DOI: 10.1002/cbic.200500294

Sequence-Specific DNA Recognition by Monomeric bZIP Basic Regions Equipped with a Tripyrrole Unit on the N-Terminal Side. Towards the Development of Synthetic Mimics of Skn-1

Juan B. Blanco, M. Eugenio Vázquez, Luis Castedo, and José L. Mascareñas*[a]

The initiation of gene transcription is highly dependent on the interaction of certain proteins called *transcription factors* (TFs) with specific DNA sequences located upstream of the coding region of the gene.[1] One of the largest families of TFs in eukaryotic cells is the basic-leucine zipper (bZIP) class of proteins. These proteins bind specific dsDNA sequences as leucine zipper-mediated homo- or heterodimers, with the N-terminal *basic region* (BR) of each monomer inserting into adjacent DNA major grooves.[2] The basic region of the proteins is largely unstructured in the absence of DNA, but it folds into an α-helix upon specific DNA binding.[3] It has been shown that monomeric, isolated bZIP basic regions exhibit low DNA affinities, except for specifically designed versions in which the key DNA-binding residues are appropriately grafted into the α-helix of an *aPP* protein.[4]

Interestingly, TFs such as the *C. elegans* developmental transcription factor Skn-1, which lacks the leucine zipper moiety but contains α-helical recognition regions similar to those of

[a] *J. B. Blanco, M. E. Vázquez, Prof. Dr. L. Castedo, Prof. Dr. J. L. Mascareñas*
Departamento de Química Orgánica y Unidad Asociada al CSIC
Universidad de Santiago de Compostela
15782 Santiago de Compostela (Spain)
Fax: (+34) 981-595-012
E-mail: qojoselm@usc.es

*Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author: synthesis and characterization of all new compounds, conditions for PAGE and CD experiments, and gel-shift titration data for the binding of **4** to CRE^{hs}/T.*

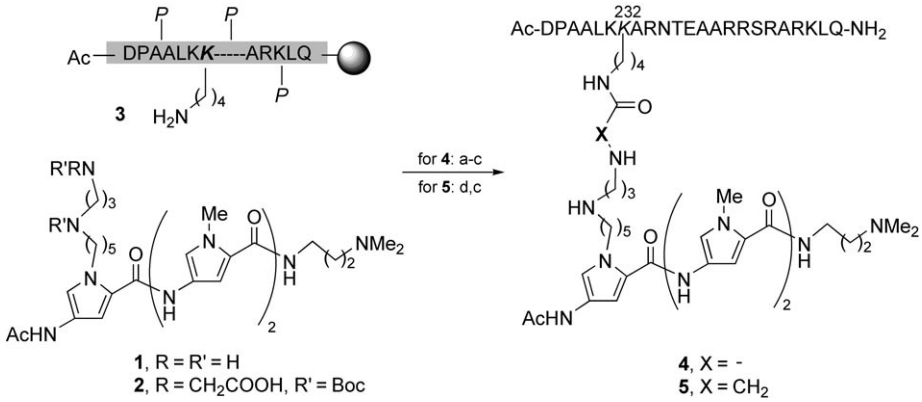
b-ZIP proteins c-Jun and GCN4, are capable of monomeric, high-affinity binding to specific 9 bp dsDNA sequences (ATGACATTG).[5,6] Relatively recent structural data indicate that the viability of this monomeric binding is to a large degree a consequence of a bipartite interaction mode involving standard basic region–major groove contacts with the DNA and ancillary interactions between an N-terminal arm of the protein and the minor groove of a contiguous sequence.[6]

We recently demonstrated that appropriate cross-linking of a tripyrrole related to distamycin A to a C-terminal amino acid of the GCN4 basic region provides a bivalent system capable of binding to designated composite sites (5′-TTTTATGAC-3′) with remarkable affinity.[7] The DNA-recognition strategy of this type of tripyrrole–peptide hybrid, which involves simultaneous minor- and major-groove interactions, therefore has a significant parallelism with the recognition mode used by the naturally occurring Skn-1 protein.

With the aim of extending the utility of this minor/major-groove-binding strategy to recognizing other DNA sequences, as well as moving forward the preparation of minimized synthetic mimics of Skn-1, we have now designed and synthesized a tripyrrole–peptide construct in which the tripyrrole moiety is connected to a residue located at the N-terminal side of the basic region.

In order to design the conjugates, we built a hypothetical model for the simultaneous interaction of the tripyrrole and the protein basic region on adjacent DNA sites, using as a reference the X-ray structures of the DNA complexes of GCN4[8] and distamycin A[9] bound to their respective cognate sequences (Figure 1). Inspection of the model suggested that an appropriate connection between the two fragments could involve the N-terminal pyrrole of the minor-groove binder and the residue 232 side chain of the BR peptide, which is an Arg in the wild-type GCN4 sequence (Figure 1 A). The resulting hybrid might be hypothetically capable of recognizing composite sites of the type 5′-ATGAXTTTT-3′.

The aminotripyrrole **1** was available from our previous studies,[7,10] and so we decided to replace Arg232 of the BR peptide by a lysine (Figure 1 B) so that coupling between **1** and the peptide could be carried out by using a biselectrophilic type of conjugating agent. This conjugation was best carried out by sequential coupling of disuccinimidyl carbonate and **1** to the selectively deprotected, resin-bound peptide **3**. We also made hybrid **5**, which has an amide instead of a urea functionality at the linking site, by direct coupling of the resin-linked peptide **3** with the tripyrrole derivative **2** (Scheme 1). Compound **2** was efficiently prepared in a convergent manner by a base-promoted coupling of tripyrrole **6** with a Boc-protected derivative of benzyl 2-[3-(5-iodopentylamino)propylamino]acetate (Scheme 2).



Scheme 1. Key steps in the synthesis of **4** and **5**. a) Disuccinimidyl carbonate, diisopropylethylamine (DIEA), DMF; b) **1**, DIEA, DMF; c) TFA, scavengers; d) **2**, *O*-(7-azabenzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate, DIEA, DMF. P: general symbol for protecting groups.

As we have already shown for related systems, mobility-shift electrophoresis (EMSA) and circular dichroism are useful methods for analyzing the DNA-binding properties of this type of molecule.

As can be deduced from the gel-shift results shown in Figure 2 A, both conjugates **4** and **5** give rise to clear retardation bands when incubated at 4 °C with ^{32}P end-labeled ds-oligonucleotides containing the designated composite sequence (CRE^{hs}/T). However, hybrid **4** exhibited a greater affinity by at least a factor of two, and we therefore decided to continue our studies with this hybrid.

Using as DNA probes the ds-oligonucleotides *m*CRE^{hs}/T or *m*′CRE^{hs}/T, which contain one

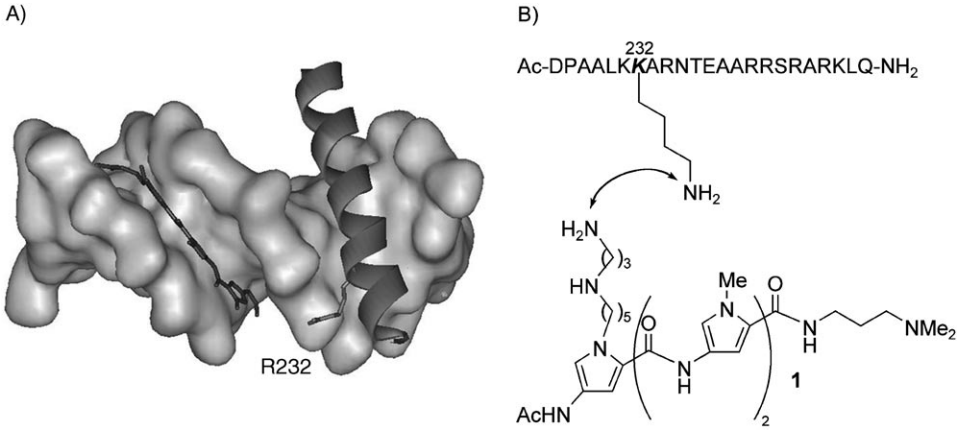


Figure 1. A) Qualitative model used for design purposes and B) structural detail of the positions to be connected.

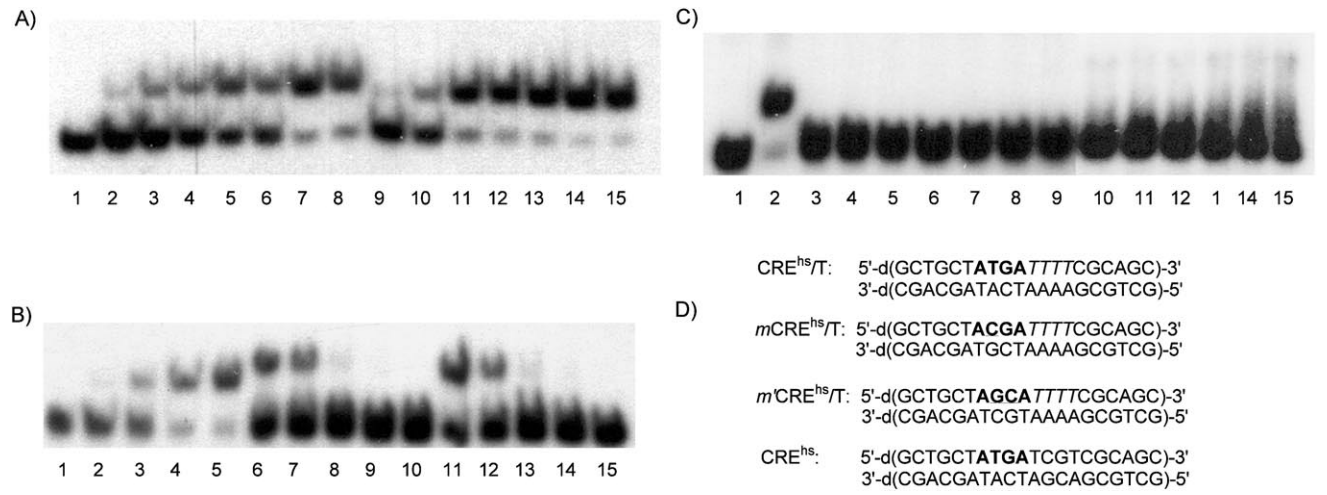


Figure 2. A) Autoradiogram showing the binding of hybrids **4** and **5** to ^{32}P-labelled CRE^{hs}/T. Lanes 1–8, [**5**] = 0, 10, 20, 40, 60, 80, 100, and 120 nM, respectively; lanes 9–15, [**4**] = 10, 20, 40, 60, 80, 100, and 120 nM, respectively. B) Autoradiogram showing the binding of hybrid **4** to ^{32}P-labelled dsDNAs: Lanes 1–5, CRE^{hs}/T, [**4**] = 0, 10, 20, 60, and 100 nM, respectively; lanes 6–10, *m'*CRE^{hs}/T, [**4**] = 100, 60, 20, 10, and 0 nM, respectively; lanes 11–15, *m*CRE^{hs}/T, [**4**] = 100, 60, 20, 10, and 0 nM, respectively. C) Autoradiogram showing the binding of hybrid **4** and peptide **7** to ^{32}P-labelled dsDNAs: Lanes 1 and 2, CRE^{hs}/T, [**4**] = 0 and 100 nM; lanes 3–9, CRE^{hs}/T, [**7**] = 0.6, 0.7, 0.8, 0.9, 1, 2, and 5 μM, respectively; lanes 10–15, CRE^{hs}, [**4**] = 0.7, 0.8, 0.9, 1, 2, and 5 μM, respectively. D) Sequences of duplex oligonucleotides used. The BR subsite (CRE^{hs}) is bold and the tripyrrole binding site (T) is in italics. Peptide **7** is: Aba-DPAALKKARNTEAARRSRARKLQ-NH$_2$ (Aba = 4-acetamidobenzoic acid).

Scheme 2. Key steps in the synthesis of **2**.

and two mutations in the BR-peptide-binding site, respectively, we observed the formation of a slightly slower migrating band than that obtained with CRE^{hs}/T (Figure 2B), this suggests the presence of a different DNA-binding mode. It might well be that these less-mobile bands arise from a complex in which the peptide moiety is not specifically inserted in the groove but electrostatically bound to the phosphate surface. The association constant of this complex is at least three times lower than that of the complex with the dsDNA containing the cognate binding site. As one would expect, incubation of **4** with a dsDNA (CRE^{hs}) lacking the AT-rich binding region characteristic of distamycin derivatives does not produce detectable retardation bands (Figure 2C, lanes 10–15). Neither were retention bands observed when CRE^{hs}/T was incubated with control peptide **7**, which lacks the tripyrrole moiety (Figure 2C, lanes 3–9).

In the absence of a nonspecific DNA competitor, we calculated an apparent dissociation constant of $35 \pm 2 \times 10^{-9}$ M at 4 °C for the complex **4**-CRE^{hs}/T (see Supporting Information).

Addition of the cognate 20-base-pair duplex oligonucleotide CRE^{hs}/T to **4** produced a significant variation in the negative intensity of the CD signal at 222 nm. Such an observation is consistent with significant α-helix formation and specific binding. There was also a positive ellipticity increase at 330 nm, which must be due to the tripyrrole moiety binding in the DNA minor groove (Figure 3). Remarkably, addition of compound **4** to the BR-mutated ds-oligonucleotide *m*CRE^{hs}/T caused a weaker, but still relevant, increase in the negative intensity of the band at 222 nm. PAGE experiments suggest that this pep-

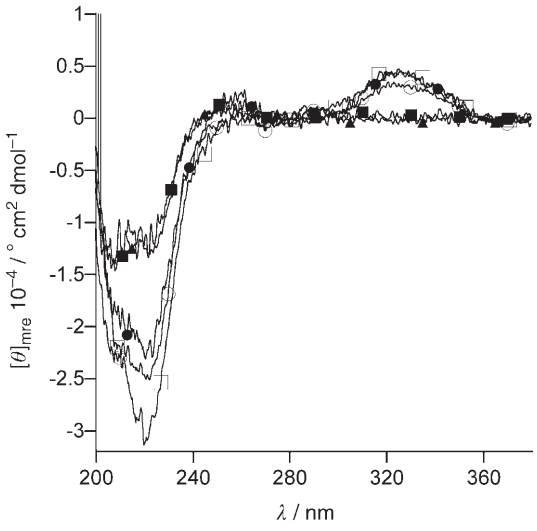


Figure 3. CD difference spectra of peptide **7** in the presence of CRE^{hs}/T (■) and of hybrid **4** in the presence or absence of ds-oligonucleotides: in the absence of DNA (▲), in the presence CRE^{hs}/T (□), in the presence *m*CRE^{hs}/T (○), in the presence of *m'*CRE^{hs}/T (●). CD spectra were obtained at 4 °C, as described in the Experimental Section (see Supporting Information) and were slightly smoothed to facilitate viewing. The difference spectra are the spectra of the ligand + DNA mixture, minus that of the DNA.

tide–DNA complex is different from that obtained with the cognate DNA, presumably because the basic region of the peptide is not specifically inserted in the DNA major groove. Therefore, it can be inferred that, in this case, even a nonspecific DNA site promotes a substantial increase in helicity.[11] We also observed significant spectral changes at 222 nm on using *m*′CRE^{hs}/T that has two mutations at the consensus peptide-binding site. Interestingly, peptide **7**, which lacks the tripyrrole moiety and the linking chain, fails to induce the α-helix transition upon addition of CRE^{hs}/T.

Therefore, in contrast with our previous constructs, in which the tripyrrole was attached to amino acid 245 of the basic region, this peptidic region of the designed hybrid seems to exhibit a great intrinsic tendency to partially fold into an α-helix upon contact with different dsDNAs.

In conclusion, we have demonstrated that appropriate linking of the basic region of a bZIP transcription factor to a Distamycin-like tripyrrole peptide produces a hybrid capable of high-affinity recognition of programmed dsDNA sequences, with a binding mode reminiscent of that for transcription factor Skn-1. Studies to further characterize the binding mode, improve the specificity of the designed conjugates, and use this DNA-recognition strategy to address the natural sequence of Skn-1 are underway.

Acknowledgements

This work was supported by the Spanish Ministry of Education and Science and the ERDF (SAF2004-01044). J.B.B. thanks the Spanish M.E.C. for a predoctoral fellowship, and M.E.V. thanks the Human Science Frontier Program. We are also grateful to Profs. J. Benavente and J. Martínez Costas for their invaluable help.

Keywords: bZIP proteins · conjugation · distamycin · DNA recognition · peptides

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

[2] a) N. M. Luscombe, S. E. Austin, H. M. Berman, J. M. Thornton, *Genome-Biology* **2000**, *1*, 1; b) J. C. Hu, R. T. Sauer, *Nucleic Acids Mol. Biol.* **1992**, *6*, 82 – 101; c) H. C. Hurst, *Protein Profile* **1995**, *2*, 101 – 168; d) T. Ellemberger, *Curr. Opin. Struct. Biol.* **1994**, *4*, 12 – 21.

[3] 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.

[4] 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; c) D. Stanojevic, G. L. Verdine, *Nat. Struct. Biol.* **1995**, *2*, 450 – 457.

[5] a) T. K. Blackwell, B. Bowerman, J. R. Priess, H. Weintraub, *Science* **1994**, *266*, 621 – 628; b) S. Pal, M.-C. Lo, D. Schmidt, I. Pelczer, S. Thurber, S. Walker, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5556 – 5561.

[6] a) M.-C. Lo, S. Ha, I. Pelczer, S. Pal, S. Walter, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8455 – 8460; b) P. B. Rupert, G. W. Daughdrill, B. Bowerman, B. Mattews, *Nat. Struct. Biol.* **1998**, *5*, 484 – 491.

[7] 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.

[8] T. E. Ellemberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **1992**, *71*, 1223 – 1237.

[9] M. Coll, C. A. Frederick, A. H. J. Wang, A. Rich, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385 – 8389.

[10] M. E. Vázquez, A. Caamaño, L. Castedo, D. Gramberg, J. L. Mascareñas, *Tetrahedron Lett.* **1999**, *40*, 3625 – 3628.

[11] There are several reports that indicate that addition of nonspecific dsDNA can also induce an increase in the helicity of the GCN4 basic region, see: J. J. Hollenbeck, M. G. Oakley, *Biochemistry* **2000**, *39*, 6380 – 6389, and references therein.

Received: July 12, 2005
Published online on October 17, 2005

