Intrinsically Disordered Proteins and Conformational Noise: Implications in Cancer

Gita Mahmoudabadi[1], Krithika Rajagopalan[2,3], Robert H. Getzenberg[2,8], Sridhar Hannenhalli[4,5] Govindan Rangarajan[6,7,*], and Prakash Kulkarni[2,8,*]

[1]Division of Bioengineering, California Institute of Technology, Pasadena, CA, [2]Department of Urology, and [8]Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD, [4]Department of Cell Biology & Molecular Genetics, and [5]Center for Bioinformatics & Computational Biology, University of Maryland, College Park, MD, and [6]Department of Mathematics, and [7]Centre for Neuroscience, Indian Institute of Science, Bangalore, India

[*]Corresponding Authors: Prakash Kulkarni
Department of Urology
The Johns Hopkins University School of Medicine
600 N Wolfe St, 105B Marburg
Baltimore, MD 21287
P-410.502.4962
F-410.502.9336
E-pkulkar4@jhmi.edu

And

Govindan Rangarajan
Department of Mathematics
Indian Institute of Science
Bangalore 560 012, India
P- +91-80-2360 0373
F- +91-80-2360 0146
E- rangaraj@math.iisc.ernet.in

[3]Present address: Department of Biological Sciences
Columbia University
New York, NY 10027

Key words: Intrinsically disordered proteins, noise, protein-protein interaction network, state-switching, cancer, evolution

Running title: Conformational noise and state-switching

Abstract

Intrinsically disordered proteins IDPs are proteins that lack a rigid 3D structure under physiological conditions at least *in vitro*. Despite the lack of structure, IDPs play important roles in biological processes and transition from disorder to order upon binding to their targets. With multiple conformational states and rapid conformational dynamics, they engage in myriad and often 'promiscuous' interactions. These stochastic interactions between proteins and their partners, defined here as conformational noise, is an inherent characteristic of IDP interactions. The collective effect of conformational noise is an ensemble of protein network configurations from which the most suitable can be explored in response to perturbations, conferring protein networks with remarkable flexibility and resilience. Moreover, the ubiquitous presence of IDPs as transcriptional factors and, more generally, as hubs in protein networks, is indicative of their role in propagation of transcriptional (genetic) noise. As effectors of transcriptional and conformational noise, IDPs rewire protein networks and unmask latent interactions in response to perturbations. Thus, noise-driven activation of latent pathways could underlie state-switching events such as cellular transformation in cancer. To test this hypothesis, we have created a model of a protein network with the topological characteristics of a cancer protein network and tested its response to a perturbation in presence of IDP hubs and conformational noise. Because numerous IDPs are found to be epigenetic modifiers and chromatin remodelers, we hypothesize that they could further channel noise into stable, heritable genotypic changes.

Introduction

From a biological perspective, cancer may be considered as a case of state-switching. Thus, the transformation of a normal cell to a transformed phenotype, and very frequently, from a non-aggressive to a highly lethal phenotype, represent state-switching by these cells. But what drives this switching and the resulting acquisition of characteristics necessary to become a cancer cell remains poorly understood. However, state-switching is not unique to cancer and in fact, is a frequently observed phenomenon in biology. In this manuscript we provide a conceptual framework to address this issue.

Intrinsically disordered proteins (IDPs) are proteins that lack rigid 3D structures either along their entire length or in localized regions when free of partners in solution.[1] With many possible conformations, and consequently many possible interactions, they play important biological roles in major cellular processes such as cell cycle regulation, transcription regulation, signal transduction, and regulation of protein self-assembly within protein networks.[1,2] Studies on protein interaction networks (PINs) from yeast to humans have demonstrated that hub proteins, defined as those that interact with many partners in the network, are significantly more disordered than end proteins, defined as those that interact with only a few partners.[3] Furthermore, a binary classification of hubs and ends into ordered and disordered subclasses has shown a significant enrichment of entirely disordered proteins and a significant depletion of entirely ordered proteins in hubs relative to ends[4] underscoring the role of IDPs as hubs within PINs. An investigation of IDPs has also revealed that they are often reused in multiple pathways to produce different physiological outcomes as they assume hub positions in signaling and

regulatory pathways.[5] Considering the widespread implication of IDPs in several pathological states[1], here we explore their possible role in state-switching by rewiring protein networks in response to perturbations. Although we have focused on cancer, we believe our thesis is not restricted to cancer and may be more generally applicable to address state-switching in biology.

Conformational Noise

The internal milieu in every living cell is abuzz with noise. Recent evidence indicates that the information transduced in cellular signaling pathways is significantly affected by noise.[6,7] In fact, it has been proposed that, noise in these pathways maybe generated by the interconnected and promiscuous nature of protein interactions that are necessary to transduce signals[8]. However, how this noise arises, and what consequences it has on cell fate, are poorly understood. We posit that the noise due to protein conformational dynamics of IDPs (conformational noise) underlies the 'promiscuous' nature of protein interactions and impacts biological information transfer.

A distinguishing feature that places IDPs at hub positions in signaling pathways, and protein networks in general, is their remarkable ability to undergo disorder-to-order transitions upon binding to their biological targets (coupled folding and binding).[9] This feature enables their interaction with a broad range of binding partners that include other proteins, membranes, nucleic acids and small molecules.[9] Each IDP contributes to network plasticity by having a rugged energy landscape with many local minima separated with low-energy barriers.[10] In contrast to energy landscapes of highly ordered proteins, this type of energy landscape enables stochastic IDP fluctuations between numerous conformational states, in response to modest

4

perturbations such as the overexpression of its binding partners.[10] Moreover, a recent comparison of k_{on} and k_{off} rates between highly ordered and disordered proteins revealed significant differences, indicative of faster IDP interaction kinetics.[11] Thus, conformational noise and fast interaction kinetics together may allow IDPs to rapidly explore the network search space and to activate previously masked interaction options in response to intrinsic and extrinsic sources of perturbation.

Transcriptional Noise Propagation in Protein Networks

It is now widely accepted that noise in gene expression underlies substantial phenotypic variations resulting in genetically identical cells to switch states and behave differently.[6,7] This is also referred to as genetic noise but we will refer to this type of noise as 'transcriptional noise'. In state-switching systems, especially those driven by positive feedback loops, transcriptional noise allows cell state choice to be probabilistic and can lead to novel phenotypes such as drug resistance in cancer cells.[12, 13] State-switching systems are widespread in biological systems as they may be integral to development, stress response, pathological states such as cancer, and evolution.[12,13]

By virtue of their prominent role as hub proteins, we posit that IDPs play special roles as direct and indirect propagators of transcriptional noise in the system. We hypothesize that in protein networks with random topologies, the effect of transcriptional noise on protein interactions would be significantly buffered due to the absence of hub proteins. However, studies on protein networks have shown that rarely are these networks randomly organized. Hubs, particularly IDP hubs, are demonstrated as critical factors in a variety of contexts: signaling in

cellular differentiation and cancer, as well as transcriptional and translational regulation in both disease and normal conditions.[14] Prominent examples of IDP hubs include Oct4, Sox2, Nanog, and others that are components of the MAP kinase, RTK signaling, and the NFkB/P53/CBP-P300 pathways. Often these IDP hubs are simultaneously employed in very different pathways, such as GSK3, which is used in both Wnt and the insulin signaling pathways.[14]

In these networks, transcriptional noise in the expression of a hub protein could influence many other interactions within the network. For example, adenylate cyclase (*cya*) in *E. coli* that has cyclic AMP as its downstream product is evolutionarily selected to have low-noise expression.[15] Under noisy *cya* expression, cyclic AMP concentrations would also be affected by noise, which in turn would have adverse effects on many regulatory processes in the cell. In at least two scenarios the effect of transcriptional noise on a limited number of genes could spread throughout protein networks: 1) in the event that noisy hubs serve as transcriptional factors, thereby directly contributing to transcriptional noise in the expression of many other proteins, and 2) indirectly, where noisy hubs exhibit high connectivity to each other, such as in bacterial metabolic networks[16] and in cancer protein networks.[17]

Interestingly, an analysis of the abundance of disordered regions in three transcription factor datasets and two control sets with several hundred proteins in each revealed that from 94.13% to 82.63% of transcriptional factors possess extended disordered regions, relative to 54.51% and 18.64% of the proteins in two control datasets, which indicates the significant prevalence of disordered regions in transcription factors.[18] One of the most interesting examples in eukaryotes is the CCCTC binding factor known as CTCF, which is responsible for regulation of

chromatin architecture by forming chromatin loops and co-localizing spatially separated DNA segments.[19] With more than 77,000 identified binding sites on the eukaryotic genome, CTCF plays a critical role in transcriptional regulation, both in promoting and in repressing the expression of many gene targets.[20] Moreover, much of CTCF functional versatility is attributed to the disordered polypeptide segments that form more than half of its amino-acid sequence and are thought to function as scaffolds for protein assembly in transcriptional regulation.[19] Given that most chromatin remodelers and transcriptional factors are IDPs[18, 21, 22], it is conceivable that transcriptional noise in these IDP hubs will directly contribute to the total transcriptional noise of the system.

This hypothesis is partly supported by recent evidence suggesting that the percentage of disordered regions is an important determinant of dosage sensitivity among proteins.[23, 24] In order to function properly, cells have developed strict spatiotemporal control over IDP expression levels.[25, 26] Consequently, many proteins associated in diseases are IDP hub proteins, including p53, p21, p27, BRCA1, kalirin, ubiquitin, and calmodulin, among many others.[3] In a study by Iakoucheva et al, 79% of cancer-associated proteins were found to contain predicted disordered segments that were 30 residues or longer[27], compared to only 13% of proteins from a set of proteins with well-defined structures that contained such long disordered regions. Consistent with these observations, a majority of the Cancer/Testis Antigens, a group of proteins that are typically restricted to the male germ cells but are aberrantly expressed in cancer, were also predicted to be intrinsically disordered and occupy hub positions underscoring the pervasiveness of IDP overexpression in cancer.[28]

Here, we have chosen to highlight the role of IDPs in propagation of transcriptional noise because this type of noise has systematically been studied and is a common feature of cellular processes in both normal and diseased conditions. However, it is worth noting that the role of IDPs is not limited to the propagation of transcriptional noise. Rather, IDPs could likely relay, and even amplify, other intrinsic and extrinsic types of noise and perturbations in the system.

Results

IDP-mediated protein network rewiring – a model: For a long time, it was tacitly assumed that most networks adopt a random architecture wherein an edge (connection) between each pair of nodes has equal probability, independent of the other edges.[29] However, pioneering work by Barabasi and colleagues[30] indicated that biological networks like many other networks they interrogated, adopt an architecture wherein the degree distribution *P(k)* exhibits a power-law behavior as a function of the degree *k*. In particular, *P(k) ~ k⁻ᵞ* with only a few nodes (hub nodes) having numerous edges while the majority of the nodes have very few edges. As such, these networks are robust to failure of random nodes, but vulnerable to failure of hubs.[30] Considering this type of network topology, we asked whether perturbation to IDP hub proteins could account for such dramatic changes as state-switching (drug resistance in cancer cells, for instance) and whether these changes could be passed onto the progeny? What are the effectors that could ultimately turn perturbations such as transcriptional noise into a driving force for transient and stable phenotypic diversity in health and disease? We present a model that predicts that, the effects of conformational noise in response to perturbations could

induce significant changes in network topology. This model employs the topological characteristics of cancer protein networks in the simulated protein network.

We begin with an undirected protein network using the Barabasi-Albert algorithm[30] with 3000 nodes. In this hypothetical protein network, the average degree per node is 4.0 and the highest degree is 103. For the network under consideration, the power law decay factor γ is observed to be 2.34. These values are typical of those found in protein networks. By defining nodes with degrees greater than or equal to 10 to be hubs, we find 160 hubs in this simulated network. It is known that a majority of the hubs in protein networks are occupied by IDPs[3]. To obtain a biologically realistic percentage, we investigated the hub proteins listed by Kar et al[17] based on the human protein-protein interaction network constructed by Jonsson and Bates[31]. For each protein sequence, we used the MobiDB database[32] of multi-source annotation of disordered regions to obtain a consensus annotation of disorder. We then calculated the fraction of protein covered by contiguous disordered regions of at least 30 consecutive amino acids. Based on this quantification of IDP, 69% of our hub proteins were classified as IDP. A similar analysis done previously by Haynes et al[33] identified ~80% of the human hub proteins as IDP. Therefore, for our simulation, we assumed 75% of the hubs to be IDP. We refer to this configuration of the protein network as the ground-state. Thus, by default, this particular configuration of the network represents its ground-state threshold.

We now introduce a source of perturbation in the protein network by increasing the expression of an IDP beyond its ground-state expression level, and representing conformational noise by adding 10 interactions to each IDP. It is now well established that several oncogenes and genes

overexpressed in cancer are IDPs.[23, 28] Indeed, in a test for a generalized involvement of intrinsic disorder in signaling and cancer, Iakoucheva et al[27] applied a neural network predictor of natural disordered regions (PONDR VL-XT) to four protein datasets: human cancer-associated proteins (HCAP), signaling proteins (AfCS), eukaryotic proteins from SWISS-PROT (EU_SW) and non-homologous protein segments with well-defined (ordered) 3D structure (O_PDB_S25). The authors observed that PONDR VL-XT predicted ≥ 30 consecutive disordered residues for 79(+/-5)%, 66(+/-6)%, 47(+/-4)% and only 13(+/-4)% of the proteins from HCAP, AfCS, EU_SW, and O_PDB_S25, respectively, indicating significantly more intrinsic disorder in cancer-associated and signaling proteins as compared to the two control sets. Furthermore, it is also the case that the IDP nodes in a cancer cell protein networks have higher degrees compared to those in protein networks of a normal cell.[31, 34, 35] In contrast to normal networks where hubs are more connected to non-hubs, cancer hubs prefer to interact with other hubs as compared to non-hub proteins.[17] Thus, it follows that from a network perspective, assortativity in cancer protein networks is increased compared to the protein networks in normal cells.

We model this increase in assortativity by preferentially connecting 50% of the new links to other IDPs. This results in several striking consequences (Fig. 1). First, the average degree per node increases (from 3.995) to 4.793 with the 95% confidence interval given by (4.791, 4.795). Further, the average degree per IDP node increases (from 18) to 34.24 with the 95% confidence interval given by (32.99, 35.33). Second, the maximum degree increases (from 103) to 116 with the 95% confidence interval given by (112,120). Third, the power law decay factor γ decreases (from 2.34) to 1.90 with the 95% confidence interval given by (1.87, 1.92). However, the most important consequence of all is the increased resilience of the network to random

perturbations. It has been mathematically[36] and empirically[37] demonstrated that resilience increases when the power law exponent decreases and the maximum degree cutoff increases. Resilience is also further enhanced by an increase in assortativity[36-38] and decrease in the shortest path length[37], another attribute of cancer protein networks. Interestingly, it has also been shown that even directed attacks on hubs are ineffective if the assortativity is sufficiently high.[38] Ultimately the threshold is reset to a higher (new ground-state) level in this new, resilient, phenotype.

Discussion

From this model we infer that conformational noise in response to a perturbation could result in topological changes in PINs. Moreover, we hypothesize that increased IDP interactions, underlying the observed increase in assortativity, could lead to the unmasking of latent pathways which drive latent phenotypes (Fig. 2). Thus, perturbations such as propagated transcriptional noise may induce profound changes in a cell's phenotype.

In summary, we have discussed the dual action of IDPs as both propagators of transcriptional noise and responsive elements to perturbations via conformational noise. Is it possible that IDPs associations with noise could provide cancer cells with novel avenues for developing drug resistance (state-switching)? Could tumor heterogeneity be a result of the stochastic activation of pathways that can lead to drug resistance? We have provided a model that serves as a conceptual framework for asking these questions; however, future models should expand on this effort by including spatial and temporal variables to better define the system of interest. Moreover, in modeling such a multi-causal phenomenon as cancer, an integrative approach

should be taken to include the contribution of factors such as post-transcriptional and epigenetic modifications. Finally, considering that IDPs undergo extensive post-translational modifications and have many alternatively spliced isoforms, it is important that future research examines the effects of these processes on IDP conformational dynamics and network re-wiring capabilities. Two recent reports on tissue-specific alternative splicing have shown that regulated alternative exons frequently remodel interactions to establish tissue-dependent protein networks.[39, 40] Future studies and more sophisticated models should help address these recent findings and provide additional insight.

But how might information be transmitted in the reverse direction from phenotype to genotype? Could changes in protein networks due to IDPs' association with noise have implications for a cell's progeny? It is now widely accepted that information that is transmitted transgenerationally can be encoded epigenetically. Interestingly, several proteins that are involved in epigenetically sculpting the chromatin are IDPs[21, 22] suggesting that re-wiring of protein networks could result in heritable epigenetic changes. For example, ~90% of proteins which recognize or interact with post-translationally modified histones in the ChromDB database that covers broad range of chromatin-related proteins, contain long intrinsically disordered regions.[21] Similarly, in a manually compiled dataset of 37 Polycomb/Trithorax-related proteins that remodel chromatin altering the accessibility of DNA to factors required for gene transcription, 31 (83%) and 27 (73%) proteins were found to contain long regions of disorder by two different algorithms, respectively[21] underscoring the propensity of chromatin remodelers to be intrinsically disordered. Moreover, emerging evidence suggests a nexus between transcription factors and chromatin remodelers[41] and between transcription factors

12

and DNA repair proteins.[42] Given the overwhelming presence of intrinsic disorder in transcription factors, it is likely that the ripple effect of transcriptional noise mediated by IDPs could be coupled to genetic changes that permanently alter the genome. Together, such changes instituted in response to information transfer from the phenotype to genotype by IDPs could potentially guide the evolution of a cell under normal and diseased conditions.

Materials and Methods

We investigated the hub proteins listed by Kar et al.[17] based on the human protein-protein interaction network constructed by Jonsson and Bates[31]. For each protein sequence, we used the MobiDB database[32] of multi-source annotation of disordered regions to obtain a consensus annotation of disorder. We then calculated the fraction of protein covered by contiguous disordered regions of at least consecutive 30 amino acids. Protein network interaction simulations were carried out using Matlab (MATLAB version 7.12. Natick, Massachusetts: The MathWorks Inc., 2011).

References

1. Uversky VN, Dunker AK. Understanding protein non-folding. Biochim Biophys Acta 2010; 1804:1231-1264; PMID: 20117254.

2. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys 2008; 37:215-246; PMID: 18573080.

3. Patil A, Kinoshita K, Nakamura H. Hub promiscuity in protein-protein interaction networks. Int J Mol Sci 2010; 11:1930-1943; PMID: 20480050.

4. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol 2006; 2:e100; PMID: 16884331

5. Gsponer J, Babu M. The rules of disorder or why disorder rules. Prog Biophysics Mol Biol 2009; 99:94-103; PMID: 19344736.

6. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, et al. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science 2010; 329: 533-538; PMID: 20671182.

7. Munsky B, Neuert G, van Oudenaarden A. Using gene expression noise to understand gene regulation. Science 2012; 336:183-187; PMID: 22499939.

8. Ladbury JE, Arold ST. Noise in cellular signaling pathways: causes and effects. Trends Biochem Sci. 2012; 37:173-178; PMID: 22341496.

9. Tompa P, Csermely P. The role of structural disorder in the function of RNA and protein chaperones. FASEB J 2004; 18: 1169-1175; PMID: 15284216.

10. Turoverov KK, Kuznetsova IM, Uversky VN. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Prog Biophy Mol Biol 2010; 102: 73-84; PMID: 20097220.

11. Huang Y, Liu Z. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism. J Mol Biol 2009; 393: 1143-1159; PMID: 19747922.

12. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature 2010; 467:167-173; PMID: 20829787.

13. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell 2010; 141: 69-80; PMID: 20371346.

14. Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: regulation and disease. Current Opinion Struct Biol 2011; 21: 432-440; PMID: 21514144.

15. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. Nature Genetics 2002; 31:69-73; PMID 11967532.

16. Yamada T, Bork, P. Evolution of biomolecular networks: lessons from metabolic and protein interactions. Nature Rev Mol Cell Biol 2009; 10: 791-803; PMID: 19851337.

17. Kar G, Gursoy, Keskin, O. Human cancer protein-protein interaction network: a structural perspective. PloS Comp Biol 2009; 5(12) e1000601; PMID: 20011507.

18. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. Biochemistry 2006; 45: 6873-6888; PMID: 16734424.

19. Martinez SR, Miranda JJ. CTCF terminal segments are unstructured. Protein Science 2010; 19:1110-1116; PMID: 20196073.

20. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res 2012; 22: 1680-1688; PMID: 22955980.

21. Sandhu KS. Intrinsic disorder explains diverse nuclear roles of chromatin remodeling proteins. J Mol Recognit 2009; 22: 1-8; PMID: 18802931.

22. Beh LY, Colwell LJ, Francis NJ. A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. Proc Natl Acad Sci USA 2012; 109:1063-71; PMID: 22517748.

23. Vavouri T, Semple JI, Garcia-Verdugo R, Lehner B. Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell 2009; 138:198-2008; PMID: 19596244.

24. Marcotte EM, Tsechansky M. Disorder, promiscuity, and toxic partnerships. Cell 2009; 138: 16-18; PMID: 19596229.

25. Gsponer J, Futschik ME, Teichmann, SA, Babu MM. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. Science 2008; 322: 1365-1368; PMID: 19039133.

26. Edwards YJ, Lobley AE, Pentony MM, Jones DT. Insights into the regulation of intrinsically disordered proteins in the human proteome by analyzing sequence and gene expression data. Genome Biol 2009; 10: R50; PMID: 19432952.

27. Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol 2002; 323: 573-584; PMID: 12381310.

28. Rajagopalan K, Mooney SM, Parekh N, Getzenberg RH, Kulkarni P. A majority of the cancer/testis antigens are intrinsically disordered proteins. J Cell Biochem 2011; 112-3256-3267; PMID: 21748782.

29. Erdős P, Rényi A. On random graphs. Publ Math 1959; 6: 290-297.

30. Barabasi AL, Albert R. Emergence of scaling in random networks. Science 1999; 286: 509-512; PMID: 10521342.

31. Jonsson PF, Bates PA. Global topological features of cancer proteins in the human interactome. Bioinformatics 2006; 22: 2291-2297; PMID: 16844706.

32. Di Domenico T, Walsh I, Martin AJM, and Tosatto, SCE. MobiDB: a comprehensive database of intrinsic protein disorder annotations. Bioinformatics. 2012; 28(15):2080-1.

33. Haynes H, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsci disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Computational Biology 2006; 2:e100.

34. Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. Genome Biol 2010; 11: R53; PMID: 20482850.

35. Xia J, Sun J, Jia P, Zhao Z. Do cancer proteins really interact strongly in the human protein-protein interaction network? Comput Biol Chem 2011; 35: 121-125; PMID: 21666777.

36. Barrat A, Barthelemy M, Vespignani A. In Dynamical Processes on Complex Networks. Cambridge University Press (2008).

37. Rodrigues FA, Costa L, Barbieri AL. Resilience of protein-protein interaction networks as determined by their large-scale topological features. Mol. Biosys 2011; 7: 1263-1269: PMID: 21298132.

38. Newman MEJ. Assortative mixing in networks. Phys Rev Lett 2002; 89: 208701; PMID: 12443515.

39. Ellis JD, Barrios-Rodiles M, Colak R, Irimia M, Kim T, Calarco JA, Wang X, Pan Q, O'Hanlon D, Kim PM, Wrana JL, Blencowe BJ. Tissue-specific alternative splicing remodels protein-protein interaction networks.Mol Cell 2012; 46: 884-892; PMID: 22749401.

40. Buljan M, Chalancon G, Eustermann S, Wagner GP, Fuxreiter M, Bateman A, Babu MM. Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. Mol Cell 2012; 46: 871-883; PMID: 22749400.

41. Murawska M, Brehm A. CHD chromatin remodelers and the transcription cycle. Transcription 2011; 2: 244-253; PMID: 22223048.

42. Fong YW, Inouye C, Yamaguchi T, Cattoglio C, Grubisic I, Tjian R. A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. Cell 2011; 147: 120-131; PMID: 21962512.

Figure Legends:

Fig.1. Degree distribution plot. The figure shows the probability *P(k)* that a given protein interacts with *k* other proteins (the so-called degree distribution) on a log-log scale. The figure compares the degree distribution of a normal protein regulatory network (black circles) with that of a network impacted by cancer (red rectangles). A majority of the hubs in protein networks are IDPs and these IDPs have aberrant expression profiles in cancer and moreover they preferentially interact with other hubs. Consequently, the slope of the straight line fitted to the points for a cancer network (red solid line) is smaller than that for a normal network (black dashed line). Further the maximum degree increases in a cancer network (the red rectangles extend further to the right as compared to the black circles). All simulations were carried out using Matlab (MATLAB version 7.12. Natick, Massachusetts: The MathWorks Inc., 2011).

Fig. 2. Rewiring of protein networks facilitates state-switching by activating latent pathways. A) The state of a cell with phenotype A is depicted in grey and shows a simple protein networks with three proteins (1-3) of which one is an IDP (indicated in dark blue), and expressed at different levels represented by the three vectors. This configuration represents the protein network's ground state threshold. B) Depicts a transition state. A perturbation causes increased IDP expression (protein 3). Overexpression of the IDP results in promiscuity and the protein network explores the network search space shown by the various dashed lines. This transition state is depicted state in yellow around the grey area. C) The state of the cell after it has transitioned to phenotype B from phenotype A represented in yellow. A particular

configuration of the protein network that increased its fitness is 'selected' which now
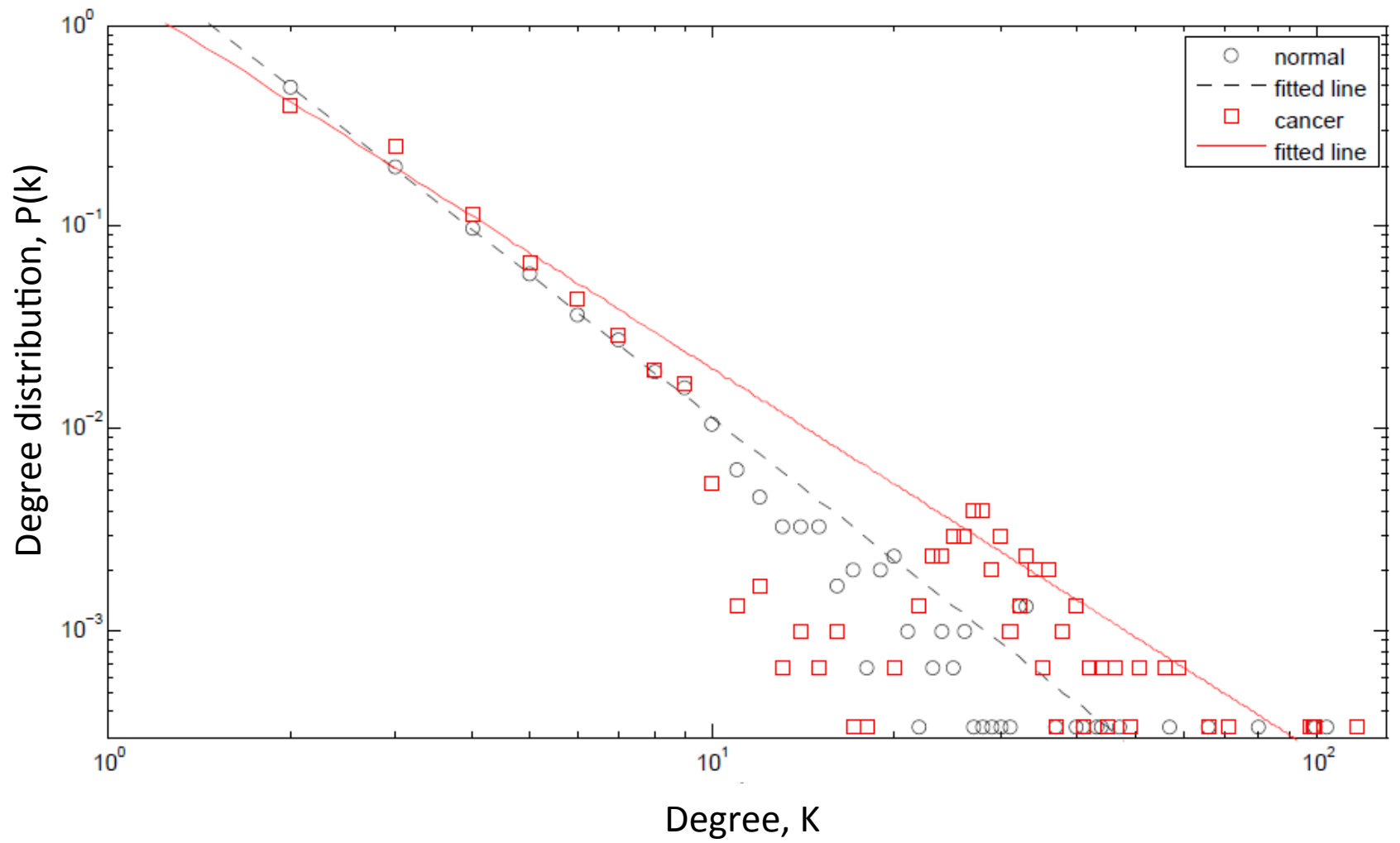
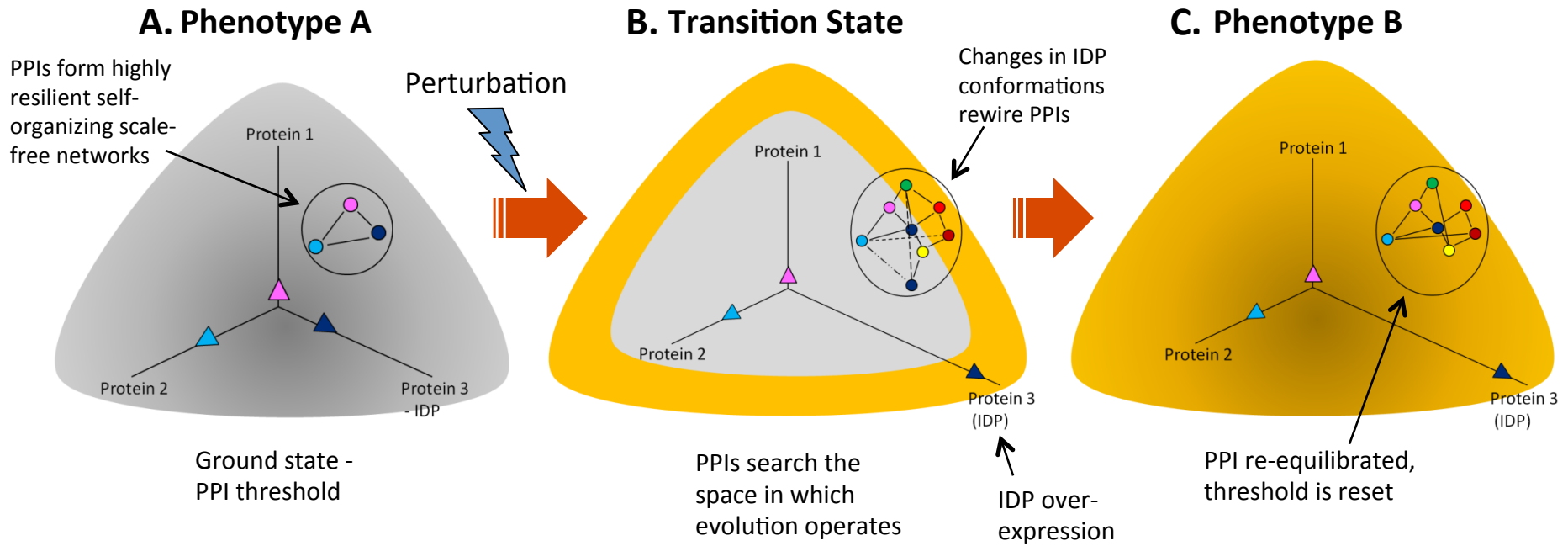represents the new ground state.

Fig.1.

A. Phenotype A

PPIs form highly resilient self-organizing scale-free networks

Protein 1

Protein 2

Protein 3 - IDP

Ground state - PPI threshold

Perturbation

B. Transition State

Protein 1

Protein 2

Protein 3 (IDP)

Changes in IDP conformations rewire PPIs

IDP over-expression

PPIs search the space in which evolution operates

C. Phenotype B

Protein 1

Protein 2

Protein 3 (IDP)

PPI re-equilibrated, threshold is reset

Fig. 2.