View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repositorio Institucional da Universidade de Santiago de Compostela

CHEMBIOCHEM
COMMUNICATIONS

ChemPubSoc Europe

NOTICE: This is the peer reviewed version of the following article: Mosquera, J., Rodríguez, J., Vázquez, M. E. and Mascareñas, J. L. (2014), Selective DNA-Binding by Designed Bisbenzamidine-Homeodomain Chimeras. ChemBioChem, 15, 1092-1095 [doi: 10.1002/cbic.201400079]. This article may be used for non-commercial purposes in accordance with Wiley "Terms and Conditions for self-archiving".

DOI: 10.1002/cbic.201((will be completed by the editorial staff))

Selective DNA-binding by designed bisbenzamidine-homeodomain chimeras

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

Dedication ((optional))

Abstract: We report the construction of conjugates between several variants of the helix 3 region of a Q50K engrailed homeodomain and bisbenzamidine minor groove DNA binders. While the hybrid featuring the sequence of the native protein fails to bind to DNA, designed modifications that increase the alpha-helical folding propensity of the peptide allowed specific DNA binding through a bipartite major-minor groove interaction.

Transcription Factors (TFs) are specialized proteins that regulate gene expression. These proteins recognize specific DNA sequences through specialized DNA-binding domains,[1] and upon binding to these sites they promote -or inhibit- the process of transcription.[2] As a result of this key role in the regulation of gene expression, it is not surprising that alterations in the activity of TFs are at the origin of many diseases, including cancer.[3] In this context, the development of non-natural agents that can mimic the dsDNA recognition properties of TFs, and potentially lead to new gene targeting tools, remains a major goal in biological chemistry.[4] TFs establish specific interactions with their DNA target sequences through relatively few residues located in their recognition motifs, in many cases alpha helices inserted in the DNA major groove. However, these motifs, when isolated from the rest of the protein, fail to show any significant DNA binding ability,[5] and therefore the preparation of small peptide models of DNA-binding proteins is a highly challenging task.[6] In recent years we have shown that the DNA binding ability of some of these regions can be recovered when appropriate tethered to minor groove binders like distamycin or propamidine, so that the resulting conjugates bind designed composite sites through a bivalent interaction.[7] This DNA binding strategy has been successfully applied in the case of the basic regions of bZIP proteins like GCN4,[7] as well as for the zinc finger module of GAGA.[8]

Homedomain proteins (HD) are highly conserved transcription factors that play key roles in eukaryotic development and work by interacting to DNA through a bipartite structure consisting of a helix-turn-helix motif and a minor groove binding N-terminal arm.[9] The major groove DNA contacts are mainly accomplished by the helix 3 of the protein (h3), which, as expected, when isolated from the rest of the protein fails to fold into an alpha-helix structure and bind to its cognate DNA.[10] Considering our previous designs with other DNA binding proteins, we wondered whether tethering of this h3 region of a homedomain to a minor groove binder might suffice for recovering a reasonable DNA interaction. In addition of the intrinsic interest of answering this question, the resulting conjugates are quite appealing as they might be considered as artificial mimics of natural homedomains because of the bipartite minor-major groove interaction.

As shown below, hybrids between a bisbenzamidine and a short DNA recognition region of the engrailed homeodomain fail to bind DNA. However, we demonstrate that a rational grafting of the peptide moiety allows for the recovery of specific binding to the designed DNA site.

The strategy was investigated using the Q50K engrailed homeodomain (En-HD), an archetypical member of this family that binds with good affinity to the QRE site (GGATTA).[11] Preliminary molecular modeling studies based on the crystal structure of En-HD bound to its target DNA,[12] suggested that a peptide consisting of residues Asn41 to Ser59 of the helix 3 of the protein could be used as reference for building our hybrids. Therefore we made the conjugate **En-HDh3-1** in which the Gln44 was mutated to a lysine to allow the connection of the minor groove binder. We have previously shown that bisbenzamidines are readily available minor groove binders that show micromolar affinities for A/T-rich segments of DNA,[13] and can be easily modified to introduce the linkers required for conjugation to the peptide.[14]

J. Mosquera, J. Rodríguez, Prof. M. E. Vázquez, Prof. J. L. Mascareñas
Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Departamento de Química Orgánica. Universidade de Santiago de Compostela.
15782 Santiago de Compostela, Spain
eugenio.vazquez@usc.es, joseluis.mascarenas@usc.es

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.20xxxxxxx.

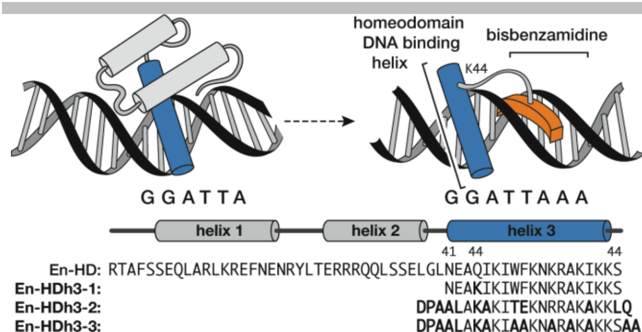
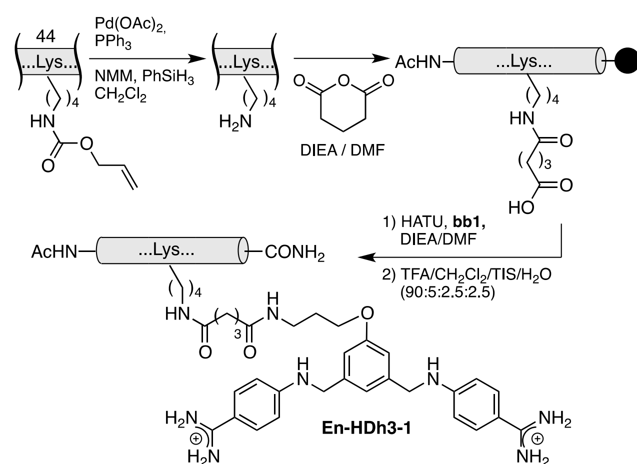


Figure 1. Left: Representation of an En-HD/DNA complex. Right: Hypothetical model of a DNA complex with a bisbenzamidine/En-DHh conjugate. Bottom: Peptide sequence of the Q44K engrailed helix 3 (En-HD) and of the conjugates synthesized in this article [Lys44, which is used to attach the minor groove binder to the peptide (darker cylinder) is highlighted].

Detailed procedures for the synthesis of the hybrid **En-HDh3-1** are given in the experimental section. In short, the peptide scaffolds were assembled following standard Fmoc/tBu solid phase peptide synthesis protocols.[15] The Lys44 residue was introduced in the sequences with its side chain protected as an alloc carbamate, which could be orthogonally removed with Pd catalysis while the peptides were still attached to the solid support.[16] A subsequent derivatization with glutaric anhydride allowed to increase the length of the linker while simultaneously installing a carboxylic acid for attaching the bisbenzamidine amine derivative **bb1** (see the Supporting Information). Coupling of the required amino-bisbenzamidine to the peptide was performed by activating the glutaric carboxylate with HATU (Scheme 1). The final conjugate was deprotected and liberated from the solid support by treatment with TFA, and purified using reverse-phase HPLC.



Scheme 1. Key steps in the synthesis of the conjugate **En-HDh3-1** based on Q50K (Asn41-Ser59, the key Lys44 required for derivatization is indicated).

The interaction of **En-HDh3-1** with the DNA was studied by electrophoretic mobility shift assays (EMSA) under non-denaturing conditions in polyacrylamide gels,[17] and using double stranded oligonucleotides containing the target composite DNA binding site GGATTAAA (**QRE-A/T**). Unfortunately, the gels did not show the expected retarded bands indicating the formation of stable **En-HDh3-1**/DNA complexes (Figure 2A). In consonance with this failing, circular dichroism (CD) spectroscopy revealed that the addition of the cognate DNA to the conjugate does not promote a significant alpha helical folding (Figure 2B). Therefore, despite the presence of the bisbenzamidine and the relatively basic peptide region of **En-HDh3**, we do not observe a significant DNA interaction. We also made a related conjugate featuring a longer connector between the peptide and the bisbenzamidine, but it also failed to give relevant interactions (**En-HDh3-1b**, see the Supporting Information).

The recognition helix of HD proteins is usually shorter and makes fewer contacts with the DNA than those present in other TFs, such as in the bZIP family, and this might be one of the reasons behind the failing of our derivatives to interact to DNA. The presence of the rest of the homeodomain protein (helix 1 and 2) is critical for promoting the required α-helix folding of h3.

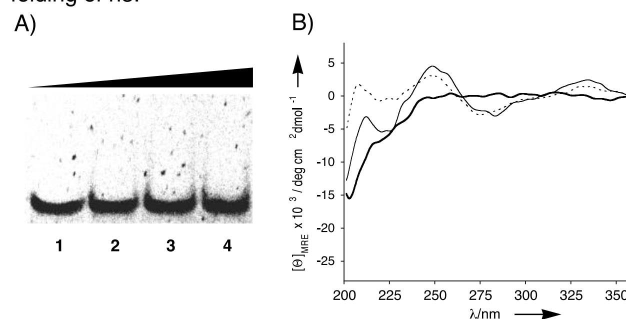
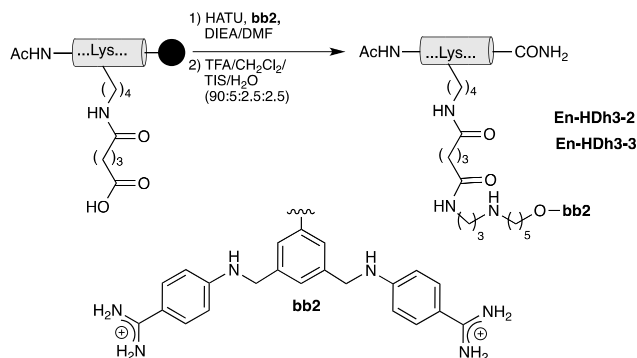


Figure 2. A) EMSA results showing the absence of retarded bands upon mixing conjugate **En-HDh3-1** and **QRE-A/T**. Lanes 1-4: [**HDh3-1**] = 0, 500, 800 and 1000 nM. Experiment was resolved by PAGE on a 10% nondenaturing polyacrylamide gel and 0.5xTBE buffer over 40 min at RT and analyzed by staining with SyBrGold (Molecular Probes: 5 μL in 50 mL of 1xTBE) for 10 min, followed by fluorescence visualization. B) Circular dichroism spectra of a 5 μM solution of **HDh3-1** (black solid line); in the presence of 1 equiv of **QRE-A/T** (solid line); same solution in the presence of 5 μM **MUT-A/T** DNA (dashed line). The CD spectra of the peptides (when measured in the presence of DNA) were calculated as the difference between the spectrum of the peptide/DNA mixture and the measured spectrum of a sample of the DNA oligonucleotide sequences (binding sites in italics, only one strand shown): **QRE-A/T**: 5'-CGTGC *GGATTAAA* AGCTGCG-3'; **MUT-A/T:** 5'-CGTGC *GACTTAAA* AGCTGCG-3'.

To compensate for the loss of the nucleating effect provided by the full HD structure, we considered the design of a modified helix based on the bZIP transcription factor GCN4. The idea consisted on grafting key DNA contacting amino acids of En-HD into the basic region of GCN4. We therefore constructed the conjugate **En-HDh3-2**, in which the peptide is a combination of the h3 of engrailed and the amino acids 226-248 of GCN4.[18] We also designed a second peptidic domain in which the residues not involved in DNA binding were replaced by alanines,[19] while keeping the short N-terminal capping motif of GCN4 (DPAAL, hybrid **En-HDh3-3**). In both cases the tether between the bisbenzamidine and the peptide includes a secondary amine in the linker, which could favor the phosphate backbone crossover.[20] The synthesis of the conjugates was achieved using a strategy similar to that previously described, but with the required aminobisbenzamidine partner for the coupling reaction (**bb2**, Scheme 2, and Supporting Information).



Scheme 2. Key steps in the synthesis of the conjugates **En-HDh3-2** and **En-HDh3-3** based on engrailed homeodomain (only the key Lys44 required for the coupling is indicated). **En-HDh3-2** and **EnHDh3-3** differ only in the peptide sequence (see Figure 1 and Supporting Information for details).

Interestingly, and in contrast to the failing of the previous constructs, EMSA experiments with conjugates **En-HDh3-2** and **En-HDh3-3** and the ds-oligonucleotide **QRE-A/T** showed the presence of a new slower-migrating band (band b), consistent with the formation of specific complexes between the peptide conjugates and the dsDNA (Figure 3A and 3B, top panels). Isothermal titrations allowed to calculate a K_D of ≈ 188 nM at 4 ºC for **En-HDh3-2** and a K_D of ≈ 131 nM for **En-HDh3-3**, which denotes a quite reasonable interaction. Importantly, incubation of the hybrids with an oligonucleotide containing a double mutation in its peptide-binding site (**MUT-A/T**) does not elicit stable complexes (Figure 3A and 3B, bottom panels), which confirms a highly selective interaction with the target DNA.[21]

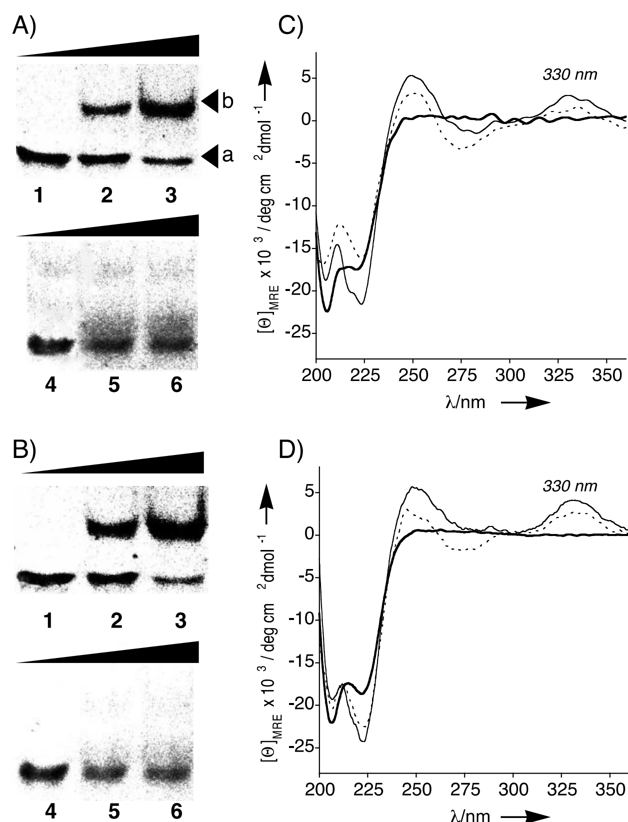


Figure 3. A) EMSA results for conjugate **En-HDh3-2**. Lanes 1-3: [**En-HDh3-2**] = 0, 200, 500 nM with 50 nM of **QRE-A/T** dsDNA. Lanes 4-6: [**En-HDh3-2**] = 0, 750, 1000 with 50 nM of **MUT-A/T** dsDNA B) EMSA results for conjugate **En-HDh3-3**. Lanes 1-3: [**En-HDh3-3**] = 0, 200, 500 nM with 50 nM of **QRE-A/T** dsDNA. Lanes 4-6: [**En-HDh3-3**] = 0, 750, 1000 with 50 nM of **MUT-A/T** dsDNA. C) Circular dichroism spectra of a 5 µM solution of **En-HDh3-2** (thick solid line) in the presence of 1 eq. of **QRE-A/T** dsDNA (solid line), and in the presence of 5 µM **MUT-A/T ds**DNA (dashed line). D) Same experiment for conjugate **En-HDh3-3**. Oligonucleotide sequences (binding sites in italics, only one strand shown): **QRE-A/T**: 5'-CGTGC *GGATTAAA* AGCTGCG-3'; **MUT-A/T**: 5'-CGTGC *GACTTAAA* AGCT GCG-3'.

The increased affinity of **En-HDh3-3** is consistent with its higher content in Ala residues, and therefore a greater tendency for alpha-helix folding.[22] Circular dichroism experiments showed that upon addition of the double stranded oligonucleotide containing the target sequence (**QRE-A/T**), both peptides experiment a significant increase in helical content, particularly for **En-HDh3-3**, which is consistent with the higher binding affinity found by EMSA. Curiously, while the addition of the mutated DNA to **En-HDh3-2** does not promote α-helical folding, addition to **En-HDh3-3** led to a notable increase in the negative intensity of the band at 222 nm. This can be interpreted in terms of the higher intrinsic α-helical propensity of the Ala-equipped peptide that favors its partial folding even in the presence of non-specific DNA.[23] On the other hand, the positive CD band centered at 330 nm, which is more intense for the better binder **En-HDh3-3**, must arise from insertion of the bisbenzamidine unit into the DNA minor groove.[24]

In summary, we have shown that by combining an appropriate amino acid grafting with the tethering of a minor groove binder, h3 regions of homedomain proteins can be induced to bind specific DNA sites with good affinity and very good selectivity. The resulting conjugates simulate the DNA recognition of native homedomains by binding to similar sequences by a bipartite major and minor group interaction. This type of constructs might open unique opportunities for interfering with the activity of homedomain transcription factors, and therefore for altering gene processes related to eukaryotic development.

Experimental Section

Electrophoretic Mobility Shift Assay. EMSA was performed with a BioRad Mini Protean gel system, powered by an electrophoresis power supplies PowerPac Basic model, maximum power 150 V, frequency 50.60 Hz at 140 V (constant V). Gel mobility shift assays binding reactions were performed over 50 min. in 18 mM Tris (pH 7.0), 50 mM KCl, 1.2 mM $MgCl_2$, 0.5 mM EDTA, 9% glycerol, 0.11 mg/mL BSA and 4.2% NP−40 at 4 ºC. In the experiments analyzed by fluorescent staining we used 50 nM of the unlabeled dsDNAs and a total incubation volume of 20 µL. Products were resolved by PAGE using a 10% non-denaturing poliacrylamide gel and 0.5X TBE buffer, and analyzed by staining with SyBrGold (Molecular Probes: 5 µL in 50 mL of 1X TBE) for 10 min and visualized by fluorescence.

Circular Dichroism spectroscopy. CD measurements were made in a 2 mm cell at 4 ºC. Samples contained 5 µM of corresponding dsDNA (when present) and 5 µM of peptides in 10 mM phosphate buffer (pH 7.5) and 100 mM of NaCl. The CD spectra of the peptides (when measured in the presence of DNA) were calculated as the difference between the spectrum of the peptide/DNA mixture and the measured spectrum of a sample of the DNA oligonucleotide.

Acknowledgements

We thank the financial support provided by the Spanish SAF2010-20822-C02, CTQ2012-31341, Consolider Ingenio 2010 CSD2007-00006, the Xunta de Galicia, GRC2013/14, the European Regional Development Fund and the European Research Council (Advanced Grant Na 340055). J.M. thanks the Spanish MCINN and J.R. thanks the Fundación Gil Dávila and the Xunta de Galicia for their PhD fellowships.

Keywords: molecular recognition • DNA recognition • peptides • oligonucleotides • supramolecular chemistry

[1] C. W. Garvie, C. Wolberger, *Mol. Cell* **2001**, *8*, 937–946.
[2] a) D. S. Latchman, *Eukaryotic Transcription Factors*, Elsevier, London, 2004; b) N. M. Luscombe, S. E. Austin, H. M. Berman, J. M. Thornton, *Genome Biol.* **2000**, *1*, 1–10.
[3] a) J. E. Darnell, *Nature Rev. Cancer* **2002**, *2*, 740–749; b) A. T. Look, *Science* **1997**, *278*, 1059–1064; c) P. Brennan, R. Donev, S. Hewamana, *Mol. Biosyst.* **2008**, *4*, 909–919.
[4] a) P. P. Pandolfi, *Oncogene* **2001**, *20*, 3116–3127; b) R. Pollock, M. Giel, K. Linher, T. Clackson, *Nature Biotechnol.* **2002**, *20*, 729–733; c) C. Denison, T. Kodadek, *Chem. Biol.* **1998**, 5, R129-45; d) R. Yaghmai, G. Cutting, *Mol Ther* **2002**, *5*, 685–694; e) A. K. Mapp, *Org. Biomol. Chem.* **2003**, *1*, 2217–2220.
[5] a) H. C. Nelson, *Curr. Opin. Genet. Dev.* **1995**, *5*, 180–189; b) R. Moretti, A. Z. Ansari, *Biochimie* **2008**, *90*, 1015–1025; c) D. J. Segal, C. F. Barbas, *Curr. Opin. Chem. Biol.* **2000**, *4*, 34–39.
[6] For the preparation of artificial DNA-binding peptides, see: a) N. J. Zondlo, A. Schepartz, *J. Am. Chem. Soc.* **1999**, *121*, 6938–6939; b) R. L. Stafford, H.-D. Arndt, M. L. Brezinski, A. Z. Ansari, P. B. Dervan, *J. Am. Chem. Soc.* **2007**, *129*, 2660–2668; c) M. E. Vazquez, A. M. Caamaño, J. L. Mascareñas, *Chem. Soc. Rev.* **2003**, *32*, 338–349; d) E. Pazos, J. Mosquera, M. E. Vazquez, J. L. Mascareñas, *ChemBioChem* **2011**, *12*, 1958–1973.
[7] a) J. B. Blanco, V. I. Dodero, M. E. Vázquez, M. Mosquera, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2006**, *45*, 8210–8214; b) J. B. Blanco, M. E. Vázquez, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003**, *10*, 713–722; c) M. E. Vazquez, A. M. Caamaño, L. Castedo, D. Gramberg, J. L. Mascareñas, *Tetrahedron Lett.* **1999**, *40*, 3625–3628; d) M. E. Vazquez, A. M. Caamaño, L. Castedo, J. L. Mascareñas, *Tetrahedron Lett.* **1999**, 40, 3621–3624.
[8] a) M. I. Sánchez, O. Vazquez, M. E. Vázquez, J. L. Mascareñas, *Chem. Eur. J.* **2013**, *19*, 9923–9929; b) O. Vazquez, M. E. Vazquez, J. B. Blanco-Canosa, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2007**, *46*, 6886–6890; J. Rodríguez, J. Mosquera, O. Vázquez, M. E. Vázquez, J. L. Mascareñas, *Chem. Commun.* **2014**, *50*, 2258–2260.
[9] a) 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; b) D. Vuzman, A. Azia, Y. Levy, *J. Mol. Biol.* **2010**, *396*, 674–684; c) S. E. Ades, R. T. Sauer, *Biochemistry* **1995**, *34*, 14601–14608.
[10] L. Guerrero, O. S. Smart, G. A. Woolley, R. K. Allemann, *J. Am. Chem. Soc.* **2005**, *127*, 15624–15629 and references therein.
[11] S. E. Ades, R. T. Sauer, *Biochemistry* **1994**, *33*, 9187–9194.
[12] L. Tucker-Kellogg, M. A. Rould, K. A. Chambers, S. E. Ades, R. T. Sauer, C. O. Pabo, *Structure* **1997**, *5*, 1047–1054; b) E. Fraenkel, M. A. Rould, K. A. Chambers, C. O. Pabo, *J. Mol. Biol.* **1998**, *284*, 351–361. (The structure is deposited in the protein data bank under the PDB ID: 2hdd or 3hdd).
[13] a) M. I. Sánchez, O. Vazquez, J. Martinez-Costas, M. E. Vazquez, J. L. Mascareñas, *Chem. Sci.* **2012**, *3*, 2383–2387. b) O. Vazquez, M. I. Sánchez, J. Martinez-Costas, M. E. Vazquez, J. L. Mascareñas, *Org. Lett.* **2010**, *12*, 216-219.
[14] M. E. Vázquez, A. M. Caamaño, J. Martínez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2001**, *40*, 4723–4725.
[15] N. Sewald, H.-D. Jakubke in Peptides: Chemistry and Biology, 558. Wiley-VCH, Weinheim, **2002**, p. 135.
[16] a) P. Gomez-Martinez, M. Dessolin, F. Guibe, F. Albericio, *J. Chem. Soc., Perkin Trans. 1* **1999**, 2871–2874; b) E. Pazos, D. Torrecilla, M. Vázquez López, L. Castedo, J. L. Mascareñas, A. Vidal, M. E. Vazquez, *J. Am. Chem. Soc.* **2008**, *130*, 9652–9653.
[17] a) L. M. Hellman, M. G. Fried, *Nature Prot.* **2007**, *2*, 1849–1861; b) D. Lane, P. Prentki, M. Chandler, *Microbiol Rev* **1992**, *56*, 509–528.
[18] a) J. J. Hollenbeck, D. L. McClain, M. G. Oakley, *Protein Sci.* **2002**, *11*, 2740–2747; b) J. Mosquera, A. Jiménez-Balsa, V. I. Dodero, M. E. Vázquez, J. L. Mascareñas, *Nature Commun.* **2013**, *4*, 1874; c) R. V. Talanian, J. McKnight, P. S. Kim, *Science* **1990**, *249*, 769–771.
[19] a) S. Marqusee, V. H. Robbins, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5286–5290; b) A. R. Chakrabartty, J. A. Schellman, R. L. Baldwin, *Nature* **1991**, *351*, 586–588; c) C. N. Pace, J. M. Scholtz, *Biophys. J.* **1998**, *75*, 422–427; d) G. H. Bird, A. R. Lajmi, J. A. Shin, *Biopolymers* **2002**, *65*, 10–20..
[20] a) J. B. Blanco, M. E. Vazquez, L. Castedo, J. L. Mascareñas, *ChemBioChem* **2005**, *6*, 2173–2176; b) J. B. Blanco, O. Vazquez, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Chem. Eur. J.* **2005**, *11*, 4171–4178.
[21] The observed smearing of bands in EMSA experiments have been attributed to the existence of multiple low affinity complexes, see for example: E. K. Liebler, U. Diederichsen, *Org. Lett.* **2004**, *6*, 2893–2896.
[22] Prediction of the helical content with the *Agadir* program suggests that h3 in **En-HDh3-2** has a higher tendency to adopt an alpha helical conformation (~18%) than the h3 sequence in **En-HDh3-1** (~12%), or the natural h3 sequence (~2%). a) V. Muñoz, L. Serrano, *Nature Struct. Biol.* **1994**, *1*, 399–409; b) V. Muñoz, L. Serrano, *J. Mol. Biol.* **1994**, *245*, 275–296.
[23] Related structuring effects have been observed with positively charged peptides as a result of low affinity and non-specific interactions with DNA: a) A. Zlotnick, S. L. Brenner, *J. Mol. Biol.* **1988**, *209*, 447–457; b) K. Niikura, H. Matsuno, Y. Okahata, *Chem. Eur. J.* **1999**, *5*, 1609–1616 ; c) N. P. Johnson, J. Lindstrom, W. A. Baase, P. H. von Hippel, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4840–4844.
[24] a) W. D. Wilson, F. A. Tanious, D. Ding, A. Kumar, D. W. Boykin, P. Colson, C. Houssier, C. Bailly, *J. Am. Chem. Soc.* **1998**, *120*, 10310–10321; b) C. Bailly, R. K. Arafa, F. A. Tanious, W. Laine, C. Tardy, A. Lansiaux, P. Colson, D. W. Boykin, W. D. Wilson, *Biochemistry* **2005**, *44*, 1941–1952.

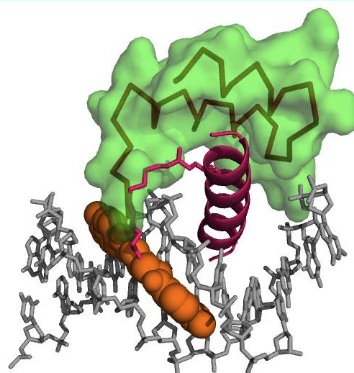
Received: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

Entry for the Table of Contents (Please choose one layout)

Layout 1:

COMMUNICATION

Conjugates between the helix 3 region of a Q50K engrailed homeodomain and bisbenzamidine minor groove DNA binders fail to bind to DNA. However, designed modifications that increase the alpha helical folding propensity of the peptide allowed specific DNA binding through a bipartite major-minor groove interaction.



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

Page No. – Page No.

Selective DNA-binding by designed bisbenzamidine-homeodomain chimeras