Design and Synthesis of a Peptide That Binds Specific DNA Sequences through Simultaneous Interaction in the Major and in the Minor Groove**

M. Eugenio Vázquez, Ana M. Caamaño, José Martínez-Costas, Luis Castedo, and José L. Mascareñas*

Most transcription factors (TFs) bind specific DNA sequences using dimeric or multimer motifs, with more or less equal binding energy contributions from each domain and usually strong cooperativity between them.[1] One of the most remarkable examples of bivalent DNA recognition is provided by the bZIP family of TFs, which bind DNA as leucine zipper mediated homo- or heterodimers, with the N-terminal basic region (BR) of each monomer inserting into adjacent DNA major grooves.[2] It has been shown that the leucine zipper unit can be replaced by other noncovalent dimerizing units, or by artificial covalent linkers, without significantly compromising the recognition properties of the system.[3, 4] However, monovalent bZIP BRs exhibit very low DNA binding affinities.[3a, 5] It was recently demonstrated that covalent attachment of one basic region of a bZIP protein (GCN4) to DNA at an appropriate position allows sequence-specific binding.[6] This results suggests that monomeric bZIP-basic region peptides might be able to fold and bind DNA specifically, provided they could be properly delivered to their cognate recognition site. On these grounds we envisaged that suitable tethering of a bZIP-BR domain to a molecule that binds with moderate-to-good affinity in the minor groove of an adjacent sequence might provide for specific binding of the peptide to its cognate major groove site (Figure 1). Most importantly, the resulting conjugate should exhibit higher affinity for its target hybrid DNA site than its individual components have for their target sites.[7] Herein we demonstrate the validity of this approach by reporting the first designed peptide derivative capable of binding to DNA with nanomolar affinity through a bivalent, major–minor groove recognition motif.[8]

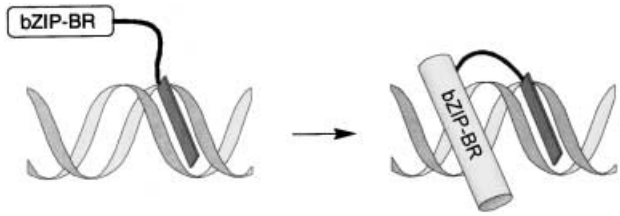


Figure 1. Strategy for DNA recognition.

For minor-groove recognition we used distamycin A,[9] a well known tripyrrole antibiotic that binds DNA preferentially at A–T rich sites. For the major-groove counterpart we chose the basic region of GCN4, a bZIP protein which binds specifically to the cAMP response element site (CRE: 5′-ATGAcgTCAT-3′). From the X-ray structures of the DNA complexes of GCN4[10] and distamycin A,[11] we built an hypothetical model for simultaneous interaction of the GCN4 BR and distamycin with adjacent DNA sites (Figure 2).[12] On

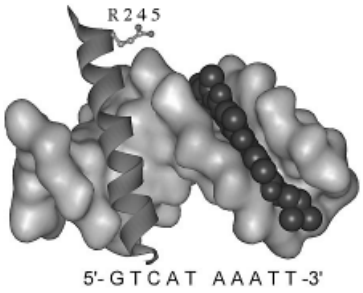


Figure 2. Model for the simultaneous interaction of the GCN4 basic region and distamycin at adjacent sites with DNA

the basis of this model we designed the hybrid **1** (Scheme 1), in which the tether connects the nitrogen atom of the N-terminal pyrrole of a distamycin analogue with the side chain of a glutamic acid. The latter replaces the Arg245 residue of the natural protein. The tether was chosen to span the required distance between the attachment points while allowing for crossover of the phosphate backbone, a process that may be facilitated by the presence of the secondary amine in the chain.

The key step for the synthesis of this hybrid consisted of the coupling of the tripyrrole derivative **2** with the peptide while it was still attached to the resin and fully protected except at the Glu245 residue (Scheme 1).[13]

The DNA binding properties of hybrid **1** were first assessed by circular dichroism, a spectroscopic technique particularly useful for studying DNA–b-ZIP interactions because of the well-established folding of the b-ZIP BRs from random coil to an α-helix upon specific DNA binding.[4] The CD spectrum of compound **1** in the absence of DNA is similar to that of peptide **3** (Figure 3a), which lacks the tripyrrole tether and hence can be used as a control in the binding studies. Unfortunately, addition of an 18 base pair (bp) duplex containing the designated hybrid DNA sequence (T/CRE^{hs})[14] to **1** at 4 °C, did not produce a significant variation of the CD signal at 222 nm, which indicates that the basic region is not inserting into the groove. The considerable

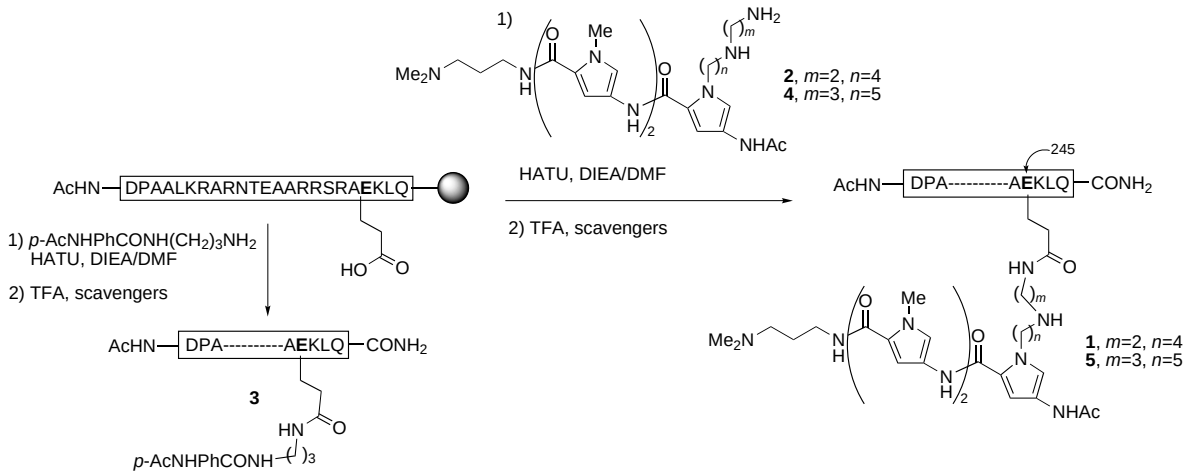
[*] Prof. Dr. J. L. Mascareñas, M. E. Vázquez, A. M. Caamaño, Prof. Dr. L. Castedo
Departamento de Química Orgánica
y Unidad Asociada al CSIC
Universidad de Santiago de Compostela
15782 Santiago de Compostela (Spain)
Fax: (+34) 981-595-012
E-mail: qojoselm@usc.es

Dr. J. Martínez-Costas
Departamento de Bioquímica y Biología Molecular
Universidad de Santiago de Compostela
15782 Santiago de Compostela (Spain)

[**] This work was supported by the Spanish M.E.C. (PB97-0524) and the Xunta de Galicia (PGIDT00PXI20912PR). M.E.V. and A.M.C. thank the Xunta de Galicia and the University of Santiago for their predoctoral fellowships. We are very grateful to Prof. G. L. Verdine for his input and support in the early phases of this work. We also thank Prof. J. Benavente for allowing us to use the radioactivity facilities and Prof. C. Abell for critical reading of the manuscript.

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.

  1433-7851/01/4024-4723 $ 17.50+.50/0



Scheme 1. Synthesis of the hybrid peptides. DIEA = diisopropylethylamine, HATU = *N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4.5-*b*]pyridin-1-ylmethylene]*N*-methylmethanaminium hexafluorophosphate oxide, DMF = *N*,*N*-dimethylformamide, TFA = trifluoroacetic acid.

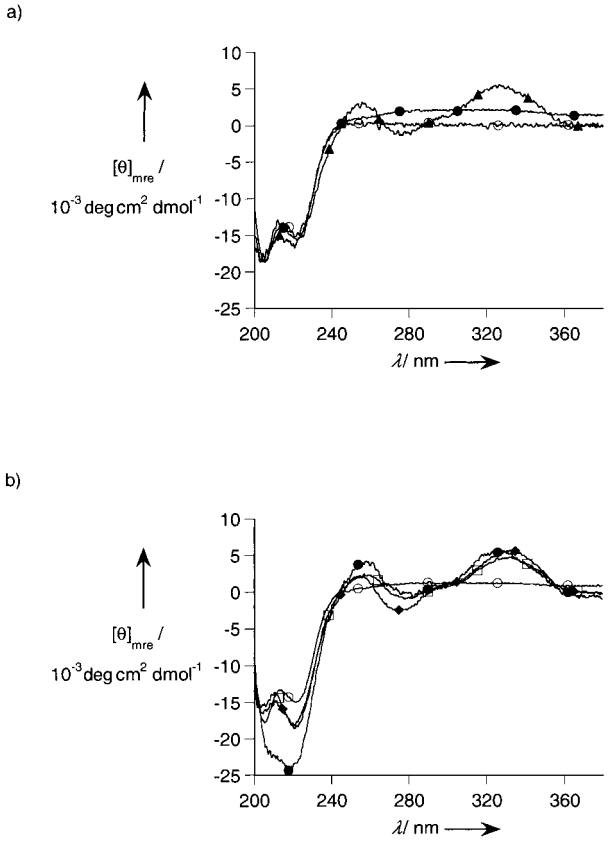


Figure 3. CD spectra of the peptides **1**, **3**, and **5** in the presence or absence of ds-oligonucleotides: a) **3** in the absence of DNA (○), **1** in the absence of DNA (●), **1** in the presence T/CRE^{hs} (▲); b) Spectra of **5** in the absence of DNA (○), in the presence of T/CRE^{hs} (●), in the presence of TcgCRE^{hs} (□), in the presence of T/CRE^{hs}*m* (◆). CD measurements were performed at 4 °C on a JASCO J-715 in a 2-mm cell. Samples contained 10 mM phosphate buffer (pH 7.5), 100 mM NaCl, 5 μM peptide, and 5 μM ds-oligo when present. Spectra of the peptides in the presence of DNA were calculated as the difference between the spectra of the peptide–DNA mixture and the spectrum of free DNA. ds-Oligonucleotides used (the BR subsite (CRE^{hs}) is underlined and the tripyrrole subsite (T) is in italics):

T/CRE^{hs}: 5′-d(CGAACGTCAT*AAAAT*CCT)-3′
3′-d(GCTTGCAGTATTTTAGGA)-5′
TcgCRE^{hs}: 5′-d(CGAACGTCATCG*AAATT*CCT)-3′
3′-d(GCTTGCAGTAGCTTTAAGGA)-5′
T/CRE^{hs}*m*: 5′-d(CGAACGTCGT*AAAAT*CCT)-3′
3′-d(GCTTGCAGCATTTTAGGA)-5′

positive signal at about 330 nm suggests that the tripyrrole does bind at the A–T rich site.[15] Indeed, a similar band was observed upon addition of the same double-stranded(ds) DNA (T/CRE^{hs}) to the tripyrrole **2**.[16] As expected, addition of T/CRE^{hs} to peptide **3** does not induce any change in its CD spectrum (results omitted in Figure 3a for clarity).

We realized from the computer model of the complex that the two parts of **1** are relatively tightly tethered and this might impose some geometric restriction on the required simultaneous binding of both segments. Therefore we synthesized the hybrid **5**, which features a slightly longer connector.[17] Addition of T/CRE^{hs} to peptide **5** induced a remarkable increase in the magnitude of the signal at 222 nm, which indicates that, in this case, the BR peptidic domain is folding and binding to the DNA (Figure 3b). A band at 330 nm in the CD spectrum corresponding to the insertion of the oligopyrrole into the minor groove is also clearly seen. As expected no significant increase in ellipticity was observed upon incubation of **5** with TcgCRE^{hs}, a ds-oligonucleotide including a bp spacing between the binding sites, nor with CRE^{hs} which features a bp mismatch at the BR peptide binding subsite (T/CRE^{hs}*m*).

With these data at hand, we further assessed the DNA-binding properties of the synthetic hybrid peptides by using gel mobility shift analysis, by titrating ^{32}P-end-labeled dsDNAs with increasing concentrations of the peptides. As illustrated in Figure 4, peptide **5** did bind to T/CRE^{hs} with nanomolar affinity, albeit producing two bands, the faster migrating one being slightly more intense. Incubation of peptide **5** with the dsDNA mutated at the BR binding site (T/CRE^{hs}*m*) gave a band with mobility similar to that of the slower migrating one in the preceding experiment. As expected, we could not detect any peptide–DNA complex at concentrations below 100 nM in the presence of a dsDNA containing exclusively the BR peptide binding site (CRE^{hs}). These results together with the CD data indicate that the more intense, faster migrating band obtained upon incubation of **5** with the designated dsDNA corresponds to the expected compact major–minor mode of binding, while the other less intense band results from a complex in which the oligopyrrole is inserted into the groove but the BR peptide

  1433-7851/01/4024-4724 $ 17.50+.50/0

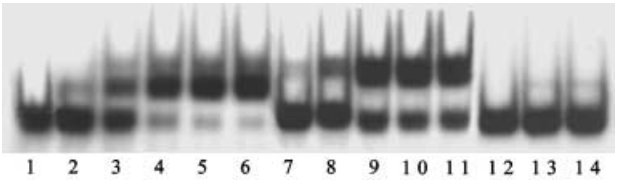


Figure 4. Autoradiograms showing the binding of hybrid **5** to ^{32}P-labeled DNAs. Lanes 1–6: T/CRE^{hs}, **5**: 0, 7.7, 19, 38, 58, 77 nM respectively; Lanes 7–11: T/CRE^{hs}*m*, **5**: 7.7, 19, 38, 58, 77 nM respectively; lanes 12–14: CRE^{hs}, **5**: 38, 77, 154 nM respectively. Binding reactions were performed over 10 min at 4 °C using < 1 nM labeled DNAs in a binding mixture (20 μL) containing 20 mM tris(hydroxymethyl)aminomethane (pH 7.5), 100 mM KCl, 2 mM $MgCl_2$, 2 mM ethylenediaminetetraacetate, 10% glycerol, 0.3 mg mL^{-1} *N,O*-bovine serum albumin (BSA), and 2% NP-40. The products were resolved by polyacrylamide gel electrophoresis using a 10% nondenaturing acrylamide gel and 0.5X TBE (25 mM tris borate and 0.5 mM EDTA) buffer.
CRE^{hs}: 5′-d(CGAC<u>GTCAT</u>CGGAGGTCCT)-3′
3′-d(GCTG<u>CAGTA</u>GCCTCCAGGA)-5′

makes non-specific electrostatic contacts to the phosphate groups.[18]

In conclusion, appropriate linking of a b-ZIP basic region to a minor groove binding tripyrrole allowed for specific binding to its cognate DNA site. The hybrid compound **5** shows considerably higher affinity for its designated target DNA sequence than that of its isolated components for their respective cognate subsites. Although further refinement of the design is necessary to obtain compounds with higher affinities and better specificities, the work described herein confirms the viability of this new type of major–minor groove DNA-binding molecules.

Received: July 23, 2001 [Z17576]

[1] For reviews on DNA-binding proteins, see: a) C. O. Pabo, R. T. Sauer, *Annu. Rev. Biochem.* **1992**, *61*, 1053–1095; b) S. K. Burley, *Curr. Opin. Struct. Biol.* **1994**, *4*, 3–11; c) B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, *Molecular Biology of the Cell*, Garland Publishing, New York, **1994**, chap. 9; d) C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York, **1999**, chap. 8–10; e) D. S. Latchman, *Eukaryotic Transcription Factors*, Academic Press, London, **1998**, chap. 8; f) N. M. Luscombe, S. E. Austin, H. M. Berman, J. M. Thornton, *Genome Biology* **2000**, *1*, 1–37.
[2] H. C. Hurst, *Protein Profile* **1996**, *3*, 1–72.
[3] 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) T. Morii, M. Shimomura, M. Morimoto, I. Saito, *J. Am. Chem. Soc.* **1993**, *115*, 1150–1151; d) T. Morii, Y. Saimei, M. Okagami, K. Makino Y. Sugiura, *J. Am. Chem. Soc.* **1997**, *119*, 3649–3655; e) Y. Aizawa, Y. Sugiura, T. Morii, *Biochemistry* **1999**, *38*, 1626–1632.
[4] The use of an azobenzene as a dimerizing unit allows the DNA affinity of the peptides to be photomodulated: 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.
[5] Recently Zondlo and Shepparzt have developed a smart strategy for achieving monovalent DNA recognition which is based on grafting the DNA-binding residues into the helix of an *aPP* protein: 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.
[6] D. Stanojevic, G. L. Verdine, *Nat. Struct. Biol.* **1995**, *2*, 450–457.
[7] This system would somewhat mimick the mode of recognition of a number of transcription factors that bind DNA by interacting with both the major and minor grooves. See for example: a) J.-A. Feng, R. C. Johnson, R. E. Dickerson, *Science* **1994**, *263*, 348–355; b) A. Draganescu, J. R. Levin, T. D. Tullius, *J. Mol. Biol.* **1995**, *250*, 595–608; c) G. Meinke, P. B. Sigler, *Nat. Struct. Biol.* **1999**, *6*, 471–477.
[8] A major–minor groove DNA-binding ligand made of a hairpin polyamide and a 11-mer oligonucleotide has been previously described: J. W. Szewczyk, E. E. Baird, P. B. Dervan, *Angew. Chem.* **1996**, *108*, 1596–1598; *Angew. Chem. Int. Ed.* **1996**, *35*, 1487–1489.
[9] C. Zimmer, U. Wahnert, *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31–112.
[10] T. E. Ellemberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **1992**, *71*, 1223–1237.
[11] M. Coll, C. A. Frederick, A. H. J. Wang, A. Rich, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385–8389.
[12] The model was constructed by superimposing appropriate base pairs, by using the INSIGHTII software package on a SGI workstation.
[13] The synthesis of **2** and **1** have been previously communicated: M. E. Vázquez, A. M. Caamaño, L. Castedo, D. Gramberg, J. L. Mascareñas, *Tetrahedron Lett.* **1999**, *40*, 3625–3629.
[14] The target site of this sequence does not match exactly with that of the model, but it was used on the basis of the reported preferred (1:1) recognition sites for distamycin, see: D. Rentzeperis, L. A. Marky, T. J. Dwyer, B. H. Geierstanger, J. G. Pelton, D. A. Wemmer, *Biochemistry* **1995**, *34*, 2937. See also reference [15].
[15] F.-M. Chen, F. Sha, *Biochemistry* **1998**, *37*, 11143–11151.
[16] We have calculated an approximate K_d value of 0.8 ± 0.5 μM by CD titration of compound **2** with this dsDNA.
[17] The synthesis of **5** was carried out by using similar procedures as for the synthesis of **1**, and is described in the Supporting Information. Electrospray mass spectrometry: *m/z* [M+H$^+$] found: 3271.4; calcd: 3270.8.
[18] A related mode of interaction has been previously observed with microgonotropens, see: D. Sengupta, A. Blaskó, T. C. Bruice, *Biorg. Med. Chem.* **1996**, *4*, 803–813, and references therein.

Solid-Phase Capture–Release Strategy Applied to Oligosaccharide Synthesis on a Soluble Polymer Support**

Hiromune Ando, Shino Manabe, Yoshiaki Nakahara, and Yukishige Ito*

Advances in the techniques for oligosaccharide synthesis have lagged behind those for other classes of biological oligomers. For the preparation of oligopeptides and oligo-

[*] Dr. Y. Ito, Dr. H. Ando, Dr. S. Manabe
RIKEN (The Institute of Physical and Chemical Research)
and CREST, Japan Science and Technology Corporation (JST)
2–1 Hirosawa, Wako-shi, Saitama 351-0198 (Japan)
Fax: (+81) 48-462-4680
E-mail: yukito@postman.riken.go.jp
Prof. Dr. Y. Nakahara
Department of Applied Biochemistry, Tokai University
and CREST, JST
Kitakaname 1117, Hiratsuka-shi, Kanagawa 259-1292 (Japan)

[**] This work was supported in part by the Mizutani Foundation for Glycoscience (grant to Y.I.), the RIKEN Presidential Fund (grant to S.M.), and the Ministry of Education, Culture, Sports, Science and Technology (Grant-in-Aid for Scientific Research to S.M.). We thank Dr. Teiji Chihara and his staff for the elemental analyses, Ms. Yasuko Tanaka for her contribution at the exploratory stage of this work, and Ms. Akemi Takahashi for her technical assistance.

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.