Ankyrin Repeat: A Unique Motif Mediating Protein−Protein Interactions[†]

Junan Li,*[,‡] Anjali Mahajan,[§] and Ming-Daw Tsai*[,‡,§,||]

Departments of Chemistry and Biochemistry and Biophysics Program, The Ohio State University, Columbus, Ohio 43210, and Genomics Research Center and Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

Received October 21, 2006; Revised Manuscript Received November 13, 2006

ABSTRACT: Ankyrin repeat, one of the most widely existing protein motifs in nature, consists of 30−34 amino acid residues and exclusively functions to mediate protein−protein interactions, some of which are directly involved in the development of human cancer and other diseases. Each ankyrin repeat exhibits a helix−turn−helix conformation, and strings of such tandem repeats are packed in a nearly linear array to form helix−turn−helix bundles with relatively flexible loops. The global structure of an ankyrin repeat protein is mainly stabilized by intra- and inter-repeat hydrophobic and hydrogen bonding interactions. The repetitive and elongated nature of ankyrin repeat proteins provides the molecular bases of the unique characteristics of ankyrin repeat proteins in protein stability, folding and unfolding, and binding specificity. Recent studies have demonstrated that ankyrin repeat proteins do not recognize specific sequences, and interacting residues are discontinuously dispersed into the whole molecules of both the ankyrin repeat protein and its partner. In addition, the availability of thousands of ankyrin repeat sequences has made it feasible to use rational design to modify the specificity and stability of physiologically important ankyrin repeat proteins and even to generate ankyrin repeat proteins with novel functions through combinatorial chemistry approaches.

Repeat proteins are the second most abundant protein classes functioning in protein−protein binding after immunoglobulins (*1*, *2*). They widely exist across most forms of life and are involved in numerous physiological processes such as cytoskeleton integrity, cell cycle control, transcriptional regulation, cell signaling, development and differentiation, apoptosis, cellular scaffolding, vesicular trafficking, inflammatory response, plant defense, and bacterial invasion (*2*, *3*). These nonglobular or modular proteins are composed of relatively short, tandem repeating motifs (of ∼20−40 amino acid residues), which exhibit a well-defined topology when present in a repeat domain but are usually unfolded on their own (*3*). Various numbers of such repeating motifs are packed together either themselves or in conjunction with other domains to form the underlying architecture of an extended, modular interface exposed to proteins for binding (*2*, *4*). To date, more than 20 classes of repeat proteins have been identified in the expanding protein sequence data bank, among which the most abundant are leucine-rich repeat (LRR)[1] proteins, ankyrin repeat (AR) proteins, armadillo/HEAT repeat proteins, and tetratricopeptide repeat proteins (*1*−*3*). While repeat proteins have been understudied in comparison with globular proteins, great effort has been made recently to investigate the unique biochemical and biophysical properties of different repeat proteins. This work reviews recent chemistry−biology interdisciplinary studies on one type of repeat proteins, the ankyrin repeat proteins, and the potential of generating AR proteins with novel binding abilities through approaches of consensus design and combinatorial library.

Ankyrin repeat, a motif of 33 amino acid residues (*5*), was first identified in the sequence of yeast Swi6p, Cdc10p, and *Drosophila melanogaster* Notch in 1987 (*6*) and later named after the cytoskeletal protein ankyrin due to the fact that it consists of 24 copies of such a sequence (*7*). The increasing availability of protein sequence data then led to the discovery of ankyrin repeats in a myriad of proteins with diverse functions (*5*). A recent sequence homology analysis study has demonstrated that there are 19 276 AR sequences in 3608 proteins identified from the nonredundant SMART protein database (*8*). These proteins include transcriptional regulators, cytoskeletal organizers, positive/negative modulators functioning in cell cycle progression, cell development, and differentiation, and toxins (*5*, *9*). A majority of such proteins are eukaryotic proteins present in both intracellular

[†] Research work on ankyrin repeat proteins in the authors' laboratory was supported by a grant from the National Institutes of Health (CA69472) to M.-D.T.

* To whom correspondence should be addressed: Department of Chemistry, 100 West 18th Ave., The Ohio State University, Columbus, OH 43210. Phone: (614) 292-6974. Fax: (614) 292-1532. E-mail: li.225@osu.edu, tsai.7@osu.edu.

[‡] Departments of Chemistry and Biochemistry, The Ohio State University.

[§] Biophysics Program, The Ohio State University.

[||] Academia Sinica.

[1] Abbreviations: AFM, atomic-force microscopy; AR, ankyrin repeat; CD, circular dichroism; CDK, cyclin-dependent kinase; CSL, CBF-1/RBP-jk, Su(H), Lag-1, named for the mammalian, *D. melanogater*, and *C. elegans* orthologous proteins, respectively, a DNA binding transcription factor required for both repression and activation of Notch target genes; HDM2, MDM2 human homologue; Ink4, inhibitors of cyclin-dependent kinase 4; LRR, leucine-rich repeat; MAGE-A4, melanoma antigen A4; MBP, maltose binding protein; NMR, nuclear magnetic resonance spectroscopy; p15, p16, p18, and p19, CDK4 inhibitors $p15^{INK4B}$, $p16^{INK4A}$, $p18^{INK4C}$, and $p19^{INK4D}$, respectively; Rb, retinoblastoma susceptible gene product; UV, ultraviolet light.

10.1021/bi062188q CCC: $33.50
Published on Web 11/30/2006

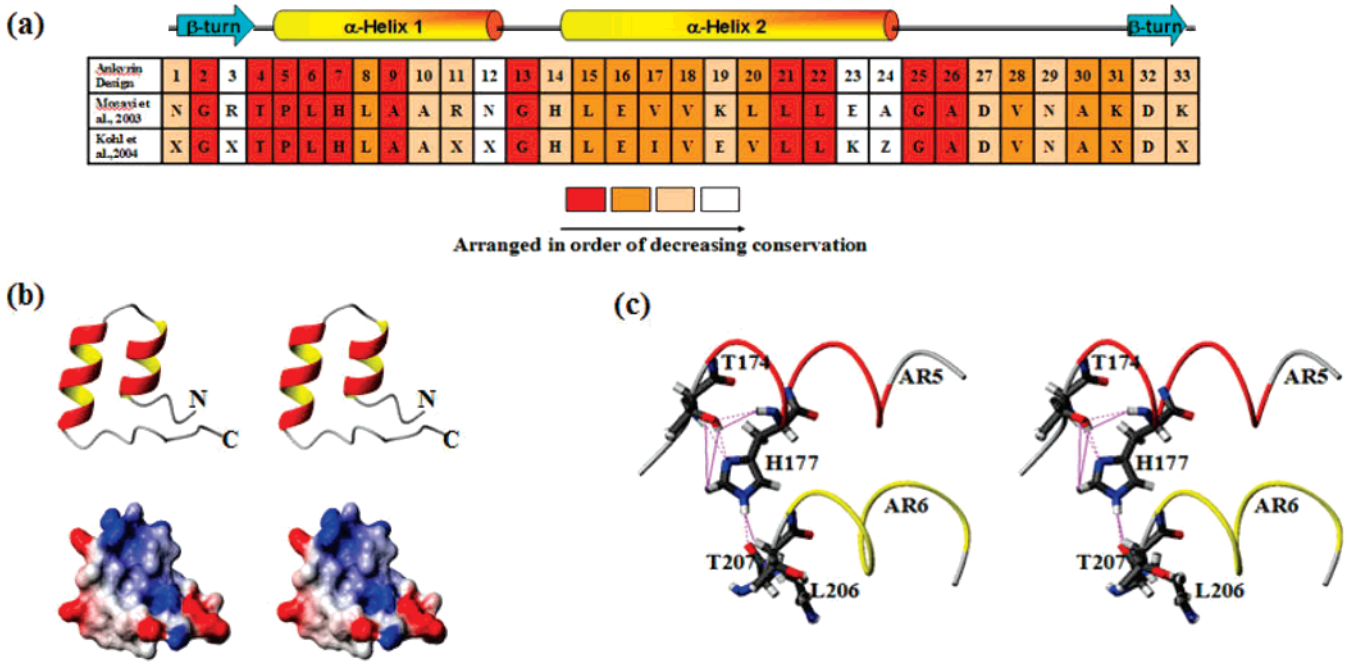

FIGURE 1: Unique structural features of an AR module. (a) Consensus sequences of the ankyrin repeat proteins as presented by Mosavi et al. (*14*) and Kohl et al. (*15*). The schematic representation of the secondary structures corresponding to this sequence is presented above the sequences. The conservation level of each sequence is color-coded. The Kohl et al. consensus sequence includes an X that denotes any amino acid except C, G, and P and a Z that can be a H, N, or Y. This figure was adapted from ref *4*. (b) Stereoview showing the ribbon diagram (top) and surface charge distribution (bottom) of the fourth ankyrin repeat of gankyrin (PDB entry 1TR4). The topology of the module resembles the letter L with the helices being the vertical arm and the N- and C-terminal stretches being the base. (c) Stereoview showing the role of ^{174}TPLH^{177} sequence in stabilizing the secondary structural modules of the AR protein gankyrin (PDB entry 1TR4) (*16*).

and extracellular milieu (*5*, *9*). It has been suggested that AR proteins are important in modulating diverse cellular pathways necessary for the evolution of a more complicated multicellular organism (*10*). In spite of the great variety of aforementioned functions, no enzymatic activity has been detected to date for any AR protein or AR domain in proteins containing both AR and non-AR domains (*5*, *8*). The AR motifs function in different pathways exclusively through mediating specific protein−protein interactions. Moreover, genetic alterations, including deletion, methylation, and point mutations, have been identified in genes encoding several AR proteins, such as tumor suppressors p16 and p18 in a variety of human tumors, indicating that dysfunction of AR proteins has a bearing on human diseases (*11*).

Structure

Ankyrin repeat is a relatively conserved motif which implies that in spite of a strong degeneracy of the repeating sequences, there is a consistent pattern of key residues to keep the structural integrity of an AR motif (*3*) (Figure 1a). Each AR motif exhibits a canonical helix−turn−helix conformation, in which two α-helices are arranged in an antiparallel fashion and the loop projects outward at an approximately 90° angle to facilitate the formation of hairpinlike β-sheets with neighboring loops (*12*, *13*). Usually, a hairpinlike β-sheet structure consists of the last three residues of the preceding AR and the first four residues of the next AR. Overall, the topology of an AR motif resembles the letter L with the helices as the vertical arm and the N- and C-terminal stretches as the base (*3*, *5*) (Figure 1b). To retain this characteristic topology, some residues are well-conserved in most AR sequences while residues at other positions vary. Figure 1a shows two AR consensus sequences derived from statistical analyses of AR sequences by Mosavi et al. (*14*) and Kohl et al. (*15*). Specifically, the Thr-Pro-Leu-His tetrapeptide motif at positions 6−9 is highly prevalent in AR sequences (*5*, *8*, *16*). It forms a tight turn and initiates the first α-helix of an AR motif. Moreover, hydrogen bonding between the hydroxyl group of Thr and the imidazole ring of His contributes to the stability of AR motif and global AR proteins (*16*) (Figure 1c). The Val/Ile-Val-Xxx (hydrophilic)-Leu/Val-Leu-Leu motif (positions 17−22) is the central piece of the second α-helix of an AR motif, and it functions to form intra- and inter-AR hydrophobic networks to stabilize the global sketch of an AR protein (*5*, *8*). In addition, glycine residues are conserved at positions 13 and 25, which terminate the helices and provide the freedom for a loop to link both helices (*5*). The key roles of these conserved residues have been further confirmed by the fact that mutations at these conserved residues in p16 and p18 have been identified in human tumor specimens, and these mutations have led to perturbation in the stability and solubility of p16 and p18 proteins (*12*, *17*, *18*).

To form a stable AR protein or AR domain capable of binding to specific targets, various numbers of AR scaffolds stack together to form a helix bundle in such a manner that the repetitive scaffolds are arrayed almost linearly with a 2−3° counterclockwise rotation between neighboring repeats (*8*, *19*). The loop regions of neighboring ARs are connected in a tail-to-head order to form hairpinlike β-sheet structures (Figure 2). While such hairpinlike structures are relatively flexible in conformation (*13*), the extended helix bundle is stabilized mainly through inter- and intrarepeat hydrophobic interactions predominantly associated with conserved non-

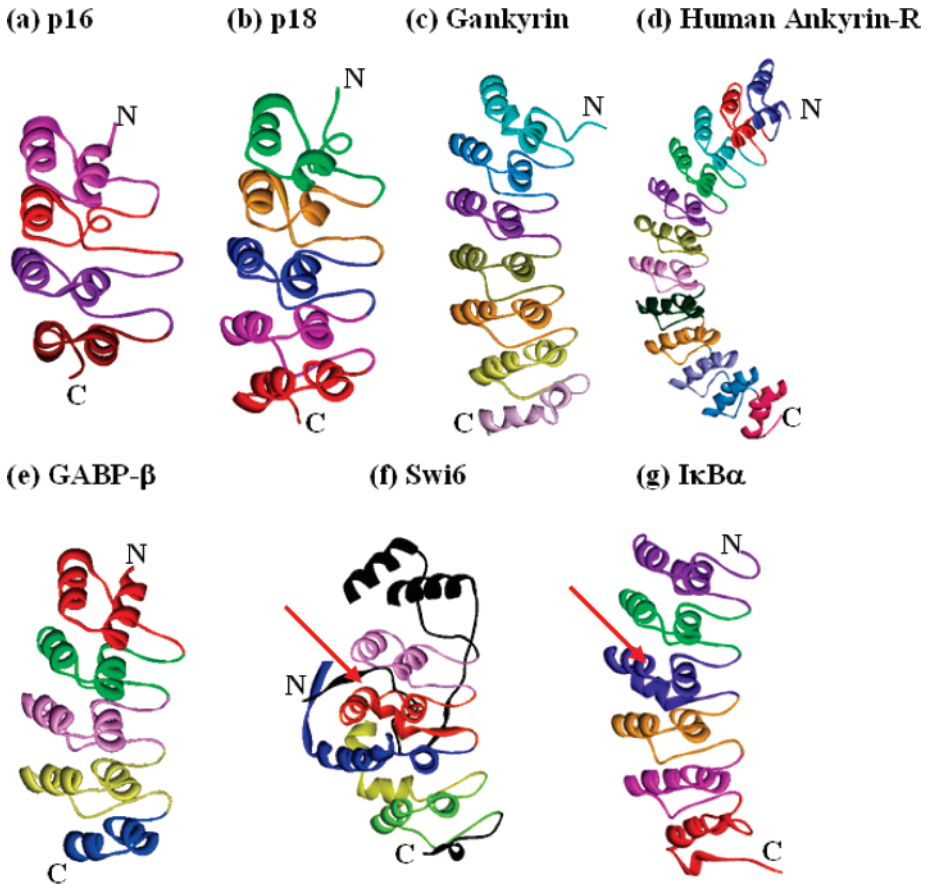

FIGURE 2: Structures of representative AR proteins: (a) p16 (PDB entry 1DC2), (b) p18 (PDB entry 1BU9), (c) human gankyrin (PDB entry 1TR4), (d) human ankyrin-R (PDB entry 1N11), (e) transcription factor GABP-β (PDB entry 1AWC), (f) transcriptional regulator Swi6 (PDB entry 1SW6), and (g) IκBα, inhibitor of NF-κB (PDB entry 1NFI). Of note, an internal loop was absent in the crystal structure of Swi6 because of its conformational flexibility. Arrows in panels f and g indicate the intervening helix in the third AR of Swi6 and IκBα, respectively.

polar residues in the helical regions (Figure 3a), as well as hydrogen bonding interactions between polar residues and the main chain atoms from neighboring repeats (*18*) (Figure 3b). Since the helices on the far side of the loop generally have larger hydrophobic side chains and are slightly longer than the helices on the near side, the molecule of an AR protein is in a concave shape.

While the architectures of AR proteins are similar in topology, the tertiary structure of each AR protein is finely adapted to the following factors.

(1) Presence of Non-AR Domains in a Protein. Some proteins contain both AR and non-AR domains, and these non-AR domains often affect the orientation of AR domains in the global structure. For example, in addition to the domain consisting of three and one-half AR motifs, the yeast ribosomal binding protein Yar-1 has a calmodulin-binding domain at the C-terminus of the AR domain (*20*).

(2) Number of AR Motifs in a Protein. While proteins with up to 29 AR motifs have been identified (*21*), the majority of AR proteins contain four to seven AR motifs (*22*), and no single AR has yet been reported to form a stable folded unit itself (*14*, *23*). In general, proteins containing more AR repeats have more compact and concave structures (Figure 2). Each AR motif causes a 2−3° counterclockwise rotation relative to the preceding AR motif as well as approximately 13° of pitch to the superhelical turn of the protein molecule (*8*, *19*). A protein of 32 AR motifs will form a complete superhelical turn with a predicted radius of 35 Å (*24*).

(3) Intrinsic Differences among AR Motifs. On one hand, the "capping" AR motifs, i.e., AR motifs at the N- and C-termini of an AR protein, are different from inner AR motifs. The capping AR motifs are more solvent-accessible, and they shield the hydrophobic core of AR proteins from the solvent (*2*, *25*). On the other hand, some AR motifs deviate considerably from the typical helix−turn−helix modules. For example, the first helix of the second AR motif of INK4 proteins is atypically short (approximately four to seven residues), but the helix−turn−helix−hairpin topology remains unchanged (Figure 2a,b) (*13*, *18*). Moreover, while GA-binding protein β (GABP-β) exhibits a global conformation as a typical AR protein with repeating helix−turn−helix−hairpin units, the "fifth" AR motif at its C-terminus is incomplete and regarded as half of an AR motif (*26*) (Figure 2e).

(4) Variance of Loops. The loops linking neighboring AR motifs are considerably flexible in size and conformation. Typically, loops assemble to form a continuous β-sheet between neighboring AR motifs. Nevertheless, some AR proteins have very long loops, and these loops exhibit more complicated conformations than plain β-sheets. For example, the loop between the third and fourth AR motifs of yeast Swi6 protein contains ~40 amino acid residues (Figure 2f), and it brings about an expanded second helix of the third AR motif, a short section of β-sheet, an unfolded loop, and two additional short helices (*27*). Similarly, a short insertion between the third and fourth AR motifs of IκBα forms an

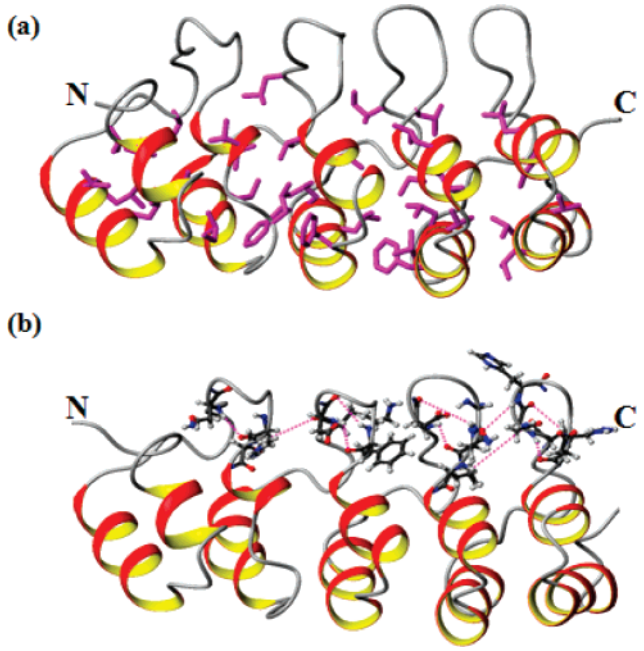

FIGURE 3: Intra- and inter-repeat interactions of p18 (PDB entry 1BU9). (a) Side chains of residues constituting the hydrophobic core of p18 protein are colored magenta. (b) The hydrogen bonding network of p18 in its β-sheets is colored magenta.

additional helical region within the loop (*28*, *29*) (Figure 2g). In contrast, there are no β-sheets formed in the loops of myotrophin (containing four ARs) (*30*), and only nascent β-sheets are present in the loops of p16 (*13*, *18*) **(**Figure 2a**)**. Since the loops protrude away from the helices at an approximately 90° angle, the variances at loops usually do not interrupt the stacking of ARs in a linear array fashion, suggesting that AR proteins have considerable tolerance in structural deviation.

Folding and Stability

Due to its modular nature, the structures of AR proteins are predominantly stabilized by local and short-range intramolecular interactions, rather than interactions between distant residues as frequently found in globular proteins (*3*, *8*). Additionally, no AR protein or an AR repeat domain has been found to contain a disulfide bond, a structural element that could stabilize a protein but may not always be advantageous (*31*). Hence, one might think that AR proteins differ from globular proteins in folding and stability and each AR in a protein could unfold and refold individually. However, folding studies on INK4 proteins (*12*, *17*, *23*, *32*, *33*), Notch AR domain (*34*, *35*), myotrophin (*22*), and some engineered AR proteins (*14*, *15*, *36*, *37*) have demonstrated that AR proteins exhibit an equilibrium two-state transition between native and unfolded proteins by spectroscopic and calorimetric criteria, similar to most globular proteins. Furthermore, no partially folded intermediates have been found in these equilibrium folding−unfolding reactions, except that a NMR study of p19 in the presence of moderate concentrations of urea revealed the possible existence of an equilibrium intermediate with poor chemical shift dispersions, suggesting that it is largely unfolded (*32*). Taken together, these findings indicate that like globular proteins, the equilibrium folding and unfolding of AR proteins is highly cooperative (*17*, *22*, *38*).

While the two-state model has been well accepted in equilibrium folding−unfolding studies, significant effort has been devoted to identification of possible kinetic folding intermediates. Tang et al. (*39*) used a Φ value analysis to investigate the kinetic folding of p16 wild-type and mutant proteins and found that Φ values on the two ARs at the C-terminus are close to 1 and generally larger than those on the two ARs at the N-terminus, suggesting that p16 was unfolded sequentially. In the transition state for kinetic folding and unfolding, the two C-terminal ARs were fully folded while the other two ARs at the N-terminus were mainly unstructured. This conclusion was further reinforced by a study using truncated p16 proteins, which demonstrated that the two ARs at the C-terminus can form a relatively stable, autonomously folded structure while the two ARs at the N-terminus were unstructured (*23*). These findings suggest that the ARs at the C-terminus fold first, forming a "core" structure to facilitate the folding of the ARs at the N-terminus. Similarly, a kinetic on-pathway intermediate has been detected in the unfolding and refolding of the AR domain (seven ARs) of Notch, a *D. melanogaster* receptor which plays a role in cell signaling (*40*−*42*). This kinetic intermediate appears to have native like Trp fluorescence properties, but it is midway between the denatured and native ensembles in denaturant sensitivity and folding free energy (*41*). The refolding of Notch AR domain fits well to a sequential three-state model, in which the formation of the intermediate is the rate-limiting step followed by a fast conversion to the native ensemble (*41*). Furthermore, ARs 3, 4, and 5 of Notch are structured in this intermediate, while ARs 2, 6, and 7 remain largely unstructured, which become structured in the subsequent conversion to the native ensemble (*42*). As for AR 1 of Notch, it is not well structured until it binds to its targets (*32*, *43*, *44*). Taken together, these studies suggested that couplings between neighboring ARs contribute to the folding cooperativity of AR proteins.

Thermodynamic and kinetic studies on the AR domain of Notch protein have also provided evidence of the likely presence of long-range coupling among ARs (*38*). Barrick et al. designed a series of truncated Notch AR proteins by removing one or two ARs from the C-terminus and analyzed their conformational stabilities through urea-induced denaturation (*34*, *38*, *45*). While all three proteins, Nank1−7*, Nank1−6*, and Nank1−5* (wild type, last AR truncated from the C-terminus, and last two ARs truncated, respectively), fold in a typical two-state manner, the reported values of ΔG_d^{water} were 8.03, 4.14, and 3.55 kcal/mol, respectively, suggesting that Nank1−7* is thermodynamically much more stable than Nank1−6* and Nank1−5* and that the seventh AR of Notch AR protein contributes more to the stability of the entire Notch AR domain than the sixth AR. Moreover, inclusion of the seventh AR shifts the unfolding transition to a higher denaturant concentration and temperature (in comparison with Nank1−6* and Nank1−5*), and the transition itself is sharpened, suggesting that the entire Notch AR domain might be stabilized by the seventh AR at the C-terminus through long-range cooperative interactions (*34*). In an earlier study, we have also shown that p18 is substantially more stable than p16 (with ΔG_d^{water} values of 1.94 and 2.98 kcal/mol, respectively), and the difference can be attributed to stabilization of the fourth AR by the additional AR in p18 (*33*), on the basis of guanidinium

hydrochloride-induced denaturation, deuterium exchange, and NMR dynamics studies.

In a recent study, Barrick and his colleagues reported that insertions of consensus ARs (see Consensus Design and Combinatorial Library of ARs) between AR 5 and AR 6 of the Notch AR domain did not dramatically change the secondary and tertiary structures of the entire AR domain as monitored by far- and near-UV CD spectroscopy (*46*) but improved the conformational stability in urea-induced denaturation studies. However, unfolding of these variants follows a three-state transition process, and the C-terminal ARs after the consensus ARs, i.e., AR 6 and AR 7, unfold at moderate urea concentrations, indicating that the insertion of consensus ARs disrupts the folding cooperativity of the Notch AR domain. Moreover, after removal of AR 6 and AR 7, variants containing Notch ARs 1−5 and consensus ARs show equilibrium two-state unfolding transitions, and the thermodynamic stability of these hybrid proteins is higher than that of parental Notch AR domain, implying that the high stability of the consensus ARs has been propagated into the N-terminal Notch ARs (*46*). Hence, while each AR motif possesses some degree of intrinsic stability and contributes to the global stability of an AR protein, long-range cooperation may exist and affect the entire molecule.

Interestingly, a recent atomic force microscopy (AFM) study on a designed protein consisting of an N-terminal cap, six identical consensus-designed ARs, and a C-terminal cap showed that this AR protein unfolds in a stepwise, repeat-by-repeat fashion under a mechanical force (*47*). Stretching the AR protein with an AFM tip led to a uniform sawtooth pattern in the force-extension curve with as many as six peaks evenly separated by ~10 nm, and the average unfolding force for each peak is 50 ± 20 pN. Evidently, this stepwise unfolding of AR proteins by mechanical force is different from the cooperative unfolding caused by temperature and chemicals such as urea. A possible explanation could be that temperature and chemicals affect the entire protein surface while the application of a stretching force in single-molecule AFM unravels a protein along a specific direction (*48*). Since some AR proteins, such as ankyrin, are involved in skeleton organization and under certain physiological conditions mechanical force may act as a denaturant, mechanical force-induced unfolding of AR proteins could be physiologically important (*47*).

Binding

The modular architecture and variable molecular surfaces generated by the assembly of multiple compatible repeats render AR proteins highly versatile in protein binding. As discussed earlier, each AR consists of framework residues important for intra- and inter-repeat interactions for the formation of the helix−turn−helix−loop sketch structure and surface-exposed residues for contacting targets. These surface-exposed residues are not restricted to any particular secondary structural element within the AR, and they vary among different ARs. When these repetitive structural units stack together to form an elongated molecule, the surface residues of several juxtaposed repeats form a large solvent-accessible surface available to specific partners. Thus, binding of an AR protein to its partner is the accumulation of contacts involving a number of residues, and these residues are discontinuously dispersed into the whole molecule. A number of structures of complexes containing AR proteins and its partners have been reported, such as the 53BP2−p53 complex (*49*), the GABPα−GABP*β*−DNA complex (*26*), the CDK6−INK4 (p16, p19, and p18) complexes (*50*−*52*), the I*κ*Bα−NF-*κ*B complexes (*28*, *29*), and the CSL−Notch−Mastermind complexes (*43*, *44*). In this review, the INK4−CDK4 (or CDK6) system is used as an example to address the binding mode and binding specificity of AR proteins.

INK4 proteins (inhibitors of cyclin-dependent kinase 4) include four AR proteins, p15 (four ARs), p16 (four ARs), p18 (five ARs), and p19 (five ARs). Their specific binding to and inhibition of cyclin-dependent kinases 4 and 6 (CDK4/6) have been found to be crucial in modulating cell progression through the G1−S transition. Binding of INK4 proteins to CDK4/6 impedes the phosphorylation of Rb (the retinoblastoma susceptible gene product), leads to the release of transcription factors E2Fs from incompetent Rb−E2F complexes, and subsequently turns on downstream genes required for entry into the S phase (*53*) (Figure 5d). In addition, several recent studies have demonstrated that p16 suppresses proliferation of stem or progenitor cells in the bone marrow, pancreas, and brain by inducing cellular senescence or other unknown mechanisms, thus playing a pivotal role in aging and longevity (*54*−*56*). Therefore, p16 functions to reduce regenerative capacity and promote aging but also reduce cancer incidences.

As revealed in the crystal structure of the p16−CDK6 complex (*50*) (Figure 4a), p16 interacts with both N-terminal and C-terminal lobes of CDK6 through a great number of interactions listed in Table 1. Binding of p16 brings about a rational misalignment at the N- and C-lobes of CDK6 relative to the conformation required for an active kinase, thus blocking the catalytic cleft for ATP binding as well as shrinking the cyclin D-binding site. In spite of the unfavorable conformational change in CDK6, structural comparison showed that there is little difference in conformation between free p16 and CDK4-bound p16, suggesting that free p16 exists in a functional conformation. Most CDK6-binding residues are located within three regions of p16 (*50*, *57*). The first region involves loops 1 and 2 (but not loop 3) of p16. Both loops 1 and 2 interact with CDK6 through the last five or six residues of the loop preceding the next AR. The second interacting region includes the first helix of both the second and the third ARs, while the second helix of both ARs makes little contact with the kinase. The third major region of interaction is located at the "turns" bridging the helices of ARs. A breakdown in the interactions by ARs has shown that the second AR of p16 has the largest surface area of contact with CDK6, the second largest contribution comes from the third AR, while the first and the fourth ARs contribute little to the interaction with CDK6 (*57*). These findings have been further confirmed by independent biochemical studies on p16 mutants (*13*, *18*, *57*). In these studies, quantitative contributions of specific p16 residues to the inhibition of the CDK4−cyclin D2 holoenzyme were evaluated using in vitro kinase assays, while the potential perturbation to the global structure by p16 mutants was analyzed using NMR. As shown in Figure 4b, functionally important residues, i.e., residues involved in inhibiting CDK4, are located mostly within the second and third ARs and the loop bridging these two ARs (loop 2). Interestingly,

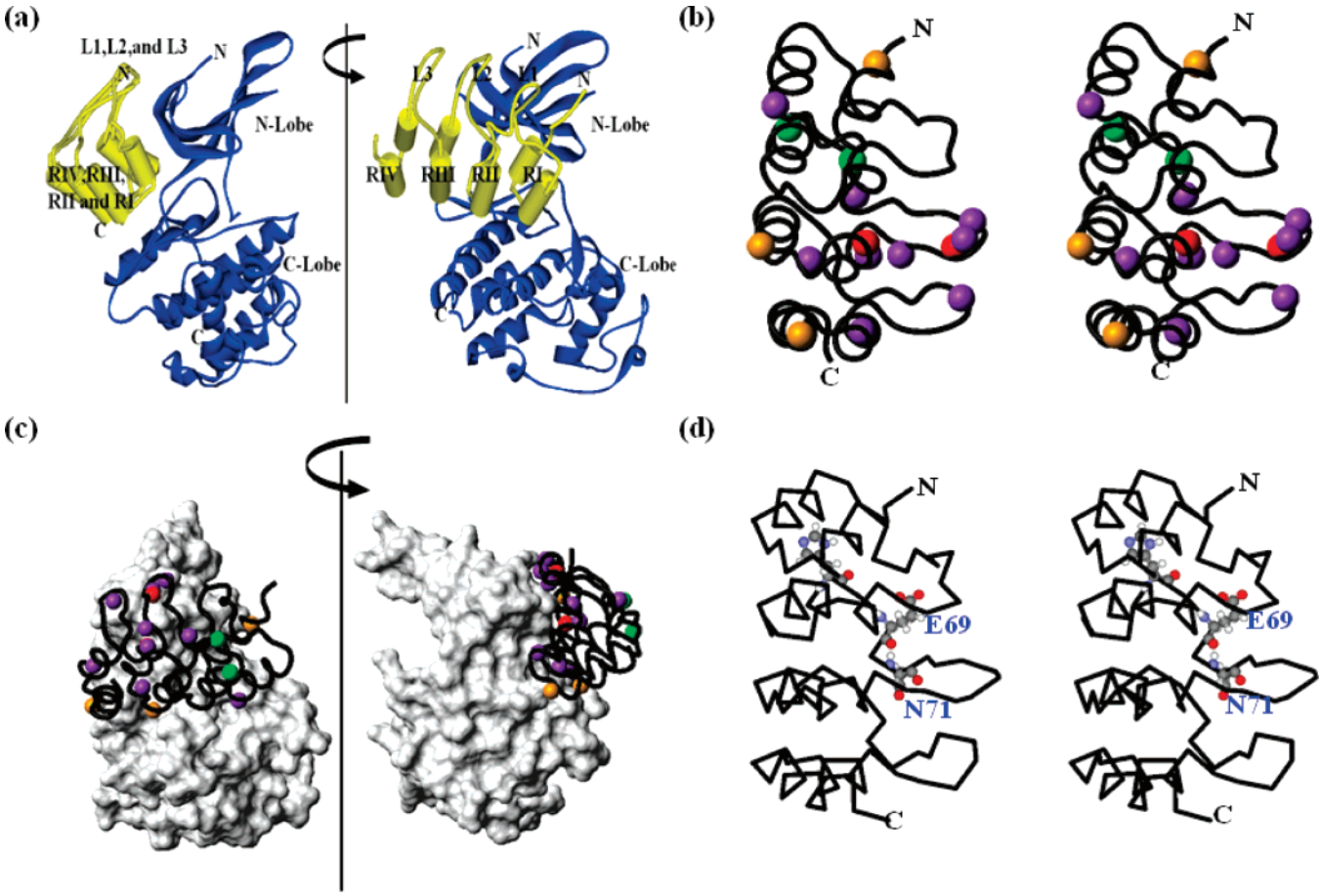

FIGURE 4: Structural bases of p16−CDK6 (or CDK4) interactions. (a) X-ray structure of human p16 (yellow) in complex with CDK6 (blue) showing the interacting regions (PDB entry 1BI7) (*50*). The ankyrin repeats are labeled as RI−RIV and the loops as L1−L3; the C- and N-lobes of CDK6 are labeled. (b) Quantitative contributions of functionally important residues of p16 as suggested by the in vitro CDK4−cyclin D2 inhibition assays (*13*, *18*). Various amounts of p16 or p16 mutants were mixed with a fixed amount of the CDK4−cyclin D2 complex, followed by incubation with [^{32}P]ATP and Rb. IC_{50} was defined as the concentration of p16 or p16 mutants required to achieve 50% of the maximum inhibition. Residues are presented in different colors on the basis of changes in the values of IC_{50} when mutated. Residues with a >20-fold increase in IC_{50} when mutated are colored red (L78 and D84), 10−20-fold orange (W15, D92, and R124), 5−10-fold green (H66 and E69), and 3−5-fold purple (E26, N71, P76, A77, T80, H83, F90, W110, and L121). (c) Structural positioning of the functionally important residues of p16 in contact with CDK6 using the crystal structure of the p16−CDK6 complex (PDB entry 1BI7). (d) E69 and N71 are part of the H-bonding network that stabilizes the global structure of p16 (*18*, *57*).

in the p16−CDK6 complex, all but three of these functionally important residues face CDK6 (Figure 4c). Two of the three residues not facing CDK6 are H66 and E69 (in green), mutations of which to Ala resulted in a 5-fold change in the CDK4 inhibitory activity (as presented by changes in the values of IC_{50}), and residue number N71 (colored purple) with an only 3.5-fold change in the inhibitory activity after mutation to Ser. These residues may indirectly bring about some "local" conformational perturbation, thus affecting other functionally important residues of p16 in the proximity. Moreover, E69 and N71 are located in the loop regions and are part of the H-bonding network stabilizing the global structure as shown in Figures 3b and 4d (*18*).

As described above, contacts between p16 (or other INK4 proteins, in general) and CDK4 occur in discontinuous patches, and a number of residues located in both loop and helical regions of p16 contribute to CDK4 binding through electrostatic, hydrogen bonding, and van der Waals interactions. While specific contacts between an AR protein and its partner vary considerably, such discontinuous, multiple-residue interacting patterns have been found in all crystal structures of complexes containing AR proteins. In the 53BP2−p53 complex (*49*), most contacts between the AR domain of 53BP2 (four ARs) and p53 are located at the C-terminal AR loop of 53BP2, including hydrogen bonds between the backbone carbonyl of S425 (53BP2) and both the side chain and backbone amide of S183 (p53), the backbone carbonyl of M422 (53BP2), and the side chain of H178 (p53), and van der Waals interactions involving W424, M427, and M433 of 53BP2 and H178, H179, S182, and R181 of p53. In the GABPα−GABPβ−DNA complex (*26*), it is principally the two residues at the tip (defined as the first residue of an AR and the last residue of the preceding AR) of each AR loop of GABPβ that interact with GABPα, while additional contacts with GABPα are mediated by GABPβ residues adjacent to the tips of the loops and in the AR helices. In the IκBα−NF-κB complex (*28*, *29*), various parts of the first two ARs of IκBα (six ARs) contact the C-terminal extension of the p65 subunit of NF-κB, including its nuclear localization signal (NLS), tips of the loops in ARs 4−6 contact the dimerization domain of the p50 subunit of NF-κB, and the inner helices of ARs 5 and 6 contact the dimerization domain of p65. AR 3 has relatively little contact with NF-κB. Specifically, contacts between the p65 subunit and the first two ARs of IκBα include salt bridges between K301 (p65) and D75 (IκBα), R302 (p65) and E85 (IκBα),

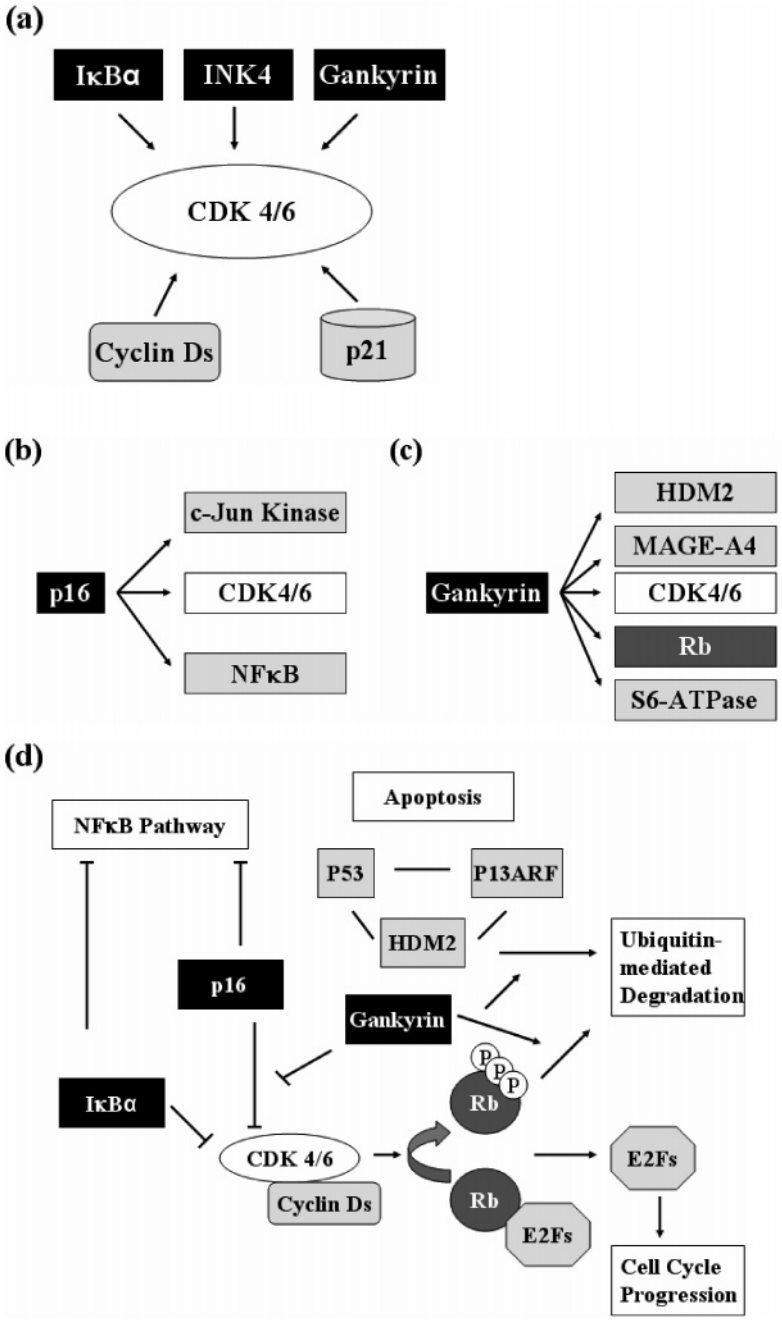

FIGURE 5: Functions of AR proteins. (a) Interactions of CDK4 and CDK6 with AR proteins (black rectangles) and non-AR proteins. (b and c) Interactions of AR proteins p16 and gankyrin with multiple partners. (d) Roles of AR proteins in cell signaling. Arrows represent positive regulation, and cross bars represent negative regulation.

and R304 (p65) and D73 (IκBα), hydrogen bonds between R302 (p65) and the main chain carbonyls of I83 and H84 (IκBα), and T305 (p65) and E85 (IκBα), and hydrophobic interactions involving F309, I312, M313, and F318 of p65 and F77, L80, A81, L89, and V93 of IκBα. In the CSL−Notch−Mastermind complexes (*43*, *44*), while van der Waals forces contribute to some degree, electrostatic interactions appear to play a predominant role in contacting CSL. Prominent electronegative patches on the CSL surface created by residues E569, D571, E580, and E607 directly interact with electropositive patches on the Notch AR domain formed by residues R1128, R1161, and R1208. The β-hairpins of the fifth and seventh ARs of Notch also contribute to binding to CSL.

Specificity

Despite the overall sequence similarity shared by the AR proteins, binding of AR proteins to their targets is specific. For the INK4−CDK4/6 system, on one hand, INK4 proteins bind to and inhibit CDK4/6, but not CDK2, even though CDK2 is similar to CDK4/6 in both structure and function (*11*, *53*). On the other hand, yeast ribosomal protein Yar 1 has three ARs but does not interact with CDK4/6 or any other CDK (*20*, *58*). Moreover, among INK4 proteins, p16 binds CDK4 more tightly than CDK6 while the reverse is true for p18 (*59*, *60*). In addition, as revealed in the crystal structures of the p16−CDK6, p19−CDK6, and p18−CDK6−viral cyclin D complexes (*50*−*52*) and other biochemical studies (*55*), loop 4 of p18 contributes to binding

Table 1: Residues of CDK4/6 Involved in Binding to p16 Protein[a]

CDK6 residue	CDK4 residue	interaction with p16
C15	P8	nonpolar contact with Y44
V16	V9	nonpolar contact with V51
A17	A10	nonpolar contact with V51
E18	E11	nonpolar contact with Y44
F28	Y21	nonpolar contact with Y44
K29	K22[b]	nonpolar contact with V51, H-bonding with M52 backbone
R31	R24[b]	nonpolar contact with T77, H-bonding with D74 and D84 side chains
L33	P26	nonpolar contact with A76, T77, and W110
G36	S28	H-bonding with R87
G37	G29	H-bonding with R87
F39	F31	nonpolar contact with M54
D102	D97	nonpolar contact with F90, H-bonding with M54 backbone
Q103	Q98	nonpolar contact with F90, H-bonding with G55 backbone
K111	K106	H-bonding with D92 side chain
S155	S150	nonpolar contact with F90, H-bonding with E88 backbone
L166	L161	nonpolar contact with R22
A167	A162	H-bonding with R22
R168	R163	H-bonding with E27
T49	N41[b]	

[a] All interactions are derived from the crystal structure of the p16−CDK6 complex (*50*). [b] CDK4 residues in which mutations have been found in human cancers (*83*).

to CDK6 while the corresponding loop in p19 does not contact CDK6 and p16 does not even have this loop.

Second, different AR proteins may bind to the same target but modulate the target differently. For example, in addition to INK4 proteins, two AR proteins, gankyrin (seven ARs) and IκBα (six ARs), have been reported to physically contact CDK4 and influence the kinase activity in vivo and in vitro (*59*−*63*) (Figure 5a). Gankyrin is an oncoprotein with multiple functions (see below) and is able to compete with p16 and p18 for CDK4 binding, thus counteracting the inhibitory activities of p16 and p18 (*59*). IκBα usually acts as a specific inhibitor of NF-κB (a transcription factor controlling vital genes involved in immune response, cell growth and differentiation, cell adhesion, and apoptosis) (*28*, *29*), but it has the potency to bind to and inhibit CDK4 (*63*). While the CDK4 binding affinity follows the order p16 > p18 > IκBα ≈ gankyrin (*59*, *63*), some CDK4 residues important for binding to p16 and p18 play only minor roles in binding to IκBα, suggesting that there are noticeable differences among p16, p18, IκBα, and gankyrin in binding to CDK4 (*63*).

Third, an AR protein may bind to multiple targets. Besides CDK4 and CDK6, p16 binds to and negatively regulates transcription factor NF-κB (*62*) and c-Jun kinase (*64*) under physiological conditions (Figure 5b). As for gankyrin, its physiological partners include CDK4 (*56*, *58*), Rb (*65*), S6 ATPase of the 26S proteasome (*61*), MAGE-A4 antigen (*66*), and HDM2 (*67*) (Figure 5c). Binding of gankyrin to Rb, S6 ATPase of the 26S proteasome, and HDM2 facilitates the ubiquitin-mediated degradation of Rb, HDM2, and other proteins (*65*, *67*−*69*), while binding of gankyrin to MAGE-A4 quenches the oncogenic activity of gankyrin through unknown mechanisms (*66*). Protein fragmentation studies have demonstrated that the first four ARs of gankyrin are involved in interacting with CDK4 and the last two ARs are responsible for binding to Rb (*59*). The multiple specificities of ARs are potentially physiologically significant, through which distinct pathways are cross-linked or coordinated (Figure 5d). For example, the dual specificities of p16 and IκBα could cross-link the CDK4/6−Rb pathway and the NF-κB pathway, thus further contributing to the coordination between cell cycle progression and cell differentiation (*62*, *63*). However, more in vivo studies are needed to address the coordination among different interactions and pathways, including whether an AR protein may bind to different target proteins and function in different pathways at different intracellular locations or at different phases of the cell cycle.

Last, ARs are different from other protein-binding domains such as SH2 and SH3. On one hand, ARs do not recognize specific sequences of targets, unlike SH2 that recognizes sequences with phosphorylated Tyr (*70*) and SH3 that binds to sequences that are usually proline rich and exhibit a polyproline II helix conformation (*71*). On the other hand, AR binding involves discontinuous patches of residues dispersed in the whole molecules of both the AR protein and its target (*18*, *57*, *72*), while SH2 or SH3 binding usually involves a localized region of SH2 or SH3 and a very short motif in the target. From this point of view, binding of AR proteins to targets is similar to the association of Fv and Fab fragments with antigens. Antibodies have similar structures but different Fv regions. A number of residues at the Fv regions are solvent-accessible, and they stack together to form a large and diverse surface exposed to antigens (*2*, *4*, *25*, *73*). One can further extend this analogy to say that AR and antibodies are similar in that they both hold and present loops for binding interactions, using helices and β-sheets, respectively (*18*).

Consensus Design and Combinatorial Library of ARs

Developed in the past decade, consensus design is an approach to improving natural proteins by replacing each residue with the corresponding consensus amino acid on the basis of statistical analyses of sequence alignments of homologous proteins (*74*, *75*). The fundamental idea underlying this approach is that as a result of diversification and selection during protein evolution, the residues that are important in maintaining the structure or function of a protein are the ones that are conserved among homologous proteins (*24*). As demonstrated in previous studies (*8*, *76*−*78*),

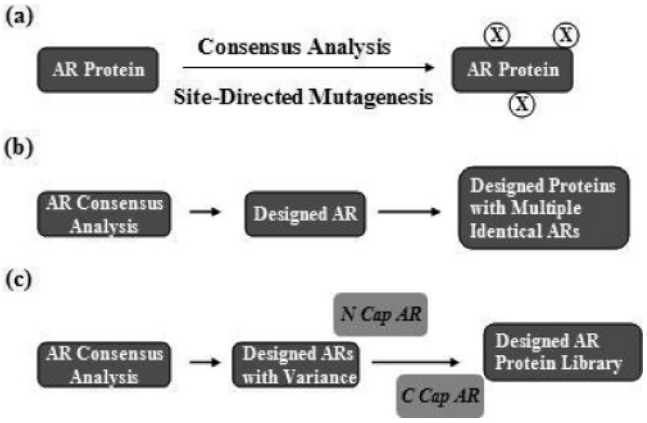

FIGURE 6: Consensus design of AR proteins. (a) Site-directed mutagenesis of p16 based on consensus sequence analyses. X represents a point mutation in p16. (b) Consensus design of AR proteins with multiple identical ankyrin repeats. (c) Design of an AR protein library based on consensus analyses.

proteins generated through consensus design usually maintain their biological function while having improved stability and folding efficiency.

The sequence abundance makes AR proteins suitable for consensus design. Consensus design in AR proteins has been pursued through different approaches, and its efficacy has been demonstrated in a number of recent studies (*14*, *15*, *36*, *79*−*82*). In one approach, information derived from sequence consensus analyses was incorporated with rational design to generate mutant AR proteins with improved thermal and chemical stability. A good example is the design of a hyperstable p16 by Cammett et al. (*79*). In this work, a series of p16 mutants based on homology-based sequence alignments and ankyrin repeat consensus analyses were generated and a p16 mutant was identified (W15D/L37S/L121R) (Figure 6a) that had a stability higher than that of wild-type p16. While this triple mutant retains its biological function, binding to CDK4, further studies showed that this triple mutant has more structural tolerance because oncogenic p16 mutations impaired the structure and function of wild-type p16, but not this hyperstable triple mutant (*79*).

In another approach, information from consensus analyses was used to generate a novel AR motif, and multiple identical ARs were stacked together to form novel AR proteins (Figure 6b). Mosavi et al. (*14*) applied a multiple-sequence alignment and statistical analysis to ~4400 identified AR sequences to calculate the probability of amino acid usage at each position of AR and designed a novel AR module with a consensus sequence as shown in Figure 1a. Subsequently, this module was used to generate proteins containing different copies (1−4) of ARs. Whereas these designed AR proteins fold and unfold in a manner similar to that of natural AR proteins, only AR proteins containing multiple consensus-designed ARs are well folded to form thermodynamically stable structures. These proteins display higher thermostabiltiy than natural proteins of the same size. Moreover, crystal structures of AR proteins containing three and four identical consensus-designed AR modules revealed that consensus-designed AR modules are packed in a linear array to form an elongated, modular molecule stabilized by inter- and intrarepeat interactions as observed in natural AR proteins.

In the third approach, consensus design was combined with a combinatorial library to generate AR proteins containing a common structural framework and novel binding specificities (*15*, *36*, *80*−*82*) (Figures 1a and 6c). For binding scaffolds such as AR proteins, its diversity in binding specificity reflects its diversity in function. While structurally important residues of AR proteins are those framework residues that mediate conserved intra- and inter-repeat interactions to maintain the defined topology, the solvent-exposed residues on the malleable and elongated surfaces contribute to the binding specificity of AR proteins and vary considerably. Hence, homologous alignment and statistical analyses of AR sequences normally result in a consensus sequence that reflects only structurally important residues. In contrast, consensus analyses of homologous globular proteins often reveal a consensus sequence which reflects both structurally and functionally important residues (*25*). However, this "weakness" in consensus design of AR proteins is also an advantage in generating AR proteins with novel binding specificity. On the basis of sequence consensus analyses, a novel design strategy was successfully developed to construct combinatorial AR protein libraries to select novel specific binders. In this approach (*15*), a consensus sequence of the AR module derived from 2220 AR sequences was designed with seven less-conserved positions (*1*, *3*, *10*, *22*, *25*) randomly filled by various amino acids (for position 26, H, N, or Y; for the other six positions, any amino acid but not G, C, or P) to generate libraries containing different AR modules. Subsequently, these AR modules were linked randomly to form AR proteins containing multiple AR modules. Furthermore, capping repeats derived from GABP-β were added to both the N- and C-termini of the above engineered AR proteins to shield the inner hydrophobic cores from the solvent, thus making AR proteins more soluble. A portion of this AR library containing members with two to four internal repeats flanked by N- and C-terminal capping repeats were selected for binding to specific target proteins using ribosomal display. Designed AR proteins with novel specificity for binding to aminoglycoside phosphotransferase (3′)-IIIa (APH), maltose binding protein (MBP), p38 AMP kinase, and JNK-2 have been successfully identified through this approach, thus proving the efficacy of consensus design-based AR combinatorial libraries (*81*, *82*). One of these designed AR proteins specifically binds to MBP with a K_d value of 4.4 nM, and the interactions mainly rely on the randomized positions of the design AR protein as revealed in the crystal structure of the designed AR protein−MBP complex (*80*). This indicates that consensus design of AR proteins not only ensures high stability but also increases the binding diversity of AR proteins. Currently, antibody libraries are widely used to generate Fv and Fab fragments with diverse binding specificities for therapeutic and diagnostic purposes, but poor solubility (when expressed in bacteria) and low stability under reducing conditions limit their intracellular applications. From this perspective, consensus-based AR combinatorial libraries could be a very attractive alternative since designed AR proteins have high solubility and stability under both extra- and intracellular conditions as well as diverse specificities (*2*, *4*).

Future Perspectives

As ubiquitous scaffolds mediating protein−protein interactions, AR proteins display a diversity of functions on a common structural framework. While this functional diversity

does not compromise the specificity of each AR-mediated protein−protein interaction, more structural and biochemical studies of AR proteins, their targets, and the AR−target complexes are needed for us to understand the molecular mechanisms underlying the strict target selection by AR proteins. Insights from these studies will not only widen and deepen our understanding of AR protein functions but also facilitate the endeavor to generate AR proteins with novel binding specificities for therapeutic or diagnostic purposes.

REFERENCES

1. Andrade, M. A., Perez-Iratxeta, C., and Ponting, C. P. (2001) Protein repeats: Structures, functions, and evolution, *J. Struct. Biol. 134*, 117−31.
2. Forrer, P., Stumpp, M. T., Binz, H. K., and Pluckthun, A. (2003) A novel strategy to design binding molecules harnessing the modular nature of repeat proteins, *FEBS Lett. 539*, 2−6.
3. Main, E. R., Jackson, S. E., and Regan, L. (2003) The folding and design of repeat proteins: Reaching a consensus, *Curr. Opin. Struct. Biol. 13*, 482−9.
4. Binz, H. K., and Pluckthun, A. (2005) Engineered proteins as specific binding reagents, *Curr. Opin. Biotechnol. 16*, 459−69.
5. Sedgwick, S. G., and Smerdon, S. J. (1999) The ankyrin repeat: A diversity of interactions on a common structural framework, *Trends Biochem. Sci. 24*, 311−6.
6. Breeden, L., and Nasmyth, K. (1987) Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*, *Nature 329*, 651−4.
7. Lux, S. E., John, K. M., and Bennett, V. (1990) Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins, *Nature 344*, 36−42.
8. Mosavi, L. K., Cammett, T. J., Desrosiers, D. C., and Peng, Z. Y. (2004) The ankyrin repeat as molecular architecture for protein recognition, *Protein Sci. 13*, 1435−48.
9. Michaely, P., and Bennett, V. (1992) The ANK repeat: A ubiquitous motif involved in macromolecular recognition, *Trends Cell Biol. 2*, 127−9.
10. Marcotte, E. M., Pellegrini, M., Yeates, T. O., and Eisenberg, D. (1999) A census of protein repeats, *J. Mol. Biol. 293*, 151−60.
11. Ortega, S., Malumbres, M., and Barbacid, M. (2002) Cyclin D-dependent kinases, INK4 inhibitors and cancer, *Biochim. Biophys. Acta 1602*, 73−87.
12. Tevelev, A., Byeon, I. J., Selby, T., Ericson, K., Kim, H. J., Kraynov, V., and Tsai, M.-D. (1996) Tumor suppressor p16INK4A: Structural characterization of wild-type and mutant proteins by NMR and circular dichroism, *Biochemistry 35*, 9475−87.
13. Byeon, I. J., Li, J., Ericson, K., Selby, T. L., Tevelev, A., Kim, H. J., O'Maille, P., and Tsai, M.-D. (1998) Tumor suppressor p16INK4A: Determination of solution structure and analyses of its interaction with cyclin-dependent kinase 4, *Mol. Cell 1*, 421−31.
14. Mosavi, L. K., Minor, D. L., Jr., and Peng, Z. Y. (2002) Consensus-derived structural determinants of the ankyrin repeat motif, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16029−34.
15. Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Pluckthun, A., and Grutter, M. G. (2003) Designed to be stable: Crystal structure of a consensus ankyrin repeat protein, *Proc. Natl. Acad. Sci. U.S.A. 100*, 1700−5.
16. Yuan, C., Li, J., Mahajan, A., Poi, M. J., Byeon, I. J., and Tsai, M.-D. (2004) Solution structure of the human oncogenic protein gankyrin containing seven ankyrin repeats and analysis of its structure−function relationship, *Biochemistry 43*, 12152−61.
17. Tang, K. S., Guralnick, B. J., Wang, W. K., Fersht, A. R., and Itzhaki, L. S. (1999) Stability and folding of the tumour suppressor protein p16, *J. Mol. Biol. 285*, 1869−86.
18. Li, J., Byeon, I. J., Ericson, K., Poi, M. J., O'Maille, P., Selby, T., and Tsai, M.-D. (1999) Tumor suppressor INK4: Determination of the solution structure of p18INK4C and demonstration of the functional significance of loops in p18INK4C and p16INK4A, *Biochemistry 38*, 2930−40.
19. Michaely, P., Tomchick, D. R., Machius, M., and Anderson, R. G. (2002) Crystal structure of a 12 ANK repeat stack from human ankyrinR, *EMBO J. 21*, 6387−96.
20. Lycan, D. E., Stafford, K. A., Bollinger, W., and Breeden, L. L. (1996) A new *Saccharomyces cerevisiae* ankyrin repeat-encoding gene required for a normal rate of cell proliferation, *Gene 171*, 33−40.
21. Walker, R. G., Willingham, A. T., and Zuker, C. S. (2000) A *Drosophila* mechanosensory transduction channel, *Science 287*, 2229−34.
22. Mosavi, L. K., Williams, S., and Peng, Z. Y. (2002) Equilibrium folding and stability of myotrophin: A model ankyrin repeat protein, *J. Mol. Biol. 320*, 165−70.
23. Zhang, B., and Peng, Z. Y. (2000) A minimum folding unit in the ankyrin repeat protein p16(INK4), *J. Mol. Biol. 299*, 1121−32.
24. Michel, F., Soler-Lopez, M., Petosa, C., Cramer, P., Siebenlist, U., and Muller, C. W. (2001) Crystal structure of the ankyrin repeat domain of Bcl-3: A unique member of the IκB protein family, *EMBO J. 20*, 6180−90.
25. Forrer, P., Binz, H. K., Stumpp, M. T., and Pluckthun, A. (2004) Consensus design of repeat proteins, *ChemBioChem 5*, 183−9.
26. Batchelor, A. H., Piper, D. E., de la Brousse, F. C., McKnight, S. L., and Wolberger, C. (1998) The structure of GABPα/β: An ETS domain-ankyrin repeat heterodimer bound to DNA, *Science 279*, 1037−41.
27. Foord, R., Taylor, I. A., Sedgwick, S. G., and Smerdon, S. J. (1999) X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function, *Nat. Struct. Biol. 6*, 157−65.
28. Jacobs, M. D., and Harrison, S. C. (1998) Structure of an IκBα/NF-κB complex, *Cell 95*, 749−58.
29. Huxford, T., Huang, D. B., Malek, S., and Ghosh, G. (1998) The crystal structure of the IκBα/NF-κB complex reveals mechanisms of NF-κB inactivation, *Cell 95*, 759−70.
30. Yang, Y., Nanduri, S., Sen, S., and Qin, J. (1998) The structural basis of ankyrin-like repeat function as revealed by the solution structure of myotrophin, *Structure 6*, 619−26.
31. Zhu, H., Dupureur, C. M., Zhang, X., and Tsai, M.-D. (1995) Phospholipase A_2 Engineering. The Roles of Disulfide Bonds in Structure, Conformational Stability, and Catalytic Function, *Biochemistry 34*, 15307−14.
32. Zeeb, M., Rosner, H., Zeslawski, W., Canet, D., Holak, T. A., and Balbach, J. (2002) Protein folding and stability of human CDK inhibitor p19(INK4d), *J. Mol. Biol. 315*, 447−57.
33. Yuan, C., Li, J., Selby, T. L., Byeon, I. J., and Tsai, M.-D. (1999) Tumor suppressor INK4: Comparisons of conformational properties between p16(INK4A) and p18(INK4C), *J. Mol. Biol. 294*, 201−11.
34. Zweifel, M. E., and Barrick, D. (2001) Studies of the ankyrin repeats of the *Drosophila melanogaster* Notch receptor. 2. Solution stability and cooperativity of unfolding, *Biochemistry 40*, 14357−67.
35. Zweifel, M. E., Leahy, D. J., Hughson, F. M., and Barrick, D. (2003) Structure and stability of the ankyrin domain of the *Drosophila* Notch receptor, *Protein Sci. 12*, 2622−32.
36. Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P., and Pluckthun, A. (2003) Designing repeat proteins: Well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins, *J. Mol. Biol. 332*, 489−503.
37. Devi, V. S., Binz, H. K., Stumpp, M. T., Pluckthun, A., Bosshard, H. R., and Jelesarov, I. (2004) Folding of a designed simple ankyrin repeat protein, *Protein Sci. 13*, 2864−70.
38. Bradley, C. M., and Barrick, D. (2002) Limits of cooperativity in a structurally modular protein: Response of the Notch ankyrin domain to analogous alanine substitutions in each repeat, *J. Mol. Biol. 324*, 373−86.
39. Tang, K. S., Fersht, A. R., and Itzhaki, L. S. (2003) Sequential unfolding of ankyrin repeats in tumor suppressor p16, *Structure 11*, 67−73.
40. Lubman, O. Y., Korolev, S. V., and Kopan, R. (2004) Anchoring notch genetics and biochemistry: Structural analysis of the ankyrin domain sheds light on existing data, *Mol. Cell 13*, 619−26.
41. Mello, C. C., Bradley, C. M., Tripp, K. W., and Barrick, D. (2005) Experimental characterization of the folding kinetics of the Notch ankyrin domain, *J. Mol. Biol. 352*, 266−81.
42. Bradley, C. M., and Barrick, D. (2006) The Notch ankyrin domain folds via a discrete, centralized pathway, *Structure 14*, 1303−12.
43. Wilson, J. J., and Kovall, R. A. (2006) Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA, *Cell 124*, 985−96.

44. Nam, Y., Sliz, P., Song, L., Aster, J. C., and Blacklow, S. C. (2006) Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes, *Cell 124*, 873−983.
45. Zweifel, M. E., and Barrick, D. (2001) Studies of the ankyrin repeats of the *Drosophila melanogaster* Notch receptor. 1. Solution conformational and hydrodynamic properties, *Biochemistry 40*, 14344−56.
46. Tripp, K. W., and Barrick, D. (2006) Enhancing the stability and folding rate of a repeat protein through the addition of consensus repeats, *J. Mol. Biol.* (in press).
47. Li, L., Wetzel, S., Pluckthun, A., and Fernandez, J. M. (2006) Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy, *Biophys. J. 90*, L30−2.
48. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999) Mechanical and chemical unfolding of a single protein: A comparison, *Proc. Natl. Acad. Sci. U.S.A. 96*, 3694−9.
49. Gorina, S., and Pavletich, N. P. (1996) Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2, *Science 274*, 1001−5.
50. Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D., and Pavletich, N. P. (1998) Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a, *Nature 395*, 237−43.
51. Brotherton, D. H., Dhanaraj, V., Wick, S., Brizuela, L., Domaille, P. J., Volyanik, E., Xu, X., Parisin, E., Smith, B. O., Archer, S. J., Serrano, M., Brenner, S. L., Blundell, T. L., and Laue, E. D. (1998) Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19INK4d, *Nature 395*, 244−50.
52. Jeffrey, P. D., Tong, L., and Pavletich, N. P. (2000) Structural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors, *Genes Dev. 14*, 3115−25.
53. Sherr, C. J., and Roberts, J. M. (2004) Living with or without cyclins and cyclin-dependent kinases, *Genes Dev. 18*, 2699−711.
54. Molofsky, A. V., Slutsky, S. G., Joseph, N. M., He, S., Pardal, R., Krishnamurthy, J., Sharpless, N. E., and Morrison, S. J. (2006) Increasing $p16^{INK4a}$ expression decreases forebrain progenitors and neurogenesis during aging, *Nature 443*, 448−52.
55. Krishnamurthy, J., Ramsey, M. R., Ligon, K. L., Torrice, C., Koh, A., Bonner-Weir, S., and Sgarpless, N. E. (2006) $p16^{INK4a}$ induces an age-dependent decline in islet regenerative potential, *Nature 443*, 453−7.
56. Janzen, V., Forkert, R., Fleming, H. E., Saito, Y., Waring, M. T., Dombkowski, D. M., Cheng, T., DePinho, R. A., Sharpless, N. E., and Scadden, D. T. (2006) Stem-cell aging modified by the cyclin-dependent kinase inhibitor $p16^{INK4a}$, *Nature 443*, 421−6.
57. Li, J., Poi, M. J., Qin, D., Selby, T. L., Byeon, I. J., and Tsai, M.-D. (2000) Tumor suppressor INK4: Quantitative structure-function analyses of p18INK4C as an inhibitor of cyclin-dependent kinase 4, *Biochemistry 39*, 649−57.
58. Loar, J. W., Seiser, R. M., Sundberg, A. E., Sagerson, H. J., Ilias, N., Zobel-Thropp, P., Craig, E. A., and Lycan, D. E. (2004) Genetic and biochemical interactions among Yar1, Ltv1 and Rps3 define novel links between environmental stress and ribosome biogenesis in *Saccharomyces cerevisiae*, *Genetics 168*, 1877−89.
59. Li, J., and Tsai, M.-D. (2002) Novel insights into the INK4-CDK4/6-Rb pathway: Counter action of gankyrin against INK4 proteins regulates the CDK4-mediated phosphorylation of Rb, *Biochemistry 41*, 3977−83.
60. Lapointe, J., Lachance, Y., Labrie, Y., and Labrie, C. (1996) A p18 mutant defective in CDK6 binding in human breast cancer cells, *Cancer Res. 56*, 4586−9.
61. Dawson, S., Apcher, S., Mee, M., Higashitsuji, H., Baker, R., Uhle, S., Dubiel, W., Fujita, J., and Mayer, R. J. (2002) Gankyrin is an ankyrin-repeat oncoprotein that interacts with CDK4 kinase and the S6 ATPase of the 26 S proteasome, *J. Biol. Chem. 277*, 10893−902.
62. Wolff, B., and Naumann, M. (1999) INK4 cell cycle inhibitors direct transcriptional inactivation of NF-κB, *Oncogene 18*, 2663−6.
63. Li, J., Joo, S. H., and Tsai, M.-D. (2003) An NF-κB-specific inhibitor, IκBα, binds to and inhibits cyclin-dependent kinase 4, *Biochemistry 42*, 13476−83.
64. Choi, B. Y., Choi, H. S., Ko, K., Cho, Y. Y., Zhu, F., Kang, B. S., Ermakova, S. P., Ma, W. Y., Bode, A. M., and Dong, Z. (2005) The tumor suppressor p16(INK4a) prevents cell transformation through inhibition of c-Jun phosphorylation and AP-1 activity, *Nat. Struct. Mol. Biol. 12*, 699−707.
65. Higashitsuji, H., Itoh, K., Nagao, T., Dawson, S., Nonoguchi, K., Kido, T., Mayer, R. J., Arii, S., and Fujita, J. (2000) Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas, *Nat. Med. 6*, 96−9.
66. Nagao, T., Higashitsuji, H., Nonoguchi, K., Sakurai, T., Dawson, S., Mayer, R. J., Itoh, K., and Fujita, J. (2003) MAGE-A4 interacts with the liver oncoprotein gankyrin and suppresses its tumorigenic activity, *J. Biol. Chem. 278*, 10668−74.
67. Higashitsuji, H., Higashitsuji, H., Itoh, K., Sakurai, T., Nagao, T., Sumitomo, Y., Masuda, T., Dawson, S., Shimada, Y., Mayer, R. J., and Fujita, J. (2005) The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53, *Cancer Cell 8*, 75−87.
68. Dawson, S., Higashitsuji, H., Wilkinson, A. J., Fujita, J., and Mayer, R. J. (2006) Gankyrin: A new oncoprotein and regulator of pRb and p53, *Trends Cell Biol. 16*, 229−33.
69. Higashitsuji, H., Liu, Y., Mayer, R. J., and Fujita, J. (2005) The oncoprotein gankyrin negatively regulates both p53 and RB by enhancing proteasomal degradation, *Cell Cycle 4*, 1335−7.
70. Malabarba, M. G., Milia, E., Faretta, M., Zamponi, R., Pelicci, P. G., and Di Fiore, P. P. (2001) A repertoire library that allows the selection of synthetic SH2s with altered binding specificities, *Oncogene 20*, 5186−94.
71. Panni, S., Dente, L., and Cesareni, G. (2002) In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules, *J. Biol. Chem. 277*, 21666−74.
72. Coleman, K. G., Wautlet, B. S., Morrissey, D., Mulheron, J., Sedman, S. A., Brinkley, P., Price, S., and Webster, K. R. (1997) Identification of CDK4 sequences involved in cyclin D1 and p16 binding, *J. Biol. Chem. 272*, 18869−74.
73. Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., Nilsson, A., Jansson, B., Ohlin, M., Wingren, C., Danielsson, L., Carlsson, R., and Borrebaeck, C. A. (2000) Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries, *Nat. Biotechnol. 18*, 852−6.
74. Lehmann, M., and Wyss, M. (2001) Engineering proteins for thermostability: The use of sequence alignments versus rational design and directed evolution, *Curr. Opin. Biotechnol. 12*, 371−5.
75. van den Burg, B., and Eijsink, V. G. (2002) Selection of mutations for increased protein stability, *Curr. Opin. Biotechnol. 13*, 333−7.
76. Nikolova, P. V., Henckel, J., Lane, D. P., and Fersht, A. R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability, *Proc. Natl. Acad. Sci. U.S.A. 95*, 14675−80.
77. Rath, A., and Davidson, A. R. (2000) The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis, *Protein Sci. 9*, 2457−69.
78. Steipe, B., Schiller, B., Pluckthun, A., and Steinbacher, S. (1994) Sequence statistics reliably predict stabilizing mutations in a protein domain, *J. Mol. Biol. 240*, 188−92.
79. Cammett, T. J., Luo, L., and Peng, Z. Y. (2003) Design and characterization of a hyperstable p16INK4a that restores Cdk4 binding activity when combined with oncogenic mutations, *J. Mol. Biol. 327*, 285−97.
80. Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., and Pluckthun, A. (2004) High-affinity binders selected from designed ankyrin repeat protein libraries, *Nat. Biotechnol. 22*, 575−82.
81. Amstutz, P., Binz, H. K., Parizek, P., Stumpp, M. T., Kohl, A., Grutter, M. G., and Pluckthun, A. (2005) Intracellular kinase inhibitors selected from combinatorial libraries of designed ankyrin repeat proteins, *J. Biol. Chem. 280*, 24715−22.
82. Kohl, A., Amstutz, P., Parizek, P., Binz, H. K., Briand, C., Capitani, G., and Pluckthun, A. (2005) Allosteric inhibition of aminoglycoside phosphotransferase by a designed ankyrin repeat protein, *Structure 13*, 1131−41.
83. Holland, E. A., Schmid, H., Kefford, R. F., and Mann, G. J. (1999) CDKN2A (P16(INK4a)) and CDK4 mutation analysis in 131 Australian melanoma probands: Effect of family history and multiple primary melanomas, *Genes, Chromosomes Cancer 25*, 339−48.

BI062188Q