Eur Biophys J (2012) 41:63–72
DOI 10.1007/s00249-011-0758-4

ORIGINAL PAPER

EBSA
Biophysics in Europe

Conformation specificity and arene binding in a peptide composed only of Lys, Ile, Ala and Gly

Fernando Diez-García · Irene Gómez-Pinto · Avijit Chakrabartty · Carlos González · Douglas V. Laurents

Received: 24 June 2011 / Revised: 21 September 2011 / Accepted: 5 October 2011 / Published online: 27 October 2011

Abstract The first life on Earth is believed to have been based on RNA, but might have taken advantage of amino acids and short peptides which form readily under conditions like those of the primitive Earth. We have shown that simple peptides adopt specifically folded four-helix bundle structures that can recognize and cleave RNA. Here, to explore the limits of conformational specificity, we characterize a simpler peptide composed of just Lys, Ile, Ala, and Gly called KIA7I. Using nuclear magnetic resonance (NMR) spectroscopy and molecular dynamics (MD) simulations, we find kinks in the helices of KIA7I and multiple C-terminal conformations. These results suggest that the C-terminal Ile residue does not completely occupy the hydrophobic pocket that is filled by aromatic side-chains in well-folded KIA7 variants. The capacity of arenes to fill this cavity was tested. Using NMR, we show that benzene and phenol can bind KIA7I, but do not bind the well-folded variant KIA7W or hen egg white lysozyme. Benzene also binds $A\beta_{1-40}$, a mostly disordered polypeptide implicated in Alzheimer's disease. 8-Anilinonaphthalene-1-sulfonate (ANS) fluorescence is further enhanced in the presence of both KIA7I and arenes relative to KIA7I alone. This ANS fluorescence enhancement is stronger for smaller and less polar arenes and less ordered KIA variants. These results suggest that arenes are not confined to the pocket, but penetrate and loosen the hydrophobic core of KIA7I.

Keywords Prebiotic chemistry · NMR · Fluorescence · Arene ligands · Four-helix bundle

Electronic supplementary material The online version of this article (doi:10.1007/s00249-011-0758-4) contains supplementary material, which is available to authorized users.

F. Diez-García · I. Gómez-Pinto · C. González · D. V. Laurents (✉)
Instituto de Química Física "Rocasolano", Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain
e-mail: dlaurents@iqfr.csic.es

A. Chakrabartty
Department of Medical Biophysical and Biochemistry, University of Toronto, Toronto MSG-2M9, Canada

Introduction

The recent elucidation of a plausible synthetic pathway for nucleotide phosphates under putative prebiotic Earth conditions (Powner et al. 2009) and the discovery of ribozymes catalyzing RNA synthesis (Lincoln and Joyce 2009; Wochner et al. 2011) have provided important new support for the RNA World hypothesis, which is the most plausible explanation for the origin of life on Earth (Orgel 1968; Gilbert 1986). These studies also rekindle interest in the possibility that RNA was not alone in the RNA World, but may have been accompanied by amino acids or peptides formed by prebiotic chemistry prior to the advent of a protoribosome (DiGiulio 1997). We have shown that peptides with simple amino acid composition can adopt rather stable and specifically folded four-helix X-bundle proteins (Boon et al. 2004; López de la Osa et al. 2007; López Alonso et al. 2010). The sequence of this series of 20 residue peptides, called KIA7, is Ac-AKAAAAAIKAIAAIIKAGGφ-NH_2, where φ is an aromatic residue that fills hydrophobic pockets formed between helix termini. These peptides interact with and specifically cleave RNA hairpins, and their stability increases with the size and hydrophobicity of the aromatic group: $His^+ < His^o < Tyr \leq Phe < Trp$. It is remarkable that the proteins with such a small alphabet (only Lys, Ala, Gly, Ile and an aromatic residue) and a set of

stabilizing interactions as limited as RNA can form specific and well-folded structures. Of these amino acids, Lys, Ala, Gly and Ile form under plausible prebiotic Earth conditions (Miller 1953; Matthews and Moser 1967; Muñoz Caro et al. 2002; Huber and Wächtershäuser 2006; Parker et al. 2011), whereas formation of the aromatic amino acids generally requires a different set of prebiotic reaction conditions (Friedman and Miller 1969). The KIA7 variant with a C-terminal Ile, called KIA7I, is chiefly unfolded in water. The addition of a stabilizing salt, sodium sulfate, induces KIA7I to adopt a tetramer with modest protection against H/D exchange, fairly cooperative thermal unfolding, and a rather tightly packed hydrophobic core (López de la Osa et al. 2007). One objective here is to explore the limit where a structure becomes so simple that it loses specificity by characterizing the structural ensemble of KIA7I using NMR and molecular dynamics simulation.

A deoxyribozyme, composed of only four bases, folds specifically, binds His, and employs it as a cofactor (Roth and Breaker 1998). This suggests that amino acids might have been used by early ribozymes to expand their catalytic repertoire. In lieu of containing aromatic residues, we hypothesize that the earliest proteins, such as KIA7I, might have been able to recognize and bind aromatic compounds. Therefore, our second objective is to test the ability of KIA7I to bind aromatic compounds using NMR and fluorescence spectroscopies.

Materials and methods

NMR spectroscopy

Peptides were purchased from Carlo Labs (Lyngby, Denmark). Deuterated acetic acid (d_3) was purchased from Sigma/Aldrich (St. Louis, USA). One-dimensional (1D) ^{1}H and two-dimensional (2D) total correlation spectroscopy (TOCSY) (60, 80, and 100 ms mixing times), nuclear Overhauser effect spectroscopy (NOESY) (80 and 150 ms mixing times), and natural-abundance ^{1}H-^{13}C and ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) NMR spectra were acquired using 600- and 800-MHz (^{1}H) Bruker instruments equipped with cryoprobes and *z*-axis gradients. NMR spectra were acquired, processed, and analyzed using TOPSPIN (Bruker). Spectra of KIA7I were recorded at 5°C in 0.50 M Na_2SO_4 and 10 mM deuterated sodium acetate/acetic acid [NaAc/DAc (d_3)] buffer, pH 5.0 in D_2O (99.9% atom D, Cambridge Isotope Labs, UK) or 90% milliQ H_2O/10% D_2O. Thin-diameter (3 mm) NMR tubes were used to minimize salt effects on the pulse efficiency. Sodium 2,2-dimethyl-2-silapentane-sulfonate (DSS; Stohler Isotope Labs, USA) was used as the internal chemical shift standard at concentration of 50 μM. The KIA7I concentration was 1.0–2.5 mM. Some samples contained aqueous benzene or phenol. Further details on the preparation of samples with aqueous benzene are given in the Electronic Supplementary Material.

Additional NMR spectra with and without benzene were acquired on the stable, well-folded KIA7 variants, KIA7W (3 mM), and CG_3KIA7 (2 mM) in 99.9% D_2O, 200 mM NaCl, 10 mM DAc/NaAc (d_3) buffer pH 5.0 at 5°C, as well as the intrinsically disordered Alzheimer's Aβ_{1-40}

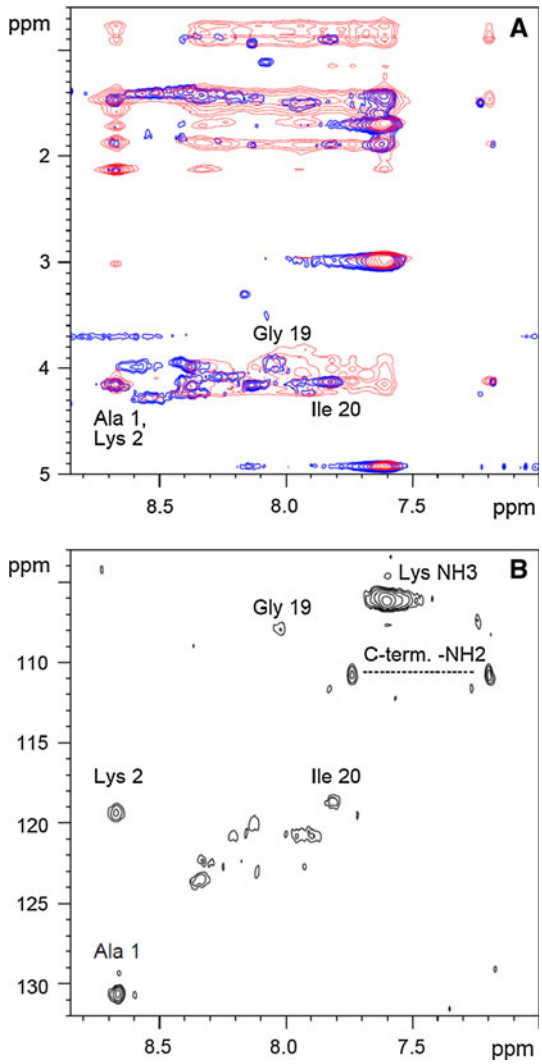


Fig. 1 **a** NMR TOCSY (*blue*, mixing time 80 ms) and NOESY (*red*, mixing time 150 ms) of KIA7I in 0.50 M Na_2SO_4, pH 5, 5°C. The downfield region of the spectra are shown, and ^{1}HN-^{1}Hα resonances that are assignable to particular residues are labeled. **b** Natural-abundance ^{1}H-^{15}N HSQC spectrum of KIA7I recorded in 90% H_2O, 10% D_2O, 0.5 M Na_2SO_4 pH 5, 5°C. The ^{1}H-^{15}N resonances that are assigned to specific residues are labeled. Note that the peaks of Ala1 and Lys2, which overlap in the ^{1}H 2D TOCSY and NOESY spectra (Fig. 1a), are well resolved here. The intense signal at approximately 7.6 ppm ^{1}H and apparently 106 ppm ^{15}N is a folded peak arising from Lys amine groups. Despite the long acquisition time (2 days) of this spectrum, not all the expected peaks appeared; this can be attributed to the low natural abundance of the ^{15}N nucleus

polypeptide (0.22 mM) in 90% H_2O/10% D_2O in 10 mM K_2HPO_4 buffer, pH 7.0 at 25°C, and hen egg white lysozyme (HEWL) (2.9 mM) in 99.9% D_2O, 10 mM DAc/NaAc (d_3) buffer, pH 4.0 at 35°C. In the case of HEWL, a ^{1}H-^{13}C HSQC spectrum with the excitation pulse centered at 120 ppm (^{13}C) was recorded to observe signals arising from aromatic groups.

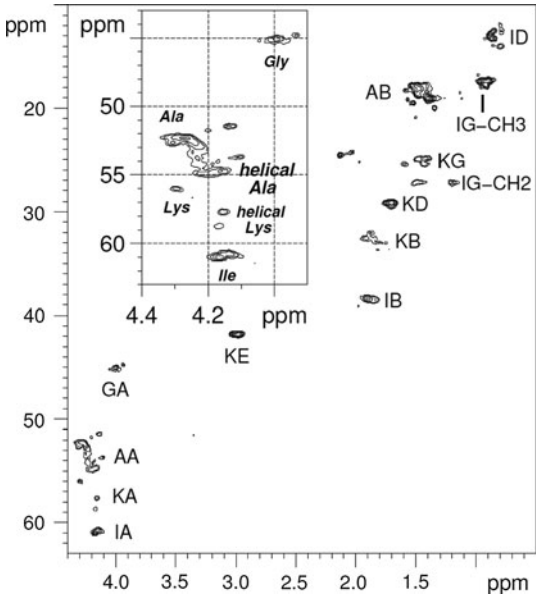


Fig. 2 Natural-abundance ^{1}H-^{13}C HSQC spectrum of KIA7I recorded in 99.9% D_2O, 0.50 M Na_2SO_4 pH 5, 5°C. Peaks are labeled with a two-letter code: the first letter stands for the amino acid type (A = Ala, G = Gly, etc.), and the second letter is the atom (A = α, B = β, G = γ, etc.) The *inset* highlights the ^{1}Hα-^{13}Cα resonances of Ala and Lys, some of which show chemical shifts characteristic of helical structure

Molecular dynamics

The program SYBYL-X (Tripos Int., St. Louis, USA) was used to prepare a set of 10 structural models of KIA7I on the basis of the NMR structure of KIA7Y. Then, the LEaP AMBER 11 module, (Case et al. 2005) was used to obtain the corresponding files in the correct AMBER format. Next, each structure, accompanied by chloride counterions, was subjected to 1,000 steps of energy minimization in vacuum using a nonbonded cutoff of 10 Å, no periodic boundary, and a distance dependence of 10 Å. Afterwards, the structures were hydrated with approximately 6,800 TIP3PBOX water molecules in a 60 × 65 × 60 Å^3 box, and were further energy-minimized for 5,000 steps until the root-mean-square (RMS) energy gradient was <0.1 kcal mol^{-1} Å^{-1}. After a further 10,000 MD steps (0.002 ps each), for energy equilibration, a long run of 1.3 ns was performed for each of the 10 starting models. At certain times, snapshots of the structures were saved and analyzed using MolMol (version 2K.2; R. Koradi; ETH Zurich) and Mage (version 6.42 for Windows; D.C. Richardson and B. Presley; Duke University). The latter was also used to prepare structural images.

Fluorescence spectroscopy

A Fluormax4 fluorescence spectrometer equipped with a Peltier temperature control module was used to register spectra using a 3 × 3 mm Hellma QS cuvette (volume

Table 1 Chemical shifts of KIA7I in 0.50 M Na_2SO_4, pH 5, 5°C

Residue	^{1}HN, ^{15}N	^{1}Hα, ^{13}Cα	^{1}Hβ, ^{13}Cβ	^{1}H$\gamma_{12,13}$; ^{13}Cγ_1	^{1}Hγ, ^{13}Cγ$_2$	^{1}Hδ, ^{13}Cδ	^{1}Hε, ^{13}Cε	^{1}Hζ
General								
Ala (helical)		4.19, **54.8**	1.46, **18.2**					
Ala (nonhelical)		4.28, **52.3**	1.39, **19.1**					
Gly		4.00, **45.1**						
Ile		4.16, **60.8**	1.90, **38.5**	1.48, 1.15, **27.2**	0.94, **17.5**	0.87, **12.9**		
Specific								
N-term Ac-CH_3		2.13, **24.6**						
Ala1	8.67, ***130.7***	4.17, **54.7**	1.48, **18.1**					
Lys2	8.68, ***119.4***	4.16, **57.7**	1.87, **32.1**	1.58, **25.4**		1.72, **29.1**	2.99, **41.8**	7.62
Ala3	8.34	4.10, **53.7**	1.43, **18.5**					
Lys9[a]		4.15, **58.7**	1.90, **32.6**	1.51, **24.9**		1.71, **29.1**	2.99, **41.8**	7.62
Lys16[b]	8.54	4.29, **56.0**	1.83, **32.9**	1.41, **24.9**			2.99, **41.8**	7.62
Gly19	8.04, ***108.0***	3.96, **45.1**						
Ile20	7.83, ***118.7***	4.13, **60.8**	1.89, **38.5**		0.94			
C-term-NH_2	7.74, 7.20, ***110.9***							

Significance of ^{13}C shifts are shown in bold, ^{15}N shifts are shown in bold italics

[a] Probable assignment based partly on the high probability that this Lys is mainly in a helical conformation

[b] Probable assignment based partly on the high likelihood that this Lys is mainly in a disordered conformation

0.10 mL). 8-Anilinonaphthalene-1-sulfonate (ANS) (Fluka) was >98% pure (thin-layer chromatography, TLC). An excitation wavelength of 365 nm was used, and emission spectra were recorded in the range of 400–650 nm, with 1 nm steps. The integration time was 0.5 s per nm, and the entrance and exit slits were set to 3 nm. The concentration of ANS was determined by absorbance using $\varepsilon_{350nm} = 5{,}000\ M^{-1}\ cm^{-1}$ (Weber and Young 1964); its concentration in these experiments was 120 μM. The concentration of KIA peptides was 120–200 μM. Samples were prepared in triplicate, and average results are shown. One series of experiments was done at pH 7.0 in 10 mM Tris buffer to match the original conditions of Boon et al. (2004). A second set of measurements was carried out in 10 mM Na/HAc buffer at pH 5.0 at various concentrations of Na_2SO_4. For some samples, a 10-μL droplet of an aromatic compound (benzene, toluene, *o*- or *p*-xylene, anisole, or aniline), a grain of phenol, or a flake of naphthalene was placed on top of the aqueous phase without mixing. Samples were incubated on ice for 18 h prior to measurement, and were placed in the sample chamber at least 3 min before measurement.

Results and discussion

NMR spectroscopy of KIA7I

The ^{1}H NMR TOCSY and NOESY spectra of KIA7I in 0.50 M Na_2SO_4, pH 5, 5°C showed rather poor chemical shift dispersion (Fig. 1), in contrast to KIA7φ variants. This is partly due to the absence of ring current effects but also reflects a lower structural uniqueness in KIA7I. Using 2D ^{1}H and ^{1}H-^{15}N and ^{1}H-^{13}C HSQC spectra (Figs. 1, 2) we assigned the resonances to the distinct spin systems of Ala, Ile, Lys, and Gly, although signal overlap thwarted assignment to individual residues in most cases. The assignments are listed in Table 1. About half of the Ala and two of three Lys residues have ^{1}Hα, ^{13}Cα, and ^{13}Cβ chemical shift values that deviate from coil values and are indicative of helical structure (Wishart et al. 1991); for example, Lys2 has a ^{13}Cα chemical shift value of 57.7 ppm, compared with a coil value of 56.2 ppm. Interestingly, whereas the Gly18 Hα and H'α resonances are degenerate in KIA7I, they are separated by 0.64–0.67 ppm in KIA7Y, KIA7F, and KIA7H and 0.90 ppm in KIA7W (López de la Osa et al. 2007; López Alonso et al. 2010). This loss of chemical shift uniqueness is good evidence that the specific turn structure adopted by Gly18 in the KIA7φ variants is not predominant in KIA7I.

Molecular dynamics of KIA7I

Restrained molecular dynamics was used to follow the evolution of energetically plausible structures of KIA7I. The structure diverges significantly after several 100 ps, suggesting that it lacks a unique conformation (Fig. 3a). The α-helices of KIA7 variants with aromatic residues are straight or slightly curved (López de la Osa et al. 2007; López Alonso et al. 2010). Analysis of the ϕ, φ angles of structures generated by the simulation shows that KIA7I tetramers retain significant helical structure but that the C-terminal G-G-I segment lacks a preferred conformation (Fig. 3b). Moreover, the KIA7I tetramers typically show two or three helices with kinks where an NH fails to

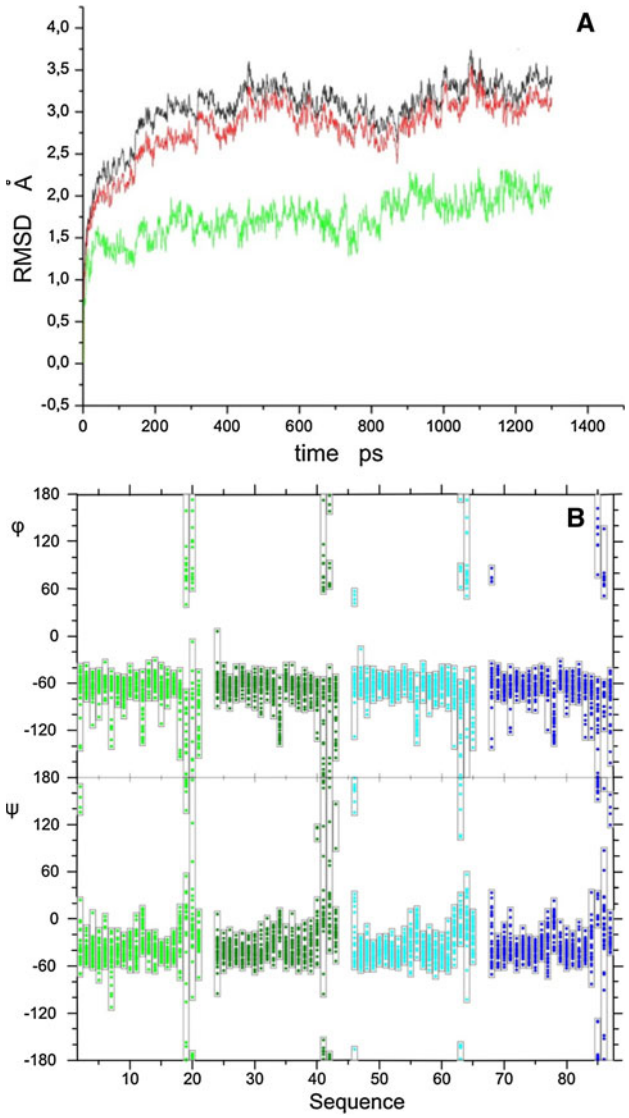


Fig. 3 MD simulations. **a** During the MD simulation of KIA7I, structural diversity, as measured by the RMS difference (RMSD) of single peptides (*green*), the main-chain atoms of the tetramer (*red*), or all atoms of the tetramer (*black*), increases significantly with time. **b** A plot of the population in ϕ, ψ space occupied by the KIA7I residues in four peptides of the tetramer shows that, whereas most residues chiefly adopt helical conformations, the C-terminal -Gly-Gly-Ile-NH_2 segments adopt a broad diversity of backbone conformations

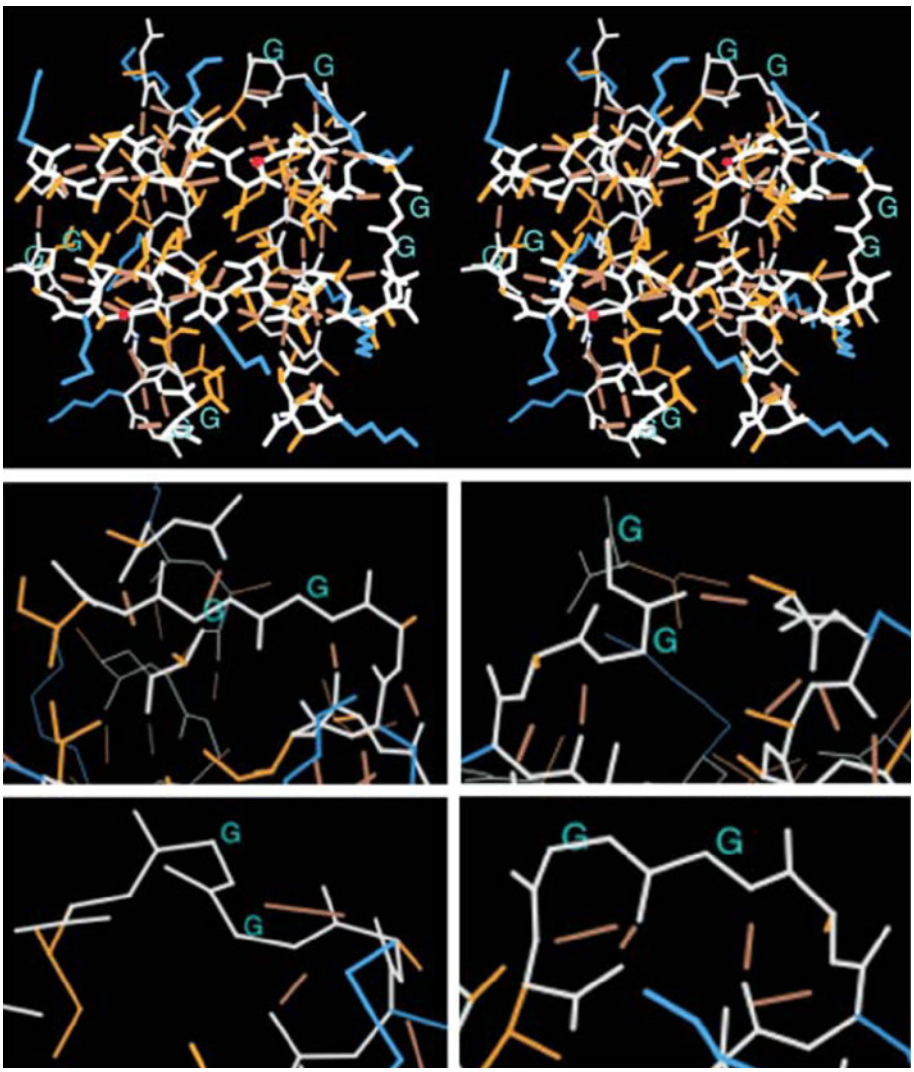


Fig. 4 MD structural models of KIA7I. *Top panel*: Cross-eyed stereo view of a representative structural model of KIA7I from MD. Backbone = *white*; Lys side-chains = *blue*; Ile and Ala side-chains = *orange*; Gly marked by a "G". H-bonds are shown as *brown lines*. *Red spheres* mark HN groups lacking H-bonds. *Bottom panels*: Zoom of four representative C-termini showing diverse conformations and H-bonding patterns

hydrogen bond to an OC acceptor (Fig. 4). These kinks seem to reduce the size of the hydrophobic pockets at helix termini to accommodate the small Ile side-chain. The C-terminal residues adopt a variety of conformations and patterns of hydrogen binding (Fig. 4). A representative KIA7I structural model has been deposited in the Protein Model Data Base: PMDB ID—PM0077491.

NMR spectroscopy reveals benzene and phenol binding to KIA7I

Aqueous benzene resonates as a fine 1.9-Hz singlet at 7.42 ppm (Fig. 5a). However, in the presence of KIA7I, this signal broadens to 16.0 Hz and produces intermolecular NOEs with resonances at 1.50 and 0.88 ppm (Fig. 5b, Electronic Supplementary Material Fig. 2). These crosspeaks, which have the same sign as KIA7I signals, are assignable to methyl H of Ala and Ile residues of KIA7I. Based on these data, we conclude that benzene binds to the hydrophobic core of KIA7I. No significant δ changes are observed for benzene or KIA7I; therefore, we conclude that benzene does not attain a fixed magnetically unique conformation when bound to KIA7I and that benzene binding does not induce a further structuralization of KIA7I. Similar benzene line width changes and benzene/KIA7I NOEs are observed in the presence or absence of DSS. Interestingly, when the three components, DSS, benzene and KIA7I are present, the signals of DSS broaden substantially, from 1.9 to 17.5 Hz for the trimethyl group, and NOE crosspeaks are also observed between the benzene resonance and the trimethyl and methylene moieties of DSS. These data are good evidence that DSS alone does not bind to KIA7I, but does so in the presence of benzene. DSS is known to bind to partly structured, positively charged polypeptides, such as the Alzheimer's Aβ polypeptide (Laurents et al. 2005), and DSS binding is a hallmark of molten globules (Shimizu et al. 1994). Therefore, we advance that benzene binding

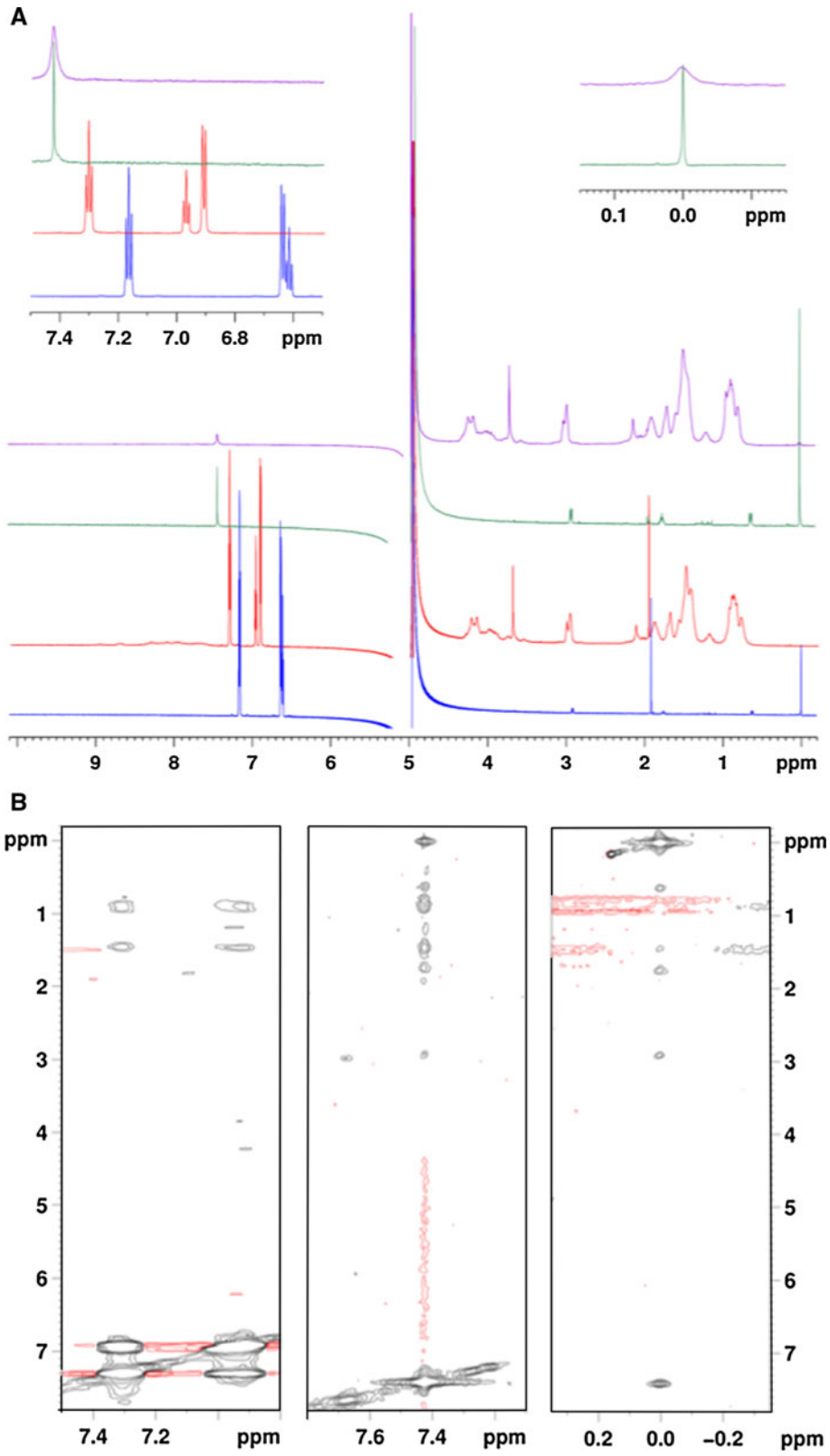
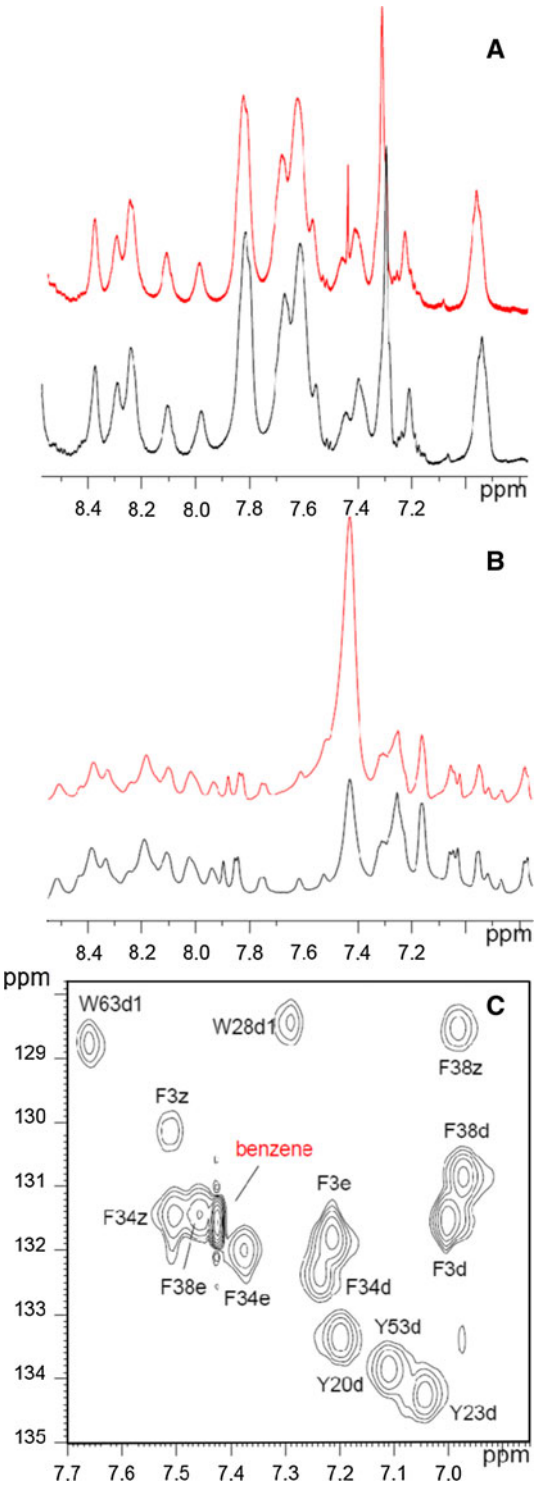


Fig. 5 KIA7I binds benzene and phenol spectra acquired at 5°C in 0.50 M Na_2SO_4, 10 mM Na/H Ac buffer, pH 5.0. **a** 1D ^{1}H NMR spectra of phenol alone (*blue*), and with KIA7I (*red*); and of benzene alone (*green*) and with KIA7I (*purple*). Magnified views highlight changes in aromatics (*left*) and DSS (*right*). The preparation of the sample with benzene is described in Electronic Supplementary Material Fig. 1. **b** Strips of the 2D ^{1}H NOESY spectra (80 ms mixing time) of KIA7I + 1.0 mM phenol (*left*), KIA7I + *aq* benzene (*middle*), and 50 μM DSS (*right panel*). The full spectra are shown in Electronic Supplementary Material Fig. 2

"lubricates" the packing of KIA7I, making it loose enough to permit DSS binding. The interaction of phenol with KIA7I was also studied. Conclusive evidence of phenol binding, namely altered phenol δ and intermolecular NOEs to Ala and Ile methyls of KIA7, was observed (Fig. 5).



◀ **Fig. 6** Benzene does not bind KIA7W or HEWL but does bind Aβ_{1-40} **a** 1D ^{1}H NMR spectra of KIA7W (3 mM, pH 5, 5°C) alone (bottom spectrum, *black*) or after incubation with benzene (top spectrum, *red*). A fine benzene singlet is observed at 7.42 ppm in the top spectrum. **b** 1D ^{1}H NMR spectra of Aβ_{1-40} (0.22 mM, pH 7, 25°C) minutes (bottom spectrum, *black*) and 1 day (top spectrum, *red*) after application of a drop of benzene to the top of the solution of Aβ_{1-40}. A 2D NOESY spectrum of this sample is shown in Electronic Supplementary Material Fig. 3. **c** 2D ^{1}H-^{13}C HSQC spectrum of HEWL (2.9 mM, pH 4, 35°C in D_2O) in the presence of benzene. The zone around the benzene peak is shown. HEWL peaks are labeled according to the assignments of Wishart and coworkers (Wang et al. 2000)

To test the generality of benzene binding to proteins, we recorded NMR spectra on four additional proteins: (1) KIA7W, a stable, well-folded variant of KIA7 containing a C-terminal Trp residue; (2) CG_3KIA7Y, another stable variant with an N-terminal disulfide bridge and a C-terminal Tyr residue; (3) the Alzheimer's Aβ_{1-40} polypeptide, which is rich in hydrophobic residues and exists as a rather disordered monomer at neutral pH; and (4) hen egg white lysozyme (HEWL), a well-folded, highly stable protein. In the presence of KIA7W, benzene resonates as fine signal (peak width at half-height = 2 Hz), which is consistent with a lack of binding (Fig. 6a). Similar results were obtained for CG_3KIA7Y (data not shown). In contrast, in the presence of Aβ_{1-40}, benzene resonates as a very broad peak (33 Hz) (Fig. 6b). The binding of benzene to Aβ_{1-40} is corroborated by the observation of intermolecular NOEs between this arene and the polypeptide (Electronic Supplementary Material Fig. 3). The benzene singlet overlaps with protein resonances in 1D and 2D ^{1}H homonuclear spectra of HEWL, but can be resolved as a fine peak in ^{1}H-^{13}C HSQC spectrum (Fig. 6c). In the same spectrum, HEWL aromatic resonances are observed to have the same chemical shift values as previously reported in the absence of benzene (Wang et al. 2000). Based on these data, we conclude that benzene binds to the mainly disordered Aβ_{1-40} polypeptide, but not to the well-folded proteins KIA7W, CG_3KIA7Y and HEWL.

Fluorescence spectroscopy corroborates that aromatics bind KIA7I

ANS fluorescence is enhanced by partly structured molten globule conformations, but not by well folded or by thoroughly unfolded "random coil" states (Semisotnov et al. 1991). ANS was used to probe KIA7I and other KIA7 variants for loosely packed nonpolar groups in 10 mM Tris buffer, pH 7.0. ANS alone in buffer has a low fluorescence emission whose maximal wavelength (λ_{em} max) is 534 nm. This value is not significantly affected by benzene or phenol (Fig. 7a). In the presence of KIA7I, the ANS emission intensity is enhanced three- to fourfold and the λ_{em} max blueshifts to about 490 nm. Therefore, we conclude that ANS binds to KIA7I and that the environment of the bound ANS is less polar. In the presence of both benzene and KIA7I, ANS shows a significant additional blueshift and emission intensity increase. This suggests that benzene is capable of intercalating within the hydrophobic core of KIA7I and loosens its packing so that more

nonpolar groups become available for ANS binding. Similar but smaller changes are observed for ANS with KIA7I and phenol or naphthalene.

Previous studies showed that, whereas the well-folded KIA7Y does not significantly enhance ANS fluorescence, a disordered variant, called KIA5, strongly enhances and blueshifts ANS fluorescence (Boon et al. 2004). The effect of different KIA variants on the fluorescence emission of ANS has been studied here in the presence or absence of benzene. Essentially no change in ANS fluorescence was observed for the stable and specifically folded KIA variants, KIA7Y, KIA7F, and CG_3KIA7Y, with or without benzene (Fig. 7b). In contrast, the poorly folded or disordered variants: KIA7Δ (which lacks residue 20), KIA7ΔΔΔ (lacking residues 18, 19, and 20), and KIA5 (in which the position of Ile and Ala residues is permutated) induced significant increases in ANS fluorescence which were further augmented in the presence of benzene. The effect of KIA7I on ANS fluorescence is midway between that observed for well-folded KIAs and the poorly structured KIAs. These results lead us to propose that the capacity of benzene to contribute to the formation of loosely packed hydrophobic structure is higher for less structured proteins. To gain further insight, the capacity of other aromatic compounds to affect ANS fluorescence in the presence and absence of KIA7I was also determined. While the aromatics had null to mild effects on the ANS λ_{em} max in the absence of KIA7I, benzene, toluene, and *o*-xylene

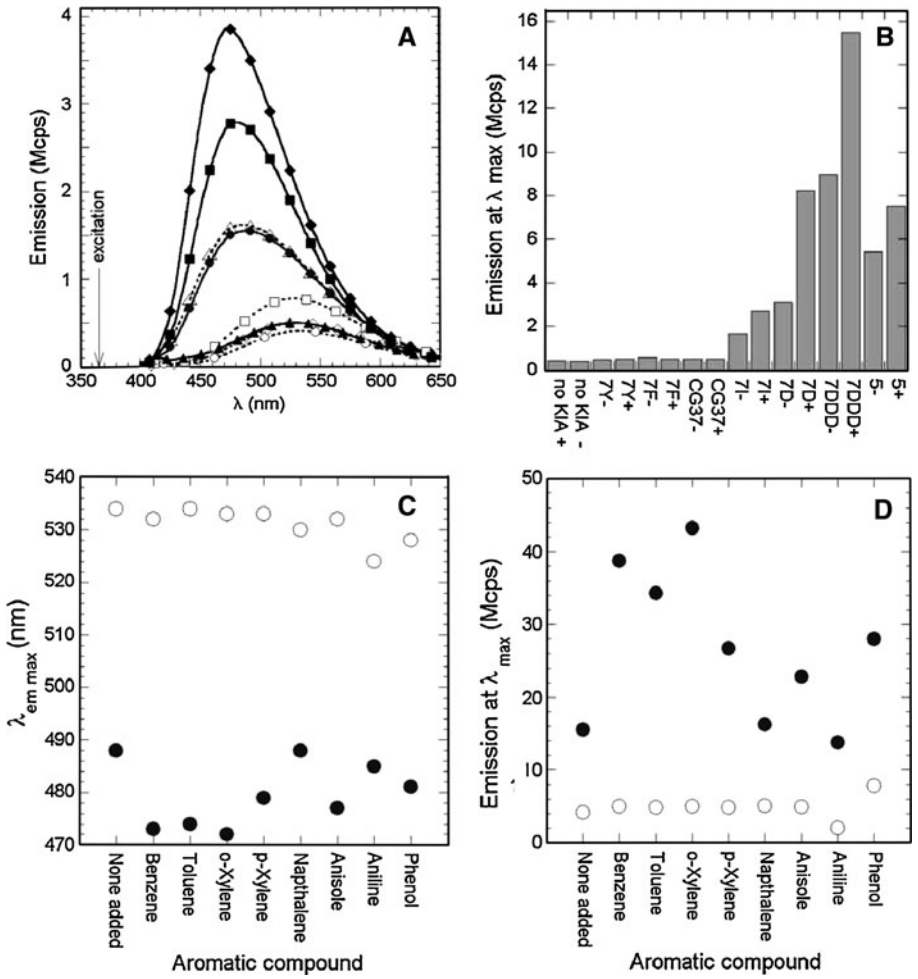


Fig. 7 ANS and aromatic compounds bind to KIA7. Spectra recorded at pH 7.0 in 10 mM Tris buffer at 25.0 or 2.0°C with 120 μM ANS and 120–200 μM peptides. **a** Fluorescence emission of ANS: *open circles*, ANS alone; *filled circles*, ANS and KIA7I; *open diamonds*, ANS and benzene; *filled diamonds*, ANS and benzene and KIA7I; *open squares*, ANS and phenol; *filled squares*, ANS and phenol and KIA7I; *filled triangles*, ANS and naphthalene; *open triangles*, ANS and naphthalene and KIA7I. Emission is plotted as millions of photons (counts) per second (Mcps). These spectra were recorded at 25°C. Similar results with larger ANS fluorescence enhancements by KIA7I are observed at lower temperatures. **b** ANS emission in the presence of KIA peptides with (+) or without (−) benzene at 2.0°C. **c** ANS λ_{em} max in the presence of various aromatic compounds, with KIA7I (*filled circles*) or without KIA7I (*open circles*). **d** ANS emission intensity at λ_{em} max in the presence of various aromatic compounds, with KIA7I (*filled circles*) or without KIA7I (*open circles*)

increased the ANS emission blueshift in the presence of KIA7I (Fig. 7c). Benzene, toluene, and *o*-xylene also induced the largest ANS emission enhancements in the presence of KIA7I (Fig. 7d). *p*-Xylene, naphthalene, anisole, aniline, and phenol had weaker effects. Based on these results, we propose that benzene, toluene, and *o*-xylene are more capable of penetrating KIA7I's hydrophobic core and loosening its structure to permit ANS binding.

These fluorescence experiments were repeated in aqueous solutions of Na_2SO_4 buffered by Na/HAc to pH 5.0, at 5°C. Without KIA7I, increasing concentrations of Na_2SO_4 were found to produce a moderate decrease in the ANS fluorescence emission intensity without shifting the peak maxima, and similar results were seen in the presence or absence of benzene (Electronic Supplementary Material Fig. 4A). In the presence of KIA7I, the ANS fluorescence emission intensity increases and blueshifts, and these changes are more pronounced in the presence of benzene (Electronic Supplementary Material Fig. 4B–D). These results corroborate the binding of KIA7I and benzene observed by NMR in these conditions and by ANS fluorescence at pH 7.0 in 10 mM Tris buffer without Na_2SO_4. Further ANS emission intensity increases and blueshifts are observed for ANS plus KIA7I at higher concentrations of Na_2SO_4, and these changes are even more significant when benzene is also present. As Na_2SO_4 increases the strength of hydrophobic interactions via the Hofmeister effect, these results can be attributed to a stronger association among ANS, benzene, and the nonpolar moieties of KIA7I.

Conclusions

We have found that KIA7I, a four-helix bundle composed of just four residues: Lys, Ile, Ala, and Gly, adopts a variety of conformations. Since their substitution for Ile leads to moderate conformational diversity, we conclude that aromatic residues are necessary to achieve specific folding in this class of putative prebiotic proteins. KIA7I binds aromatic compounds, with preference for smaller and less polar molecules. Benzene binding perturbs KIA7I's core structure and converts it into a molten globule. Benzene also binds a rather disordered polypeptide Alzheimer's $A\beta_{1-40}$, but not HEWL or KIA7W, which are well-folded, stable proteins. Based on these findings, we predict that benzene binds and perturbs the structure of partly folded and disordered proteins. Considering that one-third of eukaryotic proteins contain a disordered domain (Dunker et al. 2001), this proposal may have broad physiological significance.

Acknowledgments This work was supported by grant CTQ2010-21567-C02-02 and a fellowship (to F.D.-G.) from the Spanish Ministry of Science and Innovation and by the National Science and Engineering Research Council of Canada.

References

Boon CL, Frost D, Chakrabartty A (2004) Identification of stable helical bundles from a combinatorial library of amphipathic peptides. Biopolymers 76:244–257

Case D, Cheatham TE III, Darden T, Gohlke H et al (2005) The amber biomolecular simulation programs. J Comput Chem 26(16):1668–1688

DiGiulio M (1997) On the RNA world: evidence in favor of an early ribonucleopeptide world. J Mol Evol 45:571–578

Dunker A, Lawson J, Brown C, Williams K (2001) Intrinsically disordered protein. J Mol Graph Model 19:26–59

Friedman N, Miller SL (1969) Phenylalanine and tyrosine synthesis under primitive earth conditions. Science 166:766–767

Gilbert W (1986) The RNA world. Nature 319:618

Huber C, Wächtershäuser G (2006) Alpha-hydroxy and alpha-amino acids under possible Hadean, volcanic origin-of-life conditions. Science 314:630–632

Laurents DV, Gorman PM, Guo M, Rico M, Chakrabartty A, Bruix M (2005) Alzheimer's Abeta40 studied by NMR at low pH reveals that sodium 4, 4-dimethyl-4-silapentane-1-sulfonate (DSS) binds and promotes beta-ball oligomerization. J Biol Chem 280(5):3675–3685

Lincoln T, Joyce G (2009) Self-sustained replication of an RNA enzyme. Science 323:1229–1232

López Alonso JP, Pardo-Cea MA, Gómez-Pinto I, Fernández I, Chakrabartty A, Pedroso E, González C, Laurents DV (2010) Putative one pot prebiotic polypeptides with ribonucleolytic activity. Chemistry 16:5314–5323

López de la Osa J, Bateman DA, Ho S, González C, Chakrabartty A, Laurents DV (2007) Getting specificity from simplicity in putative proteins from the prebiotic Earth. Proc Natl Acad Sci USA 104:14941–14946

Matthews CN, Moser RE (1967) Peptide synthesis from hydrogen cyanide and water. Nature 215:1230–1234

Miller S (1953) A production of amino acids under possible primitive Earth conditions. Science 117:528–529

Muñoz Caro G, Meierhenrich U, Schutte W, Barbier B, Arcones Segovia A, Rosenbauer H, Thiemann W, Brack A, Greenberg J (2002) Amino acids from ultraviolet irradiation of interstellar ice analogues. Nature 416:403–406

Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393

Parker E, Cleaves H, Dworkin J, Glavin D, Callahan M, Aubrey A, Lazcano A, Bada J (2011) Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment. Proc Natl Acad Sci USA 108:5526–5531

Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–241

Roth A, Breaker RR (1998) An amino acid as a co-factor for a catalytic polynucleotide. Proc Natl Acad Sci USA 95:6027–6031

Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas AF, Gilmanshin RI (1991) Study of the molten globule intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers 31:119–128

Shimizu A, Ikeguchi M, Sugai S (1994) Appropriateness of DSS and TSP as internal references for 1H NMR studies of molten globule proteins in aqueous media. J Biomol NMR 4:859–862

Wang Y, Bjorndahl TC, Wishart DS (2000) Complete noncarbonylic 1H and 13C assignments of native hen egg white lysozyme. J Biomol NMR 17:83–84

Weber G, Young LB (1964) Fragmentation of bovine serum albumin by pepsin: I. The origin of the acid expansion of the albumin molecule. J Biol Chem 239:1415–1423

Wishart DS, Skyes BD, Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. J Mol Biol 222(2):311–333

Wochner A, Attwater J, Coulson A, Holliger P (2011) Ribozyme-catalyzed transcription of an active ribozyme. Science 332:209–212