WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

SHORT COMMUNICATION

Intrinsic disorder in yeast transcriptional regulatory network

Gajinder Pal Singh and Debasis Dash*

Institute of Genomics and Integrative Biology (CSIR), Delhi University Campus, Delhi, India

ABSTRACT

Intrinsic disorder has been shown to be important in mediating protein–protein and protein–DNA interactions. Proteins involved in regulatory functions are particularly enriched in intrinsic disorder. In this study we explored the role of intrinsic disorder in transcriptional regulatory network of yeast. Using disorder prediction programs we show that transcription factors (TFs) regulating large number of targets (transcriptional hubs) have significantly increased intrinsic disorder, though targets regulated by multiple TFs did not show increased intrinsic disorder. Intrinsic disorder may allow these transcriptional hubs to bind to diverse promoter regions of their targets in different contexts, and may also allow complex regulatory control of transcriptional hubs that are involved in coordinating different cellular processes.

Proteins 2007; 68:602–605.

Key words: transcriptional hubs; protein disorder; transcriptional network; unfolded proteins; unstructured proteins.

INTRODUCTION

Large scale elucidation of cellular networks has allowed global properties and design principles to be uncovered, which would not have been possible by studying individual components.[1] In yeast, manually curated data along with large scale chromatin coimmunoprecipitation-chip (ChIp-chip) experiments have been used to construct transcriptional regulatory network linking transcription factors (TFs) to their target genes.[2] This regulatory network, like many other biological and nonbiological networks, displays scale-free topology, indicating the presence of transcriptional hubs, which regulate large number of target genes.[2]

Intrinsically disordered regions or proteins lack a well-defined three-dimensional structure. Large proportion of eukaryotic proteomes has been predicted as disordered.[3–6] In particular, proteins involved in signaling and regulatory functions have been found to be enriched in intrinsic disorder.[7–9] Two recent articles highlight the role of intrinsic disorder in eukaryotic TFs.[10,11] Disordered regions have been reported to undergo disorder-order transition upon binding to their partners, which has been suggested to allow weak, yet specific interactions.[6,12–15] The structural malleability of these regions has been proposed to allow binding to multiple partners.[16–18] These regions have also been implicated in many other regulatory post-translational modifications including phosphorylation and degradation.[19,20]

In protein–protein interaction networks it was found that proteins interacting with large number of proteins (hubs) have increased intrinsic disorder,[21,22] particularly those involving transient interactions.[23] In this work, we studied the role of intrinsic disorder in yeast transcriptional regulatory network, and show increased intrinsic disorder in transcriptional hubs, which regulate large number of target genes.

MATERIALS AND METHODS

Transcriptional regulatory network

Transcriptional regulatory network was obtained from Balaji et al.[2] It was compiled from genetic, biochemical, and large scale ChIP-chip experiments. It includes 157

Grant sponsor: CSIR; Grant number: CMM0017.
*Correspondence to: Debasis Dash, G.N.R. Knowledge Centre for Genome Informatics, Institute of Genomics and Integrative Biology, Delhi University Campus, Mall Road, Delhi 110007, India.
E-mail: ddash@igib.res.in
Received 14 December 2006; Revised 26 February 2007; Accepted 6 March 2007
Published online 17 May 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21497

specific TFs, 4410 target genes, and 12,873 regulatory interactions. The connectivity distribution (the plot of number of target genes n, regulated by a TF and the number of TFs that regulate n target genes) was best fitted by a power-law equation (linear fit on log–log plot), suggesting its scale-free nature.[2] Transcriptional hubs were defined as top 20% of TFs with large number of target genes.[2] Thirty one TFs met this criterion. These TFs together accounted for more than half of all regulatory interactions.[2]

Intrinsic disorder prediction

We used modified Uversky method and IUPred to predict intrinsic disorder.

Modified Uversky method is based on the observation that disordered proteins have high mean net charge and/or low mean hydrophobicity.[24] It is a global disorder predictor and gives a Netscore to each protein.[25] This was calculated for each protein using following formula:

$$\text{Netscore} = [(\text{mean net charge}) - 2.785 \times (\text{mean net hydrophobicity}) + 1.151]/2.952$$

A protein with positive Netscore is predicted to be disordered, while negative score predicts ordered protein.[25]

IUPred method is based on the assumption that disordered regions do not form sufficient favorable interactions to fold, and thus have high estimated energy content.[26] It is a local disorder prediction program and assigns residues in proteins as disordered or ordered.

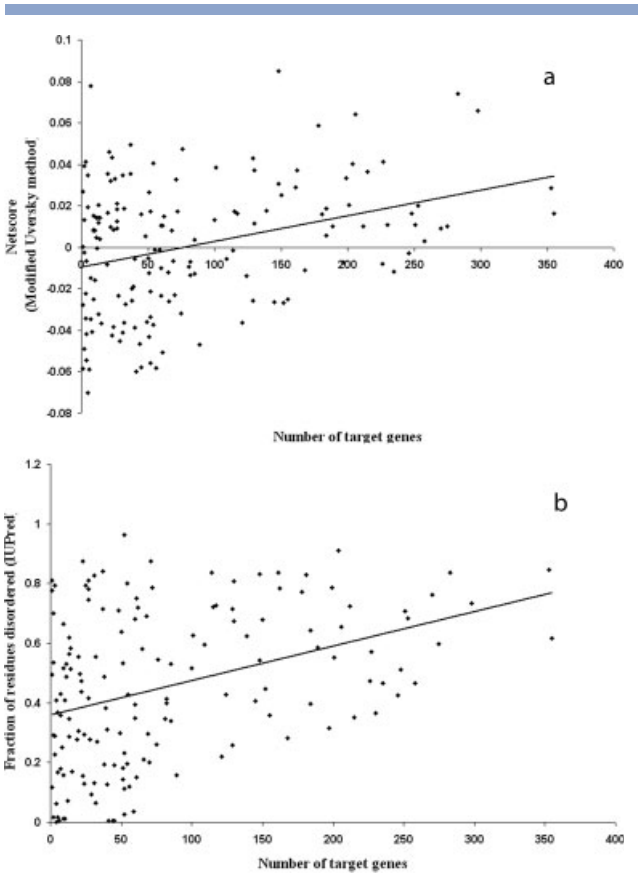

Figure 1

*Correlation between intrinsic disorder in TFs with number of target genes they regulate. (**a**) Intrinsic disorder predicted by modified Uversky method wherein positive scores predict that the protein is disordered, while negative score predict an ordered protein. (**b**) Intrinsic disorder predicted by IUPred method.*

RESULTS

Increased intrinsic disorder in yeast transcriptional hubs

We used the dataset of transcriptional regulatory network compiled by Balaji et al.[2] linking 157 TFs to 4410 target genes. We dropped one TF (YER108c) because of its ambiguous status in the Saccharomyces Genome Database. The intrinsic disorder was predicted using Uversky's modified method, which generates a Netscore from the protein sequence.[25] A positive Netscore predicts intrinsically disordered protein, whereas a negative Netscore predicts ordered protein. We examined the relationship between the predicted intrinsic disorder of a given TF with the number of target genes it regulates. The result is displayed in Figure 1(a). Although the points are somewhat scattered indicating variability, a significant positive correlation was observed (Pearson correlation coefficient 0.33, $P < 10^{-4}$). To examine this trend from a different perspective, we computed the fraction of amino acid residues predicted to be in disordered conformation in each TF using the IUPred method,[26] and investigated the relationship with the number of genes regulated by corresponding TFs. The result is displayed in Figure 1(b). A significant positive correlation was observed (Pearson correlation coefficient 0.37, $P < 10^{-5}$). The strength of correlation between predicted intrinsic disorder of TFs with number of target genes regulated is nearly same by both methods. We examined this relationship further by segregating the dataset of TFs into "ordered" set with negative Netscore (<0.0) and "disordered" set with positive Netscore (≥ 0.0). We tested the difference in the mean number of target genes regulated by the ordered set and the disordered set using t test. We observed significant difference in the mean number of target genes regulated by the two sets of TFs (two-tailed $P < 0.005$). Therefore, it is clear that on an average, TFs with intrinsically disordered conformation regulate larger number of target genes compared with ordered TFs.

TF hub proteins (comprising top 20% of the TFs with large number of targets) had 79% of its proteins with positive Netscore in contrast to the nonhubs that had 47% of its proteins with positive Netscore. The difference

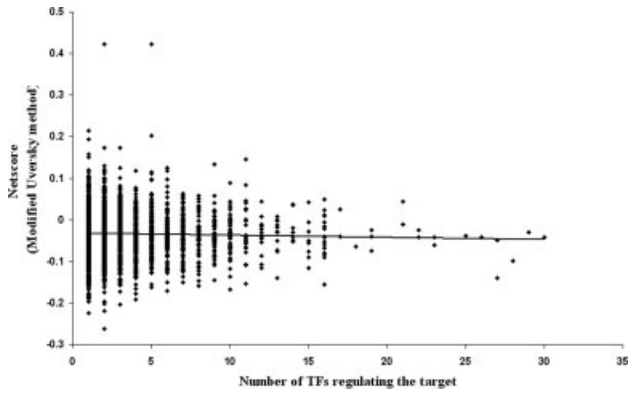

Figure 2

Lack of correlation between intrinsic disorder in target genes with number of TF it is regulated by. Intrinsic disorder was calculated by modified Uversky method.

in proportion of disordered proteins in the TF hubs dataset is significantly higher than that in the TF nonhub dataset (test for proportions, $P < 0.005$).

Target proteins regulated by multiple TFs do not show increased intrinsic disorder

We examined the Netscore of proteins encoded by genes whose transcription is regulated by multiple TFs, the underlying hypothesis is that these proteins might function through complex regulatory mechanism and might have higher intrinsic disorder. To test this possibility we examined the correlation between intrinsic disorder of the proteins encoded by 4410 target genes and the number of TFs regulating the target gene (Fig. 2). No significant correlation was evident showing that the intrinsic disorder of proteins encoded by target genes is independent of the number of TFs regulating them.

DISCUSSION

Transcription regulation in eukaryotes is viewed as a highly complex process[27] involving interactions between multiple TFs,[28–32] in a combinatorial fashion,[33–38] allowing relatively small number of TFs to generate myriad variety of gene expression states. These TFs also interact with various coregulators and chromatin regulatory proteins[39] and with basal transcription machinery. Rapid divergence of this regulatory control[40] ensures that most target genes of a transcription factor offer a unique binding environment. This combinatorial regulation of gene expression thus means that TFs bind to promoters of multiple target genes, particularly TF hubs bind to promoters of large number of genes. Conformational plasticity provided by intrinsic disorder has been reported to allow specific binding to multiple interacting partners.[16–18] For example GSK3beta interacts with both axin and FRAT through a disordered loop with significant overlap of the binding sites,[41] HMGA proteins possesses almost no secondary structure while free in solution, act as transcriptional hubs, regulating the expression of wide variety of genes.[42] Two recent reports describe the role and prevalence of intrinsic disorder in TFs and showed significant enrichment of intrinsic disorder in this class of proteins.[10,11] Activation domains of the TFs were particularly found to be disordered in these analyses, which are involved in many protein–protein interactions with other transcriptional regulators. This is consistent with experimental evidence of disorderness of activation domains in many TFs.[14,43–45] Our analysis of intrinsic disorder in transcriptional network shows increased intrinsic disorder in TFs regulating large number of target genes. Increased intrinsic disorder in TF hubs may allow binding to their different target promoters in a variety of contexts.

It is reasonable to expect that proteins encoded by genes that are regulated by large number of TFs might also be under complex posttranslational regulation. It was shown that in yeast cell-cycle genes, transcriptional and posttranslational regulations have coevolved.[46] Increase in transcriptional regulation correlated with increased posttranslational regulation. We found that intrinsic disorder of proteins is independent of the number of TFs regulating its gene.

Analysis of protein interaction networks has also shown increased intrinsic disorder in proteins interacting with large number of partners (hubs).[21–23] Thus it seems that common solution has been arrived at in both transcriptional and protein interaction networks with respect to intrinsic disorder allowing binding to multiple partners.

CONCLUSIONS

In summary, we found significantly increased intrinsic disorder in TFs regulating large number of target genes. The structural plasticity afforded by intrinsic disorder might be critical in allowing these TF hubs to regulate large number of target genes, each one of which is expected to present unique binding environment.

ACKNOWLEDGMENTS

The authors thank Dr. S. Ramachandran and Dr. Mythily Ganapathi for helpful suggestions.

REFERENCES

1. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–113.
2. Balaji S, Babu MM, Iyer LM, Luscombe NM, Aravind L. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. J Mol Biol 2006;360:213–227.

DOI 10.1002/prot

3. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. Genome Inform Ser Workshop Genome Inform 2000;11:161–171.
4. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol 2004;337:635–645.
5. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. Comparing and combining predictors of mostly disordered proteins. Biochemistry 2005;44:1989–2000.
6. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. J Mol Graph Model 2001;19:26–59.
7. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol 2002;323:573–584.
8. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J 2005;272:5129–5148.
9. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. J Mol Recognit 2005;18:343–384.
10. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. Biochemistry 2006;45: 6873–6888.
11. 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.
12. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 2005;6:197–208.
13. Tompa P. Intrinsically unstructured proteins. Trends Biochem Sci 2002;27:527–533.
14. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. Curr Opin Struct Biol 2002;12:54–60.
15. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. Protein Sci 2002;11:739–756.
16. Wright PE, Dyson HJ. Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. J Mol Biol 1999; 293:321–331.
17. Fink AL. Natively unfolded proteins. Curr Opin Struct Biol 2005; 15:35–41.
18. Tompa P, Szasz C, Buday L. Structural disorder throws new light on moonlighting. Trends Biochem Sci 2005;30:484–489.
19. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 2004;32:1037–1049.
20. Singh GP, Ganapathi M, Sandhu KS, Dash D. Intrinsic unstructuredness and abundance of PEST motifs in eukaryotic proteomes. Proteins 2006;62:309–315.
21. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol 2006;2:e100.
22. Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P. Disorder and sequence repeats in hub proteins and their implications for network evolution. J Proteome Res 2006;5:2985–2995.
23. Singh GP, Ganapathi M, Dash D. Role of intrinsic disorder in transient interactions of hub proteins. Proteins 2006;66:761–765.
24. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 2000; 41:415–427.
25. Pandey N, Ganapathi M, Kumar K, Dasgupta D, Das Sutar SK, Dash D. Comparative analysis of protein unfoldedness in human housekeeping and non-housekeeping proteins. Bioinformatics 2004;20:2904–2910.
26. Dosztanyi Z, Csizmok V, Tompa P, Simon I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J Mol Biol 2005;347:827–839.
27. Holstege FC, Young RA. Transcriptional regulation: contending with complexity. Proc Natl Acad Sci USA 1999;96:2–4.
28. Kel OV, Romaschenko AG, Kel AE, Wingender E, Kolchanov NA. A compilation of composite regulatory elements affecting gene transcription in vertebrates. Nucleic Acids Res 1995;23:4097–4103.
29. Halfon MS, Carmena A, Gisselbrecht S, Sackerson CM, Jimenez F, Baylies MK, Michelson AM. Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. Cell 2000;103:63–74.
30. Fickett JW. Coordinate positioning of MEF2 and myogenin binding sites. Gene 1996;172:GC19–GC32.
31. Zhong H, Vershon AK. The yeast homeodomain protein MAT alpha2 shows extended DNA binding specificity in complex with Mcm1. J Biol Chem 1997;272:8402–8409.
32. Newman JR, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science 2003;300:2097–2101.
33. Klinkenberg LG, Mennella TA, Luetkenhaus K, Zitomer RS. Combinatorial repression of the hypoxic genes of Saccharomyces cerevisiae by DNA binding proteins Rox1 and Mot3. Eukaryot Cell 2005; 4:649–660.
34. Wang W, Cherry JM, Nochomovitz Y, Jolly E, Botstein D, Li H. Inference of combinatorial regulation in yeast transcriptional networks: a case study of sporulation. Proc Natl Acad Sci USA 2005; 102:1998–2003.
35. Kato M, Hata N, Banerjee N, Futcher B, Zhang MQ. Identifying combinatorial regulation of transcription factors and binding motifs. Genome Biol 2004;5:R56.
36. Pilpel Y, Sudarsanam P, Church GM. Identifying regulatory networks by combinatorial analysis of promoter elements. Nat Genet 2001;29:153–159.
37. GuhaThakurta D, Stormo GD. Identifying target sites for cooperatively binding factors. Bioinformatics 2001;17:608–621.
38. Hannenhalli S, Levy S. Predicting transcription factor synergism. Nucleic Acids Res 2002;30:4278–4284.
39. Mannervik M, Nibu Y, Zhang H, Levine M. Transcriptional coregulators in development. Science 1999;284:606–609.
40. Maslov S, Sneppen K, Eriksen KA, Yan KK. Upstream plasticity and downstream robustness in evolution of molecular networks. BMC Evol Biol 2004;4:9.
41. Dajani R, Fraser E, Roe SM, Yeo M, Good VM, Thompson V, Dale TC, Pearl LH. Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-APC scaffold complex. EMBO J 2003;22: 494–501.
42. Reeves R. Molecular biology of HMGA proteins: hubs of nuclear function. Gene 2001;277:63–81.
43. Sigler PB. Transcriptional activation. Acid blobs and negative noodles. Nature 1988;333:210–212.
44. Johnson PF, Sterneck E, Williams SC. Activation domains of transcriptional regulatory proteins. J Nutr Biochem 1993;7:386–398.
45. Ferreira ME, Hermann S, Prochasson P, Workman JL, Berndt KD, Wright AP. Mechanism of transcription factor recruitment by acidic activators. J Biol Chem 2005;280:21779–21784.
46. Jensen LJ, Jensen TS, de LU, Brunak S, Bork P. Co-evolution of transcriptional and post-translational cell-cycle regulation. Nature 2006;443:594–597.