Supporting information

Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide

Jacobo Gómez-González,[a,§] Laura Martínez-Castro,[b,‡] Juan Tolosa-Barrilero,[c,d,‡] Ana Alcalde,[a] Soraya Learte-Aymamí,[a] José L. Mascareñas,[a] Joaquín C. García-Martínez,[c,d] José Martínez-Costas,[e] Jean-Didier Maréchal,[b] Miguel Vázquez López,[f] and M. Eugenio Vázquez[,a]*

a. *Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, Spain.*
b. *Insilichem, Departament de Química, Universitat Autònoma de Barcelona, 08193 Cerdanyola, Spain.*
c. *Department of Inorganic, Organic Chemistry and Biochemistry, Faculty of Pharmacy, University of Castilla-La Mancha, 02071Albacete, Spain.*
d. *Regional Center for Biomedical Research (CRIB), Almansa s/n, 02071 Albacete, Spain.*
e. *Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Bioquímica y Biología Molecular, Universidade de Santiago de Compostela, Spain.*
f. *Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Inorgánica, Universidade de Santiago de Compostela, Spain.*

‡ *These authors contributed equally.*

§ *Current address: Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP) Robert-Rössle-Strasse10, 13125 Berlin, Germany.*

Reactives

For the synthesis of the peptides all the amino acids were provided by *Sigma Aldrich*. All the amino acids were used as their Fmoc protected derivatives. The resin employed for the SPPS was H-Rink-Amide *ChemMatrix* 35-100 mesh particle size from *Sigma Aldrich*. Pentafluorophenol was acquired from *Sigma Aldrich.* DCC was acquired from *Alfa Aesar.* Solvents for SPPS and organic synthesis were provided by *Fischer Scientific.*

Mass spectrometry

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) was performed with a Bruker Autoflex MALDI-TOF model in positive scan mode by direct irradiation of the matrix-adsorbed peptide. 4-HCCA (α-cyano-4-hydroxycinnamic acid) was the selected matrix for all these experiments.

Solid phase peptide synthesis

C-terminal amide natural AT-Hook derivatives were synthesized following standard Fmoc-peptide synthesis protocols on a 0.1 mmol scale using a 0.5 mmol/g loading *H-Rink* amide *ChemMatrix* resin (35-100 mesh size particle) with a Liberty Lite automatic microwave assisted peptide synthesizer from CEM Corporation. The amino acids were coupled in 5-fold excess using oxyme as an activating agent. Couplings were conducted for 4 min at 90 °C. Deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 1 min at 75 °C. After the obtention of the monomer AT-Hook peptide, trimerization was carried out by hand employing 0.3 eq of the trispentafluorophenol ester in DIPEA/DMF 0.2M during 2h. **Cleavage from the resin and final deprotection.** Cleavage and deprotection of the peptide were simultaneously performed using standard conditions by incubating the resin for 2.5 h with an acidic mixture containing 50 µL CH_2Cl_2, 25 µL of H_2O, 25 µL of TIS (triisopropylsilane), and 900 µL of TFA. The resin was filtered, and the TFA filtrate was concentrated under a nitrogen stream to an approximate volume of 1 mL, and then added onto ice-cold diethyl ether (20 mL). After 10-30 min, the precipitate was centrifuged and washed again with 5 mL of ice-cold ether. The obtained residue was then dissolved in a mixture of H_2O and MeCN and purified by preparative HPLC for the obtention of the pure compounds.

DNA hybridization and sequences. Oligonucleotides were obtained from *Biomers*. Concentration of the oligonucleotides was measured by UV-VIS by using the extinction coefficient given by the Comercial House Biomers. For the hybridization process, a stochiometric mixture of the DNA strands was prepared in PBS buffer (10 mM, 100 mM NaCl, pH 7.0) and heated at 90 °C during 10 minutes. After this time the mixture was slowly cooled down until room temperature, obtaining the resulting 3wj/ds hybridized DNA.

TA-3WJ: 5′-CACCGTTTA[AATT]TAAAACCTC-3′; 5′-CAGGCTTTA[AATT]TAAACGGTG-3′; 5′-GAGGTTTTA[AATT]TAAAGCCTG-3′

GC-3WJ: 5′-CACCGGGGC[GCGC]GCCCCCCTC-3′; 5′-CAGGCGGGC[GCGC]GCCCCGGTG-3′; 5′-GAGGGGGGC[GCGC]GCCCGCCTG-3′

dsDNA: 5′-GACGG[AATTT]GAGAGCGTCG-3′ (only one strand shown).

HJ: 5′-CCGGAACG-GCACTCGC-3′; 5′-GCGAGTGC-AGCGTGGC-3′; 5′-GCGCGTCC-CGTTCCGG-3′; 5′-GCCACGCT-GGACGCGC-3′

The sequences between the square brackets correspond to the unpaired bases in the branching point of the 3WJ.

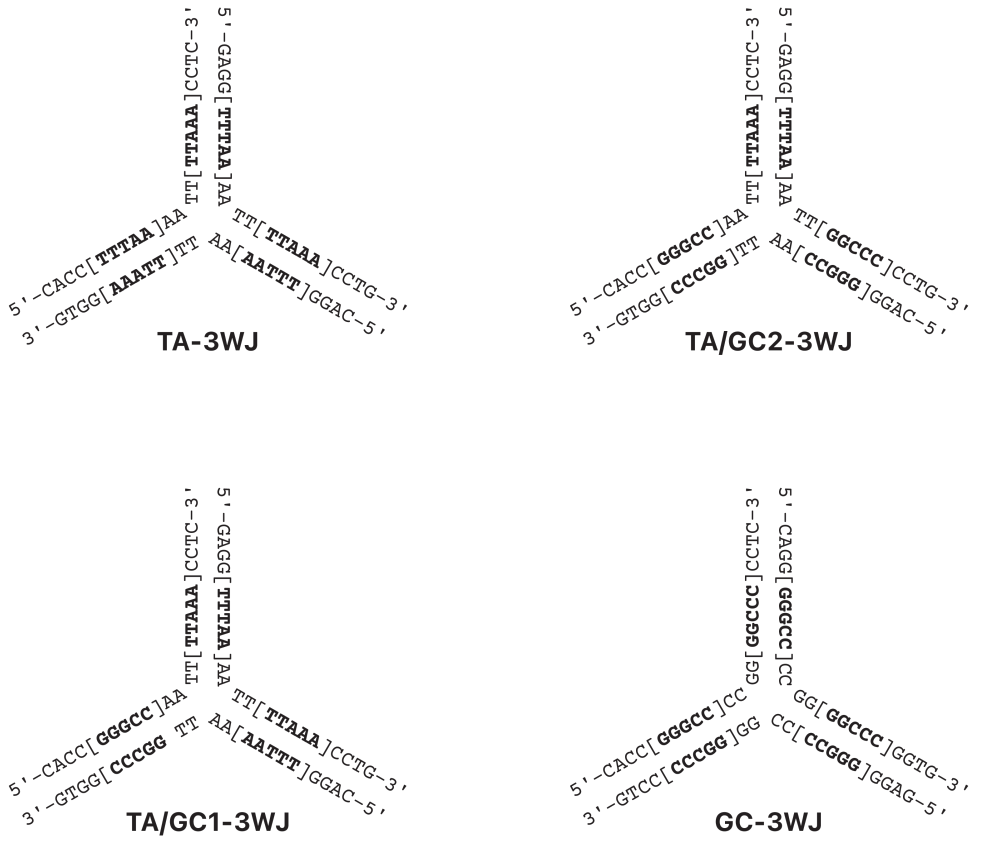


Fig. S1. 3WJ sequences used in this study. **TA-3WJ** is the highest affinity target site, **TA/GC1-3WJ**, **TA/GC2-3WJ** and **GC-3WJ** with increasing number of branches with non-target G/C-rich sequences.

HPLC: UHPLC-MS

Peptide analysis was performed by analytical UHPLC-MS with an Agilent 1200 series LC/MS using a SB C18 (1.8 μm, 2.1 × 50mm) analytical column from Phenomenex. Standard conditions for analytical UHPLC consisted on a linear gradient from 5% to 95% of solvent B in 12 min at a flow rate of 0.350 mL/min (A: water 0.1% TFA, B: acetonitrile 0.1% TFA). Compounds were detected by UV absorption at 222, 270 and 330 nm. Electrospray Ionization Mass Spectrometry (ESI/MS) was performed with an Agilent 6120 Quadrupole LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector. **Preparative HPLC.** Peptide purification was performed by preparative RP- HPLC with a Waters 1500 series Liquid Chromatograph using a Sunfire Prep C18 OBD (5 µm, 19 × 150 mm) reverse-phase column from Waters. Standard conditions for preparative RP- HPLC consisted on an isocratic regime during the first 5 min, followed by different linear gradients of solvent B during 30 min (A: water 0.1% TFA, B: acetonitrile 0.1% TFA). The gradients were adjusted for each peptide.

UHPLC-MS of the purified peptide

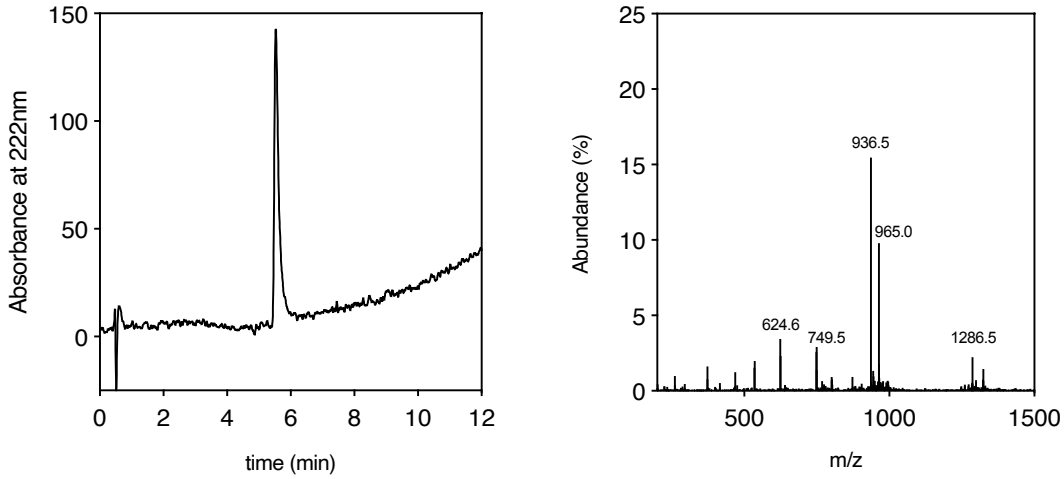


Fig. S2. Left: UHPLC of the purified peptide. Right: Mass spectra of the peak with t_R = 5.54 min.

UHPLC (ESI): (5-95% B, t_R= 5.4 min) Calculated for $C_{177}H_{294}N_{60}O_{30}$ = 3740.33; found: $[M+3H+TFA]^{3+}$ = 1286.50; $[M+4H]^{4+}$ = 936.50; $[M+4H+TFA]^{4+}$ = 965.00; $[M+5H]^{5+}$ = 749.5; $[M+6H]^{6+}$ = 624.6.

MS (MALDI-TOF): Calculated for $C_{177}H_{294}N_{60}O_{30}$ = 3740.33; found: $[M+H]^{+}$ = 3746.12.

UV-Vis Spectroscopy

UV measurements were made in a *Jasco V-630* spectrophotometer coupled to a *Jasco ETC-717* temperature controller, using a standard *Hellma* semi-micro cuvette (108.002-QS) with a light path of 10 mm. Measurements were made at 20 °C. Acquisition parameters were: 220-500 nm range, scan speed of 200 nm/min, resolution of 0.2 nm.

UV-Vis of the peptide ligand. UV-Vis absorbance of a 5 μM solution of the peptide ligand in 10 mM HEPES buffer, 100 mM NaCl, pH 7.01 was measured, showing a maximum c.a. 320 nm.

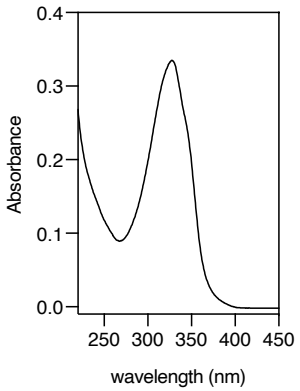


Fig. S3. UV-Vis spectra of the AT-Hook trimer **1**.

Fluorescence spectroscopy

Luminescence experiments were made with a Varian Cary Eclipse Fluorescence Spectrophotometer coupled to a *Cary* Single Cell Peltier accessory *(Agilent Technologies)* temperature controller. All measurements were made with a *Hellma* semi-micro cuvette (108F-QS) at 20 °C. The settings for these measurements were adapted depending on the studied system.

Control experiments. Control experiments carried out in the presence of the **TA-3WJ** and the aromatic core of the peptide ligand were carried out. Thus, over a 1 μM solution of the aromatic core in HEPES buffer 10 mM, NaCl 100 mM, pH 7.01, **TA-3WJ** was added. In this study no DNA-binding can be appreciated in the oscillating pattern that the profile of the titration shows.

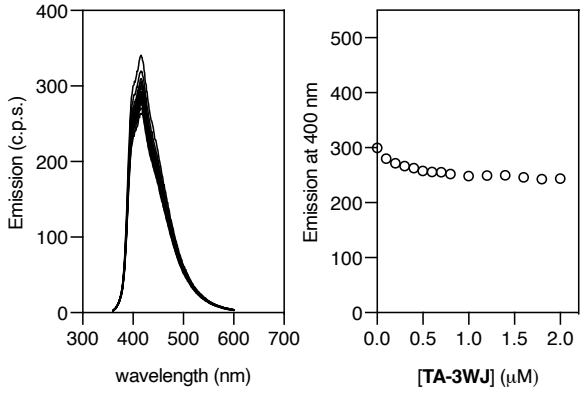


Fig. S4. Left: full emission spectra of the aromatic core (more intense) and in the presence of increasing concentrations of the **TA-3WJ** (decreasing intensity). Right: profile of the full titration.

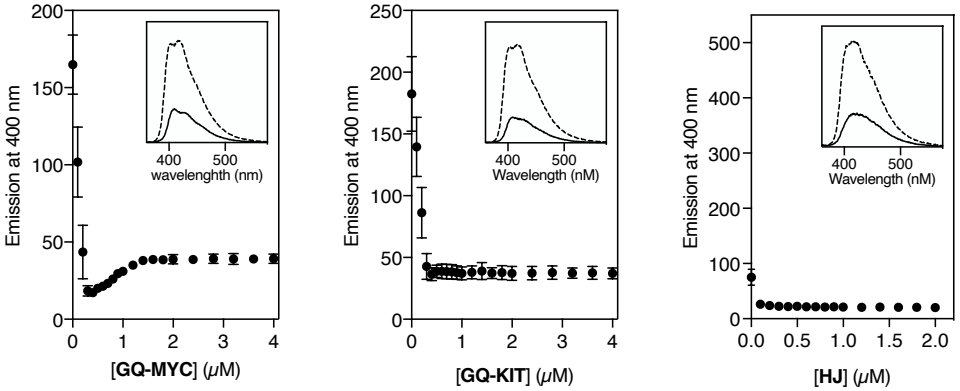


Fig. S5. Selectivity of 1 towards different non-canonical DNA structures. Fluorescence titrations of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 with: a) **GQ-MYC**; b) **GQ-KIT**; c) **HJ**. Insets show the first (dashed line) and last (solid line) emission spectra in the titrations.

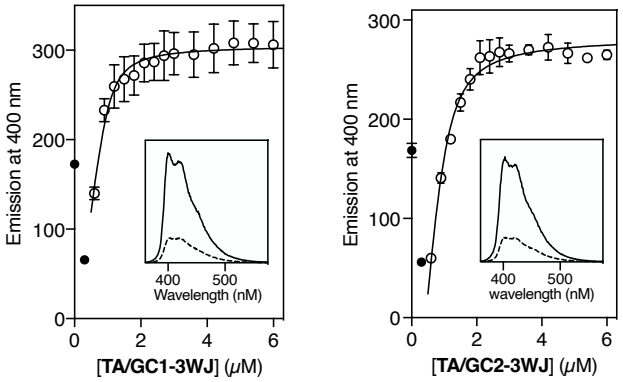


Fig. S6. Selectivity of 1 towards mixed-sequence three-way junctions. Probe 1 also binds to asymmetric 3WJs containing two AT-Hook binding sites (**TA/GC1-3WJ**) and a single AT-Hook binding site in one of the branches (**TA/GC2-3WJ**). The dissociations constants for each of these partly complementary sites is about 60 and 100 nM, for **TA/GC1-3WJ** and **TA/GC2-3WJ**, respectively. supporting the relevance of the simultaneous recognition of the three dsDNA branches in the 3WJ to achieve the best binding. The fact that non-target 3WJs still present remarkably tight binding is probably due to the high contribution of ionic (Arg/phosphate) interactions, which can also take place even with G/C-rich sequences, and has been estimated to represent up to 70% of the total energy of the AT-Hook/DNA [1].

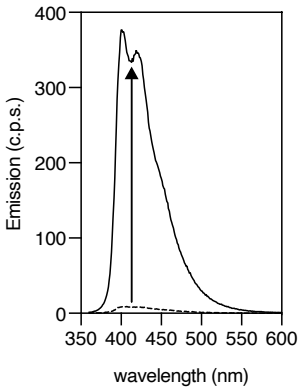


Fig. S7. Emission spectra of of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 in the presence of 2 μM **GC-3WJ** (dashed line) and after the addition of 1 eq. of **TA-3WJ** (solid line).

Data analysis in fluorescence titrations

Fluorescence titration data were analyzed using the *DynaFit 4.0* software, which characterizes the reacting system in terms of stoichiometric equations instead of mathematical notation; the chemical equations are then translated into the underlying mathematical equations wusing matrix theory [2,3].

DynaFit is available at *http://www.biokin.com/dynafit/*. *Dynafit* requires plain text files (scripts) that contain the chemical model underlying the experimental data, the values of model parameters, such as starting concentrations of reactants, as well as information about location of the files. A typical script used in the analysis titrations is included below. The file has been commented to indicate the purpose of the keywords and sections, but the reader is recommended to review the *DynaFit* scripting manual distributed along the program or available at the *DynaFit* website.

```
[task]                          ;semicolons indicate comments
 task = fit                     ;nature of the calculation to be performed
 data = equilibria

[mechanism]                     ;Free-form 1:1 binding model with Kd
 R + L <==> RL : Kd dissoc      ;to be calculated as dissociation constant

[constants]              ;Initial Kd value for iteration
 Kd = 1.0 ?              ;the "?" indicates that this will be optimized

[concentrations]
 R = 2.0                 ;Fixed conc. of the DNA during the peptide
titration

[responses]              ;contribution to the spectroscopic signal of each
 R = 0.1 ?               ;of the different components of the equilibrium
 RL = 1.5 ?              ;these will be optimized ("?" after the values).

[data]                   ;location of files and information about the data
 variable  L             ;the species that changes conc. during the
titration
 offset auto ?
 directory ./exp/brHis   ;file path (relative to DynaFit program location)
 extension txt
 file f1                 ;name of the experimental data file

[output]
 directory ./exp/brHis/out ;path indicating location of DynaFit output files

[settings]               ;cosmetic settings that control DynaFit graphics
 {Output}                ;fits were exported & finally plotted with
 XAxisUnit = uM          ;GraphPad Prism 7.0c. GraphPad Software,
 BlackBackground = n     ;La Jolla California, www.graphpad.com
 XAxisLabel = [peptide]
 YAxisLabel = anisotropy
 WriteTXT = y
```

Circular Dichroism

Circular dichroism measurements were made with a *Jasco J-715* coupled to a *Neslab RTE-111* thermostated water bath, using a *Hellma* 100-QS cuvette (2 mm light pass). Scan speed was 200 nm/min and the obtained spectra are the mean of three accumulations. CD experiments were carried out to confirm the DNA-binding properties of the ligand. Thus, a 1 μM solution of different DNA's in 10 mM HEPES buffer, 100 mM NaCl, pH 7.0 was measured before and after the addition of 1 eq. of the peptide ligand. As it can be observed in the figure below, the presence of a target TTTAA sequence in the **TA-3WJ** or the dsDNA results in DNA recognition. This can be observed in the appearance of a new induced band c.a. 325 nm in the DNA spectra. Additionally, same experiment but in this case with **GC-3WJ** resulted in no DNA-binding and as consequence, not induced bands can be observed, confirming thus the selectivity of the peptide ligand for the TA rich sequences.

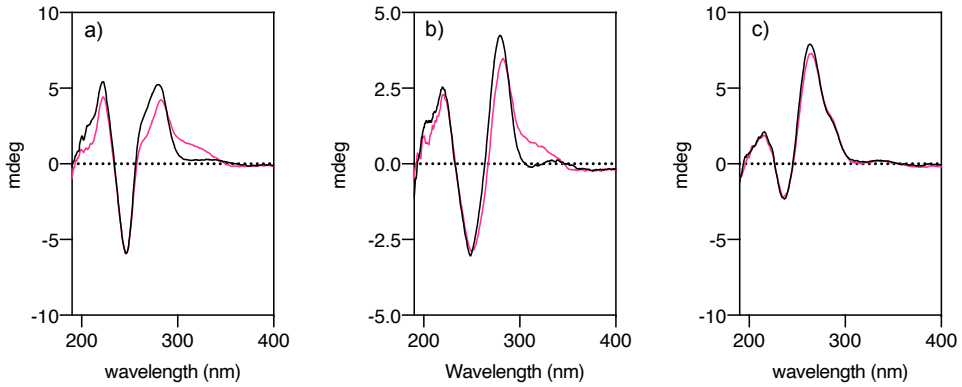
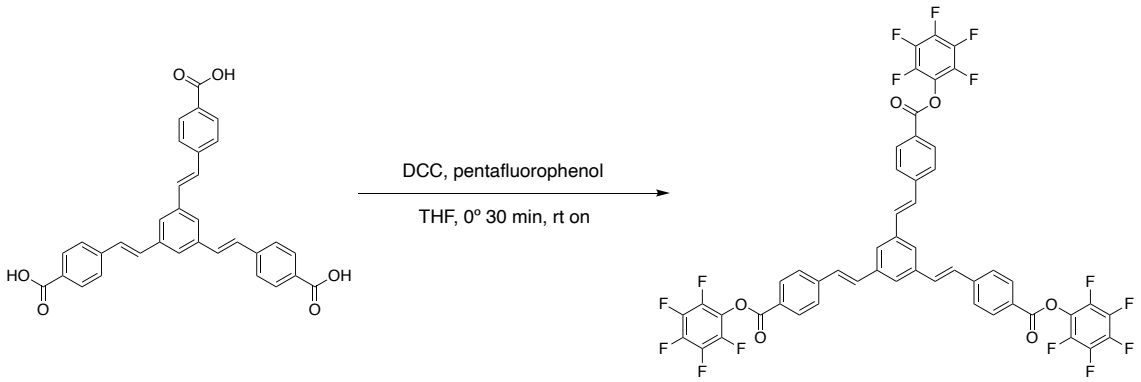


Fig. S8. Left: **TA-3WJ** before (black solid line) and after (pink solid line) the addition of 1 eq of **1**. Middle: **dsDNA** before (black solid line) and after (pink solid line) the addition of 1 eq of **1**. Right: **GC-3WJ** before (black solid line) and after (pink solid line) the addition of 1 eq of **1**.

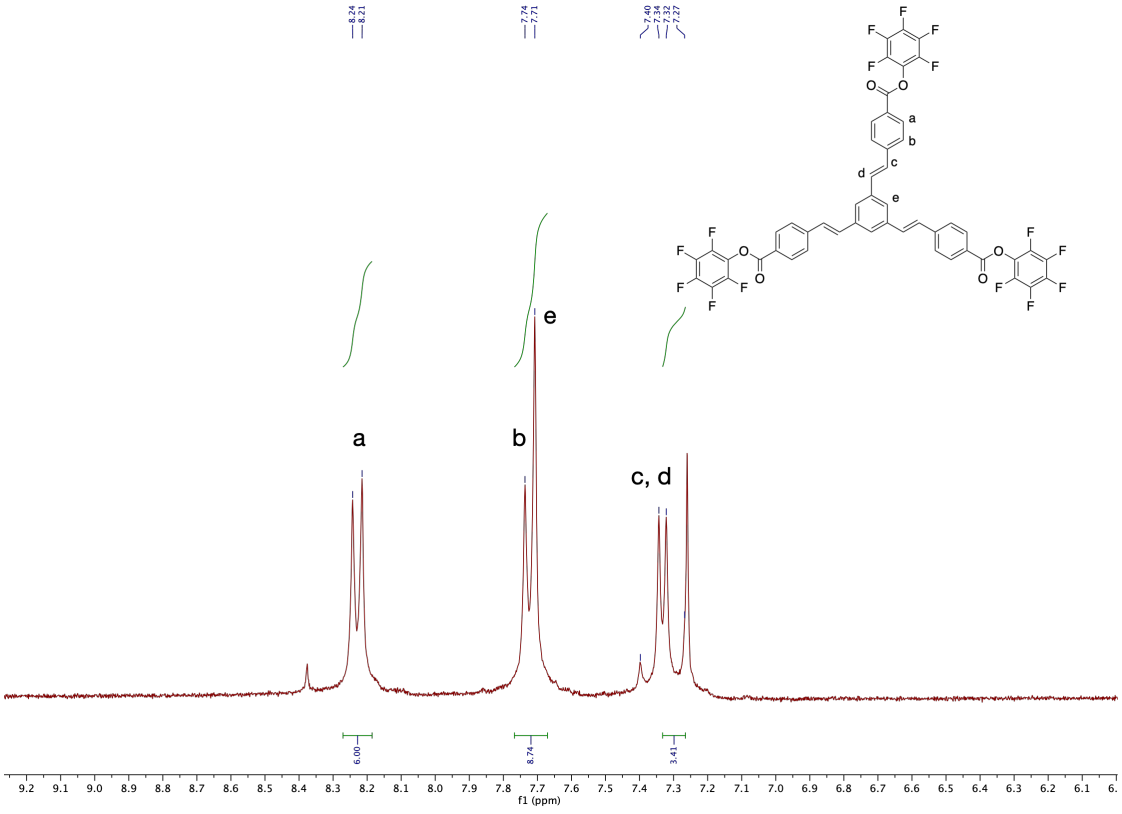
Organic synthesis

Synthesis of tris(perfluorophenyl) 4,4',4''-((1*E*,1'*E*,1''*E*)-benzene-1,3,5-triyltris(ethene-2,1-diyl)) tribenzoate (2)



1 **2**

To a stirred solution of the tricarboxylic acid **1** (50 mg, 0.096 mmol) in THF (7 mL) under argon at 0°, DCC (72 mg, 0.36 mmol) was added. The reaction was allowed to reach room temperature, after which, pentaflurophenol (53 mg, 0.288 mmol) was added. The mixture was stirred overnight. After this time, the reaction was filtered, and the solvent was evaporated under reduced pressure. The resulting solid was purified by flash chromatography using DCM as eluent to yield the compound **2** as a white powder (45 mg, 46% yield). **^{1}H-NMR** (500 MHz, CDCl3, δ): 8.23 (d, *J* = 8.3 Hz, 6H), 7.72 (m, 9H), 7.33 (dd, *J* = 16, 6.5 Hz, 6H).**MS** (**APCI-TOF**): calculated for $C_{51}H_{21}F_{15}O_6$ = 1014.11; found: $[M+H]^+$= 1015.11.



Molecular modeling

We carried out protein-ligand docking of 1,3,5-tristyrylbenzene to the 3WJ/AT-Hook model using four objectives during the simulation: the three distances between the styryl benzene carboxylates and the N-terminal Lys side chain of the AT-hook peptides, which should be set to an average of 1.35 Å, and minimization of the atomic clashes between the ligands and the receptor (Fig. S6). From those calculations, we observed 13 poses that met the distance objective with a reasonable deviation of ± 1.5 Å (an acceptable deviation considering the likely geometric rearrangements of the 3WJ/AT-Hook complexes), two of the best poses also have very good contact values (lower than 100 $Å^3$).

System set up: The structure of a crystallographic three-way junction (PDB code: 1f44) was superimposed on the AT Hook region (PDB code: 3uxw) using UCSF Chimera software.[4]. Then, the crystallographic waters and other small molecules were removed, as well as the three-way junction template after manual curation. The tristyryl benzene was set up with all carboxylic oxygens adjusted to the proper charge. Both systems were protonated using the implemented algorithm in UCSF Chimera.

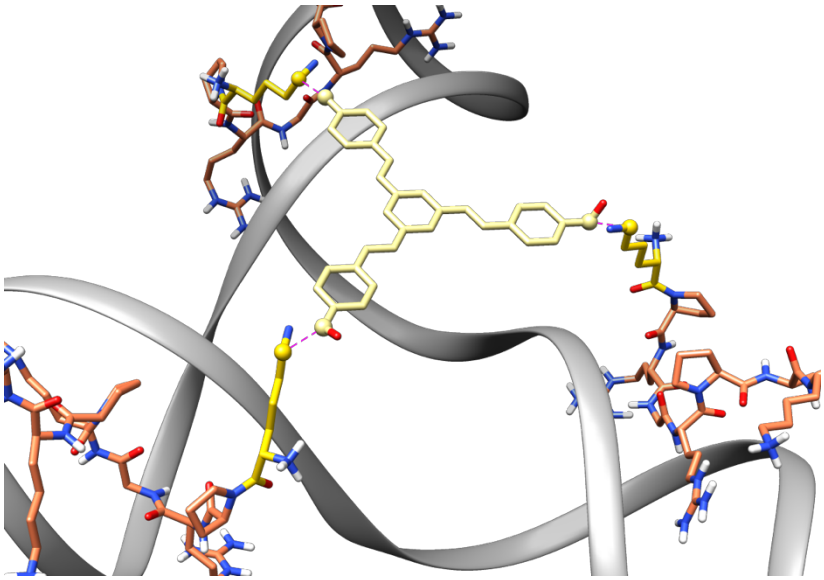


Fig. S9. Set up system. The backbone scaffold of the DNA three-way junction (PDB code: 1F44) and AT rich region (PDB code: 3UXW) are depicted in gray; AT-Hook peptides in orange with the terminal lysines in dark yellow. The 1,3,5-tritrystyrilbenzene is depicted in light yellow. Atoms with ball shape indicate the points between which the distances were restricted in the GaudiMM.

Docking. The multi-objective genetic algorithm GaudiMM was used for the simulations, with an input (input.yaml file provided at the end of the section) containing the information on the objectives and how to optimize them. Calculations were performed considering four objectives: the three distances between each carbonyl C atom on the 1,3,5-tristyrylbenzene and the N on the lysine side chain at the N-termini of the AT-hook, aiming to be matched to an average of 1.35Å and minimization of the atomic clashes between the ligands and the receptor. Moreover, the flexibility of the terminal lysines was allowed to facilitate accommodation of the linker.

GaudiMM calculations led to 13 solutions satisfying the distance objectives with ±1.5 Å (accounting for some fluctuation of the DNA structure and AT-hook peptides with respect to the original X-ray structure). Moreover, 2 of these solutions exhibit clashes lower than 100 $Å^3$; a good value as demonstrated in previous studies [5].

Input.yaml file for GaudiMM

```
path: ./input.yaml
ga:
   cx_eta: 5
   cx_pb: 0.5
   generations: 100
   lambda_: 3.0
   mu: 1.0
   mut_eta: 5
   mut_indpb: 1.0
   mut_pb: 0.5
   population: 200
genes:
-   module: gaudi.genes.molecule
   name: Protein
   path: ./protein.mol2

-   module: gaudi.genes.molecule
   name: Cofactor
   path: ./linker.mol2

-   module: gaudi.genes.torsion
   name: TorsionC
   target: Cofactor
   rotatable_bonds: [[1, 56, 1], [50, 51, 50], [53, 54, 53] ]

-   module: gaudi.genes.rotamers
   name: Rotamers
   residues: [Protein/10, Protein/1, Protein/19 ]

-   interpolation: 0.5
   module: gaudi.genes.search
   name: Search
   precision: 5
   radius: 3.0
   rotate: true
   target: Cofactor

objectives:
-   module: gaudi.objectives.contacts
   name: Clashes
   probes: [Cofactor]
   radius: 5.0
   weight: -1.0
   which: clashes

-   module: gaudi.objectives.distance
   name: Distance56-160
   target: Protein/160
   probes: [Cofactor/56]
   threshold: 1.350
   weight: -1.0

-   module: gaudi.objectives.distance
   name: Distance54-9
   target: Protein/9
   probes: [Cofactor/54]
   threshold: 1.350
   weight: -1.0

-   module: gaudi.objectives.distance
   name: Distance51-313
   target: Protein/313
   probes: [Cofactor/51]
   threshold: 1.350
   weight: -1.0
output:
   check_every: 0
   compress: true
   history: false
```

```
    name: 3uxw_dist_clash_rotK
    pareto: true
    path: ./3uxw_dist_clash_rotK
    precision: 3
    prompt_on_exception: false
    verbose: true
similarity:
    args:
    -   - Cofactor
    - 0.5
    kwargs: {}
    module: gaudi.similarity.rmsd
```

Fluorescence spectroscopy

Luminescence experiments were made with a Varian Cary Eclipse Fluorescence Spectrophotometer coupled to a *Cary* Single Cell Peltier accessory *(Agilent Technologies)* temperature controller. All measurements were made with a *Hellma* semi-micro cuvette (108F-QS) at 20 °C. The settings for these measurements were adapted depending on the studied system.

Electrophoretic gel shift assays (EMSA)

Following the spectroscopic studies we decided to characterize in more detail its DNA binding properties by electrophoretic gel shift assays under non-denaturing conditions.[6,7] Thus, incubation of **TA-3WJ** with increasing concentrations of **1** gave rise to a new slow-migrating band, consistent with the formation of the expected **1/TA-3WJ** complex. Importantly, incubation gave rise only to one new band, which indicates the formation of a single new complex, presumably that in which the three A/T-Hook arms are simultaneously interacting with the duplex branches of the junction (Fig. S7a). A similar experiment with the non-target **GC-3WJ** junction also indicated the formation of a new complex, together with a reduction in the intensity of the bands and some smearing at higher concentrations of the probe, pointing to the formation of higher-order aggregates and/or multiple non-specific complexes (Fig. S7b).[8–11] This is also consistent with the high electrostatic contribution to DNA binding with oligocationic peptides such as the AT-Hook.[12] Finally, incubation with dsDNA produced a new band only at very high concentrations of **1**, which is consistent with the lower binding affinity detected by fluorescence (Fig. S7c). No binding whatsoever was detected when incubating a single AT-Hook peptide with **TA-3WJ** or with **dsDNA**. In the case of **dsDNA**, we only observed precipitation at high peptide loadings (data not shown).

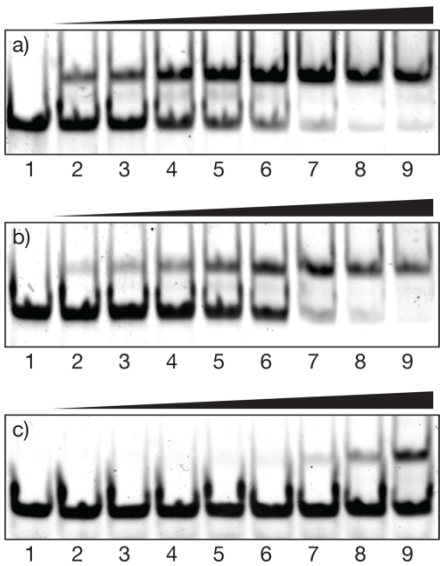


Fig. S10. DNA binding properties of **1** by EMSA. The concentrations, in all cases, are: a) Lanes 1-9, 200 nM of **TA-3WJ** and 0, 100, 200, 300, 400, 500, 600, 700 and 800 nM of **1**, respectively; b) sames as in a) byt in with **GC-3WJ**, and c) same as in a), but with **dsDNA**. EMSA experiments were resolved on a 10% nondenaturing polyacrylamide gel and 0.5× TBE buffer over 35 min at 25 °C and analyzed by staining with SYBR Gold (5 µL in 50 mL of 1× TBE) for 10 min, followed by fluorescence visualization.

Cell Internalization studies

HeLa cells were grown in DMEM (Dulbecco modified Eagle Medium) supplemented with 10% of fetal bovine serum, 1% glutamine in glass-bottom 35mm tissue culture dishes (Ibidi). Semiconfluent monolayers were transfected with plasmid GFP-PCNAL2 [13] using lipofectamine, to express proliferating cell nuclear antigen (PCNA) fused to GFP, as a label for DNA replication sites. For Digitonin permeabilization, cells were washed twice with transfer buffer (25 mM HEPES, 125 mM KOAc, 2 mM $Mg(AcO)_2$, 1 mM EDTA) and then treated with a Digitonin solution (25 μg/mL) in the same buffer for 3 minutes on ice. Cells were washed again twice with transfer buffer containing 10 mg/mL of BSA and then incubated in transfer buffer with 5 μM of **1** for 15 min at 37 ºC. After this time cells were washed three times with transfer buffer containing 10 mg/mL of BSA and covered with DMEM containing 10% of fetal bovine serum at 37 ºC. Confocal images were obtained with an Andor Dragonfly spinning disk confocal system mounted on a Nikon TiE microscope equipped with a Zyla 4.2 PLUS camera (Andor).

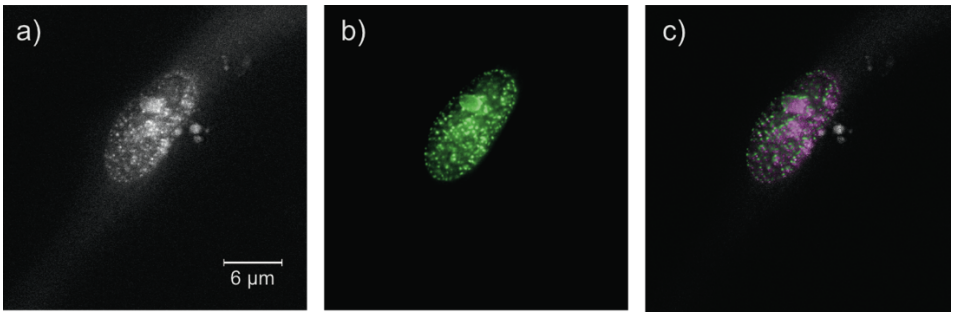


Fig. S11. DNA replication sites in HeLa cells. HeLa cells expressing protein GFP-PCNAL2 were incubated with 25 μg/ml digitonin for 3 min, then 5 μM of 1 for 30 min. a) Blue channel emission showing the distribution of **1** mainly concentrated in the nucleus (channel fake colored in white for easier view); b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci in the cell nucleus (no labeling is seen in the cytoplasm); c) overlay of the green and blue channels with the difference filter applied, (so white areas in a) invert the green color and appear as purple) to highlight the overlay between the staining of GFP-PCNAL2 and that of **1**. The image is a 1000 × magnification confocal image of a representative cell in the slide like that in Fig.4 in the manuscript.

References

1. Dragan, A. I., Liggins, J. R., Crane-Robinson, C. & Privalov, P. L. The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J. Mol. Biol.* **327**, 393–411 (2003).
2. Kuzmič, P. DynaFit—A Software Package for Enzymology. in *Methods in Enzymology* vol. 467 247–280 (Academic Press, 2009).
3. Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273 (1996).
4. Pettersen, E. F. *et al.* UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
5. Sánchez-Aparicio, J.-E. *et al.* GPathFinder: Identification of Ligand-Binding Pathways by a Multi-Objective Genetic Algorithm. *Int. J. Mol. Sci.* **20**, 3155 (2019).
6. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* **2**, 1849–1861 (2007).
7. Tuma, R. S. *et al.* Characterization of SYBR Gold nucleic acid gel stain: a dye optimized for use with 300-nm ultraviolet transilluminators. *Anal. Biochem.* **268**, 278–288 (1999).
8. Liebler, E. K. & Diederichsen, U. From IHF protein to design and synthesis of a sequence-specific DNA bending peptide. *Org. Lett.* **6**, 2893–2896 (2004).
9. He, G.-X., Browne, K. A., Blasko, A. & Bruice, T. C. Microgonotropens and Their Interactions with DNA. 4. Synthesis of the Tripyrrole Peptides Tren-Microgonotropen-a and -b and Characterization of Their Interactions with dsDNA. *J. Am. Chem. Soc.* **116**, 3716–3725 (1994).
10. Portela, C., Albericio, F., Eritja, R., Castedo, L. & Mascareñas, J. L. ds-Oligonucleotide-peptide conjugates featuring peptides from the leucine-zipper region of Fos as switchable receptors for the oncoprotein Jun. *Chembiochem* **8**, 1110–1114 (2007).
11. Vázquez, O., Vázquez, M. E., Blanco, J. B., Castedo, L. & Mascareñas, J. L. Specific DNA recognition by a synthetic, monomeric Cys2His2 zinc-finger peptide conjugated to a minor-groove binder. *Angew. Chem. Int. Ed Engl.* **46**, 6886–6890 (2007).
12. Crane-Robinson, C., Dragan, A. I. & Privalov, P. L. The extended arms of DNA-binding domains: a tale of tails. *Trends Biochem. Sci.* **31**, 547–552 (2006).
13. Leonhardt, H. *et al.* Dynamics of DNA replication factories in living cells. *J. Cell Biol.* **149**, 271–280 (2000).