PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

Folding cooperativity and allosteric function in the tandem-repeat protein class

Journal:	*Philosophical Transactions B*
Manuscript ID	RSTB-2017-0188.R1
Article Type:	Opinion piece
Date Submitted by the Author:	n/a
Complete List of Authors:	Perez-Riba, Albert; University of Cambridge, Department of Pharmacology Synakewicz , Marie; University of Cambridge, Department of Pharmacology Itzhaki, Laura; University of Cambridge, Department of Pharmacology
Issue Code (this should have already been entered but please contact the Editorial Office if it is not present):	ALLOSTERY
Subject:	Biophysics < BIOLOGY, Molecular Biology < BIOLOGY, Structural Biology < BIOLOGY, Computational Biology < BIOLOGY
Keywords:	Elastic Network Models, Rap proteins, Molecular switches, Allostery, Protein Cooperativity, Repeat proteins

Folding cooperativity and allosteric function in the tandem-repeat protein class

Albert Perez-Riba*, Marie Synakewicz* and Laura S. Itzhaki

*Equally contributing authors

Department of Pharmacology, University of Cambridge,

Tennis Court Road, Cambridge CB2 1PD, UK

Abstract

The term allostery was originally developed to describe structural changes in one binding site induced by the interaction of a partner molecule with a distant binding site [1], and it has been studied in depth in the field of enzymology [2–5]. Here we discuss the concept of action at a distance in relation to the folding and function of the solenoid class of tandem-repeat proteins such as tetratricopeptide repeats and ankyrin repeats. Distantly located repeats have a fold cooperatively, even though only nearest-neighbour interactions exist in these proteins. A number of repeat-protein scaffolds have been reported to display allosteric effects, transferred through the repeat array, that enable them to direct the activity of the multi-subunit enzymes within which they reside. We also highlight a recently identified group of tandem-repeat proteins, the RRPNN sub-class of tetratricopeptide repeats (TPRs), recent crystal structures of which indicate that they function as allosteric switches to modulate multiple bacterial quorum-sensing mechanisms. We believe that the folding cooperativity of tandem-repeat proteins and the biophysical mechanisms that transform them into allosteric switches are intimately intertwined. This opinion piece aims to combine our understanding of the two areas and develop ideas on their common underlying principles.

Key words: Repeat proteins, Protein cooperativity, Allostery, Molecular switches, Rap proteins, Elastic Network Models

Tandem-repeat protein: Folding for function?

Tandem-repeat domains are one of the most common protein architectures (Fig. 1). The frequency of these arrays is probably a result of replication slippage and recombination events on the DNA [6,7]. These mechanisms are considered sources of hypermutability and have given rise to a high polymorphism rate compared to the background rate of point

mutations [7–9]. Tandem-repeat proteins have been grouped into different classes according to the size (number of amino acids) of the individual repeats [6,7,10]. In this work, we will focus on the solenoid class comprising repeats of ~12-40 amino acids. Individual repeats are not independently stable and a minimum of two or three repeats is required for a stable unit.

The simplest solenoid proteins contain repeats of two secondary structure elements: α/α, α/β or β/β. More complex repeats have three or four secondary elements [10,11]. In all cases, the secondary structure elements and their relative arrangement give rise to a variety of tertiary structures whose geometries can readily be described using the three angles between the repeat planes: curvature, twist and lateral bending [12,13]. The "solenoid" term originally referred to a coil wound into a tightly packed helix. The repeats pack to form superhelices that differ greatly in their geometries, dependent on the structural class: some fold into planar, horse-shoe-like structures, others form spring-like helices, and others are very linear. All share a common feature in that their architectures create elongated interfaces for molecular recognition, mostly to other proteins but also for some sub-classes to DNA and RNA [14–18]. The α-solenoids, the focus of this opinion piece, comprise repeats of a hairpin of antiparallel α-helices [7]. Armadillo repeats [19], HEAT repeats [20,21], and the tetratricopeptide repeats (TPR) [22] are the most common members of this class.

Cooperativity in the folding of tandem-repeat proteins: Relationship to function?

Tandem-repeat protein structures are stabilised exclusively by local interactions either within a repeat or between adjacent repeats. In contrast, the stability of globular proteins originates from the high cooperativity between sequence-distant interactions and the burial of a large hydrophobic surface area. Nevertheless, small repeat proteins (~100 amino acids) such as p16, myotrophin and the Notch ankyrin domain, show two-state unfolding at equilibrium, like globular proteins of similar sizes [23–25]. Protein engineering analysis of p16,

myotrophin and gankyrin mapped out their kinetic folding and unfolding pathways and revealed polarised transition states in which structure was localised to a subset of repeats at one end of the protein [26–30], whereas for Notch it was the central repeats that were structured in the transition state [31]. It was shown that the order in which the repeats fold is governed by their relative stabilities, with the most stable repeats folding first, and consequently the folding pathways can be redirected relatively straightforwardly by manipulating the stability distribution across the repeat array [28,30–32]. It follows also that under any given set of conditions there may be flux through multiple alternative pathways [28], as originally predicted by Wolynes and colleagues [33]. Moreover, the cooperativity of the folding process (both at equilibrium and under kinetic conditions) can also be readily tuning using appropriate mutations [34,35].

High cooperativity is not always desirable, as non-cooperative folding may be important for biological function of some repeat proteins. A striking example is the interaction between the transcription factor NFκB and the 6-ankyrin repeat protein IκBα, which regulates NFκB by sequestering it in the cytoplasm. The two C-terminal ankyrin repeats of IκBα are intrinsically unfolded and only fold upon binding to NFκB. Not only was the folding process shown to be critical for high-affinity binding, but the large difference in intracellular stability of un-complexed IκBα compared with the NFκB–bound form was also shown to play an essential role in transcriptional regulation. Un-complexed IκBα with its unfolded repeats 5-6 is degraded in an ubiquitin-independent manner with a very short half-life, whereas NFκB-bound IκBα is stable in the cytoplasm and requires triggered ubiquitin-mediated proteolysis for its degradation and the subsequent release of NFκB [36].

Another example of the relationship between folding cooperativity and function is the 15-HEAT repeat protein PR65. PR65 is the scaffold subunit of the heterotrimeric enzyme protein phosphatase 2A (PP2A). The catalytic subunit and the substrate-bound regulatory subunit

bind at opposite ends of the PR65 repeat array, and it has been proposed that rather than being a rigid scaffold for these two subunits, PR65 functions as an elastic connector that coordinates cycles of catalysis of multiply phosphorylated substrates [37]. Our analysis suggests that the non-cooperative unfolding of the HEAT repeats, arising from the very heterogeneous distribution of stabilities across the repeat array, might also facilitate PR65's connector function [38].

The nearest-neighbour description of repeat protein folding

The simple topology of the repeat-protein architecture has enabled the use of a 1-D Ising model description to define the energetic values of each repeat under the assumption of all repeats being coupled. The Ising model was originally developed to describe interactions of atomic dipole spins in ferromagnetic materials. In such a material the atomic dipoles can adopt one of two states (spin +1 or -1). However, their states are coupled to their nearest neighbours through an exchange interaction, a potential that favours parallel alignment between states [43]. Due to this coupling, flipping of one spin can result in cascades, or so-called "spin-waves" [44]. In early work on the ankyrin-repeat domain of Notch, Barrick and co-workers recognised that the protein's stability follows a simple additive rule [23]. Regan and co-workers applied the Ising model to so-called 'consensus-designed repeat proteins, comprising identical repeats containing the most conserved residues in a protein family [39]. A single value of intrinsic stability of the repeats (ΔG_i) and of interfacial stability between repeats (ΔG_{ij}) was shown to be sufficient to describe the folding of a series of consensus-designed repeat proteins of increasing size. These energetic values are additive, and the Gibbs free energy of unfolding of a repeat protein comprising identical repeats thus follows the equation:

$$\Delta G_{D-N} = n\Delta G\mathrm{i} + (n-1)\Delta G\mathrm{i,j} = -RTln\kappa^{n}\tau^{(n-1)}$$

where n is the number of repeats, κ is the intrinsic stability and τ is the interfacial stability [39–42]. Several families of repeat proteins have been found to follow the Ising model both at equilibrium and under kinetic conditions [45] [39,46,47], and with this description one can see that the origin of cooperativity of repeat-protein folding lies in the mismatch between the intrinsic and interfacial repeat stabilities.

Folding cooperativity of repeat proteins breaks down above ~100-150 amino acids, similar to the cooperativity limit of globular proteins [48,49]. Moreover, for repeat proteins both large and small, an array with (a) fewer intrinsically stable repeats, (b) high interfacial stability relative to intrinsic stability, and (c) a more homogeneous distribution of stabilities across the array length will tend to unfold more cooperatively. Indeed, when such conditions are met, the folding of even giant repeat proteins of 300 or more amino acids has, strikingly, been shown to approximate two-state behaviour [34,50,51].

Hydrogen-exchange experiments have shown that the internal repeats of consensus-designed repeat proteins are more protected from exchange than the terminal repeats [52–54]. That is, even when the repeats are identical in sequence they are not all equally stable. The probabilistic nature of the Ising model and the higher stabilities for a greater number of repeats can be explained in simple terms. For a repeat to unfold, it requires its neighbouring repeats to unfold also. Thus, the terminal repeats of the array are the most likely to unfold, as they have only one neighbour. In natural repeat proteins, however, this simple additivity of internal and interfacial energies becomes more difficult to dissect because repeats have different sequences and therefore different stabilities. For example, analysis of the unfolding pathway of the 15-HEAT repeat protein PR65 showed that this giant repeat protein has weak central repeats, which unfold before the N- and C-terminal sub-domains [55,38].

We recently showed that extending the length of a single inter-repeat loop in a consensus-designed TPR protein can have a large effect on stability depending on the number of repeats in the array (AP, LSI et al., under revision). This is in contrast to the small effects observed upon insertion of a long loop into consensus-designed ankyrin-repeat proteins and beta-helical repeat proteins [56,57]. Although further investigation is required, there does appear to be a trend of a greater energetic cost of loop extensions when the repeat type has a smaller mismatch between intrinsic and interfacial stability as is the case for TPRs. In other words, short inter-repeat loops are required for a repeat protein that has weak inter-repeat interfaces, possibly due to a low enthalpic and high entropic contribution to the overall stability.

In summary, the rules governing the cooperativity of repeat protein structures are now well understood. More and more we are starting to see that repeat proteins are not static rods and that the natural functions of many repeat proteins require highly dynamic conformational properties. In this opinion piece, we question the relationship between the folding cooperativity and the function of repeat proteins and whether cooperativity plays a role in controlling the transmission of information across the repeat array.

The RRNPP family of molecular switches

In recent years, a new family of bacterial regulators has been gaining recognition. Known as the RRNPP family, the name of these cytosolic peptide-sensing regulators refers to the founding members of the family Rap-Rgg-NprR-PlcR-Prgx [58]. They all have the same domain organisation: An N-terminal 3-helix bundle, a flexible helical linker and a C-terminal TPR capable of binding short peptides of five to eight residues. Notably, the N-terminal domain and the helical linker form a four-helix bundle that resembles a pair of TPRs. These proteins share the following mechanism: Peptide binding to the C-terminal domain triggers a conformational change that propagates to the N-terminal domain. Here, we examine the Rap

proteins of *Bacillus subtillis*, cytosolic aspartate phosphatases that affect downstream gene expression upon binding of the quorum-sensing peptide to their C-terminal TPR domain.

Rap phosphatases and their peptide activators were originally described in *Bacillus subtilis* by Perego and co-workers [59–64]. There are 16 Rap homologs in *B. subtilis*. As an example, RapH acts as phosphatase of Spo0F in its Apo form (peptide unbound) and prevents downstream sporulation, whereas RapF binds and inhibits gene regulators such as ComA [58]. RapH and RapF were co-crystallised with Spo0F and ComA respectively, and both Rap homologs showed the same overall conformation when bound to their partner molecule (Fig. 2) [65]. In another study, the crystal structure of Apo-RapI was compared with that of RapJ in complex with the PhrC peptide. The solenoid structure of the RapJ-PhrC complex showed a higher degree of compactness relative to Apo-RapI [66] (Fig. 2).

There is not a complete set of crystal structures of the same Rap homolog in the three conformational states. Nevertheless, Parashar and co-workers used homologous structures to propose a mechanism of action for the allosteric signal transduction (Fig. 2) [66]. They concluded that quorum-sensing peptides inhibit Rap function via an allosteric mechanism. The compact solenoid was described as the inactive configuration and the extended solenoid as the active one. In its active configuration, the N-terminal helix-bundle is capable of exposing the Spo0F- or ComA-binding regions of RapF and RapH respectively. Peptide-bound Rap proteins undergo a conformational change locking the N-terminal domain in a compact configuration in which its binding sites are inaccessible.

Rap proteins are very different from their artificial counterparts, the consensus-designed TPR proteins (CTPRs). Both types of TPRs are capable of forming large cooperatively folded repeat arrays, but somehow in the amino acid sequence of the Rap proteins there is encoded the additional ability to generate an repeat array with extreme flexibility and consequent

dynamic switch-like behaviour capable of transmitting the information of an input to an output across the array. The nearest-neighbour cooperativity between repeats appears to have increased its complexity. Ultimately, allostery necessarily requires a dynamic system, providing further evidence that repeat proteins are not simply rigid rods.

Does intrinsic dynamics of Rap proteins form the underlying basis for allostery?

Given the functional relevance of the conformational changes seen for the Rap proteins crystallographically, we have conducted an extensive analysis to investigate whether they arise from the intrinsic dynamics of each protein. To model the vibrational dynamics, we have chosen to use Elastic Network Models (ENMs) for their strong dependence on the shape of the overall structure instead of atomistic detail (see SI Methods, extensively reviewed in [67] and references therein; [68]). In an ENM, protein dynamics are decomposed into different motions with specific directions, the normal modes. The lowest three normal modes of RapI are shown in Figure 3. The predominant motion is that of bending, followed by a screw-like twist of the TPR helix and more localised motions of the N- and C-terminal 3-helix bundle and TPR repeats, respectively. We compared normal modes of different Rap configurations using ENMs of a structure-based sequence alignment (Fig. S1). The dynamics we observe are very similar in all four proteins, and the lowest normal modes tend to involve the collective motion of ~40-80% of the protein (Fig. S2). Due to the high structural similarity between RapH and RapF, both proteins explore a nearly identical motional space (Fig. S3). The major difference between the extended conformation, RapI, and the active conformations is the orientation of the N-terminal 3-helix bundle, and hence the normal modes of all three proteins are very similar, especially when the TPR domains are modelled independently (Fig. S3). In contrast, when the Rap protein adopts the compact and inactive conformation, only the motion of a very few of the lowest normal modes remains conserved to some degree, most of which are dominated by the motion of the TPR domain (Fig. S3).

Comparing both peptide-bound and –unbound ENMs of RapJ reveals that most changes in dynamics of the lower modes are not due to the presence of the bound peptide but simply due to the compacted conformation (Fig. S4).

We further examined whether the motions observed in the ENMs can account for some of the conformational changes observed in crystallography by measuring how well the lowest five normal modes of a protein overlap with the conformational transition to another protein (Fig. 4). The transition between the extended and compact conformations is very well described by the normal modes of either RapI (to 0.87) or RapJ (to 0.75). Transitions between either compact or extended conformation and the active forms (Rap H and RapF) are less well described by the lowest five normal modes due to the rotation of the N-terminal 3-helix bundle and linker domain (Fig. 4). This observation is not entirely unexpected, as more localised motions are captured by normal modes of higher order. When mapping the normal modes to the conformational changes seen within the independently modelled TPR domains (small numbers in Fig. 4), agreements are > 0.7 for most.

Considering these comparisons, a mechanism emerges by which at least the TPR domain of Rap proteins on its own could potentially explore the different conformations observed. The peptide, substrate or transcription factor could then simply trap the protein energetically in a given conformation. In fact, when considering the entropic contributions of each normal mode (Fig. 5), the extended conformation is entropically the most favourable. The compact conformation of RapJ comes along with a considerable entropic cost making it energetically unfavourable (Fig. S5), for which the enthalpic contributions of multiple contacts between peptide and TPR need to compensate [66].

Lastly, by analysing the correlation of motion between different residues (Fig. 6), we can obtain an insight into why the 3-helix bundle in the compact conformation has very little

potential for rotating to bind partner proteins. The TPRs exhibit correlated motion only with their nearest neighbours, giving rise to the distinctive pattern of squares along the diagonal [50]. Binding of the peptide marginally increases nearest-neighbour correlation at the centre (purple box in Fig. 6) which understandably arises from the cross-correlations of higher modes (Figs. S4 and S6). Movements of the rotated N-terminal 3-helix bundle, linker domain and first TPR repeat (blue box in Fig. 6) are strongly correlated, suggesting that they form a subdomain relative to the rest of the TPR repeats. Some nearest-neighbour correlations are reduced in the extended conformation of RapI, whereas they are either further reduced or even reversed in the compact conformation (arrows in Fig. 6). It is notable that the presence of the peptide does not directly influence the correlated motion of the N-terminal domain, indicating that locking of the 3-helix bundle to the TPRs is entirely due to the conformational change. However, the peptide could induce indirect or allosteric effects by stabilising the TPR domain in the compact conformation. These effects could then be "transmitted" through the array via the interaction potential between repeats, that is via the altered cooperativity of individual repeat interfaces due to the TPR rearrangement in the compact conformation.

In summary, the compact and extended conformations of Rap protein have different supramolecular geometries, arising from differences in the inter-repeat packing. Consequently, they must have different values of the interfacial repeat stability. The ENMs showed that both conformations are easily accessible through the motions of the TPR domain, albeit that an extended conformation of the array may be preferred due to the entropic cost of the compact state. Ultimately, the intrinsic flexibility of the TPR array may allow for the existence of two functionally different conformations that can be locked by their respective binding partners.

Relating conformational flexibility to the allosteric mechanisms of "banana-shaped" repeat proteins in multi-protein enzyme complexes

When we look across the repeat protein class, the Rap proteins are not the only example where the repeat scaffold may contribute to allosteric mechanisms due to its dynamic flexibility. In quite a few systems the repeat protein must change its conformation to bind to a variety of partners that all differ in shape and size. We are currently investigating proteins of different repeat types to examine whether their experimentally observed dynamics can also be described by ENM normal modes. As global motions are largely determined by the over-all shape of a molecule [67], one of our leading questions is whether the dynamics of two different proteins with the same tertiary shape will exhibit the same motions, independent of repeat type.

One such protein is PR65, and, as mentioned earlier, we are interested in understanding how it regulates the activity of the heterotrimeric PP2A (Fig. 7A). From crystal structures of PR65 in complex with the catalytic C-subunit and different regulatory B-subunits, it is clear that the PR65 needs to be highly flexible in its structure to be able to form the multitude of PP2A complexes that are present in the cell [70–72]. Biophysical analysis has shown that binding of the catalytic subunit to the C-terminal repeats of PR65 increases by an order of magnitude the affinity of the N-terminal repeats of PR65 for an inhibitor, the SV40 small t antigen [73]. However, there are no obvious direct contacts between the small t antigen and the catalytic subunit [73], suggesting a process by which PR65 functions as an allosteric transmitter of catalytic-subunit binding, though the underlying mechanism involved remains to be resolved.

A number of years ago we demonstrated such an allosteric effect in the LRR protein Skp2, which is one of many variable substrate-binding subunits of the multi-subunit SCF (Skp1-Cullin-F-box) ubiquitin ligases (Fig. 7B) [74]. Skp2 has an F-box motif, with which it binds to the Skp1 subunit, thereby connecting it to the Cullin subunit and the rest of the SCF. The C-terminal “tail” of Skp2 is unstructured and folds back onto the concave face of the LRR domain (Fig. 7C) [75]. Binding of the accessory subunit Cks1 to the C-terminus of the LRR

domain results in hydrogen-deuterium exchange protection of the N-terminal LRR repeats without any direct contacts between Cks1 and N-terminal Skp2 repeats [74]. We therefore proposed that binding of Cks1 decreases fluctuations in the C-terminal tail of Skp2, thereby stabilising residues in the tail that form a β-sheet between the first LRR and the F-box. In the absence of this β-sheet the linker between the LRR domain and the F-box may constitute a hinge, which could account for deprotection of the N-terminal LRRs when Cks1 is not bound. Thus, the hinge may function as a sensor of substrate binding, tightening of which could reduce the motions of Skp2 and thereby allow for efficient ubiquitination and/or this binding event could be translated allosterically through the Cullin subunit to the E2 ubiquitin-conjugating enzyme. The Cullins themselves are highly flexible repeat proteins, and their ability to change shape is thought to be crucial for orchestrating consecutive cycles of substrate ubiquitination [76]. Considering the similar architectures of the SCF and PP2A enzyme complexes, we hypothesise that there is a common underlying mechanism exploiting flexible repeat-protein scaffolds for such catalytic processes. At this point it is not clear how exactly this scaffold flexibility arises and how it depends on the repeat types. Moreover, it remains to be seen whether repeat types with different packing interactions, interfacial energies and cooperativities will exhibit correspondingly different dynamics and macromolecular flexibility.

Bridging tandem-repeat cooperativity and allosteric transmission

In summary, we have explored the relationship between the stability and cooperativity of repeat arrays and the functional transmission of information along them. For example, Rap proteins are capable of transforming a high concentration of quorum-sensing peptides into a signalling response to downstream effectors [61,65,77,78]. This implies that Rap proteins display higher affinity for their binding partners than for the peptides since the bound conformation is only favoured at the high concentrations associated with quorum-sensing.

The Rap proteins are a nice example of a system where nearest-neighbour interactions in a repeat array can cause allosteric inactivation. Our ENM results showed that the first five normal modes of the extended conformation could account for most of the conformational changes between the extended and compact form, suggesting an equilibrium between the two that favours the extended form in the absence of the peptide. Due to different packing interactions in the extended and compact conformations, the N-terminal domain displays a varying degree of correlated movement relative to the TPR domain. This observation supports the idea of an N-terminal helix bundle reaching an energy minimum when cooperatively interacting with the compact TPR domain and thus becoming incapable of exploring partner-binding configurations.

In addition to these insights from the Rap ENMs, the examples of banana-shaped proteins discussed here suggest that repeat arrays involved in diverse cellular processes have the potential to function as allosteric modulators in multi-protein complexes and are not simply a molecular-recognition platform for multiple binding partners. In most cases, these repeat proteins are not rigid, rod-like entities, but rather they need to be flexible to function in a biological context. ENMs are a simplistic but efficient way for us to gain insights into the conformational space explored by repeat proteins. Using them, we can identify structural points of allosteric significance, such as hinges or weak points, and design experiments accordingly. Naturally, we think that repeat stability and cooperativity can be linked to distinct mechanical characteristics and therefore function as a transmission pathway for information to travel through the repeat array. Any local event, such as binding of a partner molecule and alterations of repeat packing in TPR arrays, should therefore modulate repeat stability and shape, and this change could be transmitted to nearest-neighbours by way of the interaction potential, similar to the mechanism that gives rise to spin-waves in ferromagnets. We therefore suggest that context-dependent changes in cooperativity between repeats must,

at least partially, be the basis for allosteric effects in tandem repeat proteins, and as such any repeat protein in itself could function as a switch.

Ultimately, the question of how distantly located repeats can fold cooperatively, how Rap proteins change their superhelical structure upon binding, and how information is transmitted through multi-protein complexes via the repeat protein, may be different manifestations of the same physical mechanism, namely that underlying the Ising Model. The two parameters of intrinsic repeat stability and the interaction potential (i.e. interfacial stability) are straightforward to quantify in consensus repeat arrays but are not easily determined in natural repeat proteins due to the different sequences of the repeats. Nevertheless, we believe that this parameterisation will still hold true but will just result in a model that is mathematically non-trivial. It is crucial to carefully dissect the relationship between repeat protein cooperativity and their ability to function as switches such that we can tune them artificially, thereby translating the peptide-sensing capability to the biotech industry. Last but not least, repeat proteins make up nearly one-third of the human proteome [79], and, given their widespread involvement in key signalling cascades, an understanding of allostery in repeat proteins is also necessary to shed light on the transmission of information in central cellular processes.

Competing interests

The authors declare no competing financial interests.

Authors' contributions

AP, MS and LSI conceived the opinion piece. MS carried out the ENMs. AP, MS and LSI wrote the manuscript.

Acknowledgements

LSI acknowledges the support of a Senior Fellowship from the UK Medical Research Foundation. AP was supported by a BBSRC Doctoral Training Programme scholarship and an Oliver Gatty Studentship. MS was supported by a BBSRC Doctoral Training Programme scholarship.

References

1. Monod J, Jacob F. 1961 Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 389–401.

2. Motlagh HN, Wrabl JO, Li J, Hilser VJ. 2014 The ensemble nature of allostery. *Nature* **508**, 331–339. (doi:10.1038/nature13001)

3. Cui Q, Karplus M. 2008 Allostery and cooperativity revisited. *Protein Sci.* **17**, 1295–1307. (doi:10.1110/ps.03259908)

4. Goodey NM, Benkovic SJ. 2008 Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **4**, 474–482. (doi:10.1038/nchembio.98)

5. Tsai C-J, Nussinov R. 2014 A Unified View of 'How Allostery Works'. *PLoS Comput. Biol.* **10**, e1003394. (doi:10.1371/journal.pcbi.1003394)

6. Kajava A V. 2001 Review: proteins with repeated sequence structural prediction and modeling. *J. Struct. Biol.* **134**, 132–44. (doi:10.1006/jsbi.2000.4328)

7. Kajava A V. 2012 Tandem repeats in proteins: from sequence to structure. *J. Struct. Biol.* **179**, 279–88. (doi:10.1016/j.jsb.2011.08.009)

8. Buard J, Vergnaud G. 1994 Complex recombination events at the hypermutable minisatellite CEB1 (D2S90). *EMBO J.* **13**, 3203–10.

9. Ellegren H. 2000 Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* **16**, 551–558. (doi:10.1016/S0168-9525(00)02139-9)

10. Kobe B, Kajava A V. 2000 When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends Biochem. Sci.* **25**, 509–15. (doi:10.1016/S0968-0004(00)01667-4)

11. Kajava A V. 2002 What curves alpha-solenoids? Evidence for an alpha-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome. *J. Biol. Chem.* **277**, 49791–8. (doi:10.1074/jbc.M204982200)

12. Bublitz M, Holland C, Sabet C, Reichelt J, Cossart P, Heinz DW, Bierne H, Schubert W-D. 2008 Crystal structure and standardized geometric analysis of InlJ, a listerial virulence factor and leucine-rich repeat protein with a novel cysteine ladder. *J. Mol. Biol.* **378**, 87–96.

13. Forwood JK, Lange A, Zachariae U, Marfori M, Preast C, Grubmüller H, Stewart M, Corbett AH, Kobe B. 2010 Quantitative Structural Analysis of importin-β Flexibility: Paradigm for Solenoid Protein Structures. *Structure* **18**, 1171–1183.

14. Filipovska A, Rackham O. 2012 Modular recognition of nucleic acids by PUF, TALE and PPR proteins. *Mol. Biosyst.* **8**, 699–708. (doi:10.1039/c2mb05392f)

15. Liu S, Melonek J, Boykin LM, Small I, Howell KA. 2013 PPR-SMRs: ancient proteins with enigmatic functions. *RNA Biol.* **10**, 1501–10. (doi:10.4161/rna.26172)

16. Yin P *et al.* 2013 Structural basis for the modular recognition of single-stranded RNA

by PPR proteins. *Nature* **504**, 168–71. (doi:10.1038/nature12651)

17. Barkan A, Small I. 2014 Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **65**, 415–42. (doi:10.1146/annurev-arplant-050213-040159)

18. Ke J *et al.* 2013 Structural basis for RNA recognition by a dimeric PPR-protein complex. *Nat. Struct. Mol. Biol.* **20**, 1377–82. (doi:10.1038/nsmb.2710)

19. Tewari R, Bailes E, Bunting KA, Coates JC. 2010 Armadillo-repeat protein functions: questions for little creatures. *Trends Cell Biol.* **20**, 470–81. (doi:10.1016/j.tcb.2010.05.003)

20. Andrade MA, Petosa C, O'Donoghue SI, Müller CW, Bork P. 2001 Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1–18. (doi:10.1006/jmbi.2001.4624)

21. Andrade MA, Perez-Iratxeta C, Ponting CP. 2001 Protein repeats: structures, functions, and evolution. *J. Struct. Biol.* **134**, 117–31. (doi:10.1006/jsbi.2001.4392)

22. Allan RK, Ratajczak T. 2011 Versatile TPR domains accommodate different modes of target protein recognition and function. *Cell Stress Chaperones* **16**, 353–67. (doi:10.1007/s12192-010-0248-0)

23. Zweifel ME, Leahy DJ, Hughson FM, Barrick D. 2003 Structure and stability of the ankyrin domain of the Drosophila Notch receptor. *Protein Sci.* **12**, 2622–32. (doi:10.1110/ps.03279003)

24. Zhang B, Peng Z. 1996 Defective folding of mutant p16(INK4) proteins encoded by tumor-derived alleles. *J. Biol. Chem.* **271**, 28734–7. (doi:10.1074/JBC.271.46.28734)

25. Lowe AR, Itzhaki LS. 2007 Biophysical characterisation of the small ankyrin repeat protein myotrophin. *J. Mol. Biol.* **365**, 1245–55. (doi:10.1016/j.jmb.2006.10.060)

26. DeVries I, Ferreiro DU, Sánchez IE, Komives EA. 2011 Folding Kinetics of the Cooperatively Folded Subdomain of the IκBα Ankyrin Repeat Domain. *J. Mol. Biol.* **408**, 163–176. (doi:10.1016/j.jmb.2011.02.021)

27. Hutton RD, Wilkinson J, Faccin M, Sivertsson EM, Pelizzola A, Lowe AR, Bruscolini P, Itzhaki LS. 2015 Mapping the Topography of a Protein Energy Landscape. *J. Am. Chem. Soc.* **137**, 14610–25. (doi:10.1021/jacs.5b07370)

28. Lowe AR, Itzhaki LS. 2007 Rational redesign of the folding pathway of a modular protein. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2679–84. (doi:10.1073/pnas.0604653104)

29. Tang KS, Fersht AR, Itzhaki LS. 2003 Sequential unfolding of ankyrin repeats in tumor suppressor p16. *Structure* **11**, 67–73.

30. Werbeck ND, Rowling PJE, Chellamuthu VR, Itzhaki LS. 2008 Shifting transition states in the unfolding of a large ankyrin repeat protein. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9982–7. (doi:10.1073/pnas.0705300105)

31. Tripp KW, Barrick D. 2008 Rerouting the Folding Pathway of the Notch Ankyrin Domain by Reshaping the Energy Landscape. *J. Am. Chem. Soc.* **130**, 5681–5688. (doi:10.1021/ja0763201)

32. Sivertsson EM, Itzhaki LS. 2014 Protein folding: When ribosomes pick the structure. *Nat. Chem.* **6**, 378–379. (doi:10.1038/nchem.1926)

33. Ferreiro DU, Cho SS, Komives EA, Wolynes PG. 2005 The Energy Landscape of Modular Repeat Proteins: Topology Determines Folding Mechanism in the Ankyrin Family. *J. Mol. Biol.* **354**, 679–692. (doi:10.1016/j.jmb.2005.09.078)

34. Werbeck ND, Itzhaki LS. 2007 Probing a moving target with a plastic unfolding intermediate of an ankyrin-repeat protein. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 7863–8. (doi:10.1073/pnas.0610315104)

35. Street TO, Bradley CM, Barrick D. 2007 Predicting coupling limits from an experimentally determined energy landscape. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 4907–12. (doi:10.1073/pnas.0608756104)

36. Truhlar SME, Mathes E, Cervantes CF, Ghosh G, Komives EA. 2008 Pre-folding IκBα Alters Control of NF-κB Signaling. *J. Mol. Biol.* **380**, 67–82. (doi:10.1016/j.jmb.2008.02.053)

37. Grinthal A, Adamovic I, Weiner B, Karplus M, Kleckner N. 2010 PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. *Proc. Natl. Acad. Sci.* **107**, 2467–2472.

38. Tsytlonok M, Craig PO, Sivertsson E, Serquera D, Perrett S, Best RB, Wolynes PG, Itzhaki LS. 2013 Complex Energy Landscape of a Giant Repeat Protein. *Structure* **21**, 1954–1965. (doi:10.1016/j.str.2013.08.028)

39. Kajander T, Cortajarena AL, Main ERG, Mochrie SGJ, Regan L. 2005 A new folding paradigm for repeat proteins. *J. Am. Chem. Soc.* **127**, 10188–90. (doi:10.1021/ja0524494)

40. Aksel T, Barrick D. 2009 *Biothermodynamics, Part A*. Elsevier. (doi:10.1016/S0076-6879(08)04204-3)

41. Aksel T, Majumdar A, Barrick D. 2011 The contribution of entropy, enthalpy, and hydrophobic desolvation to cooperativity in repeat-protein folding. *Structure* **19**, 349–60. (doi:10.1016/j.str.2010.12.018)

42. Millership C, Phillips JJ, Main ERG. 2016 Ising Model Reprogramming of a Repeat Protein's Equilibrium Unfolding Pathway. *J. Mol. Biol.* **428**, 1804–17. (doi:10.1016/j.jmb.2016.02.022)

43. Bowley R, Sánchez M. 1999 *Introductory statistical mechanics*. Clarendon Press.

44. Kittel C. 2005 *Introduction to solid state physics*. 8th edn. John Wiley & Sons, Inc.

45. Cortajarena AL, Mochrie SGJ, Regan L. 2011 Modulating repeat protein stability: the effect of individual helix stability on the collective behavior of the ensemble. *Protein Sci.* **20**, 1042–7. (doi:10.1002/pro.638)

46. Aksel T, Barrick D. 2014 Direct observation of parallel folding pathways revealed using a symmetric repeat protein system. *Biophys. J.* **107**, 220–32. (doi:10.1016/j.bpj.2014.04.058)

47. Kloss E, Barrick D. 2009 C-terminal deletion of leucine-rich repeats from YopM reveals a heterogeneous distribution of stability in a cooperatively folded protein. *Protein Sci.* **18**, 1948–1960. (doi:10.1002/pro.205)

48. Löw C, Weininger U, Zeeb M, Zhang W, Laue ED, Schmid FX, Balbach J. 2007 Folding Mechanism of an Ankyrin Repeat Protein: Scaffold and Active Site Formation of Human CDK Inhibitor p19INK4d. *J. Mol. Biol.* **373**, 219–231. (doi:10.1016/j.jmb.2007.07.063)

49. Low C, Weininger U, Neumann P, Klepsch M, Lilie H, Stubbs MT, Balbach J. 2008 Structural insights into an equilibrium folding intermediate of an archaeal ankyrin repeat protein. *Proc. Natl. Acad. Sci.* **105**, 3779–3784. (doi:10.1073/pnas.0710657105)

50. Kloss E, Courtemanche N, Barrick D. 2008 Repeat-protein folding: new insights into origins of cooperativity, stability, and topology. *Arch. Biochem. Biophys.* **469**, 83–99. (doi:10.1016/j.abb.2007.08.034)

51. Courtemanche N, Barrick D. 2008 The leucine-rich repeat domain of Internalin B folds along a polarized N-terminal pathway. *Structure* **16**, 705–14. (doi:10.1016/j.str.2008.02.015)

52. Wetzel SK, Ewald C, Settanni G, Jurt S, Plückthun A, Zerbe O. 2010 Residue-resolved stability of full-consensus ankyrin repeat proteins probed by NMR. *J. Mol. Biol.* **402**, 241–58. (doi:10.1016/j.jmb.2010.07.031)

53. Main ERG, Stott K, Jackson SE, Regan L. 2005 Local and long-range stability in tandemly arrayed tetratricopeptide repeats. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5721–6. (doi:10.1073/pnas.0404530102)

54. Cortajarena AL, Mochrie SGJ, Regan L. 2008 Mapping the energy landscape of repeat proteins using NMR-detected hydrogen exchange. *J. Mol. Biol.* **379**, 617–26. (doi:10.1016/j.jmb.2008.02.046)

55. Tsytlonok M, Itzhaki LS. 2012 The how's and why's of protein folding intermediates. *Arch. Biochem. Biophys.* (doi:10.1016/j.abb.2012.10.006)

56. Schilling J, Schöppe J, Plückthun A. 2014 From DARPins to LoopDARPins: novel LoopDARPin design allows the selection of low picomolar binders in a single round of ribosome display. *J. Mol. Biol.* **426**, 691–721. (doi:10.1016/j.jmb.2013.10.026)

57. MacDonald JT, Kabasakal B V, Godding D, Kraatz S, Henderson L, Barber J, Freemont PS, Murray JW. 2016 Synthetic beta-solenoid proteins with the fragment-free computational design of a beta-hairpin extension. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10346–51. (doi:10.1073/pnas.1525308113)

58. Do H, Kumaraswami M. 2016 Structural Mechanisms of Peptide Recognition and Allosteric Modulation of Gene Regulation by the RRNPP Family of Quorum-Sensing Regulators. *J. Mol. Biol.* **428**, 2793–2804. (doi:10.1016/j.jmb.2016.05.026)

59. Perego M, Hoch JA. 1996 Cell-cell communication regulates the effects of protein aspartate phosphatases on the phosphorelay controlling development in Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1549–53.

60. Tzeng Y-L, Feher VA, Cavanagh J, Perego M, Hoch JA. 1998 Characterization of

Interactions between a Two-Component Response Regulator, Spo0F, and Its Phosphatase, RapB †. *Biochemistry* **37**, 16538–16545. (doi:10.1021/bi981340o)

61. Perego M. 2001 A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis. *Mol. Microbiol.* **42**, 133–43.

62. Core LJ, Ishikawa S, Perego M. 2001 A free terminal carboxylate group is required for PhrA pentapeptide inhibition of RapA phosphatase. *Peptides* **22**, 1549–53.

63. Bongiorni C, Stoessel R, Perego M. 2007 Negative Regulation of Bacillus anthracis Sporulation by the Spo0E Family of Phosphatases. *J. Bacteriol.* **189**, 2637–2645. (doi:10.1128/JB.01798-06)

64. Diaz AR, Core LJ, Jiang M, Morelli M, Chiang CH, Szurmant H, Perego M. 2012 Bacillus subtilis RapA Phosphatase Domain Interaction with Its Substrate, Phosphorylated Spo0F, and Its Inhibitor, the PhrA Peptide. *J. Bacteriol.* **194**, 1378–1388. (doi:10.1128/JB.06747-11)

65. Baker MD, Neiditch MB. 2011 Structural Basis of Response Regulator Inhibition by a Bacterial Anti-Activator Protein. *PLoS Biol.* **9**, e1001226. (doi:10.1371/journal.pbio.1001226)

66. Parashar V, Jeffrey PD, Neiditch MB. 2013 Conformational Change-Induced Repeat Domain Expansion Regulates Rap Phosphatase Quorum-Sensing Signal Receptors. *PLoS Biol.* **11**, e1001512. (doi:10.1371/journal.pbio.1001512)

67. Bahar I, Lezon TR, Yang L-W, Eyal E. 2010 Global Dynamics of Proteins: Bridging Between Structure and Function. *Annu. Rev. Biophys.* **39**, 23–42.

68. Rodgers TL, Burnell D, Townsend PD, Pohl E, Cann MJ, Wilson MR, McLeish TCB. 2013 ΔΔPT: a comprehensive toolbox for the analysis of protein motion. *BMC Bioinformatics* **14**, 1–9.

69. Bakan A, Meireles LM, Bahar I. 2011 ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **27**, 1575–1577.

70. Cho US, Xu W. 2007 Crystal structure of a protein phosphatase 2A heheterotrimer holoenzyme. *Nature* **445**, 53–57.

71. Xu Y, Chen Y, Zhang P, Jeffrey PD, Shi Y. 2008 Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated tau dephosphorylation. *Mol. Cell* **31**, 873–885.

72. Wlodarchak N, Guo F, Satyshur KA, Jiang L, Jeffrey PD, Sun T, Stanevich V, Mumby MC, Xing Y. 2013 Structure of the Ca2+-dependent PP2A heterotrimer and insights into Cdc6 dephosphorylation. *Cell Res.* **23**, 931–946.

73. Chen Y, Xu Y, Bao Q, Xing Y, Li Z, Lin Z, Stock JB, Jeffrey PD, Shi Y. 2007 Structural and biochemical insights into the regulation of protein phosphatase 2A by small t antigen of SV40. *Nat. Struct. Mol. Biol.* **14**, 527–534.

74. Yao Z-P, Zhou M, Kelly SE, Seeliger MA, Robinson C V, Itzhaki LS. 2006 Activation of Ubiquitin Ligase SCF(Skp2) by Cks1: Insights from Hydrogen Exchange Mass

Spectrometry. *J. Mol. Biol.* **363**, 673–686.

75. Hao B, Zheng N, Schulman BA, Wu G, Miller JJ, Pagano M, Pavletich NP. 2005 Structural Basis of the Cks1-Dependent Recognition of p27Kip1 by the SCFSkp2 Ubiquitin Ligase. *Mol. Cell* **20**, 9–19.

76. Liu J, Nussinov R. 2011 Flexible cullins in cullin-RING E3 ligases allosterically regulate ubiquitination. *J. Biol. Chem.* **286**, 40934–40942.

77. Parashar V, Mirouze N, Dubnau DA, Neiditch MB. 2011 Structural Basis of Response Regulator Dephosphorylation by Rap Phosphatases. *PLoS Biol.* **9**, e1000589. (doi:10.1371/journal.pbio.1000589)

78. Mirouze N, Parashar V, Baker MD, Dubnau DA, Neiditch MB. 2011 An Atypical Phr Peptide Regulates the Developmental Switch Protein RapH. *J. Bacteriol.* **193**, 6197–6206. (doi:10.1128/JB.05860-11)

79. Pellegrini M, Marcotte EM, Yeates TO. 1999 A fast algorithm for genome-wide analysis of proteins with repeated sequences. *Proteins Struct. Funct. Bioinforma.* **35**, 440–446.

Figure 1. Schematic representation of the structures of solenoid tandem-repeat proteins. From top left to bottom right: ARM-repeat proteinβ-catenin (2z6h), HEAT-repeat protein PR65 (1b3u), Ankyrin-repeat protein gankyrin (1uoh), HEAT-repeat protein Importin-β (3nd2), leucine-rich repeat (LRR) protein Ribonuclease Inhibitor (1bnh), TPR protein RapI (4i1a), beta-helical repeat protein carbonic anhydrase (1qre).

Figure 2. Structures of different Rap proteins (C-terminus in red) depicting a possible mode of action. When the TPR domain binds to a signalling peptide, it causes the Rap protein to adopt a compact, or "closed" conformation. Upon binding an interaction partner, however, conformational changes in the TPR domain are minimal, whereas the N-terminal 3-helix bundle flips by approximately 180° [66].

Figure 3. The three lowest normal modes of RapI. The first mode (teal) describes a bending motion that alters the distance between the N- and C-termini; the second mode (magenta) tightens the superhelix in a screw-like motion; the third mode (red) twists the N-terminal 3-helix bundle and the C-terminal TPR relative to the repeat array superhelix. Models were generated using the NMWIZ plugin for VMD [69]. Movies of these modes can be found with the Supplementary Information online.

Figure 4. Quantitative comparison between the lowest five normal modes and conformational changes between different Rap proteins. The arrows represent the conformational change vector, and the values equal the corresponding cumulative overlap between the vector and the ENM of the starting structure (see SI methods). For example, the first five normal modes of RapI can account for 0.87 of the conformational change between RapI and RapJ, while the first five normal modes of RapJ only describe 0.75. Numbers in brackets correspond to the cumulative overlaps between the dynamics of truncated and independently modelled TPR domains and their respective conformational changes (see SI methods).

Figure 5. Entropy contributions of each normal mode to the total motion. The closed conformation was modelled both with and without the PhrC peptide. As the differences between both models are only small (Fig. S5), the effect of peptide binding on the entropy of the system is negligible compared to the entropic cost of the conformational change.

Figure 6. Representative cross-correlation maps for the partner-bound, open and peptide-bound conformations. Cross-correlation between residues is a measure of how much these residues move in the same direction, where values of 1 and -1 represent perfectly correlated and anti-correlated motions, respectively [67]. The TPR repeats exhibit correlated motions only with their nearest neighbours, giving rise to the distinctive pattern of squares along the diagonal. Movements of the rotated N-terminal 3-helix bundle, linker domain and first TPR motif (blue box) are non-TPR-like, exhibiting non-nearest neighbour correlations, suggesting that they form a subdomain relative to the rest of the TPR repeats. Some of these correlations are reduced in the open conformation, or even reversed, once a continuous TPR array is formed (arrows) and the distinction of this domain is lost. The global movement of peptide binding TPRs (purple box) and neighbouring repeats is only minimally affected in the presence of the peptide, which only causes a slight increase of the nearest-neighbour correlations. The N-terminal helix bundle and TPR repeats are divided by grey dashed lines and correlations are mirrored across the diagonal for clarity. The cross-correlation was summed over the lowest 25 modes, for correlation maps of lower modes see Figure S6.

Figure 7. Repeat proteins linking function in multi-protein complexes. (A) PP2A (PDB 3dw8) consisting of the HEAT-repeat scaffold subunit PR65, the catalytic subunit (C) bound to the C-terminus of PR65, and a regulatory subunit (B55) bound to the N-terminus of PR65. (B) Modelled structure of the SCF^{Skp2} ubiquitin ligase consisting of three core subunits Skp1, cullin-repeat containing protein Cul1 and Rbx1, and substrate-recognition subunit Skp2 with accessory protein Cks1, which together recruit the substrate p27. Also shown is the E2 ubiquitin-conjugating enzyme, which is recruited to the SCF by Rbx1, together forming the catalytic entity (PDBs 2AST, 1LDK and 4Q5E [75]). Thus, in both complexes the substrate-recognition subunit is bound to one end of the repeat protein and the catalytic subunit to the other end. (C) Top view of the Skp2 bound to Skp1, Cks1 and p27, highlighting the insertion of the Skp2 C-terminal tail at its N-terminus.

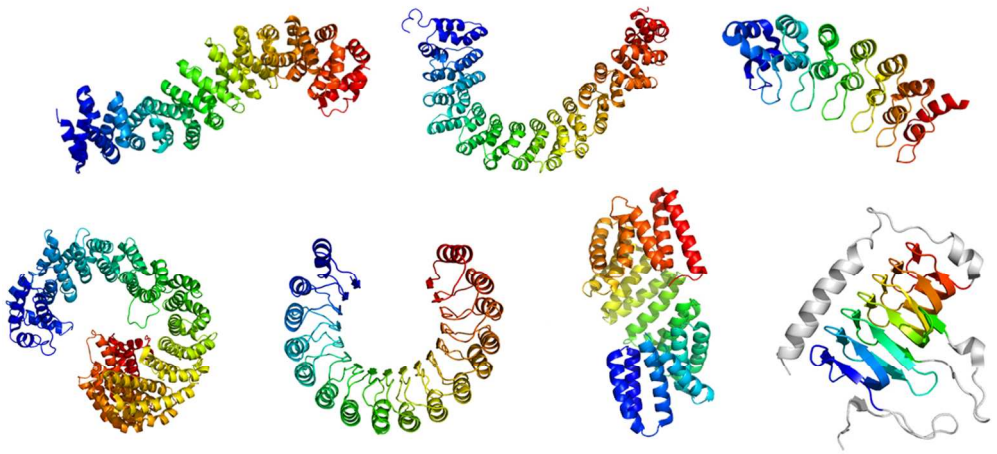

Figure 1. Schematic representation of the structures of solenoid tandem-repeat proteins. From top left to bottom right: ARM-repeat proteinβ-catenin (2z6h), HEAT-repeat protein PR65 (1b3u), Ankyrin-repeat protein gankyrin (1uoh), HEAT-repeat protein Importin-β (3nd2), leucine-rich repeat (LRR) protein Ribonuclease Inhibitor (1bnh), TPR protein RapI (4i1a), beta-helical repeat protein carbonic anhydrase (1qre).

326x150mm (95 x 95 DPI)

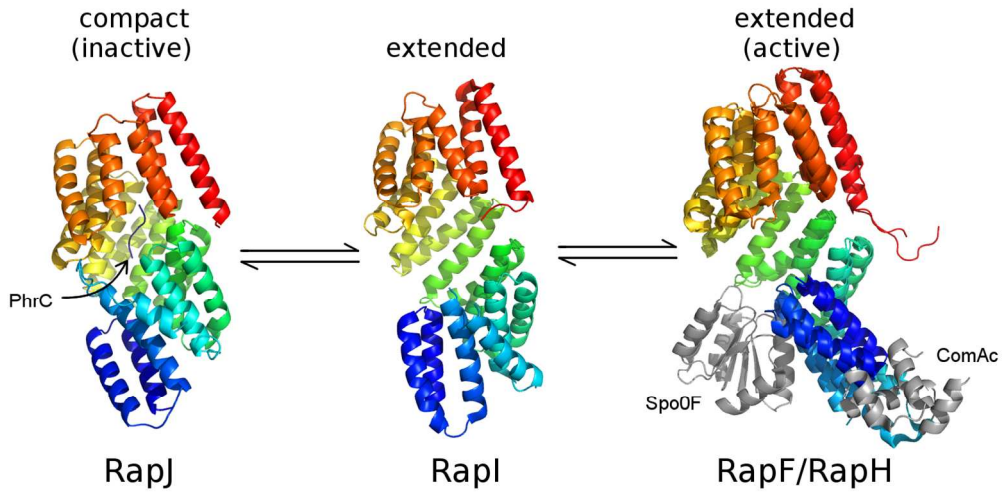

Figure 2. Structures of different Rap proteins (C-terminus in red) depicting a possible mode of action. When the TPR domain binds to a signalling peptide, it causes the Rap protein to adopt a compact, or "closed" conformation. Upon binding an interaction partner, however, conformational changes in the TPR domain are minimal, whereas the N-terminal 3-helix bundle flips by approximately 180° [66].

553x273mm (72 x 72 DPI)

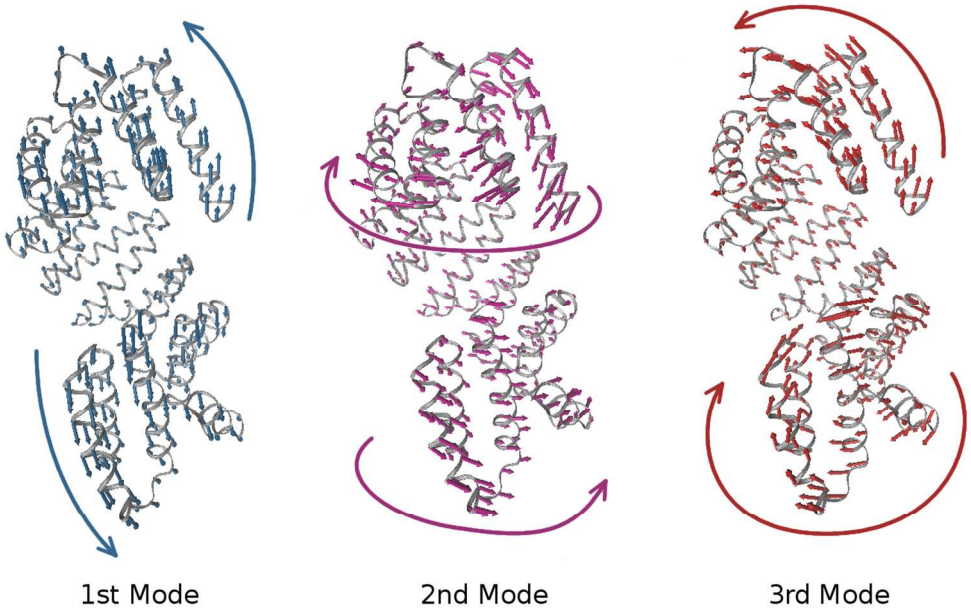

Figure 3. The three lowest normal modes of RapI. The first mode (teal) describes a bending motion that alters the distance between the N- and C-termini; the second mode (magenta) tightens the superhelix in a screw-like motion; the third mode (red) twists the N-terminal 3-helix bundle and the C-terminal TPR relative to the repeat array superhelix. Models were generated using the NMWIZ plugin for VMD [69]. Movies of these modes can be found with the Supplementary Information online.

473x294mm (72 x 72 DPI)

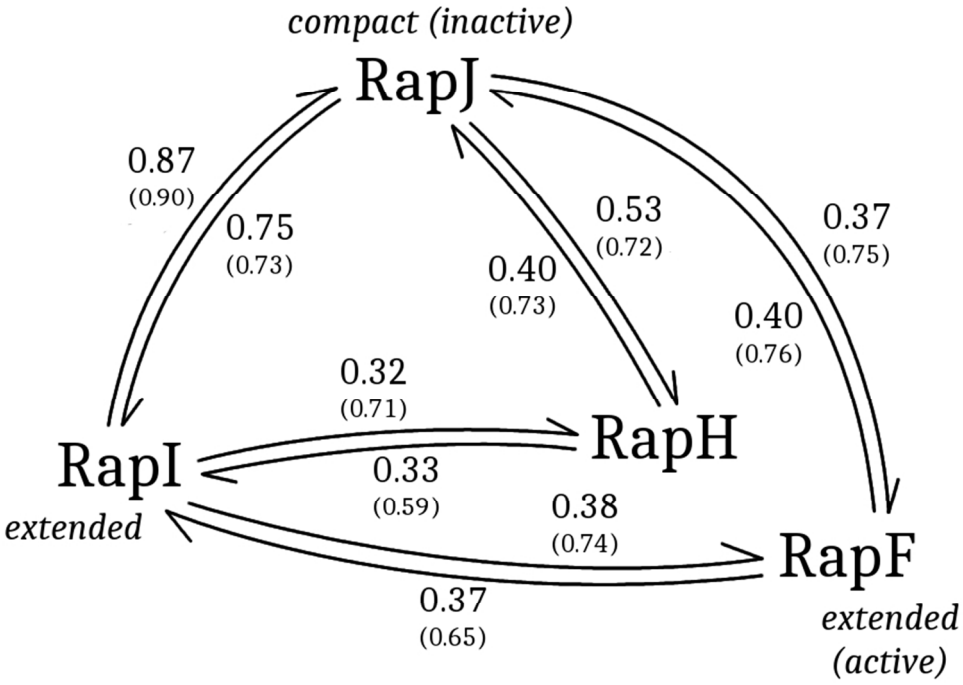

Figure 4. Quantitative comparison between the lowest five normal modes and conformational changes between different Rap proteins. The arrows represent the conformational change vector, and the values equal the corresponding cumulative overlap between the vector and the ENM of the starting structure (see SI methods). For example, the first five normal modes of RapI can account for 0.87 of the conformational change between RapI and RapJ, while the first five normal modes of RapJ only describe 0.75. Numbers in brackets correspond to the cumulative overlaps between the dynamics of truncated and independently modelled TPR domains and their respective conformational changes (see SI methods).

334x239mm (72 x 72 DPI)

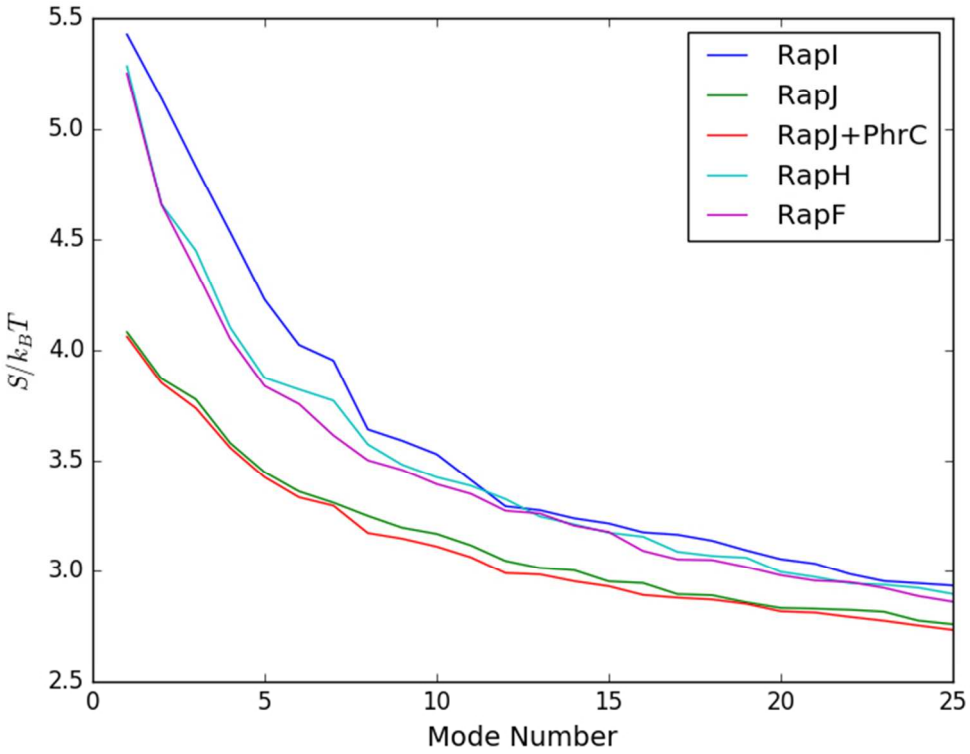

Figure 5. Entropy contributions of each normal mode to the total motion. The closed conformation was modelled both with and without the PhrC peptide. As the differences between both models are only small (Fig. S5), the effect of peptide binding on the entropy of the system is negligible compared to the entropic cost of the conformational change.

185x141mm (99 x 99 DPI)

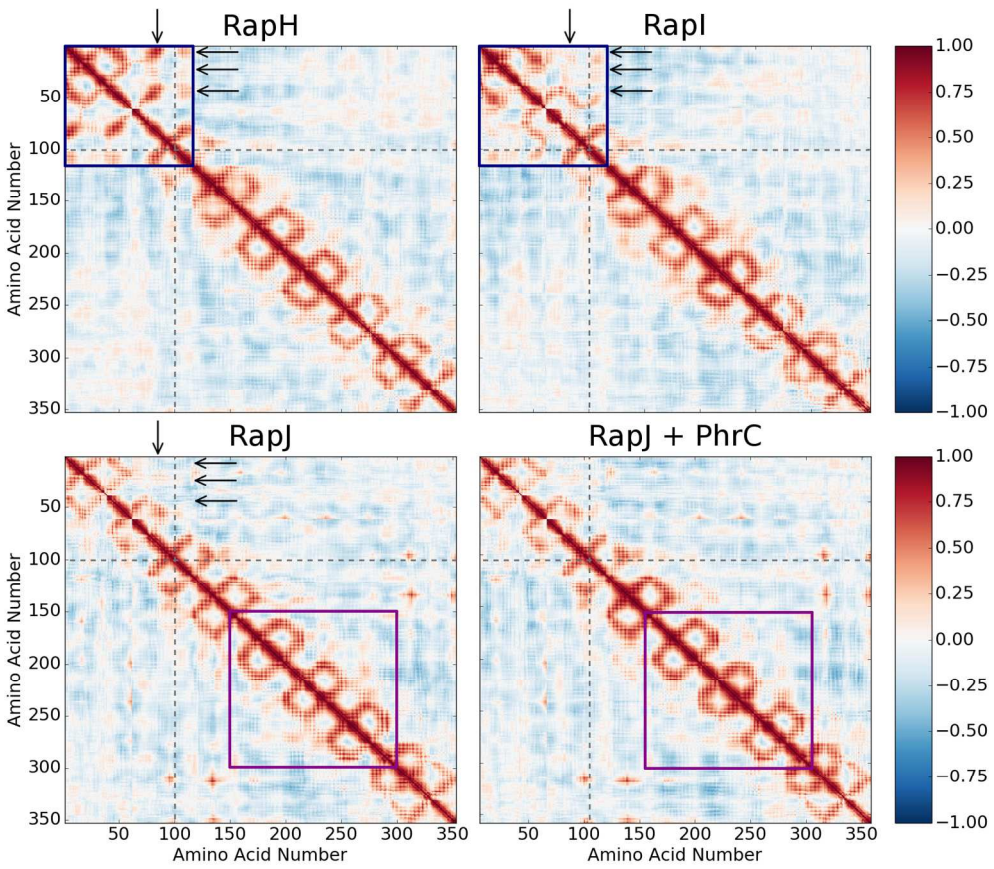

Figure 6. Representative cross-correlation maps for the partner-bound, open and peptide-bound conformations. Cross-correlation between residues is a measure of how much these residues move in the same direction, where values of 1 and -1 represent perfectly correlated and anti-correlated motions, respectively [67]. The TPR repeats exhibit correlated motions only with their nearest neighbours, giving rise to the distinctive pattern of squares along the diagonal. Movements of the rotated N-terminal 3-helix bundle, linker domain and first TPR motif (blue box) are non-TPR-like, exhibiting non-nearest neighbour correlations, suggesting that they form a subdomain relative to the rest of the TPR repeats. Some of these correlations are reduced in the open conformation, or even reversed, once a continuous TPR array is formed (arrows) and the distinction of this domain is lost. The global movement of peptide binding TPRs (purple box) and neighbouring repeats is only minimally affected in the presence of the peptide, which only causes a slight increase of the nearest-neighbour correlations. The N-terminal helix bundle and TPR repeats are divided by grey dashed lines and correlations are mirrored across the diagonal for clarity. The cross-correlation was summed over the lowest 25 modes, for correlation maps of lower modes see Figure S6.

447x388mm (99 x 99 DPI)

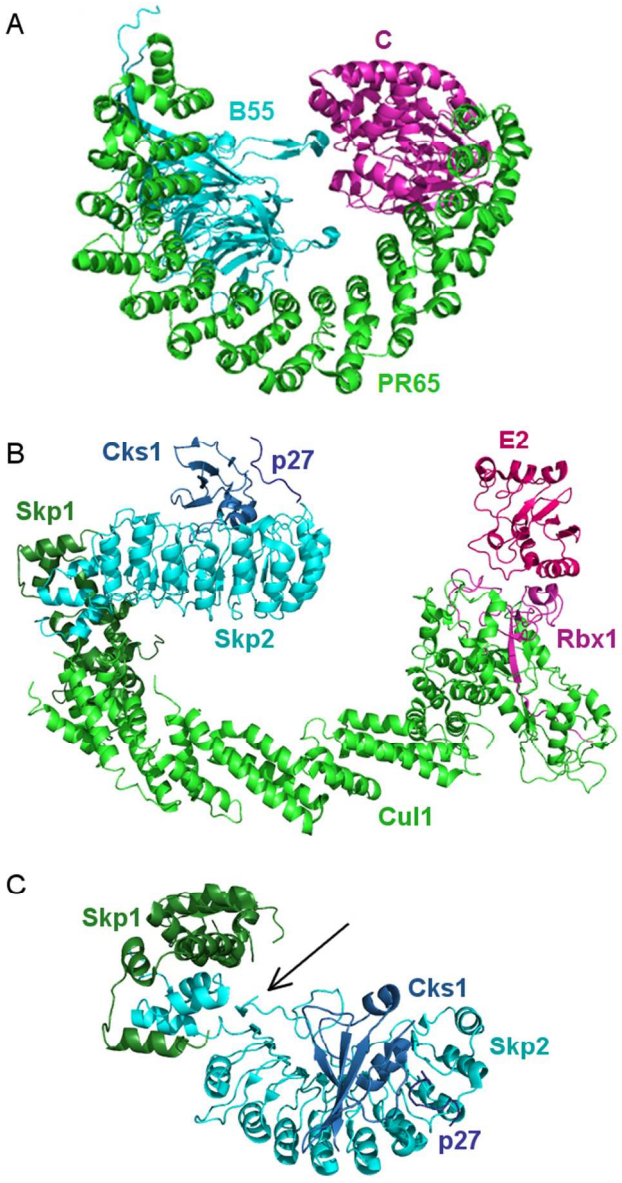

Figure 7. Repeat proteins linking function in multi-protein complexes. (A) PP2A (PDB 3dw8) consisting of the HEAT-repeat scaffold subunit PR65, the catalytic subunit (C) bound to the C-terminus of PR65, and a regulatory subunit (B55) bound to the N-terminus of PR65. (B) Modelled structure of the SCFSkp2 ubiquitin ligase consisting of three core subunits Skp1, cullin-repeat containing protein Cul1 and Rbx1, and substrate-recognition subunit Skp2 with accessory protein Cks1, which together recruit the substrate p27. Also shown is the E2 ubiquitin-conjugating enzyme, which is recruited to the SCF by Rbx1, together forming the catalytic entity (PDBs 2AST, 1LDK and 4Q5E [75]). Thus, in both complexes the substrate-recognition subunit is bound to one end of the repeat protein and the catalytic subunit to the other end. (C) Top view of the Skp2 bound to Skp1, Cks1 and p27, highlighting the insertion of the Skp2 C-terminal tail at its N-terminus.

183x344mm (95 x 95 DPI)