ORGANIC LETTERS

2004
Vol. 6, No. 17
2893−2896

From IHF Protein to Design and Synthesis of a Sequence-Specific DNA Bending Peptide

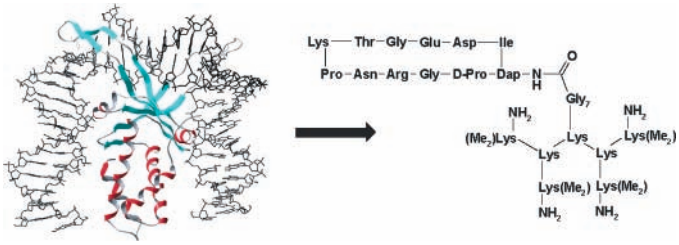
Eduard K. Liebler and Ulf Diederichsen*

Institut für Organische und Biomolekulare Chemie, Universität Göttingen, Tammannstrasse 2, D-37077 Göttingen, Germany

udieder@gwdg.de

Received May 28, 2004

ABSTRACT



The design and synthesis of a small peptide that mimics the integration host factor (IHF), a major nucleoid-associated protein, is reported. IHF induces DNA compaction by sequence-specific binding that leads to significant bending of the DNA double strand. In a modular approach a small L-lysine dendrimer responsible for nonspecific charge–charge interactions was linked to a cyclopeptide. The latter was designed for specific DNA recognition in the minor groove followed by bending of the double strand.

Unlike the well-organized eukaryotic chromatin, the bacterial chromosome is rather disordered. Nevertheless, it is highly compacted so that the almost 1 mm long chromosome of *Escherichia coli* can be stored in a cell of 1 μm diameter. In bacteria several histone-like proteins were identified to play the role of compacting DNA. Some of those protein−DNA complexes such as HU (histone-like protein) and IHF (integration host factor) have been characterized by X-ray structural analysis.[1] Besides formation of chromatin, perhaps the most prominent example of protein-induced DNA bending is the IHF−DNA interaction.

IHF is a small heterodimeric protein that consists of an α- and β-subunit, each of about 10 kDa size and with almost 30% sequence homology.[2] Both subunits are closely related to the homodimeric HU, a nonspecific DNA binding protein that also bends DNA and belongs to the same protein family.[3,4,5] On one hand, IHF acts as an architectural factor in prokaryotic cellular processes, and on the other hand it is important for the transcriptional regulation of approximately 120 *E. coli* genes[6] and initiation of replication at oriC.[7,8,9]

The IHF−DNA cocrystal structure reveals a largely α-helical body with a certain content of basic amino acids such as arginine or lysine, followed by β-sheets that extend into two flexible β-ribbon arms (Figure 1).[10] These arms wind around the DNA exclusively in the minor groove. At the tip of each arm a highly conserved proline residue intercalates between the base pairs (Figure 2). By disrupting stacking interactions between base pairs, an overall bend of the DNA double strand of almost 180° is induced. The bend is stabilized by several interactions with the body of the protein.[2] The negatively charged DNA backbone interacts

(1) Rice, P. A. *Curr. Opin. Struct. Biol.* **1994**, *7*, 86.
(2) Rice, P. A.; Yang, S.; Mizuuchi, K.; Nash, H. A. *Cell* **1996**, *87*, 1295.
(3) Megraw, T. L.; Kao, L. R.; Chae, C. B. *Biochimie* **1994**, *76*, 909.
(4) Bianchi, M. E. *Mol. Microbiol.* **1994**, *14*, 1.
(5) Oberto, J.; Drlica, K.; Rouviere-Yaniv, J. *Biochimie* **1994**, *76*, 901.
(6) Arfin, S. M.; Long, A. D.; Ito, E. T.; Tolleri, L.; Riehle, M.; Paegle, E. S.; Hatfield, G. W. *J. Biol. Chem.* **2000**, *275*, 29672.
(7) Hwang, D. S.; Kronberg, A. *J. Biol. Chem.* **1992**, *267*, 23083.
(8) Polaczek, P.; Kwan, K.; Liberies, D. A.; Campbell, J. L. *Mol. Microbiol.* **1997**, *26*, 261.
(9) Ryan, V. T.; Grimwade, J. E.; Nievera, C. J.; Leonard, A. C. *Mol. Microbiol.* **2002**, *46*, 113.
(10) Ellenberger, T.; Landy, A. *Structure* **1997**, *5*, 153.

10.1021/ol049016a CCC: $27.50
Published on Web 07/22/2004

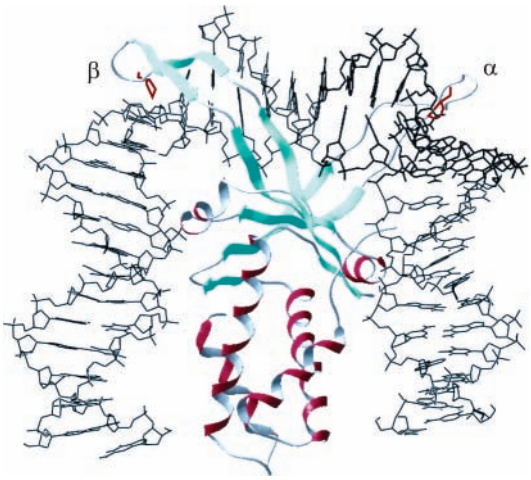


Figure 1. IHF−DNA cocrystal structure. The α-subunit is shown on the right, the *β*-unit on the left. The intercalating proline of each arm is highlighted in red (P65α/P64*β*). P64 is intercalated between base pairs 28 and 29 and P65 between base pairs 37 and 38. Image generated by *MacroModel*. PDB code 1IHF [mmdbId:5137].[2]

with 26 positively charged side chains, as well as with the N-termini of the helices in the body.[2] IHF has no contact to the DNA major groove and forms only a few hydrogen bonds to the minor groove, where all four nucleobases offer hydrogen bond donors or acceptors.[11] Therefore, the sequence specificity can only be explained by indirect readout. Specificity probably results from the sum of many small interactions such as two arginines (Arg^{60} and Arg^{63} of the α-subunit) reaching into the minor groove and forming hydrogen bonds with conserved nucleobases or with the intercalating prolines (Pro^{65} of α- and Pro^{64} of *β*-subunit).[2] The DNA used for the X-ray analysis consists of 35 base pairs and contains the H′-binding site of the *λ* phage *attP* attachment sequence (Figure 2, II), the complex in which IHF was first discovered to be necessary for the integration of the phage genome into the host chromosome of *E. coli*.

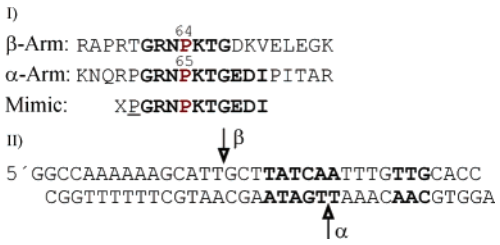


Figure 2. (I) Part of the amino acid sequences of the α- and *β*-subunits. Conserved sequences are highlighted in bold, the intercalating prolines 64 (*β*) and 65 (α) in red. X = diaminopropionic acid (Dap). (II) DNA sequence used for cocrystallization in Figure 1. The consensus sequence for IHF binding is shown in bold; arrows indicate the position of intercalating proline residues for both subunits (P64*β*/P65α).

This is the first report of our investigations to design and synthesize a peptide with specific DNA binding and bending characteristics, which is derived from the IHF−DNA complex. The minor groove binding loop of the IHF α-subunit was used as a lead structure for the design of a peptide that should specifically bind to a DNA double strand and induce bending. The positively charged body of the IHF protein was imitated by a lysine dendrimer, which was linked to the recognition unit by a glycine oligomer (Figure 3). For specific

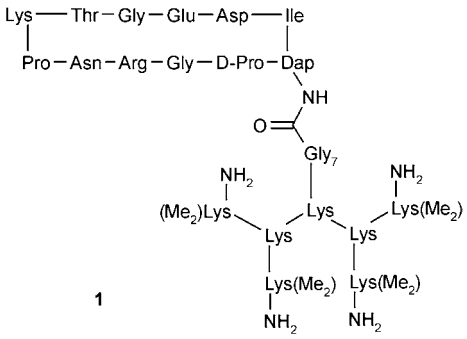


Figure 3. IHF-mimicking target peptide **1**.

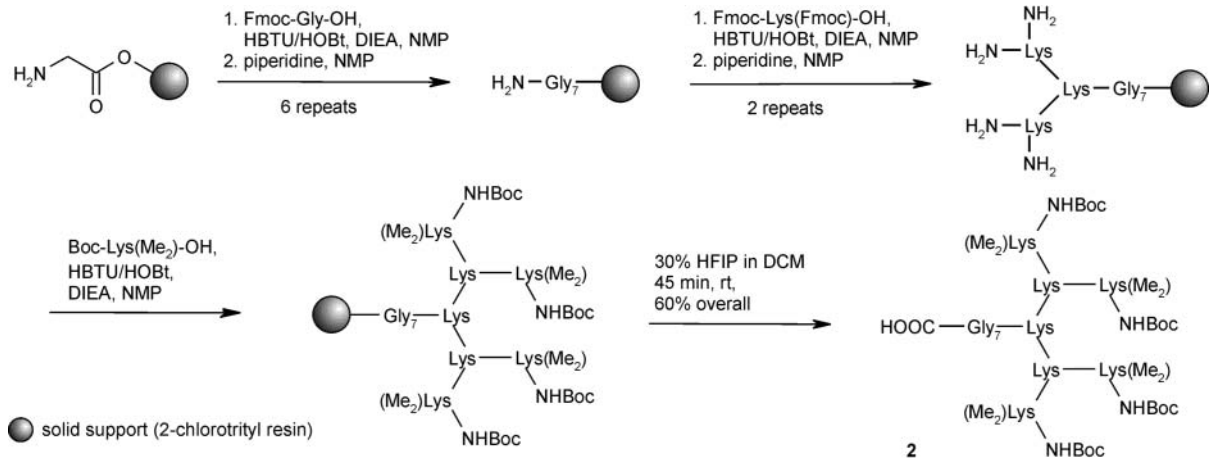
recognition and bending, the amino acid sequence of the IHF α-subunit was kept around the intercalating Pro^{65} and limited to 10 amino acids in order to cover just the contact area in the minor groove ("mimic" in Figure 2, I). Two additional amino acids were introduced for cyclization and attachment of the linker. Cyclization by introduction of a *β*-turn inducing D-proline should at least partially ensure conformational rigidity.[12] Diaminopropionic acid (Dap) was introduced for later attachment of the linker to the side chain amino group. To resemble the nonspecific DNA binding body of the IHF protein by charge neutralization, a three-generation lysine dendrimer **2** was synthesized by SPPS.[13,14] These dendrimers are known to be protonated under physiological conditions. Finally, the lysine dendrimer **2** was covalently attached to the side chain of the Dap in cyclopeptide **3** by a glycine heptapeptide linker. The linker length was set to at least seven glycines using *MacroModel* for estimating the distance between the IHF α-subunit and its body.

For synthesis of the IHF mimetic **1** the lysine dendrimer **2** was established directly continuing the SPPS of the glycine linker. Cyclopeptide **3** was prepared independently and attached to the linker-dendrimer unit **2**.

The lysine dendrimer **2** was prepared on a glycine-preloaded 2-chlorotrityl resin by automated SPPS following the Fmoc protocol. 2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorphosphate (HBTU) was used as

(11) Swinger, K. K.; Rice, P. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 28.
(12) Dittrich, B.; Koritsanszky, T.; Grosche, M.; Scherer, W.; Flaig, R.; Wagner, A.; Krane, H. G.; Kessler, H., Riemer, C.; Schreurs, A. M. M.; Luger, P. *Acta Crystallogr., Sect. B* **2002**, *58*, 721.
(13) Tam, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5409.
(14) Choi, J. S.; Lee, E. J.; Choi, Y. H.; Jeong, Y. J.; Park, J. S. *Bioconjugate Chem.* **1999**, *10*, 62.

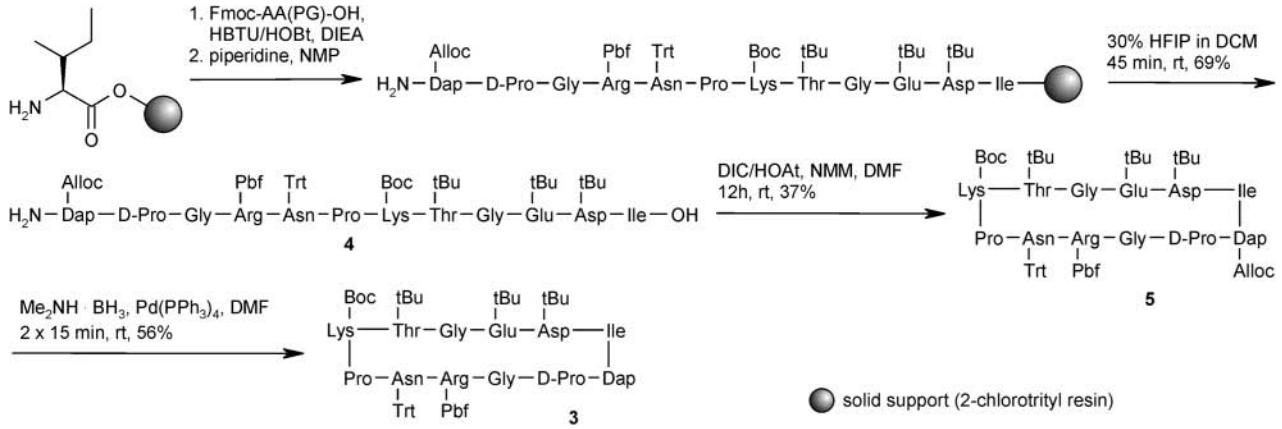
Scheme 1. Convergent Synthesis of Lysine Dendrimer Connected to an Oligoglycine Linker Performed by Automated SPPS



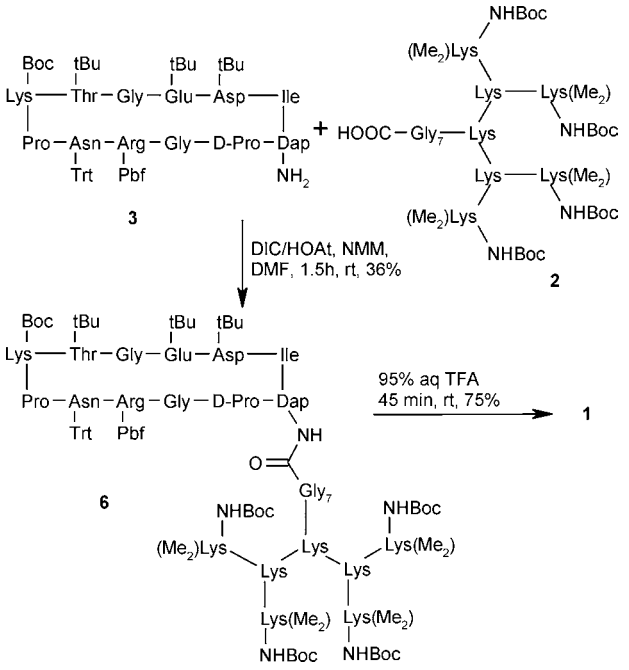
coupling reagent together with 1-hydroxy-benzotriazole (HOBt) (Scheme 1).[15] Fmoc-deprotection was carried out with piperidine in *N*-methylpyrrolidone (NMP). The lysine dendrimer **2** was synthesized using Fmoc-Lys(Fmoc)-OH as monomeric unit up to the second generation followed by Boc-Lys(Me_2)-OH for the third layer yielding [[Boc-Lys-(Me_2)]$_2$Lys]$_2$Lys-$(Gly)_7$-OH (**2**).[16,17] First attempts to generate lysine dendrimers simply using Fmoc-Lys(Fmoc)-OH for all three generations failed because of insolubility of the Fmoc-protected lysine dendrimer after cleavage from the resin. The Boc-protected dendrimer−linker construct **2** proved to be quite soluble in the coupling reaction with cyclopeptide **3**. *N*-Bismethylation still provides protonation at pH 7, and the remaining lysine amino termini should also contribute to phosphate neutralization.[2,18]

Cyclopeptide **3** was first prepared as linear dodecamer on a 2-chlorotrityl resin preloaded with isoleucine based on Fmoc chemistry. HBTU was used as coupling reagent together with HOBt in order to suppress racemization (Scheme 2). The amino acid side chains were Boc (*tert*-butoxycarbonyl)-, *t*Bu (*tert*-butyl)-, Pbf (2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)-, or Trt (trityl)-protected, respectively. Removal of the Fmoc group was accomplished with piperidine. Dap was introduced side chain protected with the Alloc (allyloxycarbonyl) group as an orthogonal protection to Fmoc and the acid-labile permanent side chain protecting groups. Furthermore, D-Pro was chosen to induce a β-turn in the cyclopeptide. After cleavage from the resin with 30% 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) in DCM the fully protected dodecapeptide (**4**) was obtained.[19] Under these conditions it was possible to avoid deprotection of the acid-labile Trt group. The crude linear precursor **4** was cyclized in solution using *N*,*N*'-diisopropylcarbodiimide (DIC) as coupling reagent together with 1-hydroxy-7-azabenzotriazole (HOAt) and *N*-methylmorpholine (NMM) as base in a highly diluted DCM/DMF solvent mixture, yielding protected cyclopeptide **5**. Orthogonal removal of the Dap-Alloc protecting group was provided with $Pd(PPh_3)_4$ and the borane complex $Me_2NH{\cdot}BH_3$ as neutral scavenger.[20] After purification the Alloc-deprotected cyclopeptide **3** was obtained in an overall yield of 56%. Coupling of the cyclopeptide **3** via the Dap side chain to the C-terminal position of the linker-dendrimer **2** was accomplished with DIC/HOAt and NMM activation in dry DMF (Scheme 3).

Scheme 2. Synthesis of Cyclopeptide **3**



Scheme 3. Coupling of Cyclopeptide **3** and Dendrimer-Linker-Construct **2** with Subsequent Side Chain Deprotection



After 2 h of reaction time, cyclopeptide **3** and dendrimer **2** were still detected. Therefore, an additional 20% DIC was added and coupling was continued for 2 h, yielding peptide **6**.

Finally, all side chain protecting groups were cleaved with TFA and target peptide **1** was isolated in 75% yield.

First binding studies of the potential IHF mimic **1** to the DNA double strand containing the λ phage consensus sequence ($^{5'}$TAAAAAAGCATTGCTTATCAA TTTGTTG-CAACGA$^{3'}$) were performed by gel electrophoresis (Figure 4). Addition of the dendrimer-linker-peptide **2** to the DNA double strand led to disappearance of the band that indicates the DNA double strand with increasing amount of peptide (Supporting Information). This is the expected result for nonspecific binding based on electrostatic interactions only creating many different species each with a concentration that is too low to be detected by gel chromatography.

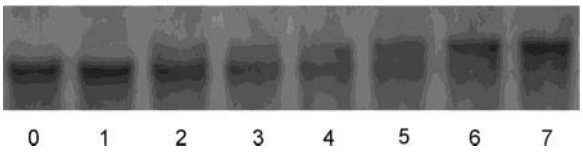


Figure 4. Gel mobility shift assay for the potential IHF-mimetic **1** binding to a 34bp ds-DNA fragment containing the consensus sequence. 0, 8, 16, 80, 96, 112, 128, 160 μM peptide **1** in lanes 0−7, respectively.

In contrast, addition of an increasing concentration of peptide **1** to the consensus DNA led to disappearance of the ds-DNA band; instead a new sharp band with lower mobility appeared (Figure 4). This result is a first indication of specific binding of our potential IHF-mimetic **1** to DNA double strands.

The remarkable bending of DNA double strands specifically induced by the IHF protein was a leading motif for the construction of an artificial peptide with comparable function. The minor groove binding IHF arm was imitated by a cyclododecapeptide. A proline was included in the sequence, which in case of the IHF protein induces the DNA bend by intercalation into the base pair plain. The cyclopeptide was linked to a three-generation lysine dendrimer that should resemble the positively charged body of the IHF protein. The synthesis is modular regarding the combination of possible cyclopeptides, linkers, and dendrimers. Structural modifications are currently under investigation as well as studies to verify our first indication of specific binding and investigate bending properties.

Acknowledgment. Financial support of the Fonds der Chemischen Industrie and of the VolkswagenStiftung is gratefully acknowledged.

Supporting Information Available: Experimental details of oligomer synthesis and gel mobility shift assays. This material is available free of charge via the Internet at http://pubs.acs.org.

OL049016A

(15) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161.
(16) Chen, F. M. F.; Benoiton, N. L. *Can. J. Biochem.* **1978**, *56*, 150.
(17) Paik, W. K.; Kim, S. *Science* **1971**, *174*, 114.
(18) Gurlie, R.; Zakrzewska, K. *Theor. Chem. Acc.* **2001**, *106*, 83.
(19) Bollhagen, R.; Schmiedberger, M.; Barlos, K.; Grell, E. *Chem. Commun.* **1994**, 2559.
(20) Fernández-Forner, D.; Casals, G.; Navarro, E.; Ryder, H.; Albericio, F. *Tetrahedron Lett.* **2001**, *42*, 4471.