

# Getting specificity from simplicity in putative proteins from the prebiotic Earth

Jaime López de la Osa\*, David A. Bateman†, Sylvia Ho†, Carlos González\*, Avijit Chakrabarty†‡, and Douglas V. Laurents\*‡

\*Instituto de Química Física "Rocasolano," Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain; and †Departments of Medical Biophysics and Biochemistry, University of Toronto, Toronto, ON, Canada M5G-2M9

Communicated by Robert L. Baldwin, Stanford University Medical Center, Stanford, CA, July 26, 2007 (received for review April 20, 2007)

**Can unique protein structures arise from a limited set of amino acids present on the prebiotic Earth? To address this question, we have determined the stability and structure of KIA7, a 20-residue polypeptide containing chiefly Lys, Ile, and Ala. NMR methods reveal that KIA7 tetramerizes and folds on the millisecond time scale to adopt a four-helix X-bundle structure with a tightly and specifically packed core. Denaturation studies and hydrogen exchange measurements of KIA7 and several variants demonstrate that ridges-into-grooves packing of Ala and Ile side chains and the packing of a C-terminal aromatic group into the hydrophobic core are sufficient to give rise to a rather stable, well folded protein structure, with no favorable electrostatic interactions or tertiary or quaternary hydrogen bonds. Both modern proteins and RNAs can adopt specific structures, but RNAs do so with a limited "alphabet" of residues and types of stabilizing interactions. The results reported here show that specific, well folded protein structures can also arise from a highly reduced set of stabilizing interactions and amino acids that are thought to have been present on the prebiotic Earth.**

four-helix bundle | NMR spectroscopy | protein stability | chemical evolution | protein folding

In the prebiotic Earth, hydrophobic and charged amino acids were thought to have been rather common, whereas polar amino acids were rare (1, 2). Could such a limited set of amino acids have given rise to unique well folded proteins? To gain insight into this question, we have studied the conformational stability and structure of KIA7, a 20-residue polypeptide containing chiefly Lys, Ile, and Ala, plus a –Gly-Gly-Tyr extension for concentration determination, which self-associates to form a well packed, four-helix bundle protein.

KIA7 was previously identified from a combinatorial peptide library as having a highly helical structure (as gauged by CD), a well packed hydrophobic core (1-anilino-8-naphthalene sulfonate fluorescence), and a specific tetrameric quaternary structure (analytical ultracentrifugation) (3, 4). The positively charged lysine side chains prevent the formation of higher-order oligomers. Because no tertiary hydrogen bonds, higher-order oligomerization or crystal packing interactions, or favorable electrostatic interactions are present in KIA7, this study constitutes a test of whether or not a specific, unique structure can arise from a highly limited set of stabilizing interactions, namely, the hydrophobic effect and van der Waals interactions.

In addition to KIA7, two other members of the KIA series of peptides, KIA5 and KIA10, were studied (Table 1). These peptides are variants of the KIA7 sequence and possess lesser degrees of ordered structure (3). A stabilized KIA7 variant with a disulfide cross-link, CG<sub>3</sub>KIA7, was also examined, and additional peptides truncating the –Gly-Gly-Tyr C-terminal extension, KIA7Δ, or carrying Tyr to Phe, KIA7F, or Tyr to Ile, KIA7I, substitutions were studied to probe the possible stabilizing or structural contributions of this extension (Table 1). An additional variant, +KIA7–, was characterized to determine whether and how terminal charges affect stability and structure

formation. The results show that the KIA7 tetrameric structure is as well packed as natural four-helix bundle proteins and is stabilized by the hydrophobic effect and van der Waals interactions, with no favorable charge–charge interactions and, in the case of KIA7F, no tertiary hydrogen bonds. Helix-capping interactions that are present in natural proteins also stabilize KIA7.

## Results

**The Sequence of Ala and Ile and a C-Terminal Aromatic Are Key for Structure Formation.** The KIA peptides were first examined by CD, a probe of secondary structure. The CD spectra at 25°C of KIA5, KIA7Δ, and KIA7I in 0.2 M NaCl, pH 5 show deep minima at 202 nm and a shallow one at 220 nm [supporting information (SI) Fig. 5], which indicate low helix contents. The CD spectra of KIA10 at 25°C in 0.2 M NaCl is typical of a mostly unstructured peptide, but upon lowering the temperature to 5°C and increasing the peptide concentration the spectra becomes characteristically α-helical (data not shown). KIA7, KIA7F, and CG<sub>3</sub>KIA in 0.2 M NaCl show strong minima at 221 and 208 nm and intense maxima <200 nm, which indicates that they adopt highly helical structures. The variant with terminal charges, +KIA7–, behaves like the corresponding uncharged peptide, KIA7, except higher concentrations of +KIA7– are needed for full helix formation.

The variation of the helix content of the KIA peptides (0.120 mM) upon heating is shown in SI Fig. 5. KIA5, KIA7Δ, KIA7I, and KIA10 in 0.2 M NaCl show small to moderate losses of helix content at elevated temperatures. In 0.2 M NaCl, KIA7, +KIA7–, KIA7F, and CG<sub>3</sub>KIA7 show broad sigmoidal transitions with apparent denaturation midpoints of 33°C, 40°C, 41°C, and 71°C, respectively, which resemble the thermal unfolding transitions of small, natural proteins. The thermodynamic parameters for KIA7, +KIA7–, KIA7F, and CG<sub>3</sub>KIA7 unfolding were obtained by applying a two-state model in which dissociation and denaturation are coupled and are shown in SI Table 2 and SI Fig. 6).

NMR spectroscopy confirms that KIA5, KIA10, KIA7Δ, and KIA7I are not well folded in 0.2 M NaCl at 5°C. The 1D <sup>1</sup>H NMR spectrum of KIA5 in 0.2 M NaCl revealed sharp lines with little or no dispersion, which is characteristic of an unfolded protein

Author contributions: J.L.d.l.O., C.G., A.C., and D.V.L. designed research; J.L.d.l.O., C.G., and D.V.L. performed research; D.A.B., S.H., and A.C. contributed new reagents/analytic tools; J.L.d.l.O., C.G., A.C., and D.V.L. analyzed data; and J.L.d.l.O., C.G., A.C., and D.V.L. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2jo4 and 2jo5).

†To whom correspondence may be addressed. E-mail: chakrab@uhhres.utoronto.ca or dlaurents@iqfr.csic.es.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0706876104/DC1](http://www.pnas.org/cgi/content/full/0706876104/DC1).

© 2007 by The National Academy of Sciences of the USA

**Table 1. Sequence of KIA molecules**

Molecule	Sequence				
	1	5	10	15	20
KIA7	Ac-AKAAA	AAIKA	IAAII	KAGGY-NH <sub>2</sub> *	
KIA5	Ac-AKAIA	AIAKA	AAAI	KAGGY-NH <sub>2</sub>	
KIA10	Ac-AKAAA	AIIKA	IAAAI	KAGGY-NH <sub>2</sub>	
KIA7Δ	Ac-AKAAA	AAIKA	IAAII	KAGG-NH <sub>2</sub>	
KIA7I	Ac-AKAAA	AAIKA	IAAII	KAGGI-NH <sub>2</sub>	
KIA7F	Ac-AKAAA	AAIKA	IAAII	KAGGF-NH <sub>2</sub>	
CG <sub>3</sub> KIA7	[Ac-CGGAKAAA	AAIKA	IAAII	KAGGY-NH <sub>2</sub> ] <sub>2</sub> <sup>†</sup>	
+KIA7-	+AKAAA	AAIKA	IAAII	KAGGY-	

\*The N-terminal amine and C-terminal carboxylate group are acetylated and amidated, respectively, and are uncharged in all the peptides, except +KIA7-. Therefore, apart from +KIA7-, no favorable charge-charge interactions can be formed, as the only charged residues present are Lys.

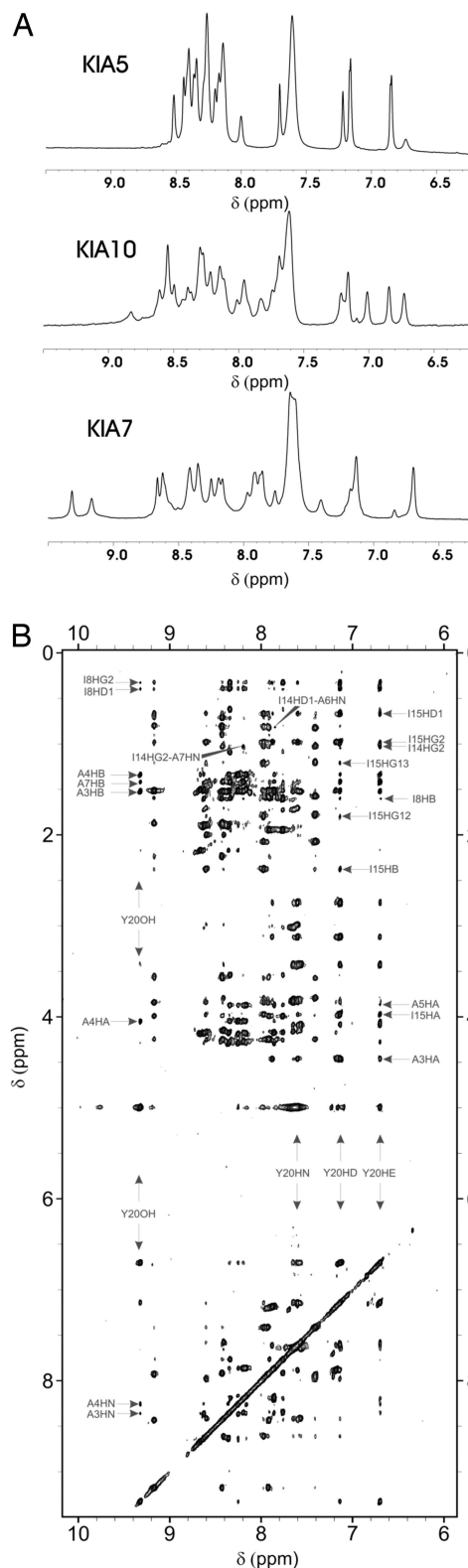
<sup>†</sup>In solution, CG<sub>3</sub>KIA7 dimerizes via disulfide bond formation. Two dimers then associate to form a tetramer.

(Fig. 1A). NOESY spectra confirmed the lack of significant conformational preferences and nonlocal contacts (SI Fig. 7). Similar results were observed for KIA7Δ, KIA7I, and +KIA7I- (data not shown). KIA10 yielded a 1D <sup>1</sup>H NMR spectrum with broadened peaks typical of a partly ordered, but fluctuating, structure (Fig. 1A).

**KIA7 Forms a Stable, Well Ordered Four-Helix Bundle Protein.** KIA7 (Fig. 1A) and +KIA7-, KIA7F, and CG<sub>3</sub>KIA7 (data not shown) in 0.2 M NaCl gave rise to highly dispersed 1D <sup>1</sup>H NMR spectra, which is a hallmark of a well ordered protein. Taking advantage of this dispersion, the <sup>1</sup>H and <sup>13</sup>C resonances of KIA7, KIA7F, and CG<sub>3</sub>KIA7 were completely assigned despite the highly repetitive sequence. The three peptides show strong sequential HN-HN NOEs extending along the sequence from the N terminus to Ala-17, and HN-H<sub>α</sub>(i, i+3, i+4) NOEs and H<sub>α</sub> and C<sub>α</sub> conformational chemical shifts, which are all indicative of a fully helical structure. Numerous side-chain NOEs, such as those between Ile-8 and Tyr-20, for example, corroborate an antiparallel orientation for the KIA7 tetramer in solution (Fig. 1B).

To study the heat-induced unfolding of KIA7, a series of 1D <sup>1</sup>H NMR spectra were recorded at different temperatures (Fig. 2). Exchange between the structured and denatured forms is on the slow-intermediate time scale. Disperse peaks are seen until 50°C; >60°C, the sharp lines characteristic of unstructured polypeptides become dominant. Interestingly, the Tyr ring protons also undergo a heat-induced transition. The Tyr ring resonances of the folded state disappear >55°C. From the largest chemical shift difference ( $\delta_{\text{denatured}} - \delta_{\text{folded}} = 0.12 \text{ ppm} \equiv 96 \text{ Hz}$  at 800 MHz), the kinetics of the folding ↔ unfolding reaction is estimated to be <600 s<sup>-1</sup>, with a kinetic lifetime of >1.7 ms, which is equivalent to a >1.1-ms half-life. Very similar results are observed for +KIA7-, except the apparent midpoint of the transition occurs near 45°C (SI Fig. 8).

The stability of KIA7, KIA7F, and CG<sub>3</sub>KIA7 at individual amide hydrogens was probed by NMR-monitored H/D exchange in 0.2 M NaCl (Fig. 3). Little or no protection against exchange was observed in KIA7I. The exchange of hydrogen-bonded amide protons was found to be slowed 20- to 80-fold in KIA7 and KIA7F and 1,000- to 10,000-fold in CG<sub>3</sub>KIA7 relative to unstructured peptides. These protection factors correspond to free energies of stability ( $\Delta G_{\text{HX}}$ ) of ≈2 kcal/mol for KIA7 and KIA7F and ≈5 kcal/mol for CG<sub>3</sub>KIA7. Chemical denaturation experiments were used to determine the free energy of conformational stability,  $\Delta G^\circ$ , and the *m*-value of KIA7, KIA7F, and +KIA7-.  $\Delta G^\circ$  was found to range from 14 kcal/mol for +KIA7- to 20 kcal/mol for KIA7. The urea *m*-value was ≈2,400 cal/mol·M for all these variants (SI Table 3, SI Text, and SI Fig. 9).

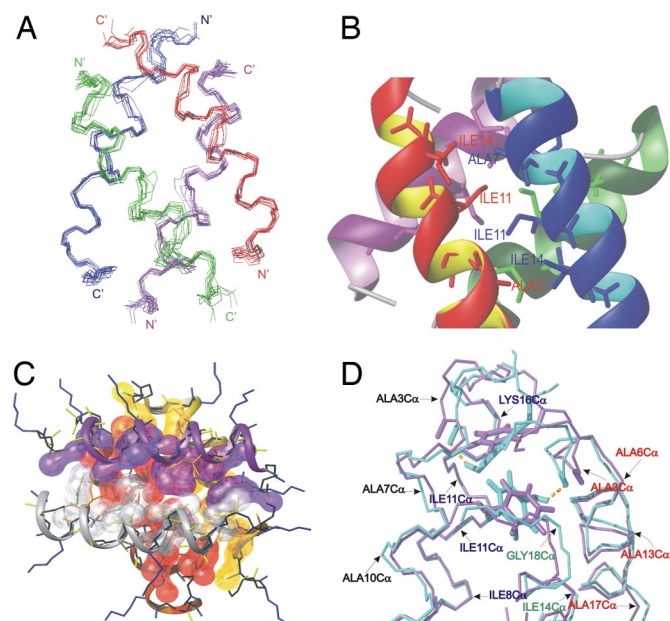


**Fig. 1.** NMR spectroscopic characterization of KIA molecules in 200 mM NaCl, 10 mM Na/HAc, pH 5.0, 5°C. (A) 1D <sup>1</sup>H NMR spectra of KIA5 (Top), KIA10 (Middle), and KIA7 (Bottom). (B) 2D <sup>1</sup>H NOESY spectrum of KIA7. Some assignments are labeled.

**KIA7I Adopts Ordered Helical Structures in Stabilizing Conditions.** The results described above indicate that KIA7, +KIA7-, KIA7F, and CG<sub>3</sub>KIA7 are folded in 0.2 M NaCl, but KIA7I is not. To







**Fig. 4.** Views of KIA7 and KIA7F NMR structures. (A) Superposition of the backbones of the 10 structures. (B and C) Two views of the packing of the Ala and Ile side chains. In C, lysine side chains are drawn as blue lines. (D) Superposition of the KIA7 (light blue) and KIA7F (violet) backbone structures and aromatic residues. The hydrogen bonds between the Tyr-20 hydroxyl and the Ala-3 carbonyl groups are shown as broken yellow lines.

## Discussion

**The Stability and Folding Kinetics of KIA7 Are Similar to Those of Natural Proteins.** The conformational stability of KIA7 and KIA7F, as determined by hydrogen exchange at 5°C, yields values of  $\approx 2$  kcal/mol. These stability values are modest; they are slightly higher than those determined for the lac repressor DNA binding domain (7) but less than the corresponding values for phage 434 Cro protein (8) and most other small, all-helical proteins. The conformational stability of CG<sub>3</sub>KIA7 is significantly higher,  $\approx 5$  kcal/mol, and similar to variants of IL-4, which also contain two disulfide bonds (9). It is remarkable that the stabilizing interactions in KIA7 are able to pay the substantial entropic penalty of oligomerization (4.5 kcal/mol) (10) and still be as stable as some single-chain proteins, which do not pay this penalty. The cost of tetramerization is also reflected in the remarkable differences in the stability values determined by hydrogen exchange near 1 mM peptide concentration and the standard free energies of stability obtained by denaturation experiments. These  $\Delta G^\circ$  values and the urea *m*-values are similar to those obtained for other oligomeric proteins (11). The lower stability of +KIA7– is probably caused by an unfavorable charge-helix macrodipole interaction at the N termini of the helices and the unfavorable desolvation energy of partially buried charges. Favorable charge-charge interactions between the charged groups at the N and C termini also probably contribute to the observed difference in stability between KIA7 and +KIA7–.

KIA7 and +KIA7– fold on the low millisecond timescale at 45°C to 55°C, as judged by the NMR chemical shift difference between the folded and unfolded states of the Tyr ring protons. This value is comparable to the folding rate of phage 434 Cro protein (12). Structural transitions in partly folded peptides are much faster; for example, the slow nucleation step in helix formation by poly L-ornithine is  $10^5$  times faster (13).

In addition to reporting on the folding rate, the Tyr ring protons give rise to a large number of NOEs, which indicates that

the ring is ordered and plays an important role in the structure. The conformation adopted by the C-terminal extension, –Gly-Gly-Tyr is that of a Gly-terminal loop seen in natural proteins (6). In contrast, this same C-terminal extension has been shown to be flexible and lack a preferred structure in monomeric helical peptides (14).

**The Presence of an Aromatic Ring at the C Terminus Is Required for Stability and Structure Formation in KIA7.** The solution structures show that the aromatic rings of Tyr-20 (KIA7) or Phe-20 (KIA7F) are well packed in the protein core, whereas peptides lacking the C-terminal aromatic group, such as KIA7I, appear to be largely unfolded in 0.2 M NaCl. The inability of KIA7I to adopt a well folded structure is probably caused by the loss of stabilizing contributions,  $\approx 0.6$  kcal/mol per aromatic >C-H group (15), from the hydrophobic effect when four Tyr (Phe) are replaced by four Ile. Nevertheless, under stabilizing conditions, i.e., 0.5 M sodium sulfate, KIA7I adopts a highly helical, tetrameric structure with hallmarks of a well packed hydrophobic center. The hydroxyl group of Tyr-20 in KIA7 forms a hydrogen bond with the carbonyl oxygen of Ala-3. The contribution of this hydrogen bond is not necessary to maintain the structure and stability of KIA7, because the KIA7F variant has essentially equivalent NMR spectra, a very similar 3D solution structure and a similar (hydrogen exchange and urea denaturation) or superior (CD-monitored thermal denaturation) conformational stability.

**Structural Comparison with Other Four-Helical Bundles.** The tertiary structure adopted by KIA7 belongs to a class of four-helix bundles described previously as the “X-bundle” (16), and it has been observed as a structural motif within natural proteins such as the globins or thermolysin (see SI Table 6 for a more complete list), but not as an independent protein, as no close matches were reported in the above cited work or in a recent searches using the Dali and BLAST programs. Therefore, to our knowledge, KIA7 appears to be the first stable isolated X-bundle structure. The stability of KIA7 is striking, because Cohen and coworkers (16) predicted the X-bundle structure to be unstable on its own, because this structure leaves many open “holes” and unsatisfied “knobs” at the helix ends. In KIA7, these open holes at the helix ends are satisfied, at least in part, by the Gly termination motif, in which the C-terminal aromatic group caps the hydrophobic core.

**The Origin of Specificity and Stability in Designed Proteins.** Many proteins that were designed by optimizing hydrophobic interactions do adopt a preferred tertiary structure but have mobile, loosely packed side chains. The ground-breaking work of De-Grado and coworkers (17–20) has led to the development of well packed helix bundle proteins with natural-like sequences and stabilizing interactions. In their work, and pioneering work from other laboratories, a specific fold has been achieved via the formation of tertiary hydrogen bonds (20–23), crystal contacts (18, 24), disulfide bonds (25, 26) or favorable charge-charge interactions (19, 20, 23, 24, 26, 27). Previous work toward simplifying proteins has shown that a native-like Src homology 3 fold can be achieved with chiefly five different residues (although 13 in total) (28), and a four-helix bundle protein can be formed from only nine different residues (29). In the latter case, from the Hecht laboratory (29), a high-resolution NMR solution structure revealed that the protein is stabilized mainly by hydrophobic interactions and van der Waals packing, although some stabilizing tertiary electrostatic interactions and hydrogen bonds are also present. Other work has demonstrated the importance of hydrophobic packing (30), particularly of Ile (31–33), in endowing proteins with a unique structure. Nakamura and colleagues (31) observed that replacing Ile by Leu

produced a loss of tight packing in a small DNA binding protein domain. Moreover, Hu and collaborators (32) showed that substituting four Ile residues endows the GCN4 leucine zipper with a unique quaternary structure. Harbury and coworkers (33) designed a protein with a novel backbone fold, taking advantage of the specific packing of Ile. The ability of Ile to afford stability and structural uniqueness stems undoubtedly from its large, hydrophobic side chain, which has few allowed rotamers caused by  $\beta$ -branching (34). In the case of KIA7F, it is shown here that ridges-into-grooves packing of nonpolar residues, Ala and Ile, and a capping C-terminal aromatic group are sufficient to give rise to a stable, unique, native-like protein structure without any favorable charge-charge interactions, crystal packing interactions, or tertiary hydrogen bonds.

Modern proteins are endowed with a large set of amino acid residues and a variety of different stabilizing interactions. However, modern RNA ribozymes adopt specific conformations and achieve efficient and selective catalysis despite being restricted to a limited “alphabet” of residues and functional groups. On the basis of the present findings, we propose that prebiotic proteins could have also adopted unique and functional structures using a much more reduced set of amino acids and stabilizing interactions.

**Implications for Prebiotic Biochemistry.** It is believed that amino acids, nucleotides, and other organic monomers emerged on Earth through either abiotic synthesis or transport by comets and meteorites (35). Nonenzymatic synthesis of oligopeptides and oligonucleotides has been accomplished under abiotic conditions (36). Although it is likely that the sequences of prebiotic peptides were generated at random, a hypothesis has been forwarded that the physicochemical environment of these peptides can exert “prebiological” selection pressure for sequences that fold into compact soluble structures (37). This hypothesis assumes that the prebiotic environment was hydrolytic toward the amide bond; hence, peptides that sequester their amide bonds avoid hydrolysis. Two methods of sequestering amide bonds are compaction of the chain into close-packed (including secondary structure formation) structure and aggregation. The concentration of the peptides in the primordial soup likely determined whether aggregated peptides or soluble folded peptides were preserved. Our results with the KIA7 series demonstrate the existence of sequences that fold into compact soluble structures using a highly reduced amino acid alphabet that is restricted to a subset of those present in the prebiotic Earth.

## Methods

**Peptide Synthesis and Purification.** Peptides were synthesized by using the Merrifield solid-phase method on a 9050 Plus Pep-Synthesizer (PerSeptive Biosystems, Framingham, MA) and later deprotected and purified by RP-HPLC as described (3). +KIA7- was purchased from the Caslo Laboratory (Lyngby, Denmark). The purity and identity of the peptides was further confirmed by MALDI-TOF mass spectroscopy. Peptide concentrations were estimated by UV absorbance for KIA peptides containing Tyr and  $\epsilon_{(280\text{ nm})} = 1,390\text{ cm}^{-1}\cdot\text{M}^{-1}$  (38) or using a bicinchoninic acid (BCA) assay for KIA peptides lacking Tyr. Tyrosine-containing KIA peptides were used as standards in the BCA assays.

**NMR Spectroscopy.** The KIA samples for NMR were dissolved in 90% H<sub>2</sub>O/10% D<sub>2</sub>O containing 0.20 M NaCl to reduce electrostatic repulsion between the charged lysine side chains, 10 mM NaAc/HAc (deuterated at the methyl group) to buffer the solution to pH 5, and 50  $\mu\text{M}$  sodium 2,2-dimethyl-2-silapentane-5-sulfonate as the internal chemical shift standard. In some experiments, 0.50 M sodium sulfate was used instead of 0.20 M sodium chloride. The final peptide concentration was 1.2 mM for KIA7, CG<sub>3</sub>KIA7, +KIA7-, and KIA5, 4.5 mM for KIA7F, 2.8

mM for KIA7 $\Delta$ , and 0.46 mM for KIA10. An additional sample of KIA7 containing only 50  $\mu\text{M}$  peptide with the same buffer solution was also studied. NMR spectra were acquired on 800- and 600-MHz AV spectrometers (Bruker, Billerica, MA), the latter equipped with a cryoprobe. NOESY (39) spectra for structure determination were acquired in 0.2 M NaCl at pH 5, 5.0°C with mixing times of 75 and 150 ms. TOCSY (40) spectra were recorded with the standard MLEV-17 spin-lock sequence and a mixing time of 80 ms. In general, water suppression was achieved by including a WATERGATE (41) module in the pulse sequence before acquisition. Spectra were processed with Bruker software and analyzed with the program SPARKY (42). This program was used to aid the manual assignment of the NOESY cross-peaks and the quantitative evaluation of the NOE intensities. The formula  $k = 2\pi\Delta\delta$ , where  $k$  is the rate constant and  $\Delta\delta$  is the chemical shift difference in Hz, was used to calculate the folding rate. NMR  $T_2$  relaxation measurements to gauge the oligomeric state of KIA molecules were performed by using a spin-echo pulse sequence (43).

**Spectral Assignments and Structure Determination.** NMR assignment was performed by standard sequence-specific methods. A large number of distance constraints could be assigned despite the intrinsic ambiguity between intermolecular and intramolecular distances in multimeric structures. Most cross-peaks of the KIA7 and KIA7F peptides could be directly assigned on the basis of a preliminary model of a tetramer consisting of four antiparallel  $\alpha$ -helices. Trial assignments were made for the few remaining cross-peaks and used in preliminary structure calculations. When a distance constraint was consistently violated in all of the resulting structures, an alternative assignment was considered. After several cycles of assignment and structure calculation, a consistent set of constraints was obtained. Any remaining cross-peaks that were still ambiguous were not included in the subsequent calculations. For KIA7, a total of 1,516 constraints, of which 232 were interhelical, were used.

Structures were calculated with the program DYANA 1.5 (44) and further refined with the molecular dynamics package AMBER 7.0 (45). The AMBER refinement consisted of an annealing protocol *in vacuo* followed by long trajectories where explicit solvent molecules were included. The temperature and the relative weights of the experimental constraints were varied during the *in vacuo* simulations according to standard annealing protocols used in our group (46, 47). The final structures were then hydrated with  $\approx 4,000$  water molecules and 12 chloride counter ions. The structures were first equilibrated for 130 ps using our standard equilibration process (48), and then subjected to 10 runs of 100-ps MD simulations in water using NMR restraints, periodic boundary conditions, and the Particle-Mesh-Ewald method (49). Simulations in water were performed in the isothermal/isobaric ensemble [ $P = 1\text{ atm}$  (1 atm = 101.3 kPa),  $T = 298\text{ K}$ ] using SHAKE in all atoms and a 2-fs time step for integration of Newton's equations. MD-averaged structures were obtained by averaging the last 250 ps of individual trajectories and further relaxation of the structure. The same procedure was used to obtain a global MD-averaged structure, resulting from averaging 10 ns of independent MD trajectories.

The programs Procheck/NMR (50) and MOLMOL (51) were used in structure analysis, molecular graphic manipulations, and detection of possible hydrogen bonds. The criteria for hydrogen bond formation were: (i) donor-acceptor distance  $< 3.0\text{ \AA}$  and (ii) the donor-proton-acceptor angle  $< 35^\circ$ . Protein structure comparisons were performed by using the Dali server (52).

**Hydrogen Exchange.** Hydrogen/deuterium exchange was initiated by adding a precooled D<sub>2</sub>O solution containing 0.20 M NaCl, 10 mM sodium acetate/acetic acid (pH 5.3), and 50  $\mu\text{M}$  sodium 2,2-dimethyl-2-silapentane-5-sulfonate to lyophilized KIA pep-

tides. Exchange was monitored by recording a series of 1D  $^1\text{H}$  spectra (16 transients with two dummy scans per spectrum, acquisition time = 45 s) at 5.0°C. The spectra were transformed, and the NH peaks were integrated by using Bruker Topspin software. The observed exchange rates of the NH protons were determined by fitting a single exponential decay function to the data by using a least-squares fitting algorithm.

The protection factor for each amide proton was calculated as the ratio of the observed exchange rate to the intrinsic exchange rate of hydrogen exchange, in short unstructured peptides using the parameters determined by Englander and coworkers (53). The free energy of stability is calculated by using the Gibbs equation,  $\Delta G = -RT \ln 1/P$ , where  $R$  is the gas constant,  $T$  is the absolute temperature, and  $P$  is the protection factor ( $= k_{rc}/k_{ex}$ ), which is the inverse of the equilibrium constant,  $K_{op}$  ( $= k_o/k_c$ ), between the open and closed states (54). This equation implicitly

assumes that hydrogen exchange proceeds via the EXII mechanism (54), which has been frequently verified for proteins <pH 7. Moreover, simulation of KIA7 hydrogen exchange at pH 5.3, 5°C using reasonable values for the opening, closing, and intrinsic rate constants are consistent with the EXII mechanism.

We thank Dr. J. M. Pérez Cañadillas and Prof. J. Santoro for assistance with NMR pulse sequences; Prof. J. Santoro for additional support (Ministerio de Ciencia y Tecnología Grant BIO2002-00720); Prof. R. M. M. Brito and Prof. M. H. Hecht for helpful discussion and unpublished data; and Drs. M. A. Jiménez, J. L. Neira, and S. Padmanabhan and Profs. R. L. Baldwin, M. Bruix, M. Levitt, M. Libonati, C. N. Pace and M. Rico for critical comments. This work was supported by Spanish Ministry of Science and Education Grant CTQ2004-08275, Alzheimer's Association Grant NIRG-04-1083, Consejo Superior de Investigaciones Científicas Grant 200580F0131, and a grant from the National Science and Engineering Research Council of Canada.

1. Miller SL (1953) *Science* 117:528–529.
2. Matthews CN, Moser RE (1967) *Nature* 215:1230–1234.
3. Boon CL, Frost D, Chakrabarty A (2004) *Biopolymers* 76:244–257.
4. Frost DWH, Yip CM, Chakrabarty A (2005) *Biopolymers* 80:26–33.
5. Chothia C, Levitt M, Richardson D (1981) *J Mol Biol* 145:215–250.
6. Aurora R, Srinivasan R, Rose GD (1994) *Science* 264:1126–1130.
7. Felitsky DJ, Record MT (2003) *Biochemistry* 42:2202–2217.
8. Padmanabhan S, Laurents DV, Fernández AM, Elias-Arnanz M, Ruiz-Sanz J, Mateo PL, Rico M, Filimonov VV (1999) *Biochemistry* 38:15536–15547.
9. Vaz DC, Rodrigues JR, Sebald W, Dobson CM, Brito RMM (2006) *Protein Sci* 15:33–44.
10. Tamura A, Privalov PL (1997) *J Mol Biol* 273:1048–1060.
11. Panse VG, Swaminathan CP, Aloor JJ, Suroliya A, Varadarajan R (2000) *Biochemistry* 39:2362–2369.
12. Laurents DV, Corrales S, Elias-Arnanz M, Sevilla P, Rico M, Padmanabhan S (2000) *Biochemistry* 39:13963–13973.
13. Hammes GG, Roberts PB (1969) *J Am Chem Soc* 91:1812–1816.
14. Chakrabarty A, Kortemme T, Padmanabhan S, Baldwin RL (1993) *Biochemistry* 32:5560–5565.
15. Pace CN (1992) *J Mol Biol* 226:29–35.
16. Harris NL, Presnell SR, Cohen FE (1994) *J Mol Biol* 236:1356–1368.
17. DeGrado WF, Wasserman ZR, Lear JD (1989) *Science* 243:622–628.
18. Hill CP, Anderson DH, Wesson L, DeGrado WF, Eisenberg D (1990) *Science* 249:543–546.
19. Walsh STR, Cheng H, Bryson JW, Roder H, DeGrado WF (1999) *Proc Natl Acad Sci USA* 96:5486–5491.
20. Geremia S, Di Costanzo L, Randaccio L, Engel DE, Lombardi A, Nistri F, DeGrado WF (2005) *J Am Chem Soc* 127:17266–17276.
21. Lumb KJ, Kim PS (1995) *Biochemistry* 34:8642–8648.
22. Oakley MG, Kim PS (1998) *Biochemistry* 37:12603–12610.
23. Rudresh U, Ramagopal A, Inai Y, Goel S, Sahal D, Chauhan VS (2004) *Structure (London)* 12:389–396.
24. Kramer RZ, Venugopal MG, Bella J, Mayville P, Brodsky B, Berman HM (2000) *J Mol Biol* 301:1191–1205.
25. Yan Y, Erickson BW (1994) *Protein Sci* 3:1069–1073.
26. Domingues H, Cregut D, Sebald W, Oschkinat H, Serrano L (1999) *Nat Struct Biol* 6:652–656.
27. Dolphin GT, Baltzer L (1997) *Folding Des* 2:319–330.
28. Riddle DS, Santiago JV, Bray-Hall ST, Doshi N, Grantcharova VP, Yi Q, Baker D (1997) *Nat Struct Biol* 4:805–809.
29. Wei Y, Kim S, Fela D, Baum J, Hecht MH (2003) *Proc Natl Acad Sci USA* 100:13270–13273.
30. Dahiyat BI, Mayo SL (1997) *Proc Natl Acad Sci USA* 94:10172–10177.
31. Furukawa K, Oda M, Nakamura H (1996) *Proc Natl Acad Sci USA* 93:13583–13588.
32. Zhu H, Celinski SA, Scholtz JM, Hu JC (2001) *Protein Sci* 10:24–33.
33. Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS (1998) *Science* 282:1462–1467.
34. Dunbrack RLJ, Karplus M (1994) *Nat Struct Biol* 1:334–339.
35. Oro J, Miller SL, Lazcano A (1990) *Annu Rev Earth Planet Sci* 18:317–356.
36. Oro J, Stephen-Sherwood E (1974) *Orig Life* 5:159–172.
37. Abkevich VI, Gutin AM, Shakhnovich EI (1996) *Proc Natl Acad Sci USA* 93:839–844.
38. Edelhoch H (1967) *Biochemistry* 6:1948–1954.
39. Kumar A, Ernst RR, Wüthrich K (1980) *Biochem Biophys Res Commun* 95:1–6.
40. Bax A, Davis DG (1985) *J Magn Reson* 65:355–360.
41. Piotto M, Saudek V, Sklenar V (1992) *J Biomol NMR* 2:661–665.
42. Goddard TD, Kneller DG (2007) SPARKY 3 (University of California, San Francisco).
43. Sklenar V, Bax A (1987) *J Magn Reson* 74:469–479.
44. Güntert P, Mumenthaler C, Wüthrich K (1997) *J Mol Biol* 273:282–298.
45. Ferguson G, Seibel GL, Kollman PA (1995) *Comput Phys Commun* 91:1–41.
46. Escaja N, Pedrosa E, Rico M, González C (2000) *J Am Chem Soc* 122:12732–12742.
47. Soliva R, Monaco V, Gómez-Pinto I, Meeuwenoord NJ, Marel GA, Boom JH, González C, Orozco M (2001) *Nucleic Acids Res* 29:2973–2985.
48. Shields GC, Laughton CA, Orozco M (1997) *J Am Chem Soc* 119:7463–7469.
49. Darden T, York D, Pedersen L (1993) *J Chem Phys* 98:10089–10092.
50. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Crystallogr* 26:283–291.
51. Koradi R, Billeter M, Wüthrich K (1996) *J Mol Graphics* 14:51–52.
52. Holm L, Sander C (1995) *Trends Biochem Sci* 20:206–209.
53. Bai Y, Milne JS, Mayne L, Englander SW (1993) *Proteins Struct Funct Genet* 17:75–86.
54. Hvidt A, Nielsen SO (1966) *Adv Protein Chem* 21:287–386.