

Molecular dynamics simulations of a small elastin-like peptide in pure water and in solutions with cosolvents.

Dissertation

zur Erlangung des Grades eines
Doktor der Naturwissenschaften
(Dr. rer. nat.)

eingereicht beim
Fachbereich Chemie der Universität Dortmund

vorgelegt von

Dipl. Phys. Aliaksei Krukau
Minsk, Weissrussland

Dortmund, 2007

Erstgutachter:

Prof. Dr. A. Geiger

Zweitgutachter:

Prof. Dr. R. Winter

Tag der mündlichen Prüfung:

18.12.2007

Danksagung

An dieser Stelle möchte ich mich der Aufgabe widmen, all jenen zu danken, die mir während meiner Promotion mit Rat und Tat zur Seite standen und somit großen Einfluss auf das Gelingen dieser Arbeit hatten.

Zuerst danke ich Herrn Prof. Dr. Alfons Geiger für die Themenstellung und seine großartige Unterstützung meiner Forschungsarbeit.

Herrn Prof. Dr. Roland Winter danke ich für die freundliche Übernahme des Zweitgutachtens und fruchtbare Diskussionen.

Für die zahlreichen Ratschläge und fruchtbare Diskussionen möchte ich an dieser Stelle ganz besonders herzlich Herrn Dr. Ivan Brovchenko und Frau Dr. Alla Oleinikova danken.

Ein besonderer Dank geht an Andreas Appelhagen für die Administration der Rechner und die Hilfe bei der Übersetzung der Zusammenfassung.

Nicht verschwiegen werden sollen auch alle übrigen wissenschaftlichen und nicht wissenschaftlichen Mitarbeiter des Lehrstuhls für Physikalische Chemie IIa, die durch ihre Kollegialität und Hilfsbereitschaft ein wunderbares Arbeitsklima geschaffen haben.

Mein ganz besonderer Dank gilt nicht zuletzt meiner Ehefrau Marina und meiner Familie, die mich auf meinem Weg begleitet, bestärkt und unterstützt haben.

List of publications

The results of the present PhD thesis were published in the following papers:

1. "Thermal breaking of spanning networks in the hydration shell of proteins." I. Brovchenko, A. Krukau, N. Smolin, A. Oleinikova, A. Geiger and R. Winter, *J.Chem.Phys.*, 123, 224905, **2005**.
2. "Percolation transition of hydration water: From planar hydrophilic surfaces to proteins." A. Oleinikova, I. Brovchenko, N. Smolin, A. Krukau, A. Geiger and R. Winter, *Phys. Rev. Lett.*, 95, 247802, **2005**.
3. "Temperature-Induced Conformational Transition of a Model Elastin-Like Peptide GVG-(VPGVG)₃ in Water", A. Krukau, I. Brovchenko and A. Geiger, *Biomacromolecules*, 8, 2196 - 2202, **2007**.

Others articles published during PhD work:

1. "Solubility of simple, nonpolar compounds in TIP4P-Ew." P.E. Krouskop, J. D. Madura, D. Paschek, A. Krukau, *J. Chem. Phys.*, 124, 016102, **2006**.
2. "Water Percolation Governs Polymorphic Transitions and Conductivity of DNA." I. Brovchenko, A. Krukau, A. Oleinikova, and A. Mazur, *Phys. Rev. Lett.*, 97, 137801, **2006**.
3. "Water Clustering and Percolation in Low Hydration DNA Shells." I. Brovchenko, A. Krukau, A. Oleinikova, and A. Mazur. *J. Phys. Chem. B.*, 111, 3258-3266, **2007**.
4. "Water Percolation Governs Polymorphic Transitions and Conductivity of DNA." I. Brovchenko, A. Krukau, A. Oleinikova, and A. Mazur, *Proceedings of the NIC Workshop 2007*, 36, 195 - 197, **2007**.
5. "Ion dynamics and water percolation effects in DNA polymorphism." I. Brovchenko, A. Krukau, A. Oleinikova, and A. Mazur. *JACS*, *in press*, **2007**.

Contents

1	<i>Introduction</i>	13
2	<i>Review</i>	15
2.1	Experimental studies of the protein <i>elastin</i> , <i>elastin</i> -based polymers and small <i>elastin</i> -like peptides.	15
2.2	Computer simulation studies of small <i>elastin</i> -like peptides.	17
2.3	Models of elasticity.	18
2.4	Aims of this work.	20
3	<i>Methods</i>	21
3.1	Molecular dynamic.	22
3.1.1	Hamiltonian of the system.	22
3.1.2	Equation of motion.	23
3.1.3	Periodic boundary conditions.	25
3.1.4	Potential truncation	25
3.2	Forcefields.	27
3.2.1	Non-bonded interactions.	27
3.2.2	Bonded interactions.	29
3.2.3	Special interactions.	30
3.3	Model systems.	32
3.4	Structural characterization of peptide.	34
3.5	Analysis of clustering and percolation of the hydration water.	36
4	<i>Results</i>	38
4.1	Temperature induced conformational changes of the Elastin-Like Peptide. . . .	39
4.1.1	Extension of the ELP.	39

4.1.2	Hydrogen bond patterns.	43
4.1.3	Occurrence of structural elements.	52
4.2	Effect of cosolvents on the structure of the elastin-like peptide.	55
4.2.1	ELP in the aqueous solution of Urea.	56
4.2.2	ELP in the aqueous solution of 2,2,2-Trifluoroethanol (TFE).	62
4.2.3	ELP in the aqueous solution of Sodium Chloride (NaCl).	68
4.2.4	ELP in the aqueous solution of Guanidinium Chloride (GdmCl).	74
4.3	Thermal breaking of the spanning network of hydration water of the elastin-like peptide.	81
5	<i>Discussion</i>	87
5.1	Temperature induced transition of the elastin-like peptide.	87
5.2	Influence of cosolvents on the structure and elasticity of the elastin-like peptide.	91
5.3	Relation between the thermal breaking of a spanning network of hydration water and structural changes of biomolecules.	94
5.4	Origin of elasticity.	98
6	<i>Summary</i>	100
7	<i>Zusammenfassung</i>	102
A	Appendix	104

List of Figures

3.1	Picture of the ELP in the <i>rigid</i> state. Left: Van der Waals representation. Right: Backbone representation.	32
3.2	Picture of the ELP in the <i>flexible</i> state. Left: Van der Waals representation. Right: Backbone representation.	32
3.3	Left panel: Regions of the secondary structure on the Ramachandran plot. Right panel: Type I and Type II β -turn regions on the Ramachandran plot.	35
3.4	Interpretation of the intra-molecular hydrogen-bond map. Blue - 2_7 -helix. Yellow - 3_{10} -helix. Green - α -helix. Red - π -helix. Black - β -sheet. Orange - disordered structure.	36
4.1	Variations of the end-to-end distance R and the radius of gyration S of the ELP with time at different temperatures. From bottom to top: $T = 280$ K, 300 K, 340 K, 380 K.	39
4.2	Probability distributions $P(R)$ of the end-to-end distance R of the ELP at various temperatures (circles) and the fits to eq. (3.40) with $\alpha = 2$ and $\beta = 2$ (lines). To separate different distributions they are shifted vertically. Note also the different scales for the left and right panel.	40
4.3	Probability distributions $P(R)$ of the end-to-end distance R of ELP at $T = 380$ K (circles) and the fits using the equation for the random chain (eq. (3.40) with $\alpha = 2$ and $\beta = 2$, solid lines) and using the equation for the worm-like chain (eq. (3.41), dashed lines). Lower panel: logarithmic scale.	41
4.4	Probability distribution $P(S)$ of the radius of gyration S of ELP at various temperatures (circles) and fits by eq. (3.40) with $\alpha = 1$ and $\beta = 1$ (lines). The different distributions are shifted vertically to avoid overlapping.	42
4.5	Probability distribution $P(L)$ of the maximal extension L of ELP at various temperatures (circles) and fits by eq. (3.40) with $\alpha = 2$ and $\beta = 1$ (lines). The different distributions are shifted vertically to avoid overlapping.	42

4.6	Deviations $\Delta P(R)$ (full circles), $\Delta P(S)$ (squares) and $\Delta P(L)$ (stars) of the probability distributions $P(R)$, $P(S)$ and $P(L)$ from the fits to the eq. (3.40) (upper panel). Parameter B_R (full circles), obtained from the fits of $P(R)$ to the eq. (3.40) (lower panel).	43
4.7	Left panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 280K. Right panel: Probability to find an intra-molecular hydrogen bond between different amino acid in the ELP at 285 K.	44
4.8	Left panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 360 K. Right panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 440 K.	44
4.9	Temperature dependence of the fraction of intra-molecular H-bonds of the ELP with $2 \leq \Delta i \leq 5$ (red circles) and with $\Delta i < 2$ or $\Delta i > 5$ (black squares).	45
4.10	Temperature dependence of the average life time (t_{av}^l) of the intra-molecular hydrogen bonds in the ELP.	46
4.11	Cross correlation between the backbone-backbone and backbone-water hydrogen bonds of the ELP molecule.	47
4.12	Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. T=280K	48
4.13	Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. T=320K	49
4.14	Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. T=340K	50
4.15	Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. T=440K	51
4.16	Dependence of the content of various structural elements on temperature. Black squares: Helices. Red circles: PPII. Green triangles: β -sheet.	52
4.17	Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in pure water. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in pure water. Dashed lines: probability distributions expected for the random formation of clusters of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 360 K, given in figure 4.16.	53

4.18	Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of urea with time at different temperatures. From bottom to top: T = 280 K, 290 K, 300 K, 320 K, 340 K.	56
4.19	Left panel: Probability distributions of the radius of gyration S of the ELP in the aqueous solution of urea at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 320 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the aqueous solution of urea at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 320 K, 340 K). Note also the shifting of the successive distributions and the different scales for the left and right panel.	57
4.20	Left panel: Probability to find intra-molecular hydrogen bonds between different amino acids for the ELP in the aqueous urea solution at 280 K. Right panel: same at 340 K.	59
4.21	Dependence of the content of various structural elements on temperature for the ELP chain in the water solution of urea. Black squares: helices. Red circles: PPII. Green triangles: β -sheet.	60
4.22	Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous urea solution. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in the hydrous urea solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.21).	61
4.23	Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of TFE with time at different temperatures. From bottom to top: T = 280 K, 290 K, 295 K, 300 K, 340 K.	62
4.24	Left panel: Probability distributions of the radius of gyration S of the ELP in the aqueous solution of TFE at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the aqueous solution of TFE at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 340 K). Note also the different scales for the left and right panel and the shift of the distributions.	63
4.25	Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous TFE solution at 280 K. Right panel: same at 340 K.	65

4.26	Dependence of the content of residues in the secondary structure on temperature for the ELP chain in the aqueous TFE solution . Black squares: helices. Red circles: PPII. Green triangles: β -sheet.	66
4.27	Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous TFE solution. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in the aqueous TFE solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.26).	67
4.28	Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of NaCl with time at different temperatures. From bottom to top: T = 280 K, 290 K, 300 K, 320 K, 340 K.	68
4.29	Left panel: Probability distributions of the radius of gyration S of the ELP in the water solution of NaCl at different temperatures (from bottom to top: T = 280 K, 290 K, 300 K, 320 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the water solution of NaCl at different temperatures (from bottom to top: T = 280 K, 290 K, 300 K, 320 K, 340 K). Note also the different scales for the left and right panel and the upward shift of the distributions.	69
4.30	Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous NaCl solution at 280 K. Right panel: same at 340 K.	71
4.31	Temperature dependence of the fraction of the intra-molecular H-bonds of the ELP chain in the aqueous solution of NaCl with $2 \leq \Delta i \leq 5$ (red circles) and with $\Delta i < 2$ and $\Delta i > 5$ (black squares).	71
4.32	Dependence of the content of residues in the secondary structure on temperature for the ELP chain in the aqueous solution of NaCl. Black squares: helices. Red circles: ppII. Green triangles: β -sheet.	72
4.33	Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous NaCl solution. Right panel: Probability n_m to find m successive residues in the ppII conformation for the ELP chain in the aqueous NaCl solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.32).	73

4.34	Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of GdmCl with time at different temperatures. From bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K.	74
4.35	Left panel: Probability distributions of the radius of gyration S of the ELP in the water solution of GdmCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the hydrous solution of GdmCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 340 K). Note also the different scales for the left and right panel and the upward shift of the distributions.	75
4.36	Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous GdmCl solution at 280 K. Right panel: same at 340 K.	77
4.37	Temperature dependence of the fraction of the intra-molecular H-bonds of the ELP chain in the aqueous solution of GdmCl with $2 \leq \Delta i \leq 5$ (red circles) and with $\text{textit}\Delta i < 2$ and $\Delta i > 5$ (black squares).	77
4.38	Dependence of the content of residues with secondary structure elements on temperature for the ELP chain in the hydrous GdmCl solution . Black squares: helices. Red circles: PPII. Green triangles: β -sheet.	78
4.39	Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous GdmCl solution. Right panel: Probability n_m to find m successive residues in the ppII conformation for the ELP chain in the water solution of GdmCl. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.38).	79
4.40	Probability distribution $P(S_{max})$ of the size S_{max} of the largest water cluster, normalized to the total number of water molecules N_w in the hydration shell of the ELP, at different temperatures. The shell width D is 4.5 Å.	81
4.41	Existence probability (SP) of a spanning network in the hydration shell of the ELP.	82
4.42	Distributions n_S of the size S of the water clusters in the hydration shell (shell width $D=4.5$ Å) of the ELP at various temperatures (from top to bottom: $T = 260 - 440$ K). The distributions are shifted subsequently, starting from the bottom. The dashed lines show the power law $n_S \sim S^{-2.05}$, as expected at a 2D percolation threshold. By red color highlighted the distribution closest to the percolation threshold.	83

4.43	Dependence of the mean size S_{mean} of the water clusters in the hydration shell (shell width $D=4.5 \text{ \AA}$) of the elastin-like peptide on the temperature.	84
4.44	Temperature dependence of the average number n_H of water-water hydrogen bonds between water molecules in the hydration shell of the ELP for various choices of the shell width D , which increases from the bottom to the top. The open circles show the values of n_H at the true percolation thresholds ($SP=95 \%$), as obtained for different D values.	84
4.45	Probability distributions of the radius of gyration S of the ELP and of the radius of gyration S_w of the largest hydrogen-bonded cluster in the hydration shell of the ELP at two different temperatures. (The distribution of the S of the ELP is shifted by $+3.5 \text{ \AA}$ at both temperatures (see text).)	85
4.46	Probability distributions $P(H_{max})$ of the distance H_{max} between the centers of mass of the ELP and the largest water cluster in the ELP hydration shell at different temperatures. . .	86
5.1	Left panel: β -turn type I. Right panel: β -turn type II.	88
A.1	Phase separation of the 4M aqueous TFE solution into a water rich (violet) and a TFE rich (green) phase.	105

List of Tables

3.1	The sizes of the systems with cosolvents.	33
4.1	Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous urea solution at different temperatures.	57
4.2	Average numbers of water molecules in the hydration shell of the ELP in the pure water and in the aqueous solution of urea and relative hydration level of the ELP in the urea solution.	58
4.3	Average values, standard deviations and relative fluctuation of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous TFE solution at different temperatures.	63
4.4	Average numbers of water molecules in the hydration shell of the ELP in pure water and in the aqueous solution of TFE and relative hydration level of the ELP in the TFE solution.	64
4.5	Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous NaCl solution at different temperatures.	69
4.6	Average numbers of water molecules in the hydration shell of the ELP in the aqueous solution of NaCl and in the pure water, and relative hydration level of the ELP in NaCl solution.	70
4.7	Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous GdmCl solution at different temperatures.	75
4.8	Average numbers of water molecules in the hydration shell of the ELP in the aqueous solution of GdmCl and in pure water, and relative hydration level of the ELP in GdmCl solution.	76
A.1	Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in pure water at different temperatures.	104

Chapter 1

Introduction.

Elastin is an extracellular matrix protein, responsible for the resilience of tissues such as skin, arteries and lung. It is an insoluble protein, forming fibers which are present in variable amounts, depending on the tissue. The insoluble elastin is an extensively cross-linked polymer of the precursor tropoelastin [1]. Tropoelastin consists of two types of domains. One of them, the cross-linking domain, is rich in lysine and provides cross-linking of the individual monomers, resulting in desmosine and isodesmosine links, which are responsible for its typical yellow color [2]. Several repeating sequences have been described for the second domain of tropoelastin. In order of their frequency of occurrence they are the pentapeptide VPGVG, the hexapeptide APGVGV, the nonapeptide VPGFGVGAG and the tetrapeptide VPGG [3, 4]. In elastin, the most frequent pentapeptide sequence, VPGVG, recurs up to 50 times in a single molecule. Synthetic polymers of $(VPGVG)_n$, where n can be as high as 150, are soluble in water below 300 K, but above a transition temperature $T_t=310$ K the solution undergoes a first order phase transition, it separates into water rich and organic rich phases. The organic rich phase, which shows viscoelastic behavior, contains approximately 63 % of water by weight [5]. Tropoelastin and its synthetic analogs have been the subject of intense investigations in the field of biopolymers and protein engineering. It has been shown that modification of the sequence of the polymers changes the transition temperature T_t , making it possible to design polymeric materials with a particular transition temperature. Additionally one can modulate the transition temperature of a given polymer solution, by other physico-chemical changes, such as change of pH [6], change of oxidation state of the side chains or functional groups [7, 8], addition of salt [9], addition of organic solutes and pressure [6]. These properties of the elastin-like polymers have the promise of providing materials for biomechanical devices [6, 10] as well

as temperature dependent molecular switches [7]. Despite intense research efforts, the detailed structure of tropoelastin and its derivatives still resists elucidation. Hence a molecular-level picture which relates the protein structure to the viscoelastic properties is still a matter of debate. The lack of detailed structural information has prevented the elucidation of the link between structure and function for the tropoelastin and elastin-like peptides. What is beyond doubt, is the determinative role of water as a plasticizing media. Still, the potentially crucial aspect of the protein hydration water remains largely unexplored with some exceptions of experiments [11] and of recent molecular dynamics (MD) simulations [12, 13].

A variety of phenomenological models was proposed to explain the elasticity of elastin and elastin-like peptides, among them: classical rubber elasticity theory [14, 15], "oiled-coils" model [16], "liquid-drop" model [17] and various librational models [6, 18, 19], but non of these models can describe all observed experimental phenomena.

Understanding the microscopic origin of the temperature induced conformational transition of ELP, as well as characterizing its conformations below and above T_t by computer simulation studies, should help to clarify the mechanism of its elasticity. Since water is important for the elasticity of the ELPs, the properties of the hydration water should be studied, to explore the interplay between the state of the hydration water and the properties of the elastin-like peptide.

Chapter 2

Review.

2.1 Experimental studies of the protein *elastin*, *elastin*-based polymers and small *elastin*-like peptides.

The history of *elastin* begins in 1873, with the pioneering work of the swiss scientist Hermann, who for the first time observed that elastic tissue contracts on heating at 65 degree Celsius[20]. In 1893, in his PhD thesis, the german scientist Emil Gotschlich had ascertained that *Ligamentum nuchae* shortens on heating and lengthens on cooling at temperatures bellow 65 degrees. If the tissue would be heated over 65 centigrades, it contracts and could not be elongated again by decreasing the temperature[21]. A lot of attempts were done to define the chemical composition of *Ligamentum nuchae*, but the most successful effort was made by Vandergrift and Gies in 1901, these scientist have found that fresh *Ligamentum nuchae* consists of 57 % of water, 32 % of elastin and 11 % of other chemical substances [22]. In 1914 McCartney, after an extensive study of *Ligamentum nuchae*, has discussed that shortening of the tissue can be due to release of water. Also he has found, for the first time, that on drying at 37 centigrades, strips of *Ligamentum nuchae*, loosing about 50 % of their weight, became hard and brittle, with a great shrinkage in length and thickness. If then immersed again in water, the strips regain all their elastic properties [23]. Twenty four years later Stein and Mueller have made an analysis of the amino-acid composition of elastin. They found that elastin contains approximately 29,4 %, 15,2 % and 13,3 % by weight of glycine, proline and valine respectively. It was not possible at that time to exactly determine whether elastin contains alanine or not. Also, they found small amounts of methionine, histidine, cystein and phenylalanine [24]. Lloyd and Garrod have

investigated the influence of different solvents on the shrinkage temperature and swelling of three proteins: silk fibroin, collagen and elastin. They have found that the elastin, in contrast to the collagen, swells in all three directions. In the judgment of the authors this fact can be an indirect evidence for the absence of complete molecular orientation in the fiber [25]. They have discussed that the elastin molecule is not self-lubricating, and water plays a crucial role for its elastic properties. The water soluble precursor of the protein elastin, tropoelastin, was found in 1955 by Partridge and coworkers [1]. In the same work it was demonstrated that tropoelastin consists of two domains which have different physical properties. In 1958 Hoeve and Flory to the first time proposed to describe the elasticity of elastin with the help of the theory of rubber elasticity [14]. In 1963 Gotte and coworkers made a more precise investigation of the amino acid composition of elastin [26]. Partridge and coworker have found that individual monomers in the tissue are connected by desmosine and isodesmosine cross-links [27]. It was found that elastin starts to show elastic properties already when the weight fraction of water is 30 %, but the full elasticity the protein achieves when the fraction of water by weight is more than 65 % [11]. Several *enthalpic* models, which tried to explain the crucial role of water for the elasticity of elastin, appeared [16, 17, 28].

Dan Urry was the first who synthesized and studied large elastin-like polymers by circular dichroism [29] and NMR [30, 31]. He has shown that aqueous solutions of large ELPs exhibit a lower critical solution temperature (LCST), where they separate into a water rich and an organic rich phase upon heating. He called this phenomenon "inverse temperature transition" [32]. It was shown that large ELPs have an isotropic structure [33] and high mobility [34]. Spectroscopic studies [35] demonstrate the conformational heterogeneity of large ELPs in liquid water, which is reflected in the simultaneous presence of various structural elements (α -helices, β -sheets, β -turns, γ -loops, and polyproline II and disordered structures). The relative population of these structures changes with temperature [35]. A macroscopic phase separation of the aqueous solutions of small ELPs (from one to several pentameric VPGVG units) [7, 36, 37] into organic-rich and water-rich phases upon heating was not detected. Their conformational changes are qualitatively similar to those detected for large ELPs, when crossing the transition temperature T_t , but are observed in a wider temperature interval [7, 36]. In particular, a minimum at 195 nm in the CD spectra, which originates from a disordered structure and/or from polyproline II structures, becomes less pronounced with increasing temperature. Such a behavior was attributed to the increase of "order" in the structure of a single ELP chain upon

heating [5, 6, 7, 35, 36]. The nature and degree of this "order" remains unclear. Of course, the elasticity of a highly ordered polypeptide chain cannot be explained, based on the entropic elasticity of a random chain, and requires another model of elasticity [5, 6, 38].

2.2 Computer simulation studies of small elastin-like peptides.

Computer simulation studies should help to understand the microscopic origin of the temperature induced conformational transition and clarify the mechanism of the ELPs elasticity. In particular, such studies should show whether an ELP exhibits a random distribution of the end-to-end distance or whether its structure is more ordered. As a phase separation of the aqueous solutions of ELPs may be expected, ideally such systems should be studied by appropriate simulation methods (such as simulations in the grand canonical or Gibbs ensemble, etc.). These methods imply the necessity to insert randomly solute molecules in a rather dense liquid phase, and therefore their applicability strongly diminishes with increasing solute size. As a result, now and in the near future, phase transitions in aqueous solutions even of small biomolecules cannot be studied by direct simulation methods. In such a situation, studies of a *single* biomolecule in liquid water remain the most popular simulation approach. They give useful information concerning the properties of the biomolecule and the surrounding water at various thermodynamic conditions. However, when comparing computational results with experiments, the intrinsic inability of single biomolecule simulation to reproduce the phase transition of a solution should be taken into account.

In the pioneering work of Wasserman and Saleme [19] the elastic properties of the polypentapeptide (VPGVG)₁₈ were studied by performing short MD simulations under constraints to mimic an external pulling force. The simulation of the polypentapeptide was performed in a water droplet. During the simulation run the starting β -spiral [35] structure remained stable. The authors found that entropic forces from the reduced librational motions of the protein were a main contributing factor in the elastic behavior. Despite the fact that the simulation times are not exceeding 1 ns, the authors were able to show, at least qualitatively, the entropic nature of the restoring elastic force. However the authors proposed another contribution to the restoring elastic force at low extensions, namely a decrease of the entropy of the water molecules, when they cover more and more hydrophobic surfaces, but this is still a matter of

debate. There are facts which are confirming [39] and contradicting [40] this point of view: indicating an important role of the curvature of the hydrophobic surface. The role of water as plasticizing media was not considered in this work.

The most detailed simulation studies of the temperature induced conformational changes of a single ELP in water were performed for a small elongated pentapeptide GVG(VPGVG) [13, 37] and a polypentapeptide (VPGVG)₁₈ [12]. Two main structures, compact and extended (with and without intra-molecular hydrogen bonds, respectively) were found for the small GVG(VPGVG) peptide. The extended structure dominates in the whole studied temperature range 280 to 380 K, the highest population of the compact structure occurs at about 330 K [13, 37]. Most important, above this temperature a peptide mode, which reflects a transition between the two main structures, noticeably speeds up. Interestingly, the temperature dependence of the dynamics of the water-peptide hydrogen bonds also qualitatively changes at about 330 K. The simulations of the much larger polypentapeptide (VPGVG)₁₈ did not reveal an ordered structure [12], despite the fact that the simulations were started from a conformation, which corresponds to an ordered β -spiral. Already after a short equilibration, this ELP adopts an "amorphous" structure, which is more compact (and has more intra-molecular H-bonds) at higher temperatures. The simulation studies of single ELPs in water did not yet answer the key question, concerning the origin of the elastic properties of the ELPs: does a single ELP show a random distribution of the end-to-end distance in water? The GVG(VPGVG) peptide [13] is too small to show distributions, which are characteristic for long chains, whereas simulation runs of 10 to 20 ns [12] are obviously not long enough for an adequate sampling of the structure of the large polypentapeptide (VPGVG)₁₈ chain in water. To avoid these problems, we have simulated a polypentapeptide of intermediate size GVG(VPGVG)₃ in water during extremely long times (up to 350 ns).

2.3 Models of elasticity.

The origin of the elastic properties of elastin and elastin-like peptides is controversially discussed in many previous works. The models of elasticity can be divided into two groups: single phase model and two phase models.

The single phase model [14, 15] is also known as a random chain model, which considers elastin to be a typical rubber with polymeric chains which are in random configuration and

any diluent present in the system is randomly distributed throughout the swollen material. In accordance with the classical theory of rubber elasticity [41], a chain with Gaussian (random) distribution of the end-to-end distance should show an elastic behavior. Deviation from this distribution causes a decrease of entropy and the appearance of a restoring elastic force. The single phase model is supported by some mechanical [42], calorimetry [43], Raman [44], NMR [34] and birefringence [33] studies. The role of water as plasticizing media essential for the elasticity of ELPs can not be explained from the point of view of the classical theory of rubber elasticity [14, 15].

In the two phase models the appearance of a restoring elastic force is caused by the interactions of the water molecules with exposed hydrophobic surfaces [16, 17, 28]. It was stated that on extension, ELPs expose their hydrophobic groups to water, which should increase the *enthalpy* of the whole "ELP-water" system, thus being the source of the restoring elastic force. The effect of exposure of the hydrophobic groups to the water can be responsible only for 10 to 15 % of the restoring force [45]. Another type of two state model considers the various librational models [6, 18, 19]. The authors of these models state, that the restoring elastic force has two sources, approximately 70 % of the restoring force is coming from the decrease of the librational entropy of the peptide chains due to extension and approximately 30 % of the restoring force is caused by the decrease of the *entropy* of the water molecules in the proximity of the exposed hydrophobic parts of the peptide. However the second source is still a matter of debate [39, 40].

Thus neither the random chain model nor various two phase models can explain the crucial role of water for the elasticity of elastin. The physical origin of the elastic properties of the hydrated elastin and elastin-like peptides is still not clarified.

2.4 Aims of this work.

The aims of the present work are, with the help of computer simulations:

1. to locate the temperature of the conformational transition of the small elastin-like pentapeptide GVG(VPGVG)₃,
2. to characterize conformation of the peptide below and above transition temperature,
3. to shed light on the influence of various cosolvents on the peptide conformation and properties,
4. to describe the properties of the hydration water and the possible interplay between the state of hydration water and the conformation of the elastin-like peptide,
5. to create a consistent model of the elasticity of the elastin-like peptides.

Chapter 3

Methods.

Computer simulations have a valuable role to play in providing essentially exact results for problems in science which would otherwise only be soluble by approximation methods, or might be quite intractable. The results of computer simulations may also be compared with those of real experiments. In the first place, this is a test of the underlying model used in a computer simulation. Eventually, if the model is a good one, the simulator hopes to offer insight to the experimentalist, and assist in the interpretation of new results. This dual role of simulations, as a bridge between models and theoretical predictions on the one hand, and between models and experimental results on the other, makes this method very valuable in science. Computer simulations provide a direct route from the microscopic details of a system to macroscopic properties of experimental interest. It may be difficult or impossible to carry out experiments under extremes of temperature or pressure, while a computer simulation of the materials in, say, a shock wave, a high-temperature plasma, a nuclear reactor, or a planetary core, would be perfectly feasible. Finally, while the speed of molecular events is itself an experimental difficulty, it presents no hindrance to the simulator. A wide range of physical, chemical and biological phenomena, from the molecular scale to the galactic, may be studied using some form of computer simulation [46].

3.1 Molecular dynamic.

3.1.1 Hamiltonian of the system.

The microscopic state of a system may be specified in terms of the positions and momenta of a constituent set of atoms. Within the Born-Oppenheimer approximation, it is possible to express the Hamiltonian of a system as a function of the nuclear variables, the motion of the electrons having been averaged out. We may write the Hamiltonian H of a system of N atoms as a sum of kinetic and potential energy functions of the set of coordinates \mathbf{q}_i and momenta \mathbf{p}_i of each atom i . Adopting condensed notation

$$q = (q_1, q_2, \dots, q_N) \quad (3.1)$$

$$p = (p_1, p_2, \dots, p_N) \quad (3.2)$$

we have

$$H(q, p) = K(p) + V(q) \quad (3.3)$$

Usually, the kinetic energy \mathbf{K} take the form

$$K = \sum_{i=1}^N \sum_{\alpha} p_{i\alpha}^2 / 2m_i \quad (3.4)$$

where m_i is the atomic mass, and the index α runs over the different (x,y,z) components of the momentum of atom i . The potential energy may be divided into terms depending on the coordinates of individual atoms, pairs, triplets etc.:

$$V = \sum_i v_1(r_i) + \sum_i \sum_{i < j} v_2(r_i, r_j) + \sum_i \sum_{j > i} \sum_{k > j} v_3(r_i, r_j, r_k) + \dots \quad (3.5)$$

The first term in the equation represents the effect of an external field on the system. The remaining terms represent particle interaction. The second term, v_2 , the pair potential, is the most important. The pair potential depends only on the magnitude of the pair separation $r_{ij} = |r_i - r_j|$, so it may be written $v_2(r_{ij})$. Three-body and higher terms are expected to be small in comparison with v_2 and can be approximately incorporated in effective pair potentials $v_2^{eff}(r_{ij})$. So we can rewrite eqn. (5) in the form

$$V \approx \sum_i v_1(r_i) + \sum_i \sum_{i < j} v_2^{eff}(r_{ij}). \quad (3.6)$$

The potential energy V contains all interesting information regarding intermolecular interactions: assuming that V is fairly sensibly behaved, it will be possible to construct, from H , an

equation of motion which governs the entire time-evolution of the system and all its mechanical properties. Solution of this equation will generally involve calculation, from V , of forces \mathbf{f}_i , and torques τ_i acting on the molecules. The Hamiltonian also dictates the equilibrium distribution function for molecular positions and momenta.

3.1.2 Equation of motion.

Molecular dynamic is a technique used to solve the classical equation of motion for a system of N molecules interacting via a potential V . The most fundamental form is the Lagrangian equation of motion

$$\frac{d}{dt}(\delta L/\delta \dot{q}_k) - (\delta L/\delta q_k) = 0 \quad (3.7)$$

where the Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}})$ is defined in terms of kinetic and potential energies

$$L = K - V \quad (3.8)$$

and is considered to be a function of the generalized coordinates q_k and their time derivatives \dot{q}_k . If we consider a system of atoms, with Cartesian coordinates r_i and the usual definitions of K and V , then eqn. (7) becomes

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad (3.9)$$

where m_i is the mass of atom i and

$$\mathbf{f}_i = \nabla_{\mathbf{r}_i} L = -\nabla_{\mathbf{r}_i} V \quad (3.10)$$

is the force on that atom. These equations also apply to the center of mass motion of a molecule, with \mathbf{f}_i representing the total force on molecule i ; the equations for rotational motion may also be expressed in the form of eqn. (7). The generalized momentum p_k , conjugate to q_k , is defined as

$$p_k = \delta L/\delta \dot{q}_k. \quad (3.11)$$

The momenta in the Hamiltonian form of the equations of motion are

$$\dot{q}_k = \delta H/\delta p_k \quad (3.12)$$

$$\dot{p}_k = -\delta H/\delta q_k \quad (3.13)$$

The Hamiltonian is strictly defined by the equation

$$H(\mathbf{p}, \mathbf{q}) = \sum_k \dot{q}_k p_k - L(\mathbf{q}, \dot{\mathbf{q}}) \quad (3.14)$$

where it is assumed that we can write \dot{q}_k on the right as some function of the momenta \mathbf{p} . Taking into account that V is independent of velocities and time, this reduces to eqn. (3), and H is automatically equal to the total energy. For Cartesian coordinates, the Hamilton's equations become

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i \quad (3.15)$$

$$\dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i} V = \mathbf{f}_i \quad (3.16)$$

A standard method for the solution of ordinary differential equations such as eqns. (7) and (15,16) is the finite difference approach. The general idea is as follows. Given the molecular positions, velocities, and other dynamic information at time t , we attempt to obtain the positions, velocities etc. at a later time $t + \delta t$, to a sufficient degree of accuracy. The equations are solved on a step-by-step basis; the choice of the time interval δt will depend somewhat on the method of solution, but δt will be significantly smaller than the typical time needed for a molecule to travel its own length. The general scheme of a stepwise MD simulation, based on a predictor-corrector algorithm, may be summarized as follows:

1. predict the positions, velocities, accelerations etc. at a time $t + \delta t$, using the current values of these quantities;
2. evaluate the forces, and hence accelerations $\mathbf{a}_i = \mathbf{f}_i/\mathbf{m}_i$, from the new positions;
3. correct the predicted positions, velocities, accelerations etc. using the new accelerations;
4. calculate any variables of interest, such as energy, virial, order parameters, ready for the accumulation of time averages, before returning to (1) for the next step.

The integration algorithm has to possess some qualities which are crucial for a successful simulation:

- It should be fast, and require little memory
- It should permit the use of a long time step δt
- It should duplicate the classical trajectory as closely as possible
- It should satisfy the known conservation laws for energy and momentum, and be time-reversible

- It should be simple in form and easy to program.

The first MD simulation of a realistic model liquid with a continuous interaction potential has been performed by A. Rahman in 1964 [47].

3.1.3 Periodic boundary conditions.

Computer simulations are usually performed on a small number of molecules, $10 \leq N \leq 10000$. The size of the system is limited by the speed of execution of the program. The time taken for the double loop, used to calculate the forces or potential energy, is proportional to N^2 . A major obstacle for many simulations is the large fraction of molecules which lie on the surface of any sample; for 1000 molecules arranged in a $10 \times 10 \times 10$ cube, no less than 488 molecules appear on the cube faces. Whether or not the cube is surrounded by a containing wall, molecules on the surface will experience quite different forces from molecules in the bulk. The problem of surface effects can be overcome by implementing *periodic boundary conditions* [48]. The cubic box is replicated through the space to form an infinite lattice. In the course of the simulation, as a molecule moves in the original box, its periodic image in each of the neighbouring boxes moves in exactly the same way. Thus, as a molecule leaves the central box, one of its images will enter through the opposite face. There are no walls at the boundary of the central box, and no surface molecules. This box forms a convenient axis system for measuring the coordinates of the N molecules. The number density in the central box is conserved. It is not necessary to store the coordinates of all the images in a simulation, just the molecules in the central box. The cubic box has been common used almost exclusively in computer simulation studies because of its geometric simplicity.

3.1.4 Potential truncation

The heart of the MD program involves the calculation of the potential energy of a particular configuration, and the forces acting on all molecules. Consider how we would calculate the force on molecule 1, or those contributions to the potential energy involving molecule 1, assuming pairwise additivity. We must include interactions between molecule 1 and every other molecule i in the simulation box. There are $N - 1$ terms in this sum. However, in principle, we must also include all interactions between molecule 1 and images i_A, i_B , etc. lying in the surrounding boxes. This is an infinite number of terms, and of course is impossible to calculate in practice.

For a short-range potential energy function, we may restrict this summation by making an approximation. Consider molecule 1 to rest at the centre of a region which has the same size and shape as the basic simulation box. Molecule 1 interacts with all the molecules whose centres lie within this region, that is with the closest periodic images of the other $N - 1$ molecules. This is called '*minimum image convention*'. This technique, which is a natural consequence of the periodic boundary conditions, was first used in simulation by Metropolis et.al. [49]. In the *minimum image convention*, then, the calculation of the potential energy due to pairwise-additive interaction involves $\frac{1}{2}N(N - 1)$ terms. This may still be a very substantial calculation for a system of 1000 particles. A further approximation significantly improves this situation. The largest contribution to the potential and forces comes from neighbors close to the molecule of interest, and for short-range forces we normally apply a spherical cutoff. This means setting the pair potential $v(r)$ to zero for $r \geq r_c$, where r_c is the cutoff distance. In a cubic simulation box of side L , the number of neighbors explicitly considered is reduced by a factor of approximately $4\pi r_c^3/3L^3$, and this can be a substantial saving. The introduction of a spherical cutoff should be a small perturbation, and the cutoff distance should be sufficiently large to ensure this. The cutoff distance must be no greater than $\frac{1}{2}L$ for consistency with the *minimum image convention*.

3.2 Forcefields.

A **force field** is built up from two distinct compositions:

- The set of equations (called the *potential functions*) used to generate the potential energies and their derivatives, the forces.
- The parameters used in this set of equations.

Within one set of equations various sets of parameters can be used. The number of parameter sets developed for application to biologically interesting molecules is great, because of the great complexity of the interactions which involve ionic and polar groups in aqueous solution. Thus, care must be taken that the combination of equations and parameters form a consistent set. It is in general dangerous to make *ad hoc* changes in a subset of parameters, because the various contributions to the total force are usually interdependent. The potential functions can be subdivided into three parts [46]:

1. *Non-bonded*: Lennard-Jones or Buckingham, Coulomb or modified Coulomb. The non-bonded interactions are computed on the basis of a neighbor list (a list of non-bonded atoms within a certain radius), in which exclusions are already removed.
2. *Bonded*: covalent bond-stretching, angle-bending, improper dihedrals, and proper dihedrals. These are computed on the basis of fixed lists.
3. *Special*: position restraints and distance restraints, based on fixed lists.

3.2.1 Non-bonded interactions.

The non-bonded interactions contain a repulsion term, a dispersion term, and a Coulomb term. The repulsion and dispersion term are combined in either the Lennard-Jones (12-6) or the Buckingham (exp-6) potential. In addition, charged atoms act through the Coulomb term. The force on a particle i is then calculated from

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} V_{ij}(\mathbf{r}_{ij}) \quad (3.17)$$

$$\mathbf{F}_i = - \sum_j \frac{dV_{ij}(r_{ij})}{dr_{ij}} \frac{\mathbf{r}_{ij}}{r_{ij}} = -\mathbf{F}_j \quad (3.18)$$

The Lennard-Jones interaction.

The Lennard-Jones potential V_{LJ} between two atoms equals

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6} \quad (3.19)$$

The parameters $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ depend on pairs of *atom types*; consequently they are taken from matrix of LJ-parameters. The force derived from this potential is:

$$\mathbf{F}_i(\mathbf{r}_{ij}) = - \left(12 \frac{C_{ij}^{(12)}}{r_{ij}^{13}} - 6 \frac{C_{ij}^{(6)}}{r_{ij}^7} \right) \frac{\mathbf{r}_{ij}}{r_{ij}} \quad (3.20)$$

The Lennard-Jones potential may also be written in the following form:

$$V_{LJ}(r_{ij}) = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \quad (3.21)$$

In constructing the parameter matrix for the non-bonded LJ-parameters, two types of combination rules are used:

$$C_{ij}^{(6)} = \left(C_{ii}^{(6)} * C_{jj}^{(6)} \right)^{1/2} \quad (3.22)$$

$$C_{ij}^{(12)} = \left(C_{ii}^{(12)} * C_{jj}^{(12)} \right)^{1/2} \quad (3.23)$$

or, alternatively,

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad (3.24)$$

$$\epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{1/2} \quad (3.25)$$

Coulomb interaction.

The Coulomb interaction between two charged particles is given by:

$$V_c(r_{ij}) = f \frac{q_i q_j}{r_{ij}} \quad (3.26)$$

where $f = \frac{1}{4\pi\epsilon_0}$. The force derived from this potential is:

$$\mathbf{F}_i(\mathbf{r}_{ij}) = -f \frac{q_i q_j}{r_{ij}^2} \frac{\mathbf{r}_{ij}}{r_{ij}} \quad (3.27)$$

3.2.2 Bonded interactions.

Bonded interactions are based on a fixed list of atoms. They are not exclusively pair interactions, but include 3- and 4-body interactions as well. There are *bond stretching* (2-body), *bond angle* (3-body), and *dihedral angle* (4-body) interactions. A special type of dihedral interaction (called *improper dihedral*) is used to force atoms to remain in a plane or to prevent transition to a configuration of opposite chirality (a mirror image).

Harmonic bond stretching.

The bond stretching between two covalently bonded atoms i and j is represented by a harmonic potential

$$V_b(r_{ij}) = \frac{1}{2}k_{ij}^b(r_{ij} - b_{ij})^2 \quad (3.28)$$

with the force

$$\mathbf{F}_i(\mathbf{r}_{ij}) = k_{ij}^b(r_{ij} - b_{ij})\frac{\mathbf{r}_{ij}}{r_{ij}} \quad (3.29)$$

Harmonic angle potential.

The bond angle vibration between a triplet of atoms $i - j - k$ is also represented by a harmonic potential on the angle θ_{ijk}

$$V_a(\theta_{ijk}) = \frac{1}{2}k_{ijk}^\theta(\theta_{ijk} - \theta_{ijk}^0)^2 \quad (3.30)$$

As the bond-angle vibration is represented by a harmonic potential, the form is the same as the bond stretching.

The force equations are given by the chain rule:

$$\mathbf{F}_i = -\frac{dV_a(\theta_{ijk})}{d\mathbf{r}_i} \quad (3.31)$$

$$\mathbf{F}_k = -\frac{dV_a(\theta_{ijk})}{d\mathbf{r}_k} \quad (3.32)$$

$$\mathbf{F}_j = -\mathbf{F}_i - \mathbf{F}_k \quad (3.33)$$

where

$$\theta_{ijk} = \arccos \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij}r_{kj}} \quad (3.34)$$

The numbering i, j, k is in sequence of covalently bonded atoms. Atom j in the middle; atom i and k are at the ends.

Improper dihedrals.

Improper dihedrals are meant to keep planar groups planar or to prevent molecules from flipping over to their mirror images:

$$V_{id}(\xi_{ijkl}) = k_{\xi}(\xi_{ijkl} - \xi_0)^2 \quad (3.35)$$

This is also a harmonic potential, since it is harmonic, periodicity not taken into account, so it is best to define improper dihedrals to have a ξ_0 as far away from $\pm 180^\circ$ as possible.

Proper dihedrals.

For the normal dihedral interaction there is a choice of either periodic function or a function based on expansion in powers of $\cos \phi$ (the so-called Ryckaert-Bellemans potential). This choice has consequences for the inclusion of special interactions between the first and the fourth atom of the dihedral quadruple.

Proper dihedrals: periodic type Proper dihedral angles are defined according to the IUPAC/IUB convention, where ϕ is the angle between the ijk and jkl planes, with **zero** corresponding to the *cis* configuration (i and l on the same side).

$$V_d(\phi_{ijkl}) = k_{\phi}(1 + \cos(n\phi - \phi_0)) \quad (3.36)$$

Proper dihedrals: Ryckaert-Bellemans function For alkanes, the following proper dihedral potential is often used:

$$V_{rb}(\phi_{ijkl}) = \sum_{n=0}^5 C_n (\cos(\psi))^n \quad (3.37)$$

where $\psi = \phi - 180^\circ$

3.2.3 Special interactions.

Special potentials are used for imposing restraints on the motion of the system, either to avoid disastrous deviations, or to include knowledge from experimental data. In either case they are not really part of the force field and the reliability of the parameters is not important.

Position restrains.

These are used to restrain particles to fixed reference positions R_i . They can be used during equilibration in order to avoid too drastic rearrangements of critical parts. Another application is the restraining of particles in a shell around a region that is simulated in detail, while the shell is only approximated because it lacks proper interaction from missing particles outside the shell. Restraining will then maintain the integrity of the inner part. It is also possible to restrain particle to a plane or a line.

Distant restrains.

Distance restrains add a penalty to the potential when the distance between specified pairs of atoms exceeds a threshold value. They are normally used to impose experimental restrains, as from experiments in NMR, on the motion of the system. Thus MD can be used for structure refinement using NMR data.

3.3 Model systems.

Our simulation systems can be divided into two different classes:

- binary mixture
- tertiary mixtures

The *binary mixture* is presented by the *elastin like peptide (ELP)* at infinite dilution.

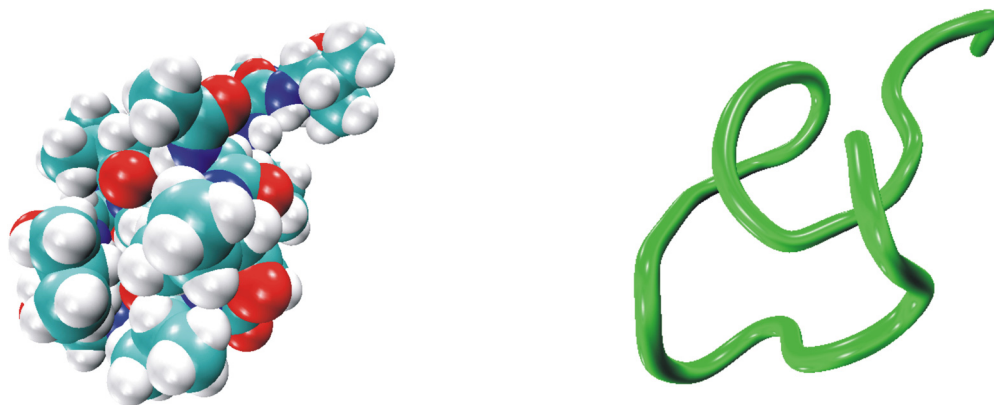


Figure 3.1: Picture of the ELP in the *rigid* state. Left: Van der Waals representation. Right: Backbone representation.

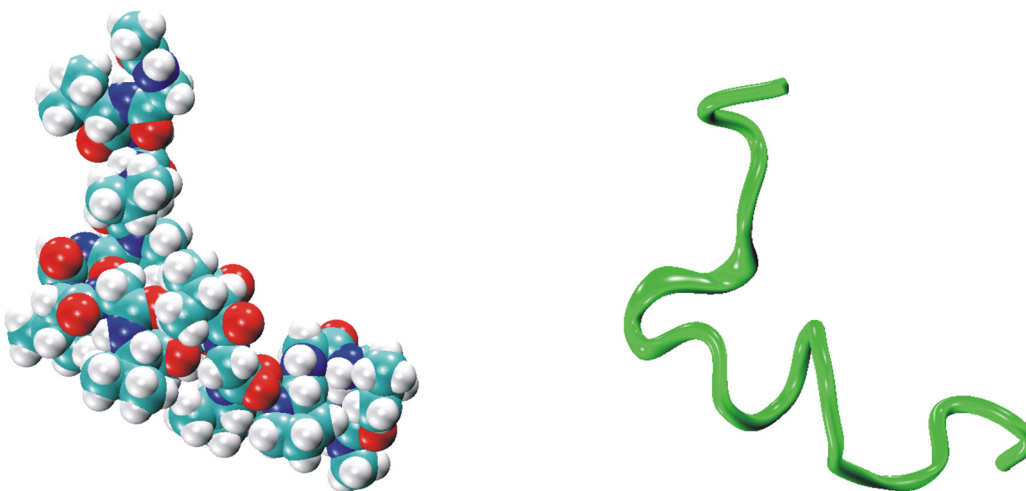


Figure 3.2: Picture of the ELP in the *flexible* state. Left: Van der Waals representation. Right: Backbone representation.

The left handed β -spiral [50] is used as an initial configuration of the small ELP $GVG-(VPGVG)_3$ capped by methylamine ($-\text{NH}-\text{CH}_3$) and acetyl ($-\text{CO}-\text{CH}_3$) at the C- and N-termini,

respectively. The all atom AMBER94 force field [51] is used for the ELP molecule parametrization. Initially ELP was equilibrated in vacuum for 10 ps at 300 K.

For the ELP (see figures (3.1,3.2)) at infinite dilution, one peptide molecule is immersed in a cubic box with 1224 water molecules (SPCE model [52]). The system was equilibrated for 5 ns at 300 K. The resulting configuration was used as an initial one for simulations at 12 temperatures between 280 and 440 K. All simulation runs were performed with 2 fs time step for 350 ns at $T=280, 300, 320, 340, 360,$ and 380 K, for 180 ns at $T=285, 290$ and 295 K, and for 120 ns at $T=400, 420$ and 440 K.

The *tertiary mixtures* are represented by solutions of one ELP in water with addition of *urea, guanidinium chloride (GdmCl), 2,2,2-trifluoroethanol (TFE) and sodium chloride*. Simulations were performed for the 8M urea solution, 1M sodium chloride solution, 4M TFE solution and 3M guanidinium chloride solution (see table (3.1)). The force field parameters for the cosolvents can be found in ref. [53]. All systems were equilibrated for 5 ns at 300 K. The resulting configurations were the starting one for simulation runs at $T=280, 290, 295, 300, 320$ and 340 K. All systems were simulated with 2fs time step for 120 ns. Periodic boundary

System	N_w	$N_{cosolvents}$	Box size / nm
ELP + water + urea	1043	226	x=y=z= 3.67
ELP + water + TFE	1658	179	x=y=z= 4.16
ELP + water + NaCl	1456	40 ion pairs	x=y=z= 3.61
ELP + water + GdmCl	1765	133 ion pairs	x=y=z= 4.07

Table 3.1: The sizes of the systems with cosolvents.

conditions were applied for all systems. The molecular dynamic simulations were performed in the NPT ensemble at constant pressure $P = 1$ bar, using the Nose-Hoover thermostat and the Parrinello-Rahman barostat. The Gromacs software package was used. A spherical cut off of 9 Å was used for the intermolecular interactions, the long-range Coulombic interactions were taken into account by particle mesh Ewald summation [54]. Molecular configurations were stored every 2 ps. Around $6 * 10^4$ to $1.7 * 10^5$ configurations were analyzed for each system.

3.4 Structural characterization of peptide.

The conformation of the ELP was characterized with the help of radius of gyration, end-to-end distance, Ramachandran plots, maps of intra-molecular hydrogen-bonds, etc.

The *radius of gyration* (S) was calculated taking into account all N heavy atoms of the ELP. The formal definition of S is:

$$S \equiv \sqrt{\frac{1}{N} \sum_{k=1}^N (\mathbf{r}_k - \mathbf{r}_{c.o.m})^2} \quad (3.38)$$

The *end-to-end distance* (R) is the distance between the two C_α atoms at both ends of the peptide.

The *maximal extension* (L) is the largest distance between two heavy atoms in the ELP chain.

The *relative fluctuations*

$$\delta_A = \frac{Sd_A}{A_{av}} * 100 \% \quad (3.39)$$

where Sd_A is a standard deviation and A denotes R , S or L , is calculated as a measure of *flexibility* of the ELP chain.

The average values of R , S and L , as well as their probability distributions $P(R)$, $P(S)$ and $P(L)$ were calculated for each temperature and for all systems. The probability distributions were fitted to the equation:

$$P(A) \sim (A - A_0)^\alpha \exp(-B_A(A - A_0)^\beta) \quad (3.40)$$

where A denotes R , S or L ; B , α and β are fitting parameters. The pre-exponential factor should reflect the increase of the configuration space due to an increase of R , S and L . The exponential term corresponds to the Boltzmann distribution of states with free energies, which vary as a power function of R , S or L . For a freely jointed chain $\alpha = \beta = 2$ in the distribution $P(R)$ [41]. To detect the temperature induced conformational changes of the ELP, the parameter B obtained from the fit, as well as the quality of the fits were analyzed as a function of temperature.

The probability distribution of the end-to-end distance R was also fitted to the equation for a semi-flexible, worm-like chain:

$$P(R) \sim (R - R_0)^\alpha \left(\exp(-B_R(R - R_0)^\beta) \right) (1 + C_R(R - R_0)^\gamma) \quad (3.41)$$

which contain one more fitting parameter (C_R) in comparison with eq. (3.40).

The probability distributions n_m to find m successive elements with the same secondary structure were compared for different temperatures with the equation

$$n_m = (1 - \theta)^2 * \theta^m \quad (3.42)$$

for the *random* cluster distribution in a one dimensional infinite chain [55], where θ is the fraction of occupied sites (content of elements with given secondary structure).

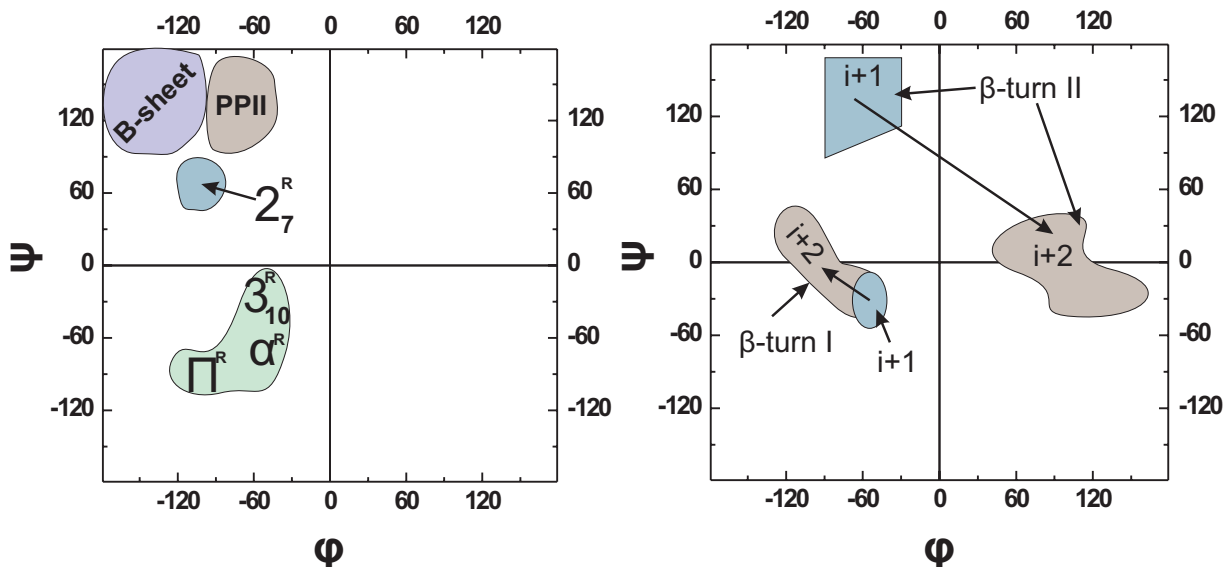


Figure 3.3: Left panel: Regions of the secondary structure on the Ramachandran plot. Right panel: Type I and Type II β -turn regions on the Ramachandran plot.

A *Ramachandran plot* (see figure (3.3)) is a probability distributions for the angles ϕ and ψ , which describe the rotation of amino acids around $N - C_\alpha$ and $C_\alpha - C'$ bonds, where ϕ is the torsion angle between $C_{n-1} - N_n - C_{\alpha n} - C'_n$ and ψ is the torsion angle between $N_n - C_{\alpha n} - C'_n - N_{n+1}$ [56]. The angles were considered positive or negative in keeping with IUPAC "Stereochemical Definitions and Notations Relating to Polymers" (1980).

An *intra-molecular hydrogen-bond map* (HB-Matrix) is a square matrix with probabilities to find hydrogen-bonds between $N - H$ and $C = O$ groups of different amino acids. $N - H$ and $C = O$ groups were decided as hydrogen-bonded if the distance r between N and O does not exceed 3,5 Å, and the $N - H - O$ angle is greater than 130°. If all non-zero values in the matrix lie on the lines parallel to the minor diagonal, we can assume that the peptide has an ordered helical structure, deviations from this pattern can be explained as a tendency of the peptide to lose it's helical structure. As a criterion to measure this deviations, Δi is calculated

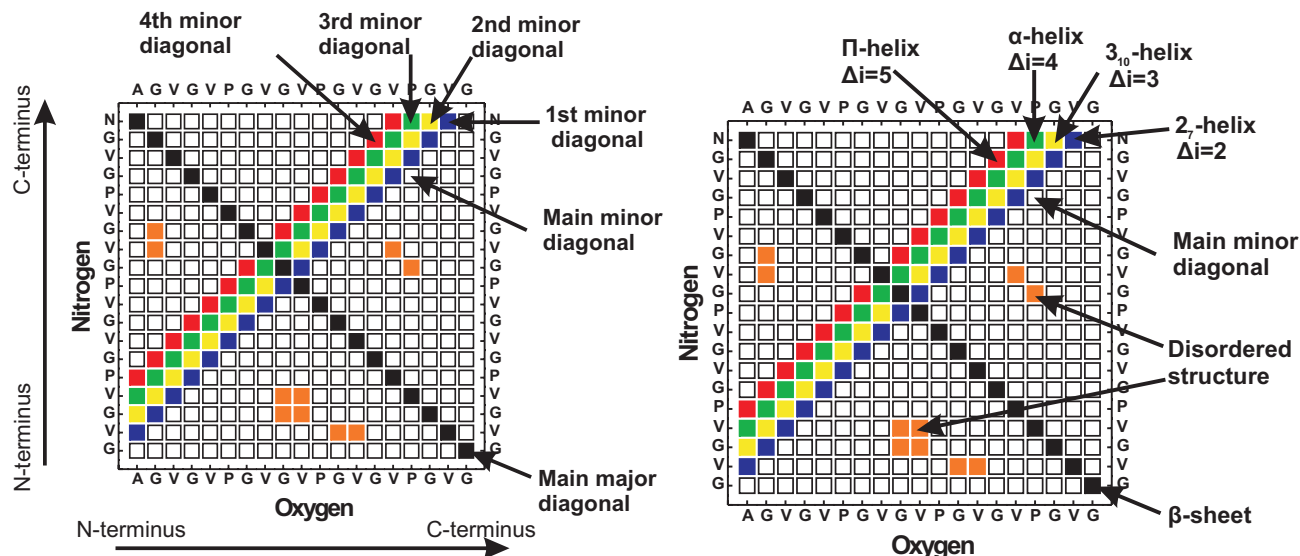


Figure 3.4: Interpretation of the intra-molecular hydrogen-bond map. Blue - 2_7 -helix. Yellow - 3_{10} -helix. Green - α -helix. Red - π -helix. Black - β -sheet. Orange - disordered structure.

as the difference between the order numbers of the C_α atoms of the hydrogen-bonded amino acids. If $2 \leq \Delta i \leq 5$, the peptide has an ordered helical structure, otherwise it has some rate of disorder. Interpretation of the *intra-molecular hydrogen-bond map* is shown in figure (3.4). H-bond motives for different types of helices, such as 2_7 (blue), 3_{10} (yellow), α (green) and π (red) lie on the 1^{st} , 2^{nd} , 3^{rd} and 4^{th} minor diagonals respectively. The hydrogen bonds pattern of the β -sheet structure (black) lies on the main major diagonal of the HB-Matrix. Hydrogen bonds which do not belong to any structural pattern are colored by orange.

3.5 Analysis of clustering and percolation of the hydration water.

The state of the hydrogen-bonded water network in the hydration shell of the ELP at various conditions was characterized by an analysis of the water clustering. A water molecule was considered as belonging to the hydration shell, when the shortest distance between its oxygen and the heavy atoms of the ELP does not exceed 4.5 \AA . This criterion is based on the analysis of the water density distribution near the ELP surface [57]. The existence of a hydrogen-bond between two SPC/E water molecules was used as connectivity criterion. Two molecules were considered as hydrogen-bonded, when the distance between their oxygens did not exceed 3.35 \AA and their pair interaction energy was below -2.7 kcal/mol . Such criteria provide about

3.3 hydrogen-bonds per molecule on average in liquid water at ambient conditions. Water molecules are considered to belong to the same cluster if they are connected by an uninterrupted path of hydrogen bonds. To locate the percolation threshold the following properties, which characterize water clustering, were studied:

- the distribution n_S of the number of clusters of size S ,
- the probability distribution $P(S_{max})$ of the size S_{max} of the largest cluster found at a given time step,
- the mean cluster size S_{mean} (calculated after exclusion of the largest cluster),
- the average number n_H of hydrogen-bonds per water molecule.

The occurrence frequency of water clusters of various sizes S is described by the cluster size distribution n_S , which at the percolation threshold obeys a power law $n_S \sim S^{-\tau}$ with critical exponent $\tau = 187/91 \sim 2.05$ and $\tau \sim 2.2$ in the cases of 2D and 3D percolation respectively [58]. The mean cluster size:

$$S_{mean} = \frac{\sum n_S S^2}{\sum n_S S} \quad (3.43)$$

calculated excluding the largest cluster, diverges at the percolation threshold in an infinite system or passes through a maximum when approaching the threshold in a finite system. The probability distribution $P(S_{max})$ of the size S_{max} of the largest water cluster was used for estimating the spanning probability SP , that is, the probability to observe a spanning cluster in an arbitrary chosen configuration. Its value was obtained as an integral of $P(S_{max})$ over $S_{max} > 0.5N_w$, where N_w is the total number of water molecules in the system (see section 4.3) [57],[59],[60].

The ability of the largest water clusters to cover the ELP molecule homogeneously was characterized by the probability distributions of three parameters, namely, the distance H_{max} between the centers of mass of the largest water cluster and of the ELP molecule, the radius of gyration S_w of the largest cluster, and its maximal extension L_w , that is, the maximal distance between two water oxygens in the largest cluster [57].

Chapter 4

Results

4.1 Temperature induced conformational changes of the Elastin-Like Peptide.

4.1.1 Extension of the ELP.

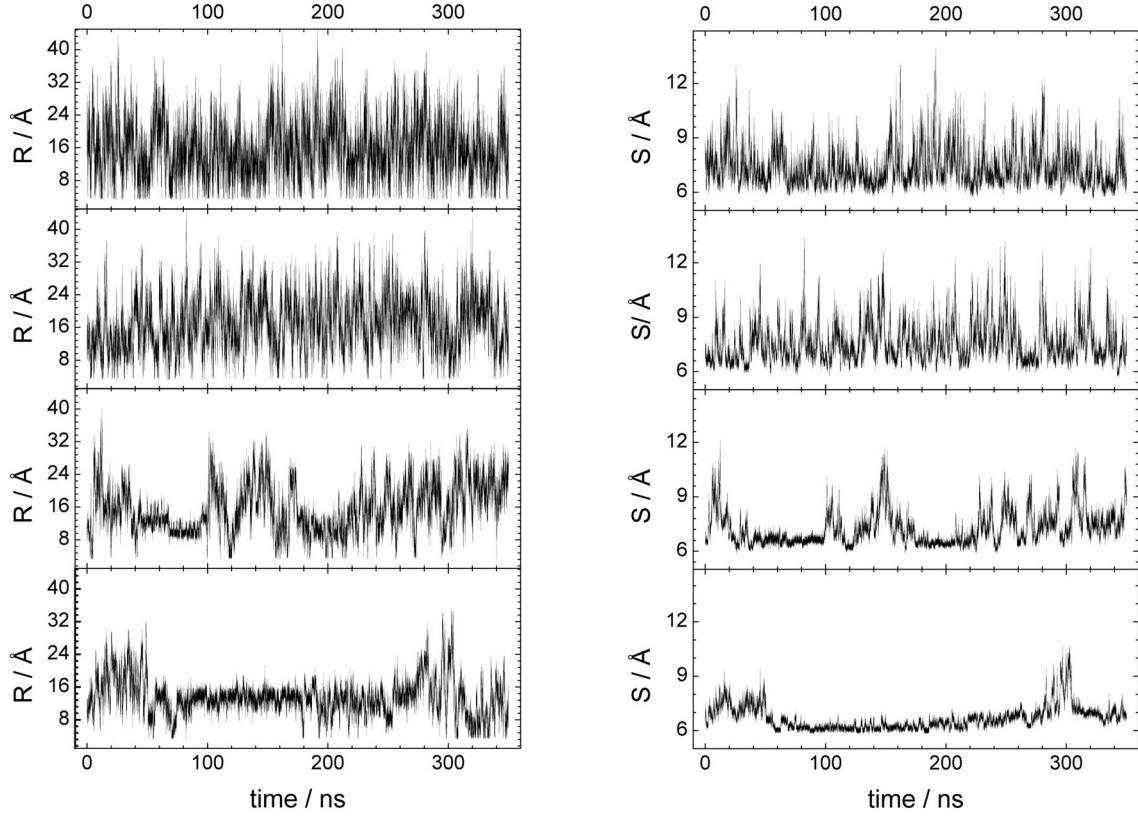


Figure 4.1: Variations of the end-to-end distance R and the radius of gyration S of the ELP with time at different temperatures. From bottom to top: $T = 280$ K, 300 K, 340 K, 380 K.

Time evolutions of the end-to-end distance R and of the radius of gyration S of the ELP are shown in figure (4.1). At $T = 280$ K two different conformational states can be distinguished. In the time interval from $80 - 180$ ns $S_{av} = 6.19 \pm 0.18$ Å, $\delta_S = 3$ % and $R_{av} = 13.03 \pm 1.83$ Å, $\delta_R = 14$ %, whereas in the time interval $200 - 300$ ns $S_{av} = 6.82 \pm 0.66$ Å, $\delta_S = 10$ % and $R_{av} = 13.73 \pm 4.34$ Å, $\delta_R = 31$ %. The first conformational state could be referred as "rigid", because of the small relative fluctuations (defined by equation (3.39)) of R and S , and the second state where the relative fluctuations of R and S are 2 - 3 times larger, could be referred as "flexible". The transitions between these two conformational states are reversible. The fraction of the "rigid" state rapidly decreases with temperature. The "rigid" conformational state of the ELP could be noticed at $T \leq 320$ K, at higher temperatures the ELP is in the "flexible" state all

the time. The time dependence of another structural characteristic of the ELP, the maximal extension L , is qualitatively similar to the time dependences of R and S , shown in figure (4.1). At $T = 280$ K, the average value $L_{av} = 20.40 \pm 0.80$ Å, $\delta_L=4$ %, in the time interval from 80 - 180 ns (rigid state), whereas in the time interval from 200 - 300 ns (flexible state) $L_{av} = 22.76 \pm 2.32$ Å, $\delta_L=10$ %.

To give a more precise definition of the "rigid" and "flexible" state of the ELP, the shapes of the probability distributions $P(R)$, $P(S)$ and $P(L)$ were analyzed at various temperatures (see circles in figures (4.2) -(4.5)). Fitting of the probability distribution $P(R)$ of the end-to-end

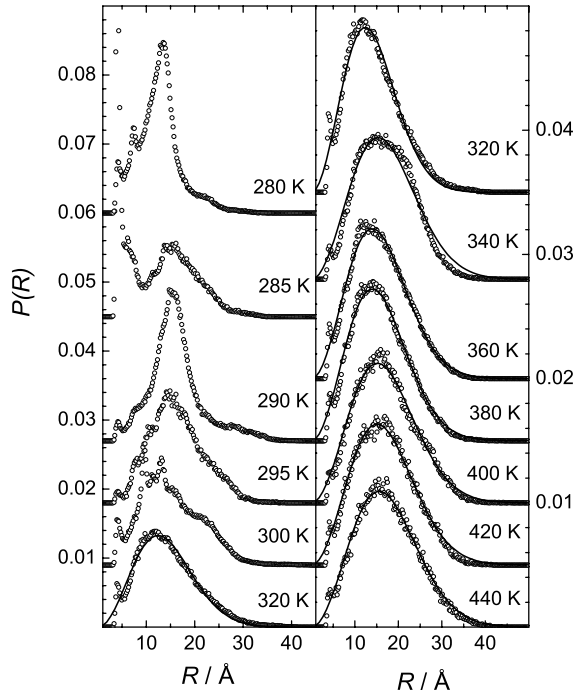


Figure 4.2: Probability distributions $P(R)$ of the end-to-end distance R of the ELP at various temperatures (circles) and the fits to eq. (3.40) with $\alpha = 2$ and $\beta = 2$ (lines). To separate different distributions they are shifted vertically. Note also the different scales for the left and right panel.

distance R of the ELP by eq. (3.40) gives values of the fitting parameters α and β , which vary with temperature. In the temperature interval from 320 to 440 K the closest integral value for both parameters is 2, whereas at lower temperatures α and β strongly deviates from this number. The fitting parameter R_0 was close to zero at all temperatures. Thus, the distribution $P(R)$ was fitted to the eq. (3.40) keeping $\alpha = \beta = 2$ and $R_0 = 0$ (see solid lines in figure (4.2)). In this case eq. (3.40) describes the distribution of the end-to-end distance for a Gaussian (random) chain. The probability distribution $P(R)$ is well fitted by eq. (3.40) at $T \geq 320$ K.

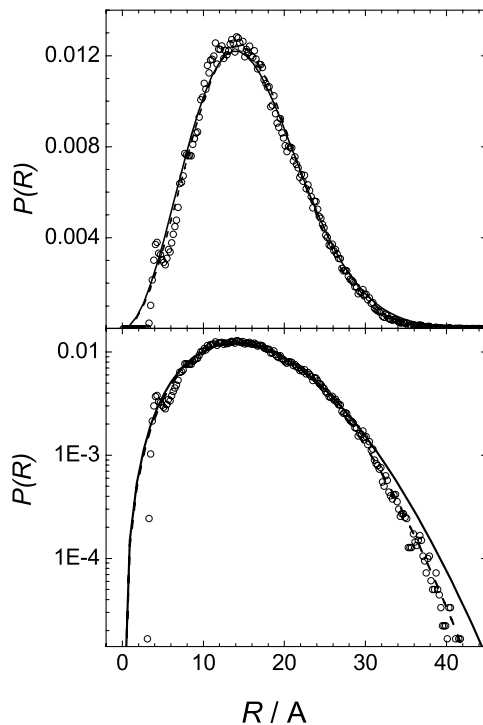


Figure 4.3: Probability distributions $P(R)$ of the end-to-end distance R of ELP at $T = 380$ K (circles) and the fits using the equation for the random chain (eq. (3.40) with $\alpha = 2$ and $\beta = 2$, solid lines) and using the equation for the worm-like chain (eq. (3.41), dashed lines). Lower panel: logarithmic scale.

The distribution $P(R)$ and the fit using eq. (3.40) with α and $\beta = 2$ at $T = 380$ K are shown in more detail in the figure (4.3). A systematic deviation of the fit from $P(R)$ appears only at $R > 30$ Å. This deviation diminishes, when $P(R)$ is fitted by eq. (3.41)(see lower panel figure (4.3)). This equation describes the distribution of the end-to-end distance for a semi-flexible worm-like chain. Note, however, that the presence of one more parameter in eq. (3.41) in comparison with eq. (3.40), improves the fit quality only for large values of R , which the ELP exhibits rather rarely. This is in agreement with experimental data, obtained by the atomic force microscopy stretching of a single tropoelastin chain, which shows worm-like chain behavior [61].

The same kind of analysis was performed for the probability distributions $P(S)$ and $P(L)$ of the radius of gyration S of the ELP and of its maximal extension L (figures (4.4) and (4.5), respectively). Fitting of the probability distribution $P(S)$ by eq. (3.40) gives the closest integral value 1 for the parameters α and β in the temperature interval from 320 to 440 K. Accordingly, the distribution $P(S)$ was fitted by eq. (3.40), keeping $\alpha = \beta = 1$ and $S_0 = 6.05$ Å (the average

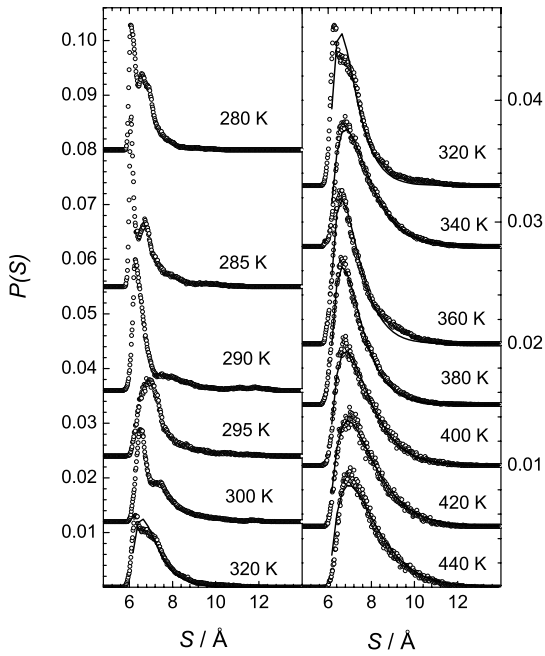


Figure 4.4: Probability distribution $P(S)$ of the radius of gyration S of ELP at various temperatures (circles) and fits by eq. (3.40) with $\alpha = 1$ and $\beta = 1$ (lines). The different distributions are shifted vertically to avoid overlapping.

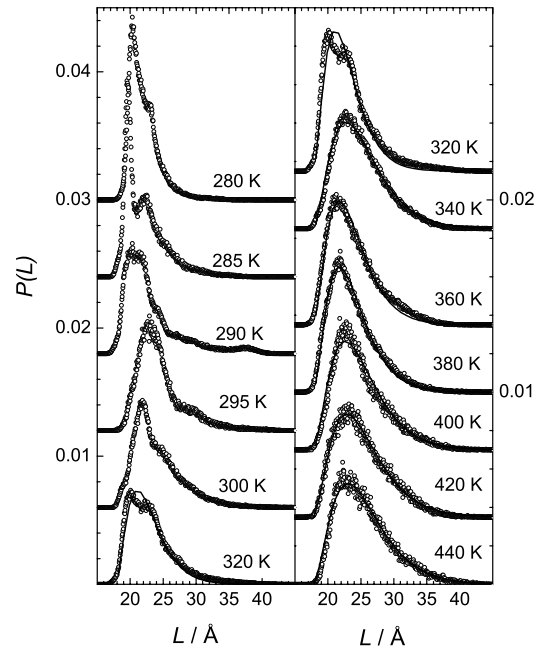


Figure 4.5: Probability distribution $P(L)$ of the maximal extension L of ELP at various temperatures (circles) and fits by eq. (3.40) with $\alpha = 2$ and $\beta = 1$ (lines). The different distributions are shifted vertically to avoid overlapping.

value of S_0 at $320 \leq T \leq 440$ K). The closest integral value 2 for the parameter α and 1 for the parameter β were obtained by fitting eq. (3.40) to the probability distribution $P(L)$ in the temperature range from 320 to 440 K. Consequently, fitting of the probability distribution $P(L)$ by eq. (3.40) was made, keeping $\alpha = 2$, $\beta = 1$ and $L_0 = 17.72$ Å (the average value of L_0 at $320 \leq T \leq 440$ K). It is clearly seen (see figures (4.4) and (4.5)), that the probability distributions $P(S)$ and $P(L)$ (open circles) are in good agreement with Gaussian (random) chain distributions (solid lines) at $T \geq 320$ K. At lower temperatures it was not possible to fit both distributions using the eq. (3.40), as they show multiple peaks.

The ability of eq. (3.40) to describe the probability distributions $P(R)$, $P(S)$ and $P(L)$ is substantially impaired below 320 K. The quality of the fits was characterized by the ratio

$$\Delta P(A) = \chi_A^2(T) / \chi_A^2(440K) \quad (4.1)$$

of the variances $\chi_A^2 = \langle (P(A) - P(A)^{fit})^2 \rangle$, which measure the deviation of the obtained distribution $P(A)$ from the fitting function $P(A)^{fit}$ at a given temperature, and A denotes R, S and L . The temperature dependences of $\Delta P(R)$, $\Delta P(S)$ and $\Delta P(L)$ are shown in the upper panel of figure (4.6). Obviously, the ability of eq. (3.40) to fit adequately the shape

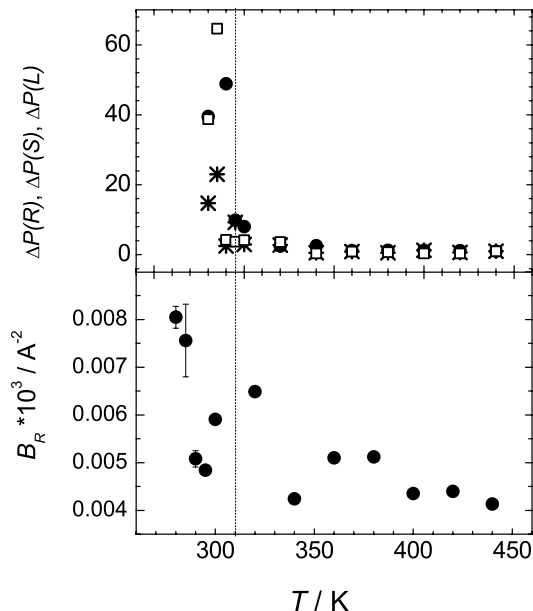


Figure 4.6: Deviations $\Delta P(R)$ (full circles), $\Delta P(S)$ (squares) and $\Delta P(L)$ (stars) of the probability distributions $P(R)$, $P(S)$ and $P(L)$ from the fits to the eq. (3.40) (upper panel). Parameter B_R (full circles), obtained from the fits of $P(R)$ to the eq. (3.40) (lower panel).

of the distributions $P(A)$ is almost independent of T in the range 320 to 440 K, whereas at lower temperatures the shape of $P(A)$ changes drastically, and, in particular, for $P(R)$ it strongly deviates from the distribution of the end-to-end distance for the random chain. The temperature dependence of the fitting parameter B_R , which reflects the behavior of the end-to-end distance R , is shown in the lower panel of the figure (4.6). This parameter varies only slightly at temperatures higher than 340 K, and increases with temperature in the range from 280 to 320 K, indicating a growth of the fraction of the rigid conformational state of the ELP.

Thus, the analysis of the shapes of the probability distributions $P(R)$, $P(S)$, and $P(L)$ at various temperatures evidences a temperature-induced conformational transition of ELP at $T_t \approx 310$ K. In conjunction with the observed fluctuations (figure (4.1)), we may conclude that, at $T > T_t$, the ELP is a flexible chain with random distribution of the end-to-end distance, whereas the fraction of a more rigid conformational state quickly grows upon cooling below T_t .

4.1.2 Hydrogen bond patterns.

Typical probability distributions for the intra-molecular hydrogen bonds (HB-matrix) (interpretation see in section 3.4), between the i -th and $(i+\Delta i)$ -th residues, in the temperature range

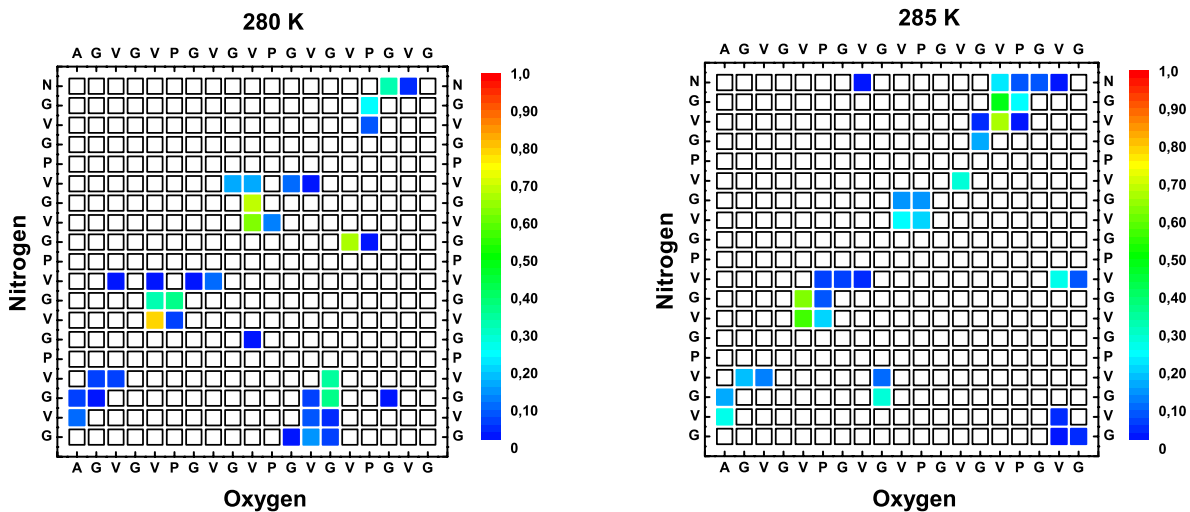


Figure 4.7: Left panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 280K. Right panel: Probability to find an intra-molecular hydrogen bond between different amino acid in the ELP at 285 K.

from 280 to 320 K are shown in figure (4.7). Two different types of hydrogen bonds (H-bonds) are seen on these matrices. First type: **regular** H-bonds, which are located on the first four minor diagonals (see figure (3.4)) of the HB-matrix, with $2 \leq \Delta i \leq 5$; they are typical for 2_7 , 3_{10} , α and π helices. Second type: **irregular** hydrogen bonds which are arranged far away from the main minor diagonal of the HB-matrix, with $\Delta i \leq 2$ or $\Delta i \geq 5$; these hydrogen bonds are typical for a random structure. It is clearly seen in figure (4.7), that at low temperatures the

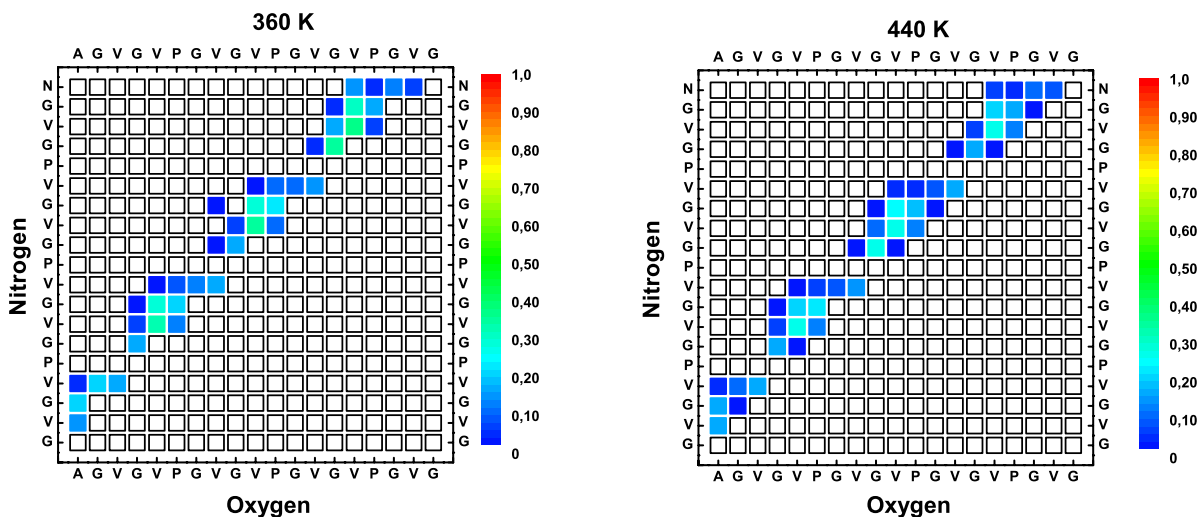


Figure 4.8: Left panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 360 K. Right panel: Probability to find an intra-molecular hydrogen bond between different amino acids in the ELP at 440 K.

majority of the intra-molecular H-bonds are distributed close to the minor diagonal (see also figure 4.9), among them some stable links with existence probability of more than 70 % of the time, which indicates, that the conformation of the ELP is quite stable at low temperatures. On the other hand, there are also some bonds with high existence probability (more than 50 % of time), which are located far away from the minor diagonal. These H-bonds connect amino acids, which are situated close to the ends of the ELP. The presence of such bonds can explain the narrow distribution of the end-to-end distance R at temperatures below 320 K, as these H-bonds can prevent fluctuations of the ends of the ELP chain, and fix them for long time in one conformation.

At the higher temperatures, intra-molecular H-bonds are located exclusively close to the minor diagonal (see figure (4.8)). Two main distinctions exist between the low and high temperature distributions. Firstly: at the temperatures higher then 320 K the probability to find on average a hydrogen bond on the lines near to and parallel to the minor diagonal of the HB-matrices is less than 30 % of the time. This tells us, that there are no long lived structures existing. Secondly: the fraction of H-bonds, for which $\Delta i < 2$ or $\Delta i > 5$, is less then 5% at the

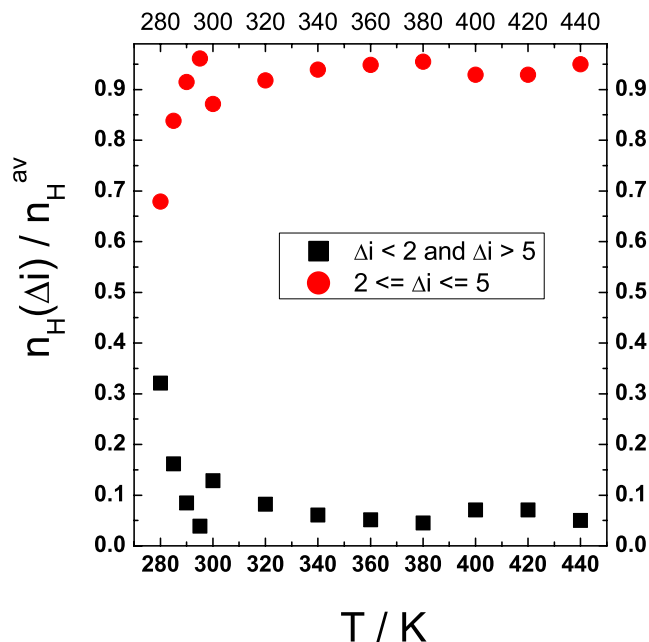


Figure 4.9: Temperature dependence of the fraction of intra-molecular H-bonds of the ELP with $2 \leq \Delta i \leq 5$ (red circles) and with $\Delta i < 2$ or $\Delta i > 5$ (black squares).

high temperatures, thus, the amino acids at the ends of the polypeptide are no more tightly bound to the peptide chain, which allows the ELP to explore configurational space better than

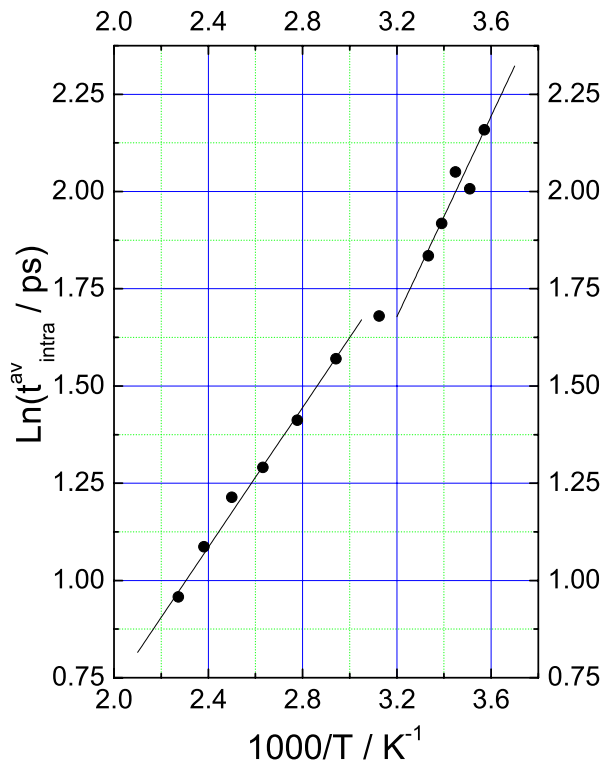


Figure 4.10: Temperature dependence of the average life time (t_{av}^l) of the intra-molecular hydrogen bonds in the ELP.

at low temperatures.

The temperature dependence of the fraction of regular and irregular H-bonds (their numbers normalized by the average number n_H^{av} of intra-molecular H-bonds) is shown in figure (4.9). At $T > T_t=310$ K, the occurrence of regular intra-molecular H-bonds does not change noticeably with temperature. At lower temperatures, their population drastically decreases, and the population of the *irregular* intra-molecular H-bonds sharply increases. Since regular H-bonds indicate the presence of ordered structural elements, the fraction of ordered structural elements also increases with increasing temperature.

The average life time t_{av}^l of the hydrogen bonds, calculated from the probability distributions of the life times of hydrogen bonds [62] at a given temperature, is used (see figure (4.10)) as a measure of the stability of the intra-molecular H-bonds of the ELP. The temperature dependence of the t_{av}^l of the H-bonds has two characteristic intervals, the slope of t_{av}^l is higher at $T < 310$ K, and weaker at $T > 330$ K. The activation energies for the two different intervals were calculated with the help of the Arrhenius equation : $k = A \cdot \exp(-E_a/RT)$. The activation energy E_a is equal to 10.7 kJ/mol for the low temperature interval, and to 7.5 kJ/mol for the

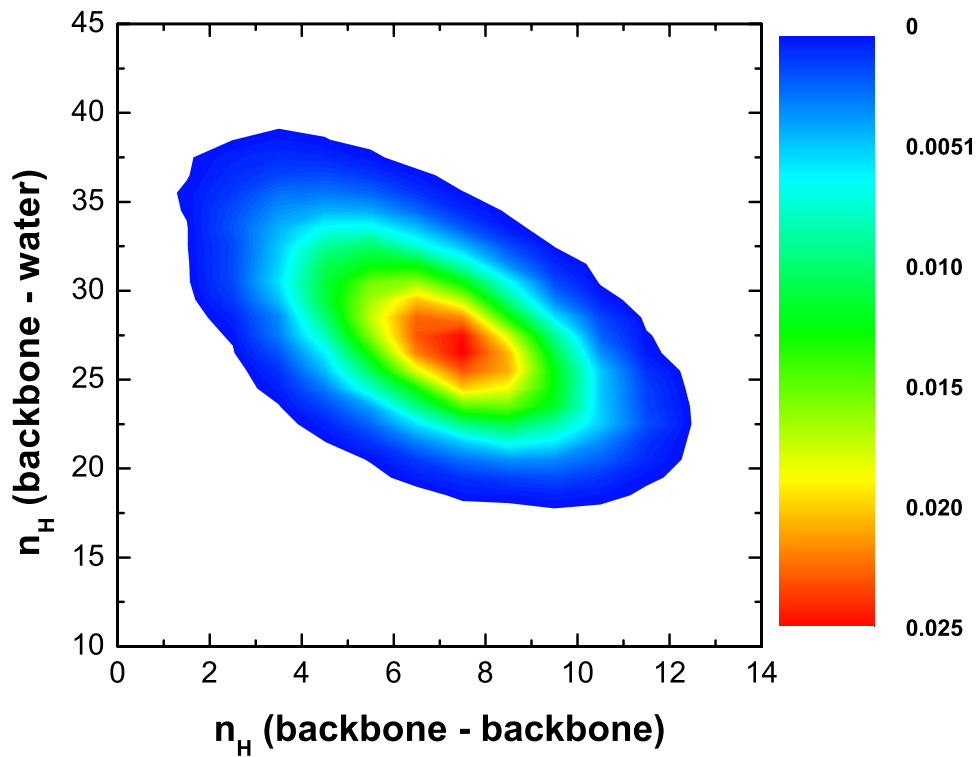


Figure 4.11: Cross correlation between the backbone-backbone and backbone-water hydrogen bonds of the ELP molecule.

high temperature interval. This indicates the presence of an additional force, which stabilizes the conformation of the peptide and its intra-molecular H-bonds, at temperatures lower than 310 K.

The correlation between intra-(backbone-backbone) and inter-molecular(backbone-water) hydrogen bonds is shown in figure (4.11). From the slope of the main axis of the distribution we can conclude that disruption of **one** *backbone-backbone* hydrogen bond leads to the formation of **two** *backbone-water* hydrogen bonds. This evidence is very important for the understanding of the role of water in the elasticity of the ELP (see section 5.4).

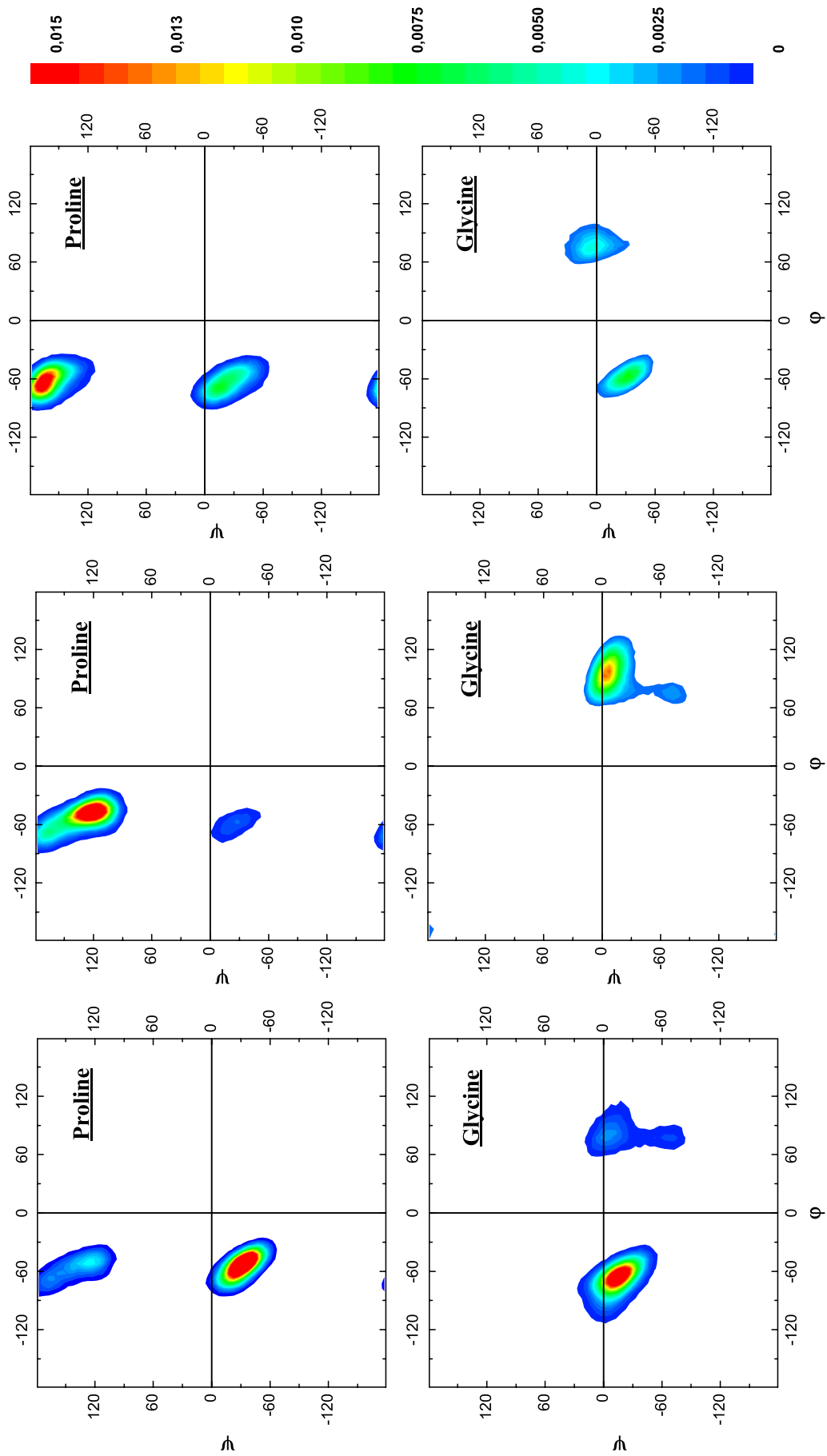


Figure 4.12: Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. $T=280\text{K}$

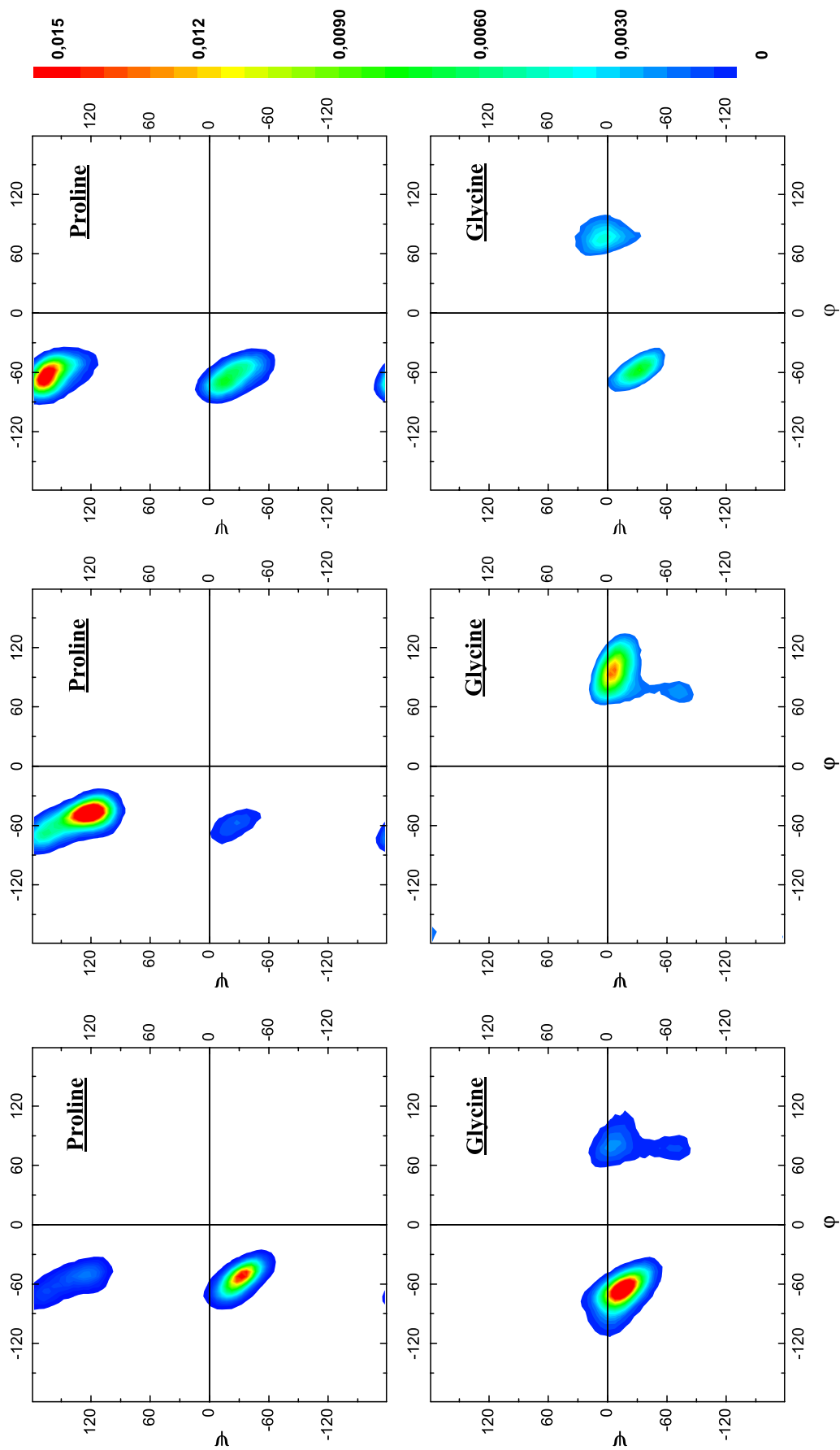


Figure 4.13: Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. $T=320K$

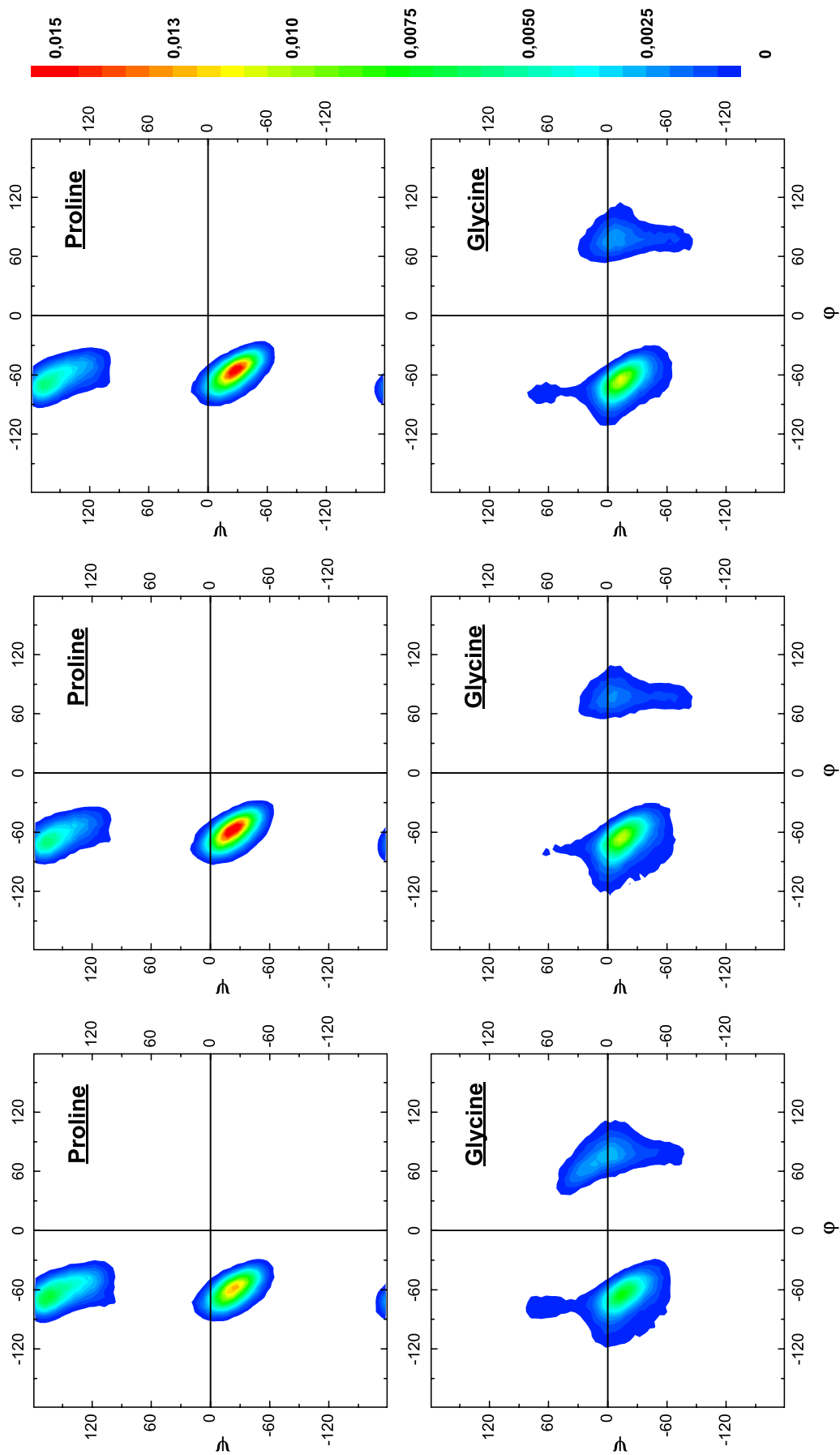


Figure 4.14: Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. T=340K

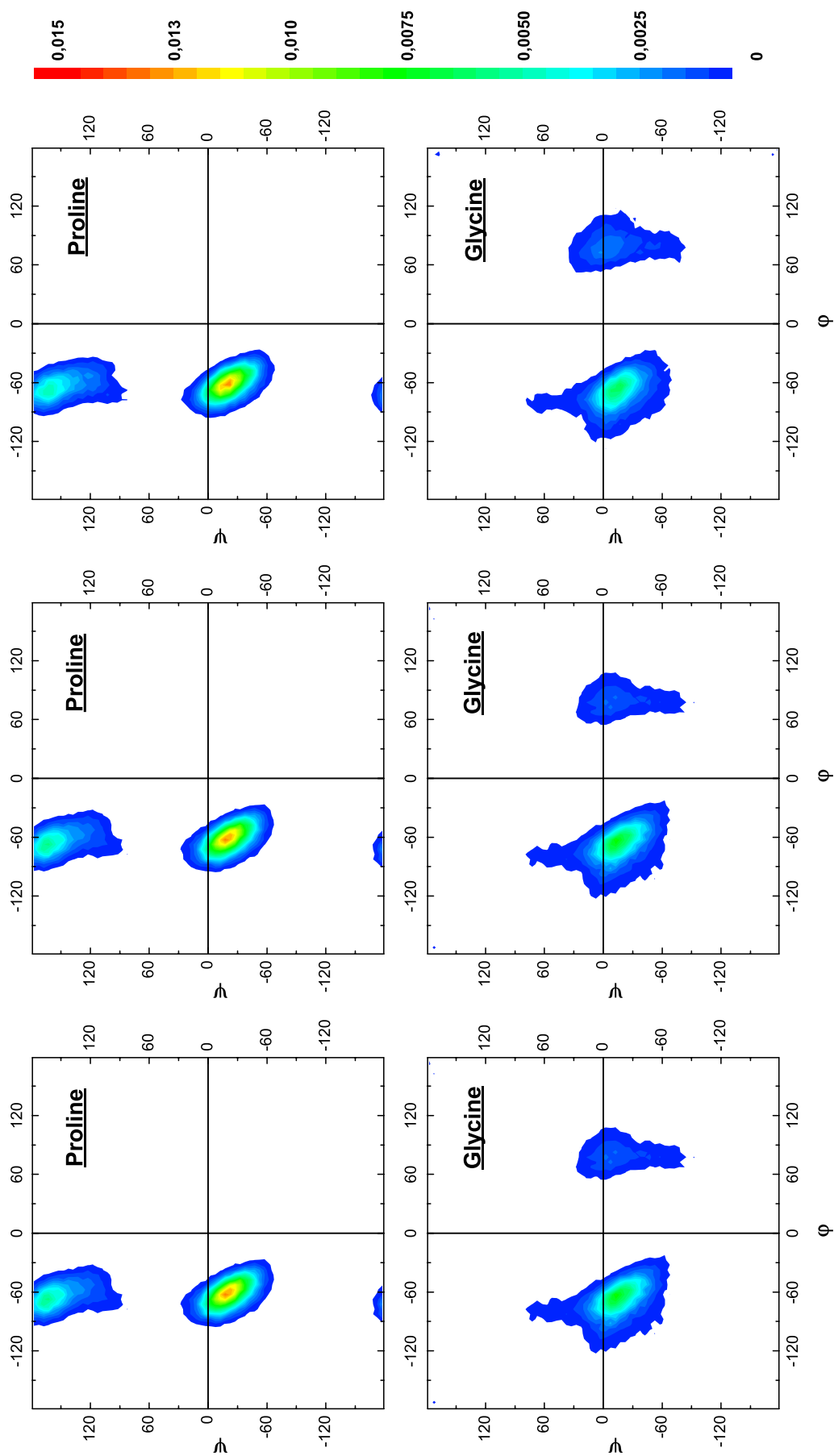


Figure 4.15: Ramachandran plots of Proline and Glycine amino acids in the three different pentameric units. $T=440K$

4.1.3 Occurrence of structural elements.

One of the most common ways to characterize the structure of proteins and peptides is to make two dimensional plots of the occurrence probability of the torsional angles ϕ and ψ for different peptide residues, first introduced by Ramachandran. The Ramachandran plots for the three proline/glycine pairs at different temperatures are shown in figures (4.12)-(4.15). At low temperatures the ϕ and ψ angles for the *first* PG pair lie mainly in the β -turn I region, the largest fraction of the angles of the second and third pairs lie in the β -turn II region, the peaks of the two-dimensional probability distributions are narrow and high (for the identification of the β -turn regions see figure (3.3) in section 3.4: here residue $i+1$ corresponds to P and residue $i+2$ to G.). In the temperature interval from 340 K to 440 K the first PG pair remains in the same β -turn I region, but the second and third pairs change from the β -turn II region to β -turn I region. Both conformations are stabilized by $i \leftarrow i+3$ H-bonds between the neighbouring valines. In fact these bonds are higher populated in the HB-matrices (fig. 4.7). On the other hand, it is difficult to distinguish between the type I β -turn and different types of helices. The

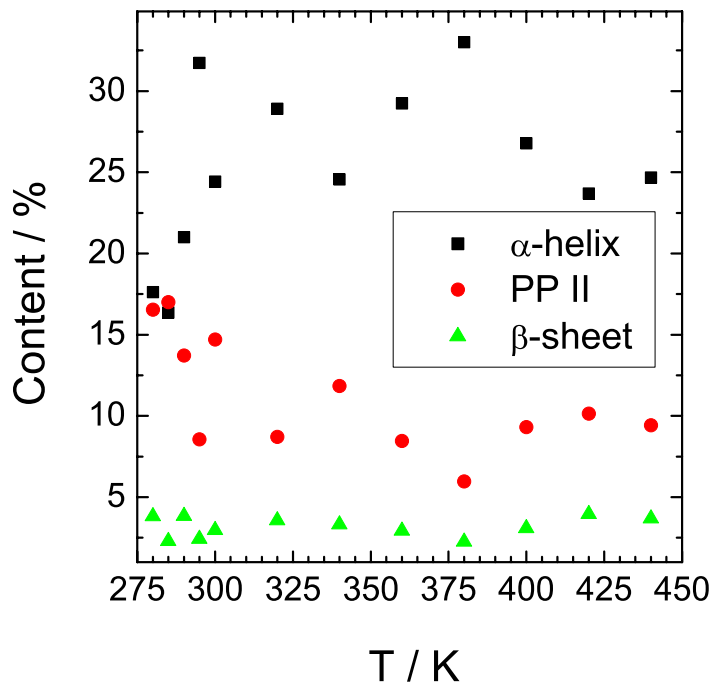


Figure 4.16: Dependence of the content of various structural elements on temperature. Black squares: Helices. Red circles: PPII. Green triangles: β -sheet.

peaks become broader and smaller. This indicates, that at high temperatures the structure of the peptide becomes more flexible and the molecule can better search the configurational

space in contrast to the low temperatures.

The content (occurrence probability per amino acids) of various structural elements, calculated from the Ramachandran plots, is shown in figure (4.16). It is easy to calculate from figure (4.16) that the fraction of all structural elements together is lower than 50 %, and it does not change with temperature. The content of amino acids in β -sheet conformation is lower than 5 % and also independent of temperature. Upon heating up to $T_t=310$ K the content of the residues in α -helical conformation grows and the content of the residues in PPII conformation reduces. At high temperatures, these occurrence frequencies varies, but do not show any clear trends.

The probability n_m to find m successive amino acids in the same conformation, normalized by the overall number of structural elements in this conformation, is shown in figure (4.17). At all temperatures which were studied, about 40-60 % of all residues do not have neighbors

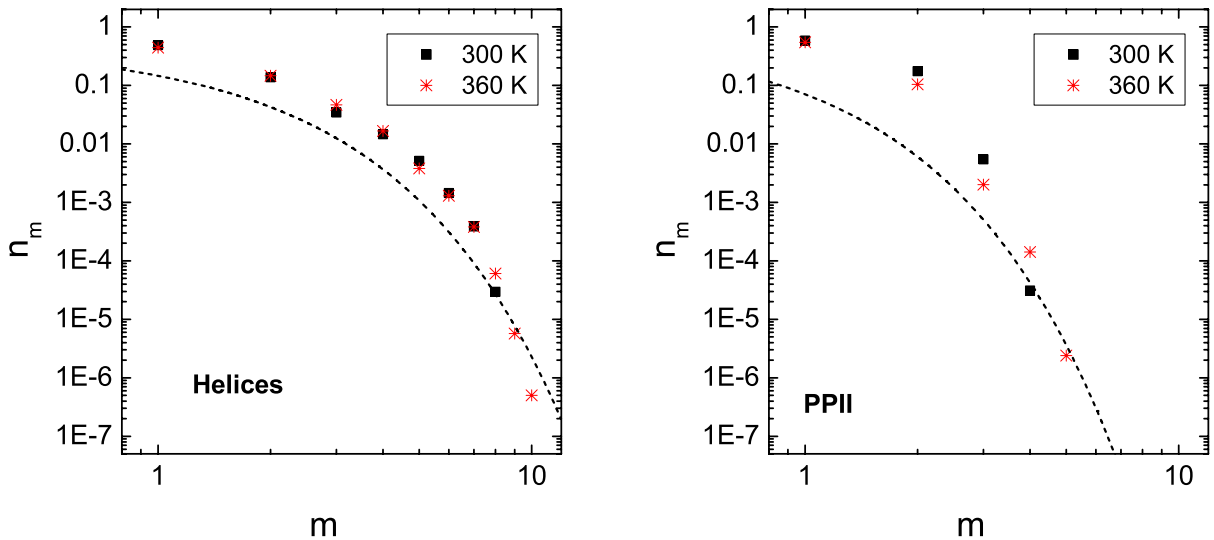


Figure 4.17: Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in pure water. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in pure water. Dashed lines: probability distributions expected for the random formation of clusters of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 360 K, given in figure 4.16.

with the same structure ($m=1$). About 20-30 % of all residues with the same structure form pairs ($m=2$). The probability to find m successive residues with the same structure drastically decreases with increasing m in strictly monotonic way, the probability n_m is less than 1 % already for $m=3$. The shapes of the probability distributions n_m agree closely with the shapes of the *random* cluster distributions in an infinite one dimensional chain, calculated using equa-

tion (3.42) (see figure (4.17)). Deviation of the probabilities n_m to higher values is caused by the finite size of the ELP chain. This indicates an essentially random distribution of the amino acids in the α -helical or PPII conformation along the peptide chain.

4.2 Effect of cosolvents on the structure of the elastin-like peptide.

The addition of cosolvents to solutions of peptides and proteins can result in a variety of effects such as denaturation, increased or decreased solubility, and secondary structure formation. Some of these effects arise as a direct consequence of relatively simple modifications of the properties of water, while others appear to be more complicated and specific to particular cosolvent. In this context, cosolvents are classified according to their ability to both stabilize the native structure and to decrease the solubility of proteins in solution. The chaotrope and kosmotrope classification is also frequently used to refer to protein structure denaturant and stabilizers, respectively. Urea and guanidinium chloride (GdmCl) are well known protein denaturants [63]. Trifluoroethanol (TFE) is known to promote the formation of secondary structure, especially helices, at low concentrations, while higher concentrations can denature protein [64].

The effect of urea, 2,2,2-trifluoroethanol, guanidinium chloride and sodium chloride (NaCl), on the conformation and elasticity of the ELP chain is poorly studied. There are only few experimental works about the influence of these cosolvents on the ELP molecules [7, 9, 65]. Unfortunately, the exact mechanism by which these cosolvents interact with the ELP is still unclear. With the help of computer simulations we tried to shed light on the mechanism of interactions of these cosolvents with the elastin-like peptide.

4.2.1 ELP in the aqueous solution of Urea.

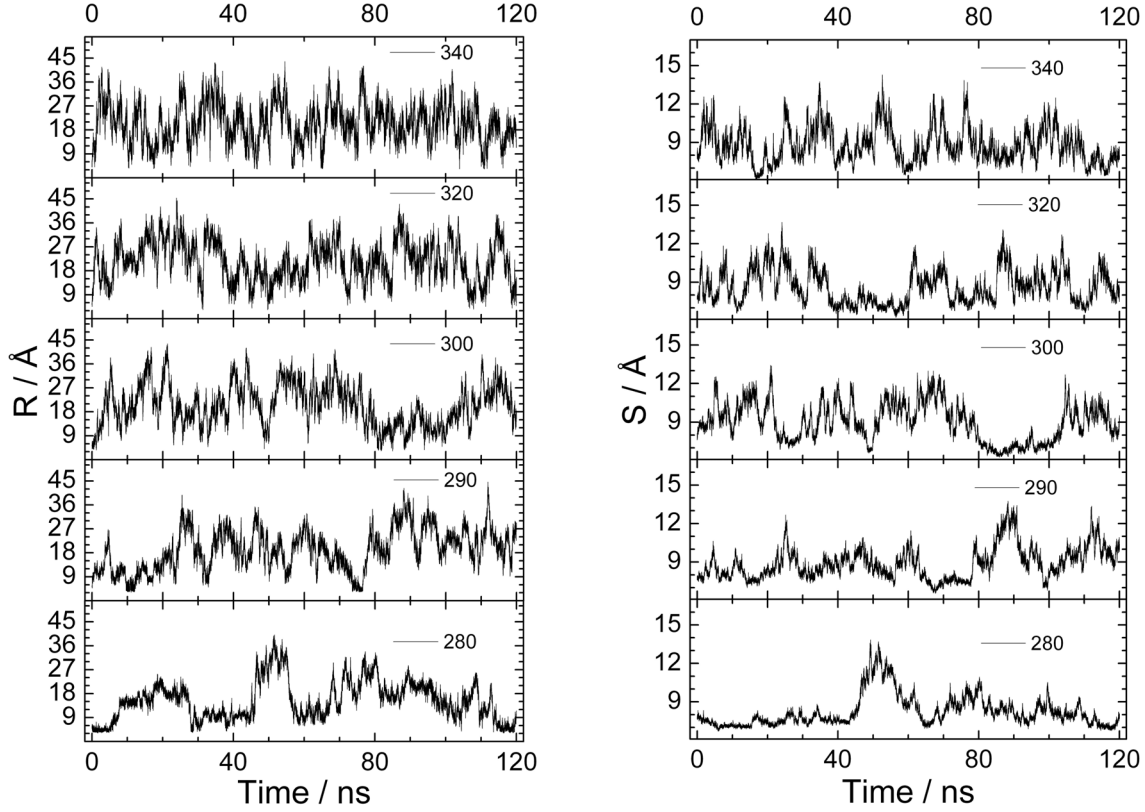


Figure 4.18: Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of urea with time at different temperatures. From bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K.

The time evolution of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous urea solution is shown in figure (4.18). Two different states, *compact* ($T = 280$ K, time interval from $30 - 35$ ns, $S_{av} = 7.73 \pm 0.38 \text{ \AA}$, $\delta_S = 5 \%$) and *extended* ($T = 280$ K, time interval $50 - 55$ ns, $S_{av} = 11.8 \pm 0.65 \text{ \AA}$, $\delta_S = 6 \%$), are also present here in the whole temperature range, the conversion from the *compact* state to the *extended* state is reversible. The fraction of *compact* states is reducing with increasing temperature, but even at $T = 340$ K *compact* states can be found.

The average values and standard deviations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous urea solution are shown in table (4.1). Comparing the average values listed in table (4.1) with the average values of the ELP in pure water at the temperature 440 K (see table (A.1) in Appendix.) it is obvious that the ELP at low temperatures in the aqueous urea solution is more extended than even the *flexible* ELP

Temp / K	R_{av} / Å	Sd / Å	δ_R / %	S_{av} / Å	Sd / Å	δ_S / %	L_{av} / Å	Sd / Å	δ_L / %
280	15.34	7.28	47	8.25	1.24	15	25.18	4.41	18
290	19.15	7.68	40	8.90	1.24	14	28.49	4.62	16
295	22.85	8.35	37	9.51	1.51	16	30.24	5.45	18
300	20.41	7.80	38	9.04	1.48	16	28.85	5.27	18
320	21.54	7.47	35	8.63	1.30	15	28.25	4.98	18
340	20.55	7.36	36	8.78	1.37	16	28.20	4.98	18

Table 4.1: Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous urea solution at different temperatures.

in the pure water at high temperatures. The average values R_{av} , S_{av} and L_{av} slightly increase between 280 K and 290 K and stay quite stable in the temperature range from 295 K to 340 K. The relative *fluctuations*, defined by equation (3.39), of the end-to-end distance R ,

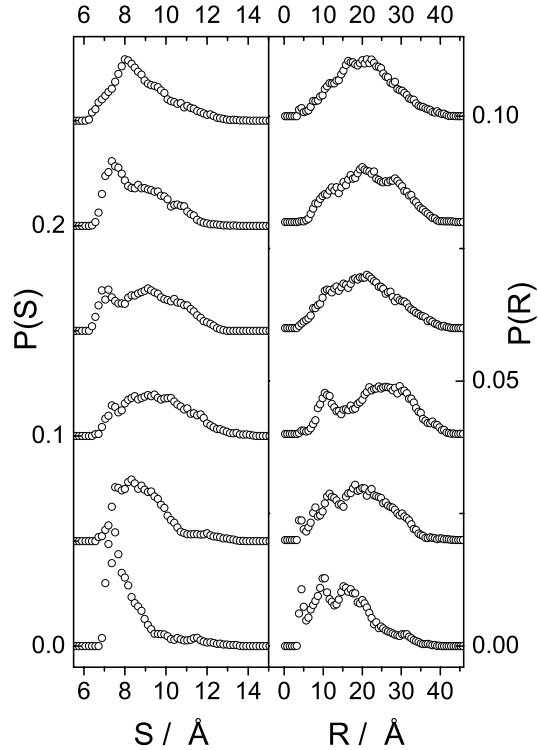


Figure 4.19: Left panel: Probability distributions of the radius of gyration S of the ELP in the aqueous solution of urea at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 320 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the aqueous solution of urea at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 320 K, 340 K). Note also the shifting of the successive distributions and the different scales for the left and right panel.

radius of gyration S and maximal extension L of the ELP chain in the aqueous urea solution are comparable with the fluctuations of R , S and L in the *flexible* state of the ELP chain in pure water. This could be interpreted such that in aqueous urea solution the ELP chain has the same flexibility like in pure water. The problem is that in the aqueous urea solution the relative fluctuations δ averaged over the whole time period reflect another fact: even at high temperatures ($T > 300$ K), both the *compact* and the *extended* states are present (see also figure (4.19)), and conversion from one state to another will enormously increase δ . That is why in the aqueous urea solution δ should be calculated for different conformational states separately. Thus for the *extended* state at $T=340$ K in the time interval from 50 - 55 ns $S_{av}=10.7\pm 1.28\text{\AA}$ and $R_{av}=27.35\pm 6.9\text{\AA}$ consequently $\delta_S=12\%$ and $\delta_R=25\%$. These values are, accordingly, 15 % and 35 % smaller than corresponding values for the *flexible* state of the ELP in pure water at high temperatures (see table (A.1) in Appendix). Based on these observations we can conclude that ELP is less flexible in the aqueous urea solution than in pure water.

The probability distributions of the radius of gyration S and of the end-to-end distance R are shown in figure (4.19). It was impossible to fit the shapes of these distributions, neither by using eq. (3.40) (random Gaussian chain), nor using eq. (3.41) (semi-flexible worm-like chain), as they show multiple peaks in the whole temperature range, reflecting the above mentioned interconversion between different states. Thus, in the aqueous urea solution the ELP does **not** show a random or semi-flexible coil behavior, even at high temperatures. It seems that urea

Temp / K	$N_{w,av}^{Urea}$	$N_{w,av}^{Pure}$	$N_{w,av}^{Urea} / N_{w,av}^{Pure} / \%$
280	74.3	132.4	56
290	79.5	134.2	59
295	83.2	136.7	61
300	79.8	134.7	59
320	79.9	129.3	62
340	80.2	129.8	62

Table 4.2: Average numbers of water molecules in the hydration shell of the ELP in the pure water and in the aqueous solution of urea and relative hydration level of the ELP in the urea solution.

molecules keep the ELP chain in two to four energetically favorable conformational states, and prevent the ELP chain from searching the whole conformational space.

The average number of water molecules in the hydration shell of the ELP in the aqueous urea solution and the relative hydration level, defined as the ratio of the average numbers of water molecules in the hydration shell (definition see in section 3.4) of the protein in the aqueous urea solution and in pure water $N_{w,av}^{cosolvent}/N_{w,av}^{pure}$ are shown in table (4.2). As it is clearly seen from the table (4.2) only 60 % of the peptide surface is covered by water. One should take into account that the solvent accessible surface of the ELP in urea solution is even larger than in pure water, as indicated by increased radius of gyration. This could be the reason for the *rigid* character of the ELP chain in the aqueous urea solution, as water, presumably, plays the role of a plasticizing medium for the ELP molecule.

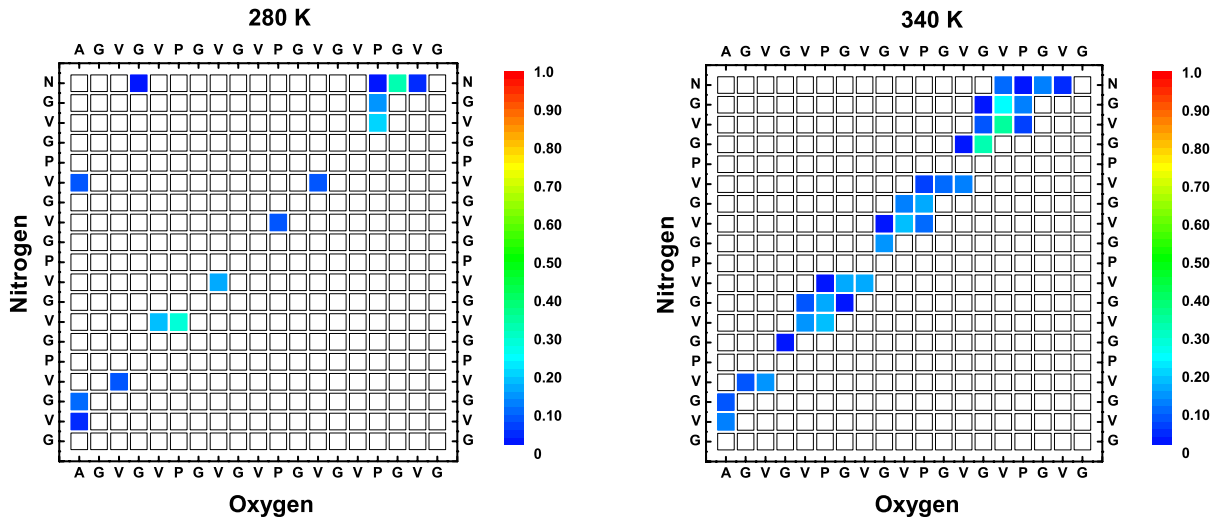


Figure 4.20: Left panel: Probability to find intra-molecular hydrogen bonds between different amino acids for the ELP in the aqueous urea solution at 280 K. Right panel: same at 340 K.

For a more detailed analysis of the ELP conformation, the probability distribution of the intra-molecular H-bonds between the residues, were calculated (HB-matrix) (for interpretation see figure (3.4) in section 3.4). Almost all H-bonds are situated on the second or third minor diagonals (see figure (4.20)). The lack of other H-bonds is characteristic for an extended chain: the *irregular* H-bonds, which are located far away from the minor diagonal, are almost not present, neither at low (figure (4.20) left panel), nor at high (figure (4.20) right panel) temperatures, thus there are no links which bind ends of the peptide and promote compact configurations. From the values of the probabilities (color coded in figure (4.20)) we can also conclude that the intra-molecular hydrogen bonds are less stable in comparison with the pure water case.

The temperature dependence of the content of the different structural elements along the

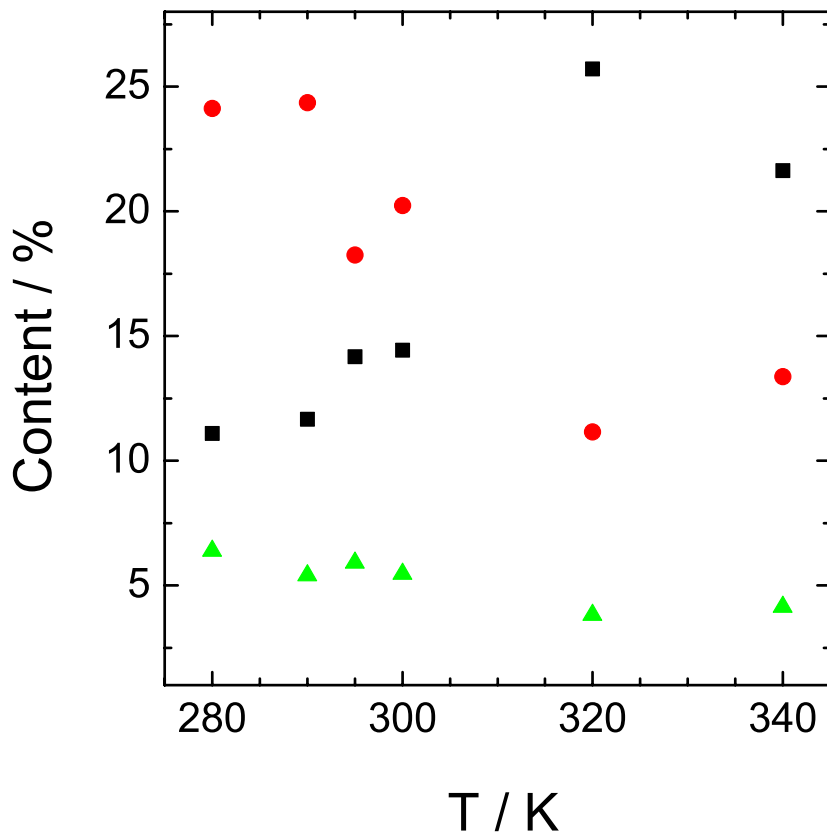


Figure 4.21: Dependence of the content of various structural elements on temperature for the ELP chain in the water solution of urea. Black squares: helices. Red circles: PPII. Green triangles: β -sheet.

ELP chain is shown in figure (4.21). As in the pure water case, the overall content of structural elements is less than 50 % and independent of temperature, which indicates the low ordered character of the ELP chain in the whole temperature range. The β -sheet content is small (less than 7 %) and also does not depend on temperature. The content of the residues in the PPII conformation is 25 % at 280 K, and is decreasing with increasing temperature till 10 - 12 %. The fraction of the amino acids in the helical conformation is 12 % at 280 K, and it grows upon heating till 25 %.

To check the presence of stable structures of the ELP chain based on identical repetitive units, the probability n_m to find m successive residues in the same conformation was calculated (figure (4.22)). This probability drastically decreases with increasing m in a strictly monotonic way. In the whole temperature range studied 50 - 60 % of all amino acids do not have neighbors in the same conformation ($m=1$), 10 to 12 % of residues make pairs ($m=2$), and only 2 - 3 % of all residues create triplets ($m=3$). This shows that the distribution of the structural elements along the ELP chain is close to random, and a pronounced structure is not present.

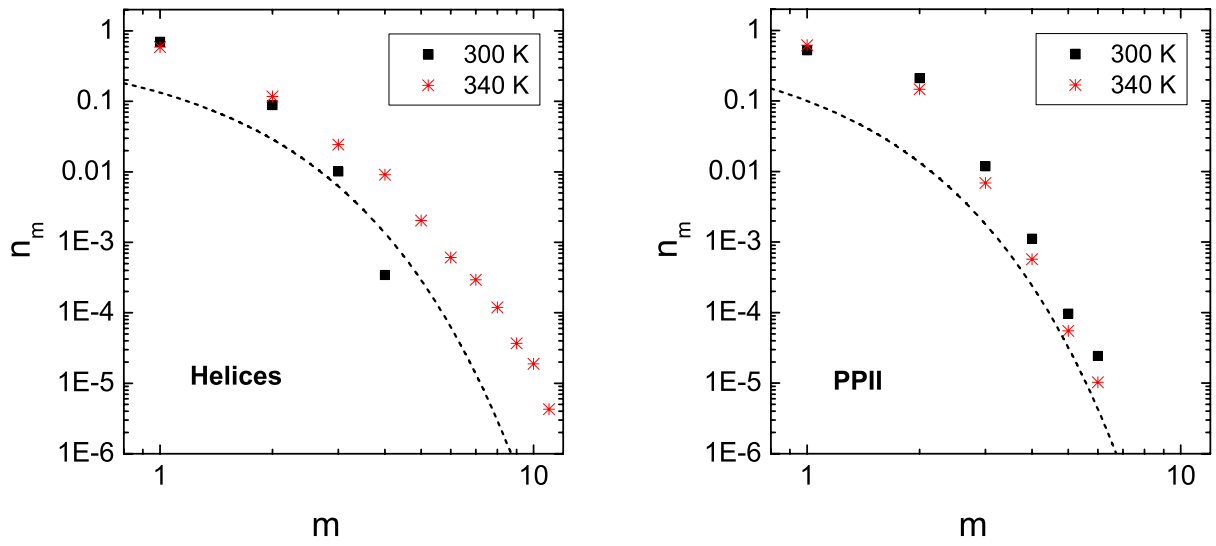


Figure 4.22: Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous urea solution. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in the hydrous urea solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.21).

The probability distributions n_m have the same shape like the *random* cluster distributions for an infinite one dimensional chain (eq. 3.42). This is an another evidence of the random character of appearance of the structural residues in the ELP chain. Deviations of the n_m distributions from the random distributions are probably caused by the small finite size of the ELP chain (see figure (4.22)).

4.2.2 ELP in the aqueous solution of 2,2,2-Trifluoroethanol (TFE).

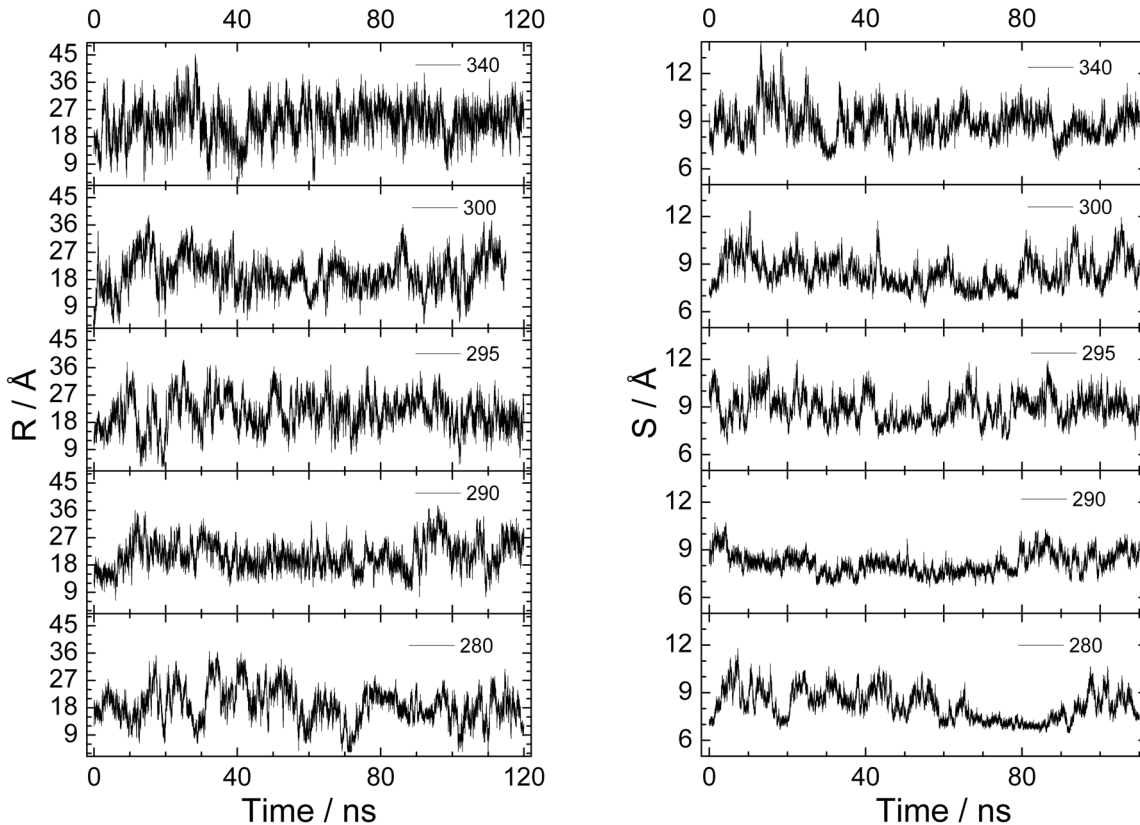


Figure 4.23: Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of TFE with time at different temperatures. From bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 340 K.

Despite the fact, that in experiments water and TFE are miscible at the concentration 4 mol/l, in our simulations the aqueous solution of TFE separates into a water rich and a TFE rich phase at the concentration 4 mol/l in the studied temperature range (see figure (A.1)) in Appendix). The interface between the water and the TFE phase is not perfectly formed, due to the small number of TFE molecules (see table (3.1)). The ELP is preferentially immersed in the TFE rich phase and only 45 % of the surface of the peptide is covered by water. The variation of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of TFE with time are shown in figure (4.23). Again two different states are observed: a *compact* ($T=280$ K, time interval from $70 - 80$ ns, $S_{av}=7.12\pm 0.2\text{\AA}$, $\delta_S=2.8$ %) and an *extended* ($T=280$ K, time interval from $30 - 40$ ns, $S_{av}=8.80\pm 0.62\text{\AA}$, $\delta_S=7$ %) one, but the difference between them are not so drastic like in the **urea** case. The fraction of the *compact* state decreases upon heating, but even at 340 K it still exists. At 280 K it is smaller than in the

Temp / K	R_{av} / Å	Sd / Å	δ_R / %	S_{av} / Å	Sd / Å	δ_S / %	L_{av} / Å	Sd / Å	δ_L / %
280	18.8	5.74	31	8.22	0.956	12	26.2	3.72	14
290	21.1	4.77	23	8.16	0.659	8	27.1	3.08	11
295	21.1	5.58	27	8.95	0.865	10	28.6	3.31	12
300	19.8	5.76	29	8.48	0.933	11	27.1	3.64	13
340	23.0	5.88	26	8.86	0.964	11	29.2	3.68	13

Table 4.3: Average values, standard deviations and relative fluctuation of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous TFE solution at different temperatures.

aqueous solution of urea or in pure water, i.e in the TFE environment the ELP prefers to be in the *extended* state.

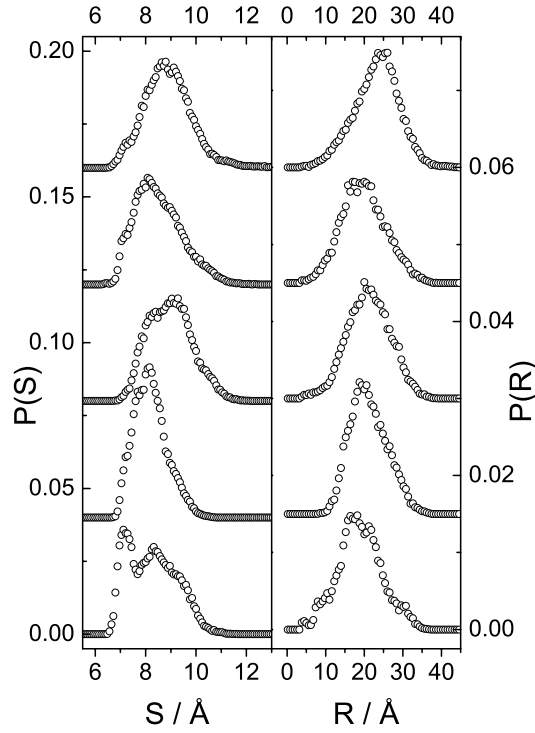


Figure 4.24: Left panel: Probability distributions of the radius of gyration S of the ELP in the aqueous solution of TFE at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the aqueous solution of TFE at different temperatures (from bottom to top: T = 280 K, 290 K, 295 K, 300 K, 340 K). Note also the different scales for the left and right panel and the shift of the distributions.

As it is seen from table (4.3) the average values of R , S and L of the ELP in the aqueous solution of TFE R_{av} , S_{av} , L_{av} are independent of temperature and in the whole temperature

range greater than average values for the ELP in pure water at the temperature 440 K (see table (A.1) in the Appendix). This indicates that the TFE environment promotes the *extended* conformation of the ELP. As it was already mentioned, the relative fluctuations δ , defined by the equation (3.39) is a good measure for the *flexibility* of the ELP chain. Again like in the urea case we can not take δ which is averaged over the hole time interval, because different, namely *compact* and *extended*, states are present in the whole temperature range studied (see also figure (4.24)). So we have calculated δ for the *extended* conformational state at T=340 K in the time interval from 80 - 90 ns. The values $\delta_S=9.7\%$ ($S_{av}=8.79\pm 0.85\text{\AA}$) and $\delta_R=17\%$ ($R_{av}=25.07\pm 4.36\text{\AA}$) are, accordingly, 44 % and 129 % smaller than the corresponding average values for the ELP in pure water in the temperature range from 320 K to 440 K (see table in Appendix). Based on these observations, we may conclude that in the TFE surroundings ELP is both *extended* and *rigid*. To make this conclusion more clear, the shapes of the probability distributions of the end-to-end distance R , the radius of gyration S and the maximal extension L were analyzed.

The probability distributions of R and S are shown in figure (4.24). The equations (3.40) and (3.41) failed to fit the shape of the probability distributions of R , S and L in the whole temperature range. This indicates the *non random* character of the ELP chain in the TFE medium.

Temp / K	$N_{w,av}^{TFE}$	$N_{w,av}^{Pure}$	$N_{w,av}^{TFE} / N_{w,av}^{Pure} / \%$
280	61.3	132.4	46
290	61.4	134.2	46
295	64.2	136.7	47
300	65.7	134.7	49
320	–	129.3	–
340	56.5	129.8	44

Table 4.4: Average numbers of water molecules in the hydration shell of the ELP in pure water and in the aqueous solution of TFE and relative hydration level of the ELP in the TFE solution.

The average number of water molecules in the hydration shell of the ELP in the aqueous solution of TFE and the relative hydration level (definition see in section 4.2.1) are shown in table (4.4). As it is clearly seen from table (4.4) only 46 % of the peptide surface is covered by water. One should consider here that the solvent accessible surface of the ELP in TFE

solution is even larger than in pure water, as indicated by increased radius of gyration. As it was already mentioned, water molecules can play the role of a plasticizing medium for the ELP chain. Thus further reducing the relative hydration level of the ELP chain makes it even stiffer, than in the aqueous solution of urea.

To characterize the conformation of the ELP in more details, the probability distributions of the intra-molecular H-bonds between different residues in the chain were calculated (see figure (4.25)). All intra-molecular H-bonds are situated on the second, third and fourth upper minor diagonals at all temperatures studied (interpretation of the HB-Matrix see in section 3.4). H-bonds with highest existence probability (around 50 %) are mainly distributed on the third

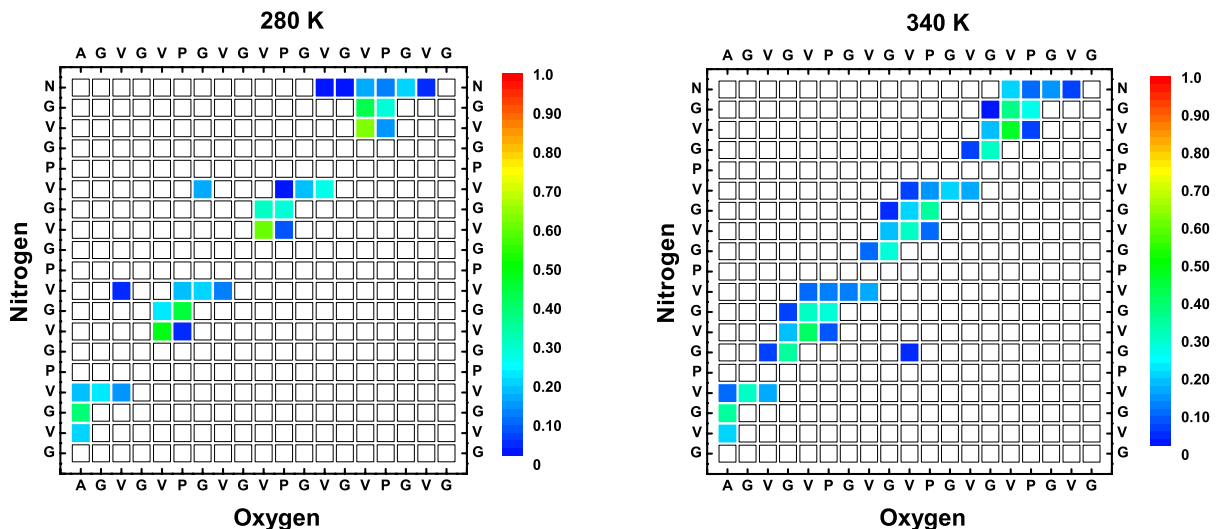


Figure 4.25: Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous TFE solution at 280 K. Right panel: same at 340 K.

minor diagonal ($\Delta i = 3$), this is typical for 3_{10} -helices. So the absence of *irregular* H-bonds and the high content of residues in the 3_{10} -helical conformation, can explain the large fraction of *extended* conformation of the ELP chain in the TFE surroundings. On the other hand, the H-bonds are quite stable and transformations between conformations with different Δi are rare, this observation can clarify the *rigid* character of the ELP chain in the aqueous TFE solution.

The temperature dependence of the fraction of residues in the different types of secondary structure is shown in figure (4.26). The overall content of the amino acids in a defined secondary structure, is less than 50 %, which indicates that the ELP chain does not have any pronounced secondary structure. The major fraction (approximately 25 %) of the structured residues is in the helical conformation at 280 K, and their content is growing upon heating. Only 10 to

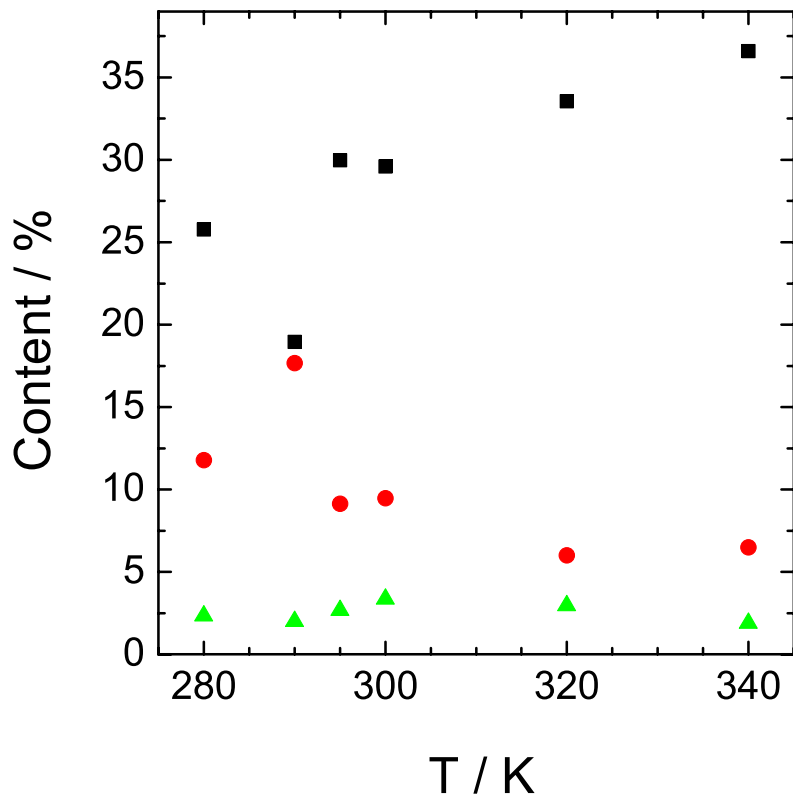


Figure 4.26: Dependence of the content of residues in the secondary structure on temperature for the ELP chain in the aqueous TFE solution . Black squares: helices. Red circles: PPII. Green triangles: β -sheet.

12 % of the structured residues are in the PPII conformation, and the number of these residues decreases with increasing temperature till 6 to 7 %. The content of amino acids in the β -sheet conformation is less than 5 % and is independent of temperature. Thus a transition of the single ELP chain in the TFE surroundings from a less ordered to a more ordered state is not observed in our simulations studies.

This is also supported by the probability distribution n_m to find m consecutive residues in the same conformation, shown for two different temperatures in figure (4.27). This probability decreases with increasing m in strictly monotonic way. The probability to find single residues without neighbors in the same secondary structure ($m=1$) is equal to 60 - 70 %. The probability to find two successive residues in the same secondary structure ($m=2$) is approximately 10 to 20 %. Only 1 to 2 % of residues in the same conformation form triplets ($m=3$). This indicates that there is no pronounced secondary structure, but structural elements are placed randomly along the ELP chain. The form of the probability distributions n_m repeat the form of the *random* cluster distribution in an infinite one dimensional chain (eq. (3.42)). This is an another clear

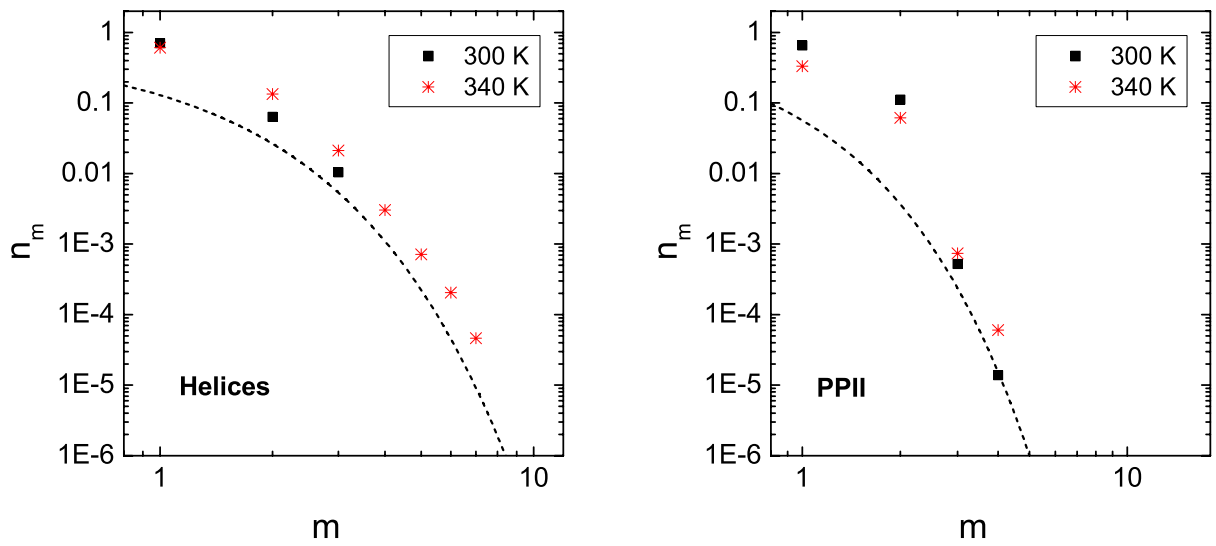


Figure 4.27: Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous TFE solution. Right panel: Probability n_m to find m successive residues in the PPII conformation for the ELP chain in the aqueous TFE solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K , given in figure (4.26).

indication of the absence of the pronounced secondary structure in the ELP chain. The shift of the probabilities n_m to higher values is due to the small finite size of the ELP chain (see figure (4.27)).

4.2.3 ELP in the aqueous solution of Sodium Chloride (NaCl).

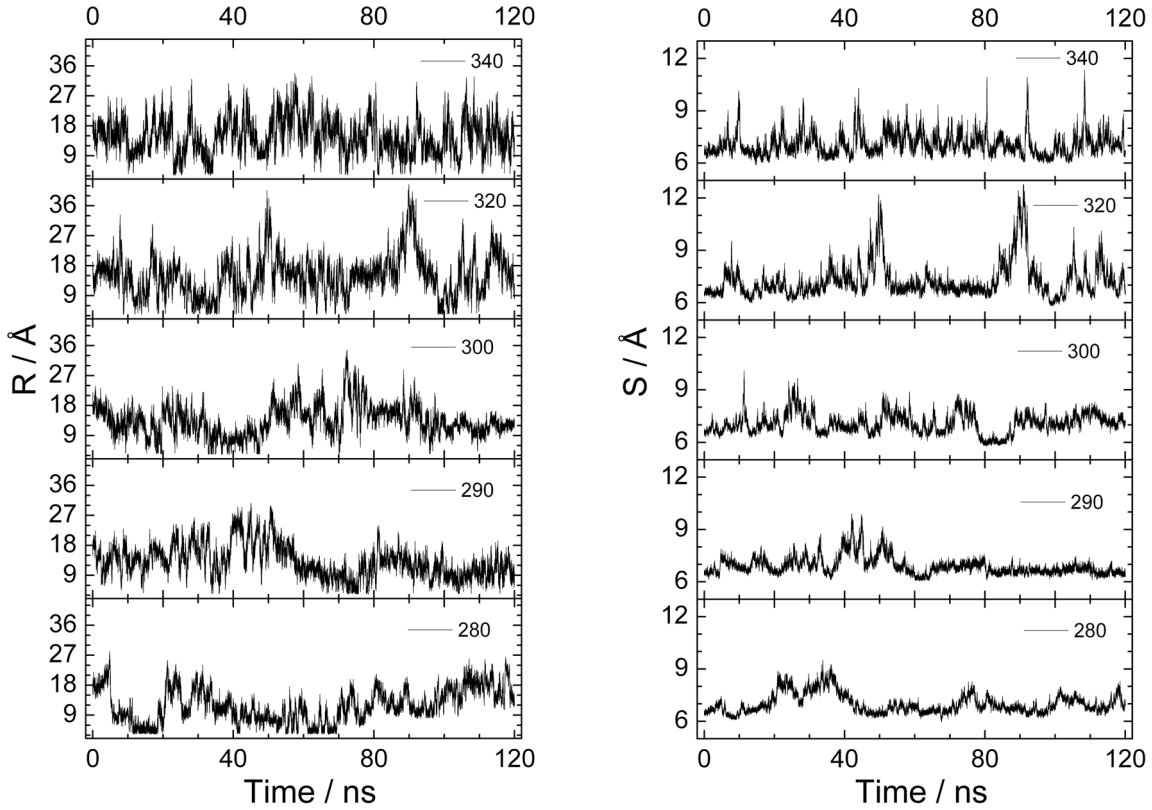


Figure 4.28: Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of NaCl with time at different temperatures. From bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K.

The time evolutions of R and S are shown in figure (4.28). The *compact* state of the ELP chain in the aqueous solution of the NaCl, which is observed at $T=280$ K in the time interval 80 - 90 ns, where $S_{av}=6.61\pm 0.21\text{\AA}$, $\delta_S=2\%$, is dominant in the whole temperature range studied. Transitions to an *extended* state, which is present at $T=280$ K in the time interval from 30 - 40 ns, where $S_{av}=8.1\pm 0.43$ and $\delta_S=5\%$, are observed very rarely and the residential time of the ELP chain in the *extended* state is very small.

The average values R_{av} , S_{av} and L_{av} , their standard deviations and relative fluctuations are shown in table (4.5). These values are slightly varying but do not show any clear trend to increase or to decrease. The average values of R , S and L are close to the corresponding values for the ELP chain in pure water at $T=280$ K ($R_{av}^{280} = 12.56\text{\AA}$, $S_{av}^{280} = 6.68\text{\AA}$, $L_{av}^{280} = 22.1\text{\AA}$). The value of the relative fluctuations δ_S of S , defined by the equation (3.39), which are observed at $T=320$ K in the *compact* state, in the time interval from 70 - 80 ns are equal to 2%. This is $\approx 50\%$ smaller than the value of the relative fluctuations of the S for the ELP

in pure water in the *rigid* state at $T=280$ K (see section 4.1.1). As far as the *compact* state is dominant in the whole temperature range, this indicates that the ELP chain is even more *rigid* in the aqueous solution of NaCl than in pure water at low temperatures. The number of Na^+ and Cl^- ions, for which the distance between the center of C_α of the ELP and the center

Temp / K	R_{av} / Å	Sd / Å	δ_R / %	S_{av} / Å	Sd / Å	δ_S / %	L_{av} / Å	Sd / Å	δ_L / %
280	11.9	5.02	42	7.00	0.557	8	23.3	2.42	10
290	13.2	4.90	37	6.90	0.520	8	22.3	2.61	12
295	13.5	4.23	31	6.92	0.520	8	22.4	2.51	11
300	13.4	4.58	34	7.08	0.567	8	23.2	2.42	10
320	15.0	6.12	41	7.25	1.00	14	23.5	3.77	16
340	14.8	5.24	35	7.08	0.635	9	23.0	2.75	12

Table 4.5: Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous NaCl solution at different temperatures.

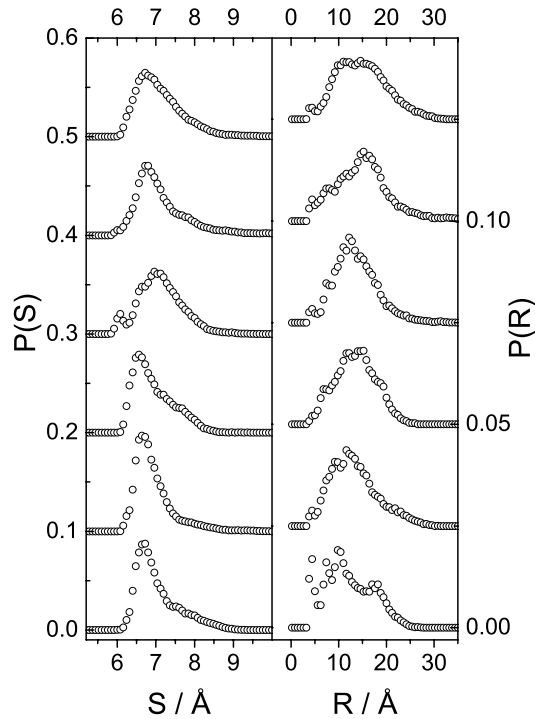


Figure 4.29: Left panel: Probability distributions of the radius of gyration S of the ELP in the water solution of NaCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the water solution of NaCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K). Note also the different scales for the left and right panel and the upward shift of the distributions.

of the ion is less than 0.55 nm, was calculated. This number is approximately equal to 3 and is constant in the whole temperature range studied. These ions strongly interact with N-H and C=O groups, and thus make the ELP chain very *rigid*. It seems that 3 such bridging ions are enough to keep the ELP chain in the *rigid* state.

The ELP chain exhibits a random distributions of the end-to-end distance R , radius of gyration S and maximal extension L only when it is *flexible* (see section 4.1.1). So the shape of these distributions is characteristic for the *rigidity/flexibility* of a single ELP chain. The probability distributions for the R and S at different temperatures are shown in figure (4.29). The inability of equations (3.40) (random Gaussian chain) or (3.41) (semi-flexible worm-like chain) to describe the shape of theses probability distributions, caused by the appearance of the multiple peaks, supports the previous conclusion about the *rigid* character of the ELP chain in the aqueous solution of the NaCl.

Temp / K	$N_{w,av}^{NaCl}$	$N_{w,av}^{Pure}$	$N_{w,av}^{NaCl} / N_{w,av}^{Pure} / \%$
280	135.5	132.4	102
290	131.3	134.2	98
295	131.0	136.7	96
300	130.9	134.7	97
320	130.5	129.3	101
340	124.3	129.8	96

Table 4.6: Average numbers of water molecules in the hydration shell of the ELP in the aqueous solution of NaCl and in the pure water, and relative hydration level of the ELP in NaCl solution.

The average number of the water molecules in the hydration shell of the ELP in the aqueous solution of NaCl and the relative hydration level (definition see in section 4.2.1) are shown in table (4.6). As it is clearly seen from the table (4.6) addition of sodium chloride does not affected the relative hydration level of the ELP. Thus we have enough water molecules to make the ELP chain *flexible*. This evidence indicates that the addition of NaCl makes the ELP chain *rigid* in a different way, in comparison with urea and TFE, namely by direct interactions of ions with the peptide.

Distributions of the probability to find an intra-molecular hydrogen bond between the i -th and $(i+\Delta i)$ -th residues are shown in figure (4.30). A large amount of *irregular* hydrogen bonds is present at 280 K, and even at 340 K *irregular* H-bonds are still observed, this indicates that

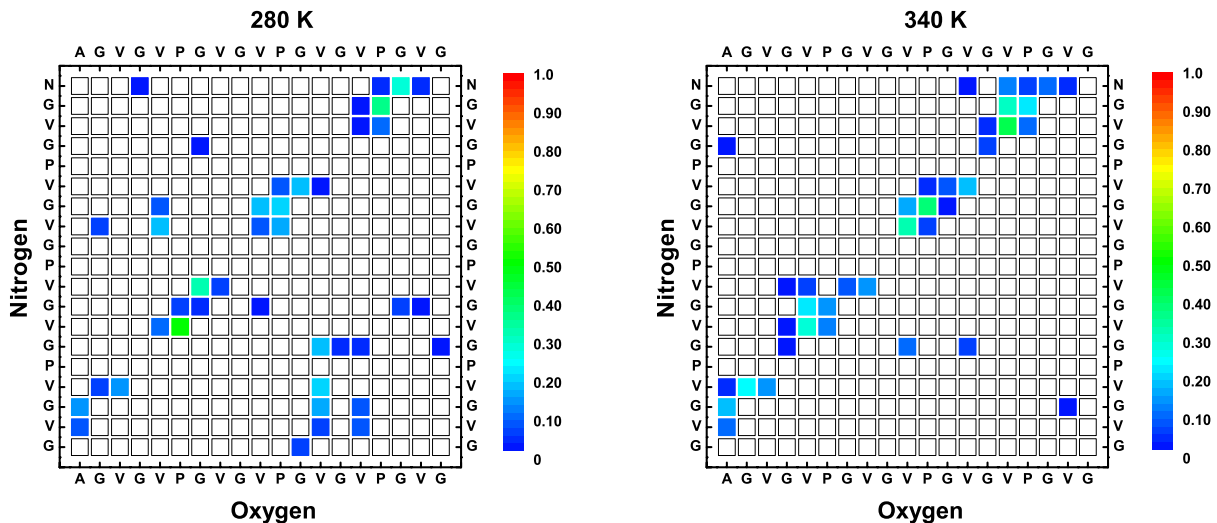


Figure 4.30: Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous NaCl solution at 280 K. Right panel: same at 340 K.

the ELP chain is *compact* and very *rigid*, because large amounts of *irregular* intra-molecular links, can prevent large fluctuations of the chain, and thus promote a *rigid* state of the ELP molecule. This view is supported by the temperature dependence of the fractions of *regular* and *irregular* intra-molecular hydrogen bonds (see figure (4.31)). The fraction of *irregular*

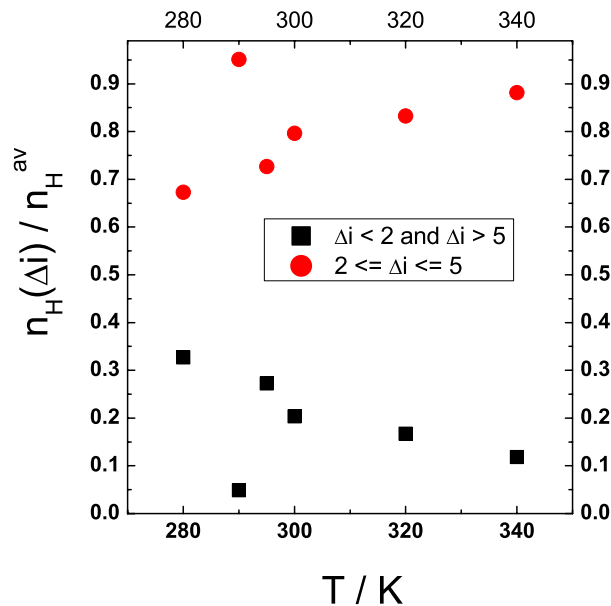


Figure 4.31: Temperature dependence of the fraction of the intra-molecular H-bonds of the ELP chain in the aqueous solution of NaCl with $2 \leq \Delta i \leq 5$ (red circles) and with $\Delta i < 2$ and $\Delta i > 5$ (black squares).

intra-molecular hydrogen bonds is approximately 32 % at 280 K and decreases with increasing temperature, but at 340 K this fraction is still 13 to 15 %, whereas in the pure water case in the

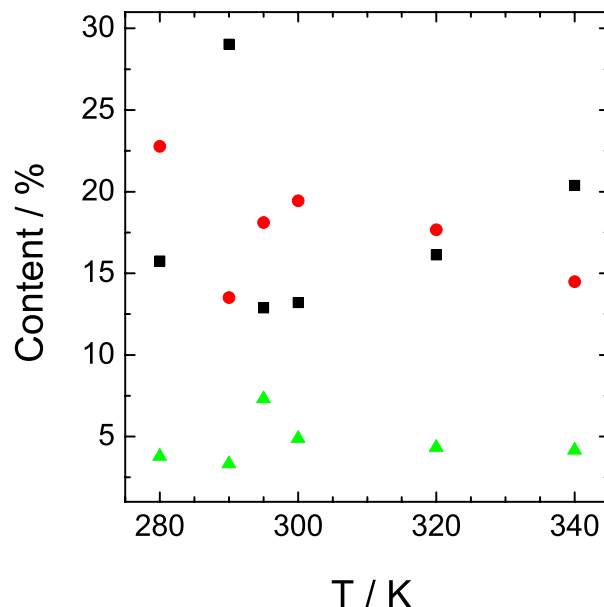


Figure 4.32: Dependence of the content of residues in the secondary structure on temperature for the ELP chain in the aqueous solution of NaCl. Black squares: helices. Red circles: ppII. Green triangles: β -sheet.

high temperature range the fraction of *irregular* H-bonds is less than 5 %. Consequently the ELP molecule in the aqueous NaCl solution has 2 to 3 times more *irregular* hydrogen bonds at 340 K in comparison with the ELP chain in pure water at high temperatures. One should also take into account ions which can also form strong connections with the ELP backbone atoms and prevent its fluctuations. So large amounts of *irregular* H-bonds and surface ions can be responsible for the *rigid* character of the ELP in the aqueous solution of sodium chloride.

To shed light on the structure of the ELP chain the Ramachandran plot of the peptide backbone was analyzed again. The temperature dependence of the content of residues in the secondary structure is shown in figure (4.32). Like in previously described cases, the ELP do not have any pronounced structure in the aqueous solution of the NaCl. The overall content of residues with distinct secondary structure is less than 45 %. The content of residues in the *beta*-sheet structure fluctuates around 5 % in the studied temperature range. The fraction of amino acids in the helical conformation fluctuates around 17.5 % and shows a slight trend to increase upon heating. The content of residues in the PPII conformation fluctuates close to 20 % and shows a weak trend to decrease with increasing temperature.

The probability n_m to find m residues with the same structure in a row is shown in figure (4.33). This figure support the view that the ELP chain is unstructured in the temperature range studied. The probability n_m to find m consecutive residues reduces with increasing m in

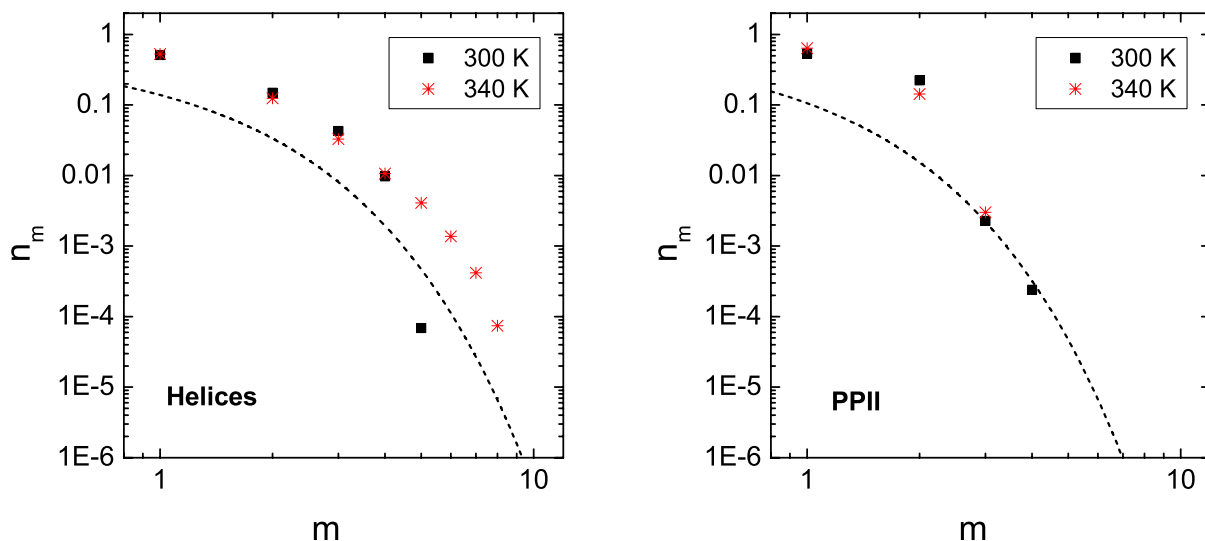


Figure 4.33: Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous NaCl solution. Right panel: Probability n_m to find m successive residues in the pPII conformation for the ELP chain in the aqueous NaCl solution. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.32).

a strictly monotonic way. Approximately 80 - 90 % of all structural elements do not have neighbors in the same secondary structure ($m=1$), the probability to find four successive residues ($m=4$) in the same conformation is less than 0.1 %. The probability distribution n_m closely follows the shape of the *random* cluster distribution in an infinite one dimensional chain, described by the equation (3.42). The shift of the probabilities n_m to higher values is due to the finite size of the ELP peptide (see figure (4.33)). Both these observations strongly support the previous claim about the absence of any pronounced structure of the ELP chain in an aqueous solution of sodium chloride. An ordering of the ELP chain upon heating is not detected in the aqueous solution of NaCl, only a slight redistribution of the structural elements is observed.

4.2.4 ELP in the aqueous solution of Guanidinium Chloride (GdmCl).

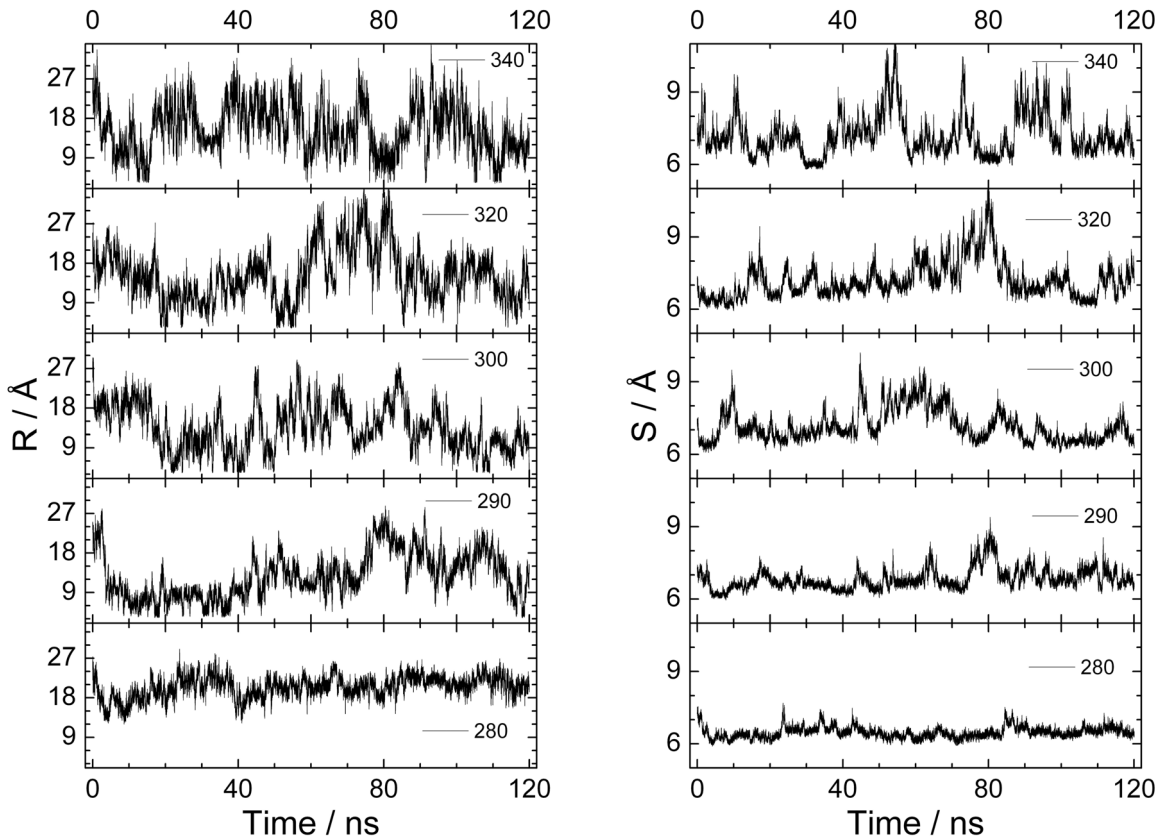


Figure 4.34: Variations of the end-to-end distance R and the radius of gyration S of the ELP in the aqueous solution of GdmCl with time at different temperatures. From bottom to top: $T = 280$ K, 290 K, 300 K, 320 K, 340 K.

The time dependences of the end-to-end distance R and radius of gyration S are shown in figure (4.34). Two different states of the ELP are found in the studied temperature range (both seen in the time dependence of S). A *compact* and *rigid* state of the ELP, which is observed at 280 K in the whole time interval with $S_{av}=6.47\pm 0.2\text{\AA}$ and $\delta_S=3\%$, dominates at temperatures from 280 K to 295 K, but its fraction drastically decreases at temperatures higher than 300 K. A small fraction of *extended* states appears already at 290 K in the time interval 80 to 83 ns with $S_{av}=8.27\pm 0.3\text{\AA}$ and $\delta_S=4\%$, but it becomes prevalent after the temperature passes 300 K.

The average value of the radius of gyration S slightly increases with temperature (see table (4.7)), the average values of the end-to-end distance R and maximal extension L remain almost constant in the whole temperature range studied. The relative fluctuations of S , defined by equation (3.39), of the ELP chain in the *compact* and *rigid* state in aqueous GdmCl solution

Temp / K	R_{av} / Å	Sd / Å	δ_R / %	S_{av} / Å	Sd / Å	δ_S / %	L_{av} / Å	Sd / Å	δ_L / %
280	20.4	2.45	12	6.47	0.200	3	23.1	1.68	7
290	12.8	5.22	41	6.83	0.440	6	22.2	1.97	9
295	16.5	3.86	23	6.47	0.350	5	21.9	2.30	10
300	13.5	5.19	39	7.16	0.660	9	22.8	2.68	12
320	15.3	5.90	39	7.19	0.757	11	23.2	3.39	15
340	15.6	5.57	36	7.21	0.863	12	23.5	3.36	14

Table 4.7: Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in the aqueous GdmCl solution at different temperatures.

are comparable with the relative fluctuations of the same parameter of the ELP chain in the pure water in the *rigid* state (see section 4.1.1). Thus at low temperatures, *rigidity* of the ELP chain in the aqueous solution of the guanidinium chloride is comparable with its *rigidity* in pure water. The relative fluctuations δ_S of S of the ELP chain in the *extended* state in the

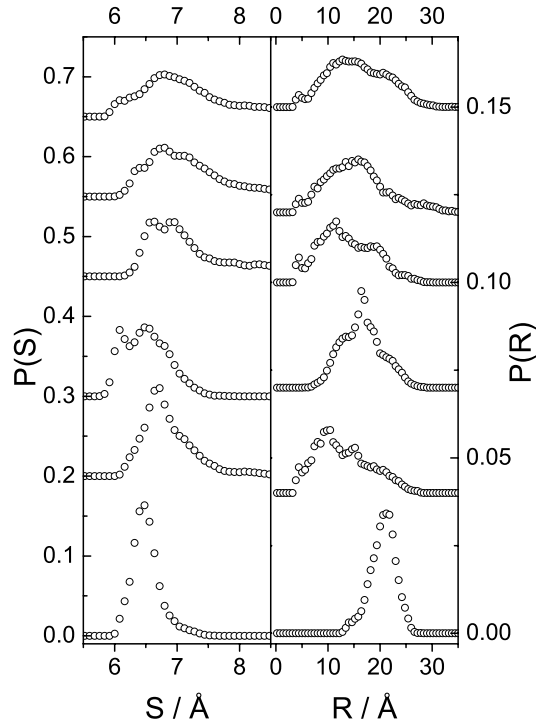


Figure 4.35: Left panel: Probability distributions of the radius of gyration S of the ELP in the water solution of GdmCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 340 K). Right panel: Probability distributions of the end-to-end distance R of the ELP in the hydrous solution of GdmCl at different temperatures (from bottom to top: $T = 280$ K, 290 K, 295 K, 300 K, 340 K). Note also the different scales for the left and right panel and the upward shift of the distributions.

aqueous solution of GdmCl at T=340 K in the time interval from 90 to 95 ns are equal to 7 %. They are two times smaller than the relative fluctuations of S of the ELP chain in pure water at high temperatures (see table A.1 Appendix). This indicates that the ELP in the aqueous solution of GdmCl remains rigid even at high temperatures. The reason for such behavior of ELP can be the same as in the case of NaCl: the ions which are close to the surface can interact with the backbone atoms of the ELP chain and thus make it highly *rigid*. To test this hypothesis the number of Gdm^+ and Cl^- ions which are closer than 0.55 nm to the C_α atoms of the peptide backbone was calculated. It is approximately equal to 3 and remains nearly constant in the whole temperature range studied. It seems that ions could be responsible for the *rigid* character of the ELP chain in the aqueous GdmCl solution.

Since the probability distributions for R and S of the ELP chain in the *flexible* state should have random (Gaussian) form, the non-Gaussian shape of these probability distributions will be another evidence of the *rigid* character of the ELP chain. It was not possible to describe the shape of the probability distributions $P(R)$ and $P(S)$ with the equation for the random Gaussian chain (3.40). The equation for the semi-flexible worm-like chain (3.41) also failed to describe the shape of these probability distributions.

Temp / K	$N_{w,av}^{GdmCl}$	$N_{w,av}^{Pure}$	$N_{w,av}^{GdmCl} / N_{w,av}^{Pure} / \%$
280	105.8	132.4	80
290	111.7	134.2	83
295	104.5	136.7	76
300	112.7	134.7	84
320	110.2	129.3	85
340	106.7	129.8	82

Table 4.8: Average numbers of water molecules in the hydration shell of the ELP in the aqueous solution of GdmCl and in pure water, and relative hydration level of the ELP in GdmCl solution.

The average number of water molecules in the hydration shell of the ELP in the aqueous solution of GdmCl and relative hydration level(definition see in section 4.2.1) are shown in table (4.8). It is clearly seen from table (4.8) that the addition of guanidinium chloride reduced the relative hydration level of the ELP to 80 %. Based on this observation we can state, that GdmCl affect the *flexibility* of the ELP chain in two ways: firstly reducing relative hydration of the peptide, secondly by direct interaction of ions with the ELP chain.

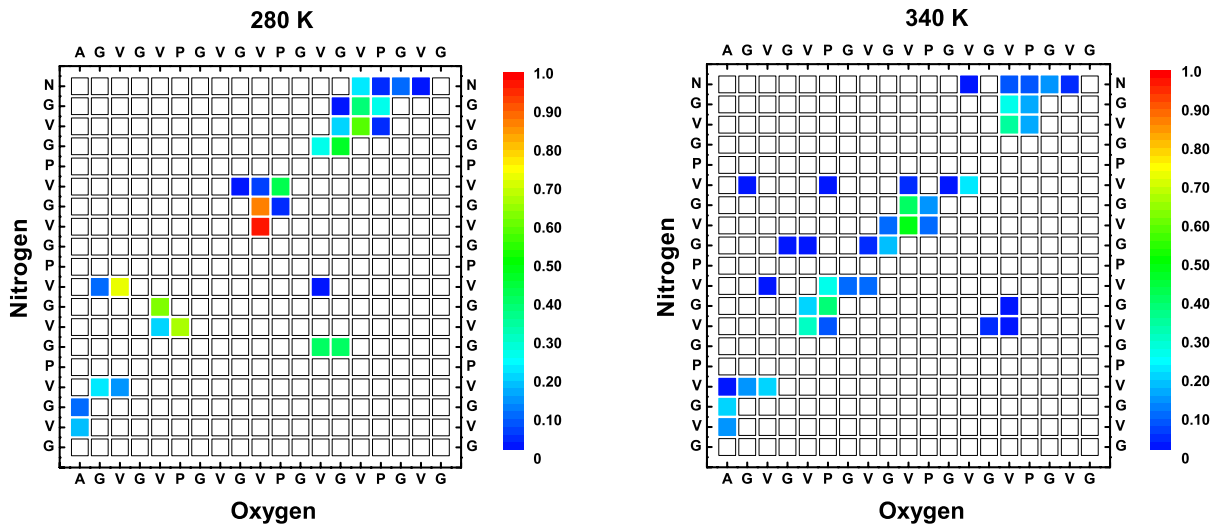


Figure 4.36: Left panel: Probability to find intra-molecular hydrogen bond between different amino acids for the ELP in the aqueous GdmCl solution at 280 K. Right panel: same at 340 K.

The probability distributions of the intra-molecular hydrogen bonds between different residues in the ELP chain were calculated to characterize the conformation of the peptide more precisely. Several *irregular* intra-molecular H-bonds with existence probability more than 50 % can be found at $T = 280$ K (see figure (4.36)). Two *regular* intra-molecular H-bonds with

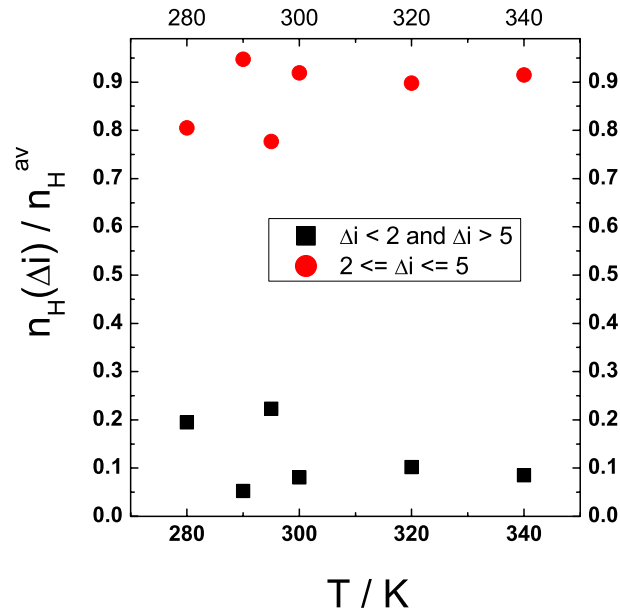


Figure 4.37: Temperature dependence of the fraction of the intra-molecular H-bonds of the ELP chain in the aqueous solution of GdmCl with $2 \leq \Delta i \leq 5$ (red circles) and with $\Delta i < 2$ and $\Delta i > 5$ (black squares).

existence probability more than 90 % are observed at the same temperature. This indicates that ELP in the aqueous solution of the GdmCl is *compact* and *rigid* at low temperatures.

Irregular H-bonds still exist at $T = 340$ K, but they are less stable in comparison with H-bonds

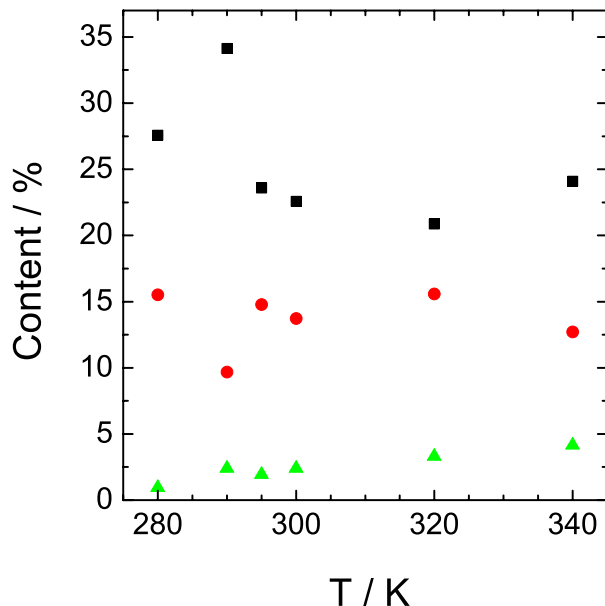


Figure 4.38: Dependence of the content of residues with secondary structure elements on temperature for the ELP chain in the hydrous GdmCl solution . Black squares: helices. Red circles: PPII. Green triangles: β -sheet.

at 280 K.

The fraction of *irregular* intra-molecular hydrogen bonds decreases with temperature from 20 % at 280 K to 10 % at 340 K (see figure (4.37)). The fraction of *regular* H-bonds increases upon heating from 80 to 90 %, but these hydrogen bonds also loose their stability. The less stable character of the intra-molecular hydrogen bonds at high temperatures can explain the slightly bigger fluctuations of the ELP chain. The high temperature fraction of *irregular* H-bonds of the ELP molecule in the aqueous GdmCl solution is 1.5-2 times higher than the high temperature fraction of *irregular* H-bonds of the ELP in the pure water. Together with the influence of ions on the ELP structure, this can explain why the ELP chain in the aqueous GdmCl solution is *rigid* even at high temperatures.

The first hint on the presence of a secondary structure is the content of structural elements in the chain. The temperature dependence of the fraction of the residues in the secondary structure is shown in figure (4.38). Like in all previous cases, none of the structural elements have content more than 50 %, which means that the peptide chain does not have any pronounced secondary structure. The content of amino acids in the helical structure fluctuates around 25 % and does not show any clear trend. The content of residues in the PPII conformation is

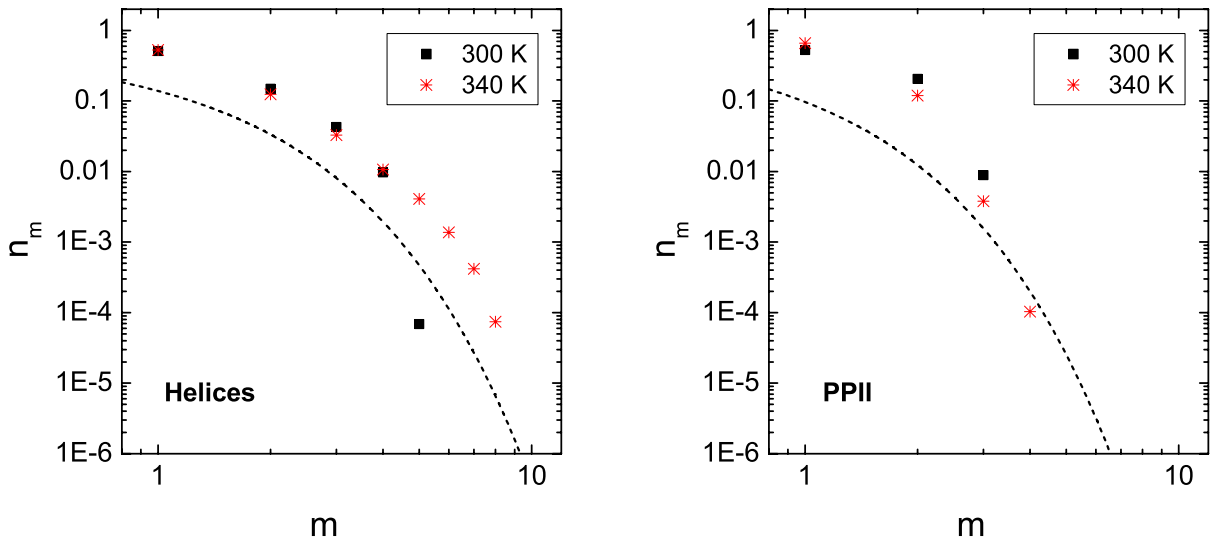


Figure 4.39: Left panel: Probability n_m to find m successive residues in the helical conformation for the ELP chain in the aqueous GdmCl solution. Right panel: Probability n_m to find m successive residues in the ppII conformation for the ELP chain in the water solution of GdmCl. Dashed lines: probability distributions expected for the random formation of the cluster of size m in an infinite one dimensional chain, where the fraction of occupied sites θ is equal to the content of structural elements in the ELP chain at 340 K, given in figure (4.38).

approximately 15 % and almost independent of temperature.

A good characteristic of the presence of the secondary structure is the probability n_m to find m consecutive residues of the same secondary structure. The indication of a stable secondary structure will be a high probability n_m at large m ($m > 4$). For the ELP in the aqueous solution of GdmCl the probability n_m to find m consecutive residues with the same structure is lower than 0.1 % when $m > 4$ (see figure (4.39)). The probability n_m reaches its highest value at $m = 1$ (single structural element without neighbor of the same structure), and then monotonically decreases with increasing m . The low occurrence probability of the clusters with $m > 4$ evidences a random distribution of the residues in the secondary conformation along the chain of the ELP. The probability distributions n_m have the same shape as the *random* cluster distribution in an infinite one dimensional chain, described by the equation (3.42). The shift of the probabilities n_m to higher values is due to finite size of the ELP peptide (see figure (4.39)). This is another confirmation of the absence of any pronounced secondary structure in the ELP molecule.

Summarizing, the addition of cosolvents make the ELP chain more rigid. We have shown that some cosolvents remove water from the first hydration shell of the ELP. Intra-molecular

hydrogen-bonded pattern have changed after the addition of cosolvents. We do not observe the appearance of a pronounced secondary structure with addition of these cosolvents. Please see section 5.2 for a detailed discussion of these results.

4.3 Thermal breaking of the spanning network of hydration water of the elastin-like peptide.

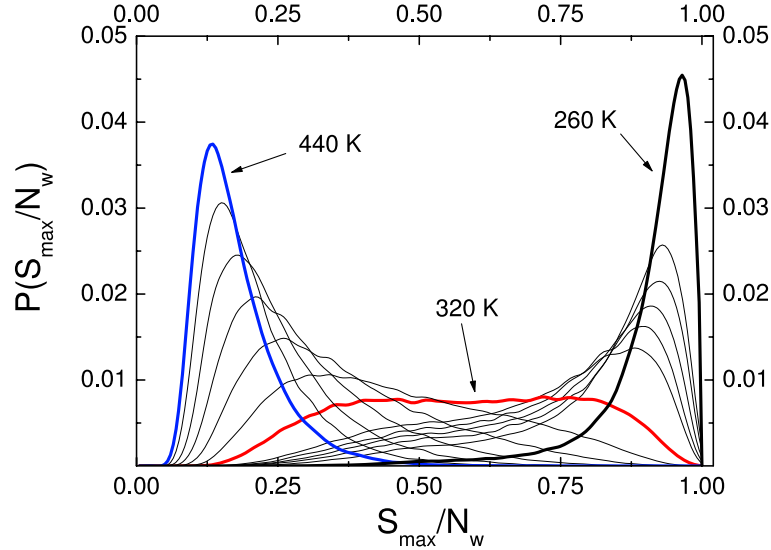


Figure 4.40: Probability distribution $P(S_{max})$ of the size S_{max} of the largest water cluster, normalized to the total number of water molecules N_w in the hydration shell of the ELP, at different temperatures. The shell width D is 4.5 Å.

Probability distributions $P(S_{max})$ of the size S_{max} of the largest hydrogen-bonded water cluster (for definitions, see section 3.5) normalized by the total number of the water molecules N_w in the hydration shell ($D=4.5$ Å) of the ELP molecule at various temperatures are shown in figure (4.40). The evolution of $P(S_{max})$ with decreasing temperature is quite similar to the one observed for hydration water at various surfaces with increasing hydration level [59, 60, 66, 67, 68, 69]. In general, the collection of probability distributions $P(S_{max})$ shows a two-peak structure with a left (small S) and right (large S) peak, corresponding to nonspanning and spanning largest clusters, respectively [59, 69]. At high temperatures, spanning clusters practically never appear and nonspanning clusters dominate (figure (4.40), $T=440$ K). With decreasing temperature, the distribution $P(S_{max})$ shifts to higher values of S , widens, and at $T=320$ K exhibits a two-peak structure with approximately equal contributions from spanning and nonspanning clusters. At low temperatures, the distribution $P(S_{max})$ becomes narrower and at $T=260$ K a spanning cluster exists almost permanently; in this case the largest water cluster in the hydration shell spans the whole molecule practically at any moment. A similar behavior of the distribution $P(S_{max})$ is also observed with increasing hydration shell width D ,

when the temperature is fixed.

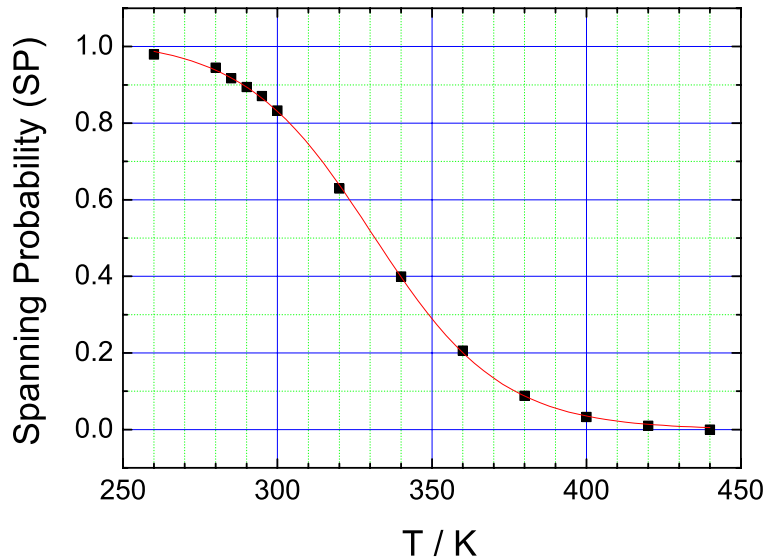


Figure 4.41: Existence probability (SP) of a spanning network in the hydration shell of the ELP.

The probability distributions $P(S_{max})$ give us the possibility to determine roughly the probability SP that a spanning cluster is present in the system (**spanning probability**). Assuming $S_{max}/N_w = 0.5$ as the minimal fraction of the water molecules, necessary for the largest cluster to be spanning (see figure (4.40)). By integrating the distributions $P(S_{max})$ from figure (4.40) for $S_{max}/N_w > 0.5$ we obtained SP as a function of temperature (see figure (4.41)), and thus could locate approximately the percolation threshold of water in the hydration shell of the ELP. In accordance with the studies of the percolation transition of water at planar surfaces [59, 69] for any system size a percolation transition occurs, when the spanning probability SP reaches values of about 95 %. The thresholds temperature obtained from the $SP=95$ % criterion is equal 280 K ($D=4.5$ Å). Due to the small size of the system (< 200 water molecules), the destruction of the spanning network of hydration water occurs in some interval of temperature or shell width. To describe this process, we determine additionally a lower characteristic point, corresponding to $SP=50$ %. A spanning cluster of hydration water of the ELP exists with probability 50 % at $T=330$ K ($D=4.5$ Å).

Distributions n_S of the size S of the water clusters in the hydration shell of the ELP for the temperature range from 260 K to 440 K are shown in figure (4.42). At the percolation threshold, the cluster size distribution follows a power law $n_S \sim S^{-\tau}$ in the widest range of cluster sizes. The exponent τ is about 2.05 for 2D systems and about 2.2 for three dimensional

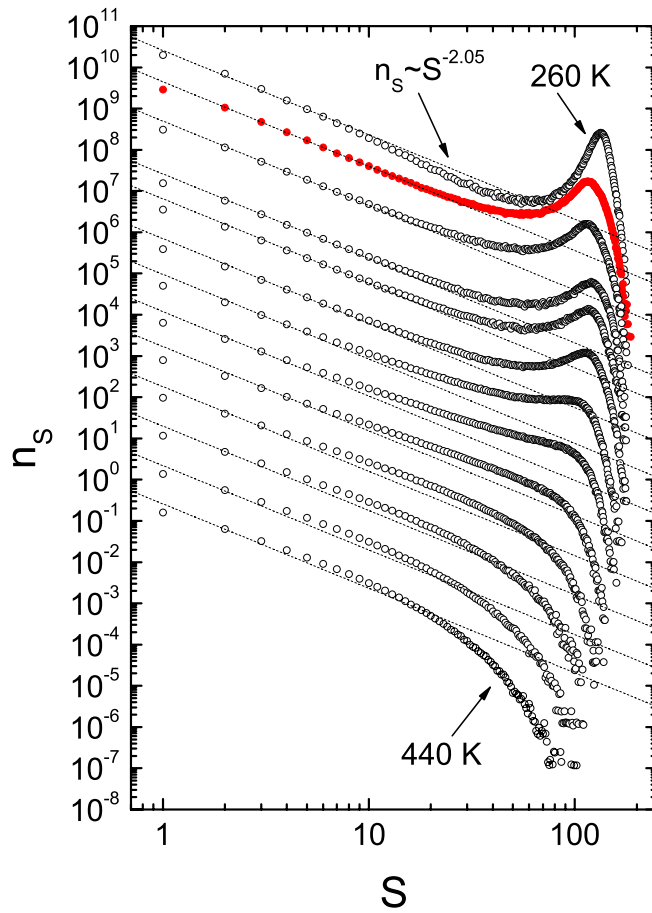


Figure 4.42: Distributions n_S of the size S of the water clusters in the hydration shell (shell width $D=4.5 \text{ \AA}$) of the ELP at various temperatures (from top to bottom: $T = 260 - 440 \text{ K}$). The distributions are shifted subsequently, starting from the bottom. The dashed lines show the power law $n_S \sim S^{-2.05}$, as expected at a 2D percolation threshold. By red color highlighted the distribution closest to the percolation threshold.

(3D) systems [58]. Thus, to locate the percolation threshold, we compare the distributions n_S with the power law $S^{-\tau}$, which is a straight line in the double logarithmic scale of figure (4.42). The distribution n_S closest to the percolation threshold is shown by red circles. Note that using the 3D value for the exponent τ does not affect the estimation of the percolation threshold noticeably. The threshold temperature, estimated based on the n_S distributions is equal 280 K. This is in a good agreement with the percolation threshold obtained from the distribution of the largest cluster size, using the 95 % criterion for the spanning probability.

The mean cluster size S_{mean} diverges at the percolation threshold in infinite systems and passes through a maximum, when approaching the threshold in finite systems [58]. The temperature dependence of S_{mean} in the hydration shell ($D=4.5 \text{ \AA}$) of the ELP is shown in figure (4.43). The mean cluster size S_{mean} passes through a maximum when approaching the true

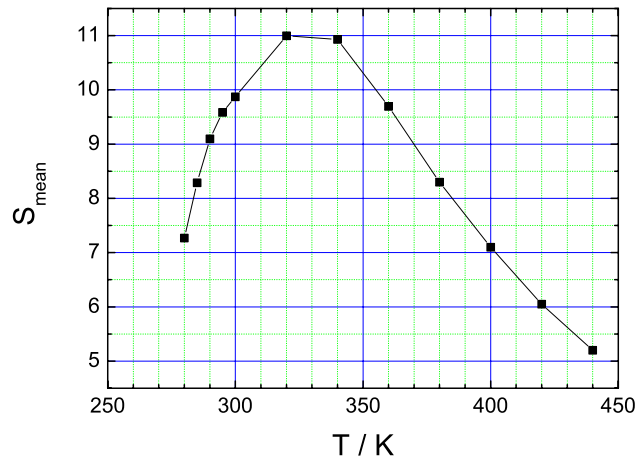


Figure 4.43: Dependence of the mean size S_{mean} of the water clusters in the hydration shell (shell width $D=4.5 \text{ \AA}$) of the elastin-like peptide on the temperature.

percolation threshold, as determined above from the 95 % criterion of SP and n_S . A similar spread of percolation thresholds has been observed in earlier studies of a small systems [66, 68].

Water clustering in various systems exhibits a highly universal behavior, when compared in terms of the average number n_H of water-water H-bonds per water molecule. In particular,

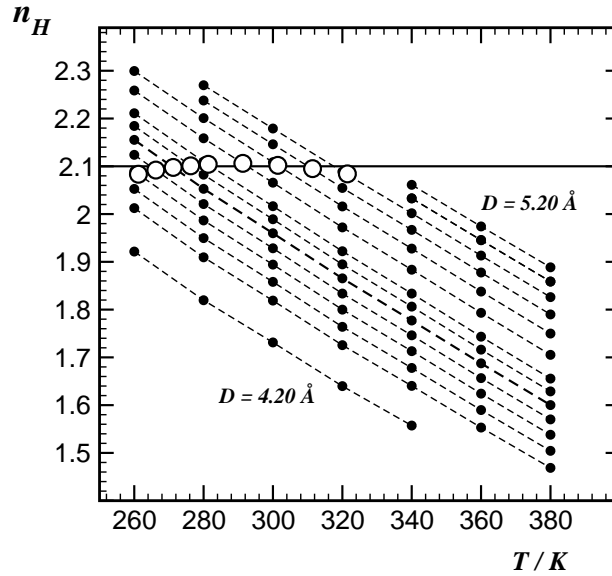


Figure 4.44: Temperature dependence of the average number n_H of water-water hydrogen bonds between water molecules in the hydration shell of the ELP for various choices of the shell width D , which increases from the bottom to the top. The open circles show the values of n_H at the true percolation thresholds ($SP=95 \%$), as obtained for different D values.

n_H is close to the value of 2.2 ± 0.1 at the percolation threshold in low-hydrated systems [66].

This value indicates that the correlated site-bond percolation of hydration water is similar to the random site or bond percolation in square and honeycomb 2D lattices with four and three neighbors, respectively [58]. The temperature dependence of the average number n_H of H-bonds (definition see in the section 3.5) between water molecules in the hydration shell of ELP for various choices of the shell width D is shown in the figure (4.44) by full circles and dashed lines. The value of n_H at the percolation thresholds, determined from the $SP \approx 95\%$ criterion and from the cluster size distribution n_S , is shown in figure (4.44) by open circles. At the percolation threshold, n_H is about 2.1 i.e., rather close to the threshold values of n_H from 2.0 to 2.3 for water percolation at smooth hydrophilic surfaces and in low-hydrated lysozyme systems [66, 67].

The distance H_{max} between the center of mass of the largest water cluster and the center of mass of the ELP as well as the radius of gyration S_w of the largest cluster of hydration water are topological parameters of the network, which directly reflect its spanning or nonspanning character. The radius of gyration S_w of the largest water cluster in the hydration shell of the

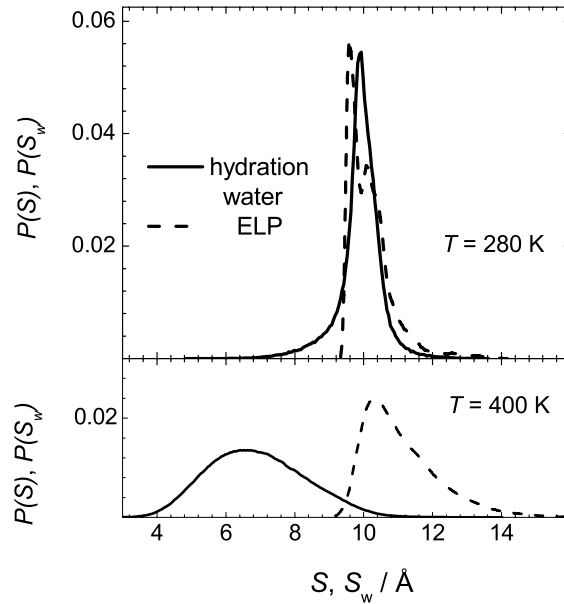


Figure 4.45: Probability distributions of the radius of gyration S of the ELP and of the radius of gyration S_w of the largest hydrogen-bonded cluster in the hydration shell of the ELP at two different temperatures. (The distribution of the S of the ELP is shifted by $+3.5\text{ \AA}$ at both temperatures (see text).)

ELP at low temperatures (see distribution in the upper panel of the figure (4.45)) exceeds the radius of gyration S of the ELP by about 3.5 \AA (which is about the width of a water layer), indicating a homogeneous coverage of the ELP by the spanning HB water network. The strong

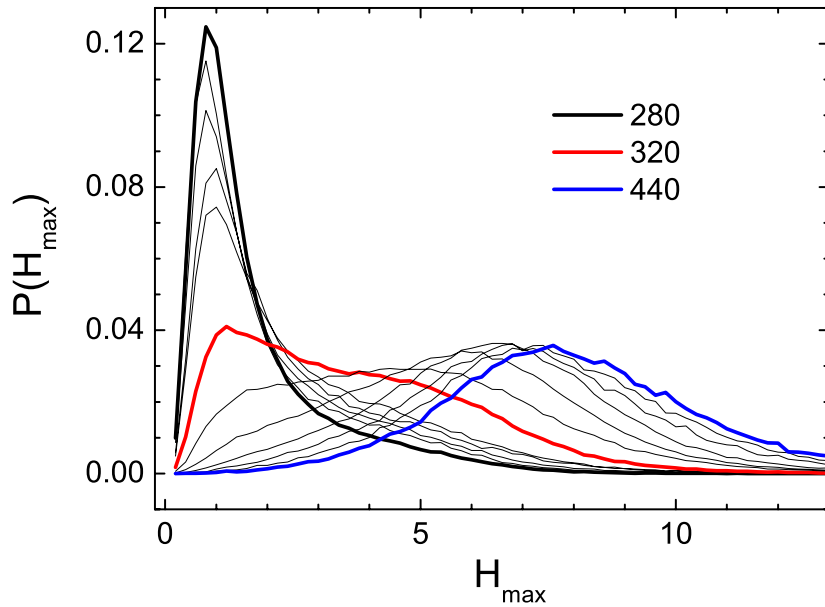


Figure 4.46: Probability distributions $P(H_{max})$ of the distance H_{max} between the centers of mass of the ELP and the largest water cluster in the ELP hydration shell at different temperatures.

difference between S and S_w at higher temperatures (see lower panel of figure (4.45)) indicates the absence of a spanning network of hydration water.

At $T=280$ K, the center of mass of the largest HB cluster of hydration water is very close to the center of mass of the ELP, and the distance H_{max} between these two centers is close to zero (see figure (4.46)). This clearly indicates the presence of a spanning HB water network, which **homogeneously envelopes** the ELP molecule. In contrast, at high temperatures the value of H_{max} is large, which indicates the nonspanning character of the largest water cluster in the hydration shell of the ELP.

We have shown that the transformation of the hydration water shell of the ELP from a spanning to non-spanning state occurs via a quasi-2D percolation transition. This transition happens in the same temperature interval as a *conformational* transition of the ELP chain. Please see detailed discussion in section 5.3.

Chapter 5

Discussion.

5.1 Temperature induced transition of the elastin-like peptide.

We distinguish the occurrence of two different conformational states of the hydrated model ELP GVG(VPGVG)₃ in the simulations. At high temperatures, the ELP is a highly flexible chain, which shows a random distribution of the end-to-end distance. Moreover, the parameter B_R in the exponent of the end-to-end distribution, equation (3.40), does not vary strongly with temperature, at least for high temperatures $T > 340$ K (see lower panel of figure (4.6)). This finding evidences the mainly entropic character of the elasticity of the ELP and supports an old idea of Hovee and Flory that the elasticity of elastin is rubber-like [14, 15]. Although the random distribution of the end-to-end distances suggests a random coil structure, the presence of various structural elements (figure (4.16)) and the essential number n_H^{av} (figure (4.9)) of intramolecular H-bonds between ELP residues (in the considered temperature range, n_H^{av} varies from about 5.5 to 6.5) seems to contradict this conclusion. However, an irregular (or even random) location of the ordered structural elements along the chain (see figure (4.17)) and/or a frequent interconversion between them may well provide a random distribution of the end-to-end distance of the chain. In general, an interconversion between the various structural elements of the chain is accompanied by a rearrangement of the intra-molecular H-bonds. Obviously, the presence of hydration water, which is a plasticizer for biomolecules, should strongly facilitate this rearrangement. This may explain why only hydrated ELPs show elastic properties [11].

The observed flexible conformational state of the ELP combines the presence of local order with a random (disordered) character of the full peptide chain. This picture is supported by the isotropic structure [33] and high mobility [34] of ELPs, seen in experiments. The presence of ordered structural elements in ELPs is clearly seen both in experiments [7, 36, 70] and in simulations (see ref. [12] and figure (4.16)). However, there is no evidence that these structures are distributed along the chain non-randomly. When the simulation starts from a highly ordered conformational state of the ELP (for example, a β -spiral [50] was used as initial state of the ELP in ref. [12] and in our studies), it vanishes quickly and the ELP becomes a disordered ("amorphous" [12]) chain. Finally, in our simulation studies, we have found an essentially random distribution of the ordered structural elements along the peptide chain (see figure (4.17)).

The temperature transition of the amino-acids proline and glycine from the β -turn II to the

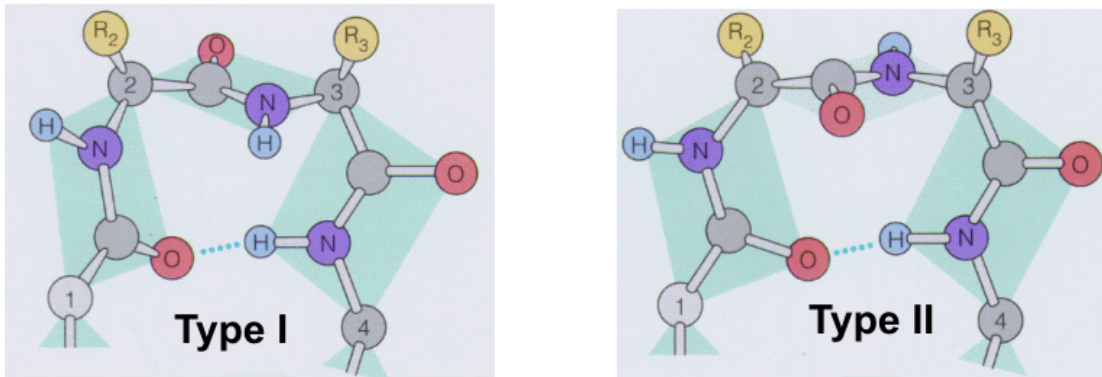


Figure 5.1: Left panel: β -turn type I. Right panel: β -turn type II.

β -turn I conformation, observed in our simulations (figures (4.12)-(4.15)), could be very important for the elasticity of the ELP chain. When these residues are in the β -turn I conformation, which actually can not be distinguished from the 3_{10} -helical conformation, all amide bonds are in one plane (see figure (5.1)), the chain can switch between different H-bond conformations, and thus can contract or extend. In contrast, when the amino-acids are in the β -turn II conformation, the orientation of the amide bond can hinder the switching of the intra-molecular H-bonds, and consequently prevent the interconversion between the local structures (see figure (5.1)).

Below $T_i \approx 310$ K another, more rigid, conformational state of the ELP, which does not show a random distribution of the end-to-end distance, appears, and its fraction drastically increases upon cooling. The low-temperature rigid state differs from the high-temperature

flexible conformational state of the ELP by the presence of a pattern of intra-molecular H-bonds between amino acids that are separated by more than 5 residues (called *irregular* H-bonds, see section 4.1.2 and figures (4.7), (4.9)). The number of such intra-molecular H-bonds drastically decreases upon heating above $T_t \approx 310$ K. The large amount and high stability of the *irregular* intra-molecular H-bonds at low temperatures, can be responsible for the narrow probability distribution of the end-to-end distance, as these H-bonds can tightly bind ends of the peptide and prevent exploration of the configurational space by the ELP chain. Unlike to the low temperature case, at temperatures higher than 310 K, *irregular* intra-molecular H-bonds are not playing an important role in the conformation of the peptide chain. The content of such H-bonds is very small and they are very unstable (see figures (4.8), (4.9)). The temperature behavior of the *regular* H-bonds is also very important for the conformation of the ELP chain. At temperatures lower than $T_t \approx 310$ K *regular* H-bonds mainly connect residues, which are further away ($4 \leq \Delta i \leq 5$) from each other (see figure (4.7)), so the ELP chain is almost all the time compact. Among them there exist H-bonds which are very stable (existence probability > 80 %, see figure (4.7)). In contrast to the low temperatures, in the high temperature range, the *regular* H-bonds are not so stable (existence probability < 30 %, see figure (4.8)), their Δi varies from 2 to 5, so the chain may extend and contract, and the H-bonds may switch more frequently (see figure (4.8)) and faster (see figure (4.10)) between different residues. These facts could help us to understand, why the elasticity of the single ELP chain is mainly entropic. During the transition from a compact to an extended conformation, *regular* H-bonds change their Δi from 5 to 2, but the number of hydrogen bonds remains unaltered, thus the *potential energy* of the chain remains constant. The *entropy* of the ELP chain in the *extended* state ($\Delta i=2$) is much smaller, than the *entropy* of the peptide chain in the *compact* state ($\Delta i=5$), consequently a reduction of the *entropy* by **stretching** the ELP chain leads to the appearance of a **restoring elastic force**. The entropy of the ELP chain at temperatures above 310 K is higher than the entropy of the peptide chain in the low temperature range. Due to the free interconversions of the *regular* H-bonds between conformations with different Δi in the high temperature range, the number of conformational states available for the ELP chain is considerably greater than at low temperatures, which gives the ELP molecule the possibility to have a Gaussian distribution of the end-to-end distance.

When crossing T_t (SP=95 %), the state of the hydration water shell also changes drastically: below T_t , most of the water molecules in the hydration shell of the ELP form a spanning

network of H-bonds; above T_t , the hydration water decomposes into small HB clusters [57]. The thermal breaking of the spanning network of the hydration water upon heating occurs via a quasi-2D percolation transition in the temperature range from 280 K to 330 K and may be considered as a transition of the hydration water from a more ordered state (large spanning network) to a less ordered state (ensemble of small clusters) (this will be discussed in more details in section 5.3). This qualitative change of the hydration water state may be responsible for the appearance of a flexible conformational state above $T_t=280$ K: contrary to the spanning water network, small water clusters in the hydration shell may not hamper interconversions between various local structures of the ELP, which seem to be important for the flexibility of the ELP chain [71].

Another very interesting transition can be localized when observing the average live time t_{av}^l of the intra-molecular hydrogen bonds. This parameter behaves differently in the low and high temperature ranges. Below a transition temperature of ≈ 320 K, the stability of the intra-molecular hydrogen bonds grows faster, than in the high temperature range ($E_a=7.5$ kJ/mol in the high temperature range, $E_a=10.7$ kJ/mol in the low temperature interval) (see figure (4.10)). This indicates the appearance of an additional stabilizing influence, which was not present at high temperatures. When we compare the temperature dependence of the average life time of the intra-molecular hydrogen bonds with the spanning probability of the water cluster, we find that the appearance of the additional stabilizing force coincides with the occurrence of the spanning network of hydration water around the peptide. This is another very important evidence, how spanning water network may stabilize a *compact* conformation of the ELP chain.

Clearly, the rigid low-temperature conformational state of the ELP, seen in our simulations, does not resemble a random coil, as it does not show a random distribution of the end-to-end distance. This contradicts the idea that the ELP is a random coil at low temperatures [5, 6, 7, 72]. This idea has been based mainly on the specific temperature behavior of a minimum in the CD spectrum of the ELP in water at 195 nm: it becomes less pronounced with increasing temperature [7, 35, 36]. This minimum was attributed to the random coil state of the ELP, and its disappearance upon heating was therefore attributed to an increasing order of the ELP chain. This interpretation of the CD spectra gave rise to the idea that ELPs, in contrast to most other biomolecules, undergo a folding upon heating ("inverse temperature transition") [5, 6, 72]. When we consider an aqueous solution of ELPs, an increase of the order at macroscopic level

upon heating appears as it separates into two phases at the lower critical solution temperature. This decrease of entropy at the macroscopic level may be overcompensated by its increase at the microscopic level upon heating [73] (breaking of intra- or inter-molecular H-bonds, disordering of hydration water, etc.). Thus, a single ELP chain does not necessarily become more ordered upon heating. The minimum of the CD spectra at 195 nm can be attributed not only to a random coil but also to an ordered polyproline II structure [74, 75, 76, 77]. This structure disappears upon thermal denaturation of some proteins [78]. This causes a change of the CD spectra which is qualitatively very similar to the one observed upon "inverse temperature transition" of ELP. Moreover, the minimum at 195 nm in the CD spectra of ELPs in water is much less negative at low temperatures than one would expect for a random coil [79]. Our simulation results also indicate a slight decrease of the content of the polyproline II structures upon heating (see figure (4.16)). Thus the results both of experiments and simulations can be explained without the assumption of folding of the ELPs upon heating. The conformational transition of hydrated ELPs upon heating appears rather as a redistribution of the populations of the various locally ordered structures. The main redistribution seems to be connected with a decrease of polyproline II structures and the appearance of other structures such as β -turns, γ -loops, etc [7, 36, 35]. Interestingly, it was observed that the hydration water network stabilizes polyproline II structures [74]. This may explain why the thermal breaking of the spanning network of hydration water and the conformational transition of ELP occur simultaneously.

5.2 Influence of cosolvents on the structure and elasticity of the elastin-like peptide.

Based on our results we can divide the cosolvents into two groups: Urea and TFE, which make the ELP chain rigid and extended; GdmCl and NaCl, which make ELP molecule rigid and compact; this indicates two different mechanisms of the interactions of the cosolvents with the ELP. It was found in experiments [5] that large ELPs show their full elastic properties only when the weight fraction of water is more than 63 %. In our simulation the weight fraction of water in the water/*urea*/elastin mixture is only 55 % (see table (3.1)) *and the relative hydration level of the peptide is only 60 %*. The molecules of *urea* partially expell water(see table (4.2)) molecules from the surface and presumably form hydrogen bonds with the peptide. As it was stated in the previous section, water is the plasticizer (for details see section 5.4) for the ELP chain. Due

to the small size, fast reorientation time and high mobility, water eliminates energy barriers between conformations of the ELP chain with different Δi and increases the probability of the transitions between them. Thus removing the water from the surface, *urea* makes ELP more rigid. One can argue, that *urea* also makes hydrogen bonds with peptides, and can play the role of plasticizer for the ELP molecule. The *urea* molecule is approximately 4 times larger than the water molecule, the reorientation time and dynamical properties of *urea* molecule should be much slower than those for water. So firstly because of its size *urea* molecules can not come close enough to the intra-molecular hydrogen bonds, due to steric hindrance. Secondly as *urea* molecules are slower than water molecules, so the energy barriers between conformations with different Δi will be much higher, the switching of the intra-molecular hydrogen bonds, which *urea* molecules should mediate, will be slower, or will not occur at all. These can explain why Gaussian and worm-like chain equations failed to describe the probability distribution of the end-to-end distance of the ELP molecule in the aqueous solution of *urea* (see figure (4.19)). In recent CD experiments on a small ELP peptide [65] the spectral minimum which in pure water was at 195 nm (polyproline II structure), is shifted, by the addition of *urea*, to higher values ≈ 205 nm. This experimental finding supports our assumptions, that *urea* distorts the polyproline II structure of the chain and makes a single ELP molecule more disordered and rigid. In our simulations the ELP molecule is more extended in the *urea* solution than in pure water. This is in agreement with previous assumptions that *urea* molecules interact more favorably with non polar groups than water [80], consequently in the *urea* environment the peptide should be more exposed to the solvent. As we do not find any pronounced secondary structure in the ELP chain in the aqueous solution of *urea* (see figures (4.21), (4.22)) neither at low nor at high temperatures, an ordering transition with increasing of temperature was not observed in our simulations.

As was already mentioned, a 4M aqueous solution of *TFE* separated into a water rich and a *TFE* rich phase. The interface between the water and the *TFE* phases are not perfectly formed (see figure (A.1) in Appendix), due to the small number of *TFE* molecules (see table (3.1)). The ELP molecule is preferentially immersed in the *TFE* rich phase (see table (4.4)). It was found experimentally [11] that at low hydration levels large ELPs are fragile and brittle. In our simulations the ELP chain became extended and very rigid in the *TFE* surroundings. The *TFE* molecules are hydrophobic, that is why in the *TFE* environment the hydrophobic ELP chain is more extended than in aqueous media. The rigidity of the ELP molecule can be

explained, using the same arguments, which was applied in the *urea* case. If we will compare the fluctuations of the end-to-end distance R and of the radius of gyration S of the ELP chain in the *TFE* and *urea* solutions, we can find, that the ELP molecule is more flexible in the *urea* than in the *TFE* solution. In the *urea* solution the number of water molecules on the surface of the ELP is 1.7 times smaller than in pure water (see table (4.2)), this amount of water can not provide full elasticity, but could supply a media for the small fluctuations. In the *TFE* environment the number of water molecules on the surface of the ELP is 2.5 times smaller than in pure water (see table (4.4)), and such an amount of water can not provide a medium even for small fluctuations for the ELP chain. Thus, due to the smaller content of water on the surface of the ELP, energy barriers between different conformational states are even more pronounced in the *TFE* media than in the aqueous *urea* solution. The minimum in the CD spectra at 195 nm, which originates from the polyproline II structure, is shifted to the higher wave length ≈ 200 nm and becomes less pronounced in the 2M aqueous *TFE* solution [65]. This fact evidences, that *TFE* deforms the polypeptide chain, makes it more disordered and stiff.

In contrast to the aqueous solutions of *urea* and *TFE*, the ELP chain is *compact* and rigid in the GdmCl and NaCl water mixtures (see figures (4.28), (4.34) and tables (4.5), (4.7)). In NaCl solution the weight fraction of water is more than 90 % and relative hydration level is 100 %. So the ELP should have enough plasticizing media for the conversions between conformations with different Δi . Thus we can assume, that the factors which are responsible for the stiffness of the ELP chain are different from those in the *urea* and the *TFE* cases. The *compactness* of the peptide chain could be explained, by the increase of the polarity of the solution, due to presence of the ions. So the non-polar amino acids have to reduce their solvent accessible area more than in the pure water. The average number of ions at the surface of the ELP is ≈ 3.0 . Bonds between N-H or C=O groups and ions are more energetically favorable, than bonds with water. So, ions at the surface of the ELP chain could successfully compete with water for the "hydrogen" bonds. One ion can bind from 2 to 5 N-H or C=O groups, this can be an additional explanation, why the ELP has a more *compact* conformation. The *inflexibility* of the ELP in the aqueous solutions of *NaCl* can be explained by the direct binding of the ions to the peptide chain. The energy barrier between configurations with bound ions and without them is extremely large, thus the peptide will almost all the time stay in "ion bound" conformation and extensions or contractions of the ELP chain will appear very rarely. The weight fraction of

water in the *GdmCl* solution is more than 70 %, but the relative hydration level (see definition in section 4.2.1) is approximately 80 %. As far as *GdmCl* is in solution in the dissociated form and we observe the same number of ions at the surface of the ELP as in NaCl solution. All statements which were made for the influence of ions on the conformation of the ELP, are also true for the solution of *GdmCl*. However due to the large size of the Guanidinium ion, the relative hydration level of the ELP chain is also changed (see table (4.8)). So we can assume, that GdmCl affects the structure and the *flexibility* of the ELP chain in two ways. Firstly it reduces the relative hydration level of the peptide and thus removes plasticizing media from the ELP surface. Secondly it makes the ELP molecule stiffer by direct interactions of ions with the peptide.

5.3 Relation between the thermal breaking of a spanning network of hydration water and structural changes of biomolecules.

Our simulation studies of water clustering and percolation in the hydration shell of the ELP polypeptide show that, at low temperatures the molecule is covered by a spanning hydrogen-bonded network, formed by water molecules from the first hydration shell. This network includes nearly all water molecules in the hydration shell and covers the surface of the biomolecule homogeneously. With increasing temperature, this spanning water network breaks up into an ensemble of small water clusters in the hydration shell, and even the largest water cluster occupies only a small surface area of the polypeptide.

The transformation of a hydration water shell from a more ordered configuration with one large cluster, which spans the whole molecule, to a less ordered configuration with small water clusters, occurs via a percolation transition. This means that such a "order-disorder transformation" of the hydration shell is a transition, which takes place in a rather narrow temperature interval. Due to the finite number of water molecules in the hydration shell, such a transformation is smeared out in a certain temperature range and can be characterized by the temperature dependence of the probability SP to find a spanning water network. Below a certain temperature, the spanning network of hydration water exists practically permanently, but when crossing the percolation threshold, the spanning probability starts to decrease sharply.

The percolation transition of hydration water (shell width $D=4.5$ Å) of the ELP molecule occurs at about 280 K (figure (4.41)), i.e., below this temperature the spanning network is present in the system practically permanently. With increasing temperature, it starts to disappear and at $T=330$ K the probability to find a spanning cluster is about 50 % only.

As it was discussed in *section 5.1*, the breaking of a spanning network of hydration water may be related to temperature-induced conformational transformations of biomolecules and phase separation of their aqueous solutions. It is difficult to say at what temperature the extend of decomposition of the spanning network of hydration water is sufficient to cause such phenomena. This temperature definitely exceeds the temperature of the percolation transition of hydration water, and it is likely below the temperature, where $SP=50$ %. The interval between the temperature of the percolation transition $SP=95$ % and the temperature, where $SP=50$ %, shrinks with increasing size of the polypeptide. For the small ELP this interval achieves 50 K (figure (4.41)), whereas for the larger SNase it is only about 20 K [57].

As it was shown in figure (4.44), the value of n_H at the percolation threshold is about 2.1 for any reasonable choice of the hydration shell width D . The value n_H for the hydration water in protein solutions decreases almost linearly with temperature (see figure (4.44) and also Ref. [81]). This gives us the possibility of estimating the value of n_H in a wide temperature range by simulations at two temperatures and, subsequently, of estimating the temperature of the thermal breaking of the spanning network of hydration water. Based on our studies of the quasi-2D percolation transition of hydration water, we propose a relatively simple method to indicate the existence of a spanning network of hydration water by an analysis of the average number n_H of hydrogen bonds between the water molecules in the hydration shell. This value can be obtained by conventional computer simulations or estimated from experimental data [82]. At ambient temperature, the threshold value of n_H is about 2.1 for fully hydrated systems. A value n_H above this threshold indicates the presence of a spanning network of hydration water.

A comparison of the absolute values of the temperatures of the percolation transition of water in the hydration shell of the ELP, obtained in simulations, with the real temperature scale needs special consideration. The phase diagrams of the available water models differ noticeably from the phase diagram of real water (see Refs. [83] and [84] for a comparative analysis of the phase diagrams of various water models). There are two main characteristic temperatures which can be used for estimating the temperature shift of the phase diagram of

model water with the behavior of real water: the critical temperature of the liquid-vapor phase transition and the temperature of the liquid density maximum. The latter temperature is the most important parameter for our study, as it is close to the considered temperature interval.

The SPCE water model shows the liquid density maximum at about 240 K, i.e., about 35 K below its location in real water [84]. Hence, to map approximately our results on the "temperature scale" of real water, we have to shift the data points, shown in figure (4.41), upwards by about 35 K. In this case, if we take $D=4.5 \text{ \AA}$, the percolation transition occurs at $\approx 315 \text{ K}$, and at $T \approx 365 \text{ K}$ the spanning cluster exists with a probability $\sim 50 \%$. An analysis of the conformation of the peptide GVG(VPGVG)₃, shows pronounced changes from a *rigid* to an *elastic* state at a temperature of about 290 – 310 K, which should correspond to the interval of 325 – 345 K if the temperature shift between SPCE and real water is taken into account. In real aqueous solutions of the large poly(VPGVG) elastin-based polymer, the phase separation into water-rich and organic-rich phases, accompanied by sharp conformational changes of the polymer (the so-called "inverse temperature transition"), occurs at about 300 K [6]. In short elastin-like peptides, where the phase separation was not detected, pronounced conformational changes of ELP molecules are still observed at temperature 300 – 310 K for the peptide GVG(VPGVG)₃ [7]. Thus, the experimentally measured conformational transition of the peptide GVG(VPGVG)₃ occurs in the same temperature range, where the spanning network of the hydration water breaks into an ensemble of small water clusters in our simulations. Both experimental [36, 37] and simulation [13, 37] studies of the even smaller GVG(VPGVG) show a conformational transition at about 310 – 330 K. Interestingly, the qualitative changes of the peptide-water dynamics detected in Ref. [13] at this temperature may reflect specific changes of the hydration shell.

The temperature of the breakup of the spanning network of hydration water depends on the definition of the hydration shell. In the present study the simplest possible definition, based on the distance between the water oxygens and heavy atoms of the biomolecule, was used. A more unambiguous and physically well grounded definition of hydration water should be developed for a more quantitative analysis of its percolation transition.

Summarizing, we have shown that the transformation of the hydration water shell of a polypeptide from an "ordered" to a "disordered" state occurs via a quasi-2D percolation transition, which happens in a biologically relevant interval of temperatures. This transition may have a direct relation to temperature-induced conformational transitions of biomolecules and

temperature-induced phase separations of their aqueous solutions. Some internal protein movements could be affected (or even activated) by the spanning water network, when their characteristic time scales become comparable with the lifetime of the spanning water network [68].

5.4 Origin of elasticity.

According to this simulation study the elasticity of the ELP is of entropic origin, this is in agreement with previously published works [14, 15, 45]. But in contrast to rubber, the presence of a plasticizing agent is necessary for the elasticity of the elastin-like peptide. Water molecules play the role of a plasticizing media for the ELP. The entropic elasticity of the chain means, that the *potential* energies of the equilibrium and perturbed states are equal. The elastic restoring force appears due to the reduction of the *entropy* of the peptide, caused by the distortion of the chain from the equilibrium Gaussian distribution, due to extension or contraction. Since polypeptides have backbone hydrogen bonds, the *potential* energies of the different conformational states could be the same only, when the number of the intra-molecular hydrogen bonds is identical in both states. Thus to keep the *potential* energy constant the Δi of the backbone hydrogen bonds should change during the extension or contraction. When the elastin-like peptide is dry, the potential barrier between conformations with different Δi is ≈ 12 kJ/mol (hydrogen bond), this barrier makes the probability of transition of the peptide between the conformations with different Δi negligible. In the fully hydrated ELP system, water molecules eliminate the potential barrier by creating an intermediate state with the peptide. From the thermodynamic point of view, the free energy of the "main" state, with one backbone hydrogen bond and one hydrogen bond between two water molecules, is equal to the free energy of the "intermediate" state, with two hydrogen bonds between peptide backbone and two water molecules (see figure (4.11) in section 4.1.2). The potential barrier between these two states is much smaller, than the potential barrier between states of the peptide chain with different Δi . So the hydration of the peptide leads to the reduction of potential barriers between conformations with different Δi and makes the ELP *flexible* and *elastic*. A similar mechanism has been observed in the liquid water, which is responsible for the high liquidity of water [85]. The state of the hydration water should be also considered. When the ELP molecule is covered by the spanning cluster of hydration water, it does not show elastic properties. Since all water molecules in the hydration shell belong to the one permanently existing cluster, this can hamper the formation of the "intermediate" state. Disruption of the spanning cluster of hydration water leads to the vanishing of the hindrance. Cosolvents interact with the ELP chain in two different ways. *Urea* and *TFE* make the peptide stiffer indirectly. They expel water molecules from the peptide surface and thus take away plasticizing media from the ELP

molecule . *NaCl* interacts directly with peptide chain. Ions tightly bind to the ELP chain, and so prevent any transition between conformational states. *GdmCl* combines both direct and indirect ways.

Chapter 6

Summary.

Molecular dynamics simulations of the small elastin-like peptide (ELP), NMeGVG(VPGVG)₃-ACE, in pure water and in solutions of various cosolvents are reported in this work.

Two different conformational states of the ELP are distinguished in pure water, namely a *rigid* and a *flexible* one. A transition between these two states is observed at about $T_t=310$ K. In the *rigid* conformational state the ELP is *compact*. The average values of the radius of gyration S , end-to-end distance R and maximal extension L and the *relative fluctuations* of these values are small. The shapes of the probability distributions of the R , S and L have non-random (non-Gaussian) form. The large amounts of *irregular* hydrogen bonds, which connect residues separated by more than 5 amino acids, and quite large average live times of the hydrogen bonds are characteristic for the *rigid* state. There is no pronounced secondary structure observed in the *rigid* conformational state. The ELP in the *flexible* conformational state is slightly more *extended* than in the *rigid* conformational state. The *relative fluctuations* of the parameters R , S and L are up to 4 times greater than those in the *rigid* state. The shapes of the probability distributions of R , S and L have random (Gaussian) form. This reflects that the *elasticity* of the ELP in the *flexible* state has a rubber like *entropic* nature. *Irregular* hydrogen bonds are not present in the *flexible* conformational state of the ELP. The activation energy of the disruption of the hydrogen bond of the ELP in the *flexible* state is an 1.5 times smaller than that in the *rigid* state. A pronounced secondary structure is not found in the *flexible* conformational state. Changes in the α -helical and PPII content are observed with increasing temperature. The conformational transition of the ELP upon heating appears as a random redistribution of the populations of the various secondary structural elements.

A drastic change of the state of the hydration water of the ELP is observed with increas-

ing temperature. The hydration water undergoes a transition from one large cluster, which homogeneously covers the surface of the peptide, into an ensemble of small clusters, so called percolation transition. The 2D percolation transition of the hydration water occurs in the same temperature interval as the *conformational* transition of the ELP.

The addition of cosolvents make the ELP chain less flexible. Cosolvents interact with the ELP in two different ways. Urea and 2,2,2-trifluoroethanol interact indirectly: they remove the water from the surface of the ELP and thus make it *rigid* and *extended*. The lack of backbone-backbone hydrogen bonds with $\Delta i=4$ or 5, evidences the *extended* structure of the ELP in aqueous solution of these cosolvents. Sodium chloride (NaCl) interacts with ELP directly: ions which come to the surface of the ELP make it *compact* and *rigid*. The large amount and the high stability of intra-molecular hydrogen bonds of the ELP in the aqueous solution of NaCl support this claim. Guanidinium chloride (GdmCl) interact with the ELP in both ways. Firstly, it removes water from the surface of ELP, due to the large size of the Guanidinium⁺ ions. Secondly, ions interact directly with the peptide backbone. A pronounced secondary structure of the ELP chain is not observed in any cosolvent solution.

The role of water as a plasticizing media is considered in our work. An alternative mechanism of the elasticity of the elastin-like peptide is proposed. It is based on the *entropic* rubber-like elasticity of the ELP chain and on the interactions of the hydration water with the peptide backbone.

Chapter 7

Zusammenfassung.

In dieser Arbeit werden Molekulardynamik-Simulationen des kleinen elastinartigen Peptids NMeGVG(VPGVG)₃ACE in reinem Wasser und in Lösungen mit unterschiedlichen Kosolventien beschrieben.

Es werden zwei unterschiedliche Konformationszustände des Peptids gefunden, und zwar ein *starrer* und ein *flexibler*. Ein Übergang zwischen diesen beiden Zuständen wird bei etwa $T_t=310$ K gefunden. Im *starren* Konformationszustand ist das Peptid *kompakt*, das heißt, sowohl die Mittelwerte als auch die *relativen Fluktuationen* des Gyrationradius S , des End-zu-End-Abstands R und der maximalen Ausdehnung L sind klein. Die Kurvenverläufe der Wahrscheinlichkeitsverteilungen dieser drei Größen haben nicht die Form einer Gauß'schen Glockenkurve. Weiterhin charakteristisch für den *starren* Zustand sind neben einer recht großen mittleren Lebensdauer der Wasserstoffbrücken auch der große Anteil an *irregulären* Wasserstoffbrücken zwischen Resten, die mehr als 5 Aminosäuren voneinander entfernt sind. Beim *starren* Konformationszustand wird keine ausgeprägte Sekundärstruktur beobachtet. Im *flexiblen* Konformationszustand besitzt das Peptid eine etwas größere Ausdehnung als im *starren* Konformationszustand: Die *relativen Fluktuationen* der Größen R , S und L sind bis zu viermal größer. Die Verläufe der entsprechenden Wahrscheinlichkeitsverteilungen weisen die Form einer Zufallsverteilung (Gauß'sche Glockenkurve) auf, das bedeutet, dass die Elastizität des Peptids im *flexiblen* Zustand eine gummiartige *entropische* Natur hat. *Irreguläre* Wasserstoffbrücken treten im *flexiblen* Konformationszustand des Peptids nicht auf. Die Aktivierungsenergie für die Spaltung einer intramolekularen Wasserstoffbrücke im *flexiblen* Zustand des Peptids ist ungefähr 1,5-mal kleiner als im *starren* Zustand. Eine ausgeprägte Sekundärstruktur wird auch in diesem Zustand nicht beobachtet. Mit steigender Temperatur

ändern sich die α -helikalen and die PPII-Anteile. Der Konformationsübergang des Peptids beim Erwärmen entspricht einer zufälligen Umverteilung der Anteile der unterschiedlichen Sekundärstrukturelemente.

Mit steigender Temperatur wird eine drastische Änderung des Zustands der Hydrathülle des Peptids beobachtet. Es handelt sich um einen so genannten Perkolationsübergang: Während die Hydrathülle bei niedrigen Temperaturen einem einzigen großen Cluster entspricht, der die Oberfläche des Peptids homogen bedeckt, besteht sie bei hohen Temperaturen aus einer Vielzahl kleiner Cluster. Der 2D-Perkolationsübergang der Hydrathülle tritt im gleichen Temperaturbereich auf wie der Konformationsübergang des Peptids.

Der Zusatz von Kosolventien verringert die Flexibilität des Peptids. Kosolventien wechselwirken mit dem Peptid auf zwei unterschiedliche Arten. Harnstoff und Trifluorethanol wechselwirken indirekt: Sie entfernen das Wasser von der Oberfläche des Peptids und sorgen so dafür, dass es *ausgedehnt*, aber *starr* wird. Das Nichtvorhandensein von Backbone-Backbone-Wasserstoffbrücken mit $\Delta i=4$ oder 5 ist ein Beweis für die *ausgedehnte* Struktur des Peptids in wässrigen Lösungen dieser Kosolventien. Natriumchlorid hingegen wechselwirkt direkt: Die Ionen nähern sich der Oberfläche des Peptids, wodurch es *kompakt* und *starr* wird. Der große Anteil und die hohe Stabilität der intramolekularen Wasserstoffbrücken in wässrigen NaCl-Lösungen unterstützen diese Vermutung. Guanidiniumchlorid wechselwirkt auf beide Arten mit dem Peptid: Zum einen entfernt es aufgrund der großen Ausdehnung des Guanidinium⁺-Ions Wasser von der Oberfläche des Peptids, zum anderen wechselwirken die Ionen direkt mit dem Backbone des Peptids. Eine ausgeprägte Sekundärstruktur des Peptids wird in Lösungen der betrachteten Kosolventien nicht beobachtet.

In dieser Arbeit wird die Rolle des Wassers als Plastifizierungsmittel beschrieben. Ein alternativer Mechanismus für die Elastizität von elastinartigen Peptiden wird vorgeschlagen. Er basiert auf der *entropischen* gummiartigen Elastizität der Peptidkette und den Wechselwirkungen der Wassermoleküle in der Hydrathülle mit dem Backbone des Peptids.

Appendix A

Temp / K	R_{av} / Å	Sd / Å	δ_R / %	S_{av} / Å	Sd / Å	δ_S / %	L_{av} / Å	Sd / Å	δ_L / %
280	12.56	4.63	37	6.68	0.65	10	22.10	2.31	10
285	11.80	6.94	59	6.80	0.89	13	22.19	3.18	14
290	16.11	5.61	35	7.08	1.26	18	23.15	4.38	19
295	16.22	5.55	34	7.40	1.06	14	24.53	3.61	15
300	14.91	6.22	42	7.33	1.06	15	24.09	3.71	15
320	14.45	6.21	43	7.18	0.98	14	23.21	3.73	16
340	17.06	6.33	37	7.57	1.06	14	25.00	3.92	16
360	15.81	6.32	40	7.30	0.99	14	23.89	3.80	16
380	15.82	6.24	39	7.31	0.97	13	23.84	3.73	16
400	17.12	6.73	39	7.59	1.08	14	25.01	4.04	16
420	16.82	6.58	39	7.71	1.08	14	25.18	4.05	16
440	17.46	6.94	40	7.81	1.12	14	25.56	4.24	17

Table A.1: Average values, standard deviations and relative fluctuations of the end-to-end distance R , radius of gyration S and maximal extension L of the ELP in pure water at different temperatures.

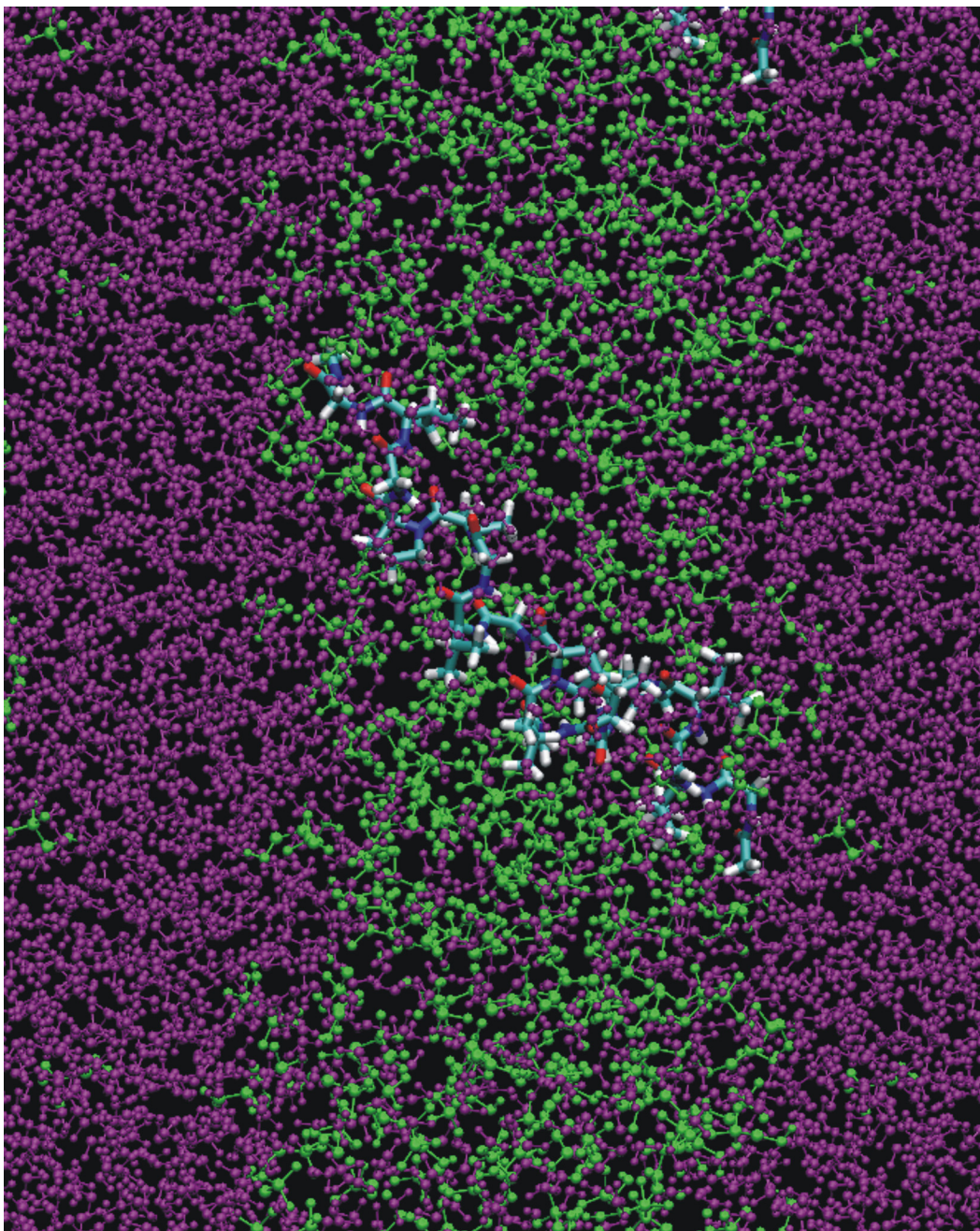


Figure A.1: Phase separation of the 4M aqueous TFE solution into a water rich (violet) and a TFE rich (green) phase.

Bibliography

- [1] S. M. Partridge, H. F. Davis, and G. S. Adair. The chemistry of connective tissues. 2. Soluble proteins derived from partial hydrolysis of elastin. *Biochem. J.*, 61(1):11–21, Sep 1955.
- [2] S.M. Partridge. Elastin. *Adv. Prot. Chem.*, 17:227 – 302, 1962.
- [3] L. B. Sandberg, J. G. Leslie, C. T. Leach, V. L. Alvarez, A. R. Torres, and D. W. Smith. Elastin covalent structure as determined by solid-phase amino-acid sequencing. *Pathol. Biol.*, 33:266–274, 1985.
- [4] H. Yeh, N. Ornstein-Goldstein, Z. Indik, P. Sheppard, N. Anderson, J. C. Rosenbloom, G. Cicila, K. G. Yoon, and J. Rosenbloom. Sequence variation of bovine elastin messenger-rna due to alternative splicing. *Collagen Relat. Res.*, 7:235–247, 1987.
- [5] D. W. Urry and T. M. Parker. Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. *J Muscle Res. Cell. Motil.*, 23(5-6):543–559, 2002.
- [6] D. W. Urry. Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *J. Phys. Chem. B*, 101:11007 – 11028, 1997.
- [7] H. Reiersen, A.R. Clarke, and A.R. Rees. Short elastin-like peptides exhibit the same temperature-induced structural transitions as elastin polymers: implications for protein engineering. *J. Mol. Biol.*, 283:255 – 264, 1998.
- [8] D. W. Urry, C. H. Luan, and S. Q. Peng. Molecular biophysics of elastin structure, function and pathology. *Ciba Found. Symp.*, 192:4–22; discussion 22–30, 1995.

- [9] C. H. Luan, T. M. Parker, K. U. Prasad, and D. W. Urry. Differential scanning calorimetry studies of NaCl effect on the inverse temperature transition of some elastin-based polytetra-, polypenta-, and polynonapeptides. *Biopolymers*, 31(5):465–475, Apr 1991.
- [10] N. Nath and A. Chilkoti. Interfacial phase transition of an environmentally responsive elastin biopolymer adsorbed on functionalized gold nanoparticles studied by colloidal surface plasmon resonance. *J. Am. Chem. Soc.*, 123(34):8197–8202, August 2001.
- [11] S. R. Kakivaya and C. A. J. Hoeve. The glass point of elastin. *Proc. Nat. Acad. Sci. USA*, 72:3505–3507, 1975.
- [12] B Li, D.O. Alonso, and V Daggett. The molecular basis of the inverse temperature transition of elastin. *J. Mol. Biol.*, 305:581 – 592, 2001.
- [13] R. Rousseau, E. Schreiner, A. Kohlmeyer, and D. Marx. Temperature-dependent conformational transitions and hydrogen-bond dynamics of the elastin-like octapeptide GVG(VPGVG): a molecular-dynamics study. *Biophys. J.*, 86:1393 – 1407, 2004.
- [14] C. A. J. Hoeve and P. J. Flory. The elastic properties of elastin. *J. Am. Chem. Soc.*, 80:6523 – 6526, 1958.
- [15] C. A. J. Hoeve and P.J. Flory. The elastic properties of elastin. *Biopolymers*, 13:677 – 686, 1974.
- [16] W. R. Gray, L.B. Sandberg, and J. A. Foster. Molecular model of elastin structure and function. *Nature*, 246:461 – 466, 1973.
- [17] T. Weis-Fogh and S. O. Andersen. New molecular model for the long-range elasticity of elastin. *Nature*, 227:718–721, 1975.
- [18] L. Debelle and A. M. Tamburro. Elastin: molecular description and function. *Int. J. Biochem. Cell Biol.*, 31(2):261–272, Feb 1999.
- [19] Z. R. Wasserman and F. R. Salemme. A molecular dynamics investigation of the elastomeric restoring force in elastin. *Biopolymers*, 29:1613–1631, 1990.
- [20] L. Hermann. Ein versuch ueber der sogenante Sehnenverkuerzung. *Archiv fuer der gesamte Physiologie.*, 7:417 – 420, 1873.

- [21] E. Gotschlich. Ueber den Einfluss der Waerme auf Laenge und Dehnbarkeit des elastisches Gewebes und des gestreiften Muskels. *Archiv fuer die gesamte Physiologie*, 3:109 – 163, 1893.
- [22] G. W. Vandergrift and W. J. Gies. The composition of yellow fibrous connective tissue. *Am. J. Physiol.*, 5:287 – 297, 1901.
- [23] J. E. McCartney. Heat contraction of elastic tissue. *Q. J. Exp. Physiol.*, 7:103–114, 1913.
- [24] W.H. Stein and E.G. Miller. The composition of elastin. *J. Biol. Chem.*, 125:599 – 614, 1938.
- [25] Dorothy Jordan Lloyd and Marjorie Garrod. A contribution to the theory fo the structure of protein fibres with thermal shrinkage of the collagen special reference to the so-called fibre. *Trans. Faraday Soc.*, 64:441 – 451, 1947.
- [26] L. Gotte, P. Stern, D. F. Elsdén, and S. M. Partridge. The chemistry of connective tissues. 8. The composition of elastin from three bovine tissues. *Biochem. J.*, 87:344–351, May 1963.
- [27] S. M. Partridge. Elastin. *The pysiology and biochemistry of Muscle as a food*, Chapter 22:327 – 339, 1966.
- [28] J.M. Gosline. Hydrophobic interaction and a model for the elasticity of elastin. *Biopolymers*, 17:677–695, 1978.
- [29] D. W. Urry, M. M. Long, T. Ohnishi, and M. Jacobs. Circular dichroism and absorption of the polytetrapeptide of elastin: a polymer model for the beta-turn. *Biochem. Biophys. Res. Commun.*, 61(4):1427–1433, Dec 1974.
- [30] D. W. Urry, L. W. Mitchell, T. Ohnishi, and M. M. Long. Proton and carbon magnetic resonance studies of the synthetic polypentapeptide of elastin. *J. Mol. Biol.*, 96(1):101–117, Jul 1975.
- [31] D. W. Urry, L. W. Mitchell, and T. Ohnishi. Studies on the conformation and interactions of elastin secondary structure of synthetic repeat hexapeptides. *Biochim. Biophys. Acta*, 393(2):296–306, Jun 1975.

- [32] D. W. Urry, K. Okamoto, R. D. Harris, C. F. Hendrix, and M. M. Long. Synthetic, cross-linked polypentapeptide fo tropoelastin: an anisotropic, fibrillar elastomer. *Biochemistry*, 15(18):4083–4089, Sep 1976.
- [33] B. B. Aaron and J. M. Gosline. Optical properties of single elastin fibres indicate random protein conformation. *Nature*, 287(5785):865–867, October 1980.
- [34] D. A. Torchia and K. A. Piez. Mobility of elastin chains as determined by ^{13}C nuclear magnetic resonance. *J. Mol. Biol.*, 76(3):419–424, May 1973.
- [35] D. W. Urry, T. L. Trapane, and K. U. Prasad. Phase-structure transitions of the elastin polypentapeptide-water system within the framework of composition-temperature studies. *Biopolymers*, 24(12):2345–2356, Dec 1985.
- [36] C. Nicolini, R. Ravindra, B. Ludolph, and R. Winter. Characterization of the temperature- and pressure-induced inverse and reentrant transition of the minimum elastin-like polypeptide GVG(VPGVG) by DSC, PPC, CD, and FT-IR spectroscopy. *Biophysical Journal*, 86:1385 – 1392, 2004.
- [37] E. Schreiner, C. Nicolini, B. Ludolph, R. Ravindra, N. Otte, R. Rousseau, R. Winter, and D. Marx. Folding and unfolding of an elastinlike oligopeptide: Inverse Temperature Transition, reentrance, and hydrogen-bond dynamics. *Phys. Rev. Lett.*, 92:148101 (1–4), 2004.
- [38] B. Li and V. Daggett. Molecular basis for the extensibility of elastin. *J. Muscle Res. Cell. Motil.*, 23:561573, 2002.
- [39] T. Kowall and A. Geiger. Molecular dynamics simulation study of 18-crown-6 in aqueous solution. 1. structure and dynamics of the hydration shell. *J. Phys. Chem.*, 98(24):6216–6224, June 1994.
- [40] I. Brovchenko, A. Geiger, A. Oleinikova, and D. Paschek. Phase coexistence and dynamic properties of water in nanopores. *Eur. Phys. J. E*, 12:69 – 76, 2003.
- [41] P. J. Flory. *Principles of Polymer Chemistry*. Cornell University Press, 1953.
- [42] J. M. Gosline and C.J. French. Dynamic mechanical properties of elastin. *Biopolymers*, 18:2091 – 2103, 1979.

- [43] A. L. Andradý and J. E. Mark. Thermoelasticity of swollen elastin networks at constant composition. *Biopolymers*, 19:849–855, 1980.
- [44] B. Prescott, V. Renugopalakrishnan, and G. J. Thomas Jr. Raman spectrum and structure of elastin in relation to type-II β -turns. *Biopolymers*, 26:934–936, 1987.
- [45] D. Volpin and A. Ciferri. Thermoelasticity of elastin. *Nature*, 225:382, 1970.
- [46] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1997.
- [47] A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136(2A):405–411, October 1964.
- [48] M. Born and Th. von Karman. Ueber Schwingungen in Raumgittern. *Physik Zeitschrift*, 13:297–309, 1912.
- [49] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.
- [50] C. M. Venkatachalam and D. W. Urry. Development of a linear helical conformation from its cyclic correlate. β -spiral model of the elastin polypentapeptide (VPGVG) $_n$. *Macromolecules*, 14:1225 – 1229, 1981.
- [51] W.D. Cornell, P. Cieplak, C. Bayly, I. Gould, K.Jr. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J.I. Caldwell, and P. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.
- [52] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentialst. *J. Phys. Chem.*, 91:6269–6271, 1987.
- [53] P.E. Smith. Computer simulation of cosolvent effects on hydrophobic hydration. *J. Phys. Chem. B*, 103(3):525 – 534, January 1999.
- [54] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19):8577–8593, November 1995.

- [55] P. J. Reynolds, H. E. Stanley, and W. Klein. Ghost fields, pair connectedness, and scaling: exact results in one-dimensional percolation. *Journal of Physics A: Mathematical and General*, 10(11):L203–L209, 1977.
- [56] G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of Polypeptide Chain Configurations. *J. Mol. Biol.*, 7:95 – 99, 1963.
- [57] I. Brovchenko, A. Krukau, N. Smolin, A. Oleinikova, A. Geiger, and R. Winter. Thermal breaking of spanning water networks in the hydration shell of proteins. *J. Chem. Phys.*, 123(22):224905, 2005.
- [58] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor and Francis Inc., 1991.
- [59] A. Oleinikova, I. Brovchenko, and A. Geiger. Percolation transition of hydration water at hydrophilic surfaces. *Physica A: Statistical Mechanics and its Applications*, 364:1–12, May 2006.
- [60] A. Oleinikova and I. Brovchenko. Percolation transition of hydration water in biosystems. *Mol. Phys.*, 104(22):3841–3855, 2006.
- [61] A. Weiss. Private communication. August 2006.
- [62] A. Geiger, P. Mausbach, J. Schnitker, R.L. Blumberg, and H. E. Stanley. Structure and dynamics of the hydrogen bond network in water by computer simulations. *Journal de Physique*, 45:7 – 13, 1984.
- [63] S. Weerasinghe and P.E. Smith. A kirkwood-buff derived force field for mixtures of urea and water. *J. Phys. Chem. B*, 107(16):3891–3898, April 2003.
- [64] M. Buck. Trifluoroethanol and colleagues: cosolvents come of age. recent studies with peptides and proteins. *Quarterly Reviews of Biophysics*, 31(03):297–355, 1998.
- [65] C. Nicolini. *Lateral Organisation Of Modelbiomembranes And Partitioning Of Dually Lipidated N-RAS Proteins Into Ternary Model Raft Lipid Mixtures*. PhD thesis, University of Dortmund, 2006.

- [66] A. Oleinikova, I. Brovchenko, N. Smolin, A. Krukau, A. Geiger, and R. Winter. Percolation transition of hydration water: From planar hydrophilic surfaces to proteins. *Phys. Rev. Lett.*, 95(24):247802, 2005.
- [67] A. Oleinikova, N. Smolin, I. Brovchenko, A. Geiger, and R. Winter. Formation of spanning water networks on protein surfaces via 2D percolation transition. *J. Phys. Chem. B*, 109:1988–1998, 2005.
- [68] N. Smolin, A. Oleinikova, I. Brovchenko, A. Geiger, and R. Winter. Properties of spanning water networks at protein surfaces. *J. Phys. Chem. B*, 109,:10995–11005, 2005,.
- [69] I. Brovchenko and A. Oleinikova. "Molecular Organisation of Gases and Liquids at Solid Surfaces" in *Handbook of Theoretical and Computational Nanotechnology.*, volume 9. American Scientific Publishers, 2006.
- [70] L. Debelle, A. Alix, S. Wei, M.-P. Jacob, J.-P. Huvenne, M. Berjot, and P. Legrand. The secondary structure and architecture of human elastin. *Eur. J. Biochem.*, 258:533 – 539, 1998.
- [71] A. Krukau, I. Brovchenko, and A. Geiger. Temperature-induced conformational transition of a model elastin-like peptide GVG(VPGVG)₃ in water. *Biomacromolecules*, 8:2196 – 2202, June 2007.
- [72] D. W. Urry, T. Hugel, M. Seitz, H. E. Gaub, L. Sheiba, J. Dea, J.Xu, and T. Parker. Elastin: a representative ideal protein elastomer. *Phil. Trans. R. Soc. Lond. B*, 357:169 – 184, 2002.
- [73] J. Hirschfelder, D. Stevenson, and H. Eyring. A theory of liquid structure. *J. Chem. Phys.*, 5(11):896–912, November 1937.
- [74] F. Eker, K. Griebenow, and R. Schweitzer-Stenner. Stable conformations of tripeptides in aqueous solution studied by UV circular dichroism spectroscopy. *J. Am. Chem. Soc.*, 125(27):8178–8185, July 2003.
- [75] N. Sreerama and R.W. Woody. Polyproline II helixes in globular proteins: Identification and circular dichroic analysis. *Biochemistry*, 33(33):10022–10025, August 1994.

- [76] I. Gokce, R.W. Woody, G. Anderluh, and J.H. Lakey. Single peptide bonds exhibit polyproline II ("random coil") circular dichroism spectra. *J. Am. Chem. Soc.*, 127(27):9700–9701, July 2005.
- [77] N. Sreerama and R.W. Woody. Structural composition of β I- and β II-proteins. *Protein Sci*, 12(2):384–388, February 2003.
- [78] Jie Wu, Jen Tsi Yang, and Chuen-Shang C. Wu. β -II conformation of all- β proteins can be distinguished from unordered form by circular dichroism. *Anal. Biochem.*, 200(2):359–364, February 1992.
- [79] L. Debelle and A. J. P. Alix. The structures of elastins and their function. *Biochimie*, 81(10):981–994, October 1999.
- [80] C. Dobson and M. Karplus. The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.*, 9:92 – 101, 1999.
- [81] S. Melchionna, G. Briganti, P. Londei, and P. Cammarano. Water induced effects on the thermal response of a protein. *Phys. Rev. Lett.*, 92(15):158101–4, April 2004.
- [82] M. Nakasako. Large-scale networks of hydration water molecules around proteins investigated by cryogenic X-ray crystallography. *Cell. Mol. Biol.*, 47:767–777, 2001.
- [83] B. Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.*, 101(1-3):219–260, November 2002.
- [84] I. Brovchenko, A. Geiger, and A. Oleinikova. Liquid-liquid phase transitions in supercooled water studied by computer simulations of various water models. *J. Chem. Phys.*, 123(4):044515–16, July 2005.
- [85] F. Sciortino, A. Geiger, and H.E. Stanley. Effect of defects on molecular mobility in liquid water. *Nature*, 354(6350):218–221, November 1991.

Lebenslauf.

Persönliche Daten

Name Aliaksei Krukau
Geburtsdatum 11 August 1980
Geburtsort Minsk, Weissrussland

Schulausbildung

1986-1995: Physic und Mathematic Liceum Nummer 116.
1995-1997: Physic und Mathematic Gymnasium Nummer 130.

Studium

09/1997 - 06/2002: International Sakharov Environmental University
04/2001 - 05/2002: Diplomarbeit: " Algorithm of Recognition of Human Blood T-Lymphocytes on the Digital Image and Its use in Diagnostics of Diseases."
27/06/2002: Abschluss: Diplom
10/2002 - z.Z.: Promotion bei Prof. Dr. Geiger and der Universität Dortmund

Berufstätigkeit

01/1999 - 01/2002: Wissenschaftliche Hilfskraft
10/2005 - 12/2007: Wissenschaftlicher Angestellter