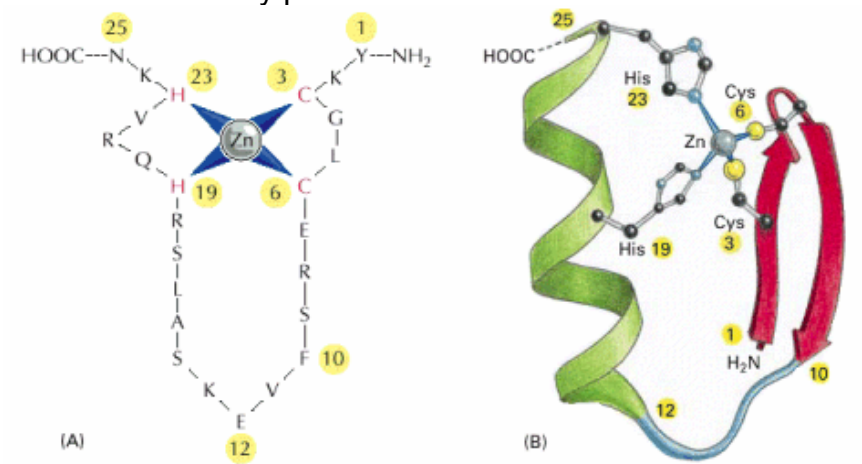
MIC/BIO/BCH522

Spring 2006

Interactions of proteins with nucleic acids
2. The Zinc Finger Motif

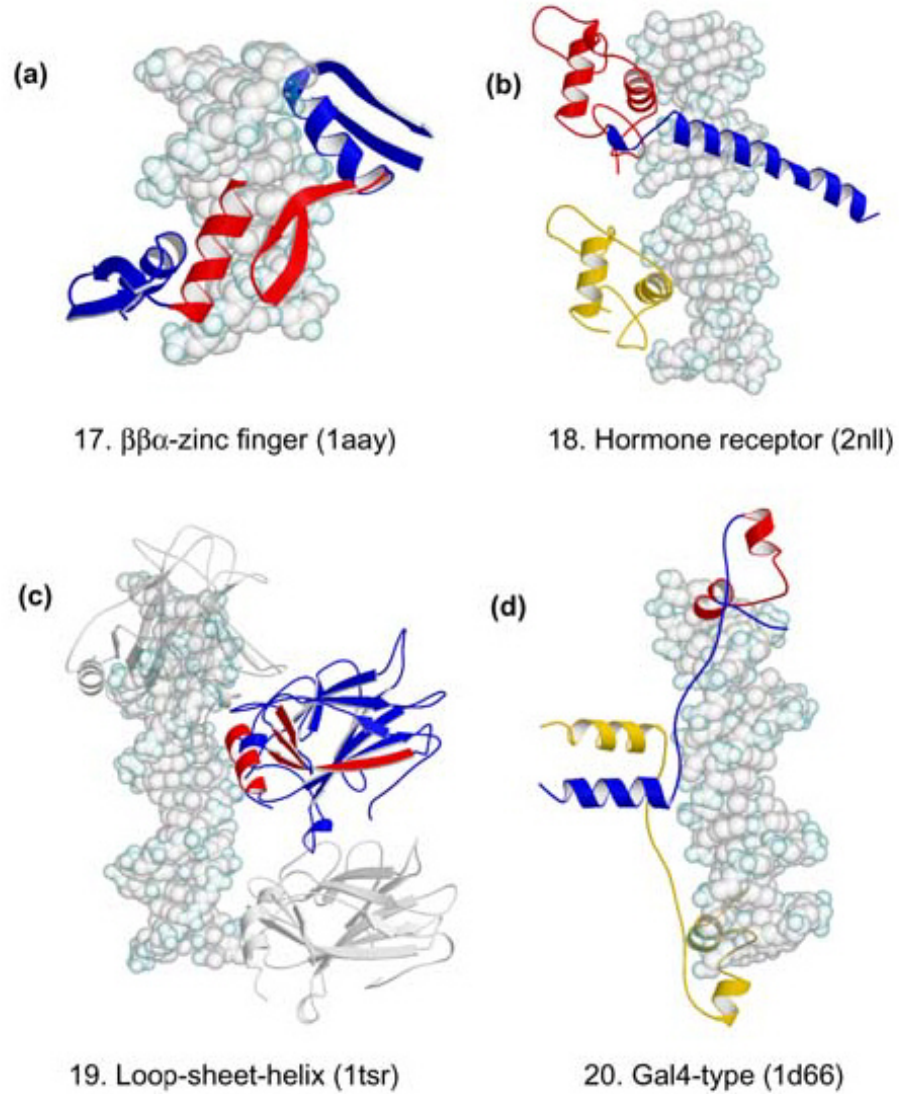
Interactions of proteins with nucleic acids
2. The Zinc Finger Motif

- This important group of DNA-binding motifs adds one or more zinc atoms as structural components.
- Although all such zinc-coordinated DNA-binding motifs are called zinc fingers, this description refers only to their appearance in schematic drawings dating from their initial discovery.
- There are now 4 broad classes
  - **ßßa zinc finger:** It is characterized by the sequence CX2-4C....HX2-4H, where C = cysteine, H = histidine, X = any amino acid. In the 3D structure, two cysteine residues and two histidine residues interact with a zinc ion
  - **Hormone receptor zinc finger:** Its consensus sequence is CX2CX13CX2CX14-15CX5CX9CX2C. The first four cysteine residues bind to a zinc ion and the last four cysteine residues bind to another zinc ion.
  - **Gal4 zinc finger.** It has the consensus sequence CX2CX6CX5-6CX2CX6C. The yeast's Gal4 contains such a motif where six cysteine residues interact with two zinc ions.
  - **Loop-sheet-helix -** characterized by p53



This protein belongs to the Cys-Cys-His-His family of zinc finger proteins, named after the amino acids that grasp the zinc. (A) Schematic drawing of the amino acid sequence of a zinc finger from Xenopus. (B) The three-dimensional structure of this type of zinc finger is constructed from an antiparallel β sheet (amino acids 1 to 10) followed by an α helix (amino acids 12 to 24). The four amino acids that bind the zinc (Cys 3, Cys 6, His 19, and His 23) hold one end of the α helix firmly to one end of the β sheet. (from M.S. Lee et al., *Science* 245:635–637, 1989.)

Summary of the Zn-binding proteins



17. ββα-zinc finger (1aay)

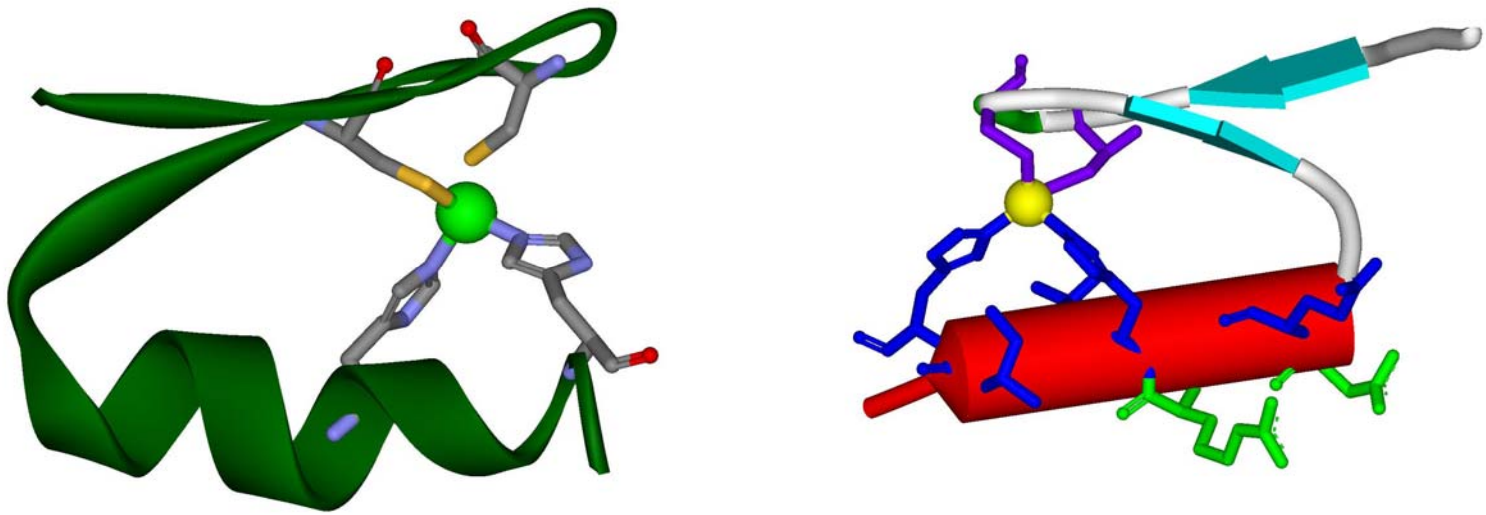
18. Hormone receptor (2nll)

19. Loop-sheet-helix (1tsr)

20. Gal4-type (1d66)

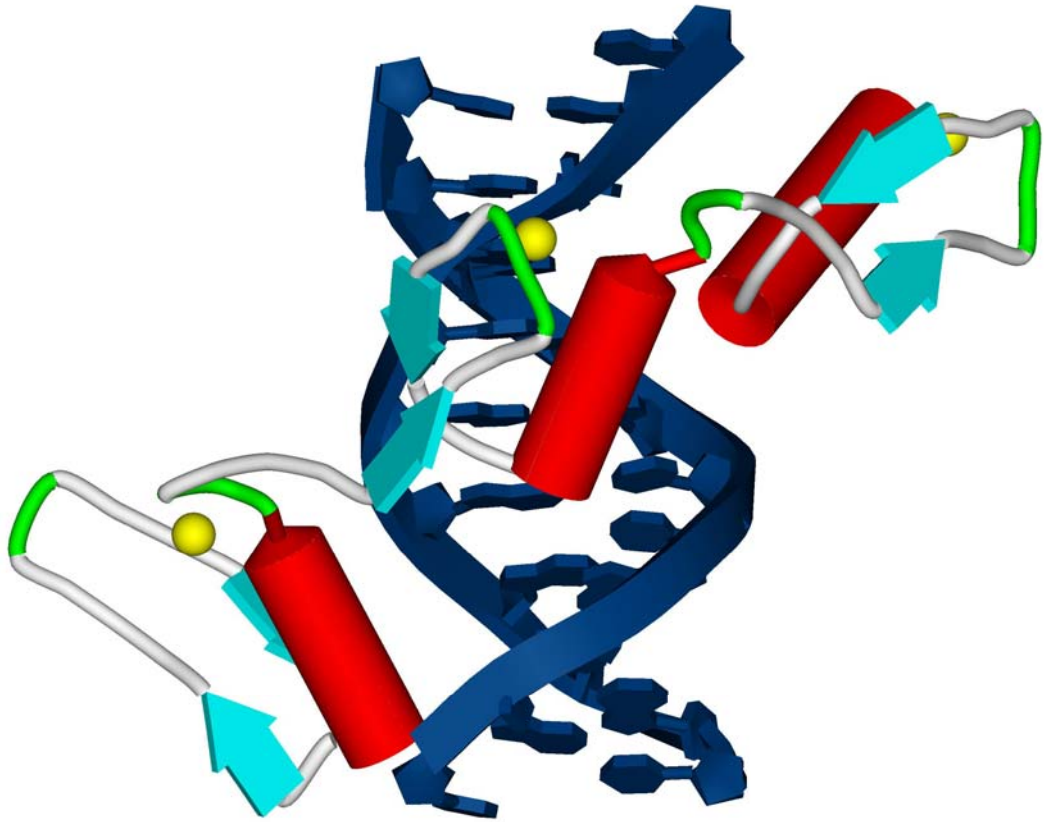
Class one zinc fingers (The ßßa zinc-finger family)

1. **Function:**
    1. The DNA-binding motif is found as part of transcription regulatory proteins.
2. **Structure:**
    1. One of the most abundant DNA-binding motifs.
    2. classic structure consisting of an approximately 30 residue motif coordinating one zinc ion via two cysteines and two histidine residues.
    3. Proteins may contain more than one finger in a single chain; each motif consists of 2 anti-parallel beta-strands followed by an alpha-helix.
    4. A single zinc ion is tetrahedrally coordinated by conserved histidine and cysteine residues, stabilising the motif



One of the zinc fingers from the mouse Zif268 protein (1ZAA)

Class 1 zinc finger bound to DNA

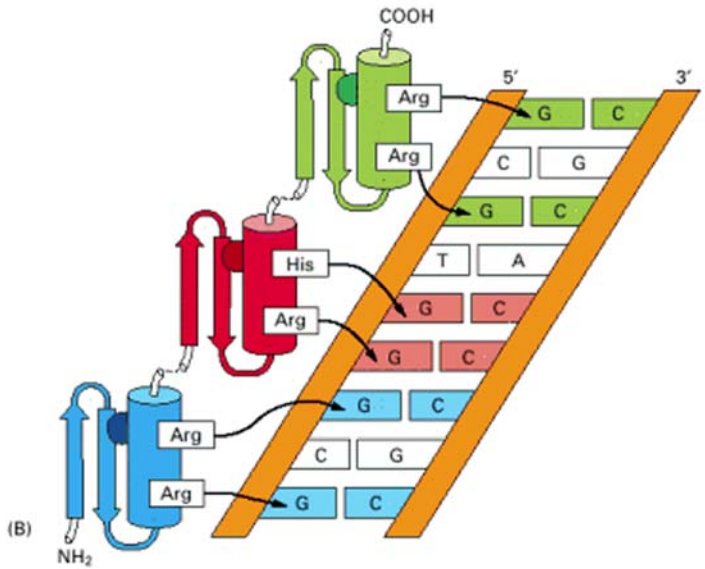


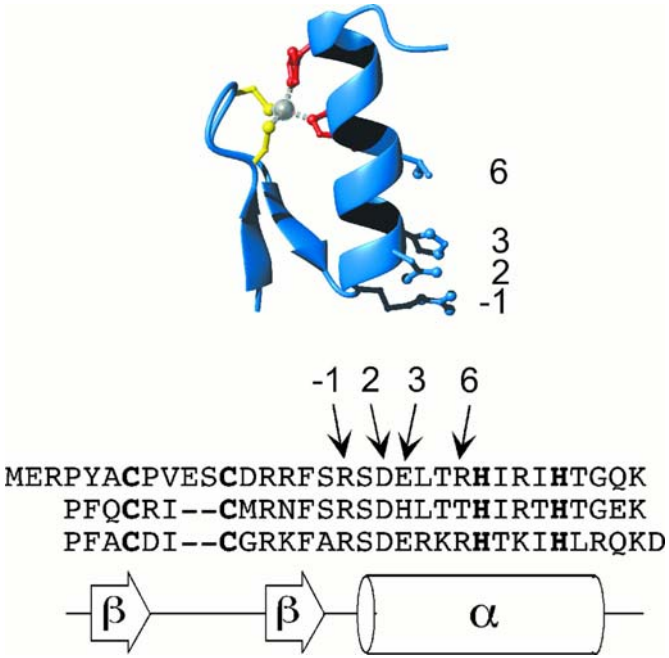
Binding of the Mouse Zif268 protein/DNA complex to DNA:

- protein binds as a monomer - unlike H-T-H and ß-sheet examples.
- Fingers bind to 3 base-pair sub-sites and specific contacts are mediated by amino acids in positions -1, 2, 3 and 6 relative to the start of the alpha-helix.
- fingers lie in the major groove and each spans about 3 bases.
    - here, successive fingers are positioned some 96^{o} of screw-like motion relative to their immediate neighbours.
    - base specific contacts are made by residues at the N-terminal end of α-helices
    - Majority of contact are with bases and BB of one strand

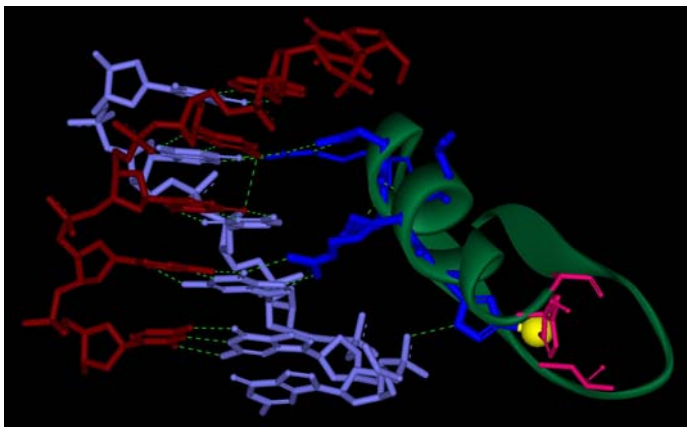
How does a Zinc finger recognize DNA?

The three fingers of **Zif268** bound to a specific DNA site have similar amino acid sequences and contact the DNA in similar ways. (N. Pavletich and C. Pabo, *Science* 252:810–817, 1991.)

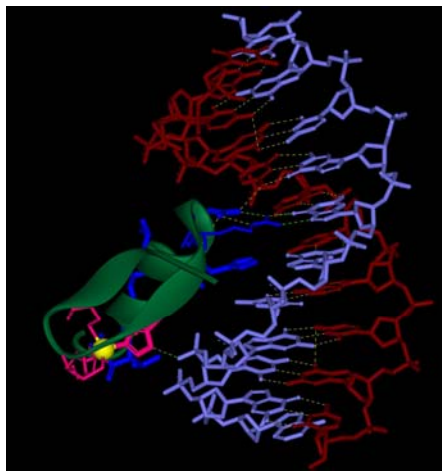




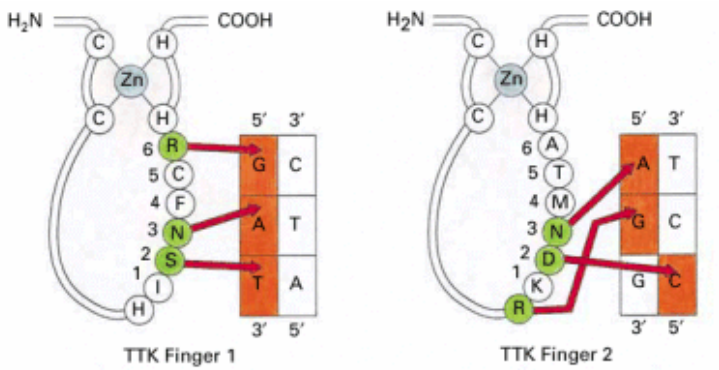
Zn-finger 1



Zn-finger 2



Sequence-specific interactions between different six zinc fingers and their DNA recognition sequences.



- all six Zn fingers have the same overall structure, but each binds to a different DNA sequence.
- The numbered amino acids from the α helix that recognizes DNA and those that make sequence-specific DNA contacts are colored *green*.
- Bases contacted by protein are *orange.*
- Arginine-guanine contacts are common
- guanine can also be recognized by serine, histidine, and lysine.
- the same amino acid (for example serine) can recognize more than one base.

Class two zinc fingers (Hormone receptor family)

Function:

Modulates transcription on binding hormone, i.e., hormone receptors

example – via the glucocorticoid receptor, cells detect and respond transcriptionally to the glucocorticoid hormones produced in the adrenal gland in response to stress.

Structure:

divided into three functional domains for **ligand binding**, **DNA binding** and **transcriptional regulation.**

DNA-binding is conducted as a pseudosymmetric homo- or heterodimer.

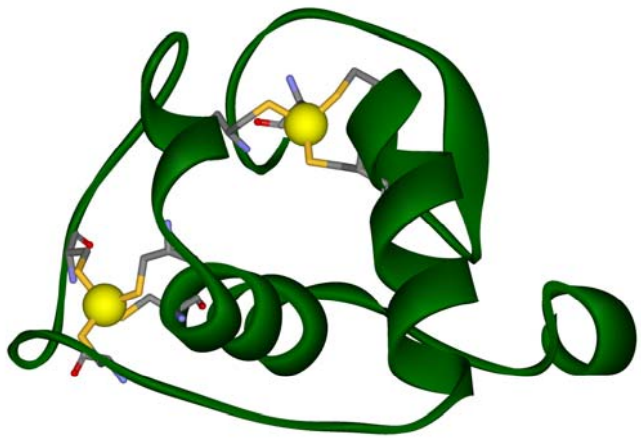
~70 residue DNA-binding domain is composed of two Zn-coordinating modules which pack against each other into a compact globule.

Recognition module: consists of a loop followed by an amphipathic alpha-helix.

The loop is often interrupted by a small antiparallel beta-sheet.

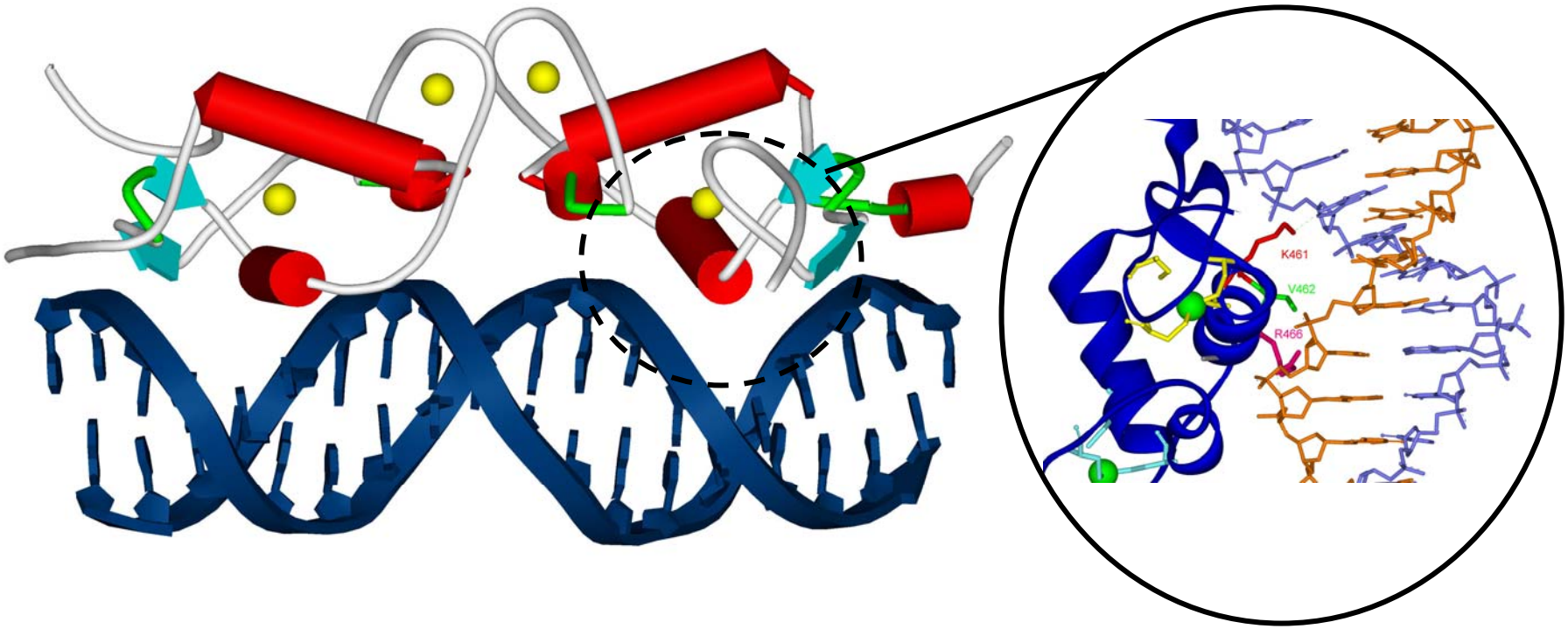
Dimerization module: similar structure to the DNA-binding module.

Zn ions are tetrahedrally coordinated by four cysteine residues; two each from the loops and helices.



Two zinc coordination sites in the rat glucocorticoid receptor (1GLU)

Class 2 zinc finger bound to DNA



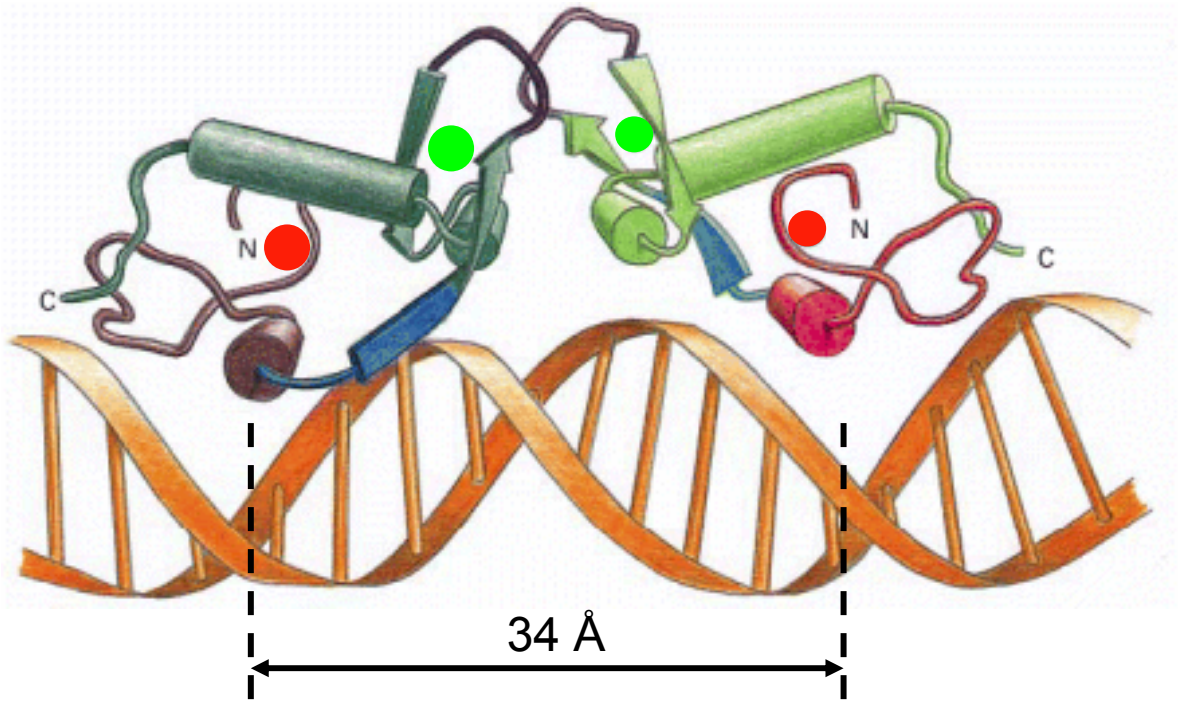
DNA binding domain of the glucocorticoid receptor.

- the N-terminal helix lies in the major groove.
-receptors bind as dimers hence dyad symmetry shown by DNA sequence.

Most inter-subunit contacts are made by loop residues between the Zn coordinating residues.

- residues near the N-terminus of this helix determine binding specificity.
-Molecular details of interaction of glucocorticoid receptor:
    - K461; terminal amino group donates H-bond to acceptors on G6'.
    - Val462; is in Van der Waals contact with T4.
    - Arg466; both terminal amino groups donate H-bonds to acceptors (O6 & N7) on G3

What are the roles of the Zn^{2+} atoms in the glucocorticoid receptor structure?



Each zinc finger domain contains two atoms of Zn (indicated by the *coloured spheres);*
one stabilizes the DNA recognition helix (shown in *red*)
one stabilizes a loop (shown in *green)* involved in dimer formation.
Like the helix-turn-helix proteins the two recognition helices of the dimer are held apart by 34 Å or one turn of B-form DNA. (B.F. Luisi et al., *Nature* 352:497–505, 1991.)

Class three zinc fingers
Gal4 family

Function:

- found exclusively in yeast transcriptional activators eg GAL4.

- a transcription activator of genes required for galactose and melibiose catabolism.

DNA-binding region in the N-terminus of GAL4

Structure:

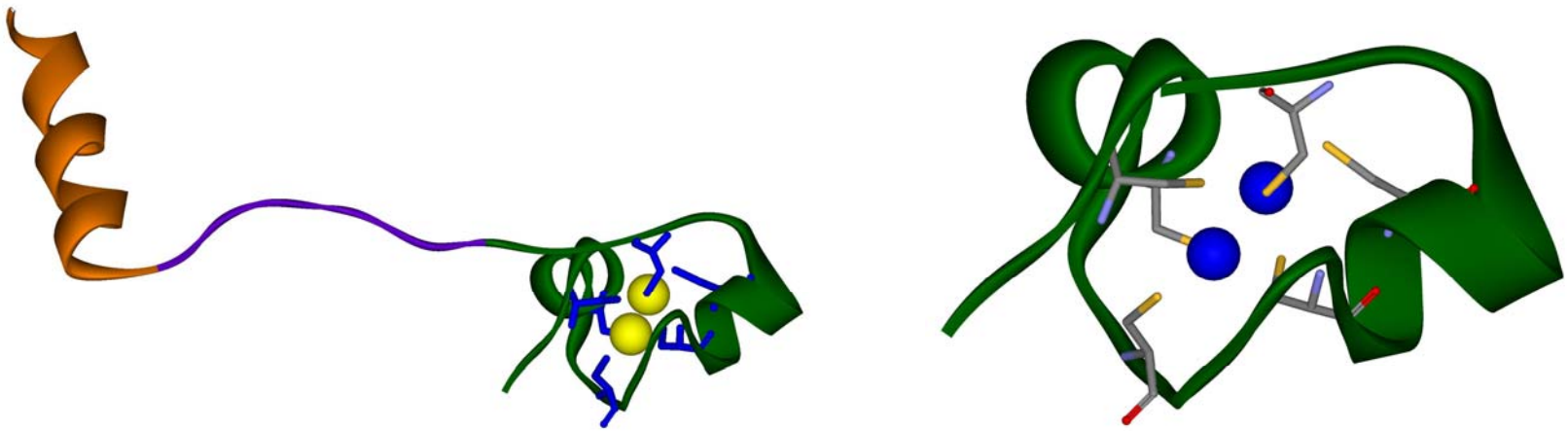
~65 residue region which binds as a symmetrical **dimer**.

Each subunit folds into three separate regions:

Zn-binding : ~35 residue segment in the N-terminus consisting of two short alpha-helices. Two zinc ions are tetrahedrally coordinated by six conserved cysteine residues, thus stabilising the formation of a compact domain.

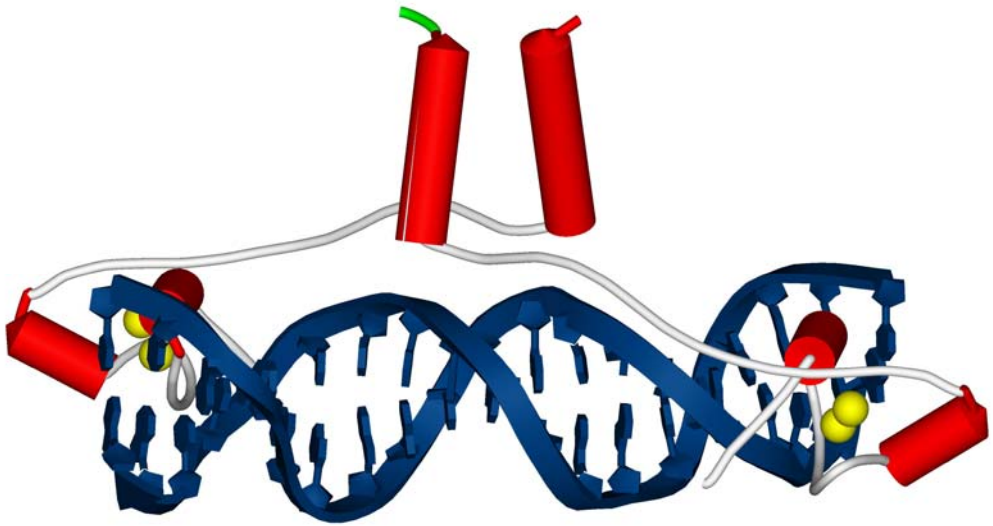
Dimerization : ~20 residue C-terminal tail containing a single amphipathic alpha-helix. The helices from the two subunits combine to form the dimer interface by packing into a parallel coiled coil.

Linker : ~10 residue section which connects the Zn-binding and dimerization regions.



Saccharomyces cerevisiae GAL4 zince binding region is shown. (A) The monomer, (B) The Zinc-binding domain The two zinc atoms are coordinated by 6 cysteine residues

Gal4 bound to DNA



File: 1dd6

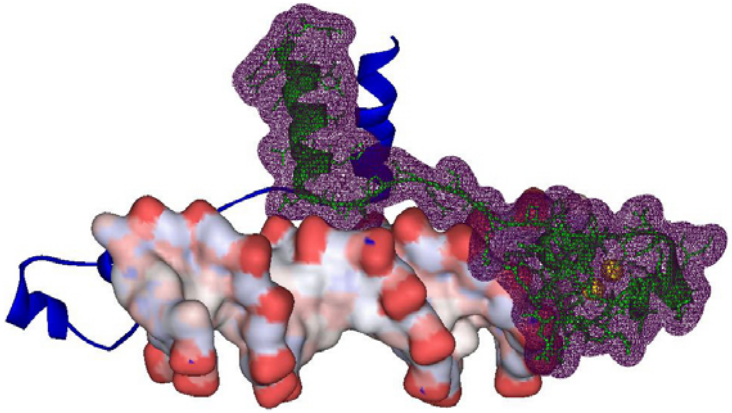
Binding:
proteins bind as dimers to a 17 bp palindromic site that shows two-fold (dyad) symmetry.

Helix 1 of the Zn-binding motif is inserted in the major groove at the outer end of the half-site.
Supporting DNA backbone contacts are made by the remainder of the region.

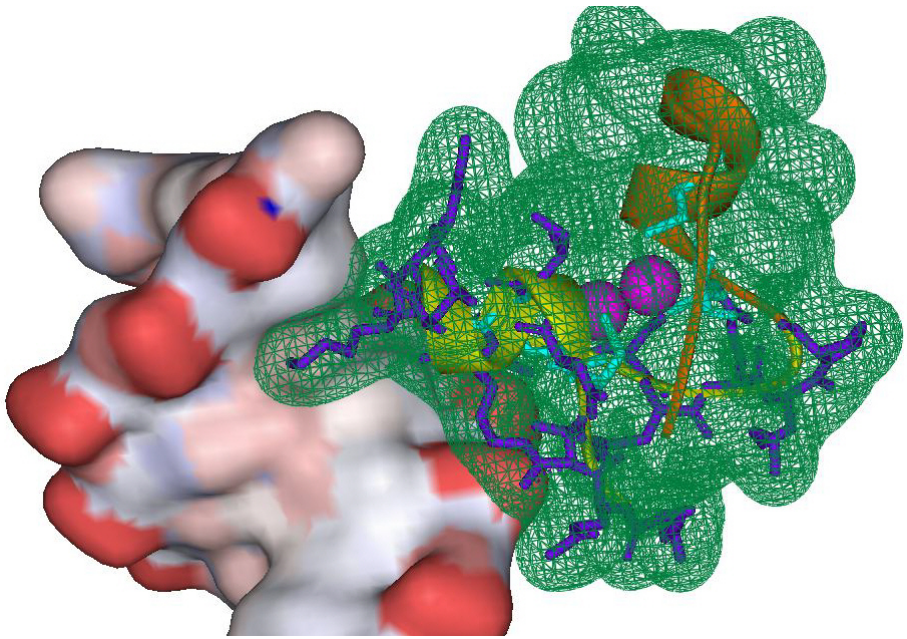
The dimerization helices interact with the phosphate backbone near the centre of the binding site.

No interactions are made by the linker.
This suggests possible binding to sites of different lengths by varying the distance between the two DNA-binding regions.

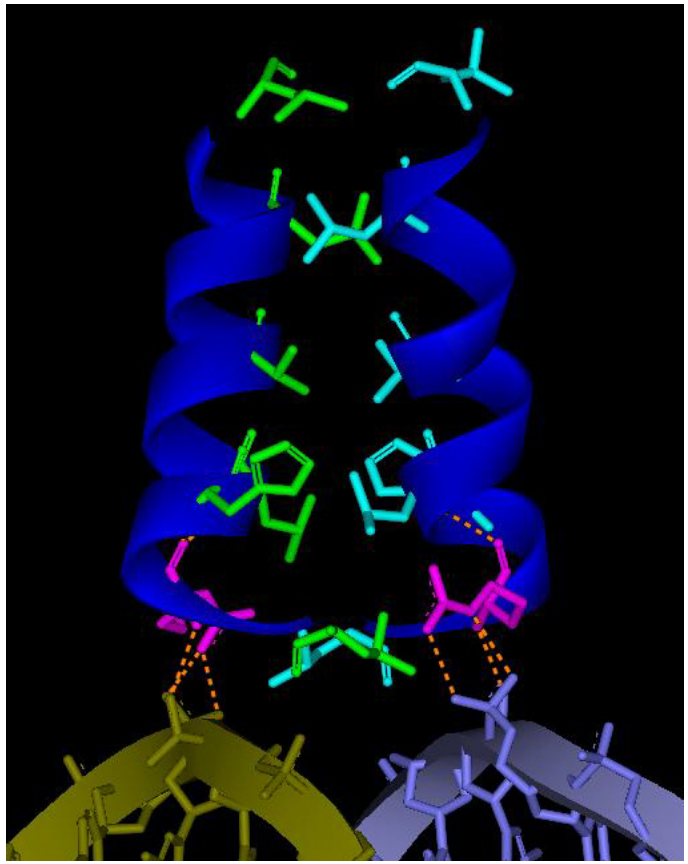
Structure of GAL4/DNA complex



The connecting loop does not contact the DNA



The recognition helix positions Lys 17 & Lys 18
in proximity of 2 conserved C's in the DNA, with which they base pair

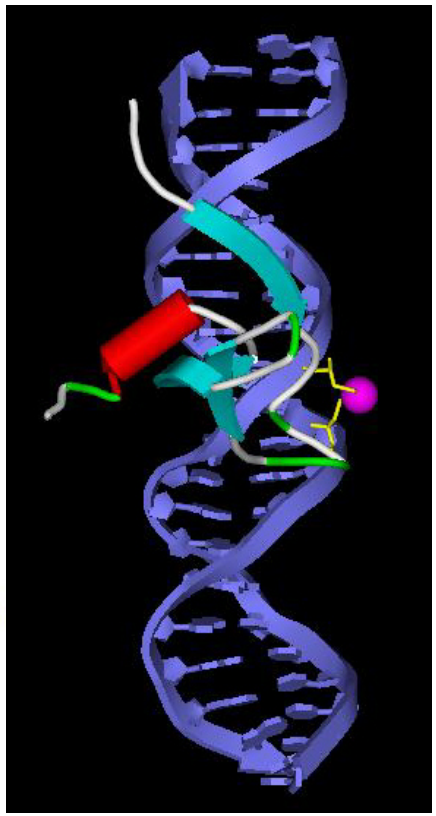


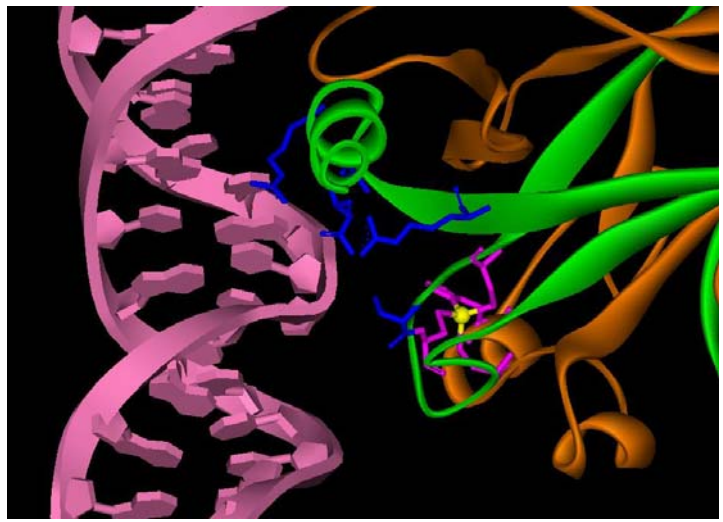
Dimerization domain

1. Two amphipathic α -helices
(polar one side hydrophobic on the other); one from each monomer which interact to form a parallel coiled-coil structure (leucine zipper?)
2. interaction between main chain nitrogens of Arg 51 residues and DNA backbone which occurs across minor groove at centre of recognition site and narrows it by about 1A.

Loop-sheet-helix family

-represented by the DNA-binding region of the protein p53, a transcriptional activator implicated in tumor suppression The protein functions as a tetramer with each subunit contacting a separate 5 bp recognition sequence positioned one after another.
-Regions outside the DNA-binding motif make the intersubunit interactions.
- As the name indicates, the DNA-binding domain consists of a loop leading out of the main body of the protein, followed by a small ß sheet, an a helix and then another loop that leads back into the protein.
-Three cysteines and a histidine in the two loop regions coordinate the zinc ion.
- The protein binds with the a helix in the DNA major groove and the loops in the minor groove, although the latter are not thought to confer specificity.





Alpha helix in groove 4 different ways

DNA complexes of four of the domains. a, The engrailed homeodomain in complex with DNA containing a TAATX 'core consensus' sequence. In the crystal structure, the sixth residue of the recognition helix (Ile) contacts the final T of this core sequence, and the tenth residue (Asn) contacts the penultimate A; minor-groove contacts to the first two core base pairs are made by residues in the N-terminal arm. All these residues, as well as the ninth recognition-helix residue (Gln), are highlighted. b, The 434 repressor DNA-binding domain (R1-69), bound to a half-site with the ACAA consensus sequence. The twofold axis relating half sites is shown by a dashed line. Highlighted residues, critical for base-pair interactions, include the last residue in the turn (Thr), and residues 1, 2, 3 and 6 in helix 3 (Gln, Gln, Ser, Gln). c, A Zn-finger bound to DNA. The N and C termini of the module project in opposition directions along the major groove. The first conserved His residue bridges between the Zn^{2+} and a DNA phosphate. The three base pairs that are contacted directly in the Zif 268 structure are highlighted. Note that the phosphate liganded by the His 'belongs' to the central base pair of the next three-base-pair set. d, The glucocorticoid receptor DNA binding domain recognizes a glococorticoid response element (GRE) with consensus sequence AGAACAXXXTGTTCT. One half of a GRE is shown, together with a bound domain; the twofold axis relating the half-sites is shown as a dashed line.