DNA-Binding Peptides

DOI: 10.1002/ange.200603115

Sequence-Specific DNA Binding by Noncovalent Peptide–Tripyrrole Conjugates**

*Juan B. Blanco, Verónica I. Dodero, M. Eugenio Vázquez, Manuel Mosquera, Luis Castedo, and José L. Mascareñas**

The initiation of gene transcription is dependent on the interaction between specific proteins and short DNA sequences that are usually located upstream of the promoter of the gene.[1] In most cases, these proteins are unable to bind DNA as monomeric modules and need to cooperate with other proteins to form high-affinity complexes with specific DNA sequences.[2] Such is the case for the bZIP family of transcription factors, which bind DNA as leucine-zipper-mediated homo- or heterodimers, with an N-terminal basic region (BR) of each monomer inserting into adjacent DNA major grooves.[2,3] Interestingly, in many of these proteins, for example, the yeast transcriptional activator GCN4, the basic region is largely unstructured in the absence of DNA but folds into an α helix upon specific DNA binding.[4] It has been shown that the leucine zipper of bZIP proteins can be replaced by other noncovalent or covalent artificial dimerizing units without considerably compromising the DNA-recognition capabilities of the system.[5,6] However, isolated monovalent bZIP BRs exhibit poor DNA-binding affinities, unless the important DNA-contacting residues are appropriately grafted into a preorganized α helix.[7]

We recently demonstrated that it is possible to promote the DNA binding of such a monomeric bZIP BR upon appropriate cross-linking to a distamycin-like tripyrrole capable of interacting with moderate to good affinity in the

[*] Dr. J. B. Blanco, Dr. V. I. Dodero, Dr. M. E. Vázquez, Prof. Dr. L. Castedo, Prof. Dr. J. L. Mascareñas
Departamento de Química Orgánica
Universidade de Santiago de Compostela
Facultade de Química
15782 Santiago de Compostela (Spain)
Fax: (+34) 981-595-012
E-mail: qojoselm@usc.es
Homepage: http://qoweb.usc.es/jlmgroup/index.htm
Prof. Dr. M. Mosquera
Departamento de Química Física
Universidade de Santiago de Compostela
Facultade de Química
15782 Santiago de Compostela (Spain)

[**] This work was supported by the Spanish MEC (SAF2004-01044). J.B.B. thanks the Spanish M.E.C. for a predoctoral fellowship and M.E.V. thanks the HSFP organization for their support with a Career Development Award (CDA0032/2005-C) and the Spanish MEC for a "Ramon y Cajal" contract. We also thank Prof. J. Benavente for access to radioactivity facilities.

Supporting information (including materials, synthetic procedures, and details of the binding experiments) for this article is available on the WWW under http://www.angewandte.org or from the author.

minor groove adjacent to the BR target site.[8] The interaction of hybrids such as **1** to designated composite DNA sites (GTCATAAAA) most probably involves a simultaneous major- and minor-groove interaction (as outlined in Figure 1),

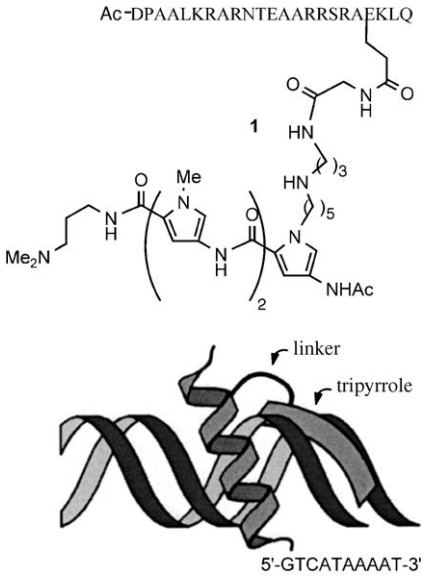
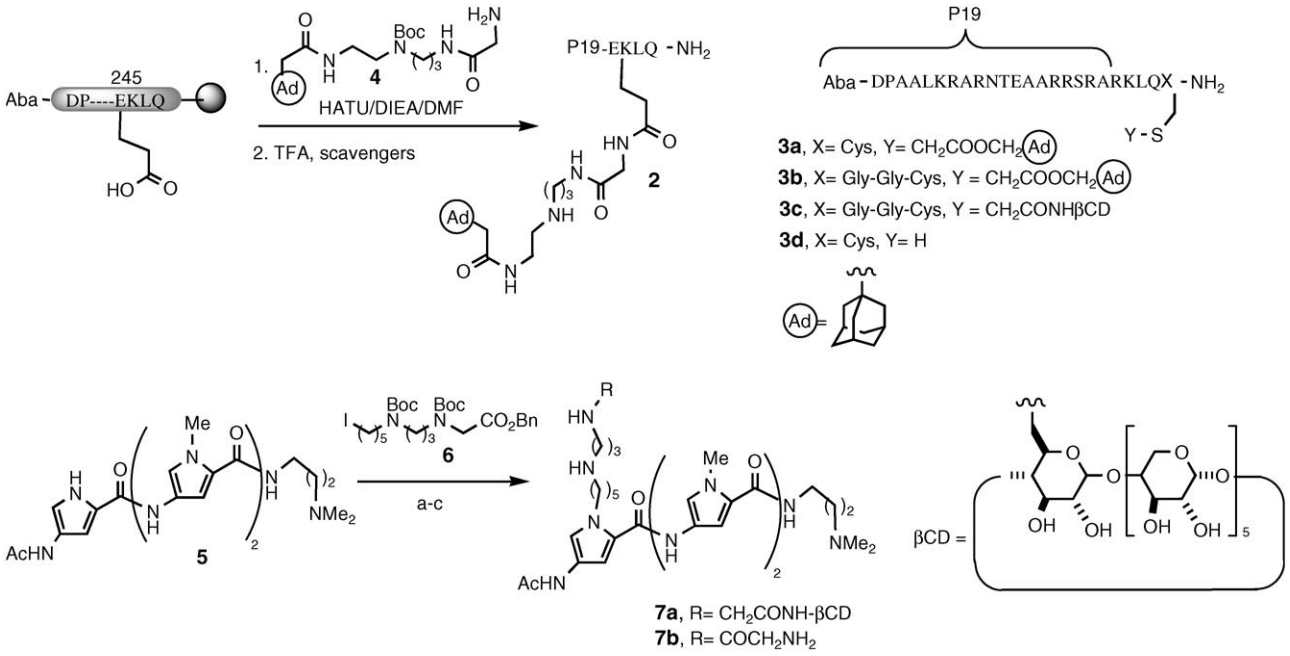


Figure 1. Structure of the tripyrrole–BR hybrid **1** and outline of its hypothetical DNA-binding mode.

and takes place with a K_d in the low nanomolar range at 4 °C. On the basis of this bipartite DNA-binding strategy, we considered it of interest to check whether both DNA-binding units could be connected by means of a noncovalent link so that the interaction would involve the formation of a stable complex among the peptide, the tripyrrole, and the DNA.[9] The feasibility of implementing this type of recognition strategy might open new and interesting opportunities to develop site-specific and ligand-responsive DNA-binding peptides. Herein we demonstrate that the attachment of an adamantyl group at the C terminus of a 23 amino acid peptide derived from a BR bZIP fragment allows its recruitment to a cognate DNA site adjacent to an A/T-rich sequence upon addition of a β-cyclodextrin-containing tripyrrole.[10]

It was not easy to deductively predict the best position of the basic region peptide to link the adamantyl group, and therefore we made peptides in which this unit is either attached to the side chain of amino acid 245 (peptide **2**) or at a C-terminal cysteine (**3a**, **3b**). In the case of peptide **2**, we chose a relatively long tether to connect the adamantane and the BR to ensure a certain degree of flexibility in the system. This peptide was readily prepared by coupling amine **4** with the required basic-region peptide while it was still bound to the resin and fully protected except at glutamic acid 245.[11] A standard TFA-promoted deprotection step led to the desired derivative **2** in good overall yield. Peptides **3a** and **3b** were readily obtained by coupling a fully deprotected basic-region peptide containing a cysteine residue at the C terminus with adamantylmethyl 2-bromoacetate. The cyclodextrin–tripyrrole unit **7a** was prepared from tripyrrole derivative **5** by following the steps indicated in Scheme 1.

As we have already shown for the covalent hybrids,[8] circular dichroism (CD) spectroscopy is a very interesting technique to extract information about the DNA-binding properties of this type of molecules. As shown in Figure 2a, the helicity of peptide **2** is not affected by the presence of 20-



Scheme 1. Synthesis of the adamantyl-containing peptides and the tripyrrole-βCD units. For the synthesis of **7a**: a) K_2CO_3, **6**, DMF/acetone, 60 °C. b) H_2, Pd/C, MeOH. c) 1) i) HATU, DIEA, ii) βCD-NH_2, DMF; 2) TFA. Aba = *p*-acetamidobenzoyl, Boc = *tert*-benzyloxycarbonyl, DIEA = *N*,*N*-diisopropylethylamine, DMF = *N*,*N*-dimethylformamide, HATU = *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, TFA = trifluoroacetic acid.

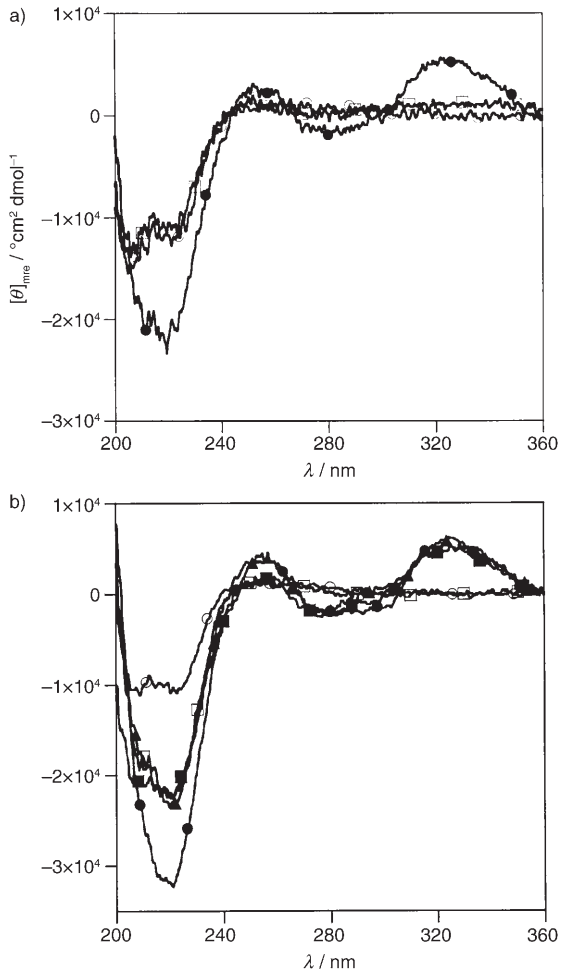


Figure 2. Circular dichroism difference spectra of peptides **2** and **3a**: a) peptide **2**: alone (○), in the presence of CRE^{hs}/A (□), and in the presence CRE^{hs}/A and **7a** (●); b) peptide **3a**: alone (○), in the presence of CRE^{hs}/A (□), in the presence CRE^{hs}/A and **7a** (●), in the presence CRE^{hs}/A and **7b** (■), and in the presence of CRE^{hs}m/A and **7a** (▲). CD spectra were obtained at 4 °C (see the Supporting Information) and are the difference between the spectra of the mixtures and the spectrum of free dsDNA. Sequences of one strand of the ds-oligonucleotides used: CRE^{hs}/A: 5′-d(ACGAACGTCAT*AAAAT*C-CTC)-3′, CRE^{hs}m/A: 5′-d(ACGAACGTCGT*AAAAT*CCTC)-3′. The BR subsite (CRE^{hs}) is shown underlined, and the tripyrrole subsite is shown in italics.

base-pair (bp) duplex oligonucleotide containing the target hybrid DNA sequence (CRE^{hs}/A). However, addition of tripyrrole **7a** to the mixture induces a significant, though not particularly large, increase in the negative intensity of the circular dichroism signal at 222 nm (Figure 2a, ●). Peptide **3a**, which exhibits the natural amino acids of the basic region of GCN4, acquires a notable degree of helicity in the presence of CRE^{hs}/A (Figure 2b, □) but becomes significantly more helicoidal upon addition of the βCD-tripyrrole **7a** (Figure 2b, ●). Peptide **3b** behaved similarly.

Interestingly, the addition of tripyrrole **7b** (that lacks the cyclodextrin unit) in place of **7a** does not promote a change in the helicity of the peptide (Figure 2b, ■). These results are consistent with significant α-helix formation and specific binding of the peptide to the DNA only when the tripyrrole equipped with the βCD (**7a**) is present. In all cases, there is also a positive ellipticity increase at 330 nm, which is consistent with the tripyrrole unit binding in the DNA minor groove. As might be expected, in the presence of the double-stranded (ds) oligonucleotide that features a bp mismatch at the BR-peptide binding subsite (CRE^{hs}m/A), the increase in the magnitude of the negative signal at 222 nm is considerably weaker (Figure 2b, ▲).

The circular dichroism information indicates that the tripyrrole-induced helical transition of peptide **3a** in the presence of the target DNA is considerably higher than for peptide **2**. Therefore, the subsequent DNA-binding studies, based on the use of gel mobility shift experiments (EMSA), were focused on the first peptide, which features the adamantane moiety at a C-terminal cysteine. As expected for an isolated monomeric bZIP BR,[12] we did not observe retard bands upon incubation of peptide **3a** with CRE^{hs}/A, a dsDNA that contains a cognate site for such peptides (Figure 3a, lane 1). However, it was gratifying to see that the addition of increasing amounts of the βCD-tripyrrole unit **7a** leads to the clear formation of a slower migrating band that must correspond to a DNA–peptide–tripyrrole complex (Figure 3a, lanes 2–5).

Interestingly, and in contrast to observations made in the case of covalent BR-peptide–tripyrrole hybrids,[8] a dsDNA molecule mutated at the BR binding site (CRE^{hs}m/A) does not induce the formation of detectable complexation bands (Figure 3a, lanes 6–10). Moreover, we could not detect mobility-shifted bands in the presence of a dsDNA molecule containing the BR-peptide binding subsite but lacking the A-rich sequence (CRE^{hs}, Figure 3a, lanes 11–12). These results are consistent with a very high sequence-specific DNA recognition, much better than in the covalent peptide–tripyrrole hybrids. That the supramolecular assembly is mediated by the host–guest cyclodextrin–adamantane interaction was demonstrated by carrying out control DNA-complexation experiments with the tripyrrole that lacks the βCD unit (**7b**) or with the peptide **3d**, which does not incorporate the adamantane group. As can be deduced from the absence of retard bands in lanes 14 and 15 in Figure 3a, both noncovalent interacting units (the βCD and the adamantane) are required for the formation of the complex between the BR peptide, the tripyrrole, and DNA. It is interesting to compare the gel migration of the above complex with that of the dimer of **3b** and **3c** with a ds-oligonucleotide (CRE) containing a cognate dimeric recognition site (Figure 3a, lane 13).[6] The relative migration is completely consistent with our complexes containing a single peptidic unit.

The specificity of the interaction was further examined by comparing the EMSA results with dsDNA molecules containing specifically designed mutations. As shown in Figure 3b, the introduction of a bp spacing between the consensus recognition sites of the tripyrrole and the BR peptide (CRE^{hs}*cg*/A) leads to a remarkable decrease in affinity (Figure 3b, lanes 6–10), and the presence of an additional bp space almost precludes the formation of the complex (CRE^{hs}*cg*/A, Figure 3b, lanes 11–15). On the other hand, interruption of the adenine (A) sequence in the ds-oligonu-

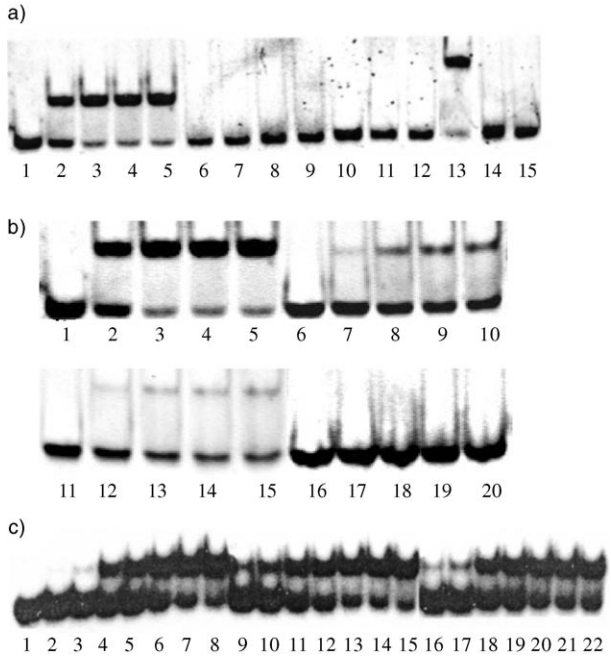


Figure 3. EMSA results showing the binding of peptide **3a** to dsDNA molecules. Experiments for (a) and (b) were analyzed by fluorescent dye staining and those of c) were analyzed by autoradiography. a) [dsDNA] ≈ 30 nM. Lanes 1–5: **7a** in the presence of a mixture of **3a** + CRE^{hs}/A, [**3a**] = 200 nM, [**7a**] = 0, 50, 100, 200, 400 nM; lanes 6–10: **7a** in the presence of a mixture of **3a** + CRE^{hs}m/A, [**3a**] = 200 nM, [**7a**] = 400, 200, 100, 50, 0 nM; lanes 11, 12: **7a** in the presence of a mixture of **3a** + CRE^{hs}, [**3a**] = 5 μM, [**7a**] = 0.5, 1 μM; lane 13: CRE + **3b** + **3c**, [**3b**] = [**3c**] = 500 nM; lane 14: **3a** + CRE^{hs}/A + **7b**, [**3a**] = 2 μM, [**7b**] = 500 nM; lane 15: **3d** + CRE^{hs}/A + **7a**, [**3d**] = 2 μM, [**7a**] = 500 nM. b) [dsDNA] ≈ 30 nM. Lanes 1–5: **7a** in the presence of a mixture of **3a** + CRE^{hs}/A, [**3a**] = 200 nM, [**7a**] = 0, 50, 100, 200, 400 nM; lanes 6–10: **7a** in the presence of a mixture of **3a** + CRE^{hs}*g*/A, [**3a**] = 200 nM, [**7a**] = 0, 50, 100, 200, 400 nM; lanes 11–15: **7a** in the presence of a mixture of **3a** + CRE^{hs}*cg*/A, [**3a**] = 200 nM, [**7a**] = 0, 50, 100, 200, 400 nM; lanes 16–20: **7a** in the presence of a mixture of **3a** + CRE^{hs}*g*/Am, [**3a**] = 200 nM, [**7a**] = 0, 50, 100, 200, 400 nM. c) Autoradiograms with ^{32}P-labeled CRE^{hs}/A (≈ 45 pM of ^{32}P-labeled + 100 nM unlabeled). Lanes 1–8: equimolecular mixture of **3a** and **7a**, [**3a**] = 0, 50, 80, 100, 150, 200, 250, 300 nM; lanes 9–15: **3a** in the presence of a mixture of **7a** + CRE^{hs}/A, [**7a**] = 200 nM, [**3a**] = 50, 80, 100, 150, 200, 250, 300 nM; lanes 16–22: **7a** in the presence of a mixture of CRE^{hs}/A + **3a**, [**3a**] = 200 nM, [**7a**] = 50, 80, 100, 150, 200, 250, 300 nM. Sequences of one strand of the ds-oligonucleotides used: CRE^{hs}: 5′-d(CGAACGTCATCGAAGGTCCT)-3′; CRE: 5′-d(TGGAGATGACGTCATCTCGT)-3′; CRE^{hs}*g*/A: 5′-d(CGAACGTCATG*AAAAT*CCTC)-3′; CRE^{hs}*cg*/A: 5′-d(CGAACGTCATCG*AAAAT*CCT)-3′; CRE^{hs}*g*/Am: 5′-d(CGAACGTCATG*AAA*GTCCTC)-3′. The BR subsite (CRE^{hs}) is shown underlined, and the tripyrrole subsite is shown in italics.

cleotide CRE^{hs}*g*/A completely abolishes the complexation (CRE^{hs}*g*/Am, Figure 3b, lanes 16–20). These data confirm the high specificity of the system and are consistent with the proposed major–minor groove interaction model.

Standard EMSA experiments by using ^{32}P-end-labeled dsDNA molecules revealed that to detect the bands of the complex it is necessary to use larger amounts of the dsDNA than commonly used in this type of assay. Well-known precedents for related systems suggest that such a requirement could probably be a consequence of a relatively high kinetic lability of the peptide–DNA complex, which may partially dissociate under the conditions of the assay.[12] Ensuring the presence of enough dsDNA (> 20 nM), we obtained clear complexation results upon titration of the target DNA with increasing amounts of equimolecular mixtures of **3a** and **7a**, or with fixed concentrations of one of the partners and increasing proportions of the other (Figure 3c). Titration assays in the presence of an excess of the tripyrrole **7a**, therefore ensuring that most of the dsDNA probe is saturated, allowed calculation of an approximate K_d value of **3a** for a DNA·**7a** complex of 63 ± 7 nM (see the Supporting Information). Assuming that the BR monomeric peptide binds DNA with affinities in the range of 1–5 μM,[12] we can infer that the presence of **7a** induces a very important binding improvement.

In conclusion, attachment of complementary noncovalent heterodimerizing units at a side chain of a distamycin-related tripyrrole and at the C terminus of a bZIP BR peptide provides for sequence-specific DNA recognition of relatively long DNA sites (8–9 bp). This strategy, which must involve a major- and minor-groove interaction, allows the DNA binding of very short peptides (23 amino acids) in a highly sequence-specific manner and represents a first step towards the development of small, highly selective and ligand-responsive DNA-binding peptides. Current studies are focused on further characterizing the binding mode, refining the system to obtain more stable complexes, and extending the recognition strategy to other transcription factor fragments.

Experimental Section

Circular dichroism measurements were made in a 2-mm cell at 4 °C. Samples contained 10 mM phosphate-buffered saline solution (pH 7.5), 100 mM NaCl, 5 μM peptide, and 5 μM ds-oligonucleotide when present. The peptide–DNA mixtures were incubated for 5 min before registering. For gel mobility shift assays, binding reactions were performed over 30 min in a binding mixture (20 or 40 μL) containing 18 mM tris(hydroxymethyl)aminomethane (Tris; pH 7.5), 90 mM KCl, 1.8 mM $MgCl_2$, 1.8 mM EDTA, 9% glycerol, 0.11 mg mL^{-1} bovine serum albumin (BSA) and 2.2% NP-40 (nonidet-P40). Products were resolved by PAGE by using a 10% nondenaturing polyacrylamide gel and 0.5XTBE buffer solution (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8) and analyzed by autoradiography (when radioactivity was used) or by staining with SyBrGold (Molecular Probes: 5 μL in 50 mL of 1XTBE) for 10 min and visualized with fluorescence.

Received: August 1, 2006
Revised: October 13, 2006
Published online: November 17, 2006

Keywords: distamycin · DNA recognition · noncovalent interactions · peptides · proteins

[1] R. J. White, *Gene Transcription, Mechanism and Control*, Blackwell Science, Oxford, **2001**.
[2] a) C. O. Pabo, R. T. Sauer, *Annu. Rev. Biochem.* **1992**, *61*, 1053–1095; b) D. S. Latchman, *Transcriptional Factors: Structural Families and Principles of DNA Recognition*, Academic Press, San Diego, **1998**; c) C. W. Garvie, C. Wolberger, *Mol. Cell* **2001**, *8*, 936–946; d) J. T. Kadonaga, *Cell* **2004**, *116*, 247–257; e) N. M.

Luscombe, S. E. Austin, H. M. Berman, J. M. Thornton, *Genome Biol.* **2000**, *1*, 1.
[3] a) J. C. Hu, R. T. Sauer, *Nucleic Acids Mol. Biol.* **1992**, *6*, 82–101; b) H. C. Hurst, *Protein Profile* **1995**, *2*, 101–168; c) T. Ellemberger, *Curr. Opin. Struct. Biol.* **1994**, *4*, 12–21.
[4] a) M. A. Weiss, *Biochemistry* **1990**, *29*, 8020–8024; b) K. T. O'Neil, R. H. Hoess, W. F. Degrado, *Science* **1990**, *249*, 774–778.
[5] For covalent dimers, see: a) R. V. Talanian, C. J. McKnight, P. S. Kim, *Science* **1990**, *249*, 769–771; b) B. Cuenoud, A. Shepartz, *Science* **1993**, *259*, 510–513; c) A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem.* **2000**, *112*, 3234–3237; *Angew. Chem. Int. Ed.* **2000**, *39*, 3104–3107; d) T. Morii, Y. Saimei, M. Okagami, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1997**, *119*, 3649–3655.
[6] For noncovalent dimers, see: a) T. Morii, M. Shimomura, M. Morimoto, I. Saito, *J. Am. Chem. Soc.* **1993**, *115*, 1150–1151; b) Y. Aizawa, Y. Sugiura, T. Morii, *Biochemistry* **1999**, *38*, 1626–1632.
[7] a) N. J. Zondlo, A. Schepartz, *J. Am. Chem. Soc.* **1999**, *121*, 6938–6939; b) J. W. Chin, A. Schepartz, *J. Am. Chem. Soc.* **2001**, *123*, 2929–2930; see also: c) D. Stanojevic, G. L. Verdine, *Nat. Struct. Biol.* **1995**, *2*, 450–457.
[8] a) M. E. Vázquez, A. M. Caamaño, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem.* **2001**, *113*, 4859–4861; *Angew. Chem. Int. Ed.* **2001**, *40*, 4723–4725; b) J. B. Blanco, M. E. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003**, *10*, 713–722; c) J. B. Blanco, O. Vázquez, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Eur. J.* **2005**, *11*, 4171–4178; d) J. B. Blanco, M. E. Vázquez, L. Castedo, J. L. Mascareñas, *ChemBioChem* **2005**, *6*, 2173–2177; For a review, see: e) M. E. Vázquez, A. M. Caamaño, J. L. Mascareñas, *Chem. Soc. Rev.* **2003**, *32*, 338–349.
[9] The formation of a cooperative DNA complex between a polyamide and the DNA-binding protein EXD (extradenticle) has been recently disclosed: a) H-D. Arndt, K. E. Hauschild, D. P. Sullivan, K. Lake, P. B. Dervan, A. Z. Ansari, *J. Am. Chem. Soc.* **2003**, *125*, 13322–13323; b) K. E. Hauschild, R. E. Metzler, H.-D. Arndt, R. Moretti, M. Raffaelle, P. B. Dervan, A. Z. Ansari, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5008–5013.
[10] We chose a cyclodextrin–adamantane interaction system because it has been successfully used to obtain DNA-binding dimeric BR bZIP peptides, see Ref. [6].
[11] The natural basic region contains an arginine at that position (245), but we changed it to glutamic acid to facilitate its synthetic manipulation (see Ref. [8a]).
[12] a) T. Morii, J. Yamane, Y. Aizawa, K. Makino, Y. Sugiura, *J. Am. Chem. Soc.* **1996**, *118*, 10012–10017; b) M. Zhang, B. Wu, H. Zhao, J. W. Taylor, *J. Pept. Sci.* **2002**, *8*, 125–136; c) D. D. Jones, P. D. Barker *Angew. Chem.* **2005**, *117*, 6495–6499; *Angew. Chem. Int. Ed.* **2005**, *44*, 6337–6341. See also Ref. [7].

