several ways. It is thus reasonable to hope that some compounds that inhibit other exchange factors for small GTP-binding proteins in similar ways could be found. The structures also reveal that exchange factors make intimate contacts with the P-loop and the switch II region, two areas that are the sites of oncogenic mutations in Ras. Thus, one could consider screening for drugs that specifically target the complex formed by oncogenic Ras and its GEFs. The structures of DH domains in complex with some Rho family proteins and of the ARF–GDP– Brefeldin A–Sec7 domain complex would also provide a more detailed molecular understanding of the action of GEFs and could help to design more effective drug-screening strategies.

Acknowledgements

We thank Bruno Antonny for his continuous collaboration in this work and Marc Chabre for critical reading of the manuscript. Coordinates of Ras–SOS, RCC1, SOS–CDC25, PIX–DH and Arf–Gea2 were kindly provided by their authors. This work is supported by the Association pour la Recherche contre le Cancer and by Zénéca Pharma.

References

1 Quilliam, L. A., Khosravi-Far, R., Huff, S. Y. and Der, C. J. (1995) *BioEssays* 17, 395–404
2 Klebe, C., Prinz, H., Wittinghofer, A. and Goody, R. S. (1995) *Biochemistry* 34, 12543–12552
3 Lai, C. C., Boguski, M., Broek, D. and Powers, S. (1993) *Mol. Cell. Biol.* 13, 1345–1352
4 Renault, L. *et al.* (1998) *Nature* 392, 97–101
5 Mossessova, E., Gulbis, J. M. and Goldberg, J. (1998) *Cell* 92, 415–423
6 Cherfils, J. *et al.* (1998) *Nature* 392, 101–105
7 Betz, S. F. *et al.* (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 7909–7914
8 Goldberg, J. (1998) *Cell* 95, 237–248
9 Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. and Kuriyan, J. (1998) *Nature* 394, 337–343
10 Soisson, S. M. *et al.* (1998) *Cell* 95, 259–268
11 Liu, X. *et al.* (1998) *Cell* 95, 269–277
12 Aghazadeh, B. *et al.* (1998) *Nat. Struct. Biol.* 5, 1098–1107
13 Sprang, S. R. and Coleman, D. E. (1998) *Cell* 95, 155–158
14 Yu, H. and Schreiber, S. L. (1995) *Nature* 376, 788–791
15 Nuoffer, C., Wu, S. K., Dascher, C. and Balch, W. E. (1997) *Mol. Biol. Cell* 8, 1305–1316
16 Bourne, H. R., Sanders, D. A. and McCormick, F. (1991) *Nature* 349, 117–127
17 Milburn, M. V. *et al.* (1990) *Science* 247, 939–945
18 Pai, E. F. *et al.* (1989) *Nature* 341, 209–214
19 Beraud-Dufour, S. *et al.* (1998) *EMBO J.* 17, 3651–3659
20 Antonny, B., Beraud-Dufour, S., Chardin, P. and Chabre, M. (1997) *Biochemistry* 36, 4675–4684
21 John, J. *et al.* (1990) *Biochemistry* 29, 6058–6065
22 Kawashima, T. *et al.* (1996) *Nature* 379, 511–518
23 Wang, Y. *et al.* (1997) *Nat. Struct. Biol.* 4, 650–656
24 Azuma, Y. *et al.* (1996) *J. Biochem. (Tokyo)* 120, 82–91
25 Cerione, R. A. and Zheng, Y. (1996) *Curr. Opin. Cell Biol.* 8, 216–222
26 Iiri, T., Farfel, Z. and Bourne, H. R. (1998) *Nature* 394, 35–38
27 Wall, M. A. *et al.* (1995) *Cell* 83, 1047–1058
28 Lambright, D. G. *et al.* (1996) *Nature* 379, 311–319
29 Peyroche, A. *et al.* (1999) *Mol. Cell* 3, 275–285
30 van den Berghe, N., Cool, R. H., Horn, G. and Wittinghofer, A. (1997) *Oncogene* 15, 845–850
31 Chardin, P. and McCormick, F. (1999) *Cell* 97, 153–155
32 Amor, J. C., Harrison, D. H., Kahn, R. A. and Ringe, D. (1994) *Nature* 372, 704–708
33 Cherfils, J. *et al.* (1997) *EMBO J.* 16, 5582–5591
34 Polekhina, G. *et al.* (1996) *Structure* 4, 1141–1151

The ankyrin repeat: a diversity of interactions on a common structural framework

Steven G. Sedgwick and Stephen J. Smerdon

The ankyrin repeat is one of the most common protein sequence motifs. Recent X-ray and NMR structures of ankyrin-repeat proteins and their complexes have provided invaluable insights into the molecular basis of the extraordinary variety of biological activities of these molecules. In particular, they have begun to reveal how a large family of structurally related proteins can interact specifically with such a diverse array of macromolecular targets.

THE EVER-EXPANDING protein-sequence database has revealed a 'menagerie' of amino acid sequence motifs that are reiterated in a bewildering variety of proteins. Indeed, proteins are often classified on the basis of the conserved sequence motifs that they contain, even though the function and structure of the motif might be uncertain. In 1987, Breeden and Nasmyth reported a ~33 residue repeating motif in the sequence of two yeast cell-cycle regulators, Swi6p and Cdc10p, and in the Notch and LIN-12 developmental regulators from *Drosophila melanogaster* and *Caenorhabditis elegans*[1]. Subsequently, the discovery of 24 copies of this sequence in the cytoskeletal protein ankyrin[2] led to the naming of this motif as the ankyrin (ANK) repeat. ANK-repeat proteins carry out a wide variety of biological activities and have been detected in organisms ranging from viruses to humans. The motif has now been recognized in >400 proteins[3], including cyclin-dependent kinase (CDK) inhibitors, transcriptional regulators, cytoskeletal organizers, developmental regulators and toxins[3,4]. These molecules are present in the nucleus, cytoplasm and the extracellular milieu. The number of repeats within any one protein is highly variable. The *Drosophila* plutonium protein contains only two repeats[5], but 20 or more are not uncommon. Some ANK-repeat proteins consist solely of ANK repeats; others are multidomain molecules in which ANK repeats are combined with other unrelated structural modules.

As argued earlier[4], the myriad of different roles for ANK-repeat proteins makes a common function such as an enzymatic activity extremely unlikely. However, the role of ANK repeats in mediating protein–protein interactions has been well documented, and their presence is often interpreted as an indicator of a similar function in otherwise uncharacterized systems. There are several examples of protein modules that recognize specific sequence motifs or

S. G. Sedgwick and **S. J. Smerdon** are at the Division of Yeast Genetics and the Division of Protein Structure, respectively, National Institute for Medical Research, The Ridgeway, Mill Hill, London, UK NW7 1AA.
Email: ssedgwi@nimr.mrc.ac.uk and s-smerdo@darwin.nimr.mrc.ac.uk

 PII: S0968-0004(99)01426-7

Table 1. X-ray and NMR structures of ANK-repeat proteins and complexes[a]

Protein[a]	Function	Structure determination	PDB ID[a]	Number of repeats	Ref.
53BP2–p53	?	X-ray	1YCS	4	6
$p16^{INK4a}$	CDK inhibitor	NMR	1A5E	4	10
$p16^{INK4a}$–CDK6		X-ray	1BI7	4	11
$p18^{INK4c}$	CDK inhibitor	X-ray	1IHB	5	8
		NMR			12
$p19^{INK4d}$	CDK inhibitor	NMR	1AP7	5	13
		X-ray	1BD8		14
$p19^{INK4d}$–CDK6		X-ray	1BI8		11
		X-ray	1BLX		15
GABPα–GABPβ–DNA	Transcription factor	X-ray	1AWC	4	9
Myotrophin	Trigger of cardiac hypertrophy?	NMR	1MYO	4	16
IκBα–NFκB	NFκB inhibitor	X-ray	1NF1	6	7
		X-ray	1IKN		17
Swi6	Transcription factor	X-ray	1SW6	5	18

[a]Abbreviations used: ANK, ankryin; CDK, cyclin-dependent kinase; GABP, GA-binding protein; PDB ID, protein database identification code.

protein modifications such as SH3 and SH2 domains. By contrast, ANK proteins do not bind selectively to a single class of protein target. Rather, the diversity of biological roles of ANK proteins is paralleled by the diversity of unrelated proteins with which they interact. Here, we use the recently determined structures of several ANK-repeat molecules and their complexes (Table 1) as a framework to review the structure and function of the ANK repeat.

Structural organization of ANK repeats

The first three-dimensional structure of an ANK-repeat molecule, that of 53BP2 bound to the p53 cell-cycle tumour suppressor[6], was determined almost ten years after the discovery of the motif. This and subsequent structures have shown that the ANK repeat consists of pairs of antiparallel α-helices stacked side by side and connected by a series of intervening β-hairpin motifs (Fig. 1a). The extended β-sheet projects away from the helical pairs almost at right angles to them, resulting in a characteristic L-shaped cross-section (Fig. 1a). This assembled structure has been likened to a cupped hand: the β-hairpins form the fingers and the concave, inner surface of the ankyrin groove, which is made up of solvent-exposed residues from the α-helical bundle, forms the palm[7].

Although some molecules appear to contain only ANK repeats, others contain insertions between the repeats or, less often, within it. The N- and C-termini of consensus sequences[3,4], based on the exon boundaries of the *ankyrin* gene (Fig. 1b; consensus 2), separated the two β-strands of the β-hairpin. This necessarily incorporated the most commonly occurring non-conserved elements within the repeating unit. The structure-based consensus that we propose (Fig. 1b; consensus 1) largely avoids these problems and reflects the architecture of the ANK domain more logically. We therefore consider the ANK-repeat motif to be defined as a β-hairpin–helix–loop–helix ($\beta_2\alpha_2$) structure.

Clearly, specific types of residues are prevalent at any one position within the consensus. However, ANK repeats can sustain a considerable variety of amino acid substitutions[3,4]. Usually, deviations from the consensus are highly conserved between orthologues; this strongly suggests that they are responsible for the observed specificity of ANK-repeat-protein function. More sophisticated homology-search algorithms have now identified ~3000 ANK repeats from ~400 proteins in the non-redundant protein database (see http://coot.embl-heidelberg.de/SMART).

The structural roles of conserved residues in ANK repeats can be divided loosely into those involved in the formation and stabilization of the individual repeating unit and those that form the interface between repeats. Although the sequence shown in Fig. 1b gives no impression of the sequence variation evident in ANK-repeat proteins, it highlights several of the most highly conserved residues that play important structural roles. The β-hairpin is stabilized through main-chain hydrogen-bonding interactions with the side-chain of the aspartate residue at position 1 of the canonical 33-residue ANK-repeat sequence (Fig. 1b, upper). Three glycine residues are conserved at consensus positions 4, 15 and 27. Gly4 terminates β_2 whereas Gly15 and Gly27 terminate the ANK-repeat helices. The characteristic Thr-Pro-Leu-His (TPLH) tetrapeptide motif forms a tight turn and initiates the first α-helix of the ANK repeat. This turn conformation is stabilized through reciprocal hydrogen-bonding interactions between the side-chain and main-chain atoms of Thr6 and His9.

Stacking of repeats relies predominantly on conserved hydrophobic side chains associated with the α-helices. These create non-polar surfaces that are highly complementary in shape and assemble in a lock-and-key fashion to produce the extended helical bundle (Fig. 1c). Packing of multiple repeats orients the β-hairpin motifs side by side, such that β_2 of ANK_n forms a β-sheet hydrogen bond to β_1 of ANK_{n+1}, etc. Stacking is further stabilized by several hydrogen-bonding interactions between conserved polar residues (His9, His16 and Asn31) and main-chain atoms from adjacent ANK-repeat modules. Together, these interactions combine to create a rather non-globular domain that, nevertheless, is sufficiently stable to permit ANK-repeat proteins to function in both intra- and extracellular locations. Conserved non-polar side chains associated with the inner helices of the bundle tend to be smaller in volume than those associated with the outer helices[7,8]. This, in combination with a left-handed twist that is intrinsic to the stacked arrangement of the adjacent helical pairs, causes a distinct curvature towards the ankyrin groove (Fig. 2). These features might be exaggerated to some extent by the presence of additional, non-ANK structural elements in the block (Fig. 2).

The ANK-domain architectures of the current set of X-ray and NMR structures vary considerably (Fig. 2). GA-binding protein β (GABPβ) contains four and a half repeats and exhibits little structural or sequence deviation from the canonical repeating unit[9]. In the three INK4 proteins, $p16^{INK4a}$ (Refs 10,11), $p18^{INK4c}$ (Refs 8,12) and $p19^{INK4d}$ (Refs 11,13–15), the first helix of the second repeat is atypically short and, as in myotrophin[16], the β-hairpins do not all interact with one another to form a continuous β-sheet. In IκBα, short insertions between repeats 1 and 2 and between repeats 3 and 4 form additional helical segments within the connecting linker regions. IκBβ is closely related to IκBα but contains an additional 47-residue insertion between repeats 3 and 4. Similarly, IκBε and the *Drosophila* homologue, Cactus, have 10 and 30 additional residues, respectively, at the same position[7,17]. Swi6 has five repeats but also

contains several additional structural elements[18]. The region between repeats 3 and 4 is particularly striking: ~40 additional residues contribute to an atypically long second helix of the third repeat, a short section of β-sheet, an unstructured loop and two additional short helices. Another extra helix is found in the linker between ANK2 and ANK3. In the cupped-hand analogy, these elements protrude from the back of the knuckles of the ANK-repeat stack. Finally, the loop joining the two helices of ANK2 also has a short helical element. In spite of these structural variations, the regular stacking of the ANK repeats is unimpaired. In fact, side-chains from the additional helical segments contribute to the hydrophobic core of the molecule. The incorporation of additional elements within a conserved structural core in Swi6 exemplifies the structural resilience of the ANK-repeat domain. However, the embellishments of the ANK domain of Swi6 are unlikely to be unique and, doubtless, even more baroque variations on the ANK-repeat domain organization will be found.

Intermolecular interactions

Eight X-ray structures of complexes involving ANK repeats have been determined (Table 1). Despite the overall sequence similarity shared by ANK proteins, the mechanism of binding to specific partner molecules varies considerably. Binding can use much of the available solvent-accessible surface, and contacts are not restricted to any particular secondary structural element within the ANK repeat. Nevertheless, these structures have revealed some common features of molecular interactions. The first of these was suggested by the structure of the 53BP2–p53 complex[6]. The C-terminal 229-residue fragment of 53BP2 used in this study comprises four ANK repeats followed by an Src-homology-3 (SH3) domain. Although the majority of the interactions between this pair of molecules are mediated through the 53BP2 SH3 domain, the β-hairpin of ANK4 of 53BP2 also makes contact with p53. The importance of the β-hairpin elements in binding and recognition is better illustrated by the GABPβ–GABPα–DNA complex (Fig. 3a). GABPα is a transcription factor of the Ets-domain family that binds to DNA in combination with an ANK-repeat accessory factor, GABPβ. GABPβ primarily contacts GABPα through two residues at the tips of all four of its β-hairpin fingers. These residues play no structural role in the

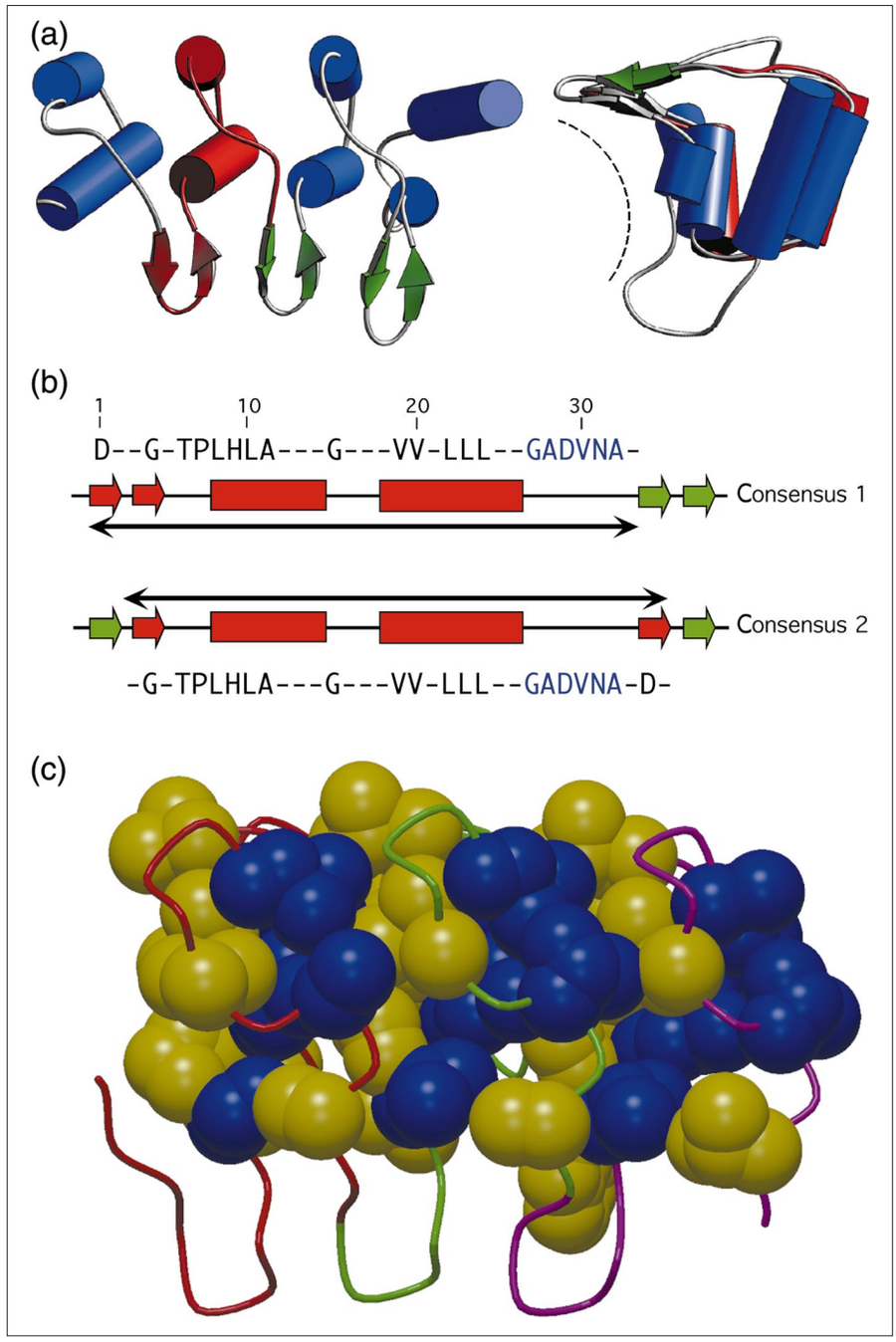

Figure 1

(a) The X-ray structure of 53BP2 shows the arrangement of α-helices (cylinders) and β-hairpins (arrows), which is characteristic of the ankyrin (ANK)-repeat architecture viewed from the 'top' of the domain (left). A single ANK repeat is highlighted in red. The continuous β-sheet projects away from the helical stack to form the ankyrin groove, which is indicated by the dotted arc (right). **(b)** The ANK-repeat consensus. ANK-repeat domains are assembled from multiple, sequential copies of a ~33 residue motif. The structure-based consensus (consensus 1) defines the ANK repeat as a $\beta_2\alpha_2$ motif highlighted in red [compare with (a)]. In many ANK-repeat molecules, sequence insertions (highlighted in blue) occur most often within the region that connects helix 2 to the first β-strand of the adjacent repeat. Many alternative definitions of the N- and C-terminal limits of the repeat have been suggested. The most common is shown in the lower panel (consensus 2) and is based on the location of exon boundaries in the *ankyrin* gene. In both consensus sequences, the extent of a single ANK repeat is indicated by a black arrow. Variations might occur at any position within the consensus. The sequence shown has been chosen only to illustrate some of the structural roles played by some of the most highly conserved residues. **(c)** Assembly of the ANK-repeat domain. The backbone of three consecutive ANK motifs is shown as red, green and purple coils. Non-polar side-chain atoms, shown in a space-filling representation, are located on both the N-terminal face (olive green) and C-terminal face (blue) of the helical pairs. These form complementary surfaces that associate and constitute the core interactions that form the ANK-repeat stack.

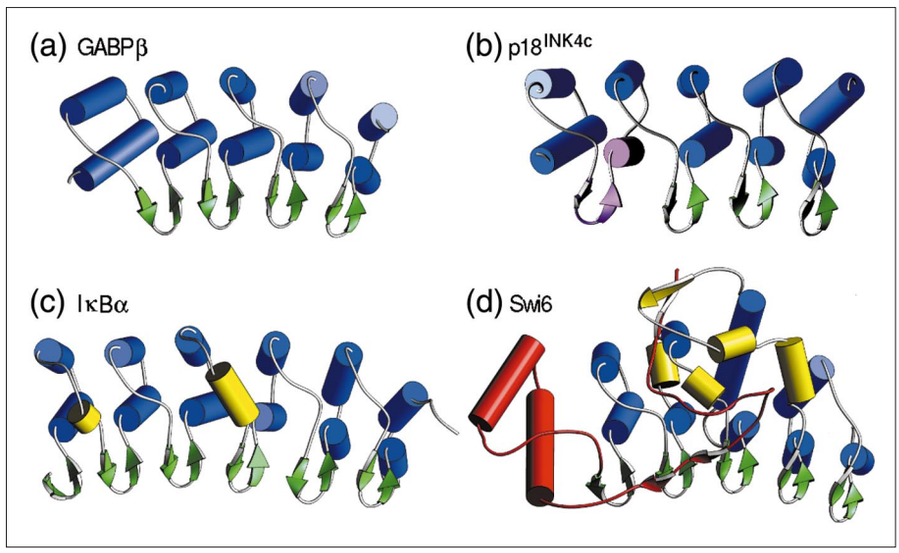

Figure 2

Schematic representations of four ankyrin (ANK)-repeat domains that illustrate the variations in architecture that might occur. **(a)** GABPβ consists only of ANK-repeat motifs. **(b)** Helix 1 of the second repeat of $p18^{INK4c}$ and the other INK4 family members is significantly shortened. Furthermore, the β-sheet structure is distorted by dislocation of the ANK1 β-hairpin (lilac). **(c)** Insertions within and between repeats are the most common variations of the basic ANK fold. Two helical insertions (yellow) occur between repeats 1 and 2, and between 3 and 4 of IκBα and the positions of these additional sequences appear to be conserved in the IκB family. **(d)** The ANK-repeat domain of Swi6 contains the largest number and variety of inserted elements observed thus far (yellow). Short helical segments are located between repeats 2 and 3 and, unusually, within the loop connecting helices 1 and 2 of ANK2. ANK3 and ANK4 are separated by a ~40-residue sequence that contains two α-helices, a short region of β-strand structure and a disordered loop. The transcriptional activation region (red) is tightly associated with the ANK-repeat domain and interacts with many of the insertion elements.

overall domain assembly, and are not conserved in the ANK consensus, which is consistent with a role in specific interactions. Further contacts involve residues on the exposed face of the α-helices in the ankyrin groove that are also not conserved in ANK-repeat proteins as a whole.

The ankyrin groove is also employed in contacts that both $p16^{INK4a}$ and $p19^{INK4d}$ make with CDK6 (Refs 11,15) to inhibit the kinase activity of CDK6 as part of a complex regulatory network in the late G1 phase of the mammalian cell cycle (Fig. 3b). CDKs contain a catalytic cleft formed by juxtaposition of N- and C-terminal domains (lobes) after activation by phosphorylation and cyclin binding. Together these events rearrange

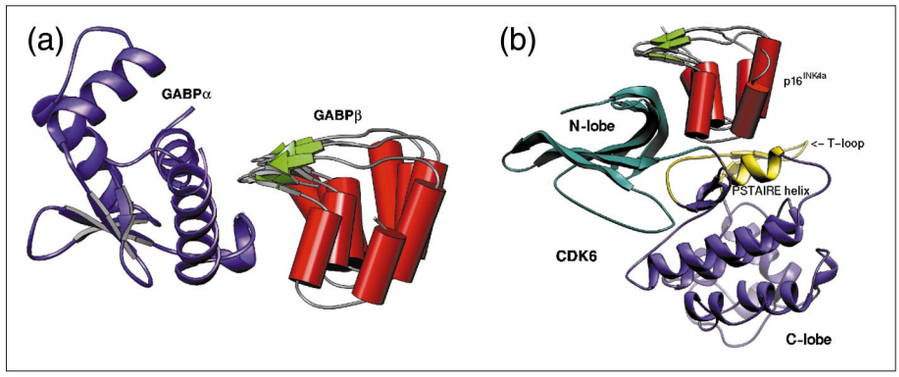

Figure 3

(a) The complex of GABPα with the ankyrin (ANK)-repeat accessory factor GABPβ clearly illustrates the use of the ankyrin groove as a binding surface. The structure was solved as a ternary complex of GABPα–GABPβ with a DNA duplex, but for clarity the DNA is not shown. **(b)** The INK4 family of ANK-repeat proteins function as cell-cycle regulators through interaction and inhibition of CDK6. The N-lobe of CDK6 (cyan) interacts with the ankyrin groove, whereas the base of the ankyrin stack binds to the PSTAIRE helix and the T-loop (yellow), preventing activation of the CDK (see text for details). CDK, cyclin-dependent kinase.

and stabilize two regulatory elements of the kinase, the T-loop and the PSTAIRE helix, and reorganize the enzyme into a catalytically competent conformation. Binding of the INK4 proteins to CDK6 locks the kinase in a conformation that is unfavourable for activation and is incapable of binding ATP. The helical face of the ankyrin groove of repeats 1, 2 and 3, together with the intervening β-hairpins, clamps the N-terminal lobe of CDK6. The connecting loops between the helix pairs of ANK1, ANK2 and ANK3 make further contacts with the T-loop located in the C-terminal lobe of CDK6. As a result, the N- and C-lobes of CDK6 are rotationally misaligned relative to the activated conformation, and movement of the highly conserved PSTAIRE helix into the catalytic cleft is blocked. Several point mutations in $p16^{INK4a}$ play a well documented role in the production of tumours. Although most of these simply destabilize the ANK-repeat structure, a subset now appears to disrupt interactions with CDK6 directly[8,10–12,15,19]. These findings emphasize the functional distinction between residues required for the structural integrity of the repeat and those involved in recognition and binding of target molecules.

The structures of two complexes formed by slightly different fragments of IκBα and the NFκB p50–p65 heterodimer (Fig. 4a) reveal the most extensive interface described thus far for any ANK-repeat domain[7,17]. IκBα antagonizes gene expression driven by NFκB transcriptional activators: it sequesters NFκB in the cytoplasm and inhibits phosphorylation events that stimulate DNA binding activity. Crystallization of these ternary complexes necessitated the use of partially deleted fragments of each component. Nevertheless, the variety of contacts and conformational changes observed in these structures have provided unprecedented insights into the multilevel regulation of NFκB by IκBα (Refs 20,21).

Cytoplasmic retention of NFκB by IκBα appears to be achieved through masking of the p65 nuclear localization signal (NLS). This region is unstructured in solution but adopts a helical conformation in the Jacobs and Harrison complex[7] (Fig. 4a). The surface of the ankyrin groove, formed by repeats 1, 2 and 3 of IκBα, contacts the NLS region of p65. Residues C-terminal to the p65 NLS wrap around the N-terminal end of IκBα, interacting with hydrophobic residues on the exposed face of the helical segments of ANK1. In addition, residues

from the ANK2 β-hairpin, which is located on the upper surface of the ankyrin domain, contact residues that are located immediately N-terminal to the NLS of p50 in the full-length molecule. Together, the observed interactions are sufficient to explain the particularly potent biological activity of IκBα compared with IκBβ and IκBκ in retaining NFκB in the cytoplasm[22].

IκBα also contacts the p65–p50 dimerization interface located within the N-terminal domain of the Rel-homology region of each subunit ($p50_{dim}$, $p65_{dim}$; Fig. 4). The β-hairpins of repeats 4, 5 and 6 contact the dimerization surface of p50, whereas the helices in the ankyrin groove of repeats 5 and 6 make contact with the corresponding region of p65. Helix 1 of ANK6 and the C-terminal PEST residues contact the N-terminal DNA-binding domain of the p65 moiety ($p65_{db}$), which is dramatically repositioned relative to the position in the p50–p65–DNA complex. Normally, the two DNA-binding domains of the p65–p50 heterodimer straddle the longitudinal axis of DNA and interact extensively with the major groove (Fig. 4b). Additional, sequence-specific contacts are supplied by residues from the dimerization domain. This 'molecular pincer' is held together by the dimerization domains of the two molecules, which are connected to the DNA-binding domains by flexible linkers. The reorientation of the DNA-binding domain, observed in the IkBα complex, disrupts this 'pincer' geometry and antagonizes DNA binding (Fig. 4c). Furthermore, it is likely that this structural rearrangement might reduce or prevent activation of p65 by cAMP-dependent protein kinase through occlusion of the phosphorylation site at Ser276 (Fig. 4a).

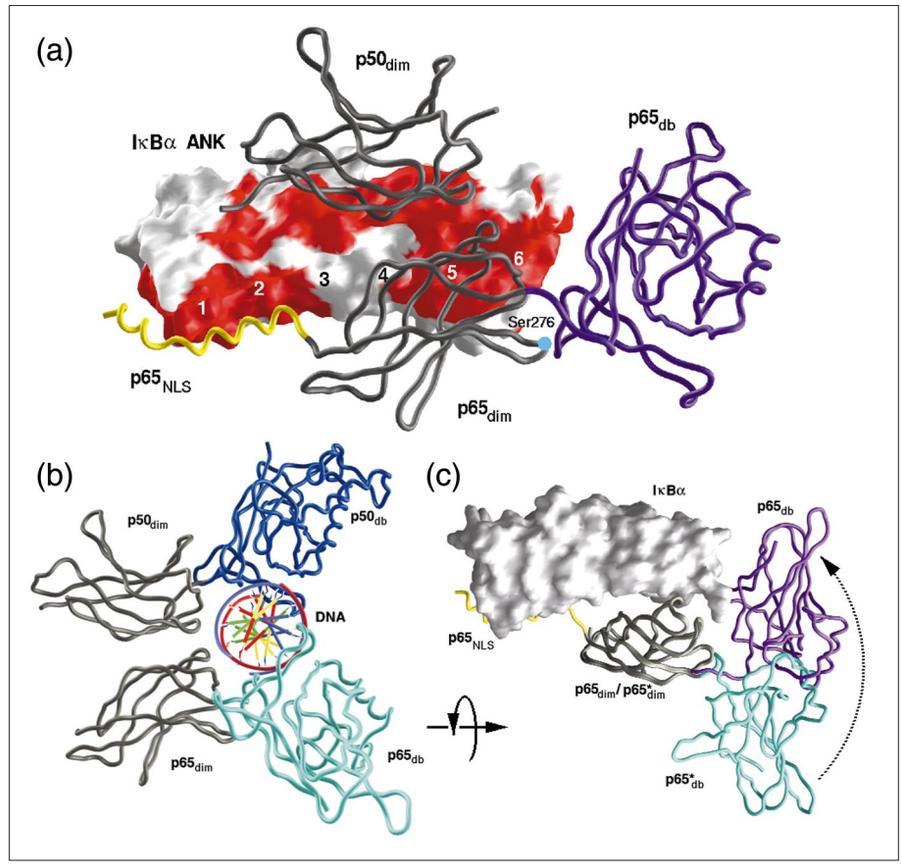

Figure 4

(a) Two independent structures of NFκB bound to the IκBα inhibitor have revealed the most extensive interface yet observed for any ankyrin (ANK)-repeat protein complex. Contacts with the IκBα ANK domain are shown as red patches on the molecular surface, which is viewed into the ankyrin groove and oriented with the β-hairpins projecting out of the plane of the page. Individual domains of the p50–p65 NFκB heterodimer and the nuclear localization signal (NLS) of p65 are shown as a backbone 'worm'. The dimerization (dim) domains of p50 (upper) and p65 (lower) are coloured grey. The DNA-binding domain (db) of p65 is shown in purple and the NLS is highlighted in yellow. **(b)** Domain arrangements as observed in NFκB–DNA co-crystal structures[28–32]. In this case, the p65–p50–DNA complex is shown with the dimerization domains (grey) oriented as shown in (a). The DNA duplex makes extensive interactions with the DNA-binding domains of both p50 (blue) and p65 (cyan). **(c)** Comparison of the domain arrangements of p65 observed in the IκBα complex (a) and the DNA complex (b). The view is towards the upper surface of the ANK domain by rotation of the molecule by ~60° about the horizontal axis with respect to the orientation in (a) and (b), as indicated. The p65 dimerization domains (grey) have been superimposed and the p50 subunits removed for clarity. The relative movement of the DNA-binding domain of p65 between the IκBα (purple) and DNA-bound (cyan) complexes is indicated by the arrow and represents a rotation of ~180° about the horizontal axis.

Intramolecular interactions

The structure of a central 36-kDa fragment of Swi6 provides the only example so far of intramolecular interactions involving ANK repeats[18] (Fig. 2d). In this case, a region that is capable of transcriptional activation *in vivo*[23] interacts with a core domain consisting of five ANK-repeat motifs. The interaction surface on the ANK-repeat domain differs considerably from those of the intermolecular complexes described above. In the cupped-hand analogy, the transcriptional activator interacts with the back of the hand and fingers, using, in part, the distinctive non-ANK structural elements that form part of this surface. In addition, the helical section of the activation domain is anchored across an extensive hydrophobic surface on the N-terminal face of the ankyrin block. Several temperature-sensitive mutations in Swi6 appear to disrupt interactions with the transactivation region rather than perturb the ANK-repeat core[24], underlining the structural significance of these intramolecular interactions. Sequence conservation indicates that similar interactions are likely to occur in the other members of the Swi6/Cdc10 family[18]. At the functional level, the ANK-repeat domains within these transcription factors might regulate, positively and negatively, the biological activities of flanking regions/domains and vice versa[23,25]. This might be a general feature of ANK-repeat domains, given that inhibition of the budding yeast Pho80p-Pho85p-cyclin–CDK complex by the ANK-repeat domain of Pho81p is modulated by distinct N- and C-terminal regions[26,27].

Summary

Protein evolution demands conservation of key residues to maintain structural integrity, but allows for sequence variations that, in turn, provide functional specificity. In fact, the complex biochemical requirements of a living cell

appear to be largely fulfilled by a relatively small number of protein architectures. In some cases, versatility has been achieved through the assembly of multiple copies of amino acid sequences to form families of structurally related but functionally diverse protein modules (e.g. Zn-finger motifs, WD40 repeats, armadillo repeats, HEAT repeats, leucine-rich repeats and ANK repeats).

Crystallographic, spectroscopic and biochemical studies of ANK-repeat proteins have now revealed how the structural and functional demands are reconciled simultaneously in these molecules. The core structure is sufficiently malleable to withstand considerable sequence variation, while still providing a stable framework for presenting surface contact residues. Binding to target molecules commonly involves contacts with the tips of the β-hairpins and the surface of the helical bundle facing into the ankyrin groove. In general, these residues are not conserved in the ANK consensus. As such, they are not structurally constrained and are ideally located to perform binding roles. The modular stacking of the repeat motifs allows incorporation of additional non-conserved elements that might contribute both to the structural stability of the domain and to the binding to the target molecule. Furthermore, as in the cases of other repetitive structural motifs, the ANK repeat lends itself to variation in overall domain size by simple sequence duplication or deletion. In this way, the ANK-repeat fold can recognize and bind to target molecules that vary considerably in size and shape. Although binding within the ankyrin groove is a common feature of known complex structures, the variations observed suggest that any part of the ANK-domain surface can be used in inter- and/or intramolecular contacts. The true diversity of these interactions will only be revealed by further structural studies of this remarkable family of proteins.

References

1 Breeden, L. and Nasmyth, K. (1987) *Nature* 329, 651–659
2 Lux, S. E., John, K. M. and Bennet, V. (1990) *Nature* 344, 36–42
3 Michaely, P. and Bennett, V. (1992) *Trends Cell Biol.* 2, 127–129
4 Bork, P. (1993) *Proteins Struct. Funct. Genet.* 17, 363–374
5 Axton, J. M. *et al.* (1994) *EMBO J.* 13, 462–470
6 Gorina, S. and Pavletich, N. P. (1996) *Science* 274, 1001–1005
7 Jacobs, M. D. and Harrison, S. C. (1998) *Cell* 95, 749–758
8 Venkataramani, R., Swaminathan, K. and Marmostein, R. (1998) *Nat. Struct. Biol.* 5, 74–81
9 Batchelor, A. H. *et al.* (1998) *Science* 279, 1037–1041
10 Byeon, I. *et al.* (1998) *Mol. Cell* 1, 421–431
11 Russo, A. *et al.* (1998) *Nature* 395, 237–243
12 Li, J. *et al.* (1999) *Biochemistry* 38, 2930–2940
13 Luh, F. Y. *et al.* (1997) *Nature* 389, 999–1003
14 Baumgartner, R. *et al.* (1999) *Structure* 6, 1279–1290
15 Brotherton, D. H. *et al.* (1998) *Nature* 395, 244–250
16 Yang, Y., Nanduri, S., Sen, S. and Qin, J. (1998) *Structure* 6, 619–626
17 Huxford, T., Huang, D-B., Malek, S. and Ghosh, G. (1998) *Cell* 95, 759–770
18 Foord, R., Taylor, I. A., Sedgwick, S. G. and Smerdon, S. J. (1999) *Nat. Struct. Biol.* 6, 157–165
19 Tang, K. S. *et al.* (1999) *J. Mol. Biol.* 285, 1869–1886
20 Baeuerle, P. A. (1998) *Cell* 95, 729–731
21 Cramer, P. and Muller, C. W. (1999) *Structure* 7, R1–R6
22 Simeonidis, S. *et al.* (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 49–54
23 Sedgwick, S. G. *et al.* (1998) *J. Mol. Biol.* 281, 763–775
24 Ewaskow, S. *et al.* (1998) *Biochemistry* 37, 4437–4450
25 Siegmund, R. F. and Nasmyth, K. (1996) *Mol. Cell. Biol.* 16, 2647–2655
26 Schneider, K. R., Smith, R. L. and O'Shea, E. K. (1994) *Science* 266, 122–126
27 Ogawa, N. *et al.* (1995) *Mol. Cell. Biol.* 15, 997–1004
28 Ghosh, G., Van Duyne, G., Ghosh, S. and Sigler, P. B. (1995) *Nature* 373, 303–310
29 Muller, C. W. *et al.* (1995) *Nature* 373, 311–317
30 Cramer, P., Larson, C. J., Verdine, G. L. and Muller, C. W. (1997) *EMBO J.* 16, 7079–7090
31 Chen, Y-Q., Ghosh, S. and Ghosh, G. (1998) *Nat. Struct. Biol.* 5, 67–73
32 Chen, F. E., Huang, D-B., Chen, Y-Q. and Ghosh, G. (1998) *Nature* 391, 410–413