FEBS Letters 586 (2012) 2740–2750

journal homepage: www.FEBSLetters.org

Review

Beyond ‘furballs’ and ‘dumpling soups’ – towards a molecular architecture of signaling complexes and networks

Marc Lewitzky, Philip C. Simister, Stephan M. Feller *

Biological Systems Architecture Group, Weatherall Institute of Molecular Medicine, Department of Oncology, University of Oxford, Oxford OX3 9DS, United Kingdom

ARTICLE INFO

Article history:
Received 24 March 2012
Accepted 16 April 2012
Available online 25 April 2012

Edited by Marius Sudol, Gianni Cesareni, Giulio Superti-Furga and Wilhelm Just

Keywords:
Molecular architecture
Signaling complex
Signalosome
Intrinsic disorder
Macromolecular crowding
Signaling network

ABSTRACT

The molecular architectures of intracellular signaling networks are largely unknown. Understanding their design principles and mechanisms of processing information is essential to grasp the molecular basis of virtually all biological processes. This is particularly challenging for human pathologies like cancers, as essentially each tumor is a unique disease with vastly deranged signaling networks. However, even in normal cells we know almost nothing. A few ‘signalosomes’, like the COP9 and the TCR signaling complexes have been described, but detailed structural information on their architectures is largely lacking. Similarly, many growth factor receptors, for example EGF receptor, insulin receptor and c-Met, signal via huge protein complexes built on large platform proteins (Gab, Irs/Dok, p130Cas[BCAR1], Frs families etc.), which are structurally not well understood. Subsequent higher order processing events remain even more enigmatic. We discuss here methods that can be employed to study signaling architectures, and the importance of too often neglected features like macromolecular crowding, intrinsic disorder in proteins and the sophisticated cellular infrastructures, which need to be carefully considered in order to develop a more mature understanding of cellular signal processing.

Abbreviations: Abl, Abelson tyrosine-protein kinase 1; ADAP1, arf-GAP with dual PH domain-containing protein 1; AFM, atomic force microscopy; cAMP, cyclic adenosine monophosphate; ApoER2, apolipoprotein E receptor 2; BCAR1, breast cancer anti-estrogen resistance 1; p130Cas, Crk-associated substrate; CASK, calcium/calmodulin-dependent serine protein kinase; Cdc4, cell division control protein 4; CDK, cyclin-dependent kinase; Cdx2, caudal type homeobox 2; COP9, constitutive photomorphogenesis protein 9; Crk, CT-10 related kinase; CRKL, Crk-like protein; cryo-EM, cryo-electron microscopy; DISC1, disrupted in schizophrenia 1; DLK, DAP-like kinase; Dok, docking protein 1; EGF, epidermal growth factor; ELM, eukaryotic linear motif; ErB2, estrogen receptor beta 2; Erk, extracellular signal-regulated kinase; Frs, fibroblast growth factor receptor substrate; Gab, Grb2-associated-binding protein; Grb2, growth factor receptor-bound protein 2; GRIP1, glutamate receptor interacting protein 1; IGFR, insulin-like growth factor 1 receptor; IR, insulin receptor; Irs, insulin receptor substrate; JIP1, c-Jun-amino-terminal kinase-interacting protein 1; LAT, linker for activation of T-cells; LMD, large multi-site docking; MALS, mammalian lin-seven protein; MEK, MAPK/Erk kinase; c-Met, hepatocyte growth-factor receptor; Mint1, neuronal munc18-1 binding protein 1; NFN hypothesis, N-terminal folding nucleation hypothesis; NMR, nuclear magnetic resonance; NR2B, *N*-methyl D-aspartate receptor subtype 2B; Pbs2, Polymyxin B resistance protein 2; PH domain, pleckstrin homology domain; PIP3, phosphatidylinositol (3,4,5)-triphosphate; PI3K, PI3 kinase; PTB domain, phosphotyrosine-binding domain; Rab, Rev/Rex activation domain-binding protein; Rack1, receptor of activated protein kinase C 1; San1, sir antagonist 1; SAXS, small angle X-ray scattering; SH2 domain, Src-homology 2 domain; SH3 domain, Src-homology 3 domain; SHP2, SH2 containing protein tyrosine phosphatase 2; Sic1, substrate/subunit/inhibitor of cyclin-dependent protein kinase 1; SLiM, short linear motif; Src, sarcoma kinase; STAT1, signal transducer and activator of transcription 1; Ste5, sterile 5 protein; Ste11, sterile 11 protein; TCR, T-cell receptor; UTR, untranslated region; ZBP1, zip code-binding protein 1

* Corresponding author.
E-mail addresses: marc.lewitzky@imm.ox.ac.uk (M. Lewitzky), philip.simister@imm.ox.ac.uk (P.C. Simister), stephan.feller@imm.ox.ac.uk (S.M. Feller).

Whether you can observe a thing or not depends on the theory which you use.

Albert Einstein (1879–1955)

1. Introduction

Representations of signaling networks in many textbooks and even in some of the most recent review articles have in their simplicity a striking similarity to children’s drawings. This is no accident: our level of sophistication when it comes to understanding cell signaling network architectures is rather limited.

Two common types of current signaling data display methods are the ‘furball’ for protein–protein interaction data and the ‘dumpling soup’ for signaling pathways and networks (Fig. 1). These schematic representations do not take into account the many layers of complexity that influence a myriad of signaling processes, which will finally funnel into seemingly simple cell decisions about their survival, growth, proliferation, migration or differentiation.

There are many reasons for the current lack of better models. On a technical level, methods that are at our disposal may commonly not be adequate for the task. Signaling molecules are often low in abundance [1] and their interactions can be highly transient [2]. This makes the endogenously expressed protein molecules hard to study with available techniques. Over-expression studies,

0014-5793
http://dx.doi.org/10.1016/j.febslet.2012.04.029

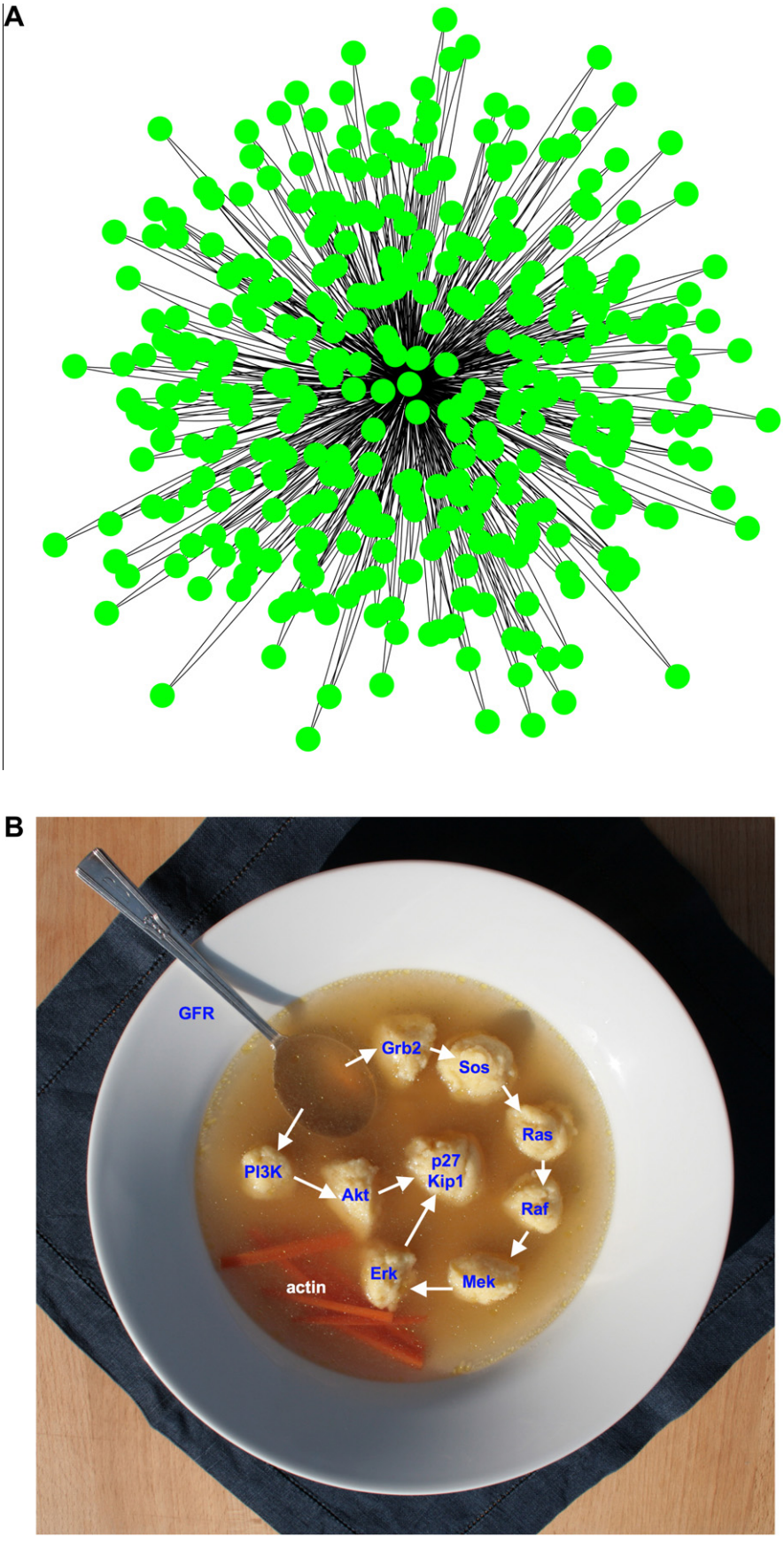

Fig. 1. (A) Example of a signaling protein 'furball' representing the numerous physical interactions of the p53 tumor suppressor protein (based on information from the BioGRID database [171] displayed in Cytoscape 2.8.2 [172]. (B) Growth factor receptor-initiated signaling depicted as a 'dumpling soup'.

a widely used tool to solve the problem of low abundance, can produce profound artifacts in terms of aberrant protein localization, as well as spill-over effects in other parts of cellular signaling networks. Epitope tags that are frequently used to mark exogenously expressed proteins can also fundamentally change the molecular behaviors of proteins.

We also simply do not understand yet many of the relevant structures of signaling architectures and we lack knowledge about some complex variables that govern signaling networks [3]. To overcome these limitations, new concepts need to be put forward and must then be rigorously scrutinized with sufficiently sophisticated experimentation.

Protein functions are frequently controlled through their associations with other proteins, which influence their activities and localizations. Groups of proteins often form functional assemblies or even large (megadalton-sized) 'signalosomes' in which the individual protein components fulfill highly specific tasks. The term 'signalosome' was possibly first used in the context of the COP9 multi-protein complex, a widely conserved and multi-functional cellular 'machine' that was initially discovered in plants, where it is regulated by light signals [4–9]. More recently other signalosomes have been proposed including the T cell receptor signalosome [10–12], the Wnt signalosome [13] and the adhesion signalosome [14,15]. Many more can be expected to emerge in the next years. Signalosomes are not at all tightly fixed in their composition and hence their architectures, and they can vary in size considerably depending on their signaling state [13].

Owing to their considerable dimensions and the 'zip codes' encoded in their components, signalosomes are expected to frequently inhabit distinct subcellular locations, sometimes also called 'microdomains', which further define their actions and functions. The same proteins, or different splice variants thereof, can therefore appear in several distinct sub-cellular compartments and engage in very different types of activities [16,17]. Many aspects of this cellular organization are still unknown. In recent years we have learned a lot about direct protein–protein interactions but relatively little about how these translate into higher-order protein complexes or sub-cellular compartmentalization. To see a more complete picture it is important to understand how this higher order of organization can fine-tune the functions of individual proteins.

We argue here that it will be necessary to thoroughly examine signaling protein functions in the context of the architectures of the multi-protein complexes in which they are incorporated and also the infrastructures provided by the cellular environment. We will touch on some of the methods that may help to further elucidate the architectures of such complexes and their networks.

2. Intracellular signal processing largely depends on protein complexes

Small secondary messengers, like phosphatidylinositols, cAMP, Ca^{2+} or NO, cannot carry a large amount of information and have to rely largely on random diffusion to reach their intended destination. They can transfer broad messages, often at a high speed. However, complex signals and signal processing are typically left to proteins.

A human cell will at any given time express a large contingent of the 22 300 or more proteins that are encoded in the human genome [18] and several thousand of these are signaling proteins. Humans have around 800 G protein-coupled receptors, over 500 protein kinases and an as yet uncounted but vast number of proteins that form specific complexes in response to post-translational protein modifications like phosphorylation, acetylation and ubiquitinylation [19–25], with many of them being involved in signaling processes. According to the RESID database, there exist more than 500 types of modification (www.ebi.ac.uk/RESID/). To match some of these, numerous 'flavors' of specialized 'reader' domains that recognize the modifications have evolved ([26]; see also EMBL database SMART). The founding family of these domains are the SH2 domains, which boast over 120 human members and are essential for the recognition of specific protein surface epitopes that contain phosphotyrosine residues ([26–28]; see also SH2 binding sites in Fig. 2). Unlike small secondary messengers, some proteins, especially signaling proteins, are often expressed in low numbers per cell [1] and their movements are substantially restricted by the overwhelming presence of other large biomolecules.

3. The impact of macromolecular crowding on signaling processes

Cells comprise extremely densely packed environments. The intracellular concentration of proteins has been estimated at around 200–300 mg/ml for mammalian cells [29] with macromolecules occupying up to 40% of the total cell volume [30]. The high concentrations of macromolecules are not limited to the cell interior but occur also in the extracellular matrix [31].

It is not the high concentration of a single species of macromolecule but the multitude of different macromolecules in a small volume that, through steric repulsion, hinders molecular diffusion and can have a profound effect on protein function and association. Hence, molecular crowding [32,33] is a key factor that shapes the molecular arrangement of cellular signaling networks.

Macromolecular crowding has several consequences for protein movement and function: it can strongly reduce the distance a protein travels (up to 100 times [31]). It can also promote the association of proteins. Equilibrium constants for macromolecular associations can be increased by several orders of magnitude [31,33]. This may help to explain why even apparently low affinity interactions, measured in aqueous solutions, may be sufficient for some biologically important protein–protein interactions [2,34]. It is one of the major challenges in the investigation of protein–protein interactions to distinguish between these functional and non-functional weak interactions.

Macromolecular crowding can also directly affect protein structure and function. In a recent example, Dhar et al. demonstrate major conformational changes of phosphoglycerate kinase and an over 15-fold increase in its viscosity-adjusted enzymatic activity under crowded conditions [35]. Two further reports suggest a role for molecular crowding in the kinetics of Erk phosphorylation [36,37]. Crowding forces a rapid re-association of MAPK/Erk kinase (MEK) and its Erk substrate in a process that mimics processive phosphorylation [38,39].

The effects of crowding are rarely considered in biochemical experiments and concentrations of macromolecules in biochemical assays often do not exceed 10 mg/ml [40]. As there is no simple way to mimic or manipulate intracellular crowding conditions, crowding effects are studied by the addition of reagents that simulate steric but not other intra-cellular interactions in in vitro experiments [40,41]. The lack of more advanced methods to assess crowding in an in vivo setting has so far made it difficult to study its effects on macromolecular assemblies in a systematic way.

4. Cellular infrastructure: localized protein expression and directional transport

Despite being restricted by macromolecular crowding, diffusion clearly plays a significant part in protein-mediated signaling. Morphogens that define cell polarity in the *Drosophila* egg are a

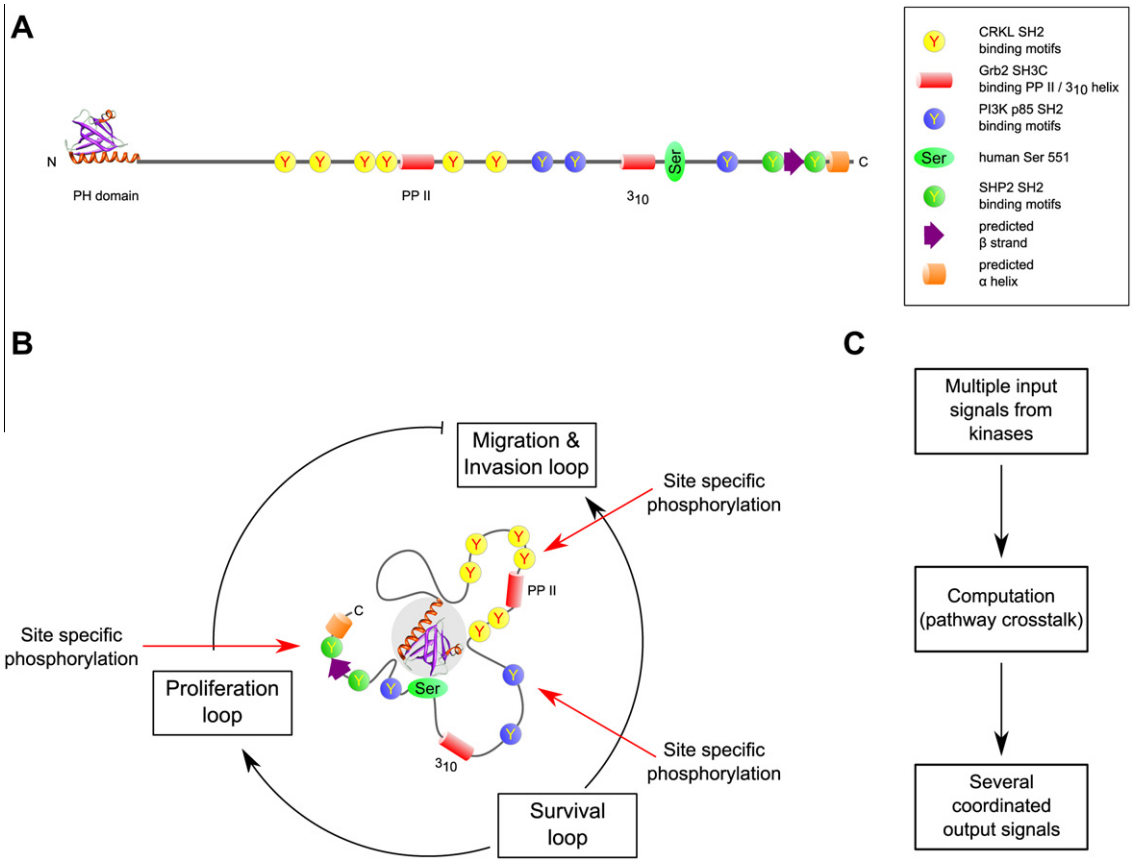

Fig. 2. Human Gab1 protein in the traditional stick representation and as a graphic display of its presumed signal computation architecture according to the N-terminal folding nucleation (NFN) hypothesis. (A) 'Classical' stick representation of human Gab1. The N-terminal phospholipid (PIP3)-binding pleckstrin homology (PH) domain is drawn as a colored ribbon structure. The ribbon representation of PDB 2X18 used here was created in UCSF Chimera 1.6 [173]. Known target sites for Gab1 phosphorylation by protein kinases (Y, Ser) are also shown. Tyrosine (Y) residues are color coded according to their presumed interactions with different SH2 domain containing proteins (see also boxed explanations of symbols). These follow site-specific Y phosphorylations by kinases like c-Met, Src and Abl. Two short helical structural elements identified through crystallographic studies are known binding sites for the C-terminal SH3 domain of the adaptor protein Grb2 [163] and couple Gab1 to various receptor proteins [104]. Serine 551 in humans (corresponding to S552 in mouse) is the target residue for phosphorylation by Erk1/2. Serine 551 phosphorylation has been proposed to disrupt the interaction of this epitope with the Gab1 PH domain, which opens up the Gab1 PH domain PIP3 binding site and allows Gab1 translocation to the membrane [162]. Predicted secondary structure elements at the C-terminus of Gab1 [161] are shown as a purple arrow and an orange cylinder. (B) Schematic representation of the proposed Gab1 signal computation architecture. Gab family proteins and their relatives (the p130Cas[BCAR1], Irs/Dok and Frs families) are large multi-site docking (LMD) proteins that are intensely studied and known to serve as assembly platforms for multi-protein complexes (signalosomes) which process signals from a wide range of transmembrane receptors as well as cytoplasmic kinases [100–104]. They are expected to function as molecular computation machines that integrate information from multiple 'input' kinases and subsequently produce coordinated output signals that steer biological cell actions. These cell actions, for example cell survival, proliferation, migration or invasion, need to be well orchestrated to avoid 'conflicts of interest', e.g. between the assembly of the mitotic spindle during cell division and cytoskeletal changes during migratory processes. The LMD protein kinases (e.g. EGFR, ErbB2, IGFR, c-Met, Src, Abl, Erk etc.) provide information about a large number of environmental and intrinsic conditions by phosphorylating, or not, one or more defined epitopes of the LMD protein tails, thereby determining whether a cell is to die, to live, to divide or to migrate. Most of the known phosphorylation sites (see www.phosphosite.org for more information) are found in proposed loop regions of the protein tails. These tail loops are suspected to form co-translationally during the emergence of the nascent protein chains from the ribosomes [143]. The N-terminal regions of the LMD proteins mentioned above contain one or two well-folded domains (PH, SH3, PTB), which form spontaneously and then could serve as 'nucleation cores' for the emerging protein tails that are several hundred amino acids in length. The first in vitro and in vivo evidence for interactions between the Gab1 PH domain and the Gab1 tail has been reported [143,162]. The architectural model shown here could explain how LMD protein tails, which were previously thought to be unstructured, may serve as organized but still dynamic platforms for an effective signal computation in cells. Although this model goes beyond traditional 'dumpling' or 'stick' representations of disordered proteins, it is still an oversimplification and would need further refinement to accommodate all existing and newly emerging data. Note that the actual tail-interacting residues on the surface of the Gab1 PH domain have not yet been defined. (C) Simple functional diagram of signal processing via the Gab1 LMD protein platform.

prominent example of protein gradients that rely heavily on protein diffusion [42,43]. Other pathways may depend on some component that diffuses freely, like STAT1 in the interferon-γ response pathway [44]. However, large signaling complexes contain a huge number of individual components which may not be very highly expressed. Random diffusion of these components, restricted by macromolecular crowding and cellular compartmentalization, would be a very unreliable and error-prone process for their assembly. It is therefore very likely that in many cases the formation of protein complexes is actively supported by the cellular transport infrastructure.

Even in prokaryotic cells diffusion may be only one way of protein distribution. Prokaryotes, sometimes wrongly considered to be little more than semi-permeable protein 'bags', actually show a high degree of spatial organization. Many of their macromolecules accumulate at specific locations where they fulfill their biochemical and physiological functions [45].

The spatial regulation of mRNA translation is one of the mechanisms that can coordinate protein localization in cells [46,47]. Site-specific and hence asymmetric mRNA localization was first observed for β-actin mRNA in ascidian embryos [48]. Recently targeting of mRNA was also shown in *Escherichia coli*, challenging the view of coupled transcription and translation in prokaryotes [49]. In mammals, mRNA localization has been observed and extensively researched in nerve cells with their often very long extensions [50]. Beyond that, it was also found to be an important

factor in ontogeny. Heterogenous distribution of Cdx2 mRNA establishes cell polarity in the early mouse embryo [51]. This is, according to very recent results, mediated by a 97 bp region preceding the Cdx2 mRNA 3'UTR (3′ untranslated region), which is responsible for the mRNA localization (Magdalena Zernicka-Goetz, personal communication).

This method of spatial regulation of protein expression may be very common. More than 70% of mRNAs of 3370 genes analyzed were asymmetrically expressed in *Drosophila* [52]. The first systematic investigation of mRNA localization in mammalian cells found that more than 50 species of mRNA preferentially accumulate in the cellular protrusions of migrating mouse NIH3T3 fibroblasts [53].

Site-specific mRNA localization can be achieved by local degradation or stabilization, by anchoring at the translation site or by active transport [54]. Active transport relies on localization signals, zip codes, in the mRNA sequence. These zip codes are recognized by specific binding proteins that regulate active transport along the cytoskeleton [55,56]. mRNAs are transported along microtubules and actin filaments, often in granules containing ribosomal proteins. In neurons their speed has been measured at 0.1 μm/s [57]. At the endpoint of their journey, translation can be initiated by local signaling events [58]. This has been demonstrated for the β-actin mRNA in zip code-binding protein 1 (ZBP1)-containing granules where phosphorylation of ZBP1 by the tyrosine kinase Src at the cell periphery induces actin mRNA translation [59].

Active localization of mRNA may be one way to allocate components of signaling complexes to a specific location in cell. However, sharing the same system of kinesin super-family protein motors running on cytoskeletal tracks, protein complexes, vesicles or even complete organelles like mitochondria can be actively transported [60]. In fact, most newly synthesized proteins are actively transported to their appropriate destinations [61]. Although kinesin molecules can bind their cargo directly, they typically interact with their cargo through adapter or scaffold proteins like ADAP1 [62], DISC1 [63], GRIP1 [64], JIP1 [65] and Mint1 [66]. Unloading is controlled here too by local signaling events and can be triggered by phosphorylation, Rab GTPase activity and Ca^{2+} signaling [67]. Scaffold proteins recognized by kinesin modules are often also involved in signaling pathways, suggesting that kinesin motors can transport pre-assembled signaling modules to specific locations [68]. This is no more clearly demonstrated than for the JNK-interacting protein 1 (JIP1) scaffold which is transported preloaded with the upstream kinase DLK and the trans-membrane receptor ApoER2 [69], or for Mint1 which is transported in a complex with its up- and down-stream interaction partners CASK, MALS and a NR2B receptor subunit [66].

5. Cell membrane properties and microdomains contribute to organizing signaling complex and network architectures

Cell membranes, another feature of the cellular infrastructure, fulfill critical roles in cellular signaling. They provide a surface for macromolecules to assemble on, and barriers to compartmentalize certain elements away from others. The movement of proteins within membranes is more restricted than previously thought as partitioning of membrane segments prevents free diffusion [70]. Many subdomains associated with signaling structures have been identified. They can contain protein scaffolds or specific lipid compositions [71].

Caveolae, a fairly well-researched class of membrane domains, are small invaginations of 60–80 nm diameter that are formed with the help of the caveolin family of integral membrane scaffold proteins [72]. These structures are thought to be a platform for signaling complexes of multiple membrane-associated signaling molecules that share binding to the caveolin proteins [72–74]. The complex role of cellular membranes in signaling is also well illustrated in the case of the T-cell receptor (TCR) signalosome [10,11]. Before TCR activation, the TCR subunits and a fraction of the membrane-associated scaffold protein LAT are separated into different clusters in the plasma membrane [75]. Another fraction of LAT is simultaneously stored in small vesicles, which can interact with components at the membrane [16]. TCR activation then leads to the large-scale reorganization at the membrane, in part through the involvement of the cytoskeleton [76].

New structural features in cells are still being discovered and we are only beginning to understand functional units like nanotubes, purinosomes (multi-enzyme complexes for purine biosynthesis), micro-compartments and the 'cellular serpents' of filament-forming enzymes [77] in terms of their impact on establishing the architectures of signaling complexes and networks. Moreover, at high concentrations, multivalent signaling protein interactions can, at least in some cases, lead to 'phase transitions' in vitro and apparently also in living cells to generate droplet-like structures [78]. This may contribute to the regulation of protein activities and can possibly complement other architecture-generating processes in cells.

6. Origins of cellular complexity and signaling architectures

The evolution of eukaryotic cells and organisms led to a great increase in structural and signaling complexity. A major contributing factor to this increase in complexity most likely has been whole genome duplications, which led to a higher number of protein families and family members, supplied more genetic material for the genesis of new functions and promoted accelerated evolution in duplicated genes [79]. In particular, cells in organisms composed of several or multiple tissues acquired many unique features [80] and increased robustness and redundancy in old signaling pathways [81]. The transition to multi-cellularity and organismal tissue diversity, favored in an environment where the smaller organisms get consumed more easily [82], was a further likely contributor to cell complexity [83]. It also led to an explosion of modular protein domains [84]. In an excellent review, Bhattacharyya et al. discuss how the modularity of proteins could have promoted evolutionary innovation, in particular in signal transduction proteins. Modularity permitted innovation by the recombination of existing elements and the establishment of novel connectivities between existing or duplicated proteins. As catalytic function and connectivity were separated, these could be recombined more easily. Moreover, the division of catalytic function and connectivity into separate proteins allowed the reuse of catalytic functionalities in different pathways. The in vivo specificities of the catalytic domains can be conferred by the combination of catalytic function with connectivity and regulatory elements on specific assembly platforms [84]. This model of separation between catalytic function and connectivity can, for example, help to understand why we observe so much promiscuity in the motifs for catalytic enzymes like kinases [85], especially when these are studied outside of their natural contexts.

It is evident that the size of signal transduction protein complexes also increased substantially as eukaryotes became much more complex. This was again partially enabled by the modularity of proteins, but also required a substantial flexibility of parts of proteins' amino acid chains to easily accommodate architectural additions and modifications. Especially in the case of signaling proteins with many transient interactions, structural 'disorder' in substantial segments of the proteins frequently became critical, as will be discussed in more detail below.

7. Hubs, adapters and scaffolds – docking points and assembly platforms for multi-molecular signal computation

Assembly platforms for signaling proteins can come in many flavors. 'Hubs' [86] are highly interconnected within the signaling network. Cytoplasmic 'adapters', 'scaffolds' and 'large multisite docking (LMD) proteins' co-localize two or more elements of a signaling pathway or enable complex signal processing. They can also appear as membrane-associated structures or may even be combined with catalytic domains, as in the case of the transmembrane receptor kinases [84,87]. The TCR component LAT and the EGF receptor are representatives and well-studied examples of two major types of membrane-associated signal complex assembly sites (see [88] and [89] for details on their architectures and interactions).

The tumor suppressor p53 is a prototypical example of a complex hub protein ([90,91]; architectural features are discussed in [92]). After decades of research into p53 new functions and involvement in different signaling processes are still being discovered. One of its important roles is to act as a transcriptional regulator for the expression of a substantial number of genes, which will in turn often provide feedback on p53 functions. p53 can also inhibit cell proliferation and, via many other pathways, contribute to tumor suppression [91]. It has numerous unstructured protein–protein interaction epitopes, some of which can adopt different secondary structures with different binding partners, thus enabling significant binding promiscuity [92–94].

The yeast Ste5 and Pbs2 proteins are prototypical examples of cytoplasmic scaffolds [95,96] and some aspects of their architecture have readily emerged [97]. Ste5 and its binding partners contribute to a pheromone response pathway that induces a mating response. Pbs2 and its binding partners contribute to an osmolarity response pathway. A common binding partner of both proteins is the Ste11 kinase that can phosphorylate different downstream effectors. These scaffold proteins, by bringing together the Ste11 kinase and different downstream targets, can add specificity to Ste11 function and direct further signaling [98]. For instance, an engineered fusion-protein of Ste5 and Pbs2 allows redirection of a pheromone signal to an osmolarity response [99].

Scaffold proteins are relatively rigid structures that coordinate direct interactions between elements of pathways in a straightforward, linear manner. By contrast, large multi-site docking (LMD) proteins like the members of the p130Cas, Gab, Irs/Dok and Frs families [100–104] allow for a much more integrative processing of signaling information from different pathways. They interact in a highly dynamic manner with a larger number of binding partners to integrate information from multiple upstream kinases and pathways, and they subsequently generate signals that regulate central biological processes like cell survival, cell proliferation, cell shape changes and migratory actions.

For example, the multifunctional LMD protein p130Cas (BCAR1) is phosphorylated as a consequence of signaling by integrins and several G protein-coupled receptors and recruits a variety of adapters and further downstream effectors, allowing it to contribute to cell survival, as well as cell shape change and migration pathways [101,105]. However, it can also act as an important component of a pro-apoptotic pathway where it contributes to cell death induction after its cleavage by proteases [106]. Beyond this, it was proposed that p130Cas can also act as a primary force sensor for cell stretching at cell-matrix contact sites [107] resulting in its increased phosphorylation, and thereby supporting cellular reorientation after mechanical stimulation [108].

The significance of LMD proteins for the integration of signals sent from distinct receptors and the resulting generation of very specific regulatory signals was highlighted again very recently by the work of Fafalios et al. (2012). They showed that only the combination of insulin receptor (IR) and c-Met receptor signals, via the same LMD protein, IRS2 (or IRS1 to a lesser degree), leads to a robust regulation of glucose levels in liver cells [109]. This clearly demonstrates that LMD proteins are not merely mediators of single-receptor signals but that they serve as computational units for a higher level of signal processing, eventually leading to enormously complex transcriptional responses, which together determine cell fate and actions.

A structurally very distinct and presumably somewhat more rigid intracellular 'meeting point' that adopts a much clearer defined, seven-bladed beta-propeller fold is the multi-functional protein Rack1 [110]. It is indeed extremely versatile and can, for example, serve in different contexts either as a shuttle protein, as an anchor protein or it can impact on the activities of its binding partners. One of its many roles seems to be to act as a bridging protein between insulin family growth factor receptors, like IGFR (IGF1R) and IR, and their effector proteins, including the Irs proteins, thereby aiding in the assembly of the architectures of signal transduction computation machines. Unfortunately most of the ultrastructural details of how Rack1 interacts with the cellular signaling machinery still remain to be defined.

8. Some methods for investigating the compositions and architectures of signaling complexes and networks

Mass spectrometry has become an important method to investigate post-translational modifications and protein–protein interactions in detail and on a global, proteomic scale [111,112]. It is highly sensitive and can thus be used to investigate many lowly abundant signaling proteins, provided that highly abundant cell proteins are sufficiently separated or depleted.

Mass spectrometry has also recently helped to unravel the fundamental relationship between mRNA expression and protein expression [113]. Moreover, it can enable researchers to simultaneously observe thousands of protein isoforms and numerous modifications in a single sample [114].

In the structural analysis of protein complexes, mass spectrometry has equally made great advances over the last years. It can now be applied to investigate the structure and dynamics of some protein assemblies [115,116]. This allows the determination of complex stoichiometries, but it is also useful to define the subunits in bigger complexes [117,118] and can even be utilized to study the kinetics and thermodynamics of complex assembly [118,119].

Mass spectrometry becomes most powerful for structural analyses when combined with other methods [120–122] as no experimental technique can decipher the full spectrum of spatial and temporal information that is necessary for a complete characterization of macromolecular assemblies [123].

Cryo-electron microscopy (cryo-EM), that is the observation of snap-frozen samples under a transmission electron microscope, is a particularly useful method to move into the realm of directly observable structures as it allows observation at near-atomic resolution. Resolutions in the range of 1 nm have been achieved for larger, symmetrical particles [124] and, at least in theory, even much higher resolution should be possible [125]. Cryo-electron tomography is a modification of cryo-EM where whole cells or sections are snap-frozen and imaged iteratively while being tilted incrementally. This permits studies of molecular assemblies in situ in 3D [126–128].

Hybrid approaches using mass spectrometry and cryo-EM have been successfully applied in determining the architecture of the 26S proteasome holocomplex [129], elucidating the intricacies of protein translation [130].

New technological developments in X-ray crystallography, such as advances in the application of hard X-ray free-electron lasers

using femtosecond pulses, may also enable structures of challenging proteins and protein complexes (i.e. those which produce only nanocrystals) to be more easily determined in the future [131]. Additionally, new computational tools have recently become available that help with the integration of data from different experimental approaches (i.e. biophysical methods) into one model [132,133].

The often rather transiently assembled signaling complexes pose additional challenges compared to ribosomes, proteasomes and similar molecular machines. However, through (i) careful stepwise biochemical assembly of signalosomes from highly purified components for their biophysical and functional analysis, (ii) the combined application of currently available methods, and (iii) the (presumed) development of additional high-resolution imaging techniques in the coming years, the emergence of important concepts regarding the architectures and mechanisms of action of multi-protein signal computation machines is expected to come finally into our reach.

9. Disordered proteins and protein segments are crucial for cellular signaling processes

Classical protein research has often focused on functional protein domains with well-folded structures. However, it has been recognized for some years now that surprisingly many proteins appear to have at least in part no apparent defined structure [134–137]. These unstructured regions are often essential for protein function. A large number of algorithms have been developed to predict regions of disorder from the amino acid sequence of a protein, commonly based on the initial finding that low hydrophobicity and high net charge is a defining feature of disordered proteins [138,139]. In mammals, 75% of signaling proteins and 50% of all cellular proteins are predicted to have regions with disorder of more than 30 amino acids in length. 25% of all proteins may even be completely disordered [135]. Not surprisingly, considering the vast number of presumably disordered proteins, some of these have been implicated in cancer development [138] as well as in neurodegenerative and cardiovascular diseases [140]. As disorder may lead to protein aggregation, a key problem in many neurodegenerative diseases, it was expected that disordered proteins may be subject to tight regulation. However, the results were not as clear-cut as previously thought [134]. On one hand, unstructured proteins were reported to be tightly regulated [141], but on the other hand studies found only a weak association of disorder with a shorter protein half-life [134,142]. Stabilization by inter- as well as intra-molecular interactions have been suggested as mechanisms to protect disordered proteins from early degradation [143,144] but other, less obvious features, for example related to their amino acid composition, or whether or not they have a structured C-terminus (M.M. Babu, personal communication), may also contribute.

Disordered regions appear to feature particularly strongly in signaling proteins [145]. They often contain short (or eukaryotic) linear motifs (SLiMs, or ELMs; [146]) that allow docking of protein domains [147,148] and are sites of post-translational modification [149]. Protein disorder may also be a characteristic feature of receptors with recognition and signaling functions distributed between separate protein chains [150]. Functionally important disordered epitopes, albeit sometimes very short, are presumably often conserved throughout evolution and can thus be detected with bioinformatics tools [151]. This discovery of novel putative functional motifs can help to then experimentally define previously unrecognized architectural elements in protein chains and complexes.

Disordered proteins are very diverse in their propensities to assume more defined conformations on their own or upon interaction with other proteins. Some proteins maintain disorder under crowded conditions [152], whereas others may gain structure [153,154]. Some disordered regions can acquire a fixed conformation upon binding [155–157] while others remain disordered [158]. Disorder also allows some very unusual interactions; the yeast ubiquitin ligase San1, for example, recognizes misfolded substrates [159] and the CDK inhibitor Sic1 interacts with a single site on Cdc4 via multiple dispersed phosphorylation sites [158]. The hub proteins already mentioned frequently contain disordered regions [160] that are used to interact with a variety of partners [93].

Research into disordered proteins, although still limited by a lack of easy-to-use methods that allow the speedy recognition of different types of disorder or their functional classification, appears to be currently gaining momentum and we expect that new details on how disorder can be used in various signaling contexts will also lead to a better understanding of signaling complex architectures.

10. Disorder is presumed to be an important prerequisite for efficient signal computation in large multi-site docking proteins

A common feature of large multi-site docking (LMD) proteins of the Gab, Irs, Frs, Dok and p130Cas families is their composition of one or two well-structured N-terminal domains followed by long disordered 'floppy tail' regions [143,161]. How such 'snake-like' appearances can effectively serve to mediate complex signal computations, which are attributed to multi-protein complexes assembled on these 'platform' proteins based on functional studies, has been puzzling us for some years. Upon closer inspection of the human proteome with domain recognition programs about 70 proteins with a similar design have been detected [143].

Recent and somewhat serendipitous findings now indicate that Gab proteins and its 'look-alikes' could actually be much 'better' organized than initially thought. For Gab1 (Fig. 2) the first evidence has emerged that the lipid binding pocket of its N-terminal pleckstrin homology (PH) domain is regulated by a specific epitope distant by over 400 amino acids in its disordered protein tail [162]. Surprisingly, upon closer inspection of this Gab1 PH domain – tail interaction, a number of additional tail regions showed binding to this domain in vitro [143]. Some of these regions overlap with already known protein interaction sites [161,163], but others are novel and beg further studies.

The detection of these interaction sites, which appear to separate functionally clustered SH2 docking motifs, has led to the proposal of a new conceptual framework for the architectures of disordered signal protein complex assembly platforms (i.e. LMD proteins) with folded N-terminal domains: the N-terminal folding nucleation (NFN) hypothesis [143] (Fig. 2). According to this hypothesis, the repeated folding back of a largely disordered protein chain of several hundred amino acids onto a well-structured N-terminal domain would generate several functionally dedicated loops. Such a model provides an intriguing explanation of how Gab family and other similarly composed LMD proteins can effectively assemble multi-protein complexes to compute multiple incoming signals and subsequently coordinate functionally interdependent downstream signals. This occurs simply by assembling pathway-specific sub-complexes in distinct tail loops that hover around the N-terminal fold and by then bringing these loops together to cross-talk [143]. Experimental validation of this architectural hypothesis is urgently required but not entirely trivial.

The conformations of disordered proteins are difficult to study with X-ray crystallography and so NMR has been applied frequently to study their structure in vitro. The development of in-cell NMR in recent years could help in future to better understand their structure in vivo [164]. In-cell NMR requires the labeling of proteins with an NMR-active isotope and their delivery into the cell.

This has been successfully applied in *E. coli* for the de novo determination of complete protein structures [165]. In eukaryotic cells protein delivery is more difficult. Microinjection has been used to deliver the disordered tau protein into *Xenopus* oocytes where it was found to maintain its disorder even under crowded conditions [166]. In other studies researchers used cell-penetrating peptides [167] or a pore complex [168] to deliver labeled proteins into human cells. Until recently, relatively high concentrations of labeled protein were required for such studies, however, newer, more rapid NMR methods may make it now possible to study proteins at more physiological levels [164]. Small angle X-ray scattering (SAXS) [169], the cryo-EM methods already discussed, as well as high-speed atomic force microscopy (AFM) [170] should prove useful for exploring the molecular organization of cellular signaling complex architectures. It will be exciting to see these advanced imaging technologies applied to their full potential on the signaling machinery of cells in the near future.

11. Perspective

We propose that the degrees of architectural and infrastructural sophistication which cells display while accomplishing their signaling tasks (and many other complex tasks) are currently still much under-appreciated and that this concept should guide us to careful, and possibly less invasive experimentation, as well as influence our attempts of data interpretation. In this dawning age of ultra-fast data generation at massive scales we need to consider carefully which concepts will enable us to conduct better experiments with the potentially powerful tools we now have and how we can best obtain meaningful information from ever-growing piles of puzzle pieces. It is time to put them together.

Without a doubt, dealing with such multi-dimensional puzzles (4D or greater: space and time, but also cell type, disease state etc.) of great complexity will also require sophisticated ways of representing the detailed molecular information on several levels of magnification, in order to make it accessible to more general audiences. These methods may not differ too much from the technologies used in '3D' films, and almost certainly will be facilitated by computer graphics and hierarchical visualization systems to represent dynamic molecular behavior and interactions, integrating cell biological, structural and biophysical knowledge. Whether the same information can be adequately translated to the 2D printed page is unclear, given the oversimplification of line drawings. Is there still an appetite for dumpling soup?

Acknowledgements

We are grateful to our colleagues from the Protein Modules Consortium (www.proteinmodules.org) for many valuable discussions, and to Keith Dunker and Richard Kriwacki for stimulating discussions on protein disorder. We also thank Bettina Staub for a delicious dumpling soup. The authors are funded by the cancer charity Heads Up (www.heads-up.org.uk), a grant from the Breast Cancer Campaign, an EU FP7 grant (TARGETBINDER) and the Oxford Cancer Research Centre Development Fund.

Note added in proof

For modeling purposes some cellular processes can be considered to occur in only one or two dimensions to simplify computation and to make the model more easily understandable. However, if understanding the geometry of a process is essential, a 3D or 4D model becomes necessary [174]. Computational approaches that take the geometry of cellular processes into account are becoming increasingly available to biologists [175] and have been successfully applied to model, for example, receptor-mediated cellular adhesion and MAPK activation in yeast [176].

References

[1] Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S. and Mann, M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. 7, 548.
[2] Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G. and Orengo, C. (2010) Transient protein-protein interactions: structural, functional, and network properties. Structure 18, 1233–1243.
[3] Gibson, T.J. (2009) Cell regulation: determined to signal discrete cooperation. Trends Biochem. Sci. 34, 471–482.
[4] Glickman, M.H. et al. (1998) A subcomplex of the proteasome regulatory particle required for Ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell 94, 615–623.
[5] Hannss, R. and Dubiel, W. (2011) COP9 signalosome function in the DDR. FEBS Lett. 585, 2845–2852.
[6] Schweitzer, K. and Naumann, M. (2010) Control of NF-kappaB activation by the COP9 signalosome. Biochem. Soc. Trans. 38, 156–161.
[7] Stratmann, J.W. and Gusmaroli, G. (2012) Many jobs for one good cop – The COP9 signalosome guards development and defense. Plant Sci. 185–186, 50–64.
[8] Wei, N., Chamovitz, D.A. and Deng, X.-W. (1994) Arabidopsis COP9 is a component of a novel signaling complex mediating light control of development. Cell 78, 117–124.
[9] Wei, N., Serino, G. and Deng, X.-W. (2008) The COP9 signalosome: more than a protease. Trends Biochem. Sci. 33, 592–600.
[10] Acuto, O., Bartolo, V.D. and Michel, F. (2008) Tailoring T-cell receptor signals by proximal negative feedback mechanisms. Nat. Rev. Immunol. 8, 699–712.
[11] Dustin, M.L. and Depoil, D. (2011) New insights into the T cell synapse from single molecule techniques. Nat. Rev. Immunol. 11, 672–684.
[12] Krummel, M.F. and Cahalan, M.D. (2010) The immunological synapse: a dynamic platform for local signaling. J. Clin. Immunol. 30, 364–372.
[13] Wang, H.-Y. and Malbon, C.C. (2011) Probing the physical nature and composition of signalsomes. J. Mol. Signal. 6, 1.
[14] Geiger, B. and Yamada, K.M. (2011) Molecular architecture and function of matrix adhesions. Cold Spring Harb. Perspect. Biol. 3, a005033.
[15] Zaidel-Bar, R. and Geiger, B. (2010) The switchable integrin adhesome. J. Cell Sci. 123, 1385–1388.
[16] Purbhoo, M.A. et al. (2010) Dynamics of subsynaptic vesicles and surface microclusters at the immunological synapse. Sci. Signal. 3, ra36.
[17] Rauch, J., Volinsky, N., Romano, D. and Kolch, W. (2011) The secret life of kinases: functions beyond catalysis. Cell Commun. Signal. 9, 23.
[18] Pertea, M. and Salzberg, S.L. (2010) Between a chicken and a grape: estimating the number of human genes. Genome Biol. 11, 206.
[19] Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V. and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325, 834–840.
[20] Deribe, Y.L., Pawson, T. and Dikic, I. (2010) Post-translational modifications in signal integration. Nat. Struct. Mol. Biol. 17, 666–672.
[21] Gnad, F., Gunawardena, J. and Mann, M. (2010) PHOSIDA 2011: the posttranslational modification database. Nucleic Acids Res. 39, D253–D260.
[22] Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P. and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127, 635–648.
[23] Seet, B.T., Dikic, I., Zhou, M.-M. and Pawson, T. (2006) Reading protein modifications with interaction domains. Nat. Rev. Mol. Cell Biol. 7, 473–483.
[24] Stasyk, T. and Huber, L.A. (2012) Mapping in vivo signal transduction defects by phosphoproteomics. Trends Mol. Med. 18, 43–51.
[25] Wagner, S.A., Beli, P., Weinert, B.T., Nielsen, M.L., Cox, J., Mann, M. and Choudhary, C. (2011). A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol. Cell. Proteomics 10, M111.013284.
[26] Pawson, T. and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. Science 300, 445–452.
[27] Huang, H. et al. (2008) Defining the specificity space of the human Src homology 2 domain. Molecular & Cellular Proteomics 7, 768–784.
[28] Liu, B.A., Shah, E., Jablonowski, K., Stergachis, A., Engelmann, B. and Nash, P.D. (2011) The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci. Signal. 4, ra83.
[29] Luby-Phelps, K. (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int. Rev. Cytol. 192, 189–221.
[30] Fulton, A.B. (1982) How crowded is the cytoplasm? Cell 30, 345–347.
[31] Ellis, R.J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment. Curr. Opin. Struct. Biol. 11, 114–119.
[32] Minton, A.P. (1980) Excluded volume as a determinant of protein structure and stability. Biophys. J. 32, 77–79.
[33] Zhou, H.-X., Rivas, G. and Minton, A.P. (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. Annu. Rev. Biophys. 37, 375–397.

[34] Vaynberg, J. et al. (2005) Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility. Mol. Cell 17, 513–523.
[35] Dhar, A., Samiotakis, A., Ebbinghaus, S., Nienhaus, L., Homouz, D., Gruebele, M. and Cheung, M.S. (2010) Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. Proc. Natl. Acad. Sci. U S A 107, 17586–17591.
[36] Aoki, K., Yamada, M., Kunida, K., Yasuda, S. and Matsuda, M. (2011) Processive phosphorylation of ERK MAP kinase in mammalian cells. Proc. Natl. Acad. Sci. U S A 108, 12675–12680.
[37] Takahashi, K., Tanase-Nicola, S. and ten Wolde, P.R. (2010) Spatio-temporal correlations can drastically change the response of a MAPK pathway. Proc. Natl. Acad. Sci. U S A 107, 2473–2478.
[38] Duyster, J., Baskaran, R. and Wang, J.Y. (1995) Src homology 2 domain as a specificity determinant in the c-Abl-mediated tyrosine phosphorylation of the RNA polymerase II carboxyl-terminal repeated domain. Proc. Natl. Acad. Sci. U S A 92, 1555–1559.
[39] Mayer, B.J., Hirai, H. and Sakai, R. (1995) Evidence that SH2 domains promote processive phosphorylation by protein-tyrosine kinases. Curr. Biol. 5, 296–305.
[40] Pielak, G.J. and Miklos, A.C. (2010) Crowding and function reunite. Proc. Natl. Acad. Sci. U S A 107, 17457–17458.
[41] Ellis, R.J. (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem. Sci. 26, 597–604.
[42] Wolpert, L. (1989) Positional information revisited. Development 107 (Suppl), 3–12.
[43] Wolpert, L. (2011) Positional information and patterning revisited. J. Theor. Biol. 269, 359–365.
[44] Kerr, I.M., Costa-Pereira, A.P., Lillemeier, B.F. and Strobl, B. (2003) Of JAKs, STATs, blind watchmakers, jeeps and trains. FEBS Lett. 546, 1–5.
[45] Spitzer, J. (2011) From water and ions to crowded biomacromolecules: in vivo structuring of a prokaryotic cell. Microbiol. Mol. Biol. Rev. 75, 491–506.
[46] Holt, C.E. and Bullock, S.L. (2009) Subcellular mRNA localization in animal cells and why it matters. Science 326, 1212–1216.
[47] Mili, S. and Macara, I.G. (2009) RNA localization and polarity: From A(PC) to Z(BP). Trends Cell Biol. 19, 156–164.
[48] Jeffery, W.R., Tomlinson, C.R. and Brodeur, R.D. (1983) Localization of actin messenger RNA during early ascidian development. Dev. Biol. 99, 408–417.
[49] Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S. and Amster-Choder, O. (2011) Translation-independent localization of mRNA in *E. coli*. Science 331, 1081–1084.
[50] Doyle, M. and Kiebler, M.A. (2011) Mechanisms of dendritic mRNA transport and its role in synaptic tagging. EMBO J. 30, 3540–3552.
[51] Jedrusik, A., Parfitt, D.-E., Guo, G., Skamagki, M., Grabarek, J.B., Johnson, M.H., Robson, P. and Zernicka-Goetz, M. (2008) Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. Genes Dev. 22, 2692–2706.
[52] Lécuyer, E. et al. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell 131, 174–187.
[53] Mili, S., Moissoglu, K. and Macara, I.G. (2008) Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. Nature 453, 115–119.
[54] Meignin, C. and Davis, I. (2010) Transmitting the message: intracellular mRNA localization. Curr. Opin. Cell Biol. 22, 112–119.
[55] Moore, M.J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. Science 309, 1514–1518.
[56] St Johnston, D. (2005) Moving messages: the intracellular localization of mRNAs. Nat. Rev. Mol. Cell Biol. 6, 363–375.
[57] Condeelis, J. and Singer, R.H. (2005) How and why does beta-actin mRNA target? Biol. Cell 97, 97–110.
[58] Besse, F. and Ephrussi, A. (2008) Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat. Rev. Mol. Cell Biol. 9, 971–980.
[59] Hüttelmaier, S. et al. (2005) Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. Nature 438, 512–515.
[60] Hirokawa, N. and Noda, Y. (2008) Intracellular transport and kinesin superfamily proteins, KIFs: structure, function, and dynamics. Physiol. Rev. 88, 1089–1118.
[61] Hirokawa, N. (1998) Kinesin and Dynein superfamily proteins and the mechanism of organelle transport. Science 279, 519–526.
[62] Horiguchi, K., Hanada, T., Fukui, Y. and Chishti, A.H. (2006) Transport of PIP3 by GAKIN, a kinesin-3 family protein, regulates neuronal cell polarity. J. Cell Biol. 174, 425–436.
[63] Taya, S. et al. (2007) DISC1 regulates the transport of the NUDEL/LIS1/14-3-3epsilon complex through kinesin-1. J. Neurosci. 27, 15–26.
[64] Setou, M., Seog, D.H., Tanaka, Y., Kanai, Y., Takei, Y., Kawagishi, M. and Hirokawa, N. (2002) Glutamate-receptor-interacting protein GRIP1 directly steers kinesin to dendrites. Nature 417, 83–87.
[65] Inomata, H., Nakamura, Y., Hayakawa, A., Takata, H., Suzuki, T., Miyazawa, K. and Kitamura, N. (2003) A scaffold protein JIP-1b enhances amyloid precursor protein phosphorylation by JNK and its association with kinesin light chain 1. J. Biol. Chem. 278, 22946–22955.
[66] Setou, M., Nakagawa, T., Seog, D.H. and Hirokawa, N. (2000) Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. Science 288, 1796–1802.
[67] Hirokawa, N., Noda, Y., Tanaka, Y. and Niwa, S. (2009) Kinesin superfamily motor proteins and intracellular transport. Nat. Rev. Mol. Cell Biol. 10, 682–696.
[68] Schnapp, B.J. (2003) Trafficking of signaling modules by kinesin motors. J. Cell Sci. 116, 2125–2135.
[69] Verhey, K.J., Meyer, D., Deehan, R., Blenis, J., Schnapp, B.J., Rapoport, T.A. and Margolis, B. (2001) Cargo of kinesin identified as JIP scaffolding proteins and associated signaling molecules. J. Cell Biol. 152, 959–970.
[70] Kusumi, A. et al. (2005) Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. Annu. Rev. Biophys. Biomol. Struct. 34, 351–378.
[71] Lindner, R. and Naim, H.Y. (2009) Domains in biological membranes. Exp. Cell Res. 315, 2871–2878.
[72] Bastiani, M. and Parton, R.G. (2010) Caveolae at a glance. J. Cell Sci. 123, 3831–3836.
[73] Li, S., Couet, J. and Lisanti, M.P. (1996) Src tyrosine kinases, Gα subunits, and H-Ras share a common membrane-anchored scaffolding protein. Caveolin. J. Biol. Chem. 271, 29182–29190.
[74] Pike, L.J. (2005) Growth factor receptors, lipid rafts and caveolae: an evolving story. Biochim. Biophys. Acta. 1746, 260–273.
[75] Lillemeier, B.F., Mörtelmaier, M.A., Forstner, M.B., Huppa, J.B., Groves, J.T. and Davis, M.M. (2009) TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. Nat. Immunol. 11, 90–96.
[76] Barda-Saad, M., Braiman, A., Titerence, R., Bunnell, S.C., Barr, V.A. and Samelson, L.E. (2004) Dynamic molecular interactions linking the T cell antigen receptor to the actin cytoskeleton. Nat. Immunol. 6, 80–89.
[77] Kwok, R. (2011) Cell biology: the new cell anatomy. Nature 480, 26–28.
[78] Li, P. et al. (2012) Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336–340.
[79] Kellis, M., Birren, B.W. and Lander, E.S. (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428, 617–624.
[80] Cavalier-Smith, T. (2009) Predation and eukaryote cell origins: a coevolutionary perspective. Int. J. Biochem. Cell Biol. 41, 307–322.
[81] Chervitz, S.A. et al. (1998) Comparison of the complete protein sets of worm and yeast: orthology and divergence. Science 282, 2022–2028.
[82] Boraas, M., Seale, D. and Boxhorn, J. (1998) Phagotrophy by a flagellate selects for colonial prey: A possible origin of multicellularity. Evol. Ecol. 12, 153–164.
[83] Rokas, A. (2008) The origins of multicellularity and the early history of the genetic toolkit for animal development. Annu. Rev. Genet. 42, 235–251.
[84] Bhattacharyya, R.P., Reményi, A., Yeh, B.J. and Lim, W.A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu. Rev. Biochem. 75, 655–680.
[85] Linding, R. et al. (2007) Systematic discovery of in vivo phosphorylation networks. Cell 129, 1415–1426.
[86] Han, J.-D.J. et al. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 430, 88–93.
[87] Pawson, T. and Scott, J.D. (1997) Signaling through scaffold, anchoring, and adaptor proteins. Science 278, 2075–2080.
[88] Balagopalan, L., Coussens, N.P., Sherman, E., Samelson, L.E. and Sommers, C.L. (2010) The LAT story: a tale of cooperativity, coordination, and choreography. Cold Spring Harb. Perspect. Biol. 2, a005512.
[89] Ferguson, K.M. (2008) A structure-based view of Epidermal Growth Factor Receptor regulation. Annu. Rev. Biophys. 37, 353–373.
[90] Horn, H.F. and Vousden, K.H. (2007) Coping with stress: multiple ways to activate p53. Oncogene 26, 1306–1316.
[91] Vousden, K.H. and Prives, C. (2009) Blinded by the light: the growing complexity of p53. Cell 137, 413–431.
[92] Uversky, V.N. and Dunker, A.K. (2010) Understanding protein non-folding. Biochim. Biophys. Acta. 1804, 1231–1264.
[93] Oldfield, C.J., Meng, J., Yang, J.Y., Yang, M.Q., Uversky, V.N. and Dunker, A.K. (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14–3-3 with their partners. BMC Genomics 9 (Suppl 1), S1.
[94] Wells, M. et al. (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc. Natl. Acad. Sci. U S A 105, 5762–5767.
[95] Dhanasekaran, D.N., Kashef, K., Lee, C.M., Xu, H. and Reddy, E.P. (2007) Scaffold proteins of MAP-kinase modules. Oncogene 26, 3185–3202.
[96] Pawson, C.T. and Scott, J.D. (2010) Signal integration through blending, bolstering and bifurcating of intracellular information. Nat. Struct. Mol. Biol. 17, 653–658.
[97] Good, M.C., Zalatan, J.G. and Lim, W.A. (2011) Scaffold proteins: Hubs for controlling the flow of cellular information. Science 332, 680–686.
[98] Harris, K., Lamson, R.E., Nelson, B., Hughes, T.R., Marton, M.J., Roberts, C.J., Boone, C. and Pryciak, P.M. (2001) Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins. Curr. Biol. 11, 1815–1824.
[99] Park, S.-H., Zarrinpar, A. and Lim, W.A. (2003) Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. Science 299, 1061–1064.
[100] Birge, R.B., Kalodimos, C., Inagaki, F. and Tanaka, S. (2009) Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. Cell Commun. Signal. 7, 13.

[101] Defilippi, P., Di Stefano, P. and Cabodi, S. (2006) P130Cas: a versatile scaffold in signaling networks. Trends Cell Biol. 16, 257–263.
[102] Mardilovich, K., Pankratz, S.L. and Shaw, L.M. (2009) Expression and function of the insulin receptor substrate proteins in cancer. Cell Commun. Signal. 7, 14.
[103] Sato, T. and Gotoh, N. (2009) The FRS2 family of docking/scaffolding adaptor proteins as therapeutic targets of cancer treatment. Expert Opin. Ther. Targets 13, 689–700.
[104] Wöhrle, F.U., Daly, R.J. and Brummer, T. (2009) Function, regulation and pathological roles of the Gab/DOS docking proteins. Cell Commun. Signal. 7, 22.
[105] Tikhmyanova, N., Little, J.L. and Golemis, E.A. (2009) CAS proteins in normal and pathological cell growth control. Cell. Mol. Life Sci. 67, 1025–1048.
[106] Kim, W., Kook, S., Kim, D.J., Teodorof, C. and Song, W.K. (2004) The 31-kDa Caspase-generated cleavage product of p130cas functions as a transcriptional repressor of E2A in apoptotic cells. J. Biol. Chem. 279, 8333–8342.
[107] Sawada, Y., Tamada, M., Dubin-Thaler, B.J., Cherniavskaya, O., Sakai, R., Tanaka, S. and Sheetz, M.P. (2006) Force sensing by mechanical extension of the Src family kinase substrate p130Cas. Cell 127, 1015–1026.
[108] Niediek, V., Born, S., Hampe, N., Kirchgeßner, N., Merkel, R. and Hoffmann, B. (2012) Cyclic stretch induces reorientation of cells in a Src family kinase- and p130Cas-dependent manner. Eur. J. Cell Biol. 91, 118–128.
[109] Fafalios, A., Ma, J., Tan, X., Stoops, J., Luo, J., Defrances, M.C. and Zarnegar, R. (2011) A hepatocyte growth factor receptor (Met)-insulin receptor hybrid governs hepatic glucose metabolism. Nat. Med. 17, 1577–1584.
[110] Adams, D.R., Ron, D. and Kiely, P.A. (2011) RACK1, a multifaceted scaffolding protein: structure and function. Cell Commun. Signal. 9, 22.
[111] Collins, M.O. and Choudhary, J.S. (2008) Mapping multiprotein complexes by affinity purification and mass spectrometry. Curr. Opin. Biotechnol. 19, 324–330.
[112] Gavin, A.-C., Maeda, K. and Kühner, S. (2011) Recent advances in charting protein–protein interaction: mass spectrometry-based approaches. Curr. Opin. Biotechnol. 22, 42–49.
[113] Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M. (2011) Global quantification of mammalian gene expression control. Nature 473, 337–342.
[114] Tran, J.C. et al. (2011) Mapping intact protein isoforms in discovery mode using top-down proteomics. Nature 480, 254–258.
[115] Benesch, J.L.P. and Ruotolo, B.T. (2011) Mass spectrometry: come of age for structural and dynamical biology. Curr. Opin. Struct. Biol. 21, 641–649.
[116] Heck, A.J. (2008) Native mass spectrometry: a bridge between interactomics and structural biology. Nat. Methods 5, 927–933.
[117] Sharon, M., Mao, H., Boeri Erba, E., Stephens, E., Zheng, N. and Robinson, C.V. (2009) Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. Structure 17, 31–40.
[118] Stengel, F., Aebersold, R. and Robinson, C.V. (2011). Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes. Mol. Cell. Proteomics 11, R111.014027.
[119] Morgner, N. and Robinson, C.V. (2012) Linking structural change with functional regulation — insights from mass spectrometry. Curr. Opin. Struct. Biol. 22, 1–8.
[120] Alber, F., Förster, F., Korkin, D., Topf, M. and Sali, A. (2008) Integrating diverse data for structure determination of macromolecular assemblies. Annu. Rev. Biochem. 77, 443–477.
[121] Robinson, C.V., Sali, A. and Baumeister, W. (2007) The molecular sociology of the cell. Nature 450, 973–982.
[122] Zhou, M. and Robinson, C.V. (2010) When proteomics meets structural biology. Trends Biochem. Sci. 35, 522–529.
[123] Russel, D., Lasker, K., Phillips, J., Schneidman-Duhovny, D., Velázquez-Muriel, J.A. and Sali, A. (2009) The structural dynamics of macromolecular processes. Curr. Opin. Cell Biol. 21, 97–108.
[124] Fotin, A., Cheng, Y., Grigorieff, N., Walz, T., Harrison, S.C. and Kirchhausen, T. (2004) Structure of an auxilin-bound clathrin coat and its implications for the mechanism of uncoating. Nature 432, 649–653.
[125] Glaeser, R.M. and Hall, R.J. (2011) Reaching the information limit in cryo-EM of biological macromolecules: Experimental aspects. Biophys. J. 100, 2331–2337.
[126] Bouchet-Marquis, C. and Hoenger, A. (2011) Cryo-electron tomography on vitrified sections: A critical analysis of benefits and limitations for structural cell biology. Micron 42, 152–162.
[127] Leis, A., Rockel, B., Andrees, L. and Baumeister, W. (2009) Visualizing cells at the nanoscale. Trends Biochem. Sci. 34, 60–70.
[128] Li, Z. and Jensen, G.J. (2009) Electron cryotomography: a new view into microbial ultrastructure. Curr. Opin. Microbiol. 12, 333–340.
[129] Lasker, K. et al. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. U S A 109, 1380–1387.
[130] Ratje, A.H. et al. (2010) Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. Nature 468, 713–716.
[131] Chapman, H.N. et al. (2011) Femtosecond X-ray protein nanocrystallography. Nature 470, 73–77.
[132] Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B. and Sali, A. (2012) Putting the pieces together: Integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. 10, e1001244.
[133] Yang, Z. et al. (2011). UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J. Struct. Biol., in press.
[134] Bellay, J., Michaut, M., Kim, T., Han, S., Colak, R., Myers, C.L. and Kim, P.M. (2012) An omics perspective of protein disorder. Mol. Biosyst. 8, 185–193.
[135] Dunker, A.K., Silman, I., Uversky, V.N. and Sussman, J.L. (2008) Function and structure of inherently disordered proteins. Curr. Opin. Struct. Biol. 18, 756–764.
[136] Dyson, H.J. (2011) Expanding the proteome: disordered and alternatively folded proteins. Q. Rev. Biophys. 44, 467–518.
[137] Galea, C.A., Wang, Y., Sivakolundu, S.G. and Kriwacki, R.W. (2008) Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. Biochemistry 47, 7598–7609.
[138] Radivojac, P., Iakoucheva, L.M., Oldfield, C.J., Obradovic, Z., Uversky, V.N. and Dunker, A.K. (2007) Intrinsic disorder and functional proteomics. Biophys. J. 92, 1439–1456.
[139] Uversky, V.N., Gillespie, J.R. and Fink, A.L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 41, 415–427.
[140] Dunker, A.K. et al. (2008) The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genomics 9 (Suppl 2), S1.
[141] Gsponer, J., Futschik, M.E., Teichmann, S.A. and Babu, M.M. (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. Science 322, 1365–1368.
[142] Tompa, P., Prilusky, J., Silman, I. and Sussman, J.L. (2008) Structural disorder serves as a weak signal for intracellular protein degradation. Proteins 71, 903–909.
[143] Simister, P.C., Schaper, F., O'Reilly, N., McGowan, S. and Feller, S.M. (2011) Self-organization and regulation of intrinsically disordered proteins with folded N-termini. PLoS Biol. 9, e1000591.
[144] Suskiewicz, M.J., Sussman, J.L., Silman, I. and Shaul, Y. (2011) Context-dependent resistance to proteolysis of intrinsically disordered proteins. Protein Sci., http://dx.doi.org/10.1002/pro.657.
[145] Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F. and Jones, D.T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J. Mol. Biol. 337, 635–645.
[146] Dinkel, H. et al. (2012) ELM–the database of eukaryotic linear motifs. Nucleic Acids Res. 40, D242–D251.
[147] Fuxreiter, M., Tompa, P. and Simon, I. (2007) Local structural disorder imparts plasticity on linear motifs. Bioinformatics 23, 950–956.
[148] Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. and Russell, R.B. (2003) Protein disorder prediction: implications for structural proteomics. Structure 11, 1453–1459.
[149] Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Uversky, V.N. and Obradovic, Z. (2007) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J. Proteome Res. 6, 1882–1898.
[150] Sigalov, A.B. and Uversky, V.N. (2011) Differential occurrence of protein intrinsic disorder in the cytoplasmic signaling domains of cell receptors. Self Nonself 2, 55–72.
[151] Nguyen Ba, A.N., Yeh, B.J., van Dyk, D., Davidson, A.R., Andrews, B.J., Weiss, E.L. and Moses, A.M. (2012) Proteome-wide discovery of evolutionary conserved sequences in disordered regions. Sci. Signal. 5, rs1.
[152] McNulty, B.C., Young, G.B. and Pielak, G.J. (2006) Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder. J. Mol. Biol. 355, 893–897.
[153] Dedmon, M.M., Patel, C.N., Young, G.B. and Pielak, G.J. (2002) FlgM gains structure in living cells. Proc. Natl. Acad. Sci. U S A 99, 12681–12684.
[154] Sugase, K., Dyson, H.J. and Wright, P.E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 447, 1021–1025.
[155] Zhang, Y., Tan, H., Chen, G. and Jia, Z. (2010) Investigating the disorder-order transition of calmodulin binding domain upon binding calmodulin using molecular dynamics simulation. J. Mol. Recognit. 23, 360–368.
[156] Ou, L., Ferreira, A.M., Otieno, S., Xiao, L., Bashford, D. and Kriwacki, R.W. (2011) Incomplete folding upon binding mediates Cdk4/cyclin D complex activation by tyrosine phosphorylation of inhibitor p27 protein. J. Biol. Chem. 286, 30142–30151.
[157] Wang, Y. et al. (2011) Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. Nat. Chem. Biol. 7, 214–221.
[158] Mittag, T. et al. (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. Proc. Natl. Acad. Sci. U S A 105, 17772–17777.
[159] Rosenbaum, J.C. et al. (2011) Disorder targets misorder in nuclear quality control degradation: A disordered Ubiquitin ligase directly recognizes its misfolded substrates. Mol. Cell 41, 93–106.
[160] Haynes, C. et al. (2006) Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput. Biol. 2, e100.
[161] Simister, P.C. and Feller, S.M. (2012) Order and disorder in large multi-site docking proteins of the Gab family-implications for signalling complex formation and inhibitor design strategies. Mol. Biosyst. 8, 33–46.
[162] Eulenfeld, R. and Schaper, F. (2009) A new mechanism for the regulation of Gab1 recruitment to the plasma membrane. J. Cell Sci. 122, 55–64.
[163] Harkiolaki, M. et al. (2009) Distinct binding modes of two epitopes in Gab2 that interact with the SH3C domain of Grb2. Structure 17, 809–822.
[164] Ito, Y. and Selenko, P. (2010) Cellular structural biology. Curr. Opin. Struct. Biol. 20, 640–648.
[165] Sakakibara, D. et al. (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. Nature 458, 102–105.

[166] Bodart, J.-F., Wieruszeski, J.-M., Amniai, L., Leroy, A., Landrieu, I., Rousseau-Lescuyer, A., Vilain, J.-P. and Lippens, G. (2008) NMR observation of Tau in Xenopus oocytes. J. Magn. Reson. 192, 252–257.
[167] Inomata, K. et al. (2009) High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. Nature 458, 106–109.
[168] Ogino, S., Kubo, S., Umemoto, R., Huang, S., Nishida, N. and Shimada, I. (2009) Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O. J. Am. Chem. Soc. 131, 10834–10835.
[169] Bernadó, P. and Svergun, D.I. (2012) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol. Biosyst. 8, 151–167.
[170] Ando, T. (2012) High-speed atomic force microscopy coming of age. Nanotechnology 23, 062001.
[171] Stark, C. et al. (2011) The BioGRID Interaction Database: 2011 update. Nucleic Acids Res. 39, D698–D704.
[172] Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.-L. and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27, 431–432.
[173] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
[174] Mogilner, A. and Odde, D. (2011) Modeling cellular processes in 3D. Trends Cell Biol. 21, 692–700.
[175] Sneddon, M.W. and Emonet, T. (2012) Modeling cellular signaling: taking space into the computation. Nat. Methods 9, 239–242.
[176] Angermann, B.R., Klauschen, F., Garcia, A.D., Prustel, T., Zhang, F., Germain, R.N. and Meier-Schellersheim, M. (2012) Computational modeling of cellular signaling processes embedded into dynamic spatial contexts. Nat. Methods 9, 283–289.