Winged helix proteins

Ketan S Gajiwala* and Stephen K Burley*†

The winged helix proteins constitute a subfamily within the large ensemble of helix-turn-helix proteins. Since the discovery of the winged helix/fork head motif in 1993, a large number of topologically related proteins with diverse biological functions have been characterized by X-ray crystallography and solution NMR spectroscopy. Recently, a winged helix transcription factor (RFX1) was shown to bind DNA using unprecedented interactions between one of its eponymous wings and the major groove. This surprising observation suggests that the winged helix proteins can be subdivided into at least two classes with radically different modes of DNA recognition.

Addresses
*Laboratories of Molecular Biophysics, †Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
†e-mail: burley@rockvax.rockefeller.edu

Current Opinion in Structural Biology 2000, **10**:110–116

0959-440X/00/$ – see front matter © 2000 Elsevier Science Ltd. All rights reserved.

Abbreviations

ADAR1	double-stranded RNA adenosine deaminase
DBD	DNA-binding domain
HFH	HNF-3/fork head
HNF-3	hepatocyte nuclear factor-3
HTH	helix-turn-helix
PDB	Protein Data Bank

Introduction

Functional studies of the winged helix proteins began with the discovery of the hepatocyte nuclear factor-3 (HNF-3) family of liver-specific transcription factors [1]. Proteins from the HNF-3 family share a highly conserved DNA-binding region with the *Drosophila* homeotic fork head proteins, which are involved in proper formation of terminal structures in fly embryos. Hence, winged helix proteins are also referred to as belonging to the HNF-3/fork head (HFH) family [2]. The co-crystal structure of the DNA-binding domain (DBD) of an HNF-3γ–DNA complex was the first to show the mechanism of DNA recognition by a winged helix protein [3]. Subsequently, X-ray and solution NMR structures of a number of winged helix proteins have been determined. This review focuses on more recent structural characterizations of winged helix proteins, their versatility in DNA recognition and diversity of biological function.

Structure of the winged helix motif

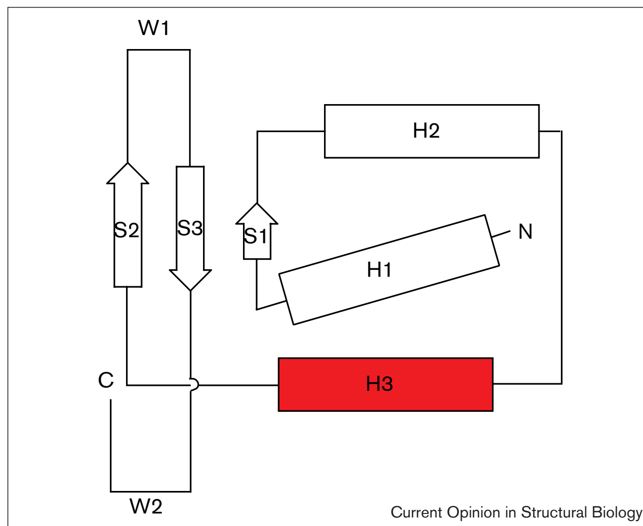
Topologically, the winged helix motif is a compact α/β structure consisting of two wings (W1 and W2), three α helices (H1, H2 and H3) and three β strands (S1, S2 and S3), arranged in order H1-S1-H2-H3-S2-W1-S3-W2 (Figure 1). The N-terminal half of the motif is largely helical, whereas the C-terminal half is composed of two of the three strands forming the twisted antiparallel β sheet and the two large loops or wings, W1 and W2 [3]. Wing W1 connects strands S2 and S3, and wing W2 extends from strand S3 to the C terminus of the DBD. These loops flank helix H3 like the wings of a butterfly, inspiring the name winged helix motif (Figure 1). HFH proteins are members of the helix-turn-helix (HTH) superfamily, but differ from canonical HTH proteins in the length of the ‘turn’ connecting helices H2 and H3 [4]. This subtle structural variation eliminates stereochemical restrictions on the precise apposition of helices H2 and H3. Among winged helix proteins, the angle between H2 and H3 ranges from 100° in the biotin operator repressor protein BirA [5] to 150° in transcription factor DP2 [6••]. In contrast, this angle is typically about 120° in HTH proteins [7]. Conserved nonpolar residues from every secondary structure element interdigitate to form the hydrophobic core of the protein (Figure 2).

Winged helix proteins frequently exhibit an exposed patch of hydrophobic residues. In HNF-3γ, six hydrophobic sidechains form a solvent-accessible rosette-like structure [3]. These sidechains, conserved among HFH homologs, are thought to mediate protein–protein interactions. Insight into the biological roles played by the hydrophobic rosette is provided by the results of structural and site-directed mutagenesis studies of the T4 transcription factor MotA [8], a bacteriophage winged helix protein. When an exposed pair of acidic/hydrophobic residues (aspartic acid and phenylalanine) in MotA is changed, T4 growth is compromised and the double mutant is defective in transcriptional activation *in vitro*. A similar effect has been observed in the context of the MDM2–p53 interaction, which is mediated by bulky hydrophobic residues on the surface of p53 [9].

DNA recognition by canonical winged helix proteins

HNF-3γ is a liver-specific transcription factor that plays an important role in cell differentiation and tissue-specific gene expression [10]. Its co-crystal structure showed that helix H3, the recognition helix of the HTH motif, is presented to the major groove of a duplex oligonucleotide derived from the transthyretin gene promoter (Figure 3) [3]. In total, 14 protein–DNA contacts are distributed throughout the length of the polypeptide chain (Figure 2). Five of the observed interactions, including all the specificity determining contacts, map to the recognition helix within the major groove, with another four involving wing W2 and the minor groove. Most of these interactions are mediated by polar sidechains, either directly or through bridging water molecules. Binding of HNF-3γ deforms the DNA by inducing a 13° bend and narrowing the major groove where it embraces helix H3. Structures of nine apo forms of winged helix domains have been published to date. For many of these proteins, there is circumstantial

Figure 1

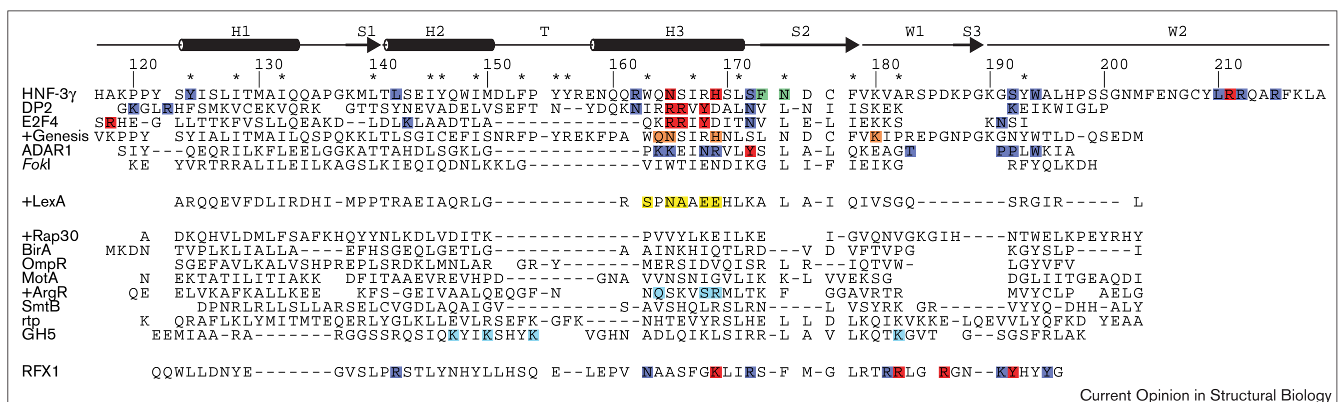


Topology of the winged helix fold.

evidence suggesting that the mode of DNA recognition is comparable to that exhibited by the HNF-3γ–DNA complex. For example, LexA is a transcriptional regulator of SOS genes in *Escherichia coli*. The solution NMR structure of its DBD has been determined [11] and a computational model for LexA interactions with DNA suggests an HNF-3γ-like mode of DNA binding [12].

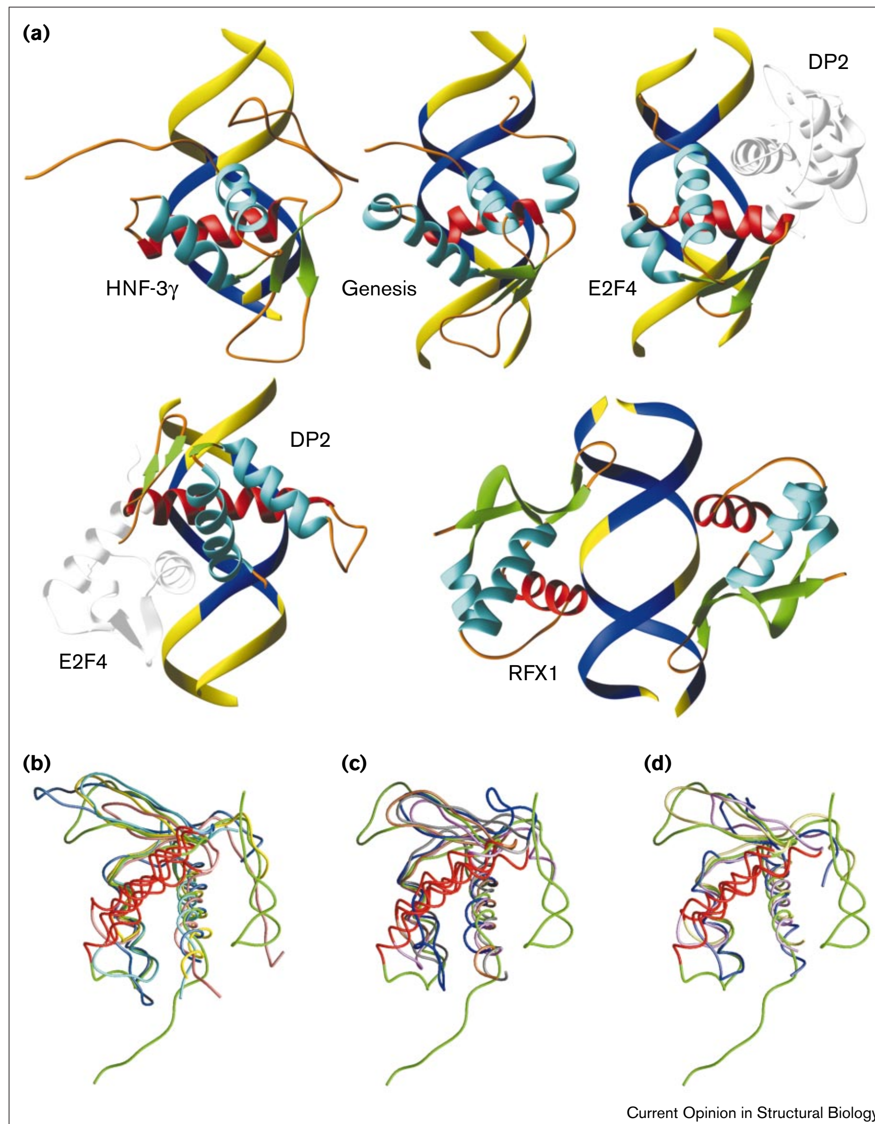
The E2F family of transcription factors controls genes involved in growth and DNA replication [13]. DNA binding by the E2F proteins is enhanced upon heterodimerization with members of the DP family, a distant relative of E2F. The co-crystal structure of the E2F4– DP2–DNA heteromeric complex demonstrated that both E2F4 and DP2 are winged helix proteins, albeit lacking wing W2. The mode of DNA recognition by each member of the heterodimer is very similar to that of HNF-3γ (Figure 3) [6••]. Conserved Arg-Arg-X-Tyr-Asp motifs found on the H3 recognition helices of both E2F4 and DP2 are presented to the major groove of the DNA, each making quasi-symmetric contacts with one half of a palindromic DNA target (5′-CGCGCG-3′). Unlike HNF-3γ, E2F4 contains an N-terminal helical extension and a shortened helix H3. An arginine sidechain from the E2F4 N-terminal extension is inserted into the minor groove of the DNA near the T-rich region of the consensus binding site, upstream of the 5′-CGCGCG-3′ palindrome. DP2, on the other hand, possesses an insertion between helices H2 and H3, extending each helix by two turns. In both the HNF-3γ and E2F4–DP2 co-crystal structures, protein–DNA contacts are primarily mediated by polar groups. Unlike HNF-3γ, however, the E2F4–DP2 complex does not contain any water molecules in the protein–DNA interface. Another significant difference in the behavior of these proteins concerns the roles of their respective wings, which are not central to DNA binding by the E2F4–DP2 heterodimer — they are far from the double helix. In the HNF-3γ complex, residues from both wings make important interactions with DNA. Moreover, there is an extensive protein–protein interface between E2F4 and DP2 formed by helices H1 and H3 of both protomers. Dimerization upon DNA utilizes reciprocal interactions, with the recognition helix of one protein packing between helices H1 and H2 of the other. This observation demonstrates the facility with which the winged helix fold can perform functions other than sequence-specific DNA binding. Additional support for this conclusion is provided by the *Fok*I–DNA co-crystal structure [14]. *Fok*I is a restriction endonuclease that recognizes a DNA consensus sequence, but cleaves nonspecifically a short distance from its binding site [15]. The DNA recognition domain of *Fok*I consists of

Figure 2



Structure-based sequence alignment of winged helix proteins. Dashes and spaces denote insertions and deletions within the alignment, respectively. Secondary structure (top, where T denotes turn region), residue numbering and residues comprising the hydrophobic core of the DBD (indicated by an asterisk) refer to HNF-3γ (see text). A plus (+) indicates that the structure was determined using solution NMR spectroscopy. Residue color coding: red, specific DNA contacts; blue, nonspecific DNA contacts; green, water-mediated DNA contacts; orange, mapped to the protein–DNA interface by solution NMR spectroscopy; yellow, DNA contacts according to LexA docking calculation; and cyan, judged to be near DNA using biochemical assays.

Figure 3



Structures of winged helix proteins.
(a) Winged helix proteins recognizing B-form DNA. RFX1 binds to the X-box as a symmetric dimer. E2F4 and DP2 bind DNA as a heterodimer. For clarity, E2F4 and DP2 are shown separately, with their heteromeric partner in gray. E2F4 and DP2 recognize contiguous sites, whereas the RFX1 homodimer binds to separated half sites within the X-box. Color coding: red, recognition helix H3; cyan, α helix; green, β strand; orange, random-coil region; yellow, DNA backbone; blue, target DNA sequence.
(b–d) Superposition of apo protein structures of winged helix DBDs on HNF-3γ (green). All recognition helices are colored red. (b) Yellow, Rap30 (PDB code 1bby); blue, replication terminator protein (PDB code 1bm9); pink, biotin repressor BirA (PDB code 1bia); cyan, globular domain of histone H5 (GH5; PDB code 1hst). (c) Gray, *Fok*I (PDB code 2fok); brown, ADAR1 (PDB code 1qbj); lilac, MotA (PDB code 1bja); blue, OmpR (PDB code 1opc). (d) Blue, ArgR (PDB code 1aoy); lilac, LexA (PDB code 1lea); yellow, SmtB (PDB code 1smt).

three HTH subdomains, denoted D1, D2 and D3. D3 is a winged helix motif, although it makes few contacts with DNA and its ‘recognition’ helix does not interact with the major groove. D3 is thought to be important for protein–protein interactions [14].

Genesis is a 465-residue transcription repressor whose expression is restricted to embryonic stem cells and certain tumor cell lines [16]. It contains a winged helix DBD that has been extensively characterized via solution NMR (Figure 3) [17•,18••,19•,20]. The major structural difference between genesis and HNF-3γ occurs in the ‘turn’ of the HTH segment; in HNF-3γ, helices H2 and H3 are connected by a random-coil segment, whereas genesis contains an additional helix between H2 and H3 (Figure 2) [21]. Solution NMR studies of a genesis–DNA complex demonstrated that genesis uses an HNF-3γ-like mechanism of DNA recognition [18••]. NMR dynamical studies of the free protein have shown that its two wings undergo collective motions [20]. The motions of wing W2 are damped upon protein–DNA complex formation, suggesting that the DNA-bound form of genesis is more ordered than its apo form. Despite contacts with the minor groove of the DNA, the dynamical properties of wing W1 remain unchanged. These results suggest that the interaction of W1 with DNA does not contribute substantially to protein–DNA interactions [18••]. This situation is reminiscent of the HNF-3γ–DNA complex, in which a single protein–DNA contact maps to the base of wing W1. Taken together, these results imply that the most important DNA-binding portion of a winged helix DBD is its H3 recognition helix. If this is true, how can we explain the fact that HNF-3γ and genesis recognize markedly different DNA sequences, when nine out of ten

residues in the recognition helix of genesis are identical to the corresponding residues in HNF-3γ (the single substitution being serine to asparagine; Figure 2).

This apparent paradox has been addressed by Overdier *et al.* [22] in studies of HFH proteins. Replacing the 20-residue segment adjacent to the N terminus of the recognition helix of HNF-3β with corresponding sequences from HFH-1 allowed the HNF-3β–HFH-1 chimera to recognize the DNA target of HFH-1. Similar observations were made using chimeras containing the HNF-3β recognition helix preceded by 20 residues derived from HFH-2. This effect is thought to be responsible for the observed differences in the DNA-binding specificities of HNF-3γ and genesis [21]. Both proteins possess an identical pentapeptide (Phe-Pro-Tyr-Tyr-Arg) in the turn between helices H2 and H3 (Figure 2); however, the three preceding residues and five residues following this segment are very different and may be involved in modulating the binding specificity of helix H3. This assertion is supported by the results of solution NMR experiments with the winged helix motif found in Rap30, the small subunit of the heterodimeric transcription factor IIF [23•]. The structure of this apo DBD is very similar to that of HNF-3γ. When incubated with an appropriate eukaryotic gene promoter, significant changes in chemical shifts were observed for residues at the head of helix H2, within the HTH turn and solvent-accessible residues of helix H3.

Double-stranded RNA adenosine deaminase (ADAR1) induces codon changes by deaminating adenine to inosine, which behaves like guanine in pre-mRNAs — giving rise to alternative forms of translated proteins [24]. ADAR1 also binds to the Z conformation of DNA through its N-terminal Zα domain. The co-crystal structure of the Zα domain of human ADAR1 bound to a short Z-DNA fragment showed that it is a winged helix protein (Figure 3) [25••]. Zα binds to Z-DNA using residues from recognition helix H3 and the C-terminal base of wing W1. The only protein–DNA major groove contact is a van der Waals interaction between a tyrosine sidechain and C8 of a guanine. This interaction is thought to specify binding to Z-DNA because it requires the *syn* conformation of guanine, as found in Z-DNA. This observation also implies that recognition of Z-DNA depends on the conformation of the double helix, whereas winged helix protein specificity for B-DNA depends on the nucleotide sequence. The solution NMR structure of the ADAR1 apo protein is not significantly different from its DNA-bound form [26,27•], effectively ruling out an induced-fit mode of Z-DNA recognition.

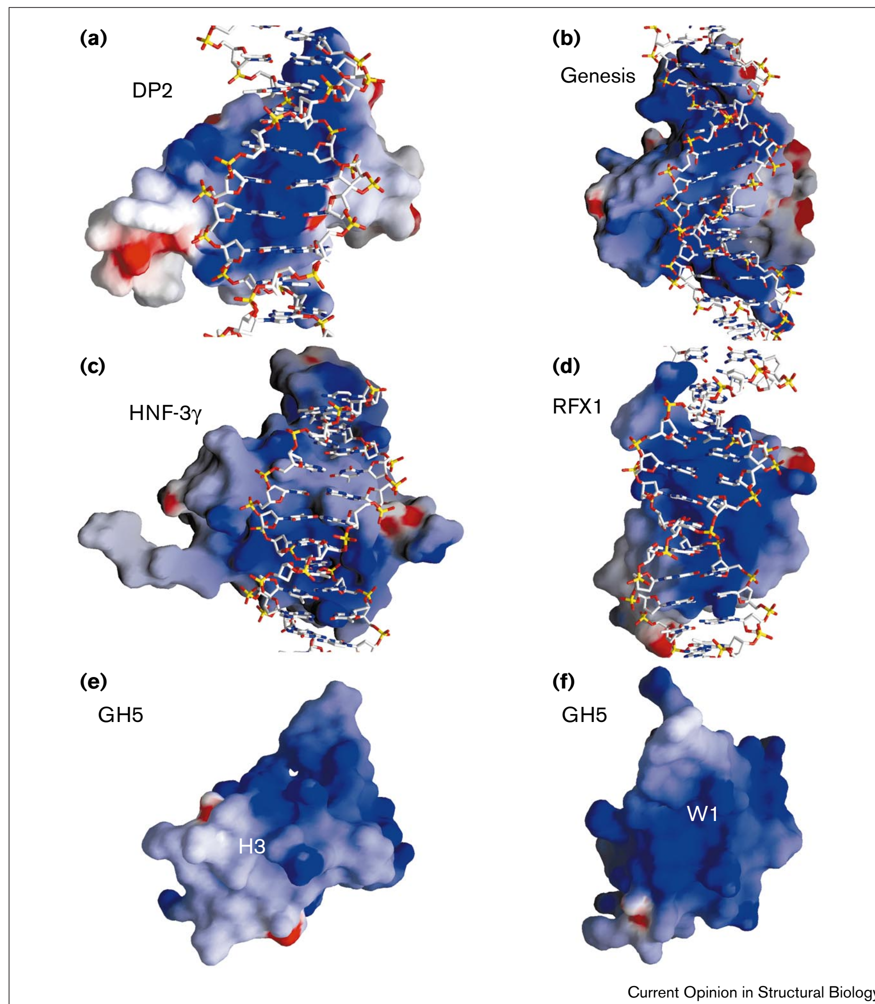
Atypical DNA recognition by a winged helix protein

The winged helix protein RFX1 uses an unprecedented wing-based strategy to recognize the bipartite DNA sequence known as an X-box (5′-CGTTACCATGGTAACG-3′) (Figure 3) (KS Gajiwala *et al.*, unpublished data). RFX1 is a cellular transcription factor that regulates the expression of a number of biologically important gene products [28,29]. The co-crystal structure of the RFX1–DNA complex revealed that wing W1 of RFX1 makes most of the contacts with DNA via its major groove. The so-called recognition helix (H3) overlies the minor groove, where it makes a single DNA contact using a lysine sidechain. All other well-characterized winged helix DBDs use their H3 recognition helices to bind the major groove. Like HNF-3γ, most RFX1–DNA interactions involve polar residues making direct or water-mediated interactions. Two van der Waals interactions with thymine methyl groups explain why RFX proteins also bind to certain methylated DNA targets.

Biochemical studies have shown that two molecules of RFX1 bind cooperatively to a single X-box [30]. In the RFX co-crystal structure, two copies of the protein make a symmetric complex in which the two monomers have no intermolecular interactions (distance of closest approach ≅10 Å). In the absence of direct protein–protein interactions, cooperativity appears to be mediated by protein-induced DNA deformation. Binding of one RFX molecule causes the minor groove to widen by over 3 Å with a reciprocal narrowing of the major groove on the opposite face, thereby inducing a better grip on wing W1 of the other copy of RFX, which binds to the other half site of the X-box (Figure 3a).

This surprising mode of RFX1 DNA recognition suggests that winged helix proteins may fall into two functional classes. Figure 4 illustrates the surface electrostatic properties [31] of DP2, genesis, HNF-3γ and RFX1. In every case, the most basic face of the protein is responsible for DNA binding. For HNF-3γ, DP2, E2F4 and genesis, this basic surface electrostatic feature corresponds to the H3 face, which interacts with the major groove of DNA. In RFX1, the W1 face displays the largest number of basic residues and is responsible for DNA binding. When we consider all winged helix DBD structures determined in the absence of DNA, transcription factor SmtB [32•] and the linker histone protein GH5 [33] are outliers. These proteins exhibit clustering of basic residues on their wing W1 surfaces and largely neutral H3 faces (Figure 4), as seen in the RFX1–DNA structure. Thus, SmtB and GH5 may well use their wing W1 surfaces to bind DNA. When these structures are docked onto B-form DNA using the RFX1–DNA complex as a guide, the resulting models are consistent with photocross-linking and chemical modification data. For example, Lys85 of GH5 is maximally protected from chemical modification in chromatin [34]. This residue is equivalent to Arg58 in wing W1 of RFX1. The RFX1-based model of GH5–DNA places Lys85 in close proximity to DNA. We propose that SmtB and GH5 are winged helix proteins that bind DNA using an RFX1-like recognition mechanism.

Figure 4



Surface electrostatic properties of winged helix proteins. Red denotes regions of negative electrostatic potential, blue denotes positive electrostatic potential and white denotes neutral. DNA is drawn as an atomic stick figure. Panels **(a–c)** show similar modes of DNA binding by winged helix proteins. Panel **(d)** shows the radically different mechanism of DNA recognition by RFX1. The basic W1 surface of RFX1 is approximated to the major groove. Panels **(e and f)** show the largely neutral H3 and basic W1 faces of GH5, respectively.

Biological importance of winged helix proteins

While structural biologists were busy characterizing winged helix proteins, considerable progress has been made in studies of the important biological roles of these molecules.

In the mouse, HNF-3α and HNF-3β were shown to be critical developmental regulators. Homozygous HNF-3α null mice exhibit a complex phenotype and die within the first two weeks of life [35,36]. During gut development, HNF-3α acts as a transcriptional activator of a genetic program leading to the orderly differentiation of the pancreatic β cells [37]. Surprisingly, HNF-3β also contributes to this process, but as a transcriptional repressor, not an activator. Mutations in either of these winged helix transcription factors lead to defects in insulin secretion and type 2 diabetes mellitus [38]. The HFH proteins act as developmental potentiators of endodermal tissue by binding to regulatory sequences in the promoters of target genes before these genes are activated in the germ layer (reviewed in [39]).

Winged helix proteins also figure prominently in the life cycle of the nematode *Caenorhabditis elegans*. Normally, these animals live for about two weeks. Reduction in the activity of the *daf-2* gene, which encodes a homolog of insulin-like growth factor, more than doubles the life span of the worm [40]. This dramatic increase in longevity depends on the activity of a gene known as *daf-16*, which encodes an HFH protein. Together, the insulin-like molecule *daf-2* and the transcription factor *daf-16* are thought to regulate the development and aging of *C. elegans* in response to external signals, such as nutrient availability [41]. In humans, insulin performs some *daf-2*-like functions, which are mediated, at least in part, by interfering with the binding of HNF-3 proteins to a DNA target known as the insulin response sequence [42].

Conclusions

X-ray crystallographic and solution NMR studies of winged helix proteins and their complexes with DNA have shown that the motif is extremely versatile. These proteins exhibit two different modes of DNA binding and appear to

be able to recognize both specific sequences in B-form DNA and distinct double helical conformations (i.e. B-form versus Z-form). Both monomeric, homodimeric and heterodimeric protein–DNA complexes have been characterized. It is also clear that this motif can participate in protein–protein interactions. The winged helix proteins play important roles in embryogenesis, development and aging. In the near future, we can expect to understand better the structural bases of alternative modes of DNA recognition by these and, possibly, other HTH proteins.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Costa RH, Grayson DR, Darnell JE Jr: **Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes.** *Mol Cell Biol* 1989, **9**:1415-1425.

2. Weigel D, Jurgens G, Kuttner F, Seifert E, Jackle H: **The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo.** *Cell* 1989, **57**:645-658.

3. Clark KL, Halay ED, Lai E, Burley SK: **Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5.** *Nature* 1993, **364**:412-420.

4. Brennan RG, Matthews BW: **The helix-turn-helix DNA binding motif.** *J Biol Chem* 1989, **264**:1903-1906.

5. Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, Matthews BW: ***Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains.** *Proc Natl Acad Sci USA* 1992, **89**:9257-9261.

6. Zheng N, Fraenkel E, Pabo CO, Pavletich NP: **Structural basis of •• DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP.** *Genes Dev* 1999, **13**:666-674.

The structure of E2F–DP shows the roles played by winged helix motifs in protein–protein interactions and cognate DNA recognition.

7. Brennan RG, Takeda Y, Kim J, Anderson WF, Matthews BW: **Crystallization of a complex of cro repressor with a 17 base-pair operator.** *J Mol Biol* 1986, **188**:115-118.

8. Finnin MS, Cicero MP, Davies C, Porter SJ, White SW, Kreuzer KN: **The activation domain of the MotA transcription factor from bacteriophage T4.** *EMBO J* 1997, **16**:1992-2003.

9. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP: **Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.** *Science* 1996, **274**:948-953.

10. Lai E, Clark KL, Burley S, James J, Darnell E: **Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: a family of transcription factors of diverse biological function.** *Proc Natl Acad Sci USA* 1993, **90**:10421-10423.

11. Fogh RH, Ottleben G, Ruterjans H, Schnarr M, Boelens R, Kaptein R: **Solution structure of the LexA repressor DNA binding domain determined by 1H NMR spectroscopy.** *EMBO J* 1994, **13**:3936-3944.

12. Knegtel RM, Fogh RH, Ottleben G, Ruterjans H, Dumoulin P, Schnarr M, Boelens R, Kaptein R: **A model for the LexA repressor DNA complex.** *Proteins* 1995, **21**:226-236.

13. Slansky JE, Farnham PJ: **Introduction to the E2F family: protein structure and gene regulation.** *Curr Top Microbiol Immunol* 1996, **208**:1-30.

14. Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK: **Structure of the multimodular endonuclease FokI bound to DNA.** *Nature* 1997, **388**:97-100.

15. Skowron P, Kaczorowski T, Tucholski J, Podhajska AJ: **Atypical DNA-binding properties of class-IIS restriction endonucleases: evidence for recognition of the cognate sequence by a FokI monomer.** *Gene* 1993, **125**:1-10. [Published erratum appears in *Gene* 1994, **141**:151.]

16. Sutton J, Costa R, Klug M, Field L, Xu D, Largaespada DA, Fletcher CF, Jenkins NA, Copeland NG, Klemsz M, Hromas R: **Genesis, a winged helix transcriptional repressor with expression restricted to embryonic stem cells.** *J Biol Chem* 1996, **271**:23126-23133.

17. Jin C, Liao X: **Backbone dynamics of a winged helix protein and its • DNA complex at different temperatures: changes of internal motions in genesis upon binding to DNA.** *J Mol Biol* 1999, **292**:641-651.

The authors show that local motions in the apo form of genesis are temperature-dependent. DNA binding damps out the motion of selected portions of the protein, implicating them in DNA contacts.

18. Jin C, Marsden I, Chen X, Liao X: **Dynamic DNA contacts observed •• in the NMR structure of winged helix protein-DNA complex.** *J Mol Biol* 1999, **289**:683-690.

This study shows that residues from both the recognition helix (H3) and wing W1 of the transcription factor genesis (HFH-2) make interactions with DNA. DNA binding induces the formation of a short helix in a region of polypeptide chain that is random coil in the apo protein structure reported in [22].

19. Marsden I, Jin C, Liao X: **Structural changes in the region directly • adjacent to the DNA-binding helix highlight a possible mechanism to explain the observed changes in the sequence-specific binding of winged helix proteins.** *J Mol Biol* 1998, **278**:293-299.

The results of NMR spectroscopy suggest that the DNA-binding specificity of winged helix proteins is determined by residues N-terminal to the recognition helix.

20. Jin C, Marsden I, Chen X, Liao X: **Sequence specific collective motions in a winged helix DNA binding domain detected by ^{15}N relaxation NMR.** *Biochemistry* 1998, **37**:6179-6187.

21. Marsden I, Chen Y, Jin C, Liao X: **Evidence that the DNA binding specificity of winged helix proteins is mediated by a structural change in the amino acid sequence adjacent to the principal DNA binding helix.** *Biochemistry* 1997, **36**:13248-13255.

22. Overdier DG, Porcella A, Costa RH: **The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix.** *Mol Cell Biol* 1994, **14**:2755-2766.

23. Groft CM, Uljon SN, Wang R, Werner MH: **Structural homology • between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly.** *Proc Natl Acad Sci USA* 1998, **95**:9117-9122.

The similarity of the Rap30 DNA-binding domain to HNF-3γ and the linker histones suggests functional similarity in transcription factor IIF binding to promoter DNA during assembly of an RNA polymerase II transcription complex.

24. Melcher T, Maas S, Herb A, Sprengel R, Seeburg PH, Higuchi M: **A mammalian RNA editing enzyme.** *Nature* 1996, **379**:460-464.

25. Schwartz T, Rould MA, Lowenhaupt K, Herbert A, Rich A: **Crystal •• structure of the Zalpha domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA.** *Science* 1999, **284**:1841-1845.

This structure shows that the winged helix motif can recognize the left-handed Z-conformation of DNA, independent of DNA sequence.

26. Schade M, Turner CJ, Lowenhaupt K, Rich A, Herbert A: **Structure-function analysis of the Z-DNA-binding domain Zalpha of dsRNA adenosine deaminase type I reveals similarity to the (alpha + beta) family of helix-turn-helix proteins.** *EMBO J* 1999, **18**:470-479.

27. Schade M, Turner CJ, Kuhne R, Schmieder P, Lowenhaupt K, Herbert A, • Rich A, Oschkinat H: **The solution structure of the zalpha domain of the human RNA editing enzyme ADAR1 reveals a prepositioned binding surface for Z-DNA.** *Proc Natl Acad Sci USA* 1999, **96**:12465-12470.

This apo protein structure shows that the Zα domain of ADAR1 does not change on binding to Z-form DNA, ruling out induced fit.

28. Iwama A, Pan J, Zhang P, Reith W, Mach B, Tenen DG, Sun Z: **Dimeric RFX proteins contribute to the activity and lineage specificity of the interleukin-5 receptor alpha promoter through activation and repression domains.** *Mol Cell Biol* 1999, **19**:3940-3950.

29. Siegrist CA, Durand B, Emery P, David E, Hearing P, Mach B, Reith W: **RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer.** *Mol Cell Biol* 1993, **13**:6375-6384.

30. Cornille F, Emery P, Schuler W, Lenoir C, Mach B, Roques BP, Reith W: **DNA binding properties of a chemically synthesized DNA binding domain of hRFX1.** *Nucleic Acids Res* 1998, **26**:2143-2149.

31. Nicholls A, Sharp K, Honig B: **Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons.** *Proteins* 1991, **11**:281-296.

32. • Cook WJ, Kar SR, Taylor KB, Hall LM: **Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins.** *J Mol Biol* 1998, **275**:337-346.
The first structure of a metal-binding DNA repressor reveals a winged helix motif.

33. Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM: **Crystal structure of globular domain of histone H5 and its implications for nucleosome binding.** *Nature* 1993, **362**:219-223.

34. Thomas JO, Wilson CM: **Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5.** *EMBO J* 1986, **5**:3531-3537.

35. Shih DQ, Navas MA, Kuwajima S, Duncan SA, Stoffel M: **Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3alpha-deficient mice.** *Proc Natl Acad Sci USA* 1999, **96**:10152-10157.

36. Kaestner KH, Katz J, Liu Y, Drucker DJ, Schutz G: **Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression *in vivo*.** *Genes Dev* 1999, **13**:495-504.

37. Vaisse C, Kim J, Espinosa R III, Le Beau MM, Stoffel M: **Pancreatic islet expression studies and polymorphic DNA markers in the genes encoding hepatocyte nuclear factor-3alpha, -3beta, -3gamma, -4gamma, and -6.** *Diabetes* 1997, **46**:1364-1367.

38. Duncan SA, Navas MA, Dufort D, Rossant J, Stoffel M: **Regulation of a transcription factor network required for differentiation and metabolism.** *Science* 1998, **281**:692-695.

39. Zaret K: **Developmental competence of the gut endoderm: genetic potentiation by GATA and HNF3/fork head proteins.** *Dev Biol* 1999, **209**:1-10.

40. Lin K, Dorman JB, Rodan A, Kenyon C: **daf-16: an HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*.** *Science* 1997, **278**:1319-1322.

41. Hsin H, Kenyon C: **Signals from the reproductive system regulate the lifespan of *C. elegans*.** *Nature* 1999, **399**:362-366.

42. O'Brien RM, Noisin EL, Suwanichkul A, Yamasaki T, Lucas PC, Wang JC, Powell DR, Granner DK: **Hepatic nuclear factor 3- and hormone-regulated expression of the phosphoenolpyruvate carboxykinase and insulin-like growth factor- binding protein 1 genes.** *Mol Cell Biol* 1995, **15**:1747-1758.