Published in final edited form as:

Mol Biosyst. 2012 January ; 8(1): 97–104. doi:10.1039/c1mb05258f.

Roles of Intrinsic Disorder in Protein-Nucleic Acid Interactions

H. Jane Dyson^a

H. Jane Dyson: dyson@scripps.edu

^aDepartment of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, U.S.A. Fax: 1 858 784 9822; Tel: 1 858 784 2223

Abstract

Interactions between proteins and nucleic acids typify the role of disordered segments, linkers, tails and other entities in the function of complexes that must form with high affinity and specificity but which must be capable of dissociating when no longer needed. While much of the emphasis in the literature has been on the interactions of disordered proteins with other proteins, disorder is also frequently observed in nucleic acids (particularly RNA) and in the proteins that interact with them. The interactions of disordered proteins with DNA most often manifest as molding of the protein onto the B-form DNA structure, although some well-known instances involve remodeling of the DNA structure that seems to require that the interacting proteins be disordered to various extents in the free state. By contrast, induced fit in RNA-protein interactions has been recognized for many years – the existence and prevalence of this phenomenon provides the clearest possible evidence that RNA and its interactions with proteins must be considered as highly dynamic, and the dynamic nature of RNA and its multiplicity of folded and unfolded states is an integral part of its nature and function.

Introduction

Examination of the amino acid sequences of eukaryotic transcriptional regulatory proteins reveals that they frequently contain regions that are predicted to be partially or completely disordered.^{1,2} Interaction with nucleic acids is a common function for disordered proteins:³ as of 2002, the number of identified intrinsically disordered proteins (IDPs) that were associated with protein-nucleic acid interactions (31) was comparable and second only to those associated with protein-protein interactions (54).⁴ The subject of intrinsic disorder in transcription factors has been reviewed previously;^{1,5} in this review I will focus more generally on the structural and dynamic implications of disorder in protein-nucleic acid interactions.

Variations on “Beads on a String”

One of the recurring themes in the study of disorder in proteins is the prevalence of disorder in linker sequences between domains that are more or less well folded. Many possible reasons exist for the presence of such linkers.^{6,7} For complexes of proteins with nucleic acids, there are frequently specific structural and dynamic requirements for particular linkers, related to the role of disorder in facilitating the search for specific sequences in long stretches of DNA and the subsequent high-affinity interaction once the correct sequence has been encountered. This phenomenon is illustrated by extensive studies on the DNA complexes of two proteins containing Cys₂His₂ zinc finger domains. The two proteins, the Wilms tumor suppressor protein WT1 and the transcription factor TFIIIA, have widely

differing functions, although the zinc finger domains are highly homologous in sequence and structure (Figure 1). For both proteins, it appears that the presence of specific linker sequences between the domains has an important function in facilitating the formation of sequence-specific DNA complexes.

Facilitated DNA Search

Published genomes contain many zinc finger genes, which are identified by patterns of potential zinc-binding side chains such as cysteine and histidine. The first zinc finger proteins to be described⁸ were of the Cys₂His₂ type, which consist of a β -hairpin and a short α -helix, held together by the presence of zinc.⁹ Although a single finger can bind DNA, significant specificity as well as greater affinity requires the presence of multiple fingers; the minimal number of finger units to promote binding to a specific DNA sequence appears to be 3.^{10,11} WT1, which contains 4 zinc finger domains with an alternative splice site in the linker between ZF3 and ZF4, binds DNA with either 3 or 4 zinc finger domains. TFIIIA has 9 zinc finger domains, which bind both DNA and RNA with different affinities, in groups of 3.

The majority of the free energy for DNA binding by the 9-finger TFIIIA is conferred by the sequence-specific binding of the first 3 fingers,¹⁰ and this affinity is finger-specific: a construct containing ZF2-4 instead of ZF1-3 binds to DNA non-specifically with lower affinity.¹⁰ A precise series of footprinting and mobility-shift assays demonstrated the binding sites of the individual zinc fingers on a 60-nucleotide double-stranded DNA representing the internal control region of the *Xenopus* 5S RNA gene,¹² and this information was used to obtain a model for the binding of the 9 zinc fingers to the DNA.¹³ Interestingly, the zinc finger domains show different modes of binding (Figure 2), with ZF1-3 binding in the major groove of the promoter C block and ZF7-9 in the major groove of the promoter A block.^{11,13} ZF5 also contacts the major groove in the intermediate element, but fingers 4 and 6 are bound across the minor groove.^{13,14} This model is consistent with 3-dimensional structures obtained for multiple-finger constructs in complex with DNA.¹⁵⁻¹⁷ Thus, in complex with DNA of the correct nucleotide sequence, the multiple zinc finger proteins form tight and specific complexes. In the presence of DNA of incorrect sequence, the zinc finger proteins associate with the DNA primarily through electrostatic interactions with the phosphate backbone, forming heterogeneous complexes. In the absence of DNA, the individual zinc finger domains remain folded, but the linkers between domains are disordered. This was shown by NMR relaxation studies both for TFIIIA ZF1-3^{18,19} and for WT1 ZF1-4,²⁰ and is illustrated in Figure 3 by the plot of ¹H-¹⁵N NOE for WT1 in the presence and absence of the cognate DNA. This loss of flexibility in the linker sequences is apparently caused by the formation of a DNA-induced α -helix cap when the protein is bound to the correct DNA sequence.²¹ These observations prompt a model for the mechanism used by multiple zinc finger proteins to find and bind the correct DNA sequence. The zinc finger protein could initially be weakly associated with DNA, but could be readily dissociated from non-cognate sites (or perhaps slide along the phosphate backbone) until the cognate site is reached. At the cognate site, the affinity of the complex is greatly increased by the protein-protein interactions, (including a significant increase in inter-finger contact surfaces, as shown in Figure 4(b)¹⁵) that result in the structuring of the linkers. This process, which invokes the concept of enthalpy-entropy compensation, has been likened to a “snap-lock”²¹ and provides a plausible explanation for the sequence homology of zinc finger linkers at the start of fingers 2 and 3 (and 4 of WT1) but not of finger 1 (Figure 1).

Sliding and Hopping

Disorder in the linkers between structural domains not only informs the model outlined in the previous paragraph, but also can help to rationalize the observation that transcription

factors are able to find their specific binding sites on DNA much faster than expected.²² Hopping between DNA binding sites was demonstrated by NMR for a multi-domain Oct-1 protein.²³ Recent computational studies have further refined these types of models by showing that both linkers²⁴ and disordered tails²⁵ play an important role in the DNA search mechanism for multidomain proteins. According to these studies, the linkers facilitate the search process by increasing the affinity to the (non-cognate) DNA, thus promoting sliding of the protein along the DNA, while the disordered tails slow the linear diffusion rate of the protein along the DNA but may promote jumping of the protein between DNA molecules by a “monkey bar” mechanism.²⁶

TFIIIA Binds both DNA and RNA

Large amounts of ribosomal RNA accumulate during oocyte development in *Xenopus*, as a direct consequence of the actions of TFIIIA both as a transcription factor favoring the production of 5S ribosomal RNA²⁷ and because TFIIIA can interact with and stabilize the product 5S RNA.^{28,29} RNA binding is localized to the middle three zinc fingers (ZF 4-6),³⁰ which clearly have a less intimate interaction with DNA than the other 6 fingers (Figure 2). The interaction of ZF4-6 of TFIIIA with RNA clearly shows the influence of intrinsic disorder in the linked (folded) zinc fingers: finger 4 binds specifically to the structured Loop E motif of the 5S RNA, while the interaction of fingers 5 and 6 with the loop A motif involves reorganization of the flexible 3-helix RNA junction together with induced structure in the ZF5-ZF6 interface.^{31,32} (Figure 5)

The ability of the 9-finger TFIIIA zinc finger domain to bind as needed to either DNA or RNA is a direct reflection of the versatility of the “beads on a string” structure of the protein. Small structured domains (structured only by virtue of the interaction with zinc) are connected by linkers that are disordered in the free protein, giving a large capture radius and the possibility of a facilitated search for a particular DNA sequence. The presence of 9 zinc fingers allows DNA interactions to be made with a selection of the fingers, primarily fingers 1–3, with additional affinity arising from the interactions of fingers 7–9. The model of Figure 2 suggests that fingers 4–6 act merely as spacers in the DNA interaction, adding very little to the binding energy of the interaction. Significantly, it is precisely these three fingers that make the primary interaction with 5S RNA. Segmental disorder in this case renders the protein capable of binding to two different nucleic acids, and suggests a mechanism whereby the protein could switch between its two functions.

Roles of Unstructured Tails

The previous section described evidence for the role of disordered regions in the facilitated search of proteins for their correct cognate binding sites on DNA. In contrast to this rather generalized function, there are a number of specific instances where the folding and unfolding of a part of a protein that is not integral to the canonical DNA-binding site is critical to the function of a DNA-protein complex. For example, the regulatory domains of transcription factors such as the nuclear hormone receptors frequently contain disordered domains.³³ Limited disorder is also a feature of the DNA-binding regions of nuclear hormone receptors. The nuclear hormone receptor DNA binding domains (DBDs) are zinc fingers, but are completely different from the Cys₂His₂ zinc finger domains described in the previous section. They consist of a highly conserved ~70-amino acid sequence containing 8 conserved cysteines that form two zinc-binding sites.³⁴ The DBDs typically bind as dimers to DNA sequences termed hormone-response elements (HREs), which consist of two half-sites. Some receptors, such as glucocorticoid receptor (GR) bind as homodimers to an inverted-repeat pair of half-sites, while others bind as heterodimers to direct repeats with various numbers of nucleotides between the half sites.³⁵ Heterodimers of hormone-

responsive receptors such as thyroid hormone receptor (TR) with the retinoid X receptor (RXR) bind with high affinity to target DNA.³⁶

RXR – Folding of a C-terminal Helix Modulates Specificity

The 3-dimensional structure of free RXR^{37,38} showed the presence of a unique C-terminal helical segment that was not present in the structures of other nuclear hormone receptor DBDs (Figure 6). This helix was shown to be required for the DNA-protein and protein-protein interactions involved in the binding of RXR to the HRE as a homodimer,^{37,39,40} but was absent from the X-ray crystal structure of a heterodimer of RXR and TR with direct-repeat half sites.³⁶ (Figure 6b) Thus it appears that, depending on the protein partner, the C-terminal helix of RXR may fold and participate in protein-protein interactions in a DNA-induced homodimer, but is unfolded in heterodimeric complexes.⁴¹ This difference may provide insights into the mechanism of exchange of subunits between homo- and heterodimeric RXR complexes.

ERR2 – Folding of a C-terminal Extension Enhances Affinity

The human estrogen-related receptors (ERR) show sequence homology to the estrogen receptors, but are constitutively active in the absence of exogenous ligands.⁴² While most nuclear hormone receptors bind DNA as homodimers or heterodimers to either inverted or direct repeats of 6 base-pair half-sites, the ERRs and other receptors such as NGFI-B⁴³ and RevErb⁴⁴ bind to single DNA half-sites as monomers. These proteins contain, in addition to a classical two-zinc Cys₄Cys₄ nuclear hormone receptor structure, a C-terminal extension that interacts with a 5' extension of the DNA half-site. The C-terminal extension was shown by NMR to be disordered in the free protein⁴⁵ but is well structured in the complex of ERR2 with an extended DNA duplex.⁴⁶ The core of the ERR2 zinc finger domain binds in the major groove of the DNA very similarly to the structures of other nuclear hormone receptors, but the C-terminal extension, including the sequence Arg179-Gly180-Gly181-Arg182 that is conserved among monomeric nuclear hormone receptors, binds in the minor groove of the DNA as an extended structure (Figure 6c). In addition, a tyrosine side chain, conserved among the ERR proteins, makes a highly specific hydrophobic contact with the core of the ERR zinc finger domain. These “folding upon binding” interactions serve to increase the affinity and specificity of the ERR family members in their monomeric DNA binding mode, forming a “pseudo-dimer interface” that may point to a mechanism through which dimerization may have evolved from an ancestral monomeric receptor.⁴⁶

Disordered Regulatory Domains

Nuclear hormone receptor proteins generally consist of 3 domains, an N-terminal regulatory domain (NTD), the DBD (as described in the previous two sections) and the ligand-binding domain. The NTD is intrinsically disordered, but in an extensive series of papers, Kumar and Thompson showed that this region of the protein can interact with a number of coregulatory proteins in a folded state,^{47,48} but only in the presence of the DBD bound to DNA.^{49,50} Disordered C-terminal domains are also the sites of regulatory interaction. The progesterone receptor (PR) coactivator Jun dimerization protein 2 (JDP2) was recently shown to enhance the transcriptional activity of PR by inducing helical structure in the C-terminal extension of PR.⁵¹

Sequence- and Structure-Specific HMG Domain Transcription Factors

The high mobility-group (HMG) proteins are a class of DNA-binding proteins that are homologous in both sequence and structure. However, there are major differences in the stability and dynamics of two sub-classes of HMG domains, those that bind DNA in a sequence-specific manner, giving complexes where the DNA is significantly bent, and the

structure-specific HMG domains that bind without sequence specificity to any pre-bent DNA sequence, such as Holliday junctions and cisplatin-bound DNA sequences.^{52,53} As a class, the structure-specific HMG domains tend to bind using multiple domains, and the free proteins are well-folded and give excellent NMR spectra.^{54–58} By contrast, the sequence-specific HMG domains bind using a single domain and are frequently incompletely folded in solution^{59–61} although the DNA complexes of these proteins are highly stable and well structured (Figure 7).^{60,62–64} The difference in the stability of the free state of the domains in the two classes likely reflects the different modes of binding. The well-formed structures of the structure-specific class^{55–58} recognize and bind to pre-bent DNA, while the flexible and partially disordered sequence-specific HMG domains bind to and distort a specific DNA sequence, bending the DNA in some cases $> 90^\circ$.^{60,62–64} The disorder and conformational heterogeneity displayed by the free sequence-specific HMG domains may be necessary to allow bending of the DNA and the simultaneous collapse to the lower-energy conformation in the final complex.

Chromatin Remodeling

Interphase chromosomes are organized into nucleosome repeating units, which consist of supercoiled DNA associated with structured core histones H2A, H2B, H3 and H4. Between the nucleosomes are 10–80 bp DNA linkers that are associated with linker histones (H1) and with the disordered N-terminal tail domains (NTDs) of the core histones.⁶⁵ Linker histones contain an ordered domain that binds to the nucleosome and a highly basic C-terminal tail that interacts with the linker DNA^{66,67} and folds upon DNA binding.⁶⁸ (Figure 8) The disordered tails of the histones play a major role in the condensation of chromatin. Similarly, chromatin remodeling proteins, which are required to reconfigure the structure of chromatin for accessibility of the DNA for transcription, frequently contain large disordered regions.⁶⁹

Disordered Proteins and RNA

Duplex DNA is relatively stable and difficult to deform, requiring a complex set of interactions and a very favorable enthalpy term to form the final folded complex. For RNA, the energy required to deform and unfold the molecule is much lower, and the characteristics of RNA-binding proteins can therefore be much more diverse. Nevertheless, both sequence and shape-dependent interactions are seen between RNA and proteins.⁷⁰ The field of RNA-protein interactions has long recognized that mutual folding of RNA and protein components is common and probably ubiquitous.^{71–73} In this brief survey, I will give some examples where the conformational diversity of both protein and RNA components has been found to be important for function.

RNA Chaperones

Because RNA molecules frequently misfold into structures that may be highly stable but non-functional, a large number of protein molecules are employed to overcome these kinetic traps and to promote correct folding and circumvent aggregation processes.^{74–76} RNA chaperones provide an extremely important quality-control function in the living cell; such molecules may have constituted the earliest proteins synthesized on the way from the ancient RNA world.⁷⁷ Disordered RNA chaperones occur even in viruses.⁷⁸ In view of the likely age of these molecules, it is perhaps unsurprising that they contain extensive disordered regions; it is clear from the retention of disorder that their contemporary function also requires the proteins to be less than fully structured, and a recent examination showed that the occurrence of disorder in established RNA chaperones is significantly higher than for almost any other class of protein.⁷⁴

Interactions of mRNA and Proteins

The product of transcription of a gene is a pre-mRNA, which is processed and transported from the nucleus to the cytoplasm associated with a varying cast of protein and ribonucleoprotein (RNP) partners, many of which utilize disordered regions to mediate their interactions. For example, the spliceosome, an RNP that processes noncoding introns from the pre-mRNA, is of necessity an extremely dynamic machine that requires the sequential interaction of a number of proteins and small RNPs with the mRNA. Many of these components undergo extensive remodeling during the splicing process, and some at least of the protein components have been shown to contain disordered regions.⁷⁴ Current knowledge on the spliceosome and its components has recently been reviewed.⁷⁹

Disordered Linkers between Folded Domains Mediate Binding to Single Stranded RNA

Binding of single-stranded RNAs frequently employs disordered regions of proteins. In particular, the specific proteins that interact with the 3' untranslated regions of mRNA often share a characteristic topology consisting of multiple folded domains linked by disordered segments that are critical in mediating the RNA sequence specificity of the interaction. For example, the disordered linker between the zinc finger domains of TIS11d makes specific contacts with the UUAUUUAUU sequence of the class II AU-rich element in the 3' untranslated region of target mRNAs,⁸⁰ and a similar role for the disordered linkers between the KH domains of KSRP^{81,82} and the RRM domains of FIR.⁸³

Ribosomal Proteins

The assembly of a ribosome involves a complex series of folding and assembly steps.⁷² The primary catalytic machinery of a ribosome is RNA-based, but there are many proteins that are intimately part of the ribosome structure. The order of the addition of the various ribosomal proteins during the assembly of a ribosome has been the subject of a great deal of interest since the original Nomura map for the assembly of the 30S ribosomal subunit.⁸⁰ Before assembly, many ribosomal proteins are found to be disordered, undergoing a classic folding upon binding process that occurs sequentially as the binding sites for the proteins on the RNA become available. Examples include the mutual folding and binding of ribosomal protein L5 to 5S rRNA^{81,82} and the folding of the N-terminal extension of protein L20 as an early step in ribosome assembly.^{83–85} (Figure 9)

Dynamic Complexes

The dynamic and varied range of interactions that occur between RNA and protein molecules have been exhaustively reviewed over many years.^{70,71,86–88} The following two examples indicate that the complexes of proteins with RNA may themselves be incompletely folded, much as protein-protein complexes can remain “fuzzy”.⁸⁹ The Vts1p sterile-alpha motif (SAM) domain interacts with RNA in a shape-dependent manner.⁹⁰ NMR relaxation measurements showed an increase in backbone mobility of some parts of the protein in the bound state compared to the free state, due to the presence of a dynamic interaction interface.⁹¹ A similar dynamic complex was observed for the complex between the Nun protein of the prophage HK022 and the λ -phage *BoxB* RNA hairpin,⁹² where an N-terminal extension that is not directly bound to the RNA is nevertheless conformationally restricted in the complex. These examples further stress the variety and versatility of RNA-protein complexes.

The Measles Virus Nucleocapsid

One of the best-characterized protein-RNA systems is the nucleocapsid structure of the measles virus. The single stranded RNA of the viral genome is enclosed in a capsid consisting of multiple copies of the nucleoprotein N. The structured core of N (N_{CORE})

mediates the interaction of N with viral RNA, while the disordered tail (N_{TAIL}) promotes interaction of the virus with the polymerase complex during transcription and replication. The Ntail polypeptide has been extensively characterized,^{93,94} most recently in the context of the entire virus capsid.⁹⁵ The molecular recognition element (MoRE) of N_{TAIL} is placed more than 90 amino acids from the structured domain; studies by NMR, EM and SAXS indicate that N_{TAIL} remains disordered even in the intact virus, and that the MoRE makes transient contact with the structured domains, thus providing a putative mechanism for the approach of the polymerase complex to the viral RNA.

Conclusions

Experimental evidence for the interactions of disordered proteins with nucleic acids has been available for a significant period. In the earlier years, the emphasis of experimental studies was frequently on the structural characterization of the complexes formed by IDPs and nucleic acids. It has only been in later years that the field has come to embrace and understand the essential role of the disordered nature of these proteins in their function. Indeed, especially for RNA interactions, it would appear that disorder and alternatively-folded structures are an important part of the nature of nucleic acids as well. Large and small biological molecules and large molecular machines are required to be dynamic to achieve their functions, and dynamic disorder takes many forms, on a continuum from random-coil-like conformational ensembles to linked domains that act as beads on a string.⁹⁶ Three-dimensional structures obtained by X-ray crystallography or NMR are valuable for understanding, but can by no means give the whole picture – instead of stereo pictures, perhaps we will in later years be presenting movies to illustrate how biological molecules interact and function.

Notes and references

1. Liu J, et al. Intrinsic disorder in transcription factors. *Biochemistry*. 2006; 45:6873–6888. [PubMed: 16734424]
2. Minezaki Y, Homma K, Kinjo AR, Nishikawa K. Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J Mol Biol*. 2006; 359:1137–1149. [PubMed: 16697407]
3. Dan A, Ofra Y, Kfiger Y. Large-scale analysis of secondary structure changes in proteins suggests a role for disorder-to-order transitions in nucleotide binding proteins. *Proteins-Structure Function and Bioinformatics*. 2010; 78:236–248.
4. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. *Biochemistry*. 2002; 41:6573–6582. [PubMed: 12022860]
5. Hager, Gordon L.; McNally, James G.; Misteli, Tom. Transcription Dynamics. *Molecular Cell*. 2009; 35:741–753. [PubMed: 19782025]
6. Wright, Peter E.; Dyson, HJ. Linking folding and binding. *Curr Opin Struct Biol*. 2009; 19:31–38. [PubMed: 19157855]
7. Dyson HJ. Expanding the proteome: disordered and alternatively-folded proteins. *Quart Rev Biophys*. 2011 in press.
8. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J*. 1985; 4:1609–1614. [PubMed: 4040853]
9. Lee MS, Gippert G, Soman KY, Case DA, Wright PE. Three-dimensional solution structure of a single zinc finger binding domain. *Science*. 1989; 245:635–637. [PubMed: 2503871]
10. Liao X, Clemens KR, Tennant L, Wright PE, Gottesfeld JM. Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the *Xenopus* 5 S RNA gene. *J Mol Biol*. 1992; 223:857–871. [PubMed: 1538401]

11. Pieler T, Hamm J, Roeder RG. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell*. 1987; 48:91–100. [PubMed: 3791417]
12. Clemens KR, et al. Relative contributions of the zinc fingers of transcription factor IIIA to the energetics of DNA binding. *J Mol Biol*. 1994; 244:23–35. [PubMed: 7966319]
13. Clemens KR, Liao X, Wolf V, Wright PE, Gottesfeld JM. Definition of the binding sites of individual zinc fingers in the TFIIIA-5S RNA gene complex. *Proc Natl Acad Sci USA*. 1992; 89:10822–10826. [PubMed: 1438283]
14. Hayes JJ, Tullius TD. Structure of the TFIIIA-5 S DNA complex. *J Mol Biol*. 1992; 227:407–417. [PubMed: 1404361]
15. Wuttke DS, Foster MP, Case DA, Gottesfeld JM, Wright PE. Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: Determinants of affinity and sequence specificity. *J Mol Biol*. 1997; 273:183–206. [PubMed: 9367756]
16. Nolte RT, Conlin RM, Harrison SC, Brown RS. Differing roles for zinc fingers in DNA recognition: Structure of a six-finger transcription factor IIIA complex. *Proc Natl Acad Sci USA*. 1998; 95:2938–2943. [PubMed: 9501194]
17. Stoll R, et al. Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. *J Mol Biol*. 2007; 372:1227–1245. [PubMed: 17716689]
18. Foster MP, et al. Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nature Struct Biol*. 1997; 4:605–608. [PubMed: 9253405]
19. Bruschweiler R, Liao X, Wright PE. Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science*. 1995; 268:886–889. [PubMed: 7754375]
20. Laity JH, Dyson HJ, Wright PE. Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proc Natl Acad Sci USA*. 2000; 97:11932–11935. [PubMed: 11050227]
21. Laity JH, Dyson HJ, Wright PE. DNA-induced α -helix capping in conserved linker sequences is a determinant of binding affinity in Cys₂-His₂ zinc fingers. *J Mol Biol*. 2000; 295:719–727. [PubMed: 10656784]
22. Berg OG, Winter RB, von Hippel PH. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*. 1981; 20:6929–6948. [PubMed: 7317363]
23. Doucleff, Michaelleen; Clore, GM. Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1. *Proc Natl Acad Sci USA*. 2008; 105:13871–13876. [PubMed: 18772384]
24. Vuzman D, Polonsky M, Levy Y. Facilitated DNA search by multidomain transcription factors: cross talk via a flexible linker. *Biophys J*. 2010; 99:1202–1211. [PubMed: 20713004]
25. Vuzman D, Azia A, Levy Y. Searching DNA via a "Monkey Bar" mechanism: the significance of disordered tails. *J Mol Biol*. 2010; 396:674–684. [PubMed: 19958775]
26. Singh GP, Ganapathi M, Dash D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins-Structure Function and Bioinformatics*. 2007; 66:761–765.
27. Engelke DR, Ng SY, Shastry BS, Roeder RG. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*. 1980; 19:717–728. [PubMed: 6153931]
28. Honda BM, Roeder RG. Cell association of a 5S gene transcription factor with 5S RNA and altered levels of factor during cell differentiation. *Cell*. 1980; 22:119–126. [PubMed: 6159099]
29. Pelham HR, Brown DD. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proc Natl Acad Sci USA*. 1980; 77:4170–4174. [PubMed: 7001457]
30. Clemens KR, et al. Molecular basis for specific recognition of both RNA and DNA by a zinc finger proteins. *Science*. 1993; 260:530–533. [PubMed: 8475383]
31. Lu D, Searles MA, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature*. 2003; 426:96–100. [PubMed: 14603324]
32. Lee BM, et al. Induced fit and "Lock and Key" recognition of 5 S RNA by zinc fingers of transcription factor IIIA. *J Mol Biol*. 2006; 357:275–291. [PubMed: 16405997]
33. Krasowski MD, Reschly EJ, Ekins S. Intrinsic disorder in nuclear hormone receptors. *J Proteome Res*. 2008; 7:4359–4372. [PubMed: 18651760]

34. Mangelsdorf DJ, Evans RM. The RXR heterodimers and orphan receptors. *Cell*. 1995; 83:841–850. [PubMed: 8521508]
35. Khorasanizadeh S, Rastinejad F. Nuclear-receptor interactions on DNA-response elements. *Trends Biochem Sci*. 2001; 26:384–390. [PubMed: 11406412]
36. Rastinejad F, Perlmann T, Evans RM, Sigler PB. Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature*. 1995; 375:203–211. [PubMed: 7746322]
37. Lee MS, Kliewer SA, Provencal J, Wright PE, Evans RM. Structure of the retinoid X receptor α DNA binding domain: a helix required for homodimeric DNA binding. *Science*. 1993; 260:1117–1121. [PubMed: 8388124]
38. Holmbeck SMA, et al. High resolution solution structure of the retinoid X receptor DNA binding domain. *J Mol Biol*. 1998; 281:271–284. [PubMed: 9698548]
39. Lee MS, et al. NMR assignments and secondary structure of the retinoid X receptor α DNA-binding domain--Evidence for the novel C-terminal helix. *Eur J Biochem*. 1994; 224:639–650. [PubMed: 7925381]
40. Wilson TE, Paulsen RE, Padgett KA, Milbrandt J. Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. *Science*. 1992; 256:107–110. [PubMed: 1314418]
41. Holmbeck SMA, Dyson HJ, Wright PE. DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. *J Mol Biol*. 1998; 284:533–539. [PubMed: 9826495]
42. Giguere V, Yang N, Segui P, Evans RM. Identification of a new class of steroid hormone receptors. *Nature*. 1988; 331:91–94. [PubMed: 3267207]
43. Meinke G, Sigler PB. DNA-binding mechanism of the monomeric orphan nuclear receptor NGFI-B. *Nature Struct Biol*. 1999; 6:471–477. [PubMed: 10331876]
44. Zhao Q, Khorasanizadeh S, Miyoshi Y, Lazar MA, Rastinejad F. Structural elements of an orphan nuclear receptor-DNA complex. *Molecular Cell*. 1998; 1:849–861. [PubMed: 9660968]
45. Sem DS, et al. NMR spectroscopic studies of the DNA-binding domain of the monomer-binding nuclear orphan receptor, human estrogen related receptor-2 - The carboxyl-terminal extension to the zinc-finger region is unstructured in the free form of the protein. *J Biol Chem*. 1997; 272:18038–18043. [PubMed: 9218433]
46. Gearhart MD, Holmbeck SMA, Evans RM, Dyson HJ, Wright PE. Monomeric complex of human orphan estrogen related receptor-2 with DNA: A pseudo-dimer interface mediates extended half-site recognition. *J Mol Biol*. 2003; 327:819–832. [PubMed: 12654265]
47. Kumar R, et al. TATA box binding protein induces structure in the recombinant glucocorticoid receptor AF1 domain. *Proc Natl Acad Sci USA*. 2004; 101:16425–16430. [PubMed: 15545613]
48. Kumar R, Betney R, Li J, Thompson EB, McEwan IJ. Induced α -helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. *Biochemistry*. 2004; 43:3008–3013. [PubMed: 15023052]
49. Kumar R, et al. Interdomain signaling in a two-domain fragment of the human glucocorticoid receptor. *J Biol Chem*. 1999; 274:24737–24741. [PubMed: 10455143]
50. Kumar R, Thompson EB. Influence of flanking sequences on signaling between the activation function AF1 and DNA-binding domain of the glucocorticoid receptor. *Arch Biochem Biophys*. 2010; 496:140–145. [PubMed: 20188692]
51. Hill KK, Roemer SC, Jones DNM, Churchill MEA, Edwards DP. A Progesterone Receptor Co-activator (JDP2) Mediates Activity through Interaction with Residues in the Carboxyl-terminal Extension of the DNA Binding Domain. *J Biol Chem*. 2009; 284:24415–24424. [PubMed: 19553667]
52. Grosschedl R, Giese K, Pagel J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet*. 1994; 10:94–100. [PubMed: 8178371]
53. Baxevanis AD, Landsman D. The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res*. 1995; 23:1604–1613. [PubMed: 7784217]
54. Broadhurst RW, Hardman CH, Thomas JO, Laue ED. Backbone dynamics of the A-domain of HMG1 as studied by ^{15}N NMR spectroscopy. *Biochemistry*. 1995; 34:16608–16617. [PubMed: 8527433]

55. Read CM, Cary PD, Crane-Robinson C, Driscoll PC, Norman DG. Solution structure of a DNA-binding domain from HMG1. *Nucleic Acids Res.* 1993; 21:3427–3436. [PubMed: 8346022]
56. Weir HM, et al. Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.* 1993; 12:1311–1319. [PubMed: 8467791]
57. Jones DNM, et al. The solution structure and dynamics of the DNA-binding domain of HMG-D from *Drosophila melanogaster*. *Structure.* 1994; 2:609–627. [PubMed: 7922039]
58. Allain FH, et al. Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding. *EMBO J.* 1999; 18:2563–2579. [PubMed: 10228169]
59. Weiss MA. Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. *Mol Endocrinol.* 2001; 15:353–362. [PubMed: 11222737]
60. Van Houte LPA, et al. Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J Biol Chem.* 1995; 270:30516–30524. [PubMed: 8530483]
61. Love JJ, Li X, Chung J, Dyson HJ, Wright PE. The LEF-1 HMG domain undergoes a disorder-to-order transition upon complex formation with cognate DNA. *Biochemistry.* 2004; 43:8725–8734. [PubMed: 15236581]
62. Love JJ, et al. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature.* 1995; 376:791–795. [PubMed: 7651541]
63. Werner MH, Huth JR, Gronenborn AM, Clore GM. Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell.* 1995; 81:705–714. [PubMed: 7774012]
64. Connor F, et al. DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res.* 1994; 22:3339–3346. [PubMed: 8078769]
65. Felsenfeld G, Groudine M. Controlling the double helix. *Nature.* 2003; 421:448–453. [PubMed: 12540921]
66. Allan J, Mitchell T, Harborne N, Bohm L, Crane-Robinson C. Roles of H1 domains in determining higher order chromatin structure and H1 location. *J Mol Biol.* 1986; 187:591–601. [PubMed: 3458926]
67. Caterino TL, Hayes JJ. Structure of the H1 C-terminal domain and function in chromatin condensation. *Biochem Cell Biol.* 2011; 89:35–44. [PubMed: 21326361]
68. Caterino TL, Fang H, Hayes JJ. Nucleosome linker DNA contacts and induces specific folding of the intrinsically disordered h1 carboxyl-terminal domain. *Mol Cell Biol.* 2011; 31:2341–2348. [PubMed: 21464206]
69. Sandhu KS. Intrinsic disorder explains diverse nuclear roles of chromatin remodeling proteins. *Journal of Molecular Recognition.* 2009; 22:1–8. [PubMed: 18802931]
70. Stefl R, Skrisovska L, Allain FH. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* 2005; 6:33–38. [PubMed: 15643449]
71. Frankel AD, Smith CA. Induced folding in RNA-protein recognition: More than a simple molecular handshake. *Cell.* 1998; 92:149–151. [PubMed: 9458038]
72. Williamson JR. Induced fit in RNA-protein recognition. *Nature Struct Biol.* 2002; 7:834–837. [PubMed: 11017187]
73. Leulliot N, Varani G. Current Topics in RNA-Protein Recognition: Control of Specificity and Biological Function through Induced Fit and Conformational Capture. *Biochemistry.* 2001; 40:7947–7956. [PubMed: 11434763]
74. Tompa P, Csermely P. The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.* 2004; 18:1169–1175. [PubMed: 15284216]
75. Cristofari G, Darlix JL. The ubiquitous nature of RNA chaperone proteins. *Progress in Nucleic Acid Research and Molecular Biology, Vol 72.* 2002; 72:223–268.
76. Herschlag D. RNA chaperones and the RNA folding problem. *J Biol Chem.* 1995; 270:20871–20874. [PubMed: 7545662]
77. Poole AM, Jeffares DC, Penny D. The path from the RNA world. *J Mol Evol.* 1998; 46:1–17. [PubMed: 9419221]

78. Ivanyi-Nagy R, Lavergne JP, Gabus C, Ficheux D, Darlix JL. RNA chaperoning and intrinsic disorder in the core proteins of Flaviviridae. *Nucleic Acids Res.* 2008; 36:712–725. [PubMed: 18033802]
79. Wahl MC, Will CL, Luhrmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell.* 2009; 136:701–718. [PubMed: 19239890]
80. Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol.* 2004; 11:257–264. [PubMed: 14981510]
81. Garcia-Mayoral MF, Diaz-Moreno I, Hollingworth D, Ramos A. The sequence selectivity of KSRP explains its flexibility in the recognition of the RNA targets. *Nucleic Acids Res.* 2008; 36:5290–5296. [PubMed: 18684992]
82. Diaz-Moreno I, et al. Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. *Nucleic Acids Res.* 2010; 38:5193–5205. [PubMed: 20385598]
83. Cukier CD, et al. Molecular basis of FIR-mediated c-myc transcriptional control. *Nat Struct Mol Biol.* 2010; 17:1058–1064. [PubMed: 20711187]
84. Held WA, Ballou B, Mizushima S, Nomura M. Assembly mapping of 30 S ribosomal proteins from *Escherichia coli*. Further studies. *J Biol Chem.* 1974; 249:3103–3111. [PubMed: 4598121]
85. DiNitto JP, Huber PW. Mutual induced fit binding of *Xenopus* ribosomal protein L5 to 5S rRNA. *J Mol Biol.* 2003; 330:979–992. [PubMed: 12860121]
86. Scripture JB, Huber PW. Binding site for *Xenopus* ribosomal protein L5 and accompanying structural changes in 5S rRNA. *Biochemistry.* 2011; 50:3827–3839. [PubMed: 21446704]
87. Raibaud S, et al. NMR structure of bacterial ribosomal protein l20: implications for ribosome assembly and translational control. *J Mol Biol.* 2002; 323:143–151. [PubMed: 12368106]
88. Timsit Y, Allemand F, Chiaruttini C, Springer M. Coexistence of two protein folding states in the crystal structure of ribosomal protein L20. *Embo Reports.* 2006; 7:1013–1018. [PubMed: 16977336]
89. Timsit Y, Acosta Z, Allemand F, Chiaruttini C, Springer M. The Role of Disordered Ribosomal Protein Extensions in the Early Steps of Eubacterial 50 S Ribosomal Subunit Assembly. *International Journal of Molecular Sciences.* 2009; 10:817–834. [PubMed: 19399222]
90. Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. *Science.* 1994; 265:615–621. [PubMed: 8036511]
91. Moras D, Poterszman A. RNA-protein interactions: Diverse modes of recognition. *Curr Biol.* 1995; 5:249–251. [PubMed: 7540101]
92. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* 2008; 582:1977–1986. [PubMed: 18342629]
93. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci.* 2008; 33:2–8. [PubMed: 18054235]
94. Oberstrass FC, et al. Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. *Nat Struct Mol Biol.* 2006; 13:160–167. [PubMed: 16429156]
95. Ravindranathan, Sapna; Oberstrass, Florian C.; Allain, Frederic H. Increase in Backbone Mobility of the VTS1p-SAM Domain on Binding to SRE-RNA. *J Mol Biol.* 2010; 396:732–746. [PubMed: 20004205]
96. Stuart AC, Gottesman ME, Palmer AG. The N-terminus is unstructured, but not dynamically disordered, in the complex between HK022 Nun protein and lambda-phage BoxB RNA hairpin. *FEBS Letters.* 2003; 553:95–98. [PubMed: 14550553]
97. Bernard, Cedric, et al. Interaction between the C-terminal domains of N and P proteins of measles virus investigated by NMR. *FEBS Letters.* 2009; 583:1084–1089. [PubMed: 19275899]
98. Gely S, et al. Solution structure of the C-terminal X domain of the measles virus phosphoprotein and interaction with the intrinsically disordered C-terminal domain of the nucleoprotein. *J Mol Recognit.* 2010; 23:435–447. [PubMed: 20058326]
99. Jensen MR, et al. Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci U S A.* 2011

100. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.* 2005; 6:197–208. [PubMed: 15738986]

Fig. 1.

Amino acid sequence alignment for representative Cys₂His₂ zinc finger proteins (adapted with permission from ref. 17).

Fig. 2.
Schematic representation of TFIIIA binding to the 5S RNA internal control region. (adapted with permission from ref. 13)

Fig. 3. Backbone $\{^1\text{H}\}-^{15}\text{N}$ nuclear Overhauser effect (NOE) for free (pink) and DNA-bound (blue) WT1 zinc finger protein. Open symbols indicate overlapped resonances, for which there are large uncertainties in the relaxation parameters (adapted with permission from ref. 20).

Fig. 4. (a) Ribbon diagram of the solution structure of ZF1-3 of TFIIIA complexed with a 15 base-pair DNA oligonucleotide.¹⁵ (b) Ribbon diagram of the protein backbone in the complex (DNA has been omitted for clarity) showing the contact regions (pink and green surfaces) present in the complex but not in the free protein. (reproduced with permission from ref. 15).

Fig. 5.

(a) Nucleotide sequence and secondary structure of the 5S RNA 55mer used to determine the structure of the TFIID-RNA complex, showing the binding of zf4, zf5 and zf6. (b) Ribbon diagram of the lowest energy refined structure of the complex. (adapted with permission from ref. 32).

Fig. 6. Disordered regions of nuclear hormone receptors. (a) Lowest-energy NMR structure of the free RXR DNA-binding domain.⁴⁰ (b) X-ray crystal structure of the complex of TR and RXR with a direct-repeat DNA duplex.³⁶ (c) Lowest-energy NMR structure of the complex of ERR2 with a single DNA half-site.⁴⁶

Fig. 7.

(a) Lowest-energy NMR structure of the complex between the LEF-1 HMG domain and its cognate DNA,⁶³ showing hydrophobic interactions in the minor groove. (b) Characteristics of the free LEF-1 HMG domain mapped onto the structure of the DNA-bound form.⁶³ Pink backbone indicates resonances missing from the spectra, light blue minimal changes between the spectra of free and DNA-bound LEF-1, and dark blue where two sets of resonances were observed in the free LEF-1 spectra. (adapted with permission from ref. 61).

Fig. 8. Models for the interaction of the linker histone H1 with the nucleosome. The globular domain binds at the nucleosome dyad and the C-terminal tail extends linearly along a linker DNA segment. (Figure reproduced with permission from ref. 68. Copyright © American Society for Microbiology)

Fig. 9.

Model for the utilization of the alternative structures observed for the ribosomal protein L20⁸⁴ in the assembly of a ribosome. (a) In the less-folded form, the attractive electrostatic surface potential is concentrated in one site, which binds to RNA1 (likely H40/41). (b) and (c) The conformational change to the more-structured form recruits RNA2 (likely H25). Boxed space-filling structures show the electrostatic surface potential for protein L20 in the two alternative structures found in the X-ray structure.⁸⁴ (adapted with permission from ref. 85).