Systematic control of protein interactions for systems biology

Nitin Bhardwaj[a], Declan Clarke[b], and Mark Gerstein[a,c,d,1]

[a]Program in Computational Biology and Bioinformatics, and Departments of [b]Chemistry, [c]Molecular Biophysics and Biochemistry, and [d]Computer Science, Yale University, New Haven, CT 06520

Since its birth, systems biology has gained a great deal from the protocols devised to study phenomena at the level of single proteins and nucleic acids. Such protocols find broad markets and utility at higher levels of biological organization, from next-generation sequencing, which uses modified nucleotides and fluorescent identifiers (1), to ChIP-seq analysis, which identifies histone modifications and binding sites in protein–DNA interactions (2). In PNAS, Sivaramakrishnan and Spudich introduce a system for interrogating interactions between pairs of proteins, domains, and peptides (3), and it is very possible that their invention will find applicability in the construction and analysis of large-scale protein–protein interaction networks (4).

The technology devised by Sivaramakrishnan and Spudich is based on Forster resonance energy transfer (FRET), which relies on the proximity-dependant excitation of an acceptor molecule by a donor (5). This excitation may be detected with spectral imaging, so FRET enables the investigator to directly probe interactions between candidate pairs of biomolecules, by using molecular cloning to tether the donor to one candidate and the acceptor to the second. For years this technology has enjoyed broad applicability (6–8), and biosensors using FRET have been applied in live cell imaging to investigate an array of biological processes, such as phosphorylation, protease activity, fluctuations in membrane potential, and changes in redox potentials, pH, and other environmental conditions (9, 10).

To engineer the effective concentrations of interacting species, Sivaramakrishnan and Spudich have recruited the ER/K single α-helix as an intradomain and intraprotein linker, in conjunction with FRET sensing, to quantitatively regulate and study interactions between proteins fused to the ends of the link. The effective concentrations are determined by the size of the linkage. Perhaps not surprisingly, the on-rate of the interaction decreases as the length of the linker increases. There is little effect on the dissociation rate. Importantly, the value of this system lies not only in the ability to modulate the interaction frequencies of two molecules without major perturbations to their structures, but also to obtain meaningful and accurate estimates of the natural
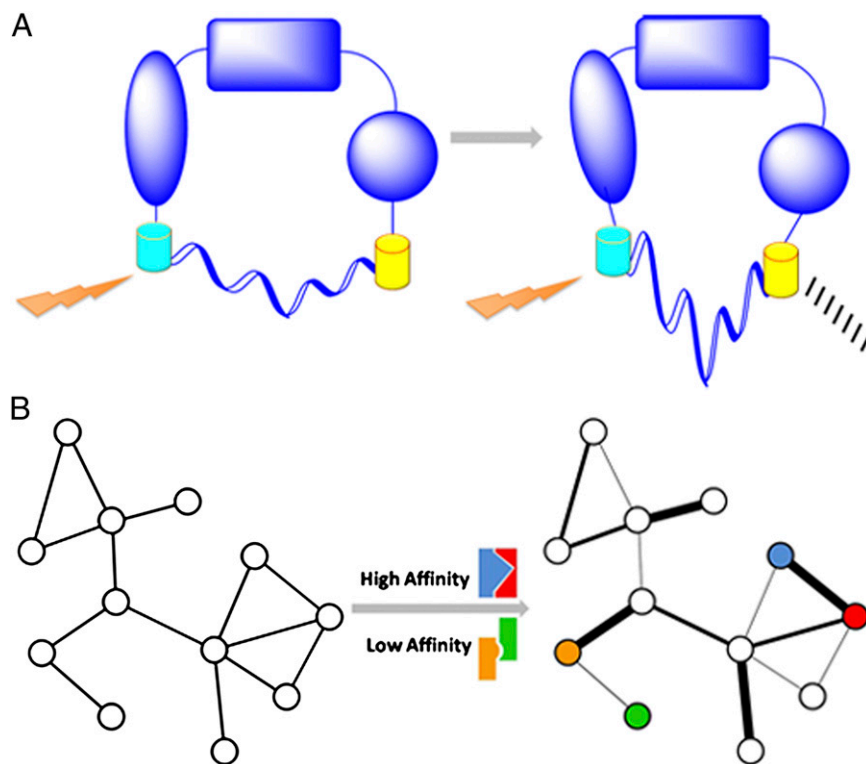

Fig. 1. Applying the ER/K α-helix to studying conformational changes and network characterization. (*A*) A protein in two possible conformations. Detection of a FRET signal may be dependent on a change in conformation. The semicircle below each conformation represents the linker. (*B*) Two representations of a protein–protein interaction network. *Left*: This network is a conventional node-and-edge depiction. Quantifying the interaction affinities may enable the addition of interaction strength as a property of the network, yielding a network with weighted edges (*Right*). Colored interaction schematics above and below the central arrow correspond to examples of mapped high- and low-affinity interactions between similarly colored nodes in the affinity-based network.

binding affinities associated with interactions. Thus, this application has far-reaching implications for the large and diverse array proteins for which interactions with other cellular species are intrinsic to their basic functionality.

There are several advantages to this linker system. The linkage motif itself has been adopted from naturally occurring proteins, thereby precluding the need of designing the construct ab initio. Second, these biosensor constructs are designed with (Gly-Ser-Gly)4 bridges separating the interacting domains and FRET species, thereby allowing the elements of the construct to fully explore rotational degrees of freedom. The unique linker system has built-in modularity as well; protease sites separate its various domains, thereby lending the system separation and purifi-

cation of its constituents, which are thus made available in stoichiometric amounts for bimolecular interactions and affinity calculations.

The interactions between proteins are inevitably tied to the motions and dynamics of the interaction participants, and a protein's conformation is an important factor in determining its interaction affinity. The linker system provides several avenues that may be of interest to the study of conformational changes. When

fused to two distinct domains of a protein, one may be able to gain a crude knowledge of potential conformational changes in response to certain stimuli (Fig. 1*A*). The existence of a FRET signal only in the presence of a given stimulus would suggest that the protein undergoes a conformational change in a way that confers the needed proximity between the donor and acceptor FRET molecules. Of course, the absence of a FRET signal may, in fact, constitute a false negative, as much of the conformational space may not allow a favorable interaction between the two FRET molecules. In addition, the preparation of such a construct would not be the exact same as that described by Sivaramakrishnan and Spudich.

Additionally, the linker may be applied to investigate the general freedom with which specific proteins explore conformational space. The authors note that this construct may be used to study the interaction affinities between distinct domains within a given protein. Careful analyses of intradomain affinities may provide valuable insights into the preferred conformations adopted by simple proteins. If two domains of a simple protein exhibit a high mutual affinity, the frequency of large-scale intramolecular motions may be low, with the closed state lying deep within an energy well. Conversely, two domains exhibiting a low mutual affinity may be one way to identify systems favoring more open configurations. Naturally, such interpretations would first entail knowledge of the protein's structure.

Questions regarding intramolecular motions and protein–protein binding affinities are relevant in the larger context of protein interaction networks (11, 12). The network representation has become an invaluable means of delineating inter-

actions within cellular systems, and this more global view of interactions is generally provided in the form of nodes, which represent biomolecules, and edges, which represent physical interactions between those biomolecules (Fig. 1*B*, *Left*).

Sivaramakrishnan and Spudich introduce a system for interrogating interactions between pairs of proteins, domains, and peptides.

Although valuable, these simple 2D renderings alone fail to relay many important biological phenomena. Thus, more recent efforts have been aimed at incorporating additional forms of data into simple node-and-edge pictures, including structural data of the constituent biomolecules (13, 14), evolutionary features (15), and expression patterns (16).

Proteins' varying conformations also constitute essential attributes in understanding networks, and how the magnitude and general nature of conformational changes influence interactions is an area of active research. Recent work has highlighted the relevance of conformational flexibility. For instance, proteins with many interaction interfaces and binding partners have been found to undergo larger conformational changes (17). In this regard, should the helical linker be implemented on a large scale, its potential ability to elucidate conformational changes may equip investigators with easier means of further studying the

role of conformational flexibility in large-scale networks.

Currently, most network edges are represented as simple binary entities, wherein the presence or absence of an edge denotes an interacting or noninteracting pair, respectively, without including information on the strength of the interaction itself (18). Notably, an additional potential benefit of recruiting this linker for network studies may be characterizing the confidence and strength with which two proteins interact. For example, it may be possible to weigh interactions on the basis of experimentally determined affinities (Fig. 1*B*), which may ultimately lead to new insights into network topology and organization. Such insights may include, for instance, the identification of highly stable or highly dynamic modules within the network. It should not be neglected that a limitation intrinsic to this line of thinking is the binary nature of the linker system itself. Many interactions within the network occur between many proteins simultaneously (16), and simultaneous interactions with a given protein are likely to influence that protein's properties in ways that may not surface in calculations on simple binary interactions.

Despite this, the genetically encoded construct to modulate interactions may very well have promise at the level of entire networks, even if only to assign varying degrees of confidence to network edges. One of the major future challenges to more fully realizing this promise will be deploying this type of system to investigate large combinatorial numbers of potential protein pairs. Although development of such high-throughput systems will constitute challenges, such a scaling up would bring huge rewards and greater dimension to network-based analyses.

1. Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46.
2. Park PJ (2009) ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet* 10:669–680.
3. Sivaramakrishnan S, Spudich JA (2011) Systematic control of protein interaction using a modular ER/K α-helix linker. *Proc Natl Acad Sci USA* 108:20467–20472.
4. Albert R (2005) Scale-free networks in cell biology. *J Cell Sci* 118:4947–4957.
5. Jares-Erijman EA, Jovin TM (2003) FRET imaging. *Nat Biotechnol* 21:1387–1395.
6. Aoki K, Kiyokawa E, Nakamura T, Matsuda M (2008) Visualization of growth signal transduction cascades in living cells with genetically encoded probes based on Förster resonance energy transfer. *Philos Trans R Soc Lond B Biol Sci* 363:2143–2151.
7. Aye-Han NN, Ni Q, Zhang J (2009) Fluorescent biosensors for real-time tracking of post-translational modification dynamics. *Curr Opin Chem Biol* 13:392–397.

8. Zhang J, Allen MD (2007) FRET-based biosensors for protein kinases: Illuminating the kinome. *Mol Biosyst* 3:759–765.
9. Gao X, Zhang J (2010) FRET-based activity biosensors to probe compartmentalized signaling. *ChemBioChem* 11:147–151.
10. Lu S, Wang Y (2010) Fluorescence resonance energy transfer biosensors for cancer detection and evaluation of drug efficacy. *Clin Cancer Res* 16:3822–3824.
11. Gursoy A, Keskin O, Nussinov R (2008) Topological properties of protein interaction networks from a structural perspective. *Biochem Soc Trans* 36:1398–1403.
12. Kiel C, Beltrao P, Serrano L (2008) Analyzing protein interaction networks using structural information. *Annu Rev Biochem* 77:415–441.
13. Aloy P, Russell RB (2002) Interrogating protein interaction networks through structural biology. *Proc Natl Acad Sci USA* 99:5896–5901.

14. Aloy P, Russell RB (2002) The third dimension for protein interactions and complexes. *Trends Biochem Sci* 27:633–638.
15. Kim PM, Korbel JO, Gerstein MB (2007) Positive selection at the protein network periphery: Evaluation in terms of structural constraints and cellular context. *Proc Natl Acad Sci USA* 104:20274–20279.
16. Han JD, et al. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430:88–93.
17. Bhardwaj N, Abyzov A, Clarke D, Shou C, Gerstein MB (2011) Integration of protein motions with molecular networks reveals different mechanisms for permanent and transient interactions. *Protein Sci* 20:1745–1754.
18. Kim PM, Lu LJ, Xia Y, Gerstein MB (2006) Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 314:1938–1941.