The GCN4 Basic Region Leucine Zipper Binds DNA as a Dimer of Uninterrupted α Helices: Crystal Structure of the Protein–DNA Complex

Research by T.E. Ellenberger, C.J. Brandl, K. Struhl, and S.C. Harrison, Cell ***1992,*** 71, *1223*

Condensation and commentary by **Thomas Dandekar** *and* **Patrick Argos,** *European Molecular Laboratory, Heidelberg*

CONDENSATION OF THE RESEARCH

PURPOSE OF THE STUDY

To characterize structurally an important protein–DNA interaction

RESEARCHERS' APPROACH

The basic region of the leucine zipper DNA binding region (bZIP) of yeast transcriptional activator GCN4 (residues 226–281) was expressed in *Escherichia coli*. Crystals were obtained by vapor diffusion of the protein complexed with a synthetic 20-mer DNA (containing an AP-1 binding site) in a 12% PEG 400 solution with salts. The three-dimensional structure of the complex was solved by X-ray crystallographic techniques relying on multiple isomorphous replacement from 5-iododeoxyuridine substituted DNAs for diffracted X-ray phasing.

OBSERVATIONS

Each monomer of the GCN4-DNA binding element forms a continuous α-helix of 52 residues with no pronounced kinks or sharp bends, with the dimer displaying the appearance of a pair of α-helical tweezers (Fig. 1). A rigid coiled coil is formed by the two helices for about 30 residues at their carboxyl terminal halves. The crossing angle between monomers (about 18°) in the coiled coil region corresponds to Crick's[1] knobs-into-holes packing scheme for two α-helices. Though van der Waals interactions provide most of the stability of the dimer, primarily through side-by-side packing of leucines in a zipper fashion, salt links also appear helpful; namely, Glu 270 and Lys 275, Glu 268 and Lys 263, Glu 270 and Arg 273, and Glu 237 and Arg 240. The coiled coil structure is similar to that observed by O'Shea et al.[2] over the entire helices for an uncomplexed leucine zipper.

From residues 245–255, the protein subunits diverge from the dimer axis in a smoothly bending forklike fashion so that the DNA double helix can be accommodated further toward the N-terminus where the two helices open as tweezers

CHEMTRACTS—BIOCHEMISTRY AND MOLECULAR BIOLOGY **4**:62–65 (1993)
 CCC 1045-2680/93/010062-04$04.00

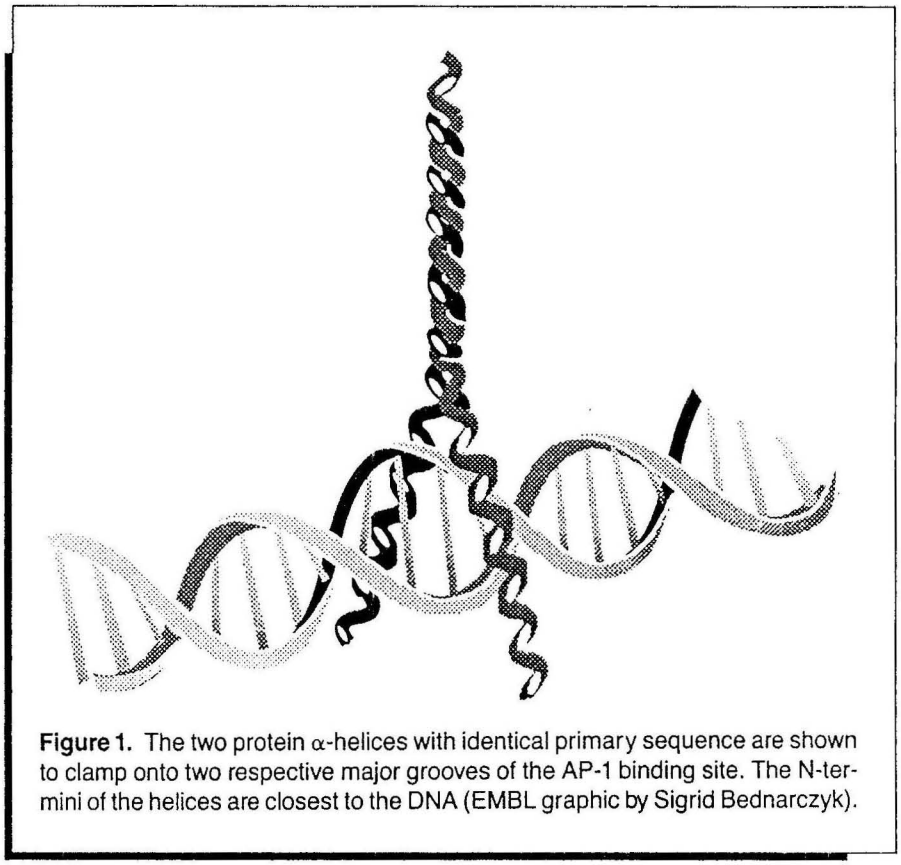

Figure 1. The two protein α-helices with identical primary sequence are shown to clamp onto two respective major grooves of the AP-1 binding site. The N-termini of the helices are closest to the DNA (EMBL graphic by Sigrid Bednarczyk).

clamping on opposite major grooves over the respective DNA half sites. At protein residue 243 the interaction with DNA begins. The first N-terminal turns of the helices do not wrap around the DNA after completing the grip but rather continue in a straight-line manner extending the tweezer tips.

The DNA AP-1 binding site ($A_{4L}T_{3L}G_{2L}A_{1L}C_0T_{1'R}C_{2'R}A_{3'R}T_{4'R}$) forms a pseudopalindrome and is straight B-form DNA in the GCN4–DNA complex. Each of the 4 bp AP-1 half sites around the central guanine–cytosine pair is contacted by one subunit of the GCN4 bZIP dimer. In each protomer the central Asn 235, conserved in the whole bZIP family, contacts Thy_3 and $Cyt_{2'}$. The β-carbons of Ala 238 and Ala 239, conserved residues in most bZIP proteins, contact Thy_3 and $Thy_{1'}$ and form a hydrophobic pocket around Asn 235. Thy_3 is also contacted by Ser 242. Arg 243, again conserved in all bZIP proteins, donates two hydrogen bonds to $Gua_{0'}$ at the center of the AP-1 site in the left site protomer and contacts Cyt_0 and Ade_{1L} in the right protomer. The base pairs known from footprint data[3] and uracil interference[4] to interact with GCN4 (1L-5L, 1′L-3′L,0, 0′,1R,2R-5R,1′R,2′R) are shown in the crystal structure to be all contacted by basic residues, notably Args (234, 241, 245, 249) and Lys 246.

COMMENTARY ON THE RESEARCH

GCN4 is an important transcriptional activator for general amino acid synthesis in yeast.[5] A crystal structure of this activator and its complex with DNA is essential for comprehension of general transcriptional activation, a key process in many types of gene regulation including differentiation, growth con-

trol, and cancer.[6] The structure is compatible with other available data. For instance, the footprint pattern[3] derived from the complete protein fits well. Thermal unfolding experiments[7] underline that residues outside of the bZIP region seem to contribute little to DNA binding.

Two models for the bZIP DNA interaction have been proposed. The scissors-grip model suggests that the basic region consists of two α-helical segments joined by a kink to permit the protein to wrap around DNA to facilitate contacts.[8] Ellenberger et al. propose, in agreement with the GCN4–DNA tertiary structure, the induced helical-fork model with complete helices in the basic region but folding only upon binding to DNA and otherwise being disordered in solution.[9,10] However, as Ellenberger et al. concede, GCN4 also recognizes specifically ATF/CREB sites, albeit in a weaker manner.[11] The latter site is a true palindrome of 5′TGAC3′ half sites and some kink or bend is required here in the bZIP region to accommodate the additional cytosine/guanosine base pair in the complex. The Fos/Jun heterodimer has absolute conservation with GCN4 in all the residues forming base specific contacts, albeit a static bend of the DNA by the dimer is observed[3] though it is still likely that the Fos/Jun contacts and structure are similar to that of the AP-1 site.

How general is the helical-fork model in protein–DNA interaction? Helices have often been observed to interact with the major groove of DNA where several repressors[12] and zinc fingers[13] afford the most prominent examples; however, they do not make use of a half zippered fork. Yet, electrostatic interactions were found not to be so important in the GCN4 bZIP. This need not be general as, for example, acidic residues in c-FOS destabilize the α-helical conformation that in turn may be stabilized by complementary residues of Jun in the Fos/Jun heterodimer.

The direct DNA–protein contacts noted from the crystal structure are not only supported by mutational data[11] but a genetic selection for altered AP-1 sites yields mutants with large hydrophobic chain substitutions to preserve exactly these interaction sites.[14] Furthermore, the straight projection of the basic region helical axis through the major groove reflects the snug fit of an α-helix in the major groove of B-form DNA. Something similar is also observed in the helix 3 from different homeodomains.[15,16]

The conserved amino acids in the bZIP family suggest a similar type of DNA recognition for members of this family. Though these base sites differ considerably, one or more central guanosine–cytosine pairs and Thy_{3L} or Gua_{2L}/Thy_{2L}, which form among the most intimate contacts with the GCN4 protein, are present in most cases. The new crystal structure should thus provide a good prototype for many similar protein–DNA interactions, a finding that will excite both theoreticians and experimentalists.

REFERENCES

1. Crick, F.H.C. *Acta Cryst.* **1953,** *6,* 689.
2. O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. *Science* **1991,** *254,* 539.
3. Gartenberg, M.R., Ampe, C., Steitz, T.A., Crothers, D.M. *Proc. Natl. Acad. Sci. USA* **1990,** *87,* 6034.
4. Pu, W.T., Struhl, K. *Nucl. Acids Res.* **1992,** *20,* 771.
5. Hinnebusch, A.G. *Microbiol. Rev.* **1988,** *52,* 248.
6. MacLeod, K., Leprince, D., Stehelin, D. *Trends Biochem. Sci.* **1992,** *17,* 251.

7. Weiss, M.A. *Biochemistry* **1990,** *29,* 8020.
8. Vinson, C.R., Sigler, P.B., McKnight, S.L. *Science* **1989,** *246,* 911.
9. Saudek, V., Pasley, H.S., Gibson, T., Gausepohl, H., Frank, R., Pastore, A. *Biochemistry* **1991,** *30,* 1310.
10. O'Neil, K.T., Shuman, J.D., Ampe, C., DeGrado, W.F. *Biochemistry* **1991,** *30,* 9030.
11. Sellers, J.W., Vincent, A.C., Struhl, K. *Mol. Cell. Biol.* **1990,** *10,* 5077.
12. Brennan, R.G. *Curr. Opinion Struct. Biol.* **1992,** *2,* 100.
13. Berg, J.M. *Curr. Opinion Struct. Biol.* **1993,** *3,* 11.
14. Tzamarias, D., Pu, W.T., Struhl, K. *Proc. Natl. Acad. Sci. USA* **1992,** *89,* 2007.
15. Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T.B., Pabo, C.O. *Cell* **1990,** *63,* 579.
16. Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D., Pabo, C.O. *Cell* **1991,** *67,* 517.