

This is the peer reviewed version of the following article: Boga, S., Bouzada, D., García Peña, D., Vázquez López, M. and Vázquez, M. E. (2018), Sequence-Specific DNA Recognition with Designed Peptides. *Eur. J. Org. Chem.*, 2018: 249-261, which has been published in final form at <https://doi.org/10.1002/ejoc.201700988>.

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions

Sequence-Specific DNA Recognition with Designed Peptides

Sonia Boga,^[a] David Bouzada,^[a] Diego García Peña,^[a] Miguel Vázquez López,^[b] M. Eugenio Vázquez^{*[a]}

Dedicated to Prof. José L. Mascareñas

Abstract: Inspired by natural TFs, researchers have explored the potential of artificial peptides for the recognition of specific DNA sequences, developing increasingly sophisticated systems that not only display excellent DNA binding properties, but also are endowed with new properties not found in their natural counterparts. Herein we review some of the developments in the field of artificial peptide-based DNA binders, focusing on the supramolecular and molecular design aspects of such systems.

1. Introduction

Gene transcription in eukaryotic cells is largely controlled by the interaction of certain proteins called transcription factors (TFs) with specific DNA sequences,¹ which modify the basal levels of transcription, either enhancing or repressing their expression.² It is known that misregulation of TFs is at the origin of a number of diseases, including cancer,³ and thus there is growing interest in understanding the molecular basis of specific DNA recognition, as well as in developing designed DNA binding agents. Furthermore, in addition to the potential applications of such synthetic binders, the development of artificial DNA binding agents represents an unmet challenge at the crossroads between supramolecular and biological chemistry and, beyond its biological relevance, it might find future applications in combination with DNA nanotechnology.⁴ Most artificial DNA binding agents described to date are small molecules, including coordination compounds, as well as small organic molecules, which typically bind to the DNA through minor groove insertion or intercalation mechanisms, and have been the subject of a number of excellent reviews.^{5,6} In addition to those classic DNA-binding agents, researchers, inspired by natural TFs, have also explored the potential of designed peptides for the specific recognition of DNA, developing increasingly sophisticated systems that not only display excellent binding properties, but also are endowed with new properties not found in their natural counterparts, such as luminescence, photocontrol, nuclease activity, or stimuli-responsive binding, which we will review in the following pages.⁷

1.1. DNA recognition from a supramolecular perspective

The B-DNA, the most relevant DNA conformation under physiological conditions, is characterized by a rather uniform right-handed double helix formed by two antiparallel oligodeoxyribonucleotide chains, held together by the combination of stacking interactions between the base pairs (bps), and hydrogen bonds between complementary Watson-Crick bps.⁸ The asymmetry of the deoxyribonucleotides generates two grooves in B-DNA double helix: the *major groove*, which is the wide side (~11.7 Å between phosphates across the groove) facing away the sugar-phosphate backbone, and the *minor groove*, which is the narrow side (~5.7 Å) of the bps facing towards the backbone (Figure 1). Importantly, the B-DNA conformation exposes the polar sugar/phosphate backbone while shielding the aromatic surface of the bases from the aqueous environment, thus giving rise to a monotonous physicochemical landscape, where particular base pairs can only be distinguished by relatively minor differences in their hydrogen bond donor/acceptor patterns at the bottom of the grooves. Thus, for example, A•T/T•A, and G•C/C•G bps are degenerate from the minor groove (Figure 1b). This configuration increases the challenge for the development of specific binders that can distinguish the different base pairs against the uniform B-DNA backdrop, and supports the use of peptide systems that can replicate the natural TFs.

Despite the apparent regularity in the B-DNA, it is now recognized that the local properties of the B-DNA double helix are highly dependent on the base pair sequence, which can be considered not only a linear code, but also a structural code that influences both the conformation and molecular dynamics.⁹ For example, A•T-rich tracts are known to induce bending of the DNA double helix, which is required for the activity of some regulatory sites.¹⁰ A•T-rich sequences also display a narrower minor groove,¹¹ with a distinct hydration pattern along the floor of the groove,¹² and is also characterized by a much more negative electrostatic potential,¹³ which can influence both reactivity,¹⁴ and specific recognition of such sequences.¹⁵

Therefore, specific DNA recognition involves two types of interactions: *direct readout* through the formation of specific hydrogen bonds, polar interactions, and hydrophobic contacts between amino acids side chains and nucleotide functional groups—typically at the exposed edges of the bps in the DNA major groove,¹⁶ and *indirect readout*, which is a more subtle mechanism related to the unique physicochemical properties of specific DNA sequences, such as readout of the stiffness of DNA double helix, its conformational flexibility, our particular electrostatic potential.¹⁷ These effects can be even observed in the DNA binding preferences of small molecules and, for example, it has been found that simple intercalators display some sequence

[a] S. Boga, D. Bouzada, Dr. D. García Peña, Dr. M. E. Vázquez. 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: eugenio.vazquez@usc.es; @ChemBioUSC

[b] Dr. M. Vázquez López. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Departamento de Química Inorgánica. Universidade de Santiago de Compostela 15782 Santiago de Compostela, Spain.

selectivity due to the local deformability and conformational preferences of particular DNA sequences.¹⁸ Evidently, natural DNA-binding proteins exploit both strategies to achieve high affinity and selectivity. Unfortunately, artificial systems are still much more rudimentary, and rational implementation of indirect readout in designed systems is still extremely complex.

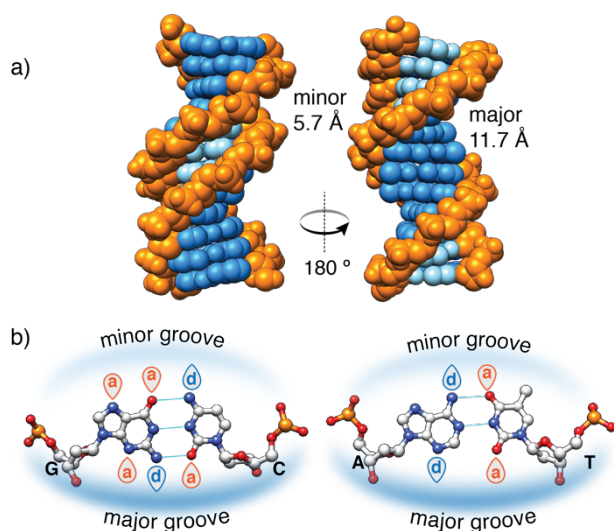


Figure 1. a) Classic Dickerson-Drew B-DNA dodecamer (PDB: 4C64)¹⁹ highlighting the major structural elements: sugar-phosphate backbone in orange, exposed bp edges in the major groove in dark blue, and exposed edges of the bps in the minor groove in light blue. b) structure of the G-C and A-T bps showing the pattern of hydrogen bond donors (d) and acceptors (a).

Recognition of specific DNA sequences in eukaryotes not only relies on the readout mechanisms that have been just outlined that mediate the interactions between the DNA and each protein, but is also orchestrated through specific protein-protein interactions that lead to the formation of higher order assemblies of multiple TFs that cooperatively bind to the DNA. This allows the integration of multiple signaling pathways into complex networks of TFs,²⁰ and also expands the number of unique DNA binding sites that can be addressed with a given number of TFs, which can combine in different ways to bind to composite sequences, thus, allowing the complex spatio-temporal control of the >30,000 human genes with a limited set of 2,000-3,000 TFs.²¹ Despite the relevance of TF cooperativity, the complexities involved in the implementation of such effects have hampered the design of artificial systems that display some level of cooperativity in their DNA binding.

2. DNA-binding peptides

2.1. GCN4, the Jack of all trades, master of designed DNA-binding peptides

Transcription factors are grouped in families according to the fold of their DNA binding domains.²² The GCN4 is an archetypical member of the Basic Leucine Zipper (bZIP) family of TFs. As all bZIP TFs, GCN4 binds to its target ATF/CREB (5'-ATGACGTCAT-3') or AP1/GCRE (5'-ATGA(c/g)TCAT-3') sites as a leucine zipper-mediated dimer of uninterrupted α -helices. The specific contacts with the edges of the bases exposed in the major groove of the DNA take place through the N-terminal basic region (br), which folds into an α -helix upon DNA binding (Figure 2).²³ Importantly, it has been shown that monomeric bZIP TFs typically display low DNA binding affinity in the μ M range,²⁴ and that dimerization is required for DNA binding by bZIP proteins, because of the high entropic cost involved in the folding of the basic region into the α -helical conformation.²⁵ Thus bZIP TFs behave as intrinsically disordered proteins that are largely unstructured under physiological conditions and display folding coupled DNA binding. This allows bZIP proteins to recognize their target DNA sites with exquisite selectivity while still forming relatively weak, and readily reversible complexes—in contrast with rigid molecules, in which stability of their complexes usually correlates with their specificity—and allows increased rates of macromolecular associations within complex interaction networks.²⁶

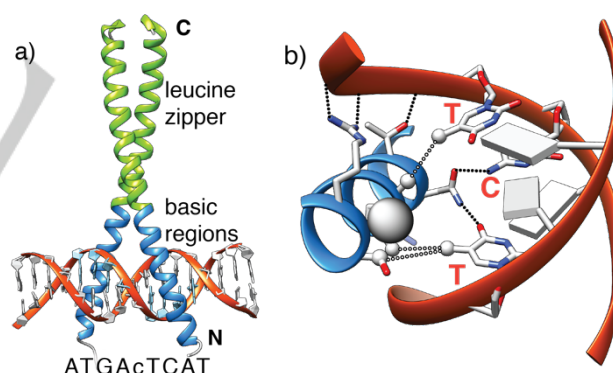


Figure 2. a) Structure of the DNA binding domain of the GCN4 TF bound to the AP1/GCRE site, highlighting the main structural elements and the key residues for specific DNA recognition. The leucine zipper in light green, and the br in blue. b) Details of the specific interactions between the GCN4 br with the AP1/GCRE half-site. Hydrophobic contacts (Ala²³⁸, Ala²³⁹, and Ser²⁴²) with white dot lines, polar/electrostatic contacts (Asn²³⁵, and Arg²⁴³) with black dot lines. Bases that do not participate in specific contacts are outlined as slabs for clarity.

2.2. Artificial GCN4 peptide dimers

The structural simplicity of bZIP TFs, added to the wealth of functional, biophysical and structural information about GCN4, made it into the preferred platform for the development of peptide-based DNA binders. A key development was reported in 1990 by

MICROREVIEW

the group of Peter S. Kim, who demonstrated in a seminal paper that the complete leucine zipper region of the GCN4 DNA binding domain could be replaced by a disulfide bond, and that the minimized 34-residue peptide corresponding largely to the GCN4 between residues 222 and 253 (plus a short Gly-Gly-Cys linker) could bind to the natural target site CREB with high (nM) affinity at 4 °C in the form of a disulfide dimer (Figure 3a). As expected, the reduced monomeric peptide did not display significant binding (Figure 3a).²⁷ Additional studies reported two years later helped to identify the sequence of the minimal GCN4 basic region fragment that displays specific DNA binding in the form of disulfide dimer. Thus, a 23-residue peptide featuring a C-terminal Gly-Gly-Cys linker (Ac-DPAALKRARNTAARRSRARKLQ-GGC) binds to both ATF/CREB and AP1/GCRE sites in a sequence-specific manner with high affinity in its oxidized form as a dimer.²⁸

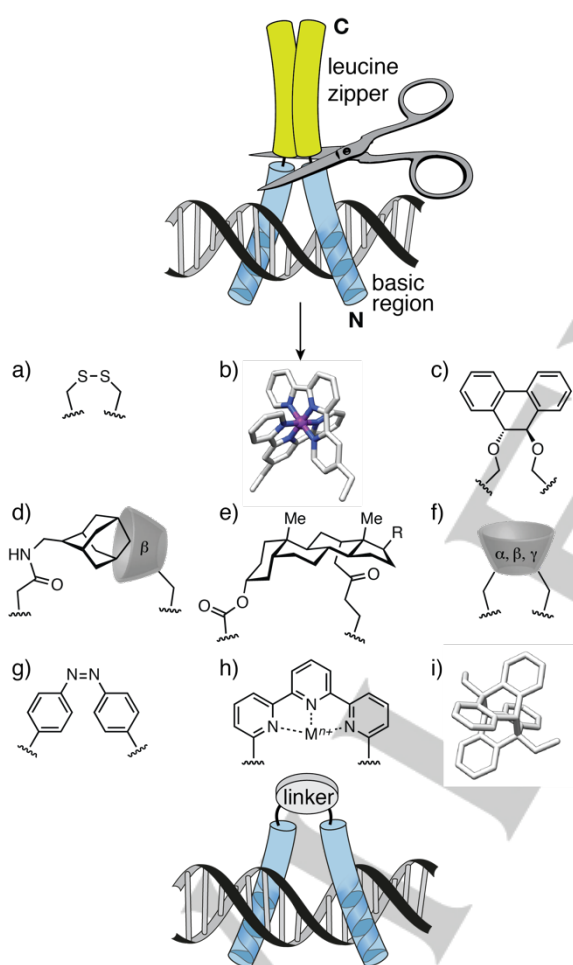


Figure 3. Schematic representation of the GCN4 dimer showing the helices as tubes, and the strategy to generate artificial GCN4 dimers by removing the leucine zipper and adding specific linkers between the br fragments. a) P. Kim's disulfide dimer,^{27,28} b) A. Schepartz's transition metal coordination compound linker,²⁹ c) T. Morii's chiral linker derived from trans-9,10-dihydrophenanthrene-9,10-diol,³¹ d) T. Morii's supramolecular linker based on a β -cyclodextrin/adamantane complex,^{33,34} e) A. Madder's steroid linker,^{36,37} f) A. Madder's cyclodextrin dimer,³⁵ g) J. L. Mascareña's cis/trans azobenzene photoswitch linker,⁴⁰ h) A. Peacock's metal switch linker,³⁰ i) A. Peacock's photocontrolled anthracene dimer.⁵³

Following Peter Kim's work, a number of other groups have reported over the years the replacement of the leucine zipper with other functionalities. The group of Alanna Schepartz reported in 1993 the use of a set of transition metal complexes as linkers to systematically explore the effect of the geometry of the linker—and thus, of the relative geometry of the attached basic regions—on the DNA binding properties of the resulting dimers (Figure 3b). Unsurprisingly, they found that dimerization was not sufficient for DNA binding, and that the geometry of the metal complexes influenced not only the binding affinity of the dimers, but also the sequence selectivity.²⁹ The effect of the geometry in metal complexes was also exploited by the group of A. Peacock, who in 2014 reported a dimer of the GCN4 basic region with a 2,2':6',2''-terpyridine linker. This linker would undergo a conformational change upon metal ion coordination, which in turn would affect the relative orientation of the basic regions, and influence the DNA binding of the dimer. CD and UV-vis spectroscopy demonstrated that coordination of this metal-chelating GCN4 dimer to Cu(II) or Zn(II) ions promoted DNA binding (Figure 3h).³⁰

The influence of geometric effects on the DNA binding properties of GCN4 dimers were also studied by T. Morii and Y. Sugiura who, based on earlier studies using a 14-residue DNA-binding peptide derived from the basic region of the helix-loop-helix transcription factor MyoD,³¹ synthesized various GCN4 dimers connected by C_2 -symmetric, trans-9,10-dihydro phenanthrene-9,10-diol linkers with different chiralities (Figure 3c). In this case, the authors found that the stability of the different complexes was unaffected by the chirality of the linkers, which was consistent with the small effect of the chirality on the relative orientation of the attached basic regions. Interestingly, the authors also reported dimers of the GCN4 br through the N-terminus, which recognized altered palindromic DNA sequences (5'-TCATC GATGA-3') in which the polarity of each half-site of the parent ATF/CREB site was reversed.³²

The group of T. Morii was the first to report the use of a non-covalent linker between GCN4 basic regions; for this, they synthesized two peptides by alkylation of a C-terminal Cys attached to the core GCN4 br peptide with *N*-(bromoacetyl)-1-adamantanemethylamine and mono-6-deoxy-6-iodo- β -cyclodextrin. Both peptides form a heterodimer, mediated by formation of an inclusion complex between β -cyclodextrin, and an adamantyl group, which specifically binds to the ATF/CREB (5'-ATGAC GTCAT-3') site. The formation of the complex was inhibited by the addition of free β -cyclodextrin, once again demonstrating that dimerization of the GCN4 basic regions is required for specific DNA binding (Figure 3d).³³ Interestingly, their design relies on the high stability of the complex between adamantane derivatives and β -cyclodextrins ($K_D \approx 1.3 \mu\text{M}$), and subsequent studies showed that efficient and cooperative DNA binding was still possible if the original adamantane was replaced with norbornyl, $K_D(\text{Nb}/\beta\text{-CD}) \approx 3.1 \mu\text{M}$, or noradamantyl groups, $K_D(\text{NrA}/\beta\text{-CD}) \approx 2.7 \mu\text{M}$, but no DNA recognition was observed with the cyclohexyl group, which displays a much lower binding affinity for β -cyclodextrin with a $K_D(\text{Ch}/\beta\text{-CD}) \approx 14 \mu\text{M}$.³⁴

A recent paper by A. Madder also makes use of (α,β,γ)-cyclodextrins for dimerization (Figure 3f). In their case, the

MICROREVIEW

cyclodextrins were not used as receptors, but as scaffolds for covalent attachment of the two GCN4 basic regions. The dimers were readily synthesized by a straightforward copper-catalyzed azide/alkyne cycloaddition (CuAAC) between the GCN4 br peptide bearing a C-terminal propargylglycine, and the corresponding diazido cyclodextrins. Curiously, while the α -cyclodextrin and β -cyclodextrin GCN4 br dimers display rather similar binding affinities for the ATF/CREB site with $K_D(\alpha\text{-CD/DNA}) = 50 \pm 20$ nM and $K_D(\beta\text{-CD/DNA}) = 30 \pm 20$ nM, the γ -cyclodextrin analog binds with weaker affinity to the same DNA with a $K_D(\gamma\text{-CD/DNA}) = 100 \pm 60$ nM, and appears to give rise to non-specific complexes in Electrophoretic Mobility Shift Assays (EMSA), suggesting that the distance between the two GCN4 basic regions in the γ -cyclodextrin dimer is not optimal for their simultaneous insertion as required in the specific complex.³⁵ Madder's group has also reported the application of steroid scaffolds for the homodimerization of GCN4 basic regions,³⁶ as well as for the heterodimerization of the basic regions of related c-Myc/Max basic Helix-Loop-Helix transcription factors, a class of TFs related to the bZIP family, containing an additional loop between the leucine zipper and the basic region.³⁷ The use of a steroidal cholic acid moiety as dimerizer element was justified because it provides a rigid, and synthetically-accessible scaffold to attach the peptides with reported benefits for improving the cell uptake of the conjugates (Figure 3e).

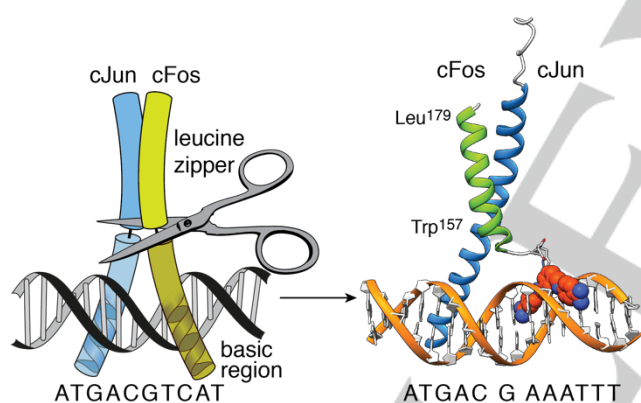


Figure 4. Left: Schematic representation of the cFos/cJun/DNA complex. Right: proposed interaction between the cFos-bisbenzamidinium hybrid and cJun, binding as a heterodimer to a composite DNA site containing the consensus recognition sites for cJun and the bisbenzamidinium. In contrast with the natural complex, the cFos Leucine Zipper is now delivered from the bisbenzamidinium minor groove anchor.

In addition to bZIP dimers that replace the leucine zipper dimerization with other functionalities, in 2012 J. L. Mascareñas reported a derivative of an heterodimeric cFos/cJun complex in which the basic region of the cFos bZIP TF was replaced by a small DNA-binding organic molecule.³⁸ Similarly to the role of cFos in the natural complex, the bisbenzamidinium hybrid of the Fos leucine zipper (featuring residues from Arg¹⁵⁷ to Leu¹⁷⁹) stabilized the cJun/DNA interaction but, in contrast with the natural cFos TF, which is bound to the DNA major groove, the cFos leucine zipper domain was delivered from the adjacent DNA

minor groove. Furthermore, the dimer between cJun and the cFos-bisbenzamidinium conjugate targeted a the composite DNA sequence different from the ATF/CREB binding site targeted by the natural cJun/cFos heterodimer (5'-ATGAC G AAATTT-3'), which contained the natural ATF/CREB half-site (ATGAC) next to the A·T-rich site preferred by the bisbenzamidinium with an extra G as spacer between both recognition elements (Figure 4).

2.3. Stimuli-responsive DNA-binding peptides

2.3.1. Light-responsive systems

Azobenzenes are well-known photoswitches that can interconvert between their two isomers: the energetically more stable (about 10–12 kcal mol⁻¹) *trans* (*E*) isomer, and the *cis* (*Z*) azobenzene—by irradiation with light of appropriate wavelength (Figure 5a). Thus, irradiation of the $\pi \rightarrow \pi^*$ band of the *trans* state at ~340 nm ($h\nu_1$) leads to the population of the metastable *cis* isomer through radiationless decay and ultrafast rotation about the N=N bond (Figure 5b). Likewise, the *cis* isomer be interconverted into the more stable *trans* isomer by irradiation of the $n \rightarrow \pi^*$ at about 450 nm ($h\nu_2$), as well as thermally. Importantly, most azobenzenes display a strong overlap in the absorption spectra of their *cis* and *trans* isomers, which typically leads to incomplete photostationary equilibria with about 80% *cis* or ~95% *trans* states upon irradiation at the $\pi \rightarrow \pi^*$ band or $n \rightarrow \pi^*$ absorption bands, respectively. On the other hand, thermal *cis*→*trans* equilibration leads to virtually complete formation of the *trans* isomer (Figure 5c).³⁹

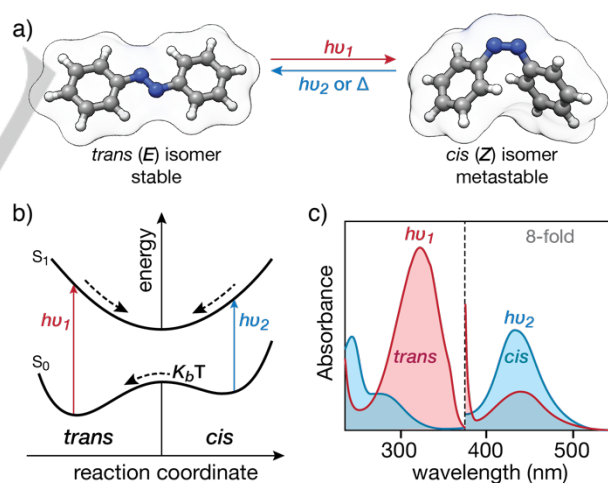


Figure 5. a) overview of the *trans/cis* isomerization of azobenzene showing the geometric changes associated with the interconversion between both states; b) The potential energy surface of the azobenzene ground state is characterized by an asymmetric double well that allows the interconversion of the *trans*-azobenzene isomer into the metastable *cis*-isomer in the ground state by irradiation with UV light ($h\nu_1$). The low energy barrier between the *cis* and the isomer allows its thermal conversion into the *trans* state. Thus, the *trans*→*cis* interconversion can only be induced by irradiation, whereas the *cis*→*trans* isomerization can be induced either by irradiation at longer wavelength ($h\nu_2$), or through a thermal transition; c) Absorption spectra of *cis* and *trans* azobenzene; the region of the $n \rightarrow \pi^*$ band at longer wavelength is zoomed 8-fold for clarity. Non-radiative processes in b are shown with dashed arrows.

MICROREVIEW

A relevant subset of GCN4 dimers are those whose DNA binding can be modulated by light. In the year 2000 the group of J. L. Mascareñas reported the first artificial DNA binding peptide that displayed light-induced DNA binding.⁴⁰ Light responsiveness was achieved by incorporating an azobenzene unit in the linker between two GCN4 basic region peptides. The design takes advantage of the large geometrical change experienced by the azobenzene upon *trans*→*cis* isomerization, so that while the *trans* configuration forces the two GCN4 br away from each other, resulting in an unfavorable geometry for simultaneous interaction of both peptides in their target sites, the *cis*- conformation presents both basic regions in the appropriate orientation for both peptides to interact in their respective half-sites at the same time (Figure 3g, and Figure 6). EMSA assays demonstrated that the *cis* isomer binds to the target ATF/CREB site with 60–70 times more affinity than the *trans* isomer. Curiously, the *trans*→*cis* isomerization can be carried out in the presence of the DNA, but in this case the isomerization is about eight times slower than in the absence of DNA, and leads to a slightly lower proportion of the *cis* isomer in the photostationary state.

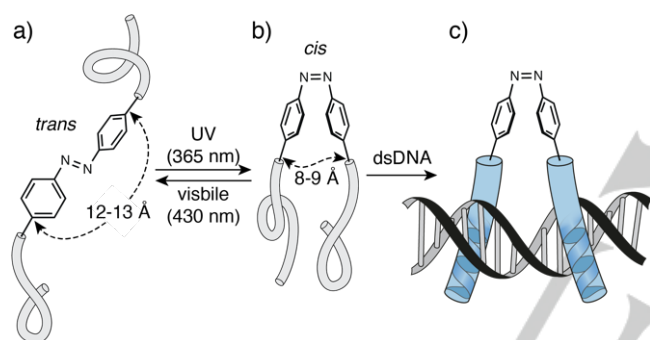


Figure 6. The DNA binding properties of GCN4 dimers featuring an azobenzene linker can be modulated by light. Thus, the opposite orientation, and large separation between the peptides in the *trans* azobenzene isomer does not support high-affinity DNA binding; a) irradiation with UV light induces the isomerization of the azobenzene unit, and b) the resulting *cis* GCN4 dimer presents the right geometry for high-affinity binding as shown in c). Unfolded peptides are represented with light-grey thin tubes.

Before presenting other examples of light-modulated systems we must introduce another of the preferred model systems in artificial DNA binding peptides, the Homeodomain (HD) proteins. HDs are highly conserved TFs that play key roles in eukaryotic development. The HD DNA binding domain is folded into a helix-turn-helix motif, with an extended minor groove-binding N-terminal arm.⁴¹ Most of the specific contacts with the DNA are made by helix 3 of the protein (h_3), which, in the complex with DNA, is inserted into the DNA major groove. As in the case of the GCN4 and bZIP proteins, the short DNA-contacting helix h_3 isolated from the rest of the protein fails to fold into the α -helical conformation and shows negligible DNA binding affinity. The engrailed HD Q50K is an altered-specificity Gln⁵⁰→Lys variant of the engrailed HD that preferentially binds to the QRE site 5'-TAATCC-3' (while the wild-type engrailed TF binds to the ERE sequence, 5'-TAATTA-3'). Furthermore, the engrailed Q50K

displays higher affinity for the QRE site ($K_D \approx 8.8$ pM) than the wild-type engrailed for its ERE target sequence ($K_D \approx 79$ pM).⁴²

After the initial work with azobenzene GCN4 dimers by J. L. Mascareñas, the group of G. A. Woolley reported in the year 2005 an alternative strategy for obtaining light-controlled DNA-binding peptides incorporating azobenzene units within DNA basic region (Figure 7).⁴³ Following their work using azobenzenes for photocontrolling the α -helical content in short peptides, their first design consisted on a 18-residue peptide fragment from the recognition helix h_3 of the Q50K engrailed homeodomain (Ac-EAQCKIAAKNARAKCKKA, mutated residues in bold typeface), in which residues Ile⁴⁵ and Ile⁵⁶ (at i , and $i+11$ positions) on the outer face of the h_3 α -helix, were mutated to Cys for introducing the 3,3'-bis(sulfo)-4,4'-bis(chloroacetamido)azobenzene cross-linker. Additionally, a number of residues that do not make direct contacts with the DNA were also replaced with Ala in order to increase the helical propensity of the peptide, and to avoid undesired steric clashes between bulky side chains with the azobenzene staple. The thermodynamically-stable *trans* azobenzene displayed significant α -helical content and formed stable complexes ($K_D \approx 7.5$ nM) with the QRE target sequence (5'-TAATCC-3'). Irradiation with UV light (360 nm) induced the *trans*→*cis* isomerization of the azobenzene, which significantly destabilized the α -helical conformation of h_3 , and resulted in decreased affinity for the DNA ($K_D \approx 140$ nM).

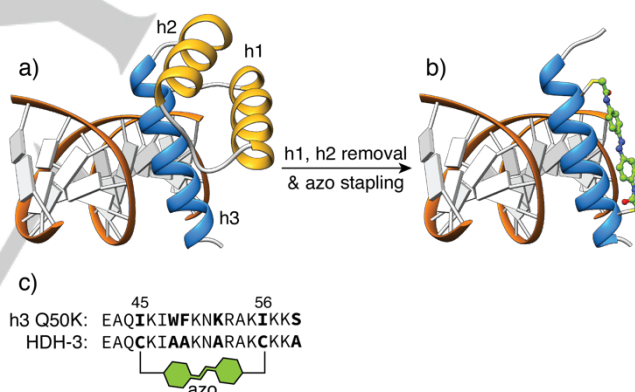


Figure 7. Miniaturized HD Q50K light-modulated DNA-binding peptide. a) Structure of the Q50K engrailed HD highlighting the DNA contacting helix (h_3) in blue, and the accessory helices h_1 and h_2 in yellow (PDB: 2HDD). b) The light-responsive peptide contains most of the HD Q50K sequence, and two Cys residues at positions 45 and 56 (i , $i+11$) for stapling with the azobenzene unit. The *trans* azobenzene stabilizes the DNA-binding α -helical conformation, while irradiation yields a low-affinity DNA configuration.

In the following years the group of G. A. Woolley refined their designs,⁴⁴ and reported modified azobenzene switches with improved optical properties for *in vivo* applications,⁴⁵ and in 2011 they demonstrated for the first time the application of artificial peptides for the photocontrol of protein expression in living cells.⁴⁶ The authors focused their attention on the oncoproteins Fos and Jun, which are also members of the bZIP family of TFs that bind to the AP1/GCRE site as heterodimeric complexes. Their design was a modified version of a dominant-negative (DN) peptide

MICROREVIEW

reported by C. Vinson,⁴⁷ which contained an optimized Fos leucine zipper dimerization domain for increased affinity for the Jun protein,⁴⁸ and replaced the normal basic region critical for DNA binding with an acidic extension that also stabilized its complex with Jun, while at the same time prevented the resulting heterodimer from binding to the DNA. This DN peptide (AFosW) was modified with a 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene cross-link between two Cys residues replacing the Asp¹⁷⁴ and Thr¹⁸¹ residues in the wild-type Fos leucine zipper sequence. The resulting photoisomerizable peptide (XAFosW) formed stable heterodimers with Jun when the azo group was in the *trans* configuration, and isomerization of the azobenzene crosslink to the *cis* configuration by irradiation with light at 365 nm reduced the stability of the XAFosW/Jun complex by about 10-fold. The *in vivo* effect of the XAFosW photoswitch was quantitatively studied by measuring the ratios of the activity between a luciferase reporter under the control of an AP1/GCRE promoter and a β -galactosidase reporter under the control of a constitutive rous sarcoma virus (RSV) promoter (pRSV-Gal). Thus, XAFosW showed concentration-dependent inhibition of luciferase activity and, more importantly, a decrease of as much as 40% in AP-1 activity upon photo-irradiation. Subsequent work has focused on fine tuning these designs, for example by identifying the optimal position for introducing crosslinks,⁴⁹ as well as on developing encodable protein chimeras with photoactive yellow protein.⁵⁰

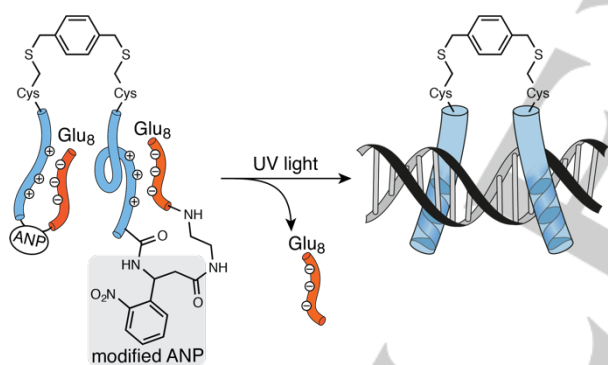


Figure 8. Strategy for electrostatic caging of the DNA binding domain of the GCN4 transcription factor. Oligoglutamic (Glu_8) domains were attached to the N-terminus of the GCN4 br through an orthogonal, photolabile ANP linker, and after irradiation, the photoreleased br dimer binds to the DNA.

In 2012 our own group reported a different approach to obtaining light-activated DNA binding peptides.⁵¹ Considering that the formation of thermodynamically stable GCN4/DNA complexes relies on non-specific electrostatic interactions between positively-charged basic residues in the GCN4 basic region with the negatively-charged DNA phosphate backbone,⁵² we envisaged that introducing repulsive interactions with short polyanionic tails would suppress the DNA interaction. Furthermore, if the polyanionic tails were attached through a photolabile linker, then the interaction between the photoreleased GCN4 basic region and the target ATF/CREB DNA could be

restored by irradiation with light. Thus, P. S. Kim's GCN4 basic region fragment was synthesized with an acidic oligo-glutamic domain (Glu_8) connected through a photolabile 3-amino-3-(2-nitrophenyl)-propionic acid (ANP) group. The resulting oligozwitterionic basic region containing a short linker with a C-terminal Cys residue (Glu_8 -(ANP)-DPAALKRARNTAARRSRAR KLQ-GGC) was dimerized by direct alkylation with 1,4-bis(bromomethyl)benzene, to yield the desired photoactivatable GCN4 derivative (Figure 8). As expected, the electrostatically-compromised dimer did not display significant DNA binding affinity for its target site, and irradiation with UV light irreversibly restored the DNA binding of the photoreleased GCN4 br dimer, as demonstrated by the appearance of a slow-migrating band in gel shift experiments. Circular dichroism assays also confirmed the folding of the photoreleased basic region in the presence of a short oligonucleotide containing the ATF/CREB target sequence.

An interesting interplay between the DNA binding of GCN4 peptides and anthracene photodimerization was described by the group of A. Peacock in 2016, when they studied monomeric GCN4 basic region peptides modified with anthracene units on their C-terminus.⁵³ While anthracene intercalation in the presence of non-target DNA sequence proved problematic, circular dichroism experiments showed a significant increase in the alpha helical content of the peptide upon irradiation in the presence of the target ATF/CREB (5'-ATGAC GTCAT-3') site. Furthermore, they found that anthracene photodimerization took place only in the presence of such target site, while dimerization of the anthracene-basic region peptide was not observed when irradiation was made in the presence of half-site or random DNAs. As expected from the previous work with GCN4 dimers, the resulting photocrosslinked dimer displayed stronger DNA binding than the monomeric peptides.

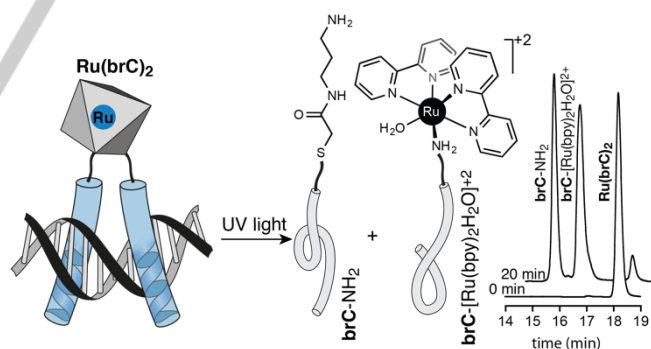


Figure 9. Photocleavable (2,2'-bipyridine)Ru(II) linkers in bZIP basic region dimers (either GCN4 homodimers or C/EBP-GCN4 heterodimers). The high-affinity, specific DNA complex can be disassembled by irradiation with long-wavelength light. The different species resulting from the asymmetric photolysis can be identified by HPLC-MS (right).

One of the major problems associated with the application of photocontrolled peptides *in vivo* is the short wavelengths at which the photoswitches and caging groups respond, typically below 400 nm, which have poor tissue penetration, and are readily absorbed by other biomolecules. In this context, photolabile

MICROREVIEW

bisbipyridyl ruthenium(II) complexes have been explored as alternative photoactive components due to their advantageous spectroscopic properties, such as long wavelength of photolysis and high uncaging quantum yields. Thus, we reported in 2014 the use of ruthenium bipyridyl complexes as photocleavable dimerizers of cysteine-containing transcription factor fragments. More specifically, we reported the synthesis of a generally applicable *cis*-bromoalkyl-(2,2'-bipyridine)Ru(II) biselectrophilic reagent, which was derivatized to obtain homodimeric GCN4 br dimers, similar to those previously described in the 90s by A. Schepartz, as well as C/EBP-GCN4 heterodimers. In both cases the dimeric constructs displayed good DNA binding affinity and selectivity for their respective target site. Moreover, their DNA binding could be suppressed by irradiation at 455 nm with a LED source and photolysis of the ruthenium dimerizer (Figure 9).⁵⁴

Ruthenium compounds have been used as photolabile protecting groups to control the DNA binding of small peptides.⁵⁵ In 2015 we reported the synthesis of a light-sensitive histidine building block for Fmoc/tBu solid-phase peptide synthesis in which the imidazole side chain is protected with a ruthenium complex. This building block was used for the synthesis of a photoactivatable Arg-Gly-His tripeptide (RGH), which has been shown to be an efficient metal-chelating sequence ($K_D < 10^{-16}$ M), endowed with DNA binding and endonuclease properties in the presence of Ni(II) ions and oxidizing agents.⁵⁶ We demonstrated that the caged His analog was unable to coordinate the Ni(II) required for the catalytic metalloprotein. Moreover, we also showed that DNA binding and nuclease activity could be recovered upon irradiation at long wavelength (455 nm LED), and uncaging of the His residue (Figure 10).

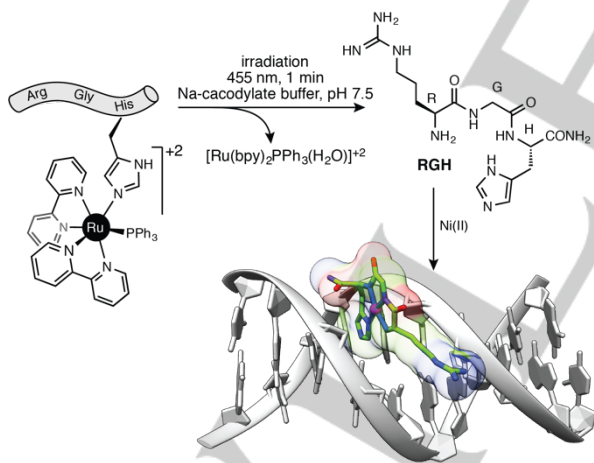


Figure 10. caged RGH peptide does not bind to Ni(II). Irradiation with red light at 455 releases the metal-chelating peptide, which folds and binds into the DNA minor groove, where it can exert its nuclease activity.

2.3.2. Metal coordination and dynamic DNA-binding

In contrast with the static designs—i.e. insensitive to external inputs—that represent the majority of DNA-binding peptides reported until now, one of the first steps into the development of stimuli-responsive peptides derived from transcription factors was reported by Y. Sugiura in 2004.⁵⁷ In this work, the Fos leucine zipper is modified with a pair of metal-chelating iminodiacetic acid derivatives of lysine (Ida) at positions *i*, and *i*+2. Incubation with Fe(III), and coordination of this metal ion, destabilizes the alpha helical conformation of the Fos(Ida)₂ peptide, thus preventing the formation of the heterodimer with the Jun leucine zipper. Unfortunately, the authors did not pursue this concept for modulating the DNA binding of the Fos-Jun network of homo- and heterodimers. In a related work, a successful DNA binding switch in response to a metal ion was reported by the group of Futaki in 2009, who expressed a recombinant GCN4 DNA binding domain (56-mer peptide) featuring two pairs of Cys residues at relative *i* and *i*+2 positions in the leucine zipper domain. These residues were selectively derivatized using *N*-(2-tosylthioethyl)iminodiacetic acid to turn them into metal-chelating residues. The resulting tetra-Ida-modified GCN4 DNA binding domain displayed significant helical structure, and high DNA binding affinity for the AP1/GCRE site ($K_D = 22 \pm 3.0$ nM) similar to that of the natural GCN4 dimer ($K_D = 15 \pm 2.6$ nM). Addition of excess of Co(II) to this metal-binding peptide induced a 33% decrease in its helical content, and also a drastic reduction in its DNA binding affinity ($K_D > 1.0$ μM). As expected, addition of EDTA to the mixture—and sequestering of the Co(II) ions—resulted in the recovery of the initial binding affinity in absence of the metal ion ($K_D = 22 \pm 1.1$ nM). Curiously, the switching effect was not observed in control peptides that contained a single pair of Ida residues in their leucine zipper domains.⁵⁸

The effect to metal ion coordination can be combined with other orthogonal signals, such as redox processes to yield more complex designs that respond to multiple stimuli. This was exemplified in 2013, when we described the synthesis of a GCN4 basic region fragment that contained a redox-sensitive Cys residue on its N-terminus, and a metal-chelating 5,5''-dimethyl-2,2':6',2''-terpyridine ligand orthogonally attached to a C-terminal Lys²³¹ residue (Cys)br(tpy).⁵⁹ Incubation of this peptide with Ni(ClO₄)₂ in the presence of the target ATF/CREB site (5'-ATGA CG TCAT-3') gave rise to slow-migrating bands consistent with the formation of a specific peptide-DNA complex between the nickel-mediated dimer of the basic region and the DNA. Under the same conditions, the peptide did not display significant affinity for the inverted sequence (5'-TCAT CG ATGA-3'). In contrast, the N-terminal disulfide dimer formed by oxidation of the peptide with DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid), only showed binding to the inverted site (Figure 11). As expected, reduction of the disulfide dimer by treatment TCEP (tris(2-carboxyethyl) phosphine) recovered the monomeric (Cys)br(tpy) peptide, and promoted the disassembly of the complex with the inverted sequence. Curiously, the disulfide dimer required the presence of Ni(II) ions to bind to its target inverted site.

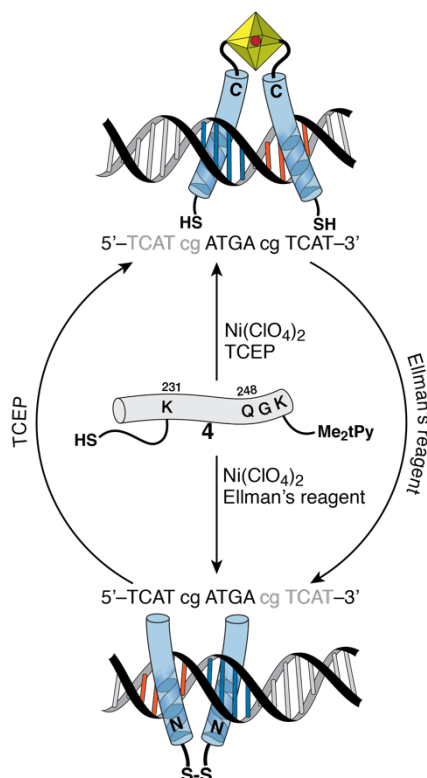


Figure 11. Dynamic DNA binding selection to multiple sites. The reduced monomeric (Cys)br(tpy) peptide in the presence of $\text{Ni}(\text{ClO}_4)_2$ binds to the consensus DNA (5'-ATGA CG TCAT-3') as a metal-mediated homodimeric complex ($K_D \approx 670$ nM at 4°C). Oxidation of the N-terminal Cys side chain with DTNB yields the N-terminal disulfide dimer, which binds to the inverted polarity sequence 5'-TCAT CG ATGA-3' in the presence of $\text{Ni}(\text{II})$ salts ($K_D \approx 758$ nM). The original C-terminal complex can be recovered by reduction with TCEP (tris(2-carboxyethyl)phosphine). In both cases the binding of the peptides for their non-target DNAs were in the μM range.

De novo designed DNA-binding metallopeptides are relatively uncommon, and typically contain an octahedral metallointercalator complex that binds in the DNA major groove attached to a short peptide that modulates its DNA recognition or internalization properties. Prominent examples of this general design have been reported by J. Barton since the 90s, when they described conjugates between short 13-mer peptides and photoactive $[\text{Rh}(\text{phi})_2(\text{phen}')]^{3+}$ ($\text{phi} = 9,10$ -phenanthrenequinone diimine; $\text{phen}' = 5$ -(amidoglutaryl)-1,10-phenanthroline) complexes. The coordination compounds intercalate through the DNA major groove with high affinity, and upon photoactivation, induce DNA strand scission. The DNA site specificity was dependent on the sequence of the appended peptide, and the authors found that a single glutamate at position was essential in directing DNA site-recognition to the sequence 5'-CCA-3'.⁶⁰ Later examples demonstrated the general applicability of this strategy by inserting different peptides with alternative sequence preferences, such as the DNA recognition helix of the phage 434 repressor (preferentially targeting the sequence 5'-ACAA-3').⁶¹ Related designs featured α -helical metallopeptides with nuclease activity tethered to the $[\text{Rh}(\text{phi})_2(\text{phen}')]^{3+}$ complex. Thus, the

peptide sequence DPDELEHAAKHEAAK, contains two His residues (in bold) in positions $i, i+4$ that create a zinc coordination site on one face of the α -helix. A number of residues were included in the sequence to increase the α -helical content, such as a C-terminal E-K salt bridge, and a number of Ala residues. The resulting metallointercalator-peptide nuclease chimera was found to convert a supercoiled pBR322 DNA plasmid to both the nicked and linear forms. Control experiments did not show any cleavage when the DNA was incubated with the unmodified rhodium intercalator, or the peptide nuclease without the intercalator, thus demonstrating the cooperative nature of the construct.⁶² Similar effects were observed with rhodium conjugates with metallopeptide hairpin nucleases.⁶³

A relatively recent example of these metallointercalator-peptide hybrids explored the effect of the peptide sequence in the internalization and intracellular localization.⁶⁴ In this study the $[\text{Ru}(\text{phen})(\text{bpy}')(\text{dppz})_2]^{2+}$ complex ($\text{bpy}' = 4$ -(3-carboxypropyl)-4'-methyl-2,2'-bipyridine) was attached to a well-known (*D*)-octaarginine (arg_8) peptide transporter, and to that same peptide modified with a fluorescein tag (arg_8 -Flu) (Figure 12). Interestingly, they found that while the simple $\text{Ru}[\text{arg}_8]$ conjugate was internalized by endocytosis, and remained in the cytoplasm in a punctuate pattern, the $\text{Ru}[\text{arg}_8$ -Flu] derivative was readily internalized into the cell nucleus under the same conditions for which the complex without fluorescein was excluded. This difference in intracellular localization is consistent with other studies that demonstrate the role of charge and hydrophobicity compensation in peptide transporters.⁶⁵ This result demonstrates the potential role of the peptide appendages not only in the DNA recognition process, but also for the modification of other properties, such as cell transport.

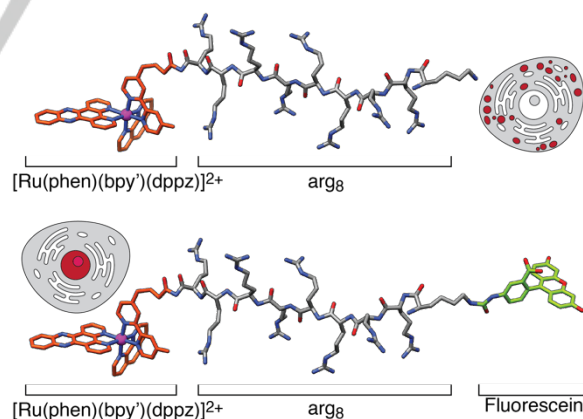


Figure 12. $[\text{Ru}(\text{phen})(\text{bpy}')(\text{dppz})_2]^{2+}$ arg_8 conjugates display different transport properties depending on their N-terminal modification with the hydrophobic fluorescein fluorophore. $[\text{Ru}(\text{phen})(\text{bpy}')(\text{dppz})_2]^{2+}$ - arg_8 is internalized in endosomes, while the analog featuring an N-terminal fluorescein is efficiently transported into the cell nucleus.

In addition to the modification of discrete metallointercalators, peptides have also been used as integral structural elements in *de novo* DNA binding metallopeptides without any resemblance

MICROREVIEW

to natural transcription factors. In this context, in 2014 we described the application of solid-phase peptide synthesis methodology to assemble peptide ligands as precursor of metal helicates, which are chiral metal complexes constituted by two (or more) metal ions. These metalocylinders have been shown to target unusual three-way DNA junctions,⁶⁶ and indeed, upon incubation with Fe(II), the peptide precursors enantioselectively assembled into $\Lambda\Lambda$ - or $\Delta\Delta$ -dimetallic helicates (Figure 13) that displayed high affinity for three-way DNA junctions ($K_D \approx 3.47 \mu\text{M}$ for the $\Lambda\Lambda$ - isomer, and $13.59 \mu\text{M}$ for the $\Delta\Delta$ -helicate). In contrast, the binding of these species to double-stranded DNA was rather weak ($K_D \approx 40 \mu\text{M}$ for A·T-rich oligonucleotides, and $K_D \approx 100 \mu\text{M}$ for G·C-rich oligos). Following the same methodology, our group has recently described a new family of oligonuclear Ru(II) and Ir(III) polypyridyl organometallopeptides that exhibit high DNA binding affinity in the low nanomolar range, sequence selectivity, and high cytotoxicity against set cancer cell lines.⁶⁷

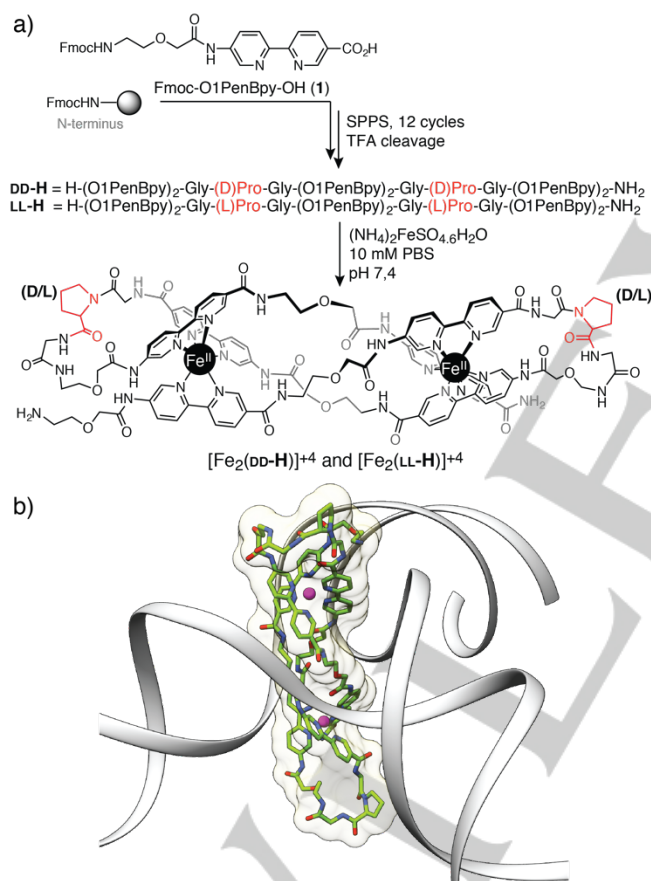


Figure 13. Solid phase peptide synthesis of the peptide helicates. Introduction of (L)-Pro or (D)-Pro residues in the loops selects a particular supramolecular chirality ($\Lambda\Lambda$ - or $\Delta\Delta$ -, respectively) in the final helicates. Bottom: Proposed structure of the peptide helicate bound to the center of a three-way DNA junction. DNA is only represented as ribbons for clarity.

2.3.3. Artificial GCN4 basic region monomers

Monomeric GCN4 basic regions do not bind to their target DNA half-site with high affinity due to the entropic cost associated to the folding of the peptide chain into the appropriate α -helical conformation, but residue grafting is a powerful strategy to obtain stabilized α -helices and thus obtain high-affinity DNA binding monomeric GCN4 peptides.⁶⁸ In this approach, selected residues that create the desired binding epitope are inserted (grafted) into a stable three-dimensional peptide scaffold. The group of A. Schepartz described in 1999 successful application of this strategy to obtaining high-affinity monomeric DNA-binding peptides.⁶⁹ Their design involved the introduction of the GCN4 residues mediating DNA binding in the basic region into the α -helix of the avian pancreatic polypeptide (aPP), a stable miniprotein consisting of a single α -helix stabilized by hydrophobic interactions with a type II polyproline helix. The resulting chimera was capable of recognizing the half site of the GCN4 DNA target sequence (5'-ATGA-3') with a K_D of 1.5 nM under physiological ionic strength. However, the mutations introduced in the aPP sequence for achieving DNA binding resulted in the disruption of the aPP hydrophobic core, and reduced structural stability of the construct. Further refinements of this initial binder using phage libraries resulted in a new peptide with extraordinary DNA affinity at 4 °C ($K_D \sim 23 \text{ pM}$), which even retained high affinity at 25 °C ($K_D \sim 1.6 \text{ nM}$).⁷⁰ Following the successful implementation of the grafting strategy to the synthesis of simplified versions of a bZIP transcription factor (GCN4), the Schepartz group demonstrated the versatility of this approach by obtaining miniature homeodomain proteins, again by dissecting the key DNA contacting residues and introducing them into the structure of the versatile aPP peptide scaffold.⁷¹

The group of G. L. Verdine demonstrated in 1995 that it was possible to drive specific DNA recognition by a monomeric GCN4 basic region if the peptide was delivered into the major groove intramolecularly.⁷² Attaching the GCN4 br to the DNA significantly increased its effective concentration, and reduced the entropic loss associated with intermolecular binding, thereby allowing the formation of the desired (intramolecular) complex. The design required a Gly-Gly-Cys connector to connect the C-terminus of the GCN4 br to an Adenine nucleotide immediately 5' to the AP1/GCRE half-site (5'-A(c)TCAT-3', modified base in bold) through a disulfide bond.

Taking the concept of intramolecular delivery for specific DNA binding one step further, the group of J. L. Mascareñas reported in 2001 the synthesis of conjugates between the GCN4 br and a small minor groove binding molecule that was capable of sequence-specific DNA recognition.⁷³ The role of the minor groove binding agent (a distamycin derivative) was analogous that of the disulfide linker in the original Verdine's intramolecular complex: to act as a supramolecular anchor that, upon binding to its target sequence in the DNA minor groove (5'-AAATT-3'), would deliver the GCN4 br into its binding site—the ATF/CREB half-site (5'-gTCAT-3')—in the adjacent major groove. Based on the X-ray structures of the GCN4 dimer and distamycin bound to their respective DNA sites,⁷⁴ the authors built a hypothetical model of the simultaneous interaction of the GCN4 basic region and the

MICROREVIEW

small DNA binding agent bound to contiguous sites, which allowed the identification of Arg²⁴⁵ as the ideal position for installing short aminoalkyl linkers connecting with the distamycin bound to the adjacent minor groove. Unlike isolated GCN4 basic regions, the resulting conjugate displayed tight (low nM affinity at 4 °C) and selective binding to a composite site containing the consensus sequences for both the GCN4 and the distamycin derivative (5'-gTCAT-AAATT-3'). As expected, control experiments with DNAs lacking the target sequence of the distamycin anchor did not show binding of the GCN4-distamycin hybrid to the DNA, thus confirming the key role of the accessory interactions established by the distamycin analog in the minor groove for stabilizing the complex.

Following that initial report, the group of J. L. Mascareñas demonstrated that the conjugation strategy could be generally applied for the stabilization of TF fragments, describing a variety of peptide hybrids that recognized extended sites containing the sequence of both the peptides and the distamycin. Thus, for example, the DNA binding helix of the Skn-1 transcription factor could be conjugated through its Lys²³² side chain to distamycin, and the resulting hybrid selectively bound to the site 5'-AAAA-TCAT-3'.⁷⁵ Importantly, this strategy could also be extended to other transcription factor families, so that conjugation to distamycin restored the DNA binding properties of a fragment of the GAGA zinc finger transcription factor, which selectively bound to the composite sequence 5'-TTTT-GAGAG-3' containing an A-T-rich site for distamycin binding (5'-TTTT-3') next to the GAGA TF consensus binding sequence (5'-GAGAG-3').⁷⁶ Similar conjugates with GCN4 and GAGA fragments, and even with homeodomain fragments,⁷⁷ were also described using bisbenzamidines as minor groove binding ligands,⁷⁸ further demonstrating the modular nature and general applicability of this strategy for obtaining sequence-selective DNA binding peptides derived from transcription factors (Figure 14).⁷⁹

The thermodynamic stabilization of TF fragments to obtain high-affinity DNA binders does not require the use of artificial DNA binding agents such as distamycin or bisbenzamidines as described before, but can also be achieved with fully peptidic anchors. The AT-Hook is a short cationic peptide (RKPRGRPCK) found in eukaryotic HMG-I(Y) nuclear proteins.⁸⁰ Although AT-Hooks bind to their target DNA sites with poor affinity (in the millimolar range),⁸¹ HMG-I(Y) proteins exploit the cooperative effect of three AT-Hook repeats to achieve high DNA binding affinity.⁸² NMR and crystallography studies have provided a detailed structural picture of the interaction of the AT-Hook with the DNA, and have shown that its central RGR core is deeply inserted into the minor groove in an extended conformation, while the various lysines in the sequence introduce additional electrostatic contacts with the phosphates of the DNA backbone.⁸³ We recently reported the use of an AT-Hook as minor groove binding element in combination with the GCN4 br. The peptidic nature of the AT-Hook allows the straightforward synthesis of the GCN4 br/AT-Hook chimera following exclusively solid-phase peptide synthesis methods, and the lower DNA binding affinity of the AT-Hook compared to bisbenzamidines or distamycin results in better selectivity profile.⁸⁴

A related example, recently published by the group of S. Roy in 2017, further demonstrates that the combination of short minor groove anchors and TF fragments is a robust approach to developing sequence-specific DNA binding peptides. In their work, S. Roy's group used a short minor groove-binding tail from the Serum Response Factor (SRF) in combination with the DNA-binding helix of Elk1 to yield a dominant negative of the two transcription factors, which upon DNA binding to the DNA upregulate the oncogene cFos.⁸⁵ The authors explored different linkers, such as GlyProGlyProGly, Ahx₃ (Ahx = aminohexanoic acid), Gly-Ahx₃, Gly₂-Ahx₃, and Gly₃-Ahx₃, between both DNA binding units, and studied as well conformationally-stabilized versions of the DNA contacting helix containing Aib residues, which are known to stabilize the α -helical conformation by restricting the available backbone conformations.⁸⁶ Importantly, in contrast with other conjugates that appear to get stuck inside endosomes upon internalization, the Gly₂-Ahx₃ chimera was readily internalized into the nuclei of A549 lung adenocarcinoma cells, and induced a specific response, down-regulating its targeted gene.

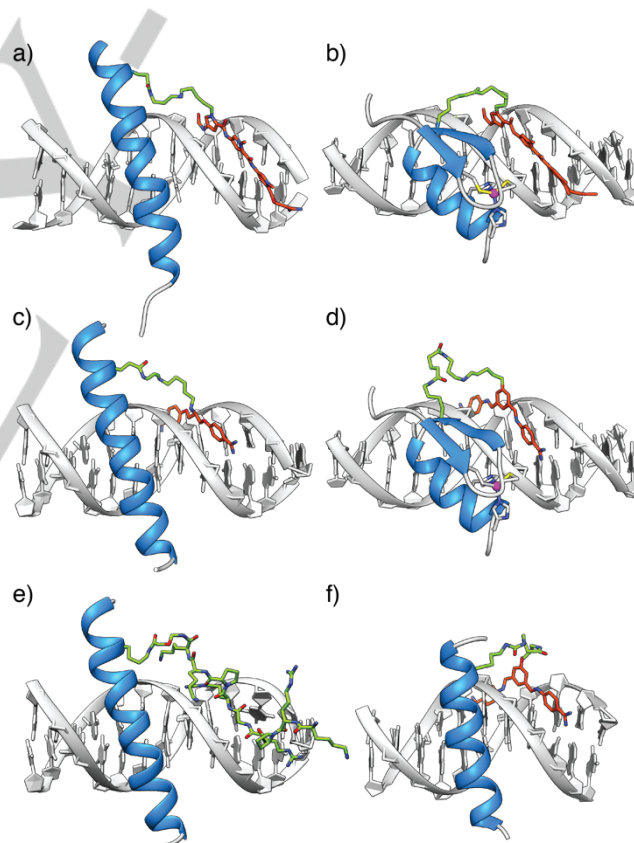


Figure 14. Proposed models of TF fragments stabilized by small minor groove binding agents. a) Original GCN4 basic region conjugate with distamycin,⁷³ b) GAGA/distamycin conjugate,⁷⁶ c) GCN4/bisbenzamide hybrid, d) GAGA/bisbenzamide hybrid; e) GCN4/AT-Hook chimera,⁸⁴ f) engrailed homeodomain/bisbenzamide conjugate.

MICROREVIEW

In an alternative approach to dominant negative DNA-binding peptides we described the synthesis of a dominant negative of the oncogenic c-Jun protein as an inactive complex by forcing its heterodimerization with a c-Fos-DNA conjugate.⁸⁷ This oligonucleotide-peptide receptor was based on the X-ray structure of the the c-Fos/c-Jun heterodimer bound to the AP1 site,⁸⁸ and included a fragment of the c-Fos leucine zipper (Arg¹⁵⁸ to Lys¹⁹²) to induce c-Jun dimerization, attached to the 5' end of a short oligonucleotide (5'-TGACT**CATCC**ATTGCGCG-3') containing the AP1 half-site (in bold) for c-Jun binding. The oligonucleotide provides increased affinity and selectivity for c-Jun, as demonstrated by its tight binding constant ($K_D \approx 200$ nM) that allowed it to efficiently compete with the natural c-Fos and disassemble the natural c-Fos/c-Jun heterodimer bound to the DNA.

In summary, since the early 90s a number of groups have developed efficient strategies for the development of miniaturized versions of transcription factors, these efforts have not only expanded the repertoire of DNA binding motifs, but also produced new systems with unique properties not found in natural systems, such as light-activation. The progress in the field, however, has been relatively slow partly due to our rudimentary understanding of peptide and protein folding, and still limited computational methods for efficient *in silico* selection in the context of peptide-DNA interactions. On the other hand, while many systems have been described, most of them are isolated and static complexes, but complex and dynamic systems involving more than one DNA binding partner are still largely unexplored, and will probably be the focus of future development. Finally, limited by their inherent internalization problems, these peptide systems have not yet found real biological applications, but new developments in delivery platforms and strategies will open the way for their use in living systems, exploiting their full potential as artificial TFs for controlling gene expression.⁸⁹

Acknowledgements

Financial support from the Spanish grant CTQ2015-70698-R, the Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2016–2019, ED431G/09) and the European Union (European Regional Development Fund - ERDF), is gratefully acknowledged. J. G.-G. thanks the Spanish MINECO for his FPI fellowship. We also wish to acknowledge the generous support by the *Fundación AECC* (IDEAS197VAZQ grant).

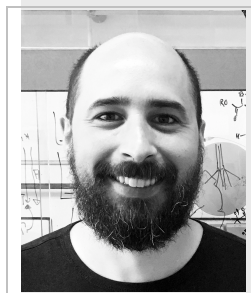
Sonia Boga was born in Teo (1990). She graduated in Chemistry from the University of Santiago de Compostela, and since February 2016 she is studying oligocationic DNA-binding metalloptides as part of her master's degree project.



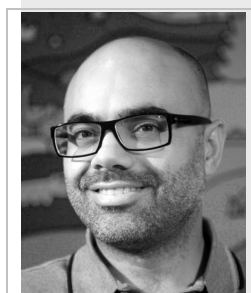
David was born in Cambados (1989). He graduated in Chemistry at the University of Santiago de Compostela in 2014, where he also obtained his Master's degree in organic chemistry. In October 2015 he joined the group of Prof. M. Eugenio Vázquez and Miguel Vázquez to start a new research line working on self-assembled ring proteins and catalytic peptides..



Dr. Diego Garcia graduated in Chemistry from the University of Santiago de Compostela in 2008. He obtained his PhD in 2015 under the supervision of Prof. Enrique Guitián and Dolores Pérez, working in the synthesis and study of new emissive polycyclic aromatic hydrocarbons (PAH's) as new molecular materials. In March 2015 he joined the group of Prof. Eugenio Vázquez and Miguel López as postdoctoral researcher to work in the development of new PAH's as mitochondria-targeted fluorescent tags and free radical sponges.



Prof. Vázquez López graduated in Chemistry from the *Universidade de Santiago de Compostela* in 1996, and finished his PhD in 2001. He then joined the group of Prof. Dante Gatteschi at the *Università degli Studi di Firenze* (Italy). In 2002 he received a TMR Marie Curie Fellowship and joined the group of Prof. Luigi Fabbrizzi at the *Università di Pavia* (Italy), where he worked for two years. In 2004 he got a *Parga Pondal* Tenure Track Position at the *Universidade de Vigo* (Spain). Prof. Vázquez López returned to the USC in 2006 as a *Ramon y Cajal* researcher. In 2009 he joined for three months the group of Prof. Jean-Claude Bünzli at the *École Polytechnique Fédérale de Lausanne* (Switzerland) as a Visitor Professor. In 2011 he became Associate Professor at the Inorganic Chemistry Department of the USC. His research interests are focused in the areas of (Metallo)Supramolecular Chemistry and Chemical Biology. He is coauthor of more than 50 research papers in peer-reviewed international journals.



Prof. Vázquez graduated in Chemistry from the *Universidad de Santiago de Compostela* in 1996. After finishing his PhD in 2001, he received the *Human Frontier Science Program* Long-Term Fellowship, and joined the group of Prof. Barbara Imperiali at the *Massachusetts Institute of Technology*, where he worked for three years. Prof. Vázquez returned to the University of Santiago in 2004 as a *Ramón y Cajal* researcher, and became Associated Professor at the Organic Chemistry Department in 2010. Since 2011 he is member of the Center for



Biological Chemistry and Molecular Materials (CIQUS) at this university. He has published more than 70 papers, and received a number of awards, including the *Spanish Royal Society of Chemistry – Lilly Organic Chemistry Group Young Researcher Award*, and the *Human Frontier Career Development Award*.

Keywords: supramolecular chemistry • peptides • DNA binding • metalloproteins • chemical biology

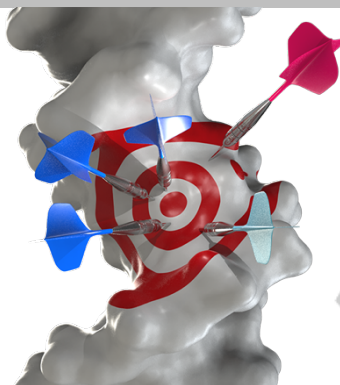
- [1] A. H. Brivanlou, J. E. Darnell Jr, *Science* **2002**, 295, 813–818.
- [2] a) A.-L. Todeschini, A. Georges, R. A. Veitia, *Trends Genet.* **2014**, 30, 211–219; b) M. Ptashne, *Nature* **1986**, 322, 697–701.
- [3] P. P. Pandolfi, *Oncogene* **2001**, 20, 3116–3127.
- [4] A. R. Chandrasekaran, *Nanoscale* **2016**, 8, 4436–4446.
- [5] a) S. Neidle, *Nat. Prod. Rep.* **2001**, 18, 291–309; b) A. Ali, S. Bhattacharya, *Bioorg. Med. Chem.* **2014**, 22, 4506–4521; c) W. Han Ang, P. J. Dyson, *Eur. J. Inorg. Chem.* **2006**, 2006, 4003–4018; d) C. Moucheron, *New J. Chem.* **2009**, 33, 235–245; e) B. M. Zeglis, V. C. Pierre, J. K. Barton, *Chem. Commun.* **2007**, 4565–4579; f) C. Y. Majmudar, A. K. Mapp, *Curr. Opin. Chem. Biol.* **2005**, 9, 467–474; g) P. B. Dervan, R. W. Bürli, *Curr. Opin. Chem. Biol.* **1999**, 3, 688–693; h) P. B. Dervan, *Bioorg. Med. Chem.* **2001**, 9, 2215–2235.
- [6] a) J. B. Chaires, *Curr. Opin. Struct. Biol.* **1998**, 8, 314–320; b) I. Romero-Canelón, P. J. Sadler, *Inorg. Chem.* **2013**, 52, 12276–12291; c) A. C. Komor, J. K. Barton, *Chem. Commun.* **2013**, 49, 3617–3630; d) G. S. Khan, A. Shah, Zia-ur-Rehman, D. Barker, *J. Photochem. Photobiol. B* **2012**, 115, 105–118.
- [7] a) E. Pazos, J. Mosquera, M. E. Vázquez, J. L. Mascareñas, *ChemBioChem* **2011**, 12, 1958–1973; b) M. E. Vázquez, A. M. Caamaño, J. L. Mascareñas, *Chem. Soc. Rev.* **2003**, 32, 338–349.
- [8] a) V. A. Bloomfield, D. M. Crothers, I. Tinoco, *Nucleic Acids: Structures, Properties, and Functions*, Sterling Publishing Company, **2000**; b) C. R. Calladine, H. Drew, *Understanding DNA: The Molecule and How It Works*, Academic Press, **1997**.
- [9] R. E. Dickerson, *J. Mol. Biol.* **1983**, 166, 419–441.
- [10] H. C. Nelson, J. T. Finch, B. F. Luisi, A. Klug, *Nature* **1987**, 330, 221–226.
- [11] D. G. Alexeev, A. A. Lipanov, Skuratovskii IYa, *Nature* **1987**, 325, 821–823.
- [12] a) S. K. Pal, L. Zhao, A. H. Zewail, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 8113–8118; b) M. L. McDermott, H. Vanselow, S. A. Corcelli, P. B. Petersen, *ACS Cent. Sci.* **2017**, DOI 10.1021/acscentsci.7b00100.
- [13] B. Jayaram, K. A. Sharp, B. Honig, *Biopolymers* **1989**, 28, 975–993.
- [14] E. P. Bishop, R. Rohs, S. C. J. Parker, S. M. West, P. Liu, R. S. Mann, B. Honig, T. D. Tullius, *ACS Chem. Biol.* **2011**, 6, 1314–1320.
- [15] V. Iyer, K. Struhl, *EMBO J.* **1995**, 14, 2570–2579.
- [16] S. A. Coulocheri, D. G. Pigis, K. A. Papavassiliou, A. G. Papavassiliou, *Biochimie* **2007**, 89, 1291–1303.
- [17] a) R. Rohs, S. M. West, P. Liu, B. Honig, *Curr. Opin. Struct. Biol.* **2009**, 19, 171–177; b) R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, B. Honig, *Nature* **2009**, 461, 1248–1253; c) M. Y. Zhitnikova, A. V. Shestopalova, *J. Biomol. Struct. Dyn.* **2016**, 1–14.
- [18] a) I. Haq, J. Ladbury, *J. Mol. Recognit.* **2000**, 13, 188–197; b) J. B. Chaires, *Biopolymers* **1997**, 44, 201–215.
- [19] L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown, C. J. Schofield, *Chem. Commun.* **2014**, 50, 1794–1796.
- [20] a) S. Neph, A. B. Stergachis, A. Reynolds, R. Sandstrom, E. Borenstein, J. A. Stamatoyannopoulos, *Cell* **2012**, 150, 1274–1286; b) T. Ravasi, H. Suzuki, C. V. Cannistraci, S. Katayama, V. B. Bajic, K. Tan, A. Akalin, S. Schmeier, M. Kanamori-Katayama, N. Bertin, et al., *Cell* **2010**, 140, 744–752.
- [21] a) L. Chen, *Curr. Opin. Struct. Biol.* **1999**, 9, 48–55; b) A. Reményi, H. R. Schöler, M. Wilmanns, *Nat. Struct. Mol. Biol.* **2004**, 11, 812–815.
- [22] C. W. Garvie, C. Wolberger, *Mol. Cell* **2001**, 8, 937–946.
- [23] a) T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **1992**, 71, 1223–1237; b) W. Keller, P. König, T. J. Richmond, *J. Mol. Biol.* **1995**, 254, 657–667.
- [24] X. Wang, W. Cao, A. Cao, L. Lai, *Biophys. J.* **2003**, 84, 1867–1875.
- [25] a) C. Park, J. L. Campbell, W. A. Goddard, *J. Am. Chem. Soc.* **1996**, 118, 4235–4239; b) M. Zhang, B. Wu, H. Zhao, J. W. Taylor, *J. Pept. Sci.* **2002**, 8, 125–136.
- [26] a) M. Miller, *Curr. Protein Pept. Sci.* **2009**, 10, 244–269; b) J. Habchi, P. Tompa, S. Longhi, V. N. Uversky, *Chem. Rev.* **2014**, 114, 6561–6588.
- [27] R. V. Talanian, C. J. McKnight, P. S. Kim, *Science* **1990**, 249, 769–771.
- [28] R. V. Talanian, C. J. McKnight, R. Rutkowski, P. S. Kim, *Biochemistry* **1992**, 31, 6871–6875.
- [29] B. Cuenoud, A. Schepartz, *Science* **1993**, 259, 510–513.
- [30] E. Oheix, A. F. A. Peacock, *Chemistry* **2014**, 20, 2829–2839.
- [31] T. Morii, M. Simomura, S. Morimoto, I. Saito, *J. Am. Chem. Soc.* **1993**, 115, 1150–1151.
- [32] T. Morii, Y. Saimei, M. Okagami, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1997**, 119, 3649–3655.
- [33] M. Ueno, A. Murakami, K. Makino, T. Morii, *J. Am. Chem. Soc.* **1993**, 115, 12575–12576.
- [34] Y. Aizawa, Y. Sugiura, M. Ueno, Y. Mori, K. Imoto, K. Makino, T. Morii, *Biochemistry* **1999**, 38, 4008–4017.
- [35] Y. Ruiz García, J. Zelenka, Y. V. Pabon, A. Iyer, M. Buděšínský, T. Kraus, C. I. E. Smith, A. Madder, *Org. Biomol. Chem.* **2015**, 13, 5273–5278.
- [36] L. L. G. Carrette, T. Morii, A. Madder, *Eur. J. Org. Chem.* **2014**, 2014, 2883–2891.
- [37] a) D. Verzele, A. Madder, *Eur. J. Org. Chem.* **2013**, 2013, 673–687; b) Y. R. García, Y. Vladimir Pabon-Martinez, C. I. Edvard Smith, A. Madder, *Chem. Commun.* **2017**, 53, 6653–6656.
- [38] M. I. Sánchez, O. Vázquez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2012**, 3, 2383.
- [39] a) R. J. Mart, R. K. Allemann, *Chem. Commun.* **2016**, 52, 12262–12277; b) H. M. D. Bandara, S. C. Burdette, *Chem. Soc. Rev.* **2012**, 41, 1809–1825; c) A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.*, **2011**, 40, 4422–4437.
- [40] A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.* **2000**, 112, 3234–3237.
- [41] W. J. Gehring, Y. Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A. F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, K. Wüthrich, *Cell* **1994**, 78, 211–223.

- [42] L. Tucker-Kellogg, M. A. Rould, K. A. Chambers, S. E. Ades, R. T. Sauer, C. O. Pabo, *Structure* **1997**, *5*, 1047–1054.
- [43] a) J. R. Kumita, O. S. Smart, G. A. Woolley, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3803–3808; b) G. A. Woolley, *Acc. Chem. Res.* **2005**, *38*, 486–493; c) J. R. Kumita, D. G. Flint, G. A. Woolley, O. S. Smart, *Faraday Discuss.* **2003**, *122*, 89–103; discussion 171–90.
- [44] 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.
- [45] a) O. Sadovski, A. A. Beharry, F. Zhang, G. A. Woolley, *Angew. Chem. Int. Ed Engl.* **2009**, *48*, 1484–1486; b) A. A. Beharry, O. Sadovski, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.
- [46] F. Zhang, K. A. Timm, K. M. Arndt, G. A. Woolley, *Angew. Chem. Int. Ed Engl.* **2010**, *49*, 3943–3946.
- [47] M. Olive, D. Krylov, D. R. Echlin, K. Gardner, E. Taparowsky, C. Vinson, *J. Biol. Chem.* **1997**, *272*, 18586–18594.
- [48] J. M. Mason, M. A. Schmitz, K. M. Müller, K. M. Arndt, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8989–8994.
- [49] A. M. Ali, M. W. Forbes, G. A. Woolley, *Chembiochem* **2015**, *16*, 1757–1763.
- [50] a) A. M. Ali, J. M. Reis, Y. Xia, A. J. Rashid, V. Mercaldo, B. J. Walters, K. E. Brechun, V. Borisenko, S. A. Josselyn, J. Karanicolas, et al., *Chem. Biol.* **2015**, *22*, 1531–1539; b) A. Kumar, A. M. Ali, G. A. Woolley, *Photochem. Photobiol. Sci.* **2015**, *14*, 1729–1736.
- [51] A. Jiménez-Balsa, E. Pazos, B. Martínez-Albardonedo, J. L. Mascareñas, M. E. Vázquez, *Angew. Chem. Int. Ed Engl.* **2012**, *51*, 8825–8829.
- [52] a) T. Härd, T. Lundbäck, *Biophys. Chem.* **1996**, *62*, 121–139; b) J. K. Strauss-Soukup, L. J. Maher 3rd, *Biochemistry* **1998**, *37*, 1060–1066.
- [53] G. A. Bullen, J. H. R. Tucker, A. F. A. Peacock, *Chem. Commun.* **2015**, *51*, 8130–8133.
- [54] J. Mosquera, M. I. Sánchez, M. Eugenio Vázquez, J. L. Mascareñas, *Chem. Commun.* **2014**, *50*, 10975–10978.
- [55] J. Mosquera, M. I. Sánchez, J. L. Mascareñas, M. Eugenio Vázquez, *Chem. Commun.* **2015**, *51*, 5501–5504.
- [56] a) Q. Liang, P. D. Eason, E. C. Long, *J. Am. Chem. Soc.* **1995**, *117*, 9625–9631; b) Y.-Y. Fang, B. D. Ray, C. A. Claussen, K. B. Lipkowitz, E. C. Long, *J. Am. Chem. Soc.* **2004**, *126*, 5403–5412.
- [57] S. Futaki, T. Kiwada, Y. Sugiura, *J. Am. Chem. Soc.* **2004**, *126*, 15762–15769.
- [58] Y. Azuma, M. Imanishi, T. Yoshimura, T. Kawabata, S. Futaki, *Angew. Chem. Int. Ed.* **2009**, *48*, 6853–6856.
- [59] J. Mosquera, A. Jiménez-Balsa, V. I. Doderio, M. E. Vázquez, J. L. Mascareñas, *Nat. Commun.* **2013**, *4*, 1874.
- [60] a) N. Y. Sardesai, S. C. Lin, K. Zimmerman, J. K. Barton, *Bioconjug. Chem.* **1995**, *6*, 302–312; b) C. A. Hastings, J. K. Barton, *Biochemistry* **1999**, *38*, 10042–10051.
- [61] N. Y. Sardesai, J. K. Barton, *J. Biol. Inorg. Chem.* **1997**, *2*, 762–771
- [62] M. P. Fitzsimons, J. K. Barton, *J. Am. Chem. Soc.* **1997**, *119*, 3379–3380.
- [63] K. D. Copeland, M. P. Fitzsimons, R. P. Houser, J. K. Barton, *Biochemistry* **2002**, *41*, 343–356.
- [64] C. A. Puckett, J. K. Barton, *Bioorg. Med. Chem.* **2010**, *18*, 3564–3569.
- [65] a) C. N. Carrigan, B. Imperiali, *Anal. Biochem.* **2005**, *341*, 290–298; b) W. B. Kauffman, T. Fuselier, J. He, W. C. Wimley, *Trends Biochem. Sci.* **2015**, *40*, 749–764.
- [66] a) A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymami, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem. Int. Ed.* **2006**, *45*, 1227–1231; b) S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger, P. Scott, *Nat. Chem.* **2012**, *4*, 31–36; c) M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3498; d) C. R. K. Glasson, G. V. Meehan, J. K. Clegg, L. F. Lindoy, J. A. Smith, F. R. Keene, C. Motti, *Chem. Eur. J.* **2008**, *14*, 10535–10538.
- [67] a) I. Gamba, I. Salvadó, R. F. Brissos, P. Gamez, J. Brea, M. I. Loza, M. E. Vázquez, M. Vázquez López, *Chem. Commun.* **2016**, *52*, 1234–1237; b) I. Gamba, I. Salvadó, G. Rama, M. Bertazzon, M. I. Sánchez, V. M. Sánchez-Pedregal, J. Martínez-Costas, R. F. Brissos, P. Gamez, J. L. Mascareñas, et al., *Chem. Eur. J.* **2013**, *19*, 13369–13375; c) G. Rama, A. Ardá, J.-D. Maréchal, I. Gamba, H. Ishida, J. Jiménez-Barbero, M. E. Vázquez, M. Vázquez López, *Chem. Eur. J.* **2012**, *18*, 7030–7035.
- [68] a) J. A. Robinson, *Chembiochem* **2009**, *10*, 971–973; b) J. Fernandez-Carneado, D. Grell, P. Durieux, J. Hauert, T. Kovacsovics, G. Tuchscherer, *Biopolymers* **2000**, *55*, 451–458.
- [69] N. J. Zondlo, A. Schepartz, *J. Am. Chem. Soc.* **1999**, *121*, 6938–6939.
- [70] J. W. Chin, A. Schepartz, *J. Am. Chem. Soc.* **2001**, *123*, 2929–2930.
- [71] J. K. Montclare, A. Schepartz, *J. Am. Chem. Soc.* **2003**, *125*, 3416–3417.
- [72] D. Stanojevic, G. L. Verdine, *Nat. Struct. Biol.* **1995**, *2*, 450–457.
- [73] M. E. Vázquez, A. M. Caamaño, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2001**, *40*, 4723–4725.
- [74] M. Coll, C. A. Frederick, A. H. Wang, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8385–8389.
- [75] J. B. Blanco, M. E. Vázquez, L. Castedo, J. L. Mascareñas, *Chembiochem* **2005**, *6*, 2173–2176.
- [76] O. Vázquez, M. E. Vázquez, J. B. Blanco, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.* **2007**, *46*, 6886–6890.
- [77] J. Mosquera, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *Chembiochem* **2014**, *15*, 1092–1095.
- [78] a) J. B. Chaires, J. Ren, D. Hamelberg, A. Kumar, V. Pandya, D. W. Boykin, W. D. Wilson, *J. Med. Chem.* **2004**, *47*, 5729–5742; b) M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, Y. Liu, L. Hu, M.-H. David-Cordonnier, A. Lansiaux, C. Bailly, et al., *J. Am. Chem. Soc.* **2007**, *129*, 13732–13743; c) O. Vázquez, M. I. Sánchez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Org. Lett.* **2010**, *12*, 216–219.
- [79] a) J. B. Blanco, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003**, *10*, 713–722; b) J. B. Blanco, O. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Eur. J.* **2005**, *11*, 4171–4178.
- [80] a) C. Crane-Robinson, A. I. Dragan, P. L. Privalov, *Trends Biochem. Sci.* **2006**, *31*, 547–552; b) L. Aravind, D. Landsman, *Nucleic Acids Res.* **1998**, *26*, 4413–4421; c) R. Reeves, M. S. Nissen, *J. Biol. Chem.* **1990**, *265*, 8573–8582; d) T. Lund, K. H. Dahl, E. Mørk, J. Holtlund, S. G. Laland, *Biochem. Biophys. Res. Commun.* **1987**, *146*, 725–730.
- [81] C.-C. Cheng, Y.-H. Jian, C.-J. Lo, J.-W. Cheng, *Jnl Chinese Chemical Soc* **1998**, *45*, 619–624.
- [82] B. H. Geierstanger, B. F. Volkman, W. Kremer, D. E. Wemmer, *Biochemistry* **1994**, *33*, 5347–5355.
- [83] a) J. R. Huth, C. A. Bewley, M. S. Nissen, J. N. Evans, R. Reeves, A. M. Gronenborn, G. M. Clore, *Nat. Struct. Biol.* **1997**, *4*, 657–665; b) E. Fonfria-Subirós, F. Acosta-Reyes, N. Saperas, J. Pous, J. A. Subirana, J. L. Campos, *PLoS One* **2012**, *7*, e37120.
- [84] J. Rodríguez, J. Mosquera, J. R. Couceiro, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2015**, *6*, 4767–4771.
- [85] M. Chakraborty, S. Roy, *Chem. Commun.* **2017**, *53*, 376–379.
- [86] De Filippis V., De Antoni F, M. Frigo, Polverino de Lauro P, A. Fontana, *Biochemistry* **1998**, *37*, 1686–1696.
- [87] E. Pazos, C. Portela, C. Penas, M. E. Vázquez, J. L. Mascareñas, *Org. Biomol. Chem.* **2015**, *13*, 5385–5390.
- [88] J. N. Glover, S. C. Harrison, *Nature* **1995**, *373*, 257–261.
- [89] a) A. K. Mapp, A. Z. Ansari, M. Ptashne and P. B. Dervan, *Proc. Natl. Acad. Sci. U. S. A.*, **2000**, *97*, 3930–3935; b) A. Z. Ansari and A. K. Mapp, *Curr. Opin. Chem. Biol.*, **2002**, *6*, 765–772; c) J. W. Højfeldt, A. R. Van Dyke and A. K. Mapp, *Chem. Soc. Rev.*, **2011**, *40*, 4286–4294.

Entry for the Table of Contents

MICROREVIEW

Researchers have tried for many years to develop sequence-specific DNA-binding peptides that can reproduce the recognition properties of natural transcription factors. We present a historical perspective of those efforts from a supramolecular perspective.

**Key Topic***

*S. Boga, D. Bouzada, D. García Peña,^[a]
M. Vázquez López,^[b] M. E. Vázquez**

Page No. – Page No.

**Sequence-Specific DNA Recognition
with Designed Peptides**

A supramolecular chemistry perspective on DNA recognition with artificial peptides