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

Jacobo Gomez-Gonzalez, Laura Martinez-Castro, Juan Tolosa-Barrilero, Ana Alcalde-Ordoñez, Soraya Learte-Aymamí, José L. Mascareñas, Joaquín C. García-Martínez, José Martínez-Costas, Jean-Didier Maréchal, Miguel Vázquez López, M. Eugenio Vázquez

COMMUNICATION

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

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

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]

Non-canonical DNA structures, particularly 3-way junctions (3WJs) that are transiently formed during DNA replication, have recently emerged as promising chemotherapeutic targets. Here, we describe a new approach to target 3WJs that relies on the cooperative and sequence selective recognition of A/T-rich duplex DNA branches by three AT-Hook peptides attached to a three-fold symmetric and fluorogenic 1,3,5-tristyrylbenzene core.

On August 25, 1942, Dr. Gustaf Lindskog administered to a patient with non-Hodgkin's lymphoma the first intravenous chemotherapy for cancer: an experimental treatment with a DNA alkylating agent.[1,2] Today, almost 80 years later and even after the spectacular developments in targeted therapies, cytotoxic agents that aim at the DNA, including cross-linking agents, are still essential weapons in the chemotherapeutic arsenal.[3] Beyond the canonical antiparallel double-stranded B-DNA conformation, non-canonical DNA structures, such as G-quadruplexes,[4,5] or 3-way junctions (3WJs),[6] have emerged as key players in cancer,[7,8] and as promising alternative DNA targets that could afford new agents with better therapeutic properties.[9–13] Thus, 3WJs are found as intermediates in homologous recombination, and in hairpin loop-outs in slipped DNA structures,[14–16] and, importantly, they are also transiently formed during DNA replication, so that 3WJ binders might lead to replication fork stalling and genotoxic replication stress of rapidly dividing cells.[17,18]

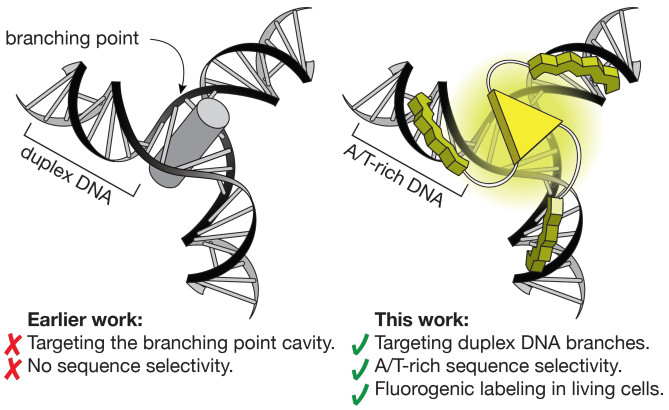


Fig. 1. Left: Current 3WJ binders are typically hydrophobic molecules that target the branching point cavity and lack sequence selectivity; Right: Our strategy to target 3WJs involves cooperative binding of three AT-Hook peptides to the minor grooves of the duplex DNAs. We achieve sequence selectivity and the use of the fluorogenic 1,3,5-tristyrylbenzene core allows us to label DNA replication sites in cells.

3WJs, the simplest branched DNA structures, are symmetric assemblies formed by three converging duplex DNA segments. The focal point of the 3WJs, where the three DNA strands are intertwined, is known as the branching point, and features a ≈12 Å diameter hydrophobic cavity that is lined up by the base pairs at the end of each of the dsDNA branches.[19] Since the group of Prof. M. Hannon fortuitously discovered that metal helicates fit into the branching point of 3WJs,[20] researchers have described a number of molecules that can also target such central cavity in 3WJs,[21] including peptide helicates,[22–25] azacryptands,[26–28] triptycenes,[29,30] cationic calix[3]carbazoles,[31] and derivatives with appended peptides for additional nonspecific interactions with the duplex branches.[32–34] However, the branching point targeted by these agents is basically a featureless hydrophobic

[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.

Electronic Supplementary Information (ESI) available: 1,3,5-Tristyrylbenzene and peptide synthesis and characterization, computational studies, DNA binding studies. See DOI: 10.1039/x0xx00000x

cavity that does not present any obvious structural elements to allow the sequence selective recognition of 3WJs. In contrast with the synthetic binders described before, we envisioned to reproduce the strategy used by natural proteins that recognize 3WJs, which ignore the branching point and interact with the duplex arms exploiting the three-fold symmetry of 3WJs.[19] Here, we show that A/T-rich 3WJs can be selectively recognized over dsDNA and G/C-rich 3WJs by a three-fold symmetric 1,3,5-tristyrylbenzene core modified with AT-Hook peptides that selectively insert in the minor grooves of the three duplex DNAs that form the junction. Moreover, we exploit the fluorogenic properties of the 1,3,5-tristyrylbenzene core to label the A/T-rich replication sites in functional cells (Fig. 1).

The AT-Hook is a 9-residue oligocationic peptide motif, first identified in HMG-I(Y) eukaryotic nuclear proteins, and now recognized as a recurrent auxiliary motif in chromatin and DNA binding proteins.[35] Monomeric AT-Hook peptides display low affinity in the order of millimolar for A/T-rich DNA sequences,[36–39] but appropriately spaced AT-Hook units in HMG-I(Y) proteins cooperatively bind to composite A/T-rich sites with high affinity in the low micromolar to nanomolar range.[39–41] Structural studies provide a detailed picture of the interaction of the AT-Hook (RKP**RGR**PKK) with the DNA, and show that the central segment (RGR) sits at the bottom of the DNA minor groove in an extended conformation making extensive polar and hydrophobic contacts, while the basic residues flanking this core sequence participate in additional electrostatic contacts with the sugar-phosphate backbone of the DNA (PDBs: 3UXW, 2EZD and 2EZF).[42,43] Researchers have exploited the simplicity of the AT-Hook module for the design of artificial duplex DNA binding agents.[44–46]

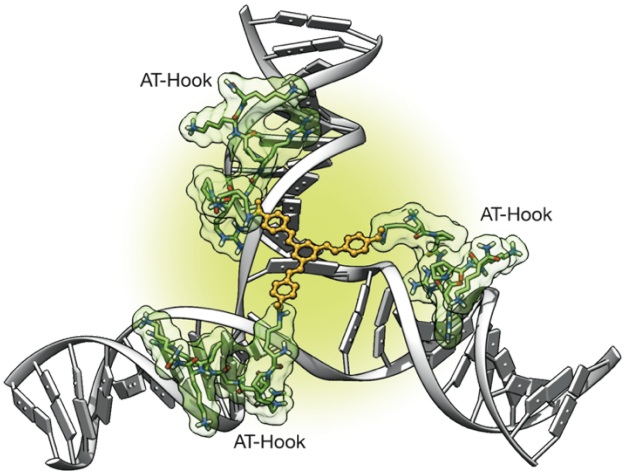
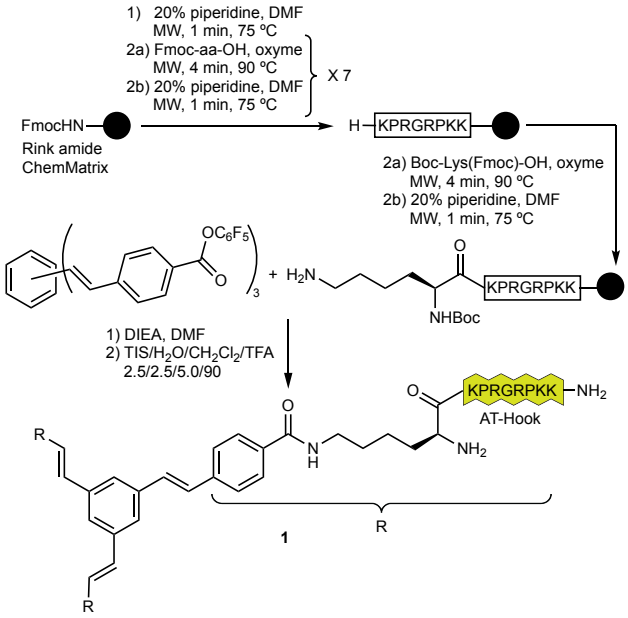


Fig. 2. Best solutions of the *GaudiMM* calculation. Atoms represented in yellow ball and stick correspond to the docked 1,3,5-tristyrylbenzene moiety (see text for explanation). AT-Hook peptides are represented in sticks and surface. The DNA scaffold in ribbons and the side chain of the nucleic acid in planks.

Taking as starting point the crystal structures of the AT-Hook peptide bound to a short A/T-rich oligonucleotide, PDB code 3UXW, and the trimeric Cre recombinase bound to a three-way Lox DNA junction, PDB code 1F44, we built a hypothetical model of three AT-Hook peptides simultaneously bound to the three DNA minor grooves adjacent to the branching point in the 3WJ. This rough model suggested that the N-terminus of the bound AT-Hook peptides would be about 4 Å from the carboxylates in a 1,3,5-tristyrylbenzene core at the center of the 3WJ, so that they could be connected through the side chain of the N-terminal Lysine residue. This hypothesis was further investigated by molecular modeling with the multi-objective genetic algorithm platform *GaudiMM*,[47] which allowed us to make an extensive conformational exploration while simultaneously evaluating several fitness functions (see ESI† for computational details). Thus, 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 ESI†). 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). The absolute best solution shows an excellent distance match between the 1,3,5-tristyrylbenzene and the AT-Hook, as well as great fitting with the 3WJ, with the central phenyl ring hovering over the large cavity of the 3-way junction branching point.



Scheme 1. Synthesis of the trimeric AT-Hook 3WJ binder, **1**, through on-resin attachment of the three AT-Hook arms.

The AT-Hook peptide was assembled following standard solid-phase MW-assisted peptide synthesis protocols.[48–50] Next, the key step in the synthesis of the C_3-symmetric AT-Hook 3WJ binder (**1**, Scheme 1) was the on-resin simultaneous attachment of three AT-Hook branches to the 1,3,5-tristyrylbenzene core. This was accomplished by incubation of the peptide—having the N-terminal lysine sidechain selectively deprotected—with 1,3,5-tristyrylbenzene pentafluorophenyl ester (Scheme 1). The

final product was purified by reverse-phase HPLC, and its identity was confirmed by ESI-MS (Fig. S1 in ESI†).

Having at hand the desired 3WJ binder **1**, we studied its interaction with the DNA by taking advantage of the intrinsic fluorogenic properties of the 1,3,5-tristyrylbenzene core. It is known that in solution, these compounds show non-radiative deactivation of their excited state by restriction of intramolecular motions (RIM) and photoisomerization processes,[51] but conformational restriction—i.e., resulting from aggregation—leads to an increase in fluorescence quantum yield.[52,53] Thus, a 1 μM solution of **1** was titrated with increasing concentrations of different DNAs, and the emission at 400 nm upon excitation at 320 nm was recorded after each addition. We selected two different 3WJs: **TA-3WJ**, containing a high-affinity site for the (TTTAA), and a control GC-rich 3WJ, **GC-3WJ**.

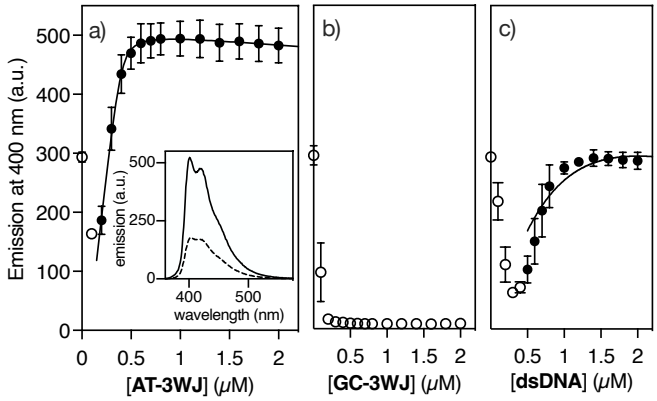


Fig 3. Fluorescence titrations of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 with: a) **TA-3WJ**; b) **GC-3WJ**; c) **dsDNA**. Graphs in a) and c) show the best fit to a mixed binding mode including 1:1 a non-specific term. The initial points of the titrations shown as open circles were not included in the K_D calculation. Inset in a) shows the third (dashed line) and last (solid line) emission spectra in the titration. The Y axis in all the graphs has the same range as in a). DNA sequences: **TA-3WJ**: 5'-CAC CTT TAA [AATT] TTA AAC CTC-3'; 5'-CAG GTT TAA [AATT] TTA AAG GTG-3'; 5'-GAG GTT TAA [AATT] TTA AAC CTG-3'; **GC-3WJ**: 5'–CAC CGG GCC [CCGG] GGC CCC CTC –3'; 5'–CAG GGG GCC [CCGG] GGC CCG GTG –3'; 5'–GAG GGG GCC [CCGG] GGC CCC CTG –3'.; **dsDNA**: 5'- GAC GGA ATT TGA GAG CGT CG-3' (only one strand shown).

In the case of **TA-3WJ** there is an initial drop in the emission, probably due to the initial non-specific association of **1** with the DNA, which disrupts the initial aggregated state of the probe (Fig. 3a, open circles at low DNA concentration); then, as the **TA-3WJ** concentration is increased and AT-Hook arms bind the DNA minor grooves in the junction, the 1,3,5-tristyrylbenzene core is rigidified, and the emission increases again because RIM and photoisomerization is no longer possible. The titration profile could be fitted to a composite binding model that combines specific 1:1 binding with K_D in the order of 20 nM and a minor contribution of nonspecific binding that is linearly proportional to the DNA concentration (Fig. 3a). The profile of the titration with the junction **GC-3WJ** shows a sharp drop in emission, probably resulting from the disassembly of the emissive aggregate by the non-specific association of the highly cationic probe with the negatively-charged **GC-3WJ** combined with the quenching in the emission of the 1,3,5-tristyrylbenzene fluorophore by the guanine bases (Fig. 3b).[54,55] Importantly, this quenching could be reverted by adding 1 equivalent of **TA-3WJ** and displacement of the probe, even in the presence of 4 eq of **GC-3WJ**, which demonstrates the selectivity of **1** for A/T-rich 3WJs (Fig. S4 in ESI†). The titration of **1** with a regular double-stranded DNA with a single A/T-Hook binding site resembles that of the junction **TA-3WJ**, although higher concentrations of dsDNA are required to disassemble the initial aggregate and to reach the titration plateau. This is reflected in a much weaker dissociation constant, approximately 10 times weaker than for the **TA-3WJ** target (Fig. 3c). The interaction of **1** with the minor groove of the duplex branches in **TA-3WJ** was confirmed by the changes in the Circular Dichroism spectrum of the junction **TA-3WJ** upon addition of **1** (Fig. S5a in ESI†), which were like those observed for a model dsDNA containing a target A/T-rich sequence (Fig. S5b in ESI†). Furthermore, no changes in the CD were observed when the **GC-3WJ** junction was incubated with the probe **1** (Fig. S5c in ESI†).

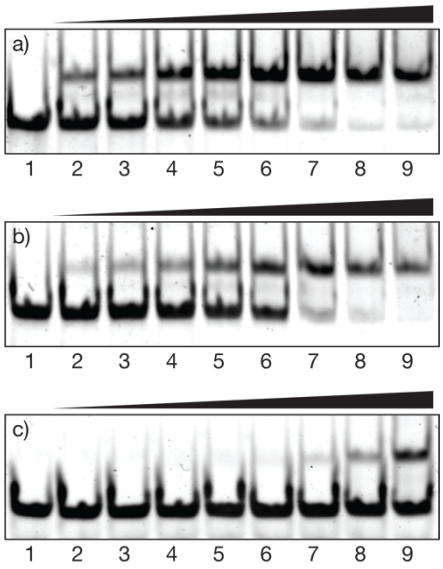


Fig. 4. DNA binding properties of **1** by EMSA. The concentrations, in all cases, are: Lanes 1-9, 200 nM of the corresponding DNA under study, and 0, 100, 200, 300, 400, 500, 600, 700 and 800 nM of **1**, respectively. EMSA experiments were resolved on a 10% nondenaturing polyacrylamide gel and 0.5× 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.

Following the spectroscopic studies we decided to characterize in more detail its DNA binding properties by electrophoretic gel shift assays under non-denaturing conditions.[56,57] 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. 4a). 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. 4b).[58–61] This is also consistent with the high electrostatic contribution to DNA binding with oligocationic peptides such as the AT-Hook.[36] 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. 4c). 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).

Having demonstrated that **1** selectively binds to A/T-rich 3WJs, we studied its application to label these structures in cells. Preliminary experiments showed that **1** was poorly internalized (incubation for 30 min of 5 μM concentration of **1**). However, if the cells were pretreated with digitonin then 1 was efficiently internalized,[62,63] and appeared in the blue emission channel as a punctuated pattern in the cytoplasm and the nucleus (Fig. 5a). Remarkably, the punctuated distribution in the cell nucleus has a strong correlation (Pearson's correlation coefficient 0.60)[64,65] with the localization of the DNA replicating sites labeled with the proliferating cell nuclear antigen (PCNA) fused to GFP,[66,67] this partial colocalization is consistent with **1** selectively targeting A/T-rich 3WJs formed during the replication process in functional cells, and not all 3WJs. (Fig. 5b and 5c). The staining in the cytoplasm and of what appears to be the cell nucleolus is consistent with the binding of **1** to RNA structures, which have been shown to be potential binding targets of the AT-Hook peptides.[68]

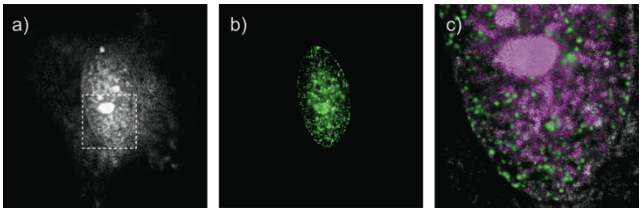


Fig. 5. Selectively stains DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated with 25 μg/ml digitonin for 3 min, then 5 μM of **1** for 30 min. a) Blue channel emission showing the distribution of **1** (channel fake colored in white for easier view); b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci; 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**.

In summary, we have demonstrated a new approach for selectively recognizing 3WJ DNAs through sequence-selective and cooperative binding of three AT-Hook peptides in the duplex branches of the junction. Moreover, we have exploited the fluorogenic properties of the 1,3,5-tristyrylbenzene core and the increased emission resulting from the conformational restrictions associated with 3WJ binding to selectively label DNA replication sites in functional cells.

This work was supported by the Spanish grants RTI2018-099877-B-I00, CTQ2017-84561-P, PID2019-105308RB-I00 and CTQ2017-87889-P. The *Xunta de Galicia* (ED431B 2018/04, Centro singular de investigación de Galicia accreditation 2016–2019, ED431G/09, and ED431B 2018/04) and the European Union (European Regional Development Fund - ERDF) are gratefully acknowledged. M. V. L. thanks the Fundación Científica de la Asociación Española Contra el Cáncer (Ideas Semilla 2021 – IDEAS211154VAZQ). J.-D.M. acknowledges the support of the *Generalitat de Catalunya* (2017SGR1323). J. G. M. thanks the Junta de Comunidades de Castilla-La Mancha (SBPLY/17/180501/000214) and the Universidad de Castilla-La Mancha (2020-GRIN-28866, 2021-GRIN-30998). J. G.-G. thanks the Spanish *Ministry of Science and Innovation/Spanish Research Agency* for his FPI fellowship. Plasmid GFP-PCNAL2 was generously provided by Dr. Cristina Cardoso. Molecular graphics with *UCSF Chimera*, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.[69]

Conflicts of interest

There are no conflicts to declare

Notes and references

1 P. Brookes, *Mutat. Res.*, 1990, **233**, 3–14.
2 K. W. Kohn, *Cancer Res.*, 1996, **56**, 5533–5546.
3 J. Sun, Q. Wei, Y. Zhou, J. Wang, Q. Liu and H. Xu, *BMC Syst. Biol.*, 2017, **11**, 87.
4 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discov.*, 2011, **10**, 261–275.
5 J. Rodríguez, J. Mosquera, J. R. Couceiro, M. E. Vázquez and J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.*, 2016, **55**, 15615–15618.
6 K. J. Neelsen and M. Lopes, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 207–220.
7 A. Bacolla and R. D. Wells, *Mol. Carcinog.*, 2009, **48**, 273–285.
8 R. D. Wells, *Trends Biochem. Sci.*, 2007, **32**, 271–278.
9 H. Tateishi-Karimata and N. Sugimoto, *Nucleic Acids Res.*, 2021, **49**, 7839–7855.
10 J. Zell, F. R. Sperti, S. Britton and D. Monchaud, *RSC Chem. Biol.*, 2021, **2**, 47–76.
11 H. Tateishi-Karimata and N. Sugimoto, *Chem. Commun.* , 2020, **56**, 2379–2390.
12 L. A. Howell and M. Searcey, *Chembiochem*, 2009, **10**, 2139–2143.
13 S. M. Nelson, L. R. Ferguson and W. A. Denny, *Cell Chromosome*, 2004, **3**, 2.
14 J. Atkinson and P. McGlynn, *Nucleic Acids Res.*, 2009, **37**, 3475–3492.
15 L. S. Shlyakhtenko, V. N. Potaman, R. R. Sinden, A. A. Gall and Y. L. Lyubchenko, *Nucleic Acids Res.*, 2000, **28**, 3472–3477.
16 R. R. Sinden, M. J. Pytlos-Sinden and V. N. Potaman, *Front. Biosci.*, 2007, **12**, 4788–4799.
17 H. Liao, F. Ji, T. Helleday and S. Ying, *EMBO Rep.*, 2018, **19**, e46263.
18 K. Duskova, P. Lejault, É. Benchimol, R. Guillot, S. Britton, A. Granzhan and D. Monchaud, *J. Am. Chem. Soc.*, 2020, **142**, 424–435.

19 K. C. Woods, S. S. Martin, V. C. Chu and E. P. Baldwin, *J. Mol. Biol.*, 2001, **313**, 49–69.

20 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, *Angew. Chem. Int. Ed.*, 2006, **45**, 1227–1231.

21 S. Vuong, L. Stefan, P. Lejault, Y. Rousselin, F. Denat and D. Monchaud, *Biochimie*, 2012, **94**, 442–450.

22 J. Gómez-González, Y. Pérez, G. Sciortino, L. Roldan-Martín, J. Martínez-Costas, J.-D. Maréchal, I. Alfonso, M. Vázquez López and M. E. Vázquez, *Angew. Chem. Int. Ed Engl.*, 2021, **60**, 8859–8866.

23 J. Gómez-González, D. Bouzada, L. A. Pérez-Márquez, G. Sciortino, J.-D. Maréchal, M. Vázquez López and M. E. Vázquez, *Bioconjug. Chem.*, 2021, **32**, 1564–1569.

24 J. Gómez-González, D. G. Peña, G. Barka, G. Sciortino, J.-D. Maréchal, M. Vázquez López and M. E. Vázquez, *Front Chem*, 2018, **6**, 520.

25 I. Gamba, G. Rama, E. Ortega-Carrasco, J.-D. Maréchal, J. Martínez-Costas, M. E. Vázquez and M. Vázquez López, *Chem. Commun.* , 2014, **50**, 11097–11100.

26 K. Duskova, J. Lamarche, S. Amor, C. Caron, N. Queyriaux, M. Gaschard, M.-J. Penouilh, G. de Robillard, D. Delmas, C. H. Devillers, A. Granzhan, M.-P. Teulade-Fichou, M. Chavarot-Kerlidou, B. Therrien, S. Britton and D. Monchaud, *J. Med. Chem.*, 2019, **62**, 4456–4466.

27 L. Guyon, M. Pirrotta, K. Duskova, A. Granzhan, M.-P. Teulade-Fichou and D. Monchaud, *Nucleic Acids Res.*, 2018, **46**, e16.

28 J. Novotna, A. Laguerre, A. Granzhan, M. Pirrotta, M.-P. Teulade-Fichou and D. Monchaud, *Org. Biomol. Chem.*, 2015, **13**, 215–222.

29 S. A. Barros and D. M. Chenoweth, *Angew. Chem. Weinheim Bergstr. Ger.*, 2014, **126**, 13966–13970.

30 S. A. Barros and D. M. Chenoweth, *Chem. Sci.*, 2015, **6**, 4752–4755.

31 Z. Yang, Y. Chen, G. Li, Z. Tian, L. Zhao, X. Wu, Q. Ma, M. Liu and P. Yang, *Chem.--Eur. J.*, 2018, **24**, 6087–6093.

32 L. Cardo and M. J. Hannon, *Inorganica Chim. Acta*, 2009, **362**, 784–792.

33 I. Yoon, S.-E. Suh, S. A. Barros and D. M. Chenoweth, *Org. Lett.*, 2016, **18**, 1096–1099.

34 L. Cardo, V. Sadovnikova, S. Phongtongpasuk, N. J. Hodges and M. J. Hannon, *Chem. Commun.* , 2011, **47**, 6575–6577.

35 L. Aravind and D. Landsman, *Nucleic Acids Res.*, 1998, **26**, 4413–4421.

36 C. Crane-Robinson, A. I. Dragan and P. L. Privalov, *Trends Biochem. Sci.*, 2006, **31**, 547–552.

37 R. Reeves and M. S. Nissen, *J. Biol. Chem.*, 1990, **265**, 8573–8582.

38 T. Lund, K. H. Dahl, E. Mørk, J. Holtlund and S. G. Laland, *Biochem. Biophys. Res. Commun.*, 1987, **146**, 725–730.

39 B. H. Geierstanger, B. F. Volkman, W. Kremer and D. E. Wemmer, *Biochemistry*, 1994, **33**, 5347–5355.

40 J. F. Maher and D. Nathans, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 6716–6720.

41 J. P. Wagner, D. M. Quill and D. E. Pettijohn, *J. Biol. Chem.*, 1995, **270**, 7394–7398.

42 J. R. Huth, C. A. Bewley, M. S. Nissen, J. N. Evans, R. Reeves, A. M. Gronenborn and G. M. Clore, *Nat. Struct. Biol.*, 1997, **4**, 657–665.

43 E. Fonfría-Subirós, F. Acosta-Reyes, N. Saperas, J. Pous, J. A. Subirana and J. L. Campos, *PLoS One*, 2012, **7**, e37120.

44 S. Learte-Aymamí, J. Rodríguez, M. E. Vázquez and J. L. Mascareñas, *Chemistry*, 2020, **26**, 8875–8878.

45 J. Rodríguez, J. Mosquera, J. R. Couceiro, M. E. Vázquez and J. L. Mascareñas, *Chem. Sci.*, 2015, **6**, 4767–4771.

46 J. Rodríguez, J. Mosquera, R. García-Fandiño, M. E. Vázquez and J. L. Mascareñas, *Chem. Sci.*, 2016, **7**, 3298–3303.

47 J. Rodríguez-Guerra Pedregal, G. Sciortino, J. Guasp, M. Municoy and J.-D. Maréchal, *J. Comput. Chem.*, 2017, **38**, 2118–2126.

48 I. Coin, M. Beyermann and M. Bienert, *Nat. Protoc.*, 2007, **2**, 3247–3256.

49 G. S. Vanier, *Methods Mol. Biol.*, 2013, **1047**, 235–249.

50 B. Bacsa, B. Desai, G. Dibó and C. O. Kappe, *J. Pept. Sci.*, 2006, **12**, 633–638.

51 R. Domínguez, M. Moral, M. P. Fernández-Liencres, T. Peña-Ruiz, J. Tolosa, J. Canales-Vázquez, J. C. García-Martínez, A. Navarro and A. Garzón-Ruiz, *Chemistry*, 2020, **26**, 3373–3384.

52 A. Garzón, M. P. Fernández-Liencres, M. Moral, T. Peña-Ruiz, A. Navarro, J. Tolosa, J. Canales-Vázquez, D. Hermida-Merino, I. Bravo, J. Albaladejo and J. C. García-Martínez, *J. Phys. Chem. C*, 2017, **121**, 4720–4733.

53 F. de Lera-Garrido, A. Sánchez-Ruiz, J. Rodríguez-López, J. Tolosa and J. C. García-Martínez, *Dyes Pigm.*, 2020, **179**, 108410.

54 K. Yabuki, T. Ouchi and Y. Ohya, *Supramol. Chem.*, 2003, **15**, 149–154.

55 T. Heinlein, J.-P. Knemeyer, O. Piestert and M. Sauer, *J. Phys. Chem. B*, 2003, **107**, 7957–7964.

56 L. M. Hellman and M. G. Fried, *Nat. Protoc.*, 2007, **2**, 1849–1861.

57 R. S. Tuma, M. P. Beaudet, X. Jin, L. J. Jones, C. Y. Cheung, S. Yue and V. L. Singer, *Anal. Biochem.*, 1999, **268**, 278–288.

58 E. K. Liebler and U. Diederichsen, *Org. Lett.*, 2004, **6**, 2893–2896.

59 G.-X. He, K. A. Browne, A. Blasko and T. C. Bruice, *J. Am. Chem. Soc.*, 1994, **116**, 3716–3725.

60 C. Portela, F. Albericio, R. Eritja, L. Castedo and J. L. Mascareñas, *Chembiochem*, 2007, **8**, 1110–1114.

61 O. Vázquez, M. E. Vázquez, J. B. Blanco, L. Castedo and J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.*, 2007, **46**, 6886–6890.

62 M. Nishikawa, S. Nojima, T. Akiyama, U. Sankawa and K. Inoue, *J. Biochem.*, 1984, **96**, 1231–1239.

63 M. P. Stewart, R. Langer and K. F. Jensen, *Chem. Rev.*, 2018, **118**, 7409–7531.

64 K. W. Dunn, M. M. Kamocka and J. H. McDonald, *Am. J. Physiol. Cell Physiol.*, 2011, **300**, C723-42.

65 J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri,

P. Tomancak and A. Cardona, *Nat. Methods*, 2012, **9**, 676–682.
66 H. Leonhardt, H. P. Rahn, P. Weinzierl, A. Sporbert, T. Cremer, D. Zink and M. C. Cardoso, *J. Cell Biol.*, 2000, **149**, 271–280.
67 R. Yokoyama, T. Hirakawa, S. Hayashi, T. Sakamoto and S. Matsunaga, *Sci. Rep.*, 2016, **6**, 29657.
68 M. Filarsky, K. Zillner, I. Araya, A. Villar-Garea, R. Merkl, G. Längst and A. Németh, *RNA Biol.*, 2015, **12**, 864–876.
69 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

