Flexibility and Disorder in Gene Regulation: LacI/GalR and Hox Proteins*

Published, JBC Papers in Press, September 4, 2015, DOI 10.1074/jbc.R115.685032

Sarah E. Bondos†,1, Liskin Swint-Kruse§1, and Kathleen S. Matthews¶2

From the †Department of Molecular and Cellular Medicine, Texas A&M Health Science Center, College Station, Texas 77843, the §Department of Biochemistry and Molecular Biology, the University of Kansas Medical Center, Kansas City, Kansas 66160, and the ¶Department of Biosciences, Rice University, Houston, Texas 77251

To modulate transcription, a variety of input signals must be sensed by genetic regulatory proteins. In these proteins, flexibility and disorder are emerging as common themes. Prokaryotic regulators generally have short, flexible segments, whereas eukaryotic regulators have extended regions that lack predicted secondary structure (intrinsic disorder). Two examples illustrate the impact of flexibility and disorder on gene regulation: the prokaryotic LacI/GalR family, with detailed information from studies on LacI, and the eukaryotic family of Hox proteins, with specific insights from investigations of Ultrabithorax (Ubx). The widespread importance of structural disorder in gene regulatory proteins may derive from the need for flexibility in signal response and, particularly, in eukaryotes, in protein partner selection.

Over the past two decades, molecular flexibility has emerged as a critical to protein function. Although not readily apparent in crystal structures, a variety of computer simulations and solution experiments, including NMR, fluorescence, and small angle x-ray scattering, demonstrate widespread flexibility in protein molecules (1). This plasticity ranges across side chain fluctuations, domain motions, and folding transitions to the extreme pliability of intrinsically disordered regions (2). Although universal, flexibility and disorder are present to different extents in prokaryotic and eukaryotic organisms (3–5). These differences are well illustrated by genetic regulatory proteins; in prokaryotic regulators, flexibility primarily occurs in short regions around specific functional sites. Examples that have been studied in some detail include the biotin repressor BirA (6), lambda repressor (7), tetracycline repressor family (8), MerR family (9), and the LacI/GalR3 family (10). In contrast, extended regions of disorder are found in genetic regulatory proteins ranging from yeast to humans (e.g. GCN4, p53, BRCA1, and Hox proteins) and are especially evident in the activation domains of eukaryotic transcription regulators (5, 11). As paradigms for these two categories—localized flexibility and intrinsic disorder—we review these roles in two families of transcription regulators that have been extensively studied: LacI/GalR (prokaryotic) and Hox (eukaryotic).

Overview of Structures and Functions

**LacI/GalR Proteins**

The LacI/GalR family of transcription regulators comprises >4000 homologs; all members of this family are found exclusively in bacteria (10, 12, 13). The common structure of this family is a homodimer that contains one DNA-binding site and two binding sites for small-molecule, allosteric ligands (10). Some members form tetramers by a variety of mechanisms, whereas other homologs bind heteroproteins as part of the regulatory cycle (10). Fig. 1 (A–C) shows the tetrameric structure for the paradigmatic lactose repressor protein (LacI), which we use here to provide an overview of the flexible regions required for transcription regulation by LacI/GalR homologs.

First, a flexible linker connects the DNA- and ligand-binding domains (Fig. 1, A–C) (14, 15). In ~60% of LacI/GalR homologs (13), this linker includes a conserved motif that forms a “hinge helix” in known structures. The side chains of the hinge helices interact with the minor groove at the center of the two DNA half-sites, bending the operator by ~45° (Fig. 1B) (14–17). In this complex, various linker side chains form specific, hydrophobic interactions with operator DNA; thus, the linker-DNA interactions appear to be critical for recognizing specific LacI/GalR operator sequences (14–17). For LacI, the hinge helices remain compact (and presumably folded) even when the LacI-operator complex is bound to its allosteric ligand, inducer IPTG (18). However, when bound to nonspecific DNA, the NMR structure of LacI DNA-binding domains/linkers shows that the hinge helix is unfolded (16). In the absence of any DNA, both NMR and small angle x-ray scattering of full-length LacI show high mobility for the N-terminal DNA-binding domain that accompanies unfolding of the hinge helix (18, 19).

The second flexible region in the LacI/GalR proteins is a three-stranded “pivot” between the N- and C-subdomains of the regulatory domains (20, 21). Changes at this pivot occur when small, allosteric ligands bind the regulatory domain. Binding therefore alters the juxtaposition of the N-subdomains, which “pulls” the hinge helices and provides a key mechanism for altering their orientation and contacts to DNA (14, 21–23).

The third flexible region is unique to *Escherichia coli* LacI. This protein has an additional C-terminal sequence that com-
MINIREVIEW: Protein Flexibility and Disorder in Gene Regulation

Flexible linkers join the tetramerization domain to the regulatory domain, allowing the angle between the two dimers to vary (18, 24, 25). For this region, freedom of motion is essential for DNA looping and is discussed further below.

The sequences and roles of these flexible regions vary significantly among LacI/GalR homologs to generate functional diversity (reviewed in Ref. 10). For example, differences in the pivot and N-subdomain interface can lead to alternative regulatory outcomes. LacI is inducible—the consequence of binding its natural allosteric effector is to reduce DNA affinity and hence relieve repression of downstream genes (Fig. 1C). In contrast, PurR is repressible—the consequence of binding its allosteric ligand, which is to enhance DNA binding and repression (15). In addition, for ~40% of homologs, the ~18-amino acid linker that connects the core domain to the DNA-binding domain appears to be completely disordered, lacking a hinge helix (13). Similar to eukaryotic intrinsically disordered proteins (26), the linker sequence in these proteins has a high density of charge and/or prolines, although the specific positions vary (Fig. 1D).

In these homologs, disorder in the linker appears to have arisen to facilitate binding DNA operators with varied spacing between half-sites (Fig. 1E) (13, 27).

Of the homologs with disordered linkers, E. coli CytR is the best studied. For high affinity DNA binding, CytR requires cooperative binding of flanking catabolite repressor proteins (CRPs) (10, 28). The unfolded linkers in CytR allow its two N-terminal DNA-binding domains to bind operators with varied half-site spacing (Fig. 1E) (28). Notably, the disordered linkers do not propagate allosteric information to the DNA-binding domains as found for LacI. Instead, the conformational change precludes simultaneous binding to catabolite repressor protein and target DNA (29).

The range of functional differences among LacI/GalR family members illustrates how sequence changes in flexible protein regions can introduce functional variation without affecting the overall fold.

**Hox Proteins**

Within multi-cellular organisms, the family of Hox transcription regulators specifies the identities of many tissues (30, 31). Each Hox homolog regulates a different set of target genes during development to specify cellular position within the organism (e.g. various head or cardiac substructures) and to determine cellular function (30). All Hox proteins contain (i) a conserved DNA-binding domain (“homeodomain”) (32–35) and (ii) a hexapeptide motif that mediates interactions with the Exd/Pbx class of Hox co-factors (Fig. 2A) (31). Hox proteins also contain transcription activation and repression domains that influence functional specificity (e.g. Ref. 36). Large regions of the Hox proteins are intrinsically disordered, as reflected by sequence analyses, striking protease sensitivity, and challenges in protein purification (32) (Fig. 2B). Unlike the LacI/GalR homologs, both domain organization and the locations of regulatory sites (e.g. phosphorylation and splicing sites) vary considerably among Hox family proteins (Fig. 2A) (33, 37–40).

In all Hox proteins, the 60-amino acid DNA-binding homeodomain accounts for only a small fraction of the total sequence (Fig. 2A). Homeodomains contain three helices, the third of which binds the DNA major groove and is stabilized by the other two helices (34). At its N terminus, the homeodomain contains a dynamic, disordered “N-terminal arm” of 9 amino acids composed of ~55% residues with charged or polar side chains (35). The dynamic nature of this region suggests that the homeodomain might explore a disordered conformational landscape before stabilizing into a functional conformation (36).

**Conclusions and Perspectives**

Interactions between DNA-protein, protein-protein, and protein-DNA structural changes are critical for regulating gene expression. These changes often involve conformational disorder, which is a significant but under-appreciated feature of the genome and proteome. The term “disorder” describes structural flexibility and is distinct from traditional concepts in nucleic acid and protein structural biology. Disorder is noteworthy for the following reasons:

- Disorder is an intrinsic property of proteins that is evolutionarily conserved.
- Disorder provides a means to modulate protein function.
- Disorder is involved in many diseases.

Disorder is dynamic and often occurs in or near the DNA-binding region of transcription factors. Disorder in the DNA-binding region of transcription factors plays a role in several aspects of gene regulation, including DNA looping, cooperative binding, and allosteric regulation. Disorder in DNA-binding proteins likely evolved to allow flexibility in transcription factor function.
MINIREVIEW: Protein Flexibility and Disorder in Gene Regulation

In addition, hopping provides a mechanism for some eukaryote gene regulation. An efficient search process (51, 52). As discussed further below, protein flexibility is key to several of these processes. Discerning the modes of transfer can be complex, giving rise to divergent views on search mechanisms (e.g. Ref. 52).

Sliding

Once a protein associates with nonspecific DNA, sliding reduces the dimensionality of the search and thereby enhances association rates for specific sites (Fig. 3A) (48, 50, 53). As a specific example from prokaryotes, in vivo experiments with LacI indicate that (i) sliding distances are ~45 bp before dissociation from DNA, consistent with theoretical analysis (52), and (ii) obstruction by other DNA-bound proteins occurs (54). The flexibility of the LacI hinge helices appears to be critical to the sliding process because these domains are unfolded when complexed with nonspecific DNA but folded in the operator-bound form in NMR studies (16).

Despite the presence of chromatin structure, sliding is also effective in eukaryotes. For Hox homeodomains, the disordered N-terminal arms play key roles in sliding, with the length and charge of this region driving sliding dynamics (55). Electrostatic interactions dominate binding in the nonspecific complex (51, 53), although the orientation and mode of homeodomain-nonspecific DNA interaction are otherwise similar to the specific complex (unlike other transcription factor families; e.g. Ref. 56).

Hopping

In this mode of transfer, proteins bind to DNA, dissociate, and then rebind DNA at another site (Fig. 3B) (53). The length of the "hop" may be quite short or can cover long distances (49). For both prokaryotic and eukaryotic transcription factors, hopping appears to increase the speed of the search process (51, 53). In addition, hopping provides a mechanism for some eukaryote...
otic transcription factors to bypass nucleosomes when sliding along DNA (49).

**Intersegment Transfer/Brachiation**

Intersegment transfer, also called “brachiation” (using appendages to swing from object to object), allows movement from one DNA segment to the next (Fig. 3C) (55). This mechanism is distinct from hopping and is more prominent at high concentrations of DNA, as found in vivo (53). Intersegment transfer facilitates searches over long stretches of DNA because regions that are far in sequence space can be close in cellular space (as occurs via extensive packing in many eukaryotic systems) (57). This mechanism requires that two segments of DNA be simultaneously bound by protein. Hence, at least two DNA-binding interfaces are needed on the protein, and sufficient protein flexibility is required (58). The two interfaces can be provided by multimeric assembly, by multiple DNA-binding domains within a monomer, or by monomers with a single, bipartite DNA-binding domain.

For tetrameric LacI, the two dimers provide the requisite two binding interfaces, and flexibility in the segments that link the regulatory domains to the C-terminal tetramerization domain allows variation in dimer orientation (18, 24, 25). For the homeodomain, the intrinsically disordered N-terminal arm, which binds the minor groove, and the third helix, which binds the major groove, provide the two protein-DNA interfaces (55). This type of interaction accelerated the rate of target recognition by the HoxD9 homeodomain by more than 3 orders of magnitude (59). Thus, the flexibility of the N-terminal arm plus the flexible “joint” between the N-terminal arm and the helical portion of the homeodomain play a critical role in enhancing the rate of searching.

**Looping**

DNA looping occurs when regulatory proteins or their complexes simultaneously bind two DNA sites (Fig. 3D). Transient looping may occur during brachiation/intersegment transfer, but stable loops persist and impact transcription (60, 61). For example, LacI looped complexes are significantly more stable than LacI bound at a single site (62). In eukaryotes, looping can place enhancers and promoters in direct physical contact (61). Loop formation is influenced by DNA sequence and/or the presence of ancillary proteins (61, 63).

For *E. coli* tetrameric LacI, distances between target operator-binding sites can vary from hundreds to more than a thousand base pairs (62, 64). The natural *lac* operon has a spacing of ~400 bp between operators O1 and O2 and ~100 bp between O1 and O3 (64). The distances between binding sites, as well as their relative rotation around the DNA helix, can greatly alter transcription (64). In addition, protein flexibility is critical to forming looped structures (24). For tetrameric LacI binding to two operators, the two dimers adopt an “open” conformation.
(i.e. the angle between the two dimers is increased relative to the crystal structure (18)). Chemically cross-linking LacI N termini across two dimers limits dimer-dimer mobility and precludes looping (24). An alternative approach to effect looping is utilized by the homolog GalR, which forms highly stable loops with the assistance of protein HU to facilitate DNA bending (65).

The substantial intrinsic disorder found in eukaryotic regulatory proteins greatly facilitates loop formation. Many eukaryotic transcription regulators, including the Hox proteins, bind to clusters of DNA sites (66). Both side-to-side cooperative Hox binding to Hox-site clusters and back-to-back Hox-Hox interactions between two clusters can enable looped structures (Fig. 2C) (66). Hox proteins can either form loops themselves or recruit large protein complexes, such as the polycomb group proteins, the cohesion complex, and the condensing complex, to bridge distant DNA sequences (67) (Fig. 3D). In addition, Ubx binds other transcription factors that have their own DNA-binding sites near those of Ubx target DNA sequences (45, 68). This arrangement provides opportunities for creating combinatorial loops that are sensitive to cellular conditions and allow response to cell-signaling stimuli. Importantly, the intrinsically disordered regions of Ubx are required for these heterologous protein interactions (69).

Regulatory Mechanisms Exploit Protein Flexibility/Disorder

Transcription regulation often requires that regulatory proteins alter their DNA binding in response to external signals. Both the LacI/GalR and the Hox proteins utilize flexibility and disorder to transmit this incoming information to the DNA-binding domain.

Allosteric Communication in the LacI/GalR Proteins

Effector binding to LacI/GalR proteins impacts several flexible regions. For *E. coli* LacI, structures of free, DNA-bound, and IPTG-bound protein (14, 70), along with molecular dynamics simulations (23, 71, 72), have been used to study these adaptable regions. The largest changes are found in the linker region and in the N-subdomain interface of LacI (14, 23, 70). Inducer binding alters the juxtaposition of the LacI N-subdomains to bring them into closer contact (Fig. 4A) (73). The
required flexibility in this region has been explored via mutation (74). A key residue is Lys-84, which is buried within the otherwise apolar interface between the N-subdomains and changes positions in the bound and unbound structures (14). When Lys-84 was substituted with Leu or Ala, the allosteric response was diminished to \( \leq 10\)-fold (as compared with \( >10^4\)-fold for wild type), the kinetics of inducer binding were greatly slowed, and protein stability was significantly enhanced (74, 75).

The motion at the N-subdomain also alters the linker/hinge helix of LacI. The point of closest approach between the two linkers of a dimer is the side chain of Val-52. When this residue was mutated to cysteine, a disulfide bond could be formed that blocked allosteric response to inducer binding (76). Other substitutions at position 52 showed that extrinsic interactions, such as interactions with operator DNA, had more influence on LacI function and allosteric response than did the intrinsic propensity of amino acids for folding the hinge helix (77). The length of the linker region is also important. When 1–3 Glu residues were inserted after the hinge helix, LacI showed progressive decreases in DNA binding affinity and allosteric response (78). Thus, this flexible linker region must be precisely positioned (i) to allow communication between the DNA-binding and regulatory domains and (ii) to align the DNA-binding domains within each dimer.

Nevertheless, linker flexibility facilitates tolerance to significant sequence diversity (Fig. 4B). In fact, fully functional hybrids were created by fusing the LacI DNA-binding domain/linker to regulatory domains from other homologs. Each chimer has the DNA binding specificity of LacI, ligand binding of the parent regulatory domain, and allosteric response defined by the regulatory domain (79). Thus, the interface between the linker and regulatory domains is highly adaptable.

Hox Regulatory Mechanisms

Prokaryotic repressors are generally designed to respond to a limited number of signals, often only one. In contrast, eukaryotic Hox proteins integrate multiple input signals to generate highly specific outcomes unique to the tissue and organism (80). Further, these proteins must differentiate a plethora of DNA sites with both cellular and tissue specificity (30). To that end, many Hox proteins have several splice isoforms (e.g. Refs. 44 and 81), a variety of modification sites (e.g. phosphorylation) (38, 82, 83), and a number of protein partners (Fig. 4C) (31, 45, 68). These regulatory mechanisms are frequently used to diversify the functions of transcription factors (84). Although these processes typically occur within intrinsically disordered regions, their locations vary among Hox proteins.

In Ubx, all of the regulatory processes are associated with intrinsically disordered regions that also regulate DNA binding (32, 38, 44). To provide an example in each category: (i) when Ubx interacts with partner protein DIP1 via these disordered regions, Ubx transcription activation is precluded \textit{in vivo} (68); (ii) the conserved hexapeptide, which alters DNA binding specificity, and the homeodomain are connected by a disordered linker that varies from 7 to 50 amino acids in length in alternatively spliced isoforms, with the result that Ubx splicing isoforms regulate different genes and construct different tissues \textit{in vivo} (39, 40, 85); and (iii) Ubx is phosphorylated within the disordered region of the transcription activation domain in a tissue-specific manner, suggesting a regulatory function (37, 38).

The disordered regions also mediate Ubx binding to a variety of heteroprotein partners, a critical element in Hox protein function (45, 68, 69, 86). Hox proteins bind to components of the transcription machinery (45, 87), as well as to other specific transcription factors, to facilitate Hox regulation of the correct subset of genes in different tissues (Fig. 4D) (45, 88, 89). For Ubx partners identified by yeast two-hybrid methods, two key elements have emerged: (i) binding to many of these partners requires the disordered regions within Ubx and (ii) partners can be classified into specific “folds” (69). Indeed, of the selected topologies, three folds include at least five Ubx partners, jointly representing more than half of known Ubx partner proteins (Fig. 4D). Different structural families preferentially bind different disordered segments and splice isoforms of Ubx (69).

These regulatory mechanisms can influence one another (80). For example, alternative splicing impacts Hox binding to other proteins (39, 69). Likewise, phosphorylation of Hox proteins can impact protein interactions and cooperative DNA binding (90). Thus, regions that exhibit intrinsic disorder have the potential to integrate multiple sources of information to regulate and coordinate Hox functions.

Interestingly, the various disordered regions of Ubx can be distorted to allow formation of biomaterials (91). Deleting the disordered regions precludes self-assembly (92, 93). Two consequences of intrinsic disorder have the potential to make these materials commercially useful: (i) Ubx fibers are remarkably strong and extensible (91, 92) and (ii) the disordered regions allow fiber formation to accommodate a wide range of other proteins fused to the Ubx sequence (94).

Conclusion

Although prokaryotic and eukaryotic proteins exhibit many unique features, flexibility has emerged as key to transcription regulation in both kingdoms. This feature of proteins permits regulatory proteins to adapt to varied spacing among DNA-binding sites and to engage multiple mechanisms of searching for and binding to DNA target sites. Flexibility allows the variety of protein interactions required to construct complex DNA structures such as loops, either by direct binding or through interactions with other proteins. Finally, flexibility, and indeed in some cases, extensive disorder are required for regulation of transcription factor function through allosteric ligand binding, protein sequence alterations (splicing and/or posttranslational modifications), and/or protein-protein interactions. The multiple modes by which flexibility enables transcription regulation generate both diverse and highly effective mechanisms for an organism to respond to a varied local cellular environment as well as features essential for the development and function of multicellular organisms.

Acknowledgments—We express our appreciation for molecular models derived from simulations provided by Justin Drake and B. Montgomery Pettit, University of Texas Medical Branch-Galveston.
References

1. DuBay, K. H., Bowman, G. R., and Geisler, P. L. (2015) Fluctuations within folded proteins: implications for thermodynamic and allosteric regulation. *Acc. Chem. Res.* 48, 1098–1105

2. Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33, 2–8

3. Pansca, R., and Tompa, P. (2012) Structural disorder in eukaryotes. *PLoS ONE* 7, e34687

4. Xue, B., Dunker, A. K., and Uversky, V. N. (2012) Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Mol. Struct. Dyn.* 30, 137–149

5. Guo, X., Bulyk, M. L., and Hartemink, A. J. (2012) Intrinsic disorder within and flanking the DNA-binding domains of human transcription factors. *Pac. Symp. Biocomput.* 104–115, 10.1142/9789814366496_0011

6. Eginton, C., Cressman, W. J., Bachas, S., Wade, H., and Beckett, D. (2015) Dynamic perspective of protein-DNA interaction. *Biol. Concepts* 5, 21–43

7. Dover, L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatturi, V., Besra, G. S., and Füterer, K. (2004) Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthB implicated in ethionamide resistance. *J. Mol. Biol.* 340, 1095–1105

8. Chang, C.-C., Lin, L.-Y., Zou, X.-W., Huang, C.-C., and Chan, N.-L. (2015) Structural basis of the mercury(II)-mediated conformational switching of the dual-function transcriptional regulator MerR. *Nucleic Acids Res.* 43, 7612–7623

9. Swint-Kruse, L., and Trewhella, J. (2008) Ligand-induced conformational changes in nonspecific and specific protein-DNA complexes. *Science* 321, 3409–3420

10. Xu, J., and Matthews, K. S. (2009) Flexibility in the inducer binding region is crucial for allosterity in the *Escherichia coli* lactose repressor. *Biochemistry* 48, 4988–4998

11. Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., and Weber, K. (1979) Differential mobility of the N-terminal headpiece in the Lac-repressor protein. *J. Mol. Biol.* 128, 259–264

12. Swint-Kruse, L., Zhan, H., Fairbanks, B. M., Maheshwari, A., and Matthews, K. S. (2003) Perturbation from a distance: mutations that alter LacI function through long-range effects. *Biochemistry* 42, 14004–14016

13. Tungtur, S., Parente, D. J., and Swint-Kruse, L. (2011) Functionally imbedded tails of homeodomains facilitate DNA recognition by providing a functional need. *J. Mol. Biol.* 294, 487–499

14. Mowbray, S. L., and Björkman, A. J. (1999) Conformational changes of ribose-binding protein and two related repressors are tailored to fit the functional need. *J. Mol. Biol.* 294, 487–499

15. Johnson, F. B., Parker, E., and Krasnow, M. A. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotic gene Ultrabithorax. *J. Biol. Chem.* 271, 910–913

16. Moens, C. B., and Selleri, L. (2006) A generic approach to identify transcription factor–specific operator motifs; inferences for LacI–family mediated regulation in *Lactobacillus plantarum* WCF11. *BMC Genomics* 9, 145

17. Tóth-Petróczy, A., Simon, I., Fuxreiter, M., and Levy, Y. (2009) Disordered tails of homeodomains facilitate DNA recognition by providing a functional need. *Proteins* 41, 415–427

18. Johnson, F. B., Parker, E., and Krasnow, M. A. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotic gene Ultrabithorax. *J. Biol. Chem.* 271, 910–913

19. Kallipolitis, B. H., and Valentín-Hansen, P. (2004) A role for the interdomain linker region of the *Escherichia coli* CytR regulator in repression complex formation. *J. Mol. Biol.* 342, 1–7

20. Pearson, J. C., Lemonds, D., and McGinnis, W. (2005) Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904

21. Mowbray, S. L., and Björkman, A. J. (1999) Conformational changes of ribose-binding protein and two related repressors are tailored to fit the functional need. *J. Mol. Biol.* 294, 487–499

22. Xu, J., and Matthews, K. S. (2009) Flexibility in the inducer binding region is crucial for allosterity in the *Escherichia coli* lactose repressor. *Biochemistry* 48, 4988–4998

23. Flynn, T. C., Swint-Kruse, L., Kong, Y., Booth, C., Matthews, K. S., and Ma, J. (2003) AllostERIC transition pathways in the lactose repressor protein core domains: asymmetric motions in a homodimer. *Protein Sci.* 12, 2523–2541

24. Tóth-Petróczy, A., Simon, I., Fuxreiter, M., and Levy, Y. (2009) Disordered tails of homeodomains facilitate DNA recognition by providing a trade-off between folding and specific binding. *J. Am. Chem. Soc.* 131, 15084–15085

MINIREVIEW: Protein Flexibility and Disorder in Gene Regulation

25. Vanzi, F., Broglio, C., Sacconi, L., and Fuxreiter, M. (2006) Lac repressor hinge flexibility and DNA looping: single molecule kinetics by tethered particle motion. *Nucleic Acids Res.* 34, 3409–3420

26. Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* 41, 415–427

27. Francke, C., Kerkhoven, R., Wels, M., and Siezen, R. J. (2008) A generic approach to identify transcription factor–specific operator motifs; inferences for LacI–family mediated regulation in *Lactobacillus plantarum* WCF11. *BMC Genomics* 9, 145

28. Velthuis, B., Cocco, M. J., and Senear, D. F. (2006) Flexibility and adaptability in binding of *E. coli* cytidine repressor to different operators suggests a role in differential gene regulation. *J. Mol. Biol.* 362, 271–286

29. Tungtur, S., Parente, D. J., and Swint-Kruse, L. (2011) Functionally important positions can comprise the majority of a protein’s architecture. *Proteins* 79, 1589–1608

30. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271, 1247–1254

31. Kalodimos, C. G., Biris, N., Bonvín, A. M., Levandoski, M. M., Guennegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* 305, 386–389

32. Schumacher, M. A., Allen, G. S., Diel, M., Seidel, G., Hillen, W., and Brennan, R. G. (1995) Mechanism of corepressor-mediated specific DNA binding by the purine repressor. *Cell* 83, 147–155

33. Vanzi, F., Broglio, C., Sacconi, L., and Fuxreiter, M. (2006) Lac repressor hinge flexibility and DNA looping: single molecule kinetics by tethered particle motion. *Nucleic Acids Res.* 34, 2523–2541

34. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* 397, 714–719

35. Johnson, F. B., Parker, E., and Krasnow, M. A. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotic gene Ultrabithorax. *J. Biol. Chem.* 278, 20874–20887

36. McGinnis, W., Garber, R. L., Wirz, I., Kuroiwa, A., and Gehring, W. J. (1984) A homologous protein–coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* 37, 403–408

37. Tóth-Petróczy, A., Simon, I., Fuxreiter, M., and Levy, Y. (2009) Disordered tails of homeodomains facilitate DNA recognition by providing a trade-off between folding and specific binding. *J. Am. Chem. Soc.* 131, 15084–15085
MINIREVIEW: Protein Flexibility and Disorder in Gene Regulation

OCTOBER 9, 2015 • VOLUME 290 • NUMBER 41
JOURNAL OF BIOLOGICAL CHEMISTRY

24677

Protein fibers self-assembled by the Drosophila Hox transcription factor Ultrabithorax. Biomacromolecules 11, 3644–3651

Greer, A. M., Huang, Z., Oriakhi, A., Lu, Y., Lou, J., Matthews, K. S., and Bondos, S. E. (2009) The Drosophila transcription factor Ultrabithorax self-assembles into protein-based biomaterials with multiple morphologies. Biomacromolecules 10, 829–837

Tsai, S.-P., Howell, D. W., Huang, Z., Hsiao, H.-C., Lu, Y., Matthews, K. S., Lou, J., and Bondos, S. E. (2015) The effect of protein fusions on the production and mechanical properties of protein-based materials. Adv. Funct. Mater. 25, 1442–1450, 10.1002/adfm.201402997

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190

Viganò, M. A., Di Rocco, G., Zappavigna, V., and Mavilio, F. (1998) Definition of the transcriptional activation domains of three human HOX proteins depends on the DNA-binding context. Mol. Cell. Biol. 18, 6201–6212

Atchison, M. L. (2014) Function of YY1 in long-distance DNA interactions. Front. Immunol. 5, 45

Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L., and Pabo, C. O. (1994) Structural studies of the engrailed homeodomain. Protein Sci. 3, 1779–1787

Xing, Y., Clements, W. K., Kimelman, D., and Xu, W. (2003) Crystal structure of a β-catenin/axin complex suggests a mechanism for the β-catenin destruction complex. Genes Dev. 17, 2753–2764

Elrod-Erickson, M., Benson, T. E., and Pabo, C. O. (1998) High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Structure 6, 451–464

85. Reed, H. C., Hoare, T., Thomsen, S., Weaver, T. A., White, R. A., Akam, M., and Alonso, C. R. (2010) Alternative splicing modulates Ubx protein function in Drosophila melanogaster. Genetics 184, 745–758

86. Baëza, M., Viala, S., Heim, M., Dard, A., Hudry, B., Duffraisse, M., Rogulja-Ortmann, A., Brun, C., and Merabet, S. (2015) Inhibitory activities of short linear motifs underlie Hox interactome specificity in vivo. eLife 4, e06034, 10.7554/eLife.06034

87. Boube, M., Hudry, B., Immarigeon, C., Carrier, Y., Bernat-Fabre, S., Merabet, S., Graba, Y., Bourbon, H. M., and Cribbs, D. L. (2014) Drosophila melanogaster Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19. PLoS Genet. 10, e1004303

88. Slattery, M., Ma, L., Négre, N., White, K. P., and Mann, R. S. (2011) Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in Drosophila melanogaster. PLoS ONE 6, e14686

89. Merabet, S., and Dard, A. (2014) Tracking context-specific transcription factors regulating Hox activity. Dev. Dyn. 243, 16–23

90. Berry, M., and Gehring, W. (2000) Phosphorylation status of the SCR homeodomain determines its functional activity: essential role for protein phosphatase 2A/ß’. EMBO J. 19, 2946–2957

91. Howell, D., Tsai, S., Churion, K., Patterson, J., Abbey, C., Atkinson, J., Porterpan, D., You, Y., Meissner, K., Bayless, K., and Bondos, S. (2015) Identification of multiple dityrosine bonds in materials composed of the Drosophila protein Ultrabithorax. Adv. Funct. Mater. 10.1002/adfm.201502852

92. Huang, Z., Lu, Y., Majithia, R., Shah, J., Meissner, K., Matthews, K. S., Bondos, S. E., and Lou, J. (2010) Size dictates mechanical properties for protein fibers self-assembled by the Drosophila Hox transcription factor Ultrabithorax. Biomacromolecules 11, 3644–3651

93. Greer, A. M., Huang, Z., Oriakhi, A., Lu, Y., Lou, J., Matthews, K. S., and Bondos, S. E. (2009) The Drosophila transcription factor Ultrabithorax self-assembles into protein-based biomaterials with multiple morphologies. Biomacromolecules 10, 829–837

94. Tsai, S.-P., Howell, D. W., Huang, Z., Hsiao, H.-C., Lu, Y., Matthews, K. S., Lou, J., and Bondos, S. E. (2015) The effect of protein fusions on the production and mechanical properties of protein-based materials. Adv. Funct. Mater. 25, 1442–1450, 10.1002/adfm.201402997

95. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190

96. Viganò, M. A., Di Rocco, G., Zappavigna, V., and Mavilio, F. (1998) Definition of the transcriptional activation domains of three human HOX proteins depends on the DNA-binding context. Mol. Cell. Biol. 18, 6201–6212

97. Atchison, M. L. (2014) Function of YY1 in long-distance DNA interactions. Front. Immunol. 5, 45

98. Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L., and Pabo, C. O. (1994) Structural studies of the engrailed homeodomain. Protein Sci. 3, 1779–1787

99. Xing, Y., Clements, W. K., Kimelman, D., and Xu, W. (2003) Crystal structure of a β-catenin/axin complex suggests a mechanism for the β-catenin destruction complex. Genes Dev. 17, 2753–2764

100. Elrod-Erickson, M., Benson, T. E., and Pabo, C. O. (1998) High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Structure 6, 451–464