PROTEINS: Structure, Function, and Bioinformatics 66:761–765 (2007)

SHORT COMMUNICATION

Role of Intrinsic Disorder in Transient Interactions of Hub Proteins

Gajinder Pal Singh, Mythily Ganapathi, and Debasis Dash*
Institute of Genomics and Integrative Biology (CSIR), Delhi University Campus, Delhi, India

ABSTRACT **Hubs in the protein–protein interaction network have been classified as "party" hubs, which are highly correlated in their mRNA expression with their partners while "date" hubs show lesser correlation. In this study, we explored the role of intrinsic disorder in date and party hub interactions. The data reveals that intrinsic disorder is significantly enriched in date hub proteins when compared with party hub proteins. Intrinsic disorder has been largely implicated in transient binding interactions. The disorder to order transition, which occurs during binding interactions in disordered regions, renders the interaction highly reversible while maintaining the high specificity. The enrichment of intrinsic disorder in date hubs may facilitate transient interactions, which might be required for date hubs to interact with different partners at different times. Proteins 2007;66:761–765.** © 2006 Wiley-Liss, Inc.

Key words: date hub; party hub; intrinsic disorder; transient interaction; reversible interaction

INTRODUCTION

Intrinsically unstructured regions in proteins almost completely lack tertiary structure under physiological conditions.[1] Such proteins or regions (also known as disordered or intrinsically disordered or natively unfolded) contain characteristically high net charge and low content of hydrophobic amino acids, and have been shown to have a biased distribution of certain amino acids.[2–4] These regions are increasingly being implicated in several important, regulated cellular functions such as binding interactions, proteolysis, post translational modifications, etc.[5]

There are several reports of the involvement of unstructured regions in binding interactions.[5–10] Intrinsically disordered proteins have been shown to be enriched in nucleic acid and protein binding classes.[11–14] Disordered domains involved in DNA/RNA/protein binding have also been shown to be conserved.[15]

Natively unfolded regions are implicated in multiple interactions as their structural plasticity allows them to efficiently interact with several regions.[2,16–18] There have been few studies wherein the multiple interactions of a protein have been correlated with the presence of intrinsically unstructured regions.[13,19,20] Unstructured regions have been reported to undergo disorder to order transition as they bind to their interaction partners. This transition is associated with a large decrease in conformational entropy, which uncouples binding strength from specificity and causes highly specific interactions to become reversible.[10,21]

In this study, our objective was to explore the role of intrinsic disorder in date and party hub interactions. The data set of Han et al. (2004), which classifies hubs in protein interaction network as "party" and "date" has been used in this study.[22] It has been recently shown that hub proteins contain more intrinsic disorder when compared with nonhub proteins,[19,20] and this has been inferred to facilitate the multiple interactions of hub proteins. To further the current understanding of the role of intrinsic disorder in protein-interaction networks, we analyzed the subclasses of hubs (date and party hubs) categorized on the basis of correlation of mRNA expression of hub with its partners.

Hubs in yeast interaction network have been defined as proteins with greater than five interaction partners.[22] For each hub, average Pearson correlation coefficient (PCC) was calculated between the expression level of the hub and with the expression level of each of its respective interaction partner. It has been inferred that hubs with lower average PCC (date hubs) bind to different partners at different times while hubs with higher PCC (party hubs) interact with most of their partners simultaneously.[22] We have used the date and party hubs dataset of Han et al. (2004) for examining the role of intrinsic disorder in date and party hub interactions.

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 11 June 2006; Revised 8 September 2006; Accepted 25 September 2006

Published online 11 December 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21281

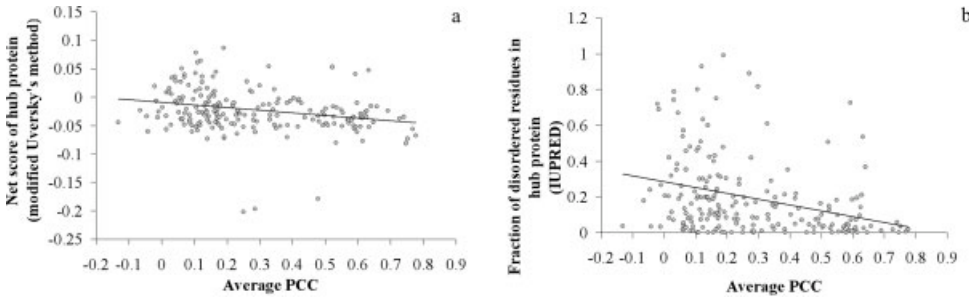

Fig. 1. Correlation between fraction of intrinsic disorder in hub proteins with the average PCC of the subnetwork. (**a**) Intrinsic disorder was calculated by modified Uversky's method wherein positive scores denote that the protein is predicted to be disordered, while negative score denote an ordered protein. (**b**) Intrinsic disorder was calculated by IUPRED method.

MATERIALS AND METHODS

Data

Protein interaction network data was obtained from Han et al. (2004). This dataset involves 1379 proteins with 2493 high-confidence interactions. Among the 2493 interactions, 931 interactions involved date hubs and 907 interactions involved party hubs. In this set, hubs were defined as proteins with more than five interacting partners.[22] The average Pearson correlation coefficient (PCC) for the subnetworks comprising the hub and its interacting partners was obtained from Han et al. (2004).

A dataset of 70,647 derived binary interactions data along with the socioaffinity index score for the interaction was obtained from the genome-wide yeast protein complex data of Gavin et al.[23] This list is available at http://yeast-complexes.embl.de/complexview.pl?rm=download.

Prediction of Disordered Regions

We have used four different methods for prediction of intrinsic disorder, namely, modified Uversky's method, IUPRED, GlobPlot, and DisEMBL (REM465).

Modified Uversky's method is based on observation that disordered proteins have high mean net charge and/or low mean hydrophobicity.[24] It is a global disorder predictor and gives a net score to each protein.[24] A protein with positive net score is predicted to be disordered, while negative score predicts ordered protein. The net score is also quantitatively related to degree of unfoldedness of the protein.

IUPRED is based on an assumption that disordered regions do not form sufficient favorable interactions to fold and thus have high estimated energy content.[25]

GlobPlot is based on propensities derived from nonglobular regions in PDB.[26]

Remark 465 method of DisEMBL is an artificial neural network-based method trained on regions with nonassigned electron densities in PDB.[27]

Secondary Structure Prediction

Consensus secondary structure for proteins was predicted using NN, SOPM, DPM, DSC, GOR4, PHD, PREDA, and SIMPA96 algorithms available at NPS server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html).

TABLE I. Prediction of Intrinsic Disorder in Date and Party Hub Proteins by Global Disorder Prediction Method-Modified Uversky's Method

Hubs	No. of unfolded proteins	No. of folded proteins
Date (91)	28 (30.8)	63
Party (108)	11 (10.2)	97

The numbers in parentheses in column 2 denote the percentage of predicted unfolded proteins.

RESULTS

Intrinsic Disorder Inversely Correlates With Average PCC of Hubs

We calculated the intrinsic disorder of the hub proteins and studied its correlation with the corresponding average PCC of the subnetwork. The plot between average PCCs and intrinsic disorder of hub proteins showed a weak yet significant negative correlation as determined by two independent disorder prediction methods [Fig. 1(a,b); modified Uverksy's method-PCC of -0.253, P value = 3.2E-04; IUPRED-PCC of -0.3155, P value = 5.6E-06].

Predicted Intrinsic Disorder in Date and Party Hub Proteins

A significantly high number of proteins were predicted to be unfolded in the date hub category when compared with the party hubs by modified Uversky's method (Table I; $\chi^2 = 13.28$, $df = 1$, P value = 0.0003). Mean net score derived from modified Uversky's method for date hubs was more than twice than that of party hubs (-0.014 and -0.033, respectively, t-test P value = 3E-04).

Date hubs also showed significantly higher number of disordered segments when compared with party hubs, by the three local disorder prediction methods used (IUPRED, Globplot, DisEMBL; Table II). When compared with party hubs, date hubs had almost twice as many long disordered regions ($\geq$30 residues) per protein, their average length was longer, and covered twice as much portion of proteins

TABLE II. Prediction of Intrinsic Disorder in Date and Party Hub Proteins by Local Disorder Predictors

Prediction method	Total no. of proteins	No. of disordered segments of ≥30 residues	No. of proteins with disordered segments of ≥30 residues	Average length of disordered segments	Average no. of disordered segments per protein	% Residues in disordered segments
IUPRED						
Date hubs	91	118	49	90.39	1.30	20.35
Party hubs	108	69	41	67.74	0.64	7.75
GlobPlot						
Date hubs	91	74	46	91.22	0.81	12.88
Party hubs	108	39	34	77.54	0.36	5.01
DISOPRED (REM465)						
Date hubs	91	48	33	43.96	0.53	4.03
Party hubs	108	32	20	42.63	0.30	2.26

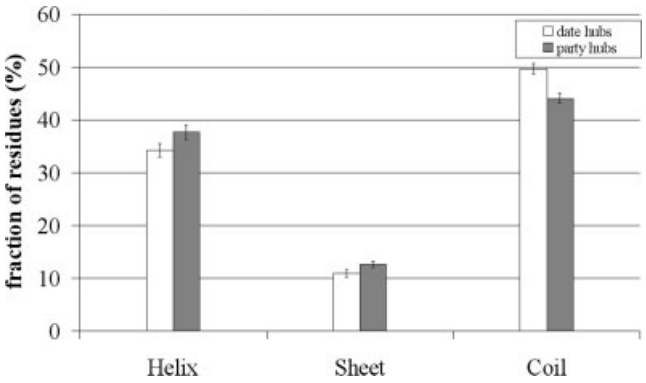

Fig. 2. Predicted secondary structure of residues in date and party hub proteins. Error bars indicate ±SE.

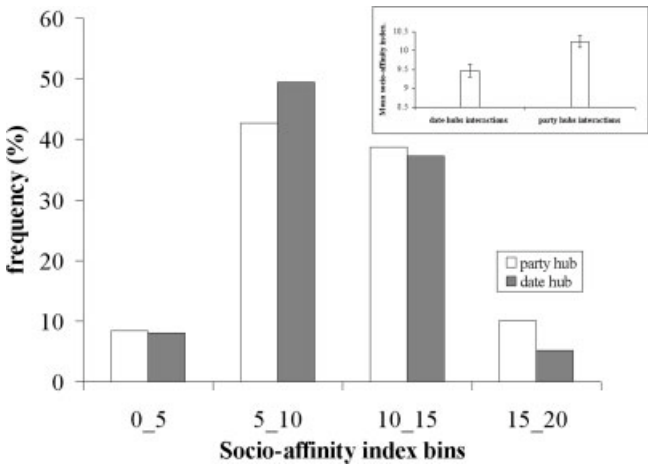

Fig. 3. Frequency distribution of socioaffinity indices of date and party hub interaction pairs. The differences in mean socioaffinity index between date and party hubs have been shown as an insertion in the figure. Error bars indicate ±SE.

(Table II). Not only the number of long disordered segments was higher in date hubs, but also the number of proteins harboring such long disordered segments was significantly higher (Table II).

We also did secondary structure analysis for the residues in date and party hub proteins. Since only few of these proteins had representation in PDB (5 full length date hub and 14 party hub proteins), we analyzed secondary structure by prediction methods. Coils are not necessarily disordered, but prevalence of coiled residues reflects prevalence of intrinsic disorder. The data shows that the fraction of coils in date hub proteins is significantly higher when compared with the party hub proteins (Fig. 2; *t*-test for fraction of coil residues-*P* value = 1E-05).

Transient Nature of Date Hub Interactions as Revealed by Protein Complex Data

We tried to analyze the distribution of date hub and party hub interactions in the recently available genome wide data of yeast protein complexes.[23] The data was compiled by searching the hub-interaction partner pairs of Han et al. (2004) in the protein–protein interaction pairs provided by Gavin et al. (2006).[22,23] We find that 42% of the interaction pairs involving date hubs (387/931) and 67% of interaction pairs involving party hubs (603/907) mapped to the derived binary interactions provided by Gavin et al. (2006) ($\chi 2$ value = 113.8, $df = 1$, P value = 1.4E-26).

Gavin et al. (2006) defined a term "socio-affinity index," which measured the propensity of proteins to form partnerships. This index was tentatively correlated to the available dissociation constant for the interaction pairs.[23] We find that the distribution of socioaffinity index was significantly different between interactions involving date hubs and party hubs (Fig. 3; *t*-test P value = 0.001, Wilcoxon rank sum test = 0.006).

DISCUSSION

Molecular recognition involving reversible but specific interaction poses particular problems to rigid body assumption of proteins. To have specific recognition, interaction partners should have large number of contacts and thus high binding affinity, but this also reduces the dissociation rate (and thus reversibility) of interaction.[28] In other words, binding affinity and binding specificity are coupled. Intrinsic disorder has been suggested to allow specific interaction in reversible/weak manner by uncoupling binding affinity and binding specificity.[28–30] Intrinsic

disorder also facilitates posttranslational modifications including phosphorylation,[31] which are frequently involved in transient signaling events. Indeed, many proteins including GSK3β,[32] BRCA1,[33] p21,[34] p53,[35] calcineurin,[36] E-cadherin,[6] CBP,[37] and HMGA[36] reversibly interact with their partners via intrinsically disordered regions.

We show that average PCC of hubs correlates weakly yet significantly with measures of intrinsic disorder in hubs. Classification of the hubs into date and party also showed clear overrepresentation of intrinsic disorder in date hubs as analyzed by a global and three local prediction methods. Intrinsic disorder might be utilized by the date hub proteins to interact with multiple partners in a transient manner. The significant prevalence of party hub interactions in the genome wide protein complexes dataset (67% compared to 42% for date for date hub interactions) supports the transient nature of date hub interactions, which makes it less likely to be identified in a screen for stable protein complexes. This is further corroborated by our data wherein date hubs show significantly low socioaffinity indices when compared with party hubs, since previously, socioaffinity index have been tentatively correlated to available dissociation constants in the literature by Gavin et al. (2006). These data suggest that date hubs bind to their interacting partners in a transient manner.

Date hubs are shown to be enriched in "cell signaling" and "transcription" categories[22] when compared with party hubs. The enrichment of disorder in date hubs might enable them to perform their functions in signaling pathways, which are extensively mediated by transient interactions.

Although, date hubs are similar in length to party hubs, the number of interaction partners of 91 date hubs (931) are significantly higher than that for 108 party hubs (907) (t-test P-value = 0.0002). Since party hubs presumably interact simultaneously with their partners and date hubs presumably interact with partners nonsimultaneously (potentially utilizing overlapping sites) the enrichment of intrinsic disorder in date hubs might also enable them to participate in multiple interactions.

It has been proposed that date hubs act as an intermodule or "higher level" connectors, and are involved in global organization of modules in the interaction network.[22] We propose that intrinsic disorder in date hubs might help in their role in organizing interaction networks.

CONCLUSIONS

The date hub proteins harbor significantly higher intrinsic disorder when compared with party hub proteins. This property might play an important role in facilitating reversible, nonsimultaneous, multiple interactions of date hub proteins, and thus help in organizing protein interaction networks.

Note added in proof: While this manuscript was under review, another study [38] using a different disorder prediction method found significantly more date hubs had longer disordered segments compared to party hubs, again supporting the robust difference between date and party hubs independent of prediction method used.

REFERENCES

1. Uversky VN. What does it mean to be natively unfolded? Eur J Biochem 2002;269(1):2–12.
2. 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.
3. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 2000; 41:415–427.
4. Williams RM, Obradovi Z, Mathura V, Braun W, Garner EC, Young J, Takayama S, Brown CJ, Dunker AK. The protein nonfolding problem: amino acid determinants of intrinsic order and disorder. Pac Symp Biocomput 2001:89–100.
5. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. Biochemistry 2002;41: 6573–6582.
6. Huber AH, Stewart DB, Laurents DV, Nelson WJ, Weis WI. The cadherin cytoplasmic domain is unstructured in the absence of β-catenin. A possible mechanism for regulating cadherin turnover. J Biol Chem 2001;276:12301–12309.
7. Weikl T, Abelmann K, Buchner J. An unstructured C-terminal region of the Hsp90 co-chaperone p23 is important for its chaperone function. J Mol Biol 1999;293:685–691.
8. Prasch S, Schwarz S, Eisenmann A, Wohrl BM, Schweimer K, Rosch P. Interaction of the intrinsically unstructured phage λ N Protein with *Escherichia coli* NusA. Biochemistry 2006;45:4542–4549.
9. Tozawa K, Macdonald CJ, Penfold CN, James R, Kleanthous C, Clayden NJ, Moore GR. Clusters in an intrinsically disordered protein create a protein-binding site: the TolB-binding region of colicin E9. Biochemistry 2005;44:11496–11507.
10. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 2005;6:197–208.
11. 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.
12. 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;25:25.
13. Liu J, Tan H, Rost B. Loopy proteins appear conserved in evolution. J Mol Biol 2002;322:53–64.
14. Bustos DM, Iglesias AA. Intrinsic disorder is a key characteristic in partners that bind 14-3-3 proteins. Proteins 2006;63:35–42.
15. Chen JW, Romero P, Uversky VN, Dunker AK. Conservation of intrinsic disorder in protein domains and families. II. Functions of conserved disorder. J Proteome Res 2006;5:888–898.
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. 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.
20. Patil A, Nakamura H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. FEBS Lett 2006;580:2041–2045.
21. Tompa P. Intrinsically unstructured proteins. Trends Biochem Sci 2002;27:527–533.
22. Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 2004;430:88–93.

23. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G. Proteome survey reveals modularity of the yeast cell machinery. Nature 2006;440: 631–636.
24. 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.
25. 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.
26. Linding R, Russell RB, Neduva V, Gibson TJ. GlobPlot: exploring protein sequences for globularity and disorder. Nucleic Acids Res 2003;31:3701–3708.
27. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. Structure 2003;11:1453–1459.
28. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with α-helix-forming molecular recognition elements. Biochemistry 2005;44:12454–12470.
29. 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.
30. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 2005;6:197–208.
31. 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.
32. Dajani R, Fraser E, Roe SM, Yeo M, Good VM, Thompson V, Dale TC, Pearl LH. Structural basis for recruitment of glycogen synthase kinase 3β to the axin-APC scaffold complex. EMBO J 2003;22: 494–501.
33. Mark WY, Liao JC, Lu Y, Ayed A, Laister R, Szymczyna B, Chakrabartty A, Arrowsmith CH. Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? J Mol Biol 2005;345:275–287.
34. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. Proc Natl Acad Sci USA93:11504–11509.
35. Rustandi RR, Baldisseri DM, Weber DJ. Structure of the negative regulatory domain of p53 bound to S100B(ββ). Nat Struct Biol 2000;7:570–574.
36. Dunker AK, Obradovic Z. The protein trinity-linking function and disorder. Nat Biotechnol 2001;19:805–806.
37. Radhakrishnan I, Perez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell 1997;91:741–752.
38. Ekman D, Light S, Bjorklund AK, Elofsson A. What properties characterize the hub proteins of the protein–protein interaction network of Saccharomyces cerevisiae? Genome Biol 2006;7:R45.