Nickel-promoted Recognition of long DNA sites by designed Peptide Derivatives

Jéssica Rodríguez,[a] Jesús Mosquera,[a] M. Eugenio Vázquez,*[a] and José L. Mascareñas*[a]

Abstract: We describe the synthesis of designed peptidic modules that self-assemble in specific DNA sequences of 12 base pairs in the presence of Ni(II) salts. The modules consist of modified fragments of transcription factors that have been appropriately engineered to include metal-chelating His and bipyridine ligands.

The initiation of gene expression requires the exquisite molecular recognition of regulatory DNA sites by specialized proteins called transcription factors (TF).[1] In general, TFs interact with DNA as non-covalent homo- or heterodimers, which ensures high affinity while allowing the selective recognition of long DNA sites.[2] As chemists, the introduction of unnatural modifications into protein scaffolds has allowed the exploration of new DNA recognition strategies [3] We have recently demonstrated that synthetic non-natural oligomers that combine several DNA binding modules can be used for the specific recognition of relatively long sites.[4]

However, the preparation of these covalent oligomers requires elaborated synthetic procedures that considerably restrict the versatility of the strategy. We therefore considered the application of metal coordination as an alternative approach to obtaining multimeric DNA binders resulting from the metal-promoted combination of different modules. Indeed, we have recently demonstrated the nickel-promoted assembly of chelating peptides and bipyridine-derivatives of bis-benzamidine minor groove binders that selectively recognize DNA sites of up to 8 base pairs.[5] Herein we demonstrate that this type of metal-directed supramolecular strategy can be used for the recognition of longer DNA sites provided one of the components presents a bivalent-DNA interacting motive that targets an 8-9 bp long site. In our new design, the bipyridine ligand is incorporated in an AT-hook peptide sequence instead of a bis-benzamidine,[6] which represents an huge advantage in terms of synthetic access and versatility.

For our design we chose GCN4, an archetypical bZIP TF that specifically binds to ATF/CREB (5'-ATGA(c/g)TCAT-3') or AP1 (5'-ATGA(c)TCAT-3') sites as a leucine zipper-mediated homodimer of uninterrupted α-helices.[7] The DNA recognition takes place through the N-terminal basic region (GCN4br), which folds into an α-helix upon DNA binding.[8] Whereas an isolated basic region peptide fails to interact with its DNA consensus site with measurable affinity, we have previously shown that appropriate conjugation to minor groove binders, included an AT-hook anchor, allows the specific recognition of composite target sites featuring the binding sequences of GCN4br and the minor groove binder in tandem.[6,9]

We envisioned that incorporating a metal-chelating bipyridine moiety at the N-terminus of fully peptidic GCN4br/AT-hook conjugates (**brHk**) might allow the implementation of new, emergent DNA binding properties. In particular, we considered that in the presence of an appropriate metal, this hybrid might trap a designed GCN4br mutant featuring two coordinating histidines (**brHis$_2$**), and bind a cognate target DNA site with an extended palindromic TCAT-A_2T_2-ATGA sequence. Both basic regions are based on the GCN4 fragment comprising residues Asp^{226} to Gln^{248}, which is the shortest peptide that retains the specific DNA binding properties of the full GCN4 DNA-binding domain in the form of a disulfide dimer.[10] **brHis$_2$** features a GCN4 basic region with residues at positions 230 and 234 mutated to histidines, which, according to the X-ray structure of the GCN4-DNA complex, would be located in the outer face of the recognition helix (Fig 1).[7a] The histidine residues in **brHis$_2$** might not only mediate the formation of the complex, but also act as coordinative staples to stabilize the α-helix required for DNA recognition.[5,11,12]

[a] J. Rodríguez, Dr. J. Mosquera, A/Prof. M. E. Vázquez, Prof. J. L. Mascareñas
Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Departamento de Química Orgánica. Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain.
E-mail: joseluis.mascarenas@usc.es

Supporting information for this article is given via a link at the end of the document.

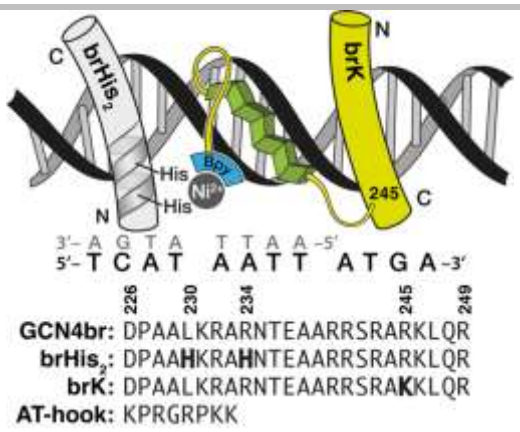
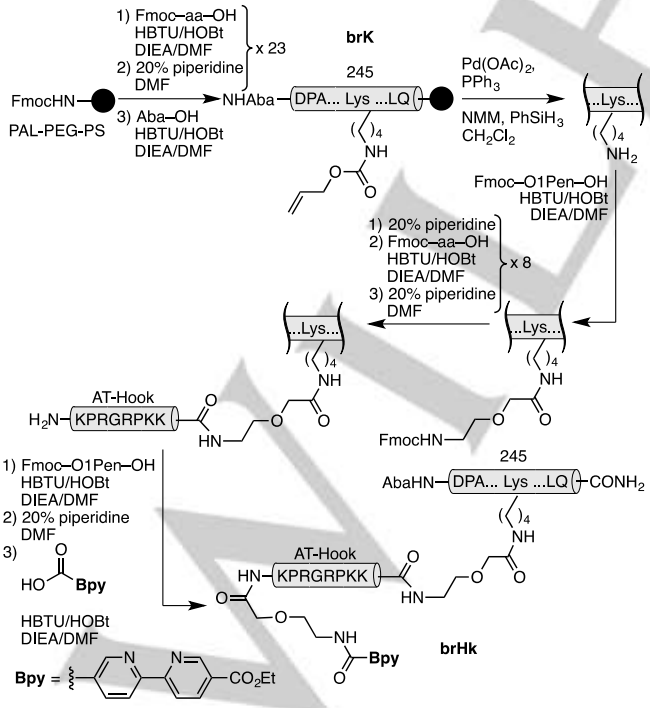


Figure 1. Top: Schematic structure of the supramolecular assembly in a composite palindromic target site. His[230] and/or His[234] residues might coordinate Ni(II) and connect **brHis$_2$** with the GCN4br/AT-hook chimera **brHk** [composed by **brK** and a bipyridine (Bpy)–containing AT-hook]. Bottom: Sequences of the peptides components of the supramolecular binder. Mutated residues with respect to natural GCN4 are in bold.

The GCN4br-based peptides **brHis$_2$** and **brHk** were synthesized following standard Fmoc/tBu solid-phase peptide synthesis (SPPS) procedures,[13] and required the orthogonal modification of the GCN4 basic region (catalytic alloc removal),[14] as well as the incorporation of the bipyridine ligand (Bpy) while the peptide is still attached to the solid support (Scheme 1).[15] The synthetic method is extremely practical as involves exclusively solid phase peptide chemistry, and therefore the product can be efficiently obtained in a couple of working days.



Scheme 1. Synthesis of the GCN4br/AT-hook chimera **brHk**. The 4-acetamidobenzoic acid chromophore (Aba) is introduced at the N-terminus of the GCN4 basic region as a spectroscopic reporter.

Having at hand the desired peptides, we studied their DNA binding by electrophoretic mobility shift assays (EMSA) in polyacrylamide gel under non-denaturing conditions,[16] and using SybrGold as DNA stain.[17] As expected, incubation of the GCN4br/AT-hook hybrid peptide (**brHk**) with a double-stranded oligonucleotide featuring the composite binding site (TCAT-A_2T_2-ATGA), induces the formation of a retarded band in the gel corresponding to the expected DNA-peptide complex (Fig. 2a, lane 2).[9] Addition of 20 equiv of $Ni(ClO_4)_2$ to this mixture does not generate new retarded bands (Fig. 2a, lane 3), but subsequent addition of the metal-chelating **brHis$_2$** gives rise to a new (more retarded) band that must correspond to the desired DNA/[(**brHis$_2$**)(**brHk**)Ni]$^{2+}$ ternary complex (Fig. 2a, lane 4). The complex can also be assembled by addition of Ni^{2+} to a mixture of **brHis$_2$** and **brHk** in the presence of the target DNA (Fig. 2b, lanes 1-2), or by addition of **brHk** to a mixture of the other components (Fig. 2b, lanes 3-4). Taken together, these results show that the system assembles dynamically regardless of the order in which the components are mixed. Therefore, the potential competitive homodimeric species [(**brHis$_2$**)$_2$Ni]$^{2+}$ or [(**brHk**)$_2$Ni]$^{2+}$ are kinetically labile and evolve towards the desired [(**brHis$_2$**)(**brHk**)Ni]$^{2+}$ complex in the presence of the target DNA oligonucleotide.

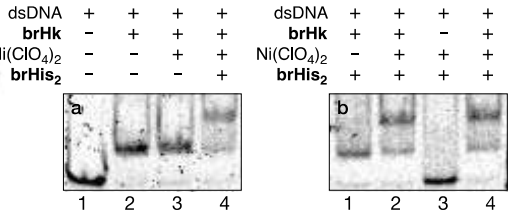


Figure 2. EMSA DNA binding studies results with target oligonucleotide 5´–GCGAG-TCAT-A_2T_2-ATGA-AGGCG–3´ (one strand shown). Concentration of the components are, when present: 75 nM dsDNA, 1 μM **brHk**, 2 μM **brHis$_2$**, and 20 μM $Ni(ClO_4)_2$. Experiment was resolved by PAGE on a 10% nondenaturing polyacrylamide gel and 0.5× TBE buffer over 40 min at 4 °C, and analyzed by staining with SyBrGold (5 μL in 50 mL of 1× TBE) for 10 min, followed by fluorescence visualization. Note: The slight smearing of the bands comes from the absence of EDTA additives, commonly used in these assays.

Regarding the selectivity in DNA binding, we found that a dsDNA mutated in the basic region binding sites (ctctc-A_2T_2-gagag), did not elicit the formation of stable complexes (Fig 3a). Using a DNA mutated in the target site of the basic region of **brHk** (TCAT-A_2T_2-gaga), we observed the formation of a retarded band (Fig 3b) that migrates faster than that arising from the specific binary complex between **brHk** and its cognate sequence (compare with lane 2, Fig. 3d). This can be explained in terms of a bipartite interaction of the basic region of **brHk** with its target site (TCAT) together with non-specific electrostatic

contacts of the AT-hook with the DNA phosphates. Indeed, the same type of band, albeit slightly more diffuse, was observed using a DNA which lacks the A/T-rich tract required for insertion of the AT-hook (TCAT-cagg-ATGA, Fig 3c). In none of the latter cases the addition of the nickel salt and **brHis$_2$** promoted the formation of any stable new complex.

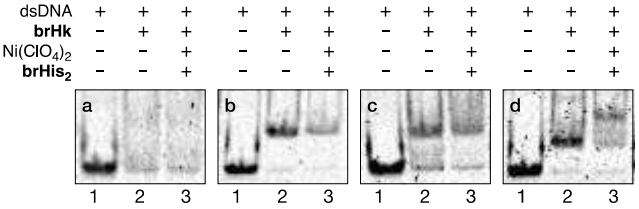


Figure 3. EMSA DNA binding studies with a series of mutated oligonucleotides (see the ESI for full sequences). **a**: tctc-A_2T_2-gaga (both peptide sites mutated); **b**: TCAT-A_2T_2-gaga (only the binding site for **brHk** mutated); **c**: TCAT-cagg-ATGA (AT-hook site mutated); **d**: TCAT-TA_2T_2-ATGA (extended AT-hook site). Mutated sites in lower case. Concentrations of the components are, when present: 75 nM dsDNA, 1 μM **brHk**, 2 μM **brHis$_2$**, and 20 μM $Ni(ClO_4)_2$. Experiment was resolved by PAGE on a 10% nondenaturing polyacrylamide gel and 0.5× TBE buffer over 40 min at 4 °C, and stained with SyBrGold.

These results support the requirement of all consensus binding sites for each component for observing the desired metal-dependent formation of the ternary complex. The design tolerated a slightly larger A/T spacer, as shown in the Fig 3d for the interaction with a DNA having a TA_2T_2 sequence between the basic region consensus sites, albeit in this case the interaction appears to be slightly weaker.

We also studied the selectivity of the system with regard to the metal ion (Co^{2+}, Ni^{2+}, Zn^{2+}, Cu^{2+}, and Fe^{2+} salts). Using comparable conditions, we only observed the formation of more slowly migrating bands in the presence of Ni^{2+} cations (see ESI).

In order to obtain more information on the DNA binding properties of the supramolecular assembly, we performed fluorescence anisotropy titrations by adding $Ni(ClO_4)_2$ to a solution containing a tetramethylrhodamine-labeled ds-oligonucleotide with the target sequence, and the constitutive peptides, **brHk** and **brHis$_2$**. As shown in Fig 4, titration with the nickel complex led to a progressive increase in the anisotropy. Considering a simplified 1:1 binding mode in which the [(**brHis$_2$**)(**brHk**)Ni]$^{2+}$ complex act as a single species binding to the DNA, we calculated an apparent dissociation constant of 330 nM at 4 °C for a 1:1 binding mode (Fig 4).

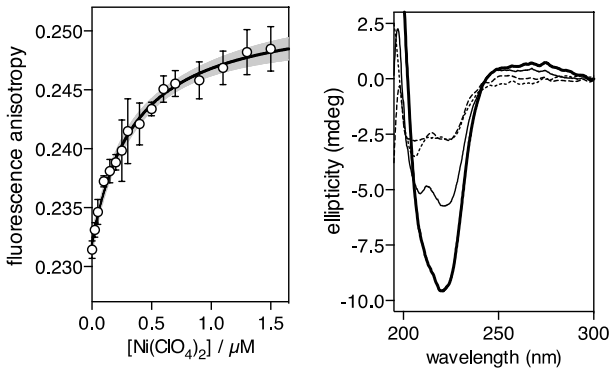


Figure 4. Left: Fluorescence anisotropy titration at 559 nm of a mixture containing 25 nM of tetramethylrhodamine (TMR) labeled ds-oligonucleotide (TMR-5'–GCGAG-TCAT-A_2T_2-ATGA-AGGCG–3'), 300 nM of **brHk** and 300 nM **brHis$_2$** with increasing concentrations of $Ni(ClO_4)_2$. The best fit to a 1:1 binding model using a simplified model is also shown, including 95% confidence of the fit shown in light grey. Experimental data correspond to the mean of three independent titration experiments. Right: Circular dichroism of a 5 μM solution of **brHis$_2$** in 10 mM phosphate buffer pH 7.5, 100 mM NaCl (dotted line), the same solution after addition of 1 equiv of the target dsDNA TCAT-A_2T_2-ATGA (dashed line), then 0.5 equiv of **brHk** (solid line), and finally 10 equiv of $Ni(ClO_4)_2$ (thick line). The contributions of the DNA to the CD havebeen subtracted for clarity. All experiments carried out at 4 °C.

In addition to the EMSA and fluorescence studies, circular dichroism provided relevant information about the binding mechanism. Thus, as expected for a poorly structured peptide, the CD spectrum of **brHis$_2$** presents a relatively weak negative signal at 222 nm, even in the presence of **brHk** and the consensus DNA (Fig 4, thin line). Addition of $Ni(ClO_4)_2$ to this DNA/**brHis$_2$**/**brHk** mixture promoted a considerable increase in the negative ellipticity at 222 nm, which is consistent with the folding of the modified basic regions into α-helices upon formation of the desired complex (Fig 4, thick line).[18]

The key role of the Ni(II) ion as an adaptor component that mediates the assembly of the supramolecular DNA binder prompted us to examine the possibility of disassembling the [(**brHis$_2$**)(**brHk**)Ni]$^{2+}$/DNA complex by using an external chelator.[19] Gratifyingly, the addition of EDTA to the preformed complex that resulted from mixing **brHis$_2$**, **brHk**, the Ni^{2+} salt, in the presence of the DNA TCAT-A_2T_2-ATGA, promoted a disassembly of the complex (lane 4, Fig. 5 left). CD also showed a decrease in the degree of helicity (Fig. 5 right). Therefore, whereas the presence of nickel is critical for the DNA interaction, the resulting multicomponent supramolecular complex can be taken apart by the addition of a nickel chelator.

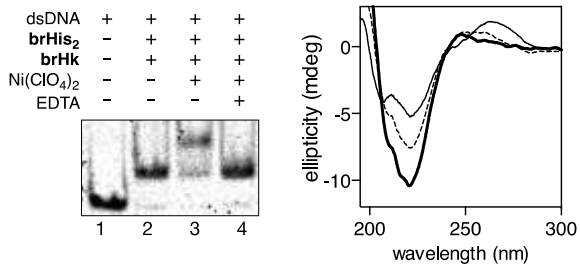


Figure 5. Left: EMSA results before and after addition of EDTA (1.8 mM). Concentrations of the components are, when present: 75 nM dsDNA, 1 μM **brHk**, 2 μM **brHis$_2$**, and 20 μM $Ni(ClO_4)_2$. Experiment was resolved by PAGE on a 10% nondenaturing polyacrylamide gel and 0.5× TBE buffer over 40 min at 4 ºC, and analyzed by staining with SyBrGold (Molecular Probes: 5 μL in 50 mL of 1× TBE) for 10 min, followed by fluorescence visualization. Right: Circular dichroism of the mixture in presence of TCAT-A_2T_2-ATGA dsDNA (5 μm) in 10 mM phosphate buffer pH 7.5 100 mM of NaCl (thick line), and after subsequent addition of 2 equiv (dashed line) or 10 equiv (solid line) of EDTA. Concentrations of the components are, when present: 5 μM dsDNA, 2.5 μM **brHk**, 5 μM **brHis$_2$**, and 50 μM $Ni(ClO_4)_2$. The contributions of the DNA to the CD have been subtracted for clarity. All experiments were carried out at 4 °C.

In summary, we have introduced a new approach for achieving a conditional and selective trivalent recognition of a long 12 base-pair DNA site. The strategy involves the nickel-promoted assembly of a supramolecular DNA binder composed by a bZIP basic region/AT-hook chimera equipped with a bipyridine ligand (**brHk**) and a bis(histidine)-modified basic region (**brHis$_2$**). Both the multicomponent nature of the system, and the kinetic lability of the metal coordination facilitate the disassembly of the supramolecular structure upon addition of external agents that sequester the nickel cation. Importantly, the trivial synthetic access to both peptide derivatives, based in solid-phase peptide synthesis, together with the modular nature of the metallo-supramolecular assembly process, promises further applications of the tactic for recognizing different sequences, or obtaining other functional derivatives.

Acknowledgements

We thank the support given by the Spanish grants SAF2013-41943-R, CTQ2015-70698-R, and CTQ2013-49317-EXP, the Xunta de Galicia GRC2013-041, the ERDF, and the European Research Council (Advanced Grant No. 340055). Support of COST Action CM1105, COST CM1306 and the orfeo-cinqa network are also kindly acknowledged. J. R. thanks the *Xunta de Galicia* for a PhD fellowship.

Keywords: DNA recognition · metallosupramolecular chemistry · peptides · self assembly · minor groove binder.

[1] C. W. Garvie, C. Wolberger, *Mol. Cell* **2001**, *8*, 937-946.

[2] a) H. C. Nelson, *Curr. Opin. Genet. Dev*. **1995**, *5*, 180-189; b) R. Moretti, A. Z. Ansari, *Biochimie* **2008**, *90*, 1015-1025; c) D. J. Segal, C. F. Barbas, *Curr. Opin. Chem. Biol*. **2000**, *4*, 34-39; d) L. Chen, *Curr. Opin. Struct. Biol*., **1999**, *9*, 48-55.

[3] a) B. Cuenoud, A. Schepartz, *Science* **1993**, *259*, 510-513; b) C. R. Palmer, S. S. Sloan, J. C. Adrian, B. Cuenoud, D. N. Paolella, A. Schepartz, *J. Am. Chem. Soc*. **1995**, *117*, 8899-8907; c) M. E. Vázquez, A. M. Caamaño, J. L. Mascareñas, *Chem. Soc. Rev*. **2003**, *32*, 338-349; d) E. Pazos, J. Mosquera, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2011**, *12*, 1958-1973; e) C. Y. Majmudar, A. K. Mapp, *Curr. Opin. Chem. Biol*. **2005**, *9*, 467-474; f)) J. B. Blanco, O. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Eur. J*. **2005**, 11, 4171-4178; i) O. Vazquez, M. E. Vazquez, J. B. Blanco-Canosa, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed*. **2007**, *46*, 6886-6890; j) A. Iyer, D. Van Lysebetten, Y. Ruiz García, B. Louage, B. G. De Geest, A. Madder, *Org. Biomol. Chem*. **2015**, *13*, 3856-3862; k) J. Mosquera, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2014**, *15*, 1092-1095; l) O. Vázquez, M. I. Sánchez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Org. Lett*. **2010**, *12*, 216-219; m) T. Morii, Y. Saimei, M. Okagami, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1997,** *119,* 3649-3655; n) S.-I. Sato, M. Hagihara, K. Sugimoto, T. Morii, *Chem. Eur. J.* **2002,** *8,* 5066-5071; o) A. Iyer, D. Van Lysebetten, Y. Ruiz García, B. Louage, B. G. De Geest, A. Madder, *Org. Biomol. Chem.* **2015,** *13,* 3856-3862; p) J. J. Hollenbeck, D. G. Gurnon, G. C. Fazio, J. J. Carlson, M. G. Oakley, *Biochemistry* **2001,** *40,* 13833-13839.

[4] J. Rodríguez, J. Mosquera, R. García-Fandiño, M. E. Vázquez, J. L. Mascareñas *Chem. Sci*. **2016**, *7*, 3298-3303.

[5] M. I. Sánchez, J. Mosquera, M. E. Vázquez and J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2014**, *53*, 9917-9921.

[6] J. Rodríguez, J. Mosquera, J. R. Couceiro, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2015**, *6*, 4767-4771.

[7] a) T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **1992**, 71, 1223-1237; b) P. König, T. J. Richmond, *J. Mol. Biol*. **1993**, *233*, 139-154; c) W. Keller, P. König, T. J. Richmond, *J. Mol. Biol*. **1995**, *254*, 657-667.

[8] M. A. Weiss, T. Ellenberger, C. R. Wobbe, J. P. Lee, S. C. Harrison, K. Struhl, *Nature* **1990**, *347*, 575-578.

[9] a) M. E. Vazquez, A. M. Caamaño, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2001,** *40,* 4723-4725; b) J. B. Blanco, M. E. Vazquez, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003,** *10,* 713-722; c) M. I. Sánchez, O. Vazquez, J. Martinez-Costas, M. E. Vazquez, J. L. Mascareñas, *Chem. Sci.* **2012,** *3,* 2383-2387.

[10] a) R. V. Talanian, C. J. McKnight, R. Rutkowski, P. S. Kim, *Biochemistry* **1992**, *31*, 6871-6875; b) R. V. Talanian, C. J. McKnight, P. S. Kim, *Science* **1990**, *249*, 769-771; c) Y. Ruiz-García, J. Zelenka, V. Pabon, A. Iyer, M. Buděšínský, T. Kraus, E. Smith, A. Madder, *Org. BIomol. Chem*. **2015**, *13*, 5273-5278; d) L. Carrette, T. Morii, A. Madder, *Eur. J. Org. Chem*. **2014**, *2014*, 2883-2891; e) G. Bullen, J. Tucker, A. Peacock, *Chem. Commun*. **2015**, *51*, 8130-8133.

[11] a) S. Futaki, T. Kiwada, Y. Sugiura, *J. Am. Chem. Soc*. **2004**, *126*, 15762-15769; b) M. J. Kelso, H. N. Hoang, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2000**, *122*, 10488-10489; c) M. R. Ghadiri, A. K. Fernholz, *J. Am. Chem. Soc*. **1990**, *112*, 9633-9635; d) M. J. Kelso, H. N. Hoang, W. V. Oliver, N. Sokolenko, D. R. March, T. G. Appleton D. P. Fairlie, *Angew. Chem. Int. Ed*. **2003**, *42*, 421-424; e) S. J. Smith, R. J. Radford, R. H. Subramanian, B. R. Barnett, J. S. Figueroa, F. A. Tezcan, *Chem. Sci*. **2016**, DOI: 10.1039/C6SC00826G.

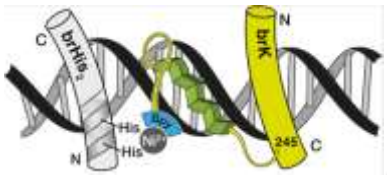
[12] To design the synthetic targets we used the X-ray structural information of the DNA complexes of GCN4 (PDB: ID 1YSA) and AT-hook (PDB: ID 3UXW).

[13] I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc*. **2007**, *2*, 3247-3256.
[14] N. Thieriet, J. Alsina, E. Giralt, F. Guibe, F. Albericio, *Tetrahedron Lett.* **1997**, *38*, 7275-7278.
[15] a) G. Rama, A. Ardá, J. –D. Maréchal, I. Gamba, H. Ishida, J. Jiménez-Barbero, M. E. Vazquez, M. Vázquez López, *Chem. Eur. J.* **2012**, *18*, 7030-7035; b) H. Ishida, M. Kyakuno S. Oishi, *Biopolymers* **2004**, *76*, 69-82.
[16] L. M. Hellman, M. G. Fried, *Nat. Protoc*. **2007**, *2*, 1849-1861.
[17] R. S. Tuma, M. P. Beaudet, X. Jin, L. J. Jones, C. Y. Cheung, S. Yue, V. L. Singer, *Ana.l Biochem.* **1999**, *268*, 278-288.
[18] J. T. Pelton, L. R. McLean, *Anal. Biochem*. **2000**, *277*, 167-176.
[19] Some examples of stimuli-responsive DNA binders: a) M. Ueno, A. Murakami, K. Makino, T. Morii, *J. Am. Chem. Soc.* **1993,** *115,* 12575-12576; b) T. Morii, J. Yamane, Y. Aizawa, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1996,** *118,* 10011-10017; c) A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed*. **2000**, *39*, 3104-3107; d) Y. Azuma, M. Imanishi, T. Yoshimura, T. Kawabata, S. Futaki, *Angew. Chem. Int. Ed*. **2009**, *48*, 6853-6856; e) A. Onoda, N. Arai, N. Shimazu, H. Yamamoto, T. Yamamura, T. *J. Am. Chem. Soc*. **2005**, *127*, 16535-16540; f) J. Mosquera, A. Jiménez, V. I. Dodero, M. E. Vázquez, J. L. Mascareñas, *Nat. Commun*. **2013**, *4*, 1874-1881; g) G. A. Woolley, A. S. I. Jaikaran, M. Berezovski, J. P. Calarco, S. N. Krylov, O. S. Smart, J. R. Kumita, *Biochemistry* **2006,** *45,* 6075-6084; h) G. A. Bullen, J. H. Tucker, A. F. Peacock, *Chem. Commun.* **2015,** *51,* 8130-8133; i) E. Oheix, A. F. A. Peacock, *Chem Eur J* **2014,** *20,* 2829-2839; j) M. I. Sánchez, O. Vázquez, M. E. Vázquez, J. L. Mascareñas, *Chem. Commun*. **2011**, *47*, 11107-11109.

Entry for the Table of Contents

COMMUNICATION

Designed peptide modules self-assemble at extended 12 base-pair DNA sites in the presence of Ni(II) salts.



J. Rodríguez, J. Mosquera, M. E. Vázquez, J. L. Mascareñas**

Page No. – Page No.

Nickel-promoted Recognition of long DNA sites by designed Peptide Derivatives