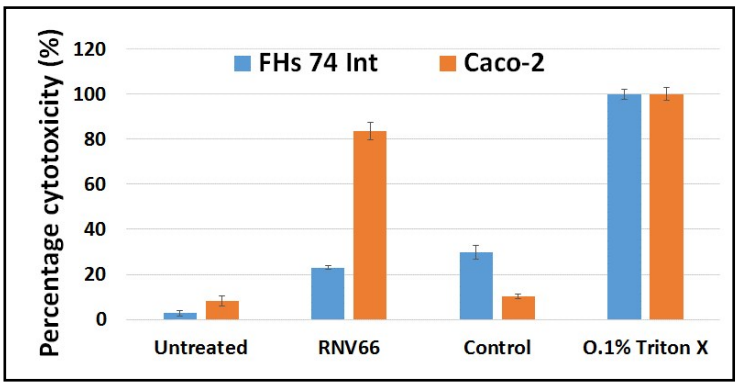


Figure S7. Cytotoxicity assay using normal normal intestinal epithelial cells (FHs 74 Int) and colon adenocarcinoma cells (Caco-2).

Serum stability analysis: The RNV66 aptamer and the scrambled aptamer were dissolved in DNAse and RNAse free water and stored at -20 °C. The serum stability of the aptamer RNV66 was tested in PBS containing 10% serum and incubated for 0 to 48h. The products were then run on 1% agarose gel and visualized using the gel-dock imaging system (Bio-Rad).

***In vivo* analysis of RNV66:**

4T1 cells were used to establish the human breast cancer model by subcutaneously injecting $2X10^5$ cells in mammary pads of female mice. Five-six week old female (*Mus musculus* nude mice-BALB/c nu/nu) were used in the study (n=6). After 45 days of tumor growth, individual groups were treated with PBS+chitosan, a non-VEGF targeting control sequence, taxol+doxorubicin (known drug combination as positive control), RNV66 (20 µg), RNV66 (40 µg) and RNV66+Taxol+doxorubicin respectively by intra-tumoral injection. Chitosan-based nanoformulation was performed for all test candidates. Regular checks were conducted for any sign of physiological or physical distress and mice were weighed thrice a week. At the end of the experimental period the mice were euthanized. Vernier callipers were used to measure the tumour size every week.

Synthesis of LNA and UNA-modified Oligonucleotides: The LNA and UNA oligonucleotides were synthesized on a DNA synthesizer via standard phosphoramidite chemistry in 10 µmol scale. The synthesized oligonucleotides were deprotected and cleaved from the solid support by treatment with NH_4OH at 55 ^oC overnight. The crude oligonucleotides were then purified by RP-HPLC, desalted and verified by MALDI-ToF MS

analysis. Oligonucleotide sequences used in our experiments are shown in Table S1.

Table S1. List of oligonucleotide sequences synthesized and tested in this study. LNA nucleotides are represented as bold underlined red letter containing the superscript 'L', and UNA nucleotides are represented as bold underlined letter containing the superscript 'u'

Name	Sequence	Length
V7t1	5′-TGTGGGGGTGGACGGGCCGGGTAGA-3′	25
RNV64	5′-TGTGGGGGTGGACGGGCCGGGTA**G^{L}**A-3′	25
RNV65	5′-TGTGGGGGTGGACGGGCCGG**G^{L}**TA**G^{L}**A-3′	25
RNV66	5′-TGTG**G^{L}**GGGTGGACGGGCCGG**G^{L}**TA**G^{L}**A-3′	25
RNV67	5′-TGTGGGG**G^{u}**TGGACGGGCCGGGTA**G^{L}**A-3′	25
RNV68	5′-TGTGGGGGT**G^{u}**GACGGGCCGGGTA**G^{L}**A-3′	25
RNV69	5′-TGTGGGGGTGG**A^{u}**CGGGCCGGGTA**G^{L}**A-3′	25
RNV70	5′-TGTGGGGGTGGA**C^{u}**GGGCCGGGTA**G^{L}**A-3′	25
RNV71	5′-TGTGGGGGTGGACGGGC**C^{u}**GGGTA**G^{L}**A-3′	25
RNV72	5′-TGTGGGGGTGGACGGGCCG**G^{u}**GTA**G^{L}**A-3′	25
Control	5′-TCATCGATGGCAGCTGCGTGTCGTT-3′	25

Characterisation of RNV66: As RNV66 was found to be more effective, we have only synthesised this oligo more often and in large scale. The ion-exchange HPLC profile and MALDI-ToF MS analysis data are given below in Figure S8 and Figure S9. The VEGF binding affinity of the DNA version of RNV66 (V7t1) was analysed by SPR and reported earlier (*Molecules,* 2010, **15**, 215). Using this published protocol, we have also perfored the binding analysis of RNV66 using SPR on a BIAcore 3000 and calculated as 8. 36 nM (Figure S10) and found no significant change in affinity.

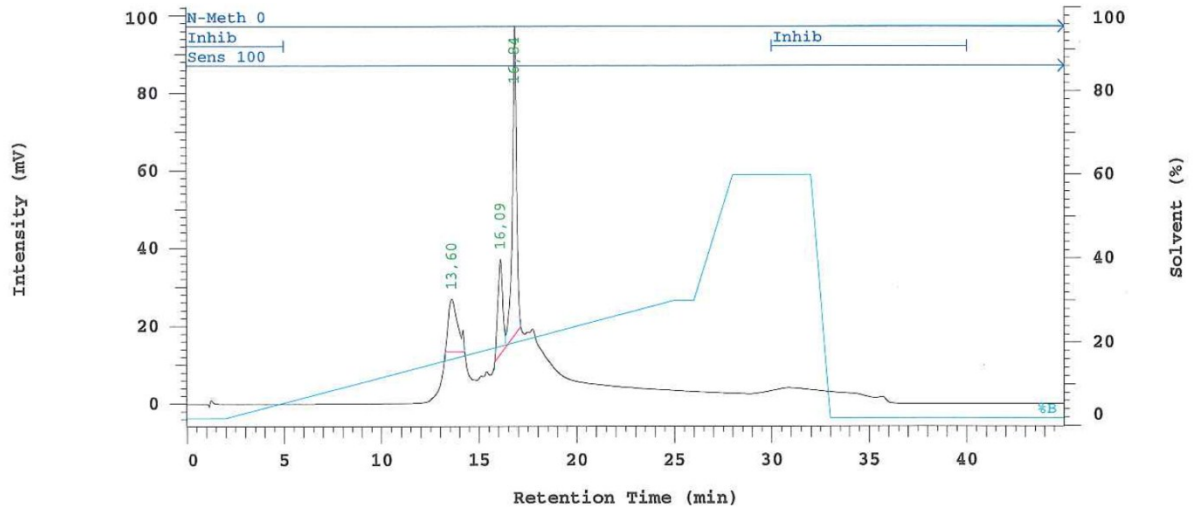


Figure S8. Ion-exchange-HPLC profile of RNV66. Multiple peaks are typical for G-quadruplex oligos.

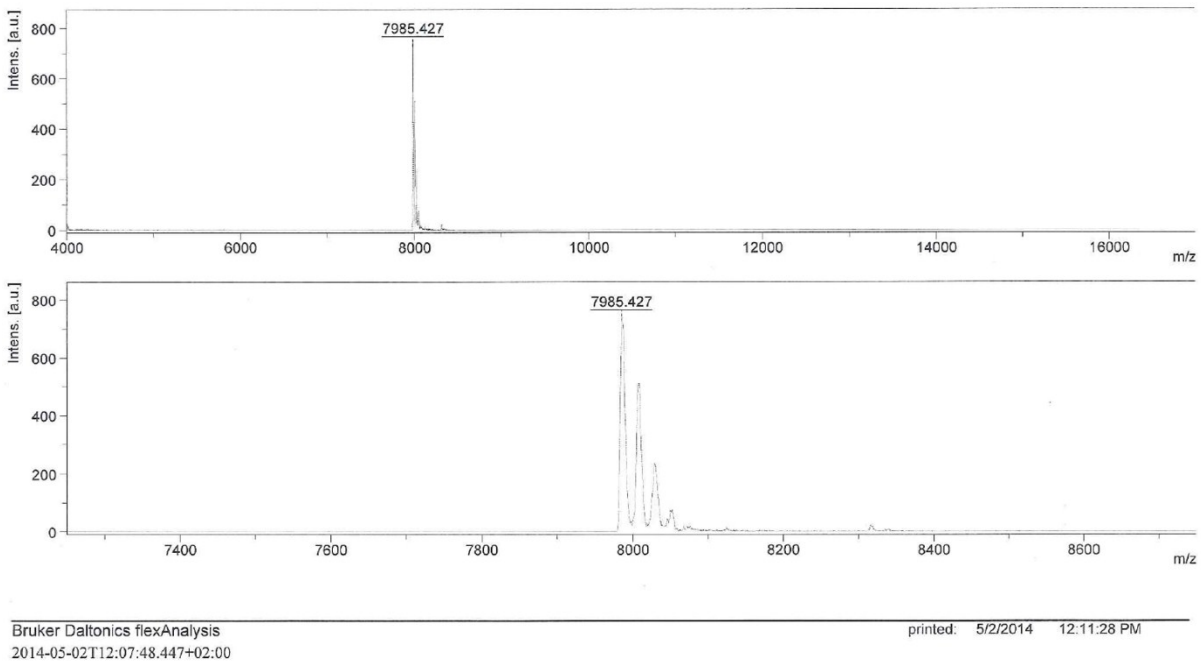


Figure S9. MALDI ToF MS analysis of RNV66. Calculated Mass: 7984.18; Found: 7985.427.

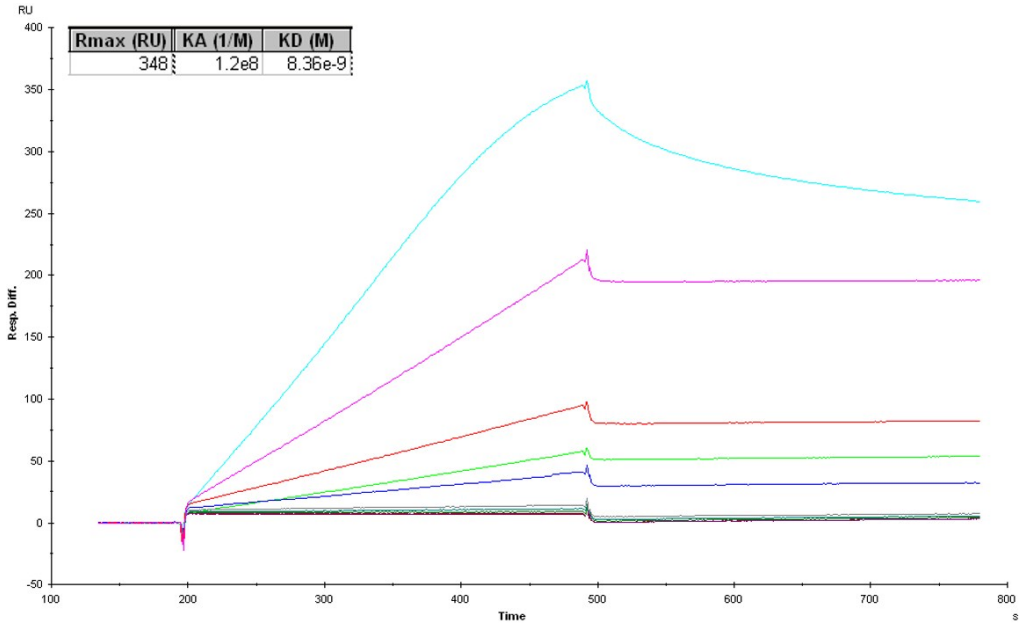


Figure S10. Affinity measurements of RNV66: VEGF binding by SPR. RU, response units.

Computational methods and materials

Binding interaction of RNV66 and VEGF-A by Molecular dynamic simulations.

Preparation of initial configuration of VEGF and RNV66: The atomic coordinates of the single 25 nt G-rich quadruplex aptamer (RNV66) targeting vascular endothelial growth factor (VEGF) were obtained from our previous NMR study (PDB ID: 2M53).[1] Of the 10

conformations of RNV66 structure, one conformation was used for the further structural refinement and molecular dynamic simulation studies. The atomic coordinates of human vascular endothelial growth factor (VEGF) complex with vascular endothelial growth factor receptor (VEGFR) as obtained by X-ray crystallography (3.20 Å resolution) were acquired from the Protein Data Bank (PDB ID: 3V2A).[2] Both structures were refined using the Protein Preparation Wizard[3] module available in the Schrödinger suite.[4] This optimization includes adding hydrogen atoms, assigning bond orders, building disulfide bonds (for VEGF and VEGFR) and placing potassium ions center in the G-quadruplex (for RNV66). The protonation states were predicted by PROPKA[5] provided in the Protein Preparation Wizard in the presence of the K^+ ions. Missing residues in VEGFR were added using the Prime 3.3 module[6] of Schrödinger suite. An optimized structure model was finally found by energy minimizing (only hydrogen atoms) using the OPLS2005 force field (Figure S6).

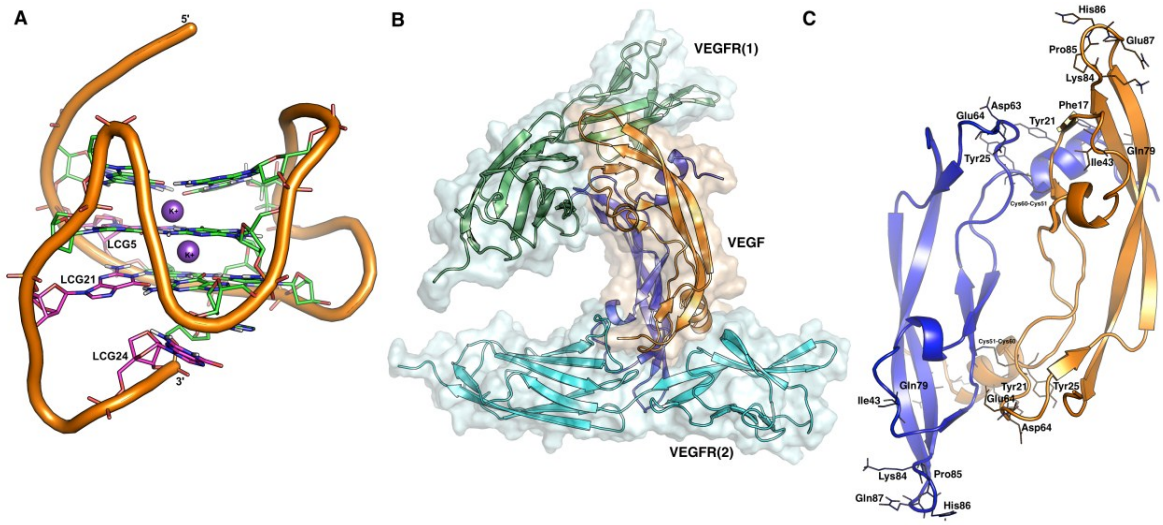


Figure S6: Cartoon representation of A: RNV66 (pdb id: 2m53), B: VEGF-VEGFR complex (pdb id: 3v2a), C: VEGF dimer (pdb id: 3v2a) are shown and important residues are highlighted. The RNV66 is highlighted with LNA and potassium ions (some residues are not shown for clarity).

Optimization of LNA residue of RNV66

Before the molecule dynamic simulations of RNV66, the LNA residue was geometry optimized with HF/6-31G** using Gaussian 09[7] and the atomic charges used for the molecular dynamics were calculated from the ESP (electrostatic potential) using B3LYP-IEFPCM/cc-pVTZ. These atomic charges were fitted using the RESP procedure as implemented in the Antechamber module of the Amber 14 software. [8]

Molecular dynamics simulation of complexes

All molecular dynamics simulations were performed and analyzed using the Amber 14 software. [8] After the geometry optimization and charge calculation for LNA residue, the *tleap* tool in the Amber suite was used to build coordinate and parameter (Amber ff14SB force field for protein and ff99bsc0 for DNA) for the RNV66. Subsequently, TIP3P water (solvent) molecules were added with a 10 Å buffering distance between the edges of the truncated octahedron box. Potassium ions were added to neutralize the charge (21 K^+). In order to avoid edge effects, periodic boundary conditions were applied during the molecular dynamics simulations. Energy minimization was carried out in two steps; first, the system was minimized using a steepest descent minimization with all heavy atoms restrained. The maximum number of minimization cycles was set to 1000. In the second stage of the minimization, the entire system was energy minimized and no positional restraints were applied at this stage. In the process of thermalization, initial velocities were generated from a Maxwell-Boltzmann distribution at 100 K and gradually increased to 300 K at constant volume over a 200 ps MD simulation. After the thermalization process, the system was equilibrated at constant temperature (300 K) and pressure (1 bar) using the Berendsen coupling algorithm[9] for another 500 ps MD simulation. After the equilibration step, the MD production run was started for 20 ns (10.000.000 MD steps with 0.002 time steps). The SHAKE algorithm was used to constrain the lengths of all bonds involving hydrogen atoms. Coordinates were saved every 10 ps from the 20 ns simulation for the analysis.

The binding free energy difference for the different complexes from molecular dynamic simulation was carried out systematically as shown in Figure S7A. Initially, the RNV66 was simulated for 20 ns and analyzed with respect to energy convergence in order to make sure that the structure is energetically stable. The RMSD (root mean square deviation) for the backbone remains low throughout the simulation compared to initial structure; therefore, we chose five random conformations for DNA-protein docking. Various docking and free energy scenarios used in this study are shown in Figure S7B**.**

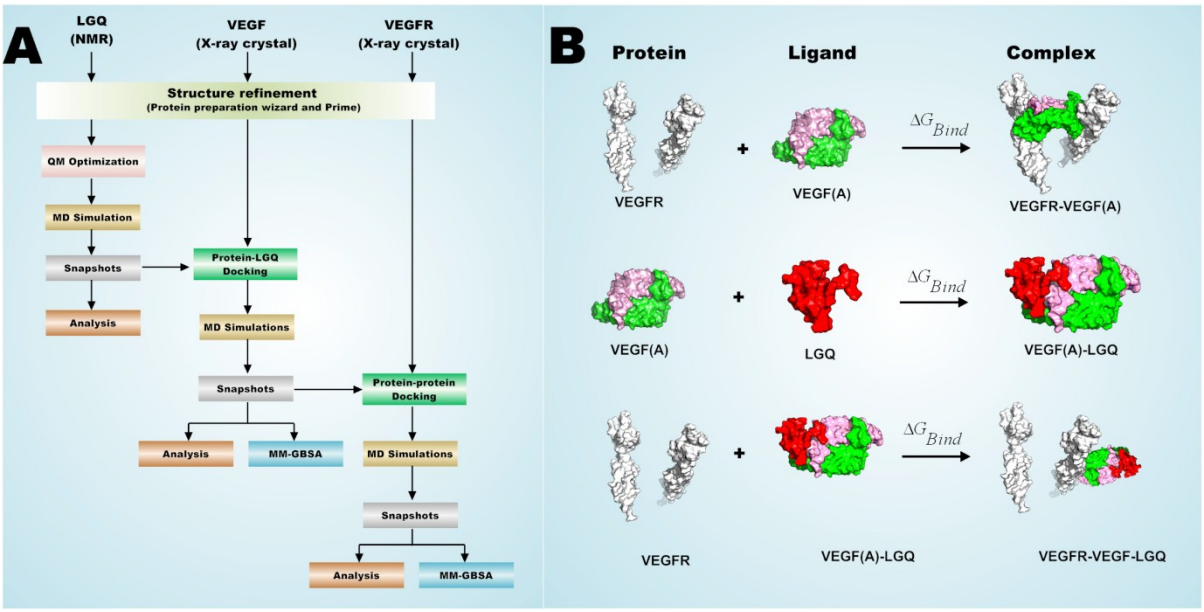


Figure S7. A: Overall workflow of computational calculations, B: Various docking and free energy scenarios used in this study are shown**.**

RNV66 docking into VEGF

The RNV66-VEGF docking was performed using the rigid-body protein-protein docking program ZDOCK.[10] Briefly, the ZDOCK program uses the Fast Fourier Transform algorithm to enable an efficient docking search on a 3D grid, and utilizes a combination of shape complementarity and electrostatics. Before RNV66 docking experiments, the docking protocol was assessed by comparing the conformation, position and orientation (the pose) of the VEGF in VEGFR as obtained from docking with the one determined experimentally with X-ray crystallography (pdb id: 3v2a). This assessment was set a minimum requirement to determine whether the program is applicable to this system or not. The root-mean-square-deviation (RMSD) was calculated between the X-ray co-crystallized conformation of VEGF and the docking solutions. Out of 10 best poses, the top-4 docking poses were able to reproduce the VEGF-VEGFR complex conformation obtained from the X-ray structure within 1Å (comparison of docking pose and experimental is shown in Figure S8).

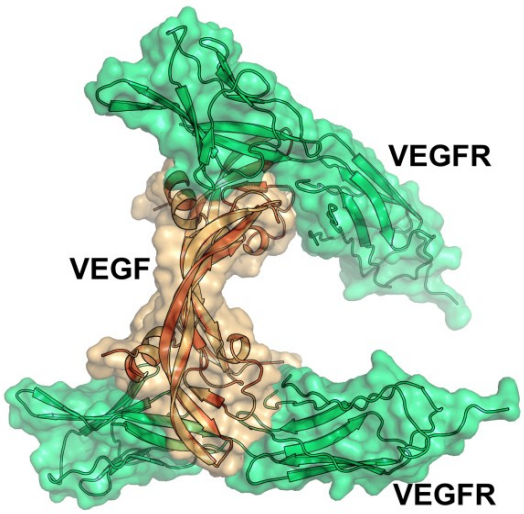


Figure S8. Comparison of docking pose (grey) with x-ray crystallized conformation (yellow) of VEGF.

Therefore, different conformations of RNV66 from MD simulations were docked into the VEGF model using the ZDOCK program. The top 10 binding poses for each conformation were saved and evaluated for the next steps. In most of the cases, the binding pose of RNV66, both 5th and 24th LNA (LNA-G5 and LNA-G24) of RNV66 binds to the VEGF. Comparing these poses, the pose with LNA-G24 of RNV66 interacts with VEGF through nucleobase and phosphate backbone. On the other hand, the pose with LNA-G5 of RNV66 binds to VEGF interacts only through phosphate backbone, since the nucleobase is involved in G-tetrad formation. However, for the binding free energy calculation, both types of poses were used. Interestingly, one of the 50 poses of RNV66-VEGF, the RNV66 binds to the other end of the VEGF (where it binds to VEGFR domain 2), therefore this pose was also chosen for free energy calculations. Various RNV66-VEGF poses investigated using the MM-GBSA method are shown in Figure S9 and summary of binding free energies obtained from various complexes is provided in Table S2.

Table S2. Summary of binding free energy calculation

Complex	Poses	E_{vdw}	E_{Ele}	ΔG_{Solv}	ΔG_{Gas}	ΔG_{Bind}
RNV66-VEGF	Pose 1	-69.36	1846.87	-1837.77	1777.51	-60.26 ± 6.54
	Pose 2	-67.56	2169.25	-2137.10	2101.69	-35.41 ± 7.21
	Pose 3	-42.57	2259.74	-2226.51	2217.17	-9.34 ± 10.34
	Pose 4	-87.35	2140.18	-2107.01	2052.83	-54.18 ± 10.85
	Pose 5	-91.25	2400.14	-2367.04	2308.89	-58.15 ± 9.30
	Pose 6	-33.08	2264.31	-2252.26	2231.24	-21.03 ± 7.09

	Pose 7	-36.33	2099.75	-2074.35	2063.42	-10.93 ± 5.67
	Pose 8	-88.46	2344.19	-2299.39	2255.73	-43.66 ± 7.63
	Pose 9	-51.40	2172.35	-2137.48	2120.95	-16.53 ± 11.74
	Pose 10	-48.48	2266.25	-2238.37	2217.77	-20.61 ± 9.14
LNA→dG	Pose 1	-44.30	2035.95	-2019.23	1991.64	-27.58 ± 10.05
VEGF-VEGFR	Native	-246.93	-2325.47	2381.86	-2572.40	-190.53 ± 17.76
VEGF-RNV66-VEGFR	Pose 1	-105.97	-4549.69	4571.19	-4655.66	-84.47 ± 22.60
	Pose 5	-87.92	-3078.33	3108.70	-3166.25	-57.55 ± 12.16
	Pose 7	-33.54	-3137.08	3136.32	-3170.61	-34.45 ± 23.09

Note*:* All energy components are extracted from the differences (average) of $\Delta G_{Complex}$-$\Delta G_{Receptor}$-ΔG_{Ligand}. The results refer to averages over 2000 frames and all units are reported in kcal mol^{-1}. E_{vdW}= van der Waals energy, E_{ele}=Electrostatic energy, ΔG_{gas}= Sum of Van der Waals energy+ Electrostatic energy+ internal energy, ΔG_{solv}= Solvation energy (polar and non-polar).

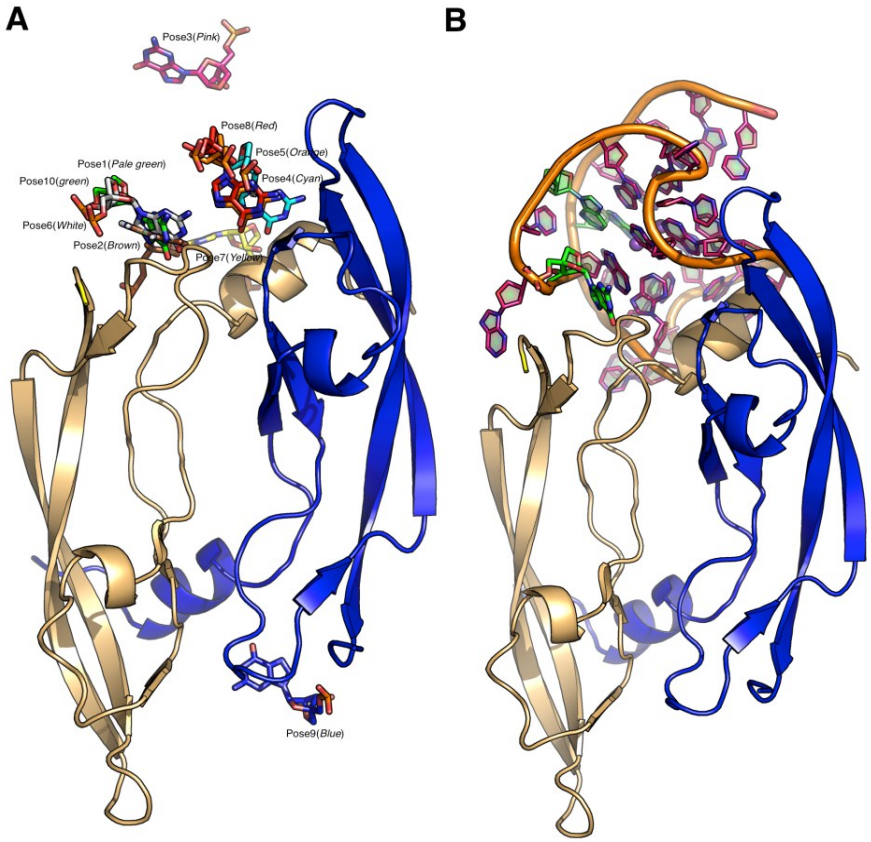


Figure S9. Comparison of different RNV66-VEGF poses. **A**. RNV66-VEGF complexes are shown only with position of LNA-G24; **B.** Representative pose of RNV66 bound to VEGF.

RNV66-VEGF complex docking into VEGF receptor

This work is devoted to study and check whether the VEGF (A) bound with RNV66 is still able to bind to the VEGFR or not by comparing the binding free energy calculation of VEGF-RNV66-VEGFR complex with native binding e.g., VEGF-VEGFR complex. Therefore, VEGF-RNV66 complex was docked into the VEGFR and top 10 poses were saved for further investigation. Results showed that out of 10 poses, none of them were able to bind to the VEGFR as previously observed. However, three complexes (poses) that showed most interaction with VEGFR were chosen for the binding affinity prediction using the MM-GBSA from MD simulation for 20 ns (Figure S10).

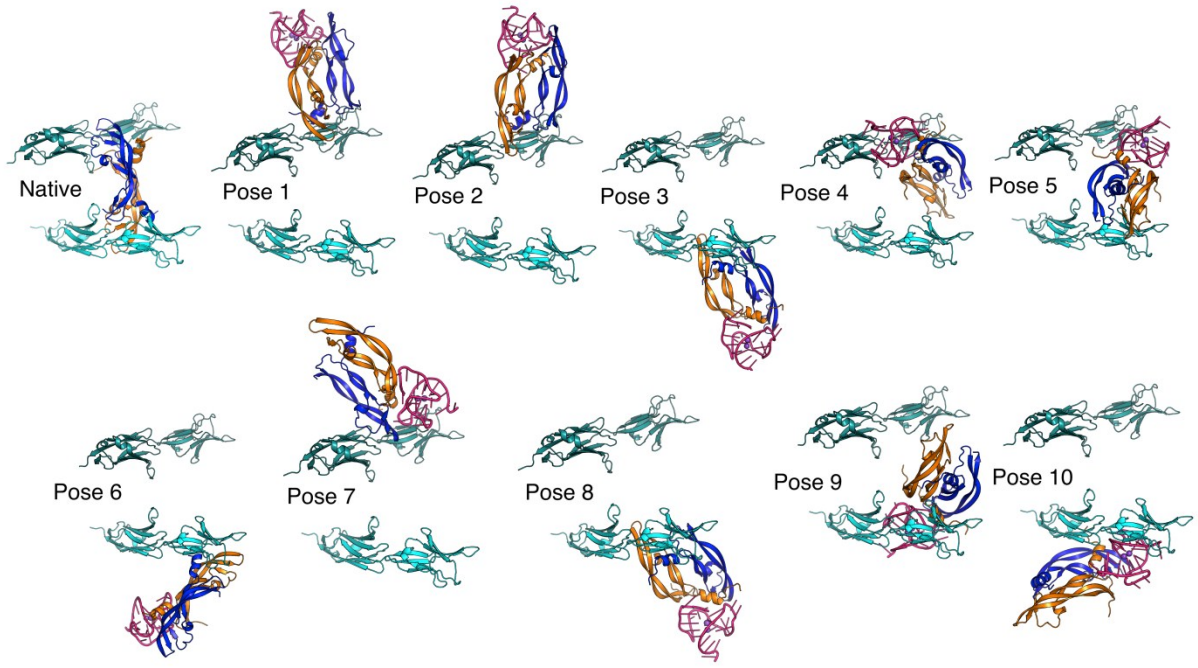


Figure S10. Comparison of poses obtained from RNV66-VEGF(A) complex docking into VEGF receptor

Binding free energy of modified RNV66-VEGF complex.

In order to better understand the role of locked nucleic acid (LNA-G24) at 24th position in RNV66 in VEGF binding, we extended our analysis by modifying LNA-G24 with dG residue in pose 1 (VEGF-RNV66 complex) that showed the best binding free energy. Subsequently, MD simulation and MM-GBSA calculations were carried out as described in the method section (Figure S11)

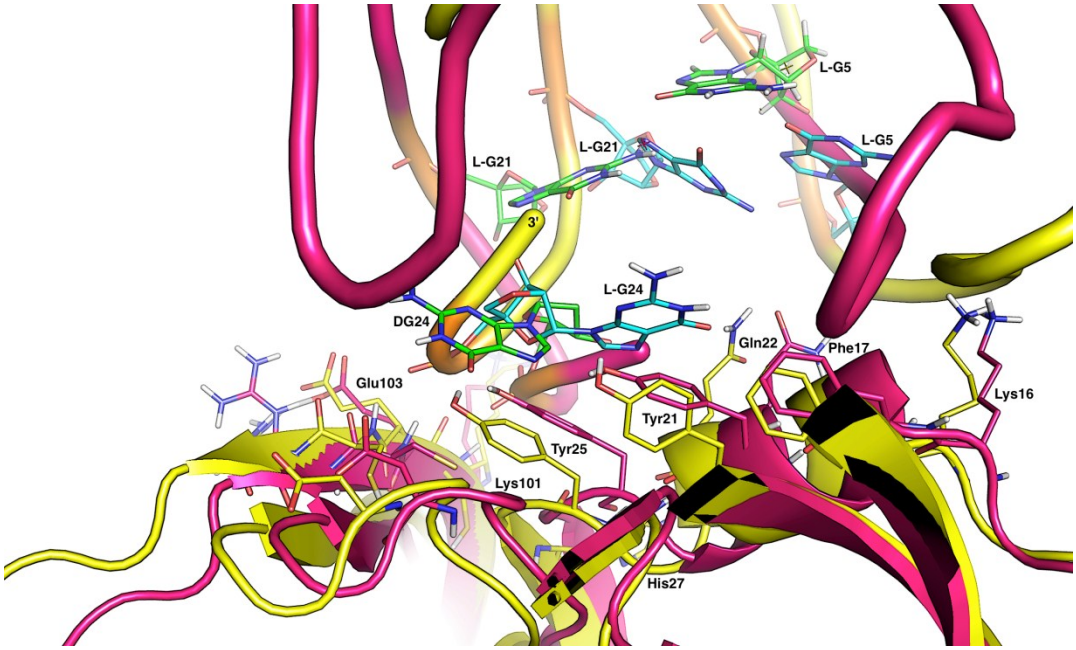


Figure S11: Comparison of snapshots of modified LNA-G24 pose 1 (pink cartoon and LNA residues are shown in green stick) with unmodified LNA-G24 pose 1 (yellow cartoon and LNA residues are shown in cyan stick). Important residues are highlighted.

Decomposition analysis of the binding free energies

The aim of this energy decomposition analysis was to examine the role of individual residue contributions to the overall binding affinity of the each complex by decomposition of the binding free energy into ligand (RNV66 or RNV66-VEGF)–residue pairs (either VEGF or VEGFR) (Figure S12a, b&c).

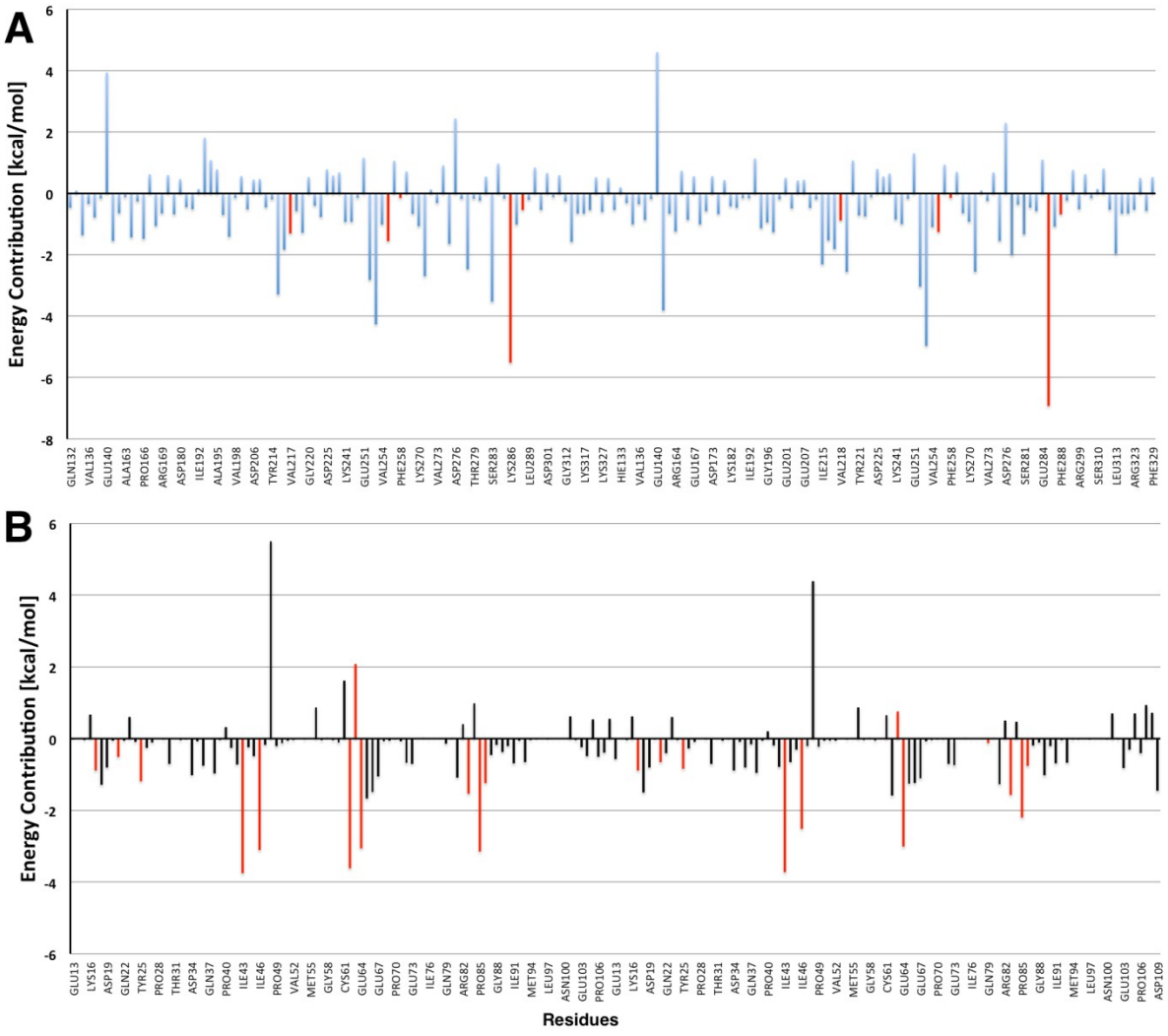


Figure S12a. Native binding mode of VEGF and VEGFR. The energy contributions to overall binding affinity of each residue of VEGF (panel A) and VEGFR (panel B) are plotted. Important residues are highlighted in red color.

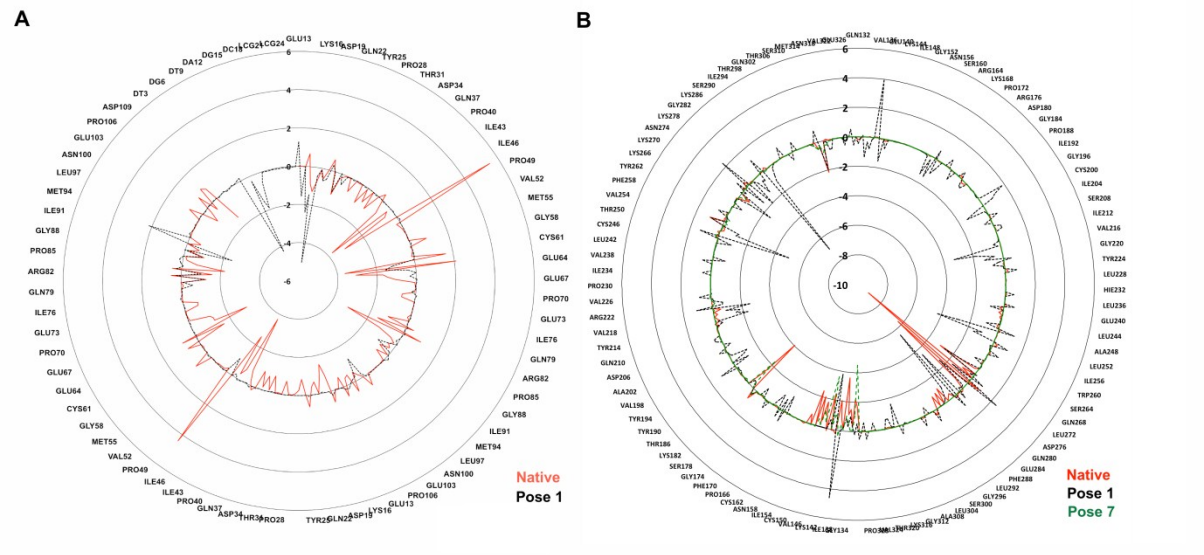
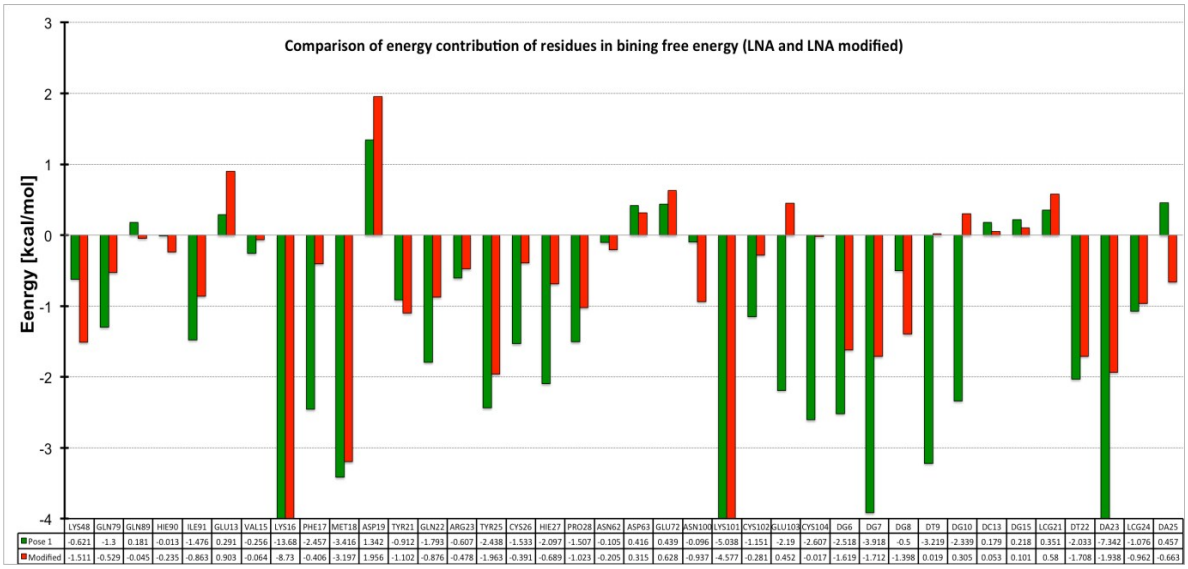


Figure S12b. A. Comparison of energy contributions of each residue of VEGF and RNV66 complex (pose 1) to overall binding affinity. B. Comparison of energy contributions of each residue of VEGF-VEGFR complex (native) and RNV66-VEGF complex bound to VEGFR .



	LYS548	GLN79	GLN89	HIE90	ILE91	GLU13	VAL15	LYS16	PHE17	MET18	ASP19	TYR21	GLN22	ARG23	TYR25	CYS26	HIE27	PRO28	ASN62	ASP63	GLU72	ASN100	LYS101	CYS102	GLU103	CYS104	DG6	DG7	DG8	DT9	DG10	DC13	DG15	LCG21	DT22	DA23	LCG24	DA25
Pose 1	-0.621	-1.3	0.181	-0.013	-1.476	0.291	-0.256	-13.68	-2.457	-3.416	1.342	-0.912	-1.793	-0.607	-2.438	-1.533	-2.097	-1.507	-0.105	0.416	0.439	-0.096	-5.038	-1.151	-2.19	-2.607	-2.518	-3.918	-0.5	-3.219	-2.339	0.179	0.218	0.351	-2.033	-7.342	-1.076	0.457
Modified	-1.511	-0.529	-0.045	-0.235	-0.863	0.903	-0.064	-8.73	-0.406	-3.197	1.956	-1.102	-0.876	-0.478	-1.963	-0.391	-0.689	-1.023	-0.205	0.315	0.628	-0.937	-4.577	-0.281	0.452	-0.017	-1.619	-1.712	-1.398	0.019	0.305	0.053	0.101	0.58	-1.708	-1.938	-0.962	-0.663

Figure S12c. Comparison of energy contributions of each residue from the pose 1 of the VEGF-RNV66 complex (green) and the modified LNA-G24 pose of the complex (red) to the overall binding affinity.

MM-GB/SA calculations

Accurate prediction of binding affinities for ligands as well as their ranking with respect to each other represents a major challenge in computer-aided drug design in particular in lead identification/optimization processes.[11, 12] Several methods have also been developed e.g., the

MM-GB/SA (Molecular Mechanics Generalized Born Solvent Accessibility). The principles of MM-GB/SA have been well described in detail in many publications.[21-24] Briefly, MM-GB/SA is force-field based method which employs molecular mechanics, the Generalized Born (GB) solvation model as well as solvent accessibility methods to approximate the free energies of binding based on snapshots extracted from MD simulations.

The binding free energy (ΔG_{Bind}) of a protein-ligand complex (PL) in an aqueous environment can be estimated from the energy difference between the bound and unbound states of the protein P and ligand L. This protocol is standard for measuring the binding affinity through computational methods:

$$\Delta G_{Bind} = G_{PL} - G_L - G_P \quad (2)$$

However, significant computational time is required to calculate the solvent-solvent interactions. Therefore, the binding free energy is calculated based on the thermodynamic cycle (Figure S13):

$$\Delta G_{Bind} = \Delta G_{Bind,Vacuum} + (\Delta G_{Solv,PL} - (\Delta G_{Solv,L} - \Delta G_{Solv,P}) \quad (3)$$

where, ΔG_{Bind} and $\Delta G_{Bind, Vacuum}$ correspond to the free energy difference between the bound and unbound states of a complex in solvent and vacuum respectively. ΔG_{Solv} ($\Delta G_{Solv,L}$, $\Delta G_{Solv,P}$ and $\Delta G_{Solv,PL}$) represent the change in free energy between the solvated and vacuum states of a ligand, receptor or complex. These different components can be calculated as a sum of three terms:

$$G = \langle E_{MM} \rangle + \langle G_{Solv(polar+nonpolar)} \rangle - T\langle S_{MM} \rangle \quad (4)$$

$$E_{MM} = E_{Int} + E_{el} + E_{vdW} \quad (5)$$

where, E_{MM} is MM energy of the molecules. E_{MM} is the sum of the internal energy (E_{Int}) of the molecules (i.e. bonded terms), E_{el} and E_{vdW} represents the intermolecular electrostatic and van der Waals interactions, respectively. In order to reduce the computational time, and to obtain stable energies, a single-trajectory is normally used for the ligand, protein and complex, i.e. only the PL complex is explicitly simulated by MD. Thereby the E_{Int} term cancels in the calculation of ΔG_{Bind}. G_{solv} is the polar and non-polar solvation energies of the molecule, estimated from the GB approximation combined with a solvent accessible surface area (SASA) calculation. T is the temperature and S_{MM} is the entropy (estimated from a normal-mode analysis calculated at the molecular mechanics level). All the components in

Eqn.4 are averages of energies obtained from snapshots from the MD simulations. A schematic representation of the thermodynamic cycle of the free energy calculation is shown in Figure S13.

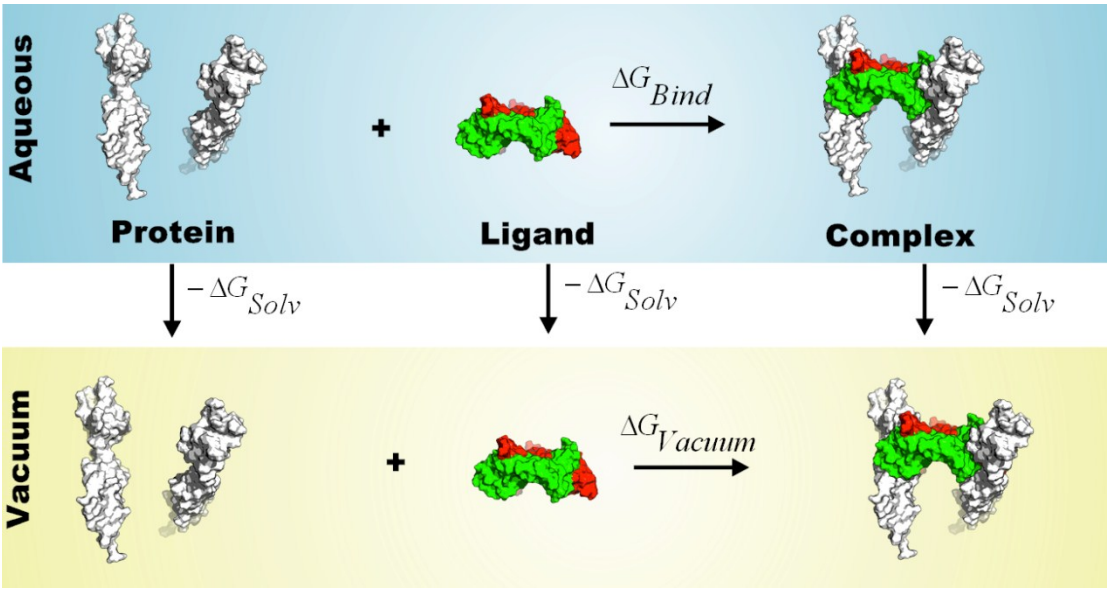


Figure S13. Schematic representation of the thermodynamic cycle of the free energy calculation.

The binding free energies (ΔG_{Bind}) for all protein-ligand complexes were calculated according to Eqn. 4 using the MMPBSA.py script in Amber 14 based on the use of 1000 snapshots of the 20 ns MD simulation. The implementation of MMPBSA.py in Amber 14 is well described in the original publication.[25] All MM-GBSA calculations are based on the "single-trajectory MD simulation", meaning that no separate MD simulations were run for free ligands or receptor. TΔS (entropy) is usually estimated from the normal mode analysis, but due to extremely slow convergence of these complexes, we excluded the entropy contribution to the free energy difference.

References

1. Marusic, M.; Veedu, R. N.; Wengel, J.; Plavec, J., G-rich VEGF aptamer with locked and unlocked nucleic acid modifications exhibits a unique G-quadruplex fold. *Nucleic acids research* **2013**, 41, 9524-36.

2. Brozzo, M. S.; Bjelic, S.; Kisko, K.; Schleier, T.; Leppanen, V. M.; Alitalo, K.; Winkler, F. K.; Ballmer-Hofer, K., Thermodynamic and structural description of allosterically regulated VEGFR-2 dimerization. *Blood* **2012**, 119, 1781-8.

3. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and

ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **2013**, 27, 221-34.

4. Schrödinger *Schrödinger LLC., Portland, USA;* http://www.schrodinger.com, Portland, USA, 2013.

5. Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005**, 61, 704-21.

6. Prime *Schrödinger, LLC, New York, NY, 2012.*, version 1.5.

7. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford CT, USA, 2009.

8. Case, D. A.; Darden, T. A.; Cheatham, I. T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER*, 12; University of California: San Francisco, 2012.

9. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, 81, 3684-3690.

10. Pierce, B. G.; Wiehe, K.; Hwang, H.; Kim, B. H.; Vreven, T.; Weng, Z., ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers.

Bioinformatics **2014**, 30, 1771-3.

11. Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L., 3rd, Assessing scoring functions for protein-ligand interactions. *J Med Chem* **2004**, 47, 3032-47.

12. Stjernschantz, E.; Oostenbrink, C., Improved ligand-protein binding affinity predictions using multiple binding modes. *Biophys J* **2010**, 98, 2682-91.

13. Nienhaus, G. U., *Protein-Ligand Interactions: Methods and Applications*. Humana Press: New Jersey, 2005.

14. Huang, S. Y.; Grinter, S. Z.; Zou, X., Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Phys Chem Chem Phys* **2010**, 12, 12899-908.

15. Reddy, M. R.; Reddy, C. R.; Rathore, R. S.; Erion, M. D.; Aparoy, P.; Reddy, R. N.; Reddanna, P., Free Energy Calculations to Estimate Ligand-Binding Affinities in Structure-Based Drug Design. *Curr Pharm Des* **2013**.

16. Gilson, M. K.; Zhou, H. X., Calculation of protein-ligand binding affinities. *Annual review of biophysics and biomolecular structure* **2007**, 36, 21-42.

17. Oostenbrink, C., Efficient free energy calculations on small molecule host-guest systems - a combined linear interaction energy/one-step perturbation approach. *J Comput Chem* **2009**, 30, 212-21.

18. Acevedo, O.; Ambrose, Z.; Flaherty, P. T.; Aamer, H.; Jain, P.; Sambasivarao, S. V., Identification of HIV inhibitors guided by free energy perturbation calculations. *Curr Pharm Des* **2012**, 18, 1199-216.

19. Gallicchio, E.; Levy, R. M., Recent theoretical and computational advances for modeling protein-ligand binding affinities. *Advances in protein chemistry and structural biology* **2011**, 85, 27-80.

20. Beveridge, D. L.; DiCapua, F. M., Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annual review of biophysics and biophysical chemistry* **1989**, 18, 431-92.

21. Hayes, J. M.; Archontis, G. MM-GB(PB)SA Calculations of Protein-Ligand Binding Free Energies. In *Molecular Dynamics - Studies of Synthetic and Biological Macromolecules*, Wang, L., Ed.; InTeck: 2012; Chapter 9.

22. Weis, A.; Katebzadeh, K.; Soderhjelm, P.; Nilsson, I.; Ryde, U., Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *J Med Chem* **2006**, 49, 6596-606.

23. Swanson, J. M.; Henchman, R. H.; McCammon, J. A., Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys J* **2004**, 86, 67-74.

24. Kongsted, J.; Ryde, U., An improved method to predict the entropy term with the MM/PBSA approach. *J Comput Aided Mol Des* **2009**, 23, 63-71.

25. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *Journal of chemical theory and computation* **2012**, 8, 3314-3321.