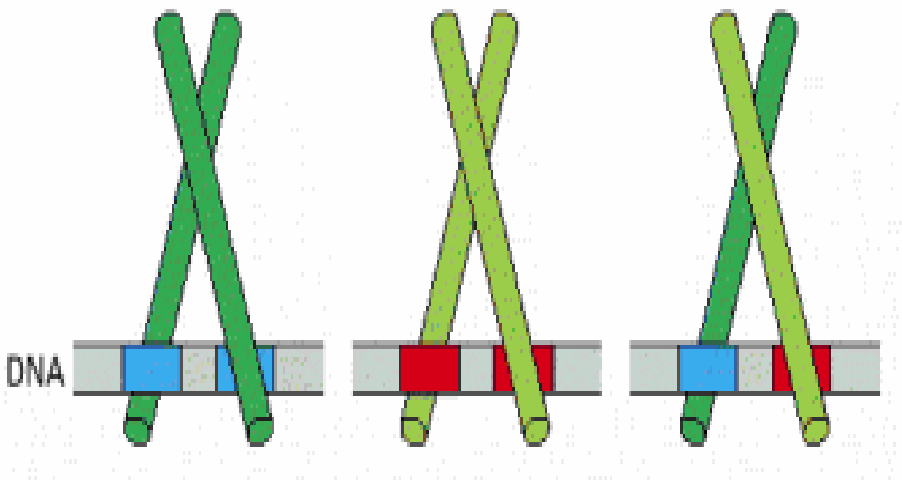
MIC/BIO/BCH522

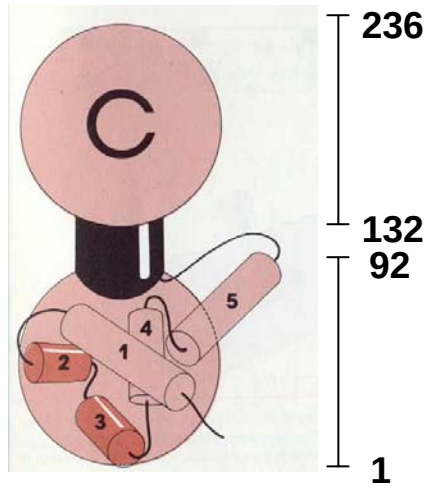
Spring 2006

Interactions of proteins with nucleic acids
3. The Leucine Zipper Motif
4. The Winged helix proteins

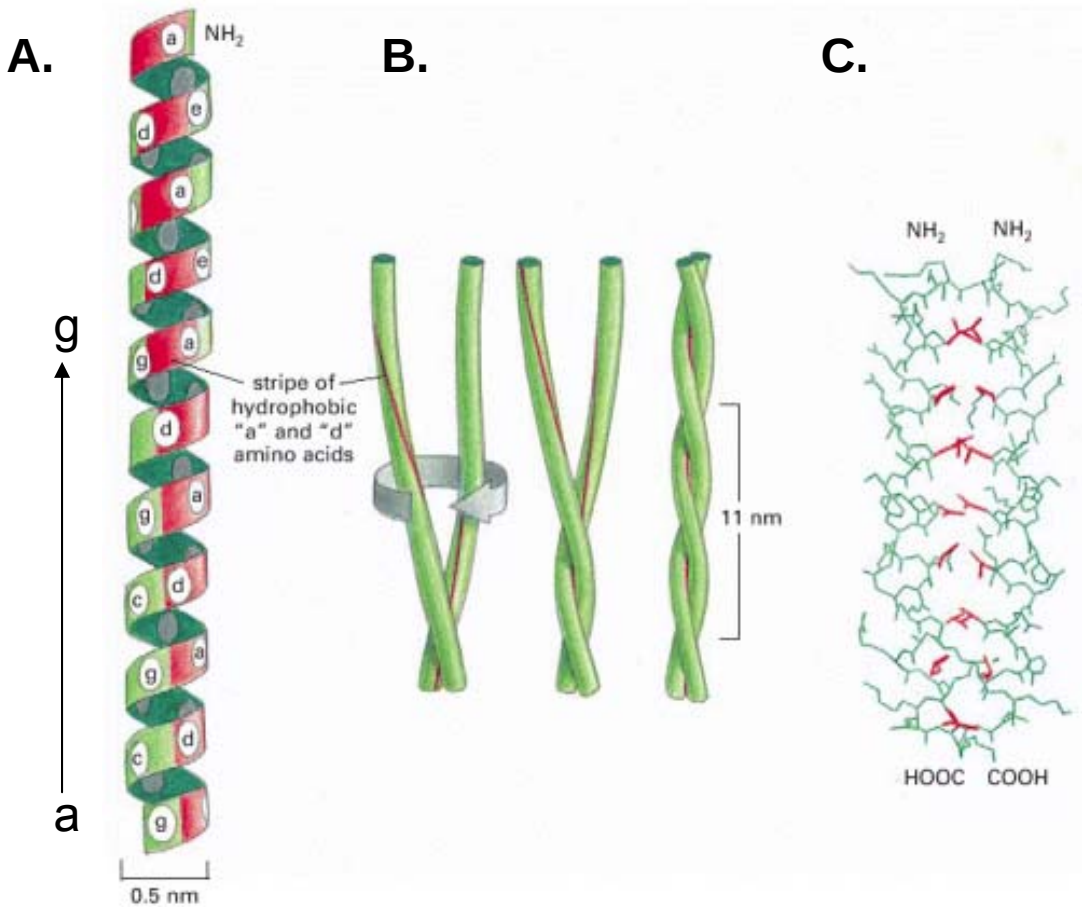
Leucine Zippers

- So far, in the proteins we've looked at, DNA binding and dimerization have been localized to separate motifs
- The leucine zipper is an efficient combination of DNA binding and dimerization motifs
- The **leucine zipper motif**, is so named because of the way the two α helices, one from each monomer, are wound around each other in a coiled-coil, held together by interactions between amino acid side chains
- The diagnostic feature of the motif is an amphipathic helix with leucine at every seventh position (**3.5 bases/turn; thus every other turn**).
- the **leucine zipper** motif itself is **not** directly involved in DNA binding.
- **B-ZIP motif** = N-terminal **basic** region + **leucine zipper** domain (with a six residue connector)
- Just beyond the dimerization interface the two α helices separate from each other to form a Y-shaped structure, which allows their side chains to contact the major groove of DNA.
- The dimer thus grips the double helix like a clothespin on a clothesline
- the helical structure of the basic region is disordered until dimerisation occurs
- Leucine zipper proteins bind to DNA as dimers hence dyad symmetry seen in their binding site sequences
- An important consequence of separating binding and dimerisation domains is the possibility that this offers for the formation of heterodimers with hybrid binding specificities.
- Leucine zipper is a general module for protein-protein interactions



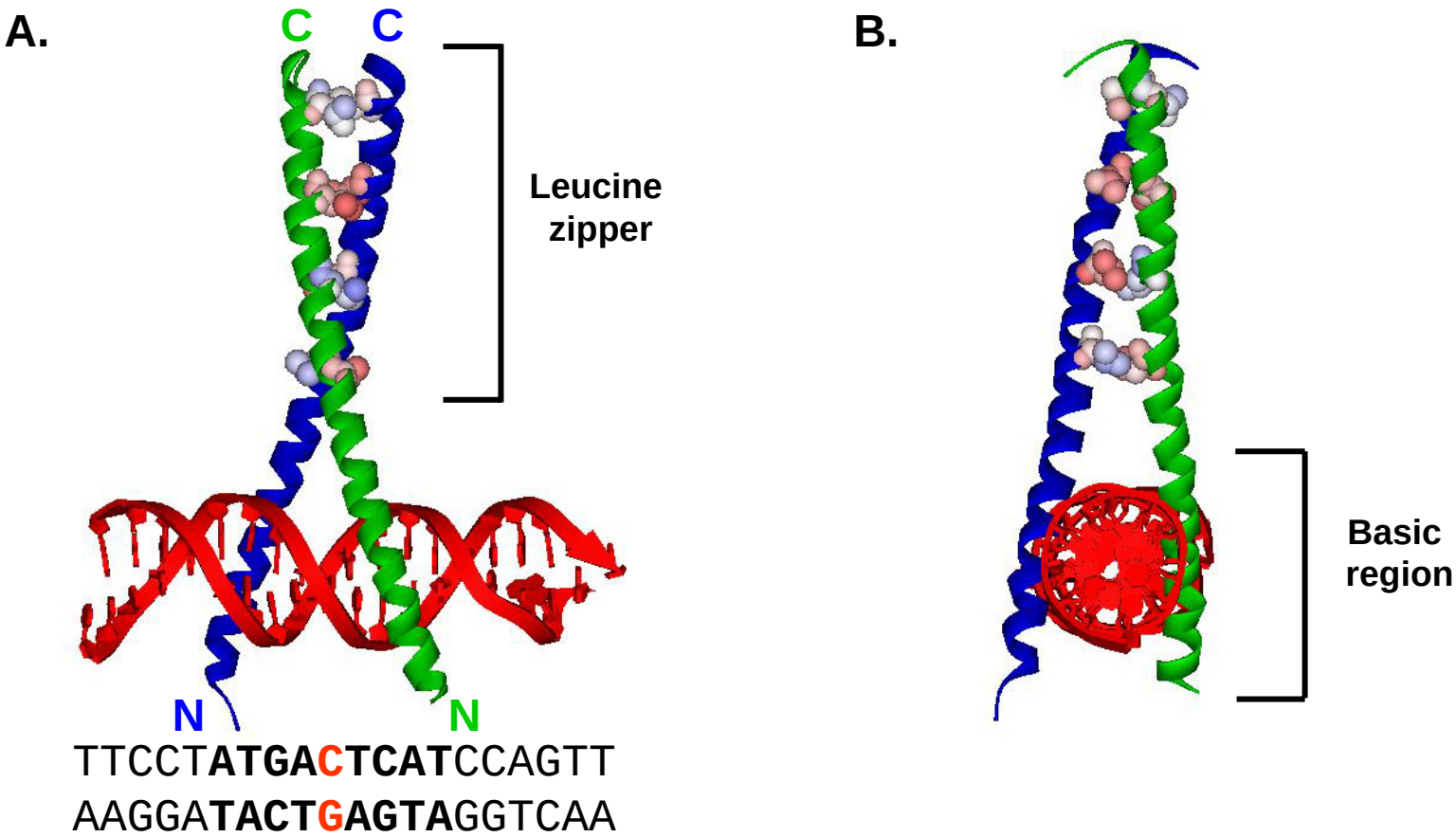


To understand leucine zippers, you need to know how α-helices pack together in a coiled-coil



(A) A single α helix, with successive amino acid side chains labeled in a **heptad repeat**, "abcdefg". Nonpolar amino acids in positions "a" and "d" in such a sequence lie close together on the cylinder surface, forming a "stripe" (*red*) that winds slowly around the α helix. Consequently, as shown in (B), the two α helices can wrap around each other with the nonpolar side chains of one α helix interacting with the nonpolar side chains of the other, while the more hydrophilic amino acid side chains are left exposed to the aqueous environment. (C) Schematic of a coiled-coil showing the positions of nonpolar amino acids (*red* side chains).

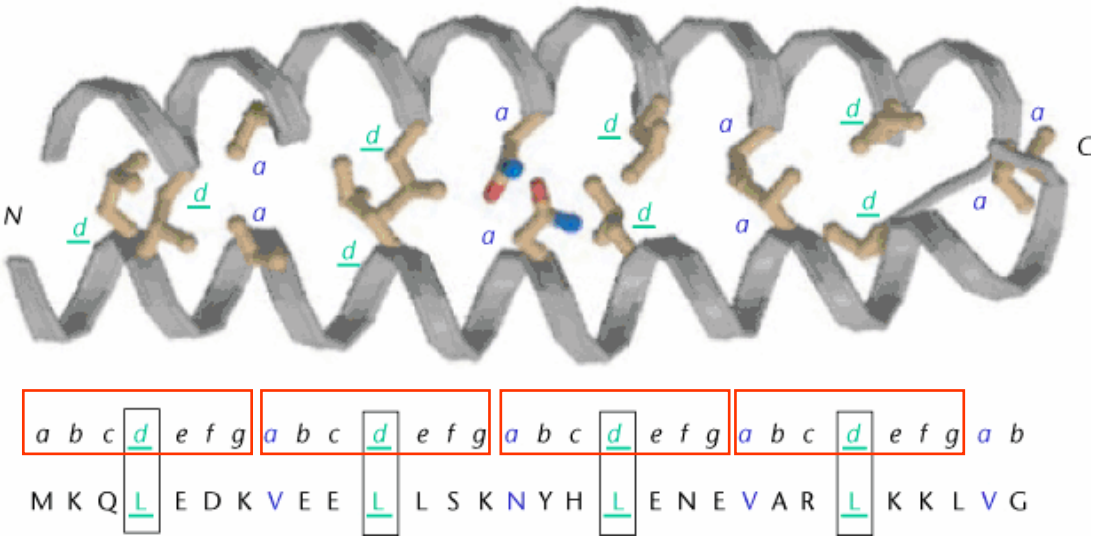
The yeast transcriptional activator GCN4



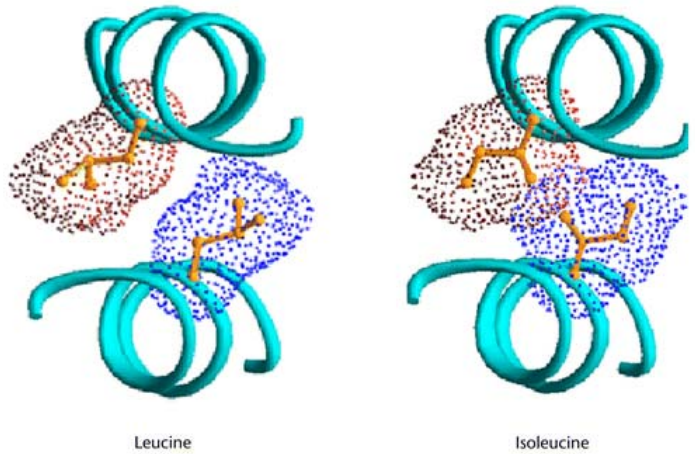
The yeast transcriptional activator GCN4 is 1 of over 30 identified eukaryotic proteins containing the basic region leucine zipper (bZIP) DNA-binding motif. The bZIP dimer is a pair of continuous α helices that form a parallel coiled coil over their carboxy-terminal 30 residues and gradually diverge toward their amino termini to pass through the major groove of the DNA-binding site. The coiled-coil dimerization interface is oriented almost perpendicular to the DNA axis, giving the DNA-protein complex the appearance of the letter T. There are no kinks or sharp bends in either bZIP monomer. Numerous contacts to DNA bases and phosphate oxygens are made by basic region residues that are conserved in the bZIP protein family. The d position amino acids (leucine)are highlighted. (PDB – 1YSA)

How do the α-helices pack together in a leucine zipper?

A.



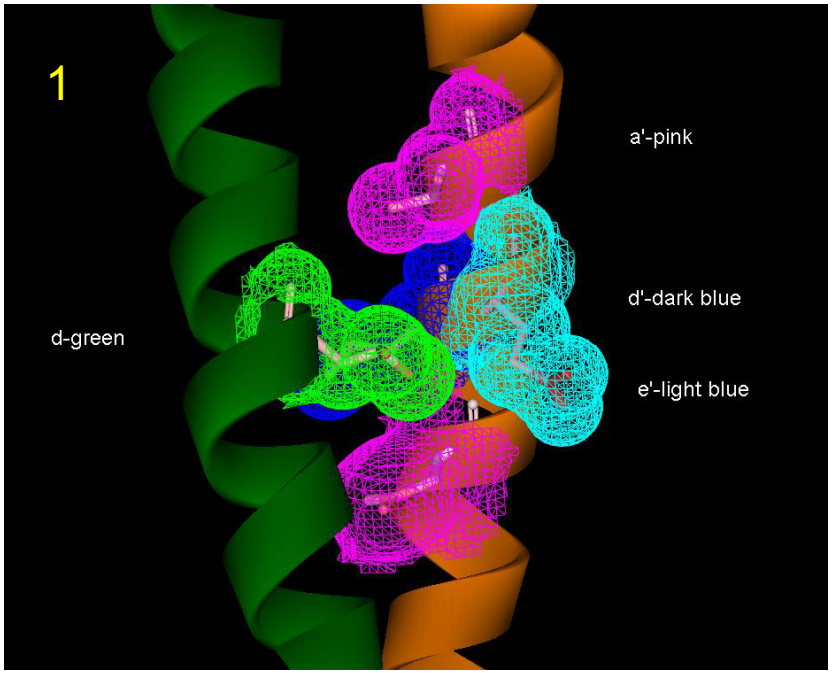
B.

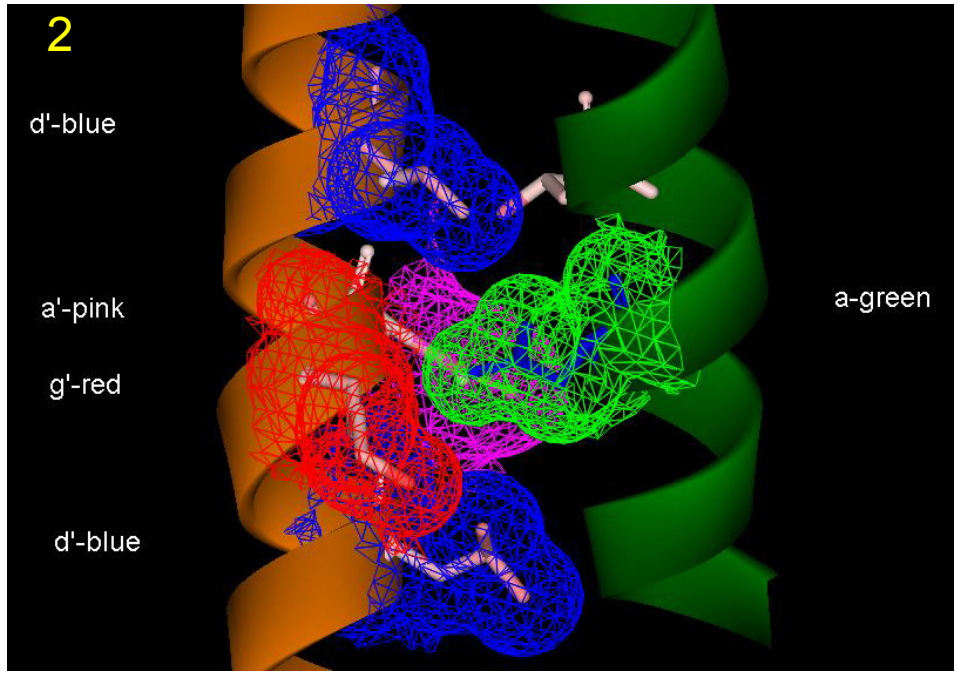


A. Parallel coiled-coil structure of GCN4 ZIP homodimer (1gd2). The main chains of the two peptide chains are represented as ribbons in gray. The side chains participating in the dimer association are represented as stick models with carbon atoms in brown, nitrogen atoms in blue and oxygen atoms in red. The positions of the heptad repeat are labeled ***a***–***g***. The ***d***-positioned leucines are boxed and highlighted in green with underline. The ***a***-positioned residues are highlighted in blue. Unusually, this GCN4 ZIP motif has a polar asparagine residue in the ***a*** position located at the middle of the ZIP motif.

B. An end view, looking from the N-terminus, of the leucine zipper interface with either leucine or isoleucine in the d and d′ positions. Leucine which packs well in the hydrophobic core to make dimers. Isoleucine is β-branched and cannot pack well and instead induces trimer formation.

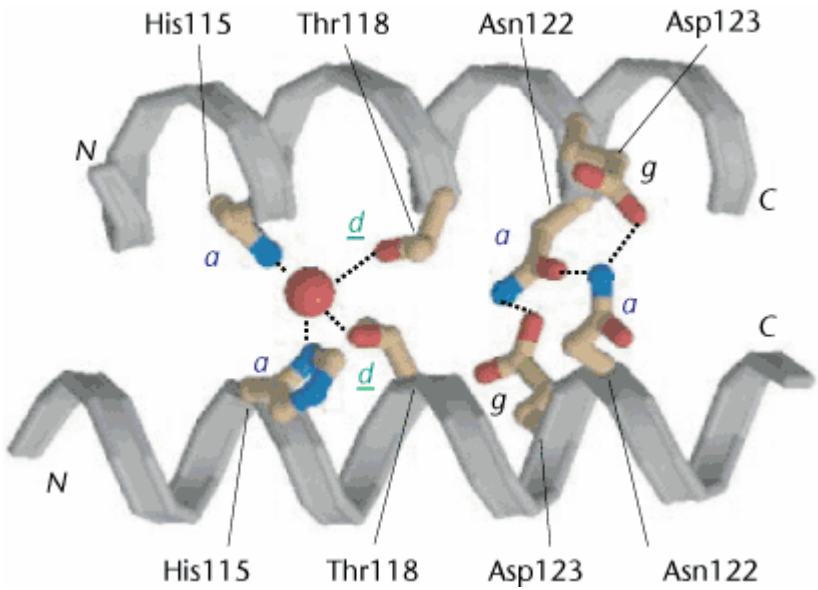
"knobs into holes" side-chain packing in leucine zippers





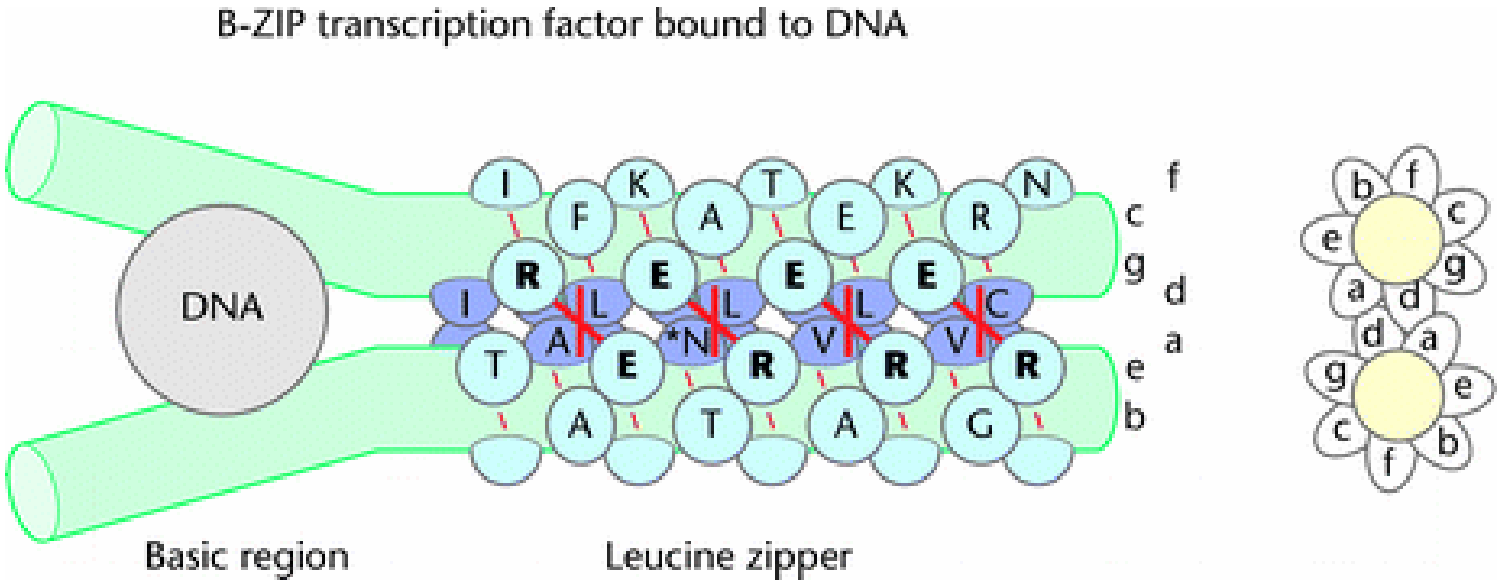
1. packing of a "d knob" into a hole formed by an a, e', d' and 2a'.
2. Packing of an "a-knob" into a hole formed by d',a',g' and second d'.

What about other residues in the α-helices ?



- Internal interactions of polar residues in ***a*** and ***d*** positions in the coiled-coil structure of the Pap1 ZIP dimer. Part of the coiled-coil structure is presented to show that His115 in the ***a*** position and Thr118 in the ***d*** position form water-mediated hydrogen bonds at the coiled-coil interface. Polar residues Asn122 in the ***a*** position and Asp123 in the ***g*** position also form interchain hydrogen bonds.

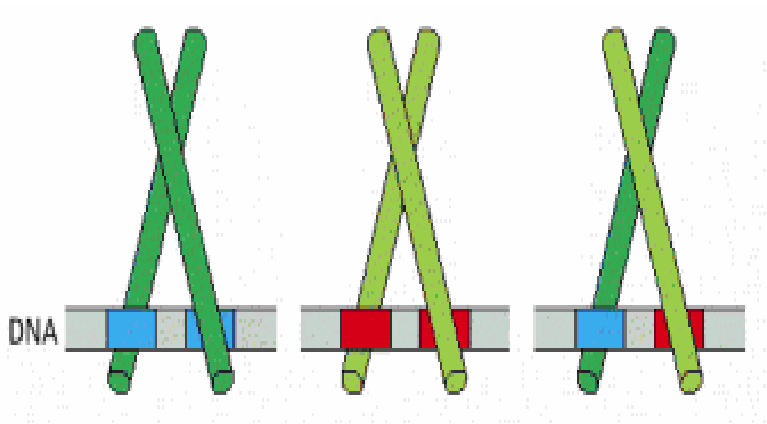
Several residues contribute to dimerization specificity



- A schematic of the B-ZIP PAR family member VBP viewed from the side with the amino acids from the VBP leucine zipper shown inside the circles which represent amino acid positions along the two α helices. Amino acids in the **e and g position are shown in bold face** and the *i*, *i*′+5 (g↔e′) interactions are connected by arrows pointing from acidic to basic. The heptad letter designations (a, b, c, d, e, f, g) are shown.
- To the right is an end view of a leucine zipper dimer looking from the N-terminus. The letters on the inside of each ellipse represents the standard nomenclature for the seven amino acids found in unique positions in a coiled-coil. The ellipses depict the orientations of the amino acid side-chains relative to the α helix. Amino acids in the a and d positions create a hydrophobic core between the interacting helices. The interaction seen between amino acids in the g and subsequent e′ position seen in X-ray structures is noted as g↔e′ pairs. Note that because of the 2-fold symmetry of the dimers, each heptad contains two g↔e′ pairs.

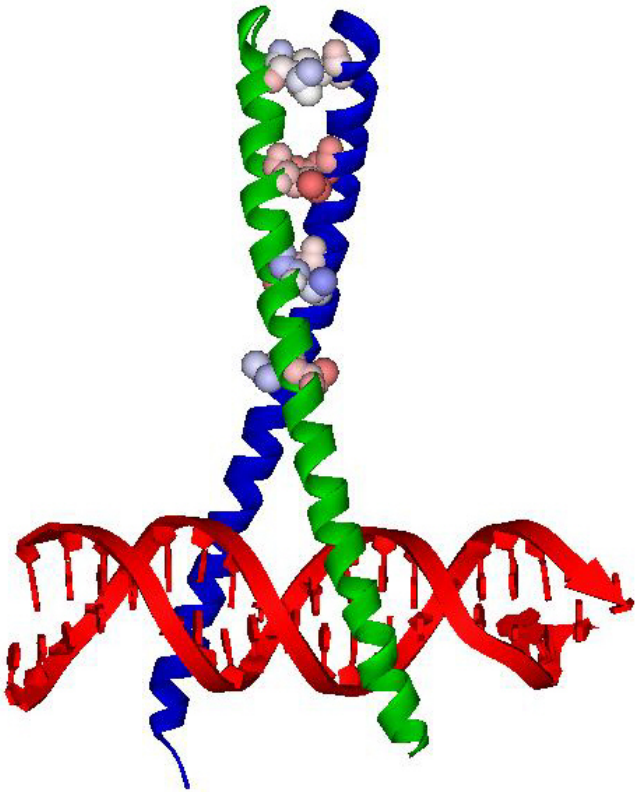
Leucine zipper heterodimers

- Many of the gene regulatory proteins we have seen thus far bind DNA as homo-dimers,
- Several of these, including leucine zipper proteins, can also associate with non-identical partners to form heterodimers composed of two different subunits.
- Because heterodimers typically form from two proteins with distinct DNA-binding specificities, the mixing and matching of gene regulatory proteins to form heterodimers greatly expands the repertoire of DNA-binding specificities that these proteins can display.
- There are, however, limits to this promiscuity: if all the many types of leucine zipper proteins in a typical eucaryotic cell formed heterodimers, the amount of “cross-talk” between the gene regulatory circuits of a cell would be so great as to cause chaos.
- Whether or not a particular heterodimer can form depends on how well the hydrophobic surfaces of the two leucine zipper α helices mesh with each other, which, in turn, depends on the exact amino acid sequences of the two zipper regions.
- Thus each leucine zipper protein in the cell can form heterodimers with only a small set of other leucine zipper proteins.
- Heterodimerization is an example of combinatorial control in which combinations of different proteins, rather than individual proteins, control a cellular process.
- Heterodimerization is one of the mechanisms used by eucaryotic cells to control gene expression in this way, and it occurs in a wide variety of different types of gene regulatory proteins

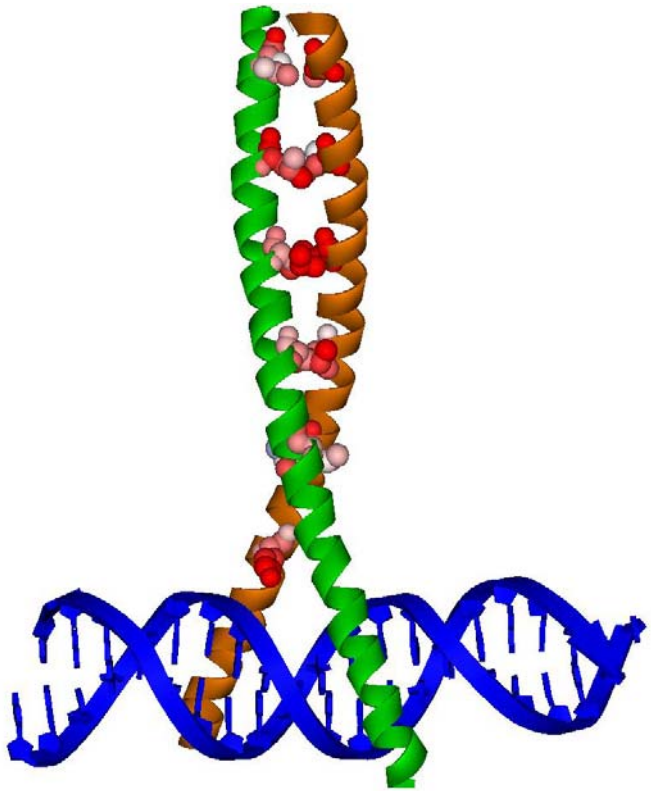


The AP-1 heterodimer: Fos-Jun

A.



B.

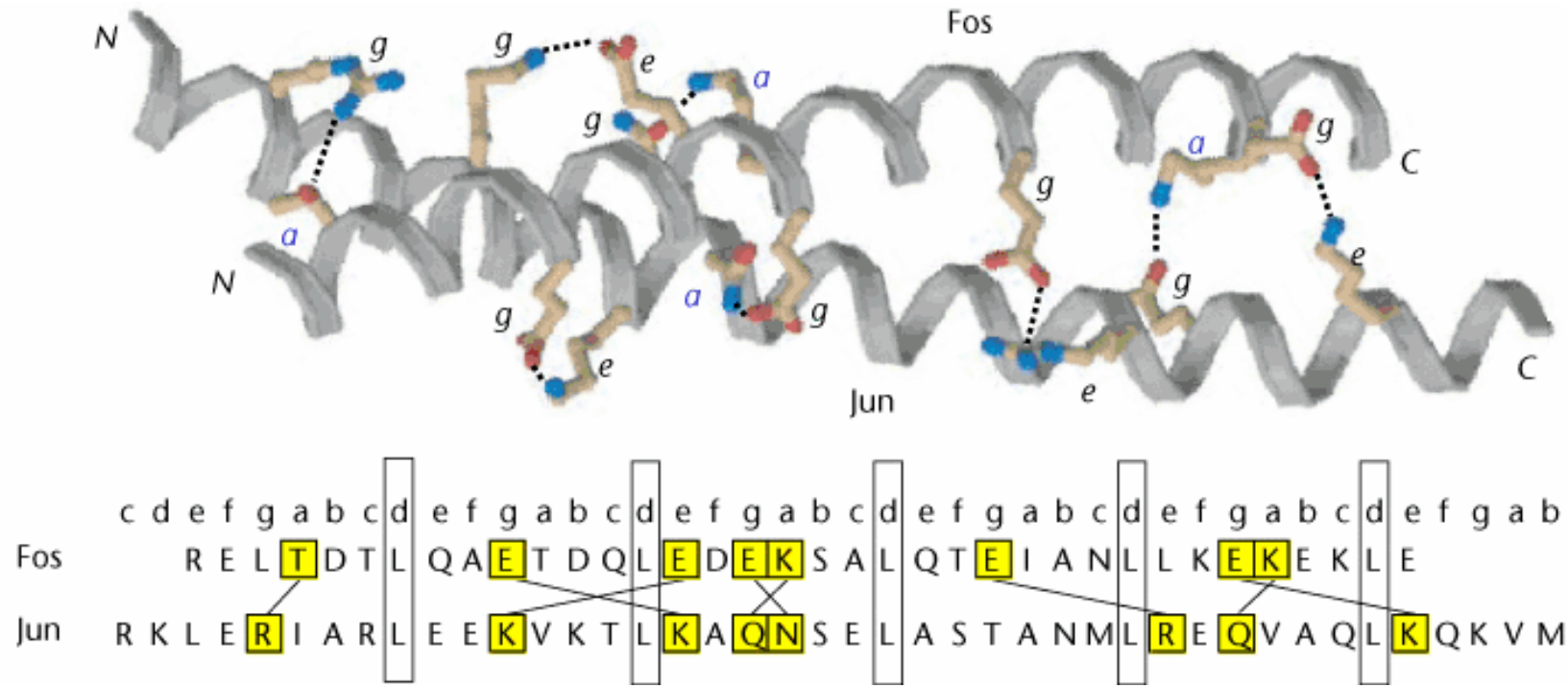


A. The yeast transcriptional activator GCN4

B. The AP-1 heterodimer Fos-Jun bound to DNA

AP = adaptor protein complex; Fos and Jun are encoded by proto-oncogenes and are involved in cell growth control

Side chain interaction in the Fos-Jun heterodimer



Interchain polar interactions found in the coiled-coil structure of the Fos–Jun ZIP heterodimer (PDB accession code 1fos). Hydrogen bonding interactions are indicated by dotted lines. The dimerization specificity is primary determined by polar interactions involving the side chains of residues at the ***e*** and ***g*** positions. A few polar residues at the ***a*** positions also contribute to the specificity.

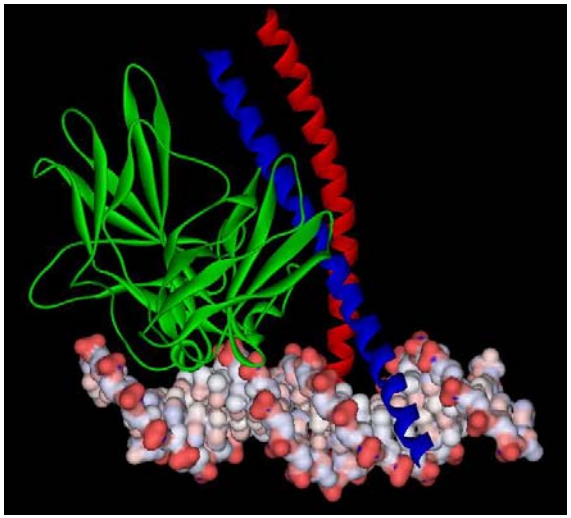
Summary for leucine zippers

- bZIP motif forms a parallel coiled-coil structure
- bZIP proteins can homodimerize and heterodimerize
- Only specific bZIP proteins heterodimerize (Fos-Jun) and some DO NOT homodimerize (eg Fos)
- **Leucine** residues at **d** position and **nonpolar** residues at **a** positions are the prerequisites for dimerization
  - This is because they provide the fundamental framework of coiled-coil structure
- Dimerization specificity involves non-conserved residues at positions e and g
- Dimerization specificity is determined primarily by electrostatic interactions between the 2 helices involving the charged termini of long charged residues at positions e and g
- These termini can form inter-helical salt bridges
- Prior to dimerization, the DNA binding region is disordered; following dimerization it is ordered
- The leucine zipper does not bind DNA, instead the basic region (~20 aa in length) contacts dsDNA in the major groove
- Contacts are made on opposite faces of the DNA duplex in adjacent turns
  - no 10 bp spacer as in HTH

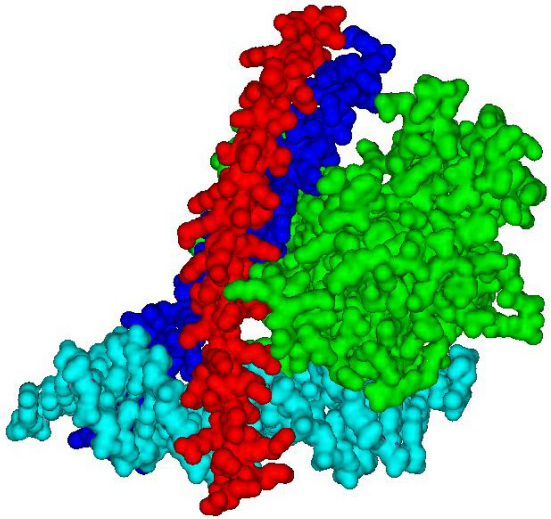
(NFAT) and the AP-1 heterodimer, Fos-Jun

cooperatively bind a composite DNA site and synergistically activate the expression of many immune response genes..

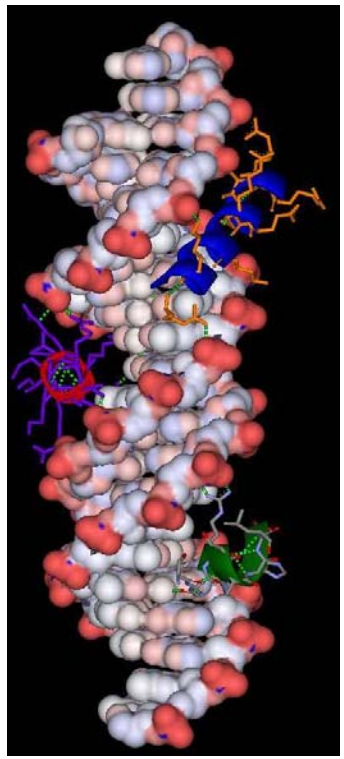
A.



B.



C.

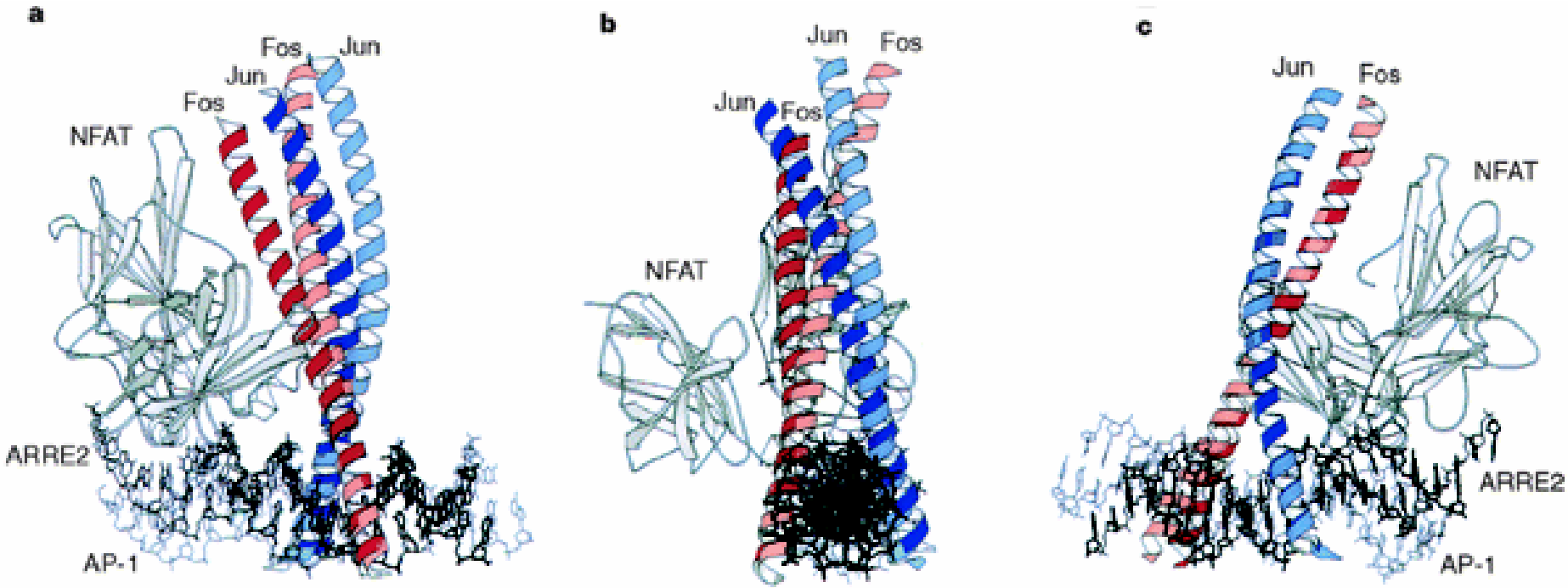


A. The nuclear factor of activated T cells (NFAT, coloured green) and the AP-1 heterodimer, Fos-Jun bound to DNA. in a quaternary complex. The DNA fragment containing the distal antigen-receptor response element from the interleukin-2 gene promoter, shows an extended interface between NFAT and AP-1, facilitated by the bending of Fos and DNA. The tight association of the three proteins on DNA creates a continuous groove for the recognition of 15 base pairs. (Nature v392 pp.42-48 , 1998).

B. Surface representation showing the partly discontinuous character of the contacts. NFAT-green, Fos –blue and Jun-red; DNA light blue

C. The alpha helices interacting with the major grooves of DNA. NFAT-green; Jun-red; Fos-blue

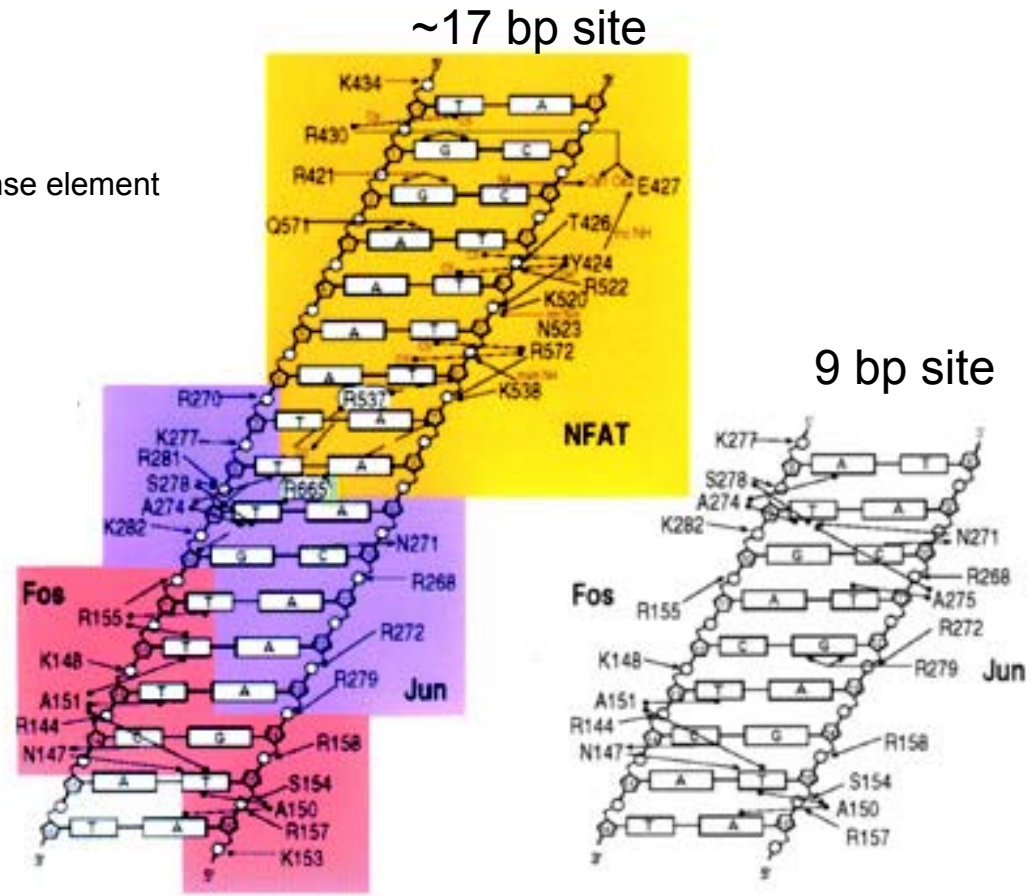
(NFAT) and the AP-1 heterodimer, Fos-Jun



- Superpositions of the NFAT–Fos–Jun–ARRE2 quaternary complex on the Fos–Jun–AP-1 site ternary complex. These comparisons illustrate that **bending of Fos and of DNA together** allow the formation of the observed NFAT–Fos–Jun interface.
- **a**, The ternary (Fos–Jun–AP-1 site) and quaternary (NFAT–Fos–Jun–ARRE2) complexes have been superposed so that the basic region of Fos and the segment of DNA that it contacts are optimally aligned. Lighter colours and thin lines indicate the ternary complex; darker colours and thick lines show the quaternary complex; NFAT is shown grey. The DNA bend can be seen by comparing the two superposed structures. **b**, End-on view of the same superposition. **c**, The two structures have been superposed so that the leucine zippers are optimally aligned. This superposition, when compared with the superposition used in **a** and **b**, shows that when shifting from the conformation in the ternary complex to that in the quaternary complex, the rest of the Fos–Jun–DNA complex can be imagined to pivot about the fork segment of Fos.

DNA recognition in the NFAT–AP 1–ARRE2 complex.

ARRE2 is the antigen receptor response element

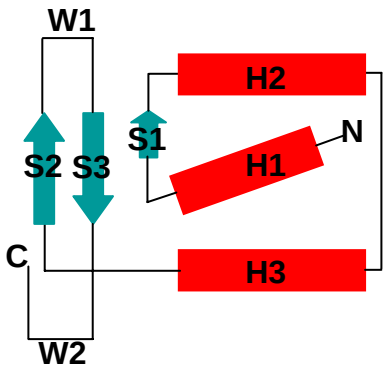


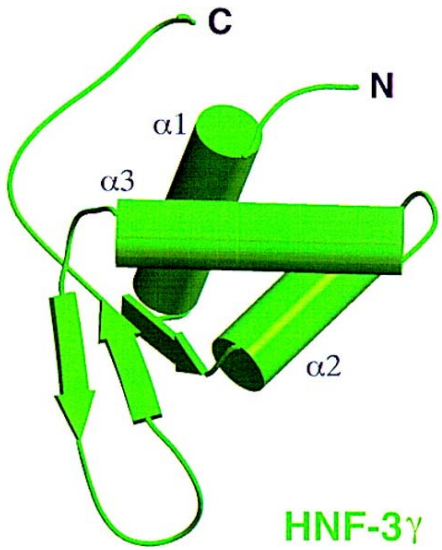
Schematic diagram of detailed protein–DNA contacts, compared the ternary Fos–Jun–AP-1 site complex (lower right).

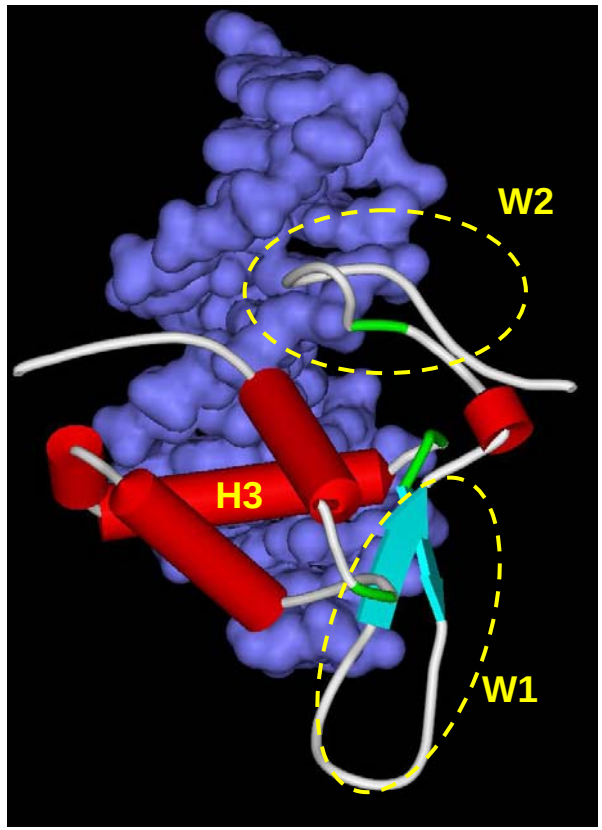
Lines with arrowheads, hydrogen bonds; lines with filled circles at either end, van der Waals contacts; solid lines, major-groove and sugar-phosphate interactions; dashed lines, minor-groove interactions. The coloured blocks indicate which protein is involved.

4. The Winged helix Motif

- A sub-family within the HTH family first discovered in liver-specific transcription factors (HNF-3)
- This motif is a compact α/β structure
- Consists of two wings (W1 and W2), three α-helices (H1-3)and three β-strands(S1-3)
- These are arranged in the following order: H1-S1-H2-H3-S2-W1-S3-W2
- Helix 3 is the recognition helix
- The loops W1 and W2 flank H3 like wings of a butterfly, hence the name winged helix
- DNA recognition also occurs between W2 and the minor groove



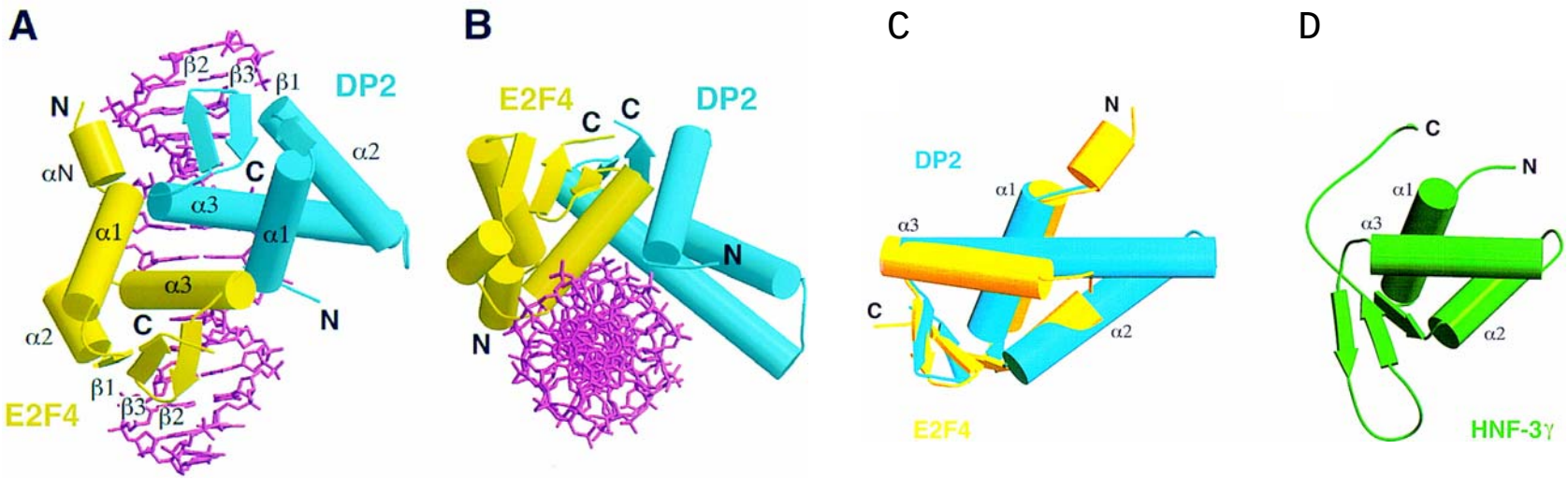




pdt013

Structure of the E2F4-DP2 heterodimer DNA complex.

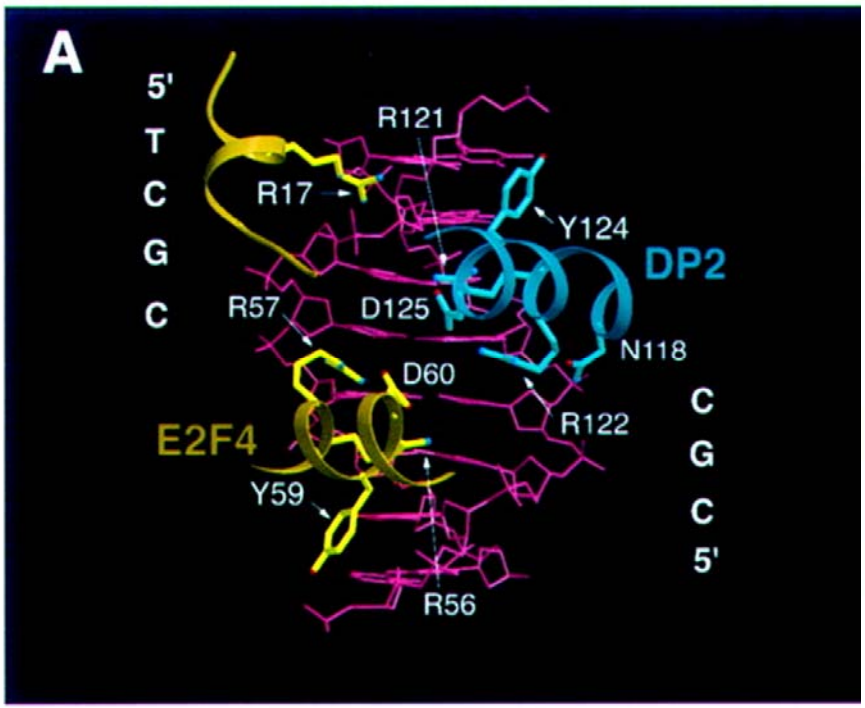
- E2F and DP2 are transcription factors control genes involved in growth and DNA replication
- DNA binding by E2F is enhanced by heterodimer formation with DP2 (distant relative)
- Both E2F and DP2 are winged helix proteins, but lack W2
- Here the wing fold participates in protein-protein interactions, not in DNA binding
- In the heterodimer, significant contacts between H1 and H3 of both proteins

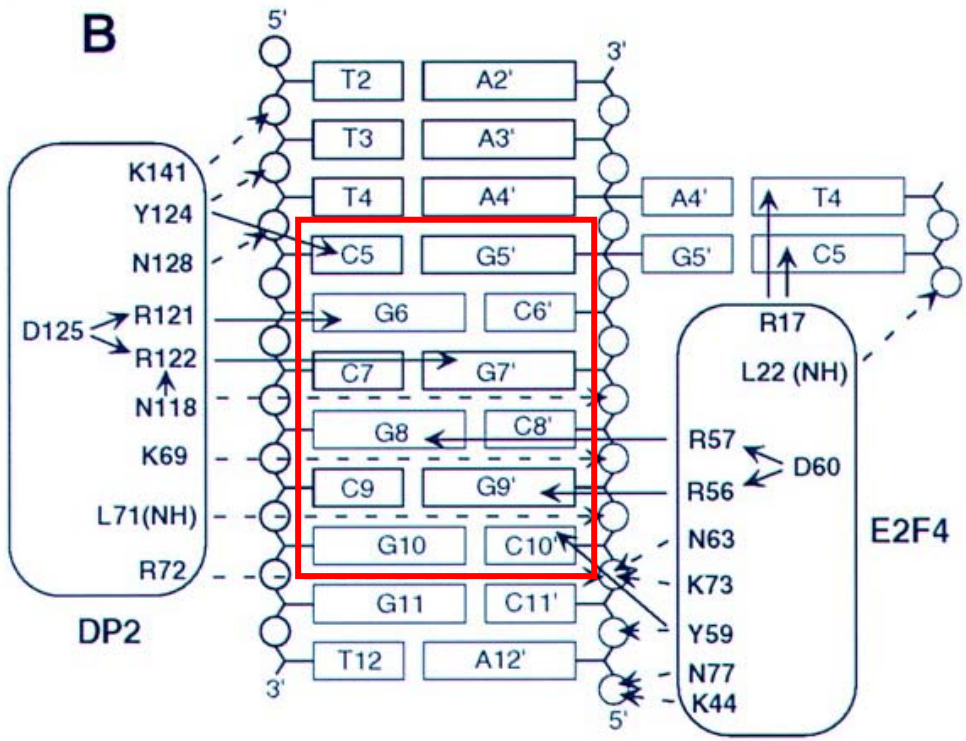


- (*A*) Schematic view looking down the approximate axis of twofold pseudo symmetry in the heterodimer. The DNA axis is vertical in this view.
- (*B*) View of the complex looking down the DNA axis.
- (*C*) Superposition of the E2F4 and DP2 winged-helix DNA-binding domains.
- (*D*) The winged-helix domain of the HNF-3 transcription factor (Clark et al. 1993) in an orientation obtained by aligning it with the DP2 winged-helix domain in *A*.

Genes and Dev. 13:666-674 (1999)

Recognition of the core DNA site by the E2F4-DP2 heterodimer is overall symmetric.



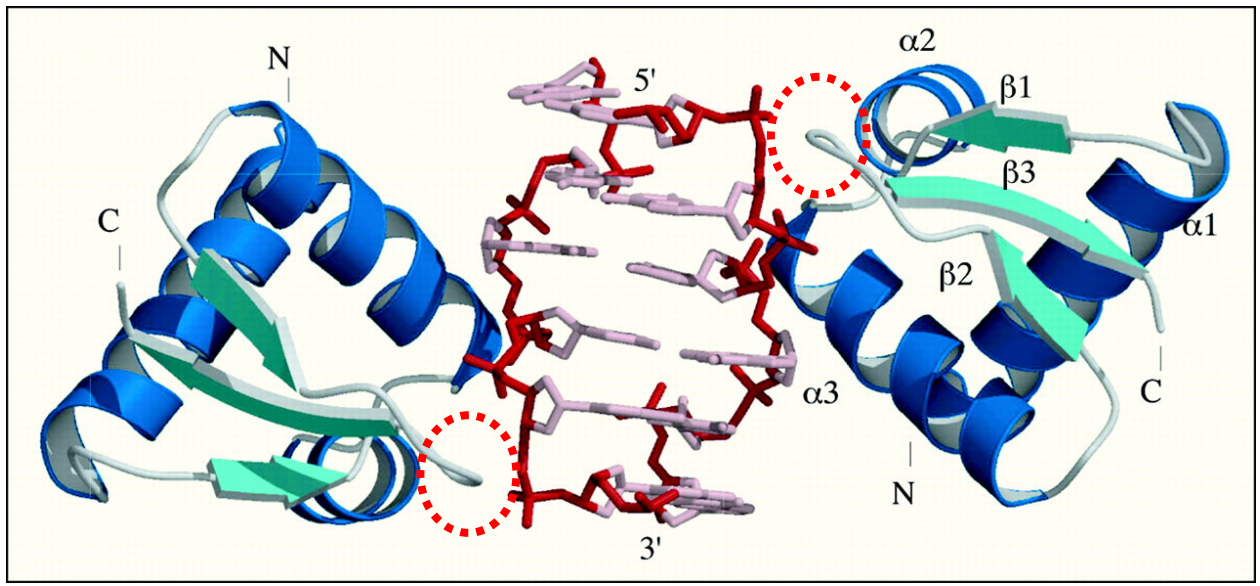


(A) The E2F4 and DP2 residues that contact the DNA bases; for clarity, only part of the α3 helices of each subunit and the amino-terminal helix of E2F4 are shown.

(B) Sketch of the DNA contacts made by the E2F4-DP2 heterodimer.

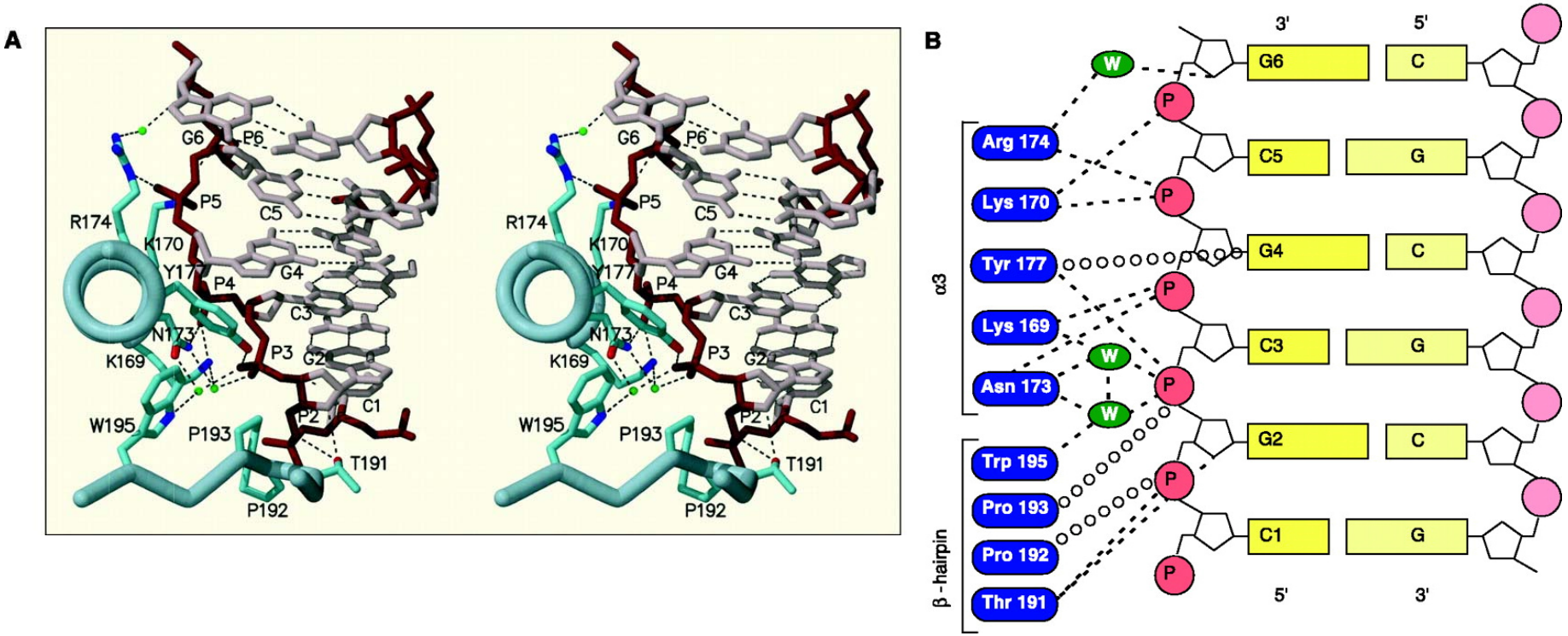
ADAR1: a Z-DNA binding, winged helix motif protein

- ADAR1 is a dsRNA editing enzyme that includes a DNA binding domain called Zα
- Zα is responsible for high affinity binding to Z-DNA (10,000-fold higher than B-DNA)
- The overall topology of the monomer is that of the compact α/β structure
- Helices 2 and 3 lie between sheets 1 and 2 and form the HTH
- Monomeric Zα binds to one strand of a DNA palindrome; a second monomer binds to the opposite strand - the two monomers do not interact
- DNA contacts are between αH3 and the C-terminal base of W1



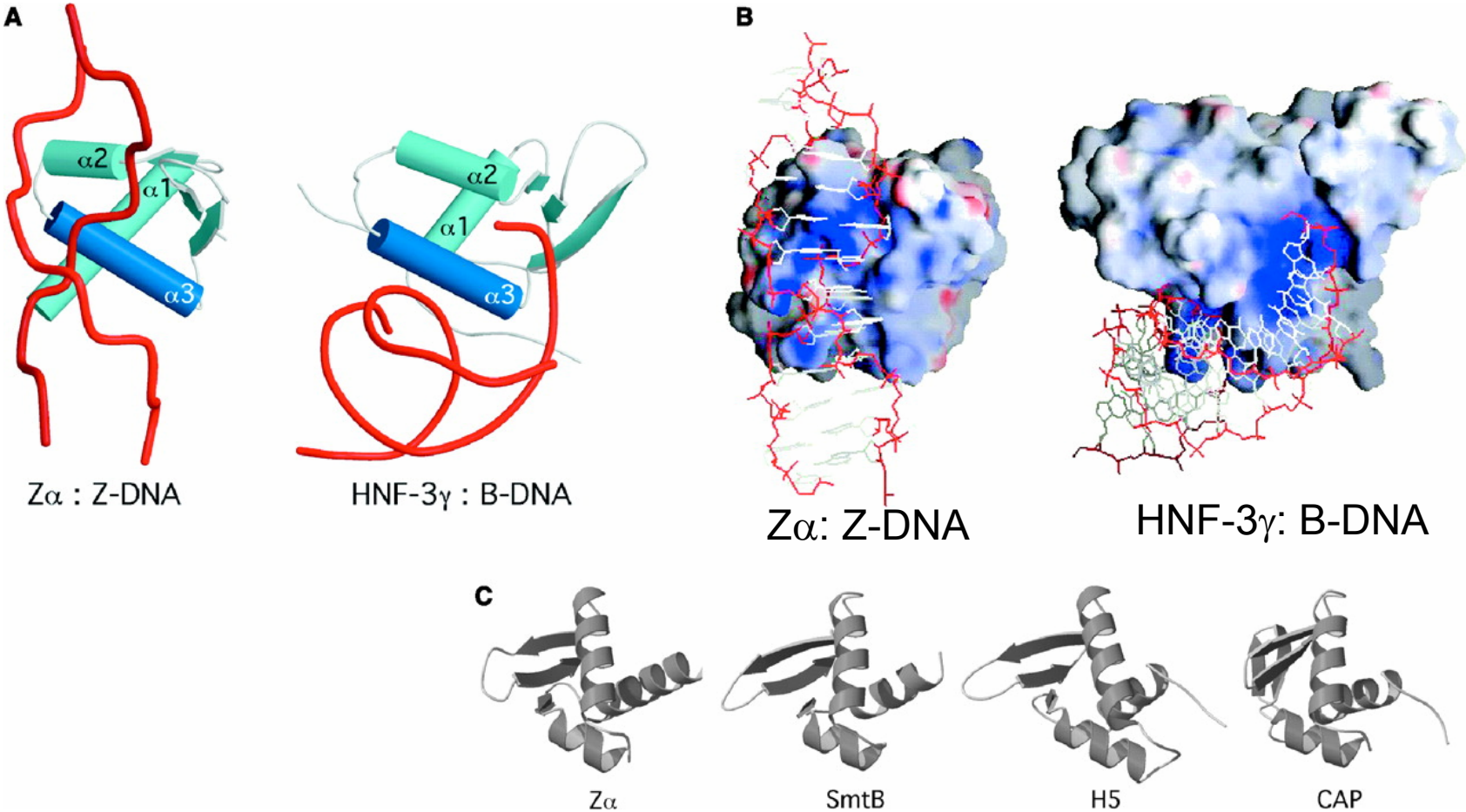
Overview of the Z domain bound to left-handed Z-DNA. Residues 134 to 198 of two symmetry-related Z monomers and the 6-bp DNA duplex d(CGCGCG)2 are shown. Labels indicate NH2- (N) and COOH-termini (C) of the proteins as well as helices (α1 to α3) and strands (β1 to β3)
Science 284: 1841-1845 (1999)

How does the recognition helix interact with Z-DNA?



- (**A**) Stereoview (down the axis of recognition helix $\alpha 3$) shows the entire region of the DNA recognized by Zα. Five consecutive backbone phosphates of the DNA are contacted by an extensive hydrogen bonding network. Protein side chains in direct or water-mediated contact with the DNA are labeled. Water molecules are green. Tyr177 is involved in the only base contact seen in the complex and is in van der Waals contact with the exposed carbon 8 of the purine base G4, characteristic of Z-DNA.
- (**B**) Protein residues involved in DNA interactions. All contacts are between the protein and one strand of the DNA duplex. Hydrogen bonds are indicated by dashed lines and van der Waals contacts are shown by open circles. Three water molecules in key positions within the protein-DNA interface are indicated by green ovals.

Comparison with other HTH proteins.



(**A**) Schematic of the Z -DNA complex (left) compared with the HNF-3 -DNA interaction (right). Proteins are oriented with HTH units in the same orientation. The recognition helix 3 of the HTH motif is dark blue and the backbone trace of the DNA is red. The DNA axes are radically different because of the different "angle of attack" both proteins use to bind DNA. (**B**) Electrostatic surfaces of the proteins in the same orientation as in (A). Positive (blue) and negative (red) surface potentials are indicated. (**C**) Schematics showing the close similarity in topology of the Z domain (left) and three members of the HTH family with α/β topology in order of increasing structural deviations from left to right: metallothionein repressor SmtB, globular domain of histone H5, and catabolite gene activator protein CAP