Published in final edited form as:

Biopolymers. 2012 April ; 97(4): 240–249. doi:10.1002/bip.22013.

Like-charged residues at the ends of oligoalanine sequences might induce a chain reversal

Joanna Makowska^{1,2,*}, Adam Liwo^{1,2}, Wioletta Żmudzińska³, Lewandowska Agnieszka³, Lech Chmurzyński¹, and Harold A. Scheraga²

¹Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland ²Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853-1301, USA

³Intercollegiate Faculty of Biotechnology University of Gdańsk - Medical University of Gdańsk, Kładki 24, 80-822 Gdańsk, Poland

Abstract

We have examined the effect of like-charged residues on the conformation of an oligoalanine sequence. This was facilitated by CD and NMR spectroscopic and differential scanning calorimetric (DSC) measurements, and molecular dynamics calculations, of the following three alanine-based peptides: Ac-K-(A)₅-K-NH₂ (KAK5), Ac-K-(A)₄-K-NH₂ (KAK4), Ac-K-(A)₃-K-NH₂ (KAK3), where A and K denote alanine and lysine residues, respectively. Our earlier studies suggested that the presence of like-charged residues at the end of a short polypeptide chain composed of nonpolar residues can induce a chain reversal. For all three peptides, canonical MD simulations with NMR-derived restraints demonstrate the presence of ensembles of structures with a tendency to form a chain reversal. The KAK3 peptide exhibits a bent shape with its ends close to each other, while KAK4 and KAK5 are more extended. In the KAK5 peptide, the lysine residues do not have any influence on each other and are very mobile. Nevertheless, the tendency to form a more or less pronounced chain reversal is observed and it seems to be stable in all three peptides. This chain reversal seems to be caused by screening of the nonpolar core from the solvent by the hydrated charged residues

Keywords

alanine-based peptides; NMR spectroscopy; molecular dynamics; conformational studies; bent-like structure

1. Introduction

Peptides, in particular those excised from sequences of native proteins, that are likely to be folding-initiation sites, often serve as models of early stages of protein folding [1,2]. This concept stems from the nucleation-condensation model of protein folding, in which local interactions control the conformational space of polypeptide chains in the search for the native state. The interactions between different parts of the sequence promote nuclei of structure leading to a cooperative rate-limiting step from which the native state emerges [3]. For example, our recent studies of six peptides excised from the β -hairpin sequence of the immunoglobulin-binding protein 1 (IGG1; PDB code: 1IGD) [4] showed that even such short sequences can preserve a crude shape of the conformation in which they exist in the context of the entire protein.

*Corresponding author: Joanna Makowska; royek@chem.univ.gda.pl, phone: +48 058 523 5427; fax: +48 058 523 5247.

Because peptides are flexible objects, the primary focus of such studies is to determine the average content of secondary structure or other conformation-averaged properties. The α -helical conformation is a common secondary structure in biopolymers and, therefore, numerous studies of the kinetics and thermodynamics of the formation of α -helical structure in peptides were carried out [5–13]. The helix–coil transition has been extensively studied, both experimentally and theoretically, as a model for conformational transitions in biopolymers and as a way to obtain information about the intermolecular forces which stabilize biopolymer structure [5]. It was shown that the replacement of interior helical residues with alanine residues has considerable influence on the stability of α -helical peptides [14]. This observation led to the successful design of isolated, monomeric helical peptides in aqueous solution, containing several salt bridges and a high alanine content [15] or a simple sequence with a high alanine content with several lysine residues to increase solubility [16].

Compared to extensive studies of the formation and stability of α -helical structure, much less is known about the formation of β -type structures, which is indispensable for a better understanding of the protein-folding process. There is a practical aspect to research on the determinants of β -hairpin-structure formation because peptides, that spontaneously form β -hairpins in solution, very often aggregate into amyloid fibrils, a phenomenon directly connected with Alzheimer, Parkinson, Creutzfeld –Jacob, and many other neurodegenerative diseases [17–19].

The β -bends or, in general, chain reversals often are nucleation sites. Formation of bends as centers of initiation of folding by association of hydrophobic groups was proposed by Matheson and Scheraga [20]. Matheson and Scheraga [20], and Muñoz and coworkers [21] proposed a mechanism of formation of turns in β – hairpins through “zipping” hydrogen bonding and hydrophobic interactions. Karplus and coworkers [22] pointed out the dominant role of hydrophobic “collapse” of that part of a sequence. In our recent work [23,24], we showed that both local conformational preferences, as well as long-range interactions (salt bridges) play a significant role in forming bent structures. The formation of hydrogen bonds by backbone peptide groups is of secondary importance compared to hydrophobic interactions [20,23,24].

Our earlier conformational studies of the alanine-based peptide XAO [25,26], with sequence Ac-(X)₂-(A)₇-(O)₂-NH₂ (where X, A, and O denote diamino-butyric acid, alanine, and ornithine, respectively), carried out by means of NMR and CD spectroscopic measurements and molecular dynamics calculations, led to the conclusion that XAO has a bent structure. This result was in agreement with an earlier small-angle X-ray diffraction study (SAXS) of Pande and coworkers [27], which indicated a small radius of gyration, $R_g=7.4$ Å, in agreement with that calculated based on our conformational studies [25,26]. Further support was gained from our potentiometric study of this peptide, in which the value of $pK_{a1}=2.3$ is extraordinarily low, which demonstrates that all four charged groups are very close to each other, which is possible only if there is a chain reversal [28].

To investigate the influence of the presence of charged residues of various kinds on peptide conformation, we carried out further conformational studies [26] of three alanine-based peptides: Ac-KK-(A)₇-KK-NH₂ (KAK), Ac-OO-(A)₇-DD-NH₂ (OAD), and Ac-KK-(A)₇-EE-NH₂ (KAE), where A, K, O, D, and E, denote alanine, lysine, ornithine, aspartic acid, and glutamic acid residues, respectively. We found that the shape of these peptides depends on the size of the charged side chains at the ends and that the peptide assumes a bent shape if the side chains of the charged groups are sufficiently short. For the OAD peptide, which has the shortest acidic-basic side chains, the pK_{a1} value in water (3.46) [29] was remarkably

lower than that of the model compound, acetic acid (4.76), which also suggests close interactions between the groups.

The results of the studies summarized above suggested that the presence of charged residues at the ends of an oligoalanine sequence and, in general, of short sequences of mainly nonpolar amino-acid residues might lead to a chain reversal because of screening of the nonpolar core by the hydrated charged residues. Consequently, the presence of flanking charged residues might be another factor leading to the formation of β -bends, which can later serve as nucleation centers. It appears that the kind of residues (like- or oppositely-charged) may be less important than the length of the spacer between the charged groups. However, in references 25 and 26 we studied peptides which had two charged groups at each ends. Such concentration of charged groups is not frequently encountered in chain reversals, which contain charged groups at the ends. Therefore, in the present study, we investigated the influence of the length of an oligoalanine chain on the conformation of alanine-based peptides with the general formula $\text{Ac-K-(A)}_n\text{-K-NH}_2$ where $n=3, 4$ and 5 .

2. Results and discussion

2.1. Circular dichroism spectroscopy

The CD spectra for the KAK3, KAK4 and KAK5 peptides in water, recorded at room temperature ($T = 298 \text{ K}$), are shown in Figure 1. The CD spectra in H_2O are characterized by a deep minimum at about 200 nm , a very small maximum for KAK3 around 215 nm which suggests that this peptide exists as a statistical-coil structure in this environment. The minimum becomes shallower with increasing length of the oligoalanine chain. The analysis by using the CONTILL [30] program showed that, for all three peptides, there is a conformational equilibrium between a statistical coil and a turn-like structure. When the alanine separator becomes longer, the percentage of the statistical-coil conformation decreases and that of the β -sheet-like structure increases (see Table 1).

2.2. Differential Scanning Calorimetry

To determine if a conformational transition occurs in KAK3, KAK4, and KAK5, we performed DSC measurements in pure water ($\sim \text{pH } 5$). For the KAK3 and KAK4 peptides, the heat capacity curves have no maximum which suggest that there is no folding-unfolding transition (data not shown). For the KAK5 peptide, a very small broad maximum region occurs in the range of temperature $60^\circ\text{C} - 85^\circ\text{C}$ (data not shown). It can, therefore, be concluded that no folding-unfolding transition occurs for the peptides studied and they occur as a mixture of interconverting conformations in a dynamic equilibrium with populations of conformations varying with temperature.

2.3. Nuclear Magnetic Resonance Spectroscopy

A more detailed structural analysis was carried out with the use of NMR spectroscopy. The chemical shifts of the proton resonances for the KAK3, KAK4 and KAK5 peptides at 303 K are listed in Tables 2–4 [31]. The values of the $^3J_{\text{HNH}\alpha}$ coupling constants are presented in Table 5 and Rotating Frame Overhauser Enhancement (ROE) signals, which were used to assign the protons, together with proton assignments, are shown in Figure 2. The ROE effects corresponding to interproton distances are shown in Figure 3 and summarized in Table 6.

For KAK3, KAK4 and KAK5, the $\text{H}^\alpha(i) - \text{H}_\text{N}(i+1)$ sequential ROEs are always observed but with different strength (Figure 3). No additional cross-peaks were observed between the Ala amino acid residues in KAK3, KAK4 and KAK5 because of quite strong overlapping of the signals of the corresponding protons.

The spin systems of the Ala2 and Ala4 residues of KAK3 were easy to distinguish from each other because of the sequential $H^{\alpha}(\text{Lys1})-H_N(\text{Ala2})$ and $H^{\alpha}(\text{Ala4})-H_N(\text{Lys5})$ ROE signals. After the Ala2 and Ala4 protons had been assigned, the assignment of the Ala3 protons could be done unambiguously, because it was the only alanine proton system remaining. Similarly, we assigned the Ala2 and Ala5 and then the Ala3 and Ala4 signals in KAK4 (it should be noted that the Ala3 and Ala4 signals in KAK4 are almost completely overlapping; see Table 3 and Figure 2). Similarly, based on the $H^{\alpha}(\text{Lys1})-H_N(\text{Ala2})$, $H^{\alpha}(\text{Ala2})-H_N(\text{Ala3})$, and $H^{\alpha}(\text{Ala6})-H_N(\text{Lys7})$ sequential ROE signals, we assigned the remaining alanine amino acid residues of KAK5.

We also observed quite strong $H_N(i) - H_N(i+1)$ and $H^{\beta}(i) - H_N(i+1)$ sequential ROEs between Lys1 and Ala2 in each investigated peptide. The $H_N(i) - H_N(i+1)$ sequential ROE connectivities can indicate either a hairpin/turn or an α -helix structure [32]. Because the chemical shift of the amide protons of Ala3, Ala4 and Lys5 of KAK3, Ala3 – Ala4 and Ala5 – Lys6 of KAK4, as well as Ala3 – Ala4 and Ala6-Lys7 of KAK5 are almost the same (see Table 2–4), we cannot definitely state the presence or absence of such signals between these residues. In the hairpin/turn structure, long-range interactions, which connect the two strands together, should also be observed. Such long-range connectivities are observed for each of the studied peptides: $H^{\alpha}(\text{Ala2}) - H^{\beta}(\text{Lys5})$, $H^{\alpha}(\text{Ala2}) - H^{\beta}(\text{Lys6})$ and $H^{\alpha}(\text{Ala2}) - H^{\beta}(\text{Lys7})$ signals for KAK3, KAK4 and KAK5, respectively. We also found quite weak $H^{\alpha}(\text{Lys2}) - H_{sc}(\text{Lys5})$, $H^{\alpha}(\text{Lys2}) - H_{sc}(\text{Lys6})$ and $H^{\alpha}(\text{Lys2}) - H_{sc}(\text{Lys7})$, $H^{\alpha}(\text{Ala3}) - H_{sc}(\text{Lys1})$ interactions for KAK3, KAK4 and KAK5, respectively [where the subscript “sc” denotes side-chain proton(s)] (see Figure 3 and Table 6). These interactions indicate that the structures of KAK3, KAK4 and KAK5 at 303 K are bent.

2.4. Conformational Analysis

The families of conformations of all three peptides, obtained by MD simulations with time-averaged restraints derived from NMR measurements and clustered by using the MOLMOL program [33] are shown in Figure 4. It can be seen that, for KAK3 (Figure 4a), the most populated family of conformations consists of structures with a tendency to form a chain-reversal structure. Chain reversal is also less pronounced for the main family of KAK5 conformations (Figure 4c). Because of increasing length of the alanine spacer, the distance between the lysine side chains increases from KAK3 to KAK5.

All three peptides are very flexible. Based on the considerations presented above, it can be suggested that the differences in the shape of the peptides studied is determined by the length of the alanine separator. This observation conforms with the results of our previous study [26] in which we found that the shape of the conformations of Ac-XX-(Ala)₇-YY-NHMe peptides (where X and Y denote aminoacid residues with ionizable side chains) depends on the length of the side chains of the end residues rather than on whether X and Y are like- or oppositely-charged. A plausible explanation for this observation is the tendency of the hydrated charged groups to shield the nonpolar sequence from the solvent.

We have also analyzed the local conformational states of the peptides studied. A collective scatter plot of the conformational states of each amino acid residue of KAK3, KAK4, and KAK5 peptides generated by using MD calculations with NMR restraints are presented in Figure 5. The most populated area on the plot for all residues of the three peptides is allowed for structures with tendency to form β -like structure. For KAK3 and KAK4, the alanine residues have the biggest propensity to populate the β -region whereas, for KAK5, in addition to the alanine residues one lysine also appears to lie in the β -region (around $\phi = -60^\circ$, $\psi = -160^\circ$).

The results reported in this study are supported by those of our accompanying potentiometric studies of the acidic-basic properties of KAK3, KAK4, and KAK5 [34]. The pK_{a1} values of KAK3, KAK4 and KAK5 sequences are 8.2 (0.12), 9.67 (0.10), and 9.93 (0.14), respectively, the numbers in parentheses denoting the standard deviations, i.e., the pK_{a1} values are lower than that of the reference compound modeling the lysine side chain (n-butylamine, $pK_a = 10.5$). The differences from the value for the reference compound are greater than $\log_{10} 4 \approx 0.60$ a value which is characteristic of bifunctional acids with non-interacting functional groups (e.g., the dicarboxylic acids with a large methylene-spacer length) [35], suggesting that the two groups interact with each other and, consequently, are close to each other, consistent with the bent shape of this peptide. As expected, the value of pK_{a1} increases with the length of the oligoalanine chain. The pK_{a2} values are, on the other hand, greater than that of the reference compound [$pK_{a2} = 11.41$ (0.29), 11.92 (0.27), and 10.99 (0.19) for KAK3, KAK4, and KAK5, respectively], demonstrating that the cationic-acid group that remains after the dissociation of the first proton from the peptide is influenced by a greater number of nonpolar groups compared to that of the reference compound, which results in the increase of pK_a values. Additionally, the fact that pK_{a2} is greater than the reference value and decreases with the length of the oligoalanine chain suggests that the charged and the deprotonated amino groups may form an intra-molecular hydrogen bond, which will be less stable for a longer chain; nevertheless, the peptide has to possess a bent shape for a hydrogen bond to form.

3. Conclusions

From the conformational study of three alanine-based peptides, Ac-K-(A)₃-K-NHMe, Ac-K-(A)₄-K-NHMe, and Ac-K-(A)₅-K-NHMe by CD and NMR spectroscopic measurements and molecular dynamics calculations with restraints derived from the experimental data, we concluded the following: the CD measurements indicate that all three peptides occur as an ensemble of statistical-coil and bend-like structures in water. This observation is confirmed by the absence of a heat-capacity peak which would appear if a transition to a well-defined conformation took place. MD simulations with NMR-derived restraints also indicate the absence of a remarkable amount of regular secondary structure (α -helix or extended).

The results of NMR measurements for the peptides studied and the subsequent MD simulations with NMR-derived restraints indicate that the peptides possess a bent shape. A plausible reason for this is that the hydrated charged residues have a tendency to shield the nonpolar sequence (alanine separator) from the solvent. When the oligoalanine chain becomes longer, the charged lysine groups are at a longer distance from each other (which is manifested as increased pK_{a1} values [34], even though the bent shape is kept Figure 4a–c), which is manifested as the presence of long-range ROE signals (Figure 3).

4. Methods

4.1 Experimental

4.1.1 Peptide Synthesis—The peptides Ac-K-(A)₃-K-NH₂ [KAK3], Ac-K-(A)₄-K-NH₂ [KAK4] and Ac-K-(A)₅-K-NH₂ [KAK5] were synthesized by standard solid-phase Fmoc-amino acid chemistry with a Milipore synthesizer. The resin Tentagel R RAM (1g, capacity 0.19 mmol/g) was treated with piperidine (20%) in DMF, and all amino acids were coupled using DIPCI/HOBt (N,N'-diisopropylcarbodiimide/1-hydroxybenzotriazol) methodology. The coupling reaction time was 2 h. Piperidine (20%) in DMF was used to remove the Fmoc group at all steps. After deprotection of the last Fmoc N-terminal group, the resins containing each of the peptides were washed with methanol and dried *in vacuo*. After a complete polypeptide chain was assembled, the resins were treated with 1 M 1-acetylimidazole in DMF at room temperature for 2 h to attach the acetyl group to the N-

terminal part of the peptide. In the final step, the resins were treated with a TFA/water/phenol/triisopropylsilane (8.8/0.5/0.5/0.2) mixture (10 ml per gram of resin) at room temperature for 2 h to remove the peptide from the resin. Cleavage of the peptide from the resin under these conditions leaves a carboxamide group at the C-terminus. Each of the resins was separated from the mother liquid; the excess of solvent was then evaporated to a volume of 2 ml, and the residue was precipitated with diethyl ether. The crude peptides were purified by reverse-phase HPLC using a Supelcosil™ SPLC-ABZ C₁₈ semi-preparative column (10 × 250 mm, 5 μm) with 4 ml/min elution with an isocratic mixture of 0.5% CH₃CN in TFA for 120 min, to adjust the pH approximately to 2.0. To identify fractions containing the pure peptides, HPLC was run first with a small amount of the crude peptide, and the absorbance at 222 nm was measured for each fraction. The time windows for the retention of purified peptides were determined as those corresponding to the large peak in the plot of absorbance vs. time. Subsequently, a semi-preparative HPLC run was carried out, and the fractions containing the pure peptide were collected and lyophilized. The purity of each peptide was confirmed by analytical HPLC and MALDI-TOF analysis. The purities of these peptides were: 99.82% (KAK3), 99.4% (KAK4), 99.85% (KAK5).

4.1.2 Circular Dichroism Spectroscopy—Circular dichroism (CD) spectra were recorded for all three peptides at room temperature (T = 295 K) in water on a Jasco-715 automatic recording spectropolarimeter. Measurements were made in 1 mm quartz cuvettes. The spectra were recorded in the 193–260 nm wavelength range, using a sensitivity of five millidegrees and a scan speed of 2 cm/min. The results were plotted as the mean residue ellipticity Θ [degree × cm² × dmol⁻¹]. The CD measurements were made at peptide concentrations of 0.1 mM in water. The secondary structure content was estimated from all the CD spectra using the CONTIN/LL program [30] provided within the CDPPro software package.

4.1.3 Differential Scanning Calorimetry—Calorimetric measurements were made with a VP-DSC microcalorimeter (MicroCal) at a scanning rate of 65°C/1h. The scans were obtained at protein concentrations of 0.188 mM, 0.166 mM and 0.15 mM, for KAK3, KAK4 and KAK5, respectively. The cell volume was 0.5 ml. All scans were run at pH 4.6 in water in the range of temperature from 5°C to 80°C. The reversibility of the transition was checked by cooling and reheating the same sample. These measurements were recorded three times. Results from DSC measurements were analyzed with the Origin 7.0 software from MicroCal using the routines of the software provided with the instrument [36].

4.1.4 ¹H-NMR spectroscopy—The NMR spectra of peptides: KAK3, KAK4 and KAK5 were measured on VARIAN 500 MHz and 600 MHz spectrometers. The following 2D H-NMR spectra were recorded: DQF-COSY [37] TOCSY [38] (80 ms), ROESY [39] (200 ms) at 298 K. The samples were dissolved in H₂O/²H₂O (9:1 by vol.). The pH value of all three peptides was ~5.5 and the concentration of each of the samples was 5 mM. The spectra were processed by using VARIAN 4.3 software (Varian Instruments, Palo Alto, CA, USA) and analyzed with the XEASY program [40]. The spectra were calibrated against DSS (sodium 4,4-dimethyl-4-silapentane-1-sulfonate) signal [31]. Proton signals were assigned based on the TOCSY spectra. The sequential analysis of the peptide was confirmed by the ROESY spectra [41]. The chemical shifts are reported in Tables 2 – 4. The coupling constants between NH and H^a protons (³J_{H^aNH^a}) of KAK3, KAK4 and KAK5 were obtained from two-dimensional DQF-COSY and one-dimensional ¹H spectra. The intensities of ROE signals were estimated from the ROESY spectra [39]. The TOCSY spectra, with peak assignments of KAK3 (a), KAK4 (b) and KAK5 (c), respectively, are shown in Figure 2.

The ROE inter-proton cross-peaks of all peptides under study were derived from 2D ¹H-NMR ROESY spectra (see Figure 3), and vicinal coupling constants ³J_{H^aNH^a} (see Table 5)

were obtained from 2D ^1H -NMR DQF-COSY spectra. In the first step, the ROESY peak volumes were converted to upper distance bounds by using CALIBA [41] of the DYANA package [42]. In the next step, torsion angles, based on the Bystrov-Karplus [43] equation, were generated using the HABAS algorithm of the DYANA package [44]. The upper distance limits and torsional angles were used as restraints in molecular dynamics calculations.

4.2 Theoretical calculations

4.2.1 Molecular Dynamics calculations with NMR-derived restraints—Molecular dynamics calculations (MD) were carried out using the AMBER 8.0 program [45,46] with the AMBER 99 force field at constant volume (the NVT scheme). All simulations were performed in a periodic box of TIP3P water [47] (with linear dimensions 59.7, 59.4, and 41.1 Å, respectively), at constant volume with the particle-mesh Ewald procedure [48,49] for long-range electrostatic interactions at $T = 303$ K. The Berendsen thermostat was implemented to maintain constant temperature. The size of the water shell was at least 10 Å from each side. The peptides studied were terminally-blocked and, for all three systems (KAK3, KAK4, KAK5), counter ions were added, two for every system. The total time of simulation was 10 ns for each trajectory, and the integration time step was 2 fs. The time-averaged restraint method (TAV) [45,50,51] was used to include experimental values for the calculations, with interproton-distance restraints (calculated from the intensities of the ROE signals or a set of anti-ROE restraints for those pairs of protons between which ROE signals were not observed) and dihedral-angle restraints (calculated from the coupling constants). The numbers of distance restraints were 62 for KAK3, 74 for KAK4, and 81 for KAK5, respectively.

In our earlier work [25, 26, 28, 29], we noticed that all- α -helical conformations are often formed if anti-ROE restraints are not imposed, even though there are no ROE signals characteristic of α -helical contacts. Including anti-ROE restraints along with ROE restraints therefore is necessary. Therefore, anti-ROEs between $j - i + 2$ residue pairs (where i and j denote residue numbers) were included for those pairs of protons for which no ROE signals were observed. This means that the protons involved were assumed to be separated by more than 6 Å, which is justified because ROE signals would be observed if the distances were smaller than 6 Å.

The numbers of dihedral-angle restraints were 8 for KAK3 (pertaining to the ϕ and the χ angles for the lysine and the ϕ angles for the alanine residues), 14 for KAK4 (pertaining to the ϕ and the χ angles for lysines and the ϕ angles for alanines), and a total of 16 for KAK5 (pertaining to the ϕ and the χ angles for lysines and the ϕ angles for alanines). The interproton distances for all three peptides were restrained with the force constant $k = 20$ kcal/(mol $\times \text{\AA}^2$), and the angles with $k = 2$ kcal/(mol $\times \text{\AA}^2$), respectively; the distances correspond to the ROE signals shown in Figure 3. The force constants corresponding to anti-ROE restraints were the same as those corresponding to ROE restraints.

The convergence of simulations was assessed by monitoring the histograms of the Lys(N^{H})-Lys(N^{H}) distances over time windows each constituting $\frac{1}{2}$ of simulation time. The histograms from the last two windows did not differ significantly, as assessed by the χ^2 test.

4000 conformations were obtained for each peptide and the last 1000 of those were analyzed. The set of the final conformations was clustered using the MOLMOL program [33]. MOLMOL uses a hierarchical minimal spanning tree method [52–55]. An rms deviation cut-off of 2.5 Å over the K1–K5, K1–K6, K1–K7 residues was used for the clustering, for the KAK3, KAK4 and KAK5 peptides, respectively.

Acknowledgments

This research was conducted by using the resources of (a) our 794-processor Beowulf cluster at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University, (b) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdansk, and (c) the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw. The publication is financed by the European Social Fund as a part of the project "Educators for the elite - integrated training program for PhD students, post-docs and professors as academic teachers at the University of Gdansk" within the framework of the Human Capital Operational Programme, Action 4.1.1, Improving the quality of educational offer of tertiary education institutions. This publication reflects the views of only the authors, and the funder cannot be held responsible for any use which may be made of the information contained therein. This work is supported by grants for Young Scientists 2011 from the University of Gdansk (JM nr 538-8232-0494-1) of DS/8232-4-0088-1 and from the National Science Centre (UMO-2011/01/D/ST4/04497).

References

1. Montelione GT, Scheraga HA. *Acc Chem Res.* 1989; 22:70–76.
2. Searle MS, Williams DH, Packman LC. *Nat Struct Biol.* 1995; 2:999–1006. [PubMed: 7583674]
3. Fersht AN. *Curr Opin Struct Biol.* 1997; 7:3–9. [PubMed: 9032066]
4. Skwierawska A, Mudzińska W, Oldziej S, Liwo A, Scheraga HA. *Proteins: Structure Function and Bioinformatics.* 2009; 76:637–654.
5. Poland D, Scheraga HA. *Journal of Chemical Physics.* 1966; 45:1456–1463. [PubMed: 5920200]
6. Brown JE, Klee WA. *Biochemistry.* 1971; 10:470–476. [PubMed: 5543977]
7. Silverman DN, Kotelchuck D, Taylor GT, Scheraga HA. *Arch Biochem Biophys.* 1972; 150:757–766. [PubMed: 5065142]
8. Jimenez MA, Herranz J, Nieto JL, Rico M, Santoro J. *J FEBS Lett.* 1987; 221:320–324.
9. Jimenez MA, Rico M, Herranz J, Santoro J, Nieto JL. *Eur J Biochem.* 1988; 175:101–109. [PubMed: 3402443]
10. Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. *J Mol Biol.* 1992; 226:795–817. [PubMed: 1507227]
11. Kuroda Y. *Biochemistry.* 1993; 32:1219–1224. [PubMed: 8383525]
12. Muñoz V, Serrano L, Jimenez MA, Rico M. *J Mol Biol.* 1995; 4:648–669.
13. Blanco FJ, Serrano L. *Eur J Biochem.* 1995; 230:634–649. [PubMed: 7607238]
14. Ramirez-Alvarado M, Kortemme T, Blanco FJ, Serrano L. *Bioorg Med Chem.* 1999; 7:93–103. [PubMed: 10199660]
15. Park SH, Shalongo W, Stellwagen E. *Biochemistry.* 1993; 32:7048–7053. [PubMed: 8334134]
16. Marqusee S, Robbins VH, Baldwin RL. *Proc Natl Acad Sci USA.* 1989; 86:5286–5290. [PubMed: 2748584]
17. Jarrett JT, Berger EP, Landsbury PT Jr. *Ann NY Acad Sci.* 1993; 695:144–146. [PubMed: 8239273]
18. Yon, JM. Protein Aggregation. In: Meyers, RA., editor. *Encyclopedia of Molecular Biology and Molecular Medicine.* Vol. V. VCH; Weinheim: 1996. p. 73-93.
19. Fink AL. *Fold Des.* 1997; 3:R9–R23. [PubMed: 9502314]
20. Matheson RP, Scheraga HA. *Macromolecules.* 1978; 11:819–829.
21. Muñoz V, Henry ER, Hofrichter J, Eaton WA. *Proc Natl Acad Sci, USA.* 1998:5872–5879. [PubMed: 9600886]
22. Dinner AR, Lazaridis T, Karplus M. *Proc Natl Acad Sci, USA.* 1999; 96:9068–9073. [PubMed: 10430896]
23. Skwierawska A, Makowska J, Oldziej S, Liwo A, Scheraga HA. *Proteins: Structure Function and Bioinformatics.* 2009; 75(4):931–953.
24. Lewandowska A, Oldziej S, Liwo A, Scheraga HA. *Biophys Chem.* 2010; 151:1–9. [PubMed: 20494507]
25. Makowska J, Rodziewicz-Motowidło S, Bagińska K, Vila JA, Liwo A, Chmurzyński L, Scheraga HA. *Proc Natl Acad Sci, USA.* 2006; 103:1744–1749. [PubMed: 16446433]

26. Makowska J, Baginska K, Skwierawska A, Liwo A, Chmurzynski L, Scheraga HA. *Biopolymers*. 2008; 90:772–782. [PubMed: 18767125]
27. Zagrovic B, Lipfert J, Sorin E, Millett IS, van Gunsteren WF, Doniach S, Pande VS. *Proc Natl Acad Sci USA*. 2005; 102:11698–11703. [PubMed: 16085707]
28. Makowska J, Baginska K, Makowski M, Jagielska A, Liwo A, Kasprzykowski F, Chmurzynski L, Scheraga HA. *J Phys Chem B*. 2006; 110:4451–4458. [PubMed: 16509748]
29. Makowska J, Baginska K, Liwo A, Chmurzynski L, Scheraga HA. *Biopolymers*. 2008; 90:724–732. [PubMed: 18618612]
30. Sreerama N, Woody RW. *Anal Biochem*. 2000; 287:252–260. [PubMed: 11112271]
31. Tiers GVD, Coon RI. *J Org Chem*. 1961; 26:2097–2098.
32. Wüthrich, K. *NMR of proteins and nucleic acids*. New York: Wiley; 1986. p. 292
33. Koradi R, Billeter M, Wüthrich K. *J Mol Graphics*. 1996; 14:51–55.
34. Makowska J, Liwo A, Chmurzynski L, Scheraga HA. *J Sol Chem*. 2011 submitted.
35. Klotz, IM. *The Proteins*. Neurath, H.; Bailey, K., editors. Vol. 1. 1953. p. 748–753.
36. Plotnikov V, Rochalski A, Brandts M, Brandts JF, Williston S, Frasca V, Lin LN. *Assay Drug Dev Technol*. 2002; 1:83–90. [PubMed: 15090159]
37. Piantini U, Sørensen OW, Ernst RR. *J Am Chem Soc*. 1982; 104:6800–6801.
38. Bax A, Freeman, Enhanced R. *J Magn Reson*. 1985b; 65:355–360.
39. Bax A, Davis DG. *J Magn Reson*. 1985a; 63:207–213.
40. Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K. *J Biomol NMR*. 1995; 6:1–10.
41. Güntert P, Braun W, Wüthrich K. *J Mol Biol*. 1991; 217:517–530. [PubMed: 1847217]
42. Güntert P, Mumenthaler C, Wüthrich K. *J Mol Biol*. 1997; 273:283–298. [PubMed: 9367762]
43. Bystrov VF. *Progr NMR Spectrosc*. 1976; 10:41–81.
44. Güntert P, Wüthrich K. *J Biomol NMR*. 1991; 1:447–456. [PubMed: 1841711]
45. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Wang, B.; Pearlman, DA.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, V.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, JW.; Rosse, WS.; Kollman, PA. *AMBER8*. University of California; San Francisco: 2004.
46. Weiner SJ, Kollman PA, Nguyen DT, Case DA. *J Comput Chem*. 1986; 7:230–252.
47. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J Chem Phys*. 1983; 79:926–935.
48. Ewald PP. *Ann Phys*. 1921; 64:253–287.
49. Darden T, York D, Pedersen L. *J Chem Phys*. 1993; 98:10089–10092.
50. Torda AE, Scheek RM, van Gunsteren WF. *Chem Phys Lett*. 1989; 157:289–294.
51. Pearlman DA, Kollman PA. *J Mol Biol*. 1991; 220:457–479. [PubMed: 1856868]
52. Shamosand MI, Hoey D. Closest-point problems of the 16th IEEE Symposium on the Foundations of Computer. Science. 1975:151–162.
53. Rohlf F. *J Information Processing Letters*. 1978; 7:44–48.
54. Bentley JL, Friedman JH. *IEEE Transactions on Computers*. 1978; C-ZI:97–105.
55. Bentley, JL.; Weide, BW.; Yao, AC. *ACM Transactions on Mathematical Software*. 1980. Optimal expectedtime algorithms for closest point problems; p. 563–580.

Figure 1. The CD spectra of KAK3 (black), KAK4 (dashed line) and KAK5 (green) in water at room temperature ($T = 298$ K).

Figure 2.

Amino acid spin systems in the diagnostic region of the TOCSY spectra (with the mixing time = 80 ms) for (a) KAK3, (b) KAK4, (c) KAK5 in water. The diagnostic region contains the cross-signals of the amide protons (ordinate) with the H α protons and the protons of the side chains (abscissa).

Figure 3. ROE effects corresponding to the interproton contacts of (a) KAK3, (b) KAK4 and (c) KAK5 measured in water at 303K. The thickness of the bars reflects the strength of the ROE correlation as strong, medium, or weak.

Figure 4. Families of the (a) KAK3, (b) KAK4, and (c) KAK5 peptides (MD at T=298K) clustered by using the MOLMOL program. For each family, the C $^{\alpha}$ -traces of all conformations are shown as blue lines and the C $^{\alpha}$ -trace of a selected representative conformation for each family is shown in stick representations. The N- and C-termini of the molecules are indicated.

Figure 5.

Scatter plot of the dihedral angles (ϕ and ψ) of all residues for the last 1000 conformations of: (a) KAK3: Lys1 (black), Ala2 (red), Ala3 (green), Ala4 (yellow), Lys 5 (purple), (b) KAK4: Lys1 (black), Ala2 (red), Ala3 (green), Ala4 (yellow), Ala5 (brown), Lys6 (purple), (c) KAK5: Lys1 (black), Ala2 (red), Ala3 (green), Ala4 (yellow), Ala5 (brown), Ala6 (purple bullets), Lys7 (purple) calculated by MD with NMR restraints.

Secondary structure content estimated from all the CD spectra using the CONTIN/LL program provided within the CDPPro software package.

Table 1

	α -helix [%]	Deformed α -helix [%]	β -sheet [%]	deformed β -sheet [%]	β -turn [%]	statistical coil [%]
KAK3	0,04	5,4	12,8	12	23,9	45,5
KAK4	0,19	0,6	21,8	12,6	23,6	41,7
KAK5	1,5	0,5	23,3	12,8	23,1	39,7

Table 2

Proton chemical shifts [ppm] using DSS as a reference [31] for **KAK3** at 303K.

Residue	Chemical shifts [ppm]						
	NH	α -CH	β -CH	γ -CH	δ -CH	ϵ -CH	ζ -CH
N-Ac		1.91					
Lys 1	8.10	4.13	1.68	1.32	1.61	2.87	7.39
Ala 2	8.20	4.17	1.26				
Ala 3	8.09	4.13	1.26				
Ala 4	8.09	4.16	1.26				
Lys 5	8.10	4.14	1.71, 1.65	1.32	1.58	2.88	7.40
C-NH ₂	7.41, 6.97						

Table 3

Proton chemical shifts [ppm] using DSS as a reference [31] for **KAK4** at 303K.

Residue	Chemical shifts [ppm]						
	NH	α -CH	β -CH	γ -CH	δ -CH	ϵ -CH	ζ -CH
N-Ac		1.92					
Lys 1	8.10	4.14	1.69	1.33	1.57	2.88	7.40
Ala 2	8.22	4.17	1.27				
Ala 3	8.07	4.15	1.28				
Ala 4	8.07	4.15	1.28				
Ala 5	8.09	4.16	1.27				
Lys 6	8.08	4.15	1.63, 1.72	1.33	1.56	2.88	7.40
C-NH ₂	7.40, 6.98						

Table 4
Proton chemical shifts [ppm] using DSS as a reference [31] for **KAK5** at 303K.

Residue	Chemical shifts [ppm]						
	NH	α -CH	β -CH	γ -CH	δ -CH	ϵ -CH	ζ -CH
N-Ac		1.92					
Lys 1	8.09	4.11	1.65	1.33	1.58	2.87	7.39
Ala 2	8.21	4.16	1.26				
Ala 3	8.02	4.13	1.26				
Ala 4	8.04	4.14	1.27				
Ala 5	8.08	4.14	1.26				
Ala 6	8.04	4.15	1.27				
Lys 7	8.05	4.13	1.72, 1.64	1.33	1.57	2.88	7.38
C-NH ₂	7.33, 6.97						

All NMR spectra were calibrated against the DSS signal [31].

Table 5

The $^3J_{HNH\alpha}$ coupling constants obtained from 2D 1H-NMR DQF-COSY and temperature-dependent 1D ^1H -NMR spectra for KAK3, KAK4 and KAK5 peptide.

$^3J_{HNH\alpha}$ vicinal coupling constants [Hz]				
	KAK3	KAK4	KAK5	
Lys	6.2	Lys	7.1	Ala
Ala	6.0	Ala	6.1	n
Ala	6.3	Ala	6.2	Ala
Ala	n	Ala	6.3	Ala
Lys	6.5	Ala	n	Ala
		Lys	6.3	Lys

n- not known

Table 6

Atoms of residues separated by at least 2 residues in sequence ($|i-j| > 1$) between which the ROE peaks were found at 303 K in KAK3, KAK4 and KAK5.

ROE peaks between residues $i-j > 1$		
KAK3	KAK4	KAK5
$\alpha\text{A2} - \beta_2\text{K5}$	$\alpha\text{A2} - \beta_2\text{K6}$	$\alpha\text{A3} - \delta\text{K1}$
		$\alpha\text{A2} - \beta_2\text{K7}$
$\alpha\text{A2} - \delta\text{K5}$	$\alpha\text{A2} - \delta\text{K6}$	$\alpha\text{A2} - \delta\text{K7}$