

## Comparisons of the Conformational Stability of Cyclin-Dependent Kinase (CDK) 4-Interacting Ankyrin Repeat (AR) Proteins<sup>†</sup>

Yi Guo,<sup>‡</sup> Anjali Mahajan,<sup>§</sup> Chunhua Yuan,<sup>||</sup> Sang Hoon Joo,<sup>§</sup> Christopher M. Weghorst,<sup>¶,§</sup> Ming-Daw Tsai,<sup>§,‡,##</sup> and Junan Li<sup>\*,¶,§</sup>

<sup>‡</sup>*Ohio State Biochemistry Program*, <sup>§</sup>*Department of Chemistry*, <sup>||</sup>*Campus Chemical Instrument Center*, <sup>¶</sup>*Division of Environmental Health Sciences, College of Public Health*, <sup>§</sup>*Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210*, and <sup>##</sup>*Genomics Research Center and Institute of Biological Chemistry, Academia Sinica, Taiwan*

Received December 9, 2008; Revised Manuscript Received March 22, 2009

**ABSTRACT:** Ankyrin repeat (AR) proteins are one of the most abundant repeat protein classes in nature, and they are involved in numerous physiological processes through mediating protein/protein interactions. The repetitive and modular architecture of these AR proteins may lead to biochemical and biophysical properties distinct from those of globular proteins. It has been demonstrated that like most globular proteins, AR proteins exhibit a two-state, cooperative transition in chemical- and heat-induced unfolding. However, the biophysical characteristics underlying such cooperative unfolding remain to be further investigated. In the present study, we evaluated the conformational stability of a group of cyclin-dependent kinase (CDK) 4-interacting AR proteins, P16, P18, I $\kappa$ B $\alpha$ , gankyrin, and their truncated mutants under different conditions, including the presence of denaturants, temperature, and pH. Our results showed that the first four N-terminal ARs are required to form a potent and stable CDK4 modulator. Moreover, in spite of their similarities in skeleton structure, CDK4 binding, and cooperative unfolding, P16, P18, I $\kappa$ B $\alpha$ , and gankyrin exhibited considerably different biophysical properties with regard to the conformational stability, and these differences mainly resulted from the discrepancies in the primary sequence of the relatively conserved AR motifs. Our results also demonstrated that these sequence discrepancies are able to influence the function of AR proteins to a certain extent. Overall, our results provide important insights into understanding the biophysical properties of AR proteins.

The ankyrin repeat (hereafter, abbreviated as AR)<sup>1</sup> is one of the few highly abundant repeat motifs found in proteins in nature (1–3). Each AR consists of 30–34 amino acid residues and exhibits a canonical helix–turn–helix conformation, in which the helices are

arranged in an antiparallel fashion and the turns (loops) point outward at an approximately 90° angle (4). Various amounts of ARs stack together in nearly linear arrays to form AR proteins in a similar topology in which  $\alpha$  helices of ARs are stacked together to form the hydrophobic core of the proteins, and the fingers in  $\beta$  turns project away from these helices (2, 3). Unlike most globular proteins, whose structures are mainly stabilized by numerous intramolecular interactions between distant residues around a central hydrophobic core, AR proteins only have local, short-range intramolecular interactions, and to date, no disulfide bond has been found in AR proteins (2). Surprisingly, unfolding/refolding studies on a number of AR proteins (5) have demonstrated that like most globular proteins, AR proteins exhibit a two-state transition in chemical- and heat-induced unfolding, indicative of a cooperative unfolding of AR proteins (6–11). The question is whether biophysical characteristics underlying such cooperative unfolding of AR proteins are different from those of globular proteins. For example, since intramolecular interactions are present only between residues in the same AR motif or in the neighboring AR motifs (2, 3), the

<sup>†</sup>Research work on ankyrin repeat proteins in the authors' laboratory was supported by a grant from NIH (CA69472) to M.-D.T.

\*To whom correspondence should be addressed. Phone: 1-614-292-4066. Fax: 1-614-292-4053. E-mail: li.225@osu.edu.

<sup>†</sup>Abbreviations: AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride; AR, ankyrin repeat; CD, circular dichroism; CDK4, cyclin-dependent kinase 4;  $D_{1/2}$ , the denaturant concentration at the midpoint of transition during chemical-induced unfolding; DTT, dithiothreitol; E2Fs, transcription factors E2F; EDTA, ethylene-diamine-tetra-acetic acid; EGTA, ethylene glycol-bis ( $\beta$ -aminoethyl ether); GdnHCl, guanidinium hydrochloride; GST, glutathione-S-transferase; HEPES, 4-(2-hydroxyethyl)-1-piperazine-ethane sulfonic acid; HSQC, heteronuclear single-quantum coherence; INK4, specific inhibitors of CDK4, including p15<sup>INK4B</sup> (P15), p16<sup>INK4A</sup> (P16), p18<sup>INK4C</sup> (P18), and p19<sup>INK4D</sup> (P19); IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; MDM2, mouse double minute 2; NMR, nuclear magnetic resonance spectroscopy; PBS, phosphate buffer saline; Rb, retinoblastoma susceptible gene product; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis;  $T_m$ , the temperature at the midpoint of transition during thermal denaturation.

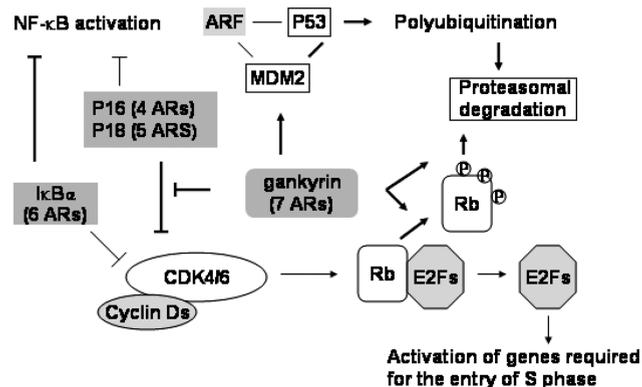


FIGURE 1: The involvement of AR proteins in the CDK4-Rb-E2F pathway. Arrows represent positive regulations, and bars represent negative regulations. E2Fs, transcription factors E2F; INK4, inhibitors of CDK4; Rb, retinoblastoma susceptible gene product; P, phosphorylated.

global stability of an AR protein is assumed to be an accumulation of the intrinsic stability of each AR motif and the interfacial stability between neighboring AR motifs (12, 13), and each sequence variance in the AR motif presumably could contribute to the stability of the AR motifs as well as the global stability of the entire protein (14). In contrast, a “minor” sequence difference generally does not affect the global stability of a globular protein as long as those crucial long-range interactions remain intact (2). From this perspective, it is of interest to compare the biophysical properties of different AR proteins, especially those with highly similar biochemical properties (function) and global structures, to understand how sequence variances in the conserved AR motif and the number of stacking AR motifs influence the global stability.

To address the above premise, we compared the conformational stability of a group of AR proteins, namely, P16, P18, IκBα, and gankyrin, under different conditions. P16, P18, IκBα, and gankyrin are exclusively composed of four, five, six, and seven ARs (15–18), respectively, and all of them are competent in physically interacting with cyclin-dependent kinase 4 (CDK4) thus modulating the tumor suppressive CDK4-Rb-E2F pathway (19) (Figure 1). P16 and P18 are specific CDK4 inhibitors and negatively regulate CDK4-mediated phosphorylation of Rb at the G1-to-S transition (20); IκBα primarily functions as a specific inhibitor of NFκB in cell signaling and differentiation (16, 17), but it is also able to interact and modulate CDK4 in a way similar to that of P16 or P18 (21). The roles of gankyrin in modulating the CDK4-Rb-E2F pathway are diverse (18, 22–25). On one hand, gankyrin binds to Rb and facilitates the proteasome-mediated degradation of Rb, which subsequently activates transcription factors E2Fs (18, 24, 25). On the other hand, gankyrin competes with P16 for CDK4 binding, thus precluding the inhibition of P16 and concomitantly triggering those E2F-controlled genes required for the entry of S phase (26). In addition, gankyrin is able to bind to MDM2, and promote ubiquitination of P53, which may indirectly influence the CDK4-Rb pathway through P53/P21 interaction (of note: P21 is a universal inhibitor of CDKs) (23, 25). Apparently, these four AR proteins

are biochemically similar in binding to CDK4, but the physiological effects after CDK4 binding are different: binding of gankyrin to CDK4 stimulates CDK4-mediated phosphorylation of Rb (26), while P16, P18, and IκBα negatively regulate the kinase activity of CDK4 as well as subsequent E2F-mediated transcription (15, 21, 27). In this study, we first compared the structures of the CDK4-binding domains in P16, P18, IκBα, and gankyrin, that is the N-terminal four ARs are very similar, which provides a common structural background for further biochemical and biophysical studies. Subsequently, we generated a series of “functional” truncated mutants of IκBα and gankyrin, and evaluated the conformational stability of these full-length and truncated AR proteins under different conditions, including the presence of denaturants, temperature, and pH. Our results demonstrated that while they follow a two-state transition in both chemical- and heat-induced unfolding, these AR proteins have different biophysical properties, such as their structural stability under different pH conditions and structural tolerance to truncation, and our studies further showed that biophysical differences of these AR proteins mainly result from the discrepancy in the primary sequence of conserved AR motifs. Further studies on P16 point mutants revealed that such sequence discrepancy also contributes to the function and/or structure of AR proteins. Taken together, our results provide important insights into understanding the unique biophysical properties of AR proteins in general.

## MATERIALS AND METHODS

*Cloning, Expression, and Purification of P16, P18, IκBα67-302, Gankyrin, and Their Mutants.* As previously described (15, 21, 26, 27), human cDNAs encoding P16, P18, IκBα67-302, and gankyrin were cloned into pGEX vectors (Amersham) and expressed in *Escherichia coli* BL21 (DE3) codon plus cells (Novagen) as glutathione-S-transferase-fusion proteins (GST) upon IPTG induction. GST-fusion proteins were purified from the cell lysate using a reduced glutathione-agarose column (Sigma) followed by a Q Fastflow column (Pharmacia). To purify free P16 and P18, 100 units of Thrombin (Sigma) was added to GST-P16 and GST-P18 proteins in PBS buffer (pH 7.4) at the final concentration of 2 units/mL (15, 27). After incubation at room temperature for 2 h, the reaction mixture was loaded onto a PBS-equilibrated reduced glutathione-agarose column to remove GST, and the flow through was further purified by an S100 column (Pharmacia) equilibrated with 5 mM HEPES, 1 μM EDTA, and 1 mM DTT (pH 7.4). Free gankyrin and IκBα67-302 proteins were purified similarly except that PreScission protease (Amersham) was used to remove the GST tag and the incubation was performed at 4 °C for 24 h (21, 26). Of note, IκBα67-302, rather than the full-length IκBα, was used in this study due to the following reasons. First, IκBα67-302 was well expressed and folded in bacteria. Secondly, IκBα67-302 is the domain responsible for binding to CDK4 (21). In addition, IκBα1-207 was used in this study to evaluate how the N-terminus influences the conformation stability of an AR protein.

All point mutants and C-terminal truncation mutants were generated through Quickchange site-directed mutagenesis (Stratagene) using corresponding pGEX-full-length AR plasmids as a template. Genes encoding N-terminal truncation AR mutants were amplified and subcloned into pGEX-6p vector as full-length AR proteins. The expression and amplification of mutant AR proteins were identical to that of wild-type AR proteins.

**In Vitro CDK4 Kinase Assay.** The in vitro CDK4 activity assay was performed as previously described (27). Briefly, each reaction mixture contained about 0.3  $\mu$ g of recombinant CDK4/cyclin D2 holoenzyme and varying concentrations of AR proteins in 15  $\mu$ L of the kinase buffer, 50 mM HEPES (pH 7.4) including 10 mM MgCl<sub>2</sub>, 2.5 mM EGTA, 0.1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM NaF, 10 mM  $\beta$ -glycerolphosphate, 1 mM DTT, 0.2 mM AEBF, 2.5 mg/mL leupeptin, and 2.5 mg/mL aprotinin. After incubation at 30 °C for 30 min, 50 ng of GST-Rb791-928 and 5  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]ATP were added in the reaction mixture which was then incubated at 30 °C for another 15 min. Proteins in the reaction mixture were separated by SDS-PAGE, and the incorporation of <sup>32</sup>P into GST-Rb791-928 was quantitatively evaluated using a PhosphorImager (Molecular Dynamics) (21). The concentrations of AR proteins were determined using absorbance at 280 nm in the kinase buffer and the molar extinction coefficients from Prot-Param at ExPASy. The IC<sub>50</sub> value was defined as the concentration of kinase inhibitor required for 50% of the maximal inhibition of CDK4 (27), and measurements were repeated in triplicate.

Recombinant CDK4/cyclin D2 holoenzyme was expressed and purified as follows (27). Human *CDK4* and *cyclin D2* cDNAs were cloned into pBacBAK8 and pBacBAK6 vectors, respectively. Of note, a His6 tag was fused to the C-terminus of CDK4 to facilitate subsequent protein purification. Each construct was cotransfected into *Spodoptera frugiperda* SF-9 cells with *Autographa California* nuclear polyiruhedrosis virus *BacPAK6/Bsu-361* DNA (BD Clontech) to generate baculovirus particles. Both baculovirus particles were cotransfected into HighFive insect cells (Invitrogen), and the CDK4/cyclin D2 holoenzyme was purified through affinity chromatography using Talon resin (BD Clontech). The final product was concentrated to approximately 0.3 mg/mL in the above kinase buffer, and aliquots were stored at -80 °C.

The substrate of CDK4, GST-Rb791-928 was expressed and purified from bacteria as previously described (21) using a pGEX-2T vector (Amersham) in which human *Rb* cDNA, corresponding to residues 791–928 was fused to the 3' end of *GST* in the open reading frame.

**P16-Counteracting Assay.** To check the counteraction between gankyrin (or gankyrin mutants) and P16 (26), 0.3  $\mu$ g of recombinant CDK4/cyclin D2 holoenzyme was mixed with varying amounts of gankyrin in 15  $\mu$ L of the kinase buffer, and incubated at 4 °C for 30 min. Subsequently, P16 was added to the reaction mixture to the final concentration of 100 nM, and after incubation at 30 °C for 30 min, the CDK4 kinase activity was evaluated as described above using 50 ng of GST-Rb791-928 and 5  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]ATP as substrates.

Of note, in the absence of gankyrin, 100 nM of P16 usually inhibits 80–90% of the CDK4 kinase activity in the reaction mixture (26).

**Circular Dichroism (CD) Analyses of AR Proteins.** For guanidinium chloride (GdnHCl)-induced unfolding, recombinant AR proteins were dissolved in 20 mM sodium borate (pH 7.4) containing 40  $\mu$ M DTT and dialyzed against this borate buffer at 4 °C overnight. Samples containing 2.5–5.0  $\mu$ M proteins were incubated with different amounts of GdnHCl (in a stock solution of 8.5 M) on ice overnight and then equilibrated at 20 °C just prior to CD analysis (28). The rotation at 222 nm was measured on an AVIV far-UV spectropolarimeter using a quartz microcell (Helma) of 0.1 cm light pass length, and the exact concentrations of GdnHCl were determined using its reflex index, and three scans were averaged. In this study, the ellipticity at 222 nm, an indicator of the existence of  $\alpha$ -helical secondary structure was taken as the measure of the degree of structure present in the protein at each GdnHCl concentration (29), and the free energy of protein denaturation in aqueous condition was obtained on the basis of two-state approximation (30). The concentrations of AR proteins in the borate buffer were determined as described earlier.

For CD analyses under different pH conditions, AR proteins were dialyzed against borate buffer (20 mM sodium borate, 40  $\mu$ M DTT) of different pH conditions (8.4, 7.4, 6.4, 5.6, and 4.4) at 4 °C overnight. 5.0  $\mu$ M of these protein solutions were then loaded in the above AVIV far-UV spectropolarimeter and 10 scans of CD spectra between 190 and 260 nm were collected at a scan speed of 50 nm per minute with a 1 nm slit and 1 nm bandwidth. The data was averaged and further normalized by baseline subtraction using the borate buffer at the same pH. Molar ellipticity was acquired using the following equation,

$$[\theta] = \theta / (10 \times n \times C \times L)$$

where  $\theta$  is the angle of circular dichroism (millidegrees),  $n$  is the number of amino acids,  $C$  is the solute molarity and  $L$  is the cuvette path length (cm).  $\alpha$ -Helical content was determined as previously described (29).

Heat-induced unfolding experiments were performed using 2.5  $\mu$ M proteins in the borate buffer (pH 7.4) with 1 nm bandwidth and 10 s response time. Thermal melting spectra were recorded at 222 nm by heating from 3 to 65 °C with a rate of 1 °C/min and a 1 °C interval and then cooling down to 3 °C at the same rate.  $T_m$  was defined as the temperature at the midpoint of transition.

**<sup>1</sup>H, <sup>15</sup>N-HSQC NMR of AR Proteins in the Presence of GdnHCl.** To perform HSQC NMR experiments under denaturing conditions, lyophilized <sup>15</sup>N-labeled AR proteins were reconstituted in 5 mM HEPES-1 mM DTT-5  $\mu$ M EDTA (pH 7.4) in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. Each denaturation mixture included a fixed amount of <sup>15</sup>N-labeled AR protein and various amounts of GdnHCl (in 8.5 M stock solution in the above HEPES buffer), and the final AR protein concentration ranged from 0.1 to 0.3 mM (8). After incubation on ice overnight, the samples were kept at 20 °C for 1 h, and subsequently, 2D <sup>1</sup>H-<sup>15</sup>N HSQC NMR experiments were

performed on a Bruker DMX-600 spectrometer at 20 °C. The exact GdnHCl concentration of each sample was determined by refractive index as before.

**Structural Modeling.** All structural modeling and comparisons were performed using DS VISUALIZER (ACCELRY) and MOLMOL.

## RESULTS

**Structural and Biochemical Comparisons of the First Four ARs in P16, P18, IκBα, and Gankyrin.** Ankyrin repeat is a relatively conserved motif, and there is a consistent pattern of key residues among different ARs to retain the characteristic topology of a canonical helix–turn–helix conformation (4, 31, 32). From this point of view, all AR proteins have a certain sequence homology and assume similar secondary structures and tertiary folds. As shown in Figure 2, the structures of the first four ARs of P16 (27), P18 (15), IκBα (16), and gankyrin (33) are almost superimposable with each other, especially in the helical regions, while most of structural variations are located within the “flexible” loop regions (linking neighboring helices) and the N- and C-termini (27). Since most of contacts between AR proteins and their binding partners are located within the helical regions (3), the high structural resemblance among P16, P18, IκBα, and gankyrin provides the basis for their potential functional similarity. In addition to residues universally conserved for the formation of the helix–turn–helix conformation of an AR motif, such as the Val/Ile-Val-Xxx (hydrophilic)-Leu/Val-Leu-Leu stretch (4, 32), residues mainly contributing to CDK4 binding, such as E26, D74, and D92 of P16, are also conserved among P16, P18, IκBα, and gankyrin (33) (Figure 3). Yet there are “notable” structural and sequence variations among these proteins. For example, some loop regions of P18, IκBα, and gankyrin are conformationally more “rigid” than those of P16 (15–17, 33), and the critical residue responsible for CDK4 inhibition (in addition to binding), D84 of P16 is not conserved in IκBα (21) and gankyrin (27, 34). Overall, with regard to CDK4 binding, P16, P18, IκBα, and gankyrin are highly similar to each other.

Table 1 summarizes the CDK4-modulating activities of P16, P18, IκBα67-302, and gankyrin. On one hand, P16, P18, and IκBα67-302 all negatively modulate CDK4 upon physical association, and they have comparable IC<sub>50</sub> values, indicating that they are biochemically indistinguishable in CDK4 inhibition (21, 27, 28). On the other hand, gankyrin binding to CDK4 does not directly influence the latter’s kinase activity but precludes binding and inhibition of P16 (as well as P18) to CDK4 (26). However, a recent study in our laboratory (34) has demonstrated that the failure of gankyrin in inhibiting CDK4 results from the lack of a negatively charged residue in the proximity of the active site of CDK4, and an I79D mutation enables gankyrin to be a CDK4 inhibitor as potent as P16. These findings, together with the fact that the first four N-terminal ARs in both IκBα67-302 and gankyrin are required and sufficient for CDK4 modulating, verified the biochemical identity of P16, P18, IκBα67-302, and gankyrin in interacting with CDK4, and enabled us to compare the

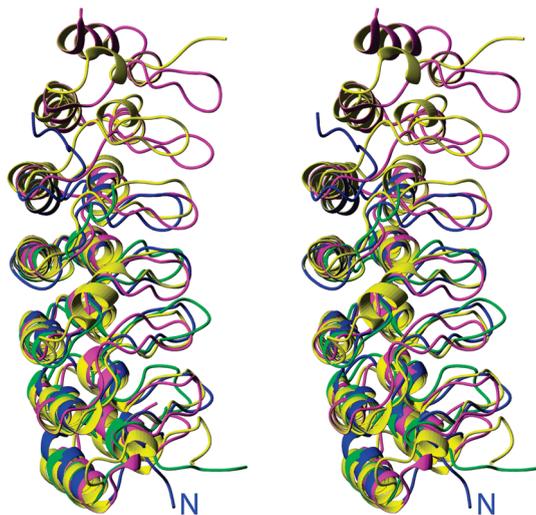


FIGURE 2: Tertiary structures of CDK4-interacting AR proteins, P16, P18, IκBα67-302, and gankyrin. The first four ARs of P18, IκBα67-302, and gankyrin are superimposed with the corresponding ARs in P16. Whereas the coordinates of IκBα67-302 (yellow) are derived from the crystal structure of IκBα/NF-κB complex (16), the structures of P16 (green) (27), P18 (blue) (15), and gankyrin (magenta) (33) are solution structures solved by NMR in our laboratory.

P16	9	MEPSADWLATAAARGRVEEVRALLEAGALP	NAP
P18	2	AEPWGNELASAAARGDLEQLTSLQLQNNVNV	NAQ
IκBα	71	TEDGDSFLHLAIIEHEEKALTMVEVIRQVKGDLAFLNFO	
Gank	3	GCVSNLMVCNLAYSGLKEELKESILADKSL	ATRT
P16	42	NSYGRRP IQVMM - GSARVAELLLLHGAEPNCA	
P18	35	NGFGR TALQVMKL - GNPEIARRLLRGANPDLK	
IκBα	107	NNLQQ TPLHLAVITNQPEIAEALLGAGCDPELR	
Gank	37	DQDSR TALHWACSAGHTEIVEFLLQLGVPVNDK	
P16	74	DPATLTRPVHDAAREGFLDTLVVLRHAGARL	DVR
P18	67	D - RTGF AVIHDAARAGFLDTLQTLLEFQADV	NIE
IκBα	140	DFR <u>GNTPLHL</u> LACEQGCLASVGVLTQSCTTPHLHSILKAT	
Gank	70	DD - AGWS <u>PLHIA</u> ASAGRDEIVKALLGKGAQV	NAV
P16	108	DAWGRLPVDLAEELGHRDVARYLRAAAGGTRGS	
P18	100	DNEG N <u>PLHL</u> AAKEGHLRVVEFLVKHTASNVGH	
IκBα	181	NYNGHT <u>CLHL</u> ASIHGYLGIVELLVSLGADVNAQ	
Gank	103	NQNGC <u>TPLH</u> YAASKNRHEIAVMLLEGGANPDAK	

FIGURE 3: Structure-based sequence alignments of P16, P18, IκBα67-302, and gankyrin. Only the four N-terminal ARs are presented here. Conserved TPLH motif and its close variants are underlined, and aspartate and glutamate residues involved in binding to CDK4 are in gray.

biophysical properties of these AR proteins in a similar biochemical and structure topological scenario. To ensure that all biophysical comparisons would be done in a similar biochemical scenario, only the truncated mutants with substantial CDK4-modulating activities (more than 10% of the CDK4-modulating activity of wild type proteins) were investigated in this study (Table 1). Moreover, since proteins expressed in inclusion bodies could be “misfolded” during purification, only truncated mutants expressed in bacteria in soluble form (like their full-length counterparts) were analyzed. The criterion of more than 10% of the CDK4-modulating activity of wild type proteins was set based on the

Table 1: CDK4-Modulating Activities of Ankyrin Repeat Proteins<sup>a</sup>

protein	size <sup>b</sup> (number of residues)	number of ARs included <sup>c</sup>	pI <sup>d</sup>	CDK4-modulating activity <sup>e</sup>
<b>P16</b>	<b>156 + 1</b>	<b>4</b>	<b>5.52</b>	<b>IC<sub>50</sub> = 72 ± 17 nM<sup>f</sup></b>
P16 1-112	112 + 1	3	5.87	IC <sub>50</sub> = 1.25 ± 0.43 μM
P16 1-79	79 + 1	2	4.91	IC <sub>50</sub> > 3.00 μM <sup>e</sup>
P16 41-156	85 + 1	3	5.67	precipitates
<b>P18</b>	<b>168 + 15</b>	<b>5</b>	<b>6.05</b>	<b>IC<sub>50</sub> = 80 ± 20 nM<sup>f</sup></b>
P18 1-137	137 + 15	4	6.28	IC <sub>50</sub> = 1.17 ± 0.28 μM
P18 1-102	102 + 15	3	4.74	IC <sub>50</sub> = 2.45 ± 0.53 μM
P181-72	72 + 15	2	8.34	IC <sub>50</sub> > 3.00 μM <sup>e</sup>
<b>IκBα67-302</b>	<b>235 + 8</b>	<b>6</b>	<b>4.67</b>	<b>IC<sub>50</sub> = 85 ± 18 nM<sup>f</sup></b>
IκBα67-243	176 + 8	5	5.32	IC <sub>50</sub> = 90 ± 25 nM
IκBα67-214	147 + 8	4	5.33	IC <sub>50</sub> = 116 ± 20 nM
IκBα1-207	207 + 8	4	4.98	IC <sub>50</sub> = 100 ± 23 nM
IκBα67-184	117 + 8	3	5.35	IC <sub>50</sub> = 1.45 ± 0.31 μM
<b>gankyrin</b>	<b>226 + 4</b>	<b>7</b>	<b>5.71</b>	<b>1.20 ± 0.25 μM requested to counteract 100 nM P16<sup>f</sup></b>
gankyrin1-201	201 + 4	6	5.45	1.48 ± 0.38 μM requested to counteract 100 nM P16
gankyrin1-171	171 + 4	5	5.88	1.42 ± 0.40 μM requested to counteract 100 nM P16
gankyrin1-138	138 + 4	4	5.41	1.63 ± 0.52 μM requested to counteract 100 nM P16
gankyrin1-105	105 + 4	3	5.00	no CDK4 binding and no counteraction against P16
gankyrin36-226	171 + 4	6	5.68	aggregates seriously

<sup>a</sup> P16, P18, and IκBα proteins bind to CDK4, and inhibit its kinase activity. Thus, their CDK4-modulating activity was represented by IC<sub>50</sub>, which was defined as the concentration of AR proteins required for 50% of the maximum inhibition. Gankyrin competes with P16 for CDK4 binding, and its CDK4-modulating activity was defined as the concentration of gankyrin proteins required to counteract against the inhibition of 100 nM P16 in the reaction mixture. The estimated error in the determination of IC<sub>50</sub> or the concentration to counteract 100 nM P16 is ± 20%, and a 3-fold change in the IC<sub>50</sub> values was regarded as biochemically significant (15, 27, 35). <sup>b</sup> After removal of the N-terminal GST tag from GST-fusion AR proteins by thrombin (for P16 and P18) and PreScission protease (for IκBα and gankyrin), there are additional residues at the N-termini of AR proteins: 1 for P16; 15 for P18; 8 for IκBα; 4 for gankyrin. It has been demonstrated that these additional residues at the N-terminus do not cause detectable changes on the structure and activity of AR proteins (27). <sup>c</sup> Determined by Pfam HMM search (www.ExPasy.com) and structural analyses of AR proteins. <sup>d</sup> Determined by ProtParam (www.ExPasy.com). <sup>e</sup> The highest concentration of AR proteins tested in the in vitro kinase assay was 3.0 μM, and an IC<sub>50</sub> value higher than 3.0 μM indicated that at the tested concentration range of AR proteins, the CDK4 kinase activity was higher than 50% of that in the absence of any modulating AR protein in the reaction mixture (26). <sup>f</sup> Similar data were previously reported by our group (8, 15, 27). However, assays were repeated simultaneously with other proteins in our current study.

following findings. Our previous studies on P16 and P18 have demonstrated that a 3-fold change in the IC<sub>50</sub> value is biochemically significant, and for those truncated mutants, a 10-fold decrease in the IC<sub>50</sub> value is mainly attributed to significant structural perturbation (15, 35). Moreover, all truncated gankyrin and IκBα mutants retaining more than 10% of the CDK4-modulating activity of wild type proteins (Table 1) are well structured as evidenced by NMR (data not shown).

**Chemical-Induced Unfolding of AR Proteins.** Guanidinium hydrochloride (GdnHCl)-induced unfolding of AR proteins was monitored by following the changes in ellipticity at 222 nm, and the unfolded fraction, defined as the ratio of the difference in ellipticity at 222 nm between the tested sample and the native state and the difference between the completely unfolded state and the native state, was derived from the raw data and plotted against the concentration of denaturant (34) (Figure 4). The unfolding curves of these AR proteins can be well fitted to a model with a two-state transition between native and unfolded proteins, indicating that GdnHCl-induced unfolding of these AR proteins is cooperative (30, 36). The values of  $\Delta G_d^{\text{water}}$  (the denaturation free energy in water),  $D_{1/2}$  (the denaturant concentration at the midpoint of transition), and the slope  $m$  were obtained on the basis of two-state approximation (30, 37) (Table 2). While the value of  $\Delta G_d^{\text{water}}$  of most proteins falls within 5–15 kcal/mol (2), all tested P16, P18, IκBα67-302, gankyrin, and their truncation mutants have fairly low values of  $\Delta G_d^{\text{water}}$  (less than 4 kcal/mol), indicating that

these AR proteins are thermodynamically unstable (3). Since the  $m$  values of these AR proteins are different from each other, the difference in the  $\Delta G_d^{\text{water}}$  values cannot be directly interpreted as the difference in conformational stability (37). However, IκBα67-302, gankyrin, and their truncation mutants have rather broad and less sigmoidal denaturation curves (in comparison with P16 and P18). Upon fitting to the two-state model, the  $D_{1/2}$  values of IκBα67-302, gankyrin, and their truncation mutants (ranging from 1.25 to 1.54 M) are significantly higher than that of P16 or P18 (0.60 and 0.79 M for P16 and P18, respectively), suggesting that P16 and P18 are less structured at high concentrations of the denaturant (GdnHCl).

We also used NMR spectroscopy to monitor the structural changes in GdnHCl-induced unfolding of P16, P18, IκBα1-207, and gankyrin (Figure 5). In the absence of GdnHCl, P16, P18, IκBα1-207, and gankyrin are well folded, and the peaks in 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra of these AR proteins are well dispersed and distinguishable. As the denaturant concentration increased, the chemical shift dispersion gradually collapsed to an extent resembling that of a random coil, indicating that the overall folded states of these proteins were perturbed. Our results showed that the unfolding transition of these AR proteins occurred in a relatively small range of GdnHCl concentrations, and no “novel” structural peaks were observed between 8.5 and 12 ppm, indicating that there is no detectable intermediate in the unfolding of these AR proteins, thus further supporting the notion that GdnHCl-induced unfolding of AR proteins exhibit

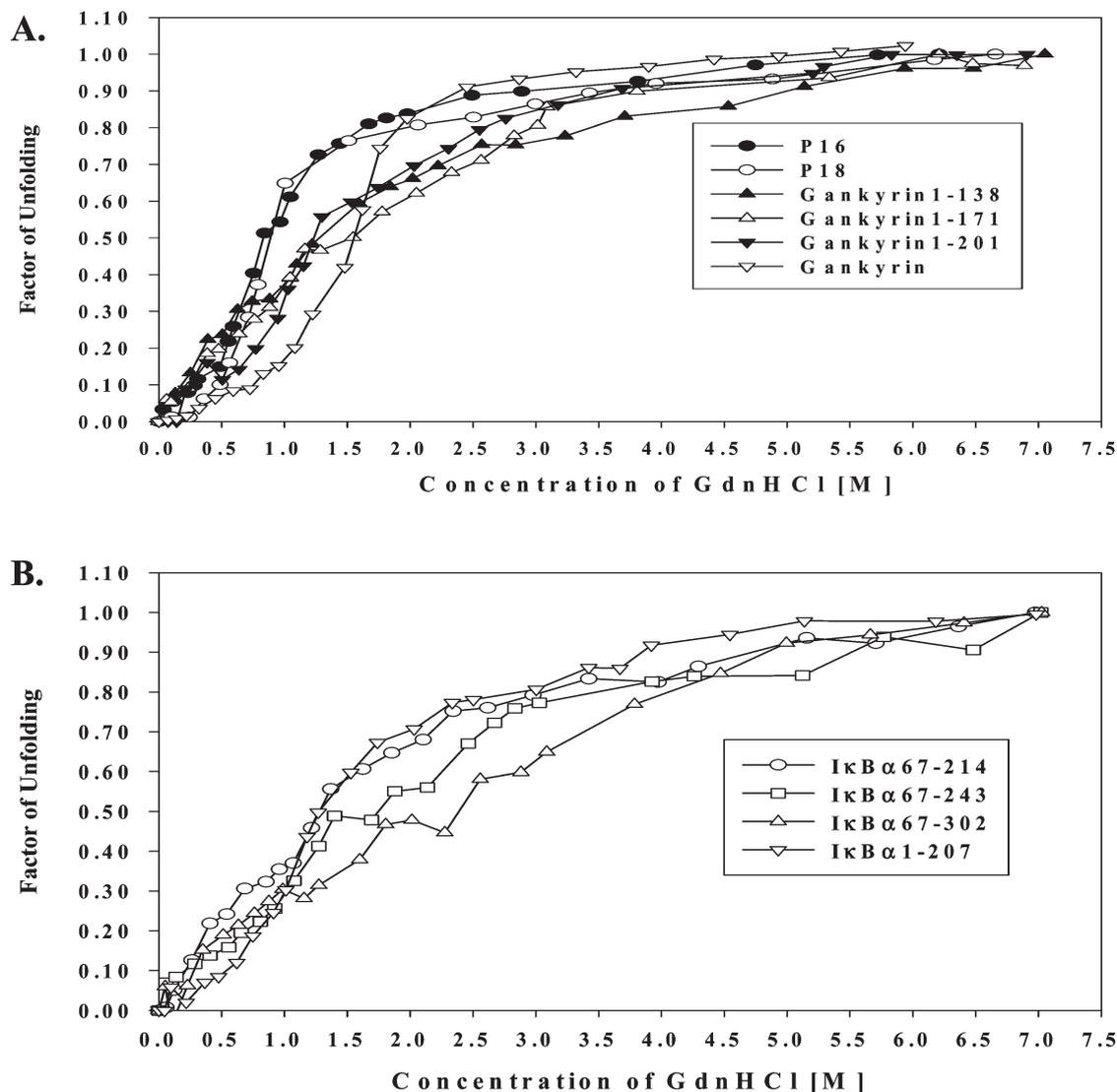


FIGURE 4: (A, B) GdnHCl-induced unfolding curves of AR proteins at 20 °C monitored by far-UV CD. Samples containing about 2.5–5.0  $\mu$ M proteins were incubated with different amounts of GdnHCl on ice overnight and the ellipticity at 222 nm was monitored by far-UV CD (190–260 nm) at 20 °C. The concentration of GdnHCl was determined using its refractive index (34). The fraction unfolded, defined as (the ellipticity at 222 nm at a denaturant concentration – the ellipticity at 222 nm at the native state)/(the ellipticity at 222 nm at the fully unfolded state – the ellipticity at 222 nm at the native state), was plotted against the GdnHCl concentration. The unfolding curve was fitted to a model of two-state approximation (30).

Table 2: Conformational Stability of CDK4-Interacting Ankyrin Proteins<sup>a</sup>

protein	size <sup>b</sup> (number of residues)	$\Delta G_d^{\text{water}}$ (kcal mol <sup>-1</sup> )	$D_{1/2}$ (M)	$m$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$T_m^c$ (°C)
P16 <sup>d</sup>	156 + 1	1.94	0.60	3.23	46.5
P18 <sup>d</sup>	168 + 15	2.98	0.79	3.77	49.9
IκBα67–302	235 + 8	1.29	1.52	0.81	46.5
IκBα67–243	176 + 8	1.56	1.38	1.06	50.2
IκBα67–214	147 + 8	1.20	1.26	0.89	52.7
IκBα1–207	207 + 8	2.21	1.28	1.72	55.1
gankyrin	226 + 4	2.77	1.54	1.80	51.5
gankyrin1–201	201 + 4	2.13	1.25	1.72	50.1
gankyrin1–171	171 + 4	1.38	1.30	1.06	46.0
gankyrin1–138	138 + 4	1.06	1.26	0.80	45.5

<sup>a</sup>  $\Delta G_d^{\text{water}}$ ,  $D_{1/2}$ , and  $m$  values were calculated according to a two-state transition model (30), and the error limit in  $\Delta G_d^{\text{water}}$  is estimated to be  $\pm 0.5$  kcal/mol (49). <sup>b</sup> See the footnote b of Table 1. <sup>c</sup>  $T_m$  was defined as the temperature at the midpoint of transition (31), and the error is estimated to be  $\pm 0.5$  °C. <sup>d</sup> Data from GdnHCl titration of P16 and P18 were cited from ref 8.

a two-state transition (30). Of note, IκBα1-207, not IκBα67-302, was used in the current NMR study mainly because IκBα67-302 has poor solution behavior within

the tested concentration range (0.2–0.4 mM) as evidenced by broadened line width and loss of major NMR signals (data not shown).

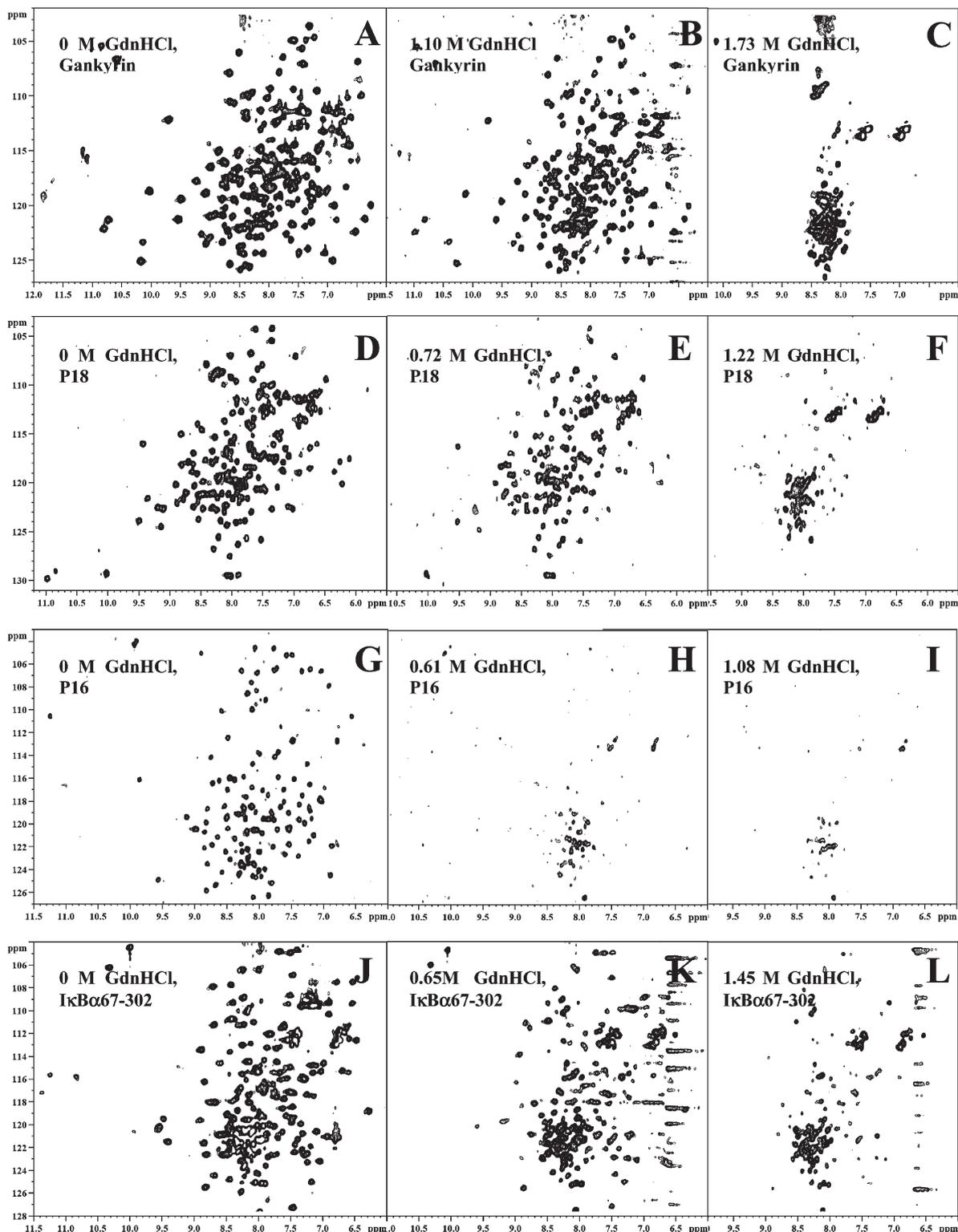


FIGURE 5: (A–L) GdnHCl-induced unfolding of AR proteins at 20 °C monitored by  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC NMR. Each sample contained a fixed amount of  $^{15}\text{N}$ -labeled AR proteins (the final concentration was about 0.2 mM) and various amounts of GdnHCl in 0.6 mL of 5 mM HEPES-1 mM DTT-5  $\mu\text{M}$  EDTA (pH 7.5) in 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$ . After incubation on ice overnight, the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiment was performed at 20 °C using a Bruker DMX-600 spectrometer. The exact GdnHCl concentration in each sample was determined by refractive index after the NMR experiment.

*Heat-Induced Unfolding of AR Proteins.* Subsequently, unfolding of AR proteins at different temperatures ranging from 3 and 65 °C was investigated, and changes in far-UV CD signals at 222 nm were monitored (34). As shown in Figure 6, heat-induced unfolding of each AR protein follows a two-state transition, implying that

heat-induced unfolding of AR proteins is also cooperative. The  $T_m$  values, temperatures at the midpoint of transition are listed in Table 2. For all these AR proteins, their heating and cooling curves between 30 and 65 °C are superimposable, and the  $T_m$  values determined from the heating and cooling curves are essentially the same

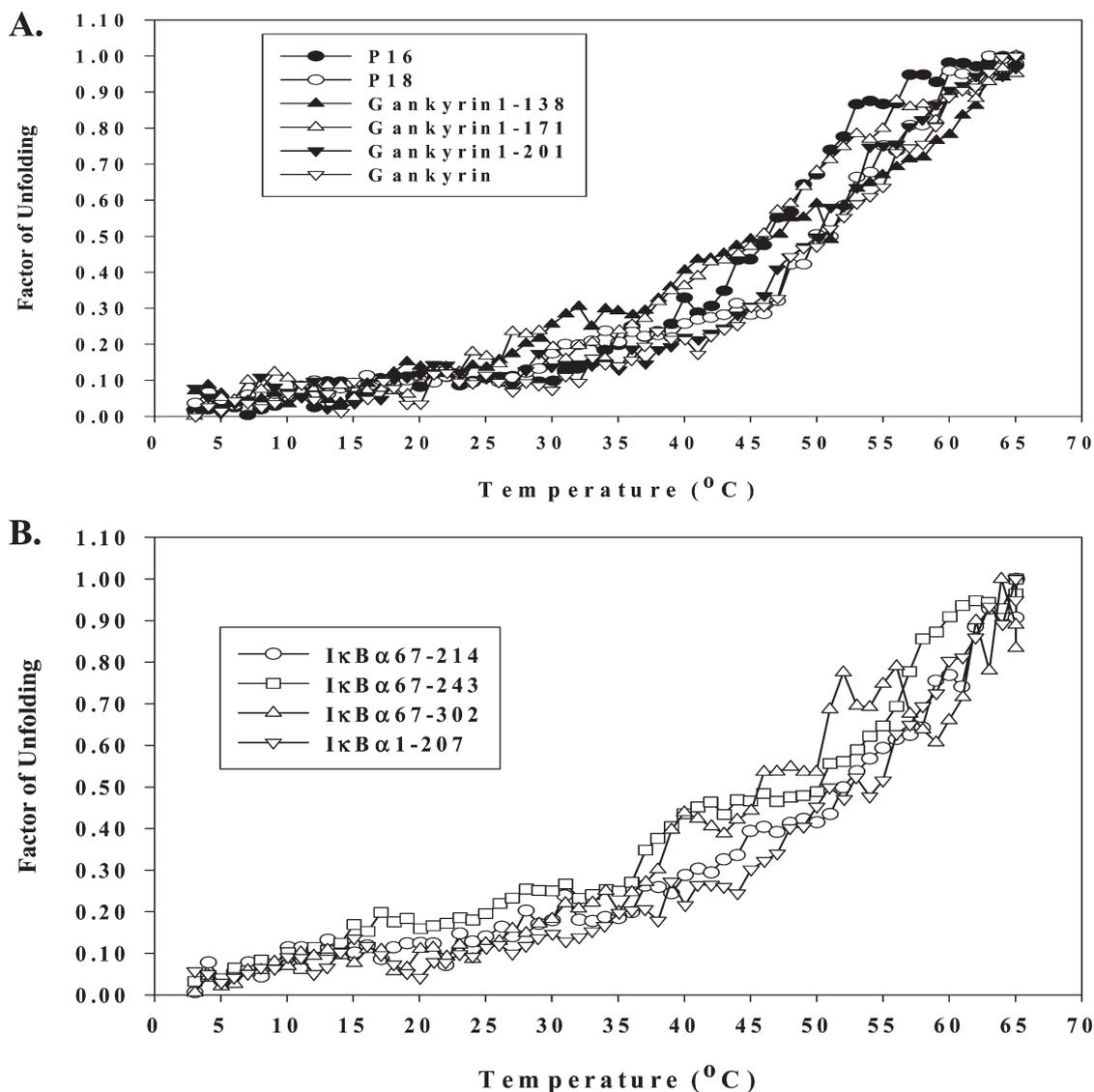


FIGURE 6: (A, B) Heat-induced unfolding curves of AR proteins monitored by far-UV CD. Thermal melting spectra were recorded at 222 nm by heating from 3 to 65 °C with a rate of 1 °C/min and a 1 °C interval (34). The fraction unfolded was defined as (ellipticity at 222 nm at certain temperature – ellipticity at 222 nm at the native state)/(ellipticity at 222 nm at the fully unfolded state – ellipticity at 222 nm at the native state).  $T_m$ , the temperature at the midpoint of transition, was obtained through fitting the melting curve to a two-state transition model (30).

within 1 °C, indicating that these proteins are at thermodynamic equilibrium and the temperature-induced folding/unfolding is reversible (data not shown) (31, 34). The  $T_m$  values of P16, P18, IκBα67-302, and gankyrin are 46.5, 49.9, 46.5, and 51.5 °C, respectively. Apparently, gankyrin and P18 are more stable than P16 and IκBα67-302 in heat-induced unfolding. As for gankyrin truncation mutants, the bigger the truncation, the lower the  $T_m$  value, implying that removal of one to three ARs from the C-terminus of gankyrin destabilize gankyrin in heat-induced unfolding. In contrast, truncated IκBα mutants, IκBα67-243, and IκBα67-214 have higher  $T_m$  values than IκBα67-302, suggesting that the last two ARs at the C-terminus of IκBα negatively contribute to its stability in heat-induced unfolding. Moreover, IκBα1-207 (55.1 °C) has a higher  $T_m$  value than IκBα67-214 (52.7 °C) indicating that the presence of an N-terminal tail (residues 1–66) positively influences the stability of IκBα.

*The Conformational Stability of AR Proteins under Different pH Conditions.* Lastly, we analyzed the

conformational stability of these AR proteins under different pH conditions. Instead of measuring the  $\Delta G_d^{\text{water}}$  values through chemical-induced denaturation under different pH conditions, changes in the secondary structure composition, that is, the  $\alpha$ -helical content, were used to evaluate the pH effect on the conformation of a protein. CD spectra from 190 nm through 260 nm of P16, P18, IκBα67-302, and gankyrin under different pH conditions were recorded, and the secondary structure composition was determined quantitatively (29). As listed in Table 3, AR proteins are more structured at high pH (pH 8.4) than at low pH (pH 4.4); however, pH effect on the conformation of AR proteins varies. For gankyrin and P18, there are no significant changes in the  $\alpha$ -helical content within a wide pH range (pH 4.4–8.4), indicating that the conformation of gankyrin or P18 is relatively stable under different pH conditions even though the CDK4-binding and modulating ability of gankyrin or P18 varies significantly with pH (data not shown). The  $\alpha$ -helical content of P16 significantly decreases at low pH

Table 3:  $\alpha$ -Helical Contents of AR Proteins at Different pH<sup>a</sup>

AR protein	$\alpha$ -helical content (%) at different pH				
	pH 8.4	pH 7.4	pH 6.4	pH 5.4	pH 4.4
<b>P16</b>	<b>32</b>	<b>29</b>	<b>35</b>	<b>34</b>	<b>15</b>
<b>P18</b>	<b>35</b>	<b>31</b>	<b>30</b>	<b>36</b>	<b>25</b>
<b>I<math>\kappa</math>B<math>\alpha</math>67-302</b>	<b>12</b>	<b>14</b>	<b>16</b>	<b>16</b>	<b>8</b>
I $\kappa$ B $\alpha$ 67-243	22	25	26	28	16
I $\kappa$ B $\alpha$ 67-214	29	36	39	36	5
I $\kappa$ B $\alpha$ 1-207	29	29	25	29	11
<b>gankyrin</b>	<b>30</b>	<b>37</b>	<b>37</b>	<b>40</b>	<b>30</b>
gankyrin1-201	26	26	29	8	13
gankyrin1-171	32	29	31	26	30
gankyrin1-138	30	30	38	29	30

<sup>a</sup>  $\alpha$ -Helical content was determined as previously described (29), and the calculation error was estimated to be  $\pm 5.0\%$ .

(pH 4.4), indicating that the tertiary structure of P16 is sensitive to low pH. In comparison with other AR proteins, I $\kappa$ B $\alpha$ 67-302 has very low molar ellipticity (data not shown) under all tested pH conditions, which makes the determination of the secondary structure composition insignificant (this issue will be further addressed in the Discussion). Interestingly, I $\kappa$ B $\alpha$ 67-243, I $\kappa$ B $\alpha$ 67-214, and I $\kappa$ B $\alpha$ 1-207 all show molar ellipticity comparable to P16, P18, and gankyrin, and they are very stable under tested pH conditions except pH 4.4. As for gankyrin truncated mutants, gankyrin1-138 and gankyrin1-171 are as stable as gankyrin wild type in a wide pH range (pH 4.4–8.4); however, gankyrin1-201 is partially unfolded at pH 4.4 and 5.4, suggesting that the removal of the last AR sensitizes gankyrin to pH. A possible explanation is that after the removal of the last (seventh) AR, the sixth AR is exposed to the aqueous environment (instead of the seventh AR), and this exposition could be conformationally unfavorable and pH-sensitive, thus destabilizing the global structure (12).

## DISCUSSION

Recently, a number of AR proteins (including natural and engineered ARs) have been biophysically characterized (2, 3, 5). While most of such characterizations focus on the unfolding and refolding of a specific AR protein and/or certain mutants, there are few studies on comparing the biophysical properties of different AR proteins, especially those natural AR proteins from different protein families. Apparently, the functional diversity of AR proteins makes it challenging to compare the biophysical properties of AR proteins under a highly similar or identical biochemical scenario (4). Here we present our biophysical investigation on the stability of four different AR proteins, P16, P18, I $\kappa$ B $\alpha$ 67-302, and gankyrin, under different experimental conditions, and their high structural and functional (CDK4 interaction) similarities enable us to distinguish factors mainly contributing to biophysical properties from factors contributing to structure and function.

*Four N-terminal ARs Are the Minimum for a Stable and Functional CDK4 Modulator.* As shown in our *in vitro* kinase assays (Table 1), gankyrin1-201, gankyrin1-171, and gankyrin1-138 (with the deletion of one, two, and three ARs at the C-terminus, respectively) retained

almost identical P16-counteracting activities, indicating that the removal of up to three ARs at the C-terminus of gankyrin does not significantly change its biochemical properties in CDK4 binding and modulating. Similarly, I $\kappa$ B $\alpha$  mutants with one or two C-terminal ARs deleted exhibited unchanged CDK4-inhibitory activities (in comparison with I $\kappa$ B $\alpha$ 67-302). Moreover, these “functional” truncated mutants (encompassing the N-terminal four ARs) remained well structured and behaved as wild type proteins in chemical-/heat-induced folding and unfolding. In contrast, significant changes in solubility (data not shown) and CDK4 binding and modulating were brought about through the following mutations: removal of more than three C-terminal ARs of gankyrin or two C-terminal ARs of I $\kappa$ B $\alpha$ 67-302, removal of only one AR at the C-terminus of P16 or P18, and deletions of the first N-terminal AR motifs of P16, P18, I $\kappa$ B $\alpha$ 67-302, and gankyrin. Taken together, these findings indicate that the N-terminal four ARs are essential and sufficient to retain a stable and functional structure for CDK4 binding and modulation. This notion is further supported by the following observations. First, small peptides or protein fragments encompassing the second and third ARs of P16 (38) or P18 (39) have been found to retain remnant CDK4-inhibiting activities, but they are not well structured. Second, while the number of AR motifs included in a protein varies, no single AR motif has been found to form a stable structure itself (2–4). Moreover, some AR proteins containing two or three ARs are structured but not active in function. For example, Peng and his co-workers (40) previously demonstrated that P16C, a truncated P16 protein with the last two ARs appears to be well folded instead of being in a molten globule-like conformation as supported by the cooperative thermal unfolding and well-dispersed NMR spectrum of P16C. However, the conformation stability of P16C is very low, and P16C does not retain the CDK4-binding ability even though a number of CDK4-binding residues are located within P16C, suggesting that P16C is somehow different from P16. Third, even though most of contacts between P16 or P18 and CDK6 (a close homologue of CDK4) are located within the second and third ARs, as revealed in the crystal structures of P16/CDK6 and P18/CDK6/viral cyclin D complexes (41–43), it has been reported that some P16 mutants in the first and fourth ARs perturb the global structure thus significantly impairing CDK4 binding and inhibition (15, 27). It is also worthwhile to note that the removal of the last (fifth) AR motif of P18 led to significant loss of CDK4-inhibitory activities, suggesting that this C-terminal AR motif is important for P18 functioning (35, 44). A possible explanation is that while this fifth AR motif does not directly contribute to CDK4 binding as revealed in the crystal structures of P16/CDK6 and P18/CDK6/viral cyclin (41–43), the removal of this AR may expose the fourth AR to aqueous environments and significantly perturb the global structure of P18 (12).

In spite of the pivotal role of the four N-terminal ARs in CDK4 interaction as well as stabilizing the global structures, other motifs in aforementioned CDK4-interacting AR proteins have considerable effect on the biophysical and biochemical properties of these AR proteins. For example, in addition to the four ARs for CDK4

binding,  $\text{I}\kappa\text{B}\alpha$  has a flexible segment (residues 1–66) at its N-terminus and two extra ARs (the fifth and sixth ARs) at its C-terminus. On one hand, the N-terminal segment of  $\text{I}\kappa\text{B}\alpha$  increased its thermal stability as evidenced by increased  $T_m$  value of  $\text{I}\kappa\text{B}\alpha$ 1-207 (in comparison with  $\text{I}\kappa\text{B}\alpha$  mutants without the N-terminal tail), which is consistent with our previous observation that full-length P16 is more stable than P16  $\Delta$ 1–8, a truncated version of P16 without the first eight residues at the N-terminus (27). Contrarily, the existence of the two C-terminal ARs of  $\text{I}\kappa\text{B}\alpha$ 67-302 destabilized the global structure of  $\text{I}\kappa\text{B}\alpha$  to a certain extent. On the other hand, most contacts between  $\text{I}\kappa\text{B}\alpha$  and  $\text{NF}\kappa\text{B}$  are located within these two C-terminal ARs, while the N-terminal segment of  $\text{I}\kappa\text{B}\alpha$  encompasses the nuclear localization sequence (NLS) and two phosphorylation sites, which are crucial for its functioning in the regulation of  $\text{NF}\kappa\text{B}$  (16, 17). Similarly, the three C-terminal ARs of gankyrin do not only considerably increase the thermodynamic stability of the global structure, but also enable gankyrin to bind to pRb (through the CxLxE motif in the fifth AR) (18, 26) and MDM2 (through the seventh AR) (23).

*Unique Biophysical Features of the AR Proteins in Folding/Unfolding.* It has been demonstrated that most AR proteins behave in the same way as their globular counterparts and follow two-state transitions in chemical- and heat-induced unfolding (30). However, AR proteins exhibit biophysical features in folding/unfolding notably different from those of globular proteins. First, all well studied natural AR proteins (not engineered AR proteins) are thermodynamically unstable with  $\Delta G_d^{\text{water}}$  values less than 5 kcal/mol (3) mainly due to lack of long-range intramolecular interactions. Second, the elongated and repetitive nature enables AR proteins to be more tolerant to structural truncation than globular proteins (12). As described earlier, after removal of three ARs at the C-terminus (88 out of 226 residues, that is, about 39% of total residues), gankyrin1-138 remains structured and functions well in CDK4 binding and modulation. For  $\text{I}\kappa\text{B}\alpha$ 67-302, the removal of two ARs (88 residues) at the C-terminus does not impair its CDK4-modulating ability but increases the conformational stability in heat-induced unfolding. Third, in chemical-induced unfolding of globular proteins, denaturant  $m$  values, which represent the dependence of free energy of unfolding on denaturant concentration (37), have been found to be strongly correlated with the protein surface exposed to solvent upon folding. Specifically, for globular proteins following two-state folding/unfolding transitions, if there is neither oligomerization nor disulfide bonding, there presumably exists a linear correlation between the protein size and the  $m$  value (Figure 7, the straight line); that is, large proteins tend to have higher  $m$  values (12, 37). However, for AR proteins, the correlation between the protein size and the  $m$  value varies considerably. On one hand, as shown in Figure 7, the  $m$  values of P16 and P18 are very close to those of globular proteins of the same size, and P18 (5 ARs) has a higher  $m$  value than P16 (4 ARs) (3.77 kcal mol<sup>-1</sup> M<sup>-1</sup> vs 3.23 kcal mol<sup>-1</sup> M<sup>-1</sup>). Nevertheless,  $\text{I}\kappa\text{B}\alpha$ 67-302 (6 ARs) and gankyrin (7 ARs) have  $m$  values of 0.81 and 1.80 kcal mol<sup>-1</sup> M<sup>-1</sup>, respectively, much lower than those of P16, P18, and their

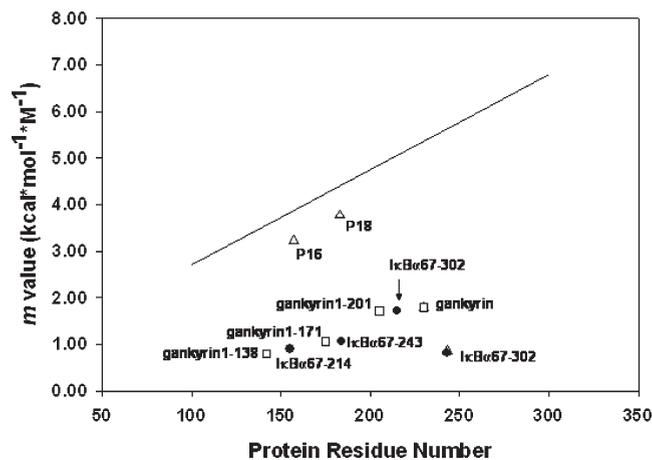


FIGURE 7: The correlation between the  $m$  value and the size of AR proteins. The  $m$  values for different AR proteins and their mutants were determined through GdnHCl titration (30). The straight line represents the linear relationship between the  $m$  value and the size of globular proteins as derived from ref 37. Empty triangles represent wild type AR proteins, and empty squares and solid circles represent truncated gankyrin and  $\text{I}\kappa\text{B}\alpha$  mutants, respectively.

globular counterparts. While the extremely low  $m$  value of  $\text{I}\kappa\text{B}\alpha$ 67-302 could be attributed to its tendency to oligomerization at tested concentrations as demonstrated in our NMR studies (data not shown), gankyrin remains a monomer in a wide range of concentrations (33). On the other hand, for gankyrin and its truncated mutants, the  $m$  value increases as the size of the protein increases; this is consistent with previous studies on Notch AR domain and its truncated mutants (5, 12). Moreover, the  $m$  values of AR proteins of nearly “identical size” differ from each other. For example, P16,  $\text{I}\kappa\text{B}\alpha$ 67-214, and gankyrin1-138 all consist of 4 ARs, but the  $m$  value of P16 is 3.23 kcal mol<sup>-1</sup> M<sup>-1</sup>, much higher than those of  $\text{I}\kappa\text{B}\alpha$ 67-214 and gankyrin1-138, 0.89 and 0.80 kcal mol<sup>-1</sup> M<sup>-1</sup>, respectively. Accordingly, the denaturant concentration at the mid point of the transition,  $D_{1/2}$  of P16 is much lower than those of  $\text{I}\kappa\text{B}\alpha$ 67-214, and gankyrin1-138 (0.60 M vs. 1.26 M, 1.26M), suggesting that  $\text{I}\kappa\text{B}\alpha$ 67-214 and gankyrin1-138 are more stable at moderate concentration of the denaturant (GdnHCl). With regard to the high similarities among P16,  $\text{I}\kappa\text{B}\alpha$ 67-214, and gankyrin1-138 in size, skeleton structure, function, and unfolding mechanism (two-state transition), the differences in their  $m$  values could mainly result from the sequence discrepancies in the conserved AR motifs. Similarly, the  $m$  value of P18 (5 ARs), 3.77 kcal mol<sup>-1</sup> M<sup>-1</sup> is much bigger than that of  $\text{I}\kappa\text{B}\alpha$ 67-243 (5 ARs) (1.06 kcal mol<sup>-1</sup> M<sup>-1</sup>) or gankyrin1-171 (5 ARs) (1.06 kcal mol<sup>-1</sup> M<sup>-1</sup>), while the  $D_{1/2}$  value of P18 (0.79 M) is much lower than that of  $\text{I}\kappa\text{B}\alpha$ 67-243 (1.38 M) or gankyrin1-171 (1.30 M). Nonetheless, it remains to be further clarified whether changes in  $m$  values reflect the differences in equilibrium unfolding mechanisms (the existence of transient and hard-to-detect intermediates) (7) or the differences in the compactness of the unfolded proteins (37). Last but not least, in addition to the sequence variances in ARs, the number of ARs also contributes to the conformational stability of AR proteins. In general, proteins containing more AR repeats have more compact and concave structures. As demonstrated earlier, there is a putative order for the

Table 4: Biochemical and Biophysical Characterization of Selected P16 Mutants

P16 mutant	IC <sub>50</sub> (nM) <sup>b</sup>	NMR structure <sup>b</sup>	CDK4-binding residue <sup>c</sup> ?	ΔG <sub>d</sub> <sup>water</sup> (kcal mol <sup>-1</sup> ) <sup>d</sup>	<i>m</i> (kcal mol <sup>-1</sup> M <sup>-1</sup> )
W15A	890 ± 140	retained	no	N/A <sup>a</sup>	N/A
E26D	280 ± 52	retained	yes	2.11	3.12
E69A	350 ± 45	retained	no	1.43	2.91
N71S	250 ± 87	retained	no	1.14	2.62
D84H	> 6000	retained	yes	1.89	3.17
F90A	200 ± 45	retained	no	1.69	3.48
D92A	850 ± 195	retained	yes	2.20	3.30
G101W	> 6000	aggregated	no	1.42	0.97
W110A	160 ± 35	retained	no	1.52	2.76
L121A	150 ± 25	retained	no	1.72	3.08

<sup>a</sup> N/A, not available due to precipitation during the CD experiment. <sup>b</sup> Data already reported in refs (27, 28), and 35. <sup>c</sup> Information derived from the crystal structures of the P16/CDK6, P19/CDK6, and P18/CDK6/viral cyclin D complexes (41–43). <sup>d</sup> See the footnote a of Table 2.

conformational stability, gankyrin (7 ARs) > P18 (5 ARs) > P16 (4 ARs). However, as the number of ARs in a protein increases, the complexity in solubility, structure, and function could increase. For example, IκBα67-302 (6 ARs) contains two “distinct” binding domains for CDK4 and NFκB, respectively, and its solubility is poor, which may lead to a conformational stability lower than that of P18 (5 ARs). Moreover, folding mechanisms for larger AR proteins could deviate from the two-state transition mechanism, which makes it challenging to evaluate the structural stability of proteins and to determine the energetic consequences of changes in proteins as well as its aqueous environments through chemical- and heat-induced unfolding. It has been reported that D34, a 426-residue fragment of ankyrinR that comprises 12 ARs, unfolds via an intermediate in which the C-terminal half is structured and N-terminal half is unfolded (45, 46).

*Sequence Variances on the Conserved AR Motif Also Contribute to the CDK4-Modulating Function.* One of the unique features of AR proteins is that the global structure of an AR protein is stabilized by the accumulation of local, short interactions originating from each individual residue (2, 3); hence, the aforementioned sequence variances may indirectly contribute to the CDK4-modulating function through destabilizing or stabilizing the global structures of these AR proteins. To address this premise, we analyzed the biochemical and biophysical properties of 10 P16 missense mutants as previously described (27, 28) (Table 4). While E26, D84, and D94 are conserved residues directly involved in CDK4 binding and modulation, W15, E69, N71, F90, G101, W110, and L121 are not conserved and belong to the aforementioned sequence variances (Figure 3). According to their mutagenic effect on function (the CDK4-inhibitory activity), structure, and conformational stability, these P16 mutants can be divided into three groups. The first group consists of E26D, D84H, and D92A. As shown in Table 4, all these three mutants retain unchanged structures, comparable conformational stabilities but decreased CDK4-inhibitory activities. While D84H does not exhibit any detectable CDK4-inhibitory activity, the activity of D92A decreases by about 10-fold in comparison with that of P16 WT. As for E26D, a moderate decrease (about 4-fold) is mainly due to the conserved nature of this substitution. These observations are in support of previous findings that these three residues mainly contribute

to CDK4 binding and modulating (27, 28, 33). The second group includes E69A, N71S, F90A, W110A, L121A, and W15A. E69A, N71S, F90A, W110A, and L121A are well structured, but both their activities and conformational stabilities decrease moderately (by 2–5-fold). As mentioned earlier, these residues are not directly involved in CDK4 binding and modulating; therefore, the partial loss of function is mainly attributed to the “destabilizing” effect caused by these substitutions, suggesting that the aforementioned sequence variances have a bearing on CDK4 binding and modulating through stabilizing/destabilizing the global conformation of these AR proteins. W15A seems to be the extreme case in this group. It remains structured, but its activity decreases significantly (by 10-fold). Moreover, W15A precipitates under the experimental condition for CD analysis (See the Materials and Methods section), indicating that this mutation brings about significant changes in the biophysical properties of P16, which consequently impair the CDK4-inhibitory activity. The last group includes only G101W. This mutant does not retain any detectable activity, and its solution structure is significantly perturbed due to aggregation. Moreover, this mutant shows a significantly lower *m* value: 0.97 vs 3.23 kcal mol<sup>-1</sup> M<sup>-1</sup> for P16 WT. These results suggest that G101W does not only alter the biophysical properties of P16, but also significantly perturbs its global structure thus eliminating its function. Taken together, our studies on P16 missense mutants demonstrate that in addition to the biophysical properties of AR proteins, sequence variances in the conserved AR motif also contribute to the functions and/or structures of these proteins. From this perspective, AR proteins are similar to their globular counterparts, in which the structural, functional, and biophysical roles of certain residues or motifs are well integrated (47). Hence, changes in the biophysical properties of AR proteins may lead to significant perturbation to their functions. This notion is of significance in biology. As shown in our current study and previous studies in our laboratory as well as other laboratories (20, 27, 28), P16 missense mutants are prevalent in human cancers, and many of these mutants do not directly disrupt its binding to CDK4, but substantially impair the conformational stability and solubility of P16 protein, thus compromising its ability in inhibiting CDK4. Interestingly, similar mutants in P18 and gankyrin do not bring about significant

changes in the conformational stability of the global structure nor the CDK4 modulating function (15, 34, 35), indicating that the low conformational stability makes P16 susceptible to mutations in cancer. In addition, from the perspective of pharmaceuticals, it is plausible to use the sequence variance to optimize the biophysical properties of P16 (as well as other AR proteins) thus enhancing its tumor-suppressing function (32, 48).

## ACKNOWLEDGMENT

The authors thank Dr. Gordon Renkes in Department of Chemistry for his help in CD spectroscopy.

## REFERENCES

- Andrade, M. A., Perez-Iratxeta, C., and Ponting, C. P. (2001) Protein repeats: structures, functions, and evolution. *J. Struct. Biol.* 134, 117–131.
- 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, 1435–1448.
- Li, J., Mahajan, A., and Tsai, M.-D. (2006) Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* 45, 15168–15178.
- Sedgwick, S. G., and Smerdon, S. J. (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem. Sci.* 24, 311–316.
- Kloss, E., Courtemanche, N., and Barrick, D. (2008) Repeat-protein folding: new insights into origins of cooperativity, stability, and topology. *Arch. Biochem. Biophys.* 469, 83–99.
- 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, 1869–1886.
- 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, 447–457.
- 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, 201–211.
- 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, 14357–14367.
- 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, 2622–2632.
- 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, 14344–14356.
- Mello, C. C., and Barrick, D. (2004) An experimentally determined protein folding energy landscape. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14102–14107.
- Tripp, K. W., and Barrick, D. (2004) The tolerance of a modular protein to duplication and deletion of internal repeats. *J. Mol. Biol.* 344, 169–178.
- 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, 373–386.
- 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, 2930–2940.
- Jacobs, M. D., and Harrison, S. C. (1998) Structure of an IkappaBalpha/NF-kappaB complex. *Cell* 95, 749–758.
- Huxford, T., Huang, D. B., Malek, S., and Ghosh, G. (1998) The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell* 95, 759–770.
- Higashitsuji, H., Itoh, K., Nagao, T., Dawson, S., Nonoguchi, K., Kido, T., Mayer, R. J., Arai, S., and Fujita, J. (2000) Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nat. Med.* 6, 96–99.
- Sherr, C. J., and Roberts, J. M. (2004) Living with or without cyclins and cyclin-dependent kinases. *Genes Dev.* 18, 2699–2711.
- Ortega, S., Malumbres, M., and Barbacid, M. (2002) Cyclin D-dependent kinases, INK4 inhibitors and cancer. *Biochim. Biophys. Acta* 1602, 73–89.
- Li, J., Joo, S. H., and Tsai, M.-D. (2003) An NF-kappaB-specific inhibitor, IkappaBalpha, binds to and inhibits cyclin-dependent kinase 4. *Biochemistry* 42, 13476–12383.
- 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, 10668–10674.
- 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.
- Dawson, S., Higashitsuji, H., Wilkinson, A. J., Fujita, J., and Mayer, R. J. Gankyrin: a new oncoprotein and regulator of pRb and p53. *Trends Cell Biol.* 16, 229–233.
- 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, 1335–1337.
- 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, 3977–3983.
- 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, 421–431.
- 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, 9475–9487.
- Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.* 6, 383–390.
- Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* 131, 266–280.
- Devi, V. S., Binz, H. K., Stumpp, M. T., Pluckthun, A., Bosshard, H. R., and Jelasarov, I. (2004) Folding of a designed simple ankyrin repeat protein. *Protein Sci.* 13, 2864–2870.
- 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, 482–489.
- 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, 12152–12161.
- Mahajan, A., Guo, Y., Yuan, C., Weghorst, C. M., Tsai, M.-D., and Li, J. (2007) Dissection of protein-protein interaction and CDK4 inhibition in the oncogenic versus tumor suppressing functions of gankyrin and p16. *J. Mol. Biol.* 373, 990–1005.
- 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, 649–657.
- 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, 165–170.
- Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4, 2138–2148.
- Fähræus, R., Paramio, J. M., Ball, K. L., Laín, S., and Lane, D. P. (1996) Inhibition of pRb phosphorylation and cell-cycle progression by a 20-residue peptide derived from p16CDKN2/INK4A. *Curr. Biol.* 6, 84–91.
- Noh, S. J., Li, Y., Xiong, Y., and Guan, K.-L. (1999) Identification of functional elements of p18INK4C essential for binding and inhibition of cyclin-dependent kinase (CDK) 4 and CDK6. *Cancer Res.* 59, 558–564.
- Zhang, B., and Peng, Z. Y. (2000) A minimum folding unit in the ankyrin repeat protein p16(INK4). *J. Mol. Biol.* 299, 1121–1132.
- 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 (London)* 395, 237–243.
42. 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 (London)* 395, 244–250.
  43. Jeffrey, P. D., Tong, L., and Pavletich, N. P. (2000) Structural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors. *Genes Dev.* 14, 3115–3125.
  44. Gump, J., Turner, S., and Koh, J. (2001) The COOH terminus of p18INK4C distinguishes function from p16INK4A. *Cancer Res.* 61, 3863–3868.
  45. Werbeck, N. D., and Itzhaki, L. S. (2007) Probing a moving target with a plastic unfolding intermediate of an ankyrin-repeat protein. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7863–7868.
  46. Werbeck, N. D., Rowling, P. J. E., Chellamuthu, V. R., and Itzhaki, L. S. (2008) Shifting transition states in the unfolding of a large ankyrin repeat protein. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9982–9987.
  47. Li, Y., and Tsai, M.-D. (1993) Phospholipase A-4 engineering. 10. The aspartate-histidine catalytic diad also plays an important structural role. *J. Am. Chem. Soc.* 115, 8523–8526.
  48. 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, 285–297.
  49. Hondal, R. J., Riddle, S. R., Kravchuk, A. V., Zhao, Z., Liao, H., Bruzik, K. S., and Tsai, M.-D. (1997) Phosphatidylinositol-specific phospholipase C: kinetic and stereochemical evidence for an interaction between arginine-69 and the phosphate group of phosphatidylinositol. *Biochemistry* 36, 6633–6642.