CELL
COMMUNICATION &
SIGNALING

COMMENTARY   **Open Access**

What's in a loop?

Stephan M Feller[*] and Marc Lewitzky

Abstract

DNAs and proteins are major classes of biomolecules that differ in many aspects. However, a considerable number of their members also share a common architectural feature that enables the assembly of multi-protein complexes and thereby permits the effective processing of signals: loop structures of substantial sizes. Here we briefly review a few representative examples and suggest a functional classification of different types of loop structures. In proteins, these loops occur in protein regions classified as intrinsically disordered. Studying such loops, their binders and their interactions with other loops should reveal much about cellular information computation and signaling network architectures. It is also expected to provide critical information for synthetic biologists and bioengineers.

DNAs are different from proteins in many ways. Our genomic DNA molecules are vast, even when compared to the largest proteins we know. DNA is essentially composed of 4 building blocks that are at best modified with a few extra side bits here and there. In proteins we find at least 20 different amino acids and more than one hundred types of posttranslational modifications.

The genomic DNAs of eukaryotes live mostly in one confined cell compartment, while proteins lurk in virtually every corner of the cell and many of them whiz about.

DNA seems to be usually just able to coil into spirals that coil into bigger spirals (30 nm fibre) that coil into even bigger spirals (200 nm fibre/chromosome), while proteins can take up a plethora of diverse and highly complex shapes, or, for intrinsically disordered proteins, no apparent shapes at all.

Most DNA seems to have the capacity to live forever, while probably all proteins have a quite limited lifespan.

Despite all of these differences, DNA and proteins have of course a number of things in common. Both are extremely important classes of biomolecules and both are, for example, able to store information. In addition, they share an architectural feature related to complex information processing: substantially sized loop structures.

In genomic DNAs, these loops have already been studied for decades and in some detail, but exciting new results are still constantly emerging [1-3].

To reveal its information, DNA must be untangled and often distant regions within one molecule or between fellow DNA molecules have to interact. These communicating loops enable promoters, enhancers and other regulatory elements, which are sometimes megabases apart, to come together in space and time in a highly dynamic process which is not entirely understood [4,5]. An early example for this type of long-distance interaction was the finding that the beta-globin enhancer, which is located far upstream of the globin genes, comes into close proximity when the genes are actively transcribed [6]. New methodologies, for example the Hi-C method [7-9], are now addressing DNA looping at the whole genome level. Here, protein-DNA complexes at interacting loci are preserved by fixation with formaldehyde, affinity purified and subsequently analyzed by high-throughput sequencing [9].

Apart from their crucial participation in information transfer, DNA loops also play an important role in DNA maintenance. Loop structures at the telomeric ends of chromosomes safeguard and prevent these ends from being treated as DNA double-strand breaks [10]. When the telomeric ends become critically short, loop structures are absent which eventually will result in cell cycle arrest [11].

It goes without saying that loops also play many critical roles in RNA molecules, although they are, to our knowledge, usually not as directly involved in signal processing by protein complex cross-talk.

Proteins use loops too, and in a gamut of contexts. Loop regions occur in inter-domain segments of otherwise well-folded proteins, where they can serve multiple

* Correspondence: stephan.feller@imm.ox.ac.uk
Biological Systems Architecture Group, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK

BioMed Central

functions: short loops sometimes feature as mere linkers or may also provide the required flexibility for the movement of the neighboring protein domains (linker loops [L-L]). Other loops serve as linkers regions, but also allow proteins to interact intramolecularly when undergoing shape changes (intramolecular docking loops [IMD-L]). The linker regions between the SH2s and catalytic domains of Src and Abl kinases [12,13] (and references therein) and the linker region around tyrosine 221, between the SH3 domains of the human c-Crk II protein, are well-studied examples proven to be essential for intramolecular protein binding events [14,15] (and Figure 1). Then there exists a vast number of loop regions which upon modification by specific enzymes serve as docking sites for a single protein interaction partner, or a couple of them (small docking loops [SD-L]). Such loops are found, for example, between the membrane-spanning helices of receptor and channel proteins that reside in cellular membranes. Short loops localized within a well-folded protein domain can also work together to form binding pockets for proteins and a range of other biomolecules (binding pocket loops [BP-L]).

In extracellular proteins and polypeptides, functionally vital loop structures, for example generated by disulfide bonds, are found in a vast range of contexts. Classical examples are the loops of the atrial natriuretic peptide hormone family members [32] (and Figure 1). These loops could be designated 'activity conferring loops' [AC-L].

Finally, the human proteome encompasses many proteins suspected to contain much larger loops with numerous putative sites for protein docking [33]. Such larger loops are thought to assemble crucial parts of molecular 'nanocomputers', which compute signaling input from environment-sensing transmembrane receptors [19,20,33] (and Figure 1). These could be designated as 'signal computation loops' [SC-L]). This type of loops is quite reminiscent of their DNA counterparts, which are, amongst other things, involved in transcriptional regulation. In proteins, they appear preferentially in the 'anarchic fraction' of proteomes; in humans approximately one third of the proteome is thought to consist of partially or mostly 'unstructured' i.e. 'intrinsically disordered' proteins.

These fickle and certainly understudied characters lack in many sections of their protein chains secondary and tertiary structure elements that would be detectable with the most commonly used current structure prediction programs. Some of those proteins display disordered regions of several to many hundred amino acids, which made it for a long time difficult to understand the molecular mechanisms of how they actually conduct their business in cells. A long and unrestrained amino acid chain that is able to move about freely can in principle adopt a virtually infinite number of conformations. So, presumably, none of these large molecules would ever take on an identical shape. Yet they do perform their usually complex duties in cells swiftly and effectively.

Protein chain loops are, of course, a feature that can make this possible [19]. Looping will substantially restrict the conformational space that protein chains can sample, but at the same time leave enough flexibility to allow multiple other proteins to bind and covalently or noncovalently modify the intrinsically disordered protein chains. Importantly, protein loop structures of sufficient size should easily be able to simultaneously interact with several other protein and/or DNA loops, thereby enabling diverse and complex signal computation operations through the cross-talk of two or more signaling pathways in a simple yet elegant manner.

At present, we still know relatively little details about large protein loops involved in signal processing. Nevertheless, it would seem to be a fair guess that the unstructured parts of many proteins 'live' in a zone of intermediate flexibility between the hypothetical structural poles of 'complete chaos' on one end and substantial rigidity imposed through extensive constraints by a highly folded protein on the other end.

That is not to say that all disordered proteins must use loops. Some proteins and in particular those that lack any well-folded regions and are relatively short may rely entirely on a more or less linear protein chain for their biological actions and, for example, adopt local structure only upon interaction with their binding partner(s). The binding of p27Kip1 to the Cdk2/cyclinB complex is a well-known point in case [34].

Nevertheless, higher order signal processing events can be expected to rely in many cases on two or more loops structures that enable the coordinated assembly of sub-complexes, which can then communicate to decide cell fates and orchestrate and direct the subsequently required cell actions. In order to generate and organize these loop structures in intrinsically disordered protein chains an 'organizing center' is probably mandatory. In many proteins, for example those of the Gab, Irs/Dok, Frs and p130Cas/BCAR1 families, well-folded N-terminal domains might serve as anchor points [33,35,36]. In principle, such an organizing center could also be more C-terminally located, or there could be more than one such center within a single protein chain.

So what about large disordered proteins without a recognizable folded region? It is conceivable that these characters could use other biomolecules, most commonly other proteins, as helpers to 'get into shape'. Stabilizing loops that are built around a folded domain could in some cases also be supported by 'loop-stabilizing intermolecular interactions in the form of

Figure 1 (See legend on next page.)
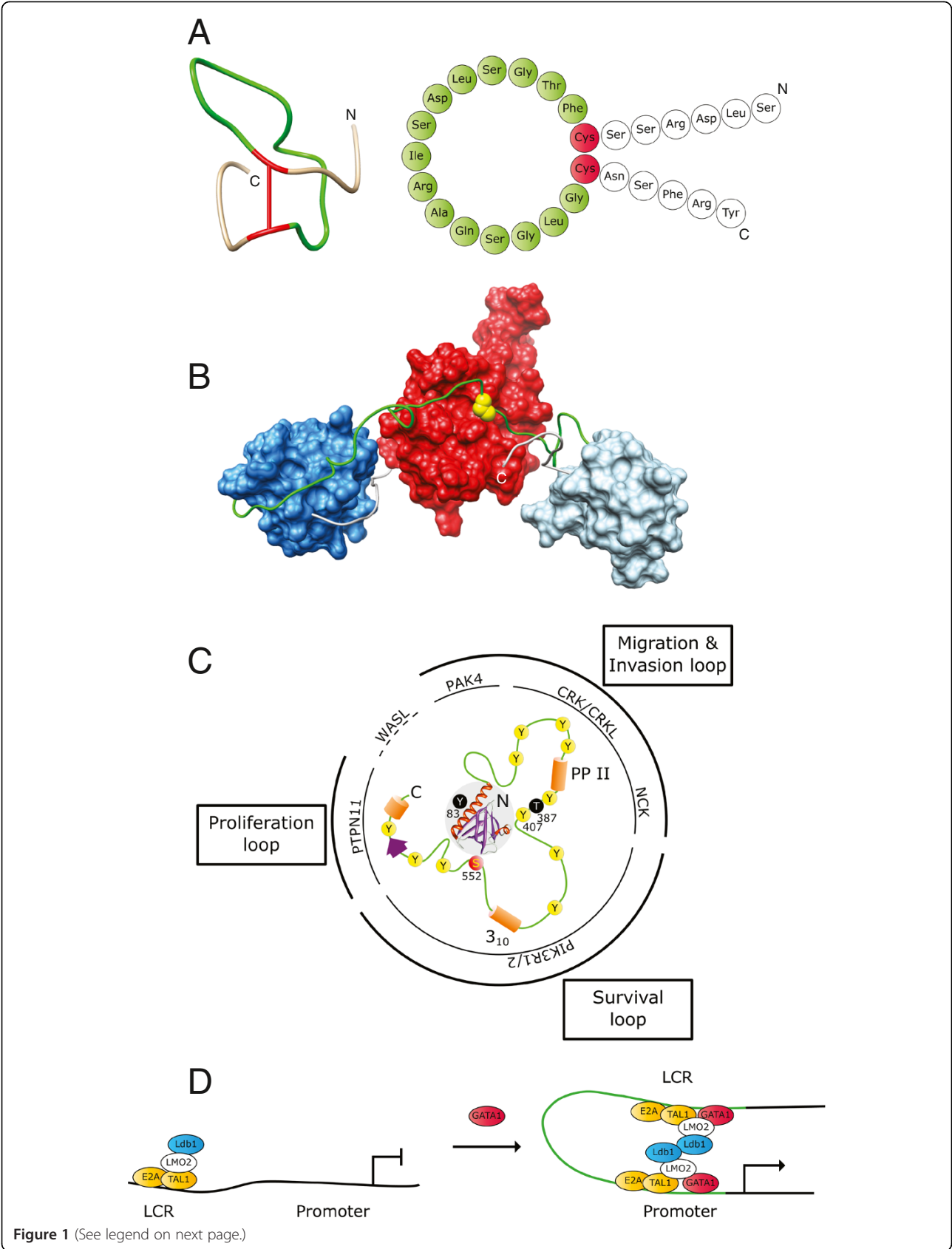
(See figure on previous page.)

Figure 1 Selected examples for different functional types of loops in large biomolecules. A. Atrial natriuretic peptide hormones contain vital 'activity conferring' loops (green) generated in each case by a single disulfide bond (red). Shown here is the solution conformation and cartoon representation of one ANP variant. Solution conformation generated from PDB database entry 1ANP [16] using Chimera [17]. Cartoon representation based on similar representation in [18]. **B**. The 'intramolecular docking' loop (green) in the c-Crk II protein regulates the overall conformation of its SH2 (red), SH3N (dark blue) and SH3C (light blue) domains by an inducible intramolecular interaction between the SH2 domain and a phosphorylated tyrosine residue (yellow) in the loop region. Structural representation generated from PDB entry 2DVJ (aa 1-228; SH2, SH3N and loop) and PDB entry 2EYZ (aa 229-304; loop and SH3C) [15]. **C**. Gab1 contains an N-terminal PH domain (grey shaded area) followed by a largely unstructured region (green) with numerous sites for potential intra- and intermolecular interactions (yellow: tyrosine-phosphorylation sites, red: serine-phosphorylation site, orange and purple: secondary structure elements). This 'signal computation' loop permits the assembly of and signaling via context-specific complexes [19,20]. The poly-proline type II helix (PP II) and the 3_{10} helix (3_{10}) in Gab1 and its close relative Gab2 can interact with the Grb2 adapter protein [21,22]. Cortactin was also reported to interact with these regions [23]. pTyr407 was mapped as binding site for NCK [24]. pSer 552 allows intramolecular interaction of the loop region with the PH domain and regulates Gab1 localization [25]. Interaction regions for PAK4 kinase [26], CRK-family proteins [27,28], PI3 kinase (PIK3R1/2; [29]) and the phosphatase SHP2 (PTPN11; [30]) have been described. WASL (N-WASP) may interact directly with the PH domain (Richard Vaillancourt, Morag Park et al.; personal communication). Elements associated with specific cellular functions like cell motility, survival or proliferation are often found co-localized in defined regions of the signaling loop. Y83 and T387 are two residues mutated in breast cancer. The Y83C mutation could interfere with PI(3,4,5)P3 binding of the PH domain [25], while a T387N mutation abolishes a threonine residue phosphorylated after EGFR or c-Met stimulation [31]. Structural representation of PH domain generated from PDB entry 2X18. **D**. A 'signal computation' loop (green) in a DNA molecule controls transcriptional activity by bringing locus control region (LCR) and promoter region together in the presence of a crucial transcription factor (red: GATA1). Based on a similar figure in [1].

homo- and hetero-oligomers. Proteomic studies of such native complexes under gentle conditions using ion mobility mass spectrometry [37] might provide first insights into the key components involved and into some details of their interactions, hence providing a basis for further detailed biophysical analyses.

We have learned a lot about the individual building blocks in cells and their potential interaction partners in recent years. Now we need to put the pieces together, to develop a much better architectural understanding of the molecular processes in cells. A functional classification of loop structures, such as we have suggested here, might help along the way, although it is clear that nature's ingenuity and use of highly versatile molecular themes that are almost endlessly varied will finally defeat our brains' apparent urge to put everything we encounter into neat little boxes.

An architectural, more holistic view of the highly dynamic living cells can, of course, be only obtained by combining a wide range of methodologies, including different super-resolution microscopy techniques [38-41], in-cell NMR [42-46], high speed in solution atomic force microscopy [47] and cryoEM [48-51], to name but a few. Studying proteins, as much as possible, in their natural habitat [52] will also substantially contribute to a more life-like model of cell architectures and cell signaling networks.

In addition, we probably need new ways to identify, describe and classify the highly dynamic, often disordered regions and arrangements of proteins, for example along the lines suggested by Peter Tompa [53] and others [54-56]. Accurate descriptions of loop structures may not be entirely trivial to obtain, but there are several algorithms and databases readily available to model loops in proteins [57-59]. We would like to argue, that

for a full description, biophysical classifications should be combined with functional descriptions, such as the one proposed in an earlier section of this manuscript.

These are exciting times for cell biologists, biochemists, biophysicists and molecular biologists, but also for chemists, physicists and engineers. As more and more research fields interconnect to answer together questions that seemed far out of reach just a few years back, we can expect to soon reach new shores on the fascinating journey into biological nanospace. The conceptual trinkets and architectural maps we will bring back from these travels should provide vital clues for better treatments of diseases, synthetic biology and a wide range of bioengineering tasks.

References
1. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA: **Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor.** *Cell* 2012, **149**:1233–1244.
2. Li GW, Berg OG, Elf J: **Effects of macromolecular crowding and DNA looping on gene regulation kinetics.** *Nature Physics* 2009, **5**:294–297.
3. Schleif R: **DNA looping.** *Annu Rev Biochem* 1992, **61**:199–223.
4. Dekker J: **Gene regulation in the third dimension.** *Science* 2008, **319**:1793–1794.
5. Sanyal A, Lajoie BR, Jain G, Dekker J: **The long-range interaction landscape of gene promoters.** *Nature* 2012, **489**:109–113.
6. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W: **Looping and interaction between hypersensitive sites in the active beta-globin locus.** *Mol Cell* 2002, **10**:1453–1465.
7. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions.** *Nature* 2012, **485**:376–380.
8. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, *et al*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome.** *Science* 2009, **326**:289–293.

9. van Berkum NL, Lieberman-Aiden E, Williams L, Imakaev M, Gnirke A, Mirny LA, Dekker J, Lander ES: **Hi-C: a method to study the three-dimensional architecture of genomes.** *J Vis Exp* 2010, **39:**e1869.

10. de Lange T: **Shelterin: the protein complex that shapes and safeguards human telomeres.** *Genes Dev* 2005, **19:**2100–2110.

11. Verdun RE, Crabbe L, Haggblom C, Karlseder J: **Functional human telomeres are recognized as DNA damage in G2 of the cell cycle.** *Mol Cell* 2005, **20:**551–561.

12. Hantschel O, Superti-Furga G: **Regulation of the c-Abl and Bcr-Abl tyrosine kinases.** *Nat Rev Mol Cell Biol* 2004, **5:**33–44.

13. Chen S, Brier S, Smithgall TE, Engen JR: **The Abl SH2-kinase linker naturally adopts a conformation competent for SH3 domain binding.** *Protein Sci* 2007, **16:**572–581.

14. Feller SM, Knudsen B, Hanafusa H: **c-Abl kinase regulates the protein binding activity of c-Crk.** *EMBO J* 1994, **13:**2341–2351.

15. Kobashigawa Y, Sakai M, Naito M, Yokochi M, Kumeta H, Makino Y, Ogura K, Tanaka S, Inagaki F: **Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK.** *Nat Struct Mol Biol* 2007, **14:**503–510.

16. Fairbrother WJ, McDowell RS, Cunningham BC: **Solution conformation of an atrial natriuretic peptide variant selective for the type A receptor.** *Biochemistry* 1994, **33:**8897–8904.

17. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF Chimera–a visualization system for exploratory research and analysis.** *J Comput Chem* 2004, **25:**1605–1612.

18. Katzung BG, Masters SB, Trevor AJ: *Basic & Clinical Pharmacology.* New York City, USA: The McGraw-Hill Companies, Inc; 2009.

19. Lewitzky M, Simister PC, Feller SM: **Beyond 'furballs' and 'dumpling soups' - towards a molecular architecture of signaling complexes and networks.** *FEBS Lett* 2012, **17:**2740–2750.

20. Simister PC, Feller SM: **Order and disorder in large multi-site docking proteins of the Gab family–implications for signalling complex formation and inhibitor design strategies.** *Mol Biosyst* 2012, **8:**33–46.

21. Harkiolaki M, Tsirka T, Lewitzky M, Simister PC, Joshi D, Bird LE, Jones EY, O'Reilly N, Feller SM: **Distinct binding modes of two epitopes in Gab2 that interact with the SH3C domain of Grb2.** *Structure* 2009, **17:**809–822.

22. McDonald CB, Seldeen KL, Deegan BJ, Bhat V, Farooq A: **Binding of the cSH3 domain of Grb2 adaptor to two distinct RXXK motifs within Gab1 docker employs differential mechanisms.** *J Mol Recognit* 2011, **24:**585–596.

23. Rajadurai CV, Havrylov S, Zaoui K, Vaillancourt R, Stuible M, Naujokas M, Zuo D, Tremblay ML, Park M: **Met receptor tyrosine kinase signals through a cortactin-Gab1 scaffold complex, to mediate invadopodia.** *J Cell Sci* 2012, **125:**1940–2953.

24. Abella JV, Vaillancourt R, Frigault MM, Ponzo MG, Zuo D, Sangwan V, Larose L, Park M: **The Gab1 scaffold regulates RTK-dependent dorsal ruffle formation through the adaptor Nck.** *J Cell Sci* 2010, **123:**1306–1319.

25. Eulenfeld R, Schaper F: **A new mechanism for the regulation of Gab1 recruitment to the plasma membrane.** *J Cell Sci* 2009, **122:**55–64.

26. Paliouras GN, Naujokas MA, Park M: **Pak4, a novel Gab1 binding partner, modulates cell migration and invasion by the Met receptor.** *Mol Cell Biol* 2009, **29:**3018–3032.

27. Garcia-Guzman M, Larsen E, Vuori K: **The proto-oncogene c-Cbl is a positive regulator of Met-induced MAP kinase activation: a role for the adaptor protein Crk.** *Oncogene* 2000, **19:**4058–4065.

28. Sakkab D, Lewitzky M, Posern G, Schaeper U, Sachs M, Birchmeier W, Feller SM: **Signaling of hepatocyte growth factor/scatter factor (HGF) to the small GTPase Rap1 via the large docking protein Gab1 and the adapter protein CRKL.** *J Biol Chem* 2000, **275:**10772–10778.

29. Holgado-Madruga M, Moscatello DK, Emlet DR, Dieterich R, Wong AJ: **Grb2-associated binder-1 mediates phosphatidylinositol 3-kinase activation and the promotion of cell survival by nerve growth factor.** *Proc Natl Acad Sci U S A* 1997, **94:**12419–12424.

30. Lecoq-Lafon C, Verdier F, Fichelson S, Chretien S, Gisselbrecht S, Lacombe C, Mayeux P: **Erythropoietin induces the tyrosine phosphorylation of GAB1 and its association with SHC, SHP2, SHIP, and phosphatidylinositol 3-kinase.** *Blood* 1999, **93:**2578–2585.

31. Ortiz-Padilla C, Gallego-Ortega D, Browne BC, Hochgrafe F, Caldon CE, Lyons RJ, Croucher DR, Rickwood D, Ormandy CJ, Brummer T, Daly RJ: **Functional characterization of cancer-associated Gab1 mutations.** *Oncogene* 2012, doi:10.1038/onc.2012.271.

32. Forssmann WG, Nokihara K, Gagelmann M, Hock D, Feller S, Schulz-Knappe P, Herbst F: **The heart is the center of a new endocrine, paracrine, and neuroendocrine system.** *Arch Histol Cytol* 1989, **52**(Suppl):293–315.

33. Simister PC, Schaper F, O'Reilly N, McGowan S, Feller SM: **Self-organization and regulation of intrinsically disordered proteins with folded N-termini.** *PLoS Biol* 2011, **9:**e1000591.

34. Lacy ER, Filippov I, Lewis WS, Otieno S, Xiao L, Weiss S, Hengst L, Kriwacki RW: **p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding.** *Nat Struct Mol Biol* 2004, **11:**358–364.

35. Mardilovich K, Pankratz SL, Shaw LM: **Expression and function of the insulin receptor substrate proteins in cancer.** *Cell Commun Signal* 2009, **7:**14.

36. Wohrle FU, Daly RJ, Brummer T: **Function, regulation and pathological roles of the Gab/DOS docking proteins.** *Cell Commun Signal* 2009, **7:**22.

37. Zhou M, Sandercock AM, Fraser CS, Ridlova G, Stephens E, Schenauer MR, Yokoi-Fong T, Barsky D, Leary JA, Hershey JW, *et al*: **Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3.** *Proc Natl Acad Sci U S A* 2008, **105:**18139–18144.

38. Huang B, Babcock H, Zhuang X: **Breaking the diffraction barrier: super-resolution imaging of cells.** *Cell* 2010, **143:**1047–1058.

39. Leung BO, Chou KC: **Review of super-resolution fluorescence microscopy for biology.** *Appl Spectrosc* 2011, **65:**967–980.

40. Toomre D, Bewersdorf J: **A new wave of cellular imaging.** *Annu Rev Cell Dev Biol* 2010, **26:**285–314.

41. Schubert W, Gieseler A, Krusche A, Serocka P, Hillert R: **Next-generation biomarkers based on 100-parameter functional super-resolution microscopy TIS.** *N Biotechnol* 2012, **29:**599–610.

42. Bodart JF, Wieruszeski JM, Amniai L, Leroy A, Landrieu I, Rousseau-Lescuyer A, Vilain JP, Lippens G: **NMR observation of Tau in Xenopus oocytes.** *J Magn Reson* 2008, **192:**252–257.

43. Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y, Hiroaki H, Shirakawa M: **High-resolution multi-dimensional NMR spectroscopy of proteins in human cells.** *Nature* 2009, **458:**106–109.

44. Ito Y, Selenko P: **Cellular structural biology.** *Curr Opin Struct Biol* 2010, **20:**640–648.

45. Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Walchli M, *et al*: **Protein structure determination in living cells by in-cell NMR spectroscopy.** *Nature* 2009, **458:**102–105.

46. Ogino S, Kubo S, Umemoto R, Huang S, Nishida N, Shimada I: **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* 2009, **131:**10834–10835.

47. Ando T: **High-speed atomic force microscopy coming of age.** *Nanotechnology* 2012, **23:**062001.

48. Bouchet-Marquis C, Hoenger A: **Cryo-electron tomography on vitrified sections: a critical analysis of benefits and limitations for structural cell biology.** *Micron* 2011, **42:**152–162.

49. Leis A, Rockel B, Andrees L, Baumeister W: **Visualizing cells at the nanoscale.** *Trends Biochem Sci* 2009, **34:**60–70.

50. Li Z, Jensen GJ: **Electron cryotomography: a new view into microbial ultrastructure.** *Curr Opin Microbiol* 2009, **12:**333–340.

51. Grigorieff N, Harrison SC: **Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy.** *Curr Opin Struct Biol* 2011, **21:**265–273.

52. Feller SM, Lewitzky M: **Very 'sticky' proteins - not too sticky after all?** *Cell Commun Signal* 2012, **10:**15.

53. Tompa P: **On the supertertiary structure of proteins.** *Nat Chem Biol* 2012, **8:**597–600.

54. Fernandez-Fuentes N, Hermoso A, Espadaler J, Querol E, Aviles FX, Oliva B: **Classification of common functional loops of kinase super-families.** *Proteins* 2004, **56:**539–555.

55. Kumar MV, Swaminathan R: **A novel approach to segregate and identify functional loop regions in protein structures using their Ramachandran maps.** *Proteins* 2010, **78:**900–916.

56. Tran TT, Kulis C, Long SM, Bryant D, Adams P, Smythe ML: **Defining scaffold geometries for interacting with proteins: geometrical classification of secondary structure linking regions.** *J Comput Aided Mol Des* 2010, **24:**917–934.

57. Hildebrand PW, Goede A, Bauer RA, Gruening B, Ismer J, Michalsky E, Preissner R: **SuperLooper–a prediction server for the modeling of loops in globular and membrane proteins.** *Nucleic Acids Res* 2009, **37**:W571–574.
58. Fiser A, Sali A: **ModLoop: automated modeling of loops in protein structures.** *Bioinformatics* 2003, **19**:2500–2501.
59. Subramani A, Floudas CA: **Structure prediction of loops with fixed and flexible stems.** *J Phys Chem B* 2012, **116**:6670–6682.